2

Βασικές Έννοιες Πιθανότητας

ΠΕΡΙΓΡΑΦΗ ΚΕΦΑΛΑΙΟΥ

Σε αυτό το κεφάλαιο θα συζητήσουμε τον τύπο των φαινομένων με τον οποίο θα ασχοληθούμε σε αυτό το βιβλίο. Επί πλέον θα αναπτύξουμε μαθηματικά μοντέλα με τα οποία θα περιγράφουμε και διερευνούμε τέτοια φαινόμενα.

Γενικά, μπορούμε να διακρίνουμε δύο ειδών μοντέλα τα ***καθοριστικά*** και τα ***στοχαστικά.*** Τα καθοριστικά μοντέλα περιγράφουν φαινόμενα των οποίων οι συνθήκες κάτω από τις οποίες πραγματοποιούνται καθορίζουν και το αποτέλεσμα της έκβασής τους. Για παράδειγμα, αν τοποθετήσουμε μία μπαταρία σε ένα απλό κύκλωμα, το μαθηματικό μοντέλο το οποίο ενδεχομένως θα περιέγραφε την παρατηρούμενη ροή ρεύματος θα ήταν ο νόμος του Ohm $I=E/R$.. Ο νόμος προβλέπει την τιμή της έντασης του ρεύματος όταν είναι γνωστά η αντίσταση R και η δύναμη της μπαταρίας E. Όσες φορές και εάν επαναλάβουμε αυτό το πείραμα, κάθε φορά παρατηρούμε την ίδια ένταση I.

Υπάρχουν πολλά φαινόμενα τα οποία περιγράφονται πλήρως από καθοριστικά μοντέλα. Παρόλα αυτά υπάρχουν επίσης πολλά άλλα φαινόμενα τα οποία περιγράφονται από ***στοχαστικά*** ή ***πιθανοτικά*** μοντέλα. Για παράδειγμα, η ένταση της βροχής που πρόκειται να ακολουθήσει σε ένα τόπο δεν μπορεί να προβλεφθεί ακριβώς. Υπάρχουν καθοριστικά μετεωρολογικά μοντέλα τα οποία περιέχουν τις επικρατούσες συνθήκες όπως βαρομετρική πίεση, διεύθυνση και ένταση ανέμου και άλλες παρατηρούμενες παραμέτρους αλλά προβλέπουν **μόνο** χαμηλή ή υψηλή ένταση βροχής στον τόπο και όχι την ακριβή της τιμή. Τέτοια φαινόμενα περιγράφονται καλύτερα από στοχαστικά μοντέλα, τα οποία λαμβάνοντας υπόψη τις επικρατούσες συνθήκες προβλέπουν **πιθανολογικά** τα διάφορα δυνατά αποτελέσματα. Με άλλα λόγια τα στοχαστικά μοντέλα προβλέπουν με κάποια αβεβαιότητα την έκβαση του φαινομένου. Την **αβεβαιότητα** αυτή μετρά μεθοδολογικά η **Θεωρία Πιθανοτήτων.**

Η στατιστική, σαν μια μέθοδο *λήψης απόφασης* κάτω από συνθήκη αβεβαιότητας, βασίζεται στην πιθανοθεωρία, αφού πιθανότητα είναι ένα μέτρο της αβεβαιότητας και των ρίσκων που σχετίζονται μ' αυτή. Προτού μάθει κανείς στατιστικές διαδικασίες απόφασης απαιτείται μία γνώση της πιθανοθεωρίας.

Ένας εύκολος και ξεκάθαρος τρόπος μεταχείρισης της *πιθανοθεωρίας* απαιτεί μερική γνώση της *συνολοθεωρίας*. Προκειμένου να μελετήσουμε τα *πιθανοκρατικά πρότυπα* που επιθυμούμε να αναπτύξουμε, θα ήταν πολύ κατάλληλο αν πρώτα γνωρίζαμε τις βασικές έννοιες από την θεωρία συνόλων.

2.1 ΑΒΕΒΑΙΟΤΗΤΑ, ΤΥΧΑΙΑ ΔΙΑΔΙΚΑΣΙΑ, ΚΑΙ ΣΥΝΑΦΕΙΣ ΕΝΝΟΙΕΣ

Βασικά, η ανάπτυξη της πιθανοθεωρίας στηρίζεται στις τρεις έννοιες του **δειγματοχώρου**, **δειγματοσημεία**, και **γεγονότα**, οι οποίες σχετίζονται με την αβεβαιότητα και το πείραμα τύχης.

2.1.1 Αβεβαιότητα και Τυχαίο Πείραμα

Η **αβεβαιότητα** αναφέρεται στην έκβαση ενός φαινομένου υπό εξέλιξη. Εάν ένα φαινόμενο υπό εξέλιξη μπορεί να οδηγήσει σε δύο ή περισσότερα πιθανά αποτελέσματα, το φαινόμενο είναι ***στοχαστικό*** και τα αποτέλεσμα λέγεται ότι είναι ***αβέβαιο.*** Έτσι, ρίχνοντας ένα ζάρι ή νόμισμα, μετρώντας την διάρκεια μιας μπαταρίας, μετρώντας τον αριθμό αυτοκινήτων πριν συμβεί ένα ατύχημα, ερωτώντας έναν ψηφοφόρο εάν προτιμά ή όχι ένα συγκεκριμένο υποψήφιο, και ούτω κάθ' εξής, όλα είναι στοχαστικά φαινόμενα, αφού σε κάθε περίπτωση , η διαδικασία μπορεί να οδηγήσει σε περισσότερα από δύο πιθανά αποτελέσματα.

Ορισμός. Στο εξής, ένα στοχαστικό φαινόμενο θα το ονομάζουμε **τυχαίο πείραμα** αν,

1. η εξέλιξη του φαινομένου μπορεί να πραγματοποιηθεί όσες φορές επιθυμούμε,
2. η εξέλιξη του φαινομένου πραγματοποιείται πάντα υπό τις ίδιες συνθήκες,
3. η προ-εκδίκαση του αποτελέσματος μιας συγκεκριμένης εξέλιξης του φαινομένου δεν είναι δυνατή,
4. το σύνολο των αβέβαιων αποτελεσμάτων της εξέλιξης του φαινομένου είναι γνωστό.

Είναι σημαντικό να σημειωθεί ότι, όταν αναφερόμαστε σε πείραμα τύχης μπορεί να εννοούμε και ένα φυσικό πείραμα το οποίο μπορεί να επαναληφθεί πολλές φορές κάτω από τις ίδιες συνθήκες. Ή, μπορεί απλά να διανοηθεί ένα σύνολο από αβέβαια αποτελέσματα , χωρίς να πραγματοποιηθεί καμία ακολουθία από επαναλαμβανόμενες δοκιμές. Για ένα πείραμα τύχης, πραγματικό ή φανταστικό, το μόνο που χρειάζεται είναι να προσδιορισθούν ακριβώς όλα τα πιθανά αποτελέσματα.

2.1.2 Δειγματοχώρος και Δειγματοσημεία

Ορισμός. Ένα σύνολο του οποίου τα στοιχεία παριστάνουν όλα τα πιθανά αποτελέσματα του πειράματος τύχης ονομάζεται **δειγματοχώρος,** ο οποίος είναι ένα γενικό (*universal*) σύνολο και συμβολίζεται με ***S***. Τα πιθανά αποτελέσματα στον δειγματοχώρο ονομάζονται **δειγματοσημεία** και συμβολίζονται με ***s***, ο δε αριθμός δειγματοσημείων σ'ένα δειγματοχώρο μπορεί να συμβολιστεί με ***N*(*S*)**.

Στη συνέχεια αναφέρονται τυχαία πειράματα ***E*** και οι αντίστοιχοι δειγματοχώροι ***S***.

Παράδειγμα 2.1

E_1: Ενα τεμάχιο επιλέγεται από ένα κιβώτιο εισερχόμενου εμπορεύματος και ελέγχεται. Το τεμάχιο μπορεί να είναι ελαττωματικό, D, ή μη ελαττωματικό, $\overline{D}$.

$$S=\{D, \overline{D}\}.$$

Παράδειγμα 2.2

E_2: Ρίχνουμε τρία νομίσματα και παρατηρούμε τις ενδείξεις τους. Ο δειγματικός χώρος S του πειράματος αυτού αποτελείται ως γνωστόν από όλα τα δυνατά ενδεχόμενα,

$$S=\{KKK, KK\Gamma, K\Gamma K, \Gamma KK, K\Gamma\Gamma, \Gamma K\Gamma, \Gamma\Gamma K, \Gamma\Gamma\Gamma\}.$$

Παράδειγμα 2.3

E_3: Ένας εργολάβος ξεκινά ένα μεγάλο χωματουργικό έργο με τρεις μπουλντόζες A_1, A_2 και A_3. Μετά από ένα χρόνο ελέγχουμε την κατάσταση των τριών μπουλντόζων. Αν A_κ και $\overline{A}_\kappa$ είναι τα ενδεχόμενα ότι η A_κ μπουλντόζα παρουσιάζει ή δεν παρουσιάζει βλάβη αντίστοιχα, τότε ο δειγματοχώρος S της κατάστασης των μπουλντόζων είναι

$$S=\{A_1A_2A_3,\ \overline{A_1}\,A_2A_3,\ A_1\overline{A_2}A_3,\ A_1A_2\overline{A_3},\ \overline{A_1}\,\overline{A_2}\,A_3,\ A_1\overline{A_2}\,\overline{A_1},\ \ \overline{A_1}\,A_2\,\overline{A_3},\ \overline{A_1}\,\overline{A_2}\,\overline{A_1}\}$$

Παράδειγμα 2.4

E_4: Ένα φτερό αεροπλάνου συναρμολογείται με ένα μεγάλο αριθμό βιδών M. Απαριθμείται ο αριθμός των ελαττωματικών βιδών.

$$S=\{1, 2, 3, \ldots, M\}.$$

Παράδειγμα 2.5

E_5: Κατασκευάζονται ηλεκτρικοί λαμπτήρες μέχρι να παραχθούν 10 ελαττωματικοί. Παρατηρείται ο αριθμός των παραγόμενων λαμπτήρων.

$$S=\{10, 11, 12, \ldots\}.$$

Παράδειγμα 2.6

E_6: Μετρούμε την αντοχή έντασης ενός τύπου χαλύβδινης δοκού.

S={x| *x μη αρνητικός πραγματικός αριθμός*}.

Παράδειγμα 2.7

E_7: Ένας θερμογράφος καταγράφει συνεχώς την θερμοκρασία σ' όλο το 24-ωρο. Σ' ένα συγκεκριμένο τόπο σε ένα ψηλό βουνό κάθε μέρα το πρωί σημειώνεται η ένδειξη του θερμογράφου σε βαθμούς κλίμακας ^0C.

S={x| $m<x<M$, *όπου m η ελάχιστη δυνατή και M η μέγιστη δυνατή θερμοκρασία στον τόπο αυτόν*}.

Στο ίδιο παράδειγμα, αν παρατηρούμε για πόσες μέρες μέσα σε ένα χρόνο η θερμοκρασία έπεσε κάτω από τους μηδέν βαθμούς ^0C, ο δειγματοχώρος είναι διαφορετικός,

S={0, 1, 2, 3, 4, . . ., 364, 365}. ■

Ένας δειγματοχώρος με μικρό αριθμό δειγματοσημείων ενδέχεται να παρασταθεί ευκολότερα με διαγράμματα δέντρων ή άλλων γραφικών παραστάσεων σε δυδιάστατο ή τριδιάστατο σύστημα αξόνων. Ανεξάρτητα από το πώς περιγράφεται το S, τα στοιχεία του S τα οποία αντιστοιχούν στα αποτελέσματα του τυχαίου πειράματος πρέπει να είναι ***αμοιβαία αποκλειόμενα*** και *συλλεκτικά εξαντλημένα*. Αυτός είναι και ο ***ορισμός του δειγματοχώρου***.

Προκειμένου να περιγράψουμε έναν δειγματοχώρο σχετιζόμενο με ένα πείραμα, πρέπει να έχουμε ξεκάθαρη ιδέα από αυτό που μετρούμε ή παρατηρούμε. Σημειώστε επίσης ότι το αποτέλεσμα ενός πειράματος δεν είναι πάντοτε αριθμός, μπορεί να είναι ακολουθία, συνάρτηση ή διάνυσμα. Ακόμη, ένα τυχαίο πείραμα μπορεί να περιγραφεί με περισσότερους από ένα δειγματοχώρο. Για παράδειγμα, στη ρίψη ενός ζαριού μπορεί να έχουμε

S={1, 2, 3, 4, 5, 6} *ή* S'={*μονός, ζυγός*}.

Τέλος, είναι απαραίτητο να συζητηθεί ο αριθμός των δειγματοσημείων σε ένα δειγματοχώρο. Διακρίνονται κυρίως τρεις περιπτώσεις:

- **πεπερασμένος** δειγματοχώρος, όταν έχει πεπερασμένο πλήθος σημείων, όπως στα παραδείγματα 2.1, 2.2, 2.3, 2.4 και 2.7 στον δεύτερο δειγματοχώρο του.
- **άπειρος αριθμήσιμος** δειγματοχώρος, όταν τα σημεία του μπορούν να αντιστοιχηθούν ένα προς ένα με το σύνολο των φυσικών αριθμών, όπως στο παράδειγμα 2.5.
- **άπειρος** μη αριθμήσιμος δειγματοχώρος, όταν έχει άπειρο πλήθος σημείων αλλά όχι αριθμήσιμο, όπως στα παραδείγματα 2.6, και 2.7 στον πρώτο δειγματοχώρο του.

2.1.3 Σύνθετος Δειγματοχώρος

Ένα ***σύνθετο πείραμα Ε*** μπορεί να συνίσταται στην εκτέλεση δύο απλών πειραμάτων E_1 και E_2 με δειγματοχώρους S_1 και S_2 αντίστοιχα. Εκτελούμε τα πειράματα αυτά από μία φορά. Τότε το καρτεσιανό γινόμενο $S_1 \times S_2$ μπορεί να χρησιμοποιηθεί για να παραστήσει όλα τα πιθανά αποτελέσματα του σύνθετου πειράματος E. Έτσι, ο ***σύνθετος δειγματοχώρος***

$$\boldsymbol{S}=\boldsymbol{S}_1 \times \boldsymbol{S}_2$$

περιέχει σαν δειγματοσημεία τα *διατεταγμένα* σύνολα

$$\{(s_{1\kappa}, s_{2\lambda}) \mid s_{1\kappa} \in S_1,\ s_{2\lambda} \in S_2,\ \kappa=1,2,\ldots,n_1,\ \lambda=1,2,\ldots,n_2\}.$$

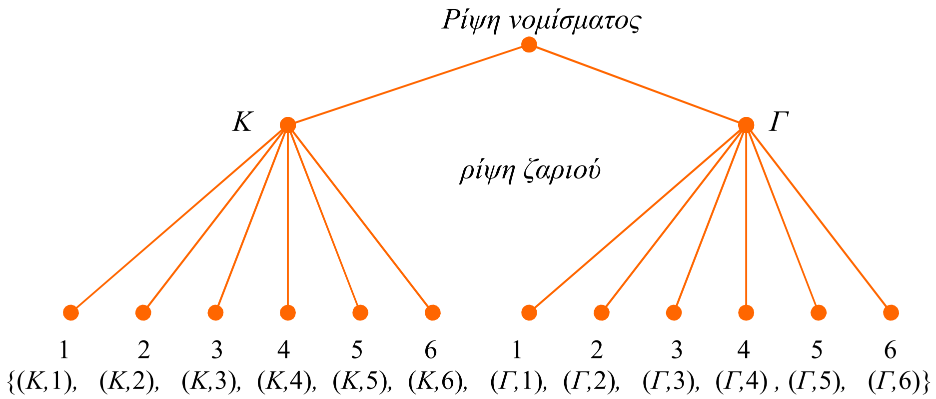
Στα παραδείγματα που ακολουθούν αναφέρονται τέτοιοι σύνθετοι δειγματοχώροι.

Παράδειγμα 2.8

Ένα νόμισμα και ένα ζάρι ρίπτονται μαζί, παρατηρείται η ένδειξη του ζαριού και νομίσματος, βλέπε Σχήμα 2.1. Εδώ μπορούμε να θεωρήσουμε ότι έχουμε ένα σύνθετο πείραμα E αποτελούμενο από την εκτέλεση δύο απλούστερων πειραμάτων E_1={*ρίψη νομίσματος*} και E_2={*ρίψη ζαριού*} με αντίστοιχους δειγματοχώρους

$$S_1=\{K, \Gamma\} \text{ και } S_2=\{1, 2, 3, 4, 5, 6\}.$$

Με βάση το παραπάνω σκεπτικό, το πείραμα E έχει σαν δειγματοχώρο το καρτεσιανό γινόμενο των S_1 και S_2,



Σχήμα 2.1. Σύνθετο πείραμα ρίψης νομίσματος και ζαριού

Για λόγους συντομίας τα διατεταγμένα υποσύνολα του σύνθετου δειγματοχώρου S γράφονται και χωρίς παρενθέσεις ή κόμματα, όπως

$$S=\{K1, K2, K3, K4, K5, K6, \Gamma1, \Gamma2, \Gamma3, \Gamma4, \Gamma5, \Gamma6\}. \ \blacksquare$$

Ακόμη, το σύνθετο πείραμα μπορεί να συνίσταται στην εκτέλεση δύο ή περισσότερων πειραμάτων με δειγματοχώρο το καρτεσιανό γινόμενο

$$S_1 \times S_2 \times \ldots \times S_n.$$

Στην πράξη συναντάμε πολλά σύνθετα πειράματα E και καταγράφουμε τον αντίστοιχο δειγματικό τους χώρο S, χωρίς να αναφερθούμε αναλυτικότερα στο καρτεσιανό γινόμενο των απλούστερων δειγματοχώρων, όπως π. χ. στα παραδείγματα που ακολουθούν.

Παράδειγμα 2.9

Ρίχνουμε τρία νομίσματα (όπως στο παράδειγμα 2.2) και παρατηρούμε τις ενδείξεις τους. Το πείραμα αυτό μπορούμε να το σκεφτούμε και σαν ένα σύνθετο πείραμα τύχης, E, το οποίο συνίσταται στην εκτέλεση των απλούστερων πειραμάτων E_1={*ρίψη νομίσματος* 1}, E_2={*ρίψη νομίσματος* 2} και E_3={*ρίψη νομίσματος* 3}, με δειγματοχώρους $S_1=\{K, \Gamma\}$, $S_2=\{K, \Gamma\}$, *και* $S_3=\{K,\Gamma\}$ αντίστοιχα. Δηλαδή, ο σύνθετος δειγματοχώρος S είναι το καρτεσιανό γινόμενο,

$$S=S_1 \times S_2 \times S_3=\{K, \Gamma\} \times \{K, \Gamma\} \times \{K, \Gamma\},$$
$$S=\{(K,K,K),(K,K,\Gamma),(K,\Gamma,K),(\Gamma,K,K), (K,\Gamma,\Gamma), (\Gamma,K,\Gamma), (\Gamma,\Gamma,K)\ \Gamma,\Gamma,\Gamma)\},$$

ο οποίος είναι ίδιος με τον δειγματοχώρο του παραδείγματος 2.2, όπως το καταγράψαμε εμπειρικά.

Παράδειγμα 2.10

Ρίχνουμε δύο ζάρια (*μια ζαριά*) και παρατηρούμε τις ενδείξεις τους. Όλα τα δυνατά ζευγάρια των ενδείξεων είναι 36 και είναι γνωστά,

$$S=\{(1,1),(1,2),(1,3),\ldots,(2,1),(2,2), (2,3),\ldots, (3,1), (3,2), (3,3),\ldots, (6,1), (6,2), (6,3), (6,4),(6,5), (6,6)\}.$$

Όμως, κάποιος μπορεί να θεωρήσει ότι E_1 και E_2 είναι τα απλά πειράματα για τη ρίψη του κάθε ζαριού με δειγματοχώρους $S_1=\{1, 2, 3, 4, 5, 6\}$ και $S_2=\{1, 2, 3, 4, 5, 6\}$ αντίστοιχα. Έτσι, το πείραμα της ζαριάς, E, συνίσταται στην εκτέλεση των E_1 και E_2 και έχει δειγματοχώρο S το γινόμενο,

$S=S_1 \times S_2=\{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$
$S=\{(1,1),(1,2),(1,3),\ldots,(2,1),(2,2), (2,3),\ldots, (3,1), (3,2), (3,3),\ldots, (6,1), (6,2), (6,3), (6,4),(6,5), (6,6)\}$.

Παρατήρηση
Στο εξής σε κάθε σύνθετο πείραμα τύχης, ο καθένας μπορεί να καταγράψει τον δειγματικό χώρο με όποιον τρόπο θεωρεί ευκολότερο. Όμως, στις περιπτώσεις όπου είναι πολλά τα δυνατά ενδεχόμενα του σύνθετου δειγματοχώρου S, συνιστάται το καρτεσιανό γινόμενο των απλούστερων δειγματοχώρων $S_1 \times S_2 \times S_3 \ldots$. Αυτή είναι και η βασικότερη ***αρχή απαρίθμησης***, όπως αναφέρεται στην Συνδυαστική.

2.1.4 Γεγονότα

Μία άλλη βασική έννοια συνδεόμενη με το τυχαίο πείραμα είναι το ***γεγονός***. Συχνά σ' ένα δειγματοχώρο δεν μας ενδιαφέρουν μόνο τα δειγματοσημεία αλλά κάποιες ομάδες δειγματοσημείων οι οποίες παριστάνουν κάποια γεγονότα.

Ορισμός. Στην ορολογία συνόλων, **γεγονός ή ενδεχόμενο** είναι ένα **υποσύνολο του δειγματοχώρου *S***. Τα γεγονότα συνηθίζεται να συμβολίζονται με κεφαλαία γράμματα, όπως A, B, Γ ή A_1, A_2, A_3, κ. τ. λ..

Σύμφωνα με τον ορισμό του γεγονότος, ο δειγματοχώρος S και το κενό σύνολο $\varnothing$ είναι και αυτά γεγονότα.

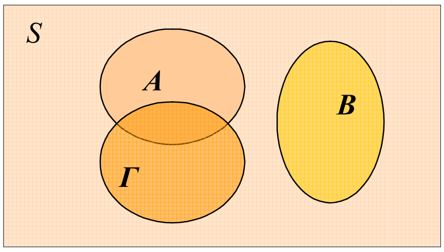
Παράδειγμα 2.11
Τα ακόλουθα είναι μερικά παραδείγματα από γεγονότα. Αναφερόμενοι στα πειράματα των παραπάνω παραδειγμάτων από 2.1 μέχρι 2.8, $\boldsymbol{A_\kappa}$ είναι ένα γεγονός που προκύπτει από το πείραμα $\boldsymbol{E_\kappa}$, το γεγονός πρώτα περιγράφεται και στη συνέχεια δίνεται το σύνολο των δειγματοσημείων του ή **ευνοϊκών δειγματοσημείων** του όπως αλλιώς αποκαλούνται:

A_1: {*Το τεμάχιο είναι ελαττωματικό*} ή $A_1=\{D\}$.
A_2: {*Τουλάχιστον δύο κεφαλές*} ή $A_2= \{KKK, KK\Gamma, K\Gamma K, \Gamma KK, \}$.
A_3: {*μία μπουλντόζα έχει βλάβη*} ή $A_3=\{ \overline{A_1}\,\overline{A_2}\,A_3,\ A_1\overline{A_2}\,\overline{A_1},\ \overline{A_1}\,A_2\,\overline{A_3}\}$.
A_4: {*Το πολύ 5 βίδες είναι ελαττωματικές*} ή $A_4=\{1, 2, 3, 4, 5\}$.
A_5: {*Όλοι οι λαμπτήρες είναι ελαττωματικοί*} ή $A_5=\{10\}$.
A_6: {*Η αντοχή είναι μικρότερη ενός θετικού αριθμού α*} ή $A_6=\{x|\ 0<x<\alpha\}$.
A_7: {*Η θερμοκρασία ξεπερνά τους* 28^0C} ή $A_7=\{x|x>28^0C\}$.
A_8: { *ένδειξη του νομίσματος κεφαλή και του ζαριού ζυγός αριθμός*}
ή $A_8 =\{(K,2),\ (K,4),\ (K,6)\}$. ■

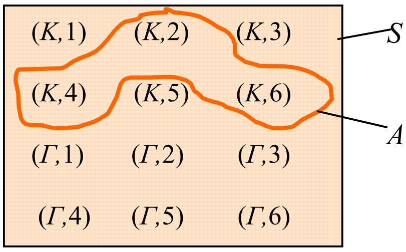
Όπως φαίνεται και στα παραπάνω παραδείγματα τα γεγονότα $A_κ$ ορίζονται είτε με την μέθοδο της ***περιγραφής*** είτε με την μέθοδο της ***αναγραφής***. Με την πρώτη περιγράφουμε το γεγονός ενώ με την δεύτερη αναφέρουμε όλα τα ευνοϊκά δειγματοσημεία του. Ένα γεγονός A οριζόμενο σ' ένα δειγματοχώρο S λέγεται ότι είναι **απλό γεγονός** εάν περιέχει μόνο ένα δειγματοσημείο. Διαφορετικά, το γεγονός A λέγεται **σύνθετο γεγονός** ή απλώς **γεγονός**.

Όταν ο δειγματοχώρος S είναι πεπερασμένος ή άπειρος αριθμήσιμος, κάθε υποσύνολό του μπορεί να θεωρηθεί σαν ένα γεγονός. Παρόλα αυτά, αν το S είναι άπειρο μη αριθμήσιμο, δεν είναι σίγουρο ότι κάθε υποσύνολο μπορεί να θεωρηθεί σαν ένα γεγονός. Ευτυχώς, τέτοια υποσύνολα συναντιόνται πολύ σπάνια για αυτό στην συνέχεια δεν θα ασχοληθούμε μ' αυτά.

Όπως οι δειγματόχωροι, έτσι και τα γεγονότα μπορούν να παρασταθούν με γραφικές παραστάσεις όπως για παράδειγμα τα **διαγράμματα Venn** (από το όνομα ενός Άγγλου μαθηματικού του 19ου αιώνα). Σ' αυτά τα διαγράμματα ο



(α) δειγματοχώρος S και γεγονότα A, B,και $Γ$



(β) Ρίψη νομίσματος και ζαριού, A={*Κεφαλή και ζυγός αριθμός*}

τα
.2.
ξύ

Σχήμα 2.2 Διαγράμματα Venn

των γεγονότων

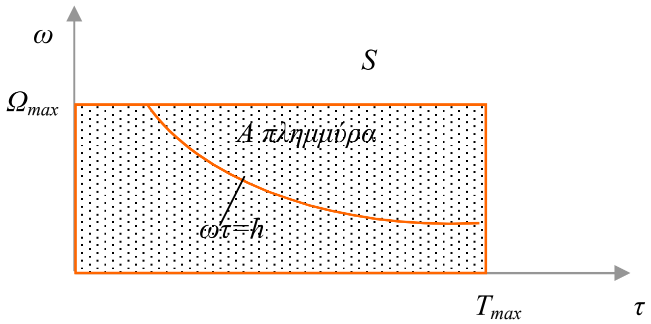
Ακόμη, μπορούν να χρησιμοποιηθούν διαγράμματα δέντρων ή άλλων γραφικών παραστάσεων σε δυδιάστατο ή τριδιάστατο σύστημα αξόνων για την παράσταση ενός δειγματοχώρου και των συσχετιζόμενων γεγονότων.

Παράδειγμα 2.12

Σε μία καταιγίδα ενδιαφέρον παρουσιάζει η έντασή της $ω$ και η διάρκειά της $τ$. Έτσι, ο δειγματοχώρος περιλαμβάνει άπειρα σημεία

$$S=\{(ω,τ)\,|\,ω<Ω_{max},\ τ<T_{max}\}.$$

Ένα απλό γεγονός A σε μία καταιγίδα είναι η πλημμύρα, δηλαδή, η ποσότητα της βροχόπτωσης ξεπερνά την μέγιστη ποσότητα αποστράγγισης στην περιοχή. Αν το γινόμενο $\omega \cdot \tau$ μετρά την ποσότητα βροχόπτωσης σε μία καταιγίδα ενώ η



Σχήμα 2.3. Το πείραμα της καταιγίδας και η πλημμύρα

δυνατότητα αποστράγγισης είναι h, το γεγονός A είναι το σύνολο

$$A=\{(\omega, \tau) \mid \omega \cdot \tau > h\},$$

βλέπε Σχήμα 2.3.

2.2 ΠΡΑΞΕΙΣ ΚΑΙ ΣΧΕΣΕΙΣ ΓΕΓΟΝΟΤΩΝ

2.2.1 Πράξεις Γεγονότων

Υπάρχουν πολλοί τρόποι με τους οποίους σχετίζονται τα γεγονότα ενός δειγματοχώρου. Συχνά, ενδιαφερόμαστε να περιγράψουμε νέα γεγονότα, συνδυάζοντας άλλα υπάρχοντα γεγονότα. Αφού τα γεγονότα είναι σύνολα, μπορούμε να χρησιμοποιήσουμε πράξεις συνόλων όπως συμπλήρωμα, ένωση, τομή, διαφορά, κ.τ.λ., για να σχηματίσουμε άλλα γεγονότα που μας ενδιαφέρουν.

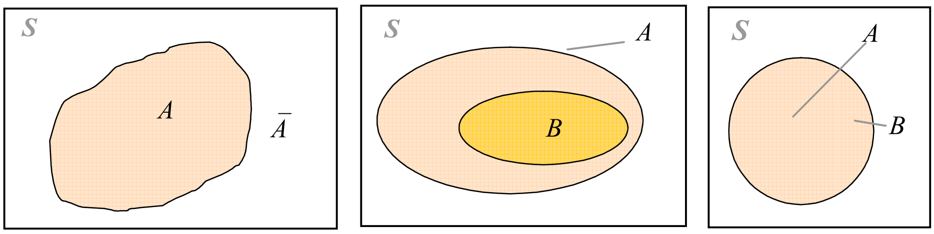
Συμπλήρωμα: Το συμπληρωματικό ενός γεγονότος A σε ένα δειγματικό χώρο S είναι το σύνολο των δειγματοσημείων του S τα οποία δεν ανήκουν στο A. Συμβολίζουμε το συμπληρωματικό του A με $\overline{A}$.

Περιεκτικότητα: Για δύο γεγονότα A και B τα οποία ανήκουν στον ίδιο δειγματικό χώρο S, λέμε ότι το A περιέχει το B, αν κάθε δειγματοσημείο του B ανήκει και στο A. Συμβολίζουμε $B \subseteq A$. Σε μία εκτέλεση του πειράματος αν πραγματοποιείται το B σίγουρα πραγματοποιείται και A.

Ισότητα: Για δύο γεγονότα A και B, τα οποία ανήκουν στον ίδιο δειγματικό χώρο S, ισχύει η ισότητα, ***A=B,*** αν $B \subseteq A$ και $A \subseteq B$. Δηλαδή, κάθε δειγματοσημείο του A είναι δειγματοσημείο του B και το αντίστροφο. Τα γεγονότα A και B λέγονται και **ισοδύναμα**, διότι σε μία εκτέλεση του πειράματος είτε πραγματοποιούνται και τα δύο μαζί ή δεν πραγματοποιείται κανένα.

Με βάση τον ορισμό της ισότητας ισχύει

$$A \cap S = A, \quad A \cap A = A, \quad A \cap \{\} = \{\}, \quad A \cup S = S,$$
$$A \cup A = A, \quad A \cup \{\} = A, \quad A \cup \overline{A} = S.$$



α) συμπλήρωμα, $\overline{A}$ β) περιεκτικότητα $B \subseteq A$ γ) ισότητα, $A=B$

Σχήμα 2.4. Σχεδιαγράμματα Venn

Ένωση: Αν A και B είναι δύο γεγονότα του ίδιου δειγματικού χώρου S, ορίζουμε σαν **ένωση** των $\boldsymbol{A}$ και $\boldsymbol{B}$ το γεγονός $\boldsymbol{\Gamma}$, το οποίο περιέχει δειγματοσημεία που ανήκουν στο A ή στο B ή και στα δυο, και γράφουμε $\boldsymbol{\Gamma = A \cup B}$.

Σε μία εκτέλεση του πειράματος τύχης, το γεγονός Γ συμβαίνει εάν και μόνον εάν συμβεί τουλάχιστον ένα από τα δύο γεγονότα A ή B, βλέπε Σχήμα 2.5.

Παραδείγματος χάριν, σε ένα εργοτάξιο σχετικά με τα αποθέματα τσιμέντου και άμμου, συμβολίζουμε με A και B τα γεγονότα ότι κάποια μέρα παρατηρείται έλλειμμα τσιμέντου και άμμου αντίστοιχα. Το γεγονός ότι κάποια μέρα παρατηρείται έλλειμμα είτε σε τσιμέντο ή σε άμμο ή μπορεί και στα δύο εκφράζεται με την ένωση $A \cup B$.

Με βάση τον παραπάνω ορισμό **ισότητας** και **συμπλήρωμα** ισχύει

$$A \cup S = S, \quad A \cup A = A, \quad A \cup \{\} = A, \quad A \cup \overline{A} = S, \quad \overline{\overline{A}} = A,$$

όπου στην τελευταία ισότητα η ερμηνεία είναι ότι το συμπλήρωμα του συμπληρώματος του A είναι το ίδιο το A.

Η ένωση μπορεί να επεκταθεί για τρία ή και περισσότερα γεγονότα,

$$A \cup B \cup \Gamma \cup \ldots .$$

Τομή: Αν A και B είναι δύο γεγονότα του ίδιου δειγματικού χώρου S, ορίζουμε σαν **τομή** των $\boldsymbol{A}$ και $\boldsymbol{B}$ το γεγονός $\boldsymbol{\Gamma}$ το οποίο περιέχει δειγματο σημεία που ανήκουν και στο A και στο B, και γράφουμε $\boldsymbol{\Gamma = A \cap B}$.

Σε μία εκτέλεση του πειράματος τύχης, το γεγονός Γ συμβαίνει εάν και μόνον εάν συμβούν ταυτόχρονα τα δύο γεγονότα A και B, βλέπε Σχήμα 2.5.

Παραδείγματος χάριν, στο παραπάνω εργοτάξιο σχετικά με τα αποθέματα τσιμέντου και άμμου, Το γεγονός ότι κάποια μέρα παρατηρείται έλλειμμα και στα δύο, και σε τσιμέντο και σε άμμο, εκφράζεται με την τομή $A \cap B$. Σαν άλλο παράδειγμα, μπορούμε να θεωρήσουμε την βλάβη στην κεντρική γραμμή ηλεκτροδότησης. Αν A είναι το γεγονός βλάβης σε κάποιο σημείο από την μονάδα παραγωγής μέχρι το 10^o χιλιόμετρο και B το γεγονός βλάβης από το 8^o μέχρι το 15^o χιλιόμετρο, η τομή $A \cap B$ παριστάνει την βλάβη από το 8^o μέχρι το 10^o χιλιόμετρο.

Με βάση τους ορισμούς των παραπάνω πράξεων ισχύει

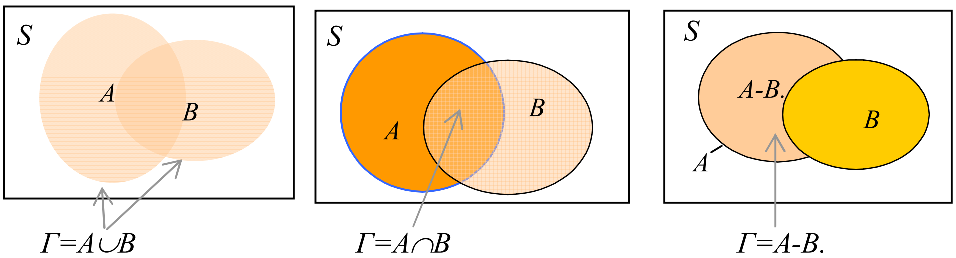
$$A \cap S = A, \quad A \cap A = A, \quad A \cap \{\} = \{\}, \quad A \cap \overline{A} = \{\}, \quad S \cap \overline{A} = \overline{A}.$$

Τέλος, Η **τομή** μπορεί να επεκταθεί για τρία ή και περισσότερα γεγονότα,

$$A \cap B \cap \Gamma \cap \ldots .$$

Διαφορά: Αν A και B είναι δύο γεγονότα του ίδιου δειγματικού χώρου S, ορίζουμε σαν διαφορά $\boldsymbol{A\text{-}B}$ το γεγονός $\boldsymbol{\Gamma}$ το οποίο περιέχει δειγματοσημεία που ανήκουν στο A και όχι στο B, και γράφουμε $\boldsymbol{\Gamma = A\text{-}B}$.

Σε μία εκτέλεση του πειράματος τύχης, το γεγονός Γ συμβαίνει εάν και μόνον εάν συμβεί το A και όχι το B, βλέπε Σχήμα 2.5.



Σχήμα 2.5. Πράξεις γεγονότων A, B.

Συνήθως, η διαφορά *Α-Β* καλείται *μόνον Α,* διότι περιέχει δειγματοσημεία που ανήκουν μόνον στο σύνολο *Α*. Επίσης, ένας ισοδύναμος συμβολισμός της διαφοράς είναι η τομή $A \cap \overline{B}$.

2.2.2 Ασυμβίβαστα γεγονότα ή αμοιβαίως αποκλειόμενα

Τέλος, από τους παραπάνω κανόνες ενδιαφέρων παρουσιάζει η περίπτωση όπου η τομή δύο γεγονότων δεν περιέχει κανένα δειγματοσημείο. Παραδείγματος χάριν αν *Α, Β* συμβολίζουν τα γεγονότα ότι κάποιος θα ταξιδέψει από Θεσσαλονίκη για Αθήνα με αεροπλάνο ή αυτοκίνητο, αντίστοιχα, η τομή των δύο γεγονότων είναι το κενό σύνολο, δηλαδή, το αδύνατο γεγονός. Ακόμη άλλο παράδειγμα, αν δύο γεγονότα παριστάνουν ότι ένα αυτοκίνητο σε μία διασταύρωση θα στρίψει αριστερά ή δεξιά, αντίστοιχα, η τομή τους δεν μπορεί να πραγματοποιηθεί.

Ορισμός: Δύο γεγονότα *Α* και *Β* οριζόμενα σε ένα δειγματικό χώρο είναι **αμοιβαία αποκλειόμενα ή ασυμβίβαστα** εάν η πραγματοποίηση του ενός αποκλείει την πραγματοποίηση του άλλου. Αυτό σημαίνει ότι τα δύο γεγονότα δεν έχουν κανένα κοινό δειγματοσημείο,

$$A \cap B = \varnothing,$$

άρα δεν μπορούν να πραγματοποιηθούν μαζί σε μία εκτέλεση του πειράματος.

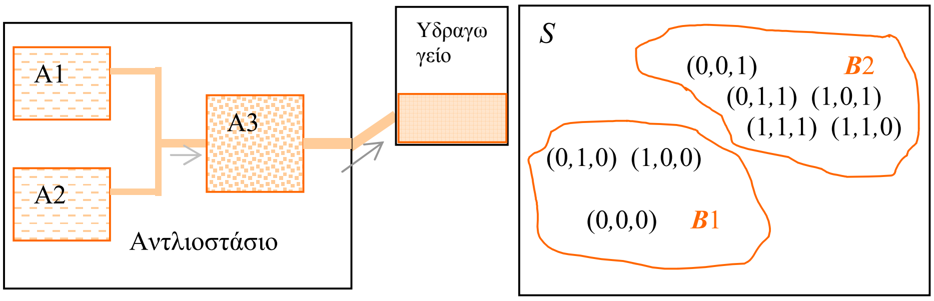
Σε ένα σχεδιάγραμμα **Venn** εύκολα διακρίνει κανείς τα ασυμβίβαστα γεγονότα, δεν έχουν κανένα κοινό δειγματοσημείο και είναι δύο σύνολα ξένα μεταξύ τους, βλέπε Σχήμα 2.6 (β).

Παράδειγμα 2.13
Ένα μεγάλο αντλιοστάσιο έχει δύο ηλεκτρικές αντλίες Α1 και Α2 εν παραλλήλω και μία τρίτη ηλεκτρική αντλία όπως φαίνεται στο Σχήμα 2.6. Εάν (N_1, N_2, N_3) είναι ένα διάνυσμα στον τριδιάστατο χώρο, όπου N_κ=1 εάν η A_κ αντλία αποτυγχάνει, και N_κ=0 σε άλλη περίπτωση, ο δειγματικός χώρος αποτελείται από 8 δειγματοσημεία

$$S=\{(0,0,0),\ (0,0,1),\ (0,1,0),(0,0,1),\ (1,1,0),\ (1,0,1),\ (0,1,1),\ (1,1,1)\}.$$

Το σύστημα λειτουργεί αν αντλεί η Α3 και μία τουλάχιστον από τις Α1 και Α2, συμβολίζουμε το γεγονός αυτό με *Β*1. Ακόμη, το σύστημα αποτυγχάνει αν δεν λειτουργεί η Α3 ή δεν λειτουργούν και οι δύο παράλληλες αντλίες, συμβολίζουμε το γεγονός αυτό με *Β*2,

$B1=\{((0,0,0), (0,1,0), (1,0,0)\}$
$B2=\{(0,0,1),(1,1,0), (1,0,1), (0,1,1), (1,1,1)\}$.



α) Σύστημα άντλησης νερού β) σχεδιάγραμμα Venn, $B1\cap B2=\varnothing$

Σχήμα 2.6. Γραφικές παραστάσεις παραδείγματος 2.13.

Τα γεγονότα $B1$ και $B2$ είναι αμοιβαίως αποκλειόμενα, διότι $B1\cap B2=\varnothing$, ή όπως αλλιώς λέγεται ασυμβίβαστα γιατί δεν μπορούν να πραγματοποιηθούν και τα δύο μαζί.

2.2.3 Κανόνες Πράξεων Γεγονότων

Αφού τα γεγονότα είναι σύνολα, οι κανόνες ή ιδιότητες των πράξεων στα σύνολα διέπουν και τις πράξεις των γεγονότων. Οι σπουδαιότερες ιδιότητες στις πράξεις γεγονότων είναι οι ακόλουθες,

Αντιμεταθετική: Η ένωση και η τομή των γεγονότων είναι αντιμεταθετικές πράξεις,

$$A\cup B=B\cup A, \quad A\cap B=B\cap A.$$

Προσεταιριστική: Η ένωση και η τομή των γεγονότων είναι προσεταιριστικές πράξεις,

$$(A\cup B)\cup \Gamma=A\cup(B\cup \Gamma), \quad (A\cap B)\cap \Gamma=A\cap(B\cap \Gamma.$$

Επιμεριστική: Η ένωση και η τομή των γεγονότων είναι επιμεριστικές πράξεις,

$$(A\cup B)\cap \Gamma=(A\cap \Gamma)\cup(B\cap \Gamma), \quad (A\cap B)\cup \Gamma=(A\cup \Gamma)\cap(B\cup \Gamma).$$

Η προτεραιότητα για τις πράξεις της ένωσης και τομής είναι όμοια με την ιεράρχηση της πρόσθεσης και πολλαπλασιασμού σε μία αλγεβρική έκφραση. Έτσι, προτεραιότητα έχουν οι πράξεις γεγονότων μέσα στις παρενθέσεις και ακολουθούν οι πράξεις της τομής και ένωσης με την σειρά που αναφέρονται.

Παρατήρηση

Πρέπει να τονισθεί ότι οι πράξεις αριθμών στην άλγεβρα δεν έχουν καμία σχέση με την ένωση, τομή ή διαφορά γεγονότων, για παράδειγμα,

- $A \cap A = A$, ενώ στο γινόμενο δύο αριθμών $\alpha \cdot \alpha = \alpha^2 \neq \alpha$,
- $(A \cup B) \cap (B \cup \Gamma) = A \cap B \cup B \cap B \cup A \cap \Gamma \cup B \cap \Gamma = A \cap \Gamma \cup B$,
  ενώ $(\alpha+\beta)(\beta+\gamma) = \alpha\beta + \alpha\gamma + \beta^2 + \beta\gamma \neq \alpha\gamma + \beta$,
- $(A-B) \cup B = A \cup B$, ενώ $(\alpha-\beta)+\beta = \alpha$. ■

Νόμος De Morgan: Εδώ έχουμε πράξεις τομής και ένωσης μεταξύ γεγονότων και των συμπληρωματικών τους,

$$\overline{A \cup B} = \overline{A} \cap \overline{B}, \quad \acute{\eta} \quad \overline{A \cap B} = \overline{A} \cup \overline{B}\,. \tag{2.1}$$

Η απόδειξη είναι προφανής αν χρησιμοποιήσει κανείς τον ορισμό της ισότητας δύο γεγονότων. Ο νόμος De Morgan μπορεί να επεκταθεί για περισσότερα των δύο γεγονότα,

$$\begin{gathered}\overline{A_1 \cup A_2 \cup A_3 \cup \ldots . \cup A_n} = \overline{A_1} \cap \overline{A_2} \cap \overline{A_3} \cap \ldots . \cap \overline{A_n} \\ \overline{A_1 \cap A_2 \cap A_3 \cap \ldots . \cap A_n} = \overline{A_1} \cup \overline{A_2} \cup \overline{A_3} \cup \ldots . \cup \overline{A_n}\end{gathered},$$

η δε απόδειξη θα μπορούσε να γίνει με την μέθοδο της μαθηματικής επαγωγής. Τέλος, αν λάβουμε υπ' όψιν ότι $\overline{\overline{A}} = A$, είναι προφανής και η ακόλουθη σχέση,

$$\overline{\overline{\overline{A_1} \cap \overline{A_2} \cap \overline{A_3} \cap \ldots . \cap \overline{A_n}}} = A_1 \cup A_2 \cup A_3 \cup \ldots . \cup A_n\,.$$

Από τις παραπάνω πράξεις και κανόνες εύκολα καταλήγει κανείς στην **δυική αρχή**, η οποία δηλώνει ότι το συμπλήρωμα ενώσεων και τομών ισούται με τις τομές και ενώσεις των συμπληρωματικών τους, παράδειγμα

$$\overline{A \cup (B \cap \Gamma)} = \overline{A} \cap \overline{(B \cap \Gamma)} = \overline{A} \cap (\overline{B} \cup \overline{\Gamma}).$$

$$\overline{A \cap B \cap \overline{\Gamma} \cup A \cap \Gamma} = \overline{(A \cap B \cap \overline{\Gamma}) \cup (A \cap \Gamma)} = \overline{A \cap B \cap \overline{\Gamma}} \cap \overline{A \cap \Gamma}$$
$$= \overline{A} \cup \overline{B} \cup \Gamma \cap \overline{A} \cup \overline{\Gamma}.$$

Παράδειγμα 2.14

Στο παράδειγμα 2.14 το νερό φθάνει στο υδραγωγείο από το αντλιοστάσιο (βλέπε Σχήμα 2.6, (α)). Συμβολίζουμε με A_1, A_2, A_3, τα γεγονότα ότι οι αντίστοιχες αντλίες διακόπτουν την άντληση λόγω βλάβης.

Το γεγονός ότι η παροχή νερού στο υδραγωγείο διακόπτεται, Δ, εκφράζεται με τις εξής πράξεις γεγονότων,

$$\Delta = A_3 \cup (A_1 \cap A_2),$$

δηλαδή, το γεγονός Δ συμβαίνει αν η A3 διακόπτει την άντληση ή και οι δύο αντλίες A1, A2 δεν αντλούν νερό.

Το γεγονός ότι το αντλιοστάσιο παρέχει νερό στο υδραγωγείο, $\overline{\Delta}$, συμβαίνει αν δεν διακόπτει την άντληση η A3 και τουλάχιστον μία από τις A2 και A1. Δηλαδή, η παροχή νερού στο υδραγωγείο εκφράζεται με την ακόλουθη άλγεβρα γεγονότων,

$$\overline{\Delta} = \overline{A_3} \cap (\overline{A_1} \cup \overline{A_2}).$$

Όμως, μπορούμε να καταλήξουμε στην ίδια άλγεβρα γεγονότων για το $\overline{\Delta}$, αν πάρουμε το συμπληρωματικό της άλγεβρας γεγονότων του Δ, σύμφωνα με τον νόμο De Morgan,

$$\overline{\Delta} = \overline{A_3 \cup (A_1 \cap A_2)} = \overline{A_3} \cap (\overline{A_1} \cup \overline{A_2}).$$

Παράδειγμα 2.15

Για την κατασκευή ενός μεγάλου εργοστασίου διακρίνουμε τρεις φάσεις: πρώτον θεμελίωση, δεύτερον δόμηση, τρίτον ηλεκτρολογική εγκατάσταση. Συμβολίζουμε με Θ , Δ *και* H τα γεγονότα καθυστέρησης θεμελίωσης, δόμησης και ηλεκτρικής εγκατάστασης αντίστοιχα. Να εκφραστούν με συμβολισμό συνόλων τα γεγονότα ότι το εργοστάσιο ολοκληρώνεται

(α) χωρίς καθυστέρηση,

(β) με καθυστέρηση μόνον μίας φάσης,

(γ) το πολύ με καθυστέρηση δύο φάσεων.

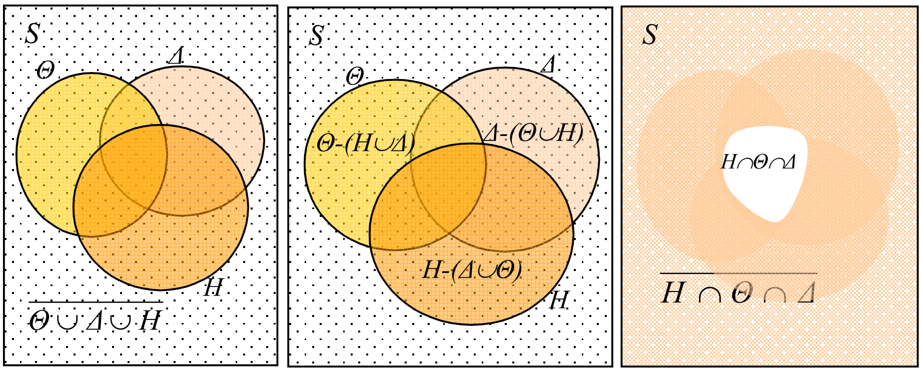
Απάντηση:

(α) Η ένωση των γεγονότων Θ, Δ και H συμβολίζει το γεγονός της καθυστέρησης του εργοστασίου. Το συμπλήρωμα της ένωσης, προφανώς, συμβολίζει την ολοκλήρωση του έργου άνευ καθυστέρησης,

$$\{\text{χωρίς καθυστέρηση}\} = \overline{(\Theta \cup \Delta \cup H)}$$

(β) Εδώ, πρέπει να πάρουμε την ένωση των γεγονότων να καθυστερήσει μόνον η θεμελίωση, μόνον η δόμηση, μόνον η ηλεκτρική εγκατάσταση, δηλαδή,

$\{καθυστέρηση\ μόνον\ μίας\ φάσης\} =$
$(\Theta - (\Delta \cup H)) \cup (\Delta - (\Theta \cup H)) \cup (H - (\Theta \cup \Delta))$,



(α) το έργο καθυστερεί (β) μόνο μία φάση καθυστερεί (γ) μέχρι δύο φάσεις

***Σχήμα* 2.7**. Διαγράμματα Venn, Παράδειγμα 2.16**.**

ή εναλλακτικά

$\{καθυστέρηση\ μόνον\ μίας\ φάσης\} =$
$(H \cap \overline{\Theta} \cap \overline{\Delta}) \cup (\overline{H} \cap \Theta \cap \overline{\Delta}) \cup (\overline{H} \cap \overline{\Theta} \cap \Delta)$,

ή

$\{καθυστέρηση\ μόνον\ μίας\ φάσης\} =$
$(H - \Theta - \Delta) \cup (\Theta - H - \Delta) \cup (\Delta - H - \Theta)$.

(γ) Το γεγονός αυτό σημαίνει όχι και τα τρία γεγονότα μαζί,

$$\overline{H \cap \Theta \cap \Delta}.$$

1.3 ΧΩΡΟΣ ΓΕΓΟΝΟΤΩΝ-ΔΥΝΑΜΟΣΥΝΟΛΟ

Ορισμός: το σύνολο ή η συλλογή από όλα τα δυνατά γεγονότα του δειγματοχώρου S ορίζεται σαν **χώρος γεγονότων** ή **δυναμοσύνολο** και συμβολίζεται με $\boldsymbol{S}^*$. Ο χώρος αυτός έχει τις εξής βασικές ιδιότητες:

- $S \in \boldsymbol{S}^*$
- $Αν\ A \in \boldsymbol{S}^* \Rightarrow \overline{A} \in \boldsymbol{S}^*$
- $Αν\ A_1 \in \boldsymbol{S}^*\ και\ A_2 \in \boldsymbol{S}^* \Rightarrow A_1 \cap A_2 \in \boldsymbol{S}^*,\ \ A_1 \cup A_2 \in \boldsymbol{S}^*,$

A_1-$A_2 \in \boldsymbol{S}^*$.

Αν ένας δειγματοχώρος S περιέχει $\boldsymbol{n}$ δειγματοσημεία, τότε όλα τα στοιχεία του $\boldsymbol{S}^*$,δηλαδή, όλα τα δυνατά γεγονότα του δειγματοχώρου είναι 2^n,

$$\boldsymbol{N(S^*)}=2^n$$

Παράδειγμα 2.16

Στη ρίψη ενός νομίσματος, ο δειγματοχώρος του πειράματος είναι $S=\{K,\Gamma\}$. Όλα τα γεγονότα που μπορούν να ορισθούν σε αυτό το πείραμα είναι τα υποσύνολα του S, όπως $\{K\}$, $\{\Gamma\}$, $\{K,\Gamma\}$ και $\varnothing$. Το δυναμοσύνολο $\boldsymbol{S}^*$ περιέχει 2^2=4 διαφορετικά υποσύνολα

$$\boldsymbol{S}^*=\{\ \{\ \},\ \{K\},\ \{\Gamma\},\ \{K,\Gamma\}\}.\ \blacksquare$$

Ενώ το S περιέχει όλα τα δυνατά αποτελέσματα του πειράματος τύχης, ο χώρος γεγονότων ή δυναμοσύνολο $\boldsymbol{S}^*$ περιέχει το S και όλα τα δυνατά υποσύνολα ή γεγονότα που ορίζονται συνδυάζοντας τα αποτελέσματα του πειράματος.

2.4. Η ΕΝΝΟΙΑ ΤΗΣ ΠΙΘΑΝΟΤΗΤΑΣ

Η ***πιθανότητα*** ενός γεγονότος είναι η έκφραση της βεβαιότητας πραγματοποίησής του. Επί πλέον, μία πιθανότητα είναι ένας αριθμός ο οποίος ξεκινά από το μηδέν, για ένα γεγονός το οποίο δεν μπορεί να συμβεί, και φθάνει στο ένα ,για ένα γεγονός που είναι σίγουρο ότι θα συμβεί. Αλλά, πώς προσδιορίζουμε τον αριθμό πιθανότητας στα άλλα γεγονότα; Ο αριθμός αυτός εξαρτάται από την δική μας ερμηνεία της πιθανότητας. Παρόλα που υπάρχουν διάφορες έννοιες και ορισμοί για την πιθανότητα στις διάφορες επιστήμες, όλοι συμφωνούν ότι ο αριθμός πιθανότητας ή η πιθανότητα ενός γεγονότος μπορεί να προσδιορισθεί με την ***θεωρία κλασσικής πιθανότητας***, ***σχετικής συχνότητας***, και ***υποκειμεν ικής πιθανότητας***..

Η Θεωρία κλασσικής πιθανότητας εφαρμόζεται μόνο σε πειράματα τύχης όπου ο δειγματοχώρος S αποτελείται από εξίσου πιθανά δειγματοσημεία. Σε τέτοια πειράματα τύχης, βασική αρχή της κλασσικής θεωρίας είναι ότι κανένα αποτέλεσμα του πειράματος δεν είναι πιθανότερο του άλλου. Με βάση αυτή την αρχή, όλα τα δυνατά αποτελέσματα του S έχουν την ίδια πιθανότητα πραγματοποίησης. Για παράδειγμα, στην ρίψη ενός ζαριού, καμία ένδειξη δεν είναι πιθανότερη της άλλης, και κατά συνέπεια η πιθανότητα εμφάνισης οιοδήποτε αριθμού (1 έως 6) είναι 1/6.

Γενικά, σε ένα πείραμα τύχης με n ισοπίθανα δειγματοσημεία,

$$S=\{s_1, s_2, \ldots, s_n\},$$

η πραγματοποίηση του s_i είναι το ίδιο πιθανή με την πραγματοποίηση οιοδήποτε άλλου ενδεχομένου s_j. Στην περίπτωση αυτή η κλασσική θεωρία ορίζει σαν πιθανότητα πραγματοποίησης του ενδεχομένου s_i, $P(s_i)$, το $1/n$. Κατ' επέκταση, αν ένα γεγονός A_r του δειγματικού χώρου S περιέχει κ δειγματοσημεία,

$$A_r = \{ s_{r1}, s_{r2}, \ldots, s_{rk} \},$$

ως πιθανότητα πραγματοποίησης του ορίζεται ο λόγος κ/n.

Η παραπάνω ανάλυση οδηγεί σε ένα απλό και κλασσικό ορισμό της πιθανότητας.

Κλασσικός ορισμός της πιθανότητας: Εάν ένα πείραμα τύχης έχει ***N***(***S***) ***ισοπίθανα*** αποτελέσματα, η πιθανότητα πραγματοποίησης ενός γεγονότος ***A*** του ***S***, το οποίο περιέχει ***N***(***A***) δειγματοσημεία, συμβολίζεται με ***P***(***A***) και είναι ο λόγος του ***N***(***A***) προς ***N***(***S***),

$$\boldsymbol{P(A)=\frac{N(A)}{N(S)}}. \tag{2.2}$$

Παράδειγμα 2.17

Στην ρίψη ενός ζαριού, αν A είναι το γεγονός ότι έχουμε ένδειξη ζυγού αριθμού, A={2,4,6}, και B είναι το γεγονός ότι η ένδειξη είναι μικρότερη του 5, B={1,2,3,4}, οι αντίστοιχες πιθανότητες πραγματοποίησης των A και B κατά την ρίψη του ζαριού είναι

$$P(A)=\frac{N(A)}{N(S)}=\frac{3}{6}=\frac{1}{2}\ , \quad P(B)=\frac{N(B)}{N(S)}=\frac{4}{6}=\frac{2}{3}.$$

Παράδειγμα 2.18

Στην ρίψη ενός νομίσματος τρεις φορές, ο δειγματοχώρος του πειράματος τύχης S έχει 8 ισοπίθανα δειγματοσημεία,

$$S=\{ΚΚΚ, ΚΚΓ, ΚΓΚ, ΓΚΚ, ΚΓΓ, ΓΚΓ, ΓΓΚ, ΓΓΓ\}.$$

Αν A είναι το γεγονός ότι έχουμε δύο Κεφαλές και ένα Γράμμα,

$$A=\{ ΚΚΓ, ΚΓΚ, ΓΚΚ\},$$

τότε στην ρίψη ενός νομίσματος 3 φορές η πιθανότητα πραγματοποίησης του γεγονότος A είναι

$$P(A)=\frac{N(A)}{N(S)}=\frac{3}{8}.$$

Παρατηρήσεις

(Α) Είναι πολύ σημαντικό να τονίσουμε ότι η κλασσική θεωρία βασίζεται στην υπόθεση ότι όλα τα δυνατά αποτελέσματα του πειράματος τύχης είναι απόλυτα ισοπίθανα και εφαρμόζεται όταν ικανοποιείται πλήρως. Έτσι, δεν δημιουργείται καμία δυσκολία όταν αναφερόμαστε σε ένα μη δολιοευμένο ζάρι, νόμισμα ή αξιόπιστο τροχό ρουλέτας, όπου όλα τα δυνατά αποτελέσματα είναι απόλυτα ισοπίθανα. Στόν πραγματικό κόσμο όμως τα δυνατά αποτελέσματα των πειραμάτων αυτών μπορεί να μην είναι ακριβώς ισοπίθανα. Για αυτό, η υπόθεση του απόλυτα ισοπίθανα θα προκαλέσει κάποια απόκλιση από την πραγματική πιθανότητα. Ακόμη, σε ένα δολιοευμένο ζάρι , νόμισμα ή μη τίμιο τροχό ρουλέτας η κλασσική θεωρία οδηγεί σε λάθος εκτίμηση πιθανότητα. Στις περιπτώσεις αυτές συνιστάται η θεωρία της σχετικής συχνότητας, όπως αναφέρεται στην επόμενη παράγραφο.

(Β) Στα παραδείγματα που αναφέρθηκαν μέχρι τώρα εύκολα απαριθμήθηκαν όλα τα δειγματοσημεία του δειγματικού χώρου S και όλοι οι δυνατοί τρόποι (ευνοϊκά δειγματοσημεία) με τους οποίους συμβαίνει το γεγονός A. Υπάρχουν όμως περισσότερο σύνθετες περιπτώσεις όπου η απαρίθμηση όλων των δειγματοσημείων του S και A απαιτεί πιο συστηματικές μεθόδους, όπως αναφέρονται στην παράγραφο της Συνδυαστικής.

■

Στην περίπτωση όπου ο δειγματοχώρος S έχει ισοπίθανα αλλά άπειρα δειγματοσημεία, ο προσδιορισμός της κλασσικής πιθανότητας ενός γεγονότος A του S ενδέχεται να είναι υπολογιστικά αδύνατος. Διότι, ενδέχεται το σύνολο A να περιέχει σε μια τέτοια περίπτωση άπειρα δειγματοσημεία και ο λόγος των δύο πληθάριθμων $N(A)$ προς $N(S)$ είναι απροσδιόριστος. Παρόλα αυτά, η κλασσική πιθανότητα μπορεί να προσδιορισθεί, αν περιγράψουμε τον δειγματοχώρο S και το γεγονός A κατά κάποιο τρόπο με πεπερασμένο αριθμό ισοπίθανων ομαδοποιημένων δειγματοσημείων.

Παράδειγμα 2.19

Μία ράβδος μήκους $α$ σπάει τυχαία σε οποιοδήποτε σημείο της. Ποια η πιθανότητα ότι το πρώτο κομμάτι της θα είναι μικρότερο από το δεύτερο;

Το πείραμα τύχης συνίσταται στο ότι η ράβδος σπάει ισοπίθανα σε κάποιο σημείο της, ο δε δειγματοχώρος S περιέχει άπειρα ισοπίθανα δειγματοσημεία, $N(S)=\infty$. Επίσης, το γεγονός A ότι το πρώτο κομμάτι είναι μικρότερο από το δεύτερο είναι ισοδύναμο με το γεγονός ότι η ράβδος σπάει στο πρώτο ήμισυ. Έτσι, τα δειγματοσημεία που περιέχει το γεγονός A

αντιστοιχούν σε όλα τα σημεία του πρώτου ήμισυ τα οποία είναι άπειρα το πλήθος, δηλαδή $N(A)=\infty$. Σύμφωνα με την κλασσική θεωρία η πιθανότητα πραγματοποίησης του A, $P(A)$, είναι ο λόγος των δύο πληθάριθμων και είναι προφανώς απροσδιόριστος.

Όμως, περιγράφοντας διαφορετικότερα τον δειγματικό χώρο, $S=\{\alpha_1, \alpha_2\}$, όπου α_1 και α_2 παριστάνουν τα ισοπίθανα ενδεχόμενα η ράβδος να σπάσει στο πρώτο ή δεύτερο ήμισυ αντίστοιχα, το γεγονός A περιέχει ένα δειγματοσημείο, $A=\{\alpha_1\}$. Έτσι, με την κλασσική θεωρία η πιθανότητα $P(A)$ είναι 1/2. ■

2.4.2 Θεωρία Σχετικής Συχνότητας

Η θεωρία αυτή υποστηρίζει ότι η μόνη έγκυρη διαδικασία προσδιορισμού πιθανότητας γεγονότος είναι η πολλαπλή επανάληψη του πειράματος E. Εάν το πείραμα διεξήχθη n φορές κάτω από τις ίδιες συνθήκες και ένα γεγονός A που σχετίζεται με το E συνέβη x_A φορές, $x_A \leq n$, τότε μία εκτίμηση της πιθανότητας πραγματοποίησης του γεγονότος A σε μία εκτέλεση του πειράματος είναι ο λόγος x_A / n. Επίσης, αν το n αυξηθεί χωρίς περιορισμό τότε ο λόγος x_A / n τείνει στην πιθανότητα του γεγονότος A, δηλαδή

$$P(A) = \lim_{n\to\infty} \frac{x_A}{n} \tag{2.3}$$

Προκειμένου να δώσουμε ένα καλύτερο μαθηματικό υπόβαθρο στην σχετική συχνότητα θεωρούμε δύο γεγονότα A και B σχετιζόμενα με το επαναλαμβανόμενο πείραμα E. Αν x_A και x_B είναι οι αριθμοί εμφάνισης των γεγονότων A και B στις n επαναλήψεις του πειράματος, αντίστοιχα, δίνουμε τον ακόλουθο ορισμό της σχετικής συχνότητας.

Ορισμός. Σχετική συχνότητα του γεγονότος A, f_A, σε n επαναλήψεις του πειράματος τύχης E καλείται ο λόγος x_A/n,

$$f_A = \frac{x_A}{n}.$$

Η σχετική συχνότητα f_A έχει τις ακόλουθες προφανείς βασικές ιδιότητες:

- $0 \leq f_A \leq 1$.
- $f_A=1$ εάν και μόνον εάν το A συμβαίνει σε κάθε επανάληψη του E.

- f_A=0 εάν και μόνον εάν το A δεν συμβαίνει σε καμία επανάληψη του E.
- Εάν τα γεγονότα A και B είναι ασυμβίβαστα, τότε η σχετική συχνότητα του γεγονότος AUB είναι
$$f_{AUB}= f_A + f_B.$$
- Το f_A είναι συνάρτηση του n και «συγκλίνει» στην πιθανότητα $P(A)$ καθώς $n\to\infty$.
$$P(A) = \lim_{n\to\infty} f_A .$$

Στην ιδιότητα 5) η σύγκλιση της σχετικής συχνότητας f_A στην πιθανότητα $P(A)$, καθώς $n\to\infty$, δεν ταυτίζεται απόλυτα με την έννοια της σύγκλισης στα μαθηματικά. Η σύγκλιση εδώ με μία πιθανολογική έννοια εκφράζει ότι η σχετική συχνότητα, καθώς αυξάνεται ο αριθμός των επαναλήψεων του πειράματος E, τείνει να «σταθεροποιηθεί» κοντά σε κάποια ορισμένη τιμή. Δηλαδή, η σύγκλιση της f_A δεν είναι ένα μαθηματικό συμπέρασμα αλλά είναι απλά ένα εμπειρικό γεγονός.

Προφανώς, στην πράξη δεν μπορούμε ποτέ να επιτύχουμε την πιθανότητα ενός γεγονότος όπως δίνεται από το όριο στην (2.3). Μπορούμε μόνο να κάνουμε μία εκτίμηση της $P(A)$ βασιζόμενοι σε ένα μεγάλο αριθμό επαναλήψεων n.

Παράδειγμα 2.20
Στην ρίψη ενός νομίσματος ποια είναι η πιθανότητα εμφάνισης της κεφαλής $P(K)$; Η θεωρία σχετικής συχνότητας προσεγγίζει το πρόβλημα ρίχνοντας το νόμισμα , ας πούμε, 100 φορές, κάτω από τις ίδιες συνθήκες, και μετά μετρούμε πόσες φορές εμφανίσθηκε η κεφαλή. Αν υποθέσουμε ότι η κεφαλή εμφανίσθηκε 45 φορές, τότε ο λόγος 45/100 χρησιμοποιείται σαν εκτίμηση της πιθανότητας $P(K)$ για το συγκεκριμένο νόμισμα. Παρόλα που το νόμισμα είναι ομοιόμορφα ζυγισμένο δεν έχουμε 50 κεφαλές στις 100 ρίψεις. Με άλλα λόγια, δεν περιμένουμε να επιτύχουμε την ακριβή πιθανότητα από τα επαναλαμβανόμενα πειράματα. Παρόλα αυτά, εάν το νόμισμα είναι ομοιόμορφα ζυγισμένο, η σχετική συχνότητα θα μπορούσε να προσεγγίσει την αληθή πιθανότητα 1/2 καθώς αυξάνει ο αριθμός δοκιμών.

2.4.3 Υποκειμενική Θεωρία

Στην υποκειμενική θεωρία σαν πιθανότητα αναφέρεται μία μέτρηση της προσωπικής εμπιστοσύνης σε μία συγκεκριμένη πρόταση, όπως για παράδειγμα, η πίστη να κερδίσει τις βουλευτικές εκλογές ένα πολιτικό κόμμα.

Σύμφωνα με αυτή την θεωρία, κάποιος αντιστοιχεί ένα βάρος μεταξύ μηδέν και ένα για ένα γεγονός, σύμφωνα με τον βαθμό της πίστης του για την πιθανή πραγματοποίηση του γεγονότος. Για παράδειγμα, εάν είναι δύο φορές σιγουρότερος για την πραγματοποίηση του γεγονότος A από ότι είναι για το γεγονός B, και εάν A και B είναι τα μόνα πιθανά γεγονότα, προκύπτουν οι πιθανότητες $P(A)=2/3$ και $P(B)=1/3$. Γενικά, οι υποκειμενικές πιθανότητες για διάφορα γεγονότα που σχετίζονται με δεδομένη πρόταση είναι βάρη με άθροισμα την μονάδα.

Συμπερασματικά, οι τρεις ερμηνείες της πιθανότητας αποτελούν η μία συμπλήρωμα της άλλης. Κάποιος χρησιμοποιεί την καταλληλότερη για τις συνθήκες και το πείραμα τύχης. Παρόλα αυτά, οι θεωρίες που έχουν αναφερθεί δεν είναι αρκετές για να στηρίξουν μία ολοκληρωμένη μαθηματική θεωρία πιθανότητας έτσι ώστε να είναι εφικτή η εύρεση πιθανότητας οιουδήποτε σύνθετου γεγονότος. Στην κλασσική θεωρία, η δυσκολία απαρίθμησης των δειγματοσημείων εξαρτάται από το πόσο σύνθετο είναι το πείραμα τύχης ή το γεγονός. Επίσης, σε μερικές περιπτώσεις δεν τεκμηριώνεται ότι τα στοιχειώδη ενδεχόμενα είναι ισοπίθανα. Στην θεωρία σχετικής συχνότητας, η εύρεση πιθανότητας εξαρτάται από ένα μεγάλο αριθμό επαναλήψεων του πειράματος. Συνήθως, δεν γνωρίζουμε πόσο μεγάλος πρέπει να είναι ο αριθμός δοκιμών έτσι ώστε η σχετική συχνότητα να προσεγγίσει την αληθή τιμή. Κατά συνέπεια, μέχρι τώρα δεν αναφέρθηκε μία ολοκληρωμένη θεωρία πιθανότητας βάση της οποίας θα υπολογίζεται τεκμηριωμένα η πιθανότητα ενός γεγονότος.

Οι ατέλειες αυτές εξέλιπαν μετά την παρουσίαση μιας αξιωματικής μεθόδου θεμελίωσης της πιθανοθεωρίας από τον Ρώσσο μαθηματικό **Kolmogorov** το **1933**, η οποία έγινε έκτοτε ευρύτερα αποδεκτή. Όμως, στην αξιωματική θεώρηση των εννοιών της πιθανοθεωρίας που θα αναπτύξουμε παρακάτω θα πρέπει κανείς, να έχει πάντα υπόψη του την πρακτική ερμηνεία τους μέσω της έννοιας της κλασικής μεθόδου και σχετικής συχνότητας.

Στην επόμενη παράγραφο αναφέρονται τρία βασικά αξιώματα πιθανότητας (αξιώματα Kolmogorov) στα οποία στηρίζεται η πλέον δυναμικότερη και αποτελεσματικότερη θεωρία πιθανότητας.

2.5 ΑΞΙΩΜΑΤΑ ΚΑΙ ΘΕΩΡΗΜΑΤΑ ΠΙΘΑΝΟΤΗΤΑΣ

Υποθέτουμε ότι E είναι ένα πείραμα τύχης με δειγματικό χώρο S και S^* το δυναμοσύνολο που περιέχει όλα τα δυνατά υποσύνολα ή γεγονότα, $S^*=\{A_1, A_2, \ldots\}$. Μια συνάρτηση $\boldsymbol{P}$ ορισμένη στο S^* καλείται **συνάρτηση πιθανότητας** αν αντιστοιχεί σε κάθε γεγονός $A_\kappa \in S^*$ ένα πραγματικό αριθμό $\boldsymbol{P(A_\kappa)}$, ο οποίος καλείται **πιθανότητα** του γεγονότος A_κ, και δεχόμαστε ότι πληροί τα ακόλουθα αξιώματα.

Αξίωμα I Η πιθανότητα για κάθε γεγονός $\boldsymbol{A_\kappa \in S^*}$ υπάρχει και περιορίζεται μεταξύ του μηδέν και ένα,

$$\mathbf{0 \leq P(A_\kappa) \leq 1.}$$

Αξίωμα II Η πιθανότητα του γεγονότος S είναι ένα,

$$\boldsymbol{P(S)=1}$$

Είναι παραδεκτό στην θεωρία πιθανότητας ότι ένα γεγονός λέγεται ότι συνέβη εάν τουλάχιστον ένα από τα δειγματοσημεία του έχει συμβεί. Όταν εκτελείται ένα πείραμα, είμαστε βέβαιοι ότι τουλάχιστον ένα από δειγματοσημεία του S πρέπει να συμβεί, και ως εκ τούτου η πιθανότητα που σχετίζεται με το S είναι ένα.

Αξίωμα III Εάν A_κ και A_λ είναι δύο γεγονότα οριζόμενα στον ίδιο δειγματικό χώρο S, $A_\kappa , A_\lambda \in S^*$ ξένα μεταξύ τους, δηλαδή, $\boldsymbol{A_\kappa \cap A_\lambda = \emptyset}$, τότε ισχύει

$$\boldsymbol{P(A_\kappa \cup A_\lambda)=P(A_\kappa)+P(A_\lambda).}$$

Η σπουδαιότητα αυτού του αξιώματος είναι προφανής. Εάν με κάποια διαδικασία, αντιστοιχούμε μια πιθανότητα σε κάθε δειγματοσημείο του S, τότε επιτυγχάνουμε την πιθανότητα οιουδήποτε γεγονότος ορισμένο στο S, προσθέτοντας τις επί μέρους πιθανότητες όλων των δειγματοσημείων που ανήκουν στο γεγονός αυτό.

Η τριάδα **($\boldsymbol{S, S^*}$, P)** λέγεται **χώρος πιθανοτήτων**.

Η επιλογή των παραπάνω αξιωμάτων πιθανότητας προφανώς προέρχεται από τα αντίστοιχα χαρακτηριστικά της σχετικής συχνότητας. Οι αριθμοί $P(A)$ και f_A είναι πολύ «κοντά» ο ένας στον άλλον, εφόσον η σχετική συχνότητα f_A βασίζεται σε ένα μεγάλο αριθμό επανάληψης του πειράματος E. Το γεγονός αυτό δικαιολογεί την χρήση του $P(A)$ για την μέτρηση του πόσο πιθανό είναι να συμβεί το A.

Μέχρι τώρα δεν γνωρίζουμε πώς να υπολογίζουμε την πιθανότητα $P(A)$. Αναφέραμε απλώς τις βασικές της ιδιότητες τις οποίες παραδεχόμαστε ως αξιώματα. Με βάση αυτά τα αξιώματα αποδεικνύονται στη συνέχεια εφτά βασικά θεωρήματα τα οποία αποτελούν τα σπουδαιότερα εργαλεία για τον υπολογισμό της πιθανότητας $P(A)$.

Θεώρημα 2.1:

Εάν $\varnothing$, $\varnothing \in S^*$, είναι το κενό σύνολο, τότε

$$P(\varnothing)=0. \tag{2.4}$$

Απόδειξη: Για κάθε γεγονός $A \in S^*$ ισχύει, $A=A \cup \varnothing$. Εφόσον τα A και $\varnothing$ είναι αμοιβαία αποκλειόμενα γεγονότα, από το τρίτο αξίωμα συνεπάγεται ότι

$$P(A)=P(A \cup \varnothing)=P(A)+P(\varnothing) \Rightarrow P(\varnothing)=0.$$

Παρατήρηση

Το γεγονός $\varnothing$ καλείται και **αδύνατο** γεγονός. Το αντίστροφο του θεωρήματος (2.1) δεν ισχύει. Όταν ένα γεγονός έχει μηδενική πιθανότητα, $P(A)=0$, δεν σημαίνει ότι το A είναι το αδύνατο γεγονός ή ότι είναι το κενό σύνολο.

Θεώρημα 2.2:

Εάν $\overline{A}$ είναι το συμπληρωματικό γεγονός του $A \in S^*$ τότε ισχύει

$$P(A) = 1 - P(\overline{A}). \tag{2.5}$$

Απόδειξη: Ο δειγματικός χώρος S μπορεί να γραφεί,

$$S = A \cup \overline{A}$$

Επειδή τα γεγονότα A και $\overline{A}$ είναι ξένα μεταξύ τους, χρησιμοποιώντας τα αξιώματα 1 και 2 επιτυγχάνουμε

$$1= P(A)+ P(\overline{A}).$$

Το θεώρημα αυτό είναι πάρα πολύ χρήσιμο, όπως θα δούμε παρακάτω. Σε πολλά προβλήματα όταν επιθυμούμε να υπολογίσουμε την πιθανότητα $P(A)$, ενδέχεται να είναι πιο εύκολος ο υπολογισμός της πιθανότητας του συμπληρωματικού γεγονότος $\overline{A}$, $P(\overline{A})$, και στη συνέχεια την αφαιρούμε από την μονάδα.

Θεώρημα 2.3:

Εάν A και B είναι δύο γεγονότα οριζόμενα στον ίδιο δειγματικό χώρο, A $B \in S^*$, έτσι ώστε $A \subset B$, τότε

$$P(A) \leq P(B).$$

Απόδειξη: Το γεγονός B είναι η ένωση των δύο ξένων γεγονότων A και B-A, B=AU(B-A), άρα χρησιμοποιώντας το τρίτο αξίωμα έχουμε

$$P(B)=P[AU(B\text{-}A)]=P(A)+P(B\text{-}A),$$

και επειδή από το πρώτο αξίωμα ισχύει $P(B\text{-}A)\geq 0$ συνεπάγεται ότι $P(A)\leq P(B)$.

Αυτό το αποτέλεσμα είναι πολύ λογικό. Διότι αυτό λέει ότι εάν το γεγονός B πρέπει να συμβαίνει οποτεδήποτε συμβαίνει το A, το B είναι τουλάχιστον πιθανό όσο και το A.

Θεώρημα 2.4:

Εάν A και B είναι δύο γεγονότα του ίδιου δειγματοχώρου S, A $B\in S^*$, ισχύει

$$\boldsymbol{P(B\text{-}A)=P(B)\text{-}P(A\cap B)}. \tag{2.6}$$

Απόδειξη:Το γεγονός B είναι η ένωση των δύο ξένων γεγονότων $A\cap B$ και B-A, B= $(A\cap B)U(B\text{-}A)$, και λόγω του τρίτου αξιώματος έχουμε

$$P(B)=P[(A\cap B)U(B\text{-}A)]=P(A\cap B)+P(B\text{-}A),$$

και μετά από αυτό το συμπέρασμα του θεωρήματος είναι προφανές.

Το θεώρημα 2.4 μπορεί να επεκταθεί και για περισσότερα γεγονότα, όπως ακολουθεί αμέσως παρακάτω.

Θεώρημα 2.5:

Εάν A, B και Γ είναι τρία γεγονότα ορισμένα στον ίδιο δειγματικό χώρο, A B $\Gamma\in S^*$, τότε

$$P[A-(B\cup\Gamma)]=P(A)-P(A\cap B)-P(A\cap\Gamma)+P(A\cap B\cap\Gamma)$$

Απόδειξη:

$$
\begin{aligned}
P[A-(B\cup \Gamma)] &= P(A)-P[A\cap(B\cup \Gamma)] \\
&= P(A)-P[(A\cap B)\cup(A\cap \Gamma)] \\
&= P(A)-[P(A\cap B)+P(A\cap \Gamma)-P(A\cap B\cap \Gamma)] \\
&= P(A)-P(A\cap B)-P(A\cap \Gamma)+P(A\cap B\cap \Gamma)
\end{aligned}
$$

Θεώρημα 2.6:

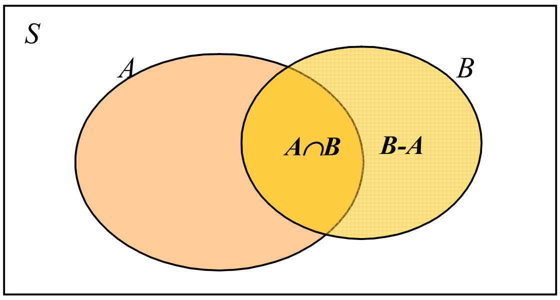
Για κάθε δύο γεγονότα A και B ορισμένα στον ίδιο δειγματικό χώρο, A $B\in S^*$, ισχύει

$$\boldsymbol{P(AUB)=P(A)+P(B)-P(A\cap B).} \tag{2.7}$$

Απόδειξη: Η ένωση του A και B, AUB, είναι ισοδύναμη με την ένωση των αμοιβαίως αποκλειόμενων γεγονότων A και B-A, βλέπε Σχήμα 2.8,

$$AUB= AU(B\text{-}A),$$

χρησιμοποιώντας το αξίωμα 3 έχουμε,



Σχήμα 2.8 Διάγραμμα Venn για δύο γεγονότα *A, B*

$$P(AUB)=P[AU(B\text{-}A)]=P(A)+P(B\text{-}A)$$

και αντικαθιστώντας την πιθανότητα της διαφοράς από το θεώρημα 2.4 προκύπτει,

$$P(AUB)=P(A)+P(B\text{-}A)= P(A)+P(B) - P(A\cap B).$$

Το αποτέλεσμα του θεωρήματος 2.6 καλείται επίσης και **προσθετικός κανόνας**. Παριστάνει μία προφανή επέκταση του αξιώματος ΙΙΙ, διότι εάν $A \cap B = \emptyset$, η τελευταία σχέση οδηγεί στο αξίωμα ΙΙΙ. Επί πλέον, με την βοήθεια του Σχήματος 2.8, ο προσθετικός κανόνας γίνεται εύκολα κατανοητός. Προσθέτοντας τις πιθανότητες $P(A)$ και $P(B)$, η πιθανότητα της τομής $P(A \cap B)$ συμπεριλαμβάνεται δύο φορές, για αυτό αφαιρείται μία φορά από το άθροισμα.

Θεώρημα 2.7:

Εάν A, B και Γ είναι τρία γεγονότα ορισμένα στον ίδιο δειγματικό χώρο, A B $\Gamma \in S^*$, τότε

$$P(A \cup B \cup \Gamma) = P(A) + P(B) + P(\Gamma) - P(A \cap B) - P(A \cap \Gamma) - P(B \cap \Gamma) + P(A \cap B \cap \Gamma) \qquad (2.8)$$

Απόδειξη: Το θεώρημα αποδεικνύεται εύκολα αν γράψει κανείς την ένωση $A \cup B \cup \Gamma$ σαν την ένωση δύο γεγονότων A και $B \cup \Gamma$, $A \cup (B \cup \Gamma)$, και εφαρμόσει στη συνέχεια τον παραπάνω προσθετικό κανόνα (2.7). Οι λεπτομέρειες αφήνονται στον αναγνώστη για εξάσκηση.

Ο **προσθετικός κανόνας** μπορεί να επεκταθεί σε όσα γεγονότα θέλουμε. Μία γενίκευση του προσθετικού κανόνα, όπως αποδεικνύεται με την μέθοδο της μαθηματικής επαγωγής, είναι η ακόλουθη. Εάν $A_1, A_2, \ldots, A_n$ είναι n γεγονότα ορισμένα στον ίδιο δειγματικό χώρο, τότε

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j=2}^{n} P(A_i \cap A_j) + \sum_{i<j<r=3} P(A_i \cap A_j \cap A_r)$$
$$+ \ldots + (-1)^{n-1} P(A_1 \cap A_2 \cap \ldots \cap A_n) \qquad (2.9)$$

Παρατήρηση

Εάν $A_1, A_2, \ldots, A_n$ είναι n γεγονότα ορισμένα στον ίδιο δειγματικό χώρο, και είναι μεταξύ τους αμοιβαίως αποκλειόμενα ή **ασυμβίβαστα** ανά δύο, ανά τρία, …, ανά n, ο προσθετικός κανόνας απλουστεύεται ως ακολούθως,

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + P(A_3) + \ldots + P(A_n)$$

Παράδειγμα 2.21

Σε ένα εργαστήριο ελέγχονται δείγματα νερού από ένα ποταμό για να διαπιστωθεί η ύπαρξη βλαβερών ουσιών *(βαρύ μέταλλο*, *μόλυβδος* ή *υδράργυρος*). Από το παρελθών είναι γνωστό ότι από τα δείγματα, τα οποία παίρνονται από μία περιοχή όπου μεγάλα εργοστάσια ρίχνουν τα απόβλητά τους,

- 36% περιέχει σε τοξικό επίπεδο μόλυβδο,
- 20% περιέχει σε τοξικό επίπεδο υδράργυρο,
- 12% περιέχει σε τοξικό επίπεδο μόλυβδο και υδράργυρο.

Παίρνουμε ένα τυχαίο δείγμα νερού από τη ίδια περιοχή του ποταμού,
(α) ποια η πιθανότητα ότι θα περιέχει σε τοξικό επίπεδο μόνο μόλυβδο;
(β) ποια η πιθανότητα ότι δεν θα περιέχει σε τοξικό επίπεδο μόλυβδο ή υδράργυρο;
(γ) ποια η πιθανότητα ότι θα περιέχει σε τοξικό επίπεδο μόνο ένα από τα επικίνδυνα μέταλλα;

Απάντηση:

Συμβολίζω με Y το γεγονός ότι το δείγμα περιέχει σε τοξικό επίπεδο υδράργυρο και με M το γεγονός ότι το δείγμα περιέχει σε τοξικό επίπεδο μόλυβδο. Μας δίνεται ότι

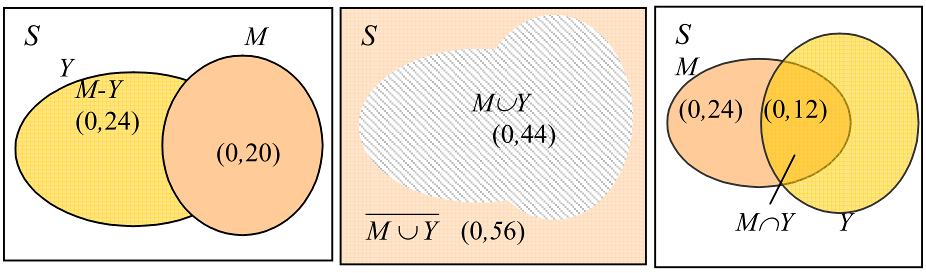
$$P(Y) = 0{,}20, \quad P(M) = 0{,}36, \quad P(M \cap Y) = 0{,}12 .$$

(α) Το γεγονός ότι το δείγμα περιέχει μόλυβδο και όχι υδράργυρο συμβολίζεται με την διαφορά M-Y. Άρα η ζητούμενη πιθανότητα, σύμφωνα με το θεώρημα 2.4, είναι

$$P(M - Y) = P(M) - P(M \cap Y) = 0{,}36 - 0{,}12 = 0{,}24$$

(β) Το γεγονός ότι το δείγμα δεν θα περιέχει σε τοξικό επίπεδο μόλυβδο ή υδράργυρο είναι το συμπλήρωμα του γεγονότος ότι θα περιέχει σε τοξικό επίπεδο μόλυβδο ή υδράργυρο. Δηλαδή, συμβολίζεται σαν

$$\overline{M \cup Y} ,$$



(α) (β) (γ)

Σχήμα 2.9 Διαγράμματα Venn και οι αντίστοιχες πιθανότητες

και η πιθανότητά του υπολογίζεται με το θεώρημα 2.2 και τον προσθετικό κανόνα.

$$P(\overline{M \cup Y}) = 1 - P(M \cup Y) = 1 - [P(M) + P(Y) - P(M \cap Y)]$$
$$= 1 - 0{,}36 - 0{,}20 + 0{,}12 = 1 - 0{,}44 = 0{,}56$$

(γ) Το γεγονός ότι το δείγμα περιέχει σε τοξικό επίπεδο μόνο ένα από τα μέταλλα συμβολίζεται με την ένωση

$$(M - Y) \cup (Y - M).$$

Άρα, ακολουθώντας τον προσθετικό κανόνα και την πιθανότητα διαφοράς δύο γεγονότων (θεώρημα 2.4), η ζητούμενη πιθανότητα είναι

$$P[(M - Y) \cup (Y - M)] = P(M - Y) + P(Y - M)$$
$$= P(M) - P(M \cap Y) + P(Y) - P(Y \cap M)$$
$$= 0{,}36 - 0{,}12 + 0{,}20 - 0{,}12 = 0{,}32$$

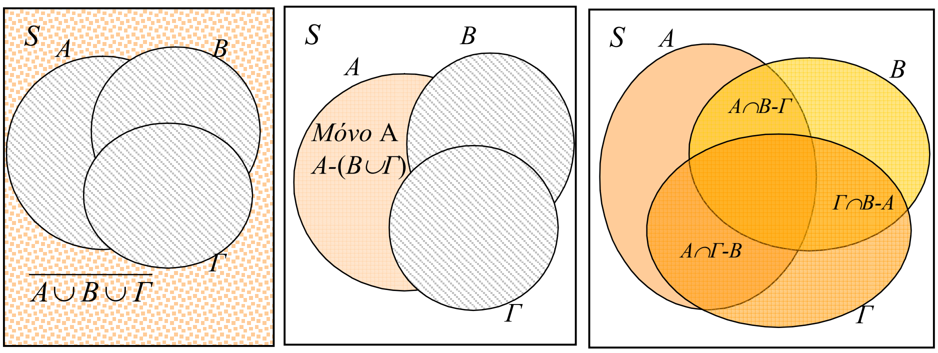
Παράδειγμα 2.22

Σε μία πόλη τρείς από τις εφημερίδες που κυκλοφορούν είναι οι Α, Β, και Γ. Είναι γνωστό ότι από τους κατοίκους :
το 20% διαβάζει την Α,
το 16% διαβάζει την Β,
το 14% διαβάζει την Γ,
το 8% διαβάζει την Α και Β, το 5% διαβάζει την Α και Γ, το 4 % διαβάζει Β και Γ, το 2% διαβάζει την Α, Β και Γ.

Για ένα άτομο που επιλέγεται τυχαία από την πόλη, υπολογίστε την πιθανότητα ότι:
(α) δεν διαβάζει καμία από τις παραπάνω εφημερίδες,
(β) διαβάζει μόνο την εφημερίδα Α,
(γ) διαβάζει μόνο δύο εφημερίδες.

Απάντηση:

Έστω τα γεγονότα



(α) Καμία από τις Α, Β, Γ (β) Μόνο Α (γ) Ακριβώς δύο από τις Α, Β, Γ

***Σχήμα* 2.10**. Παράδειγμα 2.22, επεξήγηση γεγονότων με διαγράμματα **Venn**.

A={*το άτομο διαβάζει την εφημερίδα* Α},
B={*το άτομο διαβάζει την εφημερίδα* Β},
Γ={*το άτομο διαβάζει την εφημερίδα* Γ}.

Τα παραπάνω ενδεχόμενα μπορούν να συμβολισθούν με άλγεβρα γεγονότων ώς εξής

- δεν διαβάζει καμία από τις παραπάνω εφημερίδες: $\overline{A \cup B \cup \Gamma}$
- διαβάζει μόνο την Α: $A-(B\cup\Gamma)$
- διαβάζει μόνο δύο εφημερίδες: $(A \cap B - \Gamma) \cup (A \cap \Gamma - B) \cup (B \cap \Gamma - A)$

(α) Από το θεώρημα 2.2 προκύπτει,

$$P(\overline{A\cup B\cup \Gamma}) = 1 - P(A\cup B\cup \Gamma) = 1-(0{,}20+0{,}16+0{,}14-0{,}08-0{,}04-0{,}05+0{,}02)$$
$$= 1-0{,}35 = 0{,}65.$$

(β) Πρέπει να τονίσουμε ότι στο γεγονός *A* (κάποιος διαβάζει την εφημερίδα Α) δεν σημαίνει ότι διαβάζει μόνο την Α. Ενδέχεται να διαβάζει την Β ή και την Γ κ.τ.λ.. Με βάση τον κανόνα της διαφοράς(θεώρημα 2.5) και τα δεδομένα, εύκολα μπορούμε να υπολογίσουμε την ζητούμενη πιθανότητα,

$$P[A-(B\cup\Gamma)] = P(A) - P(A\cap B) - P(A\cap\Gamma) + P(A\cap B\cap\Gamma)$$
$$= 0{,}20 - 0{,}08 - 0{,}05 + 0{,}02 = 0{,}09$$

(γ) Εδώ, αν παρατηρήσουμε το σχεδιάγραμμα Venn, καταλαβαίνουμε ότι πρόκειται για την πιθανότητα της ένωσης των γεγονότων $A\cap B-\Gamma$, $A\cap\Gamma-B$, $\Gamma\cap B-A$. Έτσι, εφαρμόζοντας τον προσθετικό κανόνα και μετά την πιθανότητα διαφοράς, έχουμε,

$$P[(A\cap B-\Gamma)\cup(A\cap\Gamma-B)\cup(B\cap\Gamma-A)] = P(A\cap B-\Gamma)+P(A\cap\Gamma-B)+P(B\cap\Gamma-A)$$
$$= P(A\cap B)-P(A\cap B\cap\Gamma) \quad + \quad P(A\cap\Gamma)-P(A\cap\Gamma\cap B) \quad + \quad p(B\cap\Gamma)-P(B\cap\Gamma\cap A)$$
$$= 0{,}08-0{,}02 \quad + \quad 0{,}05-0{,}02 \quad + \quad 0{,}04-0{,}02 = 0{,}11$$

2.6 ΑΡΧΕΣ ΑΠΑΡΙΘΜΗΣΗΣ

Στον υπολογισμό της πιθανότητας $P(A)$ με την κλασσική μέθοδο, $P(A)=n_A/n_S$, όπου n_A είναι ο αριθμός όλων των δυνατών τρόπων με τους οποίους μπορεί να συμβεί το *A* ενώ n_S είναι ο ολικός αριθμός των αποτελεσμάτων του πειράματος τύχης. Στα παραδείγματα που αναφέρθηκαν μέχρι στιγμής, σχετικά εύκολα απαριθμήθηκαν τα n_A και n_S. Όμως, σε πιο σύνθετες περιπτώσεις είναι συχνά απαραίτητο να απαριθμήσουμε ένα μεγάλο αριθμό τρόπων ή στοιχείων και χρειαζόμαστε συστηματική αρίθμηση ή διαδικασία απαρίθμησης.

Για παράδειγμα, εάν τραβήξουμε τυχαία πέντε κάρτες από μία τράπουλα, χωρίς να τις επανατοποθετούμε στην τράπουλα, ποία η πιθανότητα ότι θα πάρουμε ακριβώς δύο άσσους; Για να απαντήσουμε στην ερώτηση με την κλασσική μέθοδο πρέπει να απαριθμήσουμε όλα τα δυνατά αποτελέσματα του δειγματικού χώρου που είναι όλες οι δυνατές πεντάδες, όπως (A καρό, K κούπα, Q κούπα, J σπαθί, 10 κούπα), και στη συνέχεια να απαριθμήσουμε εκείνες που περιέχουν ακριβώς δύο άσσους.

Στο παράδειγμα αυτό δεν χρειάζεται να αναγραφούν όλες οι δυνατές πεντάδες, το οποίο θα ήταν και δύσκολο. Ευτυχώς, στον προσδιορισμό της πιθανότητας του γεγονότος, με την κλασσική μέθοδο, συχνά χρειαζόμαστε μόνο τον ολικό αριθμό των στοιχείων του συνόλου S ή του συνόλου του γεγονότος. Αποτελεσματικές τεχνικές για την απαρίθμηση τόσο πολλών στοιχείων είναι διαθέσιμες και είναι γνωστές σαν «μεταθέσεις» και «συνδυασμοί» και στηρίζονται σε ένα πολύ απλό κανόνα – ο κανόνας του γινομένου.

2.6.1 Ο κανόνας του γινομένου

Ένα σύνθετο πείραμα E συνίσταται στην εκτέλεση κ απλών πειραμάτων, $E_1, E_2, \ldots, E_\kappa$, των οποίων οι αντίστοιχοι δειγματοχώροι είναι $S_1, S_2, \ldots, S_\kappa$. Ο δειγματικός χώρος S του πειράματος E, ο οποίος περιέχει όλα τα πιθανά αποτελέσματα του σύνθετου πειράματος E (βλέπε παράγραφο 2.1, σύνθετος δειγματοχώρος) προκύπτει από το **καρτεσιανό γινόμενο**,

$$S=S_1\times S_2\times\ldots\times S_\kappa$$

Έτσι, ο **σύνθετος δειγματοχώρος *S*** περιέχει σαν δειγματοσημεία τα *διατεταγμένα* σύνολα,

$$S=\{(s_{1\lambda},s_{2\mu},\ldots,s_{\kappa\nu})|\ s_{1\lambda}\in S_1,\ s_{2\mu}\in S_2,\ldots,s_{\kappa\nu}\in S_\kappa,\ \lambda=1,2,\ldots,n_1,\ \mu=1,2,\ldots,n_2,\ \ldots,\ \nu=1,2,\ldots,n_\kappa\}.$$

Άμεση συνέπεια του παραπάνω σκεπτικού είναι ότι το πλήθος, n, όλων των δυνατών αποτελέσματα του E είναι ο **κανόνας του γινομένου**,

$$n=n_1\times n_2\times\ldots\times n_\kappa, \qquad (2.10)$$

όπου n_1, n_2,...,n_κ παριστάνουν αντίστοιχα τον αριθμό των δυνατών αποτελεσμάτων (πληθάριθμους) των κ απλών δειγματοχώρων S_1, S_2,..., S_κ.

Παράδειγμα 2.23
Εάν για την κατασκευή ενός προϊόντος απαιτούνται τρία στάδια επεξεργασίας, όπου υπάρχουν 5, 6 και 2 ίδιες μηχανές για το κάθε στάδιο αντίστοιχα, πόσες είναι όλες οι δυνατές διαδρομές που μπορεί να ακολουθήσει για την κατασκευή του μία μονάδα προϊόντος;
Απάντηση:
Θεωρούμε ότι στο κάθε στάδιο επεξεργασίας εκτελείται ένα απλό πείραμα, όπου επιλέγουμε τυχαία μία από τις μηχανές του. Έτσι, το πρώτο στάδιο επεξεργασίας μπορεί να πραγματοποιηθεί με $n_1=5$ διαφορετικούς τρόπους, το δεύτερο με $n_2=6$ διαφορετικούς τρόπους και το τρίτο με $n_3=2$ διαφορετικούς τρόπους. Η επιλογή μίας διαδρομής της επεξεργασίας του προϊόντος αποτελεί το σύνθετο πείραμα E, όπου επιλέγεται μία μηχανή από κάθε στάδιο επεξεργασίας. Όλες οι προκύπτουσες δυνατές διαδρομές σύμφωνα με τον κανόνα του γινομένου 2.10 είναι

$$n=n_1\times n_2\times n_3=5\times 6\times 2=60 \text{ διαδρομές.}$$

Παράδειγμα 2.24
Υπάρχουν σε μία τάξη 15 αγόρια και 18 κορίτσια. Πόσες διαφορετικές επιτροπές των δύο ατόμων μπορούμε να σχηματίσουμε, όπου θα υπάρχουν άτομα και των δύο φύλων σε κάθε επιτροπή;
Απάντηση:
Για να σχηματίσουμε την επιτροπή (πείραμα E), επιλέγουμε ένα αγόρι(πείραμα E_1) και μετά ένα κορίτσι(Πείραμα E_2). Προφανώς, το πλήθος όλων των δυνατών αποτελεσμάτων του πειράματος Ε είναι σύμφωνα με τον κανόνα του γινομένου,

$$n=n_1\times n_2=15\times 18=270 \text{ επιτροπές.}$$

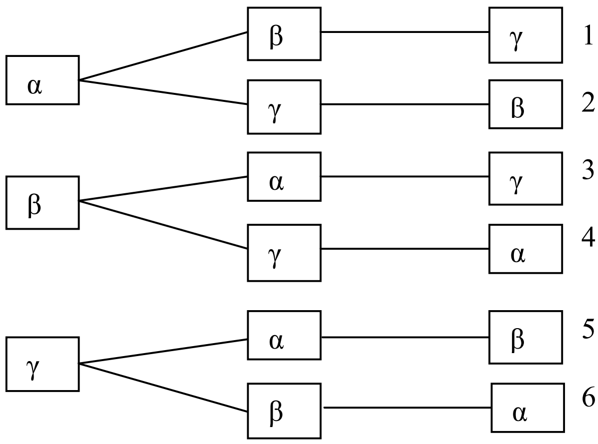
2.6.2 Μεταθέσεις

Συχνά, σε ένα σύνθετο πείραμα μας ενδιαφέρει το πλήθος όλων των δυνατών τρόπων με τους οποίους μπορούμε να τοποθετήσουμε $\boldsymbol{n}$ διαφορετικά αντικείμενα σε μία σειρά. Αυτό επιτυγχάνεται με τον κανόνα του γινομένου όπως φαίνεται στα δύο επόμενα παραδείγματα.

Παράδειγμα 2.25
Έχουμε τρία γράμματα α, β, και γ, με πόσους διαφορετικούς τρόπους μπορούμε να τα τοποθετήσουμε σε μία σειρά;

Ένα απλό σκεπτικό είναι ότι η τοποθέτηση αυτή συνίσταται στην επιλογή ενός από τα τρία γράμματα α, β, γ για την πρώτη θέση (πείραμα E_1), μετά στην επιλογή ενός από τα υπόλοιπα δύο για την δεύτερη θέση (πείραμα E_2) και τέλος την τοποθέτηση του μόνου γράμματος που μένει στην τρίτη θέση (πείραμα E_3). Από τον κανόνα του γινομένου (2.10) και την βοήθεια του Σχήματος 2.11 προκύπτει ότι όλοι οι δυνατοί είναι

$$n_1 \times n_2 \times n_3 = 3\times 2\times 1 = 6\ .$$



Σχήμα 2.11. Το διάγραμμα δένδρου για τον αριθμό τρόπων τοποθέτησης τριών γραμμάτων α, β, και γ σε μία σειρά

Παράδειγμα 2.26

Υποθέτουμε ότι έχουμε *n* διαφορετικά αντικείμενα. Με πόσους διαφορετικούς τρόπους μπορούμε να τα τοποθετήσουμε σε μία σειρά; Το πρόβλημα είναι ισοδύναμο με την τοποθέτηση των *n* αντικειμένων σε *n* κουτιά αριθμούμενα 1,2,3,…, *n,* όπου στο κάθε κουτί τοποθετούμε μόνο ένα αντικείμενο.

1	2	3	.	.	.	*n*

Στο πρώτο κουτί μπορούμε να τοποθετήσουμε ένα από τα *n* αντικείμενα (πείραμα E_1), δηλαδή με *n* διαφορετικούς τρόπους. Μετά την τοποθέτηση αυτή, στο δεύτερο κουτί τοποθετούμε ένα από τα υπόλοιπα (n-1) αντικείμενα(πείραμα E_2), παρόμοια το τρίτο κουτί μπορεί να πληρωθεί με (n-2) διαφορετικούς τρόπους (πείραμα E_3), το τέταρτο κουτί με (n-3) τρόπους (πείραμα E_4), και ούτω καθεξής. Έτσι, όλα τα δυνατά αποτελέσματα του σύνθετου πειράματος *E*, δηλαδή, ο αριθμός των τρόπων με τους οποίους μπορούμε να τοποθετήσουμε n αντικείμενα σε n κουτιά προκύπτει από τον κανόνα του γινομένου,

$$n \times (n\text{-}1) \times (n\text{-}2) \times (n\text{-}3) \times \ldots \times 2 \times 1 = n! \ (n \textit{ παραγοντικό}).$$

Γενικά, μία **μετάθεση** ενός αριθμού αντικειμένων είναι η τοποθέτηση των αντικειμένων αυτών σε μία ορισμένη σειρά. Ο αριθμός των **μεταθέσεων** ενός συνόλου από ***n*** αντικείμενα, λαμβάνοντάς τα όλα μαζί, όπως προκύπτει από την άμεση εφαρμογή του κανόνα του γινομένου είναι ***n*!.** Συμβολίζοντας αυτόν τον αριθμό και με ${}_n\boldsymbol{P}_n$ (*Permutations*), έχουμε

$${}_nP_n = n!, \tag{2.11}$$

όπου $n!$ διαβάζεται « n παραγοντικό » και είναι το γινόμενο από όλους τους αριθμούς από το 1 μέχρι το n, δηλαδή

$$n! = n\,(n\text{-}1)(n\text{-}2)\ldots 1. \tag{2.12}$$

Επίσης εδώ ορίζουμε $0! = 1$.

Παράδειγμα 2.27

Πόσους δεκαψήφιους αριθμούς χωρίς επαναλαμβανόμενα ψηφία μπορούμε να σχηματίσουμε με τα ψηφία 0, 1, 2, 3, …, 8, 9;

Απάντηση:

Ο ολικός αριθμός τοποθέτησης δέκα ψηφίων σε μία σειρά λαμβάνοντας και τα δέκα, σύμφωνα με τον κανόνα των μεταθέσεων (2.12), είναι

$$P(10,10)=10!=10\cdot 9\cdot 8\cdots 2\cdot 1=3.628.800.$$

Όμως, στον αριθμό αυτόν περιλαμβάνεται και ο αριθμός των δεκαψήφιων που αρχίζουν από μηδέν, τον οποίο πρέπει να αφαιρέσουμε από το 3.628.800. Το πλήθος των δεκαψηφίων που αρχίζουν από μηδέν είναι ο αριθμός των μεταθέσεων των ψηφίων 1, 2, 3, …, 9 που ακολουθούν μετά το μηδέν. Έτσι, όλοι οι δεκαψήφιοι που αρχίζουν με μηδέν είναι

$$P(9, 9)=9!=9\cdot 8\cdots 2\cdot 1=362.880$$

Άρα, ο ζητούμενος αριθμός είναι 3.628.800 – 362.880 = 3.265.920.
Το ίδιο αποτέλεσμα επιτυγχάνουμε αν σκεφθούμε ότι την πρώτη θέση δεν μπορεί να την πάρει ένα από τα δέκα ψηφία, αλλά ένα από τα εννέα ψηφία 1, 2, 3, …, 9, την δεύτερη θέση ένα από τα υπόλοιπα εννέα κ.ο.κ.. Έτσι, οι μεταθέσεις των δέκα ψηφίων με αυτόν τον περιορισμό είναι το γινόμενο $9\cdot 9\cdot 8\cdots 1=3.265.920$.

Θεωρούμε και πάλι n διαφορετικά αντικείμενα. Αυτή τη φορά επιθυμούμε να επιλέξουμε r από αυτά, $0 \le r \le n$, και τα τοποθετήσουμε σε μία σειρά. Με πόσους διαφορετικούς τρόπους μπορούμε να το κάνουμε αυτό. Συμβολίζουμε με $P(n, r)$ τον αριθμό των μεταθέσεων n αντικειμένων, λαμβάνοντάς τα ανά r. Λαμβάνουμε και πάλι υπόψη το τελευταίο σκεπτικό με την τοποθέτηση n αντικειμένων σε r κουτιά αριθμούμενα 1, 2, 3, …, r, όπου το κάθε κουτί χωράει μόνο ένα αντικείμενο. Έτσι το πρώτο κουτί μπορεί να δεχθεί ένα από τα n αντικείμενα, το δεύτερο ένα από τα $(n$-1), το τρίτο ένα από τα $(n$-2) , και ούτω κάθε εξής. Κατά συνέπεια η ολική διαδικασία μπορεί να εκπληρωθεί, πάλι με τον κανόνα του γινομένου, σε

$$n\,(n-1)(n-2)\ldots(n-r+1)$$

τρόπους. Χρησιμοποιώντας τον παραγοντικό συμβολισμό, μπορούμε να γράψουμε

$$P(n,r) = \frac{n!}{(n-r)!}\,. \tag{2.13}$$

Παράδειγμα 2.28

Υποθέτουμε ότι έχουμε εφτά δωμάτια και θέλουμε να δώσουμε από ένα σε τέσσερις προγραμματιστές και τα υπόλοιπα τρία να χρησιμοποιηθούν σαν εργαστήρια. Με πόσου διαφορετικούς τρόπους μπορούμε να το κάνουμε αυτό;

Απάντηση:

Αριθμούμε τα δωμάτια από ένα μέχρι εφτά και παίρνουμε τέσσερα από αυτά και τα βάζουμε σε μία σειρά. Το πρώτο δίνεται στον A προγραμματιστή, το δεύτερο στον B, κ.ο.κ.. Προφανώς, έχουμε τις μεταθέσεις εφτά στοιχείων παίρνοντάς τα ανά τέσσερα,

$$P(7,4) = \frac{7!}{(7-4)!} = \frac{7!}{3!} = 4\cdot 5\cdot 6\cdot 7 = 840 \quad \text{διαφορετικούς τρόπους.}$$

2.6.3 Συνδυασμοί

Στην περίπτωση αυτή μας ενδιαφέρει να απαριθμήσουμε με πόσους τρόπους μπορούμε να επιλέξουμε r από n αντικείμενα χωρίς να αναφερόμαστε στην σειρά. Για παράδειγμα, έχουμε τα αντικείμενα α,β, και γ, n =3,και επιλέγουμε δύο από αυτά, r=2. Επιθυμούμε να απαριθμήσουμε αβ, αγ, και βγ. Με άλλα λόγια, δεν απαριθμούμε αβ και βα εφόσον περιέχουν τα ίδια αντικείμενα και αλλάζει μόνο η σειρά.

Για να επιτύχουμε το γενικό αποτέλεσμα επικαλούμαστε τον κανόνα των μεταθέσεων *n* αντικειμένων λαμβάνοντάς τα ανά *r,* ο οποίος είναι $n!/(n-r)!$. Συμβολίζουμε με C(*n, r*) τον αριθμό των τρόπων με τους οποίους επιλέγουμε *r* αντικείμενα από *n*, ασχέτως της σειράς (δηλαδή, C(*n, r*) είναι ο ζητούμενος αριθμός). Από *r* επιλεγέντα αντικείμενα προκύπτουν *r*! μεταθέσεις αυτών. Οπότε, εφαρμόζοντας πάλι τον κανόνα γινομένου, ο αριθμός των τρόπων με τους οποίους παίρνουμε *r* από *n* αντικείμενα και τα βάζουμε σε μία σειρά είναι C(*n, r*) *r*!, και μαζί με τον παραπάνω κανόνα μεταθέσεων, επιτυγχάνουμε

$$C(n,r)r! = \frac{n!}{(n-r)!}$$

Έτσι ο αριθμός των τρόπων με τους οποίους μπορούμε να επιλέξουμε *r* από *n* αντικείμενα χωρίς να αναφερόμαστε στην σειρά, δίνεται από τον κανόνα

$$C(n,r) = \frac{n!}{(n-r)!r!}\,. \tag{2.14}$$

Ο αριθμός αυτός χρησιμοποιείται συχνά στις πιθανότητες και είναι γνωστός σαν οι συνδυασμοί *n* αντικειμένων ανά *r*. Μερικές φορές χρησιμοποιείται και ένας άλλος συμβολισμός ή μπορεί η εξίσωση 2.14 να εκφραστεί ισοδύναμα όπως,

$$C(n,r) = \binom{n}{r} = \frac{n!}{(n-r)!r!} = \frac{n(n-1)(n-2)\cdots(n-r+1)}{r!}$$

Οι αριθμοί

$$\binom{n}{r}$$

συχνά καλούνται διωνυμικοί συντελεστές, διότι εμφανίζονται ως συντελεστές στην ανάπτυξη της διωνυμικής έκφρασης $(\alpha+\beta)^n$,

$$(\alpha+\beta)^n = \sum_{r=0}^{n} \binom{n}{r} \alpha^r \beta^{n-r} \quad .$$

Άμεση σχέση με το τελευταίο αποτέλεσμα έχει και ο προσδιορισμός του αριθμού όλων των υποσυνόλων ενός δειγματικού συνόλου *S,* το οποίο περιέχει *n* δειγματοσημεία. Όπως αναφέρθηκε στην πρώτη παράγραφο, το πλήθος όλων των γεγονότων που μπορούν να ορισθούν σε ένα πείραμα τύχης είναι 2^n. Μπορούμε τώρα να τεκμηριώσουμε αυτό το αποτέλεσμα. Για να επιτύχουμε τα υποσύνολα πρέπει να επιλέξουμε το κενό σύνολο, εκείνα τα υποσύνολα που αποτελούνται από ένα δειγματοσημείο, εκείνα τα υποσύνολα που αποτελούνται από δύο δειγματοσημεία, …, και τέλος το ίδιο το σύνολο που αποτελείται από n δειγματοσημεία. Αυτό μπορεί να γίνει με

$$\binom{n}{0}+\binom{n}{1}+\binom{n}{2}+\ldots+\binom{n}{n}$$

τρόπους. Οπωσδήποτε το άθροισμα αυτών των διωνυμικών συντελεστών είναι απλά η ανάπτυξη του

$$(1+1)^n = 2^n.$$

Δύο από τις σπουδαιότερες **ιδιότητες** των διωνυμικών αριθμών αναφέρονται εδώ (τις οποίες ο αναγνώστης αν επιθυμεί μπορεί να αποδείξει σαν μία εξάσκηση):

1) $\binom{n}{r}=\binom{n}{n-r}$

2) $\binom{n}{r}=\binom{n-1}{r-1}+\binom{n-1}{r}$

Η επιβεβαίωση των παραπάνω ταυτοτήτων γίνεται εύκολα, εφαρμόζοντας και στα δύο μέρη της εξίσωσης τον κανόνα (2.14).

Παράδειγμα 2.29

Μία ομάδα καλαθοσφαίρισης διαθέτει συνολικά 10 παίκτες. Ο προπονητής της ομάδας πρέπει να επιλέξει 5 παίκτες για κάποιον αγώνα.
(α) Πόσες διαφορετικές ομάδες των πέντε παικτών μπορεί να επιλεχθούν;
(β) Εάν ο κάθε παίκτης μπορεί να πάρει οιαδήποτε από τις πέντε διαφορετικές θέσεις του παιχνιδιού, πόσες διαφορετικές πεντάδες μπορούν να σχηματισθούν;
(γ) Πόσες διαφορετικές πεντάδες μπορούν να σχηματισθούν αν δύο συγκεκριμένοι παίκτες θα συμμετέχουν οπωσδήποτε στην πεντάδα του αγώνα;
Απάντηση :
(α) Εδώ δεν μας ενδιαφέρει η θέση που παίρνει ο παίκτης που επιλέγεται για την πεντάδα. Κατά συνέπεια, αυτό είναι ένα πρόβλημα συνδυασμών,

$$C(10,5)=\frac{10!}{(10-5)!5!}=\frac{5!\cdot 6\cdot 7\cdot 8\cdot 9\cdot 10}{5!\cdot 1\cdot 2\cdot 3\cdot 4\cdot 5}=252 \text{ συνδυασμοί.}$$

(β) Στην περίπτωση αυτή μας ενδιαφέρει και η θέση που παίρνει ο παίκτης στην καλαθοσφαίριση, δηλαδή το πρόβλημα εδώ αναφέρεται στις μεταθέσεις,

$$P(10,5)=\frac{10!}{(10-5)!}=30.240 \text{ σχηματισμούς.}$$

Παρατηρούμε ότι, όταν επιλέγεται μία πεντάδα, ο αριθμός των τρόπων τοποθέτησης των παικτών είναι

$$P(5,5)=5!=120 \text{ τρόποι.}$$

Έτσι, έχουμε

30.240 σχηματισμοί=(252 πεντάδες)(120 μεταθέσεις για κάθε πεντάδα).

(γ) Επειδή δύο παίκτες από τους δέκα σίγουρα συμμετέχουν στην πεντάδα, ο προπονητής έχει να επιλέξει τρεις από τους υπόλοιπους οκτώ παίκτες,

$$C(8,3)=\frac{8!}{(8-3)!3!}=56 \text{ τρόπους.}$$

Και τέλος οι σχηματισμοί είναι,

(56 πεντάδες)(120 μεταθέσεις για κάθε πεντάδα)=6.720 σχηματισμοί.

Παράδειγμα 2.30

Ένα κιβώτιο ανταλλακτικών εξαρτημάτων περιέχει 80 καλά και 20 ελαττωματικά. Παίρνουμε τυχαία 10 ανταλλακτικά, χωρίς επανατοποθέτηση στο κιβώτιο. Ποια είναι η πιθανότητα ότι τα μισά ακριβώς από τα επιλεγόμενα ανταλλακτικά είναι ελαττωματικά;

Απάντηση:

Για να απαντήσουμε στο πρόβλημα πρέπει πρώτα να απαριθμήσουμε όλα τα δειγματοσημεία του δειγματικού χώρου S. Ο αριθμός όλων των δειγματοσημείων είναι ο αριθμός των διαφορετικών δεκάδων που μπορούν να επιλεγούν από το κιβώτιο, $C(100, 10)$. Ανάμεσα σε αυτές τις δεκάδες πόσες έχουν ακριβώς πέντε ελαττωματικά ανταλλακτικά; Δηλαδή, πόσες διαφορετικές δεκάδες μπορούμε να κάνουμε επιλέγοντας 5 από τα 20 ελαττωματικά και 5 από τα 80 καλά ανταλλακτικά; Κάνοντας χρήση τον κανόνα του γινομένου ο αριθμός διαφορετικών τέτοιων δεκάδων είναι $C(20, 5)$ φορές τους $C(80, 5)$. Επειδή όλα τα αποτελέσματα του πειράματος είναι ισοπίθανα, ακολουθώντας την κλασσική θεωρία, η πιθανότητα ακριβώς 5 ελαττωματικών και 5 καλών ανταλλακτικών στα 10 επιλεγόμενα δίνεται από τον κανόνα

$$\frac{\binom{20}{5}\binom{80}{5}}{\binom{100}{10}}=0.021\cdot$$

Παρατηρήσεις

(A) Όταν επιλέγουμε r από τα n αντικείμενα συνήθως εννοούμε χωρίς επανατοποθέτηση, στην περίπτωση αυτή ο αριθμός διαφορετικών τρόπων είναι $C(n, r)$. Εάν η επιλογή r από τα n αντικείμενα γίνει με επανάθεση, τότε προφανώς έχουμε n^r διαφορετικούς τρόπους.

(B) Όταν πρόκειται να υπολογίσουμε το παραγοντικό ενός αρκετά μεγάλου αριθμού πρέπει να γίνουν πολλές πράξεις. Μπορούμε όμως να προσεγγίσουμε το αποτέλεσμα με τον τύπο του **Stirling,**

$$n! \approx \sqrt{2\pi} e^{-n} n^{n+1/2} .$$

2.6.4 Μεταθέσεις όταν όλα τα αντικείμενα δεν είναι ίδια

Αναφερθήκαμε μέχρι τώρα σε μεταθέσεις όπου όλα τα αντικείμενα είναι διαφορετικά. Στη συνέχεια θα αναφερθούμε με τον αριθμό διαφορετικών τοποθετήσεων αντικειμένων όταν κάποια από αυτά είναι ολόιδια.

Εάν υπάρχουν n αντικείμενα, από τα οποία n_1 είναι ίδια, τότε ο αριθμός μεταθέσεων των n αντικειμένων, παίρνοντάς τα όλα μαζί, είναι $n_1!$ φορές λιγότερο από το $n!$. Διότι $n_1!$ είναι το πλήθος των μεταθέσεων των n αντικειμένων, όπου αλλάζουν θέση μεταξύ τους μόνο τα n_1 αντικείμενα, ενώ τα υπόλοιπα n-n_1 παραμένουν στις ίδιες θέσεις τους. Όμως, τα n_1 αντικείμενα είναι ίδια και οι μεταθέσεις αυτές συμπίπτουν σε μία. Έτσι για τον αριθμό μεταθέσεων n αντικειμένων, από τα οποία n_1 είναι ίδια, έχουμε

$$P(n,n) = \frac{n!}{n_1!} .$$

Παρόμοια, αν έχουμε n αντικείμενα, τα οποία χωρίζονται σε κ ομάδες ίδιων στοιχείων, δηλαδή, n_1 είναι ίδια, άλλα n_2 είναι ίδια, άλλα n_3 είναι ίδια, κ.ο.κ., έτσι ώστε $n = n_1 + n_2 + n_3 + \ldots + n_\kappa$, τότε ο αριθμός μεταθέσεων των n αντικειμένων δίνεται από τον τύπο

$$P(n,n) = \frac{n!}{n_1! n_2! n_3! \ldots n_k!} . \qquad (2.15)$$

Για παράδειγμα, ένας φοιτητής πρόκειται να τοποθετήσει δέκα από τα βιβλία του σε ένα ράφι της βιβλιοθήκης του. Τα βιβλία έχουν διαφορετικά χρώματα. Τέσσερα είναι κόκκινα, τρία είναι μπλέ, ένα είναι μαύρο, και δύο είναι λευκά. Με πόσους διαφορετικούς τρόπους μπορεί να τα τοποθετήσει; Εδώ, έχουμε n =10, n_1=4, n_2=3, n_3=1, n_4=2. Οπότε, από τον τύπο (2.15), η απάντηση είναι

$$\frac{10!}{4!3!1!2!} = 12.600 .$$

Παρατηρήσεις

(A)Εάν όλα τα n αντικείμενα είναι διαφορετικά, τότε n_1=1, n_2=1, n_3=1, …, n_κ=1 και οι μεταθέσεις των n αντικειμένων, από τον τύπο (2.15), είναι n!.

(B) Επίσης, εάν υπάρχουν δύο είδη αντικειμένων, r από τα οποία είναι ίδια και n-r από τα οποία είναι ίδια, τότε έχουμε μία σημαντική περίπτωση του τύπου (2.15):

$$P(n,n)=\frac{n!}{r!(n-r)!}=C(n,r)\ .$$

(Γ) Αυτό σημαίνει ότι ο αριθμός μεταθέσεων στην περίπτωση αυτή είναι ίδιος με τον αριθμό συνδυασμών n αντικειμένων παίρνοντάς τα ανά r. Το αποτέλεσμα αυτό επίσης τεκμηριώνει την προαναφερθείσα ιδιότητα:

$$\binom{n}{r}=\binom{n}{n-r}$$

Παράδειγμα 2.31

Επιθυμούμε να προσδιορίσουμε τον αριθμό των τρόπων με τους οποίους μπορούν να καθίσουν πέντε αγόρια σε μία σειρά από 12 καθίσματα. Το πρόβλημα μπορούμε να το δούμε και σαν την τοποθέτηση 12 αντικειμένων, από τα οποία τα 5 αντιστοιχούν στα αγόρια και διαφέρουν μεταξύ τους ενώ τα υπόλοιπα 7 αντιστοιχούν στα ελεύθερα καθίσματα και είναι ίδια μεταξύ τους. Σύμφωνα με τον τύπο (2.15), ο αριθμός των ζητούμενων τρόπων είναι

$$P(12,12)=\frac{12!}{7!1!1!1!1!1!}=\frac{12!}{7!}=95040\ .$$

2.7 ΠΕΡΙΛΗΨΗ ΚΕΦΑΛΑΙΟΥ

Ένα γεγονός είναι μία συλλογή εκβάσεων ή δειγματικών σημείων σε ένα δειγματικό χώρο, ο οποίος περιλαμβάνει όλα τα δυνατά ενδεχόμενα της έκβασης ενός πειράματος τύχης. Στον δειγματικό χώρο μπορούν να σχηματισθούν γεγονότα με όλους τους δυνατούς συνδυασμούς από ένα ή και περισσότερα δειγματοσημεία, τα οποία απαρτίζουν τον χώρο γεγονότων ή δυναμοσύνολο όπως αλλιώς λέγεται. Οι πράξεις συνόλων που παρουσιάσθηκαν για το συμπληρωματικό, αδύνατο, περιεκτικότητα, τομή, διαφορά και ένωση γεγονότων είναι απαραίτητες για τον συμβολισμό σύνθετων γεγονότων με άλλα απλούστερα ενδεχόμενα.

Βασικά, την πιθανότητα πραγματοποίησης ενός γεγονότος μπορεί κανείς να την προσδιορίσει με δύο προσεγγίσεις υποκειμενικά (διαισθητικά) ή μεθοδολογικά (πιθανοθεωρία). Η πρώτη προσέγγιση στηρίζεται στην εμπειρία του καθενός γύρω από το συγκεκριμένο γεγονός και την κρίση του. Η δεύτερη και σπουδαιότερη προσέγγιση διακρίνεται κυρίως σε δύο μεθόδους, όπως την κλασσική μέθοδο ή μέθοδο συχνότητας και στην μέθοδο έκφρασης σύνθετων γεγονότων με άλγεβρα απλούστερων γεγονότων. Η κλασσική μέθοδος απαιτεί καλή γνώση της συνδυαστικής προκειμένου να απαριθμηθούν όλα τα δυνατά ενδεχόμενα οιουδήποτε πειράματος τύχης και γεγονότος. Βασικές αρχές τέτοιων απαριθμήσεων αναφέρονται στο τέλος του κεφαλαίου. Όμως, στο κεφάλαιο αυτό αναπτύχθηκε κυρίως η μεθοδολογία δια της οποίας εκφράζουμε ένα σύνθετο γεγονός με πράξεις γνωστών γεγονότων και κατόπιν εφαρμόζουμε τους κατάλληλους κανόνες πιθανότητας.

Η θεωρία πιθανοτήτων παρέχει την μεθοδολογία για την εκτίμηση της πιθανότητας διαφόρων τύπων σύνθετων γεγονότων. Οι επιχειρησιακοί κανόνες της πιθανοθεωρίας, οι οποίοι παρέχουν την βάση για τις σχέσεις μεταξύ των πιθανοτήτων διαφορετικών γεγονότων, προέρχονται από τρία βασικά αξιώματα και τον ορισμό μιας συνάρτησης πιθανότητας. Πολλοί μηχανικοί προσεγγίζουν την πιθανότητα γεγονότων, ίσως πολλές φορές διαισθητικά. Ο σκοπός αυτού του κεφαλαίου, είναι να θέση αυστηρή θεμελίωση στις μεθόδους διεξαγωγής πιθανότητας γεγονότων με βάση το πείραμα τύχης και τον δειγματικό χώρο. Πρέπει ακόμη να τονίσουμε την σπουδαιότητα της ποιότητας των διαθέσιμων δεδομένων, επειδή ένας μηχανικός πρέπει να αναγνωρίζει και να εκτιμά ότι η σημαντικότητα και η αξιοπιστία των αποτελεσμάτων εξαρτάται και από τα δεδομένα με τα οποία εκτιμούνται οι πιθανότητες.

Οι έννοιες του κεφαλαίου αυτού αποτελούν τα θεμέλια για την παρακάτω πιθανοθεωρία, εφαρμοσμένη πιθανότητα και εφαρμοσμένη στατιστική.

ΑΣΚΗΣΕΙΣ-ΠΡΟΒΛΗΜΑΤΑ

Ασκήσεις με τις λύσεις τους

1) Ποιες από τις ακόλουθες σχέσεις είναι αληθείς;

$(\alpha)\ (A\cup B)\cap(A\cup C)=A\cup(B\cap C)$ $\quad(\beta)\ (A\cup B)=(A\cap\overline{B})\cup B$

$(\gamma)\ \overline{A}\cap B=A\cup B$ $\quad(\delta)\ \overline{(A\cup B)}\cap C=\overline{A}\cap\overline{B}\cap\overline{C}$ $\quad(\varepsilon)\ (A\cap B)\cap(\overline{B}\cap C)=\varnothing$

Λύση

Αληθείς είναι μόνο οι (α), (β), (ε).

(α) Εφαρμόζοντας πρώτα την επιμεριστική ιδιότητα και μετά την προσεταιριστική έχουμε

$$(A\cup B)\cap(A\cup C)=[A\cap(A\cup C)]\cup[B\cap(A\cup C)]$$
$$=[A]\cup[(B\cap A)\cup(B\cap C)]=[A\cup(B\cap A)]\cup(B\cap C)=A\cup(B\cap C).$$

(β) Παρόμοια με τις γνωστές ιδιότητες των πράξεων συνόλων,

$$(A\cap\overline{B})\cup B=(A\cup B)\cap(B\cup\overline{B})=(A\cup B)\cap S=(A\cup B)$$

(ε) χρησιμοποιώντας την προσεταιριστική ιδιότητα,

$$(A\cap B)\cap(\overline{B}\cap C)=A\cap(B\cap\overline{B})\cap C=A\cap\{\}\cap C=\{\}$$

Ψευδείς οι (γ) και (δ).

(γ) $\overline{A}\cap B\neq A\cup B$

Δότι το αριστερό μέρος της ανίσωσης περιέχει δειγματοσημεία που δεν ανήκουν στο A ενώ το δεξιό μέρος περιέχει δειγματοσημεία που ανήκουν στο A.

(δ) Χρησιμοποιώ τον νόμο De Morgan,

$$\overline{(A\cup B)}\cap C=(\overline{A}\cap\overline{B})\cap C\neq\overline{A}\cap\overline{B}\cap\overline{C}\,.$$

2) Ένα νόμισμα ρίχνεται μέχρι να εμφανισθεί η ένδειξη `*Γράμμα*'. Να ευρεθεί ο Δειγματικός Χώρος του πειράματος.

Λύση

Ο δειγματικός χώρος είναι άπειρος αλλά αριθμήσιμος,

$$S=\{Γ, ΚΓ, ΚΚΓ, ΚΚΚΓ,.....\}.$$

3) Θεωρούμε ότι η σειρά με την οποία αναφέρονται τα τέσσερα στοιχεία *a, b, c,* και *d* παριστάνουν το αποτέλεσμα ενός πειράματος. Αν *A* και *B* είναι τα γεγονότα που ορίζονται ως ακολούθως *: A={α είναι στην πρώτη θέση}, B={b είναι στην δεύτερη θέση}*

(α) Να καταγραφούν όλα τα δειγματοσημεία,

(β) Να καταγραφούν τα στοιχεία που ανήκουν στα γεγονότα $A \cap B$ και $A \cup B$.

Λύση

(α) Όλοι οι δυνατοί τρόποι με τους οποίους μπορώ να τοποθετήσω 4 διαφορετικά πράγματα στη σειρά είναι 4!=24. Άρα ο δειγματικός χώρος περιέχει 24 δειγματοσημεία,

S=

{abcd, abdc, adbc, adcb, acbd, acdb, dabc, dacb, dbac, dbca, dcab, dcba, bacd, badc. bcad, bcda, bdca, bdac, cbad, cbda, cdab, cdba, cabd, cadb}

(β)

A={abcd, abdc, adbc, adcb, acbd, acdb},
B={abcd, abdc, dbac, dbca, cbad, cbda},
$A \cap B$ *={ abcd, abdc},*
$A \cup B$ *={ abcd, abdc, adbc, adcb, acbd, acdb, dbac, dbca, cbad, cbda}.*

4) Ένα κουτί περιέχει *N* λάμπες από τις οποίες οι *κ* (*κ<N*)λάμπες είναι ελαττωματικές.

(α) Ελέγχονται οι λάμπες μία προς μία, μέχρι να βρεθεί μία ελαττωματική. Να αναγραφεί ο δειγματικός χώρος του πειράματος αυτού.

(β) Υποτίθεται ότι οι λάμπες ελέγχονται μία προς μία, μέχρι να βρεθούν όλες οι ελαττωματικές. Να αναγραφεί ο νέος δειγματικός χώρος.

Λύση

Έστω το ενδεχόμενο E_i={ *η i λάμπα είναι ελαττωματική*}

(α) $S=\{ E_1, \overline{E}_1 E_2, \overline{E}_1 \overline{E}_2 E_3, ..., \overline{E}_1 \overline{E}_2 ... \overline{E}_{N-\kappa} E_{N-\kappa+1} \}$

(β) $S=\{ E_1 E_2 ... E_\kappa,\ \overline{E}_1 E_2 ... E_{\kappa+1},\ \overline{E}_1 \overline{E}_2 E_3 ... E_{\kappa+1}, ...,\ \overline{E}_1 \overline{E}_2 ... \overline{E}_{N-\kappa} E_{N-\kappa+1} .... E_N \}$.

5) Κάλπη περιέχει πέντε σφαίρες αριθμημένες με 1 έως 5. Στα πλαίσια τυχαίου πειράματος επιλέγονται τρεις σφαίρες και σημειώνεται η μέγιστη ένδειξη. Έστω *A, B, Γ* τα γεγονότα ότι η μέγιστη ένδειξη είναι 3, 4, 5 αντίστοιχα.

(α) Να αναγραφούν ο δειγματικός χώρος του πειράματος καθώς και τα σύνολα των γεγονότων *Α, Β, Γ*.
(β) Να υπολογισθούν οι πιθανότητες των γεγονότων *Α, Β, Γ*.

Λύση
(α) Θα μπορούσε κανείς να αναφέρει σαν δειγματικό χώρο το σύνολο όλων των δυνατών μέγιστων ενδείξεων του πειράματος S={3, 4, 5}. Τα σημεία όμως του δειγματοχώρου αυτού δεν είναι ισοπίθανα, γι'αυτό μια λεπτομερέστερη καταγραφή των ενδεχομένων του πειράματος θα ήταν προτιμότερη:

$$S=\{(1,2,3), (1,2,4), (2,3,4), (1,3,4), (1,2,5), (1,3,5),\ldots,(3,4,5)\}$$

Προφανώς τα γεγονότα *Α, Β* και *Γ* αντιστοιχούν στα ακόλουθα σύνολα,
$A=\{(1,2,3)\}$,
$B=\{(1,2,4), (1,3,4), (2,3,4)\}$,
$\Gamma=\{(1,2,5), (1,3,5),\ldots, (3,4,5)\}$.
(β) Όλα τα δειγματοσημεία του S είναι όλες οι διαφορετικές τετράδες, δηλαδή οι συνδυασμοί πέντε πραγμάτων ανά τρία,

$$C(5,3)\frac{5!}{(5-3)!\,3!}=\frac{20}{2}=10$$

Αφού όλα τα δειγματοσημεία είναι ισοπίθανα, οι πιθανότητες των παραπάνω γεγονότων υπολογίζονται με βάση την κλασσική μέθοδο,

$$P(A)=\frac{1}{10} \qquad P(B)=\frac{4}{10} \qquad P(\Gamma)=\frac{6}{10}.$$

6) Ένα τετράγωνο χωρίζεται σε 4 ίσες ζώνες. Σε κάθε ζώνη ρίχνουμε τυχαία ένα σημείο. Στη συνέχεια χωρίζουμε το τετράγωνο σε 4 ίσες στήλες. Ποια η πιθανότητα ότι κάθε στήλη περιέχει μόνο ένα σημείο;

•			
		•	
			•
	•		

Λύση
Κάθε ζώνη του πίνακα περιέχει ένα μόνο σημείο. Όλοι οι δυνατοί τρόποι με τους οποίους μπορούμε να τοποθετήσουμε ένα σημείο σε κάθε ζώνη του πίνακα είναι τέσσερις. Άρα, σύμφωνα με τον **κανόνα του γινομένου** (2.17) όλοι οι δυνατοί τρόποι με τους οποίους τοποθετούμε σε κάθε ζώνη ένα σημείο είναι

$N(S)=4\times 4\times 4\times 4=256$

Παρόμοια, όπως φαίνεται και στον πίνακα, όλοι οι δυνατοί τρόποι με τους οποίους μπορεί να εμφανισθεί και σε κάθε στήλη ένα μόνο σημείο είναι σύμφωνα με τον **νόμο του γινομένου**,

$N(A)=4\times 3\times 2\times 1=24.$

Διότι, στην πρώτη ζώνη το σημείο μπορεί να βρίσκεται σε οιαδήποτε από τις 4 στήλες. Στην δεύτερη ζώνη το σημείο πρέπει να κατέχει μία από τις υπόλοιπες 3 στήλες (εκτός της στήλης που κατέχει το σημείο της πρώτης ζώνης). Για τον ίδιο λόγο, στην τρίτη ζώνη το σημείο πρέπει να βρίσκεται σε μία από τις υπόλοιπες 2 στήλες, και τέλος στην τέταρτη ζώνη το σημείο μπορεί να βρίσκεται μόνο στην στήλη που δεν έχει κανένα σημείο.

Έτσι, με την κλασσική μέθοδο η ζητούμενη πιθανότητα υπολογίζεται εύκολα,

$$p=\frac{N(A)}{N(S)}=\frac{24}{256}.$$

Προβλήματα με τις λύσεις τους

1) Καθυστέρηση έργου. Στην κατασκευή ενός δρόμου έχουμε τρεις φάσεις:: *εσκαφή*, *ασφαλτόστρωση*, και *ηλεκτρική σηματοδότηση*.

Να εκφραστούν με συμβολισμό συνόλων τα γεγονότα ότι η κατασκευή του δρόμου τελειώνει

(α) χωρίς καθυστέρηση,

(β) με καθυστέρηση,

(γ) χωρίς καθυστέρηση της εσκαφής και τουλάχιστον μία από τις άλλες δύο φάσεις καθυστερεί.

Λύση

Έστω τα γεγονότα

A=\{*καθυστερεί η εσκαφή*\},

B=\{*καθυστερεί η ασφαλτόστρωση*\},

Γ=\{*καθυστερεί η ηλεκτρική σηματοδότηση*\}.

Με τον νόμο De Morgan ή άλλες πράξεις προκύπτουν τα γεγονότα,

(α) \{ *το έργο συμπληρώνεται χωρίς καθυστέρηση*\}= $\overline{(A\cup B\cup \Gamma)}$,

(β) \{*το έργο συμπληρώνεται με καθυστέρηση*\}= $A\cup B\cup \Gamma$,

(γ) \{*το έργο συμπληρώνεται χωρίς καθυστέρηση της εσκαφής και τουλάχιστον μία από τις άλλες δύο φάσεις καθυστερεί*\}= $\overline{A}\cap(B\cup \Gamma)$.

2) Μειοδοτικός διαγωνισμός.Ένας εργολάβος υποβάλλει προσφορές για τρία έργα Α, Β, και Γ. Να εκφραστούν με συμβολισμό συνόλων όλα τα δυνατά ενδεχόμενα του διαγωνισμού.

Λύση

Έστω τα γεγονότα

Α={*ο εργολάβος αναλαμβάνει το* Α *έργο*}
Β={ *ο εργολάβος αναλαμβάνει το* Β *έργο* }
Γ={ *ο εργολάβος αναλαμβάνει το* Γ *έργο* }
E_κ={*ο εργολάβος αναλαμβάνει κ έργα*}.

Όλα τα δυνατά ενδεχόμενα του διαγωνισμού αντιστοιχούν στην ανάληψη 0, 1, 2 ή 3 έργα, δηλαδή,

$$E_0 = \overline{(A \cup B \cup \Gamma)}$$

$$E_1 = (A \cap \overline{B \cup \Gamma}) \cup (B \cap \overline{A \cup \Gamma}) \cup (\Gamma \cap \overline{A \cup B})$$

$$E_2 = (A \cap B \cap \overline{\Gamma}) \cup (A \cap \overline{B} \cap \Gamma) \cup (\overline{A} \cap B \cap \Gamma)$$

$$E_3 = A \cap B \cap \Gamma$$

3) Τεμαχισμός ράβδου. Μια ράβδος μήκους *α* χωρίζεται τυχαία σε δύο κομμάτια. Υπολογίστε τις εξής πιθανότητες:
(α) Το πρώτο κομμάτι να είναι μεγαλύτερο από *α*/4.
(β) Ένα από τα δύο κομμάτια να είναι μεγαλύτερο από *α*/4.
(γ) Ένα από τα δύο κομμάτια να είναι τουλάχιστον διπλάσιο του άλλου.

Λύση

(α) Το σημείο χωρισμού της ράβδου μπορεί να βρίσκεται ισοπίθανα σε ένα από τα τέσσερα ίσα διαστήματά της, *S*={*α1, α2, α3, α4*}.

Οι ευνοικές περιπτώσεις του γεγονότος που μας ενδιαφέρει είναι ότι το σημείο χωρισμού βρίσκεται κάπου στα τρία τελευταία διαστήματα, $A=\{\alpha_2, \alpha_3, \alpha_4\}$. Με την κλασσική μέθοδο η πιθανότητα είναι

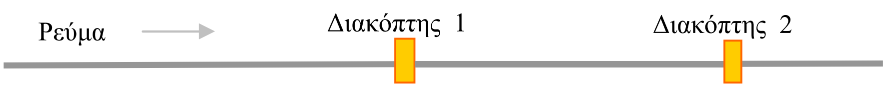
$$p = \frac{N(A)}{N(S)} = \frac{3}{4} = 0{,}75$$

Παρόμοια έχουμε,
(β) πιθανότητα *p*=0,5,
(γ) πιθανότητα *p* =0,6666.

4) Διακόπτες στη σειρά Σένα ηλεκτρικό κύκλωμα έχουμε δύο διακόπτες Δ1 και Δ2 εν σειρά. Ο κάθε διακόπτης ενδέχεται να είναι κλειστός ή ανοικτός. Κάποια στιγμή το κύκλωμα τίθεται σε λειτουργία.

(α) Να περιγραφεί με συμβολισμό συνόλων ο δειγματικός χώρος του πειράματος καθώς και το γεγονός ότι έχουμε διακοπή ρεύματος.

(β) Αν δίνονται οι πιθανότητες $P(\Delta_1)=0{,}2$, $P(\Delta_2)=0{,}3$ και $P(\Delta_1 \cap \Delta_2)=0{,}1$, να υπολογισθεί η πιθανότητα διακοπής ρεύματος.



Λύση

(α) Έστω το ενδεχόμενο Δ_κ={ο κ διακόπτης είναι ανοικτός}. Παρατηρούμε ότι σε αυτό το πρόβλημα ο δειγματικός χώρος σαν σύνολο περιέχει τα δειγματοσημεία του προκύπτουν από το καρτεσιανό γινόμενο των

$$S_1 = \{ \Delta_1 , \overline{\Delta}_2 \} \text{ και } S_2 = \{ \Delta_2 , \overline{\Delta}_2 \},$$

$$S = S_1 \times S_2 = \{(\Delta_1, \Delta_2), \quad (\overline{\Delta}, \Delta_2), \quad (\Delta_1, \overline{\Delta}), \quad (\overline{\Delta}_1, \overline{\Delta}_2)\}.$$

Εναλλακτικά ο δειγματοχώρος S είναι και αυτός ένα γεγονός (σίγουρο γεγονός) και μπορούμε να τον περιγράψουμε με άλγεβρα γεγονότων ως εξής:

$$S = (\Delta_1 \cap \Delta_2) \cup (\overline{\Delta}_1 \cap \Delta_2) \cup (\Delta_1 \cap \overline{\Delta}_2) \cup (\overline{\Delta}_1 \cap \overline{\Delta}_2)$$

Αν A είναι το γεγονός ότι έχουμε διακοπή ρεύματος, τότε

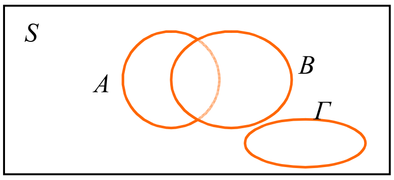
$$A = (\Delta_1 \cap \Delta_2) \cup (\overline{\Delta}_1 \cap \Delta_2) \cup (\Delta_1 \cap \overline{\Delta}_2) \quad \text{ή} \quad A = (\Delta_1 \cup \Delta_2).$$

(β) Στην τελευταία σχέση εφαρμόζουμε το προσθετικό θεώρημα,

$$P(A) = P(\Delta_1 \cup \Delta_2) = P(\Delta_1) + P(\Delta_2) - P(\Delta_1 \cap \Delta_2) = 0{,}2 + 0{,}3 - 0{,}1 = 0{,}4$$

Ασκήσεις προς λύση

1) Τρία γεγονότα ενός δειγματικού χώρου φαίνονται στο ακόλουθο διάγραμμα Venn.



Να αναπαράγεται το διάγραμμα σκιαγραφώντας τα ακόλουθα γεγονότα:

$$\overline{A}, \quad (A\cap B)\cup(A\cap\overline{B}), \quad (A\cap B)\cup\Gamma, \quad \overline{B\cup\Gamma}\,.$$

2) Έστω ότι με A συμβολίζω το γεγονός ότι ένα βάρος ξεπερνά τα 11 κιλά, έστω ότι B συμβολίζει το γεγονός ότι αυτό το βάρος είναι λιγότερο από 15 κιλά, και έστω ότι Γ συμβολίζει το γεγονός ότι το βάρος είναι μεταξύ 8 και 12 κιλά. Περιγράψατε τα ακόλουθα γεγονότα:

$$A\cup B, \quad A\cap B, \quad \overline{A}, \quad A\cup B\cup\Gamma, \quad \overline{A\cup\Gamma}, \quad A\cap B\cap\Gamma, \quad \Gamma\cap\overline{B}, \quad A\cup(B\cap\Gamma).$$

3) Τρία γεγονότα Α, Β, Γ ανήκουν σε κάποιο δειγματοχώρο. Να συμβολίσετε με άλγεβρα γεγονότων τα γεγονότα,

- τουλάχιστον ένα από τα γεγονότα
- ακριβώς ένα από τα γεγονότα
- το πολύ δύο από τα γεγονότα.

4) Δύο σημεία επιλέγονται τυχαία στην περιφέρεια ενός κύκλου ακτίνας r. Να βρεθεί η πιθανότητα ότι η χορδή που τα ενώνει έχει μήκος μικρότερο του r.

5) Πόσες στήλες ΠΡΟΠΟ πρέπει να συμπληρώσει κανείς για να έχει σίγουρη επιτυχία;

6) Σε ένα *πάρτυ* υπάρχουν 20 παιδιά. Ποια η πιθανότητα τουλάχιστον δύο από αυτά έχουν γενέθλια την ίδια μέρα;

7) Ένα κουτί περιέχει 8 ζευγάρια παπούτσια. Επιλέγονται τυχαία έξι παπούτσια. Ποια η πιθανότητα να περιέχεται ένα σωστό ζευγάρι ανάμεσα στα έξι;

8) Ένα τετράγωνο χωρίζεται με οριζόντιες γραμμές σε κ ίσες ζώνες. Σε κάθε μία από αυτές τοποθετείται τυχαία ένα σημείο. Στη συνέχεια φέρονται κάθετες γραμμές έτσι ώστε το τετράγωνο χωρίζεται σε κ ίσες στήλες. Ποία η πιθανότητα ότι κάθε στήλη θα περιέχει μόνο ένα σημείο;

Προβλήματα προς λύση

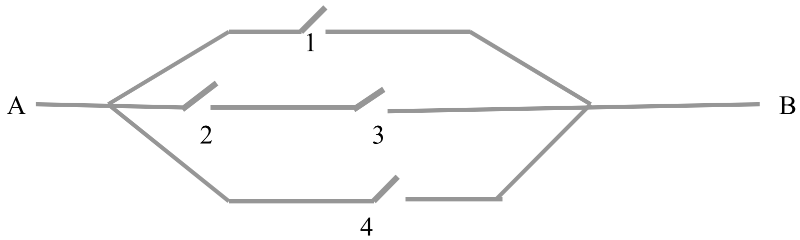
1) Ποιότητα τσιμέντου. Δείγματα τσιμέντου από τρεις βιομηχανίες ελέγχονται για να διαπιστωθεί η ποιότητά τους. Τα δείγματα ταξινομούνται σε δύο κατηγορίες, σχετικά με το αν πληρούν ή όχι τις τεχνικές προδιαγραφές. Τα αποτελέσματα από 100 δείγματα έχουν ως εξής:

	Βιομηχανία 1	*Βιομηχανία* 2	*Βιομηχανία* 3
Δείγματα πληρούν τις προδιαγραφές	18	17	50
Δείγματα δεν πληρούν τις προδιαγραφές	2	3	10

Συμβολίζουμε με A το γεγονός ότι το δείγμα είναι από την βιομηχανία 1, και με B το γεγονός ότι το δείγμα πληροί τις προδιαγραφές. Να προσδιορίσετε τον αριθμό δειγμάτων που αντιστοιχούν στα γεγονότα $\overline{A}\cap B$, $\overline{B}$, και $A\cup B$.

2) Διακοπή συγκοινωνίας. Σε μία σχετικά μικρή απόσταση ενός αυτοκινητόδρομου έχουμε δύο γέφυρες. Η κάθε γέφυρα, λόγω πλημμύρας, ενδέχεται να είναι κλειστή με πιθανότητα 0,3 ή 0,5 αντίστοιχα. Η πιθανότητα και οι δύο γέφυρες να είναι κλειστές είναι 0,2. Κάποια στιγμή ένα αυτοκίνητο ξεκινά να ταξιδέψει πάνω σε αυτόν τον αυτοκινητόδρομο. Να περιγραφεί με συμβολισμό συνόλων ο δειγματικός χώρος του πειράματος (σχετικά με την κατάσταση των γεφυρών ανοικτή-κλειστή) καθώς και η πιθανότητα ότι το αυτοκίνητο δεν θα μπορέσει να ταξιδέψει, λόγω πλημμύρας.

3) Μεταφορά ρεύματος. Στην γραμμή μεταφοράς ρεύματος συμβολίζουμε με Δ_κ το γεγονός ότι ο κ διακόπτης είναι ανοικτός και έστω ότι $P(\Delta_\kappa)$=0, 60.
(α) Να εκφράσετε με συμβολισμό γεγονότων το γεγονός {*τουλάχιστον ένας διακόπτης είναι ανοικτός*}.
(β) Να εκφράσετε με συμβολισμό γεγονότων το γεγονός μεταφοράς ρεύματος από το A στο B..



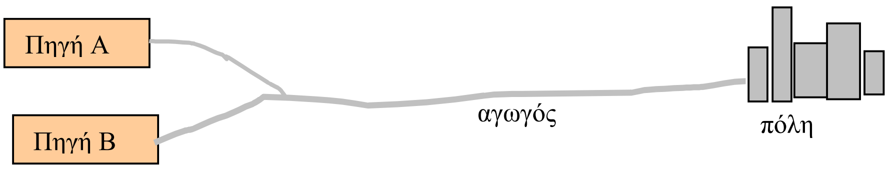
4) Διακοπή ύδρευσης. Το νερό στην πόλη φτάνει από τις δύο πηγές Α και Β, όπως φαίνεται στο παρακάτω σχήμα, δια μέσου του συστήματος των αγωγών 1, 2 και 3. Υποθέτουμε ότι η κάθε πηγή ξεχωριστά ικανοποιεί τις ανάγκες της πόλης. Έστω τα γεγονότα,

E_1={*στερεύει η πηγή* Α},
E_2={*στερεύει η πηγή* Β}
E_3={*διακοπή λόγω βλάβης στον αγωγό* }.

Το γεγονός

E={*υπάρχει διακοπή νερού στην πόλη*}

να το συμβολίσετε με άλγεβρα των γεγονότων E_1, E_2 και E_3.



5) Μειοδοτικός διαγωνισμός. Τρεις κατασκευαστικές εταιρείες, Α, Β και Γ κάνουν συχνά προσφορές σε μειοδοτικούς δαγωνισμούς για κατασκευή έργων. Έστω ότι *Α* είναι το ενδεχόμενο η Α εταιρία να κερδίσει ένα έργο, *Β* το ενδεχόμενο η εταιρία Β να κερδίσει ένα έργο, και *Γ* το ενδεχόμενο η εταιρία Γ να κερδίσει ένα έργο. Να δειχθούν τα διαγράμματα Venn που βασίζονται στις εξής περιπτώσεις:

- Οι εταιρίες Α, Β και Γ είναι οι μόνες που προσφέρονται για το ίδιο έργο.
- Οι εταιρίες Α, Β και Γ κάνουν προσφορά για το ίδιο έργο, για το οποίο υπάρχουν προσφορές και από άλλες εταιρίες.
- Οι εταιρείες Α, Β και Γ κάνουν προσφορά για διαφορετικά έργα.

6) Κίνηση σωματιδίου. Ένα σωματίδιο κινείται πίσω μπρος ασταμάτητα στο διάστημα (0, 1). Κατά την διάρκεια της πορείας του από το 0 στο 1, ο νόμος κίνησής του δίδεται με την σχέση $s_0=t^2$, όπου s_0 η εκάστοτε απόσταση του από το 0 και *t* ο χρόνος που διαρέει από την στιγμή που εγκατέλειψε το σημείο 0.

Κατά την διάρκεια της πορείας του από το 1 στο 0 ο νόμος κίνησής του δίδεται με την σχέση $s_1=t^3$, όπου s_1 η απόσταση του από το σημείο 1 και t ο χρόνος που διαρρέει από την στιγμή που εγκατέλειψε το σημείο 1. Έστω ότι κάποια τυχαία χρονική στιγμή παρατηρείτε τη θέση του σωματιδίου. Είναι πιθανότερο να το βρείτε στο διάστημα (0, 0,5) ή (0,5, 1).

7) Αντίσταση βάρους. Ένας βαρύς άνδρας 120 κιλών κινείται πάνω σε μία δοκό, όπως φαίνεται στο σύστημα του παρακάτω σχήματος. Στο σύστημα αυτό προκαλούνται αντιδράσεις στα ακραία σημεία της δοκού Α και Β αντίστοιχα, έτσι ώστε $R_A + R_B = 120$ κιλά.

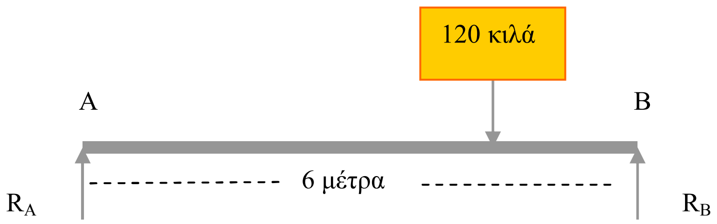
(α) Να ευρεθεί ο δειγματοχώρος όλων των δυνατών ζευγών (R_A, R_B) και να δοθεί μία γραφική του παράσταση με συντεταγμένες R_A και R_B.

Επίσης, στην περίπτωση που το βάρος του άνδρα είναι τυχαίο από 80 μέχρι 120 κιλά,

(β) να ευρεθεί ο νέος δειγματοχώρος του πειράματος, και να δοθεί η γραφική του παράσταση,

(γ) να σχεδιαστεί στον δειγματοχώρο το γεγονός,

$E1$={ *ο άνδρας βρίσκεται σε μία απόσταση* 2 *με* 4 *μέτρα από το σημείο* Α},



8) Καθίζηση σε γέφυρα. Οι πιθανότητες καθίζησης των στηριγμάτων a b και c σε μία γέφυρα είναι $P(A)=0.30$, $P(B)=0.20$ και $P(C)=0.15$ αντίστοιχα. Ακόμη είναι γνωστές οι πιθανότητες $P(A\cap B)=0.08$, $P(A\cap C)=0.05$, $P(B\cap C)=0.05$ και $P(A\cap B\cap C)=0.03$. Να ευρεθούν οι πιθανότητες:

- μόνο ένα στήριγμα θα υποστεί καθίζηση,
- ακριβώς δύο στηρίγματα θα υποστούν καθίζηση,
- θα υπάρξει τουλάχιστον μία καθίζηση.

3

Δεσμευμένη Πιθανότητα, Ολική Πιθανότητα-Θεώρημα Bayes, Ανεξαρτησία και συναφείς έννοιες

ΠΕΡΙΓΡΑΦΗ ΚΕΦΑΛΑΙΟΥ

Συχνά όταν εκτιμούμε την πιθανότητα ενός γεγονότος έχουμε κάποια πληροφορία για το αποτέλεσμα του πειράματος τύχης. Η διάθεση τέτοιας πληροφορίας έχει σαν συνέπεια τον περιορισμό του αρχικού δειγματοχώρου σε ένα από τα υποσύνολά του. Δηλαδή, η πληροφορία υποδεικνύει ότι το αποτέλεσμα του πειράματος σίγουρα βρίσκεται σε ένα τμήμα του δειγματικού χώρου παρά κάπου αλλού στο υπόλοιπο μέρος του. Συχνά, η πιθανότητα ενός γεγονότος είναι διαφορετική όταν έχουμε κάποια πληροφορία. Για παράδειγμα, η πιθανότητα ότι μία κάρτα που τραβάμε από μία τράπουλα να είναι *Άσσος* είναι μεγαλύτερη αν γνωρίζουμε ότι η κάρτα δεν είναι *Φιγούρα*. Ακόμη, αν A και B είναι δύο αμοιβαία αποκλειόμενα γεγονότα, ορισμένα στον ίδιο δειγματικό χώρο με πιθανότητες διάφορες του μηδενός, $P(A)\neq 0$, $P(B)\neq 0$, η πιθανότητα να συμβαίνει το A, όταν έχουμε την πληροφορία ότι συμβαίνει το B κατά την εκτέλεση του πειράματος, είναι προφανώς μηδέν. Αυτά τα παραδείγματα δείχνουν την ανάγκη να εισάγουμε μία λίγο διαφορετική επίσης σπουδαία έννοια της πιθανότητας.

3.1 ΥΠΟ ΣΥΝΘΗΚΗ Ή ΔΕΣΜΕΥΜΕΝΗ ΠΙΘΑΝΟΤΗΤΑ

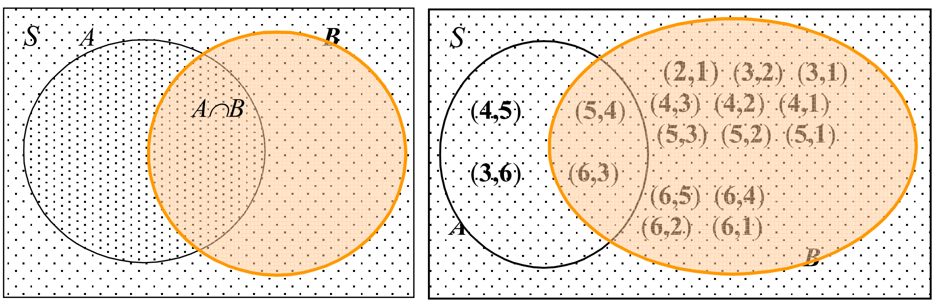
Αν A και B είναι δύο γεγονότα ορισμένα στον ίδιο δειγματικό χώρο, δηλώνουμε με $P(A|B)$ την *υπό συνθήκη ή δεσμευμένη πιθανότητα* του γεγονότος A, δεδομένου ότι έχει συμβεί το B. Για να προσδιορίσουμε την πιθανότητα $P(A|B)$, βασικά υπολογίζουμε την πιθανότητα $P(A)$ αναφερόμενοι στον συρρικνωμένο δειγματόχωρο B, και όχι στον αρχικό δειγματικό χώρο S.

Όπως φαίνεται και στο διάγραμμα Venn, όταν υπολογίζουμε την πιθανότητα $P(A)$ θέτουμε την ερώτηση πόσο πιθανό είναι το αποτέλεσμα του πειράματος να βρίσκεται στην περιοχή του A, γνωρίζοντας ότι σίγουρα βρίσκεται μέσα στο S. Και όταν υπολογίζουμε την υπό συνθήκη πιθανότητα $P(A|B)$ θέτουμε την ερώτηση πόσο πιθανόν είναι το αποτέλεσμα του πειράματος να βρίσκεται μέσα στο A, γνωρίζοντας ότι βρίσκεται σίγουρα μέσα στο B. Ή γνωρίζοντας ότι συμβαίνει το γεγονός B το αποτέλεσμα του πειράματος να είναι ένα δειγματοσημείο που ανήκει και στην τομή $A\cap B$. Με άλλα λόγια στην υπό συνθήκη πιθανότητα ο δειγματικός χώρος έχει ελαττωθεί από το S στο B.

Για να καταλάβουμε καλύτερα το παραπάνω σκεπτικό υπολογισμού της υπό συνθήκη πιθανότητας θα αναφέρουμε αναλυτικότερα ένα παράδειγμα και στη συνέχεια θα διατυπώσουμε τον ορισμό.

Παράδειγμα 3.1

Στην ρίψη δύο ζαριών όλα τα δυνατά ενδεχόμενα του δειγματικού χώρου S είναι ισοπίθανα και είναι 36, $S=\{(x_1,x_2)|\ x_1=1,2,\dots,6,\ x_2=1,2,\dots,6\}$. Θεωρούμε



α) συρρίκνωση δειγματικού χώρου S στο B β) παράδειγμα 3.1

Σχήμα 3.1. Υπό συνθήκη πιθανότητα, A|B, διαγράμματα Venn.

δύο γεγονότα A και B, όπου A είναι το γεγονός ότι το άθροισμα των δύο ενδείξεων είναι 9, $A=\{(x_1,x_2)|\ x_1+x_2=9\}$. και B είναι το γεγονός ότι η πρώτη ένδειξη είναι μεγαλύτερη από την δεύτερη, $B=\{(x_1,x_2)|\ x_1>x_2\}$. Προφανώς τα δειγματοσημεία του A είναι τέσσερα, του B είναι δεκαπέντε (Σχήμα 3.1, (β)) και σύμφωνα με κλασσική θεωρία έχουμε τις πιθανότητες,

$$P(A)=\frac{N(A)}{N(S)}=\frac{4}{36},\quad P(B)=\frac{N(B)}{N(S)}=\frac{15}{36}.$$

Ακόμη, όπως φαίνεται και στο σχεδιάγραμμα Venn, το γεγονός ότι το άθροισμα των ενδείξεων είναι εννέα και η πρώτη ένδειξη μεγαλύτερη της δεύτερης, δηλαδή η τομή $A\cap B$, περιέχει δύο δειγματοσημεία. Συνεπώς, η πιθανότητα ότι σε μία ρίψη θα συμβαίνουν και τα δύο γεγονότα, πάλι με την κλασσική θεωρία, είναι

$$P(A\cap B)=\frac{N(A\cap B)}{N(S)}=\frac{2}{36}$$

Τελικά, ας υπολογίσουμε την υπό συνθήκη πιθανότητα, $P(A|B)$. Εάν γνωρίζουμε ότι μετά την ρίψη των δύο ζαριών η πρώτη ένδειξη είναι μεγαλύτερη της δεύτερης, μπορούμε να θεωρήσουμε ότι όλα τα δυνατά ενδεχόμενα πλέον είναι 15, δηλαδή ο αρχικός δειγματοχώρος συρρικνώνεται στο σύνολο του B. Έτσι, το γεγονός A συμβαίνει μόνον αν το αποτέλεσμα του περάματος βρίσκεται στην περιοχή της τομής $A\cap B$, δηλαδή αν το αποτέλεσμα είναι (6,3) ή (5,4) όπως φαίνεται και στο σχήμα. Με την κλασσική θεωρία υπολογίζουμε την υπό-συνθήκη πιθανότητα

$$P(A|B)=\frac{N(A\cap B)}{N(B)}=\frac{2}{15}.$$ ■

Εάν στην τελευταία σχέση διαιρέσουμε τον αριθμητή και τον παρανομαστή με τον αριθμό δειγματοσημείων του S, η υπό συνθήκη πιθανότητα γενικεύεται ως εξής

$$P(A|B)=\frac{N(A\cap B)}{N(B)}=\frac{\frac{N(A\cap B)}{N(S)}}{\frac{N(B)}{N(S)}}=\frac{P(A\cap B)}{P(B)} \tag{3.1}$$

Η τελευταία σχέση δεν ισχύει μόνο για το συγκεκριμένο παράδειγμα, είναι μία γενική σχέση και μας υποδεικνύει πως μπορούμε να διατυπώσουμε τον ορισμό της υπό συνθήκη πιθανότητας.

Ορισμός. Εάν A και B είναι δύο γεγονότα τα οποία σχετίζονται με το ίδιο πείραμα τύχης, ως **υπό συνθήκη πιθανότητα $P(A|B)$** καλείται η πιθανότητα να συμβεί το A δεδομένου ότι συνέβη το B και ορίζεται από την σχέση,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad \text{για} \quad P(B) > 0 . \tag{3.2}$$

Παρατηρήσεις

(Α) Η σχέση (3.2) δεν είναι κάποιο θεώρημα ούτε κάποιο αξίωμα. Με τη σχέση αυτή απλά εισάγουμε διαισθητικά τον συμβολισμό της υπό συνθήκη πιθανότητας και διατυπώνουμε τον ορισμό της υπό συνθήκη πιθανότητας έπειτα από το παραπάνω σκεπτικό.

(Β) Για κάθε γεγονός A και B του ίδιου δειγματοχώρου, οι πιθανότητες $P(A|B)$ *και* $P(A)$ αναφέρονται στην πραγματοποίηση του γεγονότος A και είναι δύο διαφορετικές τιμές. Πότε οι πιθανότητες $P(A|B)$ *και* $P(A)$ έχουν ίσες τιμές αναφέρεται στην μεθεπόμενη παράγραφο.

(Γ) Όλα τα θεωρήματα ή κανόνες που αναφέρονται στην πιθανοθεωρία ισχύουν και διά την υπό συνθήκη πιθανότητα. Για παράδειγμα, αν A, B και Γ είναι τρία γεγονότα του ίδιου δειγματικού χώρου, τότε το προσθετικό θεώρημα 2.5 ισχύει και για την υπό συνθήκη πιθανότητα,

$$P[(AUB)|\Gamma]=P(A|\Gamma)+P(B|\Gamma)-P[(A\cap B)|\Gamma], \text{ κ.ο.κ..}$$

Εάν $B=S$, μπορούμε να έχουμε

$$P(A|S) = P(A\cap S)\,|\,P(S) = P(A)/P(\ S\ )= P(A)$$

επειδή $P(S)=1$. και $A\cap S=A$. Λογικά, έτσι πρέπει να είναι η πιθανότητα αυτή, διότι η πληροφορία ότι συμβαίνει το γεγονός S είναι μόνο η πληροφορία ότι εκτελέσθηκε το πείραμα τύχης. ■

Τέλος, ο υπολογισμός της υπό συνθήκη πιθανότητας $\boldsymbol{P(A|B)}$ μπορεί να γίνει με δύο τρόπους:

- Άμεσα, υπολογίζοντας την πιθανότητα του γεγονότος A με την κλασσική θεωρία συρρικνώνοντας τον δειγματικό χώρο από S σε B.
- Χρησιμοποιώντας τον ορισμό (3.2), όπου οι πιθανότητες των γεγονότων $A\cap B$ και B υπολογίζονται σε σχέση με τον αρχικό δειγματικό χώρο S.

Παράδειγμα 3.2

Στο παράδειγμα 2.22, όπου σε μία πόλη τρεις από τις εφημερίδες που κυκλοφορούν είναι η Α, Β, και Γ. Για ένα άτομο που επιλέγεται τυχαία από την πόλη,

(α) είναι γνωστό ότι διαβάζει τουλάχιστον μία από τις τρεις εφημερίδες, να υπολογισθεί η πιθανότητα ότι διαβάζει την εφημερίδα Α,
(β) είναι γνωστό ότι διαβάζει την Α και την Β, να υπολογισθεί η πιθανότητα ότι διαβάζει και την εφημερίδα Γ.

Απάντηση

Έστω τα γεγονότα

A={*το άτομο διαβάζει την εφημερίδα* Α}
B={*το άτομο διαβάζει την εφημερίδα* Β}
Γ={*το άτομο διαβάζει την εφημερίδα* Γ}.

Είναι γνωστές οι πιθανότητες, $P(A)=0{,}20$, $P(B)=0{,}16$, $P(\Gamma)=0{,}14$, $P(A\cap B)=0{,}08$, $P(A\cap B\cap \Gamma)=0{,}02$, $P(A\cup B\cup \Gamma)=0{,}35$.

(α) Η πιθανότητα σε αυτή την περίπτωση είναι δεσμευμένη πιθανότητα,

$$P(A|(A\cup B\cup \Gamma))=\frac{P(A\cap(A\cup B\cup \Gamma))}{P(A\cup B\cup \Gamma)}=\frac{P(A)}{P(A\cup B\cup \Gamma)}=\frac{0{,}20}{0{,}35}=0{,}57$$

Δηλαδή, η πιθανότητα ότι κάποιος διαβάζει την Α, με την πληροφορία ότι διαβάζει τουλάχιστον μία από τις τρεις εφημερίδες, είναι μεγαλύτερη από την αρχική πιθανότητα $P(A)$.

(β)Εδώ, ο αναγνώστης σίγουρα ανήκει στην τομή των γεγονότων A και B, $A\cap B$, ζητάμε την πιθανότητα να ανήκει και στην τομή των τριών γεγονότων, $A\cap B\cap \Gamma$,

$$P[(A\cap B\cap \Gamma)|(A\cap B)]=\frac{P[(A\cap B\cap \Gamma)\cap(A\cap B)]}{P(A\cap B)}=\frac{P(A\cap B\cap \Gamma)}{P(A\cap B)}=\frac{0{,}02}{0{,}08}=0{,}25.$$

Παρατηρούμε ότι η πιθανότητα ένα τυχαίο άτομο να διαβάζει και τις τρεις εφημερίδες είναι $P(A\cap B\cap \Gamma)=0{,}02$, με την πληροφορία ότι είναι αναγνώστης των εφημερίδων Α και Β η πιθανότητα αυτή αυξάνει σε 0,25.

Η Πιθανότητα της Τομής και το Πολλαπλασιαστικό Θεώρημα

Η ποιο σπουδαία συνέπεια του ορισμού της υπό συνθήκη πιθανότητας είναι ότι από την σχέση (3.2) μπορεί να προκύψει η εξίσωση

$$\boldsymbol{P(A\cap B)=P(B)P(A|B)}$$

ή ισοδύναμα

$$\boldsymbol{P(A\cap B)=P(A)P(B|A)} \qquad (3.3)$$

Η σχέση αυτή είναι γνωστή και σαν **Πολλαπλασιαστικό Θεώρημα**. Στο εξής, συχνά εφαρμόζεται αυτό το θεώρημα για να υπολογίσουμε την πιθανότητα της ταυτόχρονης πραγματοποίησης δύο γεγονότων A και B, $A \cap B$, τα οποία ανήκουν στον ίδιο δειγματοχώρο.

Παράδειγμα 3.3

Ένα μεγάλο κατάστημα έχει 100 φορητούς ηλεκτρονικούς υπολογιστές. Κάποιοι από αυτούς τους φορητούς είναι τύπου *Pendium* (P) ενώ οι άλλοι είναι *Celeron* (C). Και κάποιοι από τους φορητούς είναι καινούργιοι (K) ενώ οι άλλοι είναι μεταχειρισμένοι (M). Ο Πίνακας 3.1 δείχνει αναλυτικότερα τον αριθμό των φορητών σε κάθε κατηγορία.

Πίνακας 3.1

	καινούργιοι (K)	μεταχειρισμένοι (M)	**Συνολικά**
Pentium (***P***)	30	10	40
Celeron (***C***)	40	20	60
Συνολικά	70	30	100

Ένας πελάτης εισέρχεται στο κατάστημα και επιλέγει τυχαία έναν φορητό.
(α) Ποια είναι η πιθανότητα ότι ο φορητός είναι ένας καινούργιος *Celeron;*
(β) Διαπιστώνει ότι ο φορητός είναι καινούργιος. Ποια η πιθανότητα ότι ο φορητός είναι και *Celeron;*
Απάντηση:
Το πείραμα τύχης συνίσταται στην τυχαία επιλογή ενός φορητού. Προφανώς, ο δειγματοχώρος του πειράματος περιέχει εκατό ισοπίθανα δειγματοσημεία. Ορίζουμε τα γεγονότα:

- K={ *ο φορητός είναι καινούργιος*}, περιέχει εβδομήντα δειγματο σημεία, $N(K)$=70,
- C={*ο φορητός είναι Celeron*}, περιέχει εξήντα δειγματοσημεία, $N(C)$=60,
- $K \cap C$={*ο φορητός είναι ένας καινούργιος Celeron*}, περιέχει σαράντα δειγματοσημεία, $N(K \cap C)$=40.

(α) Σύμφωνα με την **κλασσική θεωρία** η πιθανότητα του γεγονότος $K \cap C$ είναι ο λόγος 40 προς 100,

$$P(K \cap C) = \frac{N(K \cap C)}{N(S)} = \frac{40}{100} = 0.40$$

Επίσης, μπορούμε να εφαρμόσουμε και τον πολλαπλασιαστικό κανόνα, όπως παρακάτω.
(β) Επιθυμούμε να υπολογίσουμε την πιθανότητα $P(C|K)$. Περιορίζοντας τον αρχικό δειγματοχώρο (100 φορητοί) στην στήλη των καινούργιων φορητών (70 φορητοί), σύμφωνα με την **κλασσική θεωρία** έχουμε

$$P(C|K)=40/70=4/7.$$

Χρησιμοποιώντας τον **ορισμό της υπό συνθήκη πιθανότητας** έχουμε πάλι το ίδιο αποτέλεσμα,

$$P(C|K)=\frac{P(K\cap C)}{P(K)}=\frac{0.4}{0.7}=\frac{4}{7}$$

Η απάντηση στο ερώτημα α) δίνεται και με τον **πολλαπλασιαστικό κανόνα.** Αν λάβουμε υπόψη το τελευταίο αποτέλεσμα και ότι $P(K)=70/100$, έχουμε

$$P(K\cap C)=P(K)P(C|K)=\frac{7}{10}\cdot\frac{4}{7}=\frac{4}{10}=0.4 \quad .$$

■

Το Πολλαπλασιαστικό θεώρημα, σχέση 3.3, μπορεί να επεκταθεί και για περισσότερα από δύο γεγονότα όπως ακολουθεί:

$$\begin{aligned}&\boldsymbol{P(A_1\cap A_2\cap A_3\ldots\cap A_n)}\\&\boldsymbol{=P(A_1)P(A_2|A_1)P(A_3|(A_1\cap A_2))\ldots P(A_n|(A_1\cap A_2\cap A_3\ldots\cap A_{n-1}))}\end{aligned} \quad (3.4)$$

Η απόδειξη του θεωρήματος είναι περισσότερο αλγεβρική παρά πιθανολογική, Βιβλίο Θεωρίας Πιθανοτήτων, Λ. Καμαρινόπουλου.

Παράδειγμα 3.4

Σε ένα μεγάλο κιβώτιο 100 ηλεκτρικών εξαρτημάτων υπάρχουν 20 ελαττωματικά. Παίρνουμε τυχαία τρία εξαρτήματα και τα ελέγχουμε. Ποια η πιθανότητα ότι και τα τρία εξαρτήματα είναι ελαττωματικά;

Απάντηση:

Ορίζουμε τα γεγονότα,

A_κ=*{το κ εξάρτημα είναι ελαττωματικό}, για κ=1,2,3.*

Επίσης, ορίζουμε και τα υπό συνθήκη γεγονότα,

$(A_2 \,|\, A_1)$={*το δεύτερο εξάρτημα είναι ελαττωματικό* ***δεδομένου*** *ότι το πρώτο ήταν ελαττωματικό*}

$(A_3 \,|\, A_1 \cap A_2)$={*το τρίτο εξάρτημα είναι ελαττωματικό* ***δεδομένου*** *ότι τα δύο πρώτα ήταν ελαττωματικά*}.

Επειδή η επιλογή των εξαρτημάτων γίνεται χωρίς επανάθεση, προφανώς με την κλασσική μέθοδο προκύπτουν οι πιθανότητες:

$$P(A_1) = \frac{20}{100}, \qquad P(A_2|A_1) = \frac{19}{99}, \qquad P(A_3|A_1 \cap A_2) = \frac{18}{98}.$$

Το γεγονός ότι και τα τρία γεγονότα είναι ελαττωματικά ορίζεται με την τομή των τριών γεγονότων, A_1, A_2 και A_3, $A_1 \cap A_2 \cap A_3$, και η πιθανότητά του υπολογίζεται σύμφωνα με τον πολλαπλασιαστικό κανόνα,

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2|A_1)P(A_3|(A_1 \cap A_2)) = \frac{20}{100} \cdot \frac{19}{99} \cdot \frac{18}{98}.$$

3.2 ΟΛΙΚΗ ΠΙΘΑΝΟΤΗΤΑ

Στην προηγούμενη παράγραφο χρησιμοποιήσαμε την έννοια της υπό συνθήκη πιθανότητας προκειμένου να εκτιμήσουμε την πιθανότητα της ταυτόχρονης πραγματοποίησης δύο γεγονότων. Μπορούμε να εφαρμόσουμε αυτή την έννοια και με έναν άλλο τρόπο για να προσδιορίσουμε την πιθανότητα ενός σύνθετου γεγονότος B, το οποίο είναι η ένωση των τομών του με άλλα γεγονότα (βλέπε Σχήμα 3.3). Χρειαζόμαστε όμως τον ακόλουθο ορισμό.

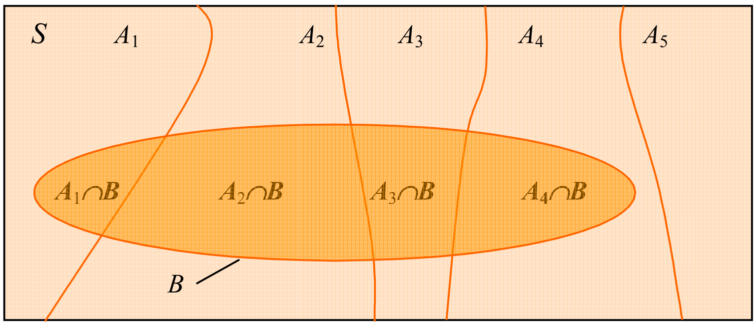
Ορισμός: Τα γεγονότα $A_1, A_2, \ldots, A_\kappa$ αποτελούν **διαμέριση** του δειγματικού χώρου S, αν

1. είναι ασυμβίβαστα, $\boldsymbol{A_i} \cap \boldsymbol{A_j} = \varnothing$ για κάθε $i \neq j$,
2. η ένωσή τους καταλαμβάνει όλο τον δειγματικό χώρο, $\bigcup_{i=1}^{k} \boldsymbol{A_i} = \boldsymbol{S}$,
3. $\boldsymbol{P(A_i)>0}$ για κάθε i.

Με άλλα λόγια όταν διεξάγεται ένα πείραμα τύχης, πραγματοποιείται *ένα και μόνον ένα* από τα γεγονότα A_i . Για παράδειγμα στην ρίψη του ζαριού τα γεγονότα A_1={1,2,3}, A_2={4}, A_3={5,6} αποτελούν διαμέριση του δειγματικού χώρου S={1,2,3,4,5,6}, ενώ τα γεγονότα $Γ_1$={1,2,3,4} και $Γ_2$={3,4,5,6} δεν αποτελούν.

Ας υποθέσουμε ότι B είναι ένα γεγονός σε κάποιο πείραμα τύχης με δειγματικό χώρο S και ότι τα γεγονότα $A_1, A_2, \ldots, A_κ$ αποτελούν διαμέριση του S. Στο Σχήμα 3.3 το διάγραμμα Venn παριστάνει μία τέτοια διαμέριση με $κ$=5, καθώς και την τομή του B με γεγονότα της διαμέρισης. Το γεγονός B αποτελείται από την ένωση όλων των τομών του με τα γεγονότα της διαμέρισης,

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup \ldots \cup (A_κ \cap B) \tag{3.5}$$



Σχήμα 3.3. Διάγραμμα Venn, τομή του γεγονότος B με μία διαμέριση $A_ι$ του δειγματικού χώρου S.

Όπως φαίνεται και στο Σχήμα 3.3 μερικές τομές $A_ι \cap B$ ενδέχεται να είναι το κενό σύνολο αλλά αυτό δεν δημιουργεί κανένα πρόβλημα στην παραπάνω διαμέριση του B. Το ποιο σημαντικό είναι ότι όλα τα γεγονότα $A_ι \cap B$ είναι μεταξύ τους αμοιβαία αποκλειόμενα και αυτό μας οδηγεί στην διατύπωση και απόδειξη του θεωρήματος της **Ολικής Πιθανότητας**.

Θεώρημα Ολικής Πιθανότητας

Εάν τα γεγονότα $A_1, A_2, \ldots, A_κ$ αποτελούν διαμέριση του δειγματικού χώρου S, η πιθανότητα ενός γεγονότος B, οριζόμενο στον ίδιο δειγματικό χώρο, ορίζεται από την σχέση

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \ldots + P(B|A_κ)P(A_κ)\,. \tag{3.6}$$

Απόδειξη: Από την σχέση (3.5) η πιθανότητα του B είναι,

$$P(B) = P[(A_1 \cap B) \cup (A_2 \cap B) \cup .... \cup (A_k \cap B)] \,.$$

Εφαρμόζοντας το προσθετικό θεώρημα 2.6 και λαμβάνοντας υπόψη ότι τα γεγονότα $A_i \cap B$ είναι αμοιβαία αποκλειόμενα, έχουμε

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + ... + P(A_k \cap B) \,.$$

Αλλά κάθε όρος $P(A_i \cap B)$ με χρήση τον πολλαπλασιαστικό κανόνα μπορεί να γραφεί

$$P(A_i \cap B) = P(B|A_i)P(A_i) \,,$$

και με αντικατάσταση στην προηγούμενη σχέση επιτυγχάνουμε το επιθυμητό αποτέλεσμα. ■

Το θεώρημα ολικής πιθανότητας ή η **σχέση της ολικής πιθανότητας** όπως συνηθίζεται να λέγεται είναι πάρα πολύ χρήσιμη. Διότι, όταν είναι δύσκολο να εκτιμήσουμε άμεσα την πιθανότητα $P(B)$, η πληροφορία πραγματοποίησης του A_i μπορεί να οδηγεί στην πιθανότητα $P(B | A_i)$ και εφαρμόζοντας το κανόνα 3.6 υπολογίζουμε ευκολότερα την πιθανότητα του B.

Παράδειγμα 3.5

Θεωρούμε και πάλι (παράδειγμα 2.30) το κιβώτιο με 20 ελαττωματικά και 80 καλά εξαρτήματα από το οποίο επιλέγουμε τυχαία δύο εξαρτήματα *χωρίς επανάθεση*. Ορίζουμε τα γεγονότα,

A_1=*{το πρώτο εξάρτημα είναι ελαττωματικό}*,
A_2=*{το πρώτο εξάρτημα είναι καλό}*,
B=*{το δεύτερο εξάρτημα είναι καλό}*.

Προφανώς, τα γεγονότα A_1 και A_2 αποτελούν διαμέριση του δειγματικού χώρου S. Έτσι, μπορούμε να υπολογίσουμε την πιθανότητα του B από τον κανόνα της ολικής πιθανότητας,

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) \,.$$

Παρόμοια με τους υπολογισμούς στο παράδειγμα 2.30 βρίσκουμε ότι

$$P(B) = \frac{80}{99} \cdot \frac{20}{100} + \frac{79}{99} \cdot \frac{80}{100} = \frac{80}{100}$$

Το αποτέλεσμα εδώ τυχαίνει να είναι το ίδιο έστω και εάν ακόμη τα εξαρτήματα τα επιλέγαμε τυχαία αλλά με επανάθεση.

Παράδειγμα 3.6

Μία βιομηχανία αυτοκινήτων προμηθεύεται τους κινητήρες από τρεις διαφορετικές εταιρείες. Από το σύνολο των κινητήρων της βιομηχανίας το 50% προέρχεται από την πρώτη εταιρεία Α1, ενώ η δεύτερη Α2 και τρίτη Α3 προμηθεύουν από 25% έκαστη. Είναι επίσης γνωστό ότι από την πρώτη εταιρεία όλοι οι κινητήρες είναι καλοί, ενώ το ποσοστό των ελαττωματικών για την Α2 και Α3 είναι 2% και 6%, αντίστοιχα. Ποια είναι η πιθανότητα ότι ένα αυτοκίνητο της βιομηχανίας φοράει ελαττωματικό κινητήρα;

Ας ορίσουμε τα εξής γεγονότα:

A_1=*{ο κινητήρας προέρχεται από την πρώτη εταιρεία},*
A_2=*{ο κινητήρας προέρχεται από την δεύτερη εταιρεία},*
A_3=*{ο κινητήρας προέρχεται από την τρίτη εταιρεία},*
B=*{ο κινητήρας είναι ελαττωματικός}.*

Από τα δεδομένα του προβλήματος είναι γνωστές οι πιθανότητες,

$$P(A_1)=0{,}50, \quad P(A_2)=0{,}25, \quad P(A_3)=0{,}25,$$
$$P(B|A_1)=0{,}0, \quad P(B|A_2)=0{,}02, \quad P(B|A_3)=0{,}06.$$

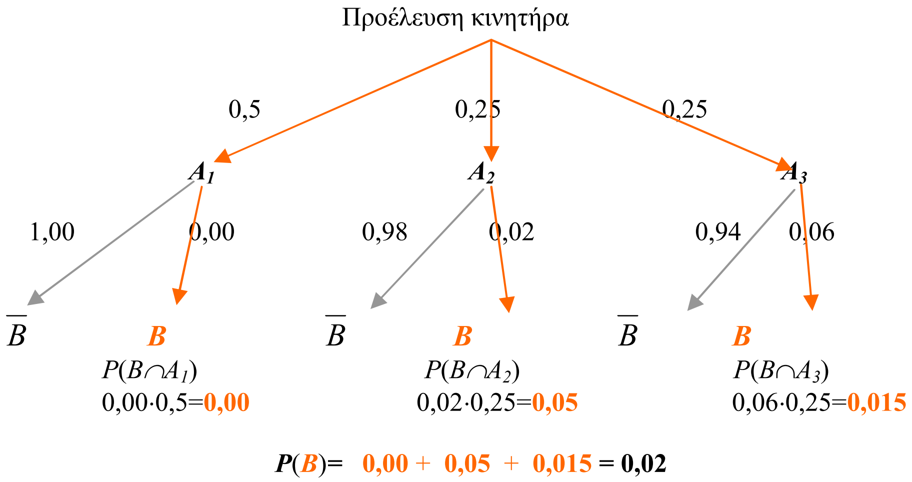
Το πείραμα τύχης ισοδυναμεί με την τυχαία επιλογή κινητήρα στην βιομηχανία και έλεγχος αυτού. Αναφερόμενοι στην προέλευση του κινητήρα τα τρία ενδεχόμενα A_1, A_2, A_3 αποτελούν προφανώς διαμέριση του δειγματικού χώρου. Εφαρμόζοντας τον παραπάνω κανόνα, μπορούμε να γράψουμε για την πιθανότητα του B

$$P(B)=P(B|A_1)P(A_1)+P(B|A_2)P(A_2)+P(B|A_3)P(A_3).$$

Αντικαθιστώντας τις παραπάνω τιμές στην τελευταία εξίσωση, η ζητούμενη πιθανότητα είναι

$$P(B)=0{,}50 \cdot 0 + 0{,}25 \cdot 0{,}02 + 0{,}25 \cdot 0.06.=0{,}02$$

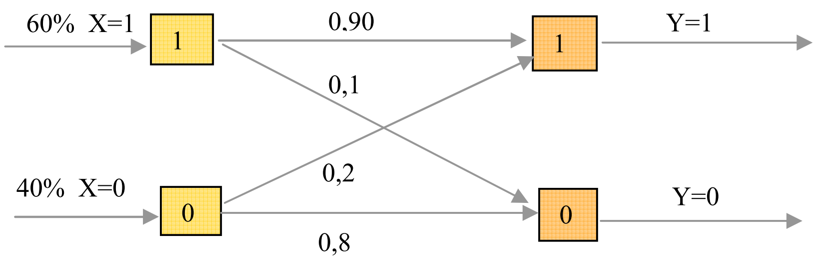
Η λύση του παραδείγματος 3.7 επίσης διευκολύνεται και με διάγραμμα δένδρου, βλέπε Σχήμα 3.4.



Σχήμα 3.4. Διάγραμμα δένδρου για το Παράδειγμα 3.6

Παράδειγμα 3.7

Ένα σύστημα τηλεπικοινωνίας δέχεται στην είσοδο του δυαδικό σήμα X που παίρνει τιμές 0 και 1. Το σήμα εξόδου Y είναι επίσης δυαδικό με τιμές 0 και 1. Λόγου θορύβου στο κανάλι μπορεί κατά την έξοδο να υπάρξει αναστροφή του σήματος από 0 σε 1 και αντίστροφα. Το 60% των σημάτων εισόδου είναι 1 και το 40% είναι 0. Η πιθανότητα αναστροφής του σήματος κατά την έξοδο είναι 0,10 από 1 σε 0, και 0,2 από 0 σε 1.
Ποια είναι τα αντίστοιχα ποσοστά των σημάτων 0 και 1 στην έξοδο του καναλιού:



Σχήμα 3.5 Σύστημα τηλεπικοινωνίας

Ορίζω σαν {X=0} και {X=1} τα γεγονότα ότι στην είσοδο καταφθάνουν σήματα 0 και 1 αντίστοιχα. Παρόμοια ορίζω και για την έξοδο τα γεγονότα {Y=0} και {Y=1}. Η πιθανολογική συμπεριφορά του καναλιού περιγράφεται και στο Σχήμα 3.5.

Με συμβολισμούς παρόμοιους του θεωρήματος ολικής πιθανότητας έχουμε A_1={X=0}, A_2={X=1} και B={Y=1}. Τα γεγονότα A_1 και A_2 αποτελούν διαμέριση του δειγματικού χώρου σχετικά με τα σήματα εισόδου. Εφαρμόζοντας τον τύπο της ολικής πιθανότητας,

$$\begin{aligned}P(Y=1) &= P(\{Y=1\} | \{X=1\})P(\{X=1\}) + P(\{Y=1\} | \{X=0\})P(\{X=0\}) \\ &= 0{,}9 \cdot 0{,}60 + 0{,}20 \cdot 0{,}40 = 0{,}62\end{aligned}$$

Συμβολίζοντας με B={Y=0}και ακολουθώντας την ίδια διαδικασία έχουμε για το ποσοστό του σήματος εξόδου 0,

$$\begin{aligned}P(Y=0) &= P(\{Y=0\} | \{X=1\})P(\{X=1\}) + P(\{Y=0\} | \{X=0\})P(\{X=0\}) \\ &= 0{,}1 \cdot 0{,}60 + 0{,}80 \cdot 0{,}40 = 0{,}38\end{aligned}$$

3.3 ΘΕΩΡΗΜΑ BAYES

Αναφερόμενοι στο παραπάνω παράδειγμα 3.6, θα μπορούσαν να τεθούν και άλλες ενδιαφέρουσες ερωτήσεις. Να ελέγξουμε τον κινητήρα ενός αυτοκινήτου και διαπιστώσουμε ότι είναι ελαττωματικός, ποια η πιθανότητα ότι ο κινητήρας προέρχεται από την πρώτη εταιρεία;

Χρησιμοποιώντας τους ίδιους συμβολισμούς, όπως στο πιο πάνω **παράδειγμα 3.7**, θέλουμε να υπολογίσουμε την υπό συνθήκη πιθανότητα

$$P(A_1 | B).$$

Έχοντας την πληροφορία ότι ο κινητήρας είναι ελαττωματικός η βεβαιότητα να προέρχεται από την πρώτη εταιρεία μηδενίζεται. Σίγουρα, ο κινητήρας αφού είναι ελαττωματικός προέρχεται από την A2 ή A3 εταιρεία. Ακόμη, λογικά σκεπτόμενοι, επειδή η A3 έχει τριπλάσιο ποσοστό ελαττωματικών κινητήρων απ' ότι η A2 , ένας ελαττωματικός κινητήρας προέρχεται από την A3 με τριπλάσια βεβαιότητα. Πράγματι, το σκεπτικό αυτό επιβεβαιώνεται εφαρμόζοντας τον κανόνα της υπό συνθήκη πιθανότητας και χρησιμοποιώντας τα παραπάνω δεδομένα ,

$$P(A_1|B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{P(A_1)P(B|A_1)}{P(B)} = \frac{0.5 \cdot 0.00}{0.02} = 0.00$$

$$P(A_3|B) = \frac{P(A_3 \cap B)}{P(B)} = \frac{P(A_3)P(B|A_3)}{P(B)} = \frac{0.25 \cdot 0.06}{0.02} = 0{,}75$$

Τελικά, παρατηρούμε ότι με την πληροφορία ότι κινητήρας είναι ελαττωματικός, η πιθανότητα να προέρχεται από κάποια από τις τρεις εταιρείες έχει αλλάξει σημαντικά.

Επίσης, στην παραπάνω εξίσωση η πιθανότητα της τομής, $P(A_1 \cap B)$, ορίσθηκε με τον πολλαπλασιαστικό κανόνα και την επιλεγόμενη σειρά, η οποία οδηγεί σε γνωστές πιθανότητες, $P(A_1)$, $P(B|A_1)$.

Γενικεύοντας το τελευταίο αποτέλεσμα οδηγούμαστε στο γνωστό θεώρημα του **Bayes**.

Θεώρημα Bayes

Αν τα γεγονότα $\boldsymbol{A_1, A_2, \ldots, A_\kappa}$ αποτελούν **διαμέριση** του δειγματικού χώρου S, και $\boldsymbol{B}$ είναι ένα γεγονός το οποίο σχετίζεται με τον ίδιο χώρο S, η υπό συνθήκη πιθανότητα $\boldsymbol{P(A_i|B)}$ ορίζεται από την σχέση

$$\boldsymbol{P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(B)}} \qquad i = 1,2,\ldots,\kappa \tag{3.7}$$

Η απόδειξη του **θεωρήματος Bayes** είναι εύκολη αν ξεκινήσει κανείς με την υπό-συνθήκη πιθανότητα,

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)P(B|A_i)}{P(B)} \qquad i = 1,2,\ldots,\kappa \,.$$

Η παραπάνω φόρμουλα είναι γνωστή και σαν ***εκ των υστέρων πιθανότητα***. Δεδομένου ότι έχει πραγματοποιηθεί το γεγονός B, η φόρμουλα αυτή προσδιορίζει την πιθανότητα πραγματοποίησης ενός γεγονότος της διαμέρισης, A_i. Επειδή, ένα και μόνον ένα από τα κ γεγονότα της διαμέρισης πραγματοποιείται, η πιθανότητα αυτή εκφράζει την πιθανότητα ότι η πραγματοποίηση του B είχε σαν ''αίτιο'' την πραγματοποίηση του A_i.

Μία εικονογραφημένη παρουσίαση της **ολικής πιθανότητας** και του θεωρήματος **Bayes** μπορεί να δοθεί με τα **διαγράμματα δέντρου**, τα οποία είναι χρήσιμα σε πολλές περιπτώσεις για την ανάλυση τέτοιων προβλημάτων.

Παράδειγμα 3.8

Συνεχίζοντας το παράδειγμα 3.7, έστω ότι ένα σήμα στην έξοδό του είναι 1. Ποια η πιθανότητα ότι είναι σωστή η μετάδοσή του;

Η ζητούμενη πιθανότητα μπορεί να διατυπωθεί και σαν την πιθανότητα του υπό συνθήκη γεγονότος

$$\{X=1\}\big|\{Y=1\}$$

ότι, δηλαδή, δεδομένου ότι το σήμα εξόδου είναι 1, ήταν και στην είσοδό του 1. Χρησιμοποιώντας τους ίδιους συμβολισμούς, A_1={X=0}, A_2={X=1} και B={Y=1}, η ζητούμενη πιθανότητα μπορεί να υπολογισθεί εφαρμόζοντας το θεώρημα του Bayes,

$$P(\{X=1\}\big|\{Y=1\})=\frac{P(\{Y=1\}\big|\{X=1\})P(\{X=1\})}{P(\{Y=1\})}=\frac{0{,}9\cdot 0{,}6}{0{,}62}=\frac{0{,}54}{0{,}62}=0{,}87$$

Παρατηρούμε ότι η εκ των *υστέρων* πιθανότητα για ένα σήμα να είναι στην είσοδό του 1, δεν είναι 0,60 όπως είναι η εκ των *προτέρων* πιθανότητα. Με την πληροφορία ότι το σήμα κατά την έξοδό του είναι 1 αυξάνει και την πιθανότητα ότι και στην είσοδό του ήταν 1, αφού με πολύ μεγάλη πιθανότητα (0,90) τα σήματα μεταδίδονται σωστά.

Παράδειγμα 3.9

Σε μία παραγωγής μεταλλικών ράβδων, το 20% δεν είναι καλής ποιότητας διότι παρουσιάζουν πολύ μικρές ρωγμές, οι οποίες βέβαια δεν είναι ευδιάκριτες.. Για τον σκοπό αυτόν κάθε ράβδος μετά την παραγωγή ελέγχεται από ένα μηχάνημα με ακτίνες leizer, και αν ο έλεγχος είναι θετικός η ράβδος καταστρέφεται. Η αξιοπιστία όμως του ελέγχου

- είναι 90%, όταν η ράβδος δεν έχει ρωγμές ,
- είναι 80%, όταν η ράβδος είναι ελαττωματική.

(α) Ποιο ποσοστό της παραγωγής ράβδων καταστρέφεται;

(β)Ποιο ποσοστό από τις ράβδους που καταστρέφονται είναι στην πραγματικότητα καλές;

Απάντηση

Μπορούμε να ορίσουμε τα γεγονότα,

A_1={η ράβδος είναι ελαττωματική, έχει ρωγμές},
A_2={η ράβδος είναι καλή, δεν έχει ρωγμές},
B={ο έλεγχος είναι θετικός,, η ράβδος καταστρέφεται},

Σύμφωνα με τα δεδομένα του προβλήματος ισχύουν οι ακόλουθες πιθανότητες,

$$P(A_1)=0{,}20,\quad P(A_2)=0{,}80,\quad P(B\big|A_1)=0{,}80,\quad P(B\big|A_2)=0{,}10,$$

.

(α) Το ποσοστό των ράβδων που καταστρέφεται είναι η πιθανότητα $P(B)$. Επειδή τα γεγονότα A_1 και A_2 αποτελούν διαμέριση του δειγματικού χώρου, σχετικά με την ποιότητα της ράβδου, εφαρμόζω τον προσθετικό κανόνα,

$$\begin{aligned} P(B) &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) \\ &= 0{,}80 \cdot 0{,}20 + 0{,}10 \cdot 0{,}80 = 0{,}24 \end{aligned}$$

Έτσι, από όλη την παραγωγή το 24% καταστρέφεται.
(β) Εδώ, ζητάμε την πιθανότητα η ράβδος να είναι καλή A_2 με την πληροφορία ότι έλεγχος είναι θετικός Β, $P(A_2| B)$. Προφανώς, θα εφαρμόσω το θεώρημα του Bayes,

$$P(A_2|B) = \frac{P(B|A_2)P(A_2)}{P(B)} = \frac{0{,}10 \cdot 0{,}80}{0{,}24} = \frac{1}{3}$$

Άρα, από τις ράβδους που καταστρέφονται το ένα τρίτον δεν έχουν ρωγμές, και άδικα καταστρέφονται. Για αυτό προτείνεται και δεύτερος έλεγχος προκειμένου να περιορίσουμε το ποσοστό των καλών ράβδων που καταστρέφονται, βλέπε στην παράγραφο προβλήματα με τις λύσεις τους, πρόβλημα 11.

3.4 ΣΤΑΤΙΣΤΙΚΗ ΑΝΕΞΑΡΤΗΣΙΑ ΚΑΙ ΣΥΝΑΦΕΙΣ ΕΝΝΟΙΕΣ

Έχουμε διαπιστώσει μέχρι τώρα ότι η πραγματοποίηση ενός γεγονότος B σε ένα δειγματικό χώρο S, ενδεχομένως να επιδρά στην πιθανότητα πραγματοποίησης ενός άλλου γεγονότος A που σχετίζεται με τον ίδιο δειγματικό χώρο. Για παράδειγμα, αν τα γεγονότα A και B είναι ασυμβίβαστα, $A \cap B = \varnothing$, τότε η πραγματοποίηση του ενός μηδενίζει την πιθανότητα πραγματοποίησης του άλλου στην ίδια εκτέλεση του πειράματος, $P(A|B)=0$. Δηλαδή, η πραγματοποίηση του B αποκλείει την πραγματοποίηση του A, γι' αυτό ονομάζονται και αμοιβαία αποκλειόμενα γεγονότα. Άλλες φορές πάλι, όταν το γεγονός B είναι υποσύνολο του A, $B \subset A$, η πραγματοποίηση του A είναι σίγουρη όταν συμβαίνει το B, $P(A|B)=1$.

3.4.1 Στατιστικά Ανεξάρτητα Γεγονότα

Στις επάνω περιπτώσεις η πιθανότητα πραγματοποίησης του A εξαρτάται από την πραγματοποίηση του B. Όμως, σε άλλες περιπτώσεις η πιθανότητα πραγματοποίησης του A δεν εξαρτάται από την πραγματοποίηση B. Για παράδειγμα, στην ρίψη δύο ζαριών, ενός πράσινου και ενός κόκκινου, αν A και B είναι τα γεγονότα

A={*το πράσινο ζάρι φέρνει άσσο*},
B={*το κόκκινο ζάρι φέρνει ζυγό αριθμό*},

λογικά, τα δύο γεγονότα είναι ασυσχέτιστα. Γνωρίζοντας ότι έχει πραγματοποιηθεί το A δεν παρέχει καμία πληροφορία σχετικά με την πιθανότητα πραγματοποίησης του B. Αυτό επιβεβαιώνεται και με την εκτίμηση των υπό συνθήκη πιθανοτήτων τους. Ο δειγματικός χώρος του πειράματος έχει 36 ισοπίθανα δειγματοσημεία, οι ευνοικές περιπτώσεις για το A είναι 6 , για το B είναι 18 και για την τομή τους, $A\cap B$, είναι 3, και από την κλασσική θεωρία οι πιθανότητες τους είναι $P(A)=6/36=1/6$, $P(B)=18/36=1/2$, $P(A\cap B)=3/36=1/12$.

Οπότε, $P(A|B)=\dfrac{P(A\cap B)}{P(B)}=\dfrac{1/12}{1/2}=\dfrac{1}{6}$.

Έτσι, βρίσκουμε ότι η υπό συνθήκη πιθανότητα $P(A|B)$ είναι ίδια με την πιθανότητα $P(A)$. Παρόμοια, υπολογίζουμε και την υπό συνθήκη πιθανότητα του B,

$$P(B|A)=\frac{P(A\cap B)}{P(A)}=\frac{1/12}{1/6}=\frac{1}{2}=P(B).$$

Παρατηρούμε ότι η πληροφορία πραγματοποίησης του γεγονότος A δεν έχει καμία επίδραση στην πιθανότητα πραγματοποίησης του γεγονότος B και το αντίστροφο. Με άλλα λόγια, η παραπάνω συζήτηση οδηγεί στο ότι δύο γεγονότα A και B είναι ανεξάρτητα εάν και μόνον εάν $P(A|B)= P(A)$ και $P(B|A)= P(B)$.

Θεωρώντας ότι οι παραπάνω υπό συνθήκη πιθανότητες είναι ίδιες με τις χωρίς συνθήκη πιθανότητες, η πιθανότητα της τομής, $P(B\cap A)$ είναι μία ειδική περίπτωση του πολλαπλασιαστικού κανόνα και μπορεί να γραφεί

$$P(A\cap B)=P(B)P(A|B)=P(B)P(A)$$

ή ισοδύναμα

$$P(A\cap B)=P(A)P(B|A)=P(A)P(B).$$

Έτσι, στην περίπτωση που οι πιθανότητες $P(B)$ και $P(A)$ είναι διάφορες του μηδενός η πραγματοποίηση του ενός γεγονότος δεν επιδρά στην πιθανότητα πραγματοποίησης του άλλου εάν και μόνον εάν $P(B\cap A)= P(B)P(A)$. Ποιο μεθοδικά μπορούμε να διατυπώσουμε τον ακόλουθο ορισμό της ***ανεξαρτησίας***.

Ορισμός: Δύο γεγονότα, $\boldsymbol{A}$ και $\boldsymbol{B}$, οριζόμενα στον ίδιο δειγματικό χώρο με πιθανότητες $P(B)\neq 0$ και $P(A)\neq 0$, είναι **ανεξάρτητα** εάν και μόνον εάν

$$\boldsymbol{P(A\cap B)=P(B)P(A)}. \tag{3.8}$$

Παρατηρήσεις

(Α) Αυτός ο ορισμός είναι ισοδύναμος με τον παραπάνω, όπου δύο γεγονότα είναι **ανεξάρτητα** εάν και μόνον εάν $\boldsymbol{P(A|B)=P(A)}$ και $\boldsymbol{P(B|A)=P(B)}$, δηλαδή, η πραγματοποίηση του ενός δεν επιδρά στην πιθανότητα πραγματοποίησης του άλλου.

(Β) Η ανεξαρτησία δύο γεγονότων είναι περισσότερο πιθανολογικός ορισμός, δηλαδή, δύο γεγονότα είναι ανεξάρτητα αν ισχύει η εξίσωση πιθανότητας (3.8).

(Γ) Τονίζουμε ακόμη μία φορά ότι και οι δύο ορισμοί ισχύουν μόνον όταν οι πιθανότητες των γεγονότων είναι διάφορες του μηδενός, $P(B)\neq 0$ και $P(A)\neq 0$. Αυτό επιβεβαιώνεται και από το παράδειγμα που ακολουθεί.

Παράδειγμα 3.10

Ράβδος σπάει τυχαία σε κάποιο σημείο της. Συμβολίζουμε με A το γεγονός ότι η ράβδος σπάει σε ένα συγκεκριμένο σημείο α, το οποίο ανήκει στο πρώτο ήμισυ μήκος της, και με B το γεγονός ότι η ράβδος σπάει σε οποιοδήποτε σημείο στο δεύτερο ήμισυ μήκος της. Είναι τα γεγονότα A και B ανεξάρτητα;

Η πιθανότητα του A σύμφωνα με την κλασσική θεωρία είναι ο λόγος ένα προς άπειρον, $P(A)=1/\infty=0$. Διότι, όλα τα δειγματοσημεία του S είναι ισοπίθανα και άπειρα το πλήθος, ο δε πληθάριθμος του συνόλου A είναι ένα. Παρόμοια υπολογίζεται και η πιθανότητα του B, $P(B)=1/2$, βλέπε παράδειγμα 2.19. Τα γεγονότα A και B είναι προφανώς αμοιβαίως αποκλειόμενα και συνεπώς $P(A\cap B)=0$.

Παρατηρούμε ότι παρόλα που ισχύει η συνθήκη ανεξαρτησίας 3.8

$$P(A\cap B)=0$$

$$P(A)P(B)=0\cdot 1/2=0$$

τα γεγονότα A και B δεν είναι ανεξάρτητα, διότι όπως αναφέρθηκε και στην αρχή της παραγράφου, $P(B|A)=0\neq P(B)$. Αυτό επιβεβαιώνει ότι για να είναι δύο γεγονότα ανεξάρτητα δεν αρκεί μόνο να ισχύει η συνθήκη 3.8, αλλά πρέπει και οι πιθανότητες των γεγονότων να είναι διάφορες του μηδέν, $P(A)\neq 0$ και $P(B)\neq 0$.

■

Παράδειγμα 3.11

Στη ρίψη ενός άσπρου και ενός κόκκινου ζαριού, συμβολίζουμε με α και κ την ένδειξη του άσπρου και κόκκινου ζαριού αντίστοιχα. Αν στην ρίψη των δύο ζαριών ορίσουμε τα γεγονότα $A=\{\alpha\geq 5\}$ και $B=\{\kappa\leq 4\}$, τα γεγονότα A και B

πρέπει να είναι ανεξάρτητα διότι η πραγματοποίηση του ενός δεν επιδρά στην πιθανότητα πραγματοποίησης του άλλου. Αυτό μπορεί να τεκμηριωθεί με τον ορισμό 2.15.

Παρατηρούμε ότι ο δειγματικός χώρος του πειράματος αποτελείται από 36 ισοπίθανα δειγματοσημεία, όλοι οι δυνατοί συνδυασμοί των ενδείξεων του άσπρου και κόκκινου ζαριού, Τα δειγματοσημεία του A είναι 12 τα δειγματοσημεία του B είναι 24 και συνεπώς οι πιθανότητες πραγματοποίησης των γεγονότων αυτών σε μια ρίψη των ζαριών είναι, $P(A)=12/36=1/3$, $P(B)=24/36=2/3$. Ακόμη, τα δειγματοσημεία της τομής των A και B, $A\cap B$,, είναι 8 και έτσι έχουμε

$$P(A\cap B)=\frac{8}{36}=\frac{2}{9}$$

$$P(A)P(B)=\frac{1}{3}\cdot\frac{2}{3}=\frac{2}{9}$$

Σύμφωνα με τον ορισμό ανεξαρτησίας τα γεγονότα A και B είναι πράγματι ανεξάρτητα.

Η **επέκταση του ορισμού ανεξαρτησίας** για τρία η περισσότερα γεγονότα δεν είναι άμεσα προφανής, και απαιτείται μία επί πλέον προσθήκη. Γενικά, ***n*** γεγονότα $A_1, A_2, \ldots, A_n$ είναι **ανεξάρτητα** εάν και μόνον εάν είναι ανά δύο ανεξάρτητα, είναι ανά τρία ανεξάρτητα, είναι ανά τέσσερα ανεξάρτητα και ούτω καθεξής.

Ορισμός: *N* γεγονότα $A_1\ A_2, \ldots\ A_n$ οριζόμενα στον ίδιο δειγματικό χώρο με πιθανότητες $P(A_\kappa)\neq 0$ για $\kappa=1,2,\ldots,$ n, είναι **ανεξάρτητα** εάν και μόνον εάν είναι ανά **δύο** ανεξάρτητα, είναι ανά **τρία** ανεξάρτητα, **…,** είναι ανά **n-1** ανεξάρτητα και είναι ανά **n ανεξάρτητα**.

Ειδικά, τρία γεγονότα A, B και Γ είναι ανεξάρτητα αν και μόνον αν

$$\begin{aligned}&P(A\cap B)=P(A)P(B),\\&P(A\cap\Gamma)=P(A)P(\Gamma),\\&P(B\cap\Gamma)=P(B)P(\Gamma),\\&P(A\cap B\cap\Gamma)=P(A)P(B)P(\Gamma).\end{aligned}\tag{3.9}$$

3.4.2 Ανεξάρτητα και Αμοιβαίως Αποκλειόμενα Γεγονότα

Είναι πολύ σημαντικό να διασαφηνίσουμε τώρα την διαφορά μεταξύ των αμοιβαίως αποκλειόμενων γεγονότων με τα ανεξάρτητα γεγονότα.

- Θέλουμε να ξέρουμε εάν τα γεγονότα είναι **αμοιβαίως αποκλειόμενα** όταν σχετιζόμαστε με την πιθανότητα της **ένωσης** γεγονότων ή την πιθανότητα ότι τουλάχιστον ένα από τα γεγονότα θα συμβεί.

- Ενδιαφερόμαστε για την **ανεξαρτησία** όταν θεωρούμε την **τομή** των γεγονότων ή την πιθανότητα ότι όλα τα γεγονότα θα συμβούν.

Επί πλέον, εάν δύο γεγονότα είναι **αμοιβαίως αποκλειόμενα**, τότε τα δύο γεγονότα είναι **στατιστικώς εξαρτημένα**. Το αντίστροφο δεν ισχύει πάντα, εάν δύο γεγονότα είναι εξαρτημένα, δεν είναι υποχρεωτικά αμοιβαίως αποκλειόμενα.

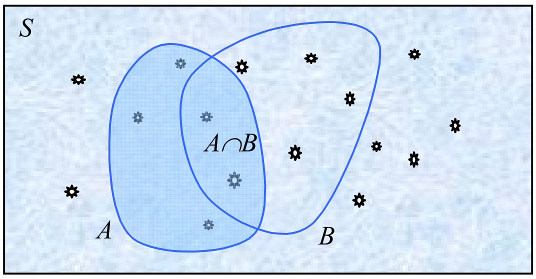
Σαν μία γενική διευκρίνιση αυτών των εννοιών, θεωρούμε τα δύο γεγονότα A και B, όπως στο Σχήμα 3.7, όπου τα 15 αποτελέσματα του δειγματικού χώρου S υποτίθεται ότι είναι ισοπίθανα. Βεβαίως, τα A και B δεν είναι αμοιβαίως αποκλειόμενα, εφόσον $A \cap B \neq Ø$. Οπωσδήποτε, αυτά είναι. ανεξάρτητα, αφού

$$P(A) = P(A \mid B) = 5/15,$$

$$P(B) = P(B \mid A) = 6/15,$$

$$P(A \cap B) = P(A)P(B) = 2/15.$$

Είναι ενδιαφέρον να αναφέρουμε ότι μία μικρή αλλαγή στον ορισμό των συγκεκριμένων γεγονότων A και B – ας πούμε, μετακινούμε ένα δειγματοσημείο από το σύνολο $\overline{A} \cap \overline{B}$, στο σύνολο $A \cap \overline{B}$ – θα καθιστούσε τα A και B στο εξής όχι ανεξάρτητα.



Σχήμα 3.7. **Διάγραμμα Venn επεξηγώντας την ανεξαρτησία**

Τελικά, εάν A και B είναι ανεξάρτητα και εάν $P(A) \neq 0$ και $P(B) \neq 0$, τότε A και B δεν μπορούν να είναι αμοιβαίως αποκλειόμενα, αφού, για δύο ανεξάρτητα γεγονότα, η πιθανότητα ότι και τα δύο γεγονότα θα πραγματοποιηθούν είναι το γινόμενο των πιθανοτήτων τους. Δεδομένου ότι $P(A) > 0$ και $P(B) > 0$, το γινόμενο πρέπει να είναι μεγαλύτερο του μηδενός. Εάν A και B είναι αμοιβαίως αποκλειόμενα, τότε $A \cap B \neq Ø$, το οποίο έχει μηδέν πιθανότητα. Οπότε, η τομή δεν πρέπει να είναι το κενό σύνολο, το οποίο σημαίνει ότι τα γεγονότα δεν είναι αμοιβαίως αποκλειόμενα.

Για να κλείσουμε την παράγραφο της ανεξαρτησίας, ας παρουσιάσουμε τα ακόλουθα γενικά αποτελέσματα (τα οποία ο αναγνώστης αν επιθυμεί μπορεί να τα αποδείξει σαν μία εξάσκηση):

1. Για κάθε γεγονός *A* στο *S, το A* και το αδύνατο γεγονός *Ø,* καθώς και το *A* και το *S* είναι ανεξάρτητα.
2. Εάν *A* και *B* οριζόμενα στον ίδιο δειγματικό χώρο *S* είναι ανεξάρτητα, το ίδιο είναι και τα
   α. $\overline{A}$ και $\overline{B}$,
   β. $\overline{A}$ και B,

3.5 ΠΕΡΙΛΗΨΗ ΚΕΦΑΛΑΙΟΥ

Η θεωρία πιθανοτήτων παρέχει την μεθοδολογία για την εκτίμηση της πιθανότητας διαφόρων τύπων σύνθετων γεγονότων. Οι επιχειρησιακοί κανόνες της πιθανοθεωρίας, οι οποίοι παρέχουν την βάση για τις σχέσεις μεταξύ των πιθανοτήτων διαφορετικών γεγονότων, προέρχονται από τρία βασικά αξιώματα και τον ορισμό μιας συνάρτησης πιθανότητας. Οι κανόνες αυτοί οδηγούν σε νόμους που θεμελιώνουν την υπό συνθήκη πιθανότητα, στατιστική ανεξαρτησία, προσθετικό και πολλαπλασιαστικό κανόνα, το θεώρημα της ολικής πιθανότητας και το θεώρημα του Bayes.

Πολλοί μηχανικοί προσεγγίζουν την πιθανότητα γεγονότων, ίσως πολλές φορές διαισθητικά. Ο σκοπός αυτού του κεφαλαίου, το οποίο κλείνει κυρίως με την ολική πιθανότητα και το θεώρημα του Bayes, είναι να θέση αυστηρή θεμελίωση στις μεθόδους διεξαγωγής πιθανότητας γεγονότων με βάση το πείραμα τύχης και τον δειγματικό χώρο. Πρέπει ακόμη να τονίσουμε την σπουδαιότητα της ποιότητας των διαθέσιμων δεδομένων, επειδή ένας μηχανικός πρέπει να αναγνωρίζει και να εκτιμά ότι η σημαντικότητα και η αξιοπιστία των αποτελεσμάτων εξαρτάται και από τα δεδομένα με τα οποία εκτιμούνται οι πιθανότητες.

Στο κεφάλαιο αυτό όλοι οι κανόνες που αναπτύχθηκαν για την εκτίμηση πιθανότητας γεγονότων επεξηγήθηκαν και πραγματικά προβλήματα στον χώρο του μηχανικού. Οι έννοιες του κεφαλαίου αυτού αποτελούν τα θεμέλια για την παρακάτω πιθανοθεωρία, εφαρμοσμένη πιθανότητα και εφαρμοσμένη στατιστική.

ΑΣΚΗΣΕΙΣ-ΠΡΟΒΛΗΜΑΤΑ

Ασκήσεις με τις λύσεις τους

1) Τρεις παίκτες Α, Β, Γ ρίχνουν ένα νόμισμα με τη σειρά που τους αναφέραμε. Ο πρώτος που θα φέρει *κεφαλή* κερδίζει το παιχνίδι. Να συμβολισθεί με άλγεβρα συνόλων το γεγονός A, ότι ο παίκτης Α θα κερδίσει το παιχνίδι. Να ευρεθεί η πιθανότητα του γεγονότος A.

Λύση

Έστω τα γεγονότα

A_ν={*ο* Α *παίκτης φέρει στη ν ρίψη 'κεφαλή'*},
B_ν={*ο* Β *παίκτης φέρει στη ν ρίψη 'κεφαλή'*},
Γ_ν={*ο* Γ *παίκτης φέρει στη ν ρίψη 'κεφαλή'*}.

Το γεγονός ότι κερδίζει ο Α παίκτης μπορεί να πραγματοποιηθεί στην πρώτη ρίψη ή στην τέταρτη ή στην έβδομη κ.τ.λ., ως ακολούθως:

$$A = A_1 \cup (\overline{A}_1 \cap \overline{B}_2 \cap \overline{\Gamma}_3 \cap A_4) \cup .....$$
$$P(A) = P(A_1) + P(\overline{A}_1 \cap \overline{B}_2 \cap \overline{\Gamma}_3 \cap A_4) + ...$$
$$P(A) = \frac{1}{2} + P(\overline{A})P(\overline{B}_2)P(\overline{\Gamma}_3)P(A_4) + ...$$
$$P(A) = \frac{1}{2} + \frac{1}{2^4} + \frac{1}{2^7} + ... = \frac{1/2}{1-1/8} = \frac{8}{14}$$

διότι είναι γεωμετρική πρόοδος με λόγο 1/8 και αρχικό όρο 1/2.

2) Δύο ακέραιοι αριθμοί διαλέγονται τυχαία και υπολογίζεται το γινόμενό τους. Ποια η πιθανότητα να λήγει σε τέσσερα; Πώς αλλάζει η παραπάνω πιθανότητα αν γνωρίζαμε ότι ο δεύτερος είναι μεγαλύτερος του πέντε;

Λύση

Σε αυτό το πείραμα τύχης παρατηρούμε το τελευταίο ψηφίο του γινομένου δύο ακεραίων και ο δειγματικός χώρος είναι S={0, 1,2 ,3 ,4 ,5 ,6 ,7 ,8 ,9}, όπου τα δειγματοσημεία του **δεν** είναι ισοπίθανα.

Όμως, μπορούμε να θεωρήσουμε ότι ένα ισοδύναμο με το παραπάνω πείραμα τύχης συνίσταται στην τυχαία επιλογή δύο ακεραίων και καταγραφή των τελευταίων ψηφίων τους. Ο δειγματικός χώρος του πειράματος αυτού περιέχει 100 ισοπίθανα δειγματοσημεία όπως φαίνεται στον παρακάτω πίνακα(σε κάθε στοιχείο του πίνακα αναφέρονται τα ψηφία στα οποία ενδέχεται να λήγουν οι δύο ακέραιοι αριθμοί).

0, 0	0, 1	0, 2	0, 3	0, 4	0, 5	0, 6	0, 7	0, 8	0, 9
1, 0				**1,4**					
2, 0		**2,2**							
3, 0								**3,8**	
4, 0	**4,1**					**4,6**			
5, 0									
6, 0				**6,4**					
7, 0									
8, 0			**8,3**					**8,8**	
9, 0									

Αν A είναι το γεγονός ότι το γινόμενο των δύο ακεραίων λήγει σε τέσσερα τότε το σύνολο του γεγονότος αυτού περιέχει τα εξής σημεία από τον παραπάνω δειγματοχώρο:

A={(2,2), (1,4), (4,1), (7,2), (2,7), (8,3), (3,8), (4,6), (6,4), (8,8), (6,9), (9,6) }.

Σύμφωνα με την κλασσική μέθοδο εκτίμησης πιθανότητας έχουμε

$$P(A) = \frac{N(A)}{N(S)} = \frac{12}{100}.$$

Δοθέντος ότι ο δεύτερος αριθμός λήγει σε ψηφίο μεγαλύτερο του 5 ο δειγματοχώρος περιορίζεται στις τέσσερις τελευταίες στήλες του παραπάνω πίνακα. Ο νέος δειγματοχώρος περιέχει 40 δειγματοσημεία και μόνο τρία από αυτά, (3, 8), (4, 6) και (8, 8) ανήκουν στο A. Αρα η ζητούμενη πιθανότητα είναι

$$P(A) = \frac{6}{40}.$$

3) Ποιο γεγονός είναι πιο πιθανό;

- Τουλάχιστον μια φορά *άσσος* σε 3 ζαριές με ένα ζάρι;
- Τουλάχιστον μια φορά *άσσοι* σε 12 ζαριές με δύο ζάρια;

Λύση

Θεωρώ τα γεγονότα A_κ={ *στην κ ρίψη έχουμε την ένδειξη άσσο*}. Τότε το πρώτο από ταπαραπάνω γεγονότα συμβολίζεται ώς εξής:

$$A = A_1 \cup A_2 \cup A_3 .$$

Εφαρμόζοντας γνωστά θεωρήματα της πιθανότητας γεγονότων έχουμε

$$P(A) = P(A_1 \cup A_2 \cup A_3) = 1 - P(\overline{A_1 \cup A_2 \cup A_3}) = 1 - P(\overline{A}_1 \cap \overline{A}_2 \cap \overline{A}_3)$$

$$P(A) = 1 - P(\overline{A}_1)P(\overline{A}_2)P(\overline{A}_3) = 1 - \left(\frac{5}{6}\right)^3 = 0{,}42 \quad .(\overline{A}_1, \overline{A}_2, \overline{A}_3 \ \textit{ανεξαρτητα})$$

Παρόμοια, η πιθανότητα του δεύτερου γεγονότος είναι

$$P(B) = 1 - \left(\frac{35}{36}\right)^{12} = 0{,}287 \ .$$

Συγκρίνοντας τις παραπάνω πιθανότητες, μία φορά άσσοι σε 12 ρίψεις είναι λιγότερο πιθανό από ένα άσσο σε 3ρίψεις.

Παρατήρηση: Η πιθανότητα της ένωσης πολλών γεγονότων, όπως τριών ή δώδεκα γεγονότων κ.τ.λ. απαιτεί πολλές πράξεις σύμφωνα με το θεώρημα της ένωσης. Στις περιπτώσεις αυτές υπολογίζουμε πιο εύκολα την πιθανότητα του συμπληρωματικού της ένωσης,

$$P(\overline{A_1 \cup A_2 \cup A_3 \cup \ldots}) = P(\overline{A}_1 \cap \overline{A}_2 \cap \overline{A}_3 \cap \ldots) = P(\overline{A}_1)P(\overline{A}_2)P(\overline{A}_3)\ldots$$

και την αφαιρούμε από την μονάδα.

4) Στη ρίψη δύο ζαριών γνωρίζεται ότι το άθροισμα είναι πολλαπλάσιο του 3. Ποια η πιθανότητα

(α) κάθε ζάρι να έχει φέρει 3;

(β) τουλάχιστον ένα ζάρι να έχει φέρει 3;

Στη ρίψη τριών ζαριών γνωρίζουμε ότι το άθροισμα είναι 10. Ποια η πιθανότητα να έχουμε τουλάχιστον ένα 3;

Λύση

Δεδομένου ότι το άθροισμα είναι πολλαπλάσιο του 3 ο *υπό συνθήκη* δειγματοχώρος περιορίζεται στα εξής σημεία:

S={(1,2), (2,1), (3,3), (2,4), (4,2), (5,1), (1,5), (6,6), (6,3), (3,6), (5,4), (4,5)}.

(α) Προφανώς το γεγονός A, ότι κάθε ζάρι έχει φέρει 3 είναι A={(3,3)}. Με την κλασσική μέθοδο υπολογίζεται η πιθανότητα

$$P(A)=\frac{n(A)}{n(S)}=\frac{1}{12}.$$

(β) Έστω το γεγονός {*τουλάχιστον ένα ζάρι φέρνει* 3} συμβολίζεται με το *B*, *B*={(3,), (3, 6). (6, 3)}. Η πιθανότητα του *B* υπολογίζεται όπως παραπάνω,

$$P(B)=\frac{N(B)}{N(S)}=\frac{3}{12}.$$

(γ) Όταν στην ρίψη τριών ζαριών το άθροισμα είναι 10, όλα τα δυνατά ενδεχόμενα είναι 28,

S={(1, 4, 5), (1, 5, 4), (1, 3, 6), (1, 6, 3), (2, 4, 4), (2, 3, 5), (2, 5, 3), (2, 2, 6), (2, 6, 2), (3, 6, 1), (3, 1, 6), (3, 5, 2), (3, 2, 5), (3, 3, 4), (3, 4, 3), (4, 5, 1), (4, 1, 5), (4, 4, 2), (4, 2, 4), (4, 3, 3), (5, 1, 4), (5, 4, 1), (5, 3, 2), (5, 2, 3), (5, 4, 1), (5, 1, 4), (6, 3, 1), (6, 2, 1)}.

Σε δεκαπέντε δειγματοσημεία, *αν δεν έχουμε κάνει λάθος στην μέτρηση*, υπάρχει τουλάχιστον μία ένδειξη με 3. Εφαρμόζουμε και πάλι την κλασσική μέθοδο, αφού τα 27 δειγματοσημεία είναι ισοπίθανα,

$$P(\{\text{τουλάχιστον ένα } 3\})=\frac{15}{27}.$$

5) Θεωρούμε τα γεγονότα *A* και *B*, να συγκριθούν οι πιθανότητες *P*(*A*) και *P*(*A*|*B*) σε κάθε μία από τις παρακάτω περιπτώσεις α, β, γ του Σχήματος 3.2. Πώς μεταβάλλεται η πιθανότητα πραγματοποίησης του *A*, όταν γνωρίζουμε ότι συμβαίνει το γεγονός *B* σε κάποια εκτέλεση του πειράματος;

Απάντηση:

(α) $P(A|B)=0\le P(A)$, επειδή το *A* δεν μπορεί να συμβεί όταν συμβαίνει το *B*.

(β) $P(A|B)=P(A\cap B)/P(B)=P(A)/P(B)\ge P(A)$, επειδή $A\cap B=A$ και $P(B)\le 1$.

(γ) $P(A|B)=P(A\cap B)/P(B)=P(B)/P(B)=1\ge P(A)$.

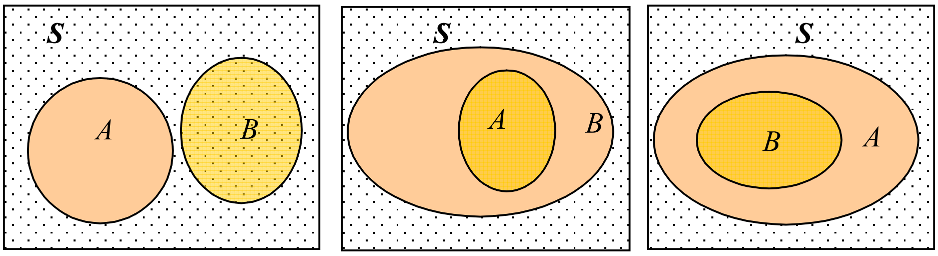
Παρατηρούμε ότι η πιθανότητα του υπό συνθήκη γεγονότος {*A*|*B*} στην πρώτη περίπτωση είναι μικρότερη από την πιθανότητα του *A*,

$$P(A|B)\le P(A),$$

και στις δύο άλλες περιπτώσεις είναι μεγαλύτερη,

$$P(A|B)\ge P(A).$$

Ανάλογα με την περίπτωση, η πληροφορία ότι συμβαίνει το γεγονός B ενδέχεται να αυξήσει ή να ελαττώσει ή και να αφήσει αμετάβλητη την πιθανότητα πραγματοποίησης του γεγονότος A.



(α) (β) (γ)

Σχήμα Δειγματοχώροι και γεγονότα

6) Στόχος αποτελείται από δύο τμήματα I και II. Οι αντίστοιχες πιθανότητες ευστοχίας είναι p_1 και p_2. Σε συγκεκριμένη βολή δεν πλήττεται το τμήμα I. Να ευρεθεί η πιθανότητα να πληγεί το τμήμα II.

Λύση

Ονομάζω με A και B τα γεγονότα ότι πλήττονται τα τμήματα I και II αντίστοιχα. Ζητείται η υπό συνθήκη πιθανότητα

$$P(B/\overline{A}) = \frac{P(B \cap \overline{A})}{P(\overline{A})} = \frac{P(B)P(\overline{A}/B)}{P(\overline{A})} = \frac{p_2 \cdot 1}{1-p_1} = \frac{p_2}{1-p_1}$$

7) Ένα ζάρι ρίχνεται 6 φορές. Ποια η πιθανότητα να έρθουν έξι διαφορετικά νούμερα;

Λύση

Ορίζουμε τα γεγονότα

A_1={ *στην πρώτη ρίψη έχουμε οιαδήποτε ένδειξη*}

$A_κ$={*στην κ ρίψη έχουμε ένδειξη διάφορη από την ένδειξη της κ-1, κ-2, …, 1 ρίψης*}, κ=2, 3, 4, 5, 6.

Έτσι το γεγονός R να έρθουν 6 διαφορετικά νούμερα συμβολίζεται με την ακόλουθη άλγεβρα,

$$R = A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5 \cap A_6$$

Για να υπολογίσω την πιθανότητα της τομής ακολουθώ τον **Πολλαπλασιαστικό νόμο,**

$$P(R)=P(A_1)P(A_2|A_1)P(A_3|A_1\cap A_2)....P(A_6|A_1\cap A_2\cap A_3\cap A_4\cap A_5)$$

όπου

$$P(A_1)=1$$

$$P(A_2|A_1)=\frac{5}{6}$$

$$P(A_3|A_1\cap A_2)=\frac{4}{6}$$

.....

$$.P(A_6|A_1\cap A_2\cap A_3\cap A_4\cap A_5)=\frac{1}{6}$$

Άρα

$$P(R)=P(A_1\cap A_2\cap A_3\cap A_4\cap A_5\cap A_6)=\frac{6}{6}\frac{5}{6}\frac{4}{6}\frac{3}{6}\frac{2}{6}\frac{1}{6}=\frac{6!}{6^6}$$

8) Μία κάλπη περιέχει δύο άσπρες και τρεις μαύρες σφαίρες. Δύο παίκτες τραβάνε ο ένας μετά τον άλλον από μία σφαίρα χωρίς επανατοποθέτηση στην κάλπη. Κερδίζει ο παίκτης που τραβά πρώτος μία άσπρη σφαίρα. Να υπολογιστεί η πιθανότητα να κερδίσει ο πρώτος παίκτης.

Λύση

Ορίζουμε τα γεγονότα

$A_κ$={ο πρώτος παίκτης διαλέγει για πρώτη φορά στο κ τράβηγμα μία άσπρη σφαίρα}

$B_κ$={ ο δεύτερος παίκτης διαλέγει για πρώτη φορά στο κ τράβηγμα μία άσπρη σφαίρα}.

Το γεγονός R ότι κερδίζει ο πρώτος παίκτης συμβολίζεται με την ακόλουθη άλγεβρα συνόλων,

$$R=A_1\cup(\overline{A_1}\cap\overline{B_2}\cap A_3)$$

Εφαρμογή του θεωρήματος της ένωσης δύο γεγονότων και μετά του πολλαπλασιαστικού νόμου δίνει

$$P(R)=P(A_1\cup(\overline{A_1}\cap\overline{B_2}\cap A_2))=P(A_1)+P(\overline{A_1}\cap\overline{B_2}\cap A_3)$$

$$=P(A_1)+P(\overline{A_1})P(\overline{B_2}|\overline{A_1})P(A_3|\overline{A_1}\cap\overline{B_2})$$

όπου

$$P(A_1) = \frac{2}{5}$$

$$P(\overline{A_1}) = \frac{3}{5}$$

$$P(\overline{B_2} | \overline{A_1}) = \frac{2}{4}$$

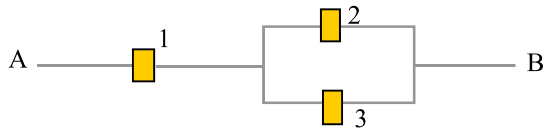
$$P(A_3 | \overline{A_1} \cap \overline{B_2}) = \frac{2}{3}$$

Οπότε

$$P(R) = \frac{2}{5} + \frac{3}{5}\frac{2}{4}\frac{2}{3} = \frac{3}{5}$$

Προβλήματα με τις λύσεις τους

1) Μεταφορά ρεύματος. Στη γραμμή μεταφοράς του παρακάτω κυκλώματος καθένας από τους διακόπτες μπορεί να είναι ανοικτός (διακοπή) με πιθανότητα p<0,5. Συμβολίζοντας με A_κ το γεγονός ότι ο κ διακόπτης είναι ανοικτός, όπου κ=1, 2, 3, να βρεθούν τα παρακάτω:



(α) το γεγονός R={ *διακοπή ρεύματος*} να διατυπωθεί με συμβολισμό συνόλων,
(β) η πιθανότητα του γεγονότος R αν γνωρίζουμε ότι διακοπή στον διακόπτη 2 διπλασιάζει την πιθανότητα διακοπής στον διακόπτη 3,
(γ) αν έχουμε διακοπή ρεύματος και ο διακόπτης 2 είναι κλειστός ποια η πιθανότητα ότι ο διακόπτης 1 είναι ανοικτός.

Λύση

(α) Διακοπή ρεύματος συμβαίνει στο κύκλωμα όταν ο διακόπτης 1 είναι ανοικτός ή όταν οι διακόπτες 2 και 3 είναι ανοικτοί ή όταν όλοι είναι ανοικτοί. Αυτό μπορεί να διατυπωθεί με την ακόλουθη άλγεβρα των συνόλων A_κ,

$$R = A_1 \cup (A_2 \cap A_3)$$

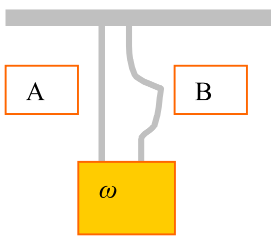
(β) Εφαρμόζω τον τύπο της πιθανότητας ένωσης ή τομής δύο ή περισσοτέρων γεγονότων. Τονίζουμε εδώ ότι το γεγονός A_1 δεν επιδρά στην πιθανότητα πραγματοποίησης των A_2 και A_3. Έτσι, έχουμε

$$\begin{aligned} P(R) &= P(A_1 \cup (A_2 \cap A_3)) = P(A_1) + P(A_2 \cap A_3) - P(A_1 \cap A_2 \cap A_3) \\ &= P(A_1) + P(A_2)P(A_3 | A_2) - P(A_1)P(A_2)P(A_3 | A_2) \\ &= p + p \cdot 2p - p \cdot p \cdot 2p = p + 2p^2 - 2p^3 \end{aligned}$$

(γ) Λογικά, αφού έχουμε διακοπή και ο διακόπτης 2 είναι κλειστός, ο διακόπτης 1 σίγουρα είναι ανοικτός, διαφορετικά δεν θα υπήρχε διακοπή στο κύκλωμα. Αυτό μπορούμε να το τεκμηριώσουμε και μεθοδολογικά,

$$\begin{aligned} P(A_1 | R \cap \overline{A_2}) &= \frac{P(A_1 \cap (R \cap \overline{A_2}))}{P(R \cap \overline{A_2})} = \frac{P((A_1 \cap R) \cap \overline{A_2})}{P(R \cap \overline{A_2})} = \\ &= \frac{P(R \cap \overline{A_2})}{P(R \cap \overline{A_2})} = 1 \end{aligned}$$

2) Ανύψωση φορτίου. Για την ανύψωση ενός μεγάλου βάρους ω πρόκειται να χρησιμοποιηθεί ένα χονδρό σχοινί A. Στην περίπτωση που κοπεί το σχοινί A, υπάρχει ένα εφεδρικό σχοινί B, το οποίο θα συνεχίσει την ανύψωση του φορτίου ω.
Η πιθανότητα να κοπεί το σχοινί A κατά την ανύψωση του φορτίου είναι 0,04. Στην περίπτωση αποτυχίας του A , η πιθανότητα ανύψωσης του βάρους από το B σχοινί είναι 0,80.



(α) Ποιά είναι η πιθανότητα να ανυψωθεί το φορτίο;

(β) Αν ανυψωθεί το φορτίο, ποια είναι η πιθανότητα ότι κανένα σχοινί δεν έχει κοπεί;

Λύση

Συμβολίζουμε με A και B τα γεγονότα ότι σπάει το Α και Β σχοινί αντίστοιχα. Ακόμη το γεγονός R παριστάνει την ανύψωση του φορτίου ω

(α) Η ανύψωση του φορτίου R μπορεί να παρασταθεί με την ακόλουθη άλγεβρα γεγονότων,

$$R = \overline{A} \cup (A \cap \overline{B}) .$$

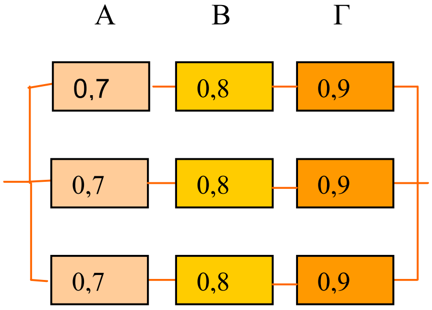
Κατά συνέπεια η ζητούμενη πιθανότητα προκύπτει εύκολα από το θεώρημα της ένωσης δύο γεγονότων

$$P(R) = P(\overline{A} \cup (A \cap \overline{B})) = P(\overline{A}) + P(A \cap \overline{B}) = P(\overline{A}) + P(A)P(\overline{B}|A)$$
$$= 0{,}96 + 0{,}04 \cdot 0{,}80 = 0{,}992$$

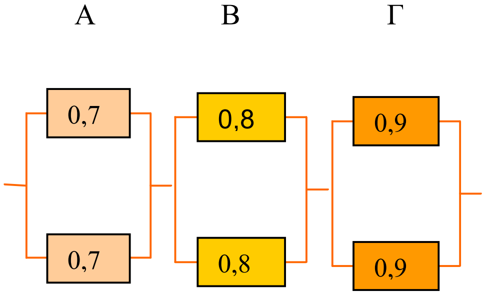
(β) Το γεγονός ότι κανένα σχοινί δεν έχει κοπεί είναι ισοδύναμο με το γεγονός ότι δεν κόπηκε το σχοινί Α. Άρα η ζητούμενη πιθανότητα εύκολα υπολογίζεται από τον κανόνα της υπό συνθήκης πιθανότητας

$$P(\overline{A}|R) = \frac{P(\overline{A} \cap R)}{P(R)} = \frac{P(\overline{A})P(R|\overline{A})}{P(R)} = \frac{0{,}96 \cdot 1}{0{,}932} = 0{,}968$$

3) Στάδια επεξεργασίας. Στη μελέτη για την κατασκευή ενός εργοστασίου, το οποίο θα παράγει ένα προϊόν με τρία βασικά στάδια επεξεργασίας, Α, Β και Γ, ο μηχανικός προβληματίζεται με δύο εναλλακτικές τοποθετήσεις των μηχανών. Σε κάθε στάδιο επεξεργασίας υπάρχουν δύο ή τρεις ίδιες μηχανές, όλες μεταξύ τους *στατιστικώς ανεξάρτητες*, με αντίστοιχες πιθανότητες καλής λειτουργίας όπως φαίνεται στο σχήμα.



Σχέδιο α



Σχέδιο β

Το κόστος κατασκευής και στα δύο σχέδια είναι το ίδιο, θα προτιμηθεί όμως αυτό που είναι πιο παραγωγικότερο. Δεδομένου ότι για να παραχθεί το προϊόν πρέπει να περάσει από μία μηχανή του κάθε σταδίου επεξεργασίας πάντοτε από αριστερά προς τα δεξιά όπως βλέπετε το σχήμα.

Λύση

Στο **Σχέδιο α** για να παραχθεί το προϊόν πρέπει τουλάχιστον σε μία από τις διαδρομές να λειτουργούν και οι τρεις μηχανές. Έτσι, αν R είναι το γεγονός ότι μπορεί το προϊόν να επεξεργασθεί και στα τρία στάδια (κινούμενο από αριστερά προς δεξιά) και αν A_i, B_i, Γ_i, είναι τα γεγονότα ότι λειτουργούν σωστά οι μηχανές i (i=1, 2, 3) των φάσεων Α, Β, και Γ, έχουμε

$$
\begin{aligned}
&R = (A_1 \cap B_1 \cap \Gamma_1) \cup (A_2 \cap B_2 \cap \Gamma_2) \cup (A_3 \cap B_3 \cap \Gamma_3) \\
&P(R) = P[(A_1 \cap B_1 \cap \Gamma_1) \cup (A_2 \cap B_2 \cap \Gamma_2) \cup (A_3 \cap B_3 \cap \Gamma_3)] \\
&P(R) = P(A_1 \cap B_1 \cap \Gamma_1) + P(A_2 \cap B_2 \cap \Gamma_2) + P(A_3 \cap B_3 \cap \Gamma_3) \\
&- P(A_1 \cap B_1 \cap \Gamma_1 \cap A_2 \cap B_2 \cap \Gamma_2) - P(A_1 \cap B_1 \cap \Gamma_1 \cap A_3 \cap B_3 \cap \Gamma_3) \\
&- P(A_2 \cap B_2 \cap \Gamma_2 \cap A_3 \cap B_3 \cap \Gamma_3) + P(A_1 \cap B_1 \cap \Gamma_1 \cap A_2 \cap B_2 \cap \Gamma_2 \cap A_3 \cap B_3 \cap \Gamma_3) \\
&= 3 \cdot 0{,}7 \cdot 0{,}8 \cdot 0{,}9 - 3 \cdot 0{,}7^2 \cdot 0{,}8^2 \cdot 0{,}9^2 + 0{,}7^3 \cdot 0{,}8^3 \cdot 0{,}9^3 = 1{,}512 - 0{,}762 + 0{,}128 = 0{,}878
\end{aligned}
$$

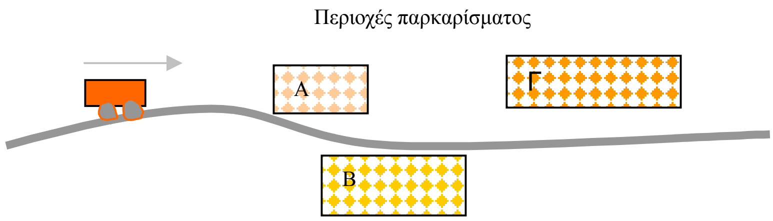
Στο **Σχέδιο** β η επεξεργασία του προϊόντος μπορεί να γίνει σε μία από τις δύο μηχανές του κάθε σταδίου, και με παρόμοιο συμβολισμό έχουμε την εξής άλγεβρα γεγονότων,

$$
\begin{aligned}
&R = (A_1 \cup A_2) \cap (B_1 \cup B_2) \cap (\Gamma_1 \cup \Gamma_2) \\
&P(R) = P[(A_1 \cup A_2) \cap (B_1 \cup B_2) \cap (\Gamma_1 \cup \Gamma_2)] = P(A_1 \cup A_2) P(B_1 \cup B_2) P(\Gamma_1 \cup \Gamma_2) \\
&P(R) = [P(A_1) + P(A_2) - P(A_1 \cap A_2)][P(B_1) + P(B_2) - P(B_1 \cap B_2)] \times \\
&\qquad \times [P(\Gamma_1) + P(\Gamma_2) - P(\Gamma_1 \cap \Gamma_2)] \\
&\qquad = (0{,}7 + 0{,}7 - 0{,}49)(0{,}8 + 0{,}8 - 0{,}64)(0{,}9 + 0{,}9 - 0{,}81) = 0{,}91 \cdot 0{,}96 \cdot 0{,}99 = 0{,}8648
\end{aligned}
$$

Επειδή 0,878 > 0,8648, ο μηχανικός ***προτιμά*** το **Σχέδιο α**.

4) Χώρος στάθμευσης. Για την μελέτη του χώρου στάθμευσης των αυτοκινήτων σε μία πόλη, επιλέγεται ένας μηχανικός που επισκέπτεται καθημερινά την πόλη και παρατηρούνται οι ευκολίες ή δυσκολίες του στην εύρεση χώρου παρκαρίσματος. Ο μηχανικός κάθε πρωί φτάνει στην πόλη, διέρχεται με την σειρά από τις περιοχές Α, Β και Γ, και παρκάρει αμέσως μόλις βρεί κενό χώρο σε κάποια από αυτές. Το παρκάρισμα στις περιοχές Α και Β

είναι δωρεάν, ενώ στην Γ υπάρχουν παρκόμετρα. Δεν υπάρχει άλλη περιοχή για παρκάρισμα εκτός από τις περιοχές Α, Β και Γ.
Από τα στατιστικά στοιχεία που συλλέγονται προκύπτει ότι, καθώς έρχεται το πρωί ο μηχανικός, οι πιθανότητες να βρεθεί χώρος στις περιοχές Α, Β και Γ είναι, 0,4, 0,2 και 0,6 αντίστοιχα. Αν, τώρα, η Α είναι γεμάτη, η πιθανότητα να βρεθεί χώρος στη Β είναι μόνο 0,10. Επίσης, αν οι Α και Β είναι γεμάτες, η πιθανότητα κενού χώρου στη Γ είναι 0,30. Για μια συγκεκριμένη μέρα να υπολογιστούν τα εξής:
(α) Η πιθανότητα ότι ο μηχανικός θα παρκάρει χωρίς να πληρώσει.



(β) Η πιθανότητα ότι ο μηχανικός θα βρεί μέρος να παρκάρει το αυτοκίνητο.
(γ) Αν ο μηχανικός βρήκε χώρο, κάποιο πρωί, ποια η πιθανότητα ότι πλήρωσε για παρκάρισμα;

Λύση
Ορίζω τα γεγονότα,

A={Υπάρχει κενός χώρος στην περιοχή Α, την συγκεκριμένη μέρα},
B={Υπάρχει κενός χώρος στην περιοχή Β, την συγκεκριμένη μέρα},
Γ={Υπάρχει κενός χώρος στην περιοχή Γ, την συγκεκριμένη μέρα}.

(α) Το γεγονός ότι δεν πλήρωσε για παρκάρισμα είναι ισοδύναμο με την ύπαρξη θέσης στις περιοχές Α ή Β, και συμβολίζεται με την ένωση $A \cup B$.
Ζητούμε την πιθανότητα $P(A \cup B)$, η οποία υπολογίζεται εύκολα με τον νόμο De Morgan και πολλαπλασιαστικό κανόνα,

$$P(A \cup B) = 1 - P(\overline{A \cup B}) = 1 - P(\overline{A} \cap \overline{B}),$$

αλλά

$$P(\overline{A} \cap \overline{B}) = P(\overline{A})P(\overline{B}|\overline{A}) = 0,8[1 - P(B|\overline{A})] = 0,6 \cdot (1 - 0,1) = 0,6 \cdot 0,9 = 0,54$$

άρα

$$P(A \cup B) = 1 - 0,54 = 0,46.$$

(β) Το γεγονός του οποίου ζητάμε την πιθανότητα συμβολίζεται με την ένωση $A \cup B \cup \Gamma$. Θα υπολογίσω την πιθανότητα του συμπληρωματικού γεγονότος, και θα την αφαιρέσω από την μονάδα.

$$P(A \cup B \cup \Gamma) = 1 - P(\overline{A \cup B \cup \Gamma}) = 1 - P(\overline{A} \cap \overline{B} \cap \overline{\Gamma})$$
$$= 1 - P(\overline{A})\, P(\overline{B} | \overline{A})\, P(\overline{\Gamma} | \overline{A} \cap \overline{B})$$
$$= 1 - 0{,}6 \cdot 0{,}09 \cdot 0{,}70 = 1 - 0{,}378 = 0{,}622$$

(γ) Το γεγονός ότι ο μηχανικός πλήρωσε για παρκάρισμα ταυτίζεται με το γεγονός

$$\overline{A} \cap \overline{B} \cap \Gamma .$$

Βρέθηκε χώρος για παρκάρισμα συμβολίζεται με $A \cup B \cup \Gamma$. Ζητάμε την υπό συνθήκη πιθανότητα,

$$P[(\overline{A} \cap \overline{B} \cap \Gamma) | (A \cup B \cup \Gamma)] = \frac{P[(\overline{A} \cap \overline{B} \cap \Gamma) \cap (A \cup B \cup \Gamma)]}{P(A \cup B \cup \Gamma)}.$$
$$= \frac{P(\overline{A} \cap \overline{B} \cap \Gamma)}{P(A \cup B \cup \Gamma)} = \frac{0{,}162}{0{,}622} = 0{,}26$$

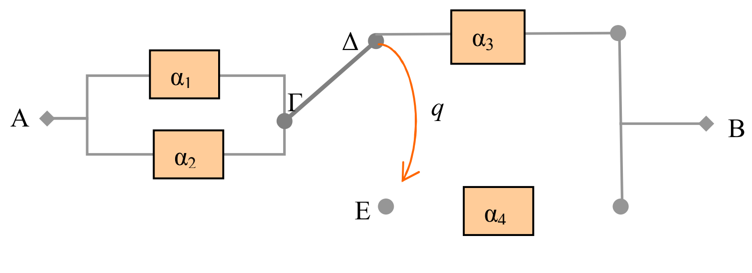
5) Ηλεκτρικό κύκλωμα. Στο παρακάτω κύκλωμα ο κάθε διακόπτης α_1, α_2, α_3, και α_4 λειτουργεί ανεξάρτητα με πιθανότητα p να είναι κλειστός. Παριστάνουμε με A_κ το γεγονός ότι ο α_κ διακόπτης είναι κλειστός, όπου κ=1, 2, 3, 4.

Αν ΓΔ είναι μεταλλική ράβδος η οποία μπορεί να μετακινηθεί στο Ε με πιθανότητα q, τότε

(α) να παρασταθεί αλγεβρικά το γεγονός R={ *δεν υπάρχει διακοπή ρεύματος*},

(β) να βρεθεί η πιθανότητα του γεγονότος R,

(γ) αν δεν έχουμε διακοπή ρεύματος, ποια η πιθανότητα ότι ο διακόπτης α_4 είναι κλειστός;



(α) Το γεγονός R συμβαίνει όταν δεν έχουμε διακοπή από το Α στο Γ και από το Γ στο Β. Αυτό μπορώ να το εκφράσω με την ακόλουθη άλγεβρα,

$$R = (A_1 \cup A_2) \cap (A_3 \cup (\mathit{\Gamma E} \cap A_4))$$

Τονίζουμε εδώ ότι το ενδεχόμενο μεταφοράς ρεύματος από το Γ στο Δ είναι το σίγουρο γεγονός, ενώ από το Γ στο Ε συμβολίζεται με το γεγονός *ΓΕ* με πιθανότητα πραγματοποίησης q.
(β) Εφαρμόζουμε την πιθανότητα της τομής των δύο παρενθέσεων, οι οποίες είναι ανεξάρτητες, και λόγω ανεξαρτησίας των όλων των γεγονότων έχουμε

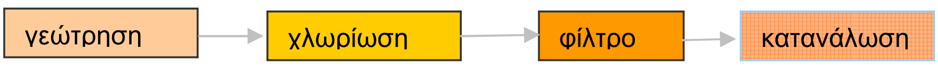
$$P(R) = P((A_1 \cup A_2) \cap (A_3 \cup (\mathit{\Gamma E} \cap A_4))) = P(A_1 \cup A_2)P(A_3 \cup (\mathit{\Gamma E} \cap A_4))$$
$$= [P(A_1) + P(A_2) - P(A_1 \cap A_2)] \cdot [P(A_3) + P(\mathit{\Gamma E} \cap A_4) - P(A_3 \cap \mathit{\Gamma E} \cap A_4)]$$

$$= [P(A_1) + P(A_2) - P(A_1)P(A_2)] \cdot [P(A_3) + P(\mathit{\Gamma E})P(A_4) - P(A_3)P(\mathit{\Gamma E})P(A_4)]$$
$$= [p + p - p \cdot p] \cdot [p + q \cdot p - p \cdot q \cdot p] = (2p - p^2)(p(1 + q - qp))$$

(γ) Έχουμε υπό συνθήκη πιθανότητα, και ύστερα από απλές αλγεβρικές πράξεις έχουμε

$$P(A_4 | R) = \frac{P(A_4 \cap R)}{P(R)} = \frac{P[A_4 \cap [(A_1 \cup A_2) \cap (A_3 \cup (\Gamma\mathrm{E} \cap \mathrm{A}_4))]]}{P(R)}$$
$$= \frac{P[(A_1 \cup A_2) \cap ((A_3 \cap A_4) \cup (\Gamma\mathrm{E} \cap \mathrm{A}_4))]}{P(R)} = \frac{P(A_1 \cup A_2)P[(A_3 \cap A_4) \cup (\Gamma\mathrm{E} \cap \mathrm{A}_4)]}{P(A_1 \cup A_2)P[(A_3 \cup (\Gamma\mathrm{E} \cap \mathrm{A}_4)]}$$
$$\frac{P((A_3 \cap A_4) \cup (\Gamma\mathrm{E} \cap \mathrm{A}_4))}{P((A_3 \cup (\Gamma\mathrm{E} \cap \mathrm{A}_4))} = \frac{p^2 + pq - p^2 q}{p + qp - p^2 q} = \frac{p(p + q - p^2 q)}{p(1 + q - pq)} = \frac{p + q - p^2 q}{1 + q - pq}$$

6) Χλωρίωση νερού. Το νερό για μία κατοικημένη περιοχή πρώτα αντλείται από μία γεώτρηση, μετά περνά από ένα σύστημα χλωρίωσης, και τέλος μέσα από ένα φίλτρο νερού.



Η πιθανότητα αποτυχίας της αντλίας, του συστήματος χλωρίωσης, και φιλτραρίσματος μέσα σε ένα χρόνο είναι 0,1, 0,2 και 0,1 αντίστοιχα. Η πιθανότητα να αστοχήσουν μέσα στον χρόνο και τα δύο η χλωρίωση και το φιλτράρισμα είναι 0,02. Η αποτυχία άντλησης προκαλεί έλλειψη ικανοποιητικής παροχής νερού, ενώ η αποτυχία της χλωρίωσης ή φιλτραρίσματος ελαττώνει την ποιότητα νερού κάτω από τα επιτρεπτά όρια.

Το γεγονότα αποτυχίας άντλησης νερού, χλωρίωσης, και φιλτραρίσματος θεωρούνται στατιστικά ανεξάρτητα.
(α) Ποια η πιθανότητα αυτόν τον χρόνο η κατοικημένη περιοχή να έχει ικανοποιητική παροχή νερού παραδεκτής ποιότητας;
(β) Όταν η κατοικημένη περιοχή πίνει νερό χαμηλής ποιότητας (κάτω από τα επιτρεπτά όρια), ποια η πιθανότητα να οφείλεται στην αποτυχία της χλωρίωσης;
(γ) Εάν το σύστημα χλωρίωσης αντικατασταθεί με ένα ποιο αξιόπιστο, πιθανότητα αποτυχίας 0,1, κατά πόσο περιορίζεται το ποσοστό ευθύνης της ερώτησης (β);

Λύση
Ορίζουμε τα γεγονότα:

$A1$={ *αποτυχία άντλησης νερού μέσα στον χρόνο*},
$A2$={ *αποτυχία χλωρίωσης νερού μέσα στον χρόνο*},
$A3$={ *αποτυχία φιλτραρίσματος νερού μέσα στον χρόνο*}.

Ακόμη, μπορούμε να ορίσουμε τα γεγονότα,

$E1$={ *έλλειψη ικανοποιητικής παροχής νερού*}=$A1$,
$E2$={ *ποιότητα νερού κάτω από τα επιτρεπτά όρια*}$=A3\cup A2$.
(α) Για να έχει η πόλη ικανοποιητική παροχή νερού παραδεκτής ποιότητας πρέπει να μην συμβαίνει ούτε το $E1$ ούτε το $E2$. Δηλαδή, το γεγονός εδώ, R, συμβολίζεται σαν

$$R=\overline{E1}\cap\overline{E2}.$$

Άρα η ζητούμενη πιθανότητα μπορεί να υπολογισθεί από τον πολλαπλασιαστικό κανόνα,

$$P(R)=P(\overline{E1}\cap\overline{E2})=P(\overline{E1})P(\overline{E2})=P(\overline{A1})P(\overline{A2\cup A3})=$$
$$=0{,}90P(\overline{A2}\cap\overline{A3})=0{,}9P(\overline{A2})P(\overline{A3})=0{,}9\cdot 0{,}8\cdot 0{,}9=0{,}648$$

αφού τα γεγονότα $\overline{A1}, \overline{A2}, \overline{A3}$ είναι ανεξάρτητα.
(β) Είναι προφανές ότι ζητάμε την υπό συνθήκη πιθανότητα

$$P(A2\,|\,A2\cup A3)=\frac{P[A2\cap(A2\cup A3)]}{P(A2\cup A3)}=\frac{P(A2)}{P(A2\cup A3)}=\frac{0{,}2}{0{,}2+0{,}1-0{,}2\cdot 0{,}1}=0{,}714$$

(γ) Για να εκτιμήσουμε το ποσοστό ευθύνης της χλωρίωσης όταν η κατοικημένη περιοχή πίνει νερό χαμηλής ποιότητας αρκεί να υπολογίσουμε την υπό συνθήκη πιθανότητα όπως στην ερώτηση (β) αλλά τώρα θα θέσουμε όπου $P(A2)$=0,1. Έτσι έχουμε,

$$P(A2 \mid A2 \cup A3) = \frac{0{,}1}{0{,}1 + 0{,}1 - 0{,}1 \cdot 0{,}1} = \frac{0{,}1}{0{,}19} = 0{,}526\,.$$

Δηλαδή, με το νέο σύστημα χλωρίωσης, όταν το νερό είναι χαμηλής ποιότητας το ποσοστό ευθύνης της χλωρίωσης περιορίζεται από 0,714 σε 0,526.

7) Αστοχία θεμελίων. Τα θεμέλια ενός μεγάλου κτιρίου μπορεί να αστοχήσουν είτε από υπερβολική καθίζηση ή από θραύση του εδάφους που συμβαίνουν με πιθανότητες 0,01 και 0,02 αντίστοιχα. Ακόμη, είναι γνωστό ότι η εμφάνιση υπερβολικής καθίζησης αυξάνει την πιθανότητα θραύσης του εδάφους από 0,02 σε 0,08. Κάποια στιγμή αστοχούν τα θεμέλια του συγκεκριμένου κτιρίου, ποια η πιθανότητα ότι η αστοχία προκλήθηκε μόνο από θραύση εδάφους;

Λύση

Ορίζουμε τα γεγονότα

K={*υπερβολική καθίζηση εδάφους*},

Θ={*θραύση εδάφους*},

οι δε δοθείσες πιθανότητες είναι, $P(K) = 0{,}01, \quad P(\Theta) = 0{,}02, \quad P(\Theta \mid K) = 0{,}08 \cdot$

Ακόμη, συμβολίζουμε με πράξεις τα γεγονότα

$\Theta \cap \overline{K}$ ={ *μόνο θραύση εδάφους*}

$K \cup \Theta$ ={ *αστοχία θεμελίων*}.

Ζητάμε την πιθανότητα P({*μόνο θραύση εδάφους*| *αστοχία θεμελίων*}), η οποία είναι μία υπό συνθήκη πιθανότητα,

$$P[(\Theta \cap \overline{K}) \mid (\Theta \cup K)] = \frac{P[(\Theta \cap \overline{K}) \cap (\Theta \cup K)]}{P(\Theta \cup K)} = \frac{P(\Theta \cap \overline{K})}{P(\Theta \cup K)} =$$

$$= \frac{P(\Theta) P(\overline{K} \mid \Theta)}{P(\Theta) + P(K) - P(\Theta \cap K)} = \frac{P(\Theta)(1 - P(K \mid \Theta))}{P(\Theta) + P(K) - P(\Theta \cap K)}$$

όπου οι πιθανότητες $P(K \mid \Theta)$ και $P(K \cap \Theta)$ μπορούν να υπολογισθούν από τα δεδομένα του προβλήματος,

$$P(K \cap \Theta) = P(K) P(\Theta \mid K) = 0{,}01 \cdot 0{,}08 = 0{,}0008$$

$$P(K \mid \Theta) = \frac{P(K \cap \Theta)}{P(\Theta)} = \frac{0{,}0008}{0{,}02} = 0{,}04$$

Αντικαθιστώντας στην υπό συνθήκη πιθανότητα έχουμε

$$P[(\Theta \cap \overline{K}) | (\Theta \cup K)] = \frac{0{,}02 \cdot 0{,}96}{0{,}01 + 0{,}02 - 0{,}0016} = \frac{0{,}0192}{0{,}0284} = 0{,}676$$

8) Υλοποίηση έργου. Τρείς κατασκευαστικές εταιρείες Α, Β και Γ, υποβάλλουν προσφορές για ένα μεγάλο δρόμο. Από προηγούμενα έργα που έχουν αναλάβει οι εταιρείες έχει διαμορφωθεί η γνώμη ότι οι πιθανότητες καθυστέρησης της υλοποίησης του έργου είναι 0,6, 0,4 και 0,3 για τις τρείς εταιρείες αντίστοιχα. Εάν οι ομάδες Α και Β έχουν την ίδια πιθανότητα να κερδίσουν το έργο, ενώ η ομάδα Γ έχει τριπλάσια πιθανότητα από τις Α ή Β, να υπολογιστούν:
(α) Η πιθανότητα ότι το έργο θα υλοποιηθεί χωρίς καθυστέρηση.
(β) Η πιθανότητα ότι το έργο το ανέλαβε η ομάδα Γ, εάν γνωρίζουμε ότι δεν καθυστέρησε η υλοποίησή του.

Λύση

Συμβολίζοντας με *Α, Β, Γ* τα γεγονότα ότι το έργο το κέρδισε η Α, Β, Γ εταιρεία αντίστοιχα, από τα δεδομένα του προβλήματος οι *εκ των προτέρων* πιθανότητες των γεγονότων αυτών είναι, $P(A)=P(B)=p$, $P(\Gamma)=3p$.
Αφού τα γεγονότα *Α, Β, Γ* αποτελούν διαμέριση του δειγματικού χώρου, όσο αφορά ποια εταιρεία κερδίζει το έργο, έχουμε

$$P(A)+P(B)+P(\Gamma)=1 \text{ ή } p+p+3p=1$$

και λύνοντας την εξίσωση έχουμε p=0,20 ή $P(A)$=0,20, $P(B)$=0,20 και $P(\Gamma)$=0,6. Εδώ, μία γραφική παράσταση με δένδρο φαίνεται στο ακόλουθο Σχήμα. Όπου,

K={καθυστερεί η υλοποίηση του έργου},

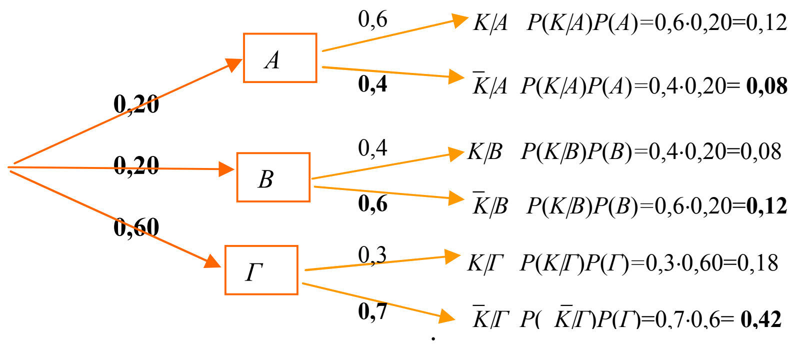
K={δεν καθυστερεί η υλοποίηση του έργου}.

Συνεπώς,
(α) η πιθανότητα να ολοκληρωθεί το έργο χωρίς καθυστέρηση, σύμφωνα με την ολική πιθανότητα είναι

$$\begin{aligned} P(\overline{K}) &= P(\overline{K}|A)P(A) + P(\overline{K}|B)P(B) + P(\overline{K}|\Gamma)P(\Gamma) \\ &= 0{,}4 \cdot 0{,}20 + 0{,}6 \cdot 0{,}20 + 0{,}7 \cdot 0{,}6 = 0{,}08 + 0{,}12 + 0{,}42 = 0{,}62, \end{aligned}$$

(β) η δε *εκ των υστέρων* εκτίμηση της πιθανότητας να ανέλαβε το έργο η Γ εταιρεία με την πληροφορία ότι το έργο ολοκληρώθηκε χωρίς καθυστέρηση, δίνεται από το τύπο του Bayes,

$$P(\Gamma|\overline{K}) = \frac{P(\overline{K}|\Gamma)P(\Gamma)}{P(\overline{K})} = \frac{0{,}7 \cdot 0{,}6}{0{,}62} = 0{,}677$$



Πράγματι, παρατηρούμε ότι όταν το έργο δεν καθυστερεί στην υλοποίησή του αυξάνεται η υπόνοια ή η πιθανότητα να το ανέλαβε η Γ εταιρεία, από 0,6 σε 0,677, αφού η Γ είναι η πιο αξιόπιστη απ' όλες.

9) Μονάδες ηλεκτρικού ρεύματος. Δύο παρόμοιες μονάδες Α και Β, παραγωγής ηλεκτρικής ενέργειας, τροφοδοτούν μία πόλη με ηλεκτρισμό. Η πιθανότητα βλάβης για κάθε μονάδα είναι 0,10 ενώ η πιθανότητα ταυτόχρονης βλάβης και στις δύο μονάδες είναι 0,02.
(α) Εκφράστε το γεγονός ότι μόνο μία μονάδα έχει βλάβη με συμβολισμό συνόλων και υπολογίστε την πιθανότητά του.
(β) Η ζήτηση ηλεκτρικού ρεύματος ικανοποιείται με πιθανότητα 90% , 70% , 40% όταν, δεν υπάρχει βλάβη στις μονάδες, υπάρχει βλάβη μόνο σε μία μονάδα, έχουν βλάβη και οι δύο μονάδες , αντίστοιχα. Κάποια στιγμή η ζήτηση ηλεκτρικού ρεύματος δεν ικανοποιείται, ποια η πιθανότητα ότι έχουν βλάβη και οι δύο μονάδες

Λύση

Συμβολίζω τα εξής γεγονότα:

A={*βλάβη στην μονάδα* Α},
B={ *βλάβη στην μονάδα* Β},
R={*βλάβη μόνο σε μία μονάδα*}.

(α) Το γεγονός *R* συμβαίνει στην περίπτωση που μόνο η μονάδα Α έχει βλάβη ή μόνο η μονάδα Β, δηλαδή

$$R = (A \cap \overline{B}) \cup (\overline{A} \cap B).$$

Έτσι, η πιθανότητα του γεγονότος R προκύπτει εύκολα με τους κανόνες της ένωσης και τομής.

$$P(R) = P[(A \cap \overline{B}) \cup (\overline{A} \cap B)] = P(A \cap \overline{B}) + P(\overline{A} \cap B) = P(A)P(\overline{B}|A) + P(B)P(\overline{A}|B)$$

αλλά

$$P(\overline{B}|A) = 1 - P(B|A) = 1 - \frac{P(A \cap B)}{P(A)} = 1 - \frac{0{,}02}{0{,}1} = 1 - 0{,}2 = 0{,}8$$

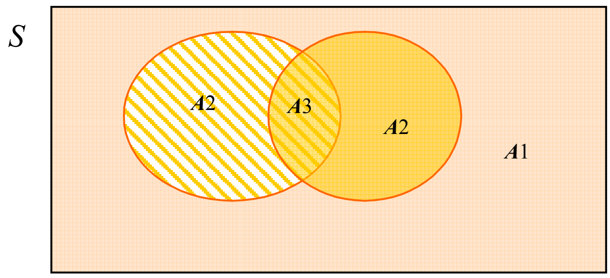
παρόμοια

$P(\overline{A}|B) = 0{,}8$. Άρα, $P(R)$ = 0,1·0,8+0,1·0,8=0,16.

(β) Εδώ μπορώ για λόγους ευκολίας να ορίσω τα γεγονότα,

E={ *Η πόλη έχει ικανοποιητική παροχή ρεύματος*},
*A*1={*Καμία μονάδα δεν έχει βλάβη*},
*A*2={*Μόνο μία μονάδα έχει βλάβη*},
*A*3={ *Και οι δύο μονάδες έχουν βλάβη*}.

Τα γεγονότα *A*1, *A*2, *A*3 αποτελούν διαμέριση του δειγματικού χώρου που αντιστοιχεί στον αριθμό μονάδων με βλάβη,



με αντίστοιχες πιθανότητες

$$P(A3) = P(A \cap B) = 0{,}02$$
$$P(A2) = 0{,}16$$
$$P(A1) = 1 - P(A2) - P(A3) = 0{,}82$$

Επίσης, συμβολίζοντας με Ε το γεγονός να έχουμε ικανοποιητική παροχή ρεύματος υπολογίζουμε την πιθανότητά του από τον νόμο της ολικής πιθανότητας

$$P(E)=P(E\,|A1)P(A1)+P(E\,|A2)P(A2)+P(E\,|A3)P(A3)$$
$$=0{,}90\cdot 0{,}82+0{,}70\cdot 0{,}16+0{,}40\cdot 0{,}02=0{,}738+0{,}112+0{,}008=0{,}858$$

Τέλος, η ζητούμενη πιθανότητα διατυπώνεται και ως $P(A3|\overline{E})$ και υπολογίζεται με το γνωστό θεώρημα του Bayes:

$$P(A3|\overline{E})=\frac{P(\overline{E}|A3)P(A3)}{P(\overline{E})}=\frac{0{,}60\cdot 0{,}02}{1-P(E)}=\frac{0{,}012}{0{,}142}=0{,}084$$

Παρατήρηση. Δεν είναι απαραίτητο να εκτιμήσω και την $P(\overline{E})$ με τον νόμο της ολικής πιθανότητας αφού έχω υπολογίσει την $P(E)$.

10) Ατμοσφαιρική ρύπανση. Η ρύπανση της ατμόσφαιρας σε μία μεγάλη πόλη οφείλεται κυρίως στην βιομηχανία, στα αυτοκίνητα και στους καυστήρες θέρμανσης. Μία νέα κυβέρνηση θέλει να ελέγξει αυτές τις πηγές ρύπανσης μέσα στα επόμενα τέσσερα χρόνια. Οι ευκαιρίες για επιτυχή έλεγχο αυτών των πηγών είναι 80%, 70%, και 50% αντίστοιχα. Ακόμη, η κυβέρνηση πιστεύει ότι εάν μόνο μία από τις πηγές ελεγχθούν επιτυχώς, η πιθανότητα να φέρει την ρύπανση κάτω από τα επιτρεπτά όρια είναι μόνο 50%, αλλά η πιθανότητα αυτή αυξάνει σε 80% αν ελεγχθούν επιτυχώς δύο από τις πηγές ρύπανσης. Προφανώς, σε περίπτωση επιτυχούς ελέγχου όλων των πηγών σίγουρα θα υπάρξει ρύπανση πολύ πιο κάτω από τα επιτρεπτά όρια.

Με την υπόθεση ότι υπάρχει στατιστική ανεξαρτησία μεταξύ των επιτυχών ελέγχων ρύπανσης στην βιομηχανία, αυτοκίνητα και θέρμανση, ποια η πιθανότητα ότι η νέα κυβέρνηση θα επιτύχει πτώση της ρύπανσης κάτω από το επιτρεπτό όριο;

Λύση

Ορίζουμε τα ακόλουθα γεγονότα,

B={*επιτυχής έλεγχος στην βιομηχανική ρύπανση*},
A={ “ “ “ *ρύπανση αυτοκινήτου*},
K={ “ “ “ “ *καυστήρα*},
E={*πτώση της ρύπανσης κάτω από το επιτρεπτό όριο*}.

Το γεγονότα

A_0={ *κανένας επιτυχής έλεγχος*},
A_1={*ακριβώς ένας επιτυχής έλεγχος*},
A_2={*ακριβώς δύο επιτυχείς έλεγχοι*},
A_3={*ακριβώς τρεις επιτυχείς έλεγχοι*}.

αποτελούν διαμέριση του δειγματικού χώρου *S*, με αντίστοιχες πιθανότητες

$$
\begin{aligned}
P(A_0) &= P(\overline{B}\cap\overline{A}\cap\overline{K}) = P(\overline{B})P(\overline{A})P(\overline{K}) = 0{,}03 \\
P(A_1) &= P[(B\cap\overline{A}\cap\overline{K})\cup(\overline{B}\cap A\cap\overline{K})\cup(\overline{B}\cap\overline{A}\cap K)] \\
&= P(B\cap\overline{A}\cap\overline{K}) + P(\overline{B}\cap A\cap\overline{K}) + P(\overline{B}\cap\overline{A}\cap K) \\
&= P(B)P(\overline{A})P(\overline{K}) + P(\overline{B})P(A)P(\overline{K}) + P(\overline{B})P(\overline{A})P(K) \\
&= 0{,}80\cdot 0{,}30\cdot 0{,}50 + 0{,}20\cdot 0{,}70\cdot 0{,}50 + 0{,}20\cdot 0{,}30\cdot 0{,}50 = 0{,}22 \\
P(A_2) &= P[(B\cap A\cap\overline{K})\cup(\overline{B}\cap A\cap K)\cup(B\cap\overline{A}\cap K)] \\
&= P(B\cap A\cap\overline{K}) + P(\overline{B}\cap A\cap K) + P(B\cap\overline{A}\cap K)] \\
&= 0{,}80\cdot 0{,}70\cdot 0{,}50 + 0{,}20\cdot 0{,}70\cdot 0{,}50 + 0{,}80\cdot 0{,}30\cdot 0{,}50 = 0{,}47 \\
P(A_3) &= P(A\cap B\cap K) = 0{,}28
\end{aligned}
$$

Εφαρμόζοντας την ολική πιθανότητα επιτυγχάνουμε το ζητούμενο,

$$
\begin{aligned}
P(E) &= P(E|A_0)P(A_0) + P(E|A_1)P(A_1) + P(E|A_2)P(A_2) + P(E|A_3)P(A_3) \\
&= 0\cdot 0{,}03 + 0{,}50\cdot 0{,}22 + 0{,}80\cdot 0{,}47 + 1\cdot 0{,}28 = 0 + 0{,}11 + 0{,}376 + 0{,}28 = 0{,}766
\end{aligned}
$$

11) Έλεγχος ποιότητας. Συνεχίζοντας το παράδειγμα 3.9 κατά πόσο περιορίζεται το ποσοστό της ερώτησης (β), δηλαδή των ράβδων που καταστρέφονται ενώ στην πραγματικότητα είναι καλές, αν σε κάθε ράβδο γίνονται δύο έλεγχοι από το ίδιο μηχάνημα, και μία ράβδος καταστρέφεται μόνο όταν και οι δύο έλεγχοι είναι θετικοί;

Λύση

Συμβολίζοντας τα γεγονότα όπως και στο παράδειγμα 3.9,

A_1={η ράβδος είναι ελαττωματική, έχει ρωγμές},
A_2={η ράβδος είναι καλή, δεν έχει ρωγμές},
B_1 ={ο πρώτος έλεγχος είναι θετικός },
B_2 ={ο δεύτερος έλεγχος είναι θετικός },

Υπολογίζω την ολική πιθανότητα οι δύο έλεγχοι να είναι θετικοί ,

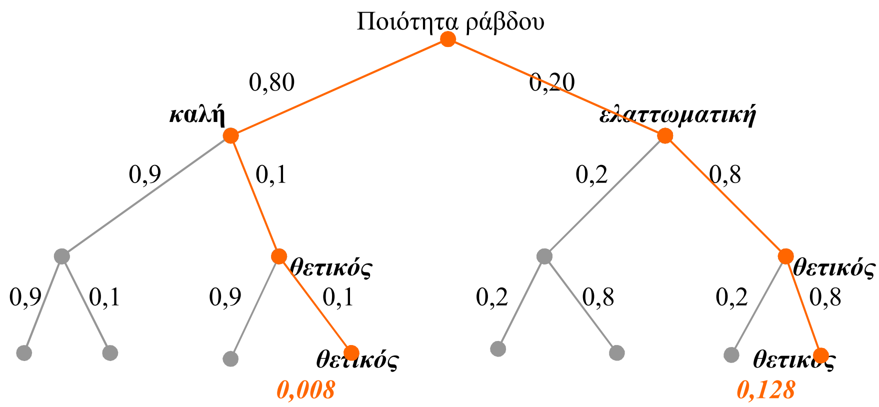
$$
\begin{aligned}
P(B_1\cap B_2) &= P(B_1\cap B_2|A_1)P(A_1) + P(B_1\cap B_2|A_2)P(A_2) \\
&= 0{,}8\cdot 0{,}8\cdot 0{,}2 + 0{,}1\cdot 0{,}1\cdot 0{,}8 = 0{,}128 + 0{,}008 = 0{,}136
\end{aligned}
$$

όπου $P(B_1 \cap B_2 | A_1) = P(B_1 | A_1) P(B_2 | A_1)$, διότι οι δύο έλεγχοι είναι ανεξάρτητοι για δεδομένη την ποιότητα των ράβδων. Καταλήγουμε λοιπόν, ότι με δύο ελέγχους καταστρέφεται το 13,6% της παραγωγής.

Ζητούμε την πιθανότητα ότι η ράβδος είναι καλής ποιότητας δεδομένου ότι και οι δύο έλεγχοι βγήκαν θετικοί, $P(A_2 | B_1 \cap B_2)$. Εφαρμόζοντας το θεώρημα Bayes έχω

$$P(A_2 | B_1 \cap B_2) = \frac{P(B_1 \cap B_2 | A_2) P(A_2)}{P(B_1 \cap B_2)} = \frac{0{,}1 \cdot 0{,}1 \cdot 0{,}8}{0{,}136} = \frac{0{,}008}{0{,}136} = \frac{8}{136} \cong 0{,}06$$

Με έναν διάγραμμα δένδρου γίνεται ευκολότερη η κατανόηση της απάντησης



Με δύο ελέγχους μόνο το 6% από αυτές που καταστρέφονται είναι καλής ποιότητας. Με έναν έλεγχο (βλέπε παράδειγμα 3.9) το 33% από αυτές που καταστρέφονται είναι κακής ποιότητας

Ασκήσεις προς λύση

1) Ένα κουτί περιέχει 4 ελαττωματικά και 6 καλά ανταλλακτικά.
(α) Παίρνουμε δύο ανταλλακτικά τυχαία από το κουτί και το πρώτο από αυτά ελέγχεται και βρίσκεται ότι είναι καλό, ποια η πιθανότητα ότι και το δεύτερο ανταλλακτικό είναι καλό;
(β) Επιλέγουμε ένα ανταλλακτικό από το κουτί και ελέγχουμε. Η διαδικασία αυτή επαναλαμβάνεται χωρίς επανάθεση έως ότου βρεθούν 4 ελαττωματικά. Ποια η πιθανότητα ότι το $4^{ο}$ ελαττωματικό θα βρεθεί στη $5^{η}$ επιλογή;

2) Έστω δύο γεγονότα A και B ορισμένα στον ίδιο δειγματικό χώρο. Δεδομένου ότι η πραγματοποίηση του A αυξάνει την πιθανότητα πραγματοποίησης του B ($P(B/A)>P(B)$), αποφανθείτε αν η πραγματοποίηση του B αυξάνει την πιθανότητα πραγματοποίησης του A ($P(A|B) > P(A)$).

3) Ένα κουτί περιέχει κ μαύρες και λ κόκκινες σφαίρες. Επιλέγουμε τυχαία μία σφαίρα από το κουτί και την επανατοποθετούμε μαζί με άλλες μ σφαίρες του ιδίου χρώματος. Επιλέγουμε για δεύτερη φορά μία σφαίρα από το κουτί και διαπιστώνουμε ότι η σφαίρα είναι κόκκινη. Ποια η πιθανότητα ότι η σφαίρα στην πρώτη επιλογή ήταν κόκκινη;

4) Δύο κάλπες Α και Β περιέχουν ν σφαίρες αριθμημένες με 1, 2, 3, …, ν εκάστη. Από κάθε κάλπη επιλέγεται τυχαία μία σφαίρα. Ποια η πιθανότητα η αρίθμηση της σφαίρας που επιλέγεται από την κάλπη Β να είναι μικρότερη αυτής της Α.
Βοήθημα: $1+2+3+\ldots+\nu = \nu(\nu+1)/2$

5) Το 60% ενός πληθυσμού έχει αυτοκίνητο, ενώ το 80% έχει μοτοσυκλέτα. Να βρεθεί το ελάχιστο ποσοστό ιδιοκτητών που αυτοκινήτου που διαθέτει και μοτοσυκλέτα.

6) Το *τέστ Παπανικολάου* έχει πιθανότητα σωστής διάγνωσης 95%. Δοθέντος ότι 8% του πληθυσμού των γυναικών έχει την ασθένεια και ότι το *τέστ* έδωσε αρνητικό αποτέλεσμα σε μία περίπτωση, ποια η πιθανότητα η εξετασθείς να έχει παρόλα αυτά την ασθένεια:

Για λόγους καλύτερης διάγνωσης ένα δεύτερο *τέστ* έδωσε αρνητικό αποτέλεσμα στην ίδια περίπτωση. Πόσο σίγουροι είμαστε ότι η γυναίκα δεν πάσχει από την ασθένεια;

7) Σε μία έκθεση ζωγραφικής υπάρχουν 12 πίνακες από τους οποίους μόνο οι 10 είναι αυθεντικοί. Ένας επισκέπτης επιλέγει τυχαία έναν πίνακα και πριν αποφασίσει να τον αγοράσει ζητά την γνώμη από έναν εμπειρογνώμονα, ο οποίος κρίνει σωστά για την αυθεντικότητα του πίνακα με πιθανότητα 80%.
(α) Δεδομένου ότι ο εμπειρογνώμων κρίνει τον πίνακα ως αυθεντικό , ποια η πιθανότητα ότι πράγματι είναι;
(β) Εάν ο εμπειρογνώμων κρίνει τον πίνακα ως αντίγραφο, τότε ο επισκέπτης επιστρέφει τον πίνακα και επιλέγει έναν από τους υπόλοιπους 11 πίνακες. Ποια η πιθανότητα ότι η νέα του επιλογή είναι αυθεντικός πίνακας;

8) Για δύο γεγονότα A και B ισχύει

$$P(A \cup B) = 0{,}7, P(A) = 0{,}4 \, και \, P(B) = p$$

(α) Για ποια τιμή του p τα γεγονότα A και B είναι ανεξάρτητα;
(β) Για ποια τιμή του p τα γεγονότα A και B είναι ξένα;

9) Τρία αντικείμενα Α, Β και Γ τοποθετούνται τυχαία στην σειρά. Συμβολίζουμε με

W: Το γεγονός, ότι το αντικείμενο Β βρίσκεται δεξιά (όχι κατ' ανάγκη δίπλα) του Α

R: Το γεγονός, ότι το αντικείμενο Γ βρίσκεται δεξιά (όχι κατ' ανάγκη δίπλα) του Α.

Εξετάστε αν τα γεγονότα W και R είναι ανεξάρτητα.

10) Δύο αφηρημένοι μηχανικοί φεύγουν από το γραφείο τους ένα πρωί με τις ομπρέλες τους για να επισκεφθούν τρία καταστήματα. Αν σε κάθε κατάστημα που επισκέπτονται ο καθένας από αυτούς(ανεξάρτητα από τον άλλον) ενδέχεται να ξεχάσει την ομπρέλα του με πιθανότητα 0,25, ποια η πιθανότητα ότι τουλάχιστον ένας από αυτούς επιστρέφει στο γραφείο χωρίς την ομπρέλα του;

11) Ο ακόλουθος πίνακας περιέχει το ιστορικό των χαρακτηριστικών από 940 εκτελέσεις προγραμμάτων σε ένα υπολογιστή

		Έξτρα μνήμη	
		Όχι	*Ναι*
Επιλογή υψηλής ταχύτητας επεξεργαστή	Ναι	514	68
	Όχι	112	246

Αν A είναι το γεγονός ότι μία εκτέλεση απαιτεί υψηλή ταχύτητα επεξεργαστή, και αν B είναι το γεγονός ότι μία εκτέλεση απαιτεί έξτρα μνήμη, προσδιορίστε τις ακόλουθες πιθανότητες.

(α) $P(A \cup B)$, $P(A \cap B)$, $P(\overline{A} \cap B)$, $P(\overline{A} \cap \overline{B})$.

(β) Ποια η πιθανότητα ότι μία εκτέλεση απαιτεί έναν υψηλής ταχύτητας επεξεργαστή δεδομένου ότι η εκτέλεση απαιτεί έξτρα μνήμη;

12) Από τρία νομίσματα εκ των οποίων το ένα φέρει και στις δύο όψεις 'Γράμμα' (Γ) ενώ τα άλλα είναι κανονικά, επιλέγουμε τυχαία ένα και το ρίχνουμε τρίς φορές. Ποια η πιθανότητα:

(α) Να εμφανιστούν 3 γράμματα ($\Gamma\Gamma\Gamma$).

(β) Δοθέντος ότι οι 3 ρίψεις έφεραν ($\Gamma\Gamma\Gamma$), ποια η πιθανότητα το νόμισμα που επιλέγει να φέρει 'Γράμμα' και στις δύο όψεις;

13) Ένας οδηγός μοτοσικλέτας αλλάζει λωρίδα κυκλοφορίας κάθε 15 δευτερόλεπτα. έστω ότι σε κάποια χρονική στιγμή εισέρχεται σε κάποια λωρίδα. Αν ο δρόμος έχει 4 λωρίδες, ποια η πιθανότητα να βρεθεί στην ίδια λωρίδα ένα λεπτό αργότερα;

14) Ένα τετράγωνο χωρίζεται με οριζόντιες γραμμές σε κ ίσες ζώνες. Σε κάθε μία από αυτές τοποθετείται τυχαία ένα σημείο. Στη συνέχεια φέρονται κάθετες γραμμές έτσι ώστε το τετράγωνο χωρίζεται σε κ ίσες στήλες. Ποία η πιθανότητα ότι κάθε στήλη θα περιέχει μόνο ένα σημείο;

15) Η πιθανότητα το παιδί μίας οικογένειας να είναι αγόρι είναι 50%. Πόσα παιδιά πρέπει να έχει μία οικογένεια , ώστε με πιθανότητα πάνω από 90% να έχει παιδιά και των δύο φύλων.

16) Ένα ζάρι ρίχνεται 10 φορές και έστω M η μέγιστη και m η ελάχιστη ένδειξη. Υπολογίστε την πιθανότητα: $\mathrm{P}(\{\mathrm{m}=2\} \cap \{\mathrm{M} \leq 5\})$.

17) Σε ένα κύκλο ακτίνας $\lambda N^{1/2}$ κατανέμουμε τυχαία N σημεία. Να αποδειχθεί ότι για σταθερό λ και μεγάλο N, η πιθανότητα ότι κανένα σημείο δεν πέφτει σε δοθείσα περιοχή εμβαδού A είναι προσεγγιστικά

$$e^{-A/\pi\lambda^2} .$$

Προβλήματα προς λύση

1) Αστοχία κρεμαστής γέφυρας. Ένα προκαταρκτικό σχέδιο μίας κρεμαστής γέφυρας αποτελείται κυρίως από 3 κατακόρυφα σχοινιά και 4 δοκούς. Από μελέτη των συνθηκών φόρτισης και της αντοχής των διάφορων στατικών μελετών, προκύπτει ότι η πιθανότητα αστοχίας για κύρια δοκό είναι 0,002 και για κατακόρυφο σχοινί 0,006. Υποθέτουμε ότι οι αστοχίες μεταξύ των στατικών μελών είναι στατιστικά ανεξάρτητες.

Να υπολογιστούν,

(α) η πιθανότητα αστοχίας της γέφυρας λόγω δοκού ή δοκών,

(β) η πιθανότητα αστοχίας της γέφυρας λόγω σχοινιού ή σχοινιών,

(γ) η πιθανότητα αστοχίας της γέφυρας.

2) Συχνότητα στροφής σε διασταύρωση. Από παρατηρήσεις αυτοκινήτων που περνούν μία διασταύρωση, προκύπτει ότι, για ένα αυτοκίνητο που πλησιάζει τυχαία στην διασταύρωση η πιθανότητα να προχωρήσει ευθεία είναι ίση με την πιθανότητα να στρίψει. Αριστερές στροφές συμβαίνουν με διπλάσια συχνότητα των δεξιών στροφών.

Ένα αυτοκίνητο πλησιάζει στην διασταύρωση,

(α) ποια η πιθανότητα να στρίψει δεξιά;

(β) αν γνωρίζουμε ότι θα στρίψει, ποια η πιθανότητα να στρίψει δεξιά;

3) Εντόπιση τηλεφωνικής κλήσης. Μία τηλεφωνική κλήση τυχαία στο διάστημα (0, T). Να υπολογιστεί η πιθανότητα η τηλεφωνική κλήση να συμβεί στο διάστημα $[t_1, t_2]$

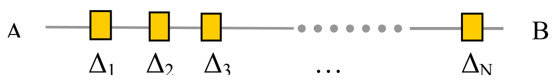
- αν δεν έχουμε άλλη πληροφορία
- αν γνωρίζουμε ότι δεν έγινε στο διάστημα $[0, t_0]$.

4) Επιλογή ελαττωματικής σιδηροδοκού. Ένα κιβώτιο περιέχει 25 σιδερένιες σιδηροδοκούς από τις οποίες οι 5 φέρουν ράγισμα.

(α) Εάν πάρουμε τυχαία δύο σιδηροδοκούς (χωρίς αντικατάσταση), ποια η πιθανότητα ότι η δεύτερη φέρει ράγισμα;

(β) Εάν πάρουμε τυχαία τρεις σιδηροδοκούς (χωρίς αντικατάσταση), ποια η πιθανότητα ότι η τρίτη φέρει ράγισμα;

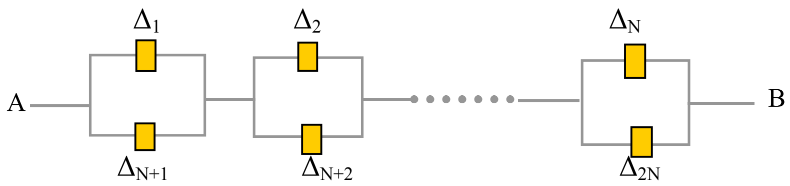
7) Ν όμοιοι διακόπτες Δ_1, Δ_2, Δ_3, …, Δ_N λειτουργούν ανεξάρτητα και συνδέουν τα σημεία Α και Β σύμφωνα με το παρακάτω σχήμα.

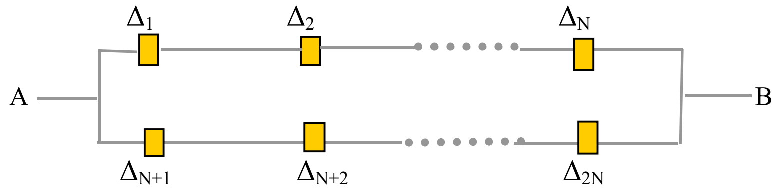


(α) Να υπολογιστεί η πιθανότητα διακοπής των σημείων Α και Β, αν η πιθανότητα διακοπής για κάθε διακόπτη ισούται με p.
(β) Για να μειωθεί η πιθανότητα διακοπής ρεύματος μεταξύ των σημείων Α και Β προτείνονται δύο λύσεις:

- κάθε διακόπτης διπλασιάζεται,
- όλο το σύστημα διπλασιάζεται,

όπως φαίνεται και στο σχήμα που ακολουθεί.





Να υπολογιστούν οι πιθανότητες διακοπής για τις δύο προτάσεις και να συγκριθούν τα αποτελέσματα βάσει των τύπων και φαινομενολογικά.

5) Αξιοπιστία επικοινωνίας. Ένα μήνυμα μπορεί να μεταδοθεί μέσα από n κανάλια επικοινωνίας, όπου n_1 κανάλια είναι σε άριστη κατάσταση, n_2 είναι σε καλή κατάσταση, n_3 είναι σε μέτρια κατάσταση και n_4 κανάλια δεν είναι σε καλή κατάσταση ($n= n_1+ n_2+ n_3+ n_4$). Η πιθανότητα ένα μήνυμα να μεταδοθεί σωστά είναι p_1, p_2, p_3, p_4 για τα τέσσερα είδη καναλιών αντίστοιχα. Για να αυξήσουμε την αξιοπιστία της μετάδοσης των μηνυμάτων, κάθε μήνυμα στέλνεται από δύο διαφορετικά κανάλια, τα οποία επιλέγονται τυχαία.

Να ευρεθεί η πιθανότητα σωστής μετάδοσης του μηνύματος.

6) Σύστημα ασφαλείας. Σε σύγχρονες μονάδες παραγωγής υπάρχουν συστήματα ασφαλείας που αποτελούνται από κυκλώματα διακοπτών, τα οποία μπορούν να θέσουν τέρμα στη λειτουργία της μονάδας διακόπτοντας την ροή του ρεύματος σε επικίνδυνες καταστάσεις.

ΟΙ διακόπτες λειτουργούν ανεξάρτητα και αποτυγχάνουν να λειτουργήσουν σωστά κατά δύο τρόπους

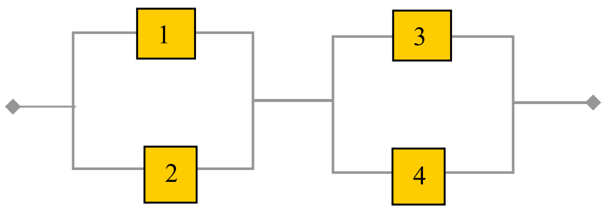
1. όταν δεν κλείνουν το ρεύμα σε περίπτωση βλάβης της μονάδας με πιθανότητα q_1 έκαστος
2. όταν κλείνουν το ρεύμα σε περίπτωση που δεν υπάρχει βλάβη με πιθανότητα q_2 έκαστος

Το όλο σύστημα αποτυγχάνει δηλαδή, όταν

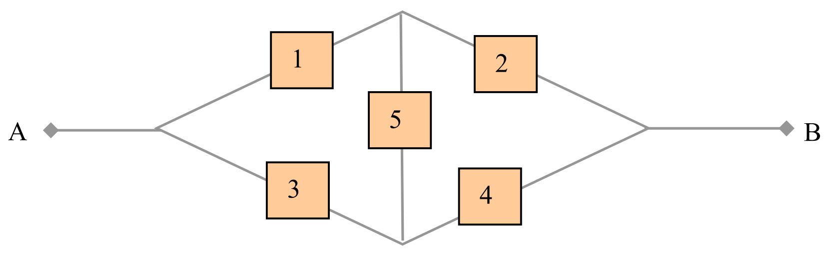
- δεν επισημαίνεται ο κίνδυνος όταν πράγματι υπάρχει, η πιθανότητα του γεγονότος αυτού συμβολίζεται με Q_1,
- ή εσφαλμένα επισημαίνεται στην περίπτωση που δεν υπάρχει η πιθανότητα του γεγονότος αυτού συμβολίζεται με Q_2.

(α) Να βρεθούν οι πιθανότητες Q_1, Q_2 για το σύστημα με δύο διακόπτες εν σειρά.

(β) Να βρεθούν οι πιθανότητες Q_1, Q_2 του συστήματος ασφαλείας με δύο μπλόκ εν σειρά αλλά διακόπτες εν παραλλήλω.



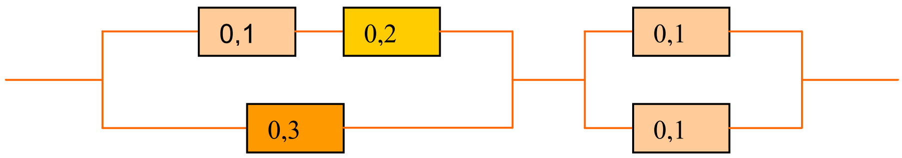
7) Κύκλωμα ρεύματος. Ο κάθε διακόπτης στο κύκλωμα του σχήματος έχει πιθανότητα p να είναι κλειστός (διέρχεται ρεύμα).



(α) εκφράστε το γεγονός « διέρχεται ρεύμα από το Α στο Β « συναρτήσει των γεγονότων Δ_κ, « ο διακόπτης κ είναι κλειστός», όπου κ=1, 2, 3, 4, 5.
(β) Με την προϋπόθεση ότι τα γεγονότα Δ_κ είναι ανεξάρτητα, ποια η πιθανότητα ροής ρεύματος από το Α στο Β;

8) Δίκτυο ύδρευσης. Η τροφοδοσία νερού σε μία αγροτική περιοχή το καλοκαίρι διενεργείται μέσω των κεντρικών παροχών Π1, Π2 και Π3. Κάθε παροχή εάν λειτουργεί επιτυχώς καλύπτει το 50% των αναγκών στις θερινές καλλιέργειες. Είναι γνωστό ότι οι επιτυχείς λειτουργίες των παροχών είναι γεγονότα στατιστικά ανεξάρτητα με αντίστοιχες πιθανότητες $P(\Pi 1)=0{,}9$, $P(\Pi 2)=0{,}7$ $P(\Pi 3)=0{,}6$.
(α) Ποια η πιθανότητα ότι το καλοκαίρι θα υπάρξει ικανοποιητική παροχή νερού για το πότισμα των αγροτικών καλλιεργειών;
(β) Κάποιο καλοκαίρι δεν καλύπτονται ικανοποιητικά οι ανάγκες ποτίσματος στην περιοχή. Ποια η πιθανότητα ότι ευθύνονται για αυτό οι παροχές Π2 και Π3;

9) Λειτουργία παραγωγικού κυκλώματος. Το κύκλωμα παρακάτω λειτουργεί αν υπάρχει τουλάχιστον ένα μονοπάτι από ενεργούντα εξαρτήματα από αριστερά προς τα δεξιά. Είναι γνωστό ότι τα εξαρτήματα αποτυγχάνουν να



λειτουργήσουν ανεξάρτητα μεταξύ τους με πιθανότητα όπως φαίνεται στο σχήμα. Ποια η πιθανότητα ότι το κύκλωμα λειτουργεί.;

10) Αξιοπιστία δομικών μηχανών. Κατά την διάρκεια ενός μεγάλου κατασκευαστικού έργου εκτιμάται ότι το 20% από τις δομικές μηχανές νέου τύπου και το 80% από τις δομικές μηχανές παλαιού τύπου θα παρουσιάσουν βλάβη. Από τις δομικές του εργοταξίου το 40% είναι νέου τύπου. Ποιο ποσοστό δομικών μηχανών θα λειτουργήσει κατά την διάρκεια του έργου χωρίς βλάβη;

11) Λήψη σήματος. Τηλεγραφικά σήματα (•) και (–) στέλνονται με αναλογία 1:4 αντίστοιχα. Λόγω θορύβου στη γραμμή μεταφοράς μια τελεία (•) καταφθάνει σαν παύλα (–) με πιθανότητα 1/4 και αντίστροφα μία παύλα (–) καταφθάνει σαν τελεία (•) με πιθανότητα 1/3. Εάν ένα σήμα καταφθάνει σαν παύλα(–), να βρεθεί η πιθανότητα να έχει μεταφερθεί σωστά (να εστάλη παύλα (–)).

12) Ανίχνευση ραγίσματος. Από τις συγκολλήσεις σε μία μεγάλη κατασκευή το 10% έχουν ράγισμα. Όταν υπάρχει υπόνοια ότι μία συγκόλληση έχει ράγισμα γίνονται δύο ανεξάρτητοι έλεγχοι με ανιχνευτικό μηχάνημα. Σε κάθε

έλεγχο συγκόλλησης το μηχάνημα επισημαίνει ράγισμα με πιθανότητα 0,9 ή 0,1 όταν υπάρχει πράγματι ράγισμα ή δεν υπάρχει αντίστοιχα. Να ευρεθεί η πιθανότητα ότι σε μία συγκόλληση
(α) και στους δύο ελέγχους το μηχάνημα επισημαίνει ράγισμα,
(β) δεδομένου ότι και οι δύο έλεγχοι επισημαίνουν ράγισμα, η συγκόλληση έχει πράγματι ράγισμα.

13) Διάσπαση σωματιδίου. Στα πλαίσια μιας φυσικής αντίδρασης ένα αρχικό σωματίδιο μπορεί να διασπασθεί σε δύο άλλα σωματίδια, να μεταβληθεί σε νέο σωματίδιο , ή τέλος, να εξαφανισθεί. Οι πιθανότητες είναι 1/3, 1/6 και 1/2 αντίστοιχα. Τα σωματίδια που δημιουργούνται συμπεριφέρονται όπως και το αρχικό.

Έστω ότι η μεταβλητή X_κ συμβολίζει το πλήθος των σωματιδίων της κ γενιάς(μετά από κ αντιδράσεις). Υπολογίστε την πιθανότητα $P(X_1 = 1 \mid X_2 = 1)$.

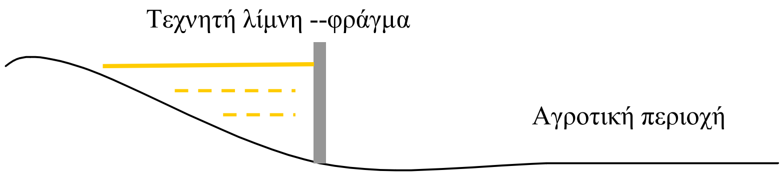
14) Ποιότητα πεζοδρομίου. Η Δημαρχία μιας πόλης πριν πληρώσει την κατασκευαστική εταιρεία για το έργο ενός πεζοδρομίου 250 μέτρων , ελέγχει με ένα όργανο το πάχος του πεζοδρομίου ανά τμήμα μήκους 0,5 μέτρα, για να επιβεβαιώσει ότι πληροί τις προδιαγραφές. Αν το μετρούμενο πάχος ενός τμήματος είναι μικρότερο από 10 πόντους δεν γίνεται αποδεκτό. Είναι γνωστό από το παρελθόν ότι το 85 % των τμημάτων πεζοδρομίου που κατασκευάζει η εταιρεία πληρούν τις προδιαγραφές. Ακόμη, το όργανο που μετρά το πάχος του πεζοδρομίου είναι αξιόπιστο μόνο 75%, δηλαδή, σε κάθε έλεγχο για το πάχος ενός τμήματος πεζοδρομίου το συμπέρασμα είναι λανθασμένο με πιθανότητα 0,25.
(α) Ποια η πιθανότητα ότι ένα τμήμα πεζοδρομίου γίνεται αποδεκτό με βάση την μέτρηση του οργάνου;
(β) Ποια η πιθανότητα ότι πληροί τις προδιαγραφές κάθε τμήμα πεζοδρομίου το οποίο γίνεται αποδεκτό σύμφωνα με τον έλεγχο του οργάνου,

15) Αξιολόγηση έργων. Μία επιτροπή με τρεις κριτές Κ1, Κ2, και Κ3 πρόκειται να αξιολογήσουν τις προσφορές τεσσάρων κατασκευαστικών εταιριών (Ε1, Ε2, Ε3, και Ε4) για ένα μεγάλο έργο. Ο κάθε κριτής ιεραρχεί τις προσφορές θέτοντας 3 για την καλύτερη προσφορά και 2, 1, 0 για τις άλλες, και έτσι, αθροίζοντας για κάθε εταιρία τους βαθμούς που παίρνει από τους κριτές, αναδεικνύεται η επιτυχούσα προσφορά. Υποθέτουμε ότι οι κριτές δεν είναι σε θέση να ξεχωρίσουν τις καλές προσφορές και βαθμολογούν τυχαία.
(α) Ποια η πιθανότητα ότι η εταιρία Ε1 θα έχει συνολικό βαθμό 4;
(β) Ποια η πιθανότητα ότι κάθε εταιρία θα πάρει συνολικά 4;
(γ) Ποια η πιθανότητα ότι κάθε εταιρία θα πάρει συνολικά 5 ή λιγότερο;

16) Ανίχνευση μετάλλου. Για την ανίχνευση ενός ειδικού μετάλλου σε μία αχανή ερημική έκταση χρησιμοποιούνται δύο συσκευές Α και Β πάνω σε ένα δορυφόρο που περνά από την περιοχή αυτή. Οι συσκευές ευαισθητοποιούνται και δίνουν σήμα όταν ο δορυφόρος περνά από περιοχή η οποία εμπεριέχει αυτό το ειδικό μέταλλο, όχι όμως με πλήρη επιτυχία. Οι δύο συσκευές Α, Β ευαισθητοποιούνται με πιθανότητα 0,80 όταν στην περιοχή υπάρχει το μέταλλο. Ακόμη, όταν οι συσκευές αυτές δίνουν σήμα ύπαρξης του μετάλλου η ακριβής εντόπιση δεν είναι πάντοτε επιτυχής. Έτσι, η Α συσκευή εντοπίζει ακριβώς το σημείο ύπαρξης μετάλλου με πιθανότητα 70%, ενώ η αντίστοιχη πιθανότητα για την Β συσκευή είναι 50%. Όταν και οι δύο συσκευές δίνουν σήμα για το μέταλλο η ακριβής εντόπισή του είναι 0,90.
(α) Ποια είναι η πιθανότητα ότι καθώς ο δορυφόρος περνά από μία περιοχή με το μέταλλο θα ευαισθητοποιηθεί μόνο μία συσκευή;
(β) Ποια είναι η πιθανότητα ότι θα ανακαλυφθεί μία περιοχή που περιέχει το μέταλλο;
(γ) Ποια είναι η πιθανότητα να εντοπισθεί η ακριβής θέση ύπαρξης του μετάλλου;

17) Αντιπλημμυρικό φράγμα. Μία αγροτική περιοχή προστατεύεται από τις πιθανές πλημμύρες καταιγίδων από ένα φράγμα το οποίο έχει σχεδιασθεί για την πλημμύρα των 20 ετών (δηλαδή, η πιθανότητα ότι το φράγμα θα ξεχειλίσει από καταιγίδες μέσα σε ένα χρόνο είναι 1/20).

Η περιοχή αυτή είναι σεισμογενής και κατά μέσο όρο συμβαίνουν 4 ισχυροί σεισμοί ανά 100 χρόνια (η πιθανότητα να συμβεί σεισμός μέσα σε ένα χρόνο είναι 1/25). Έχει παρατηρηθεί ότι μία στις δέκα φορές που συμβαίνει ισχυρός σεισμός το φράγμα υποχωρεί και η περιοχή πλημμυρίζει.
(α) Ποια είναι η πιθανότητα ότι δεν θα πλημμυρίσει η περιοχή μέσα στον επόμενο χρόνο;
(β) Αν μέσα στον επόμενο χρόνο συμβεί σεισμός, ποια η πιθανότητα ότι θα πλημμυρίσει η περιοχή;



20) Υδροδότηση. Η υδροδότηση μιας πόλης το καλοκαίρι γίνεται από δύο λίμνες Α και Β. Η αποθηκευμένη ποσότητα νερού σε κάθε λίμνη εξαρτάται από την βροχόπτωση κατά την διάρκεια του χειμώνα. Έτσι, η μέση ποσότητα νερού στην λίμνη Α και στην λίμνη Β ξεπερνιέται με πιθανότητα 0,70 και 0,60 αντίστοιχα. Ακόμη, είναι γνωστό ότι η πιθανότητα να ξεπερασθεί η μέση ποσότητα και στις δύο λίμνες είναι 0,30.

Η ζήτηση νερού στην πόλη εξαρτάται προφανώς από αρκετούς παράγοντες, είναι γνωστό ότι η πόλη έχει ικανοποιητική παροχή νερού το καλοκαίρι με πιθανότητα,

- 0,95 αν και οι δύο λίμνες ξεπεράσουν την μέση ποσότητα,
- 0,6 αν μόνο μία λίμνη ξεπεράσει την μέση ποσότητα,
- 0,2 αν σε καμία λίμνη δεν ξεπερασθεί η μέση ποσότητα.

Να υπολογιστεί η πιθανότητα ότι η πόλη το επόμενο καλοκαίρι θα έχει ικανοποιητική παροχή νερού.

21) Κίνδυνος πυρκαγιάς σε δάσος. Σένα δάσος η πιθανότητα να συμβεί πυρκαγιά το καλοκαίρι είναι 0,20. Ενδέχεται, ακόμη, να συμβούν και δύο πυρκαγιές μέσα στο ίδιο καλοκαίρι με πιθανότητα 0,05. Όταν συμβεί πυρκαγιά, ανάλογα με τον άνεμο που φυσά στην περιοχή και από την ετοιμότητα της πυροσβεστικής υπηρεσίας, το δάσος ενδέχεται να καεί ολοσχερώς με πιθανότητα 0,30.
(α) Ποια η πιθανότητα ότι το δάσος δεν θα καεί το επόμενο καλοκαίρι;
(β) Στην περιοχή υπάρχουν τρία παρόμοια δάση όπου επικρατούν οι ίδιες συνθήκες. Ποια η πιθανότητα ότι κανένα δάσος δεν θα καεί το επόμενο καλοκαίρι;

22) Ζημίες στους δοκούς από σεισμό. Όταν συμβαίνει σεισμός οι δοκοί μιας πολυκατοικίας ενδέχεται να υποστούν ζημίες, οι οποίες μπορεί να είναι καταστροφικές για την πολυκατοικία. Ανάλογα με την ένταση του σεισμού, οι δοκοί παθαίνουν ζημιά με πιθανότητες 0,10, 0,25, και 0,40 όταν η ένταση του σεισμού είναι μικρή, μέτρια και ισχυρή αντίστοιχα. Η συχνότητα εμφάνισης ισχυρού σεισμού είναι 5 φορές μικρότερη από την εμφάνιση μέτριου σεισμού, ενώ ασθενείς σεισμοί συμβαίνουν στην περιοχή δύο φορές συχνότερα απ’ ότι οι μέτριοι σεισμοί.
(α) Ποια η πιθανότητα ότι ο επόμενος σεισμός θα είναι ισχυρής έντασης;
(β) Ποια η πιθανότητα ότι στον επόμενο σεισμό θα έχουμε ζημίες στους δοκούς της πολυκατοικίας;
(γ) Γίνονται δύο σεισμοί. Ποια η πιθανότητα ότι δεν θα υπάρξουν ζημίες σους δοκούς της πολυκατοικίας;

22) Συγκολλήσεις μετάλλου. Μία μεταλλική κατασκευή αποτελείται κυρίως από δύο μέρη Α και Β. Το Α μέρος περιλαμβάνει δύο τμήματα , το τμήμα Α1(μήκος 20 μέτρα) και το τμήμα Α2(μήκος 30 μέτρα), ενώ το Β είναι μόνο ένα τμήμα μήκους 50 μέτρων. Στις συγκολλήσεις των στοιχείων της κατασκευής υπάρχουν ραγίσματα με πιθανότητα $p=0,005\cdot μήκος$ για τα τμήματα του Α και $p=0,008\cdot μήκος$ για το τμήμα Β.

Αν βρεθούν ραγίσματα σε ένα τμήμα του Α, διπλασιάζεται η αρχική πιθανότητα ύπαρξης ραγισμάτων στο άλλο. Ακόμη, αν βρεθούν ραγίσματα στο Α μέρος της κατασκευής, τετραπλασιάζεται η αρχική πιθανότητα ύπαρξης ραγισμάτων στο Β μέρος.

Έστω ότι $A1$, $A2$, B συμβολίζουν τα γεγονότα ύπαρξης ραγισμάτων στα αντίστοιχα τμήματα.

(α) Το γεγονός ότι υπάρχει ράγισμα στο μέρος Α να το συμβολίσετε με άλγεβρα των γεγονότων $A1$, $A2$, και να υπολογίσετε την πιθανότητά του.

(β) Το γεγονός ότι υπάρχει ράγισμα στην κατασκευή να το συμβολίσετε με άλγεβρα των γεγονότων $A1$, $A2$, B, και να υπολογίσετε την πιθανότητά του.

(γ) Αν βρεθούν ραγίσματα στην κατασκευή, ποια η πιθανότητα ότι βρέθηκαν μόνο στο μέρος Α;

Άλλα ενδιαφέροντα προβλήματα μηχανικού

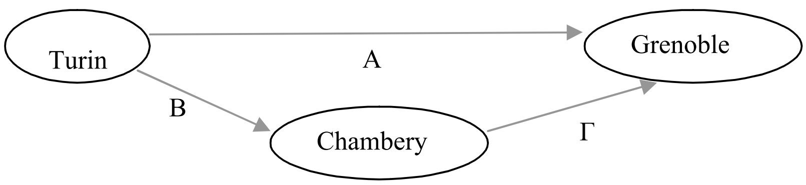
1) Σύστημα μεγάλων δρόμων. Για να φθάσει κανείς στο Grenoble της Γαλλίας από το Turin της Ιταλίας μπορεί να ακολουθήσει δύο δρόμους. Ο πρώτος συνδέει απ' ευθείας Turin και Grenoble, ενώ ο δεύτερος περνά από το Chambery της Γαλλίας. Κατά την διάρκεια του χειμώνα λόγω κακών καιρικών συνθηκών το ταξίδι από Turin σε Grenoble δεν είναι πάντα δυνατό επειδή κάποια τμήματα του δρόμου μπορεί να είναι κλειστά. Δηλώνουμε με A, B, και Γ αντίστοιχα τα γεγονότα ότι οι δρόμοι από Turin για Grenoble, Turin για Chambery, και Chambery για Grenoble είναι ανοικτοί.

Ένας ταξιδιώτης που πρόκειται να ταξιδέψει από Turin για Grenoble, ακούει το δελτίο καιρού. Εάν προβλέπεται χιόνι στις Άλπεις, μπορεί κανείς να υποθέσει (με βάση δεδομένων στο παρελθών) ότι $P(A)$=0,6, $P(B)$=0,7, $P(\Gamma)$=0,4, $P(\Gamma/B)$=0,5, και $P(A/B\Gamma)$=0,4.

(α) Ποια είναι η Πιθανότητα ότι ο ταξιδιώτης θα φθάσει την επομένη ημέρα στο Grenoble της Γαλλίας;

(β) Ποια είναι η πιθανότητα ότι ο ταξιδιώτης μπορεί να οδηγήσει από Turin για Grenoble δια μέσου του Chambery;

(γ) Ποιο δρόμο πρέπει να ακολουθήσει για να μεγιστοποιήσει την πιθανότητα να φθάσει στο Grenoble ;.



Λύση

(α) Το γεγονός ότι θα φθάσει ο ταξιδιώτης από το Turi Grenoble το συμβολίζω με R και μπορεί να γραφεί με την ακόλουθη άλγεβρα,

$$R = A \cup (B \cap \Gamma).$$

Έτσι, μπορώ να υπολογίσω την πιθανότητα του R, εφαρμόζοντας το θεώρημα της ένωσης,

$$\begin{aligned} P(R) = P[A \cup (B \cap \Gamma)] &= P(A) + P(B \cap \Gamma) - P(A \cap B \cap \Gamma) \\ &= 0,6 + P(\Gamma)P(\Gamma|B) - P(B)P(\Gamma|B\}P(A|B\Gamma) \\ &= 0,6 + 0,7 \cdot 0,5 - 0,7 \cdot 0,5 \cdot 0,4 = 0,6 + 0,35 - 0,14 = 0,81 \end{aligned}$$

(β) Προφανώς, η πιθανότητα είναι

$$P(B \cap \Gamma) = P(\Gamma)P(\Gamma | B) = 0{,}7 \cdot 0{,}5 = 0{,}35$$

(γ) Συγκρίνοντας τις αντίστοιχες πιθανότητες, ο δρόμος ευθεία στο Grenoble έχει μεγαλύτερη πιθανότητα να είναι ανοικτός.

2) Φράγμα τεχνητής λίμνης. Ένα φράγμα τεχνητής λίμνης βρίσκεται σε μία σεισμογενή περιοχή. Όταν συμβεί σεισμός, η πιθανότητα αστοχίας του φράγματος εξαρτάται από την ένταση του σεισμού και από το ύψος της επιφάνειας του νερού στη λίμνη την ώρα του σεισμού. Από παρατηρήσεις στο παρελθών, μπορούμε να υποθέσουμε ότι το ύψος της επιφάνειας του νερού στην λίμνη ενδέχεται να είναι ψηλά ή χαμηλά με αντίστοιχες πιθανότητες 0,2 και 0,80. Ο σεισμός ενδέχεται να είναι ισχυρός, μέτριος ή ασθενής με αντίστοιχες πιθανότητες 0,10, 0,30 και 0,60.

Όταν συμβαίνει ισχυρός σεισμός, το φράγμα αστοχεί, άσχετα από το ύψος της επιφάνειας νερού στην λίμνη. Σε περίπτωση σεισμού μετρίας έντασης το φράγμα αστοχεί, αν το ύψος της επιφάνειας νερού βρίσκεται ψηλά, με πιθανότητα 60%,. Στον ασθενή σεισμό το φράγμα δεν διατρέχει κανένα κίνδυνο.
(α) Ποια η πιθανότητα ότι το φράγμα θα υποχωρήσει στον επόμενο σεισμό;
(β) Γίνεται σεισμός και το φράγμα υποχωρεί. Ποια η πιθανότητα ότι η επιφάνεια του νερού της λίμνης ήταν ψηλά την ώρα του σεισμού;

Λύση
Σχετικά με την επιφάνεια νερού στην λίμνη , ορίζω τα γεγονότα,

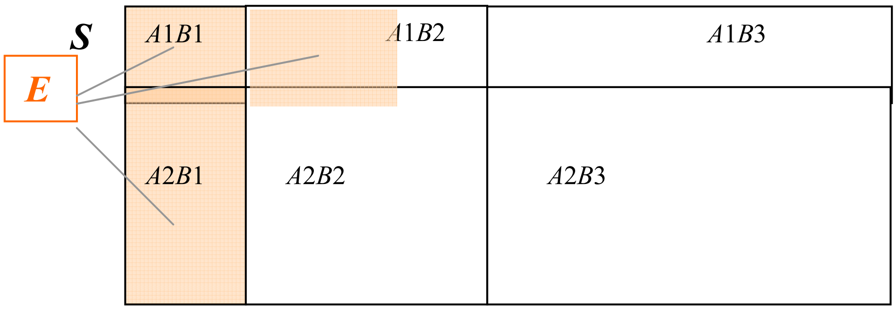
*A*1=*[επιφάνεια νερού ψηλά}*,
*A*2={*επιφάνεια νερού χαμηλά*}.

Τα γεγονότα *A*1, *A*2 αποτελούν διαμέριση του δειγματικού Χώρου *S*1.
Ακόμη, σε έναν άλλο δειγματικό χώρο *S*2 ο οποίος σχετίζεται με την ένταση ενός σεισμού, τα γεγονότα,

*B*1={ *ο σεισμός είναι ασθενής*}
*B*2={*ο σεισμός είναι μέτριος*}
*B*3={*ο σεισμός είναι ασθενής*}.

Αν θεωρήσω τώρα ένα σύνθετο πείραμα τύχης όπου την στιγμή που γίνεται ο σεισμός παρατηρώ την ένταση του καθώς και το ύψος νερού στην λίμνη, ο σύνθετος δειγματοχώρος *S* είναι το καρτεσιανό γινόμενο των *S*1 και *S*2,

S=$S1\times S2$={$A1$, $A2$} $\times$ {$B1$, $B2$, $B3$}={($A1$,$B1$), ($A1$,$B2$), ($A1$,$B3$), ($A2$,$B1$), ($A2$,$B2$), ($A2$, $B3$)}



Τα 6 ενδεχόμενα του σύνθετου δειγματοχώρου S αποτελούν διαμέριση του χώρου αυτού, όπως φαίνεται και στο σχεδιάγραμμα *Venn*. Επίσης, στο ίδιο σχεδιάγραμμα φαίνεται το γεγονός E, το οποίο παριστάνει την αστοχία του φράγματος έναν σεισμό.

Τέλος από τα δεδομένα του προβλήματος και τον κανόνα της πιθανότητας της τομής ανεξάρτητων γεγονότων ισχύει

$$P(A1B1) = P(A1 \cap B1) = P(A1)P(B1) = 0{,}2 \cdot 0{,}1 = 0{,}02$$

παρόμοια

$$P(A1B2) = 0{,}06, \quad P(A1B3) = 0{,}12$$

$$P(A2B1) = 0{,}08, \quad P(A2B2) = 0{,}24. \quad P(A2B3) = 0{,}48$$

(α) Έτσι, η πιθανότητα να υποχωρήσει το φράγμα σε έναν σεισμό είναι η πιθανότητα του E, και με τον νόμο της ολικής πιθανότητας έχω,

$$\begin{aligned} P(E) &= P(E\,|A1B1)P(A1B1) + P(E\,|A1B2)P(A1B2) + P(E\,|A1B3)P(A1B3) + \\ &+ P(E\,|A2B1)P(A2B1) + P(E\,|A2B2)P(A2B2) + P(E\,|A2B3)P(A2B3) \\ &= 1 \cdot 0{,}02 + 0{,}6 \cdot 0{,}06 + 0 \cdot 0{,}12 + 1 \cdot 0{,}08 + 0 \cdot 0{,}24 + 0 \cdot 0{,}48 = 0{,}136 \end{aligned}$$

(β) Εδώ έχω την πληροφορία ότι σε έναν σεισμό υποχωρεί το φράγμα, δηλαδή, συμβαίνει το γεγονός E. Ζητώ την πιθανότητα να συμβαίνει και το A1 ή ισοδύναμα η ένωση $A1B1 \cup A1B2 \cup A1B3$. Έτσι, η πιθανότητα εκφράζεται με το μοντέλο,

$$P[(A1B1\cup A1B2\cup A1B3)\,|E] = P(A1B1\,|E) \quad + P(A1B2\,|E) \quad + P(A1B3)\,|E)$$
$$= \frac{P(E\,|A1B1)P(A1B1)}{P(E)} + \frac{P(E\,|A1B2)P(A1B2)}{P(E)} + \frac{P(E\,|A1B3)P(A1B3)}{P(E)}$$
$$= \frac{1\cdot 0{,}02}{0{,}136} + \frac{0{,}6\cdot 0{,}06}{0{,}136} + \frac{0\cdot 0{,}12}{0{,}136} = \frac{0{,}056}{0{,}136} = 0.412$$

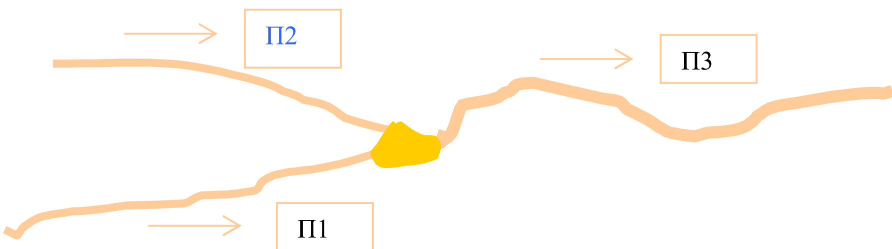
διότι τα γεγονότα $A1B1,\ A1B2,\ A1B3$ είναι ξένα μεταξύ τους, και ακόμη θυμηθείτε ότι στον κανόνα της ένωσης ισχύει και η *υπό συνθήκη πιθανότητα* (βλέπε στις *παρατηρήσεις* της παραγράφου 3.4)

3) Ξεχείλισμα ποταμού. Δύο ποταμοί Π1 και Π2 ενώνονται και σχηματίζουν έναν μεγαλύτερο ποταμό Π3, όπως φαίνεται στο Σχήμα που ακολουθεί. Σε πολύ βροχερές περιόδους ενδέχεται τα νερά των ποταμών να ανέβουν και να υπάρξουν ξεχειλίσματα. Ο όγκος νερού που ρέει στους ποταμούς Π1 και Π2 σε μεγάλες νεροποντές καθώς και η χωρητικότητα των ποταμών διαφέρει μεταξύ τους. Αν $\Pi1$, $\Pi2$, και $\Pi3$ συμβολίζουν τα γεγονότα ξεχειλισμάτων στους αντίστοιχους ποταμούς σε περίοδο νεροποντής, είναι γνωστές οι εξής πιθανότητες,

$$P(\Pi1) = 0{,}15, \quad P(\Pi2) = 0{,}30, \quad P(\Pi1|\Pi2) = 0{,}60, \quad P(\Pi2|\Pi1) = 1.$$

Επίσης, είναι γνωστό ότι αν ξεχειλίσουν και οι δύο ποταμοί Π1 και Π2, σίγουρα θα ξεχειλίσει και ο ποταμός Π3. Ακόμη, σε περίπτωση που μόνο ο Π1 ή μόνο ο Π2 ξεχειλίσει, ο ποταμός Π3 θα ξεχειλίσει με πιθανότητα 0,25.

Συμβαίνει μία μεγάλη νεροποντή στην περιοχή των ποταμών, ποια η πιθανότητα να ξεχειλίσει ο ποταμός Π3.



Λύση

Όσο αφορά την κατάσταση των ποταμών Π1 και Π2 στην περίοδο της νεροποντής, τα ακόλουθα γεγονότα αποτελούν διαμέριση του δειγματικού χώρου,

$$\Pi1\cap\overline{\Pi2}, \quad \Pi2\cap\overline{\Pi1}, \quad \Pi1\cap\Pi2, \quad \overline{\Pi1}\cap\overline{\Pi2}.$$

Έτσι, η πιθανότητα του γεγονότος $\Pi 3$ μπορεί να υπολογιστεί με τον νόμο της ολικής πιθανότητας, όπως

$$P(\Pi 3) = P(\Pi 3 | \Pi 1 \cap \overline{\Pi 2}) P(\Pi 1 \cap \overline{\Pi 2}) + P(\Pi 3 | \Pi 2 \cap \overline{\Pi 1}) P(\Pi 2 \cap \overline{\Pi 1}) +$$

$$+ P(P3 | \Pi 1 \cap \Pi 2) P(\Pi 1 \cap \Pi 2) + P(\Pi 3 | \overline{\Pi 1} \cap \overline{\Pi 2}) P \overline{\Pi 1} \cap \overline{\Pi 2}).$$

Όμως, πρώτα έχω να υπολογίσω τις πιθανότητες των γεγονότων της διαμέρισης:

$$P(\Pi 1 \cap \overline{\Pi 2}) = P(\Pi 1) P(\overline{\Pi 2} | \Pi 1) = 0{,}15 \cdot (1 - P(\Pi 2 | \Pi 1)) = 0{,}15 \cdot 0 = 0,$$

$$P(\Pi 2 \cap \overline{\Pi 1}) = P(\Pi 2) P(\overline{\Pi 1} | \Pi 2) = 0{,}30 \cdot (1 - P(\Pi 1 | \Pi 2)) = 0{,}30 \cdot 0{,}40 = 0{,}12,$$

$$P(\Pi 1 \cap \Pi 2) = P(\Pi 1) P(\Pi 2 | \Pi 1) = 0{,}15 \cdot 1 = 0{,}15,$$

$$P(\overline{\Pi 1} \cap \overline{\Pi 2}) = 1 - (0 + 0{,}12 + 0{,}15) = 1 - 0{,}27 = 0{,}73.$$

Έτσι, καταλήγω στο ακόλουθο αποτέλεσμα,

$$P(\Pi 3) = 0{,}25 \cdot 0 + 0{,}25 \cdot 0{,}12 + 1 \cdot 0{,}15 + 0 \cdot 0{,}73 = 0{,}03 + 0{,}15 = 0{,}153$$

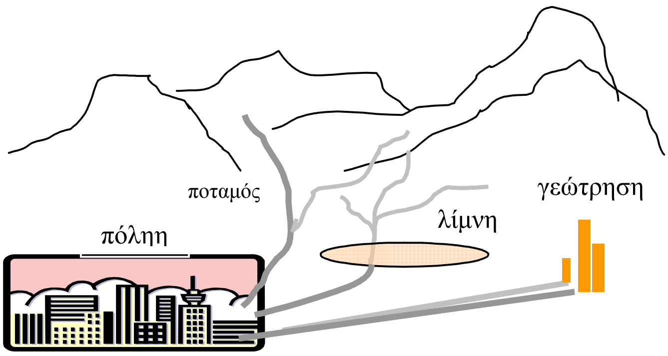
4) Έλλειμμα νερού. Η προμήθεια νερού σε μία πόλη, κοντά στην περιοχή του Ολύμπου, στηρίζεται και στα δύο επιφανειακό και υπόγειο νερό. Το επιφανειακό νερό αποτελείται από ένα ποταμό που καταβαίνει από το βουνό και μία μεγάλη δεξαμενή (τεχνητή λίμνη). Το υπόγειο νερό αντλείται από κάποια μεγάλη γεώτρηση.

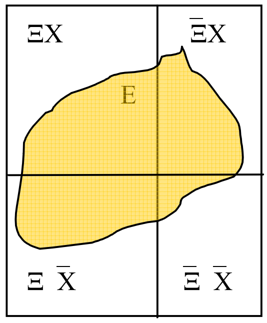
Η λειτουργία του συστήματος επιρεάζεται από κάποιες έντονες ξηρασίες οι οποίες συμβαίνουν κατά μέσο όρο μία φορά στα πέντε χρόνια. Όταν συμβαίνει ξηρασία υπάρχει πιθανότητα 0,40 να ελαττωθεί κατά πολύ η ροή των υδάτων του ποταμού και των άλλων ρυακιών που γεμίζουν την τεχνητή λίμνη.

Η ξηρασία και η χαμηλή ροή των επιφανειακών υδάτων μπορεί να προκαλέσει έλλειψη νερού στην πόλη με πιθανότητα 0,30. Ακόμη, όταν υπάρχει ξηρασία ενδέχεται η ροή των υδάτων στον ποταμό και στα ρυάκια να παραμένει ικανοποιητική, αλλά το απόθεμα στην τεχνητή λίμνη να είναι λίγο και να προκαλεί ανεπάρκεια νερού στην πόλη με πιθανότητα 0,10. Τέλος, σε χρονιά με ξηρασία όπου η ροή στον ποταμό και ρυάκια είναι χαμηλή αλλά υπάρχει αρκετό απόθεμα, ενδέχεται να προκληθεί μόλυνση στο υπόγειο νερό και η υδροδότηση να μην είναι ικανοποιητική με πιθανότητα 0,25.

(α) Να ευρεθεί η πιθανότητα ότι μέσα σε ένα χρόνο θα υπάρξει έλλειμμα νερού στην πόλη.

(β) Αν μία χρονιά έχουμε ξηρασία, ποια η πιθανότητα ότι δεν θα υπάρξει έλλειμμα νερού στην πόλη.





Λύση.
Ορίζουμε τα ακόλουθα γεγονότα

Ξ={*Ξηρασία μέσα στον χρόνο*},
X={*Χαμηλή ροή στον ποταμό και ρυάκια*},
E={*Έλλειψη νερού στην πόλη*}.

Προφανώς, έχουμε τις εξής πιθανότητες,

$$P(\Xi \cap X) = P(\Xi)P(X|\Xi) = 0{,}20 \cdot 0{,}40 = 0{,}08, \quad P(\Xi \cap \overline{X}) = P(\Xi)P(\overline{X}|\Xi) = 0{,}20 \cdot 0{,}60 = 0{,}12$$

$$P(\overline{\Xi} \cap X) = P(\overline{\Xi})P(X|\overline{\Xi}) = 0{,}80 \cdot 0{,}0 = 0{,}0 \quad P(\overline{\Xi} \cap \overline{X}) = P(\overline{\Xi})P(\overline{X}|\overline{\Xi}) = 0{,}80 \cdot 1{,}0 = 0{,}80.$$

Ακόμη, είναι γνωστές οι υπό συνθήκη πιθανότητες

$$P(E|(\Xi \cap \overline{X})) = 0{,}10,$$
$$P(E|(\Xi \cap X)) = 0{,}30 + 0{,}70 \cdot 0{,}25 = 0{,}475$$

Η τελευταία πιθανότητα προκύπτει από το σκεπτικό ότι στην περίπτωση ξηρασίας και χαμηλής ροής, το γεγονός *E* συμβαίνει λόγω έλλειψης νερού (με πιθανότητα 0,30) ή αρκετό απόθεμα νερού αλλά μολυσμένο (με πιθανότητα 0,70·0,25=0,175).

(α) Έτσι, αφού τα γεγονότα

$$\Xi \cap X, \quad \Xi \cap \overline{X}, \quad \overline{\Xi} \cap X, \quad \overline{\Xi} \cap \overline{X}$$

αποτελούν διαμέριση του δειγματικού χώρου *S*, η ολική πιθανότητα έλλειψης νερού στην πόλη μπορεί να υπολογισθεί από το άθροισμα

$$\begin{aligned} P(E) &= P(E|(\Xi \cap X))P(\Xi \cap X) + P(E|(\Xi \cap \overline{X}))P(\Xi \cap \overline{X}) \\ &\quad + P(E|(\overline{\Xi} \cap \overline{X}))P(\overline{\Xi} \cap \overline{X}) + P(E|(\overline{\Xi} \cap X))P(\overline{\Xi} \cap X) \\ &= 0{,}475 \cdot 0{,}080 + 0{,}10 \cdot 0{,}12 + 0{,}0 \cdot 0{,}80 + 0{,}0 \cdot 0{,}0 = 0{,}05 \end{aligned}$$

(β) Μπορούμε να απαντήσουμε εδώ αν στην τελευταία ολική πιθανότητα P(E) θεωρήσουμε σίγουρο το γεγονός ξηρασίας *Ξ*.

$$P(E) = P(E|(\Xi \cap X))P(\Xi \cap X) + P(E|(\Xi \cap \overline{X}))P(\Xi \cap \overline{X}) +$$

$$+ P(E|(\overline{\Xi} \cap \overline{X}))P(\overline{\Xi} \cap \overline{X}) + P(E|(\overline{\Xi} \cap X))P(\overline{\Xi} \cap X)$$

$$= 0{,}475 \cdot 0{,}40 + 0{,}10 \cdot 0{,}60 + 0{,}0 \cdot 0{,}0 + 0{,}0 \cdot 0{,}0 = 0{,}25$$

αφού

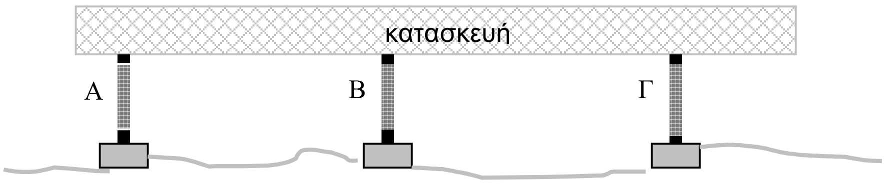
$$P(\Xi \cap X) = P(X) = 0{,}40 \quad P(\Xi \cap \overline{X}) = P(\overline{X}) = 0{,}60$$

$$P(\overline{\Xi} \cap X) = P(\{\}) = 0{,}0 \quad P(\overline{\Xi} \cap \overline{X}) = P(\{\}) = 0{,}0.$$

5) Καθίζηση εδάφους. Μία κατασκευή στηρίζεται σε τρία πέλματα, όπως φαίνεται στο σχήμα που ακολουθεί. Το κάθε πέλμα ενδέχεται να παραμείνει στη αρχική του θέση ή να υποχωρήσει κατά 4 πόντους με πιθανότητα 0,20. Όταν υπάρχει καθίζηση σε ένα πέλμα τότε η πιθανότητα καθίζησης στο γειτονικό του αυξάνει σε 0,60.

(α) Εάν το πέλμα Α δεν υπέστη καθίζηση, ποια η πιθανότητα ότι θα συμβεί καθίζηση στο πέλμα Β;

(β) Εάν το πέλμα Α υπέστη καθίζηση, ποια η πιθανότητα ότι θα υπάρξει καθίζηση στο πέλμα Γ;



Λύση

Συμβολίζουμε με *A, B*, και *Γ* τα γεγονότα ότι τα αντίστοιχα πέλματα θα υποστούν καθίζηση. Γνωρίζουμε τις εξής πιθανότητες με βάση τα δεδομένα του προβλήματος,

$$P(A) = 0{,}2, \quad P(B) = 0{,}2, \quad P(\Gamma) = 0{,}2$$

$$P(A|B) = 0{,}6, \quad P(B|A) = 0{,}6, \quad P(B|\Gamma) = 0{,}6, \quad P(\Gamma|B) = 0{,}6$$

$$P(\overline{A}|B) = 1 - P(A|B) = 1 - 0{,}6 = 0{,}4, ...$$

(α) Πρέπει να τονίσουμε ότι αν δύο γεγονότα π.χ. A, B είναι εξαρτημένα τότε τα γεγονότα A, $\overline{B}$
είναι επίσης εξαρτημένα. Έτσι, η πιθανότητα $P(B|\overline{A})$ δεν είναι 0,2. Μπορώ όμως να την υπολογίσω με την υπό συνθήκη πιθανότητα.

$$P(B|\overline{A}) = \frac{P(B \cap \overline{A})}{P(\overline{A})} = \frac{P(B)P(\overline{A}|B)}{0{,}80} = \frac{0{,}2 \cdot 0{,}4}{0{,}80} = 0{,}10$$

Το ίδιο προκύπτει και για παρόμοιες πιθανότητες

$$P(\Gamma|\overline{B}) = 0{,}1, \quad P(A|\overline{B}) = 0{,}1.$$

(β) Ζητείται η υπό συνθήκη πιθανότητα,

$$P(\Gamma|A) = \frac{P(\Gamma \cap A)}{P(A)} = \frac{P(\Gamma \cap A)}{0{,}20}$$

Παρατηρούμε στο σχεδιάγραμμα Venn ότι το γεγονός $A \cap \Gamma$ αποτελείται από την ένωση των δύο τεμαχίων $A \cap B \cap \Gamma$ και $A \cap \overline{B} \cap \Gamma$. Έτσι η πιθανότητα $P(A \cap \Gamma)$ μπορεί να υπολογιστεί με τον κανόνα της ένωσης.

$$\begin{aligned} P(A \cap \Gamma) &= P[(A \cap B \cap \Gamma) \cup (A \cap \overline{B} \cap \Gamma)] = P(A \cap B \cap \Gamma) + P(A \cap \overline{B} \cap \Gamma) = \\ &= P(A)P(B|A)P(\Gamma|A \cap B) + P(A)P(\overline{B}|A)P(\Gamma|\overline{B} \cap A) \\ &= 0{,}2 \cdot 0{,}6 \cdot 0{,}6 + 0{,}2 \cdot 0{,}4 \cdot 0{,}1 = 0{,}08 \end{aligned}$$

Αντικαθιστώντας στην υπό συνθήκη πιθανότητα έχουμε

$$P(\Gamma|A) = \frac{P(\Gamma \cap A)}{P(A)} = \frac{0{,}08}{0{,}20} = 0{,}40$$

6) Γεωτρήσεις νερού. ***Μία μικρή πόλη υδροδοτείται από δύο μεγάλες γεωτρήσεις που βρίσκονται σχετικά κοντά στην πόλη. Η κάθε γεώτρηση μπορεί να ικανοποιήσει την ζήτηση νερού με πιθανότητα 60%. Όταν υπάρχει ανομβρία η στάθμη των υπόγειων υδάτων πέφτει χαμηλά, και ενδέχεται για την κάθε γεώτρηση να αποτύχει στην άντληση με πιθανότητα 0,10. Είναι γνωστό ότι αν η μία γεώτρηση σταματήσει την άντληση λόγω χαμηλής στάθμης των υπόγειων υδάτων, ενδέχεται να σταματήσει και η άλλη***

γεώτρηση με πιθανότητα 0,30. Κάποιο καλοκαίρι παρατηρείται διαρκής ξηρασία και οι εμπειρογνώμονες εκτιμούν ότι τουλάχιστον μία από τις γεωτρήσεις θα σταματήσει την άντληση λόγω πτώσης της στάθμης των υπόγειων υδάτων. Ποια η πιθανότητα να έχουμε ικανοποιητική παροχή νερού στην πόλη;

Λύση.

Ορίζουμε τα γεγονότα,

$A1$={*η πρώτη γεώτρηση διακόπτει την άντληση νερού λόγω χαμηλής στάθμης*}
$A2$={ *η δεύτερη γεώτρηση διακόπτει την άντληση νερού λόγω χαμηλής στάθμης*},
$A1 \cup A2$={*τουλάχιστον μία από τις γεωτρήσεις διακόπτει την άντληση*},
B={*μόνο μία γεώτρηση αντλεί νερό*}$=(A1\cap \overline{A2})\cup(\overline{A1}\cap A2)$.
R={*η πόλη έχει ικανοποιητική παροχή νερού*}

Από τα δεδομένα του προβλήματος προκύπτουν εύκολα οι πιθανότητες,

$$P(A1)=0{,}10,\quad P(A2)=0{,}10,\quad P(A1|A2)=0{,}30,\quad P(A2|A1)=0{,}30,$$
$$P(A1\cap A2)=P(A1)P(A2|A1)=0{,}10\cdot 0{,}30=0{,}03$$

Θεωρούμε ότι συμβαίνει το γεγονός $A1\cup A2$, ότι τουλάχιστον μία από τις γεωτρήσεις διακόπτει την άντληση. Το γεγονός R μπορεί να πραγματοποιηθεί αν μόνο μία από τις γεωτρήσεις διακόπτει την άντληση (γεγονός B) και το νερό που αντλείται μόνο από μία γεώτρηση επαρκή για την πόλη. Δηλαδή, θα υπολογίσω την υπό συνθήκη πιθανότητα, $P(B|A1\cup A2)$ και θα το πολλαπλασιάσω με την πιθανότητα 0,60.

$$\begin{aligned}
P(B|A1\cup A2) &= P[[(A1\cap \overline{A2})\cup(\overline{A1}\cap A2)]|A1\cup A2)] \\
&= P[(A1\cap \overline{A2})|A1\cup A2)] + P[(\overline{A1}\cap A2)|A1\cup A2)] \\
&= \frac{P[(A1\cap \overline{A2})\cap(A1\cup A2)]}{P(A1\cup A2)} + \frac{P[(A2\cap \overline{A1})\cap(A1\cup A2)]}{P(A1\cup A2)} \\
&= \frac{P(A1\cap \overline{A2})}{P(A1\cup A2)} + \frac{P(A2\cap \overline{A1})}{P(A1\cup A2)}
\end{aligned}$$

$$= \frac{P(A1)P(\overline{A2}|A1)}{P(A1 \cup A2)} + \frac{P(A2)P(\overline{A1}|A2)}{P(A1 \cup A2)}$$

$$= \frac{P(A1)(1 - P(A2|A1)}{P(A1) + P(A2) - P(A1 \cap A2)} + \frac{P(A2)(1 - P(A1|A2)}{P(A1) + P(A2) - P(A1 \cap A2)}$$

$$= \frac{0{,}10 \cdot 0{,}70 + 0{,}10 \cdot 0{,}70}{0{,}10 + 0{,}10 - 0{,}03} = \frac{0{,}14}{0{,}17} = 0{,}824$$

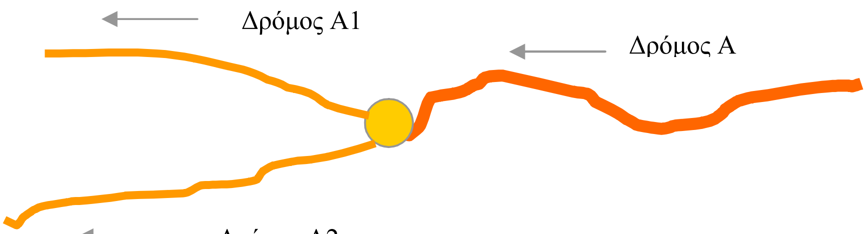
Και η πιθανότητα ότι η πόλη θα έχει ικανοποιητική παροχή νερού,

$$P(R) = P(B|A1 \cup A2) \cdot 0{,}60 = 0{,}824 \cdot 0{,}60 = 0{,}494$$

7) Κυκλοφοριακή συμφόρηση. Ένας μεγάλος αυτοκινητόδρομος ο Α οδηγεί σε ένα κόμβο και στη συνέχεια διακλαδίζεται σε δύο μικρότερους δρόμους τους Α1 και Α2, όπως φαίνεται στο σχήμα που ακολουθεί. Ο κυκλοφοριακός όγκος του Α μοιράζεται, όχι κατ' ανάγκη εξίσου, στους δρόμους Α1 και Α2. Αν A, $A1$, και $A2$ συμβολίζουν τα γεγονότα κυκλοφοριακής συμφόρησης σε περίοδο αιχμής στους αντίστοιχους δρόμους, από παρατηρήσεις στο παρελθόν είναι γνωστό ότι

$$P(A) = 0{,}60,\ P(A1) = 0{,}30,\ P(A2) = 0{,}40,\ P(A1|A) = 0{,}40,$$
$$P(A2|A) = 0{,}50,\ P(A2|A \cap A1) = 0{,}50,\ P(A1|A \cap A2) = 0{,}40 \text{ .}$$

(α) Σε περίοδο κυκλοφοριακής αιχμής δεν συμβαίνει συμφόρηση στον δρόμο Α, ποια η πιθανότητα να υπάρχει πρόβλημα συμφόρησης στον δρόμο Α1;
(β) Ποια η πιθανότητα συμφόρησης και στους τρείς δρόμους σε περίοδο κυκλοφοριακής αιχμής;
(γ) Σε περίοδο κυκλοφοριακής αιχμής παρατηρείται συμφόρηση στον Α2, ποια η πιθανότητα να έχει πρόβλημα συμφόρησης και ο Α1.



Λύση
(α) Σε περίοδο κυκλοφοριακής αιχμής ο δρόμος Α ενδέχεται να έχει συμφόρηση ή να μην έχει, άλλο ενδεχόμενο δεν υπάρχει. Έτσι, σε περίοδο αιχμής τα γεγονότα A και $\overline{A}$ αποτελούν διαμέριση του δειγματικού χώρου S.

Κατά συνέπεια τις πιθανότητες των γεγονότων $A1$ και $A2$ μπορώ να τις εκφράσω με τον νόμο της ολικής πιθανότητας ως εξής:

$$P(A1) = P(A1|A)P(A) + P(A1|\overline{A})P(\overline{A}) = 0{,}30$$

$$= 0{,}40 \cdot 0{,}60 + P(A1|\overline{A}) \cdot 0{,}4 = 0{,}30$$

Από την παραπάνω εξίσωση συνεπάγεται ότι

$$P(A1|\overline{A}) = 0{,}15$$

(β) Προφανώς, ζητάμε την πιθανότητα της τομής τριών γεγονότων και με τον πολλαπλασιαστικό κανόνα έχουμε,

$$P(A \cap A1 \cap A2) = P(A)P(A1|A)P(A2|A \cap A1) = 0{,}6 \cdot 0{,}40 \cdot 0{,}50 = 0{,}12 .$$

(γ) Από τα δεδομένα του προβλήματος μόνο η πραγματοποίηση του A μεταβάλλει την πιθανότητα του $A1$. Δεδομένου ότι συμβαίνει το γεγονός $A2$ μπορούμε με το θεώρημα του Bayes να εκτιμήσουμε την πιθανότητα να συμβαίνει το A,

$$P(A|A2) = \frac{P(A2|A)P(A)}{P(A2)} = \frac{0{,}50 \cdot 0{,}60}{0{,}40} = 0{,}75$$

αφού η αναθεωρημένη (ή εκ των υστέρων) πιθανότητα του A είναι $P(A|A2)$=0,75, με τον κανόνα της ολικής πιθανότητας υπολογίζω την πιθανότητα του $A1$,

$$P(A1|A2) = P((A1|A2)|A)P(A|A2) + P((A1|A2)|\overline{A})P(\overline{A}|A2)$$

$$= P(A1|A \cap A2)P(A|A2) + P(A1|A2 \cap \overline{A})P(\overline{A}|A2) = 0{,}40 \cdot 0{,}75 + 0{,}15 \cdot 0{,}25 = 0{,}3375$$

8) Τακτική διαχείρισης υδραγωγείου. Η αποθήκευση νερού σε ένα μεγάλο υδραγωγείο εξαρτάται από

- το αρχικό απόθεμα νερού S στην αρχή του χρόνου,
- την ετήσια εισροή νερού I,
- την δυνατότητα αποθήκευσης νερού c,
- και την ετήσια ζήτηση d.

Ο διαχειριστής πρέπει να αφήνει ετησίως μία ποσότητα νερού R, η οποία εξαρτάται από το απόθεμα S στην αρχή της χρονιάς, από την ετήσια ποσότητα εισροής νερού I, και από την ετήσια ποσότητα d κατανάλωσης (ζήτηση). Ο διαχειριστής ακολουθεί την ακόλουθη τακτική για την έξοδο νερού R,

$R=d$ εάν $d<I+S<d+c$
$R=I+S$, εάν $I+S<d$

R=d+(I+S-c), εάν I+S>d+c

Αν είναι γνωστό ότι

$P(\{d \le I+S \le d+c\}) = 0{,}60, \quad P(\{I+S \le\}) = 0{,}10, \quad P(\{I+S > d+c\}) = 0{,}30,$

να ευρεθεί η πιθανότητα ότι αυτόν τον χρόνο θα ικανοποιηθεί η ζήτηση.

9) Βλάβες στο δίκτυο υδροδότησης. Η ικανοποιητική παροχή νερού σε μία πόλη συνίσταται στην καλή λειτουργία του συστήματος άντλησης νερού και του υπόγειου δικτύου σωλήνων. Λόγω κακής συντήρησης ή περιοδικών βλαβών το σύστημα άντλησης ενδέχεται μέσα σε κάθε εβδομάδα να ελαττώσει την άντληση με πιθανότητα 0,02 (το ενδεχόμενο αυτό συμβολίζεται με B_1). Για το δίκτυο των σωλήνων υπάρχουν πολλές αιτίες, όπως διέλευση βαρέων οχημάτων πάνω από κεντρικούς αγωγούς, καθίζηση εδάφους, σκουρίαση σωλήνων, και άλλα οι οποίες μπορούν να προκαλέσουν ζημία στους σωλήνες και να υπάρξει μερική διακοπή παροχής νερού, (*το γεγονός βλάβης στο δίκτυο μέσα σε μία εβδομάδα συμβολίζεται* B_2). Από το τμήμα συντήρησης του δικτύου είναι γνωστό ότι τέτοιες βλάβες στο δίκτυο εμφανίζονται κατά μέσο όρο κάθε 20 εβδομάδες.

Κατά την διάρκεια της εβδομάδας το καλοκαίρι υπάρχει πιθανότητα 0,08 να συμβεί πυρκαιά στο γύρω δάσος (*το γεγονός αυτό συμβολίζεται με* A), και για την κατάσβεση αυτής καταναλώνεται μεγάλη ποσότητα νερού, ώστε να υπάρξει απώλεια πίεσης του νερού στο δίκτυο.

Έτσι, η συμπεριφορά του συστήματος ύδρευσης το καλοκαίρι κατά την διάρκεια μιας εβδομάδας, σχετικά με την απαιτούμενη πίεση του νερού στους καταναλωτές μπορεί να είναι

επιτυχής, E, αν δεν υπάρχει βλάβη στην άντληση και στο δίκτυο και δεν έχουμε πυρκαιά.

μη επιτυχής, F, αν υπάρχει βλάβη στο δίκτυο ή σύστημα άντλησης και συμβεί πυρκαιά,

μέτρια, M, αν συμβεί βλάβη μόνο στο σύστημα άντλησης ή μόνο στο δίκτυο και δεν υπάρξει πυρκαιά στο δάσος.

Με την υπόθεση ότι τα γεγονότα B_1 , B_2 και A είναι ανεξάρτητα,

(α) να εκφράστε το κάθε γεγονός E, F, M με άλγεβρα των γεγονότων B_1 , B_2 και A,

(β) να υπολογίστε την πιθανότητα ότι βλάβη θα συμβεί μόνο στο δίκτυο των σωλήνων,

(γ) να υπολογιστούν οι πιθανότητες των γεγονότων E, F και M.

4

ΤΥΧΑΙΕΣ ΜΕΤΑΒΛΗΤΕΣ, ΣΥΝΑΡΤΗΣΕΙΣ ΚΑΤΑΝΟΜΗΣ ΠΙΘΑΝΟΤΗΤΑΣ

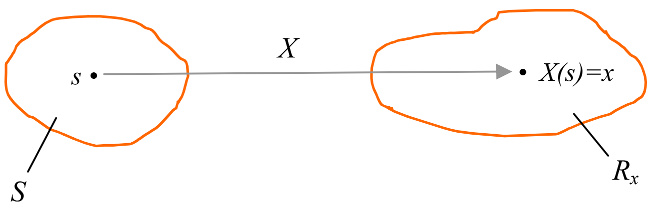
ΠΕΡΙΓΡΑΦΗ ΚΕΦΑΛΑΙΟΥ

Όταν εκτελείται ένα πείραμα τύχης, συχνά δεν μας ενδιαφέρουν όλες οι λεπτομέρειες του πειραματικού αποτελέσματος αλλά μόνο η τιμή κάποιας αριθμητικής ποσότητας που προσδιορίζεται από το αποτέλεσμα. Για παράδειγμα, στη ρίψη δύο ζαριών συχνά μας ενδιαφέρει το άθροισμα των ενδείξεων των δύο ζαριών και δεν σχετιζόμαστε καθόλου με τις ενδείξεις του κάθε ζαριού ξεχωριστά. Δηλαδή, μπορεί να μας ενδιαφέρει αν το άθροισμα είναι εφτά, άσχετα αν το αποτέλεσμα της ρίψης ήταν (1,6) ή (6,1) ή (5,2) ή (2,5) ή (3,4) ή (4,3). Τέτοιες μετρήσιμες ποσότητες, οι οποίες προσδιορίζονται από το αποτέλεσμα του πειράματος τύχης είναι γνωστές σαν ***τυχαίες μεταβλητές***.

4.1 ΕΝΝΟΙΑ ΤΥΧΑΙΑΣ ΜΕΤΑΒΛΗΤΗΣ

Στα πειράματα τύχης μέχρι τώρα το αποτέλεσμα ενδέχεται να ήταν ένας αριθμός ή να μην είναι κάποια αριθμήσιμη ποσότητα. Για παράδειγμα στην ρίψη ενός ζαριού το αποτέλεσμα του πειράματος είναι ένας αριθμός, ενώ στην ρίψη ενός νομίσματος η εμφάνιση της «Κεφαλής» ή «Γράμματος» δεν είναι αριθμός. Στο εξής, σε πολλά πειράματα τύχης πρόκειται να ασχοληθούμε με μετρήσιμες ποσότητες και στο αποτέλεσμα θα αντιστοιχούμε έναν αριθμό, π.χ., στην ρίψη του νομίσματος μπορούμε να αντιστοιχίσουμε στην εμφάνιση «Κεφαλής» ή «Γράμματος» τους αριθμούς μηδέν ή ένα αντίστοιχα.

Γενικά, θέλουμε να μετρούμε το κάθε αποτέλεσμα του πειράματος τύχης ή με άλλα λόγια θα αντιστοιχούμε έναν πραγματικό αριθμό x σε κάθε δειγματοσημείο s του δειγματικού χώρου S. Δηλαδή, $x=X(s)$ είναι μία τιμή της συνάρτησης X από τον δειγματικό χώρο S στους πραγματικούς αριθμούς R, όπως φαίνεται και στο σχήμα που ακολουθεί.



Σχήμα 4.1 Τυχαία μεταβλητή X

Ορισμός Σε ένα πείραμα τύχης E το οποίο σχετίζεται με τον δειγματικό χώρο S , ονομάζουμε **τυχαία μεταβλητή** μία συνάρτηση $\boldsymbol{X}$, η οποία αντιστοιχεί σε κάθε δειγματοσημείο $s \in S$ έναν πραγματικό αριθμό $X(s)=x$.

Παρατηρήσεις

(Α) Σε μερικές περιπτώσεις το αποτέλεσμα s σε ένα δειγματικό χώρο έχει ήδη αριθμητικό χαρακτήρα. Έχουμε απλά $X(s)= s$, δηλαδή η X είναι μία ταυτοτική συνάρτηση.

(Β) Προς το παρών θα ασχοληθούμε με την μονοδιάστατη τυχαία μεταβλητή X. Όπως ακολουθεί και από τις βασικές απαιτήσεις μίας συνάρτησης, σε κάθε $s \in S$ αντιστοιχεί ακριβώς μία τιμή $X(s)$, όπως

φαίνεται και στο Σχήμα 4.1. Διαφορετικά δειγματοσημεία s μπορεί να οδηγούν σε διαφορετικές τιμές της X, δεν ισχύει όμως το αντίστροφο. Η παρατήρηση αυτή δεν ισχύει στις δυδιάστατες ή πολυδιάστατες τυχαίες μεταβλητές όπου σε κάθε δειγματοσημείο $s \in S$ η συνάρτηση X αντιστοιχεί ακριβώς δύο ή και περισσότερες τιμές στον πραγματικό χώρο. Τέτοιες πολυδιάστατες μεταβλητές μελετώνται σε άλλα κεφάλαια της Στατιστικής.
(Γ) Μία τυχαία μεταβλητή συμβολίζουμε συνήθως με ένα κεφαλαίο γράμμα, X, ενώ με μικρό γράμμα x συμβολίζουμε την εκάστοτε τιμή που ενδέχεται να πάρει η μεταβλητή, $X(s)=x$. Σε κάθε πείραμα τύχης E μπορούμε να ορίσουμε περισσότερες από μία τυχαίες μεταβλητές, τις οποίες πάντα συμβολίζουμε με κεφαλαία γράμματα όπως X, Y, Z, κ.τ.λ., ενώ τις τιμές που παίρνουν με μικρά, x, y, z, αντίστοιχα.

Παράδειγμα 4.1
Τρία νομίσματα ρίχνονται πάνω σε ένα τραπέζι. Μία έκβαση s του πειράματος τύχης είναι *πώς* και *που* πέσανε τα νομίσματα πάνω στο τραπέζι. Προφανώς μας ενδιαφέρουν κάποια αριθμητικά χαρακτηριστικά που σχετίζονται με αυτό το πείραμα. Παραδείγματος χάριν, ίσως επιθυμούμε να μετρήσουμε

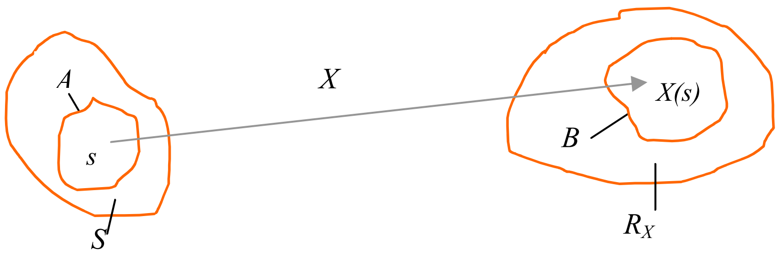
- $X(s)$= *τον αριθμό κεφαλών*,
- $Y(s)$=*την μέγιστη απόσταση μεταξύ των δύο νομισμάτων,*
- $Z(s)$= *την μικρότερη απόσταση των νομισμάτων από τις άκρες του τραπεζίου.*

Εάν αναφερθούμε στην τυχαία μεταβλητή X, για κάποιο αποτέλεσμα του πειράματος s, η τιμή $X(s)$ είναι μία τιμή από το σύνολο $\{0, 1, 2, 3\}$. Παρόμοια, οι τιμές $Y(s)$, $Z(s)$ που παίρνουν οι άλλες δύο μεταβλητές Y, Z για κάποια έκβαση s του πειράματος, ανήκουν σε ένα τμήμα των πραγματικών αριθμών. ■

Ο πραγματικός χώρος R_x περιέχει το σύνολο όλων των τιμών της X, και παριστάνει με αριθμητικούς χαρακτήρες όλα τα αποτελέσματα ενός πειράματος τύχης E, που περιέχονται στον δειγματικό χώρο S. Εάν $X(s)= s$, έχουμε $S= R_x$.

Μία **τυχαία μεταβλητή** είναι μία **αριθμητική συνάρτηση** ορισμένη στον δειγματικό χώρο S. Κάθε τιμή που παίρνει η τυχαία μεταβλητή είναι ένα γεγονός, αφού υπάρχει πάντα κάποιο στοιχειώδες γεγονός ή ένα σύνολο από στοιχειώδη γεγονότα στον δειγματικό χώρο , που οδήγησαν σε αυτήν την τιμή. Όπως αναφερθήκαμε στα γεγονότα που σχετίζονται με τον δειγματικό χώρο S, έτσι το ίδιο θεωρούμε αναγκαίο να καθιερώσουμε γεγονότα που αφορούν την τυχαία μεταβλητή X, δηλαδή, υποσύνολα του χώρου R_X. Πολύ συχνά συγκεκριμένα γεγονότα του δειγματικού χώρου S σχετίζονται ή περιγράφονται ισοδύναμα με γεγονότα του χώρου R_X.

Αν υποθέσουμε ότι A είναι ένα γεγονός στο δειγματικό χώρο S το οποίο περιέχει εκείνα τα δειγματοσημεία s για τα οποία $X(s) \in$ B, όπως φαίνεται και στο Σχήμα 4.2,



Σχήμα 4.2. Ισοδύναμα γεγονότα *A, B*

τότε τα γεγονότα *A* και *B* είναι ισοδύναμα. Αυτό σημαίνει με άλλα λόγια ότι όταν συμβαίνει το γεγονός *A* , συμβαίνει το γεγονός *B* και αντίστροφα. Διότι εάν συνέβη το *A*, τότε κάποιο αποτέλεσμα συνέβη για το οποίο $X(s) \in B$ και οπότε συμβαίνει και το *B*, Αντίστροφα, εάν συνέβη το *B*, παρατηρείται μία τιμή $X(s)$ για την οποία $s \in A$ και άρα συμβαίνει και το *A*. Πρέπει να τονίσουμε ότι τα γεγονότα *A* και *B* είναι ισοδύναμα αλλά αναφέρονται σε διαφορετικούς δειγματικούς χώρους, S, R_X.

Παράδειγμα 4.2

Ενδιαφερόμαστε για τον αριθμό των αγοριών, *X*, σε τρία νεογέννητα παιδιά. Το *X* προφανώς παριστάνει εδώ την τυχαία μεταβλητή, η οποία είναι μια αριθμητική συνάρτηση ορισμένη στον δειγματικό χώρο όπως φαίνεται παρακάτω,

$$S=\{AAA, AAK, AKA, KAA, AKK, KAK, KKA, KKK\}$$

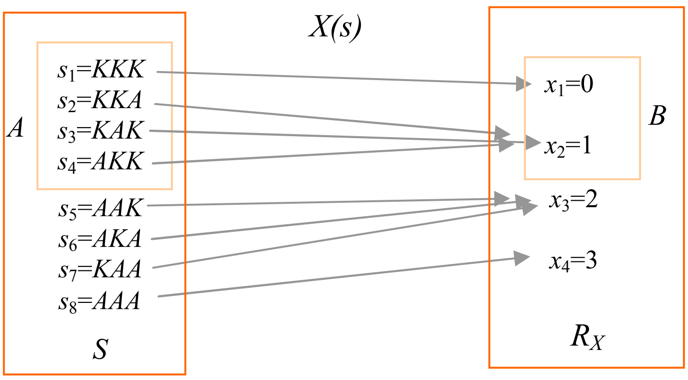
όπου *A* και *K* συμβολίζουν αγόρι ή κορίτσι αντίστοιχα και το κάθε δειγματοσημείο αναφέρει την σειρά που γεννήθηκαν. Για να ορίσουμε την *X* στον παραπάνω δειγματικό χώρο, θα μπορούσαμε πρώτα να χωρίσουμε το *S* στα ακόλουθα ξένα συλλεκτικά εξαντλητικά γεγονότα (υποσύνολα)

A0={*καθόλου αγόρια*}={*KKK*}
*A*1={*ένα αγόρι*}={ *AKK, KAK, KKA* }
*A*2={*δύο αγόρια*}={ *AAK, AKA, KAA*}
*A*3={*τρία αγόρια*}={*AAA*}

Η συνάρτηση της τυχαίας μεταβλητής αντιστοιχεί σε κάθε δειγματοσημείο των γεγονότων *A0*, *A*1, *A*2, *A*3 την τιμή $x_1=0$, $x_2=1$, $x_3=2$, $x_4=3$ αντίστοιχα, όπως φαίνεται στο Σχήμα 4.3. Τα γεγονότα

$A=\{\text{το πολύ ένα αγόρι}\}=\{s\,|\,X(s)\leq 1\}=\{s_1,s_2,s_3,s_4\}\in S$
και
$B=\{X(s)\leq 1\}=\{x_1,\ x_2\}\ \in R_X$

είναι ισοδύναμα.



Σχήμα 4.3 Γραφική παράσταση της συνάρτησης *X*.

Παρατηρούμε τελικά ότι κάθε τιμή της τυχαίας μεταβλητής είναι ένα γεγονός, ισοδύναμο με ένα γεγονός στον δειγματικό χώρο, το οποίο περιέχει ένα ή περισσότερα στοιχειώδη γεγονότα. Μία τιμή της τυχαίας μεταβλητής, είναι ένα γεγονός στον χώρο R_X,

$$\{X=x_1\},\ \{X\leq x_1\},\ \{x_1\leq X\leq x_2\}$$

και μπορούμε να του αντιστοιχίσουμε και μία πιθανότητα,

$$P(X=x_1)\ \text{ή}\ P(X\leq x_1)\ \text{ή}\ P(x_1\leq X\leq x_2).$$

Γενικά αν B είναι ένα γεγονός στον χώρο R_X μπορούμε να ορίσουμε την πιθανότητα του B

$$P(B)=P(A)$$

την πιθανότητα $P(A)$ του ισοδύναμου γεγονότος A στον δειγματικό χώρο S, $A\in S$.

Καταλήγουμε ότι όταν σε ένα πείραμα τύχης εισάγουμε μία τυχαία μεταβλητή X και τον σχετικό χώρο τιμών της R_X, οι πιθανότητες για γεγονότα όπως

$$\{X=x_1\},\ \{X\leq x_1\},\ \{x_1\leq X\leq x_2\}$$

που ανήκουν στο R_X σχετίζονται με γεγονότα στον δειγματικό χώρο S. Οπότε, σύμφωνα και με τον παραπάνω ορισμό μπορούμε να ορίσουμε τις πιθανότητες

$$P(X=x_1),\ P(X \le x_1),\ P(x_1 \le X \le x_2),$$

αν εκτιμήσουμε τις πιθανότητες των αντίστοιχων ισοδύναμων γεγονότων στον δειγματικό χώρο S.

Παράδειγμα 4.3

Εάν στο παράδειγμα 4.2, το ενδεχόμενο ότι το νεογέννητο παιδί είναι αγόρι είναι ισοπίθανο με το ενδεχόμενο ότι το νεογέννητο είναι κορίτσι, έχουμε σύμφωνα με την κλασσική μέθοδο,

$$P(\{AAA\})=1/8,$$
$$P(\{AKK, KAK, KKA\})=3/8,$$
$$P(\{AAK, AKA, KAA\})=3/8,$$
$$P(\{KKK\})=1/8.$$

Αφού η τυχαία μεταβλητή παριστάνει αριθμό αγοριών σε τρία νεογέννητα, το γεγονός $\{X=0\}$ (στον χώρο R_X) είναι ισοδύναμο με το γεγονός $\{KKK\}$ (στον δειγματικό χώρο S), το γεγονός $\{X=1\}$ είναι ισοδύναμο με το γεγονός $\{AKK, KAK, KKA\}$, το γεγονός $\{X \le 1\}$ είναι ισοδύναμο με το $\{AAA, AKK, KAK, KKA\}$ κ.ο.κ.. Κατά συνέπεια μπορούμε να ορίσουμε τις πιθανότητες για γεγονότα στον χώρο R_X,

$$P(X=0)=P(\{KKK\})=1/8,$$
$$P(X=1)=P(\{AKK, KAK, KKA\})=3/8,$$
$$P(X>1)=P(\{AAA, AAK, AKA, KAA\})= 4/8.$$

Παρόμοια,

$$P(1 \le X \le 2)=P(\{AKK, KAK, KKA , AAK, AKA, KAA\})=6/8,$$
$$P(X \ge 2)=P(\{AAK, AKA, KAA, AAA\})=4/8,$$

κ.τ.λ..

Έτσι έχουμε εισάγει μία συνάρτηση πιθανότητας στο πεδίο τιμών της X. Στο εξής θα γράφουμε απλούστερα (όπως στο επάνω παράδειγμα) $P(X=1)=3/8$, και εννοούμε την πιθανότητα ενός συγκεκριμένου γεγονότος

$$\{AKK, KAK, KKA\}=\{s \mid X(s)=1\}$$

στον δειγματικό χώρο S, το οποίο συμβαίνει με πιθανότητα 3/8. Θα συνεχίσουμε να γράφουμε εκφράσεις όπως $P(X=1)$, $P(X\leq 2)$, κ.τ.λ., αλλά είναι πολύ σημαντικό να γνωρίζει ο αναγνώστης τι πραγματικά παριστάνουν αυτές οι εκφράσεις. Αφού οι πιθανότητες που σχετίζονται με το πεδίο R_X έχουν προσδιορισθεί, συχνά θα *ξεχνούμε* τον δειγματικό χώρο S από τον οποίο προκύπτουν αυτές οι πιθανότητες. Έτσι, στο παραπάνω παράδειγμα, αν μας ενδιαφέρει να μελετήσουμε μόνο τις τιμές της τυχαίας μεταβλητής X, θα ασχοληθούμε με το πεδίο τιμών της, $R_X=\{0, 1, 2, 3\}$, και τις σχετικές πιθανότητες (1/8, 3/8, 3/8, 1/8), χωρίς να αναφερόμαστε αναλυτικότερα στον δειγματικό χώρο S.

Για να μελετήσουμε λεπτομερέστερα τις ***συναρτήσεις πιθανότητας*** που σχετίζονται με την τυχαία μεταβλητή, ξεχωρίζουμε την μεταβλητή σε τρεις βασικές περιπτώσεις: την ***διακριτή***, την ***συνεχή*** και την ***μικτή*** τυχαία μεταβλητή.

4.2 ΣΥΝΑΡΤΗΣΕΙΣ ΜΑΖΑΣ Ή ΠΥΚΝΟΤΗΤΑΣ ΠΙΘΑΝΟΤΗΤΑΣ

4.2.1 Διακριτή Τυχαία Μεταβλητή

Ορισμός. Εάν το σύνολο όλων των δυνατών τιμών της τυχαίας μεταβλητής $\boldsymbol{X}$ (δηλαδή το πεδίο τιμών R_X) είναι **πεπερασμένο** ή **άπειρο αριθμήσιμο**, ονομάζουμε την $\boldsymbol{X}$ **διακριτή τυχαία μεταβλητή**. Όλες οι δυνατές τιμές της $\boldsymbol{X}$ μπορούν να απαριθμηθούν σαν $\boldsymbol{x_1, x_2, x_3, \ldots, x_n, \ldots}$, όπου στην πεπερασμένη περίπτωση η αρίθμηση σταματά σε κάποιο n ενώ στην άπειρη αριθμήσιμη συνεχίζει ατελείωτα.

Παράδειγμα 4.4

Ένα τηλεφωνικό κέντρο έχει 48 γραμμές για τηλεφωνική επικοινωνία. Κάθε φορά που παρατηρούμε το κέντρο, κάποιες από τις γραμμές του είναι κατειλημμένες. Αν X είναι μία τυχαία μεταβλητή που δηλώνει πόσες γραμμές χρησιμοποιούνται ταυτόχρονα, τότε η μεταβλητή μπορεί να πάρει τιμές 0, 1, 2, 3, …48. Δηλαδή η τυχαία μεταβλητή X είναι διακριτή με πεδίο τιμών

$$R_X=\{0,1,2,\ldots,48\}.$$

Κάποια στιγμή παρατηρούμε το σύστημα και διαπιστώνουμε ότι δέκα γραμμές είναι κατειλημμένες, σημειώνουμε $X=x=10$. ■

Μέχρι τώρα έχουμε αναφερθεί σε αρκετές διακριτές μεταβλητές όπως

X={*ο αριθμός κεφαλών στην ρίψη ενός νομίσματος τρεις φορές*},
X={*ο αριθμός αγοριών σε τρία νεογέννητα παιδιά*},
X={*το άθροισμα των ενδείξεων της ρίψης ζαριού δύο φορές*},

X={*ο αριθμός ελαττωματικών εξαρτημάτων σε μία παραγωγή* }

και πολλές άλλες περιπτώσεις που θα αναφερθούν στην συνέχεια. Η πιθανοκρατική περιγραφή τέτοιων διακριτών τυχαίων μεταβλητών δεν παρουσιάζει καμία δυσκολία όπως φαίνεται παρακάτω.

Το μεγαλύτερο ενδιαφέρον μας επικεντρώνεται στην πιθανότητα η τυχαία μεταβλητή X να πάρει ειδικά μία ή κάποιες άλλες τιμές.

Η κατανομή πιθανότητας μιας τυχαίας μεταβλητής X είναι μία περιγραφή των πιθανοτήτων που σχετίζονται με τις δυνατές τιμές της X. Για μία διακριτή τυχαία μεταβλητή, η κατανομή ορίζεται συχνά από μία λίστα όλων των δυνατών τιμών της μαζί με τις αντίστοιχες πιθανότητες.

Ορισμός Για μία τυχαία μεταβλητή X με πεδίο τιμών $R_X=\{x_1, x_2, x_3,,\ldots, x_n,\ldots\}$ για κάθε δυνατή τιμή x_i σχετίζομε μία πιθανότητα $\boldsymbol{p(x_i)=P(X=x_i)}$**.** Το σύνολο που περιέχει τα διατεταγμένα ζευγάρια $\boldsymbol{(x_i\ ,\ p(x_i))}$, i=1, 2, 3, …, n,…, καλείται **συνάρτηση μάζας πιθανότητας (*σμπ*)** της X.

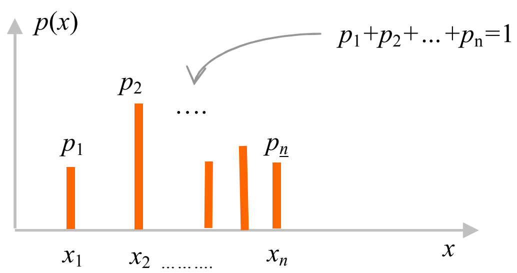
Μπορούμε να δηλώσουμε την συνάρτηση μάζας πιθανότητας και με έναν πιο απλό και πρακτικότερο τρόπο όπως

$$\boldsymbol{f(x_i)=P(X=x_i)=p(x_i)=p_i.}$$

Επειδή η συνάρτηση μάζας πιθανότητας ορίζεται σαν πιθανότητα , οι ακόλουθες **ιδιότητές της** είναι σχετικά προφανείς

$$\begin{aligned} &1)\quad f(x_i)\geq 0 \qquad \gamma\iota\alpha \quad \kappa\alpha\theta\varepsilon \quad i,\\ &2)\quad \sum_i f(x_i)=1. \end{aligned} \tag{4.1}$$

Θα μπορούσαμε να δώσουμε μία πραγματιστική ερμηνεία της συνάρτησης μάζας πιθανότητας, αν θεωρήσουμε μία ολική μάζα μιας μονάδας κατανεμημένη πάνω στην γραμμή των πραγματικών αριθμών με όλη την μάζα τοποθετημένη στα σημεία x_1, x_2, x_3,,…, x_n, και με $p(x_i)$ να δηλώνει την ποσότητα μάζας στα αντίστοιχα σημεία (Σχήμα 4.4)



Σχήμα 4.4 Συνάρτηση μάζας πιθανότητας

Παράδειγμα 4.5

Έστω ότι η τυχαία μεταβλητή X παριστάνει τον αριθμό των καλωδίων που χρειάζεται να ελεγχθούν μέχρι να εντοπισθεί ένα ελαττωματικό. Υποθέτουμε ότι η πιθανότητα ότι ένα καλώδιο είναι ελαττωματικό είναι 0.01 και ότι τα καλώδια είναι ανεξάρτητα. Προσδιορίστε την συνάρτηση μάζας πιθανότητας της X.

Έστω ότι κ δηλώνει το ενδεχόμενο το καλώδιο να είναι καλό, και έστω ε δηλώνει ένα ελαττωματικό καλώδιο. Ο δειγματικός χώρος του περάματος τύχης είναι άπειρος αριθμήσιμος και μπορεί να παρασταθεί από όλες τις δυνατές ακολουθίες που ξεκινούν με μία αλυσίδα από κ και τελειώνουν με ένα ε. Δηλαδή,

$$S=\{\varepsilon, \kappa\varepsilon, \kappa\kappa\varepsilon, \kappa\kappa\kappa\varepsilon, \kappa\kappa\kappa\kappa\varepsilon, \kappa\kappa\kappa\kappa\kappa\varepsilon, \ldots\}.$$

Παρατηρούμε μερικές περιπτώσεις. Έχουμε $P(X=1)=P(\varepsilon)=0{,}01$. Επίσης, χρησιμοποιώντας την υπόθεση της ανεξαρτησίας των καλωδίων

$$P(X=2)=P(\kappa\varepsilon)=0{,}99(0{,}01)=0{,}0099.$$

Μία γενική διατύπωση είναι

$$P(X=x)=P(\kappa\kappa\kappa\kappa\ldots\kappa\varepsilon)=0{,}99^{x-1}\cdot 0{,}01, \text{ για κάθε } x=1,2,3,\ldots$$

Περιγράφοντας τις πιθανότητες σχετικά με τις τιμές της X όπως στην παραπάνω φόρμουλα είναι ο απλούστερος τρόπος περιγραφής της συνάρτησης μάζας πιθανότητας της X σε αυτό το παράδειγμα.

Προφανώς $f(x)\geq 0$. Για να ελέγξουμε ότι οι πιθανότητες p_i ικανοποιούν την δεύτερη βασική ιδιότητα παρατηρούμε ότι

$$\sum_{x=1}^{\infty} P(X=x) = 0{,}01 + 0{,}99\cdot 0{,}01 + 0{,}99^2\cdot 0{,}01 + \ldots = 0{,}01(1+0{,}99+0{,}99^2+\ldots)$$

$$=0{,}01\frac{1}{1-0{,}99}=1$$

Χρησιμοποιούμε εδώ το αποτέλεσμα ότι η *γεωμετρική σειρά*

$$1+r+r^2+\ldots$$

συγκλίνει στο 1/(1-r) όταν $|r|<1$. ■

Οι τυχαίες μεταβλητές είναι τόσο σπουδαίες στα πειράματα τύχης έτσι που μερικές φορές βασικά παραμελούμε τον αρχικό δειγματοχώρο του πειράματος και συγκεντρωνόμαστε στην κατανομή μάζας πιθανότητας της τυχαίας μεταβλητής. Όπως π.χ., στο Παράδειγμα 4.2 μπορούμε περιληπτικά να αναφερθούμε στο πείραμα τύχης με δυνατές τιμές της *X*, {0, 1, 2, 3}. Στο Παράδειγμα 4.4 η ανάλυσή μας μπορεί να συγκεντρωθεί μόνο στο πεδίο τιμών της *X*, {0, 1, 2, . . ., 48}. Με αυτόν τον τρόπο, η τυχαία μεταβλητή μπορεί να απλουστεύσει την περιγραφή και ανάλυση ενός πειράματος τύχης.

Παρατηρήσεις

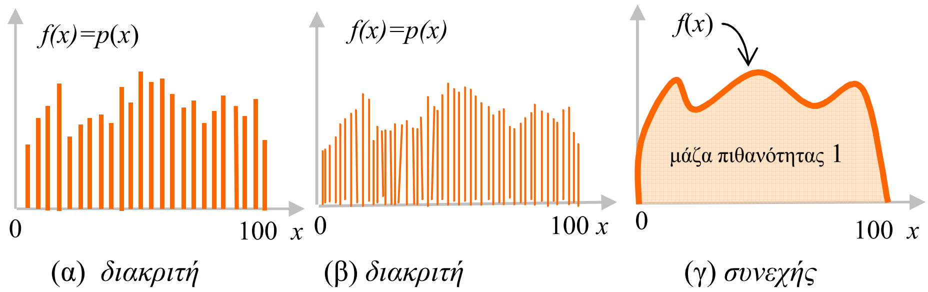
(A) Αν μία τυχαία μεταβλητή έχει σαν πεδίο τιμών της τον πεπερασμένο αριθμό τιμών $x_1, x_2, \ldots, x_n$ και κάθε ενδεχόμενο είναι ισοπίθανο, τότε προφανώς έχουμε $p(x_1)=p(x_2)=\ldots=p(x_n)=1/n$

(B) Εάν έχουμε μία τυχαία μεταβλητή με άπειρο αριθμήσιμο αριθμό τιμών, τότε είναι αδύνατο να έχουμε ισοπίθανα όλα τα ενδεχόμενα . Διότι δεν ικανοποιείται η συνθήκη $\sum_{i=1}^{\infty} p(x_i)=1$ εάν πρέπει να έχουμε $p(x_i)=c$ για κάθε *i*.

4.2.2 Συνεχής Τυχαία Μεταβλητή

Σε αρκετούς τύπους πειραμάτων, η ενδιαφέρουσα μέτρηση-αντοχή τάσης ενός υλικού, διάρκεια λειτουργίας μίας μηχανής, το μήκος ενός καλωδίου σε μία παραγωγή-μπορούν να παρασταθούν από μία τυχαία μεταβλητή. Είναι λογικό να περιγράψουμε το πεδίο όλων των δυνατών τιμών της τυχαίας μεταβλητής από ένα διάστημα (πεπερασμένο, απεριόριστο) πραγματικών αριθμών. Για παράδειγμα, ο χρόνος λειτουργίας μίας μηχανής θα μπορούσε να ήταν οιοσδήποτε αριθμός εντός κάποιου διαστήματος πραγματικών αριθμών. Επειδή το πεδίο είναι κάθε τιμή μέσα στο διάστημα, ο χρόνος λειτουργίας μετριέται με οιαδήποτε ακρίβεια. Ως εκ τούτου, επειδή ο αριθμός όλων των δυνατών τιμών της τυχαίας μεταβλητής *X* είναι άπειρος μη μετρήσιμος, η *X* έχει μία ξεχωριστή διαφορετική κατανομή από τις διακριτές τυχαίες μεταβλητές που μελετήθηκαν μέχρι τώρα. Το πεδίο τιμών της *X* περιλαμβάνει όλες τις τιμές σε κάποιο διάστημα των πραγματικών αριθμών, δηλαδή, το πεδίο της *X* μπορεί να θεωρηθεί σαν ένα συνεχές.

Έστω ότι έχουμε μία διακριτή τυχαία μεταβλητή *X* η οποία μπορεί να πάρει 25 διακριτές τιμές από το 0 μέχρι το 100 με συνάρτηση μάζας πιθανότητας όπως φαίνεται στο Σχήμα 4.5 (α). Ας πούμε τώρα ότι η τυχαία μεταβλητή X μπορεί να πάρει περισσότερες τιμές στο διάστημα $0 \le x \le 100$, όπου ο αριθμός τιμών είναι μεγάλος αλλά πεπερασμένος, όπως 1, 2, 3, . . ., 98, 99, 100. Η μάζα πιθανότητας των 25 μπάρων απλώνεται τώρα σε περισσότερα σημεία της X, όπως παριστάνεται στην περίπτωση (β) του Σχήματος 4.5.



Σχήμα 4.5 (α), (β), συναρτήσεις μάζας πιθανότητας, (γ) συνάρτηση πυκνότητας πιθανότητας

Σε κάθε μία από τις δυνατές τιμές x_i αντιστοιχεί μία πιθανότητα $p_i = p(x_i) = P(X=x_i)$, όπου i=1,2,…100, των οποίων το άθροισμα ισούται με την μονάδα.

Αναφέραμε παραπάνω ότι η τυχαία μεταβλητή X ενδέχεται να πάρει οιαδήποτε τιμή μέσα στο διάστημα , $0 \le x \le 100$. Αφού οι τιμές που μπορεί να πάρει η X είναι άπειρες μη μετρήσιμες, τι θα συμβεί με τις πιθανότητες $p_i = p(x_i)$; Οι πιθανότητες p_i είναι άπειρες τον αριθμό και το άθροισμά τους θα οδηγούσε σε τιμή μεγαλύτερη της μονάδας. Οι πιθανότητες αυτές έχουν μηδενική τιμή και δεν έχουν νόημα. Εκείνο που πρέπει να κάνουμε είναι να σκορπίσουμε την μάζα των πιθανοτήτων $p(x_i)$ που ορίζονται μόνο για τις τιμές $x_1, x_2, x_3, \ldots, x_n$ σε όλο το διάστημα , $0 \le x \le 100$, όπως φαίνεται στο Σχήμα 4.5 (γ). Προκύπτει έτσι μία συνάρτηση f ορισμένη (για το παρών παράδειγμα) για όλες τις τιμές x, $0 \le x \le 100$, η οποία εκφράζει την πυκνότητα της μάζας πιθανότητας σε όλο το πεδίο τιμών της X.

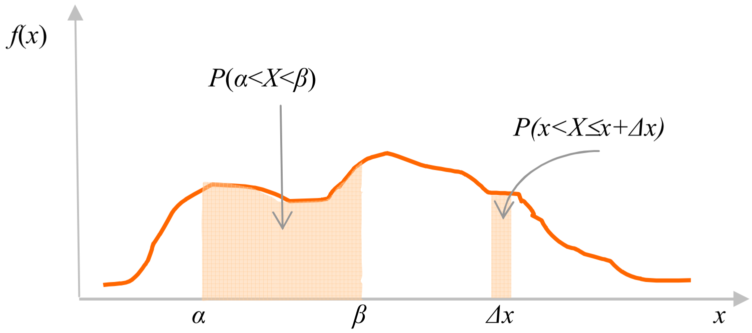
Ορισμός. Για μία **συνεχή** τυχαία X, μία συνάρτηση ***f(x)*** ονομάζεται **συνάρτηση πυκνότητας πιθανότητας** (***σππ***) αν πληροί τις ακόλουθες συνθήκες:

1) $f(x) \ge 0$ (4.2)

2) $$\int_{-\infty}^{\infty} f(x)dx = 1$$ (4.3)

3) $$P(a \le X \le \beta) = \int_{a}^{\beta} f(x)dx \ .$$ (4.4)

Παρόμοια με την διακριτή και στην συνεχή τυχαία μεταβλητή η συνάρτηση πυκνότητας πιθανότητας, $f(x)$, χρησιμοποιείται για να περιγράψει την κατανομή πιθανότητας της X. Η τιμή της *σππ* $f(x)$ αυξάνει στις πιο πιθανές περιοχές του πεδίου τιμών της X, και μηδενίζεται για εκείνες τις τιμές x που δεν μπορούν να συμβούν.



***Σχήμα* 4.6** Πιθανότητα προσδιοριζόμενη από την επιφάνεια κάτω από την $f(x)$

Η **συνάρτηση πυκνότητας πιθανότητας,** $f(x)$, είναι ένα μαθηματικό εργαλείο αρκετά χρήσιμο για την διερεύνηση και μελέτη της τυχαίας μεταβλητής X.

Παρατηρήσεις

(A) ΣΕ μία συνεχή τυχαία μεταβλητή X υποθέτουμε ότι **μάζα πιθανότητας 1** κατανέμεται σε όλο το πεδίο τιμών της και κάτω από την επιφάνεια της $f(x)$. Η πιθανότητα, ότι η μεταβλητή X πέφτει μεταξύ $α$ και $β$, προσδιορίζεται από το ολοκλήρωμα της $f(x)$ από το $α$ έως $β$,

$$P(a \leq X \leq \beta) = \int_{\alpha}^{\beta} f(x)dx$$

και παριστάνει την μάζα πιθανότητας που περιλαμβάνεται κάτω από την επιφάνεια της $f(x)$ στο διάστημα $α \leq x \leq β$ όπως φαίνεται στο Σχήμα 4.6.

(B) Η *σππ* $f(x)$ παριστάνει την πυκνότητα της μάζας πιθανότητας στο σημείο x, με την οποία μπορούμε να προσδιορίσουμε την πιθανότητα η X να πέσει σε ένα διάστημα αν την ολοκληρώσουμε. Η συνάρτηση $\boldsymbol{f(x)}$ **δεν παριστάνει πιθανότητα,** μόνον όταν η συνάρτηση ολοκληρώνεται μεταξύ δύο ορίων προκύπτει πιθανότητα. Αυτό μπορούμε να το τεκμηριώσουμε εύκολα με το ακόλουθο σκεπτικό. Από τον ορισμό της *σππ* προκύπτει

$$P(x \leq X \leq x + \Delta x) = \int_{x}^{x+\Delta x} f(u)du = \Delta x \cdot f(x),$$

όπου Δx μία πολύ μικρή ποσότητα, βλέπε Σχήμα 4.6. Δηλαδή, το γινόμενο $\Delta x{\cdot}f(x)$ παριστάνει την πιθανότητα $P(x \leq X \leq x{+}\Delta x)$, όχι όμως το $f(x)$ από μόνο του.

Σαν συνέπεια της παραπάνω πιθανοκρατικής περιγραφής της X είναι ότι για κάθε συγκεκριμένη τιμή της μεταβλητή X, ας πούμε x_0 έχουμε

$$P(X = x_0) = P(x_0 \leq X \leq x_0) = \int_{x_0}^{x_0} f(x)dx = 0\,.$$

Το πλάτος της επιφάνειας κάτω από την $f(x)$ στο σημείο x_0 είναι μηδέν και για αυτό η πιθανότητα η X να πάρει την τιμή x_0 είναι μηδέν. Το αποτέλεσμα αυτό ακούγεται λίγο περίεργο. Πρέπει όμως να γίνει κατανοητό ότι εάν η X μπορεί να πάρει τιμές μέσα σε κάποιο διάστημα, και αν x_0 είναι μία τιμή του διαστήματος αυτού, τότε το ότι $P(X{=}x_0){=}0$ δεν σημαίνει ότι το γεγονός $\{X{=}x_0\}$ είναι αδύνατο. Το γεγονός $\{X{=}x_0\}$ μπορεί να συμβεί αλλά με μία πάρα πολύ μικρή πιθανότητα που τείνει στο μηδέν, μηδενική όπως λέγεται διαφορετικά. Εξάλλου έχει τονισθεί στο κεφάλαιο της πιθανοθεωρίας ότι όταν ένα γεγονός είναι αδύνατο $A{=}\varnothing$, τότε η πιθανότητα να πραγματοποιηθεί σε κάποια εκτέλεση του πειράματος τύχης είναι μηδέν, $P(A){=}0$. Δεν συμβαίνει όμως το αντίστροφο. Κάθε γεγονός A με μηδενική πιθανότητα πραγματοποίησης δεν είναι πάντα το αδύνατο γεγονός $\varnothing$.

Παραδειγματικά θα μπορούσαμε να αναφέρουμε την περίπτωση μίας τυχαίας μεταβλητής X με πεδίο τιμών όλους τους πραγματικούς του διαστήματος (0, 2), $\{x \mid 0 \leq x \leq 2\}$. Όλοι συμφωνούμε πως όταν εκτελείται το πείραμα τύχης είναι βέβαιο ότι κάθε σημείο του διαστήματος (0, 2) μπορεί να είναι το αποτέλεσμα του πειράματος. Θα μας φαινόταν όμως εξαιρετικά καταπληκτικό αν σε μία εκτέλεση η X πάρει ακριβώς την μεσαία τιμή του διαστήματος ή κάποιο άλλο συγκεκριμένο σημείο του διαστήματος, διότι το διάστημα (0, 2) έχει πάρα πολλά σημεία άπειρα το πλήθος.

Πρέπει να τονίσουμε εδώ ότι το μοντέλο μιας συνεχούς μεταβλητής X δεν είναι λιγότερο χρήσιμο επειδή $P(X{=}x_0){=}0$. Εάν στην πράξη μας ενδιαφέρει η πιθανότητα η X να πάρει τιμή πολύ κοντά στο x_0, μπορούμε να εκτιμήσουμε την πιθανότητα $P(x_0{-}\varepsilon \leq X \leq x_0{+}\varepsilon)$, όπου ε είναι μία αυθαίρετα μικρή τιμή. Αν για παράδειγμα στην παραπάνω περίπτωση επιθυμούσαμε να εκτιμήσουμε την πιθανότητα η τυχαία μεταβλητή X να πάρει τιμή κοντά στο 1,47, τότε πρέπει να υπολογίσουμε $P(1{,}465 \leq X \leq 1{,}475)$.

Γενικά, εάν η X είναι συνεχής τυχαία μεταβλητή ισχύει

$$P(a \le X \le \beta) = P(a \le X < \beta) = P(a < X \le \beta) = P(a < X < \beta).$$

Παράδειγμα 4.6

Υποθέτουμε ότι η τυχαία μεταβλητή X μπορεί να πάρει ισοπίθανα τιμές στο διάστημα πραγματικών αριθμών $[a,\beta]$, Σχήμα 4.7. Δηλαδή, έχει σταθερή ή ομοιόμορφη συνάρτηση πυκνότητας πιθανότητας στο διάστημα $[a,\beta]$. Η αλγεβρική της έκφραση είναι

$$f(x) = \begin{cases} c & \gamma\iota\alpha \quad a \le x \le \beta \\ 0 & \alpha\lambda\lambda o\upsilon \end{cases}$$

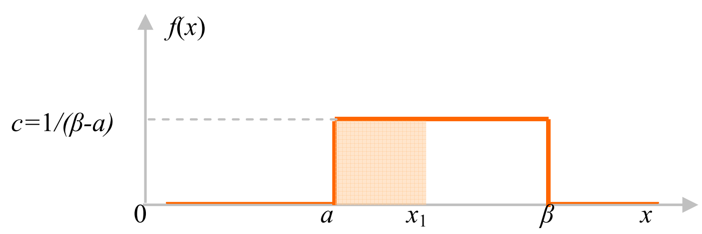
Από την βασική ιδιότητα της συνάρτησης πυκνότητας πιθανότητας (συνθήκη 2) υπολογίζουμε την σταθερά c

$$\int_a^\beta f(x)dx = \int_a^\beta c dx = c[x]_a^\beta = c(\beta - a) = 1$$

$$\Rightarrow c = \frac{1}{\beta - a}$$

Η πιθανότητα ότι η $X \le x_1$ φαίνεται παριστάνεται από την σκιαγραφημένη επιφάνεια στο Σχήμα 4.7 και υπολογίζεται από το ολοκλήρωμα,

$$P(X \le x_1) = \int_{-\infty}^{x_1} \frac{1}{\beta - a} dx = \frac{1}{\beta - a}\int_a^{x_1} dx = \frac{1}{\beta - a}[x]_a^{x_1} = \frac{x_1 - a}{\beta - a}$$



Σχήμα 4.7 Συνάρτηση πυκνότητας πιθανότητας για μία ομοιόμορφη κατανομή

Παράδειγμα 4.7

Έστω ότι X είναι η τυχαία μεταβλητή η οποία παριστάνει την διάμετρο της τρύπας που ανοίγεται σε ένα μεταλλικό εξάρτημα. Η επιθυμητή διάμετρος είναι 12.5 mm. Αν και το τρυπάνι φέρει αρίδα για τρύπα 12,5 mm, ο θόρυβος, οι κραδασμοί και άλλοι στοχαστικοί παράγοντες προκαλούν διάμετρο μεγαλύτερη από 12,5 mm. Έτσι, σε μία ανοιγμένη τρύπα η συνάρτηση πυκνότητας πιθανότητας για την διάμετρο X είναι

$$f(x) = \begin{cases} ce^{-10(x-12,5)} & \text{για} \quad x > 12,5\ mm \\ 0 \quad \text{αλλού} & \end{cases}$$

(α) Να προσδιορίσετε την παράμετρο c.
(β) Εάν μία τρύπα ξεπερνά τα 12,6 mm το μεταλλικό εξάρτημα αχρηστεύεται. Ποιο ποσοστό μεταλλικών εξαρτημάτων αχρηστεύεται;

Πρέπει να έχουμε

$$\int_{12,5}^{\infty} ce^{-10(x-12,5)}dx = 1, \text{ όπου } c\int_{12,5}^{\infty} e^{-10(x-12,5)}dx = c(-\frac{1}{10})e^{-10(x-12,5)}\Big|_{12,5}^{\infty} = \frac{c}{10}$$

επομένως c=10.
Έτσι,

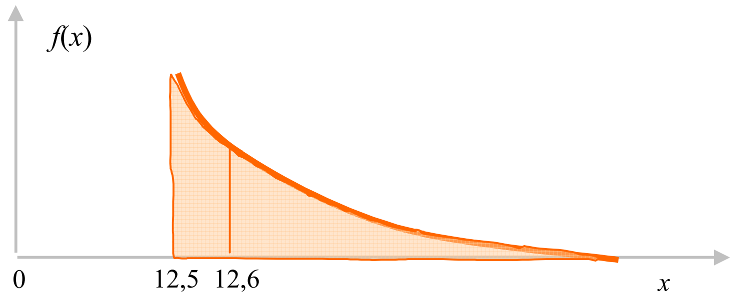
$$P(X > 12,6) = 10\int_{12,6}^{\infty} e^{-10(x-12,5)}dx = 10(-\frac{1}{10})e^{-10(x-12,5)}\Big|_{12,6}^{\infty} = e^{-10(12,6-12,5)} - 0 = e^{-1}$$

Ποιο ποσοστό από τις τρύπες είναι μεταξύ 12,5 και 12,6 mm;

$$P(12,5 < X < 12,6) = 10\int_{12,5}^{12,6} e^{-10(x-12,5)}dx = 10(-\frac{1}{10})e^{-10(x-12,5)}\Big|_{12,5}^{12,6}$$
$$= e^{-10(12,5-12,5)} - e^{-10(12,6-12,5)} = 1 - e^{-1}$$

Επειδή η ολική επιφάνεια κάτω από την *σππ f(x)* ισούται με ένα, μπορούμε επίσης να υπολογίσουμε

$$P(12.5<X<12,6)=1-P(X>12,6)=1-e^{-1}$$



Σχήμα 4.8 Συνάρτηση πυκνότητας πιθανότητας για το

4.3 ΑΘΡΟΙΣΤΙΚΗ ΣΥΝΑΡΤΗΣΗ ΠΙΘΑΝΟΤΗΤΑΣ

Στα παραδείγματα 4.2, 4.3, 4.4 διαφαίνεται ότι θα ήταν χρήσιμο να εκφράσουμε την αθροιστική πιθανότητα όπως $P(X \leq x)$ με κάποια φόρμουλα. Ακόμη, μία εναλλακτική μέθοδο περιγραφής της κατανομής πιθανότητας μιας τυχαίας μεταβλητής είναι η αθροιστική συνάρτηση πιθανότητας.

Ορισμός. Έστω μία τυχαία μεταβλητή X. Η συνάρτηση $\boldsymbol{F(x)}$ η οποία πληροί την ακόλουθη ισότητα

$$\boldsymbol{F(x)=P(X \leq x) \quad \text{για} -\infty < x < \infty} \qquad (4.5)$$

ονομάζεται **αθροιστική συνάρτηση πιθανότητας (*ασπ*)** ή απλά **συνάρτηση κατανομής**.

Η αθροιστική συνάρτηση πιθανότητας μπορεί να ορισθεί για κάθε πραγματικό αριθμό x, $-\infty < x < \infty$, έστω και αν δεν ανήκει στο πεδίο τιμών της X. Λεκτικά, η αθροιστική συνάρτηση πιθανότητας, $F(x)=P(X \leq x)$, εκφράζει την πιθανότητα ότι η τυχαία μεταβλητή X παίρνει μία τιμή, η οποία είναι μικρότερη ή ίση με την τιμή x. Η αθροιστική συνάρτηση F μπορεί να εκφρασθεί από την συνάρτηση πυκνότητας πιθανότητας f, όπως ακολουθεί για τις περιπτώσεις **διακριτής** και **συνεχούς** τυχαίας μεταβλητής αντίστοιχα.

4.3.1 Διακριτή Τυχαία Μεταβλητή

Εάν η X είναι **διακριτή τυχαία μεταβλητή**, το γεγονός $\{X \leq x\}$ είναι ισοδύναμο με το γεγονός ότι η τυχαία μεταβλητή X θα πάρει μία οιαδήποτε τιμή x_i μικρότερη του x, δηλαδή, με την ένωση των ασυμβίβαστων μεταξύ τους γεγονότων,

$$\{X \leq x\} = \{(X = x_1) \cup (X = x_2) \cup \ldots \cup (X = x_i)\}.$$

Σύμφωνα με το τρίτο αξίωμα του Kolmogorov καταλήγουμε στο παρακάτω αποτέλεσμα,

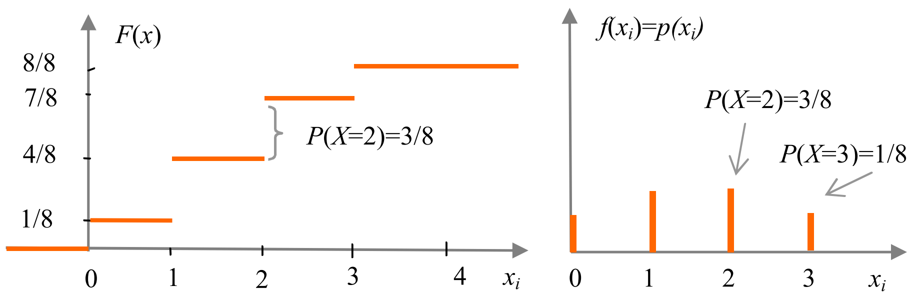
$$F(x) = P(X \le x) = \sum_{x_i \le x} P(X = x_i) = \sum_{x_i \le x} f(x_i) = \sum_{x_i \le x} p(x_i). \qquad (4.6)$$

Παράδειγμα 4.8

Στο παράδειγμα 4.2 η αθροιστική συνάρτηση πιθανότητας της **διακριτής** τυχαίας μεταβλητής X, η οποία παριστάνει τον αριθμό αγοριών σε τέσσερα νεογέννητα παιδιά, είναι

$$F(x) = \begin{cases} P(X \le x) = 0 & \text{για} \quad x < 0 \\ P(X \le 0) = \dfrac{1}{8} & \text{για} \quad 0 \le x < 1 \\ P(X \le 1) = \dfrac{4}{8} & \text{για} \quad 1 \le x < 2 \\ P(X \le 2) = \dfrac{7}{8} & \text{για} \quad 2 \le x < 3 \\ P(X \le 3) = \dfrac{8}{8} & \text{για} \quad x \ge 3 \end{cases}$$

Η γραφική παράσταση της συνάρτησης κατανομής αυτού του παραδείγματος δίνεται στο Σχήμα 4.9 .
Παρατηρούμε ότι η $F(x)$ ορίζεται για όλα τα x από $-\infty < x < \infty$ -όχι μόνο για 0, 1, 2, και 3. Για παράδειγμα η τιμή της αθροιστικής στο σημείο x=1,5 είναι F(1,5)=4/8. Στην διακριτή τυχαία μεταβλητή X με πεπερασμένο αριθμό τιμών η γραφική παράσταση της αθροιστικής $F(x)$ αποτελείται από οριζόντια ευθύγραμμα τμήματα, γι' αυτό καλείται και **σκαλωτή** συνάρτηση. Η συνάρτηση $F(x)$ είναι συνεχής εκτός των πιθανών τιμών της X, όπως x_1, x_2, x_3,,…, x_n. Στο σημείο x_i η γραφική παράσταση παρουσιάζει ένα ''άλμα'' ύψους



Σχήμα 4.9. Γραφικές παραστάσεις αθροιστικής συνάρτησης πιθανότητας και συνάρτησης μάζας πιθανότητας διακριτής τυχαίας μεταβλητής.

$p(x_i)=P(X=x_i)$. Για παράδειγμα, στο Σχήμα 4.9, η αθροιστική $F(x)$ στο σημείο μηδέν παρουσιάζει άλμα ίσο με την πιθανότητα $p(0)=P(X=0)=1/8$. Παρόμοια και για τις άλλες τιμές της X στα σημεία 1, 2 και 3 τα ‘’άλματα’’ στην αθροιστική είναι ίσα με τις αντίστοιχες πιθανότητες της συνάρτησης μάζας πιθανότητας $f(x_i)$, δηλαδή, $p(1)=3/8$, $p(2)=3/8$ και $p(3)=1/8$, όπως φαίνεται και στις γραφικές παραστάσεις του Σχήματος 4.9.

4.3.2 Συνεχής Τυχαία Μεταβλητή

Εάν η X είναι **συνεχής τυχαία μεταβλητή**,

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(u)du\,. \tag{4.7}$$

Το αποτέλεσμα αυτό προκύπτει άμεσα από τον ορισμό της αθροιστικής συνάρτησης F και της συνάρτησης πυκνότητας πιθανότητας f.

Παράδειγμα 4.9

Η διάρκεια καλής λειτουργίας ενός τύπου ηλεκτρικής γεννήτριας, X, είναι **συνεχής** τυχαία μεταβλητή, και μπορεί να πάρει τιμή ισοπίθανα στο διάστημα [100, 200 μήνες]. Ποια η πιθανότητα ότι η διάρκεια μιας τέτοιας γεννήτριας θα βρεθεί στο διάστημα από 150 μέχρι 180 μήνες;

Η συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής X, $f(x)$, δείχνεται στο Σχήμα 4.10. Η $f(x)$ είναι σταθερή (*ομοιόμορφη*) σε όλο το διάστημα από 100 μέχρι 200 μήνες. Η σταθερή τιμή c, όπως φαίνεται και στο Σχήμα 4.10, μπορεί να υπολογισθεί από την βασικότερη ιδιότητα της $f(x)$,

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{100} 0 \cdot dx + \int_{100}^{200} c \cdot dx + \int_{200}^{\infty} 0 \cdot dx = 1$$

$$\int_{100}^{200} c \cdot dx = c \cdot 100 = 1, \quad \Rightarrow \quad c = \frac{1}{100}.$$

Έτσι, μπορούμε να υπολογίσουμε την ζητούμενη πιθανότητα με βάση την ιδιότητα 3 της $f(x)$

$$P(150 \le X \le 180) = \int_{150}^{180} \frac{1}{100} dx = \frac{1}{100}[x]_{150}^{180} = \frac{30}{100} = 0{,}30$$

Η πιθανότητα που υπολογίσαμε φαίνεται στο Σχήμα 4.10 με την σκιασμένη επιφάνεια κάτω από την $f(x)$.

Η αθροιστική συνάρτηση κατανομής της X, $F(x)$, αποτελείται από τρία σκέλη,

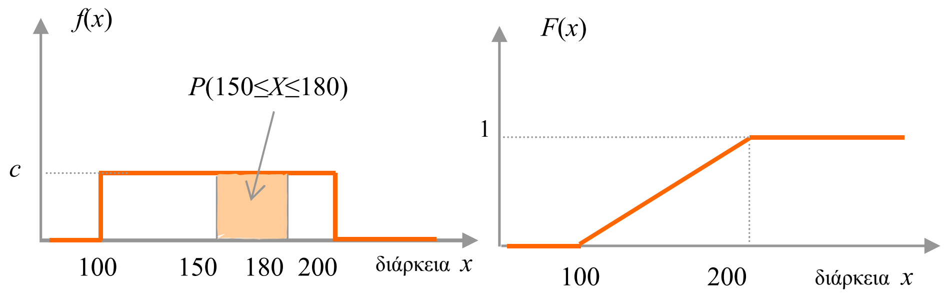
α) $F(x)=P(X\le x)=0$ για $x<100$

β) $F(x) = \int_{100}^{x} f(u)du = \int_{100}^{x} \frac{1}{100} du = \frac{x-100}{100}$ για $100\le x\le 200$

γ) $F(x) = \int_{100}^{200} f(x)dx = 1$ για $x\ge 200$

Τελικά η αθροιστική συνάρτηση κατανομής της διάρκειας X δίνεται από την ακόλουθη μαθηματική φόρμουλα:

$$F(x) = \begin{cases} 0 & \gamma\iota\alpha \quad x < 100 \\ \dfrac{x-100}{100} & \gamma\iota\alpha \quad 100 \le x \le 200 \\ 1 & \gamma\iota\alpha \quad x > 200 \end{cases}$$



Σχήμα 4.10 Γραφική παράσταση συνάρτησης πυκνότητας πιθανότητας και αθροιστικής συνάρτησης κατανομής της διάρκειας X.

Η γραφική παράσταση στο Σχήμα 4.10 είναι μία κλασσική περίπτωση μίας ομοιόμορφης συνάρτησης πυκνότητας συνεχούς τυχαίας μεταβλητής. Παρατηρούμε ότι η αθροιστική συνάρτηση $F(x)$ είναι παντού συνεχής.

4.3.3 Ιδιότητες Αθροιστικής Συνάρτησης Πιθανότητας $F(x)$

Υπάρχουν αρκετές σημαντικές ιδιότητες της αθροιστικής συνάρτησης πιθανότητας οι οποίες αναφέρονται περιληπτικά στη συνέχεια

Ιδιότητες της αθροιστικής κατανομής $F(x)$

α) Η συνάρτηση $F(x)$ είναι **αύξουσα.** Δηλαδή, αν $x_1 \le x_2$ έχουμε $F(x_1) \le F(x_2)$.

β) $F(-\infty)=0$ και $F(\infty)=1$.

γ) Η $F(x)$ είναι πάντα **συνεχής από τα δεξιά**. Δηλαδή,

$$\lim_{e \to 0} F(x+e) = F(x)$$

δ) για $x_1 \le x_2$ ισχύει

$$P(x_1 < X \le x_2) = F(x_2) - F(x_1).$$

ε) $P(X=x) = F(x) - \lim_{e \to 0} F(x-e)$

- αν X συνεχής, $P(X=x)=0$
- αν X διακριτή στο σημείο x, $P(X=x) \neq 0$.

στ)

- Αν η X είναι **διακριτή** τυχαία μεταβλητή με τιμές $x_1 < x_2 < \ldots$ τότε

$$p(x_i) = P(X=x_i) = F(x_i) - F(x_{i-1})$$

- Αν η X είναι **συνεχής** τυχαία μεταβλητή τότε ισχύει

$$f(x) = \frac{d}{dx} F(x)$$

Για τον αναγνώστη θα ήταν μία καλή εξάσκηση αν αποδείκνυε τις παραπάνω ιδιότητες. Σε περίπτωση όμως δυσκολίας, είναι πολύ σημαντικό να μελετήσει κανείς προσεκτικά τις αποδείξεις που ακολουθούν.

α) αν $\boldsymbol{x_1 \le x_2}$ έχουμε $\boldsymbol{F(x_1) \le F(x_2)}$

Απόδειξη

Ορίζουμε τα γεγονότα $A=\{X \le x_1\}$ και $B=\{X \le x_2\}$. Επειδή $x_1 \le x_2$ το γεγονός B είναι υπερσύνολο του A, $A \subset B$ και από το θεώρημα 2.5, $P(A) \le P(B)$ ή $P(X \le x_1) \le P(X \le x_2)$.

β) $\boldsymbol{F(-\infty)=0}$ **και** $\boldsymbol{F(\infty)=1}$

Απόδειξη

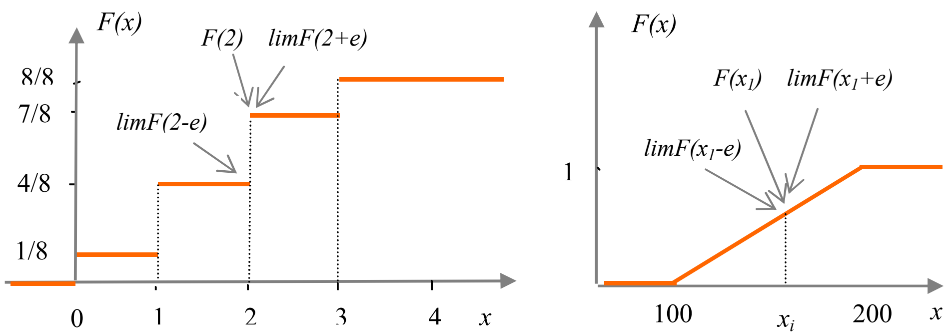
$F(-\infty)=P(X\le -\infty)$. Ως γνωστόν $\{X\le -\infty\}=\emptyset$, άρα $F(-\infty)=P(\emptyset)=0$.
$F(\infty)=P(X \le \infty)$. Ως γνωστόν $\{X \le \infty\}=S$, άρα $F(\infty)=P(S)=1$.

γ) Η $\boldsymbol{F(x)}$ είναι **πάντα** (για διακριτή και συνεχή τυχαία μεταβλητή) **συνεχής από τα δεξιά.** Δηλαδή, $\lim_{e\to 0} F(x+e)=F(x)$ **.**

Απόδειξη

Παρατηρώντας στο Σχήμα 4.11 τις γραμμές της αθροιστικής τόσο για την διακριτή μεταβλητή όσο και για την συνεχή, η απόδειξη σχηματικά είναι προφανής. Η πλήρη απόδειξη είναι

$$\lim_{e\to 0} F(x+e)=\lim_{e\to 0} P(X \le x+e)=\lim_{e\to 0} P(\{X \le x\}\cup\{x < X \le x+e\})=$$
$$P\{X \le x\}+\lim_{e\to 0} P\{x < X \le x+e\}=P(X \le x)+0=F(x)$$



Σχήμα 4.11 Γραφικές παραστάσεις αθροιστικών συναρτήσεων των παραδειγμάτων 4.8 και 4.9

δ) για $\mathbf{x_1 \leq x_2}$ ισχύει $\boldsymbol{P(x_1 < X \leq x_2)=F(x_2) - F(x_1)}$

Απόδειξη

Αφού υποθέσαμε $x_1 \leq x_2$, το γεγονός $\{x_1 < X \leq x_2\}$ περιλαμβάνει εκείνες τις τιμές της X που ανήκουν στο σύνολο $B=\{X \leq x_2\}$ και όχι στο $A=\{X \leq x_1\}$. Άρα το γεγονός $\{x_1 < X \leq x_2\}$ είναι η *Διαφορά* των γεγονότων B και A και σύμφωνα με το θεώρημα 2.4 έχουμε,

$$P(x_1 < X \leq x_2) = P(A) - P(A \cap B) = P\{X \leq x_2\} - P\{X \leq x_1\} = F(x_2) - F(x_1)$$

ε) $\boldsymbol{P(X = x) = F(x) - \lim_{e \to 0} F(x-e)}$.

Απόδειξη

Από την δ) για $x_2=x$ και $x_1=x-e$ έχουμε

$$\lim_{e \to 0} P(x-e < X \leq x) = P(X = x) = F(x) - \lim_{e \to 0} F(x-e)$$

Εδώ διακρίνουμε δύο περιπτώσεις:

- Αν η X είναι **συνεχής** τότε $\boldsymbol{P(X=x)=0}$ και προφανώς η αθροιστική είναι **συνεχής και από τα αριστερά**.
- Αν η X είναι διακριτού τύπου τότε ενδέχεται $P(X=x) \neq 0$ και κατά συνέπεια η αθροιστική στο σημείο x **δεν είναι συνεχής από τα αριστερά**.

στ)

- Αν η X είναι **διακριτή** τυχαία μεταβλητή με τιμές $x_1 < x_2 < \ldots$ τότε

  $$\boldsymbol{p(x_i)=P(X=x_i)=F(x_i)-F(x_{i-1})}$$

  Απόδειξη

  Για την διακριτή τυχαία μεταβλητή X έχουμε

$$F(x_i) = P(X \le x_i) = P((X = x_i) \cup (X = x_{i-1}) \cup ... \cup (X = x_1))$$
$$= P(X = x_i) + P(X = x_{i-1}) + ... + P(X = x_1)$$
και
$$F(x_{i-1}) = P(X \le x_{i-1}) = P((X = x_{i-1}) \cup (X = x_{i-2}) \cup ... \cup (X = x_1))$$
$$= P(X = x_{i-1}) + P(X = x_{i-2}) + ... + P(X = x_1)$$
οποτε
$$F(x_i) - F(x_{i-1}) = P(X = x_i) = p(x_i)$$

- Αν η X είναι **συνεχής** τυχαία μεταβλητή τότε ισχύει

$$\boldsymbol{f(x) = \frac{d}{dx} F(x)}$$

Απόδειξη

Για την **συνεχή** τυχαία μεταβλητή X έχουμε

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(u)du \,.$$

Έτσι εφαρμόζοντας το βασικό θεώρημα της ολοκλήρωσης επιτυγχάνουμε, $f(x) = F'(x)$.

Παρατηρήσεις

(A) Η συνάρτηση κατανομής $F(x)$ και η συνάρτηση πυκνότητας $f(x)$ αποτελούν πλήρεις και ισοδύναμες περιγραφές μιας τυχαίας μεταβλητής X. Έτσι η γνώση της $f(x)$ μας δίνει μέσω της σχέσης (4.7) την αντίστοιχη $F(x)$. Το αντίστροφο επιβεβαιώνεται από την ιδιότητα στ).

(B) Αν μία συνάρτηση $F(x)$ πληροί τις ιδιότητες α), β) και γ), τότε είναι πράγματι αθροιστική συνάρτηση πιθανότητας μιας τυχαίας μεταβλητής X.

Παράδειγμα 4.10

Ο χρόνος μέχρι την ολοκλήρωση μιας χημικής αντίδρασης προσεγγίζεται από την ακόλουθη αθροιστική συνάρτηση κατανομής

$$F(x) = \begin{cases} 0 & \text{για} \quad x < 0 \\ 1 - e^{-0,01x} & \text{για} \quad 0 \le x \end{cases}$$

Προσδιορίστε την συνάρτηση πυκνότητας πιθανότητας του χρόνου ολοκλήρωσης X, $f(x)$. Ποιο ποσοστό των χημικών αντιδράσεων ολοκληρώνεται εντός χρόνου 200 milliseconds; Χρησιμοποιώντας το αποτέλεσμα της ζ) ιδιότητας, όπου η συνάρτηση πυκνότητας πιθανότητας είναι η παράγωγος της $F(x)$, έχουμε

$$f(x) = \frac{dF(x)}{dx} = \begin{cases} 0 & \gamma\iota\alpha \quad x < 0 \\ 0{,}01e^{-0{,}01x} & \gamma\iota\alpha \quad 0 \le x \end{cases}$$

Το ποσοστό των χημικών αντιδράσεων που ολοκληρώνεται εντός χρόνου 200 milliseconds, είναι η πιθανότητα

$$P(X \le 200) = F(200) = 1 - e^{-0{,}01 \cdot 200} = 1 - e^{-2} = 0{,}8647 .$$ ■

Παράδειγμα 4.11

Η αθροιστική συνάρτηση πιθανότητας σε μία τυχαία μεταβλητή Χ είναι

$$F(x) = \begin{cases} 0 & \gamma\iota\alpha \quad x < 0 \\ 0{,}2x & \gamma\iota\alpha \quad 0 \le x < 5 \\ 1 & \gamma\iota\alpha \quad x \ge 5 \end{cases}$$

Προσδιορίστε τις πιθανότητες α) $P(X=4)$, β) $P(X < 2{,}8)$, γ) $P(X > 1{,}5)$, δ) $P(X < -2)$, ε) $P(X>6)$, στ) $P(1{,}2 < X < 4)$. Παρατηρούμε ότι η αθροιστική $F(x)$ είναι παντού συνεχής, η γραφική της παράσταση είναι παρόμοια του Σχήματος 4.10, και κατά συνέπεια η τυχαία μεταβλητή είναι συνεχής. Άρα λοιπόν, κάνοντας χρήση τις παραπάνω ιδιότητες έχουμε

α) $P(X=4)=0$
β) $P(X< 2{,}8)= P(X\le 2{,}8)= F(2{,}8)=0{,}2\cdot 2{,}8=0{,}56$
γ) $P(X> 1{,}5)=1- P(X\le 1{,}5)= 1- F(1{,}5)=1-0{,}2\cdot 1{,}5=1-0{,}3=0{,}7$
δ) $P(X< -2)= P(X\le -2)= F(-2)=0$
ε) $P(X> 6)=1- P(X\le 6)=1- F(6)=1-1=0$
στ) $P(1{,}2 < X < 4)= P(1{,}2 < X \le 4)= F(4)- F(1{,}2)=0{,}8-0{,}24=0{,}56$. ■

4.4 ΜΙΚΤΗ ΤΥΧΑΙΑ ΜΕΤΑΒΛΗΤΗ

Μέχρι τώρα στο παρών Κεφάλαιο επικεντρωθήκαμε σε τυχαίες μεταβλητές οι οποίες είναι εξ ολοκλήρου είτε συνεχείς ή διακριτές. Αυτές οι τυχαίες μεταβλητές είναι οι ποιο σημαντικές που συναντώνται στις περισσότερες εφαρμογές. Παρόλα αυτά όμως συναντούμε μερικές φορές και τυχαίες μεταβλητές ***μικτού τύπου***.

Ορισμός. Μία τυχαία μεταβλητή είναι **μικτού τύπου,** εάν το σύνολο όλων των δυνατών τιμών της τυχαίας μεταβλητής X (δηλαδή το πεδίο τιμών R_X) σε κάποια περιοχή είναι πεπερασμένο $x_1, x_2, x_3, \ldots, x_n$ με πιθανότητα

$$\boldsymbol{P(X=x_i)=p(x_i)\neq 0} \quad (\mathrm{i}=1,2,\ldots,n)$$

και κάπου αλλού επίσης μπορεί να πάρει τιμή μέσα σε ένα διάστημα πραγματικών αριθμών, ας πούμε $\alpha \leq x \leq \beta$ με συνάρτηση πυκνότητας πιθανότητας $\boldsymbol{f(x)}$.

Η συνάρτηση μάζας πιθανότητας για μία τέτοια τυχαία μεταβλητή μπορεί να επιτευχθεί από τον συνδυασμό των συναρτήσεων διακριτής και συνεχούς τυχαίας μεταβλητής.

$$f(x)=\begin{cases} P(X=x_i)=p(x_i) & \gamma\iota\alpha \quad x=x_i \ i=1,2,..n \\ f_1(x) & \gamma\iota\alpha \quad \alpha\leq x\leq\beta \\ 0 & \alpha\lambda\lambda o\upsilon \end{cases} \tag{4.8}$$

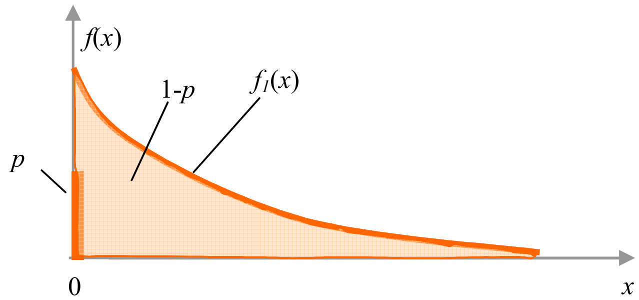
Σε κάθε τιμή x_i αντιστοιχούμε μία πιθανότητα $p(x_i)>0$, ενώ σε κάθε τιμή του διαστήματος $\alpha\leq x\leq\beta$ αντιστοιχεί η συνάρτηση $f_1(x)$, έτσι ώστε

$$\sum_{i=1,\ldots,n} p(x_i)+\int_{\alpha}^{\beta} f_1(x)dx=1 \tag{4.9}$$

Όπως και στους άλλους τύπους τυχαίας μεταβλητής, έτσι και εδώ, η αθροιστική συνάρτηση πιθανότητας δίνεται από την εξίσωση

$$F(x)=P(X\leq x)=\sum_{x_i\leq x} p(x_i)+\int_{-\infty}^{x} f_1(u)du\,. \tag{4.10}$$

Μία τυχαία μεταβλητή μικτού τύπου συναντάται αρκετές φορές όταν η X παριστάνει χρόνο λειτουργίας μιας μηχανής. Σε πολλά προβλήματα ο χρόνος καλής λειτουργίας περιγράφεται από μία τυχαία μεταβλητή X συνεχούς τύπου με πιθανές τιμές $x\geq 0$. Παρόλα αυτά, σε μερικούς τύπους μηχανών, η μηχανή ενδέχεται να πάθει βλάβη την στιγμή της εκκίνησης της λειτουργίας της. Δηλαδή, υπάρχει μία πιθανότητα ότι η μηχανή δεν λειτουργεί καθόλου, αποτυγχάνει στο χρόνο $X=0$, με πιθανότητα $P(X=0)=p>0$. Προφανώς υπάρχει μία πιθανότητα $1-p$ ότι η μηχανή θα λειτουργήσει κάποιο χρονικό διάστημα, $P(X>0)=1-p$. Έτσι, η πιθανότητα p περιγράφει την συνάρτηση κατανομής πιθανότητας της X στο σημείο $x=0$, ενώ η μάζα πιθανότητας $1-p$ κατανέμεται για τις υπόλοιπες τιμές της X, $X>0$, σύμφωνα με μία συνάρτηση πυκνότητας πιθανότητας $f_1(x)$, ενδεχομένως εκθετικής κατανομής όπως φαίνεται και στο Σχήμα 4.12.



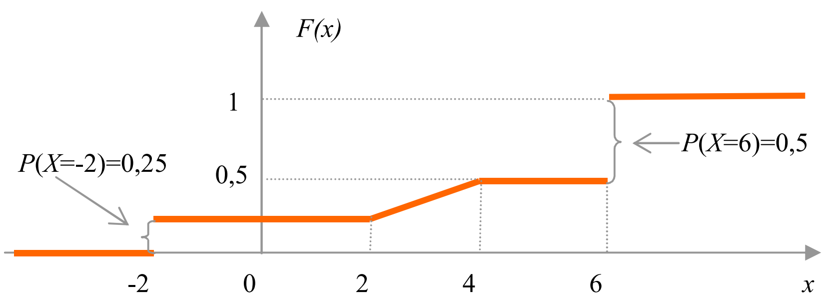
Σχήμα 4.12 Συνάρτηση πυκνότητας πιθανότητας για μια μικτού τύπου τυχαία μεταβλητή, παράδειγμα 4.11.

Παράδειγμα 4.12

Έστω ότι η αθροιστική συνάρτηση πιθανότητας μιας τυχαίας μεταβλητής X δίνεται από την ακόλουθη μορφή,

$$F(x)=\begin{cases}0 & \text{για} \quad x<-2\\ 0{,}25 & \text{για} \quad -2\le x<2\\ 0{,}125x & \text{για} \quad 2\le x<4\\ 0{,}5 & \text{για} \quad 4\le x<6\\ 1 & \text{για} \quad x\ge 6\end{cases}$$

Η γραφική παράσταση της αθροιστικής $F(x)$ δίνεται στο Σχήμα 4.13. Από το σχήμα αναγνωρίζεται ο τύπος της τυχαίας μεταβλητής X. Είναι προφανές ότι η τυχαία μεταβλητή X στα σημεία x=-2 και x=6, όπου η $F(x)$



Σχήμα 4.13 Αθροιστική συνάρτηση πιθανότητας για την μικτή τυχαία μεταβλητή, παράδειγμα 4.12

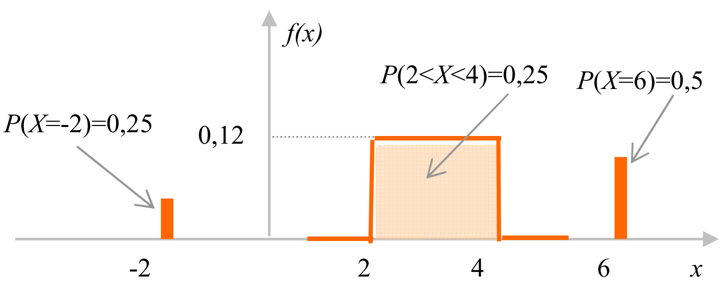
παρουσιάζει ασυνέχεια, είναι διακριτού τύπου με πιθανότητες $P(X=-2)=0,25$ και $P(X=6)=0,5$ αντίστοιχα (ίσες με τα βήματα της ασυνέχειας), στο δε διάστημα $2 \le x < 6$ η αθροιστική είναι συνεχής άρα και η μεταβλητή X στο διάστημα αυτό είναι συνεχής. Η συνάρτηση πυκνότητας πιθανότητας $f_1(x)$ στο διάστημα $2 \le x < 6$ επιτυγχάνεται με παραγώγιση της αθροιστικής $F(x)$. Έτσι, η συνάρτηση κατανομής πιθανότητας $f(x)$ για την μικτή τυχαία μεταβλητή διαμορφώνεται ως εξής

$$f(x) = \begin{cases} 0 & \text{για} \quad x < -2 \\ P(X=-2) = p(-2) = 0,25 & \text{για} \quad x = -2 \\ f_1(x) = \dfrac{dF(x)}{dx} = 0 & \text{για} \quad -2 < x < 2 \\ f_1(x) = \dfrac{dF(x)}{dx} = 0,125 & \text{για} \quad 2 \le x < 4 \\ f_1(x) = \dfrac{dF(x)}{dx} = 0 & \text{για} \quad 4 \le x < 6 \\ P(X=6) = 0,5 & \text{για} \quad x = 6 \\ 0 & \text{αλλου} \end{cases}$$

Αν επιθυμώ να υπολογίσω την πιθανότητα $P(X<3)$, έχω

$$P(X<3) = P(X=-2) + \int_{-2}^{2} 0dx + \int_{2}^{3} 0,125dx = 0,25 + 0 + 0,125 = 0,325$$

Η αντίστοιχη γραφική παράσταση της $f(x)$ φαίνεται στο Σχήμα 4.14



Σχήμα 4.14 Συνάρτηση μάζας πιθανότητας μικτής τυχαίας μεταβλητής του παραδείγματος 4.12

4.5 ΠΕΡΙΛΗΨΗ ΚΕΦΑΛΑΙΟΥ

Η έννοια της τυχαίας μεταβλητής , η οποία αρχικά ξεκίνησε στο πρώτο κεφάλαιο, παρουσιάσθηκε μεθοδικότερα και αναλυτικότερα σε αυτό το κεφάλαιο. Οι διακριτές μεταβλητές σχετίζονται με διαδικασία μέτρησης και οι συνεχείς μεταβλητές χρησιμοποιούνται να μοντελοποιήσουν διάφορες παρατηρήσεις του μηχανικού.

Προσδιορίσθηκε η έννοια της συνάρτησης πιθανότητας για διακριτή και συνεχή τυχαία μεταβλητή. Περιγράφθηκαν οι ιδιότητες των κατανομών πιθανότητας και επισημάνθηκαν οι πρακτικές ερμηνείες στην διερεύνηση των διάφορων τυχαίων μεταβλητών. Αναπτύχθηκαν περισσότερες λεπτομέρειες στον προσδιορισμό ή αναγνώριση του τύπου της τυχαίας μεταβλητής, διακριτής, συνεχούς, και μικτού τύπου, καθώς και στην αντίστοιχη συνάρτηση κατανομής πιθανότητας.

Τα πιο βασικά εργαλεία που απαιτούνται στις μεθόδους πιθανότητας και στατιστικής αναπτύχθηκαν σε αυτό το κεφάλαιο, με αρκετά λεπτομερή μεθοδική εφαρμογή. Τέλος, το κεφάλαιο αυτό ήταν βασικά επακόλουθο του δεύτερου κεφαλαίου, όπου αναπτύχθηκαν οι βασικές έννοιες της πιθανότητας.

ΑΣΚΗΣΕΙΣ-ΠΡΟΒΛΗΜΑΤΑ

Ασκήσεις με τις λύσεις τους

1) Υποθέστε ότι με X συμβολίζουμε την τυχαία μεταβλητή, η οποία παριστάνει το άθροισμα των ενδείξεων δύο ζαριών. Ποιο είναι το πεδίο τιμών της μεταβλητής X, και ποια η συνάρτηση κατανομής πιθανότητας;

Λύση

$P(X=2)=P\{(1,1)\}=1/36$
$P(X=3)=P\{(1,2), (2,1)\}=2/36$
$P(X=4)=P\{(1,3), (3,1), (2,2)\}=3/36$
$P(X=5)=P\{(3,2), (2,3), (4,1), (1,4)\}=4/36$
$P(X=6)=P\{(3,3), (2,4), (4,2), (1,5), (1,5)\}=5/36$
$P(X=7)=P\{(3,4), (4,3), (5,2), (2,5), (1,6), (6,1)\}=6/36$
$P(X=8)=P\{(4,4), (5,3), (5,3), (2,6), (2,6)\}=5/36$
$P(X=9)=P\{(4,5), (5,4), (6,3), (3,6)\}=4/36$
$P(X=10)=P\{(5,5), (6,4), (6,4)\}=3/36$
$P(X=11)=P\{(6,5), (5,6)\}=2/36$
$P(X=12)=P\{(6,6)\}=1/36$
Με άλλα λόγια, η τυχαία μεταβλητή μπορεί να πάρει κάθε ακέραια τιμή από 2 μέχρι 12 με τις παραπάνω πιθανότητες αντίστοιχα.

2) Υποθέστε ότι η τυχαία μεταβλητή X έχει αθροιστική κατανομή πιθανότητας

$$F(x)=\begin{cases}1-\exp(-x^2) & x>0\\ 0 & x\le 0\end{cases}$$

Ποια η πιθανότητα ότι η X ξεπερνά το 1;

Λύση

Η ζητούμενη πιθανότητα υπολογίζεται ως ακολούθως:

$$\begin{aligned}P(X>1)&=1-P(X\le 1)\\ &=1-F(1)=e^{-1}\\ &=0{,}368\end{aligned}$$

3) Υποθέτουμε ότι η τυχαία μεταβλητή X είναι συνεχής με συνάρτηση πυκνότητας πιθανότητας,

$$f(x)=\begin{cases} c(4x-2x^2) & 0<x<2 \\ 0 & \alpha\lambda\lambda o\upsilon \end{cases}$$

(α) Ποια είναι η τιμή της *c;*
(β) Να υπολογιστεί η πιθανότητα $P\{X>1\}$.

Λύση
(α) Αφού *f* είναι μία συνάρτηση πυκνότητας πιθανότητας, πρέπει να ισχύει ότι

$$\int_{-\infty}^{\infty} f(x)dx=1.$$

Καταλήγουμε έτσι στην εξίσωση,

$$c\int_0^2 (4x-2x^2)dx=1$$

ή

$$c\left[2x^2-2x^3/3\right]_{x=0}^{x=2}=1$$

ή

c=3/8.

(β) Οπότε,

$$P(X>1)=\int_1^{\infty} f(x)dx=\frac{3}{8}\int_1^2 (4x-2x^2)dx=\frac{1}{2}$$

4) Κάθε μία από τις ακόλουθες συναρτήσεις παριστάνει την αθροιστική συνάρτηση κατανομής πιθανότητας από μία τυχαία μεταβλητή *X*. Για κάθε μία από τις ακόλουθες περιπτώσεις προσδιορίστε την συνάρτηση πυκνότητας πιθανότητας της *X*, και επιβεβαιώστε ότι είναι πράγματι συνάρτηση πυκνότητας πιθανότητας.

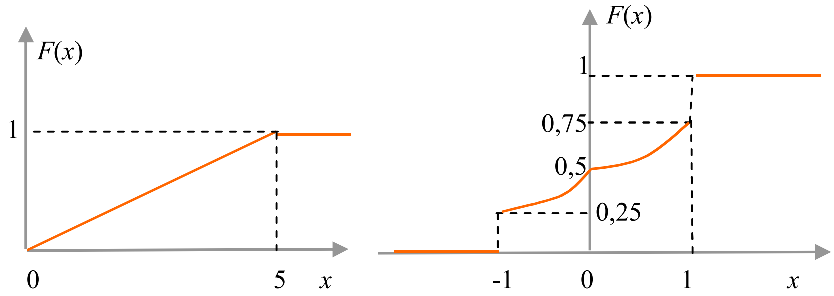
$$(\alpha)\ \ F(x)=\begin{cases} 0 & \gamma\iota\alpha \quad x<0 \\ x/\kappa & \gamma\iota\alpha \quad 0\le x<5 \\ 1 & \gamma\iota\alpha \quad x\ge 5 \end{cases}$$

$$(\beta)\ \ F(x)=\begin{cases} 0 & \gamma\iota\alpha \quad x<-1 \\ x^3/4+0{,}5 & \gamma\iota\alpha \quad -1\le x<1 \\ 1 & \gamma\iota\alpha \quad x\ge 1 \end{cases}$$

Λύση
Είναι πολύ χρήσιμο να σχεδιάσω την γραφική παράσταση της αθροιστικής πιθανότητας, προκειμένου να διευκρινισθεί αν είναι συνεχής ή υπάρχουν σημεία όπου η αθροιστική πραγματοποιεί *άλματα.*

(α) Εδώ η αθροιστική $F(x)$ είναι συνεχής, για να προσδιορίσω την συνάρτηση πυκνότητας πιθανότητας $f(x)$ παίρνω την παράγωγο,

$$f(x) = \frac{d(F(x))}{dx} = \frac{1}{\kappa}.$$



(α) (β)

Προφανώς, πρόκειται για μία σταθερή *σππ* $f(x)$, η σταθερά της κ μπορεί να εκτιμηθεί από την βασική της ιδιότητα, ότι το ολοκλήρωμά της από 0 μέχρι 5 ισούται με 1, ή πιο εύκολα από την αθροιστική όπου για x=5 έχουμε x/κ=1, δηλαδή, κ=5. Έτσι, η αναλυτική μορφή της *σπκ* είναι,

$$f(x) = \begin{cases} \dfrac{1}{5} & \gamma\iota\alpha \quad 0 \le x \le 5 \\ 0 & \alpha\lambda\lambda o\upsilon \end{cases}$$

Η $f(x)$ είναι πράγματι *σ.π.π.*, διότι είναι θετική και το ολοκλήρωμά της ισούται με ένα.

(β) Στην περίπτωση αυτή η αθροιστική στα σημεία x=-1 και x=1 πραγματοποιεί *άλματα*=1/4, άρα υπάρχει συγκεντρωμένη μάζα πιθανότητας στα σημεία αυτά, $P(X=-1)=1/4$ και $P(X=1)=1/4$. Σε οποιοδήποτε άλλο σημείο $-1<x<1$ η αθροιστική είναι συνεχής και η *σ.π.π.* δίνεται από την παράγωγο

$$f(x) = \frac{d(F(x))}{dx} = \frac{3x^2}{4}$$

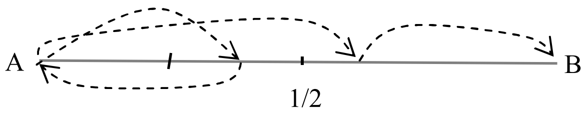
Με άλλα λόγια, η τυχαία μεταβλητή X είναι μικτού τύπου, διακριτή στα σημεία x=-1, x=1, και συνεχής στο διάστημα $-1<x<1$. Τέλος, η αναλυτική μορφή της *σ.π.π.* $f(x)$ είναι,

$$f(x)=\begin{cases} P(X=-1)=0{,}25 \quad \gamma\iota\alpha \quad x=-1 \\ \dfrac{3x^2}{4} \quad \gamma\iota\alpha \quad -1<x<-1 \\ P(X=1)=0{,}25 \quad \gamma\iota\alpha \quad x=1 \\ 0 \quad \alpha\lambda\lambda o\upsilon \end{cases}$$

Η $f(x)$ είναι πράγματι *σππ,* διότι είναι θετική και η συνολική μάζα πιθανότητάς της, η οποία είναι το άθροισμα των δύο πιθανοτήτων στα διακριτά σημεία της μαζί με το ολοκλήρωμά της στο συνεχές διάστημα, δίνει την μονάδα.

5) Δρομέας στοχεύει να διανύσει μια απόσταση ΑΒ μήκους ενός χιλιομέτρου. Σε τυχαίο σημείο μετά τα 250 μέτρα της διαδρομής του αποφασίζει να τερματίσει στο κοντινότερο από τα δύο άκρα Α ή Β. Να βρεθεί η συνάρτηση πυκνότητας της συνολικής διανυόμενης απόστασης.

Λύση



Το τυχαίο σημείο της απόφασης μπορεί να είναι στο πρώτο ή στο δεύτερο ήμισυ της απόστασης ΑΒ με πιθανότητα 1/3 και 2/3 αντίστοιχα (σύμφωνα με την κλασσική μέθοδο).

Συμβολίζω με X την τυχαία μεταβλητή, η οποία παριστάνει την συνολική διανυόμενη απόσταση.

Παρατηρώ ότι αν το τυχαίο σημείο της απόφασης του δρομέα βρίσκεται στο δεύτερο ήμισυ της απόστασης ΑΒ, ο δρομέας διανύει συνολικά ένα χιλιόμετρο, X=1. Στην τιμή αυτή η τυχαία μεταβλητή είναι διακριτή διότι $P(X=1)=2/3$.

Εάν το σημείο απόφασης βρίσκεται στο πρώτο ήμισυ της απόστασης ΑΒ η συνολική διανυόμενη απόσταση X είναι μικρότερη από ένα χιλιόμετρο. Στην περίπτωση αυτή η X παίρνει άπειρες τιμές στο διάστημα [0,5, 1), είναι συνεχής με μία σταθερή συνάρτηση πυκνότητας πιθανότητας.

Από το παραπάνω σκεπτικό προκύπτει η ακόλουθη μικτή συνάρτηση μάζας πιθανότητας,

$$f(x)=\begin{cases} f_1(x)=P(X=1)=\frac{2}{3} \quad \gamma\iota\alpha \quad x=1 \\ f_2(x)=c \quad \gamma\iota\alpha \quad 0{,}5\le x<1 \\ 0 \quad \alpha\lambda\lambda o\upsilon \end{cases}$$

Η σταθερά c μπορεί να εκτιμηθεί από την βασική ιδιότητα της συνάρτησης μάζας,

$$P(X=1)+\int_{0,5}^{1} c\,dx = 1$$

$$\frac{2}{3}+c[x]_{0,5}^{1}=\frac{2}{3}+0,5c=1 \quad \Rightarrow \quad c=\frac{2}{3}$$

Προβλήματα με τις λύσεις του

1) Αντοχή σχοινιού. Η αντοχή X (N/mm^2) ενός τύπου σχοινιού δεν είναι σταθερή, είναι τυχαία μεταβλητή. Από μία μεγάλη δειγματοληψία όπου δοκιμάσθηκε η δύναμη αντοχής του σχοινιού καταλήξαμε με την βοήθεια της περιγραφικής στατιστικής ότι η μεταβλητή X έχει συνάρτηση πυκνότητας πιθανότητας $f_X(x)$,

$$f_X(x)=\begin{cases} \dfrac{x}{1400}, & \text{για} \quad 0 \le x \le 40 \\ \dfrac{70-x}{1050}, & \text{για} \quad 40 < x \le 70 \\ 0 & \text{για} \quad 70 < x \quad \text{και} \quad x < 0 \end{cases}$$

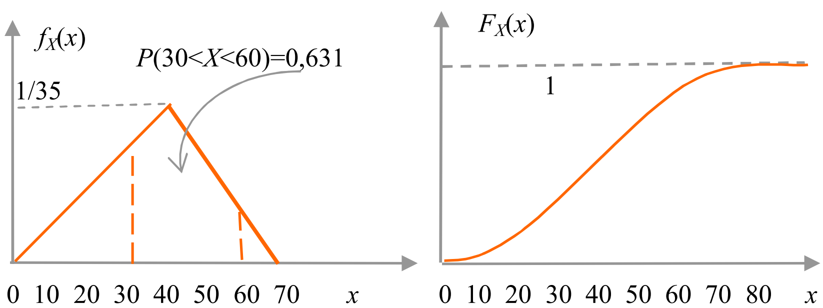
(α) Να γίνει η γραφική παράσταση της συνάρτησης πυκνότητας $f_X(x)$.
(β) Να προσδιορίσετε την αθροιστική κατανομή πιθανότητας $F_x(x)$, και σχεδιάσετε την γραφική της παράσταση.
(γ) Δοκιμάζουμε το σχοινί σε τυχαίο σημείο του. Να υπολογισθεί η πιθανότητα ότι η αντοχή βρίσκεται μεταξύ 30 και 60 (N/mm^2).

Λύση
(α) Η συνάρτηση πυκνότητας πιθανότητας $f_X(x)$ είναι τριγωνικής μορφής, όπως φαίνεται και στην (α) γραφική παράσταση.
(β) Η αθροιστική συνάρτηση κατανομής έχει ως εξής,

$$F_X(x) = \begin{cases} 0 & \text{για} \quad x < 0 \\ \int_0^x \frac{u}{1400}\,du = \frac{x^2}{2800}, & \text{για} \quad 0 \le x \le 40 \\ \frac{40^2}{2800} + \int_{40}^{x} \frac{70-u}{1050}\,du = \frac{4}{7} + \frac{70(x-40)}{1050} - \frac{x^2-40^2}{2100} & \text{για} \quad 40 < x \le 70 \\ 1 & \text{για} \quad 70 < x \end{cases}$$

και οι γραφικές παραστάσεις φαίνονται στο ακόλουθο σχήμα.

(γ) Η ζητούμενη πιθανότητα μπορεί να ευρεθεί αν ολοκληρώσουμε την *σππ* $f_X(x)$ από 30 μέχρι 60, όπως φαίνεται και στο σχήμα (α). Επίσης, ευκολότερα υπολογίζεται άμεσα και από γνωστή ιδιότητα της αθροιστικής πιθανότητας,



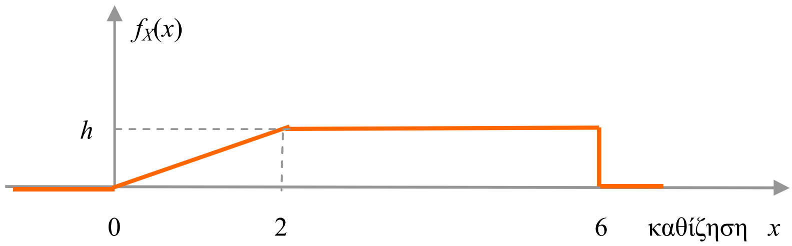
(α) (β)

$$P(30 < X \le 60) = F_X(60) - F_X(30)$$

$$= \left[\frac{4}{7} + \frac{70(60-40)}{1050} - \frac{(60-40)^2}{2100}\right] - \left[\frac{30^2}{2800}\right]$$

$$= 0{,}952 - 0{,}321 = 0{,}631.$$

2) Καθίζηση. Η καθίζηση σε μία κατασκευή είναι τυχαία μεταβλητή, Χ, με συνάρτηση πυκνότητας πιθανότητας x όπως φαίνεται στο ακόλουθο σχήμα,



(α) Να ευρεθεί η $f_X(x)$ (η αναλυτική της μορφή).
(β) Να ευρεθεί η πιθανότητα ότι η καθίζηση θα ξεπεράσει τους 2 πόντους.
(γ) Αν η καθίζηση είναι πάνω από 2 πόντους, ποια η πιθανότητα να ξεπεράσει τους πέντε πόντους;

Λύση

(α) Η συνάρτηση πυκνότητας πιθανότητας από το 0 μέχρι το 2 προφανώς είναι της μορφής $\alpha x+\beta$, ενώ, από 2 μέχρι 6 είναι μία σταθερά h. Έχουμε τρεις παραμέτρους α, β, h, άγνωστες τις οποίες μπορούμε να εκτιμήσουμε από τις προκύπτουσες εξισώσεις,

$$\alpha \cdot 0+\beta=0$$
$$\alpha \cdot 2+\beta=h$$
$$\int_0^2 (\alpha x+\beta)dx+\int_2^6 hdx=1$$

Λύνοντας τις δύο πρώτες εξισώσεις προκύπτει $\beta=0$, $\alpha=h/2$. Αντικαθιστώντας στην τρίτη εξίσωση έχουμε,

$$\int_0^2 (\alpha x+\beta)dx+\int_2^6 hdx=\int_0^2 \frac{h}{2}xdx+\int_2^6 hdx=5h=1$$
$$\Rightarrow \quad h=1/5$$

Έτσι, η αναλυτική μορφή της συνάρτησης $f_X(x)$ είναι

$$f_X(x)=\begin{cases} \frac{1}{10}x & \gamma\iota\alpha \quad 0<x\le 2 \\ \frac{1}{5} & \gamma\iota\alpha \quad 2<x\le 6 \\ 0 & \alpha\lambda\lambda o\upsilon \end{cases}$$

(β) Η πιθανότητα υπολογίζεται εύκολα, αφού γνωρίζουμε την συνάρτηση πυκνότητας πιθανότητας.

$$P(X\ge 2)=1-P(X<2)=1-\int_0^2 \frac{1}{10}xdx=1-\frac{1}{5}=\frac{4}{5}$$

(γ) Εδώ έχουμε υπό συνθήκη πιθανότητα

$$P(X \geq 5 | X > 2) = \frac{P[(X \geq 5) \cap (X > 2)]}{P(X > 2)} = \frac{P(X \geq 5)}{P(X > 2)}$$

αλλά

$$P(X \geq 5) = \int_5^6 \frac{1}{5} dx = \frac{1}{5}$$

και το αποτέλεσμα είναι

$$P(X \geq 5 | X > 2) = \frac{P(X \geq 5)}{P(X > 2)} = \frac{1/5}{4/5} = \frac{1}{4}$$

3) Αντοχή εδάφους. Η αντοχή του εδάφους κάτω από το πέδιλο μιας κολώνας είναι τυχαία μεταβλητή, X, και παίρνει τιμές από 8 μέχρι 16 *kips/ft*^{2}, με την ακόλουθη συνάρτηση πυκνότητας πιθανότητας,

$$f_X(x) = \begin{cases} \kappa(1 - x/16) & \gamma\iota\alpha \quad 8 < x \leq 16 \\ 0 \quad \alpha\lambda\lambda o\upsilon & \end{cases}$$

Ο μηχανικός έχει σχεδιάσει έτσι την κολώνα ώστε να δέχεται φορτίο 10 *kips/ft*^{2}. Ποια η πιθανότητα αστοχίας της κολώνας;

Λύση

Η *σππ* $f_X(x)$ είναι πάντα θετική και περιέχει μία παράμετρο κ, την οποία μπορούμε να υπολογίσουμε από την βασική ιδιότητά της,

$$\int_8^{16} f_X(x) dx = 1$$

$$\int_8^{16} \kappa(1 - x/16) dx = 1$$

$$\kappa\left[[x]_8^{16} - \left[\frac{x^2}{32}\right]_8^{16} \right] = \kappa[8 - (8 - 2)] = 2\kappa = 1$$

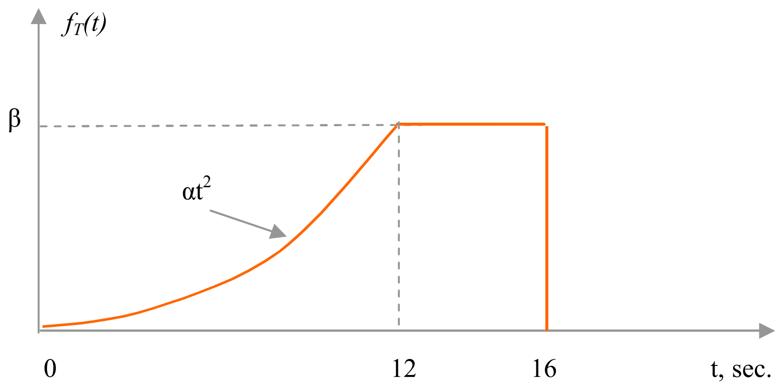
$$\Rightarrow \kappa = 1/2.$$

Αστοχία της κολώνας συμβαίνει όταν η αντοχή εδάφους *X* είναι μικρότερη από το βάρος της κολώνας 10 *kips/ft*^{2}. Η πιθανότητα υπολογίζεται *αθροίζοντας* την μάζα πιθανότητας μέχρι το 10.

$$P(X<10)=\int_8^{10} f_X(x)dx=\int_8^{10} 1/2(1-x/16)dx$$

$$1/2\left[[x]_8^{10}-\left[\frac{x^2}{32}\right]_8^{10}\right]=1/2(2-36/32)=7/16.$$

4) Η διάρκεια επίδρασης μιας δύναμης πάνω σε ένα σημείο μιας μηχανής είναι τυχαία μεταβλητή είναι τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας όπως φαίνεται στο σχήμα.



(α) Να βρεθούν οι τιμές των α και β για την συνάρτηση πυκνότητας πιθανότητας.
(β) Να υπολογιστεί η πιθανότητα P(X>6).

Λύση

(α) δημιουργώ δύο εξισώσεις ως προς α και β εφαρμόζοντας την βασική ιδιότητας της σππ και εξισώνοντας τις δύο μορφές στο σημείο t=12.

$$\int_0^{12} \alpha t^2 dt+\int_{12}^{16}\beta dt=1$$

$$\alpha 12^2=\beta$$

Λύνω το σύστημα ως προς α και β,

$$\alpha\frac{t^3}{3}\Big|_0^{12}+\beta t\Big|_{12}^{16}=a\frac{12^3}{3}+4\beta=1$$

$$\alpha=\frac{\beta}{144}$$

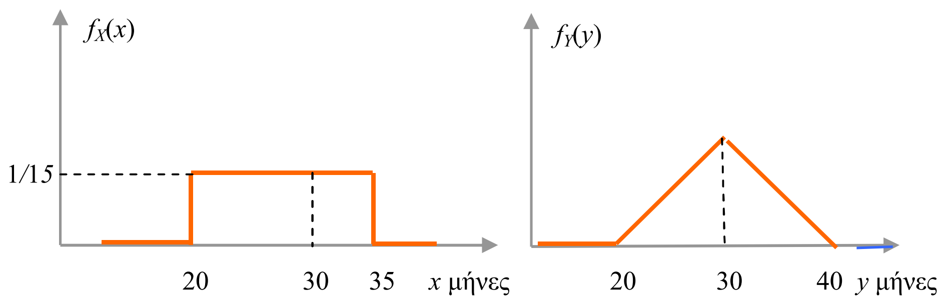
ή τελικά α=1/1152 και β=0,125

(β) Ένας τρόπος είναι ο εξής

$$\text{P(X>6)=1-P(X<6)=}1-\int_0^6 \alpha t^2 dt=1-\alpha\left[\frac{t^3}{3}\right]_0^6=1-\frac{3}{32}=\frac{29}{32}$$

4) Κατασκευή αεροδρομίου. Μία μεγάλη εταιρία αναλαμβάνει την κατασκευή ενός αεροδρομίου, με την υπόσχεση να ολοκληρωθεί μέσα σε 30 μήνες. Το τεχνικό έργο αποτελείται από την κατασκευή του διαδρόμου και του κτιριακού συγκροτήματος. Οι εργασίες μπορούν να αρχίσουν παράλληλα, και εκτιμάται ότι η διάρκεια, X, κατασκευής του διαδρόμου κυμαίνεται από 20 μέχρι 35 μήνες και η διάρκεια κατασκευής, Y, των κτιρίων από 20 μέχρι 40 μήνες με αντίστοιχες συναρτήσεις πυκνότητας όπως φαίνεται στα Σχήματα (α) και (β).

Ποια η πιθανότητα ότι το αεροδρόμιο θα παραδοθεί εγκαίρως;



Σχήμα (α) Σχήμα (β)

Λύση

Η συνάρτηση πυκνότητας πιθανότητας $f_X(x)$ είναι σταθερή με τιμή c=1/15, όπως προκύπτει από την βασική ιδιότητα της. Δηλαδή, το ολοκλήρωμα της σταθεράς από το 20 μέχρι 35 ισούται με 1.

Η συνάρτηση πυκνότητας πιθανότητας $f_Y(y)$ είναι τριγωνικής μορφής. Δεν είναι απαραίτητο να προσδιορίσουμε την αναλυτική της μορφή, διότι το ολοκλήρωμα της που θέλουμε να εκτιμήσουμε στη συνέχεια είναι προφανές από το σχήμα.

Συμβολίζοντας με

A={*η κατασκευή των κτιρίων ολοκληρώνεται μέσα σε* 30 *μήνες*},
B={*η κατασκευή του διαδρόμου ολοκληρώνεται μέσα σε* 30 *μήνες*},
το γεγονός της έγκαιρης ολοκλήρωσης του αεροδρομίου (*E*) μπορώ να το συμβολίσω με την τομή των *A* και *B*, $E=A\cap B$.

Κατά συνέπεια, η πιθανότητα πραγματοποίησης του *E* είναι

$$P(E)=P(A\cap B)=P(A)P(B)$$

αφού μπορώ να υποθέσω ότι τα γεγονότα *A, B* είναι στατιστικά ανεξάρτητα. Όμως

$$P(A)=\int_{20}^{30} f_X(x)dx=\int_{20}^{30}(1/15)dx=2/3$$

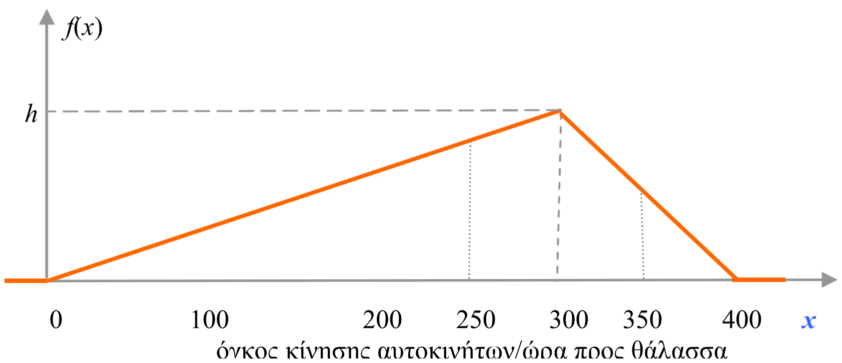
$$P(B)=\int_{20}^{30} f_Y(y)dy=1/2$$

Άρα,

$$P(E)=(2/3)\cdot(1/2)=1/3\,.$$

5) Κυκλοφορία αιχμής. Το καλοκαίρι**,** σε ένα μεγάλο αυτοκινητόδρομο ο όγκος κυκλοφορίας αυτοκινήτων/ώρα προς την θάλασσα είναι τυχαία μεταβλητή , X, με συνάρτηση πυκνότητας πιθανότητας όπως φαίνεται στο Σχήμα. Ο αυτοκινητόδρομος έχει τρεις λωρίδες ανά κατεύθυνση. Όταν ο όγκος κυκλοφορίας προς την θάλασσα ξεπεράσει τα 300 αυτοκίνητα/ώρα τότε δίνεται ακόμη μία λωρίδα για να διευκολύνει τα αυτοκίνητα προς την κατεύθυνση αυτή.

Είναι γνωστό ότι με τρεις λωρίδες προς μία κατεύθυνση εξυπηρετούνται χωρίς συνωστισμό μέχρι 250 αυτοκίνητα. Αν χρησιμοποιηθεί και τέταρτη λωρίδα μποτιλιάρισμα συμβαίνει μόνο στην περίπτωση που η κυκλοφορία ξεπεράσει τα 350 αυτοκίνητα/ώρα προς την κατεύθυνση αυτή.



Μία μέρα το καλοκαίρι,

(α) ποια η πιθανότητα να χρησιμοποιηθεί και τέταρτη λωρίδα προς την κατεύθυνση της θάλασσας;
(β) ποια η πιθανότητα μποτιλιαρίσματος της κίνησης των αυτοκινήτων προς την θάλασσα;

Λύση

(α) Προφανώς, έχω να υπολογίσω την πιθανότητα $P(X>300)$. Η αναλυτική μορφή της σππ $f(x)$ είναι

$$f(x)=\begin{cases} 0 \quad \gamma\iota\alpha \quad x<0 \\ \alpha x \quad \gamma\iota\alpha \quad 0\le x<300 \\ \beta_1\chi+\beta_2 \quad \gamma\iota\alpha \quad 300\le x<400 \\ 0 \quad \gamma\iota\alpha \quad x>400 \end{cases}$$

Το ύψος h της τριγωνικής *σ.π.π* μπορεί να υπολογισθεί αν χρησιμοποιήσει κανείς την βασική ιδιότητά της,

$$\int_0^{400} f(x)dx=1$$

$$400h=2\Rightarrow \quad h=2/400=0{,}005$$

Μπορώ να εκτιμήσω τις παραμέτρους α, β_1, β_2 λύνοντας το σύστημα τριών εξισώσεων,

$\alpha\cdot 300=0{,}005$
$\beta_1\cdot 300+\beta_2=0{,}005$
$\beta_1\cdot 400+\beta_2=0{,}0$

Τέλος, οι λύσεις του συστήματος είναι, $\alpha=1/60.000$, $\beta_2=0{,}02$, και $\beta_1=-0{,}00005$. Άρα,

$$P(X>300)=1-P(X\le 300)=1-\int_0^{300}(1/60000)\cdot x\cdot dx$$

$$=1-(1/60000)\left[\frac{x^2}{2}\right]_0^{300}=1-0{,}75=0{,}25$$

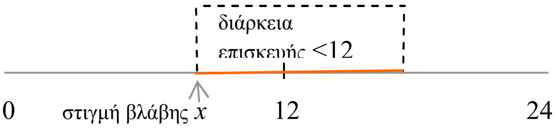
(β) Εδώ, συμβολίζω με A_1, A_2 τα γεγονότα ότι ο όγκος κίνησης είναι $250<X\le 300$ ή $300<X\le 350$ αντίστοιχα. Ως εκ τούτου το γεγονός να υπάρξει συνωστισμός στην κίνηση (γεγονός E) το συμβολίζω με την ένωση των A_1, A_2, $E=A_1\cup A_2$. Έτσι, η ζητούμενη πιθανότητα είναι

$$P(E)=P(A_1\cup A_2)=P(A_1)+P(A_2)$$
$$=\int_{250}^{300}(1/120000)x dx+\int_{350}^{400}(-0{,}000025x+0{,}01)dx$$
$$=0{,}3646+0{,}03125=0{,}39585$$

6) Λειτουργία συσκευής. Μία ηλεκτρική συσκευή αστοχεί σε τυχαία χρονική στιγμή κατά την διάρκεια ενός εικοσιτετραώρου. Επισκευάζεται και τίθεται πάλι σε λειτουργία. Η διάρκεια επισκευής της είναι τυχαία μεταβλητή ομοιόμορφα κατανεμημένη στο διάστημα (x, 24), όπου x η χρονική στιγμή βλάβης.

Ποια η πιθανότητα ότι η συσκευή λειτουργεί συνολικά τουλάχιστον 12 ώρες το εικοσιτετράωρο;(δυνατότητες πέραν της μία βλάβης κατά την διάρκεια του εικοσιτετραώρου αποκλείονται).

Λύση



Διακρίνω εξαρχής δύο γεγονότα.

A_1={ *Η συσκευή να αστοχήσει πριν την παρέλευση των 12 ωρών*}

A_2={ *Η συσκευή να αστοχήσει μετά την παρέλευση των 12 ωρών*}

Αν συμβολίσω με Ε το γεγονός του οποίου ζητώ την πιθανότητα, τότε αφού τα γεγονότα A_1, A_2 αποτελούν διαμέριση του δειγματικού χώρου κάνω εφαρμογή του νόμου ολικής πιθανότητας

$$P(E)=P(E|A_1)P(A_1)+P(E|A_2)P(A_2)$$

Προφανώς, αφού η συσκευή αστοχεί ισοπίθανα μέσα στο εικοσιτρέαωρο έχουμε

$$P(A_1)=0{,}5 \quad \kappa\alpha\iota \quad P(A_2)=0{,}5$$

Ακόμη στην δεύτερη περίπτωση η συσκευή λειτουργεί σίγουρα πάνω από 12 ώρες, $P(E|A_2)=1$. Για να υπολογίσω την πιθανότητα $P(E|A_1)$ σκέφτομαι ως εξής:

Έστω ότι η συσκευή αστοχεί την χρονική στιγμή x πριν τις 12 ώρες. Για να συμπληρώσει συνολικό χρόνο λειτουργίας πάνω από 12 ώρες πρέπει η διάρκεια επισκευής της Y να είναι μικρότερη του 12, όπως φαίνεται και στο σχήμα. Η πιθανότητα του ενδεχομένου η συσκευή να αστοχήσει την στιγμή x και επισκευασθεί σε λιγότερο από 12 ώρες είναι

$$P(\{x \le X \le x+dx\} \cap \{Y \le 12-x\}) = P(x \le X \le x+dx)P(Y \le 12-x)$$
$$= f_X(x)dx \cdot \frac{12}{24-x}$$

Τέλος για να υπολογίσω την πιθανότητα $P(E|A_1)$ πρέπει να ενώσω το παραπάνω ενδεχόμενο άπειρες φορές, αφού υπάρχουν άπειρες στιγμές όπου μπορεί να αστοχήσει η συσκευή πριν τις 12 ώρες. Άρα

$$P(E|A_1) = \int_0^{12} f_X(x) \cdot \frac{12}{24-x} dx = \int_0^{12} \frac{1}{12} \cdot \frac{12}{24-x} dx = \int_0^{12} \frac{1}{24-x} dx = ...$$

Επομένως καταλήγουμε στην ζητούμενη πιθανότητα

$$P(E) = P(E|A_1)0{,}5 + 1 \cdot 0{,}5$$

Ασκήσεις προς λύση

1) Κάλπη περιέχει πέντε σφαίρες αριθμημένες με 1 έως 5. Στα πλαίσια τυχαίου πειράματος επιλέγονται τρεις σφαίρες και έστω X ο ελάχιστος παρατηρούμενος αριθμός. Ποια είναι η συνάρτηση πυκνότητας $f(x)$ και κατανομής πιθανότητας $F(x)$ της τυχαίας μεταβλητής X.

2) Τρεις ολόιδοι βώλοι, αριθμούμενοι 1, 2, και 3 τοποθετούνται μέσα σε ένα σάκο. Εάν δύο βώλοι επιλέγονται τυχαία με επανατοποθέτηση, ποια είναι η συνάρτηση πιθανότητας της τυχαίας μεταβλητής X, όπου X είναι το άθροισμα των αριθμών πάνω στους βώλους;

3) Από μία σωρό δώδεκα ηλεκτρονικών εξαρτημάτων όπου τα πέντε είναι ελαττωματικά, επιλέγεται ένα τυχαίο δείγμα με τρία εξαρτήματα χωρίς αντικατάσταση. Αποδείξτε ότι η συνάρτηση και η αθροιστική συνάρτηση κατανομής πιθανότητας των ελαττωματικών εξαρτημάτων στο δείγμα είναι όπως ακολουθεί:

ελαττωματικά x_i	**0**	**1**	**2**	**3**
$f_X(x_i)$	7/44	21/44	14/44	2/44
$F_X(x_i)$	7/44	28/44	42/44	44/44

4) Να αποδείξετε ότι η ακόλουθη συνάρτηση είναι πράγματι συνάρτηση μάζας πιθανότητας,

$$f_X(x)=(3/4)(1/4)^x \text{ για } x=0, 1, 2, 3, \ldots$$

και να προσδιορίσετε τις πιθανότητες, $P(X=2)$, $P(X\leq 2)$, $P(X>2)$, $P(X\geq 1)$.

5) Υποθέτουμε ότι η τυχαία μεταβλητή X έχει την ακόλουθη αθροιστική συνάρτηση πιθανότητας

$$F(x)=\begin{cases} 0 & \gamma\iota\alpha \quad x<-2 \\ 0{,}2 & \gamma\iota\alpha \quad -2\leq x<0 \\ 0{,}7 & \gamma\iota\alpha \quad 0\leq x<2 \\ 1 & \gamma\iota\alpha \quad 2\leq x \end{cases}$$

Να προσδιορίσετε την συνάρτηση μάζας πιθανότητας $f(x)$.

6) Ένα κιβώτιο με δέκα σιδηροδοκούς, εννέα καλούς και μία ραγισμένη, δοκιμάζονται με τυχαία σειρά μία προς μία. Εάν X είναι ο αριθμός της δοκιμής όπου εντοπίζεται η ραγισμένη σιδηροδοκός, ποιες είναι οι πιθανές τιμές της X, και ποιες οι αντίστοιχες πιθανότητες;

7) Η διάμετρος X ενός καλωδίου είναι συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας

$$f(x)=6x(1-x),\ 0<x<1.$$

Να ελεγχθεί αν η παραπάνω συνάρτηση είναι όντως συνάρτηση πυκνότητας πιθανότητας και να βρεθεί η αθροιστική συνάρτηση πιθανότητας $F(x)$.

8) Η τυχαία μεταβλητή X είναι ομοιόμορφα κατανεμημένη στο διάστημα (α,β).
(α)Υπολογίστε τις τιμές α και β αν γνωρίζετε ότι η μέση τιμή της X είναι μηδέν και ότι $P(X<1)=2/3$.
(β)Υπολογίστε τις πιθανότητες

$$P(X=0{,}5),\ P(X<0{,}5),\ P(X<4),\ P(-0{,}5<X<0{,}5).$$

9) Δύο φίλοι δίνουν ένα ραντεβού σε ένα καφενείο μεταξύ 4 και 5 το απόγευμα. Όποιος πάει πρώτος στο καφενείο περιμένει 10 λεπτά και αν στο διάστημα αυτό δεν έρθει ο φίλος του φεύγει από το καφενείο. Ποια η πιθανότητα να συναντηθούν οι δύο φίλοι;

10) Σε ένα κύκλο ακτίνας $\lambda N^{1/2}$ κατανέμουμε τυχαία N σημεία. Για σταθερό λ και μεγάλο N, η πιθανότητα ότι κανένα σημείο δεν πέφτει σε δοθείσα περιοχή εμβαδού A είναι προσεγγιστικά

$\exp(-A/\pi\lambda^2)$.

Να αποδειχθεί ότι η απόσταση του κάθε σημείου από το πιο γειτονικό του είναι τυχαία μεταβλητή X με συνάρτηση πυκνότητας πιθανότητας

$$f_X(x) = 2x\lambda^{-2} \exp(\frac{-x^2}{\lambda^2})$$

Προβλήματα προς λύση

1) Προδιαγραφές μεταλλικής ράβδου. Η συνάρτηση πυκνότητας πιθανότητας του μήκους μεταλλικής ράβδου είναι $f(x)$=2 για 2,3<x<2,8 μέτρα.
(α) Εάν οι προδιαγραφές για αυτήν την παραγωγή είναι από 2,25 μέχρι 2,75 μέτρα, ποιο ποσοστό μεταλλικών ράβδων δεν πληροί τις προδιαγραφές;
(β) Πάνω σε ποια τιμή θα έπρεπε να επικεντρωθεί η συνάρτηση πυκνότητας πιθανότητας $f(x)$ για να έχουμε το μέγιστο ποσοστό ράβδων εντός προδιαγραφών;

2) Λειτουργία ηλεκτρονικού εξαρτήματος. Η συνάρτηση πυκνότητας πιθανότητας του χρόνου λειτουργίας ενός τύπου ηλεκτρονικού εξαρτήματος (σε ώρες) είναι $f(x) = (1/1000)e^{-x/1000}$ για x>0. Προσδιορίστε την πιθανότητα ότι
(α) ένα εξάρτημα λειτουργεί περισσότερο από 3000 ώρες,
(β) ένα εξάρτημα αποτυγχάνει πριν τις 1000 ώρες,
(γ) ένα εξάρτημα αποτυγχάνει στο διάστημα από 1000 μέχρι 2000 ώρες.
Προσδιορίστε τον αριθμό ωρών λειτουργίας μέχρι τον οποίο το 10% των εξαρτημάτων θα έχουν αποτύχει.

3) Χημική αντίδραση. Ο χρόνος X (σε milliseconds) μέχρι να ολοκληρωθεί μία χημική αντίδραση είναι τυχαία μεταβλητή με αθροιστική συνάρτηση κατανομής

$$F(x) = \begin{cases} 0 \quad \gamma\iota\alpha \quad x < 0 \\ 1 - e^{-0,01x} \quad \gamma\iota\alpha \quad x \geq 0 \end{cases}$$

(α) Να προσδιορίσετε την συνάρτηση πυκνότητας πιθανότητας της X.
(β) Ποιο ποσοστό χημικών αντιδράσεων ολοκληρώνετε μέσα σε 200 milliseconds΄

4) Ξύλινο panelling. Το πάχος ξύλινου panelling (*σε πόντους*), το οποίο παραγγέλνει πελάτης σε μία βιομηχανία, είναι τυχαία μεταβλητή με αθροιστική συνάρτηση πιθανότητας

$$F_X(x) = \begin{cases} 0 \quad για \quad x < 1/2 \\ 0{,}2 \quad για \quad 1/2 \le x < 1 \\ 0{,}9 \quad για \quad 1 \le x < 2 \\ 1 \quad για \quad x \ge 2 \end{cases}$$

(α) Να προσδιορισθούν οι ακόλουθες πιθανότητες,
$P(X \le 1/2)$, $P(X \le 1)$, $P(X \le 1{,}8)$, $P(X > 1)$, $P(X \le 2{,}5)$
(β) Να προσδιορισθεί η συνάρτηση πυκνότητας πιθανότητας $f_X(x)$.

5) Φορτίο ανέμου. Ένας πύργος εκτίθεται στην οριζόντια δύναμη των δυνατών ανέμων. Ένας σημαντικός παράγοντας ο οποίος θα έπρεπε να ληφθεί υπ' όψη στην αντοχή του πύργου είναι η διάρκεια των ανέμων. Η διάρκεια T τέτοιων δυνατών ανέμων που επικρατούν στην περιοχή είναι τυχαία μεταβλητή με μέγιστη τιμή 18 ώρες. Από παρατηρήσεις των δεδομένων των ανέμων, η *σ.π.π.* της T μπορεί να προσεγγισθεί με την συνάρτηση $f_T(t) = ct^{1.5}$.
(α) Να προσδιορίσετε την παράμετρο c στην *σ.π.π.* της T.
(β) Ποια η πιθανότητα ότι ένας άνεμος διαρκεί πάνω από 9 ώρες;

6) Προδιαγραφές εξαρτημάτων. Ένα σύστημα αποτελείται από τρία εξαρτήματα. Έστω ότι οι πιθανότητες ότι το πρώτο, το δεύτερο, και το τρίτο εξάρτημα πληρούν τις προδιαγραφές είναι αντίστοιχα 0,95, 0,98, και 0,99. Υποθέτουμε ότι τα εξαρτήματα είναι ανεξάρτητα.

Προσδιορίστε την συνάρτηση μάζας πιθανότητας του αριθμού των εξαρτημάτων στο σύστημα τα οποία πληρούν τις προδιαγραφές. Επίσης, να κάνετε την γραφική παράσταση αυτής της συνάρτησης.

7) Διαγωνισμός. Ένας εργολάβος υποβάλλει προσφορές για τρία έργα Α, Β και Γ. οι πιθανότητες επιτυχίας του στους διαγωνισμούς είναι $P(A)$=0,5, $P(B)$=0,8, $P(\Gamma)$=0,2, αντίστοιχα. Υποθέστε ότι τα ενδεχόμενα επιτυχίας στα έργα Α, Β, και Γ είναι ανεξάρτητα μεταξύ τους. Έστω X ο αριθμός επιτυχιών του εργολάβου.
(α) Ποιες είναι οι δυνατές τιμές του X; Να υπολογιστεί και να απεικονιστεί γραφικά η συνάρτηση μάζας πιθανότητας της τ.μ. X.
(β) Να απεικονιστεί γραφικά η αθροιστική συνάρτηση κατανομής της X.
(γ) Να υπολογιστούν οι πιθανότητες $P(X<2)$, $P(0<X<2)$.

8) Ένταση ρεύματος. Η ένταση X του ρεύματος που διέρχεται από μία αντίσταση R σ' ένα κύκλωμα, είναι τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας

$$f(x) = \begin{cases} c & \text{για } x \in [-\alpha, -\alpha/2] \\ c & \text{για } x \in [\alpha/2, \alpha] \\ 0 & \text{αλλού} \end{cases}$$

και ισχύει $P(X \le 3) = 0,75$

(α) Να βρεθούν οι πιθανότητες $P(X \ge 2)$, $P(X > 1,5 / X > 0)$, $P(X = 2)$.

(β) Να βρεθούν η αθροιστική συνάρτηση πιθανότητας.

9) Φανάρι διασταύρωσης. Το φανάρι μιας διασταύρωσης δείχνει διαδοχικά ένα λεπτό πράσινο και μισό λεπτό κόκκινο. Ένα αυτοκίνητο πλησιάζει τυχαία την διασταύρωση. Να υπολογιστούν:

(α) Η πιθανότητα το αυτοκίνητο να περάσει την διασταύρωση με πράσινο.

(β) Η συνάρτηση πυκνότητας $f(x)$ και η αθροιστική συνάρτηση κατανομής $F(x)$ του χρόνου αναμονής X του αυτοκινήτου στην διασταύρωση.

10) Κατανάλωση πετρελαίου . Η μηνιαία κατανάλωση πετρελαίου για θέρμανση μιας πολυκατοικίας σε χιλιάδες γαλόνια είναι τυχαία μεταβλητή με πυκνότητα

$$f(x) = \begin{cases} \alpha(1-x)^4 \gamma\iota\alpha 0 \le x \le 1 \\ 0 \alpha\lambda\lambda o\upsilon \end{cases}$$

Ποια χωρητικότητα πρέ πει να έχει ο λέβητας, ώστε η πιθανότητα να εξαντληθεί το πετρέλαιο σ' ένα μήνα να είναι p=0,01.

11) Δομική μηχανή. Μία ηλεκτρική μηχανή αστοχεί με πιθανότητα 0,20 κατά την έναρξη της λειτουργίας της. Εάν η μηχανή δεν πάθει βλάβη στην αρχή, ο χρόνος καλής λειτουργίας της, X, είναι τυχαία μεταβλητή και ακολουθεί την εξής συνάρτηση πυκνότητας πιθανότητας,

$$f(x) = \begin{cases} 1,6 - \kappa x \quad \gamma\iota\alpha \quad 0 < x < 1 \, \varepsilon\tau o\varsigma \\ 0 \quad \gamma\iota\alpha \quad x \ge 1 \, \varepsilon\tau o\varsigma \end{cases}$$

(α) Να βρεθεί η συνάρτηση κατανομής πιθανότητας του χρόνου καλής λειτουργίας της μηχανής.

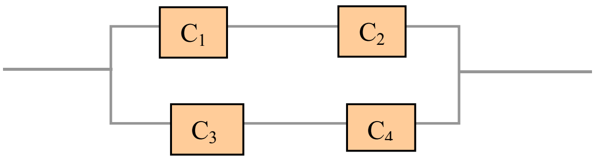
(β) Ποια είναι η πιθανότητα ότι η μηχανή θα λειτουργήσει σωστά για έξι μήνες;

12) Στάση λεωφορείου. Ένα λεωφορείο μεταφέρει n επιβάτες. Στην επόμενη στάση του, ο κάθε επιβάτης κατεβαίνει με πιθανότητα p ανεξάρτητα από τους άλλους. Ο αριθμός k των επιβατών που ανεβαίνουν στη στάση αυτή ακολουθεί την κατανομή:

$$p(k=i)=\begin{cases} 1/6 & \gamma\iota\alpha \quad i=0 \\ \dfrac{1}{i+1} & \gamma\iota\alpha \quad i=1,2 \\ 0 & \alpha\lambda\lambda o\upsilon \end{cases}$$

Ποια η πιθανότητα μετά την αναχώρηση του λεωφορείου από την στάση ο αριθμός των μεταφερομένων επιβατών να είναι πάλι n;

13) Συνάρτηση πυκνότητας πιθανότητας συστήματος. Στο παρακάτω σύστημα ο χρόνος καλής λειτουργίας, T, των διακοπτών C_1, C_2, C_3, C_4 είναι τυχαία μεταβλητή με *σ.π.π.* εκθετικής μορφής, $f(x)=(1/50)e^{-(1/50)x}$. Οι διακόπτες C_1 και C_2 αποτυγχάνουν ταυτόχρονα(όταν αποτυγχάνει ο C_1 αμέσως αποτυγχάνει ο C_2 και αντίστροφα) ενώ στους υπόλοιπους συνδυασμούς οι διακόπτες είναι στατιστικά ανεξάρτητοι μεταξύ τους.
(α) Ποια η πιθανότητα ότι η λειτουργία του συστήματος δεν έχει σταματήσει μετά τις 80 ώρες;
(β) Να ευρεθεί η *σ.π.π.* $f_T(t)$ του χρόνου λειτουργίας του συστήματος.



5

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΥΧΑΙΩΝ ΜΕΤΑΒΛΗΤΩΝ

ΠΕΡΙΓΡΑΦΗ ΚΕΦΑΛΑΙΟΥ

Στο προηγούμενο κεφάλαιο, διαπιστώσαμε την χρησιμότητα της κατανομής πιθανότητας $f(x)$ για μία τυχαία μεταβλητή X. Με βάση την κατανομή πιθανότητας, ο Μηχανικός εκτιμά την πιθανότητα ότι κάποιο μέγεθος θα ξεπεράσει το όριο ασφαλείας ή θα βρεθεί μέσα σε ένα επιθυμητό διάστημα κ.τ.λ., πληροφορίες αρκετά χρήσιμες στη μελέτη του. Ατυχώς όμως, μερικές φορές η κατανομή πιθανότητας $f(x)$ δεν είναι γνωστή, για αυτό και αναζητούνται άλλα **χρήσιμα χαρακτηριστικά** των τιμών που παίρνει η τυχαία μεταβλητή. Για μία τυχαία μεταβλητή, εκτός της συνάρτησης κατανομής πιθανότητας, οι παράμετροί της ***μέση τιμή, διακύμανση, διάμεσος, πιθανότερη τιμή*** και άλλες περιγράφουν περιληπτικά την μεταβλητή.

Αλλά, και αν ακόμη είναι γνωστή η συνάρτηση κατανομής $f(x)$, όπως είδαμε για να εκτιμηθούν οι διάφορες πιθανότητες της X πρέπει να χρησιμοποιηθούν και κάποιες παράμετροι που χαρακτηρίζουν την κατανομή πιθανότητας. Με κάθε συνάρτηση πιθανότητας σχετίζεται τουλάχιστον μία παράμετρος, η οποία προσφέρει πολύτιμες πληροφορίες για την κατανομή π.χ. την κλίση της γραμμής, την περιοχή του πεδίου τιμών της X και άλλες. Στο παράδειγμα 4.8, όπου η τυχαία μεταβλητή X παριστάνει την διάρκεια λειτουργίας μίας μηχανής, η συνάρτηση κατανομής πιθανότητας είναι εκθετικής μορφής $f(x) = Ae^{-Ax}$ και είναι απαραίτητη η εκτίμηση της παραμέτρου A. Η τιμή της παραμέτρου A μπορεί να διαφέρει για τους διάφορους τύπους μηχανών και όπως θα διαπιστώσουμε παρακάτω η παράμετρος αυτή είναι χαρακτηριστική για την τυχαία μεταβλητή X, γιατί παριστάνει την μέση τιμή του χρόνου λειτουργίας της μηχανής.

5.1. ΜΕΣΗ ΤΙΜΗ

Μία από τις πιο σπουδαίες έννοιες στην στατιστική είναι η μέση τιμή μιας τυχαίας μεταβλητής. Όπως είναι γνωστό στην περιγραφική Στατιστική, σε ένα δείγμα τιμών $x_1, x_2, \ldots, x_n$ η δειγματική μέση τιμή

$$\bar{x} = \frac{1}{n}x_1 + \frac{1}{n}x_2 + \frac{1}{n}x_3 + \ldots + \frac{1}{n}x_n$$

δίνει μία περιληπτική πληροφορία για τις τιμές του δείγματος. Παρόμοια, και για μία τυχαία μεταβλητή X ένα πιθανοκρατικό μοντέλο, το οποίο αθροίζει τις πιθανές τιμές της μεταβλητής με βάρος όσο η αντίστοιχη πιθανότητα της συνάρτησης κατανομής, προσδιορίζει την *μέση τιμή ή την αναμενόμενη τιμή* της X.

Παράδειγμα 5.1

Το **μήκος** ενός είδους σύρματος δεν είναι σταθερό λόγω ατέλειας του μηχανισμού κοπής. Για αυτό, το μήκος $\boldsymbol{X}$ του σύρματος είναι τυχαία μεταβλητή και μπορεί να θεωρηθεί ότι η συνάρτηση κατανομής πιθανότητας $f(x)$ είναι ομοιόμορφη στο διάστημα [24,5, 25,5]. Επειδή οι προδιαγραφές ορίζουν μήκος 25,

- αν $24{,}7 < X < 25{,}3$ το κέρδος ανά σύρμα είναι 250,
- αν $X>25{,}3$ το σύρμα ξανά-κόβεται και το κέρδος περιορίζεται σε 150
- αν το σύρμα είναι πολύ κοντό, $X < 24{,}7$, το σύρμα καταστρέφεται με ζημία 100.

Αφού η συνάρτηση κατανομής πιθανότητας είναι ομοιόμορφη, εύκολα υπολογίζονται οι πιθανότητες

$P(X < 24{,}7)=0{,}2$,

$P(24{,}7 < X < 25{,}3)=0{,}6$,

$P(X > 25{,}8)=0{,}2$.

Το **κέρδος** του κάθε σύρματος ***Z*** είναι διακριτή τυχαία μεταβλητή γιατί προφανώς μπορεί να πάρει τιμές –100, 250, 150 με αντίστοιχες πιθανότητες $p(-100)=0{,}2$, $p(150)=0{,}6$, και $p(250)=0{,}2$.

Εάν από τα N σύρματα που παράγει η εταιρεία, N_1 έχουν μήκος $x < 24{,}7$, N_2 έχουν μήκος $24{,}7 < x < 25{,}8$ και N_3 έχουν μήκος $x > 25{,}8$ το ολικό κέρδος λογικά είναι

$$\text{Ολικό κέρδος } T = N_1\cdot(-100) + N_2\cdot 250 + N_3\cdot 150 \ \text{ δραχμές.}$$

Ακόμη, το μέσο κέρδος ανά σύρμα $E(Z)$ μπορεί να εκτιμηθεί διαιρώντας με N,

$$E(Z) = T/N = (N_1/N)\cdot(-100) + (N_2/N)\cdot 250 + (N_3/N)\cdot 150 \ \text{ δραχμές.}$$

Αν υποθέσει κανείς ότι το N είναι αρκετά μεγάλο($N\rightarrow\infty$), σύμφωνα με την θεωρία της σχετικής συχνότητας οι λόγοι (N_1/N), (N_2/N), (N_3/N) προσεγγίζουν τις αντίστοιχες πιθανότητες

$$(N_1/N) \sim P(Z=-100)=0{,}2,\ \ (N_2/N) \sim P(Z=250)=0{,}6, \text{ και } (N_3/N) \sim P(Z=150)=0{,}2.$$

Έτσι το μέσο ή αναμενόμενο κέρδος ανά σύρμα μπορεί να γραφεί σαν το άθροισμα των πιθανών τιμών της Z με βάρη $P(Z=-100)$, $P(Z=250)$, $P(Z=150)$,

$$\begin{aligned} E(Z) &= P(Z=-100)\cdot(-100) + P(Z=250)\cdot 250 + P(Z=150)\cdot 150 \\ &= 0{,}2\cdot(-100) + 0{,}6\cdot 250 + 0{,}2\cdot 150 = 160\,. \end{aligned}$$

Τελικά, σε αυτό το παράδειγμα καταλήγουμε ότι το κέρδος του κάθε σύρματος Z είναι τυχαία μεταβλητή και το αναμενόμενο μέσο κέρδος ανά σύρμα είναι η ***μέση τιμή*** *της μεταβλητής Z,* η οποία εκτιμήθηκε σε 160, αθροίζοντας τις πιθανές τιμές της Z με βάρη τις αντίστοιχες πιθανότητες. ■

Ορισμός: Αν X είναι **διακριτή** τυχαία μεταβλητή με πεδίο τιμών $x_1, x_2, x_3, \ldots, x_{n+}, \ldots$ και συνάρτηση κατανομής πιθανότητας $f(x_i)=p(x_i)$ (για $\iota=1, 2, 3, \ldots, n, \ldots$), ορίζουμε **μέση τιμή** ή **αναμενόμενη τιμή της *X* ή μαθηματική ελπίδα της *X*** και δηλώνουμε ***μ*** ή ***E*(*X*),** την τιμή

$$\mu = E(X) = \sum_{i=1}^{\infty} x_i\, p(x_i) \qquad (5.1)$$

Αν η τυχαία μεταβλητή X είναι **συνεχής** με συνάρτηση πυκνότητας πιθανότητας $f(x)$, μπορούμε να ορίσουμε παρόμοια την αναμενόμενη τιμή ή μαθηματική ελπίδα ***E*(*X*).** Η πιθανότητα η μεταβλητή X να βρεθεί πολύ κοντά σε μία τιμή x είναι

P(x < X < x+dx) ~ f(x)dx

Όπου dx είναι μία πολύ μικρή τιμή. Αλλά η X μπορεί να βρεθεί κοντά σε άπειρες τέτοιες τιμές x με αντίστοιχη πιθανότητα $f(x)dx$. Έτσι, η μέση τιμή της X, σύμφωνα με το παραπάνω σκεπτικό της διακριτής μεταβλητής, θα προσέγγιζε το άθροισμα ενός πολύ μεγάλου αριθμού γινομένων $x\cdot f(x)dx$. Βέβαια, αυτό το άθροισμα είναι το ολοκλήρωμα του γινομένου $x\cdot f(x)$ ως προς x

Ορισμός. Αν η τυχαία μεταβλητή X είναι **συνεχής** με συνάρτηση πυκνότητας πιθανότητας $f(x)$, ορίζουμε παρόμοια την **αναμενόμενη τιμή ή μαθηματική ελπίδα $E(X)$,**

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx\,. \qquad (5.2)$$

Επίσης για μία **μικτή** τυχαία μεταβλητή X με διακριτές τιμές $x_1, x_2, x_3, \ldots, x_n$ και συνεχής σε κάποιο διάστημα των πραγματικών αριθμών ορίζουμε παρόμοια την αναμενόμενη τιμή ή μαθηματική ελπίδα **$E(X)$,**

$$\mu = E(X) = \sum_{i}^{n} x_i p(x_i) + \int_{-\infty}^{\infty} xf(x)dx \qquad (5.3)$$

Παρατηρήσεις

(A) Για να υπάρχει μέση τιμή $E(X)$ της διακριτής τυχαίας X πρέπει το άθροισμα της σειράς

$\sum_{i=1}^{\infty} x_i p(x_i)$ να συγκλίνει απόλυτα, δηλαδή

$\sum_{i=1}^{\infty} |x_i| p(x_i) < \infty$. Διαφορετικά η μέση τιμή είναι απροσδιόριστη. Κάτι ανάλογο ισχύει για την συνεχή τυχαία μεταβλητή.

(B) Εάν μία διακριτή τυχαία μεταβλητή παίρνει πεπερασμένο αριθμό τιμών $x_1, x_2, x_3, \ldots, x_n$, η μέση της τιμή

$$E(X) = \sum_{i=1}^{n} x_i p(x_i)$$

μπορεί να θεωρηθεί και σαν *ζυγιζόμενο* άθροισμα όλων των πιθανών τιμών $x_1, x_2, x_3, \ldots, x_n$ με βάρη τις αντίστοιχες πιθανότητες $p_1, p_2, \ldots, p_n$.

Εάν η μεταβλητή X παίρνει ισοπίθανα όλες αυτές τις πιθανές τιμές, Τότε

$$E(X) = \frac{1}{n}\sum_{i=1}^{n} x_i\,,$$

δηλαδή, η μέση τιμή $E(X)$ παριστάνει τον γνωστό αριθμητικό μέσο όρο των n πιθανών τιμών.

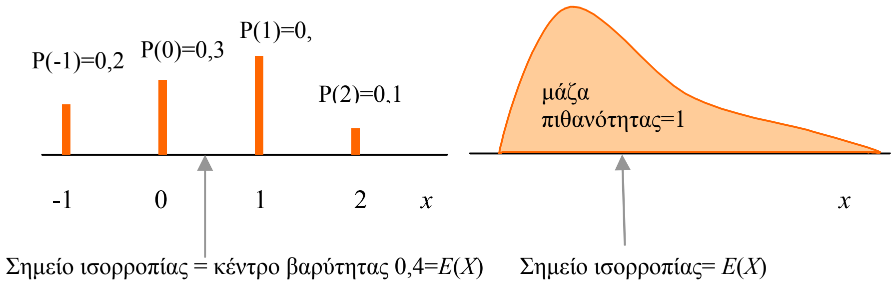
(Γ) Η μέση τιμή $E(X)$ ενδέχεται να μην συμπίπτει με καμία από τις πιθανές τιμές της X. Για παράδειγμα, η μέση τιμή της τυχαίας μεταβλητής που παριστάνει την ένδειξη στην ρίψη ενός ζαριού είναι

$$E(X)=1\cdot\frac{1}{6}+2\cdot\frac{1}{6}+3\cdot\frac{1}{6}+4\cdot\frac{1}{6}+5\cdot\frac{1}{6}+6\cdot\frac{1}{6}=3{,}5.$$

Ο αριθμός 3,5 δεν ανήκει στο πεδίο τιμών της X. Αυτό όμως δεν σημαίνει ότι η τιμή αυτή δεν έχει νόημα. Εάν κάποιος π.χ. ρίξει πολλές φορές(n φορές)το ζάρι, το άθροισμα όλων των ενδείξεων αναμένεται να είναι κοντά στον αριθμό $n\cdot 3{,}5$. ■

Η έννοια της μέσης τιμής μ= $E(X)$ είναι ανάλογος με την έννοια του κέντρου βαρύτητας σε μία κατανομή μάζας στην Φυσική. Θεωρείστε μία διακριτή τυχαία μεταβλητή X η οποία έχει συνάρτηση μάζας πιθανότητας $p(x_i)$, για $i=1,2,\ldots,n$. Εάν σκεφθούμε τώρα μά βέργα χωρίς βάρος πάνω στην οποία τοποθετούμε βάρη $p_1, p_2,\ldots, p_n$ στα αντίστοιχα σημεία $x_1, x_2, x_3,\ldots, x_n$ (βλέπε Σχήμα 5.1), τότε το σημείο στο οποίο ισορροπεί η βέργα είναι το κέντρο βαρύτητας της. Με απλή γνώση στατικής είναι απλό να αποδείξει κανείς ότι αυτό το σημείο είναι η μέση τιμή $E(X)$.

Στη συνεχή μεταβλητή, μέση τιμή είναι επίσης το σημείο όπου ισορροπεί ο άξονας X εάν επάνω του συσσωρεύσουμε όλη την μάζα πιθανότητας όπως υποδεικνύει η συνάρτηση πυκνότητας πιθανότητας. (βλέπε Σχήμα 5.1)



Σχήμα 5.1 Η μέση τιμή συμπίπτει με το κέντρο βαρύτητας της μάζας πιθανότητας

Παράδειγμα 5.2
Υποθέτουμε ότι η τυχαία μεταβλητή X έχει ομοιόμορφη συνάρτηση κατανομής πιθανότητας στο διάστημα $[a,\beta]$. Έχουμε υπολογίσει προηγούμενα ότι η συνάρτηση κατανομής πιθανότητας είναι

$$f(x)=\begin{cases}\dfrac{1}{\beta-a} & \gamma\iota\alpha \quad a\le x\le\beta\\ 0 & \alpha\lambda\lambda o\upsilon\end{cases}$$

Οπότε,

$$E(X)=\int_{-\infty}^{\infty}x\frac{1}{\beta-a}dx=\int_{a}^{b}x\frac{1}{\beta-a}dx=\frac{1}{\beta-a}\left[\frac{x^2}{2}\right]_a^b=\frac{1}{\beta-a}\frac{\beta^2-a^2}{2}=\frac{\beta+a}{2}$$

ύστερα από εύκολες πράξεις.. ■

Ιδιότητες Μέσης Τιμής

Έστω ότι X είναι μία τυχαία μεταβλητή με συνάρτηση κατανομής πιθανότητας $f(x)$. Έστω επίσης, ότι δεν ενδιαφερόμαστε μόνο για την μέση τιμή της X αλλά και για την μέση τιμή μιας συνάρτησης της X, ας πούμε $Y=g(X)$. Πώς μπορούμε να προχωρήσουμε εδώ; Ένας τρόπος είναι ο εξής. Αφού η Y είναι τυχαία μεταβλητή(ως συνάρτηση τυχαίας μεταβλητής), πρέπει να έχει συνάρτηση κατανομής πιθανότητας $f_Y(y)$, η οποία μπορεί να προσδιορισθεί από την συνάρτηση κατανομής της X, $f(x)$. Από την στιγμή που θα προσδιορίσουμε την συνάρτηση κατανομής πιθανότητας της Y, $f_Y(y)$, μπορούμε να υπολογίσουμε την $E(Y)$ από τον ορισμό της μέσης τιμής,

$$E(Y)=\int_{-\infty}^{\infty}yf_Y(y)dy\,.$$

Όμως, μέχρι τώρα, δεν έχουμε αναφερθεί στον προσδιορισμό της συνάρτησης κατανομής πιθανότητας $f_Y(y)$ της μεταβλητής $Y=g(X)$ όταν είναι γνωστή η $f(x)$. Αυτό μελετάται σε άλλο κεφάλαιο των πιθανοτήτων. Για αυτό προτιμούμε να υπολογίσουμε την μέση τιμή της μεταβλητής $g(x\ )$ *με* έναν ευκολότερο τρόπο.

Αφού η τυχαία μεταβλητή $Y=g(X)$ παίρνει την τιμή $g(x)$ όταν η τυχαία μεταβλητή X παίρνει την τιμή x, $X=x$, φαίνεται λογικό ότι $E(g(X))$ θα έπρεπε είναι το ζυγιζόμενο άθροισμα όλων των πιθανών τιμών $g(x)$ με βάρος την αντίστοιχη πιθανότητα η X να είναι ίση με x.

Α) Εάν X είναι διακριτή τυχαία μεταβλητή με συνάρτηση κατανομής πιθανότητας $f(x)=p(x)$, τότε για κάθε πραγματική συνάρτηση $g(x)$ έχουμε

$$E(Y)=E(g(X))=\sum_x g(x)p(x) \ .$$

Επίσης, εάν η X είναι συνεχής ισχύει παρόμοια

$$E(Y)=E(g(X))=\int_{-\infty}^{\infty} g(x)f(x)dx \, .$$

Με αυτή την ιδιότητα ο υπολογισμός της $E(Y)$ γίνεται απλούστερος, διότι δεν χρειάζεται η εύρεση της συνάρτησης κατανομής πιθανότητας $f_Y(y)$. Η γνώση της κατανομής πιθανότητας της X, $f(x)$, είναι επαρκής.

Β) Εάν α και β είναι κάποιες σταθερές, τότε

E(αX+β)=αE(X)+β

Απόδειξη: Στη συνεχή περίπτωση,

$$E(\alpha X+\beta)=\int_{-\infty}^{\infty}(\alpha x+\beta)f(x)dx=\int_{-\infty}^{\infty}(\alpha x)f(x)dx+\int_{-\infty}^{\infty}\beta f(x)dx=\alpha\int_{-\infty}^{\infty}xf(x)dx$$

$$+\beta\int_{-\infty}^{\infty}f(x)dx=\alpha E(X)+\beta\cdot 1=\alpha E(X)+\beta$$

Παρόμοια αποδεικνύεται και για την διακριτή περίπτωση.

Γ) Είναι προφανής η ιδιότητα $E(\beta)=\beta$.

Δ) Η μέση τιμή του αθροίσματος τυχαίων μεταβλητών ισούται με το άθροισμα των μέσων τιμών τους,

$$E(X_1+X_2+X_3+...+X_n)=E(X_1)+E(X_2)+E(X_3)+...+E(X_n)$$

Η απόδειξη της τελευταίας σχέσης στηρίζεται στην πρώτη ιδιότητα, αλλά απαιτείται και η γνώση της πολυδιάστατης τυχαίας μεταβλητής, για αυτό δεν αναφέρεται εδώ. ■

Παράδειγμα 5.3

Υποθέτουμε ότι η τυχαία μεταβλητή X έχει συνάρτηση κατανομής πιθανότητας

$$P(X=0)=p(0)=0{,}2 \quad P(X=1)=p(1)=0{,}5 \quad P(X=2)=p(2)=0{,}3$$

Να υπολογισθεί η $E(X^2)$.

Μπορούμε να θέσουμε $g(X)=X^2$, οπότε από την ιδιότητα (Α) προκύπτει

$$E(X^2)=0^2\cdot 0{,}2+1^2\cdot 0{,}5+2^2\cdot 0{,}3=1{,}7.$$ ■

Παράδειγμα 5.4

Ας πούμε ότι η ταχύτητα του ανέμου V(χιλμ.) είναι ομοιόμορφα κατανεμημένη στο διάστημα (0, 20). Η πίεση Z(lb/ft) που ασκείται από τον άνεμο στην επιφάνεια μιας κεραίας δίνεται από την σχέση Z=0.005V^2. Για να υπολογίσουμε την αναμενόμενη πίεση στην κεραία χρησιμοποιούμε την ιδιότητα (Α). Όμως, πρώτα προσδιορίζουμε την ακριβή μορφή της κατανομής $f(v)$. Έτσι έχουμε

$$f(v)=\begin{cases}\dfrac{1}{20} & \text{για} \quad 0<v<20\\ 0 & \text{αλλου}\end{cases}$$

$$E(Z)=E(g(V))=\int_0^{20}0{,}005v^2\frac{1}{20}dv=0{,}00025\left[\frac{v^3}{3}\right]_0^{20}=0{,}667\ lb/ft^2$$ ■

Παράδειγμα 5.5

Ο χρόνος X, σε ώρες, που παίρνει για να επισκευασθεί μία χαλασμένη μηχανή είναι τυχαία μεταβλητή, της οποίας η συνάρτηση πυκνότητας πιθανότητας είναι

$$f(x)=\begin{cases}1 & \text{για} \quad 0<x<1\\ 0 & \text{αλλου}\end{cases}$$

Εάν το κόστος επισκευής περιλαμβάνει ένα πάγιο έξοδο 100.000 και ο χρόνος επισκευής x χρεώνεται με $300.000\cdot x^3$, ποιος είναι ο μέσος χρόνος επισκευής και ποιο το αναμενόμενο κόστος για μία χαλασμένη μηχανή;

Εφαρμόζοντας την ιδιότητα (Β) έχουμε

$$\text{Μέσος χρόνος επισκευής } E(X)=\int_0^1 x\cdot 1\cdot dx=\left[\frac{x^2}{2}\right]_0^1=\frac{1}{2}$$

Το κόστος επισκευής είναι τυχαία μεταβλητή διότι είναι συνάρτηση του χρόνου επισκευής. Μπορούμε να γράψουμε

$$Y = X^3 300.000 + 100.000\,,$$

και εφαρμόζοντας τις παραπάνω ιδιότητες προκύπτει

$$E(Y) = 300.000E(X^3) + 100.000 = 300.000\int_0^1 x^3 \cdot 1 \cdot dx + 100.000 =$$

$$= 300.000\left[\frac{x^4}{4}\right]_0^1 + 100.000 = 300.000 \cdot \frac{1}{4} + 100.000 = 175.000 \quad ■$$

5.2 ΔΙΑΚΥΜΑΝΣΗ

Όταν δίνεται μία τυχαία μεταβλητή X με την συνάρτηση κατανομής πιθανότητας, θα ήταν εξαιρετικά χρήσιμο αν μπορούμε με κάποιες κατάλληλες μετρήσεις να παρουσιάζουμε τις βασικές ιδιότητες της μάζας πιθανότητας. Μία τέτοια μέτρηση είναι η μέση τιμή $E(X)$. Παρόλα που η $E(X)$ προσδιορίζει το άθροισμα των επιβαρούμενων δυνατών τιμών της X, δεν μας πληροφορεί σχετικά με το πόσο απέχουν μεταξύ τους αυτές οι τιμές.

Αν σε μία τυχαία μεταβλητή X βρούμε ότι η μέση τιμή $E(X)$ είναι 1000, ποια η πληροφορία αυτής της τιμής; Μία χρήσιμη πληροφορία είναι ότι αν εκτελέσουμε n φορές το πείραμα τύχης, η X παίρνει τις τιμές $x_1, x_2, x_3, \ldots, x_n$, όπου ο μέσος όρος αυτών προσεγγίζει το 1000 για πολύ μεγάλο n. Παρόλα αυτά, η μέση τιμή δεν περιέχει καμία πληροφόρηση για την διασπορά των τιμών. Για παράδειγμα, ας υποθέσουμε ότι η μεταβλητή X παριστάνει την διάρκεια ζωής δύο τύπων τρανζίστορα που προμηθεύεται ένας κατασκευαστής. Στον πρώτο τύπο για το μεγαλύτερο ποσοστό, η διάρκεια κυμαίνεται από 900 έως 1100 ώρες η δε μέση διάρκεια είναι 1000 ώρες. Στον δεύτερο τύπο η μέση διάρκεια είναι επίσης 1000 ώρες, αλλά το 50 % τρανζιστόρων είναι κακής ποιότητας με διάρκεια κοντά στις 700 ώρες, ενώ το άλλο 50% είναι καλής ποιότητας με διάρκεια γύρω στις 13000 ώρες. Είναι προφανές εδώ ότι για την σύγκριση της ποιότητας των δύο τύπων η γνώση της μέσης τιμής της διάρκειας ζωής του τρανζίστορα δεν είναι αρκετή.

Υπάρχει προφανώς ανάγκη να εισάγουμε μία ακόμη ποσοτική μέτρηση η οποία να μας πληροφορεί πόσο απομακρύνονται οι πιθανές τιμές της X από την μέση τιμή της ή αν υπάρχει μεγάλη ανομοιογένεια στις τιμές αυτές κ.τ.λ. Έχουν προταθεί πολλές τέτοιες μετρήσεις. Ένας καλός τρόπος για να μετρήσουμε την διασπορά των τιμών της X είναι να παρατηρήσουμε κατά μέσο όρο πόσο απέχουν οι τιμές της από την μέση τιμή. Αυτό μπορεί να επιτευχθεί αν θεωρήσουμε την ποσότητα της μέσης απόλυτης απόκλισης

$$E(|X-\mu|).$$

Αυτή η μέτρηση είναι χρήσιμη αλλά από μαθηματικής πλευράς όχι η καταλληλότερη. Η πιο συνηθισμένη στην Στατιστική και χρήσιμη μέτρηση της διασποράς των τιμών της X είναι η μέση τιμή του τετραγώνου των αποκλίσεων (X-μ). Έτσι, καταλήγουμε στον ακόλουθο ορισμό.

Ορισμός. **Εάν X είναι τυχαία μεταβλητή με μέση τιμή μ, τότε η** διακύμανση **της X, η οποία συμβολίζεται με $Var(X)$ ή σ_X^2, ορίζεται ως ακολούθως:**

$$Var(X)=E[(X-\mu)^2]. \tag{5.4}$$

Η τετραγωνική ρίζα της διακύμανσης $Var(X)$, $\sigma_X=\sqrt{Var(X)}$, ονομάζεται **τυπική απόκλιση** της X.

Σε μία τυχαία μεταβλητή X με συνάρτηση πυκνότητας πιθανότητας $f(x)$ και μέση τιμή μ, για υπολογίσουμε την διακύμανση, θεωρούμε $g(X)=(X-\mu)^2$ και εφαρμόζουμε την (Α) ιδιότητα της μέσης τιμής. Δηλαδή, πρέπει να ολοκληρώσουμε ώς ακολούθως

$$Var(X)=E[(X-\mu)^2]=\int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx \tag{5.5}$$

Ευκολότερα, μία εναλλακτική φόρμουλα της διακύμανσης $Var(X)$ μπορεί να προκύψει ως ακολούθως:

$$\begin{aligned} Var(X) &= E[(X-\mu)^2] \\ &= E(X^2-2X\mu+\mu^2) \\ &= E(X^2)-2\mu E(X)+E(\mu^2) \\ &= E(X^2)-2\mu^2+\mu^2 \\ &= E(X^2)-\mu^2. \end{aligned}$$

Δηλαδή,

$$Var(X) = E(X^2) - [E(X)]^2 \qquad (5.6)$$

ή με άλλα λόγια, η διακύμανση της X είναι η μέση τιμή του τετραγώνου της X μείον το τετράγωνο της μέσης τιμής της. Πρακτικά, αυτός είναι και ο ποιο εύκολος τρόπος να υπολογίσουμε την $Var(X)$.

Παρατηρήσεις

(A) Η τιμή της διακύμανσης $Var(X)$ εκφράζεται σε *τετραγωνικές μονάδες,* π.χ. κιλά^2, ώρες^2. Αυτός είναι ο λόγος που στην πράξη χρησιμοποιούμε την τυπική απόκλιση σ_X, η οποία εκφράζεται με τις ίδιες μονάδες όπως η X. Η διακύμανση χρησιμοποιείται περισσότερο στην θεωρητική μελέτη.

(B) Η διακύμανση όπως ορίσθηκε παραπάνω είναι μία ειδική περίπτωση της *κ ροπής* της τυχαίας μεταβλητής X ως προς την μέση τιμή της που ορίζεται από τον τύπο $\mu_\kappa = E[(X-\mu)^\kappa]$. Προφανώς για κ=2 προκύπτει η διακύμανση. ■

Ιδιότητες Διακύμανσης

Υπάρχουν αρκετά σπουδαίες ιδιότητες της διακύμανσης. Παρόμοια με τις ιδιότητες της μέσης τιμής αναφέρουμε τις ακόλουθες.

A) Εάν η β είναι μία σταθερά τότε

$$\begin{aligned} Var(X+\beta) &= E[(X+\beta) - E(X+\beta)]^2 \\ &= E[X + \beta - E(X) - \beta]^2 \\ &= E[X - E(X)]^2 \\ &= Var(X)\,. \end{aligned}$$

Η ερμηνεία της ιδιότητας αυτής είναι ότι η μεταβλητότητα των τιμών της X, δηλαδή αυτό που μετρά η διακύμανση, δεν αλλάζει αν όλες οι τιμές *μετατοπισθούν* κατά μία σταθερά ποσότητα β.

B) Εάν α είναι σταθερά τότε

$$\begin{aligned} Var(aX) &= E[(aX - E(aX)]^2 = E[aX - aE(X)]^2 \\ &= E[a(X - E(X))]^2 = E[a^2(X - E(X))^2]) \\ &= a^2 E[X - E(X)]^2 = a^2 Var(X). \end{aligned}$$

Γ) Είναι σημαντικό να αναφέρουμε ότι γενικά η διακύμανση δεν είναι *προσθετική*, όπως η μέση τιμή. Μόνο εάν X_1, X_2, X_3, ..., X_n είναι τυχαίες ανεξάρτητες μεταβλητές, τότε

$$Var(X_1 + X_2 + X_3 + ... + X_n) = Var(X_1) + Var(X_2) + Var(X_3) + ... + Var(X_n)$$

Η απόδειξη παρατίθεται σε άλλο κεφάλαιο της πολυδιάστατης τυχαίας μεταβλητής. ■

Πρέπει να αναφέρουμε ότι η διακύμανση δεν κατέχει την γραμμική ιδιότητα όπως αναφέραμε στην μέση τιμή, ιδιότητα (Β) παρ. 5.1. Δηλαδή,

$$Var(aX + \beta) \neq aVar(X) + \beta \ .$$

Αντί αυτού έχουμε

$$Var(aX + \beta) = a^2 Var(X) \ .$$

Άμεσα, από την σχέση αυτή προκύπτει και η ιδιότητα $Var(\beta)=0$.

Παράδειγμα 5.6

Υποθέτουμε ότι η τυχαία μεταβλητή X έχει ομοιόμορφη συνάρτηση κατανομής πιθανότητας στο διάστημα $[a,\beta]$. Έχουμε υπολογίσει προηγούμενα ότι

$$f(x) = \frac{1}{\beta - a} \quad \kappa\alpha\iota \quad E(X) = \frac{\beta + a}{2} \quad .$$

Για να υπολογίσουμε την $Var(X)$ πρώτα βρίσκουμε

$$E(X^2) = \int_a^b x^2 \frac{1}{\beta - a} dx = \frac{\beta^3 - a^3}{3(\beta - a)} \, .$$

Οπότε,

$$Var(X) = E(X^2) - [E(X)]^2 = \frac{\beta^3 - a^3}{3(\beta - a)} - \frac{(\beta + a)^2}{4} = \frac{(\beta - a)^2}{12}$$

ύστερα από απλές πράξεις. ■

Παρατηρούμε ότι η διακύμανση εξαρτάται από το τετράγωνο της διαφοράς $(b - a)^2$, και όχι από τις τιμές a και β μόνες τους. Έτσι, για δύο τυχαίες μεταβλητές όπου η κάθε μία έχει ομοιόμορφη κατανομή σε διαφορετικό διάστημα, αν το μήκος των διαστημάτων είναι ίδιο, ίδιες θα είναι και οι διακυμάνσεις των μεταβλητών.

Παράδειγμα 5.7

Ο χρόνος λειτουργίας X ενός μετασχηματιστή είναι τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας εκθετικής μορφής,

$$f_X(x) = \lambda e^{-\lambda x}.$$

Να ευρεθεί η τυπική απόκλιση του χρόνου λειτουργίας του μετασχηματιστή.
Θα χρησιμοποιήσω την σχέση 5.6, γι' αυτό πρώτα βρίσκω την μέση τιμή,

$$E(X) = \int_0^{\infty} x \cdot \lambda e^{-\lambda x} dx$$

Ολοκληρώνοντας μερικά και θέτοντας $\lambda e^{-\lambda x} dx = dv$ και $x=u$, επιτυγχάνουμε $v=-e^{-\lambda x}$, $du=dx$. Έτσι, η μέση τιμή είναι

$$E(X) = [-xe^{-\lambda x}]_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx = 0 + \frac{1}{\lambda} = \frac{1}{\lambda}.$$

Με παρόμοια ολοκλήρωση μπορούμε να υπολογίσουμε

$$E(X^2) = \int_0^{\infty} x^2 \cdot \lambda e^{-\lambda x} dx = [-x^2 e^{-\lambda x}]_0^{\infty} + \int_0^{\infty} 2xe^{-\lambda x} dx = 2\left(\frac{1}{\lambda}\right)^2$$

Έτσι από την σχέση 5.6 έχουμε

$$Var(X) = E(X^2) - [E(X)]^2 = 2\left(\frac{1}{\lambda}\right)^2 - \left(\frac{1}{\lambda}\right)^2 = \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

Τέλος, η τυπική απόκλιση της X είναι

$$\sigma_X = \frac{1}{\lambda}$$

δηλαδή η σταθερά που εμφανίζεται στον εκθέτη της συνάρτησης πυκνότητας πιθανότητας είναι το αντίστροφο της μέσης τιμής ή της τυπικής απόκλισης του χρόνου λειτουργίας του μετασχηματιστή.

5.3 ΤΥΠΙΚΗ ΤΥΧΑΙΑ ΜΕΤΑΒΛΗΤΗ

Όπως θα δούμε, πολλές φορές είναι ανάγκη να χρησιμοποιούμε την **τυπική** τυχαία μεταβλητή. Για παράδειγμα, αν οι τυχαίες μεταβλητές X και Y παριστάνουν τους μισθούς των μηχανικών στην Ελλάδα και Γερμανία αντίστοιχα, τα δύο μεγέθη δεν είναι άμεσα συγκρίσιμα. Ας πούμε, ότι κάποιος έλληνας μηχανικός πληρώνεται μηνιαίως 600,000 δραχμές, X=600,000, και κάποιος γερμανός μηχανικός 3000 μάρκα, Y=3000. Ποιος από τους δύο αμείβεται καλύτερα; Πρώτον, οι αμοιβές εκφράζονται σε διαφορετικές μονάδες, και δεύτερον το σχετικό οικονομικό επίπεδο των δύο χωρών είναι διαφορετικό. Δηλαδή, στην Ελλάδα η συγκεκριμένη τιμή x=600,000 δρχς θεωρείται πολύ καλή αμοιβή γιατί υπερβαίνει τον μέσο μισθό μ_X του μηχανικού, ενώ στη Γερμανία η τιμή y=3000 μάρκα είναι αρκετά χαμηλή σχετικά με την μέση αμοιβή μ_Y του μηχανικού.

Μία **τυπική τυχαία μεταβλητή** είναι μία μεταβλητή με μηδέν μέση τιμή και διακύμανση ίση με ένα. Κάθε τυχαία μεταβλητή X μπορεί να τυποποιηθεί αφαιρώντας από αυτήν την μέση τιμή της μ_X και διαιρώντας με την τυπική της απόκλιση σ_X. Η τυπική τυχαία μεταβλητή δηλώνεται συνήθως σαν ***Z***, και είναι

$$Z = \frac{X - \mu_X}{\sigma_X}$$

Η μέση τιμή μιας τυπικής τυχαίας μεταβλητής είναι πάντα 0, αφού εφαρμόζοντας τις ιδιότητες της μέσης τιμής έχουμε

$$E(Z) = E\left(\frac{X - \mu_X}{\sigma_X}\right) = \frac{E(X) - \mu_X}{\sigma_X} = 0 \quad ,$$

όπου μ_X και σ_X υποτίθεται ότι είναι σταθερές.

Η διακύμανση της τυπικής τυχαίας μεταβλητής Z είναι 1, αφού

$$Var(Z) = E(Z^2) - [E(Z)]^2 = E\left[\left(\frac{X - \mu_X}{\sigma_X}\right)^2\right] - 0^2$$

$$= \frac{E\left[(X - \mu_X)^2\right]}{\sigma_X^2} = \frac{\sigma_X^2}{\sigma_X^2} = 1.$$

Έτσι και η τυπική απόκλιση της Z είναι ένα, σ_Z=1.

Προφανώς, ο Z μετασχηματισμός δεν αλλάζει με κανένα τρόπο το σχήμα της αρχικής συνάρτησης κατανομής πιθανότητας. Η πιθανότητα για κάθε τιμή της Z είναι απλά η πιθανότητα της αντίστοιχης τιμής της αρχικής μεταβλητής X. Μεταξύ Z και X υπάρχει αντιστοιχία ένα προς ένα.

Η μέτρηση της τυπικής τυχαίας μεταβλητής Z δεν εκφράζεται σε μονάδες.

5.4 ΑΝΙΣΟΤΗΤΑ CHEBYSHEV

Για να περιγράψουμε την κατανομή πιθανότητας σε μία τυχαία μεταβλητή X πρέπει απαραίτητα να γνωρίζουμε την συνάρτηση κατανομής πιθανότητας $f(x)$. Παρόλα αυτά, χρήσιμες πληροφορίες γύρω από την κατανομή πιθανότητας της X παρέχεται εάν είναι γνωστή η μέση τιμή της μ_X και η τυπική της απόκλιση σ_X. Υπάρχει το θεώρημα γνωστό σαν Ανισότητα **Chebyshev** το οποίο μπορεί να προσδιορίσει τα όρια της πιθανότητας η τυχαία μεταβλητή να πάρει τιμές σε διάφορα διαστήματα. Το θεώρημα αυτό δηλώνει ότι για πεπερασμένη μέση τιμή μ_X , πεπερασμένη τυπική απόκλιση σ_X και h>0, ισχύει

$$P(|X-\mu_X|<h\sigma_X)\geq 1-\frac{1}{h^2}.$$

Λεκτικά, η πιθανότητα του γεγονότος ότι η τυχαία μεταβλητή X παίρνει μία τιμή x, η οποία απέχει από την μέση τιμή h τυπικές αποκλίσεις, είναι μεγαλύτερη από $p=1-(1/h)^2$, ανεξάρτητα από το πόσο μεγάλο είναι το h.

Ύστερα από απλές πράξεις η ανισότητα του Chebyshev μπορεί να γραφεί απλούστερα,

$$P(|X-\mu_X|\geq h\sigma_X)\leq\frac{1}{h^2},$$

το οποίο σημαίνει ότι η πιθανότητα η τιμή της τυχαίας μεταβλητής X να απέχει από την μέση τιμή μ_X περισσότερο από h φορές την τυπική απόκλιση σ_X έχει ανώτερο όριο το $(1/h)^2$. Στην πράξη η ανισότητα χρησιμοποιείται για h>1, διαφορετικά η τιμή $(1/h)^2$ είναι μεγαλύτερη της μονάδας και η πληροφορία της ανισότητας είναι περιττή.

Παράδειγμα 5.8

Σε ένα εργοστάσιο ο αριθμός τεμαχίων X ενός προϊόντος που παράγονται την εβδομάδα είναι τυχαία μεταβλητή. Η συνάρτηση κατανομής πιθανότητας είναι άγνωστη, όμως είναι γνωστό ότι κατά μέσο όρο παράγονται 50 τεμάχια την εβδομάδα με μία τυπική απόκλιση 5. Ποια η πιθανότητα αυτήν την εβδομάδα ο αριθμός των παραγόμενων τεμαχίων να βρεθεί μεταξύ 40 και 60.

Χρησιμοποιώντας την ανισότητα **Chebyshev** μπορούμε να επιτύχουμε μόνο όριο της ζητούμενης πιθανότητας. Φυσικά αν γνωρίζαμε την συνάρτηση κατανομής πιθανότητας θα υπολογίζαμε ακριβώς την πιθανότητα αυτή.

Από την πρώτη ανισότητα Chebyshev,

$$P(|X-50|<10)=P(|X-50|<2\cdot 5)\ \geq\ 1-\frac{1}{2^2}=\frac{3}{4},$$

και έτσι η πιθανότητα ότι η παραγωγή της εβδομάδας αυτής θα είναι μεταξύ 40 και 60 είναι μεγαλύτερη από 0,75. ■

Παράδειγμα 5.9

Με το παράδειγμα αυτό επιδεικνύουμε την χρησιμότητα της ανισότητας Chebyshev. Θεωρούμε μία τυχαία μεταβλητή X με την ακόλουθη συνάρτηση κατανομής πιθανότητας:

x_i	1	2	3	4	5
$f(x_i)$	0,1	0,1	0,2	0,3	0,3

Για την τυχαία μεταβλητή υπολογίζουμε μέση τιμή και τυπική απόκλιση,

$$\mu_X=E(X)=3{,}6,$$
$$E(X^2)=14{,}6\ \Rightarrow\ \sigma_X^2=E(X^2)-\mu_X^2=14{,}6-3{,}6^2=1{,}64\ \Rightarrow \sigma_X=1{,}281.$$

Με την κατανομή $f(x_i)$	Με την ανισότητα Chebyshev.
$P(\lvert X-3{,}6\rvert\geq 1\cdot 1{,}281)=0{,}5$	$P(\lvert X-3{,}6\rvert\geq 1\cdot 1{,}281)\ \leq\ \frac{1}{1^2}=1$
$P(\lvert X-3{,}6\rvert\geq 2\cdot 1{,}281)=0{,}1$	$P(\lvert X-3{,}6\rvert\geq 2\cdot 1{,}281)\ \leq\ \frac{1}{2^2}=0{,}25$
$P(\lvert X-3{,}6\rvert\geq 3\cdot 1{,}281)=0$	$P(\lvert X-3{,}6\rvert\geq 3\cdot 1{,}281)\ \leq\ \frac{1}{3^2}=\frac{1}{9}$

Υπολογίζουμε την πιθανότητα η X να πάρει τιμές πέρα από την μέση τιμή 3,6 κατά h·1,281 για h=1, 2, 3. Στον πίνακα μπορούμε να συγκρίνουμε τις πιθανότητες από την συνάρτηση κατανομής με αυτές που προκύπτουν από την ανισότητα Chebyshev.

5.6. p-ΠΟΣΟΣΤΙΑΙΟ ΣΗΜΕΙΟ, ΔΙΑΜΕΣΟΣ, ΕΠΙΚΡΑΤΕΣΤΕΡΗ ΤΙΜΗ

Η μέση τιμή και διακύμανση μιας τυχαίας μεταβλητής X δεν είναι οι μόνες παράμετροι κεντρικής τάσης ή διασποράς. Μία ακόμη παράμετρος κεντρικής τάσης είναι η γνωστή *διάμεσος*. Προτού δώσουμε τον ορισμό της διαμέσου θα αναφερθούμε στο p-ποσοστιαίο σημείο μίας τυχαίας μεταβλητής X.

Ορισμός. Κάθε πραγματικός αριθμός $\boldsymbol{x_p}$ ονομάζεται ***p*-ποσοστιαίο** σημείο της τυχαίας μεταβλητής X αν ισχύει

$$P(X\leq x_p)=p \text{ και } P(X\geq x_p)=1-p,$$

Όπως φαίνεται και στο σχήμα 5.3

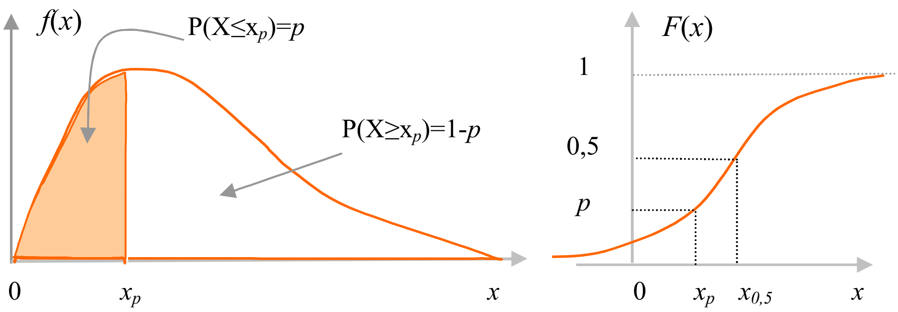
Όταν συμβαίνει σε ένα p-ποσοστιαίο σημείο να έχουμε p=0,5, τότε το σημείο x_p είναι η γνωστή **διάμεσος**.

Ορισμός. Κάθε πραγματικός αριθμός ονομάζεται **διάμεσος** της τυχαίας μεταβλητής X, και συμβολίζεται με $\boldsymbol{m}$ ή $\boldsymbol{M}$, εάν ισχύει

$$\boldsymbol{P(X\leq m)=0,5} \text{ και } \boldsymbol{P(X\geq m)=0,5}.$$

Ενδιαφέρων συχνά παρουσιάζεται και για την τιμή της X, η οποία εμφανίζεται συχνότερα από οιαδήποτε άλλη.

Ορισμός. Σε μία τυχαία μεταβλητή X, ονομάζουμε **επικρατέστερη ή συχνότερη τιμή** μία τιμή του πεδίου τιμών της X, η οποία **μεγιστοποιεί** την συνάρτηση κατανομής πιθανότητας $\boldsymbol{f(x)}$**.** Συνήθως η επικρατέστερη τιμή της X συμβολίζεται με το γράμμα $\boldsymbol{T}$.



***Σχήμα* 5.3** Γραφική παράσταση του p-ποσοστιαίου σημείου, x_p

Παράδειγμα 5.10
Μία διακριτή τυχαία μεταβλητή X μπορεί να πάρει τις τιμές {1, 2, 3, 4, ….}, άπειρες το πλήθος αλλά αριθμήσιμες. Η συνάρτηση κατανομής πιθανότητας δίνεται από την εξίσωση

$$f(x)=\frac{1}{2^x}\,.$$

Να ευρεθούν από τις παραμέτρους κεντρικής τάσης της X, η διάμεσός της M, και η επικρατέστερη τιμή της T.

Παρατηρώ ότι η διάμεσος της X δεν μπορεί να είναι το 1, διότι

$$P(X\leq 1)=P(X=1)=0{,}5 \text{ αλλά } P(X\geq 1)=P(X=1)+P(X=2)+P(X=3)+\ldots=1$$
$$\text{άρα } P(X\geq 1)\neq 0{,}5.$$

Επίσης, το 2 δεν είναι διάμεσος της X, διότι

$$P(X\geq 2)=P(X=2)+P(X=3)+P(X=4)+\ldots=0{,}5 \text{ αλλά } P(X\leq 2)=P(X=1)+P(X=2)=0{,}75 \neq 0{,}5.$$

Τέλος, κάθε αριθμός στο ανοικτό διάστημα (1, 2) είναι διάμεσος της X. Αν πάρω π.χ. M=1,5 ισχύει

$$P(X\leq 1{,}5)=P(X=1)=0{,}5 \text{ και } P(X\geq 1{,}5)=P(X=2)+P(X=3)+\ldots=0{,}5,$$

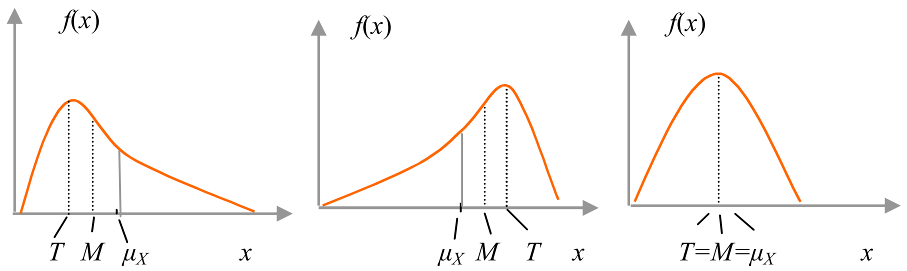
Άρα η διάμεσος είναι M=1,5, αλλά και κάθε άλλος αριθμός στο διάστημα (1, 2). Έχουμε άπειρο το πλήθος διαμέσους.

Η επικρατέστερη τιμή T, προφανώς είναι το 1, T=1, αφού η συνάρτηση κατανομής στο σημείο αυτό έχει την μέγιστη τιμή της,

$$f(1)=\frac{1}{2^1}=0{,}5\,.$$

Παρατήρηση

Κάθε τυχαία μεταβλητή έχει μόνο μία μέση τιμή. Η διάμεσος ή η επικρατέστερη τιμή μπορεί να είναι και περισσότερες από μία. Η σειρά με την οποία εμφανίζονται οι τρεις παράμετροι κεντρικής τάσης σε μία μη συμμετρική συνάρτηση κατανομής είναι μ_X , M, T ή αντίστροφα T, M, μ_X. Συμπίπτουν όμως σε μία συμμετρική κατανομή της X όπως φαίνεται και στο Σχήμα 5.4



Σχήμα 5.4. Θέση μεταξύ των παραμέτρων κεντρικής τάσης σε ασύμμετρο και συμμετρική κατανομή.

Παράδειγμα 5.11

Η ένταση ρεύματος X σε ένα κύκλωμα είναι τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας

$$f_X(x)=\begin{cases}\dfrac{x(4-x^2)}{4} & \textit{για} \quad 0\le x\le 2\\ 0 & \textit{αλλού}\end{cases}$$

Να συγκριθούν μεταξύ τους η **μέση** τιμή η **διάμεσος** και η **επικρατέστερη** τιμή.

Η **μέση** τιμή υπολογίζεται από το ολοκλήρωμα,

$$E(X)=\mu=\frac{1}{4}\int_0^2 x\cdot x(4-x^2)dx=\frac{1}{4}\left(\frac{4}{3}x^3-\frac{x^5}{5}\right)_0^2=\frac{16}{15}=1{,}07\,.$$

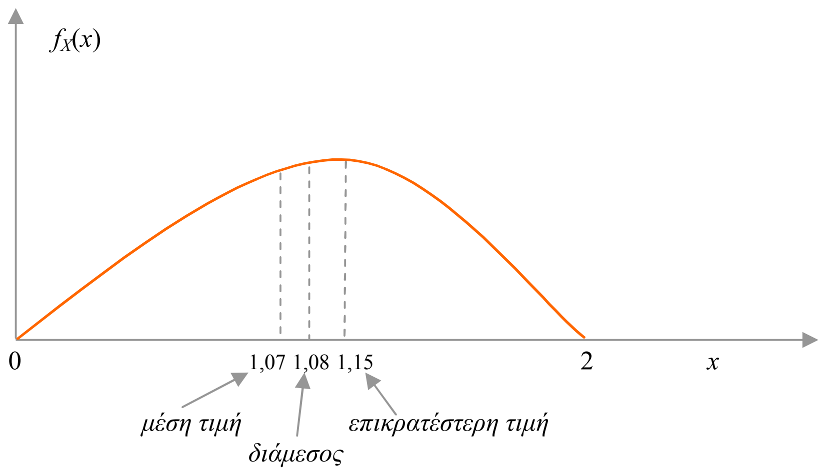
Για την **διάμεσο** έχουμε

$$\frac{1}{4}\int_0^M x(4-x^2)dx = 0{,}5 \Rightarrow \frac{1}{4}\left(2x^2 - \frac{x^4}{4}\right)_0^M = 0{,}5$$

$$\Rightarrow M^4 - 8M^2 + 8 = 0 \;\Rightarrow\; M^2 = \frac{8 \pm \sqrt{32}}{2}$$

και δεδομένου ότι η X βρίσκεται ανάμεσα σε 0 και 2 κρατούμε μόνο την ρίζα

$$M^2 = \frac{8-\sqrt{32}}{2} \Rightarrow M = 1{,}08$$



Σχήμα 5.5 Χαρακτηριστικά της μεταβλητής X του παραδείγματος 5.10

Η **πιθανότερη** τιμή της X είναι εκείνη η τιμή για την οποία μεγιστοποιείται η $f_X(x)$. Για να βρούμε τα μέγιστα της $f_X(x)$ μηδενίζουμε την παράγωγό της.

$$\frac{d}{dx}\left[\frac{x(4-x^2)}{4}\right] = \frac{4-3x^2}{4} = 0 \Rightarrow x = \sqrt{\frac{4}{3}} = 1{,}155$$

δεδομένου ότι η δεύτερη παράγωγος είναι μηδέν.

Τέλος, η σύγκριση των εκτιμηθέντων παραμέτρων φαίνεται στο Σχήμα 5.5, όπου επιβεβαιώνεται ότι η διάμεσος βρίσκεται ανάμεσα στην μέση και την επικρατέστερη τιμή. ■

5.6 ΑΛΛΕΣ ΠΑΡΑΜΕΤΡΟΙ ΚΑΙ ΡΟΠΕΣ

Εκτός της κεντρική τάση και διασπορά, η συνάρτηση μάζας πιθανότητας μπορεί να χαρακτηρισθεί και από τον βαθμό λοξότητας και από τον βαθμό κορυφότητας ή κυρτότητας. Περιγραφικές μετρήσεις αυτών των ιδιοτήτων μπορούν να προσδιορισθούν με την ορολογία των *ροπών* μιας συνάρτησης κατανομής. *Ροπές* είναι απλά η μέση τιμή διαφορετικών δυνάμεων της τυχαίας μεταβλητής. Έτσι, η ***κ-στη* ροπή** της τυχαίας μεταβλητής X γύρω από την αρχή των αξόνων ορίζεται ως η μαθηματική ελπίδα της μεταβλητής

$$X^{\kappa},\ E(X^{\kappa}).$$

Παρατηρούμε ότι η **πρώτη ροπή** ως προς την αρχή των αξόνων είναι η **μέση τιμή** της τυχαίας μεταβλητής X.

Ροπές μπορούν επίσης να προσδιορισθούν γύρω από οιοδήποτε άλλο σταθερό σημείο. Ειδικότερα, ροπές γύρω από την μέση τιμή μ παρουσιάζουν ιδιαίτερο ενδιαφέρον. Για παράδειγμα, η **δεύτερη ροπή** της X γύρω από την $\boldsymbol{\mu_X}$

$$\mu_2 = E\left[(X-\mu)^2\right],$$

είναι η **διακύμανση** της X. Όπως ακριβώς η μέση τιμή και διακύμανση μετρούν την τοποθεσία και την διασπορά μιας κατανομής, αντίστοιχα, έτσι μεγαλύτερες ροπές μετρούν άλλες ιδιότητες μιας κατανομής πιθανότητας. Η **τρίτη ροπή** γύρω από την μέση τιμή

$$\mu_3 = E\left[(X-\mu)^3\right]$$

μπορεί να χρησιμοποιηθεί για να προσδιορίσει εάν μία κατανομή είναι συμμετρική ή λοξή. Αφού όλες οι αποκλίσεις στην μ_3 είναι στον κύβο, αρνητικές και θετικές αποκλίσεις αναιρεί η μία την άλλη, δίνοντας $\mu_3=0$, εάν η κατανομή είναι συμμετρική γύρω από την μέση τιμή μ. Εάν η κατανομή είναι λοξή προς τα δεξιά, τότε $\mu_3>0$. Από την άλλη πλευρά, για μία αριστερά λοξή κατανομή, θα έχουμε $\mu_3<0$. Η τρίτη ροπή μ_3 από μόνη της δεν αποτελεί ικανοποιητική μέτρηση της λοξότητας, αφού το μέγεθος της μ_3 εξαρτάται από τις μονάδες μέτρησης των τιμών της X. Για να απομακρύνουμε το πρόβλημα αυτό, μπορούμε να σχηματίσουμε μία **σχετική μέτρηση** οριζόμενη ως

$$SK = \frac{\mu_3}{\sigma^3}.$$

Όταν $SK=\pm 1$, η κατανομή είναι πολύ λοξή. Όταν , η κατανομή είναι μέτρια λοξή, και όταν $0<|SK|<0{,}5$, η κατανομή είναι περίπου συμμετρική.

Η **τέταρτη ροπή** γύρω από την μέση τιμή,

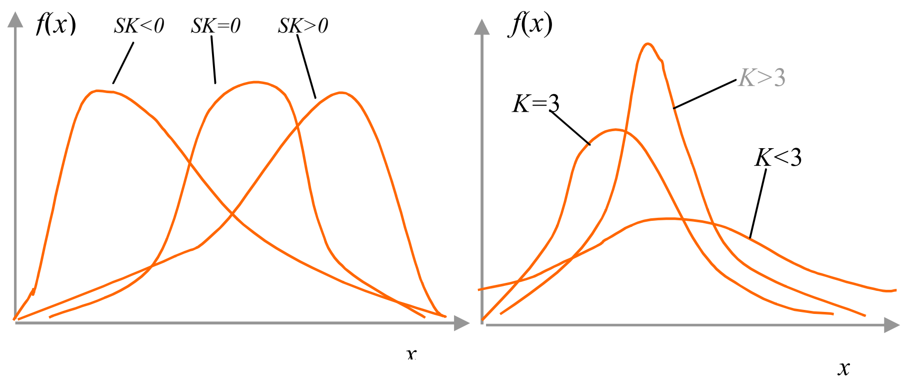
$$\mu_4 = E\left[(X-\mu_X)^4\right],$$

είναι πάντα θετική και μπορεί να χρησιμοποιηθεί για την μέτρηση της κορυφότητας ή κυρτότητας της κατανομής. Εισάγουμε πάλι μία σχετική μέτρηση όπως

$$K = \frac{\mu_4}{\sigma^4}$$

η οποία δείχνει το μέγεθος της **κυρτότητας** της κατανομής. Όταν $K=3$ η κατανομή είναι μεσόκυρτη σχετικά με μία Κανονική κατανομή όπως θα δούμε στο επόμενο κεφάλαιο. Όταν K<3, η κατανομή λέγεται πλατύκυρτη, και όταν K>3, η κατανομή λέγεται λεπτόκυρτη, (βλέπε Σχήμα 5.5).

Ροπές υψηλότερες από την μ_4 παρουσιάζουν μόνο θεωρητικό ενδιαφέρων και δεν θα μας απασχολήσουν προς το παρών.



Σχήμα 5.5. Γραφικές παραστάσεις μέτρησης λοξότητας και κυρτότητας κατανομών.

5.7 ΠΕΡΙΛΗΨΗ ΚΕΦΑΛΑΙΟΥ

Η μέση τιμή, η διακύμανση, η διάμεσος, η επικρατέστερη τιμή και άλλα χαρακτηριστικά περιγράφθηκαν και αναπτύχθηκαν σε αυτό το κεφάλαιο. Οι παράμετροι αυτοί δίνουν χρήσιμες πληροφορίες. Ακόμη, στην περίπτωση που είναι άγνωστη η συνάρτηση κατανομής πιθανότητας μία προσέγγιση της πιθανότητας για την τυχαία μεταβλητή μπορεί να δοθεί από την ανισότητα Chebyshev . Αναφέρθηκαν οι ιδιότητες των βασικότερων χαρακτηριστικών, καθώς και η πρακτική τους ερμηνεία. Σε αρκετές εφαρμογές του μηχανικού, η γνώση των χαρακτηριστικών των τυχαίων μεταβλητών που εμπλέκονται στην μελέτη του θα μπορούσε προσεγγιστικά να αναπληρώσει την πληροφορία από κάποια άγνωστη κατανομή πιθανότητας, και να τον διευκολύνει στην διερεύνηση των επιπτώσεων και λήψη απόφασης σχετικά με το έργου του.

ΑΣΚΗΣΕΙΣ-ΠΡΟΒΛΗΜΑΤΑ

Ασκήσεις με τις λύσεις τους

1) Κάλπη περιέχει πέντε σφαίρες αριθμημένες με 1 έως 5. Στα πλαίσια τυχαίου πειράματος επιλέγονται τρεις σφαίρες και έστω X ο ελάχιστος παρατηρούμενος αριθμός. Ποια είναι η μέση τιμή και η διακύμανση της τυχαίας μεταβλητής X.

Λύση

Η τυχαία μεταβλητή έχει πεδίο τιμών, Χ={1, 2, 3}, με συνάρτηση μάζας πιθανότητας $f(x)$

$$P(X=3)=P(\{345\})=\frac{1}{10}$$

$$P(X=2)=P(\{234,245,235\})=\frac{3}{10}$$

$$P(X=1)=P(\{123,124,125,134,135,145\})=\frac{6}{10}$$

Η μέση τιμή της X είναι

$$E(X)=3\cdot P(X=3)+2\cdot P(X=2)+1\cdot P(X=1)=1{,}5.$$

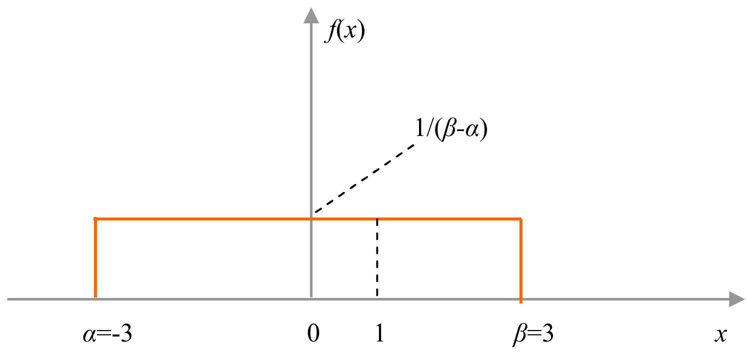
Η διακύμανση της X είναι

$$Var(X)=E(X^2)-[E(X)]^2=3^2\cdot\frac{1}{10}+2^2\cdot\frac{3}{10}+1^2\cdot\frac{6}{10}-1{,}5^2=0{,}45$$

2) Η τυχαία μεταβλητή X είναι ομοιόμορφα κατανεμημένη στο διάστημα (α,β).

(α) Υπολογίστε τις τιμές α και β αν γνωρίζετε ότι η μέση τιμή της X είναι μηδέν και ότι $P(X<1)=2/3$.

(β) Υπολογίστε τις πιθανότητες: $P(X=0{,}5)$, $P(X<0{,}5)$, $P(X<4)$, $P(-0{,}5<X<0{,}5)$.



Λύση

(α) Από το παράδειγμα 5.2 έχει προσδιορισθεί η αναλυτική μορφή της *σ.π.π.* σε μία ομοιόμορφη κατανομή, $f(x)=1/(\beta-\alpha)$, όπως και η μέση τιμή $E(X)=(\alpha+\beta)/2$. Άρα, μπορώ με βάση τα δεδομένα να έχω δύο εξισώσεις,

$$P(X>1)=\int_1^\beta f(x)dx=\int_1^\beta 1/(\beta-\alpha)dx=\frac{\beta-1}{\beta-\alpha}=\frac{1}{3}$$

$$E(X)=\frac{\alpha+\beta}{2}=0$$

Λύνοντας το παραπάνω σύστημα ως προς α και β βρίσκουμε α=-3 και β=3.

(β) Γνωρίζοντας πλήρως την *σ.π.π.* μπορώ να υπολογίσω τις πιθανότητες, $P(X=0,5)=0$ διότι η X είναι συνεχής,

$$P(X<0,5)=\int_{-3}^{0,5} f(x)dx=\int_{-3}^{0,5} 1/6dx=\frac{3,5}{6}$$

$$P(X<4)=1$$

$$P(-0,5<X<0,5)=\int_{-0,5}^{0,5} 1/6dx=\frac{1}{6}$$

3) Εάν η αθροιστική συνάρτηση πιθανότητας μιας τυχαίας μεταβλητής X δίνεται από την ακόλουθη μορφή

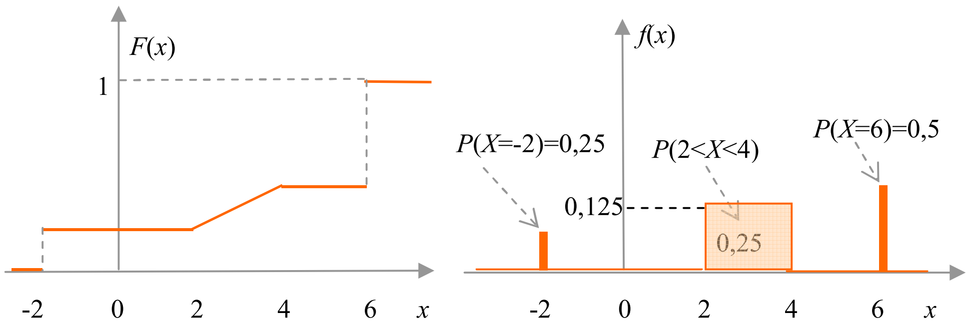
$$F(x)=\begin{cases} 0 & \gamma\iota\alpha \quad x<-2 \\ 0,25 & \gamma\iota\alpha \quad -2\le x<2 \\ 0,125x & \gamma\iota\alpha \quad 2\le x<4 \\ 0,5 & \gamma\iota\alpha \quad 4\le x<6 \\ 1 & \gamma\iota\alpha \quad x\ge 6 \end{cases}$$

να ευρεθεί η συνάρτηση πυκνότητας πιθανότητας της X, $f(x)$, η διάμεσος και η επικρατέστερη τιμή της X.

Λύση

Η τυχαία μεταβλητή X είναι προφανώς μικτού τύπου, αφού η αθροιστική συνάρτηση πιθανότητας κάνει άλματα στα σημεία –2 και 6. και είναι συνεχής και παραγωγίσιμη στο διάστημα $2<x<4$. Έτσι, η συνάρτηση μάζας πιθανότητας $f(x)$ θα ισούται με τα αντίστοιχα άλματα στα σημεία –2 και 6, ενώ στο διάστημα

$2<x<4$ ορίζεται από την παράγωγο της αθροιστικής $F(x($, όπως φαίνεται και στο σχήμα που ακολουθεί,



Τέλος, η αναλυτική μορφή της *σ.π.π.* $f(x)$ είναι

$$f(x)=\begin{cases} 0 \quad \gamma\iota\alpha \quad x<-2 \\ P(X=-2)=0{,}25 \quad \gamma\iota\alpha \quad x=-2 \\ 0 \quad \gamma\iota\alpha \quad -2<x<2 \\ c=0{,}125 \quad \gamma\iota\alpha \quad 2\le x\le 4 \\ 0 \quad \gamma\iota\alpha \quad 4<x<6 \\ P(X=6)=0{,}5 \quad \gamma\iota\alpha \quad x=6 \\ 0 \quad \gamma\iota\alpha \quad x>6 \end{cases}$$

Η διάμεσος M της X έχει τέτοια τιμή ώστε να ισχύει $P(X\le M)=P(X\ge M)=0{,}50$. Κάθε τιμή μεταξύ $4\le M<6$ ικανοποιεί την συνθήκη αυτή. Άρα, υπάρχουν άπειρες τιμές για την διάμεσο της μεταβλητής X.

Η επικρατέστερη τιμή της X είναι προφανώς η τιμή $T=6$, διότι η X παίρνει αυτήν την τιμή με την μεγαλύτερη πιθανότητα απ' ότι αλλού.

4) Ένα αυτοκίνητο στοχεύει να διανύσει μια απόσταση ΑΒ εκατό χιλιομέτρων. Σε τυχαίο σημείο της διαδρομής του ενδέχεται να παρουσιάσει κάποια βλάβη, και αποφασίζει να επιστρέψει ή να συνεχίσει αν το Α ή το Β είναι πλησιέστερα αντίστοιχα. Να βρεθεί η συνάρτηση πυκνότητας της συνολικής διανύομενης απόστασης. Ποια είναι η μέση συνολική διανύομενη απόσταση;

Λύση

Η συνολική διανυόμενη απόσταση X είναι τυχαία μεταβλητή μικτού τύπου με συνάρτηση μάζας πιθανότητας,

$$f(x)=\begin{cases} f_1(x)=P(X=100)=\frac{1}{2} & \gamma\iota\alpha \quad x=100 \\ f_2(x)=\dfrac{1}{2} \quad \gamma\iota\alpha \quad 0\le x<100 & \\ 0 \qquad \alpha\lambda\lambda o\upsilon & \end{cases}$$

Από τον ορισμό της μέσης τιμής $E(X)$ μιας τυχαίας μεταβλητής μικτού τύπου ισχύει

$$E(X)=100\cdot P(X=1)+\int_0^{100} xf_2(x)dx$$

$$=100\cdot\frac{1}{2}+\int_0^{100} x\frac{1}{2}dx=\frac{1}{2}+\frac{1}{2}\left[\frac{x^2}{2}\right]_0^{100}=75 \quad \chi\lambda\mu.$$

5) Υποθέστε ότι $f(x)$ είναι η συνάρτηση πυκνότητας πιθανότητας μιας τυχαίας μεταβλητής X. Αν η *σ.π.π.* είναι συμμετρική γύρω από το α, δηλαδή, $f(\alpha+x)=f(\alpha-x)$ για κάθε x. Να αποδειχθεί ότι η μέση τιμή $E(X)$ ισούται με α, δεδομένου ότι υπάρχει.

Λύση

Ξεκινώντας από τον ορισμό της μέσης τιμής μπορούμε να γράψουμε,

$$E(X)=\int_{-\infty}^{\infty} xf(x)dx=\int_{-\infty}^{\infty}(x-\alpha)f(x)dx+\alpha\int_{\alpha}^{\infty} xf(x)dx=\int_{-\infty}^{\infty}(x-\alpha)f(x)dx+\alpha$$

$$=\int_{-\infty}^{\alpha}(x-\alpha)f(x)dx+\int_{\alpha}^{\infty}(x-\alpha)f(x)dx+\alpha$$

Αλλά ισχύει

$$\int_{-\infty}^{\alpha}(x-\alpha)f(x)dx=-\int_{\alpha}^{\infty}(x-\alpha)f(x)dx$$

Άρα

$E(X)=\alpha$.

6) Ανισότητα Chebychev's. Έστω X μια τυχαία μεταβλητή με μέση τιμή μ και διασπορά σ^2. Τότε για κάθε θετικό αριθμό h, η πιθανότητα να εμφανιστεί τιμή που απέχει από τη μέση τιμή τουλάχιστον $h\cdot\sigma$ είναι το πολύ $\frac{1}{h^2}$,

$$P(|X-\mu_X|\ge h\sigma_X)\le\frac{1}{h^2}.$$

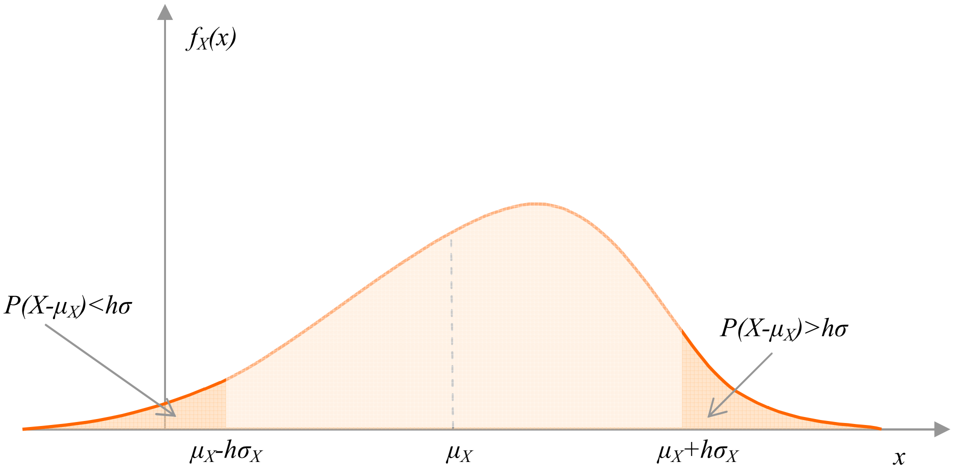
Απόδειξη: Θα περιορισθούμε στην περίπτωση της συνεχούς τυχαίας μεταβλητής, η απόδειξη για την διακριτή τυχαία μεταβλητή είναι παρόμοια.
Εφαρμόζουμε τον ορισμό της διακύμανσης,

$$\sigma_X^2 = \int_{-\infty}^{\infty}(x-\mu_X)^2 f_X(x)dx \geq \int_{-\infty}^{\mu_X - h\sigma}(x-\mu_X)^2 f_X(x)dx + \int_{\mu_X + h\sigma}^{\infty}(x-\mu_X)^2 f_X(x)dx$$

Δεδομένου ότι για τα όρια των δύο τελευταίων ολοκληρωμάτων ισχύει

$$(x-\mu_X)^2 \geq (h\cdot\sigma_X)^2$$

έχουμε



$$\sigma_X^2 \geq (h\sigma_X)^2\left[\int_{-\infty}^{\mu_X - h\sigma} f_X(x)dx \;+\; \int_{\mu_X + h\sigma}^{\infty} f_X(x)dx\right] = (h\sigma_X)^2 P(|X-\mu_X| \geq h\cdot\sigma)$$

ή

$$P(|X-\mu_X| \geq h\cdot\sigma) \leq \frac{1}{h^2}$$

Προβλήματα με τις λύσεις τους

1) Μειοδοτικός διαγωνισμός. Ένας εργολάβος υποβάλλει προσφορές για τρία μεγάλα οικοδομικά έργα Α, Β και Γ. Οι πιθανότητες επιτυχίας για το κάθε έργο είναι αντίστοιχα $P(A)$=0,5, $P(B)$=0,8 και $P(\Gamma)$=0,2, όπου *A, B* και *Γ* παριστάνουν τα γεγονότα ότι ο εργολάβος θα επιτύχει στα αντίστοιχα έργα. Ποιος ο αναμενόμενος αριθμός των έργων που θα επιτύχει ο εργολάβος σε αυτόν τον διαγωνισμό, αν οι επιτυχίες των έργων είναι στατιστικώς ανεξάρτητες;

Λύση

Η τυχαία μεταβλητή *X* η οποία παριστάνει τον αριθμό έργων που θα κερδίσει ο εργολάβος από τον διαγωνισμό έχει σαν πεδίο τιμών το σύνολο {0, 1, 2, 3}. Η συνάρτηση κατανομής πιθανότητας ορίζεται από τις πιθανότητες $p(0)$, $p(1)$, $p(2)$, $p(3)$, οι οποίες εύκολα μπορούν να εκτιμηθούν, διότι έχουμε αντιστοιχία με τα ακόλουθα γεγονότα στον δειγματικό χώρο:

{*X*=0}={*κανένα από τα γεγονότα A, B, Γ*}
{*X*=1}={ *μόνο ένα από τα γεγονότα A, B, Γ*}
{*X*=2}={*μόνο δύο από τα γεγονότα A, B, Γ*}
{*X*=3}={*και τα τρία γεγονότα A, B, Γ*}

Τα γεγονότα αυτά μπορούν να συμβολιστούν με την άλγεβρα συνόλων,

$$\{X=0\} = \overline{(A \cup B \cup \Gamma)}$$
$$\{X=1\} = (A \cap \overline{B \cup \Gamma}) \cup (B \cap \overline{A \cup \Gamma}) \cup (\Gamma \cap \overline{A \cup B})$$
$$\{X=2\} = (A \cap B \cap \overline{\Gamma}) \cup (A \cap \overline{B} \cap \Gamma) \cup (\overline{A} \cap B \cap \Gamma)$$
$$\{X=3\} = A \cap B \cap \Gamma$$

Δεδομένου ότι τα γεγονότα *A, B, Γ* είναι ανεξάρτητα έχουμε

$$P(X=0) = P(\overline{A} \cap \overline{B} \cap \overline{\Gamma}) = P(\overline{A})P(\overline{B})P(\overline{\Gamma}) = 0{,}5 \cdot 0{,}2 \cdot 0{,}8 = 0{,}08$$

$$\begin{aligned} P(X=1) &= P[(A \cap \overline{B} \cap \overline{\Gamma}) \cup (\overline{A} \cap B \cap \overline{\Gamma}) \cup (\overline{A} \cap \overline{B} \cap \Gamma)] = P(A)P(\overline{B})P(\overline{\Gamma}) \\ &\quad + P(\overline{A})P(B)P(\overline{\Gamma}) + P(\overline{A})P(\overline{B})P(\Gamma) \\ &= 0{,}5 \cdot 0{,}2 \cdot 0{,}8 + 0{,}5 \cdot 0{,}8 \cdot 0{,}8 + 0{,}5 \cdot 0{,}2 \cdot 0{,}2 = 0{,}08 + 0{,}32 + 0{,}02 = 0{,}42 \end{aligned}$$

Παρόμοια εκτιμούμε και τις άλλες πιθανότητες, $P(X{=}2){=}0{,}42$, $P(X{=}3){=}0{,}08$.

Έτσι, έχουμε μία τυχαία μεταβλητή $X{=}\{0,\ 1,\ 2,\ 3\}$ με συνάρτηση μάζας πιθανότητας,

$$p(0){=}0{,}08,\ \ p(1){=}0{,}42,\ \ p(2){=}0{,}42,\ \ p(3){=}0{,}08.$$

Εύκολα υπολογίζουμε και την μέση τιμή της X,

$$\mu = \sum_{x_i=0}^{3} x_i\, p(x_i) = 0\cdot 0{,}08 + 1\cdot 0{,}42 + 2\cdot 0{,}42 + 3\cdot 0{,}08 = 1{,}58$$

‘Όπως βλέπουμε η τιμή 1,58 δεν ανήκει στο πεδίο τιμών της X. Παρόλα αυτά είναι χρήσιμη, π.χ. σε 100 παρόμοιους διαγωνισμούς ο εργολάβος ελπίζει να αναλάβει περίπου 158 έργα!

2) Διάμετρος σύρματος. Η διάμετρος X ενός σύρματος είναι τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας $f(x){=}6x(1{-}x)$, $0{<}x{\le}1$.

(α) Να ελεγχθεί αν η παραπάνω συνάρτηση είναι πράγματι συνάρτηση πυκνότητας πιθανότητας και να βρεθεί η αθροιστική συνάρτηση πιθανότητας $F(x)$.

(β) Να βρεθεί η μέση τιμή και η πιθανότερη τιμή της X.

Λύση

(α) Μία συνάρτηση μιας τυχαίας* X *είναι συνάρτηση πυκνότητας πιθανότητας αν πληροί δύο βασικές ιδιότητες:

$$f(x)\ge 0,\quad \int_{-\infty}^{\infty} f(x)dx = 1.$$

Πράγματι η τιμή της συνάρτησης $f(x){=}6x(1{-}x)$, για $0{<}x{\le}1$ είναι πάντα θετική. Για την δεύτερη ιδιότητα, βρίσκουμε την τιμή του ολοκληρώματος,

$$\int_{-\infty}^{\infty} 6x(1-x)dx = \int_{0}^{1} 6x(1-x)dx = 6\left[\frac{x^2}{2}\right]_0^1 - 6\left[\frac{x^3}{3}\right]_0^1 = 3-2=1.$$

Άρα η συνάρτηση $f(x){=}6x(1{-}x)$, για $0{<}x{\le}1$ είναι πράγματι συνάρτηση πυκνότητας πιθανότητας της X. Εύκολα μπορούμε να προσδιορίσουμε την αθροιστική συνάρτηση πιθανότητας,

$$F(x) = \int_0^x 6u(1-u)du = 6\left[\frac{u^2}{2}\right]_0^x - 6\left[\frac{u^3}{3}\right]_0^x = 3x^2 - 2x^3 = x^2(3-2x)$$

Παρατηρούμε ότι για x=1 η τιμή της αθροιστικής είναι 1, δηλαδή

$$F(x)=\begin{cases} x^2(3-2x) & \text{για} \quad 0<x\le 1 \\ 1 & \text{για} \quad x\ge 1 \\ 0 & \text{αλλου} \end{cases}$$

(β) Από τον ορισμό της μέσης τιμής έχουμε

$$E(X)=\int_0^1 x\cdot 6x(1-x)dx = 6\left[\frac{x^3}{3}\right]_0^1 - 6\left[\frac{x^4}{4}\right]_0^1 = 2-1{,}5=0{,}5$$

Η συχνότερη τιμή της τυχαίας μεταβλητής X βρίσκεται αν παραγωγίσουμε την συνάρτηση $f(x)$ ως προς x και λύσουμε την εξίσωση,

$$\frac{\partial f(x)}{\partial x}=6-12x=0 \quad \Rightarrow \quad x=0{,}5$$

Δηλαδή, σύρματα διαμέτρου γύρω από το 0,5 εμφανίζονται συχνότερα από οιαδήποτε άλλη τιμή.

3) Κατανάλωση πετρελαίου. Η μηνιαία κατανάλωση πετρελαίου X για θέρμανση μιας πολυκατοικίας σε χιλιάδες γαλόνια είναι τυχαία μεταβλητή με πυκνότητα

$$f(x)=\begin{cases} 5(1-x)^4 & \text{για} \quad 0\le x\le 1 \\ 0 & \text{αλλου} \end{cases}$$

Ποια η αναμενόμενη κατανάλωση πετρελαίου για τους πέντε μήνες της χειμερινής περιόδου όπου οι μήνες είναι το ίδιο ψυχροί;

Λύση

Η μέση κατανάλωση πετρελαίου τον μήνα μπορεί να υπολογισθεί από τον ορισμό της

$$E(X)=\int_0^1 x\cdot 5(1-x)^4\,dx$$

Θέτω t=1-x και ολοκληρώνουμε

$$E(X)=\int_0^1 x\cdot 5(1-x)^4\,dx=-5\int_1^0 t^4dt+5\int_1^0 t^5dt.=-5\left[\frac{t^5}{5}\right]_1^0+5\left[\frac{t^6}{6}\right]_1^0=1-\frac{5}{6}=\frac{1}{6}$$

Μπορώ να υποθέσω ότι X_i (i=1, 2, 3, 4, 5) είναι ανεξάρτητες τυχαίες μεταβλητές οι οποίες παριστάνουν την κατανάλωση του κάθε μήνα αντίστοιχα και ακολουθούν την ίδια κατανομή πιθανότητας με την X. Το άθροισμα των μεταβλητών αυτών

$$X_1+X_2+X_3+X_4+X_5$$

είναι τυχαία μεταβλητή και παριστάνει την συνολική κατανάλωση των πέντε μηνών. Η αναμενόμενη συνολική κατανάλωση είναι

$$E(X_1+X_2+X_3+X_4+X_5)=E(X_1)+....+E(X_5)=5E(X)=5\frac{1}{6}=\frac{5}{6},$$

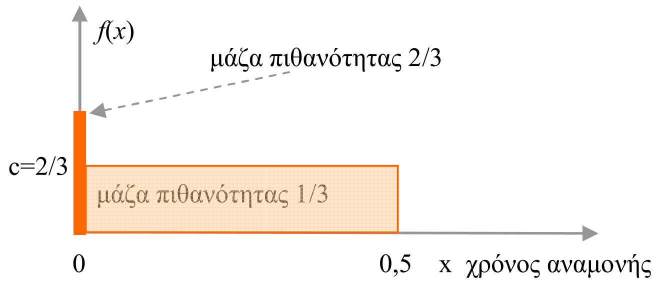
σύμφωνα με την **(Δ)** ιδιότητα της μέσης τιμής.

4) Χρόνος αναμονής στο φανάρι. Το φανάρι μιας διασταύρωσης δείχνει διαδοχικά ένα λεπτό πράσινο και μισό λεπτό κόκκινο. Να υπολογιστούν:

- η μέση τιμή του χρόνου αναμονής του αυτοκινήτου στο φανάρι,
- η τυπική απόκλιση του χρόνου αναμονής,
- πόσο χρόνο το πολύ περιμένουν στο φανάρι το 50% των αυτοκινήτων που περνούν από αυτήν την διασταύρωση;

Λύση

ο χρόνος αναμονής X ενός αυτοκινήτου που πλησιάζει τυχαία στο φανάρι είναι μικτή τυχαία μεταβλητή. Διότι, η πιθανότητα ένα αυτοκίνητο που πλησιάζει τυχαία την διασταύρωση να συναντήσει πράσινο είναι 2/3. Έτσι, στο σημείο x=0 η τυχαία μεταβλητή είναι διακριτή αφού $P(X=0)=2/3$. Ακόμη, αν το αυτοκίνητο συναντήσει κόκκινο, ο χρόνος αναμονής x μπορεί να πάρει άπειρες τιμές, $0< x \leq 0{,}5$ λεπτού, με μία σταθερή συνάρτηση πυκνότητας πιθανότητας, όπως φαίνεται και στο ακόλουθο Σχήμα.



Η σταθερά c πρέπει να είναι 2/3, λόγω της βασικής ιδιότητας της συνάρτησης πυκνότητας πιθανότητας *f*(*x*).

$$f(x) = \begin{cases} P(X=0) = \dfrac{2}{3} & \gamma\iota\alpha \quad x = 0 \\ c = \dfrac{2}{3} & \gamma\iota\alpha \quad 0 < x \le 0{,}5 \\ 0 \quad \alpha\lambda\lambda o\upsilon & \end{cases}$$

Έτσι, ο μέσος χρόνος αναμονής του αυτοκινήτου στο φανάρι είναι η μέση τιμή της *X*,

$$E(X) = 0 \cdot P(X=0) + \int_0^{0,5} x \cdot c \cdot dx = 0 \cdot \frac{2}{3} + \int_0^{0,5} x \cdot \frac{2}{3} dx = 0 \cdot \frac{2}{3} + \frac{2}{3}\left[\frac{x^2}{2}\right]_0^{0,5} = \frac{1}{12}.$$

Η τυπική απόκλιση του χρόνου αναμονής μπορεί να υπολογισθεί παρόμοια,

$$Var(X) = E(X^2) - [E(X)]^2 = 0^2 \cdot \frac{2}{3} + \int_0^{0,5} x^2 \cdot \frac{2}{3} dx - \left(\frac{1}{12}\right)^2 = \frac{2}{3}\left[\frac{x^3}{3}\right]_0^{0,5} - \left(\frac{1}{12}\right)^2$$

$$= \frac{1}{36} - \frac{1}{144} = \frac{3}{144}$$

και συνεπώς η τυπική απόκλιση είναι

$$\sigma_X = \sqrt{\frac{3}{144}} \text{ λεπτά.}$$

Στο τρίτο ερώτημα πρέπει να βρούμε την διάμεσο *M* έτσι ώστε

$$P(X \le M) = P(X \ge M) = 0{,}50\ .$$

Προφανώς, δεν υπάρχει τέτοιο σημείο, όπως φαίνεται και από την συνάρτηση μάζας πιθανότητας. Παρόλα αυτά το 50% των αυτοκινήτων έχει χρόνο αναμονής μηδέν.

6) Πλάτος δρόμου. Η απαιτούμενη ημερήσια κυκλοφοριακή κίνηση, *X*, για ένα προτεινόμενο δρόμο έχει την ακόλουθη *σ.π.π.*:

$$f(x)=\begin{cases} c & \gamma\iota\alpha \quad 2<x\le 4 \\ 0 & \alpha\lambda\lambda o\upsilon \end{cases}$$

Ο μηχανικός μπορεί να επιλέξει τέτοιο πλάτος για τον δρόμο ώστε η δυνατότητα ημερήσιας κυκλοφοριακής κίνησης να είναι ίση με ένα από τα εξής:

- το 0,8-*ποσοστιαίο σημείο* της *X*,
- την διάμεσο της *X*,

με πιθανότητες 0,65 και 0,35 αντίστοιχα. Να υπολογισθεί η πιθανότητα ότι ο δρόμος θα καλύπτει την ημερήσια ανάγκη της κυκλοφοριακής κίνησης.

Λύση

Το 0,8-*ποσοστιαίο σημείο* της *X* είναι ένα σημείο $x_{0,8}$ έτσι ώστε $P(X< x_{0,8})$=0,8. Αφού η κατανομή της *X* είναι ομοιόμορφη το 0,8-*ποσοστιαίο σημείο* της είναι η τιμή 3,6. Προφανώς η διάμεσος της *X*, *M*, είναι η τιμή 3.

Συμβολίζω τα γεγονότα

E={ *ο δρόμος θα καλύπτει την κυκλοφοριακή κίνηση*}
A_1={*δυνατότητα του δρόμου ίση με* 3,6}
A_2={ *δυνατότητα του δρόμου ίση με* 3}.

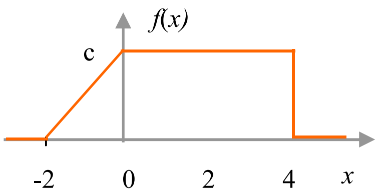
Σχετικά, με τις επιλογές της χωρητικότητας του δρόμου, τα γεγονότα A_1 και A_2 αποτελούν διαμέριση του δειγματικού χώρου, και σύμφωνα με το θεώρημα της ολικής πιθανότητας έχουμε,

$$P(E)=P(E|A_1)P(A_1)+P(E|A_2)P(A_2)=0{,}8\cdot 0{,}65+0{,}5\cdot 0{,}35=0{,}695\ .$$

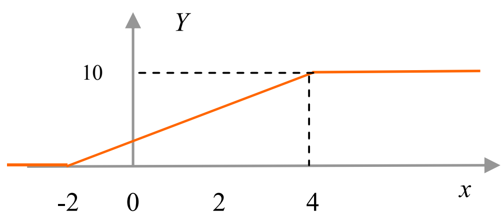
Όπου,

$$P(E|A_1)=P(X\le x_{0,8})=0{,}8$$
$$P(E|A_2)=P(X\le M)=0{,}50$$

7) Απόδοση μηχανήματος. Μία ηλεκτρική τάση *X* ακολουθεί την κατανομή του Σχήματος *α*. Η τάση *X* κινεί μηχάνημα του οποίου η απόδοση *Y* εξαρτάται από την τιμή της *X*, όπως φαίνεται στο Σχήμα *β*.



Σχήμα α



Σχήμα β

(α) Να ευρεθεί η αθροιστική συνάρτηση πιθανότητας $F(x)$ και η διάμεσος M της X.
(β) Να ευρεθεί η μέση απόδοση του μηχανήματος.

Λύση
Από το Σχήμα *α* μπορώ να προσδιορίσω την αναλυτική μορφή της *σ.π.π.* $f(x)$. Η σταθερά c πρέπει να έχει τιμή c=0,2, διότι το εμβαδόν κάτω από την *σ.π.π.* πρέπει να ισούται με την μονάδα. Ακόμη, η αναλυτική μορφή της γραμμής στο διάστημα $-2 \le x \le 0$ εύκολα προσδιορίζεται αν σχηματίσουμε δύο εξισώσεις για x=-2 και για x=0. Έτσι έχω

$$f_X(x) = \begin{cases} 0 & \gamma\iota\alpha \quad x < -2 \\ 0{,}1x + 0{,}2 & \gamma\iota\alpha \quad -2 \le x < 0 \\ 0{,}2 & \gamma\iota\alpha \quad 0 \le x \le 4 \\ 0 & \gamma\iota\alpha \quad x > 4 \end{cases}$$

(α) Η αθροιστική συνάρτηση πιθανότητας προσδιορίζεται με βάση τον ορισμό της και την *σ.π.π.* $f(x)$

$$F_X(x) = \begin{cases} 0 & \gamma\iota\alpha \quad x < -2 \\ \int_{-2}^{x} (0{,}1u + 0{,}2)du = 0{,}05x^2 + 0{,}2x + 0{,}2 & \gamma\iota\alpha \quad -2 \le x < 0 \\ 0{,}2 + \int_0^x 0{,}2du = 0{,}2 + 0{,}2x & \gamma\iota\alpha \quad 0 \le x \le 4 \\ 1 & \gamma\iota\alpha \quad x > 4 \end{cases}$$

Η διάμεσος M, όπως φαίνεται και από το Σχήμα *α*, βρίσκεται μετά το μηδέν, και από τον ορισμό της έχουμε

$$0{,}2 + \int_0^M 0{,}2dx = 0{,}5 \quad \Rightarrow \quad 0{,}2 + 0{,}2M = 0{,}5 \quad \Rightarrow \quad M = 1{,}5$$

(β) Η τυχαία μεταβλητή Y, όπως φαίνεται και από το Σχήμα *β*, είναι συνάρτηση της X με αναλυτική μορφή Y=(10/6)X+20/6. Σύμφωνα με βασική ιδιότητα της $E(X)$, έχουμε για την μέση απόδοση του μηχανήματος,

$$E(Y) = \int_{-2}^{0}((10/6)x + 20/6)(0{,}1x + 0{,}2)dx + \int_{0}^{4}(10/6)x + 20/6)0{,}2dx$$

$$= \frac{1}{6}\left[\frac{x^3}{3}\right]_{-2}^{0} + \frac{2}{6}\left[\frac{x^2}{2}\right]_{-2}^{0} + \frac{2}{6}\left[\frac{x^2}{2}\right]_{-2}^{0} + \frac{4}{6}[x]_{-2}^{0} + \frac{2}{6}\left[\frac{x^2}{2}\right]_{0}^{4} + \frac{4}{6}[x]_{0}^{4}$$

$$= \frac{88}{18}$$

Ασκήσεις προς λύση

1) Ένας φοιτητής απαντά σε ένα *τέστ* πολλαπλής επιλογής, το οποίο αποτελείται από δύο προβλήματα. Το πρώτο πρόβλημα έχει τρεις δυνατές απαντήσεις και το δεύτερο πέντε. Ο φοιτητής επιλέγει τυχαία μία από κάθε πρόβλημα. Να βρεθεί :
(α) Η μέση τιμή, $E(X)$, των σωστών απαντήσεων X του φοιτητή.
(β) Η διακύμανση $Var(X)$.
Να γενικευθούν οι απαντήσεις.

2) Το πάχος σε μία διαδικασία επικάλυψης είναι τυχαίο μέγεθος και μπορεί να πάρει ισοπίθανα τις τιμές 0,15, 0,16, 0,17,0,18, και 0,19. Να προσδιορίσετε την μέση τιμή και διακύμανση τους πάχους επικάλυψης.

3) Ένας παίκτης πληρώνει 1 ευρώ για κάθε συμμετοχή στο ακόλουθο παιχνίδι: Ρίχνονται τρία ζάρια, εάν έρθει ένας άσσος ο παίκτης παίρνει 1 ευρώ, εάν έρθουν δύο άσσοι παίρνει 2 ευρώ, και εάν έρθουν τρεις άσσοι παίρνει 8 ευρώ, διαφορετικά δεν παίρνει τίποτα. Είναι το παιχνίδι τίμιο; π.χ., είναι το μέσο κέρδος του παίκτη μηδέν; Εάν όχι, πόσο θα έπρεπε να πληρώνεται ο παίκτης όταν έρχονται τρεις άσσοι;

4) Μικτή κατανομή. Έστω ότι X είναι τυχαία μεταβλητή με αθροιστική κατανομή

$$F(x) = \begin{cases} 1 - 0{,}8e^{-x} & \gamma\iota\alpha \quad x \geq 0 \\ 0 \quad \gamma\iota\alpha \quad x < 0 \end{cases}$$

Να σχεδιάσετε την $F(x)$ και μετά να υπολογίσετε την μέση τιμή $E(X)$. Να δώσετε ένα παράδειγμα όπου μία τυχαία μεταβλητή X θα μπορούσε να έχει την παραπάνω κατανομή.

5) *Ανισότητα Chebyshev's*. Να συγκρίνεται το άνω όριο της πιθανότητας

$$P[|X - E(X)| \geq 2\sqrt{V(X)}]$$

το οποίο επιτυγχάνεται από την ανισότητα Chebyshev's με την ακριβή πιθανότητα εάν η X είναι ομοιόμορφα κατανεμημένη ($f(x)=c$) στο διάστημα (-1, 3).

6) Να αποδειχθεί ότι για κάθε τυχαία μεταβλητή X ισχύει,
(α) εάν $E(X)=E(X^2)=0$ τότε $P(X=0)=1$,
(β) εάν $Var(X)=0$ τότε υπάρχει σταθερά c έτσι ώστε $P(X=c)=1$,
(γ) $|E(X)| \leq E(|X|)$.

Προβλήματα προς λύση

1) μόλυνση Το μέγεθος μόλυνσης μορίου σε *micrometers* είναι τυχαία μεταβλητή X με συνάρτηση πυκνότητας πιθανότητας

$$f(x) = \begin{cases} 2x^{-3} & \textit{για} \quad 1 < x \\ 0 & \textit{αλλου} \end{cases}$$

Να προσδιορίσετε την *μέση τιμή*, την *επικρατέστερη τιμή* και την *διάμεσο* της X.

2) επικάλυψη αγωγού Το πάχος επικάλυψης σε έναν αγωγό είναι τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας

$$f(x) = \begin{cases} cx^{-2} & \textit{για} \quad 100 < x < 120 \quad \mu m \\ 0 & \textit{αλλου} \end{cases}$$

(α) Προσδιορίστε την μέση τιμή και τυπική απόκλιση του πάχους επικάλυψης.
(β) Εάν η επικάλυψη κοστίζει 0,50 ευρώ ανά micrometer, ποιο είναι το *μέσο* κόστος της επικάλυψης;

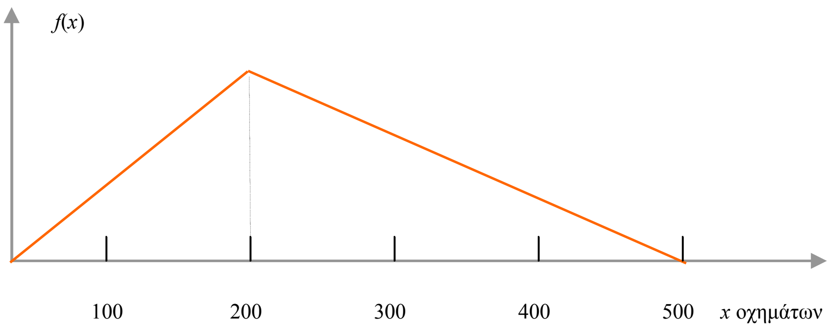
3) τρυπάνι Αν και το τρυπάνι φέρει αρίδα για τρύπα 5mm, ο θόρυβος, οι κραδασμοί και άλλοι στοχαστικοί παράγοντες προκαλούν διάμετρο μεγαλύτερη από 5 mm. Έτσι, σε μία ανοιγμένη τρύπα η συνάρτηση πυκνότητας πιθανότητας για την διάμετρο X είναι

$$f(x) = \begin{cases} cx^{-10(x-5)} & \gamma\iota\alpha \quad x > 5mm \\ 0 \quad \alpha\lambda\lambda o\upsilon & \end{cases}$$

(α) Να προσδιορίσετε την *μέση τιμή* και το $x_{0,90}$ (0,90-*ποσοστιαίο σημείο*) της διαμέτρου μιας ανοιγμένης τρύπας.
(β) Να εκτιμήστε την πιθανότητα ότι σε μία ανοιγμένη τρύπα η διάμετρός ξεπερνά τα 5.1 mm

4) χρόνος συναρμολόγησης Ο χρόνος συναρμολόγησης ενός συστήματος είναι τυχαίο μέγεθος μεταξύ $30<x<40$ δευτερόλεπτα με σταθερή συνάρτηση πυκνότητας πιθανότητας.
(α) Να προσδιορίσετε το ποσοστό των συναρμολογήσεων που απαιτεί περισσότερο από 35 δευτερόλεπτα.
(β) Ποιος χρόνος ξεπερνιέται από το 90% των συναρμολογήσεων; (0,10-*ποσοστιαίο σημείο*)
(γ) Να υπολογίσετε την μέση τιμή του χρόνου συναρμολόγησης.

5) όγκος κυκλοφορίας. Ο αναμενόμενος ωριαίος όγκος κυκλοφορίας για μία υπό μελέτη γέφυρα έχει τη *σ.π.π.* του Σχήματος.



Ο μηχανικός μπορεί να σχεδιάσει την γέφυρα έτσι ώστε η πραγματική χωρητικότητά της να είναι ίση με ένα από τα εξής:

- Την *πιθανότερη τιμή* της *τ. μ. X.*
- Την *διάμεσο* της *X.*
- Την *μέση τιμή* της *X.*

(α) Να ευρεθεί η συνάρτηση πυκνότητας πιθανότητας και η αθροιστική κατανομή της τυχαίας μεταβλητής *X.*

(β) Να υπολογιστεί η χωρητικότητα του δρόμου σε κάθε μία από τις παραπάνω περιπτώσεις, καθώς και η πιθανότητα ότι η χωρητικότητα αυτή δεν θα ικανοποιεί τις κυκλοφοριακές ανάγκες.

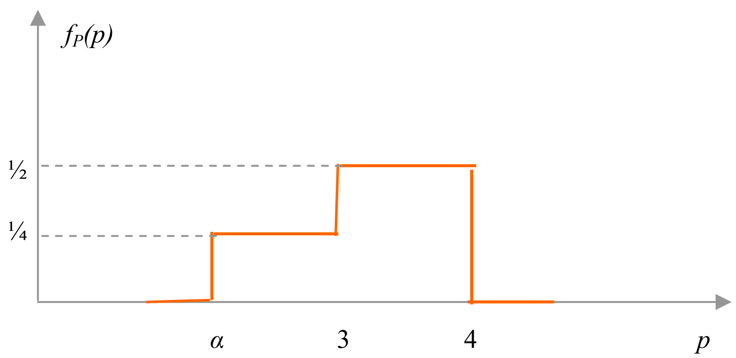
6) κατανάλωση ενέργειας Η μηνιαία κατανάλωση ηλεκτρικής ενέργειας X σε ένα εργοστάσιο είναι τυχαία μεταβλητή με πυκνότητα

$$f(x) = \begin{cases} \alpha(1-x)^4 & \gamma\iota\alpha \quad 0 \le x \le 1 \\ 0 & \alpha\lambda\lambda o\upsilon \end{cases}$$

(α) Να υπολογισθεί η *μέση τιμή* της μηνιαίας κατανάλωσης ηλεκτρικής ενέργειας.
(β) Ποια κατανάλωση ενέργειας (10/12-*ποσοστιαίο σημείο*) δεν ξεπερνιέται στους 10 μήνες του χρόνου;

7) Μετρήσεις της ισχύος $P=J^2R$ στη σταθερή αντίσταση R του παρακάτω κυκλώματος

Οδηγούν στην συνάρτηση πυκνότητας πιθανότητας του παρακάτω σχήματος.



Να υπολογισθεί η σταθερά α καθώς και η μέση τιμή και διασπορά της ισχύος P.

8) Υποβάθμιση ποιότητας. Το μήκος X των παραγόμενων σιδηροδοκών σε μία βιομηχανία είναι τυχαία μεταβλητή. Η διάσταση του μήκους των σιδηροδοκών όπως προβλέπεται από τις προδιαγραφές είναι 6 μέτρα. Όπως είναι λογικό κάθε απόκλιση από προδιαγραφόμενο μήκος επιφέρει απώλεια ποιότητας στην παραγωγή. Έτσι, η απώλεια ποιότητας προσδιορίζεται από την αναμενόμενη τιμή του στοχαστικού μεγέθους $K(X-6)^2$, όπου K είναι μία σταθερά παράμετρος.

(α) Αν το μήκος X της παραγόμενης σιδηροδοκού έχει μέση τιμή 6 μέτρα και τυπική απόκλιση 0,02 μέτρα, Ποια είναι η απώλεια ποιότητας στην παραγωγή;

(β) Αν το μήκος X της παραγόμενης σιδηροδοκού έχει μέση τιμή μ μέτρα και τυπική απόκλιση σ μέτρα, Ποια είναι η απώλεια ποιότητας στην παραγωγή;

6

Θεωρητικές Κατανομές Πιθανότητας για Διακριτή Τυχαία Μεταβλητή.

ΠΕΡΙΓΡΑΦΗ ΚΕΦΑΛΑΙΟΥ

Σε αυτό το κεφάλαιο θα παρουσιάσουμε τις *θεωρητικές κατανομές πιθανότητας* (μερικές από αυτές τις συναντήσαμε στα προηγούμενα κεφάλαια-ομοιόμορφη, εκθετική κ.τ.λ.), με την έννοια ότι αυτές προέρχονται από λογικές αιτίες παρά από πραγματικές εργαστηριακές παρατηρήσεις.

Θα επικεντρωθούμε περισσότερο σε ειδικές κατανομές πιθανότητας τυχαίων μεταβλητών, οι οποίες είναι κατάλληλα μαθηματικά μοντέλα για καταστάσεις του πραγματικού κόσμου κάτω από ειδικές συνθήκες. Τέτοια μοντέλα είναι σπουδαία με την ευρύτερη έννοια ότι μπορούν να μας βοηθήσουν να προβλέψουμε την συμπεριφορά των μελλοντικών επαναλήψεων του πειράματος τύχης. Έχουν επίσης πρακτικό ενδιαφέρων, αφού εάν έχουμε ακριβή μαθηματική φόρμουλα των διαφόρων καταστάσεων που συχνά συναντούνται στην στατιστική ανάλυση, υπολογισμοί των πιθανοτήτων που εμπλέκονται μπορούν να περιορισθούν σε μία τυπική διαδικασία. Επίσης, μία μαθηματική ανάλυση αυτών των ειδικών μοντέλων κατανομών μπορεί να μας βοηθήσει σε μία πληρέστερη κατανόηση των διαφόρων σχέσεων μεταξύ στοχαστικών γεγονότων ή μεγεθών.

Στην ανάλυση που θα ακολουθήσει θα προσδιορίσουμε την ακριβή μαθηματική έκφραση της κάθε ειδικής κατανομής και τις υποθέσεις κάτω από τις οποίες αυτά τα μοντέλα περιγράφουν μία κατάσταση στον πραγματικό κόσμο. Θα αναφερθούμε επίσης σε στατιστικούς πίνακες, έτσι ώστε να περιορίσουμε την υπολογιστική δουλειά. Ο αναγνώστης σε αυτή τη φάση πρέπει να παρατηρήσει ότι είναι πολύ σπουδαίο γι' αυτόν να συνδέσει το πρόβλημα από τον πραγματικό κόσμο με το σωστό μοντέλο κατανομής παρά να αποστηθίζει την λεπτομερή μαθηματική ανάπτυξη του μοντέλου.

6.1 Η ΚΑΤΑΝΟΜΗ BERNOULLI

Αρχίζουμε την ανάλυση των θεωρητικών κατανομών με το μοντέλο Bernoulli- όχι μόνο επειδή είναι το πιο απλό, αλλά επίσης διότι αποτελεί την βάση για τον προσδιορισμό της Διωνυμικής κατανομής, η οποία είναι κατάλληλη για ένα μεγάλο αριθμό πραγματικών διαδικασιών.

Η κατανομή **Bernoulli** ταιριάζει όταν ενδιαφερόμαστε για πειράματα τα οποία οδηγούν μόνο σε δύο αποτελέσματα, ας πούμε το γεγονός A και το συμπληρωματικό του $\overline{A}$, όπως επιτυχία ή αποτυχία, άρρεν ή θήλυ, ελαττωματικό ή μη ελαττωματικό και ούτω καθεξής. Προφανώς, η τυχαία μεταβλητή μπορεί να πάρει μόνο δύο τιμές. Για λόγους απλούστευσης, οι δύο τιμές της τυχαίας μεταβλητής είναι 1 ή 0 όταν συμβαίνει το γεγονός A ή $\overline{A}$ αντίστοιχα.

Ορισμός. Ένα πείραμα τύχης ή ένας απλός έλεγχος λέγεται **δοκιμή Bernoulli** εάν

- όλα τα δυνατά αποτελέσματα είναι δύο $\boldsymbol{S=\{A, \overline{A}\}}$, A=«επιτυχία» και $\overline{A}$ = «αποτυχία»,
- υπάρχει σταθερή πιθανότητα $\boldsymbol{p}$ για την «επιτυχία» και **1-$\boldsymbol{p}$** για την «αποτυχία», $\boldsymbol{P(A)=p}$ και $P(\overline{A})=1-p$.

και καλείται **κατανομή Bernoulli**

Η τυχαία μεταβλητή ***X***, η οποία αντιστοιχεί την τιμή 1 στο A και 0 στο $\overline{A}$, καλείται **μεταβλητή Bernoulli,** η δε συνάρτηση κατανομής πιθανότητας για την X είναι

x_i	$f(x_i)$	
1	p	(6.1)
0	$q=1-p$	

Για την μεταβλητή **Bernoulli** εύκολα υπολογίζουμε την μέση τιμή και διακύμανσή της.

$$E(X)=\sum_i x_i f(x_i)=p,$$

$$Var(X)=E(X^2)-[E(X)]^2=\sum_i x_i^2 f(x_i)-p^2=p-p^2 \qquad (6.2)$$

$$=p(1-p)=pq.$$

Παρατηρούμε εδώ ότι η κατανομή Bernoulli εξαρτάται μόνο από μία παράμετρο, την πιθανότητα p. Ακόμη, σε μία δοκιμή **Bernoulli** το αποτέλεσμα A στην Στατιστική καλείται και «*επιτυχία*», χωρίς να σημαίνει απαραίτητα ότι το αποτέλεσμα αυτό είναι επιθυμητό στην πράξη.

Παράδειγμα 6.1

Η ρίψη ενός ζαριού είναι μία δοκιμή Bernoulli εάν ενδιαφερόμαστε για παράδειγμα μόνο για την ένδειξη του *Άσσου*. Τα δυνατά γεγονότα είναι A={*ένδειξη Άσσος*} και $\overline{A}$={*οιαδήποτε άλλη ένδειξη*}. Αν υποθέσουμε ότι έχουμε κανονικό ζάρι σε μία δοκιμή η πιθανότητα *«επιτυχίας»* είναι $p=P(A)=1/6$ και η πιθανότητα «*αποτυχίας*» (άλλος αριθμός) είναι $q=5/6$.

Εάν το ζάρι δεν είναι κανονικό έχουμε πάλι δοκιμή Bernoulli, αλλά δεν γνωρίζουμε ποια είναι η πιθανότητα «επιτυχίας» p.

Παράδειγμα 6.2

Σε μία διαδικασία παραγωγής το 10% των τεμαχίων του προϊόντος είναι ελαττωματικό και το 90% είναι καλό. Ο έλεγχος σε κάθε τεμάχιο παραγωγής είναι μία δοκιμή Bernoulli με πιθανότητα» *«επιτυχίας* (ελαττωματικό) p=0,10 και *«αποτυχίας»* q=0,90.

Ακόμη, μία σειρά από δοκιμές Bernoulli αποτελεί μία διαδικασία ή ακολουθία Bernoulli, η οποία παρουσιάζει ιδιαίτερο ρόλο σε πολλά προβλήματα της Στατιστικής

Ορισμός. Μία σειρά από n δοκιμές Bernoulli λέγεται **διαδικασία ή ακολουθία Bernoulli** εάν

- σε κάθε δοκιμή έχουμε σταθερή πιθανότητα *«επιτυχίας»* p,
- οι n δοκιμές είναι στατιστικώς ανεξάρτητες

Έστω μία ακολουθία Bernoulli με n ανεξάρτητες δοκιμές, π.χ.,

$$AA\overline{A}\,AAA\overline{A}\;\overline{A}\,AA\overline{A}\,....A$$

όπου έχουμε ακριβώς *κ επιτυχίες* (*κ* φορές *A*) και κατά συνέπεια ακριβώς *n-κ αποτυχίες (n-κ* φορές $\overline{A}$*).* Η πιθανότητα να συμβεί αυτή η συγκεκριμένη **ακολουθία Bernoulli** είναι

$$P(AA\overline{A}\,AAA\overline{A}\;\overline{A}\,AA\overline{A}\,....A)=P(A)P(A)P(\overline{A})P(A)P(A)P(A)P(\overline{A})P(\overline{A})P(A)...$$
$$= p^{\kappa}(1-p)^{n-\kappa},$$

εφαρμόζοντας τον πολλαπλασιαστικό κανόνα και λαμβάνοντας υπ' όψη ότι οι δοκιμές είναι ανεξάρτητες.

6.2 Η ΔΙΩΝΥΜΙΚΗ ΚΑΤΑΝΟΜΗ

Σε μία δοκιμή Bernoulli ο δειγματοχώρος του πειράματος τύχης είναι $S=\{A, \overline{A}\}$. Εάν εκτελέσουμε το πείραμα Bernoulli n φορές προκύπτει ένας σύνθετος δειγματικός χώρος με όλες τις δυνατές ακολουθίες $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ (υπάρχουν 2^n τέτοιες ακολουθίες), όπου κάθε α_i παριστάνει το αποτέλεσμα A ή $\overline{A}$ στην i δοκιμή Bernoulli. Η τυχαία μεταβλητή X η οποία δηλώνει πόσες «επιτυχίες» (πόσες φορές το A) υπάρχουν σε κάθε ακολουθία n δοκιμών Bernoulli είναι γνωστή στην Στατιστική με το όνομα της Διωνυμικής μεταβλητής.

Ορισμός Μία τυχαία μεταβλητή X λέγεται **Διωνυμική μεταβλητή** αν πληρούνται οι ακόλουθες τρείς συνθήκες:

1) Υπάρχουν n επαναλαμβανόμενες **δοκιμές Bernoulli** οι οποίες είναι στατιστικά ανεξάρτητες.
2) Όλες οι δοκιμές έχουν την ίδια **σταθερή πιθανότητα** «επιτυχίας», $P(A)=\boldsymbol{p}$.
3) Η τυχαία μεταβλητή X ορίζεται ώς ακολούθως:
   ***X*={αριθμός επιτυχιών (φορές το γεγονός *A*) στις n δοκιμές Bernoulli }.**

Προφανώς, η Διωνυμική μεταβλητή X έχει πεδίο τιμών το σύνολο {0, 1, 2, …, n}. Ισοδύναμα λέμε ότι η X ακολουθεί την Διωνυμική κατανομή πιθανότητας, η οποία προκύπτει με βάση την πιθανότητα μίας συγκεκριμένης ακολουθίας n δοκιμών Bernoulli.

Θεωρούμε n επαναλαμβανόμενες δοκιμές Bernoulli και επιθυμούμε να υπολογίσουμε την πιθανότητα ότι η Διωνυμική μεταβλητή X θα πάρει μία τιμή ,ας πούμε κ, $P(X=\kappa)$. Ένα αποτέλεσμα του πειράματος το οποίο οδηγεί την X στην τιμή κ είναι, για παράδειγμα, στις πρώτες κ δοκιμές να πραγματοποιείται το A και στις υπόλοιπες n-κ να έχουμε συνεχώς $\overline{A}$,

$$\underbrace{A\;A\;A\ldots A}_{\kappa\ \text{φορές}}\;\underbrace{\overline{A}\;\overline{A}\;\overline{A}\ldots\overline{A}}_{n-\kappa\ \text{φορές}}$$

Αφού οι δοκιμές είναι στατιστικώς ανεξάρτητες, η πιθανότητα να συμβεί αυτή η συγκεκριμένη ακολουθία είναι

$$p^{\kappa}(1-p)^{n-\kappa}\,.$$

Αλλά με την ίδια πιθανότητα υπάρχει και άλλη ακολουθία δοκιμών όπου κ φορές συμβαίνει το A και n-κ συμβαίνει το $\overline{A}$. Ακόμη, ο ολικός αριθμός τέτοιων ακολουθιών είναι οι συνδυασμοί n πραγμάτων ανά κ,

$$\binom{n}{\kappa} = \frac{n!}{(n-\kappa)!\kappa!},$$

διότι είναι σαν να επιλέγουμε κ θέσεις (συνολικά από τις n) για να τοποθετήσουμε τα A. Ο αριθμός των συνδυασμών αυτών εκφράζει όλους τους δυνατούς τρόπους με τους οποίους συμβαίνει το γεγονός (X=κ), και επειδή όλοι οι τρόποι αυτοί είναι αμοιβαίως αποκλειόμενοι έχουμε

$$P(X=\kappa) = \binom{n}{\kappa} p^{\kappa}(1-p)^{n-\kappa} \tag{6.3}$$

Για να γενικεύσουμε, η παραπάνω συνάρτηση (6.3) είναι η Διωνυμική συνάρτηση μάζας πιθανότητας για να κατορθώσουμε κ *«επιτυχίες» σε* n *ανεξάρτητες δοκιμές, ενός πειράματος τύχης με πιθανότητα «επιτυχίας»* p *σε κάθε δοκιμή.*

Παρατηρήσεις

1. Για να επιβεβαιώσουμε ότι η παραπάνω συνάρτηση (6.3) είναι πράγματι συνάρτηση κατανομής πιθανότητας παρατηρούμε ότι

$$\sum_{0}^{n} P(X=\kappa) = \sum_{0}^{n} \binom{n}{\kappa} p^{\kappa}(1-p)^{n-\kappa} = [p+(1-p)]^{n} = 1^{n} = 1 \tag{6.4}$$

όπως θα έπρεπε είναι. Αφού οι πιθανότητες της συνάρτησης (6.3) επιτυγχάνονται και με το ανάπτυγμα της **διωνυμικής έκφρασης**

$$[p+(1-p)]^{n},$$

η συνάρτηση κατανομής πιθανότητας ονομάζεται **Διωνυμική** κατανομή.

2. Εάν ο αριθμός δοκιμών n είναι μικρός, οι πιθανότητες της Διωνυμικής κατανομής είναι εύκολο να υπολογισθούν. Στην περίπτωση όπου το n είναι μεγάλο, το παραγοντικό του n, $n!$, είναι δύσκολο να υπολογισθεί, για αυτό χρησιμοποιούμε τον τύπο του ***Stirling,***

$$n! \cong \sqrt{2\pi n} \cdot n^{n} e^{-n}.$$

Παράδειγμα 6.3

Σε ένα μεγάλο κιβώτιο υπάρχουν 12 τηλεοράσεις εκ των οποίων οι 3 είναι ελαττωματικές. Εάν επιλέξουμε τυχαία τρεις από αυτές με επανατοποθέτηση, ποια η πιθανότητα ότι ακριβώς μία είναι ελαττωματική; το πολύ μία είναι ελαττωματική;

Η επιλογή με επανατοποθέτηση εξασφαλίζει ότι σε κάθε επιλογή (δοκιμή Bernoulli) η πιθανότητα η τηλεόραση να είναι ελαττωματική (πιθανότητα «επιτυχίας»)είναι σταθερή και ίση με p=3/12=0,25. Έτσι μπορούμε να θεωρήσουμε ότι έχουμε

- n=3, τρεις ανεξάρτητες δοκιμές,
- p=0,25, σταθερή πιθανότητα «επιτυχίας»,
- X={ αριθμός των ελαττωματικών τηλεοράσεων στις 3 επιλογές(αριθμός «επιτυχιών» σε 3 δοκιμές)}.

Χρησιμοποιώντας την Διωνυμική κατανομή, έχουμε

$$P(X=1)=\binom{3}{1}0{,}25^1 0{,}75^2=0{,}42$$

$$P(X\le 1)=P(X=0)+P(X=1)=\binom{3}{0}0{,}25^0 0{,}75^3+\binom{3}{1}0{,}25^1 0{,}75^2$$

$$=0{,}75^3+0{,}42\cong 0{,}84$$

■

Η **αθροιστική συνάρτηση** κατανομής σε μία Διωνυμική μεταβλητή X, είναι παρόμοια με κάθε άλλη αθροιστική σε μία διακριτή μεταβλητή, και παριστάνει την πιθανότητα μέχρι και x «επιτυχίες» σε μία σειρά από n ανεξάρτητες δοκιμές,

$$F(x)=P(X\le x)=\sum_{i=0}^{x}\binom{n}{i}p^i(1-p)^{n-i} \qquad (6.4)$$

Παράδειγμα 6.4

Κάθε δείγμα από τον αέρα έχει πιθανότητα 0,10 να περιέχει ένα *σπάνιο μόριο.* Υποτίθεται ότι τα δείγματα είναι ανεξάρτητα όσο αφορά την παρουσία του *σπάνιου μορίου.* Να βρεθεί η πιθανότητα ότι στα επόμενα 18 δείγματα, ακριβώς 2 περιέχουν το σπάνιο μόριο.

Αν X είναι η τυχαία μεταβλητή η οποία παριστάνει τον αριθμό δειγμάτων που περιέχουν το σπάνιο μόριο στα επόμενα 18 δείγματα που θα ελεγχθούν, τότε η X είναι διωνυμική μεταβλητή με παραμέτρους n=18 και p=0,10. Έτσι έχουμε,

$$P(X=2)=\binom{18}{2}0{,}10^2(1-0{,}10)^{16}=153(0{,}10)^2(0{,}90)^{16}=0{,}284.$$

Προσδιορίστε την πιθανότητα ότι τουλάχιστον τρία δείγματα περιέχουν το *σπάνιο μόριο* .Η ζητούμενη πιθανότητα είναι

$$P(X\geq 3)=\sum_{x=3}^{18}\binom{18}{x}0{,}10^x 0{,}90^{18-x}\ .$$

Παρόλα αυτά, η παραπάνω πιθανότητα μπορεί πιο εύκολα να υπολογισθεί αν αφαιρέσουμε την πιθανότητα του συμπληρωματικού γεγονότος από την μονάδα,

$$P(X\geq 3)=1-P(X\leq 2)=1-F(2)=1-\sum_{x=0}^{2}\binom{18}{x}0{,}1^x 0{,}9^{18-x}$$
$$=1-(0{,}150+0{,}300+0{,}284)=1-0{,}734=0{,}166$$

Να υπολογισθεί η πιθανότητα ότι αριθμός δειγμάτων που περιέχουν το *σπάνιο μόριο* είναι μεταξύ 3 και 5. Έχουμε

$$P(3\leq X\leq 5)=P(2<X\leq 5)=F(5)-F(2).$$

Μπορεί να υπολογισθεί ευκολότερα αθροίζοντας,

$$P(3\leq X\leq 5)=\sum_{x=3}^{5}\binom{18}{x}0{,}10^x 0{,}90^{18-x}=0{,}168+0{,}070+0{,}022=0{,}260.$$

Μία διωνυμική διαδικασία μπορεί να θεωρηθεί το άθροισμα n ανεξάρτητων μεταβλητών Bernoulli. Πιο αναλυτικά, μία διωνυμική μεταβλητή X, η οποία παριστάνει τον αριθμό «επιτυχιών» σε n επαναλαμβανόμενες στατιστικώς ανεξάρτητες δοκιμές, μπορεί να θεωρηθεί σαν το άθροισμα

$$X=X_1+X_2+\ldots+X_n,$$

όπου X_i είναι μία Bernoulli μεταβλητή με μέση τιμή p. Έτσι, η **μέση τιμή** μίας διωνυμικής μεταβλητής μπορεί να υπολογιστεί από το άθροισμα των μέσων τιμών των n Bernoulli μεταβλητών. Δηλαδή, έχουμε

$$\boldsymbol{E(X)}=E(X_1)+E(X_2)+\ldots+E(X_n)=p+p+\ldots+p=\boldsymbol{np}, \tag{6.5}$$

εφαρμόζοντας την (Δ) ιδιότητα της μέσης τιμής στο άθροισμα ανεξάρτητων μεταβλητών.

Επί πλέον, αφού οι μεταβλητές Bernoulli είναι ίδιες και ανεξάρτητες, οι διακυμάνσεις τους είναι αθροισιμες. Δηλαδή η διωνυμική σαν άθροισμα n Bernoulli μεταβλητών, πρέπει να έχει σαν **διακύμανση** το άθροισμα των διακυμάνσεών των. Δηλαδή, έχουμε

$$\begin{aligned}\boldsymbol{Var(X)}&=Var(X_1)+Var(X_2)+\ldots+Var(X_n)=p(1-p)+p(1-p)+\ldots+p(1-p)\\&=\boldsymbol{np(1-p)}.\end{aligned} \tag{6.6}$$

Παρατήρηση στην ορολογία. Η μέση τιμή και διακύμανση εξαρτώνται μόνο από τις παραμέτρους n, p. Γενικά, εάν X είναι μία διωνυμική τυχαία μεταβλητή με παραμέτρους (n, p), τότε λέμε ότι η X ακολουθεί διωνυμική κατανομή με παραμέτρους (n, p).

Παράδειγμα 6.5

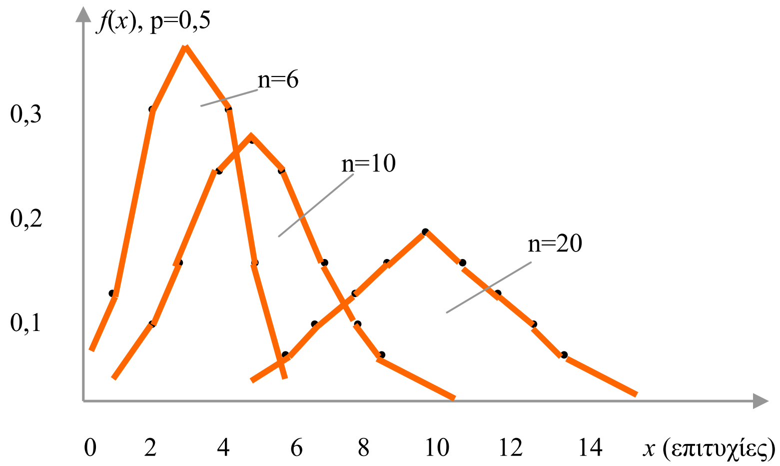
Η πιθανότητα ότι ένα bit, το οποίο στέλνεται από ένα κανάλι, λαμβάνεται εσφαλμένα είναι 0,05. Επίσης, είναι γνωστό ότι οι αποστολές των bit είναι ανεξάρτητες. Να ευρεθεί ο αναμενόμενος αριθμός των λανθασμένων bit που λαμβάνονται σε μία μετάδοση ενός κειμένου 3000 bit από το παραπάνω κανάλι. Ακόμη να ευρεθεί η τυπική απόκλιση του αριθμού των λανθασμένων bit από κείμενο σε κείμενο όταν όλα τα κείμενα που μεταδίδονται είναι τα ίδια.

Μπορούμε να υποθέσουμε ότι έχουμε 3000 ανεξάρτητες δοκιμές Bernoulli, όπου ελέγχουμε αν κάθε bit ελήφθη σωστά ή λανθασμένα. Η τυχαία μεταβλητή X, η οποία παριστάνει τον αριθμό των λανθασμένων bit σε όλο το κείμενο, είναι διωνυμική με παραμέτρους n=3000 και p=0,05. Ουσιαστικά ζητούμε την μέση τιμή και την διακύμανση της X. Οπότε, σύμφωνα με τους παραπάνω κανόνες έχουμε,

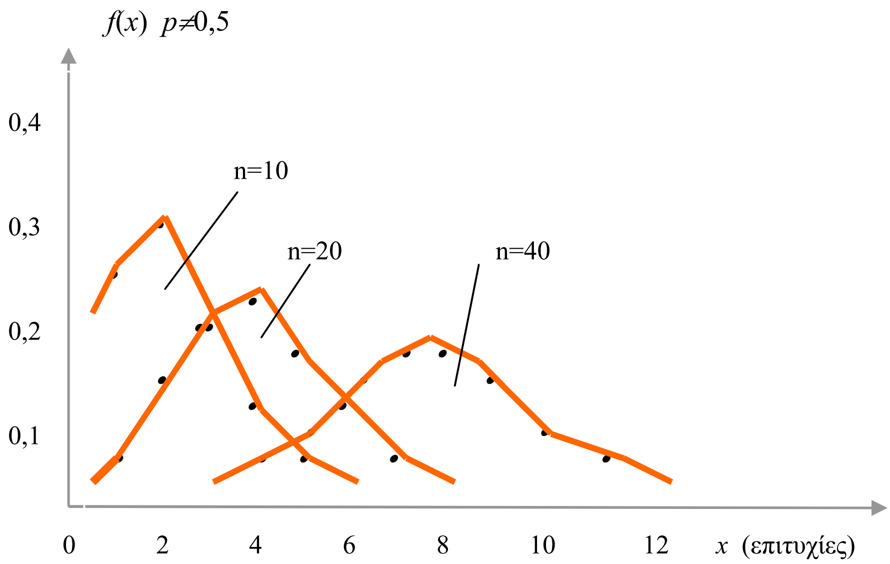
$$\begin{aligned}&E(X)=np=3000\cdot 0{,}05=150\\&Var(X)=np(1-p)=3000\cdot 0{,}05\cdot 0{,}95=142{,}5\\&\Rightarrow\quad \sigma_X=\sqrt{142{,}5}=11{,}937\approx 12.\end{aligned}$$

Έτσι, αναμένουμε σε κάθε κείμενο που λαμβάνεται τα 150 bit να είναι λανθασμένα. Βέβαια, από κείμενο σε κείμενο υπάρχει μία απόκλιση στον αριθμό των λανθασμένων bit, το μέγεθος της οποίας δίνεται από την τυπική απόκλιση της X που βρέθηκε περίπου 12 bit.

Πριν ολοκληρώσουμε την ανάλυσή μας στη διωνυμική κατανομή, πρέπει



Σχήμα 6.1. Γραφικές παραστάσεις διωνυμικής κατανομής, p=0,5, n=6, n=10, n-20.



Σχήμα 6.2. Γραφικές παραστάσεις διωνυμικής κατανομής, p≠0,5, n=10, n=20, n=40.

να παρατηρήσουμε ότι η διωνυμική για p=0,5 είναι συμμετρική και ασύμμετρη

να παρατηρήσουμε ότι η διωνυμική για p=0,5 είναι συμμετρική και ασύμμετρη ταν $p \neq$,0,5. Η ασυμμετρία σε μία διωνυμική κατανομή, ανεξάρτητα από την τιμή του p, γίνεται ασθενέστερη καθώς αυξάνει το n. Κάποιες από αυτές τις ιδιότητες διαφαίνονται στο Σχήμα 6.1, 6.2.

6.3 Η ΓΕΩΜΕΤΡΙΚΗ ΚΑΤΑΝΟΜΗ

Η κατανομή αυτή ταιριάζει στις περιπτώσεις όπου εκτελείται ένα πείραμα και μας ενδιαφέρει αν πραγματοποιείται ή δεν πραγματοποιείται το γεγονός A. Υποθέτουμε, όπως και στην περίπτωση της διωνυμικής κατανομής, ότι το πείραμα ή η δοκιμή Bernoulli επαναλαμβάνεται ανεξάρτητα η μία από την άλλη, πάντα με σταθερή πιθανότητα «επιτυχίας» , $P(A)=p$ και $P(\overline{A})=1-p=q$. Εδώ, οι δοκιμές επαναλαμβάνονται μέχρι πραγματοποίησης για πρώτη φορά του A. Ο αριθμός επανάληψης του πειράματος ή δοκιμών στην διωνυμική ήταν προκαθορισμένος, ενώ εδώ στην γεωμετρική ο αριθμός δοκιμών είναι η τυχαία μεταβλητή.

Ορισμός Μία τυχαία μεταβλητή X λέγεται ότι έχει **γεωμετρική κατανομή** αν πληρούνται οι ακόλουθες τρείς συνθήκες:

1) Υπάρχουν επαναλαμβανόμενες δοκιμές Bernoulli οι οποίες είναι στατιστικά **ανεξάρτητες**.
2) Όλες οι δοκιμές έχουν την ίδια σταθερή πιθανότητα «επιτυχίας» ***P(A)=p***.
3) Η τυχαία μεταβλητή X ορίζεται ώς ακολούθως:
   ***X*={αριθμός δοκιμών μέχρι πραγματοποίησης για πρώτη φορά του γεγονότος *A* }.**

Οι τιμές οι οποίες μπορεί να πάρει η X σε μία γεωμετρική κατανομή είναι άπειρες αριθμήσιμες, $X=\{1, 2, 3, \ldots\}$. Προκειμένου η X να πάρει μία συγκεκριμένη τιμή κ, $X=\kappa$, είναι αναγκαίο και ικανό ότι στις πρώτες κ-1 δοκιμές συμβαίνει το γεγονός $\overline{A}$ και στην κ-οστή ακριβώς δοκιμή πραγματοποιείται το A, δηλαδή έχουμε,

$$p(\kappa) = P(X=\kappa) = P(\underbrace{\overline{A} \cap \overline{A} \cap \overline{A} \cap \ldots \cap \overline{A}}_{\kappa-1\ \varphi o\rho\varepsilon\varsigma} \cap A)$$
$$= P(\overline{A})P(\overline{A})P(\overline{A})\ldots P(\overline{A})P(A) = (1-p)^{\kappa-1} p. \qquad (6.7)$$

αφού οι δοκιμές είναι ανεξάρτητες. Η εξίσωση (6.7) αποτελεί και την συνάρτηση μάζας πιθανότητας για κ=1, 2, 3, … . Για να ελέγξουμε ότι η $p(\kappa)$ είναι πράγματι συνάρτηση μάζας πιθανότητας, παρατηρούμε ότι

$$\sum_{\kappa=1}^{\infty} p(\kappa) = p\sum_{\kappa=1}^{\infty}(1-p)^{\kappa-1} = p(1+q\ +q^2+\ldots) = p\frac{1}{1-q} = 1$$

και επειδή $p(\kappa) \geq 0$, σίγουρα η $p(\kappa)$ της εξίσωσης (6.7) είναι συνάρτηση μάζας πιθανότητας.

Ο αναμενόμενος αριθμός των επαναλαμβανόμενων δοκιμών μέχρι της πρώτης «επιτυχίας» είναι το αντίστροφο της σταθερής πιθανότητας p, όπως προκύπτει από τον ορισμό της μαθηματικής ελπίδας της γεωμετρικής μεταβλητής X,

$$E(X) = \sum_{\kappa=1}^{\infty} \kappa(1-p)^{\kappa-1} p = p\sum_{\kappa=1}^{\infty} \kappa q^{\kappa-1}$$
$$= p\sum_{\kappa=1}^{\infty} \frac{d}{dq} q^{\kappa} = p\frac{d}{dq}\sum_{\kappa=1}^{\infty} q^{\kappa} \qquad (6.8)$$
$$= p\frac{d}{dq}\left(\frac{q}{1-q}\right) = \frac{p}{(1-q)^2} = \frac{1}{p}.$$

Παρόμοια, μπορούμε να αποδείξουμε ότι και η διακύμανση της γεωμετρικής μεταβλητής X είναι

$$Var(X) = \frac{q}{p^2}. \qquad (6.9)$$

Περίοδος Επαναφοράς.

Ιδιαίτερο ενδιαφέρων για τους μηχανικούς παρουσιάζει η μέση τιμή της γεωμετρικής μεταβλητής, ***E*(*X*)**, η οποία είναι γνωστή και σαν **περίοδος επαναφοράς**, όταν η «επιτυχία» ή το γεγονός A όπως το έχουμε αναφέρει παριστάνει μία πλημμύρα, ένα ισχυρό σεισμό, ένα τυφώνα, μία ισχυρή καταιγίδα, ή και γενικά ένα καταστροφικό φαινόμενο. Στις περιπτώσεις αυτές σαν δοκιμή θεωρείται η πραγματοποίηση του καταστροφικού φαινομένου μέσα στην μονάδα του χρόνου, π. χ. ένα έτος ή ένας μήνας. Έτσι, η μέση τιμή της *X, E(X),* παριστάνει τον μέσο αριθμό μηνών ή χρόνων αντίστοιχα που μεσολαβούν ανάμεσα σε δύο διαδοχικά τέτοια φαινόμενα.

Με άλλα λόγια, η μέση τιμή της γεωμετρικής μεταβλητής ή η περίοδος επαναφοράς παριστάνει σε πόσο χρόνο αναμένεται να ξαναφανεί το γεγονός *A*.

Παράδειγμα 6.6

Μελετάται το ύψος μίας γέφυρας που κατασκευάζεται τώρα, έτσι ώστε να μην πλημμυρίζει το οδόστρωμά της από ισχυρές βροχοπτώσεις, οι οποίες ανυψώνουν την επιφάνεια του ποταμού. Όμως, το κόστος κατασκευής δεν επιτρέπει την σχεδίαση της γέφυρας έτσι ώστε να αντέξει στην **πλημμύρα των 50 ετών.** Κατασκευάζεται η γέφυρα,

- ποια η πιθανότητα ότι στον δέκατο ακριβώς χρόνο θα αστοχήσει η γέφυρα από μία τέτοια πλημμύρα;
- ποια η πιθανότητα ότι το πολύ μέσα σε 5 χρόνια θα καταστραφεί η γέφυρα από την μεγάλη πλημμύρα;

Υποθέτουμε ότι η *πλημμύρα των πενήντα ετών* εμφανίζεται κατά μέσο όρο κάθε 50 έτη (περίοδος επαναφοράς), μπορεί να πραγματοποιηθεί το πολύ μία φορά μέσα σε ένα χρόνο, και η εμφάνισή της είναι στατιστικώς ανεξάρτητη από τις προηγούμενες χρονιές. Έτσι σε κάθε χρόνο αντιστοιχεί μία δοκιμή Bernoulli όπου το γεγονός *A* («επιτυχία») είναι η *πλημμύρα των 50 ετών*. Με αυτές τις λογικές παραδοχές, η τυχαία μεταβλητή *X* που παριστάνει αριθμό δοκιμών ή χρόνων μέχρι εμφάνισης της μεγάλης πλημμύρας είναι γεωμετρική μεταβλητή. Η μέση τιμή της *X* είναι η περίοδος επαναφοράς της πλημμύρας, δηλαδή

$$E(X) = \frac{1}{p} = 50 \quad \Rightarrow \quad p = 0{,}02\,.$$

Εύκολα, υπολογίζουμε την πρώτη ζητούμενη πιθανότητα,

$$P(X = 10) = 0{,}98^9 0{,}02 = 0{,}0167\,.$$

Στην δεύτερη ερώτηση η γέφυρα θα καταστραφεί σε πέντε χρόνια αν η *X* πάρει μία τιμή μέχρι και το 5. Πρέπει να χρησιμοποιήσουμε την αθροιστική συνάρτηση κατανομής *F(x)* της γεωμετρικής μεταβλητής, η οποία εύκολα υπολογίζεται αθροίζοντας την συνάρτηση μάζας πιθανότητας ,

$$P(X \le 5) = F(5) = \sum_{x=1}^{5} (1-p)^{x-1} p = \sum_{x=1}^{5} 0{,}98^{x-1} 0{,}02 \quad = 0{,}096$$

Παράδειγμα 6.7

Σε μία περιοχή και κατά τους θερινούς μήνες Ιουλίου και Αυγούστου, η πιθανότητα ότι μία μέρα θα ξεσπάσει ισχυρή καταιγίδα είναι σταθερή και ίση με p=0,05. Υποθέτουμε ότι υπάρχει στατιστική ανεξαρτησία από μέρα σε μέρα. Ποια η πιθανότητα ότι για πρώτη φορά θα ξεσπάσει ισχυρή καταιγίδα στις 25 Ιουλίου;

Εάν X είναι η τυχαία μεταβλητή που παριστάνει των αριθμό ημερών, αρχίζοντας από την 1^{η} Ιουλίου, που θα περάσουν μέχρι να έχουμε για πρώτη φορά τέτοια καταιγίδα, μπορούμε να υποθέσουμε ότι η X είναι γεωμετρικής κατανομής. Άρα έχουμε

$$P(X = 25) = (1-p)^{25-1} p = 0{,}95^{24} \cdot 0{,}05 = 0{,}292 \cdot 0{,}05 = 0{,}0146$$

Πόσες μέρες κατά μέσο όρο μεσολαβούν ανάμεσα σε δύο συνεχόμενες ισχυρές καταιγίδες; Ενδιαφερόμαστε, δηλαδή, για την μέση τιμή της γεωμετρικής μεταβλητής X,

$$E(X) = \frac{1}{p} = \frac{1}{0{,}05} = 20 \text{ ημέρες.}$$

Εδώ μπορούμε να εκφρασθούμε και διαφορετικότερα, λέγοντας ότι η *περίοδος επαναφοράς* μίας τέτοιας καταιγίδας το καλοκαίρι είναι 20 ημέρες, ή αναμένουμε κάθε 20 ημέρες να έχουμε καταιγίδα.

Παράδειγμα 6.8

Υποτίθεται ότι το κόστος της εκτέλεσης ενός πειράματος είναι 100000 δρχμς. Εάν το πείραμα αποτύχει, ένα επί πλέον κόστος των 30000 δραχμών επιβαρύνει την επόμενη εκτέλεση του πειράματος. Εάν το πείραμα εκτελείται μέχρι να επιτύχει και εάν η κάθε εκτέλεση έχει σταθερή πιθανότητα επιτυχίας p=0,4, ποιο το αναμενόμενο κόστος της συνολικής διαδικασίας αν υπάρχει στατιστική ανεξαρτησία μεταξύ των εκτελέσεων του πειράματος.

Εάν X είναι η μεταβλητή που παριστάνει τον αριθμό εκτελέσεων του πειράματος μέχρι επιτυχίας του τότε το συνολικό κόστος της όλης διαδικασίας είναι K=100000·X+30000(X-1), διότι όταν το πείραμα εκτελείται συνολικά X φορές, υπάρχουν σίγουρα (X-1) αποτυχίες. Έτσι το αναμενόμενο κόστος είναι

$$E(K) = E[100000X + 30000(X-1)] = 100000E(X) + 30000E(X) - 30000 =$$
$$= 100000\frac{1}{0{,}4} + 30000\frac{1}{0{,}4} - 30000 = 295000 \quad \delta\rho\chi\mu\varsigma$$

Έλλειψη μνήμης

Μία γεωμετρική τυχαία μεταβλητή παριστάνει τον αριθμό δοκιμών μέχρι την πρώτη «επιτυχία». Παρόλα αυτά, επειδή οι δοκιμές είναι ανεξάρτητες, η μέτρηση του αριθμού των δοκιμών μέχρι την επόμενη «επιτυχία» μπορεί να αρχίσει από οποιαδήποτε δοκιμή χωρίς να αλλάξει την κατανομή πιθανότητας της μεταβλητής. Για παράδειγμα, στην περίπτωση με τις καταιγίδες (παράδειγμα 6.6), η πιθανότητα να ξεσπάσει καταιγίδα σε 8 ημέρες ακριβώς είναι

$$(1-p)^7 p = 0{,}95^7 0{,}05 = \ 0{,}0349.$$

Αυτή η πιθανότητα δεν εξαρτάται από τον αριθμό ημερών που περάσανε χωρίς να ξεσπάσει καταιγίδα. Δηλαδή, αν βρισκόμαστε στην 1η Ιουλίου η πιθανότητα να ξεσπάσει καταιγίδα για πρώτη φορά στις 8 Ιουλίου είναι 0,0349, εάν βρισκόμαστε στις 10 Ιουλίου η πιθανότητα να ξεσπάσει καταιγίδα σε 8 μέρες δηλαδή την 18η Ιουλίου είναι πάλι 0,0349 ανεξάρτητα από το πόσες καταιγίδες συνέβηκαν πριν τις 10 Ιουλίου.

Γενικά, η **ιδιότητα έλλειψης μνήμης** της γεωμετρικής κατανομής διατυπώνεται και ως εξής: Εάν το γεγονός A δεν συνέβη μέχρι την κ δοκιμή, τότε η πιθανότητα ότι δεν θα συμβεί και στις επόμενες λ δοκιμές, είναι ίδια με την πιθανότητα ότι το γεγονός A δεν θα συμβεί στις πρώτες λ δοκιμές. Δηλαδή, ισχύει

$$P(X > \kappa + \lambda | X > \kappa) = P(X > \lambda). \tag{6.10}$$

απόδειξη:

$$P(X > \kappa + \lambda | X > \kappa) = \frac{P(\{X > \kappa + \lambda\} \cap \{X > \kappa\})}{P(X > \kappa)} = \frac{P(X > \kappa + \lambda)}{P(X > \kappa)}$$
$$= \frac{(1-p)^{\kappa+\lambda}}{(1-p)^{\kappa}} = (1-p)^{\lambda} = P(X > \lambda).$$

6.4 Η ΑΡΝΗΤΙΚΗ ΔΙΩΝΥΜΙΚΗ ΚΑΤΑΝΟΜΗ (PASCAL)

Μία γενίκευση της γεωμετρικής μεταβλητής είναι η Αρνητική Διωνυμική *Pascal* μεταβλητή *X*, η οποία παριστάνει τον αριθμό δοκιμών που απαιτούνται μέχρι να συμπληρωθούν ακριβώς για πρώτη φορά ***r*** εμφανίσεις του γεγονότος *A* (*r* «επιτυχίες»).

Ορισμός Μία τυχαία μεταβλητή *X* ακολουθεί **Αρνητική Διωνυμική** ή **Pascal κατανομή** αν πληρούνται οι ακόλουθες τρείς συνθήκες:

1) Υπάρχουν επαναλαμβανόμενες δοκιμές Bernoulli οι οποίες είναι στατιστικά ανεξάρτητες.
2) Όλες οι δοκιμές έχουν την ίδια σταθερή πιθανότητα «επιτυχίας» ***P*(*A*)=*p*.**
3) Η τυχαία μεταβλητή *X* ορίζεται ώς ακολούθως: ***X*={αριθμός δοκιμών μέχρι να πραγματοποιηθεί *r* φορές το γεγονός *A*.}.**

Επειδή τουλάχιστον *r* δοκιμές απαιτούνται για να φθάσουμε τις *r επιτυχίες,* το πεδίο τιμών της *X* είναι από το *r* μέχρι το ∝. Είναι προφανές ότι αν *r*=1, η Pascal μεταβλητή *X* συμπίπτει με την γεωμετρική μεταβλητή και ακολουθεί την γεωμετρική κατανομή (6.7).

Στην γενική περίπτωση {*X*=*x*} εάν και μόνον εάν το *A* συμβαίνει στην *x*-οστή δοκιμή και επίσης το *A* συνέβη *r*-1 φορές στις προηγούμενες *x*-1 δοκιμές. Η πιθανότητα να συμβεί το {*X*=*x*} είναι η πιθανότητα να συμβεί η τομή των δύο τελευταίων γεγονότων, και επειδή είναι ανεξάρτητα έχουμε,

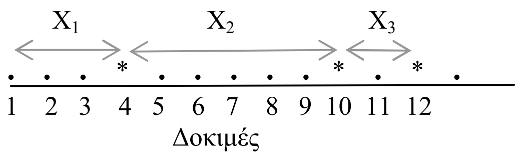
$$P(X=x)=\underbrace{\binom{x-1}{r-1}p^{r-1}(1-p)^{x-1-(r-1)}}_{\substack{\text{πιθανότητα εμφάνισης του A r-1}\\ \text{φορές στις x-1 δοκιμές (διωνυμική)}}}\cdot p=\binom{x-1}{r-1}p^{r}(1-p)^{x-r} \qquad (6.11)$$

Η εξίσωση (6.11) ισχύει προφανώς για τιμές της *X* ={*r*, *r*+1, *r*+2, … } και αποτελεί την **συνάρτηση μάζας πιθανότητας *f(x)*** της μεταβλητής **Pascal**, διότι

$$\sum_{x=r}^{\infty} f(x)=\sum_{x=r}^{\infty}\binom{x-1}{r-1}p^{r}(1-p)^{x-r}=p^{r}\sum_{x=r}^{\infty}\binom{x-1}{r-1}(1-p)^{x-r}=p^{r}p^{-r}=1$$

Μπορούμε να θεωρήσουμε ότι η γέννηση της αρνητικής διωνυμικής κατανομής οφείλεται στην ιδιότητα *έλλειψη μνήμης* της γεωμετρικής κατανομής. Ας υποθέσουμε ότι η X δηλώνει τον αριθμό δοκιμών μέχρι την *r επιτυχία.* Ακόμη, ας υποθέσουμε ότι X_1 δηλώνει τον αριθμό δοκιμών μέχρι την πρώτη επιτυχία, X_2 δηλώνει τον αριθμό των επιπλέον δοκιμών που απαιτούνται μέχρι την δεύτερη επιτυχία, X_3 δηλώνει τον αριθμό των επιπλέον δοκιμών που απαιτούνται μέχρι την τρίτη επιτυχία, και ούτω καθεξής. Τότε ο ολικός αριθμός δοκιμών που απαιτούνται μέχρι την r επιτυχία είναι

$$X = X_1 + X_2 + \ldots + X_r$$



*: *δείχνει ότι η δοκιμή αυτή καταλήγει σε επιτυχία*

Σχήμα 6.3 Αρνητική διωνυμική τυχαία μεταβλητή σαν ένα άθροισμα τυχαίων μεταβλητών γεωμετρικής κατανομής

Λόγω της ιδιότητας έλλειψης μνήμης, κάθε μία από τις τυχαίες μεταβλητές $X_1, X_2, \ldots, X_r$ έχει γεωμετρική κατανομή με την ίδια τιμή της πιθανότητας p. Κατά συνέπεια, μία αρνητική διωνυμική τυχαία μεταβλητή μπορεί να ερμηνευθεί σαν το άθροισμα από r γεωμετρικές τυχαίες μεταβλητές. Η ερμηνεία αυτή παριστάνεται στο Σχήμα 6.3.

Παρατήρηση

Υπενθυμίζουμε ότι η διωνυμική τυχαία μεταβλητή είναι μία μέτρηση του αριθμού *επιτυχιών* σε n δοκιμές Bernoulli. Δηλαδή ο αριθμός δοκιμών προσδιορίζεται από την αρχή, και ο αριθμός επιτυχιών είναι στοχαστικό μέγεθος. Μία αρνητική διωνυμική μεταβλητή είναι η μέτρηση του αριθμού των δοκιμών που απαιτούνται για να φθάσουμε τις *r επιτυχίες.* Δηλαδή, ο αριθμός επιτυχιών προσδιορίζεται από την αρχή, ενώ ο αριθμός δοκιμών είναι στοχαστικό μέγεθος. Με αυτή την έννοια, μία αρνητική διωνυμική τυχαία μεταβλητή μπορεί να θεωρηθεί το αντίθετο , ή το *αρνητικό*, από μία δυωνυμική τυχαία μεταβλητή. ■

Η περιγραφή της αρνητικής διωνυμικής τυχαίας μεταβλητής σαν ένα άθροισμα τυχαίων μεταβλητών γεωμετρικής κατανομής οδηγεί στα ακόλουθα αποτελέσματα σχετικά με την μέση τιμή και την διακύμανση.

Εάν X είναι τυχαία μεταβλητή **αρνητικής διωνυμικής** κατανομής με παραμέτρους, p και r, τότε η **μέση τιμή** και **διακύμανση** της **X** είναι

$$\boldsymbol{\mu = E(X) =} \frac{r}{p} \quad \text{και} \quad \boldsymbol{\sigma^2 = Var(X)} = \frac{r(1-p)}{p^2} \tag{6.12}$$

Παράδειγμα 6.9

Ένα αεροπλάνο υψηλής αξιοπιστίας περιέχει τρεις ίδιους ηλεκτρονικούς υπολογιστές. Μόνο ο ένας χρησιμοποιείται για την λειτουργία του αεροπλάνου, οι άλλοι δύο είναι διαθέσιμοι στην περίπτωση βλάβης του αρχικού. Η πιθανότητα αποτυχίας του υπολογιστή για μία ώρα λειτουργίας του είναι 0,0005. Με την υπόθεση ότι κάθε ώρα λειτουργίας παριστάνει και μία δοκιμή Bernoulli, ποιος είναι ο μέσος χρόνος λειτουργίας του αεροπλάνου;

Εάν X παριστάνει τον αριθμό ωρών μέχρι να αποτύχουν και οι τρεις υπολογιστές, η X είναι τυχαία μεταβλητή αρνητικής διωνυμικής κατανομής με p=0,0005 και r=3. Κατά συνέπεια,

$$E(X)=3/0{,}0005=6000 \text{ ώρες.}$$

Ποια είναι η πιθανότητα ότι όλοι οι υπολογιστές αποτυγχάνουν σε μία πτήση 6 ωρών; Η απαιτούμενη πιθανότητα είναι $P(X \le 6)$ και έχουμε

$$\begin{aligned} P(X \le 6) &= P(X=3)+P(X=4)+P(X=5)+P(X=6) \\ &= 0{,}0005^3 + \binom{3}{2}0{,}0005^3(0{,}9995) + \binom{4}{2}0{,}0005^3(0{,}9995)^2 + \binom{5}{2}0{,}0005^3(0{,}9995)^3 \\ &= 1{,}25\cdot 10^{-10} + 3{,}75\cdot 10^{-10} + 7{,}49\cdot 10^{-10} + 15\cdot 10^{-10} = 27{,}49\cdot 10^{-10} \end{aligned}$$

6.5 Η ΥΠΕΡΓΕΩΜΕΤΡΙΚΗ ΚΑΤΑΝΟΜΗ

Το μοντέλο της υπεργεωμετρικής κατανομής εφαρμόζεται σε δειγματοληψία χωρίς αντικατάσταση από ένα σύνολο πεπερασμένου πληθυσμού του οποίου τα στοιχεία μπορούν να χωρισθούν σε δύο κατηγορίες. Θεωρούμε ένα πληθυσμό από N στοιχεία, εκ των οποίων τα κ κατέχουν μία συγκεκριμένη ιδιότητα και N-κ δεν κατέχουν τέτοια ιδιότητα. Προφανώς, σε μία τυχαία επιλογή (δοκιμή) ενός στοιχείου από έναν τέτοιο πληθυσμό, το αποτέλεσμα θα είναι είτε ένα από τα κ («επιτυχία») ή ένα από τα N-κ(«αποτυχία»). Είναι επίσης ξεκάθαρο ότι εάν n τυχαίες επιλογές γίνονται χωρίς αντικατάσταση από τον πληθυσμό, κάθε επιλογή (δοκιμή) είναι εξαρτώμενη από τις προηγούμενές της και η πιθανότητα «επιτυχίας» αλλάζει σε κάθε επιλογή. Κάτω από αυτές τις συνθήκες, μπορεί να επιθυμούμε να βρούμε την πιθανότητα να επιτύχουμε x στοιχεία τύπου κ (να έχουμε x «επιτυχίες»), σε ένα τυχαίο δείγμα μεγέθους n.

Ορισμός. ***Θεωρούμε ένα πεπερασμένο πληθυσμό*** N ***στοιχείων, εκ των οποίων τα*** κ ***κατέχουν μία συγκεκριμένη ιδιότητα. Σε ένα τυχαίο δείγμα*** n ***στοιχείων (*** **n** ***δοκιμών) χωρίς αντικατάσταση, καλούμε*** *υπεργεωμετρική μεταβλητή* X ***τον αριθμό των στοιχείων («επιτυχιών») στο δείγμα που κατέχουν την συγκεκριμένη ιδιότητα των*** **κ.**

Η υπεργεωμετρική μεταβλητή δημιουργείται από τις τρείς παραμέτρους N, κ***, και*** n.

Η συνάρτηση μάζας πιθανότητας για την υπεργεωμετρική μεταβλητή στηρίζεται μάλλον στην κλασσική μέθοδο ορισμού πιθανότητας. Ο ολικός αριθμός διαφορετικών τρόπων επιλογής n στοιχείων από ένα σύνολο N στοιχείων είναι οι συνδυασμοί

$$\binom{N}{n}.$$

Παρατηρούμε ότι ο αριθμός διαφορετικών επιλογών x στοιχείων από τα κ στοιχεία με την συγκεκριμένη ιδιότητα είναι ο ίδιος με τον αριθμό διαφορετικών επιλογών n-x στοιχείων από τα N-κ, τα οποία δεν κατέχουν την ιδιότητα αυτή. Έτσι ο συνολικός αριθμός διαφορετικών τρόπων επιλογής n στοιχείων, όπου τα x κατέχουν την ιδιότητα των κ και τα n-x δεν την κατέχουν, είναι

$$\binom{\kappa}{x}\binom{N-\kappa}{n-x}.$$

Τελικά, εφόσον η διαδικασία δειγματοληψίας είναι τυχαία, όλα τα διαφορετικά δείγματα των n στοιχείων είναι ισοπίθανα και ο δειγματικός χώρος έχει δηλαδή

$$\binom{N}{n}$$

ισοπίθανα δειγματοσημεία.

Έτσι η **συνάρτηση μάζας ποιθανότητας** σε μία **υπεργεωμετρική μεταβλητή *X*** είναι

$$P(X=x)=\frac{\binom{\kappa}{x}\binom{N-\kappa}{n-x}}{\binom{N}{n}}, \tag{6.13}$$

όπου η μεταβλητή *X* μπορεί να πάρει τιμές από *max*{*0, n*+*κ*-*N*} έως και *min* {*κ*,*n*}

Παράδειγμα 6.10

Ένα κιβώτιο περιέχει 100 βίδες εγχώριες και άλλες 200 ίδιου τύπου αλλά παραγόμενες σε άλλη χώρα στο εξωτερικό. Εάν τέσσερις βίδες επιλέγονται τυχαία από το κιβώτιο και χωρίς αντικατάσταση, ποια η πιθανότητα ότι όλες είναι εγχώριες;

Εάν X παριστάνει τον αριθμό εγχώριων βιδών που υπάρχουν στο δείγμα των τεσσάρων βιδών, τότε η τυχαία μεταβλητή X ακολουθεί υπεργεωμετρική κατανομή, και η ζητούμενη πιθανότητα είναι

$$P(X=4)=\frac{\binom{100}{4}\binom{200}{0}}{\binom{300}{4}}=0{,}0119\,.$$

Ποια είναι η πιθανότητα ότι τουλάχιστον δύο είναι εγχώριες;

$$P(X\geq 2)=\frac{\binom{100}{2}\binom{200}{2}}{\binom{300}{4}}+\frac{\binom{100}{3}\binom{200}{1}}{\binom{300}{4}}+\frac{\binom{100}{4}\binom{200}{0}}{\binom{300}{4}}$$
$$=0{,}298+0{,}098+0{,}0119$$

■

Η **μέση τιμή και διακύμανση** μιας υπεργεωμετρικής τυχαίας μεταβλητής μπορεί να προσδιορισθεί θεωρώντας τις δοκιμές που απαρτίζουν το πείραμα. Εντούτοις όμως, οι δοκιμές δεν είναι ανεξάρτητες, και έτσι οι υπολογισμοί που χρησιμοποιήθηκαν στην διωνυμική τυχαία μεταβλητή δεν είναι κατάλληλοι.

Για να εξάγουμε την **μέση τιμή** μίας υπεργεωμετρικής μεταβλητής, παρατηρούμε ότι η υπεργεωμετρική μεταβλητή X μπορεί να θεωρηθεί σαν το άθροισμα n μεταβλητών X_i, οι οποίες δεν είναι ανεξάρτητες. Αλλά, επειδή η προσθετική ιδιότητα της μέσης τιμής δεν απαιτεί την ανεξαρτησία των X_i , ισχύει η εξίσωση

$$\begin{aligned} E(X) &= E(X_1) + E(X_2) + \ldots + E(X_n) \\ &= n\frac{\kappa}{N} \end{aligned} \tag{6.14}$$

διότι η μέση τιμή $E(X_i)$ είναι η πιθανότητα «επιτυχίας» στην i δοκιμή ίση με κ/N, όταν δεν είναι γνωστό τι συνέβη στις προηγούμενες δοκιμές (αυτό αφήνεται στον αναγνώστη σαν μία καλή εξάσκηση).

Χωρίς απόδειξη, παρόμοια μπορούμε να εξάγουμε και την **διακύμανση** μίας υπεργεωμετρικής μεταβλητής, η οποία είναι

$$Var(X) = \left(\frac{N-n}{N-1}\right) n \left(\frac{\kappa}{N}\right)\left(\frac{N-\kappa}{N}\right). \tag{6.15}$$

Παρατήρηση

Αν p=κ/N, τότε p ερμηνεύεται σαν πιθανότητα *επιτυχίας* στο σύνολο απ' όπου επιλέγεται το δείγμα, και έχουμε για την μέση τιμή και διακύμανση της X,

$$E(X) = n\frac{\kappa}{N} = np, \qquad Var(X) = \left(\frac{N-n}{N-1}\right) n\, p(1-p)\,.$$

Η μέση τιμή E(X) της υπεργεωμετρικής είναι παρόμοιο αποτέλεσμα. Επίσης, Var(X) διαφέρει από το αποτέλεσμα για μία διωνυμική τυχαία μεταβλητή μόνο ως προς τον όρο

$$\frac{N-n}{N-1},$$

γνωστός σαν **συντελεστής διόρθωσης πεπερασμένου πληθυσμού**. ■

Δειγματοληψία με επανάθεση

Δειγματοληψία με επανάθεση είναι ισοδύναμη με δειγματοληψία χωρίς επανάθεση από ένα άπειρο σύνολο διότι η αναλογία επιτυχίας, p=κ/N, παραμένει σταθερή για κάθε δοκιμή στο πείραμα. Όπως αναφέρθηκε παραπάνω, εάν η δειγματοληψία γίνεται με επανάθεση, τότε η μεταβλητή X είναι διωνυμική και η διακύμανσή της είναι np(1-p). Κατά συνέπεια, ο **συντελεστής διόρθωσης** παριστάνει την διόρθωση στην διωνυμική διακύμανση η οποία προκύπτει επειδή η δειγματοληψία γίνεται χωρίς επανάθεση από ένα πεπερασμένο σύνολο μεγέθους N.

Όταν ο πληθυσμός N είναι πολύ μεγάλος ή το μέγεθος του δείγματος n μικρό τότε το υπεργεωμετρικό μοντέλο προσεγγίζει το μοντέλο της διωνυμικής κατανομής. Γενικά, η προσέγγιση της υπεργεωμετρικής κατανομής από την διωνυμική είναι αρκετά καλή όταν

$$n/N \leq 0{,}1.$$

Παράδειγμα 6.11

Μία λίστα από το πελατολόγιο μίας μεγάλης εταιρίας περιέχει 1000 πελάτες. Από αυτούς οι 600 έχουν αγοράσει τουλάχιστον ένα προϊόν από την εταιρία τους τελευταίους τέσσερις μήνες. Για να σχεδιάσουμε ένα καινούργιο προϊόν ένα δείγμα από 40 πελάτες επιλέγεται τυχαία. Ποια η πιθανότητα ότι τουλάχιστον 35 από τους πελάτες του δείγματος έχουν αγοράσει τουλάχιστον ένα προϊόν από την εταιρία το τελευταίο τετράμηνο;

Η δειγματοληψία γίνεται χωρίς επανάθεση. Εάν X είναι η μεταβλητή η οποία παριστάνει τον αριθμό πελατών με την πρόσφατη αγορά μέσα στο δείγμα των 40 πελατών, τότε οι παράμετροι της υπεργεωμετρικής μεταβλητής είναι N=1000, n=40, και κ=600. Έτσι, μπορούμε να χρησιμοποιήσουμε την συνάρτηση μάζας πιθανότητας ή την αθροιστική κατανομή της υπεργεωμετρικής κατανομής, οπότε έχουμε,

$$P(X \geq 35) = \sum_{x=35}^{40} \frac{\binom{600}{x}\binom{400}{40-x}}{\binom{1000}{40}}$$

Επειδή το δείγμα είναι σχετικά μικρό ως προς το μέγεθος του πληθυσμού, n/N<0,10, μπορούμε να προσεγγίσουμε την ζητούμενη πιθανότητα με διωνυμική κατανομή θέτοντας p=κ/N=0,60.

$$P(X \geq 35) = \sum_{x=35}^{40} \binom{40}{x} 0{,}6^x (1-0{,}6)^{40-x}$$

6.6 ΔΙΑΔΙΚΑΣΙΑ POISSON

6.6.1 Κατανομή Poisson

Η διωνυμική μεταβλητή συμβολίζει πόσες φορές πραγματοποιείται ένα γεγονός *Α* σε *n* δοκιμές Bernoulli, η *Poisson* μεταβλητή την οποία μελετούμε σ' αυτήν την παράγραφο παριστάνει τον αριθμό πραγματοποίησης του *Α* σε ένα χρονικό ή τοπικό διάστημα.

Η *κατανομή* **Poisson** χρησιμοποιείται συχνά σε προβλήματα μηχανικού τα οποία εμπεριέχουν την πραγματοποίηση τυχαίων γεγονότων στο **συνεχές διάστημα χρόνου ή χώρου**. Για παράδειγμα, ο αριθμός

- ατυχημάτων σε ένα εργοστάσιο μέσα σε ένα έξι μήνες,
- σεισμών σε μία χώρα μέσα σε πέντε χρόνια,
- άφιξης αυτοκινήτων σε ένα μεγάλο πρατήριο καυσίμων μέσα σε μία ώρα
- βλαβών σε μία ηλεκτρικό σύστημα μέσα σε ένα μήνα λειτουργίας του,
- τηλεφωνικών κλίσεων σε ένα τηλεφωνικό κέντρο μέσα σε 20 λεπτά,
- ισχυρών καταιγίδων σε ένα τόπο το καλοκαίρι,
- εξυπηρέτησης πελατών την ημέρα σε ένα μεγάλο κατάστημα,

είναι διακριτή τυχαία μεταβλητή και παριστάνει πόσες φορές συμβαίνει ή εμφανίζεται ένα γεγονός *Α* σε κάποιο χρονικό διάστημα. Ακόμη, ο αριθμός

- τυπογραφικών λαθών σε μία σελίδα σε ένα βιβλίο,
- αυτοκινητιστικών ατυχημάτων σε ένα διάστημα 100 χιλιομέτρων
- ελαττωμάτων σε μία μεγάλη επιφάνεια υαλο-κατασκευής,
- ραγισμάτων ανά τετραγωνικό μέτρο σε μία μεταλλική επιφάνεια,

είναι διακριτή τυχαία μεταβλητή και παριστάνει πόσες φορές συμβαίνει ή εμφανίζεται ένα γεγονός *Α* σε κάποιο διάστημα μήκους ή επιφάνειας ή ακόμη και χώρου.

Τέτοια γεγονότα μπορούν να συμβούν τυχαία σε οποιοδήποτε σημείο του χρόνου ή του τόπου. Θα μπορούσε να μελετήσει κανείς αυτά τα γεγονότα με ακολουθίες Bernoulli αν χωρίσει τον χρόνο ή τον τόπο σε πολύ μικρά διαστήματα Δt, έτσι ώστε σε κάθε τέτοιο μικρό διάστημα να αντιστοιχεί μία δοκιμή Bernoulli, δηλαδή, αν συνέβη ή δεν συνέβη το γεγονός *A*. Έτσι ο αριθμός πραγματοποίησης του *A* μέσα σε κάποιο χρονικό διάστημα ισοδυναμεί με τον αριθμό «επιτυχιών» σε ένα ίσως μεγάλο αλλά πεπερασμένο αριθμό δοκιμών Bernoulli, και την συνάρτηση μάζας πιθανότητας των προαναφερθέντων τυχαίων μεταβλητών θα την μελετούσαμε με την διωνυμική κατανομή. Όμως, στην προκειμένη περίπτωση το γεγονός *A* μπορεί να συμβεί σε οποιαδήποτε χρονική στιγμή ή τοπικό σημείο, και αυτό σημαίνει ότι σε ένα μικρό διάστημα Δt, όσο μικρό και αν είναι, ενδέχεται να συμβεί περισσότερες από μία φορές το *A*, το οποίο αντιφάσκει με την διαδικασία Bernoulli.

Η κατηγορία των παραπάνω τυχαίων μεταβλητών X_t, οι οποίες παριστάνουν πόσες φορές εμφανίζονται τα τυχαία γεγονότα ή φαινόμενα σε κάποιο χρονικό διάστημα, μελετάται με την ***διαδικασία Poisson.***

Ορισμός Μία διαδικασία με την οποία συμβαίνει ένα τυχαίο γεγονός *A* πάνω στον άξονα του χρόνου *t* λέγεται **διαδικασία Poisson,** αν ισχύουν οι παρακάτω υποθέσεις:

1) Ο αριθμός πραγματοποίησης ενός τυχαίου γεγονότος *A* σε ένα χρονικό διάστημα *t* εξαρτάται μόνο από το μήκος του χρονικού διαστήματος και όχι από τα άκρα του διαστήματος(τοποθεσία του διαστήματος πάνω στον άξονα του χρόνου).
2) Υπάρχει ένα μικρό διάστημα $\boldsymbol{\Delta t}$ μέσα στο οποίο η εμφάνιση του τυχαίου γεγονότος *A* περισσότερο από μία φορά είναι αμελητέα.
3) Η πιθανότητα να συμβεί το γεγονός *A* μέσα στο διάστημα $\boldsymbol{\Delta t}$ είναι το γινόμενο $\boldsymbol{\lambda\cdot\Delta t}$, όπου $\boldsymbol{\lambda}$ είναι μία σταθερή παράμετρος. Αυτό σημαίνει ότι εάν το διάστημα Δt είναι ικανοποιητικά μικρό, η πιθανότητα να συμβεί το γεγονός μέσα στο μικρό διάστημα Δt εξαρτάται από το μήκος του διαστήματος αυτού.

Ορισμός. Η τυχαία μεταβλητή $\boldsymbol{X_t}$, η οποία παριστάνει τον αριθμό εμφάνισης του *A* στο χρονικό διάστημα ***t*** σε μία **διαδικασία Poisson**, ονομάζεται **μεταβλητή Poisson**, η δε συνάρτηση κατανομής μάζας πιθανότητας δίνεται από το **μοντέλο Poisson**

$$P(X_t = x) = e^{-\lambda t}\frac{(\lambda t)^x}{x!} \quad \text{για} \quad x=0,1,2,3,\dots \qquad (6.16)$$

όπου η παράμετρος $\boldsymbol{\lambda}$ παριστάνει τον μέσο αριθμό εμφάνισης του γεγονότος *A* στην μονάδα του χρόνου.

Η συνάρτηση (6.16) είναι πράγματι συνάρτηση μάζας πιθανότητας αφού

$$\sum_{x=0}^{\infty} p(x) = e^{-\lambda t} \sum_{x=0}^{\infty} \frac{\lambda t^{x}}{x!} = e^{-\lambda t} e^{\lambda t} = 1.$$

Η κατανομή πιθανότητας (6.16) για την τυχαία μεταβλητή X_t δημοσιεύθηκε από τον S.D. Poisson το 1837.

Ο προσδιορισμός της **μέσης τιμής** μίας μεταβλητής Poisson X_t επιτυγχάνεται ως συνήθως από τον ορισμό,

$$E(X_t) = \sum_{x=0}^{\infty} x e^{-\lambda t} \frac{(\lambda t)^{x}}{x!} = \lambda t \sum_{x=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^{x-1}}{(x-1)!}$$

θέτοντας s=x-1, έχουμε

$$E(X_t) = \lambda t \sum_{s=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^{s}}{s!} = \lambda t \qquad (6.17)$$

Δηλαδή, η μέση τιμή της τυχαίας μεταβλητής X_t η οποία αναφέρεται στην μονάδα του χρόνου είναι η παράμετρος ***λ*.**

Χωρίς απόδειξη, αναφέρουμε ότι και η **διακύμανση** της μεταβλητής Poisson είναι πάλι το γινόμενο *λt,*

$$Var(X_t) = \lambda t. \qquad (6.18)$$

Παρατηρήσεις

1) Πρέπει να τονίσουμε εδώ ότι μία Poisson κατανομή προκύπτει ύστερα από τις βασικές υποθέσεις που κάναμε. Αυτό σημαίνει ότι όταν ικανοποιούνται αυτές οι υποθέσεις, το μοντέλο της Poisson είναι το πιο κατάλληλο. Όπως προκύπτει παρακάτω ένας μεγάλος αριθμός φαινομένων ταιριάζει στο μοντέλο της Poisson κατανομής.
2) Η Poisson μεταβλητή κατέχει την ιδιότητα ότι η μέση τιμή είναι ίση με την διακύμανσή της.

3) Η συνάρτηση μάζας πιθανότητας της Poisson μεταβλητής X_t προσδιορίζεται πλήρως από την παράμετρο λ, η οποία παριστάνει τον μέσο αριθμό εμφάνισης του γεγονότος A στην μονάδα του χρόνου ή με άλλα λόγια την *συχνότητα* εμφάνισης του A. Ενώ, λt παριστάνει την μέση τιμή του αριθμού εμφάνισης του A σε χρόνο t,

$$E(X_t)= \lambda t.$$

Παράδειγμα 6.12

Από το 1900 έως το 1980 σε μία σεισμογενή περιοχή στην Ελλάδα έγιναν 16 ισχυροί σεισμοί. Ποια η πιθανότητα ότι τα επόμενα 10 χρόνια θα συμβούν δύο σεισμοί; Πόσοι ισχυροί σεισμοί αναμένονται για τα επόμενα 20 χρόνια;

Αφού στα 80 χρόνια έγιναν 16 σεισμοί η συχνότητα ή ο μέσος αριθμός σεισμών τον χρόνο είναι λ=16/80=0,2 σεισμοί /χρόνο. Συμβολίζουμε με X_{10} την μεταβλητή η οποία παριστάνει τον αριθμό σεισμών που θα συμβούν στα επόμενα 10 χρόνια, οπότε από το μοντέλο του Poisson έχουμε

$$P(X_{10} = 2) = e^{-0,2\cdot 10}\,\frac{(0,2\cdot 10)^2}{2!} = e^{-2}\cdot 2 = 0,2723$$

Η τυχαία μεταβλητή X_{20} παριστάνει τον αριθμό σεισμών στα 20 χρόνια. Άρα η μέση της τιμή είναι

$$E(X_{20}) = \lambda\cdot 20 = 0,2\cdot 20 = 4 \text{ σεισμοί}$$

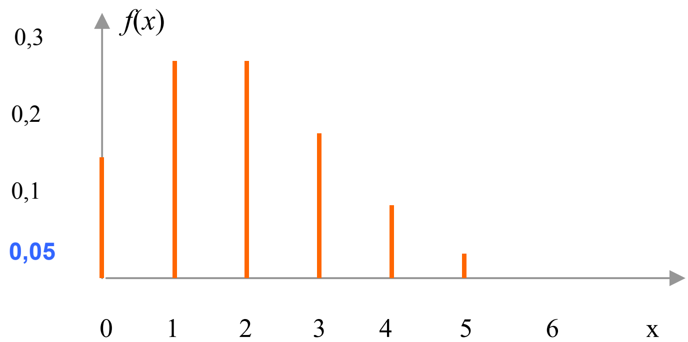
Παράδειγμα 6.13

Όλες οι ηλεκτρικές αντλίες νερού που χρησιμοποιούνται για την ύδρευση μιας πόλης είναι του ιδίου τύπου. Είναι γνωστό ότι την θερινή περίοδο κατά μέσο όρο 2 αντλίες παρουσιάζουν βλάβη και δεν λειτουργούν. Ο νέος διευθυντής ύδρευσης της πόλης θεωρεί ότι τους θερινούς μήνες και μέχρι 4 αντλίες να χαλάσουν η παροχή νερού στην πόλη είναι ικανοποιητική. Ποια η πιθανότητα ότι αυτό το καλοκαίρι δεν θα δημιουργηθεί πρόβλημα στην παροχή νερού στην πόλη λόγω βλάβης αντλιών.

Θεωρούμε ότι η μεταβλητή X_3 παριστάνει τον αριθμό αντλιών που παρουσιάζουν βλάβη μέσα σε χρόνο 3 μηνών το καλοκαίρι. Η συχνότητα βλάβης είναι λ=2/3 βλάβες/μήνα. Ακόμη, υποθέτουμε ότι πληρούνται οι συνθήκες για μία διαδικασία Poisson (και είναι λογικό αυτό) Επομένως, η ζητούμενη πιθανότητα σύμφωνα με το αθροιστικό μοντέλο Poisson είναι

$$P(X_3 \le 4) = \sum_{x=0}^{4} e^{-(2/3)\cdot 3} \frac{[(2/3)\cdot 3]^x}{x!} = e^{-2} + e^{-2}\frac{2^1}{1!} + e^{-2}\frac{2^2}{2!} + e^{-2}\frac{2^3}{3!} + e^{-2}\frac{2^4}{4!}$$
$$= 0{,}135 + 0{,}271 + 0{,}271 + 0{,}180 + 0{,}090 = 0{,}947$$

Στο παράδειγμα 6.13, η γραφική παράσταση της συνάρτησης μάζας πιθανότητας της Poisson μεταβλητής X_3 δίνεται στο Σχήμα 6.4. Πρέπει να τονίσουμε ότι στην Poisson κατανομή θα περίμενε κανείς ότι η πιθανότητα P(X_t=λt) (ή τον ακέραιο πλησιέστερα στο λt) είναι αρκετά μεγαλύτερη από όλες τις άλλες. Αυτό όμως δεν συμβαίνει γιατί η X_t μπορεί να πάρει πολλές άλλες τιμές και η μάζα πιθανότητας μοιράζεται σε πολλά σημεία. Πιο συγκεκριμένα στο παράδειγμα, η πιθανότητα $P(X_3 = \lambda t = 2) = P(X_3 = 1) = 0{,}271$ παρατηρούμε ότι δεν είναι αρκετά μεγάλη όπως θα περίμενε κανείς. Σημαντική μάζα πιθανότητας κατανέμεται και στα σημεία x=0, x=3, ενώ στα υπόλοιπα σημεία x=4, x=5, x=6, … η μάζα ολοένα και ελαττώνεται.



Σχήμα 6.4. Συνάρτηση μάζας πιθανότητας για την Poisson μεταβλητή X_3 με μέση τιμή λt=2

Poisson διαδικασία ως προς το μήκος, επιφάνεια και χώρο

Η διαδικασία Poisson αρχικά έχει χρησιμοποιηθεί για γεγονότα που εμφανίζονται με την πάροδο του χρόνου. Επίσης, η διαδικασία Poisson εφαρμόζεται και σε φαινόμενα που εμφανίζονται ως προς το **μήκος** ,**επιφάνεια** ή ακόμη και **χώρο**, και μπορούμε βεβαίως να χρησιμοποιήσουμε την συνάρτηση κατανομής Poisson. Για παράδειγμα, εκτός των περιπτώσεων που αναφέρθηκαν στην αρχή αυτής της παραγράφου,

- ο αριθμός ελαττωματικών στηλών σε κάποιο μήκος ηλεκτρικής γραμμής.
- ο αριθμός αποτελεσματικών γεωτρήσεων σε έναν Νομό,
- ο αριθμός σωματιδίων σε μία ποσότητα πρώτης ύλης για την κατασκευή κρυστάλλινων ποτηριών και άλλων.

Παρατηρούμε ότι η Poisson κατανομή εφαρμόζεται και σε καταστάσεις όπου ένας μεγάλος αριθμός αντικειμένων είναι σκορπισμένος πάνω σε μία μεγάλη επιφάνεια ή μέσα σε ένα μεγάλο χώρο. Για παράδειγμα,

- ο αριθμός ελαττωμάτων σε κάποιο υλικό που χρησιμοποιεί ο μηχανικός,
- ο αριθμός βλαβερών οργανισμών σε κάποιο όγκο υγρού ή αέρα,
- ο αριθμός ελαττωμάτων πάνω σε μία επιφάνεια δαπέδου και άλλα.

Όμως, η Poisson διαδικασία δεν ισχύει όταν τα αντικείμενα ομαδοποιούνται πάνω σε μία επιφάνεια και η ευκαιρία να τα ευρείς σε μία περιοχή είναι διαφορετική από ότι σε κάποια άλλη. Όπως η Poisson κατανομή δεν εφαρμόζεται σε αλληλοεξαρτώμενα γεγονότα που συμβαίνουν μέσα στον χρόνο, έτσι και γεγονότα που ενδιαφέρουν τον μηχανικό και συμβαίνουν ομαδικά-όπως αλληλοεξαρτώμενες βροχοπτώσεις σε μία περιοχή ή ατέλειες σε μία επιφάνεια ενός υλικού που βρίσκονται ευκολότερα σε μία περιοχή παρά σε κάποια άλλη- θα πρέπει να αποκλειστούν από την Poisson.

Παράδειγμα 6.14

Τα σωματίδια σκόνης στην ατμόσφαιρα προκαλούν περιβαλλοντικό πρόβλημα. Με ένα ειδικό μικροσκόπιο παρατηρούμε σε πολλά δείγματα τον αριθμό τέτοιων σωματιδίων ανά μονάδα όγκου, όπου ο μέσος αριθμός ανά μονάδα όγκου βρέθηκε 4,2 σωματίδια. Ποια η πιθανότητα σε μία μονάδα όγκου ατμόσφαιρας να μην βρεθεί σωματίδιο; Ποια η πιθανότητα να βρεθούν 4 σωματίδια;

Με την παραδοχή ότι τα σωματίδια υπάρχουν τυχαία μέσα στην ατμόσφαιρα, ο αριθμός σωματιδίων ανά μονάδα όγκου X ακολουθεί Poisson κατανομή με παράμετρο λ=4,2. Έτσι, εύκολα υπολογίζουμε τις ζητούμενες πιθανότητες,

$$P(X=0)=e^{-4,2}\frac{4,2^{0}}{0!}=e^{-4,2}=0,0152$$
$$P(X=4)=e^{-4,2}\frac{4,2^{4}}{4!}=0,0152\frac{311,17}{24}=0,197$$

6.6.2 Η Poisson σαν μία προσέγγιση στην διωνυμική κατανομή

Η μεταβλητή Poisson έχει ένα μεγάλο πεδίο εφαρμογών σε διάφορες περιοχές διότι μπορεί να χρησιμοποιηθεί σαν μία προσέγγιση για την διωνυμική μεταβλητή με παραμέτρους (n, p) όταν n είναι πολύ μεγάλο και p είναι πολύ μικρό.

Για να καταλάβουμε καλύτερα την προσέγγιση των δύο κατανομών αρχίζουμε με το εξής σκεπτικό: Η μεταβλητή Poisson X_t είναι διακριτή και συμβολίζει τον αριθμό πραγματοποίησης του A σε κάποιο συνεχές διάστημα t με συχνότητα λ(μέσος αριθμός πραγματοποίησης του A στην μονάδα του χρόνου). Την κατανομή μάζας πιθανότητας της X_t την προσεγγίζουμε με την διωνυμική αν διαιρέσουμε το χρονικό διάστημα t σε πολύ μικρά χρονικά διαστήματα πλάτους Δt, έτσι ώστε η πιθανότητα να συμβεί το A περισσότερες από μία φορές μέσα στο Δt να είναι μηδενική. Αυτό σημαίνει ότι το Δt είναι πάρα πολύ μικρό και έτσι το διάστημα t χωρίζεται σε ένα πολύ μεγάλο αριθμό n υπο-διαστημάτων, $n=(t/\Delta t)$, του ιδίου πλάτους. Σε κάθε διάστημα Δt θεωρούμε μία δοκιμή Bernoulli η οποία αφορά την εμφάνιση του υπό μελέτη γεγονότος A ή $\overline{A}$. Η πιθανότητα πραγματοποίησης του A σε μία δοκιμή ή μέσα στο διάστημα Δt είναι σταθερή και ίση με

$$p=\lambda\cdot\Delta t,$$

αφού το λ παριστάνει την μέση τιμή του αριθμού εμφάνισης του A στην μονάδα του χρόνου, ή ισοδύναμα σε αριθμό δοκιμών $n=(1/\Delta t)$, και σε μία διωνυμική κατανομή ισχύει $\lambda=(1/\Delta t)\cdot p$. Αφού το Δt είναι πολύ μικρό και η προκύπτουσα πιθανότητα p θα είναι επίσης πολύ μικρή.

Συμπερασματικά, σε μία μεταβλητή Poisson την πιθανότητα $P(X_t=x)$ μπορούμε να την προσεγγίσουμε με το μοντέλο της διωνυμικής όπου το πλήθος n δοκιμών είναι πολύ μεγάλο ($n\to\infty$) και αντίστροφα η σταθερή πιθανότητα «επιτυχίας» p πολύ μικρή($p\to 0$), δεδομένου ότι η παράμετρος λ είναι σταθερή και ισχύει, $\lambda t=np$. Έτσι, έχουμε

$$P(X_t = x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{(n-x)!\,x!} p^x (1-p)^{n-x}$$

$$= \frac{n!}{(n-x)!\,x!}\left(\frac{\lambda t}{n}\right)^x \left(1-\frac{\lambda t}{n}\right)^{n-x} = \frac{n(n-1)\ldots(n-x+1)}{n^x}\frac{(\lambda t)^x}{x!}\frac{(1-\lambda t/n)^n}{(1-\lambda t/n)^x}$$

και επειδή για πολύ μεγάλο n και πολύ μικρό p ισχύει

$$\left(1-\frac{\lambda t}{n}\right)^n \approx 1, \qquad \frac{n(n-1)\ldots(n-x+1)}{n^x} \approx 1, \qquad \left(1-\frac{\lambda t}{n}\right)^x \approx 1$$

καταλήγουμε στην εξίσωση

$$P(X_t = x) \approx e^{-\lambda t}\frac{(\lambda t)^x}{x!}.$$

Με άλλα λόγια, εάν εκτελεσθούν n ανεξάρτητες δοκιμές Bernoulli με πιθανότητα «επιτυχίας» *p,* και εάν το *n* είναι αρκετά μεγάλο και το *p* μικρό, ο αριθμός «επιτυχιών» είναι προσεγγιστικά Poisson τυχαία μεταβλητή με μέση τιμή $\lambda t = np$.

Μερικές περιπτώσεις διωνυμικών μεταβλητών, όπου έχουμε μεγάλο αριθμό δοκιμών και μικρή πιθανότητα «επιτυχίας» και ταιριάζει μία καλή προσέγγιση με Poisson μοντέλο, είναι οι ακόλουθες,

- Ο αριθμός λανθασμένων τηλεφωνικών κλήσεων σε μία μέρα
- Ο αριθμός τυπογραφικών λαθών σε ένα βιβλίο
- Ο αριθμός πελατών που αγοράζουν ένα συγκεκριμένο προϊόν σε ένα μεγάλο κατάστημα
- Ο αριθμός ανδρών που ξεπερνούν την ηλικία των 90 ετών σε μία πόλη

Οι μεταβλητές αυτές και αρκετές άλλες είναι προσεγγιστικά Poisson επειδή έχουμε προσέγγιση της Poisson στην διωνυμική.

Για παράδειγμα, μπορούμε να υποθέσουμε ότι υπάρχει μία μικρή πιθανότητα *p* ότι κάθε γράμμα που γράφεται σε ένα βιβλίο να είναι λανθασμένο, και έτσι ο αριθμός λαθών σε ένα βιβλίο είναι προσεγγιστικά Poisson με μέση τιμή $\lambda = np$ όπου *n* είναι προφανώς ένας μεγάλος αριθμός των γραμμάτων που περιέχει το βιβλίο.

Παρόμοια, μπορούμε να υποθέσουμε ότι κάθε άνδρας σε μία πόλη, ανεξάρτητα ο ένας από τον άλλον, έχει μία μικρή πιθανότητα *p* να ζήσει περισσότερο από 90 έτη και ο αριθμός των ανδρών οι οποίοι θα ζήσουν πάνω από 90 έτη είναι προσεγγιστικά Poisson μεταβλητή με μέση τιμή $\lambda = np$, όπου *n* είναι ο αριθμός των ανδρών της πόλης.

Παρατηρήσεις

1) Η διωνυμική κατανομή χαρακτηρίζεται από δύο παραμέτρους, n και p, ενώ η Poisson κατανομή χαρακτηρίζεται από την μοναδική παράμετρο λ,η οποία παριστάνει τον μέσο αριθμό «επιτυχιών» ή συμβάντων ανά μονάδα χρόνου ή χώρου και είναι γνωστή σαν *συχνότητα*. Σε μία Poisson μεταβλητή X_t πρέπει να ξεχωρίζουμε την συχνότητά της λ από την μέση τιμή της $\lambda t=np$, ή οποία είναι ο μέσος αριθμός των συμβάντων στο προκαθορισμένο χρονικό διάστημα t.
2) Μπορούμε να εκτιμήσουμε την μέση τιμή και διακύμανση μίας Poisson μεταβλητής X_t από την προσέγγισή της με την διωνυμική. Μία διωνυμική μεταβλητή Y με πολύ μεγάλο n και πολύ μικρό p προσεγγιστικά ταυτίζεται με την X_t οπότε έχουμε

$$E(X_t) = E(Y)=np=\lambda t, \text{ και } Var(X_t)=Var(Y)=np(1-p)=np=\lambda t,$$

διότι όταν η πιθανότητα p είναι πολύ μικρή προσεγγιστικά $(1-p)=1$.

Παράδειγμα 6.15

Από μία διασταύρωση ενός αυτοκινητόδρομου περνούν την ημέρα (σε 12 ώρες) κατά μέσο όρο 10800 αυτοκίνητα, από τα οποία το 10% στρίβει αριστερά. Στην διασταύρωση υπάρχει ένα φανάρι όπου το κόκκινο βέλος για αριστερά διαρκεί δύο λεπτά. Ανάβει πράσινο βέλος για αριστερά, ποια η πιθανότητα ότι τέσσερα αυτοκίνητα περιμένανε να στρίψουν αριστερά;

Θεωρούμε σαν γεγονός A την τυχαία άφιξη ενός αυτοκινήτου στην διασταύρωση, το οποίο επιθυμεί να στρίψει αριστερά. Το γεγονός αυτό μπορεί να συμβεί τυχαία σε οιαδήποτε χρονική στιγμή μέσα στην ημέρα. Η δυσκολία είναι ότι σε οσοδήποτε μικρό υποδιάστημα του χρόνου υπάρχει όχι μόνο μία χρονική στιγμή αλλά πολλές που αντιστοιχούν στις αφίξεις των αυτοκινήτων για αριστερά.

Ο μέσος όρος αυτοκινήτων που επιθυμούν να στρίψουν αριστερά είναι προφανώς λ=1080/720=1,5 αυτοκίνητα/λεπτό. Χωρίζουμε το διάστημα t=δύο λεπτά σε n=12 υποδιαστήματα των Δt=10 δεύτερων λεπτών. Με την παραδοχή ότι μέσα στο υποδιάστημα των Δt=10’’ το πολύ ένα αυτοκίνητο πλησιάζει το φανάρι για αριστερά, μπορούμε να απαντήσουμε στην ερώτηση με διωνυμική κατανομή, διότι

- σε κάθε δοκιμή Bernoulli ερωτούμε αν μέσα στο υποδιάστημα Δt=10 δεύτερα ήλθε (γεγονός A) ή δεν ήλθε (γεγονός $\overline{A}$) αυτοκίνητο για αριστερά,
- υπάρχουν συνολικά n=12 δοκιμές Bernoulli στατιστικώς ανεξάρτητες, μέσα στο διάστημα t=2 λεπτά,
- υπάρχει σταθερή πιθανότητα p να αφιχθεί αυτοκίνητο για αριστερή στροφή «επιτυχία» μέσα στο Δt=10 δεύτερα, η οποία είναι $p=\lambda t/n$=(1,5·2)/12=0,25. Διότι, η μέση τιμή του αριθμού των «επιτυχιών» στις n δοκιμές Bernoulli είναι np και αυτό αντιστοιχεί στο μέσο αριθμό αυτοκινήτων που επιθυμούν να στρίψουν αριστερά μέσα σε χρόνο t=2 λεπτά, δηλαδή $np=\lambda t$ ή $p=\lambda t/n$.

Έτσι η ζητούμενη πιθανότητα προσεγγιστικά είναι

$$P(X=4)=\binom{12}{4}p^4(1-p)^{12-4}=1980(0{,}25)^4 0{,}75^8=0{,}774$$

Εάν θεωρούμε ότι το υπό-διάστημα Δt=10 δευτερόλεπτα είναι μεγάλο και ενδέχεται μέσα σε αυτό να υπάρξουν περισσότερες από μία αφίξεις για αριστερά, χωρίζουμε το διάστημα t=2 λεπτών σε μικρότερα υπό-διαστήματα των Δt=5 δευτερολέπτων, οπότε έχουμε 24 δοκιμές Bernoulli με n=24, και σταθερή πιθανότητα επιτυχίας $p=\lambda t/n$=3/24=1/8. Επιχειρούμε και πάλι να υπολογίσουμε την πιθανότητα με διωνυμική κατανομή,

$$P(X=4)=\binom{24}{4}(1/8)^4(1-1/8)^{12-4}=10626(1/8)^4(7/8)^{20}=0{,}1795$$

Παρόμοια μπορούμε για καλύτερη προσέγγιση να πάρουμε μικρότερο υποδιάστημα, Δt=2 δεύτερα και κατά συνέπεια έχουμε n=60, p=3/60=1/20

$$P(X=4)=\binom{60}{4}(1/20)^4(1-1/20)^{12-4}=403560(0{,}05)^4(0{,}95)^8=0{,}14254$$

Καθώς ελαττώνουμε το υπό-διάστημα Δt=10'', 5'', 2'', … παρατηρούμε ότι αυξάνει ο αριθμός δοκιμών n=12, 24, 60, …, αντίστροφα ελαττώνεται η πιθανότητα «επιτυχίας» p=1/4, 1/8, 1/20, …, επειδή η μέση τιμή του αριθμού αυτοκινήτων για αριστερά παραμένει σταθερά, np=3. Η ζητούμενη πιθανότητα με το διωνυμικό μοντέλο είναι αντίστοιχα $P(X$=4)=0,774, 0,1795, 0,1426, … και άν συνεχίσουμε για πολύ μεγάλο n και πολύ μικρό p θα προσεγγίσει την πιθανότητα της Poisson κατανομής

$$P(X=4)=e^{-\lambda t}\frac{(\lambda t)^4}{4!}=0{,}168750$$

Παράδειγμα 6.16

Σε ένα σύνθετο σύστημα υπάρχουν 15000 ηλεκτρικά εξαρτήματα. Η πιθανότητα να συμβεί βλάβη σε κάθε εξάρτημα μέσα σε ένα μήνα είναι p=0,0001. Το σύστημα επιθεωρείται στο τέλος από κάθε μήνα και εξαρτήματα με βλάβη αντικαθίστανται από καινούργια. Κατά την διάρκεια της λειτουργίας του μέσα στον μήνα εάν δύο η περισσότερα εξαρτήματα παρουσιάσουν βλάβη το σύστημα διακόπτει την λειτουργία του. Ποια η πιθανότητα ότι το σύστημα θα λειτουργήσει τον επόμενο μήνα χωρίς διακοπή:

Υποθέτουμε ότι η πιθανότητα p να συμβεί βλάβη σε κάθε εξάρτημα είναι σταθερή. Εάν X είναι τυχαία μεταβλητή, η οποία παριστάνει τον αριθμό εξαρτημάτων που παρουσιάζουν βλάβη μέσα στον μήνα, μπορούμε να πούμε ότι είναι διωνυμική μεταβλητή με παραμέτρους n=15000 και p=0,0001. Έτσι, η ζητούμενη πιθανότητα είναι

$$P(X<2)=P(X=0)+P(X=1)$$
$$=\binom{15000}{0}\cdot 0{,}0001^0\cdot 0{,}9999^{15000}+\binom{15000}{1}\cdot 0{,}0001^1\cdot 0{,}9999^{14999}=\ldots$$

Παρατηρούμε ότι ο υπολογισμός της παραπάνω πιθανότητας παρουσιάζει κάποια δυσκολία με τους μεγάλους αριθμούς. Εφόσον η πιθανότητα p είναι πολύ μικρή και ο αριθμός n είναι πολύ μεγάλος μπορούμε να προσεγγίσουμε πολύ καλά την ζητούμενη πιθανότητα με Poisson κατανομή όπου $\lambda t=np$=15000(0,0001)=1,5 βλάβες/μήνα,

$$P(X_t=x)=e^{-\lambda t}\frac{(\lambda t)^x}{x!}=e^{-1{,}5}\frac{(1{,}5)^x}{x!}$$

οπότε,

$$P(X_t < 2) = e^{-1,5}\frac{(1,5)^0}{0!} + e^{-1,5}\frac{(1,5)^1}{1!} = e^{-1,5}(1+1,5) = 0,56$$

6.7 ΠΟΛΥΩΝΥΜΙΚΗ ΚΑΤΑΝΟΜΗ

Τελικά, αναφέρουμε και το πολυωνυμικό μοντέλο το οποίο είναι αρκετά σημαντικό στις διακριτές μεταβλητές, και μπορεί να θεωρηθεί σαν μία γενίκευση της διωνυμικής κατανομής. Θεωρούμε ένα πείραμα τύχης όπου όλα τα δυνατά αποτελέσματα ή ο δειγματικός χώρος S χωρίζεται σε κ αμοιβαία αποκλειόμενα γεγονότα A_1, A_2, ..., A_κ. Για κάθε γεγονός A_i η πιθανότητα πραγματοποίησής του p_i, $P(A_i)=p_i$, είναι σταθερή σε κάθε εκτέλεση του πειράματος και προφανώς η άθροισή τους μας δίνει την μονάδα, $p_1+p_2+\ldots+p_\kappa=1$. Ορίζουμε τις κ μεταβλητές X_1, X_2, ..., X_κ οι οποίες παριστάνουν τον αριθμό πραγματοποίησης των γεγονότων A_1, A_2, ..., A_κ αντίστοιχα στις n επαναλήψεις του πειράματος τύχης. Οι μεταβλητές
X_1, X_2, ...,X_κ δεν είναι ανεξάρτητες διότι το άθροισμά τους είναι σταθερό $X_1+X_2+\ldots+X_\kappa=n$ και όταν η μία παίρνει μία τιμή προσδιορίζονται έμμεσα και οι τιμές των άλλων.

Εάν X_1, X_2, ...,X_κ είναι μεταβλητές όπως ορίσθηκαν παραπάνω, ισχύει το ακόλουθο αποτέλεσμα

$$P[(X_1=x_1)\cap(X_2=x_2)\cap\ldots\cap(X_\kappa=x_\kappa)] = \frac{n!}{x_1!x_2!\ldots x_\kappa!}p_1^{x_1}p_2^{x_2}\ldots p_k^{x_\kappa} \qquad (6.19)$$

όπου $x_1+x_2+\ldots+x_\kappa=n$.

Η απόδειξη του πολυωνυμικού μοντέλου (6.19) προκύπτει όπως και στην διωνυμική κατανομή. Όλοι οι δυνατοί τρόποι με τους οποίους θα πραγματοποιηθούν x_1, x_2,...,x_κ φορές τα γεγονότα A_1, A_2, ..., A_κ αντίστοιχα είναι όλοι οι δυνατοί τρόποι με τους οποίους μπορούμε να τοποθετήσουμε n πράγματα σε μία σειρά, τα οποία χωρίζονται σε ομάδες των x_1, x_2,...,x_κ όμοιων πραγμάτων. Από την παράγραφο της Συνδυαστικής του Κεφαλαίου 2 οι τρόποι αυτοί είναι τον αριθμό,

$$\frac{n!}{x_1!x_2!\ldots x_\kappa!}.$$

Παρατηρήσεις

1) Εάν κ=2 το πολυωνυμικό μοντέλο καταλήγει στο διωνυμικό.

2) Το παραπάνω μοντέλο είναι γνωστό σαν πολυωνυμική κατανομή πιθανότητας. Υπενθυμίζουμε ότι η διωνυμική ήταν ανάπτυγμα του διωνύμου $(p+(1-p))^n$. Ανάλογα και οι πιθανότητες στην η πολυωνυμική κατανομή προκύπτουν από το ανάπτυγμα του πολυωνύμου $(p_1+p_2+...+p_n)^n$.

Η κάθε μία μεταβλητή X_i στο πολυωνυμικό μοντέλο μπορεί να θεωρηθεί σαν διωνυμική μεταβλητή , με πιθανότητα «επιτυχίας» p_i, $P(A_i)=p_i$. Συνεπώς, η **μέση τιμή** και **διακύμανση** της κάθε μεταβλητής στο πολυωνυμικό μοντέλο είναι

$$E(X_i)=np_i, \qquad Var(X_i)=np_i(1-p_i)$$

Παράδειγμα 6.17

Σε μία μεγάλη αποθήκη ηλεκτρολογικού υλικού υπάρχουν καλώδια των οποίων η διάμετρος X δεν είναι σταθερό και βρίσκεται μέσα στο διάστημα ($5<X<5{,}8$) χιλιοστών. Ο μηχανικός ενδιαφέρεται να ξέρει αν το καλώδιο ανήκει σε μία από τις τρεις κατηγορίες:

A_1={X<5,2 χιλιοστά}, A_2={5,2≤X≤5,5 χιλιοστά } και A_3={X>5,5 χιλιοστά}.

Από άλλες πληροφορίες είναι περίπου γνωστό ότι στην αποθήκη το 25% των καλωδίων ανήκουν στην ομάδα A_1, το 60% στην ομάδα A_2 και 15% στην ομάδα A_3. Επιλέγει τυχαία 20 καλώδια για ένα μία συσκευή, ποια η πιθανότητα δύο καλώδια να έχουν διάμετρο <5,2 χιλιοστά και τέσσερα >5,5 χιλιοστά.

Εφαρμόζω το πολυωνυμικό μοντέλο

$$P[(X_1=4)\cap(X_2=14)\cap(X_3=2)]=\frac{20!}{4!14!2!}0{,}25^4 0{,}60^{14} 0{,}15^2=0{,}008$$

6.8 ΣΧΕΣΕΙΣ ΜΕΤΑΞΥ ΔΙΑΚΡΙΤΩΝ ΚΑΤΑΝΟΜΩΝ

Πριν περάσουμε στις κατανομές πιθανοτήτων συνεχών τυχαίων μεταβλητών, πρέπει να τονίσουμε ότι όλες οι διακριτές κατανομές που εισήχθηκαν προηγουμένως σχετίζονται η μία με την άλλη. Όπως, έχει αναφερθεί η διωνυμική είναι το άθροισμα από n ανεξάρτητες Bernoulli μεταβλητές, και ακόμη η διωνυμική είναι μία ειδική περίπτωση του πολυωνυμικού μοντέλου.

Ας δούμε όμως λίγο περισσότερο τις **σχέσεις** μεταξύ **διωνυμικής**, **υπεργεωμετρικής** και **Poisson** κατανομής.

Πρώτα, ας δούμε την σχέση μεταξύ **διωνυμικής** και **υπεργεωμετρικής** κατανομής. Υποθέτουμε ότι 30.000 από τις 100.000 μέλη ενός συνδικάτου προτίθενται να απεργήσουν. Εάν 10 μέλη επιλέγονται τυχαία χωρίς αντικατάσταση, τότε η πιθανότητα ότι ακριβώς x μέλη του δείγματος πρόκειται να απεργήσουν δίνεται από την υπεργεωμετρική κατανομή,

$$P(X=x)=\frac{\binom{\kappa}{x}\binom{N-\kappa}{n-x}}{\binom{N}{x}}=\frac{\binom{30.000}{x}\binom{100.000-30.000}{x}}{\binom{100.000}{x}}$$

Εάν, όμως, η επιλογή των 10 μελών γίνεται τυχαία με αντικατάσταση, τότε η πιθανότητα «επιτυχίας» σε κάθε ανεξάρτητη δοκιμή είναι σταθερή, p=30.000/100.000=0,3, η διωνυμική κατανομή είναι κατάλληλη για τον αριθμό μελών x στο δείγμα που πρόκειται να απεργήσουν, και η πιθανότητα υπολογίζεται από τον τύπο της,

$$P(X=x)=\binom{n}{x}p^x(1-p)^{n-x}=\binom{10}{x}0{,}3^x 0{,}7^{10-x}.$$

Τώρα, ο αριθμός μελών του συνδικάτου είναι πολύ μεγάλος N=100.000, δεν θα αλλάξει και πολύ άν πάρουμε από αυτόν ένα μικρό δείγμα n=10. Δηλαδή, όταν το n είναι πολύ μικρό και το N πολύ μεγάλο σε κάθε δοκιμή μπορούμε να υποθέσουμε ότι η πιθανότητα «επιτυχίας» παραμένει σταθερή έστω και αν η επιλογή γίνεται χωρίς αντικατάσταση. Για παράδειγμα, σε επιλογή χωρίς αντικατάσταση, η πιθανότητα «επιτυχίας» στην πρώτη επιλογή είναι p=30.000/100.000=0,3 και στην δεύτερη p=29999/100.000≈0,3 ή p=30.000/100.000=0,3, εάν η πρώτη έφερε «επιτυχία» ή «αποτυχία» αντίστοιχα. Έτσι, θα περιμέναμε η διωνυμική πιθανότητα να προσεγγίσει την υπεργεωμετρική πιθανότητα,

$$P(X=x)=\frac{\binom{\kappa}{x}\binom{N-\kappa}{n-x}}{\binom{N}{x}}\cong\binom{n}{x}p^{x}(1-p)^{n-x} \tag{6.20}$$

το οποίο σημαίνει ότι, εάν μικρό δείγμα n επιλέγεται από μεγάλο πληθυσμό N, τότε χωρίς αντικατάσταση(η υπεργεωμετρική κατανομή) και με αντικατάσταση (η διωνυμική κατανομή) δίνουν περίπου τις ίδιες πιθανότητες. Ένας εμπειρικός κανόνας χαρακτηρίζει το n μικρό όταν n<0,05N.

Όταν μπορούμε να προσεγγίσουμε την υπεργεωμετρική με την διωνυμική υπάρχει ευκολία στον υπολογισμό πιθανότητας. Καθώς το N τείνει στο άπειρο οι δύο κατανομές ταυτίζονται. Για αυτό μπορούμε να πούμε ότι η διωνυμική είναι μία περίπτωση της υπεργεωμετρικής ή ο περιορισμός της υπεργεωμετρικής, καθώς το N προσεγγίζει το άπειρο.

Ενώ η διωνυμική μπορεί να χρησιμοποιηθεί σαν μία προσέγγιση στην υπεργεωμετρική κατανομή, η **Poisson** μπορεί να χρησιμοποιηθεί σαν μία προσέγγιση στην **διωνυμική κατανομή**. Όπως έχει αποδειχθεί, εάν n είναι μεγάλο και p είναι μικρό, τότε το μοντέλο της διωνυμικής και της Poisson σχετίζονται με την ακόλουθη εξίσωση

$$P(X=x)=\binom{n}{x}p^{x}(1-p)^{n-x}=e^{-np}\frac{(np)^{x}}{x!}, \tag{6.21}$$

όπου η πιθανότητα σε μία διωνυμική μεταβλητή μπορεί να αντικατασταθεί από την πιθανότητα της αντίστοιχης Poisson άν $\lambda=np$.

Όταν, όμως βρίσκουμε κατάλληλο να προσεγγίσουμε την διωνυμική με την Poisson πρέπει να γνωρίζουμε πότε το n είναι μεγάλο και το p μικρό. Σαν αριθμητικός κανόνας, όταν n>100 και p<0,01, η διωνυμική και η Poisson συμφωνούν κατά προσέγγιση στο τρίτο δεκαδικό ψηφίο για κάθε τιμή της τυχαίας μεταβλητής. Ακόμη, για εκείνες τις τιμές x για τις οποίες ισχύει $[(x-np)^2/n]<0{,}01$, τα δύο μοντέλα συμφωνούν με ακρίβεια εκατοστού. Τελικά, εάν η πιθανότητα p είναι πολύ μικρή τα δύο μοντέλα έχουν ικανοποιητική προσέγγιση ακόμη και εάν το n είναι μικρό όπως π.χ. n=10.

Ο Πίνακας 6.1 δίνει τις διωνυμικές πιθανότητες για x=1, 2, 3, 4, και 5 με n=10 και p=0.1, και με n=20 και p=0,05 αντίστοιχα. Σε κάθε περίπτωση οι

Poisson προσεγγίσεις είναι το ίδιο αφού οι δύο διωνυμικές κατανομές έχουν την ίδια μέση τιμή, $\lambda=np= 10(0,01)=20(0,05)=1$.

Πίνακας 6.1 **Διωνυμικές κατανομές και οι Poisson προσεγγίσεις τους**

x	0	1	2	3	4	5
Διωνυμική, n=10, p=0,10	0,349	0,387	0,194	0,057	0,011	0,0015
Διωνυμική, n=20, p=0,05	0,358	0,377	0,189	0,060	0,013	0,0022
Poisson , λ=1	0,368	0,368	0,184	0,061	0,015	0,0031

Από την επίδειξη στον Πίνακα 6.1 φαίνεται ότι υπάρχει μία ικανοποιητική προσέγγιση της διωνυμικής από την Poisson κατανομή. Παρόλα αυτά, όταν έχουμε την ευχέρεια να υπολογίσουμε τις πιθανότητες σε μία διωνυμική ή να τις πάρουμε από στατιστικούς πίνακες, αποφεύγουμε την προσέγγιση με την Poisson. Όμως, στις περιπτώσεις όπου το n είναι πολύ μεγάλο και το p ικανοποιητικά μικρό και δεν μπορούμε να επιτύχουμε τις διωνυμικές πιθανότητες από πίνακες, μπορούμε να κερδίσουμε χρόνο με Poisson προσέγγιση, όπως στο παράδειγμα 6.15 που ακολουθεί.

Παράδειγμα 6.18

Υποθέτουμε ότι η πιθανότητα μία στήλη ρεύματος να αστοχήσει λόγω κακοκαιρίας είναι 0,00002. Σε μία περιοχή όπου υπάρχουν 100.000 στήλες , ποια η πιθανότητα να αστοχήσουν λόγω κακοκαιρίας:

(α) τουλάχιστον 4 στήλες

Η πιθανότητα ότι τέσσερις ή περισσότερες στήλες θα αστοχήσουν με την διωνυμική κατανομή είναι

$$P(X \geq 4) = \sum_{x=4}^{100.000} \binom{100.000}{x} 0,00002^{x} 0,99998^{100.000-x} .$$

Θα προτιμούσαμε όμως να προσεγγίσουμε την παραπάνω πιθανότητα με Poisson μοντέλο, παρά να κάνουμε τόσο χρονοβόρες πράξεις. Έτσι μπορούμε με Poisson κατανομή και $\lambda=np$=100.000(0,00002)= 2 να έχουμε

$$P(X \geq 4) = 1 - P(X \leq 3) = 1 - \sum_{x=0}^{3} e^{-2} \frac{2^x}{x!}$$
$$= 1 - 0{,}85712$$
$$= 0{,}14288.$$

(β) *ακριβώς τέσσερις στήλες*

Παρόμοια, η πιθανότητα για ακριβώς τέσσερις αστοχίες σε μία κακοκαιρία υπολογίζεται με προσέγγιση Poisson,

$$P(X = 4) = e^{-2} \frac{2^4}{4!} = 0{,}09023.$$

Συμπερασματικά, η **υπεργεωμετρική** έχει σάν όριο της την **διωνυμική** και η **διωνυμική** έχει σαν όριο της την **Poisson**. Κατά συνέπεια και η **Poisson** μπορεί να χρησιμοποιηθεί για να προσεγγίσει την **υπεργεωμετρική** μέσου της **διωνυμικής** κατανομής, κάτω από τις **κατάλληλες συνθήκες**. Όπως θα δούμε στη συνέχεια, όλες αυτές οι πιθανότητες μπορούν να προσεγγισθούν από μία **συνεχή** κατανομή την **κανονική κατανομή-** κάτω από ορισμένες προυποθέσεις.

6.9 ΠΕΡΙΛΗΨΗ ΚΕΦΑΛΑΙΟΥ

Οι διακριτές κατανομές πιθανότητας που περιγράφθηκαν και αναπτύχθηκαν σε αυτό το κεφάλαιο παίζουν σημαντικό ρόλο σε πολλές εφαρμογές του μηχανικού. Ο μηχανικός πριν επιλέξει την θεωρητική κατανομή για την τυχαία μεταβλητή ή κάποιο χαρακτηριστικό που διερευνά, πρέπει να επιβεβαιώσει ότι λογικά πληρούνται οι συνθήκες για την αντίστοιχη επιλογή.

Έχουμε δείξει ότι υπάρχει θεωρητική σύνδεση μεταξύ των διακριτών κατανομών. Η Bernoulli είναι η βασικότερη για την δημιουργία της διωνυμικής, γεωμετρικής, αρνητικής διωνυμικής, υπεργεωμετρικής. Ακόμη, σημαντική για τον μηχανικό θεωρείται και η προσέγγιση της διωνυνικής με την Poisson, στην περίπτωση που υπάρχει μεγάλος αριθμός δοκιμών.

ΑΣΚΗΣΕΙΣ-ΠΡΟΒΛΗΜΑΤΑ

Με τις λύσεις τους

1) Πόσες δοκιμές Bernoulli με p=0,01 πρέπει να εκτελέσουμε για να βεβαιώσουμε , ότι η πιθανότητα εμφάνισης του A είναι 0,5 και πάνω;

Λύση

Αν η τυχαία μεταβλητή X παριστάνει τον αριθμό εμφάνισης του A στις n **δοκιμές Bernoulli** , ζητάμε να ισχύει η σχέση $P(X\geq 1)\geq 0,5$. Έχουμε

$$P(X\geq 1)=1-P(X<1)=1-P(X=0)=1-(1-p)^n=1-0,99^n\geq \frac{1}{2}$$

$$-n\ln(0,99)\geq \ln 2$$

ή τελικά $n\geq 69$.

2) Ρίχνουμε ένα ζάρι 20 φορές. Αν X ο αριθμός των εμφανίσεων του 1 τότε να βρεθεί

(α) η πιθανότητα εμφάνισης του 1 πέντε φορές

(β) η μέση τιμή και απόκλιση της X.

Λύση

¨εχουμε μία τυχαία μεταβλητή η οποία ακολουθεί την **διωνυμική κατανομή** με n=20 και p=1/6, οπότε

$$P(X=5)=\binom{20}{5}\left(\frac{1}{6}\right)^5\left(\frac{5}{6}\right)^{15}=0,129$$

$$E(X)=np=20\cdot\frac{1}{6}=3,33$$

$$Var(X)=np(1-p)=20\cdot\frac{1}{6}\cdot\frac{5}{6}=\frac{120}{36}=1,66^2$$

$$ή\quad \sigma=1,66$$

3) Αναβάθμιση τηλεφωνικού κέντρου. Δέκα συνδρομητές επικοινωνούν μέσω του κόμβου A με την βοήθεια *κ* γραμμών με το γενικό τηλεφωνικό δίκτυο, όπως φαίνεται στο Σχήμα. Κάθε συνδρομητής τηλεφωνεί κατά μέσο όρο 12 λεπτά την ώρα. Τα τηλεφωνήματα είναι ανεξάρτητα μεταξύ τους. Υπολογίστε τον αριθμό *κ* απαιτούμενων γραμμών, ώστε η πιθανότητα για δεδομένη στιγμή ένας συνδρομητής να βρει τις γραμμές κατειλημμένες να είναι μικρότερη του 1 στα 10.

Λύση

Θεωρώ ότι σε οιαδήποτε χρονική στιγμή η πιθανότητα ένας συνδρομητής να ομιλεί στο τηλέφωνο είναι p=12/60=0,20. Έτσι, μπορώ να υποθέσω ότι σε μία τυχούσα στιγμή κάνω n=10 δοκιμές, ελέγχω έναν προς έναν τους 10 συνδρομητές αν απασχολούν ή όχι μία γραμμή. Αν X είναι η τυχαία μεταβλητή η οποία παριστάνει τον αριθμό των συνδρομητών που απασχολούν τηλεφωνική γραμμή για την δεδομένη στιγμή, τότε το πεδίο τιμών της είναι X={0, 1, 2, 3, …, 10}.

Η τυχαία μεταβλητή X ακολουθεί διωνυμική κατανομή και ζητούμε το μικρότερο κ για το οποίο

$$P(X>\kappa)\leq 0{,}10.$$

Η ανίσωση μπορεί να λυθεί ξεκινώντας από την εξίσωση

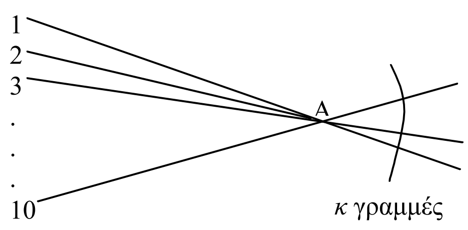
$$P(X>\kappa)=0{,}10$$

$$\text{ή}\quad P(X\leq\kappa)=0{,}90$$

Η πιθανότητα αυτή δίνεται από την αθροιστική κατανομή της διωνυμικής κατανομής. Δοκιμάζουμε για ποια τιμή του κ η παραπάνω πιθανότητα ξεπερνά για πρώτη φορά το 0,90. Έχουμε

Γι α κ=1 $P(X\leq\kappa)=0{,}3758$
Γι α κ=2 $P(X\leq\kappa)= 0{,}6778$
Γι α κ=3 $P(X\leq\kappa)= 0{,}8791$
Γι α κ=4 $P(X\leq\kappa)= 0{,}9672$.

Άρα ο απαιτούμενος αριθμός γραμμών στο τηλεφωνικό κέντρο είναι τουλάχιστον 4.



4) Τηλεφωνικές κλήσεις. Υποθέστε ότι κάθε τηλεφωνική κλήση σας προς ένα δημοφιλή ραδιοφωνικό σταθμό έχει πιθανότητα 0,02 για σύνδεση, δηλαδή να μην είναι απασχολημένη η γραμμή. Υποθέτουμε, ακόμη, ότι οι τηλ. κλήσεις είναι στατιστικά ανεξάρτητες.
(α) Ποια η πιθανότητα να συνδεθείτε με το ραδιοφωνικό σταθμό στην πέμπτη τηλεφωνική σας προσπάθεια;
(β) Ποια η πιθανότητα ότι για να συνδεθείτε με τον ραδιοφωνικό σταθμό απαιτούνται πάνω από 5 τηλεφωνικές κλήσεις;
(γ) Ποιος ο μέσος αριθμός κλήσεων για να συνδεθείτε με τον ραδιοφωνικό σταθμό;

Λύση
(α) Σε κάθε τηλεφωνική προσπάθεια η πιθανότητα σύνδεσης με τον σταθμό είναι σταθερή p=0,02. Θεωρώ την κάθε προσπάθεια σαν μία δοκιμή Bernoulli, και αν η τυχαία μεταβλητή X παριστάνει τον αριθμό δοκιμών μέχρι σύνδεσης, προφανώς η X ακολουθεί **γεωμετρική** κατανομή, οπότε έχουμε

$$P(X=5)(1-p)^4 p=(1-0{,}02)^4 0{,}02=0{,}98^4 0{,}02=0{,}018$$

(β) Εδώ ζητούμε $P(X>5)$. Να μην επιτευχθεί σύνδεση στις 5 πρώτες τηλεφωνικές προσπάθειες, είναι το ισοδύναμο γεγονός του {X>5}. Έτσι έχουμε

$$P(X>5)=(1-p)^5=0{,}98^5=0{,}904\,.$$

(γ) Η μέση τιμή της X ή περίοδος επαναφοράς όπως λέγεται στην γεωμετρική κατανομή είναι

$$E(X)=\frac{1}{p}=\frac{1}{0{,}02}=50 \text{ δοκιμές.}$$

5) Αστοχία ρουκέτας. Αποδείξτε ότι η **γεωμετρική** κατανομή δεν έχει μνήμη. Εάν για παράδειγμα X παριστάνει την φορά που μια ρουκέτα πλήττει τον στόχο της, τότε η ιδιότητα "*έλλειψης-μνήμης*" είναι ισοδύναμη με το γεγονός ότι η επιτυχία σε μια ρουκέτα δεν εξαρτάται από τις προηγούμενες αποτυχίες. Δηλαδή να δειχθεί ότι

$$P(X>s+t\,|\,X>s)=P(X>t)$$

Λύση
Θεωρούμε ότι σε κάθε δοκιμή υπάρχει σταθερή πιθανότητα p η ρουκέτα να πλήξει τον στόχο της. Εφαρμόζουμε τον τύπο της υπό συνθήκη πιθανότητας,

$$P(X>s+t|X>s)=\frac{P(\{X>s+t\}\cap\{X>s\})}{P(X>s)}=\frac{P(X>s+t)}{P(X>s)}$$

$$=\frac{(1-p)^{s+t}}{(1-p)^{s}}=(1-p)^{t}=P(X>t)$$

Τονίζουμε εδώ ότι η πιθανότητα $P(X>s)$ είναι η πιθανότητα του γεγονότος ότι στις πρώτες s δοκιμές δεν έχει πληγεί ο στόχος, και σύμφωνα με το σκεπτικό της **γεωμετρικής** κατανομής είναι $P(X>s)=(1-p)^s$.

6) Το πρόβλημα Banach. Ένας καπνιστής μηχανικός φέρνει μαζί του δύο κουτιά σπίρτα που περιέχουν 50 σπίρτα το καθένα. Για να ανάψει το τσιγάρο του διαλέγει πάντα στην τύχη ένα από τα δύο κουτιά. Ποια η πιθανότητα να διαλέξει για πρώτη φορά ένα άδειο κουτί και το άλλο να είναι επίσης άδειο;

Λύση

Στο πρόβλημα αυτό η εκάστοτε επιλογή ενός κουτιού είναι μία δοκιμή Bernoulli. Επειδή γίνεται τυχαία έχουμε p=0,5. Εάν την πρώτη φορά που διαλέγουμε ένα άδειο κουτί το άλλο έχει κ σπίρτα, τότε έχουμε διαλέξει 50+1 φορές το άδειο και 50-κ το άλλο κουτί. Ας ονομάσουμε με A την επιλογή του ενός και με $\overline{A}$ την επιλογή του άλλου κουτιού. Έστω X={ *το πλήθος των δοκιμών μέχρι της n+1 εμφάνισης του A*}. Η τυχαία μεταβλητή X ακολουθεί **Pascal κατανομή** και η ζητούμενη πιθανότητα δίνεται από τον αντίστοιχο τύπο

$$P(X=2\cdot 50-\kappa+1)=2\cdot\binom{2\cdot 50-\kappa+1-1}{50+1-1}\left(\frac{1}{2}\right)^{50+1}\left(\frac{1}{2}\right)^{50-\kappa}=\binom{2\cdot 50-\kappa}{50}\cdot 2^{\kappa-2\cdot 50}$$

όπου ο συντελεστής 2 εισέρχεται λόγω συμμετρίας των δύο κουτιών.

Τέλος, η ζητούμενη πιθανότητα αντιστοιχεί για κ=0 και το αποτέλεσμα ύστερα από εφαρμογή του τύπου του **Stirling** για τον υπολογισμό του παραγοντικού του 50, 50!, είναι p=0,079

7) Μετρητής Geiger-Muller Σε έναν μετρητή Geiger-Muller καταφθάνουν κατά μέσο όρο 100 σωματίδια το λεπτό. Έστω ότι ο αριθμός των σωματιδίων ακολουθεί κατανομή **Poisson**. Να βρεθούν

(α) η πιθανότητα καταγρ αφής 100 σωματιδίων σε ένα λεπτό

(β) η πιθανότητα καταγραφής 10 σωματιδίων σε ένα δευτερόλεπτο.

Λύση

Συμβολίζω με X_t τον αριθμό σωματιδίων που καταφθάνουν στο χρονικό διάστημα (0, t λεπτά), όπου μονάδα μέτρησης του χρόνου είναι το λεπτό.

Δεδομένου ότι ο μέσος αριθμός σωματιδίων ανά λεπτό είναι λ=100σωματίδια/λεπτό, έχουμε

(α)

$$P(X_{1\lambda\varepsilon\pi\tau\acute{o}}=100)=e^{-\lambda\cdot t}\frac{(\lambda\cdot t)^{100}}{100!}=e^{-100\cdot 1}\frac{(100\cdot 1)^{100}}{100!}$$
$$=e^{-100}\frac{100^{100}}{e^{-100}100^{100}\sqrt{2\pi 100}}=0{,}04$$

όπου το παραγοντικό του 100, 100! υπολογίσθηκε κατά προσέγγιση από τον τύπο του Stirling.

(β)

$$P(X_{1\,\delta\varepsilon\upsilon\tau\varepsilon\rho\acute{o}\lambda\varepsilon\pi\tau o}=10)=e^{-\lambda t}\frac{(\lambda t)^{10}}{10!}=e^{-100\cdot\frac{1}{60}}\frac{(100\cdot\frac{1}{60})^{10}}{10!}$$
$$=e^{-\frac{100}{60}}\frac{\frac{100^{10}}{6^{10}}}{e^{-10}100^{10}\sqrt{2\pi 10}}\approx 0.$$

8) Άφιξη μηνυμάτων. **Μηνύματα καταφθάνουν σε έναν υπολογιστή σύμφωνα με μία διαδικασία Poisson με συχνότητα 10 την ώρα. Προσδιορίστε το μήκος του χρονικού διαστήματος έτσι ώστε η πιθανότητα να μην αφιχθεί μήνυμα στην διάρκεια αυτή να είναι 0,90.**

Λύση

Συμβολίζω με t το ζητούμενο χρονικό διάστημα, και αν X_t είναι η Poisson μεταβλητή που παριστάνει τον αριθμό μηνυμάτων σε χρόνο t, έχουμε,

$$P(X_t=0)=0{,}90$$

$$P(X_t=0)=e^{-10\cdot t}\frac{(10\cdot t)^0}{0!}=e^{-10\cdot t}=0{,}90$$

$$\Rightarrow -10t=\ln(0{,}90)\Rightarrow t=\frac{\ln(0{,}90)}{-10}=1{,}053 \quad \acute{\omega}\rho\varepsilon\varsigma$$

9) Εσοχή για στροφή σε διασταύρωση. Ο αριθμός οχημάτων που φθάνουν τυχαία σε μια διασταύρωση και επιθυμούν να στρίψουν αριστερά είναι τυχαία μεταβλητή και ακολουθεί Poisson κατανομή με μέση τιμή 150 αυτοκίνητα την

ώρα. Η διάρκεια του κόκκινου σήματος για αριστερές στροφές είναι 1 λεπτό και το μήκος της αριστερής λωρίδας χωράει 3 αυτοκίνητα.
(α) ποιες υποθέσεις πρέπει να πληροί μια τυχαία μεταβλητή X για να ακολουθεί Poisson κατανομή;
(β) Ποια η πιθανότητα ότι η αριστερή λωρίδα θα χωρέσει τα αυτοκίνητα που επιθυμούν να στρίψουν αριστερά κατά την διάρκεια του κόκκινου σήματος;
(γ) Κατά πόσο πρέπει να αυξηθεί το μήκος της αριστερής λωρίδας έτσι ώστε με πιθανότητα πάνω από 98% θα χωρούν στην λωρίδα αυτή όλα τα αυτοκίνητα που επιθυμούν να στρίψουν αριστερά;

Λύση

(α) Οι συνθήκες για μία Poisson κατανομή αναφέρονται στην παράγραφο της Poisson κατανομής, κεφάλαιο 6.
(β) Ο μέσος αριθμός αυτοκινήτων για αριστερές στροφές είναι λ=150/60=2,5 αυτοκίνητα/λεπτό. Συμβολίζοντας με X_t τον αριθμό αυτοκινήτων τα οποία θέλουν να στρίψουν αριστερά στο χρονικό διάστημα t, έχουμε

$$P(X_{1\,\lambda\varepsilon\pi\tau\acute{o}} \leq 3) = \sum_{x=0}^{3} P(X_{1\,\lambda\varepsilon\pi\tau\acute{o}} = x) = \sum_{x=0}^{3} e^{-2,5\cdot 1}\frac{(2,5\cdot 1)^x}{x!}$$
$$= 0,0596 + 0,1490 + 0,2240 + 0,2240 = 0,6566$$

(γ) Συμβολίζοντας με κ το ζητούμενο μήκος της εσοχής έχουμε

$$P(X_{1\,\lambda\varepsilon\pi\tau\acute{o}} \leq \kappa) \geq 0,98.$$

Για κ=3 η παραπάνω πιθανότητα δεν ξεπερνά την τιμή 0,98, όπως φαίνεται από την απάντηση στο (β). Δοκιμάζουμε για κ=4, 5, 6, …και επιλέγουμε εκείνο το κ για το οποίο η πιθανότητα ξεπερνά την τιμή 0,98.

Για κ=4 $P(X_{1\,\lambda\varepsilon\pi\tau\acute{o}} \leq 4)=0,82$
Για κ=5 $P(X_{1\,\lambda\varepsilon\pi\tau\acute{o}} \leq 5)=0,91$
Για κ=6 $P(X_{1\,\lambda\varepsilon\pi\tau\acute{o}} \leq 6)=0,96$
Για κ=7 $P(X_{1\,\lambda\varepsilon\pi\tau\acute{o}} \leq 5)=0,9816$.

Άρα για μήκος εσοχής κ=7 αυτοκίνητα, η πιθανότητα ότι χωράνε στην εσοχή τα αυτοκίνητα που επιθυμούν να στρίψουν αριστερά είναι τουλάχιστον 0,98. Το μήκος της εσοχής πρέπει να μεγαλώσει κατά 4 αυτοκίνητα.

Ασκήσεις προς λύση

Διωνυμική κατανομή

1) Μία τυχαία μεταβλητή X έχει διωνυμική κατανομή με n=10 και p=0,5. Σχεδιάστε την συνάρτηση μάζας πιθανότητας f(x).
(α) Ποια τιμή της X είναι η επικρατέστερη;
(β) Προσδιορίστε τις πιθανότητες $P(X=5)$, $P(X<3)$, $P(X>8)$, και $P(3 \leq X<5)$.

2) Μία τυχαία μεταβλητή X έχει διωνυμική κατανομή με n=10 και p=0,01. Σχεδιάστε την συνάρτηση μάζας πιθανότητας f(x) και σχολιάστε για την κυρτότητα της συνάρτησης.
(α) Ποια τιμή της X είναι η επικρατέστερη;
(β) Προσδιορίστε τις πιθανότητες $P(X=5)$, $P(X<3)$, $P(X>8)$, και $P(3 \leq X<5)$.

3) Στο φανάρι της διασταύρωσης, από το οποίο περνάτε κάθε πρωί με το αυτοκίνητό σας, 20% του χρόνου δείχνει πράσινο για την δική σας κατεύθυνση. Υποθέστε ότι το κάθε πρωί παριστάνει μία ανεξάρτητη δοκιμή.
(α) Για πέντε πρωινά, ποια η πιθανότητα ότι μόνο μία φορά δεν σταματήσατε στο φανάρι;
(β) Για είκοσι πρωινά, ποια η πιθανότητα ότι δεν σταματήσατε στο φανάρι περισσότερο από 4 φορές;

4) Σε ένα τεστ πολλαπλής επιλογής, υπάρχουν 25 ερωτήσεις, και για την κάθε ερώτηση αντιστοιχούν 4 απαντήσεις. Υποθέτουμε ότι ο φοιτητής για κάθε ερώτηση επιλέγει τυχαία μία από τι 4 απαντήσεις.
(α) Ποια η πιθανότητα ότι ο φοιτητής απαντά σωστά σε περισσότερες από 20 ερωτήσεις;
(β) Ποια η πιθανότητα ότι ο φοιτητής απαντά σωστά σε λιγότερες από 5 ερωτήσεις;

5) Σε ένα μεγάλο αριθμό καλωδίων, το 80% έχουν διάμετρο 8mm και το 20% έχουν διάμετρο 6mm. Επιλέγονται στην τύχη 10 από τα παραπάνω καλώδια.
(α) Ποια η πιθανότητα ότι ακριβώς 5 έχουν διάμετρο 8mm και 5 έχουν διάμετρο 6mm;
(β) Ποια η πιθανότητα ότι το δείγμα αποτελείται από καλώδια ενός είδους;
(γ) Ποια η πιθανότητα ότι υπάρχουν 6 καλώδια ενός είδους και 4 καλώδια του άλλου είδους;

6) Ένας μαθητής, από πλούσιους γονείς, διατηρεί στο συρτάρι του πολλές σοκολάτες. Περίπου 50% από αυτές είναι *αμυγδάλου*, 30% *γάλακτος* και 20% *φρούτων*. Πια η πιθανότητα ότι από τις δέκα σοκολάτες τις οποίες παίρνει μαζί του μία μέρα στο σχολείο
(α) οι 5 είναι *αμυγδάλου;*
(β) οι 4 είναι *γάλακτος;*
(γ) οι 5 είναι αμυγδάλου , οι 4 είναι *γάλακτος* και 1 είναι *φρούτων*;

Γεωμετρική κατανομή

7) Υποτίθεται ότι μία τυχαία μεταβλητή X ακολουθεί μία γεωμετρική κατανομή με p=0.5. Προσδιορίστε τις ακόλουθες πιθανότητες.
(α) P(X=1)
(β) P(X=4)
(γ) P(X≤2)
(δ) P(X<2)

8) Υποτίθεται ότι μία τυχαία μεταβλητή ακολουθεί γεωμετρική κατανομή με μέση τιμή (περίοδο επαναφοράς) 2,5. Προσδιορίστε τις ακόλουθες πιθανότητες.
(α) P(X=1)
(β) P(X=5)
(γ) P(X≤3)
(δ) P(X>3)

Υπεργεωμετρική κατανομή

9) Υποθέστε ότι μία τυχαία μεταβλητή είναι υπεργεωμετρικής κατανομής με N=100, n=4, και κ=20. Προσδιορίστε τα ακόλουθα.
(α) P(X=1)
(β) P(X=6)
(γ) P(X=4)
(δ) Την μέση τιμή και διακύμανση της X.

10) Κάθε κεντρικό καπάκι ενός τύπου μηχανής περιέχει 4 βίδες. Οι βίδες επιλέγονται τυχαία από ένα κιβώτιο το οποίο περιέχει 30 βίδες ενός προμηθευτή και 70 βίδες από κάποιον άλλον προμηθευτή.
(α) Ποια η πιθανότητα ότι οι βίδες στο καπάκι της μηχανής προέρχονται από τον ίδιο προμηθευτή;
(β) Ποια η πιθανότητα ότι ακριβώς 3 βίδες είναι του ίδιου προμηθευτή;

11) Μία μεγάλη κατασκευαστική εταιρία διαθέτει 75 εκσκαφείς από τους οποίους οι 5 είναι ελαττωματικοί. Επιλέγεται τυχαία ένα σύνολο από δέκα εκσκαφείς για να ξεκινήσουν ένα χωματουργικό έργο.
(α) Ποια η πιθανότητα ότι κανένας εκσκαφέας από το σύνολο αυτό δεν είναι ελαττωματικός;
(β) Ποιος ο μέσος αριθμός ελαττωματικών από τους εκσκαφείς που ξεκίνησαν το έργο;

12) Χρησιμοποιείστε τον διωνυμική προσέγγιση στην υπεργεωμετρική κατανομή για να προσεγγίσετε τις πιθανότητες (α) και (β) της άσκησης 12. Ποιος είναι ο συντελεστής διόρθωσης σε αυτήν την άσκηση;

Αρνητική Διωνυμική ή Pascal κατανομή

13) Ας υποθέσουμε ότι η τυχαία μεταβλητή X ακολουθεί αρνητική διωνυμική (Pascal) κατανομή με $p=0,2$ και $r=4$. Προσδιορίστε τα ακόλουθα.
(α) $E(X)$
(β) $P(X=20)$
(γ) $P(X=19)$
(δ) Την επικρατέστερη τιμή της X.

14) Μία ηλεκτρονική ζυγαριά σε ένα σύστημα αυτόματης διαδικασίας γεμίσματος, σταματά την διαδικασία παραγωγής μετά από την ανίχνευση τριών ελλειποβαρών πακέτων. Υποθέτουμε ότι η πιθανότητα ένα πακέτο να είναι ελλειποβαρές είναι 0,01, και ότι υπάρχει στατιστική ανεξαρτησία στο γέμισμα των πακέτων.
(α) Ποιος είναι ο μέσος αριθμός γεμισμάτων πριν σταματήσει η παραγωγή;
(β) Ποια είναι η τυπική απόκλιση του αριθμού γεμισμάτων πριν σταματήσει η παραγωγή;

Poisson κατανομή

15) Ο αριθμός των πελατών X που εισέρχονται σε μία τράπεζα ανά ώρα είναι τυχαία μεταβλητή, υποθέτουμε ότι ακολουθεί Poisson κατανομή με μέση τιμή $\lambda=10$ ανά ώρα.
(α) Ποια είναι η πιθανότητα άφιξης 5 πελατών σε μία ώρα;
(β) Ποια είναι η πιθανότητα άφιξης το πολύ 3 πελατών σε μία ώρα;
(γ) Ποια η πιθανότητα άφιξης ακριβώς 15 πελατών σε 2 ώρες;

16) Ο αριθμός χαραμάδων που υπάρχουν σε ένα τμήμα ενός διεθνούς δρόμου και χρειάζονται επισκευή ακολουθεί Poisson κατανομή με μέση τιμή 2 χαραμάδες ανά 3 χιλιόμετρα.
(α) Ποια η πιθανότητα ότι δεν υπάρχουν χαραμάδες που χρειάζονται επισκευή σε ένα τμήμα του δρόμου 5 χιλιομέτρων;
(β) Ποια η πιθανότητα ότι υπάρχει τουλάχιστον μία χαραμάδα που χρειάζεται επισκευή σε ένα τμήμα 1500 μέτρων;

17) Ο αριθμός ψεγαδιών που υπάρχουν στην πλαστική επιφάνεια, για την επένδυση του εσωτερικού αυτοκινήτου μιας εταιρείας, ακολουθεί Poisson κατανομή με μέσο όρο 0,05 ανά τετραγωνικό μέτρο. Υποθέτουμε ότι το εσωτερικό του αυτοκινήτου αυτού του τύπου χρειάζεται 5 τετραγωνικά μέτρα πλαστική επιφάνεια για την επένδυσή του.
(α) Ποια η πιθανότητα ότι δεν υπάρχουν ψεγάδια στο εσωτερικό ενός αυτοκινήτου;
(β) Εάν 10 αυτοκίνητα πουλήθηκαν σε μία εταιρεία ενοικίασης αυτοκινήτων, ποια η πιθανότητα ότι κανένα δεν φέρει ψεγάδι στο εσωτερικό του;
(γ) Εάν 10 αυτοκίνητα πουλήθηκαν σε μία εταιρεία ενοικίασης αυτοκινήτων, ποια η πιθανότητα ότι το πολύ ένα αυτοκίνητο φέρει ψεγάδι στο εσωτερικό του;

18) Ο αριθμός μικρών ραγισμάτων που υπάρχουν σε κάποιον τύπο σιδηροδοκού σε ένα μεγάλο κατασκευαστικό έργο ακολουθεί Poisson κατανομή με μέσο όρο 0,02 ανά δοκό.
(α) Εάν 50 σιδηροδοκοί ελέγχονται προσεκτικά, ποια η πιθανότητα ότι δεν υπάρχουν ραγίσματα;
(β) Ποιος ο αναμενόμενος αριθμός σιδηροδοκών που απαιτείται να ελεγχθεί μέχρι να βρεθεί ράγισμα;
(γ) Εάν 50 σιδηροδοκοί ελέγχονται, ποια είναι η πιθανότητα ότι ο αριθμός των σιδηροδοκών που έχουν δύο ή περισσότερα ραγίσματα είναι λιγότερο ή ίσο με δύο;

19) Μηνύματα καταφθάνουν σε έναν υπολογιστή σύμφωνα με μία διαδικασία Poisson με μέση συχνότητα 10 την ώρα. Προσδιορίστε το μήκος ενός χρονικού διαστήματος έτσι ώστε η πιθανότητα να μην αφιχθεί μήνυμα στην διάρκεια αυτού του διαστήματος είναι 0,90.

Προσέγγιση διωνυμικής με Poisson κατανομή

20) Υποτίθεται ότι η πιθανότητα ένα ανταλλακτικό να είναι ελαττωματικό είναι 0,2. Χρησιμοποιείστε την Διωνυμική κατανομή για να υπολογίσετε την πιθανότητα ότι όχι πάνω από ένα ελαττωματικό βρίσκεται σε μια τυχαία επιλογή 10 τέτοιων ανταλλακτικών.

Να εκτιμηθεί η παραπάνω πιθανότητα με Poisson προσέγγιση, και να συγκριθούν τα αποτελέσματα με την Διωνυμική κατανομή.

Προβλήματα προς λύση

1) Ποσότητα αποθήκευσης. Ένας βιομήχανος αποθηκεύει ενός είδους εξαρτήματα τα οποία αγοράζει από έναν προμηθευτή. Υποθέτουμε ότι 2% των εξαρτημάτων αυτών είναι ελαττωματικά και ότι η εμφάνιση ελαττωματικών εξαρτημάτων σχετίζεται με στατιστικώς ανεξάρτητα γεγονότα.
(α) Εάν ο βιομήχανος έχει στην αποθήκη 102 εξαρτήματα, ποια η πιθανότητα ότι οι 100 παραγγελίες θα ικανοποιηθούν χωρίς να ξανά-παραγγείλει εξαρτήματα;
(β) Πόσα εξαρτήματα πρέπει να έχει στην αποθήκη έτσι ώστε η πιθανότητα ότι 100 παραγγελίες μπορούν να ολοκληρωθούν χωρίς να παραγγελθούν άλλα εξαρτήματα είναι τουλάχιστον 0,95;

2) Έλεγχος φορτίου. Μία εταιρεία ανέλαβε την επιθεώρηση των φορτίων που στέλνουν προμηθευτές σε μία βιομηχανία. Υποθέτουμε ότι κάθε φορτίο περιέχει 1000 μονάδες του προιόντος και ότι 1% είναι ελαττωματικό. Ποιο μέγεθος δείγματος(σε μονάδες προιόντος) απαιτείται από κάθε φορτίο έτσι ώστε η πιθανότητα να υπάρξει τουλάχιστον ένα ελαττωματικό στο δείγμα είναι τουλάχιστον 0,90;
Υποθέτουμε ότι η προσέγγιση της διωνυμικής στην υπεργεωμετρική κατανομή είναι επαρκής.

3) Μετασεισμικός έλεγχος δοκαριών. Ο αριθμός ραγισμένων δοκαριών στις πολυκατοικίες μετά από ένα σεισμό ακολουθεί Poisson κατανομή με μέση συχνότητα 0,1 ανά πολυκατοικία.
(α) Εάν ελέγξουμε 100 πολυκατοικίες ποια η πιθανότητα ότι το πολύ σε 5 πολυκατοικίες θα παρατηρηθούν ραγίσματα στα δοκάρια;
(β) Ποιος ο ελάχιστος αριθμός πολυκατοικιών που πρέπει να ελέγξει ο μηχανικός έτσι ώστε να βρεθούν 8 πολυκατοικίες με ραγισμένα δοκάρια;

4) Καταστροφικοί σεισμοί. Σε μία σεισμογενή περιοχή έχουμε κατά μέσο όρο 2 σεισμούς κάθε 10 χρόνια. Υποθέτουμε ότι οι σεισμοί συμβαίνουν με μία διαδικασία Poisson. Ακόμη, είναι γνωστό ότι το 60% των σεισμών είναι καταστρεπτικοί.

(α) Ποια είναι η πιθανότητα να συμβούν το πολύ 2 σεισμοί στα 10 επόμενα χρόνια;
(β) Ποια είναι η πιθανότητα να υπάρξει καταστρεπτικός σεισμός στα επόμενα 20 χρόνια:

5) Νεροποντή και πλημμύρα. Σε μία περιοχή ο αριθμός ισχυρών νεροποντών ακολουθεί Poisson κατανομή με μέσο αριθμό 1 νεροποντή στα 5 χρόνια,
(α) ποια είναι η πιθανότητα ότι στα επόμενα 10 χρόνια θα έχουμε καμία νεροποντή; μία νεροποντή; πάνω από 2 νεροποντές;

Στην περιοχή αυτή βρίσκεται ένας ποταμός ο οποίος όταν συμβαίνει ισχυρή νεροποντή ξεχειλίζει με πιθανότητα 0,08.
(β) Να υπολογιστεί η πιθανότητα πλημμύρας του ποταμού σε 5 ισχυρές νεροποντές.
(γ) Να υπολογιστεί η πιθανότητα πλημμύρας του ποταμού στα επόμενα 5 χρόνια.

6) Ζημιές από απεργίες. Οι απεργίες στον κλάδο των οικοδόμων συμβαίνουν με μία διαδικασία Poisson, με μέσο όρο μία κάθε 4 χρόνια . Μία απεργία μπορεί να διαρκέσει ισοπίθανα από 5 μέχρι 15 ημέρες και εκτιμάται ότι η ζημία για κάθε ημέρα απεργίας ανέρχεται σε 1000 ευρώ.
(α) Ποια η αναμενόμενη ζημία στον εργολάβο όταν συμβαίνει απεργία;
(β) Ποια η αναμενόμενη ζημία στον εργολάβο από απεργίες για τα επόμενα 2 χρόνια:

7) Απεργία στα πρατήρια υγρών καυσίμων. Κατά μήκος ενός αυτοκινητόδρομου ο αριθμός πρατηρίων υγρών καυσίμων ακολουθεί Poisson κατανομή, με 2 πρατήρια ανά 25 χιλιόμετρα κατά μέσο όρο. Λόγω απεργιακών κινητοποιήσεων, η πιθανότητα ένα πρατήριο να έχει διαθέσιμη βενζίνη είναι 0,80. Υποθέτουμε ότι σχετικά με την διαθεσιμότητα της βενζίνης υπάρχει στατιστική ανεξαρτησία μεταξύ των πρατηρίων.
(α) Ποια είναι η πιθανότητα ότι στα επόμενα 40 χιλιόμετρα θα συναντήσουμε πρατήριο βενζίνης;
(β) Ποια η πιθανότητα ότι από τα επόμενα 3 πρατήρια ένα μόνο θα έχει βενζίνη;
(γ) Ένα αυτοκίνητο ταξιδεύει στον αυτοκινητόδρομο και απέχει από τον προορισμό του 200 χιλιόμετρα, ενώ το απόθεμα βενζίνης αρκεί για 60 χιλιόμετρα. Ποια η πιθανότητα να φθάσει στον προορισμό του χωρίς πρόβλημα βενζίνης;

8) Αντοχή κατασκευής σε σφοδρούς ανέμους. Από τις πληροφορίες που έχουμε για την μέγιστη ετήσια ταχύτητα V του ανέμου σε μία πόλη,

συμπεραίνουμε ότι μέσα σε ένα χρόνο υπάρχει σταθερή πιθανότητα $P(V>90$ χιλμ./ώρα$)=0{,}02$.
(α) Ποια η πιθανότητα, ότι στα επόμενα 100 χρόνια στην πόλη αυτή, θα υπάρξουν το πολύ τέσσερα χρόνια όπου η ετήσια μέγιστη ταχύτητα θα ξεπεράσει τα 90 χιλμ./ώρα;
(β) Αν μία κατασκευή στην πόλη αυτή αντέχει σε μία ταχύτητα ανέμου μέχρι 90 χιλμ./ώρα, ποια η πιθανότητα ότι η κατασκευή θα αντέξει το πολύ 20 χρόνια;
(γ) Πόσα χρόνια θα αντέξει στους σφοδρούς ανέμους η κατασκευή της ερώτησης (β) με πιθανότητα πάνω από 90%;

9) Αριθμός επιβίβασης και αποβίβασης σε λεωφορείο. Ένα λεωφορείο μεταφέρει n το πλήθος επιβάτες. Στην επόμενη στάση του ο κάθε επιβάτης κατεβαίνει με πιθανότητα p ανεξάρτητα από τους άλλους. Ο αριθμός κ των επιβατών που ανεβαίνουν στη στάση αυτή ακολουθεί την κατανομή:

$$P(\kappa=i)=\begin{cases}\dfrac{5}{12} & \gamma\iota\alpha \quad i=0\\ \dfrac{1}{i+2} & \gamma\iota\alpha \quad i=1,2\\ 0 & \alpha\lambda\lambda o\upsilon\end{cases}$$

Ποια η πιθανότητα μετά την αναχώρηση από την στάση αυτή ο αριθμός των μεταφερομένων επιβατών να είναι πάλι n;

10) Βλάβες σε δομικές μηχανές. Σε ένα ηλεκτρικό σύστημα Α ενός εργοστασίου ο αριθμός βλαβών του συστήματος είναι τυχαία μεταβλητή και ακολουθεί Poisson κατανομή με μέση τιμή $\lambda_A=1{,}1$ βλάβες/έτος. Υποτίθεται, ότι όταν το σύστημα παρουσιάζει βλάβη επισκευάζεται αμέσως και τίθεται πάλι σε λειτουργία.
(α) Ποια η πιθανότητα ότι το σύστημα Α θα παρουσιάσει 2 βλάβες σ ένα εξάμηνο;
(β) Στο ίδιο εροστάσιο λειτουργεί και ένα άλλο ηλεκτρικό σύστημα Β παρόμοια με το Α. Ο αριθμός βλαβών του Β ακολουθεί επίσης Poisson κατανομή με μέση τιμή $\lambda_B=1{,}4$ βλάβες/έτος. Ποια η πιθανότητα ότι το επόμενο έτος το σύστημα Β θα παρουσιάσει λιγότερες βλάβες από ότι το σύστημα Α;

11) Αστοχία από πλημμύρες. Σε μία περιοχή οι πλημμύρες συμβαίνουν σύμφωνα με μία διαδικασία *Poisson* με μέσο χρονικό διάστημα, μεταξύ δύο διαδοχικών πλημμύρων, 10 χρόνια.

(α) Ποια η πιθανότητα να μην έχουμε πλημμύρα τα επόμενα 10 χρόνια στην περιοχή αυτή;
(β) Στην περιοχή αυτή βρίσκεται μια κατασκευή η οποία σε περίπτωση πλημμύρας ενδέχεται να αστοχήσει με πιθανότητα 0,08. Να υπολογιστεί η πιθανότητα αστοχίας της κατασκευής σε διάστημα 8 ετών.

12) Σχεδιασμός γέφυρας. Μία μεγάλη γέφυρα πρόκειται να κατασκευασθεί σένα ποταμό. Το κριτήριο σχεδίασης είναι ότι η ανύψωση επιφάνειας του νερό λόγω πλημμύρας ξεπερνά τις βάσεις της γέφυρας το πολύ μία φορά στα 25 χρόνια με πιθανότητα 0.15.
(α) Ποια η περίοδος επιστροφής μίας τέτοιας πλημμύρας που πρέπει να χρησιμοποιηθεί στον σχεδιασμό της γέφυρας;
(β) Ποια η πιθανότητα ότι η ανύψωση της επιφάνειας του ποταμού λόγω πλημμύρας δεν θα ξεπεράσει τις βάσεις της γέφυρας σε χρονικό διάστημα ίσο με την περίοδο επιστροφής που συμπεριλαμβάνεται στον σχεδιασμό.;

13) Αξιοπιστία ύδρευσης. Η ύδρευση μιας πόλης εξασφαλίζεται από τρεις ανεξάρτητες παροχές, Π_1, Π_2 και Π_3, όπου η κάθε παροχή δύναται να καλύψει το 50%, 60% και 70% των αναγκών της πόλης αντίστοιχα. Από παρατηρήσεις σχετικά με την αξιοπιστία των παροχών, εκτιμάται ότι η πιθανότητα για την κάθε παροχή να τροφοδοτήσει με νερό την πόλη για κάθε έτος είναι 0,90, 0,80 και 0,80 αντίστοιχα.
(α) Ποια η πιθανότητα ότι η παροχή της πόλης θα αστοχήσει (δηλ. καλύπτεται λιγότερο από 100%) κατά την διάρκεια ενός έτους.
(β) Ποια η πιθανότητα ότι στα επόμενα 10 χρόνια η πόλη θα έχει ικανοποιητική παροχή τουλάχιστον σε 8 από αυτά;
(γ) Για πόσα χρόνια δεν θα συναντήσουμε πρόβλημα ύδρευσης στην πόλη με πιθανότητα πάνω από 90%;

14) Αντοχή στα κύματα. Μία θαλάσσια κατασκευή έχει κατασκευαστεί έτσι ώστε να αντιμετωπίζει κύματα μέχρι 6 μέτρα πάνω από την μέση επιφάνεια της θάλασσας. Αν το ύψος των κυμάτων ξεπεράσει τα 6 μέτρα η κατασκευή αστοχεί .

Κάθε φορά που συμβαίνει τρικυμία στην περιοχή το ύψος των κυμάτων πάνω από την μέση επιφάνεια της θάλασσας, H, είναι τυχαίο μέγεθος και εκτιμάται $P(H>6$ μέτρα$)=0,25$. Ακόμη, είναι γνωστό ότι στην περιοχή ο αριθμός τρικυμιών ακολουθεί Poisson κατανομή με μέση τιμή 2 τρικυμίες/έτος.
(α) Να υπολογιστεί η πιθανότητα ότι σε 5 τρικυμίες που θα συμβούν στην περιοχή, το πολύ σε δύο από αυτές το ύψος των κυμάτων θα ξεπεράσει τα 6 μέτρα.

(β) Να υπολογιστεί η πιθανότητα ότι η κατασκευή δεν θα αστοχήσει λόγω τρικυμίας μέσα στον επόμενο χρόνο.

15) Αντικατάσταση εξαρτημάτων. Η διάρκεια λειτουργίας T ενός ηλεκτρικού εξαρτήματος είναι τυχαία μεταβλητή. Στο τέλος της κάθε έτους γίνεται επισκευή του εξαρτήματος και εκτιμάται ότι η πιθανότητα να λειτουργήσει χωρίς διακοπή στον επόμενο χρόνο είναι $P(T>1$ έτος$)= 0{,}90$.

Για ένα σύστημα που αποτελείται από 6 τέτοια ηλεκτρικά εξαρτήματα,
(α) ποια η πιθανότητα ότι δεν θα προκύψει ανάγκη αντικατάστασης ηλεκτρικού εξαρτήματος μέσα σε ένα έτος;
(β) ποια η πιθανότητα ότι θα προκύψει ανάγκη αντικατάστασης ηλεκτρικού εξαρτήματος (ή ηλεκτρικών εξαρτημάτων) για πρώτη φορά εντός του τρίτου έτους λειτουργίας του συστήματος;

16) Επισκευές σε οδόστρωμα. Η πιθανότητα βλάβης ενός χιλιομέτρου οδοστρώματος μέσα σε ένα χρόνο είναι πάντα σταθερή και ίση με 0,10. Όταν το οδόστρωμα παρουσιάσει βλάβη επισκευάζεται αμέσως. Υποθέτουμε ακόμη ότι οι βλάβες διαφορετικών χιλιομέτρων οδοστρώματος είναι στατιστικά ανεξάρτητες.
(α) Να υπολογισθεί η πιθανότητα ότι 10 χιλιόμετρα οδοστρώματος θα χρειαστούν επισκευές μέσα σε ένα χρόνο.
(β) Να υπολογισθεί η πιθανότητα ότι σε ένα τμήμα οδοστρώματος 4 χιλιομέτρων, μόνο 2 από αυτά θα χρειαστούν επισκευές μέσα σε 3 χρόνια.
(γ) Να υπολογισθεί η πιθανότητα ότι για ένα τμήμα οδοστρώματος 4 χιλιομέτρων η πρώτη επισκευή θα γίνει μέσα στον δεύτερο χρόνο.

17) Ασφαλέστερη πτήση. Θεωρούμε ότι σε μία πτήση, οι μηχανές του αεροπλάνου αποτυγχάνουν με πιθανότητα q και ότι οι αποτυχίες συμβαίνουν ανεξάρτητα από μηχανή σε μηχανή. Υποθέτοντας ότι το αεροπλάνο μπορεί να πραγματοποιήσει επιτυχή πτήση αν λειτουργούν τουλάχιστον οι μισές από τις μηχανές του, για ποια τιμή του q είναι
(α) το μονοκινητήριο αεροπλάνο προτιμότερο από ένα τρικινητήριο;
(β) το δικινητήριο αεροπλάνο προτιμότερο από ένα τετρακινητήριο;

18) Κατειλημμένες τηλεφωνικές γραμμές. Οι τηλεφωνικές γραμμές για κράτηση θέσης σε μία αεροπορική εταιρία είναι κατειλημμένες 40% του χρόνου. Υποθέτουμε ότι τα γεγονότα οι γραμμές να είναι κατειλημμένες ή όχι σε κάθε κλήση είναι στατιστικά ανεξάρτητα. Γίνονται δέκα τηλεφωνικές κλήσεις προς την εταιρία.
(α) Ποια η πιθανότητα ότι για τις τρεις κλήσεις οι γραμμές είναι κατειλημμένες
(β) Ποια η πιθανότητα ότι για τουλάχιστον μία κλήση οι γραμμές δεν είναι κατειλημμένες;

(γ) Ποιος ο αναμενόμενος αριθμός κλήσεων που βρίσκουν κατειλημμένες τις γραμμές;

19) Πώληση εισιτηρίων στις αεροπορικές πτήσεις. Μία αεροπορική εταιρεία γνωρίζει ότι το 5% από τους ταξιδιώτες που αγοράζουν εισιτήριο για την πτήση μιας συγκεκριμένης γραμμής δεν θα ταξιδέψουν. Έτσι πουλάει για κάθε πτήση 52 εισιτήρια ενώ υπάρχουν πάντα διαθέσιμες 50 θέσεις.
(α) Ποια η πιθανότητα ότι σε μία πτήση θα υπάρξει θέση για κάθε ταξιδιώτη;
(β) Πόσα το πολύ εισιτήρια πρέπει να πουλάει η εταιρεία για κάθε πτήση, έτσι ώστε το πρόβλημα με τις θέσεις θα εμφανίζεται κατά μέσο όρο κάθε 10(ή και >10)πτήσεις;

20) Αξιοπιστία ηλεκτρικού συστήματος. Υποθέτουμε ότι η διάρκεια ζωής (σε ώρες) ενός τύπου ηλεκτρικής λυχνίας είναι συνεχής τυχαία μεταβλητή Χ με συνάρτηση πυκνότητας πιθανότητας

$$f(x)=\begin{cases}\dfrac{\kappa}{x^2} & \gamma\iota\alpha \quad x>100\\ 0 & \alpha\lambda\lambda o\upsilon\end{cases}$$

(α) Ποια η πιθανότητα ότι η λυχνία θα λειτουργήσει λιγότερο από 200 ώρες, αν είναι γνωστό ότι λειτουργεί επιτυχώς χωρίς διακοπή για 150 ώρες;
(β) Ποια η πιθανότητα ότι σε μία συσκευή με 3 τέτοιες λυχνίες ακριβώς μία θα πρέπει να αντικατασταθεί σε διάστημα λειτουργίας (0, 150);
(γ) Δίδεται συσκευή που περιέχει n από τις ανωτέρω λυχνίες. Υπολογίστε το μέγιστο n, ώστε η πιθανότητα μετά από 150 ώρες λειτουργίας της συσκευής συνεχίζουν να λειτουργούν όλες οι λυχνίες χωρίς βλάβη είναι τουλάχιστον 0,5.

21) Ανελκυστήρας. Τέσσερα άτομα εισέρχονται σε ανελκυστήρα στο ισόγειο μιας πενταόροφης πολυκατοικίας. Έστω A_{ij} το γεγονός: {το άτομο i εξέρχεται στον όροφο j}, όπου i=1, 2, 3, 4, και j=1, 2, 3, 4, 5. Με την υπόθεση ότι τα γεγονότα A_{ij} είναι ισοπίθανα και ανεξάρτητα, υπολογίστε
(α) την πιθανότητα όλα τα άτομα να εξέλθουν στον πρώτο όροφο,
(β) την πιθανότητα όλα τα άτομα να εξέλθουν στον ίδιο όροφο,
(γ) την πιθανότητα τα άτομα i=1, i=2, και i=3 να εξέλθουν στον ίδιο όροφο.

22) Πάσσαλοι σε πετρώδες έδαφος. Για την θεμελίωση ενός οικοδομικού συγκροτήματος σκοπεύουμε να μπήξουμε μεγάλους σιδερένιους πασσάλους μέχρι 20 μέτρα. Το υπέδαφος μέχρι αυτό το βάθος διακρίνεται σε τέσσερα στρώματα των 5 μέτρων έκαστο, και η πιθανότητα ο κάθε πάσσαλος να συναντήσει μεγάλη πέτρα σε κάθε στρώμα είναι σταθερή p=0,08. Υποθέτουμε ότι μεταξύ στρωμάτων υπάρχει στατιστική ανεξαρτησία.

(α) Ποια η πιθανότητα ένας πάσσαλος να φτάσει το βάθος 20 μέτρων χωρίς να συναντήσει πέτρα;
(β) Ποια η πιθανότητα ότι ο πάσσαλος θα συναντήσει πέτρα το πολύ σε δύο από τα στρώματα:
(γ) Αν για τη θεμελίωση του συγκροτήματος χρειάζονται 10 πάσσαλοι, ποια η πιθανότητα ότι πέτρα θα συναντήσουν το πολύ τρεις από αυτούς;

23) Εκφόρτωση πλοίων. Πλοία φθάνουν σε ένα μεγάλο λιμάνι για εκφόρτωση με μέση συχνότητα 6 πλοία την ημέρα. Έχει παρατηρηθεί ότι οι εγκαταστάσεις των γερανών επιτρέπουν την δυνατότητα εκφόρτωσης 6 πλοίων την ημέρα. Εάν έρθουν μία μέρα περισσότερα από 6 πλοία, αναχωρούν για άλλο λιμάνι. Υποθέτουμε ότι ο αριθμός άφιξης πλοίων στο λιμάνι την ημέρα είναι τυχαία μεταβλητή Poisson κατανομής.
(α) Ποια η πιθανότητα μία μέρα να ξεφορτώσουν όλα τα πλοία που ήρθαν στο λιμάνι;
(β) Ποιος ο μέσος αριθμός πλοίων που ξεφορτώνουν την ημέρα στο λιμάνι αυτό;
(γ) Ποιος ο μέσος αριθμός πλοίων τον μήνα αναχωρεί για εκφόρτωση σε άλλο λιμάνι;

24) Αναχώρηση τουριστικού λεωφορείο. Ένα μικρό τουριστικό λεωφορείο ξεκινά την ξενάγηση σε μία ωραία πόλη από την αφετηρία του μόλις συμπληρωθούν 12 αφίξεις τουριστών. Υποθέτουμε ότι οι αφίξεις τουριστών στην αφετηρία είναι ανεξάρτητες μεταξύ τους και η μέση συχνότητα άφιξης είναι 9 την ώρα. Ποια η πιθανότητα ότι ο χρόνος μεταξύ δύο διαδοχικών αναχωρήσεων είναι μεγαλύτερος από 60 λεπτά;

25) Αξιοπιστία μηχανής παραγωγής πασσάλων. Η πιθανότητα ότι μία μηχανή που κατασκευάζει πασσάλους παρουσιάζει βλάβη είναι 0,0002 για κάθε 100 μέτρα παραγόμενων πασσάλων.
(α) Ποια η πιθανότητα να παρουσιάσει βλάβη για πρώτη φορά μεταξύ της παραγωγής 1000 με 1100 μέτρων πασσάλων;
(β) Αν η μηχανή έχει την δυνατότητα να παράγει 1000 μέτρα πασσάλους την εβδομάδα, ποια η πιθανότητα ότι η μηχανή θα παρουσιάσει βλάβη για πρώτη φορά μετά από τρεις μήνες;

26) Έλεγχος διαδικασίας παραγωγής. Δείγματα των 20 τεμαχίων από μία διαδικασία διάτρησης μετάλλου επιλέγεται κάθε ώρα. Τυπικά το 10% των παραγόμενων τεμαχίων απαιτεί πάλι επεξεργασία. Συμβολίστε με X τον αριθμό που δηλώνει πόσα τεμάχια από τα 20 απαιτούν πάλι επεξεργασία. Πρόβλημα στην διαδικασία παραγωγής υπάρχει αν το X ξεπερνά την μέση τιμή του περισσότερο από τρις φορές την τυπική του απόκλιση.

(α) Αν το ποσοστό των τεμαχίων που απαιτούν πάλι επεξεργασία παραμένει στο 10%, ποια η πιθανότητα να υπάρξει πρόβλημα στην διαδικασία παραγωγής.
(β) Αν το ποσοστό των τεμαχίων που απαιτούν πάλι επεξεργασία ελαττωθεί στο 4%, ποια η πιθανότητα να μην υπάρξει πρόβλημα στην διαδικασία παραγωγής στις επόμενες 6 ώρες;.

7

Θεωρητικές Κατανομές Πιθανότητας για Συνεχή Τυχαία Μεταβλητή.

ΠΕΡΙΓΡΑΦΗ ΚΕΦΑΛΑΙΟΥ

Σε αυτό το κεφάλαιο θα παρουσιάσουμε τις *θεωρητικές κατανομές πιθανότητας* (μερικές από_αυτές τις συναντήσαμε στα προηγούμενα κεφάλαια-ομοιόμορφη, εκθετική κ.τ.λ.), με την έννοια ότι αυτές προέρχονται από λογικές αιτίες παρά από πραγματικές εργαστηριακές παρατηρήσεις.

Θα επικεντρωθούμε περισσότερο σε ειδικές κατανομές πιθανότητας τυχαίων μεταβλητών, οι οποίες είναι κατάλληλα μαθηματικά μοντέλα για καταστάσεις του πραγματικού κόσμου κάτω από ειδικές συνθήκες. Στην ανάλυση που θα ακολουθήσει θα προσδιορίσουμε την ακριβή μαθηματική έκφραση της κάθε ειδικής κατανομής και τις υποθέσεις κάτω από τις οποίες αυτά τα μοντέλα περιγράφουν μία κατάσταση στον πραγματικό κόσμο. Θα αναφερθούμε επίσης σε στατιστικούς πίνακες, έτσι ώστε να περιορίσουμε την υπολογιστική δουλειά.

Ο αναγνώστης σε αυτή τη φάση πρέπει να παρατηρήσει ότι είναι πολύ σπουδαίο γι' αυτόν να συνδέσει το πρόβλημα από τον πραγματικό κόσμο με το σωστό μοντέλο κατανομής παρά να αποστηθίζει την λεπτομερή μαθηματική ανάπτυξη του μοντέλου.

Στο κεφάλαιο αυτό θα αναπτύξουμε πέντε συνεχείς κατανομές: την ομοιόμορφη, την εκθετική, την κανονική, την λογαριθμοκανονική, και βήτα κατανομή. Η ομοιόμορφη κατανομή είναι πολύ χρήσιμη σε πολλές εφαρμογές του μηχανικού αλλά και πολύ σπουδαία γιατί χρησιμοποιείται για την δημιουργία τυχαίων αριθμών στην προσομοίωση. Οι άλλες συνεχείς κατανομές όπως η κανονική και λοραριθμοκανονική είναι επίσης σπουδαίες αφού χρησιμοποιούνται στα περισσότερα στοχαστικά μεγέθη στις μελέτες του μηχανικού. Ακόμη, η εκθετική κατανομή έχει ιδιαίτερο ενδιαφέρων διότι προέρχεται από την Poisson. Τέλος, αναφέρονται περιληπτικά και άλλες κατανομές όπως η βήτα, η Weibull, η t-student, και chi-square οι οποίες είναι σπουδαίες επειδή χρησιμοποιούνται σε πολλά κεφάλαια στην στατιστική, όπως αξιοπιστία, εκτιμητική, έλεγχο στατιστικών υποθέσεων και αλλού.

7.1 ΟΜΟΙΟΜΟΡΦΗ ΚΑΤΑΝΟΜΗ

Η ομοιόμορφη κατανομή είναι η απλούστερη από όλες τις κατανομές, και μας είναι γνώριμη από μερικά παραδείγματα στο κεφάλαιο 5. Μία τυχαία μεταβλητή X της οποίας η τιμή μπορεί να βρεθεί **τυχαία** ή **ισοπίθανα** σε οποιοδήποτε σημείο ενός πεπερασμένου διαστήματος $[\alpha, \beta]$-για παράδειγμα, X={πραγματικός αριθμός x: $\alpha<x<\beta$} – έχει μία ομοιόμορφη κατανομή ή η συνάρτηση πυκνότητας πιθανότητας είναι σταθερή σε όλο το διάστημα από α μέχρι β. Εύκολα, μπορεί κανείς να προσδιορίσει την συνάρτηση πυκνότητας πιθανότητας $f(x)$ μίας ομοιόμορφης τυχαίας μεταβλητής X, αφού είναι μία σταθερά c και το ολοκλήρωμά της από α μέχρι β πρέπει να ισούται με 1,

$$\int_{\alpha}^{\beta} f(x)dx = \int_{\alpha}^{\beta} c dx = 1 \quad \Rightarrow \quad c = \frac{1}{\beta-\alpha}$$

Ορισμός: Μία συνεχής τυχαία μεταβλητή ***X***, η οποία μπορεί να πάρει τιμές παντού στο διάστημα [***α,β***], όπου ***α*** και ***β*** πεπερασμένοι αριθμοί, ονομάζεται **ομοιόμορφη μεταβλητή** εάν η συνάρτηση πυκνότητας πιθανότητας δίνεται από την εξίσωση,

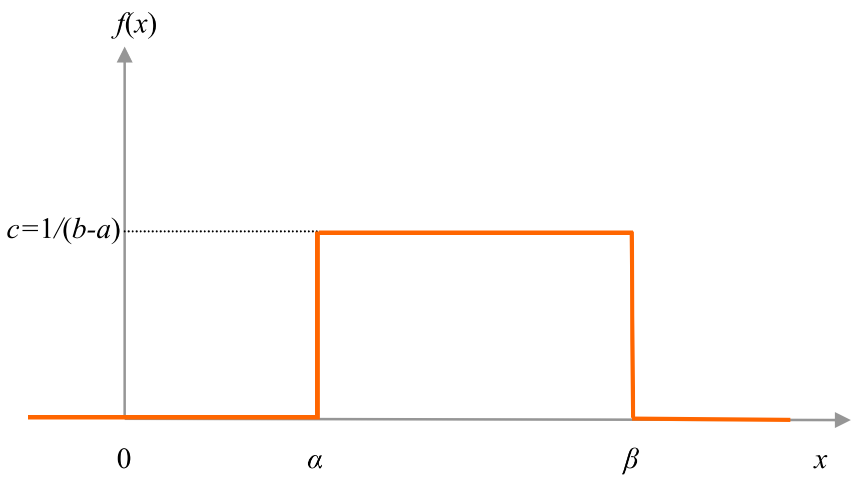
$$f(x) = \begin{cases} \dfrac{1}{\beta-\alpha} & \gamma\iota\alpha \quad \alpha \le x \le \beta \\ 0 & \alpha\lambda\lambda o\upsilon \end{cases} \tag{7.1}$$

Λέμε, επίσης, ότι η X είναι ομοιόμορφα κατανεμημένη στο διάστημα $[\alpha, \beta]$.

Εύκολα επιτυγχάνουμε και την αθροιστική συνάρτηση πιθανότητας για το ομοιόμορφο μοντέλο

$$F(x)=\begin{cases}0 & \alpha\nu \quad x<\alpha \\ \dfrac{x-\alpha}{\beta-\alpha} & \alpha\nu \quad \alpha\le x\le\beta \\ 1 & \alpha\nu \quad x>\beta.\end{cases} \tag{7.2}$$

Το Σχήμα 7.1 δίνει την συνάρτηση πυκνότητας της συνεχούς ομοιόμορφης κατανομής.



Σχήμα 7.1 Συνάρτηση πυκνότητας πιθανότητας f(x), για μία ομοιόμορφη τυχαία μεταβλητή X στο διάστημα (α, β).

Παρατήρηση

Μία ομοιόμορφη τυχαία μεταβλητή X παίρνει ισοπίθανα τιμές μέσα στο διάστημα $[\alpha, \beta]$, η έννοια αυτή μπορεί να ερμηνευθεί ώς ακολούθως. Η πιθανότητα η X να πάρει τιμή μέσα σε οποιοδήποτε υπό-διάστημα (γ, δ), όπου $\alpha<\gamma<\delta<\beta$, είναι η ίδια, εφόσον τα υπό-διαστήματα αυτά έχουν το ίδιο μήκος δ-γ. Η πιθανότητα, δηλαδή, εξαρτάται από το εύρος του διαστήματος (γ, δ) και όχι από πού αρχίζει και που τελειώνει το υπό-διάστημα, όπως φαίνεται στην εξίσωση

$$P(\gamma < X \leq \delta) = \int_{\gamma}^{\delta} f(x)dx = \frac{\delta - \gamma}{\beta - \alpha}$$

Έτσι, η πιθανότητα για κάθε γεγονός το οποίο κατέχει ένα υποδιάστημα του διαστήματος [*α, β*] είναι απλά ο λόγος του μήκους του υποδιαστήματος προς το μήκος του [*α, β*]. ■

Η ομοιόμορφη κατανομή έχει σαν παραμέτρους το *α* και *β*. Η μέση τιμή και διακύμανσή της *X*, με βάση τους ορισμούς αυτών, μπορούν να προσδιορισθούν όπως στο κεφάλαιο 5 ,

$$E(X) = \frac{\alpha + \beta}{2}$$
$$Var(X) = \frac{(\beta - \alpha)^2}{12} \qquad (7.3)$$

Σε μία ομοιόμορφη κατανομή η μέση τιμή της *X* είναι το ενδιάμεσο σημείο του διαστήματος [*α, β*] και συμπίπτει με την διάμεσό της.

Παράδειγμα 7.1

Κατά την διάρκεια της ημέρας, τα τραίνα μίας συγκεκριμένης γραμμής σε έναν υπόγειο σιδηρόδρομο περνούν κάθε μισή ώρα. Ποία η πιθανότητα ένας άνδρας που εισέρχεται τυχαία κατά την διάρκεια της ημέρας στον υπόγειο να περιμένει τουλάχιστον 20 λεπτά;

Η τυχαία μεταβλητή *X*={*χρόνος μέχρι το επόμενο τραίνο*}, υπό την προϋπόθεση ότι ο άνδρας φθάνει τυχαία στον σταθμό, ακολουθεί την ομοιόμορφη κατανομή και μπορεί να πάρει τιμές μέσα στο διάστημα [0, 30] λεπτά. Έτσι η πιθανότητα την οποία ζητάμε να βρούμε είναι

$$P(X \geq 20) = 1 - F(20) = 1 - \frac{20 - 0}{30} = \frac{1}{3}.$$

Ο μέσος χρόνος αναμονής, προφανώς, είναι 15 λεπτά. ■

Παράδειγμα 7.2

Από προηγούμενες πληροφορίες ή στατιστικά δεδομένα ένας μηχανικός πιστεύει ότι η αντοχή του σκυροδέματος που χρησιμοποιείται για την κατασκευή μίας γέφυρας βρίσκεται ισοπίθανα σε οποιοδήποτε σημείο μεταξύ 3000 και 4000 psi.

Εδώ, μπορεί να χρησιμοποιηθεί η ομοιόμορφη κατανομή αφού η αντοχή *X* είναι τυχαία μεταβλητή και μπορεί να πάρει ισοπίθανα τιμές στο διάστημα

[3000, 4000] psi. Έτσι, μπορεί να προσδιορισθεί η συνάρτηση πυκνότητας πιθανότητας *f(x)* ,

$$f(x) = \begin{cases} \dfrac{1}{4000 - 3000} = 0{,}001 & \textit{για} \quad 3000 \le x \le 4000\,psi \\ 0 & \textit{αλλου} \end{cases}.$$

Η μέση αντοχή του σκυροδέματος είναι 3500 psi, και η τυπική της απόκλιση 288,7 psi. Η πιθανότητα ότι η αντοχή του σκυροδέματος θα ξεπεράσει τα 3900 psi είναι

$$P(X > 3900) = 1 - P(X \le 3900) = 1 - \frac{3900 - 3000}{4000 - 3000} = 0{,}100$$

■

7.1.1 Δημιουργία τυχαίων αριθμών

Εδώ αναφέρουμε μία πολύ σημαντική εφαρμογή της ομοιόμορφης κατανομής στην δημιουργία ενός δείγματος τιμών x_1, x_2, x_3, …, x_n από μία τυχαία μεταβλητή *X* οιασδήποτε θεωρητικής κατανομής.

Είναι γνωστές αρκετές μέθοδοι οι οποίες δημιουργούν τυχαίους αριθμούς από μία ομοιόμορφη κατανομή (ομοιόμορφα κατανεμημένους) στο διάστημα [0, 1]. Κάθε ηλεκτρονικός υπολογιστής, π.χ., έχει πρόγραμμα το οποίο δίνει τυχαία δείγματα από μία ομοιόμορφη κατανομή στο [0, 1].

Ας υποθέσουμε ότι μία συνεχής τυχαία μεταβλητή *X* έχει συνάρτηση πυκνότητας πιθανότητας $f_X(x)$ και αθροιστική κατανομή πιθανότητας $F_X(x)$. Ακόμη, θεωρούμε και μία ομοιόμορφα κατανεμημένη μεταβλητή *U* στο διάστημα (0, 1).

Επιθυμούμε να βρούμε μία συνάρτηση *G* η οποία για κάθε τυχαίο αριθμό *u* της ομοιόμορφης να δίνει ένα τυχαίο αριθμό *x* της $f_X(x)$, δηλαδή, $x = G(u)$, έτσι ώστε

$$P(X \le x) = P(U \le u)\,.$$

Αλλά, είναι γνωστό ότι

$$P(X \le x) = F_X(x)$$

$$P(U \le u) = F_U(u) = \frac{u - 0}{1 - 0} = u$$

δηλαδή, θέλουμε να ισχύει

$$F_X(x) = F_U(u) = u$$

$$F_X(x) = u$$

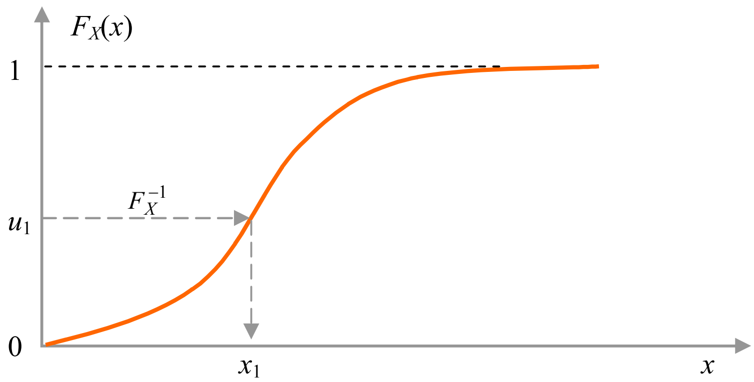
$$x = F_X^{-1}(u)$$

Με άλλα λόγια, η συνάρτηση που επιθυμώ να προσδιορίσω είναι η, $G = F_X^{-1}$,δηλαδή η αντίστροφη της αθροιστικής πιθανότητας της *X*.

Έτσι, για να δημιουργήσω ένα δείγμα τιμών x_1, x_2, x_3, …, x_n από μία τυχαία μεταβλητή με αθροιστική κατανομή πιθανότητας $F_X(x)$, ακολουθώ τα εξής βήματα:

- Βρίσκω την αναλυτική μορφή της $F_X(x)$,
- Ορίζω την αντίστροφη $G = F_X^{-1}$
- Επιλέγω τυχαία μία τιμή *u* από την ομοιόμορφη [0, 1],
- θέτω $x = F_X^{-1}(u)$

Ο μηχανισμός δημιουργίας τυχαίου δείγματος που μόλις αναπτύξαμε περιγράφεται γραφικά και στο Σχήμα 7.2. Επιλέγεται τυχαία μία τιμή u_1 από την ομοιόμορφη [0, 1], στη συνέχεια με την χρήση της F_X^{-1} μεταφέρεται στο άξονα x, δίνοντας μία τυχαία τιμή x_1 για την *X*.



Σχήμα 7.2. Γραφική παράσταση δημιουργίας τυχαίας τιμής x_1 από μία τυχαία τιμή της ομοιόμορφης στο διάστημα [0, 1].

Παράδειγμα 7.3

Σε μία τυχαία μεταβλητή X, η οποία παριστάνει τον χρόνο καλής λειτουργίας ενός συστήματος η αθροιστική κατανομή πιθανότητας είναι $F_X(x) = 1 - e^{-\lambda x}$. Για να δημιουργήσουμε ένα τυχαίο δείγμα το οποίο θα παριστάνει τους χρόνους λειτουργίας *n* τέτοιων συστημάτων ακολουθούμε τα εξής βήματα:

- Με τον υπολογιστή μας επιτυγχάνουμε από την ομοιόμορφη [0, 1] τυχαίο δείγμα τιμών u_1, u_2, …, u_n.

- *Θέτουμε για κάθε τιμή* $u_κ$, $u_κ = 1 - e^{-\lambda x_κ}$,
- Μετασχηματίζουμε ως προς $x_κ$, $x_κ = -\frac{1}{\lambda}\ln(1-u_κ)$.

Οι προκύπτουσες τιμές x_1 , x_2 , .. ,x_n, είναι τυχαίο δείγμα τιμών από την μεταβλητή *X,* η οποία ακολουθεί την κατανομή $F_X(x) = 1 - e^{-\lambda x}$.

7.2 ΕΚΘΕΤΙΚΗ ΚΑΤΑΝΟΜΗ

Η εκθετική κατανομή έχει σαν βάση την διαδικασία Poisson. Υπενθυμίζουμε ότι η Poisson πιθανότητα αναφέρεται στον αριθμό «επιτυχιών» σε συγκεκριμένο πεπερασμένο χρονικό διάστημα, όπου ο αριθμός «επιτυχιών» είναι η τυχαία μεταβλητή. Τώρα, εάν αντιστρέψουμε τους ρόλους της Poisson μεταβλητής με τον χρόνο, έχουμε το ονομαζόμενο **εκθετικό μοντέλο**. Ειδικότερα, μία **εκθετική τυχαία μεταβλητή** *X* είναι το χρονικό διάστημα μεταξύ δύο διαδοχικών «επιτυχιών». Για παράδειγμα, εάν ο αριθμός αφίξεων αυτοκινήτων στα διόδια ενός δρόμου ακολουθεί Poisson κατανομή, τότε ο χρόνος μεταξύ δύο διαδοχικών αφίξεων είναι εκθετική μεταβλητή.

Το εκθετικό μοντέλο είναι πολύ σημαντικό στην εφαρμοσμένη πιθανότητα και στατιστική επειδή μπορεί να χρησιμοποιηθεί για να περιγράψει μία μεγάλη ποικιλία στοχαστικών καταστάσεων. Το εκθετικό μοντέλο είναι το ποιό κατάλληλο για φαινόμενα όπως η διάρκεια ζωής μιας ηλεκτρικής λάμπας, η χρονική διάρκεια μεταξύ δύο διαδοχικών βλαβών ενός ηλεκτρικού μηχανισμού, το χρονικό διάστημα μέχρι τον επόμενο σεισμό, ο χρόνος ανάμεσα σε δύο συνεχόμενα αυτοκινητιστικά ατυχήματα ή η απόσταση ανάμεσα σε δύο συνεχόμενα ατυχήματα σε ένα δρόμο και πολλά άλλα.

Από τα παραδείγματα παραπάνω βλέπουμε ότι το εκθετικό μοντέλο πιθανότητας προκύπτει από το εξής ερώτημα: Εάν μία ακολουθία από γεγονότα συμβαίνουν στον χρόνο με τον νόμο του Poisson με συχνότητα λ γεγονότα στην μονάδα του χρόνου, πόσο θα περιμένουμε μέχρι την πρώτη εμφάνιση του γεγονότος; Η απάντηση σε αυτό το ερώτημα προτείνει μία μέθοδο κατασκευής του εκθετικού μοντέλου πιθανότητας από την Poisson διαδικασία. Ας πούμε ότι η τυχαία μεταβλητή *X* παριστάνει τον χρόνο μεταξύ δύο διαδοχικών βλαβών μίας μηχανής, και επιχειρούμε να προσδιορίσουμε την συνάρτηση κατανομής πυκνότητας από τον υπολογισμό της πιθανότητας $P(X>x)$, για κάθε συγκεκριμένο διάστημα χρόνου x.Το γεγονός $\{X>x\}$ είναι ισοδύναμο με το γεγονός {*καμία βλάβη στο διάστημα x* }, και κατά συνέπεια από τον νόμο Poisson έχουμε

$$P(X > x) = P(\{0\ \beta\lambda\alpha\beta\varepsilon\varsigma\ \sigma\tau o\ \delta\iota\alpha\sigma\tau\eta\mu\alpha\ (0, x)\}$$
$$= e^{-\lambda x}\frac{(\lambda x)^0}{0!} = e^{-\lambda x}.$$

Έτσι, η αθροιστική συνάρτηση πιθανότητας της εκθετικής μεταβλητής X μπορεί να γραφεί

$$F(x) = P(X \le x) = 1 - e^{-\lambda x}. \tag{7.4}$$

Απ' την (7.4) με παράγωγο ως προς x, εύκολα μπορεί να προσδιορισθεί η συνάρτηση πυκνότητας πιθανότητας $f(x)$,

$$f(x) = \begin{cases} \dfrac{d}{dx}F(x) = \lambda e^{-\lambda x} & \gamma\iota\alpha \quad x \ge 0 \\ 0 & \alpha\lambda\lambda o\upsilon \end{cases}$$

Ορισμός: Η τυχαία μεταβλητή ***X***, η οποία παριστάνει την απόσταση μεταξύ διαδοχικών εμφανίσεων Poisson γεγονότων, ακολουθεί **εκθετική κατανομή** με παράμετρο $\boldsymbol{\lambda}$,

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \gamma\iota\alpha \quad 0 \le x < \infty \\ 0 & \alpha\lambda\lambda o\upsilon \end{cases} \tag{7.5}$$

Η συνάρτηση $f(x)$ στην εξίσωση (7.5) είναι πράγματι συνάρτηση πυκνότητας πιθανότητας, αφού μία απευθείας ολοκλήρωση βεβαιώνει ότι

$$\int_0^\infty f(x)dx = \int_0^\infty \lambda e^{-\lambda x}dx = 1.$$

Η μέση τιμή της X επιτυγχάνεται από την σχέση του ορισμού της,

$$E(X) = \int_0^\infty x \cdot \lambda e^{-\lambda x}dx.$$

Ολοκληρώνοντας μερικά, όπως στο παράδειγμα 5.7 έχουμε

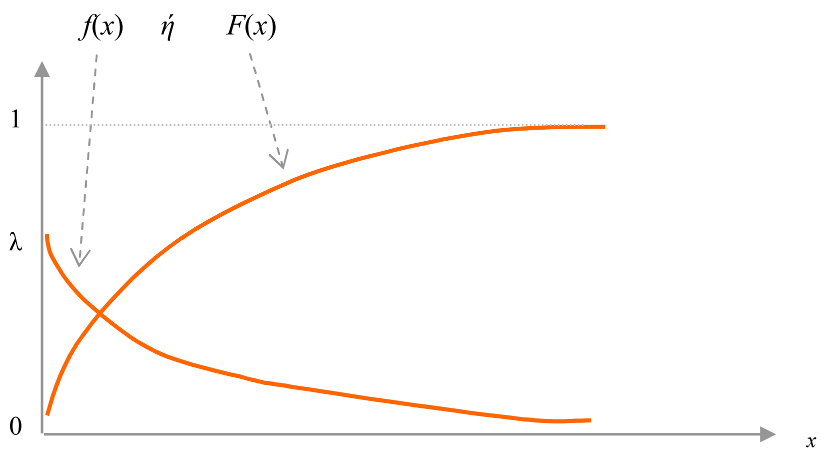
$$E(X) = [-xe^{-\lambda x}]_0^\infty + \int_0^\infty e^{-\lambda x} dx = 0 + \frac{1}{\lambda} = \frac{1}{\lambda}. \quad (7.6)$$

το οποίο σημαίνει ότι ο μέσος χρόνος επαναφοράς ή **περίοδος επαναφοράς** ενός γεγονότος σε μία διαδικασία **Poisson** είναι **1/*λ***. Το αποτέλεσμα αυτό είναι παρόμοιο εκείνου σε μία ακολουθία **Bernoulli** όπου η **περίοδος επαναφοράς** είναι **1/*p***. Το γεγονός ότι $E(X)=1/\lambda$ είναι ένα φυσιολογικό αποτέλεσμα, αφού λ είναι η συχνότητα ή ο αριθμός εμφάνισης Poisson γεγονότων στην μονάδα χρόνου. Η εκθετική μεταβλητή X είναι ο χρόνος μεταξύ διαδοχικών Poisson γεγονότων και το πηλίκο $1/\lambda$ παριστάνει την μέση τιμή του χρόνου αυτού όταν λ γεγονότα πραγματοποιούνται στο χρονικό διάστημα 1. Τέλος μία πρακτική ερμηνεία της μέσης τιμής μίας εκθετικής μεταβλητής σε μία διαδικασία Poisson είναι ότι παριστάνει, για παράδειγμα, τον μέσο χρόνο μέχρι την πρώτη βλάβη μίας μηχανής, ή τον μέσο χρόνο μεταξύ διαδοχικών βλαβών.

Η **διακύμανση** της X μπορεί να επιτευχθεί με παρόμοια ολοκλήρωση. βρίσκουμε $E(X^2) = 2/\lambda^2$, όπως στο παράδειγμα 5.7, και ως εκ τούτου

$$Var(X) = \sigma_X^2 = E(X^2) - [E(X)]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}. \quad (7.7)$$

Στο Σχήμα 7.2 φαίνονται οι συναρτήσεις πυκνότητας και αθροιστικής κατανομής πιθανότητας μίας εκθετικής μεταβλητής X.



Σχήμα 7.3. Συναρτήσεις πυκνότητας και αθροιστικής κατανομής ιθανότητας μιας εκθετικής μεταβλητής X.

Παράδειγμα 7.4

Είναι γνωστό από υπάρχοντα δεδομένα ότι από το 1900 μέχρι 1980 σε μία περιοχή στην Αμερική συνέβησαν 120 ισχυροί τυφώνες. Αν θεωρήσουμε ότι οι τυφώνες συμβαίνουν με μία διαδικασία Poisson, ποια η πιθανότητα να μην έχουμε τυφώνα τα επόμενα 5 χρόνια;

Η συχνότητα των ισχυρών τυφώνων είναι

$$\lambda = \frac{120}{80} = 1{,}5 \text{ τυφώνες /χρόνο.}$$

Το χρονικό διάστημα X μεταξύ διαδοχικών τυφώνων είναι εκθετική μεταβλητή, για αυτό έχουμε

$$P(X > 5) = e^{-\lambda 5} = e^{-1{,}5(5)} = 0{,}000565$$

Ποια η περίοδος επαναφοράς ενός ισχυρού τυφώνα στην περιοχή αυτή;
Η μέση τιμή του χρόνου μεταξύ δύο διαδοχικών τυφώνων είναι

$$E(X) = \frac{1}{1{,}5} = \frac{2}{3} \text{ έτη=8 μήνες.}$$

Ποια η πιθανότητα να εμφανισθεί ισχυρός τυφώνας στους επόμενους 8 μήνες (διάστημα ίσο με την περίοδο επαναφοράς);
Επειδή, το χρονικό διάστημα αναφέρεται σε μήνες υπολογίζουμε την συχνότητα λ ως προς μήνα,

$$\lambda = \frac{120}{960} = 0{,}125 \text{ τυφώνες / μήνα.}$$

Έτσι έχουμε,

$$P(X < 8) = 1 - e^{-0{,}125(8)} = 1 - e^{-1} = 0{,}632.$$

Αξίζει να τονίσουμε εδώ ότι σε κάθε Poisson διαδικασία, το γεγονός μπορεί να συμβεί σε λιγότερο από την περίοδο επαναφοράς του με πιθανότητα πάντοτε 0,632. Με άλλα λόγια, αν ο χρόνος καλής λειτουργίας ενός μηχανήματος, για παράδειγμα, είναι εκθετική μεταβλητή, η πιθανότητα το μηχάνημα να παρουσιάσει βλάβη πριν παρέλθει η μέση(αναμενόμενη) διάρκεια λειτουργίας του είναι 0,632. Διότι,

$$P(X < E(X)) = P\left(X < \tfrac{1}{\lambda}\right) = 1 - e^{-\lambda\left(\frac{1}{\lambda}\right)} = 1 - e^{-1} = 0{,}632.$$ ■

Έλλειψη μνήμης

Η εκθετική κατανομή έχει μία πολύ ενδιαφέρουσα ιδιότητα, ανάλογη με την ιδιότητα της γεωμετρικής κατανομής (6.10). Για να κατανοήσουμε την ιδιότητα αυτή αναφέρουμε πρώτα ένα παράδειγμα.

Αν X είναι ο χρόνος καλής λειτουργίας ενός υπολογιστή, ο οποίος ακολουθεί εκθετική κατανομή με μέση τιμή 360 ώρες, η συχνότητα βλαβών (παράμετρος λ) είναι λ=1/360 βλάβες/ώρα. Η πιθανότητα ο υπολογιστής να λειτουργήσει 100 ώρες από την έναρξη της λειτουργίας του, χωρίς καμία διακοπή είναι

$$P(X > 100) = 1 - F(100) = e^{-\lambda 100} = e^{-\frac{1}{360}100} = 0{,}757.$$

Υποθέτουμε, τώρα, ότι ανοίγουμε τον υπολογιστή, ο οποίος λειτουργεί χωρίς καμία βλάβη για 200 ώρες. Ποια η πιθανότητα να λειτουργήσει άλλες 100 ώρες χωρίς βλάβη;

Επειδή παρήλθαν 200 ώρες χωρίς βλάβη, διαισθάνεται κανείς ότι βλάβη θα εμφανισθεί στις επόμενες ώρες και με μικρότερη βεβαιότητα θα λειτουργήσει ο υπολογιστής στις 100 ώρες που ακολουθούν. Δηλαδή, η πιθανότητα ο υπολογιστής να λειτουργήσει άλλες 100 ώρες χωρίς βλάβη είναι μικρότερη από 0,757. Παρόλα αυτά, σε μία εκθετική κατανομή αυτό δεν είναι αλήθεια. Η ζητούμενη πιθανότητα μπορεί να εκφρασθεί σαν μία υπό συνθήκη πιθανότητα και εύκολα να υπολογισθεί,

$$P(X > 200 + 100 \mid X > 200) = \frac{P[(X > 300) \cap (X > 200)]}{P(X > 200)} = \frac{P(X > 300)}{P(X > 200)}$$

$$= \frac{e^{-\frac{300}{360}}}{e^{-\frac{200}{360}}} = e^{-\frac{100}{360}} = P(X > 100).$$

Μετά την συνεχή λειτουργία των 200 ωρών, η πιθανότητα ο υπολογιστής να συνεχίσει άλλες 100 ώρες ακόμη χωρίς βλάβη, ισούται με την πιθανότητα καλής λειτουργίας του υπολογιστή για περισσότερο από 100 ώρες μετά την έναρξη της λειτουργίας του. Το γεγονός ότι ο υπολογιστής λειτούργησε 200 ώρες ή η παλαιότητα του υπολογιστή δεν επιδρά στην πιθανότητα να λειτουργήσει στις ακόλουθες 100 ώρες.

Το παραπάνω παράδειγμα περιέγραψε την ιδιότητα έλλειψης μνήμης στην εκθετική κατανομή. Την ιδιότητα αυτή, από τις συνεχείς τυχαίες μεταβλητές την κατέχει μόνο η εκθετική κατανομή.

$$P(X > s+t \mid X > s) = P(X > t) \tag{7.8}$$

Η απόδειξη της (7.8) είναι παρόμοια με την απόδειξη στο παράδειγμα, και θα ήταν αρκετά συναρπαστική για τον αναγνώστη!.

Η ιδιότητα έλλειψης μνήμης δεν αποτελεί έκπληξη, αφού η εκθετική μεταβλητή X παριστάνει χρόνο μεταξύ διαδοχικών επαναλήψεων του γεγονότος σε μία Poisson διαδικασία. Στην διαδικασία αυτή, το πόσες φορές θα πραγματοποιηθεί το γεγονός στο επόμενο διάστημα δεν εξαρτάται από πόσες φορές πραγματοποιήθηκε προηγουμένως. Μία πρακτική ερμηνεία της έλλειψης μνήμης είναι ότι οι βλάβες, π.χ. στον υπολογιστή του παραπάνω παραδείγματος, δεν οφείλονται στην παλαιότητα του μηχανήματος, αλλά σε τυχαίες αιτίες πιθανόν άγνωστες, οι οποίες εμφανίζονται με μία διαδικασία Poisson.

Παρατηρήσεις

1) Η εκθετική κατανομή είναι **ανάλογη** της γεωμετρικής κατανομής. Υπενθυμίζουμε ότι η γεωμετρική μεταβλητή παριστάνει τον αριθμό δοκιμών Bernoulli μέχρι την εμφάνιση «επιτυχίας», ανάλογα και η εκθετική μεταβλητή σαν συνεχής παριστάνει χρόνο ή γενικά απόσταση μέχρι την πραγματοποίηση «επιτυχίας» ή γεγονότος σε μία Poisson διαδικασία.
2) Η μέση τιμή (**περίοδος επαναφοράς**) τόσο στην **γεωμετρική** όσο και στην **εκθετική** παριστάνουν γενικά την απόσταση μεταξύ διαδοχικών πραγματοποιήσεων του συγκεκριμένου γεγονότος. Βέβαια, στην γεωμετρική μετράμε την απόσταση με αριθμό δοκιμών Bernoulli, ενώ στην εκθετική η απόσταση είναι συνεχής μεταβλητή παριστάνοντας χρόνο, μήκος, επιφάνεια, όγκο, κ.τ.λ..
3) Η εκθετική κατανομή χρησιμοποιείται συχνά στην μελέτη **αξιοπιστίας** σαν μοντέλο για τον χρόνο μέχρι αποτυχίας ενός συστήματος, αποτελούμενο από πολλά μηχανήματα. Σέ ένα μηχάνημα του οποίου ο χρόνος λειτουργίας του είναι εκθετική μεταβλητή, με την ιδιότητα **έλλειψης μνήμης**, η πιθανότητα να λειτουργήσει, για παράδειγμα, 1000 ώρες όσο παλιό και αν είναι, είναι ίδια με την πιθανότητα να λειτουργήσει 1000 ώρες μετά την έναρξη της λειτουργίας του. Αυτό συμβαίνει επειδή το μηχάνημα δεν φθείρεται με την πάροδο του χρόνου, η δε βλάβη του οφείλεται στην εμφάνιση τυχαίων άλλων παραγόντων. Εάν ένα μηχάνημα έχει φθορά με την πάροδο του χρόνου λειτουργίας του, λογικά, ο χρόνος καλής λειτουργίας του X δεν ακολουθεί εκθετική κατανομή.
4) Σε μία διαδικασία Poisson ενδέχεται η συχνότητα λ εμφάνισης του γεγονότος A να **μην είναι σταθερή**. Στην περίπτωση αυτή τεμαχίζουμε το διάστημα σε μικρότερα διαστήματα, έτσι ώστε να προκύψουν

σταθερές συχνότητες στα αντίστοιχα υπό-διαστήματα. Για παράδειγμα, εάν η εκθετική μεταβλητή παριστάνει τον χρόνο μεταξύ δύο διαδοχικών αφίξεων αεροπλάνων σε ένα αεροδρόμιο, μπορούμε λογικά να χρησιμοποιήσουμε μεγαλύτερη συχνότητα άφιξης αεροπλάνων την ημέρα από ότι την νύχτα.

5) Μία γενικότερη μορφή εκθετικής κατανομής προκύπτει αν η μικρότερη τιμή της X είναι μεγαλύτερη από ε, x>ε. Σε αυτή την περίπτωση, η κατανομή καλείται **μετατοπισμένη εκθετική κατανομή** με συνάρτηση πυκνότητας πιθανότητας

$$f(x)=\begin{cases}\lambda e^{-\lambda(x-\varepsilon)} & \gamma\iota\alpha \quad \varepsilon\le x<\infty\\ 0 \quad \alpha\lambda\lambda o\upsilon & \end{cases}$$

Παράδειγμα 7.5

Υποθέτουμε ότι η εμφάνιση ισχυρών σεισμών σε μία συγκεκριμένη περιοχή ακολουθεί Poisson κατανομή με συχνότητα λ=0,04 σεισμοί/ έτος.

Ποια η πιθανότητα να συμβεί ακριβώς ένας ισχυρός σεισμός σε ένα χρόνο; Η ζητούμενη πιθανότητα δίνεται από το μοντέλο της Poisson κατανομής,

$$p=(\lambda t)\frac{(\lambda t)^1}{1!}=(0{,}04\cdot 1)\frac{(0{,}04\cdot 1)^1}{1!}=0{,}0016.$$

Ποια η πιθανότητα να έχουμε τουλάχιστον ένα σεισμό σε ένα χρόνο; Μπορούμε να χρησιμοποιήσουμε και πάλι το μοντέλο Poisson. Όμως, αν X είναι η τυχαία μεταβλητή που παριστάνει τον χρόνο μέχρι την εμφάνιση ισχυρού σεισμού, τότε η X ακολουθεί εκθετική κατανομή και η ζητούμενη πιθανότητα υπολογίζεται από την αθροιστική της συνάρτηση,

$$P(X<1)=1-e^{-0{,}04(1)}=1\text{-}0{,}9609\text{=}0{,}031.$$

Έχουν περάσει 50 χρόνια χωρίς να συμβεί σεισμός. Ποια η πιθανότητα να συμβεί σεισμός τον επόμενο χρόνο; Η ζητούμενη πιθανότητα είναι $P(X<1)$=0,031, όπως και προηγούμενα. Η πιθανότητα στην εκθετική κατανομή δεν εξαρτάται από πόσα έτη περάσανε χωρίς να συμβεί σεισμός. Αν για παράδειγμα, συμβεί τώρα ισχυρός σεισμός, η πιθανότητα να φανεί σεισμός τον επόμενο χρόνο είναι πάλι η ίδια, δηλαδή $P(X<1)$=0,031.

Ο Μέσος χρόνος μεταξύ διαδοχικών ισχυρών σεισμών είναι

$$E(X)=\frac{1}{\lambda}=\frac{1}{0{,}04}=25 \text{ έτη}.$$

Δηλαδή, η περίοδος επαναφοράς ενός ισχυρού σεισμού στην περιοχή αυτή είναι 25 έτη. Αυτό δεν σημαίνει ότι κάθε 25 χρόνια εμφανίζεται μόνο ένας σεισμός, στην πραγματικότητα μπορεί να μην συμβεί ή να συμβούν περισσότεροι του ενός ισχυροί σεισμοί σε διάστημα 25 ετών. Παρόλα αυτά, κατά μέσο όρο θα έχουμε ένα σεισμό κάθε 25 χρόνια. ■

Παράδειγμα 7.6

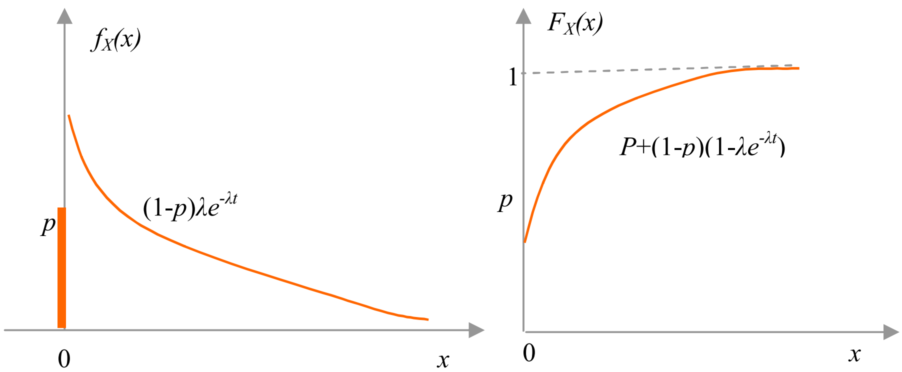
(Θεωρία ουρών) Ο χρόνος αναμονής X ενός πελάτη σε ένα σύστημα ουράς είναι μηδέν εάν αυτός βρίσκει το σύστημα κενό, και κατανέμεται εκθετικά εάν αυτός βρίσκει το σύστημα απασχολημένο. Οι πιθανότητες ότι αυτός βρίσκει το σύστημα απασχολημένο είναι p και 1-p, αντίστοιχα. Να βρεθεί η *α.σ.π.* της X, $F_X(x)$.

$$
\begin{aligned}
F_X(x) &= P(X \le x) \\
&= P(X \le x | \sigma\acute{\upsilon}\sigma\tau.\, \kappa\varepsilon\nu\acute{o})\, p + P(X \le x | \sigma\acute{\upsilon}\sigma\tau.\, \alpha\pi\alpha\sigma\chi o\lambda\eta\mu\acute{\varepsilon}\nu o)(1-p)
\end{aligned}
$$

όπου η τελευταία εξίσωση ισχύει λόγω του θεωρήματος ολικής πιθανότητας. Δεδομένου ότι $P(X \le x |$ σύστ. κενό$)=1$ όταν $x \ge 0$ και 0 σε άλλη περίπτωση, έχουμε

$$
F_X(x) = \begin{cases} 0 & x < 0 \\ p + (1-p)(1-e^{\lambda t}) & x \ge 0. \end{cases}
$$

Η *α.σ.π.* δείχνεται στο ακόλουθο σχήμα. Παρατηρούμε ότι η $F_X(x)$ μπορεί να εκφρασθεί σαν ένα άθροισμα σκαλωτής συνάρτησης με άλμα p και συνεχής συνάρτηση της x.



Σχήμα 7.4 **Γραφικές παραστάσεις για την *σ.π.π.* και *α.σ.π.* της τυχαίας μεταβλητής *X* για το πρόβλημα της αναμονής του πελάτη, παράδειγμα 7.6**

7.3 ΚΑΤΑΝΟΜΕΣ ERLANG και ΓΑΜΜΑ

Στην περίπτωση του αριθμού των ανεξάρτητων δοκιμών Bernoulli μέχρι την πραγματοποίηση ακριβώς r «επιτυχιών» είχαμε μία επέκταση από την γεωμετρική, η οποία σχετίζεται μόνο με μία «επιτυχία», στην αρνητική διωνυμική μεταβλητή. Παρόμοια, όταν σχετιζόμαστε με συνεχή τυχαία μεταβλητή, ενδεχομένως να μας ενδιαφέρει το χρονικό διάστημα μέχρι την πραγματοποίηση ακριβώς r γεγονότων Α σε μία διαδικασία Poisson. Αυτό οδηγεί σε μία κατανομή γνωστή σαν Erlang κατανομή(O Erlang ήταν ο πρώτος μηχανικός από την Δανία , ο οποίος πριν 100 περίπου χρόνια ήταν γνωστός για την δουλειά του στις τηλεφωνικές επικοινωνίες).

Έστω X_r ο χρόνος μέχρι την πραγματοποίηση r γεγονότων σε μία διαδικασία Poisson. Με βάση την Poisson κατανομή μπορούμε να προσδιορίσουμε την αθροιστική κατανομή πιθανότητας για την Erlang μεταβλητή X_r, υπολογίζοντας την πιθανότητα $P(X_r<x)$:

$$\begin{aligned} F_{X_r}(x) &= P(X_r < x) = \sum_{\kappa=r}^{\infty} P\{\kappa \text{ γεγονοτα } A \text{ σε χρονο } x\} \\ &= 1 - \sum_{\kappa=0}^{r-1} \frac{(\lambda x)^{\kappa}}{\kappa!} e^{-\lambda x} \end{aligned} \tag{7.9}$$

Από την αθροιστική συνάρτηση κατανομής με παραγώγιση ως προς x προκύπτει η συνάρτηση πυκνότητας όπως φαίνεται στην εξίσωση (7.10).

Ορισμός: Μία τυχαία μεταβλητή $\boldsymbol{X_r}$, η οποία παριστάνει το χρονικό διάστημα μέχρι την εμφάνιση ακριβώς $\boldsymbol{r}$ γεγονότων σε μία διαδικασία Poisson, καλείται **Erlang μεταβλητή.**

Σε μία **Erlang κατανομή** η συνάρτηση πυκνότητας πιθανότητας είναι:

$$f(x) = \begin{cases} \dfrac{\lambda(\lambda x)^{r-1} e^{-\lambda x}}{(r-1)!} & \text{για} \quad x>0 \quad \text{και} \quad r = 1, 2, 3, \ldots, \\ 0 & \text{αλλου} \end{cases} \tag{7.10}$$

Ο παρανομαστής $(r$-1)! στην (7.10), είναι το γινόμενο από τους πρώτους (r-1) φυσικούς αριθμούς. Μπορεί να γραφεί με την φόρμα $\boldsymbol{\Gamma(r)}$. Αυτό εφαρμόζεται επίσης και σε **μη ακέραιες τιμές του $\boldsymbol{r}$** και καλείται **συνάρτηση Γάμμα** όταν ορίζεται ως

$$\Gamma(r)=\int_0^{\infty} t^{r-1}e^{-t}dt \quad \gamma\iota\alpha \quad r>0 \tag{7.11}$$

Ολοκληρώνοντας κατά μέρη, οδηγούμαστε στο αποτέλεσμα $\Gamma(r+1)=r\Gamma(r)$ για κάθε $r > 0$. Έτσι, φαίνεται ότι η Γάμμα συνάρτηση έχει μία ενδιαφέρουσα επαναληπτική σχέση. Αν υποθέσουμε ότι r είναι ακέραιος αριθμός, έχουμε

$$\begin{aligned}\Gamma(r)&=(r-1)\Gamma(r-1)\\&=(r-1)(r-2)\Gamma(r-2)\\&=(r-1)(r-2)\ldots\Gamma(1)\end{aligned}$$

Αλλά επειδή από την (7.11) για r=1 προκύπτει εύκολα ότι $\Gamma(1)=1$, έχουμε

$$\Gamma(r)=(r-1)!$$

Άρα λοιπόν, αν r είναι ακέραιος αριθμός, η Γάμμα συνάρτηση είναι μία γενίκευση του παραγοντικού.

Ορισμός: Αν ***r*** είναι ένας πραγματικός αριθμός, μία τυχαία μεταβλητή $\boldsymbol{X_r}$ ακολουθεί την **Γάμμα κατανομή** εάν έχει συνάρτηση πυκνότητας πιθανότητας την ακόλουθη,

$$f(x)=\begin{cases}\dfrac{\lambda(\lambda x)^{r-1}e^{-\lambda x}}{\Gamma(r)} & \gamma\iota\alpha \quad x>0 \quad \kappa\alpha\iota \quad r>0\\ 0 & \alpha\lambda\lambda o\upsilon\end{cases} \tag{7.12}$$

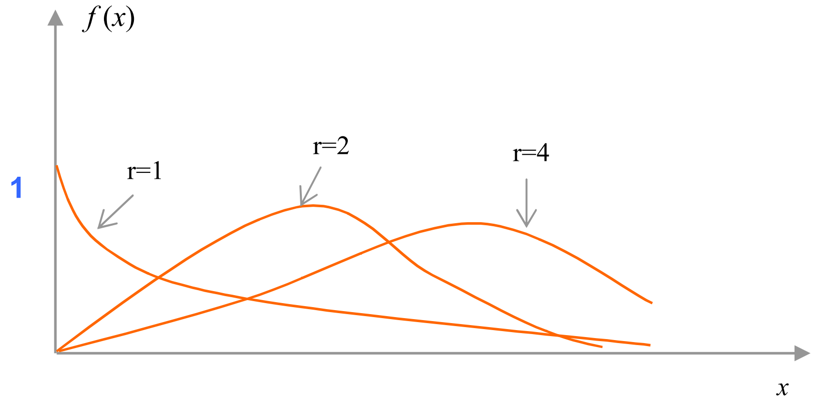
Η Γάμμα κατανομή είναι μία γενίκευση της Erlang κατανομής, αφού είναι ίδιες με την μόνη διαφορά ότι η Γάμμα ισχύει για όλες τις θετικές τιμές του *r,* ενώ η Erlang μόνο για θετικές ακέραιες. Στην πράξη συνήθως αναφερόμαστε στην Γάμμα κατανομή χρησιμοποιώντας ακέραιες τιμές για το *r*.

Η κατανομή Γάμμα εξαρτάται από δύο παραμέτρους, r και λ από τις οποίες απαιτούμε $r > 0$ και $\lambda > 0$. Από τον ορισμό (7.11) της Γάμμα συνάρτησης, προκύπτει εύκολα ότι η (7.12) είναι πράγματι συνάρτηση πυκνότητας πιθανότητας, δηλαδή

$$\int_{-\infty}^{\infty} f(x)dx=1.$$

Η παράμετρος ***r*** είναι γνωστή και σαν **παράμετρος μορφής,** διότι σχετίζεται με την μορφή της συνάρτησης πυκνότητας πιθανότητας. Για r=1 η συνάρτηση πυκνότητας πιθανότητας (7.12) γίνεται εκθετική, $f(x)=\lambda e^{-\lambda x}$.

Καθώς αυξάνει η τιμή του r η συνάρτηση πυκνότητας αρχίζει να γίνεται πιο συμμετρική. Στο Σχήμα 7.4 φαίνονται οι συναρτήσεις πυκνότητας πιθανότητας για διαφορετικές τιμές του r και λ=1.



Σχήμα 7.5. Συναρτήσεις πυκνότητας πιθανότητας Γάμμα κατανομής για r=1, r=2, r=4 και λ=1.

Παρατηρήσεις

1) Για r=1 η συνάρτηση πυκνότητας της Γάμμα γίνεται $f(x)=\lambda e^{-\lambda x}$. Επομένως η εκθετική κατανομή είναι μία ειδική περίπτωση της Γάμμα κατανομής

2) Στις περισσότερες εφαρμογές, η παράμετρος r είναι θετικός ακέραιος. Στις περιπτώσεις αυτές, η αθροιστική συνάρτηση κατανομής της Γάμμα κατανομής μπορεί να εκφρασθεί με όρους της αθροιστικής συνάρτησης πιθανότητας της Poisson κατανομής, όπως φαίνεται και από την (7.9)

$$F(x)=1-\sum_{\kappa=1}^{r-1}\frac{e^{-\lambda x}(\lambda x)^{\kappa}}{\kappa!}$$

3) Η Γάμμα κατανομή προκύπτει όταν ζητάμε την κατανομή του χρόνου που απαιτείται για την πραγματοποίηση r φορές το γεγονός A, σε μία διαδικασία Poisson.

Σε μία Γάμμα τυχαία μεταβλητή X η μέση τιμή της, δηλαδή, ο μέσος χρόνος μέχρι να πραγματοποιηθεί r φορές το γεγονός A σε μία διαδικασία Poisson, είναι

$$E(X)=\frac{r}{\lambda} \quad \text{και} \quad Var(X)=\frac{r}{\lambda^2}. \tag{7.13}$$

Η απόδειξη είναι εύκολη αν σκεφθεί κανείς ότι ο χρόνος μέχρι την πραγματοποίηση r φορές το A είναι το άθροισμα r ανεξαρτήτων τυχαίων μεταβλητών εκθετικής κατανομής. Ο αναγνώστης μπορεί να προχωρήσει στην ολοκλήρωση της απόδειξης, θα ήταν μία καλή εξάσκηση.

Παράδειγμα 7.7

Σε ένα μεγάλο αντλιοστάσιο ύδρευσης, υπάρχούν τρεις αντλίες ιδίου τύπου. Όταν η μία αντλία παρουσιάσει βλάβη αχρηστεύεται, τίθεται αυτόματα σε λειτουργία η δεύτερη αντλία, παρόμοια και για την τρίτη. Αν η εμφάνιση βλάβης ακολουθεί Poisson κατανομή και η μέση τιμή του χρόνου καλής λειτουργίας της κάθε αντλίας είναι 15 μήνες, ποια η πιθανότητα το αντλιοστάσιο να αντλεί νερό αδιάκοπα τουλάχιστον 5 χρόνια από την έναρξη της λειτουργίας της πρώτης αντλίας;

Αν X παριστάνει τον χρόνο λειτουργίας της αντλίας, η συνάρτηση πυκνότητας πιθανότητας είναι,

$$f(x)=\frac{1}{15}e^{-\frac{x}{15}} \quad \text{για } x\geq 0.$$

Επίσης, η λειτουργία της κάθε αντλίας είναι ανεξάρτητη από τις άλλες, έτσι ώστε ο χρόνος λειτουργίας του συστήματος T, ο οποίος είναι το άθροισμα τριών ανεξαρτήτων εκθετικών μεταβλητών, είναι τυχαία μεταβλητή Γάμμα κατανομής. Με άλλα λόγια, η μεταβλητή T παριστάνει τον χρόνο μέχρι την εμφάνιση τριών βλαβών, άρα είναι Γάμμα κατανομής με παράμετρο $r=3$.Έτσι, η συνάρτηση πυκνότητας πιθανότητας για την Γάμμα μεταβλητή T είναι,

$$f(t)=\frac{\lambda^r t^{r-1}e^{-\lambda t}}{\Gamma(r)}=\frac{1}{3\cdot 2\cdot 1}\left(\frac{1}{15}\right)^3 t^2 e^{\frac{t}{15}} \quad \text{για } t\geq 0.$$

Ο υπολογισμός της πιθανότητας $P(T>60$ μήνες$)$ μπορεί να επιτευχθεί με ολοκλήρωση της συνάρτησης πυκνότητας $f(t)$. Όμως, την ζητούμενη πιθανότητα μπορούμε πιο εύκολα να την υπολογίσουμε και από την άθροιστική $F(t)$, όπως αναφέρεται στην (7.9),

$$P(T>60)=1-F(60)=1-\left(1-\sum_{x=0}^{2}\frac{4^x}{x!}e^{-4}\right)=\sum_{x=0}^{2}\frac{4^x}{x!}e^{-4}$$

$$=0{,}0185+0{,}074+0{,}1480{,}407.$$

■

7.4 ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ (Normal ή Gauss)

Αναμφίβολα, η πιο συνηθισμένη κατανομή για μία τυχαία μεταβλητή είναι η **κανονική κατανομή**, λόγω απλότητας και χρησιμότητάς της. Αυτή η κατανομή είναι η βάση για πολλές στατιστικές μεθόδους και χρησιμοποιείται επίσης σαν προσέγγιση για πολλές άλλες κατανομές οι οποίες έχουν πρακτική εφαρμογή.

Όταν επαναλαμβάνεται ένα τυχαίο πείραμα, η τυχαία μεταβλητή που ισούται με το μέσο ή ολικό αποτέλεσμα των επαναλήψεων τείνει να έχει κανονική κατανομή καθώς ο αριθμός επαναλήψεων γίνεται μεγάλος. Το αποτέλεσμα αυτό διατυπώθηκε από τον De Moivre το 1733. Αργότερα ο Gauss συνέχισε την διερεύνηση με τις κανονικές μεταβλητές και έτσι η κανονική κατανομή αναφέρεται επίσης και σαν **Gaussian** κατανομή.

Ακόμη, το άθροισμα τυχαίων ανεξάρτητων μεταβλητών διαφόρων κατανομών είναι τυχαία μεταβλητή και προσεγγίζει την κανονική κατανομή. Για παράδειγμα, η απόκλιση (λάθος) στο μήκος ενός παραγόμενου σύρματος είναι το άθροισμα μικρότερων άλλων στοχαστικών παραγόντων όπως, θερμοκρασία και υγρασία της πρώτης ύλης, δονήσεις, διακυμάνσεις στην γωνία κοπής, διακυμάνσεις στα χαρακτηριστικά της πρώτης ύλης και άλλα. Εάν οι επί μέρους αποκλίσεις είναι ανεξάρτητες και ισοπίθανα αρνητικές ή θετικές, τότε η ολική απόκλιση μπορεί να αποδειχθεί ότι είναι μεταβλητή και ακολουθεί κατά προσέγγιση κανονική κατανομή. Επί πλέον, η κανονική κατανομή προσαρμόζεται βασικά σε φυσικά φαινόμενα. Η μοριακή ταχύτητα , για παράδειγμα, είναι τυχαία μεταβλητή και θα μπορούσαμε να υποθέσουμε ότι ακολουθεί κανονική κατανομή.

Η θεωρητική βάση της κανονικής κατανομής αναφέρεται σε σύνθεση άλλων κατανομών, σύμφωνα με το **Κεντρικό Οριακό Θεώρημα** όπως θα μελετήσουμε αργότερα. Προς το παρών μας ενδιαφέρει να υπολογίσουμε την πιθανότητα μία κανονική τυχαία μεταβλητή να πάρει τιμή μέσα σε κάποιο διάστημα ή κάτω από κάποιο όριο κ.τ.λ.

Ορισμός: Η τυχαία μεταβλητή ***X*** η οποία υποτίθεται ότι μπορεί να πάρει τιμές **-∞<*x*<∞,** έχει **κανονική κατανομή** ή κατανομή **Gauss** εάν η συνάρτηση πυκνότητας πιθανότητάς της είναι

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad για \quad -\infty<x<\infty. \tag{7.14}$$

όπου ***μ*** και ***σ*** είναι οι παράμετροι της κανονικής κατανομής, και όπως αποδεικνύεται στην συνέχεια, είναι αντίστοιχα η **μέση τιμή** και **τυπική απόκλιση** της κανονικής τυχαίας μεταβλητής *X*. Χρησιμοποιούμε τον συμβολισμό $X \sim N\left(\mu,\sigma^2\right)$ εάν και μόνον εάν η συνάρτηση πυκνότητας πιθανότητας της X δίνεται από εξίσωση (7.14).

Ας ελέγξουμε αν πράγματι η παραπάνω συνάρτηση $f(x)$ είναι όντως συνάρτηση πυκνότητας πιθανότητας. Η $f(x)$ είναι προφανές ότι είναι θετική. Ακόμη πρέπει το ολοκλήρωμα της συνάρτησης να δώσει την τιμή 1. Πράγματι, θέτοντας $t=(x-\mu)/\sigma$ έχουμε,

$$\int_{-\infty}^{\infty} f(x)dx = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}dt = \frac{1}{\sqrt{2\pi}}\sqrt{2\pi} = 1$$

Η μέση τιμή της κανονικής τυχαίας μεταβλητής X, $E(X)$, συμπίπτει με την παράμετρο μ της συνάρτησης (7.14), αφού,

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} xe^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

και θέτοντας πάλι $t=(x-\mu)/\sigma$ έχουμε

$$E(X) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\sigma t+\mu)e^{-\frac{t^2}{2}}dt$$

$$= \frac{1}{\sqrt{2\pi}}\sigma\int_{-\infty}^{\infty} te^{-\frac{t^2}{2}}dt + \mu^2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}dt$$

Το πρώτο ολοκλήρωμα είναι μηδέν επειδή η συνάρτηση $e^{-t^2/2}$ είναι συμμετρική ως προς το μηδέν. Το δεύτερο ολοκλήρωμα (εκτός του παράγοντα μ) παρατηρούμε ότι είναι η μάζα πιθανότητας της κανονικής μεταβλητής X και ισούται ασφαλώς με την μονάδα. Έτσι, $E(X)=\mu$.

Επίσης, η διακύμανση της τυχαίας μεταβλητής X είναι

$$Var(X) = E\left(X^2\right) - [E(X)]^2 =$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} x^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} - \mu^2$$

ακολουθώντας πάλι τον ίδιο μετασχηματισμό $t=(x-\mu)/\sigma$ και την παραπάνω σκέψη επιτυγχάνουμε,

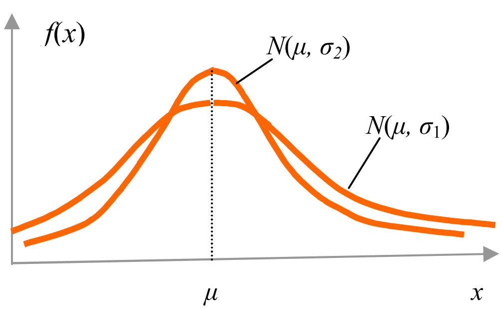
$$Var(X)=\sigma^2+\mu^2-\mu^2=\sigma^2 .$$

Έτσι επιβεβαιώνουμε ότι οι δύο παράμετροι μ και σ οι οποίες χαρακτηρίζουν την κανονική κατανομή είναι η μέση τιμή και η τυπική απόκλιση της X, αντίστοιχα. Με άλλα λόγια, εάν γνωρίζουμε ότι η X ακολουθεί κανονική κατανομή, γνωρίζουμε μόνο ότι η κατανομή πιθανότητάς της ανήκει σε μία οικογένεια κατανομών. Εάν επί πλέον γνωρίζουμε την μέση τιμή $E(X)$ και διακύμανση $Var(X)$, η κατανομή πιθανότητας της X είναι πλήρως ορισμένη.

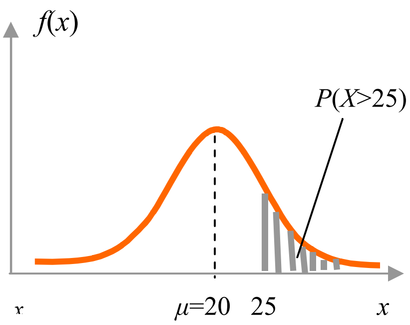
Ιδιότητες κανονικής κατανομής

Η κανονική κατανομή έχει αρκετές σημαντικές **ιδιότητες** όπως:

1. $f(x)$ προσεγγίζει το 0 καθώς το x προσεγγίζει είτε το $-\infty$ ή το ∞.
2. $f(x+\mu)=f(-x+\mu)$ για κάθε x, δηλαδή, η συνάρτηση πυκνότητας είναι συμμετρική γύρω από την μέση τιμή μ.
3. η μέγιστη τιμή της $f(x)$ συμβαίνει όταν $x=\mu$.
4. η γεωμετρική ερμηνεία της παραμέτρου σ έχει ως ακολούθως.
   a. Στο σημείο $x=\mu$ η συνάρτηση $f(x)$ είναι *κυρτή* προς τα *κάτω*.
   b. Για $x \to \pm\infty$, $f(x) \to 0$, ασυμπτωτικά. Αυτό σημαίνει ότι για πολύ μεγάλες θετικές ή αρνητικές τιμές της X η συνάρτηση $f(x)$ γίνεται κυρτή προς τα επάνω.
   c. Το σημείο στο οποίο αρχίζει να αλλάζει η κυρτότητα της συνάρτησης $f(x)$ μπορεί να προσδιορισθεί λύνοντας την εξίσωση $f''(x)=0$. Εάν λύσουμε την εξίσωση, το σημείο $x=\mu\pm\sigma$ είναι σημείο *κάμψης* ή *αλλαγής κυρτότητας*.
5. η συνάρτηση πυκνότητας πιθανότητας $f(x)$ έχει το σχήμα *καμπάνας*. ■



Σχήμα 7.5 Απεικόνιση συναρτήσεων πυκνότητας κανονικής κατανομής, $\sigma_1>\sigma_2$



Σχήμα 7.6 Απεικόνιση κατανομής N(20, 9), του παραδείγματος 7.9

Όπως προκύπτει από τα παραπάνω, η γραφική παράσταση της $f(x)$ σε μία τυχαία μεταβλητή κανονικής κατανομής είναι συμμετρική περί την μέση τιμή μ. Το πλάτος και η κυρτότητά της προσδιορίζονται από την διακύμανσή της σ. Έτσι, αν δύο τυχαίες μεταβλητές X_1 και X_2 έχουν κατανομές $N(\mu, \sigma_1)$ και $N(\mu, \sigma_2)$ αντίστοιχα, όπου $\sigma_1 > \sigma_2$, τότε οι συναρτήσεις πυκνότητας πιθανότητάς των έχουν τις σχετικές μορφές, όπως φαίνεται στο Σχήμα 7.5

Παράδειγμα 7.8

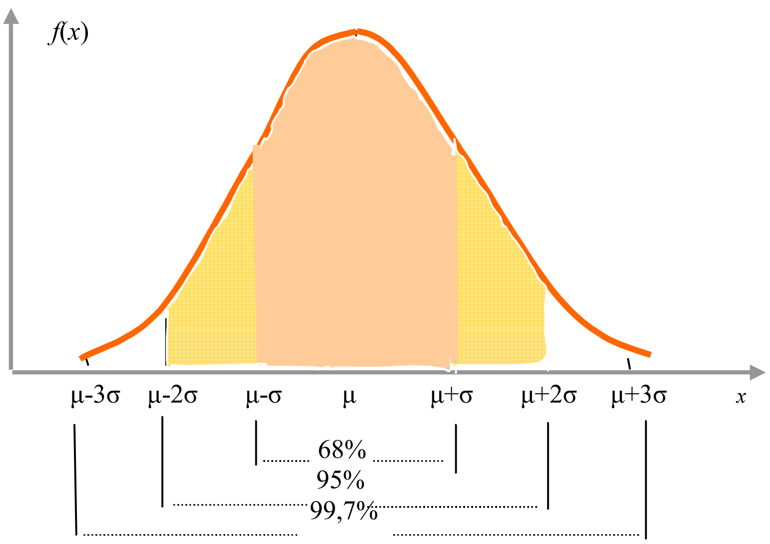
Υποτίθεται ότι η μέτρηση ρεύματος σε ένα σύρμα ακολουθεί κανονική κατανομή με μέση τιμή 20 *milliamperes* και διακύμανση 9 *milliamperes*^{2}. Ποια είναι η πιθανότητα ότι μία μέτρηση ξεπερνά τα 25 *milliamperes*.

Έστω ότι X είναι η τυχαία μεταβλητή που δηλώνει το ρεύμα σε *milliamperes*. Η ζητούμενη πιθανότητα

$$P(X > 25) = 1 - F(25) = \frac{1}{\sqrt{2\pi}\cdot 3}\int_{-\infty}^{25} e^{-\frac{(x-20)^2}{2\cdot 3^2}}\,dx$$

επιδεικνύεται με την σκιασμένη επιφάνεια στο Σχήμα 7.6. Ατυχώς δεν υπάρχει κλειστός τύπος για να υπολογίσουμε το παραπάνω ολοκλήρωμα. Η πιθανότητα στην κανονική κατανομή υπολογίζονται αριθμητικά ή από τον στατιστικό πίνακα αφού γίνει μετασχηματισμός σε *τυπική κανονική κατανομή*, όπως αναλύεται στην επόμενη παράγραφο. ■

Μερικά χρήσιμα αποτελέσματα που αφορούν την κανονική κατανομή αναφέρονται περιληπτικά στο Σχήμα 7.7. Για κάθε κανονική τυχαία μεταβλητή ισχύει



Σχήμα 7.7. Πιθανότητες σε μία κανονική κατανομή.

$P(\mu\text{-}\sigma < X < \mu+\sigma) = 0{,}6827$,
$P(\mu\text{-}2\sigma < X < \mu+2\sigma) = 0{,}9545$,
$P(\mu\text{-}3\sigma < X < \mu+3\sigma) = 0{,}9973$.

Στο Σχήμα 7.7, παρατηρούμε ότι η πιθανότητα μία κανονική τυχαία μεταβλητή X να πάρει τιμές εκτός του διαστήματος [μ-3σ, μ+3σ] είναι πολύ μικρή, και γι αυτό το διάστημα αυτό συνήθως λέγεται και *πλάτος της κανονικής κατανομής*.

7.4.1 Τυπική Κανονική Κατανομή

Εάν η τυχαία μεταβλητή X έχει κανονική κατανομή $N(\mu,\ \sigma)$, η τυχαία μεταβλητή Z η οποία είναι η τυποποίηση της X,

$$Z = \frac{X - \mu_X}{\sigma_X},$$

σύμφωνα με γνωστό θεώρημα στη στατιστική, ακολουθεί κανονική κατανομή. Όπως έχει αναφερθεί και στην παράγραφο 5.3, μία τυποποιημένη τυχαία μεταβλητή Z έχει προφανώς μέση τιμή το μηδέν και τυπική απόκλιση την μονάδα, $E(Z)$=0 και $Var(Z)$=1.

Ορισμός: Μία τυχαία μεταβλητή κανονικής κατανομής με μέση τιμή $\boldsymbol{\mu}$**=0** και τυπική απόκλιση $\boldsymbol{\sigma}$**=1** λέγεται **τυπική κανονική κατανομή** και συμβολίζεται με ***Z***. Η μεν συνάρτηση πυκνότητας πιθανότητας της ***Z*** συμβολίζεται με $\boldsymbol{\varphi(z)}$ και είναι

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad \gamma\iota\alpha \quad -\infty < z < \infty$$

η δε αθροιστική πιθανότητα της ***Z*** συμβολίζεται παρόμοια με κεφαλαίο $\boldsymbol{\Phi}$,

$$\Phi(z) = P(Z \le z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{u^2}{2}} du \quad \gamma\iota\alpha \quad -\infty < z < \infty \qquad (7.15)$$

Η πιθανότητα στην αθροιστική $\Phi(z)$ της (7.15) δεν μπορεί να προσδιορισθεί με αναλυτική ολοκλήρωση, αφού δεν μπορούμε να βρούμε μία συνάρτηση της οποίας η παράγωγος να ισούται με $e^{-u^2/2}$. Παρόλα αυτά, μέθοδοι με αριθμητική ολοκλήρωση μπορούν να χρησιμοποιηθούν για να υπολογίσουν το παραπάνω ολοκλήρωμα της (7.15), και για πολλές τιμές του z οι πιθανότητες

$\Phi(z)=P(Z\leq z)$ περιλαμβάνονται σε στατιστικό πίνακα τυπικής κανονικής κατανομής. Αυτός ο πίνακας, όπως και πίνακες για άλλες κατανομές εμφανίζονται στις τελευταίες σελίδες σε όλα σχεδόν τα βιβλία Πιθανοτήτων και Στατιστικής.

Με άλλα λόγια, μπορούμε τώρα να χρησιμοποιήσουμε τους στατιστικούς πίνακες για να προσδιορίσουμε την πιθανότητα

$$P(z_1 < Z \leq z_2)=\Phi(z_2)-\Phi(z_1),$$

όπου Z είναι τυχαία μεταβλητή τυπικής κανονικής κατανομής, $Z\sim N(\mu=0, \sigma=1)$.

Η ιδιαίτερη χρησιμότητα του στατιστικού πίνακα που περιλαμβάνει τιμές για την αθροιστική πιθανότητα $\Phi(z)$ οφείλεται στο γεγονός ότι εάν η X είναι κανονική τυχαία μεταβλητή, $X\sim N(\mu,\sigma)$, μπορούμε να χρησιμοποιήσουμε την Φ για να υπολογίσουμε πιθανότητες που αφορούν την X. Έστω ότι θέλουμε να υπολογίσουμε την πιθανότητα $P(X\leq x)$, με απλό μετασχηματισμό έχουμε,

$$\begin{aligned} P(X \leq x) &= P(X-\mu \leq x-\mu) \\ &= P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) \\ &= P\left(Z \leq \frac{x-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{x-\mu}{\sigma}\right) \end{aligned} \tag{7.16}$$

αφού η μεταβλητή $Z=(X-\mu)/\sigma$ ακολουθεί τυπική κανονική κατανομή $N(0,1)$, σύμφωνα με γνωστό θεώρημα.

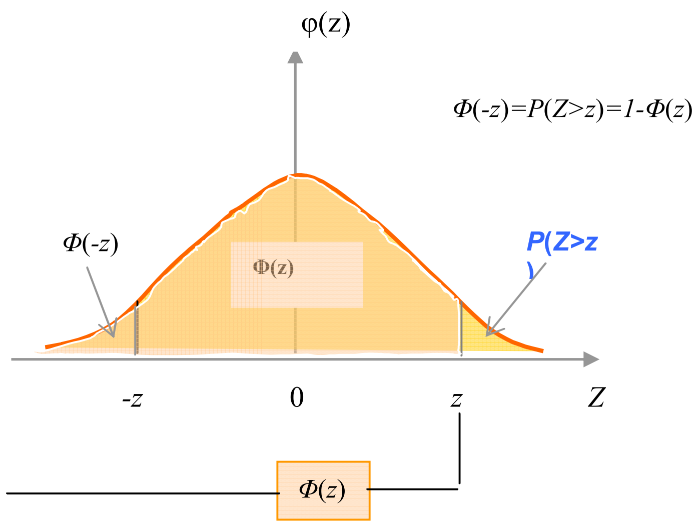
Μετασχηματισμός σε τυπική κανονική. Τελικά, εάν για μία κανονική μεταβλητή ***X~N(μ, σ)*** θέλουμε να υπολογίσουμε την πιθανότητα ***P(X≤x),*** με τυποποίηση της ***X*** μεταφερόμαστε στην τυπική κανονική μεταβλητή ***Z~N*(0, 1)** και βρίσκουμε από τους στατιστικούς πίνακες την ισοδύναμη πιθανότητα ***P(Z≤z)=Φ(z)*** για ***z=(x-μ)/σ,***

$$P(X \leq x)=\Phi\left(\frac{x-\mu}{\sigma}\right) \tag{7.17}$$

Είναι επίσης προφανές ότι από τον ορισμό της αθροιστικής $\Phi(z)$ ισχύει

$$\Phi(-z)=P(Z\leq -z)=P(Z>z)=1-P(Z\leq z)=1-\Phi(z)$$

λόγω συμμετρίας όπως φαίνεται και στο σχήμα 7.8. Αυτή η σχέση είναι ιδιαίτερα χρήσιμη αφού οι στατιστικοί πίνακες δίνουν τιμές για την αθροιστική πιθανότητα $\Phi(z)$, η οποία αντιστοιχεί μόνο σε θετικές τιμές z της τυπικής κανονικής κατανομής.



Σχήμα 7.8. Συνάρτηση πυκνότητας πιθανότητας τυπικής κανονικής κατανομής

Παράδειγμα 7.9

Υποθέτουμε ότι η διάρκεια κατασκευής μιας γέφυρας X είναι τυχαία μεταβλητή και ακολουθεί κανονική κατανομή με αναμενόμενη τιμή 180 ημέρες και τυπική απόκλιση 10 ημέρες, X~N(μ=180 ημέρες, σ=10ημέρες). Θεωρούμε ακόμη ότι η κατασκευή της γέφυρας θα αποπερατωθεί εγκαίρως αν τελειώσει η κατασκευή της μέσα σε 195 ημέρες. Ποια η πιθανότητα ότι δεν θα καθυστερήσει η κατασκευή της γέφυρας.

Πρέπει να υπολογίσουμε την πιθανότητα $P(X<195)$. Έτσι έχουμε

$$P(X<195)=P\left(\frac{X-180}{10}<\frac{195-180}{10}\right)$$
$$=P\left(Z<\frac{195-180}{10}\right)=\Phi\left(\frac{195-180}{10}\right)$$
$$=\Phi(1{,}5)=0{,}9332$$

■

Η κανονική κατανομή στην πράξη

Στο παράδειγμα 7.9 θα μπορούσε να υπάρξει αντίρρηση στην χρήση κανονικής κατανομής, διότι η τυχαία μεταβλητή X παριστάνει χρόνο λειτουργίας και δεν μπορεί να πάρει αρνητικές τιμές. Όμως, ως γνωστόν κάθε κανονική κατανομή θεωρείται ότι παίρνει τιμές από -∞ μέχρι ∞. Παρόλα αυτά, το μοντέλο της κανονικής κατανομής στην συγκεκριμένη περίπτωση (*ολόκληρη η καμπάνα*) πρακτικά βρίσκεται στην πλευρά των θετικών αριθμών, αφού η πιθανότητα ότι $\{X<0\}$ είναι αμελητέα. Δηλαδή,

$$P(X<0)=P\left(\frac{X-180}{10}<\frac{0-180}{10}\right)=\Phi\left(\frac{0-180}{10}\right)=\Phi(-18)=1-\Phi(18)=1-(\approx 1)\cong 0$$

Ο ενδοιασμός που αναφέρθηκε εδώ συναντιέται συχνά, και τίθεται το ακόλουθο ερώτημα: Για μία τυχαία μεταβλητή X, η οποία από την φύση της παίρνει μόνο θετικές τιμές, μπορούμε να υποθέσουμε ότι ακολουθεί κανονική κατανομή; Δηλαδή, η υπόθεση ότι η μεταβλητή X μπορεί να πάρει και αρνητικές τιμές, $-\infty<x<\infty$, έρχεται σε αντίθεση με την πραγματικότητα;

Εφ' όσον, οι τιμές των παραμέτρων μ και σ είναι τέτοιες έτσι ώστε η πιθανότητα $P(X<0)$ είναι ουσιαστικά μηδέν *(σχεδόν ολόκληρη η καμπάνα βρίσκεται στο μέρος των θετικών αριθμών*), η υπόθεση για την κανονική κατανομή είναι αρκετά λογική.

■

7.4.2 Κανονική Προσέγγιση στην Διωνυμική και Poisson κατανομή

Όπως αναφέρθηκε και στην εισαγωγή της κανονικής κατανομής η κανονική μπορεί να χρησιμοποιηθεί και σαν προσέγγιση σε άλλες κατανομές, όπως π.χ. στην διωνυμική όταν ο αριθμός δοκιμών είναι αρκετά μεγάλος. Έτσι λοιπόν, σε πολλά προβλήματα όπου ο αριθμός δοκιμών n σε μία διωνυμική κατανομή είναι μεγάλος, και υπολογιστικά είναι δύσκολο να εκτιμήσουμε την πιθανότητα, χρησιμοποιούμε την τυπική κανονική κατανομή.

Εάν **X** είναι μία **διωνυμική** τυχαία μεταβλητή

$$Z=\frac{X-np}{\sqrt{np(1-p)}} \qquad (7.18)$$

είναι κατά **προσέγγιση** μία τυπική **κανονική** κατανομή. Η προσέγγιση είναι καλή για

$np>5$ και $n(1-p)>5$

Υπενθιμίζουμε ότι στην διωνυμική μεταβλητή *X*, $E(X)=np$ και $V(X)=np(1\text{-}p)$. Κατά συνέπεια, η έκφραση στην εξίσωση 7.18 δεν είναι τίποτε άλλο εκτός από μία τυποποίηση της τυχαίας μεταβλητής *X*. Πιθανότητες σχετικά με την *X* μπορούν να προσεγγισθούν με την τυπική κανονική κατανομή. Η προσέγγιση είναι καλή όταν το *n* είναι μεγάλο σχετικά με το *p*.

Παράδειγμα 7. 10

Σε ένα ψηφιακό κανάλι επικοινωνίας, υποθέτουμε ότι ο αριθμός των λανθασμένων bits ακολουθεί διωνυμική κατανομή, και ότι ένα bit που λαμβάνουμε είναι λανθασμένο με πιθανότητα 10^{-5}. Εάν μεταδίδονται 16 εκατομμύρια bits, ποια η πιθανότητα ότι συμβαίνουν περισσότερα από 150 λάθη;

Δηλώνουμε με την τυχαία μεταβλητή X τα λάθη που συμβαίνουν κατά την μετάδοση. Η χ ακολουθεί διωνυμική κατανομή και η ζητούμενη πιθανότητα είναι

$$P(X>150)=1-P(X\leq 150)$$

$$=1-\sum_{x=0}^{150}\binom{16.000.000}{x}(10^{-5})^{x}(1-10^{-5})^{16.000.000-x}$$

Προφανώς η πιθανότητα αυτή είναι δύσκολο να υπολογισθεί. Ευτυχώς , η κανονική κατανομή μπορεί να χρησιμοποιηθεί για να δώσει μία πολύ καλή προσέγγιση στο παράδειγμα, όπως ακολουθεί,

$$P(X>150)=P\left(\frac{X-160}{\sqrt{160(1-10^{-5})}}>\frac{150-160}{\sqrt{160(1-10^{-5})}}\right)$$

$$=P(Z>-0,79)=P(Z<0,79)=0,785$$

επειδή $np=(16\times 10^{6})(10^{-5}=160$.

Υπενθυμίζουμε ότι η Poisson κατανομή αναπτύχθηκε σαν το όριο μιας διωνυμικής κατανομής καθώς ο αριθμός δοκιμών αυξανότανε στο άπειρο. Κατά συνέπεια, δεν αποτελεί έκπληξη ότι η κανονική κατανομή μπορεί να προσεγγίσει επίσης και την Poisson κατανομή.

Εάν η ***X*** είναι **Poisson** τυχαία μεταβλητή με μέση τιμή ***E*(*X*)=*λ*** και διακύμανση ***V*(*X*)=*λ***, τότε

$$Z = \frac{X-\lambda}{\sqrt{\lambda}}$$

είναι κατά προσέγγιση τυπική κανονική κατανομή. Η προσέγγιση είναι καλή για ***λ>5.***

Παράδειγμα 7. 11

Υποθέτουμε ότι τα σωματίδια της άσβεστου σε ένα τετραγωνικό εκατοστό σκόνης είναι κατά μέσο όρο 1000. Εάν αναλύσουμε ένα τυχαίο τετραγωνικό σκόνης ποια η πιθανότητα να ευρεθούν λιγότερο από 950 σωματίδια άσβεστου;

Η πιθανότητα αυτή μπορεί να εκφρασθεί ακριβώς με τον κανόνα Poisson,

$$P(X \le 950) = \sum_{x=0}^{950} \frac{e^{-1000} 1000^x}{x!}$$

Ο υπολογισμός της πιθανότητας αυτής είναι δύσκολος. Η πιθανότητα μπορεί να προσεγγισθεί σαν

$$P(X \le 950) = P\left(Z \le \frac{950-1000}{\sqrt{1000}} \right) = P(Z \le -1.58) = 0{,}057$$

7.5 ΛΟΓΑΡΙΘΜΟ-ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ

Στην παράγραφο 7.4 παρατηρήσαμε ότι η άθροιση μεγάλου αριθμού από μικρές τυχαίες επιδράσεις προσεγγίζει την κανονική κατανομή. Παρόμοια, εάν ένα φαινόμενο διεγείρεται από πολλαπλές επιδράσεις ενός μεγάλου αριθμού ασυσχέτιστων παραγόντων, η προκύπτουσα κατανομή τείνει να είναι λογάριθμο-κανονική (ή λογαριθμικά κανονική), δηλαδή, ο λογάριθμος της μεταβλητής ακολουθεί κανονική κατανομή. Υπάρχουν πολλά παραδείγματα στην φύση, όπως το μέγεθος και τα χρονικά διαστήματα μεταξύ διαδοχικών σεισμών, η αντοχή τάσης κάποιου υλικού που χρησιμοποιείται σε μία κατασκευή, η ένταση βροχόπτωσης, η διάρκεια ενός έργου, ο χρόνος εμφάνισης μιας πλημμύρας ή ξηρασίας και άλλα.

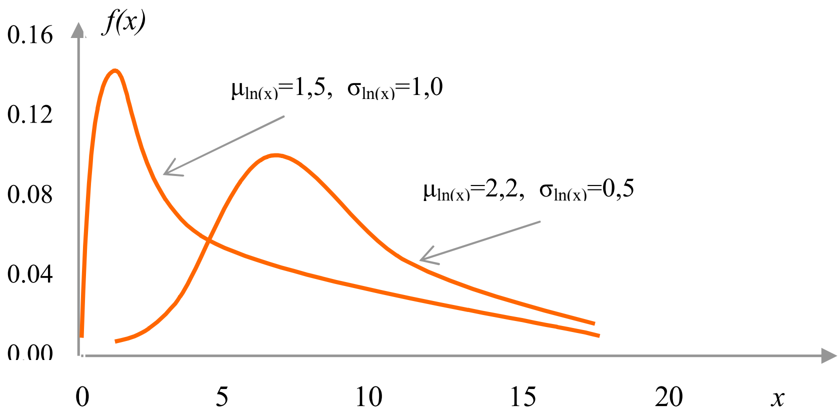
Μία θετική τυχαία μεταβλητή X θεωρείται ότι ακολουθεί λογάριθμο κανονική κατανομή εάν η τυχαία μεταβλητή Y,

$$Y = \ln(X)$$

ακολουθεί κανονική κατανομή, $N(\mu_Y, \sigma_Y)$, όπου $ln(x)$ είναι ο φυσικός λογάριθμος με βάση το e. Η συνάρτηση πυκνότητας πιθανότητας για την λογάριθμο κανονική κατανομή της X δίνεται από την εξίσωση,

$$f_X(x) = \frac{1}{x\sigma_Y \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln x - \mu_Y}{\sigma_Y} \right)^2 \right] \quad \text{για } 0 < x < \infty. \tag{7.18}$$

Συνηθίζεται ο συμβολισμός $X \sim LN(\mu_Y, \sigma_Y)$, ο οποίος περιγράφει περιληπτικά την λογαριθμοκανονική κατανομή. Ο συμβολισμός αυτός δηλώνει ότι η X είναι λογαριθμοκανονικά κατανεμημένη με παραμέτρους μ_Y και σ_Y.



***Σχήμα* 7.9** Συναρτήσεις πυκνότητας πιθανότητας λογαριθμικής κατανομής

Γραφικές παραστάσεις της λογαριθμικής συνάρτησης πυκνότητας πιθανότητας δίνονται στο Σχήμα 7.9 για διαφορετικές τιμές των παραμέτρων της. Επειδή $Y=ln(X)$ έχει κανονική κατανομή, μπορούμε να χρησιμοποιήσουμε τους πίνακες της κανονικής κατανομής για να προσδιορίσουμε μία πιθανότητα.

Ιδιότητες λογαριθμοκανονικής κατανομής

Η λογαριθμοκανονική κατανομή έχει τις ακόλουθες ιδιότητες:

1) Οι τιμές της τυχαίας μεταβλητής X είναι πάντα θετικές, $x>0$.
2) $f_X(x)$ δεν είναι συμμετρική συνάρτηση πυκνότητας πιθανότητας γύρω από την μέση τιμή μ_X.
3) Η μέση τιμή μ_X και η διακύμανση σ_X^2 δεν ισούνται με τις παραμέτρους της κατανομής μ_Y και σ_Y^2. Παρόλα αυτά, αυτές σχετίζονται μεταξύ τους ως ακολούθως:

$$\sigma_Y^2 = \ln\left[1 + \left(\frac{\sigma_X}{\mu_X}\right)^2\right]$$
$$\mu_Y = \ln(\mu_X) - \frac{1}{2}\sigma_Y^2 \qquad (7.19)$$

Οι σχέσεις (7.19) μπορούν να αντιστραφούν,

$$\mu_X = \exp\left(\mu_Y + \frac{1}{2}\sigma_Y^2\right)$$
$$\sigma_X^2 = \mu_X^2\left[\exp\left(\sigma_Y^2\right) - 1\right] \qquad (7.20)$$

Η αθροιστική κατανομή πιθανότητας μιας λογαριθμοκανονικής μεταβλητής X, μπορεί να προσδιορισθεί από την σχέση της με την κανονική κατανομή. Παρόμοια με την κανονική κατανομή, η δυσκολία υπολογισμού του ολοκληρώματος της αθροιστικής κατανομής μπορεί να περιορισθεί μετασχηματίζοντας το αποτέλεσμα σε μία τυπική κανονική κατανομή με μ=0 και σ=1, Z~(0,1). Ως εκ τούτου, χρησιμοποιώντας τον μετασχηματισμό μεταξύ των κατανομών $X \sim LN(\mu_Y, \sigma_Y)$ και Z~(0,1), και τις τιμές πιθανότητας από τους πίνακες τυπικής κανονικής κατανομής, η αθροιστική κατανομή για την λογαριθμοκανονική κατανομή μπορεί να υπολογισθεί χρησιμοποιώντας τον μετασχηματισμό:

$$Z = \frac{\ln X - \mu_Y}{\sigma_Y} \qquad (7.21)$$

Από τον μετασχηματισμό (7.21), πιθανότητες, σχετικά με οποιαδήποτε λογαριθμοκανονική κατανομή, μπορούν εύκολα να υπολογισθούν με την βοήθεια του πίνακα τιμών της Φ κατανομής. Για παράδειγμα, η αθροιστική πιθανότητα $P(X \leq x)$ για κάθε $X \sim LN(\mu_Y, \sigma_Y)$ μπορεί να υπολογισθεί αλλάζοντας την μεταβλητή ολοκλήρωσης σύμφωνα με την 7.21. Έτσι, η αθροιστική πιθανότητα δίνεται από την ακόλουθη εξίσωση,

$$P(X \leq x) = \int_{-\infty}^{\frac{\ln x - \mu_Y}{\sigma_Y}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz = \Phi\left(\frac{\ln x - \mu_Y}{\sigma_Y}\right) \qquad (7.22)$$

Μπορεί ακόμη να αποδειχθεί ότι

$$P(a < X \le b) = F_X(b) - F_X(a) = \Phi\left(\frac{\ln b - \mu_Y}{\sigma_Y}\right) - \Phi\left(\frac{\ln a - \mu_Y}{\sigma_Y}\right).$$

Για πληρέστερη κατανόηση της κανονικής και λογαριθμοκανονικής κατανομής θα ήταν ιδιαίτερα χρήσιμο στον αναγνώστη να ακολουθήσει την απόδειξη της σχέσης (7.19).

Αφού *Y=ln(X),* προφανώς έχουμε $X = e^Y$, και η μέση τιμή της *X* δίνεται από την ακόλουθη σχέση,

$$\mu_X = E(X) = E\left(e^Y\right) = \frac{1}{\sigma_Y\sqrt{2\pi}}\int_{-\infty}^{\infty} e^y \exp\left[-\frac{1}{2}\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]dy$$

$$= \frac{1}{\sigma_Y\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left[y - \frac{1}{2}\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]dy$$

Ο εκθέτης μέσα στο ολοκλήρωμα μπορεί να συμπληρωθεί σαν τετράγωνο αν προσθέσουμε και αφαιρέσουμε το $\mu_Y + \frac{1}{2}\sigma_Y^2$,

$$\mu_X = \left[\frac{1}{\sigma_Y\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{y-\left(\mu_Y+\sigma_Y^2\right)}{\sigma_Y}\right)^2\right]dy\right]\exp\left(\mu_Y + \frac{1}{2}\sigma_Y^2\right)$$

Το ολοκλήρωμα μέσα στην αγκύλη παριστάνει την συνολική μάζα πιθανότητας της κατανομής $N\left(\mu_Y + \sigma_Y^2, \sigma_Y^2\right)$ και ισούται με 1. Άρα έχουμε

$$\mu_X = \exp\left(\mu_Y + \frac{1}{2}\sigma_Y^2\right)$$

Λογαριθμίζοντας έχουμε το ζητούμενο,

$$\mu_Y = \ln(\mu_X) - \frac{1}{2}\sigma_Y^2$$

Παρόμοια εργαζόμαστε και για την διασπορά της *X*.

$$E\left(X^2\right)=E\left(e^{2Y}\right)=\frac{1}{\sigma_Y\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{2y}\exp\left[-\frac{1}{2}\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]dy$$

$$=\frac{1}{\sigma_Y\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left[2y-\frac{1}{2}\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]dy$$

Προσθέτοντας και αφαιρώντας $2\left(\mu_Y+\sigma_Y^2\right)$ στον εκθέτη του ολοκληρώματος, συμπληρώνεται τετράγωνο και προκύπτει

$$E\left(X^2\right)=\left[\frac{1}{\sigma_Y\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}\left(\frac{y-\left(\mu+2\sigma_Y^2\right)}{\sigma_Y}\right)^2\right]dy\right]\exp 2\left(\mu_Y+\sigma_Y^2\right)$$

$$=e^{2\left(\mu_Y+\sigma_Y^2\right)}$$

Τελικά η διακύμανση της X προκύπτει και με την βοήθεια της παραπάνω μ_X,

$$\sigma_X^2=E\left(X^2\right)-\mu_X^2=e^{2\left(\mu_Y+\sigma_Y^2\right)}-e^{2\left(\mu_Y+\frac{1}{2}\sigma_Y^2\right)}$$

$$=e^{2\left(\mu_Y+\frac{1}{2}\sigma_Y^2\right)}\left(e^{\sigma_Y^2}-1\right)=\mu_X^2\left(e^{\sigma_Y^2}-1\right)$$

Λογαριθμίζοντας έχουμε το ζητούμενο,

$$\sigma_Y^2=\ln\left[1+\left(\frac{\sigma_X}{\mu_X}\right)^2\right].$$

Παράδειγμα 7.12

Η αντοχή X ενός τύπου δοκού σε N/mm^2 μέχρι να παρουσιάσει ρωγμή είναι τυχαία μεταβλητή και μάλλον από στατιστικά δεδομένα ταιριάζει σε μία λογαριθμική κατανομή με παραμέτρους μ_X=39,33 N/mm^2 και σ_X=0,24 N/mm^2. Να προσδιορισθεί η ελάχιστη αντοχή του 95% των δοκών. Ακόμη, ποια η πιθανότητα ότι μία δοκός που επιλέξαμε τυχαία η αντοχή της να είναι τουλάχιστον 20 N/mm^2;.

Ως γνωστόν, η μεταβλητή $Y=ln(X)$ ακολουθεί κανονική κατανομή, $Y\sim N(\mu_Y, \sigma_X)$. Από τις εξισώσεις (7.20) έχουμε

$$\sigma_Y^2 = \ln\left[1+\left(\frac{\sigma_X}{\mu_X}\right)^2\right] = \ln\left[1+\left(\frac{0,24}{39,33}\right)^2\right] = 0,237^{\ 2}$$

$$\mu_Y = \ln(\mu_X) - \frac{1}{2}\sigma_Y^2 = \ln(39,33) - \frac{1}{2}0,237^{\ 2} = 3,62$$

Για να προσδιορίσουμε την ελάχιστη αντοχή του 95% των δοκών, βρίσκουμε το αντίστροφο δηλαδή την μέγιστη αντοχή $x_{0,05}$ του 5% των δοκών,

$$P(X \le x_{0,05}) = P(Y \le \ln(x_{0,05})) = \Phi\left(\frac{\ln(x_{0,05})-3,62}{2,37}\right) = 0,05$$

Από τους στατιστικούς πίνακες δίνεται ότι $\Phi(-1,65)=0,05$. Έτσι έχουμε

$$\frac{\ln(x_{0,05})-3,62}{2,37} = -1,65$$
$$\ln(x_{0,05}) = -1,65\cdot 2,37+3,62 = 3,898+3,62 = -0,278$$
$$x_{0,05} = e^{-0,278} =$$

7.6 ΒΗΤΑ ΚΑΤΑΝΟΜΗ

Η Βήτα κατανομή σχετίζεται με μεταβλητές οι οποίες παίρνουν τιμές μεταξύ 0 και 1. Παρουσιάζουν ιδιαίτερο ενδιαφέρων για τους μηχανικούς στην λήψη αποφάσεων.

Ορισμός: Μία τυχαία μεταβλητή ***X*** ακολουθεί **Βήτα** κατανομή αν η συνάρτηση πυκνότητας πιθανότητας είναι

$$f(x\ \ \alpha,\beta) = \begin{cases} \dfrac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1} & \gamma\iota\alpha \quad 0<x<1, \quad \alpha>0, \quad \beta>0 \\ 0 \quad \alpha\lambda\lambda o\upsilon & \end{cases} \tag{7.23}$$

Η έκφραση

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \tag{7.24}$$

δηλώνει την Βήτα συνάρτηση με παραμέτρους α και β. Η μέση τιμή και διασπορά μίας Βήτα μεταβλητής X έχει αποδειχθεί ότι είναι

$$E(X) = \frac{\alpha}{\alpha+\beta}, \quad Var(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{7.25}$$

7.7 ΠΕΡΙΛΗΨΗ ΚΕΦΑΛΑΙΟΥ

Συχνά επιθυμούμε να παρουσιάσουμε ένα σύνολο παρατηρήσεων που αφορούν μία τυχαία μεταβλητή με μία ειδική συνάρτηση κατανομής πιθανότητας (θεωρητική κατανομή), παρά με ένα ιστόγραμμα, το οποίο είναι μόνο μια αρχική γραφική παράσταση. Με αυτές τις συναρτήσεις κατανομών μπορούμε ευκολότερα να διερευνήσουμε τις παραμέτρους , να εκτιμήσουμε πιθανότητες και γενικά να έχουμε πιο μεθοδική και αποτελεσματική στατιστική ανάλυση.

Σ' αυτό το κεφάλαιο, περιγράφθηκαν και αναπτύχθηκαν μερικές από τις κατανομές πιθανότητας συνεχούς τυχαίας μεταβλητής χρήσιμες για τους μηχανικούς. Έχουν αναφερθεί σημαντικές ιδιότητες , οι συνθήκες που πρέπει να πληρούνται για την κάθε κατανομή, καθώς και η φυσική της ή πρακτική της ερμηνεία στα προβλήματα του μηχανικού.

Έχουμε δείξει σε αυτό το κεφάλαιο ότι υπάρχουν θεωρητικές συνδέσεις μεταξύ κάποιων διακριτών και συνεχών κατανομών. Για παράδειγμα, η εκθετική κατανομή σχετίζεται με την Poisson, την Weibull και Γάμμα, ενώ η κανονική είναι η κεντρική για πολλούς τύπους όπως Poisson, διωνυμική, λογάριθμο-κανονική κ.τ.λ..

ΑΣΚΗΣΕΙΣ-ΠΡΟΒΛΗΜΑΤΑ

Με τις λύσεις τους

1) Γραμμές ηλεκτρικής παροχής. Η ηλεκτρική παροχή ενός συστήματος περιλαμβάνει δύο ανεξάρτητες γραμμές παροχής. Η παροχή του συστήματος θεωρείται εξασφαλισμένη εφ' όσον λειτουργεί τουλάχιστον μία γραμμή. Οι χρόνοι ζωής T_1, T_2 των δύο γραμμών είναι κατανεμημένοι εκθετικά με σ.π.π.

$$f_{T_1}(t) = f_{T_2}(t) = \lambda \cdot e^{-\lambda t}$$

Να υπολογιστεί η πιθανότητα εξασφάλισης της παροχής του συστήματος στο διάστημα $(0, t_0)$.

Λύση

Δεδομένου ότι η παροχή δεν εξασφαλίζεται μόνον αν χαλάσουν και οι δύο γραμμές, ορίζοντας με A={*εξασφάλιση της παροχής του συστήματος στο* $(0, t_0)$} έχουμε

$$\overline{A} = \{T_1 \le t_0\} \cap \{T_2 \le t_0\}$$

οπότε για την ζητούμενη πιθανότητα $P(A)$ βρίσκουμε λόγω ανεξαρτησίας των γεγονότων $\{T_1 \le t_0\}$ και $\{T_2 \le t_0\}$

$$\begin{aligned} P(A) &= 1 - P(\overline{A}) = 1 - P(T_1 \le t_0) \cdot P(T_2 \le t_0) \\ &= 1 - F_{T_1}(t_0) \cdot F_{T_2}(t_0) = 1 - (1 - e^{-\lambda t_0})^2 \\ &= 2e^{-\lambda t_0} - e^{-2\lambda t_0}. \end{aligned}$$

2) Επιλογή μεθόδου παραγωγής. Υποθέτουμε ότι ο χρόνος ζωής X μιας ηλεκτρικής ασφάλειας ακολουθεί εκθετική κατανομή. Υπάρχουν δύο μέθοδοι με τις οποίες μπορεί να κατασκευαστεί η ηλεκτρική ασφάλεια. Η μέθοδος I ενδίδει μέσο χρόνο ζωής 150 ώρες, ενώ η μέθοδος II ενδίδει μέσο χρόνο ζωής 100 ώρες.

Εάν μία ασφάλεια διαρκέσει λιγότερο από 200 ώρες ο κατασκευαστής χρεώνεται με ένα ποσό *K*. Ποια μέθοδο πρέπει να χρησιμοποιήσει ο κατασκευαστής αν το κόστος *C* της μεθόδου Ι είναι διπλάσιο από αυτό της ΙΙ;

Λύση

Θέλουμε να επιλέξουμε εκείνη την μέθοδο έτσι ώστε να έχουμε το μικρότερο συνολικό κόστος *SK* το οποίο είναι για την μέθοδο Ι

$$SK_I = K\cdot P(X<200)+C = K(1-e^{-\frac{1}{150}200})+C = K(1-e^{-\frac{2}{3}})+C$$

και για την μέθοδο ΙΙ συγκρίνουμε τα δύο μεγέθη

$$SK_I = K(1-e^{-\frac{2}{3}})+C = 0{,}49K+C$$

$$SK_{II} = K(1-e^{-2})+\frac{C}{2} = 2K(1-e^{-2})+C = 1{,}73K+C,$$

προφανώς φθηνότερη ή προτιμότερη είναι η μέθοδος Ι , 0,49*K*<1,73*K*.

3) Ηλεκτρονική αντίσταση. Η διάρκεια ζωής *T* ενός τύπου ηλεκτρονικών αντιστάσεων είναι τυχαία μεταβλητή κατανεμημένη εκθετικά σύμφωνα με την *α.σ.π.*

$$F_T(t) = 1-e^{-\lambda t}.$$

(α) Ένας καταναλωτής επιθυμεί το 90% των αντιστάσεων του να έχει διάρκεια ζωής μεγαλύτερη των 100.000 ωρών. Υπολογίστε το μέγιστο λ που ικανοποιεί την απαίτηση του.

(β) Έστω ότι οι ανωτέρω αντιστάσεις προ της διανομής στους καταναλωτές υπόκεινται σε ποιοτικό έλεγχο, ο οποίος ανιχνεύει και απορρίπτει κάθε αντίσταση με διάρκεια ζωής μικρότερη των 80.000 ωρών. Πόσο μεταβάλλεται στην περίπτωση αυτή το μέγιστο λ που ικανοποιεί την απαίτηση του καταναλωτή;

Λύση

(α) Η απαίτηση του καταναλωτή μπορεί να διατυπωθεί με την ακόλουθη μαθηματική φόρμουλα,

$$P(X>100.000) = e^{-\lambda 100.000} \geq 0{,}90$$

Λύνουμε την παραπάνω ανισότητα ως προς *λ*,

$$-100.000\lambda=\ln(0{,}90)=-0{,}105 \Rightarrow \lambda=0{,}00000105$$

(β) Στην περίπτωση αυτή είναι δεδομένο ότι οι αντιστάσεις έχουν διάρκεια πάνω από 80.000 ώρες, έτσι η παραπάνω φόρμουλα διαμορφώνεται σε υπό συνθήκη πιθανότητα

$$P(X > 100.000 \mid X > 80.000) \geq 0{,}90$$

Εφαρμόζοντας τον τύπο της υπό συνθήκη πιθανότητας έχουμε

$$P(X > 100.000 \mid X > 80.000) = \frac{P(\{X > 100.000\} \cap \{X > 80.000\})}{P(X > 80.000)}$$

$$= \frac{P(X > 100.000)}{P(X > 80.000)} = e^{-\lambda \cdot 20.000} \geq 0{,}90.$$

Τέλος λύνουμε την ανισότητα ως προς λ,

$$-20.000\lambda = \ln(0{,}90) = -0{,}105 \Rightarrow \lambda = 0{,}00000503.$$

4) Λειτουργία ηλεκτρικού συστήματος. Σ’ ένα ηλεκτρικό κύκλωμα υπάρχουν δύο διακόπτες Α και Β εν σειρά. Ο χρόνος καλής λειτουργίας του κάθε διακόπτη X_A, X_B, είναι τυχαία μεταβλητή και ακολουθεί εκθετική κατανομή με μέση τιμή $E(X_A)$=100 ώρες, $E(X_B)$=140 ώρες για τους διακόπτες Α, Β αντίστοιχα.
(α) Έχουν παρέλθει 100 ώρες λειτουργίας του κυκλώματος χωρίς διακοπή. Ποια η πιθανότητα να μην διακοπεί το ρεύμα στο κύκλωμα για άλλες 140 ώρες, λόγω βλάβης διακόπτη;
(β) Σ’ ένα σύστημα υπάρχουν 8 ίδια με τα παραπάνω ηλεκτρικά κυκλώματα. Για να λειτουργεί σωστά το σύστημα χρειάζεται η σωστή λειτουργία τουλάχιστον 6 κυκλωμάτων. Ποια η πιθανότητα ότι το σύστημα θα λειτουργήσει σωστά τουλάχιστον 140 ώρες;

Λύση
(α) Αφού οι διακόπτες είναι εν σειρά, το κύκλωμα λειτουργεί όταν λειτουργούν αμφότεροι. Έτσι η ερώτηση μπορεί να διατυπωθεί με την αναζήτηση της ακόλουθης δεσμευμένης πιθανότητας,

$$P(\{X_A > 140\} \cap \{X_B > 140\} \mid \{X_A > 100\} \cap \{X_B > 100\})$$

Επειδή η **εκθετική** κατανομή ***δεν έχει μνήμη*** η ζητούμενη πιθανότητα μπορεί να γραφεί

$$P(\{X_A > 140\} \cap \{X_B > 140\} = P(X_A > 140)P(X_B > 140)$$

λόγω ανεξαρτησίας των γεγονότων { X_A >140}, { X_B >140}. Οι παραπάνω πιθανότητες υπολογίζονται από την αθροιστική της εκθετικής αν θέσουμε

λ_A=1/ $E(X_A)$=1/100=0,010
λ_B=1/ $E(X_B)$=1/140=0,007.

Έτσι έχουμε

$$P(\{X_A > 140\} \cap \{X_B > 140\} = e^{-\lambda_A 140} e^{-\lambda_B 140} = e^{-0,01 \cdot 140} e^{-0,007 \cdot 140}$$
$$= e^{-1,4} e^{-0,98} = 0,095$$

(β) Για κάθε κύκλωμα του συστήματος υπάρχει σταθερή πιθανότητα να λειτουργήσει σωστά τις επόμενες 140 ώρες και ίση με p=0,095. Θεωρώ ότι για τα 8 κυκλώματα έχω n=8 δοκιμές Bernoulli (*επιτυχίας* ή *αποτυχίας* στις 140 ώρες) και τυχαία μεταβλητή Y η οποία παριστάνει τις επιτυχίες στις 8 δοκιμές, Y=[0, 1, 2, …, 8}. Ζητώ την πιθανότητα

$$P(Y \geq 6) = P(Y = 6) + P(Y = 7) + P(Y = 8) =$$
$$= \binom{8}{6} 0,095^6 \cdot 0,905^2 + \binom{8}{7} 0,095^7 \cdot 0,905^1 \binom{8}{8} 0,095^8 \cdot 0,9050^0$$

5) Προδιαγραφές εξαρτήματος. Η διάρκεια ζωής X ενός εξαρτήματος κατανέμεται κανονικά με μέση τιμή 200 ώρες. Εάν ένας αγοραστής απαιτεί τουλάχιστον το 90% των εξαρτημάτων του να έχουν διάρκεια ζωής πάνω από 150 ώρες, ποια είναι η μέγιστη δυνατή τιμή της τυπικής απόκλισης σ της μεταβλητής X ώστε να ικανοποιείται αυτή η απαίτηση;

Λύση
Αν η τυπική απόκλιση σ ήταν πολύ κοντά στο μηδέν τότε σχεδόν 100% των εξαρτημάτων θα είχαν διάρκεια ζωής πάνω από 150 ώρες. Καθώς όμως αυξάνει η τυπική απόκλιση σ τότε αυξάνει και το ποσοστό των εξαρτημάτων των οποίων η τιμή της X είναι μικρότερη από 150. Για να ικανοποιήσουμε την απαίτηση του αγοραστή πρέπει

$$P(X > 150) = 0,90$$

ή

$$P(X > 150) = 1 - P(X < 150) = 0{,}90$$

$$\Rightarrow \Phi(\frac{150-200}{\sigma}) = \Phi(\frac{-50}{\sigma}) = 0{,}10 \Rightarrow \frac{-50}{\sigma} = -1{,}28 \Rightarrow \sigma = \frac{50}{1{,}28} = 39{,}06$$

Έτσι για σ=39,06 τουλάχιστον 90% των εξαρτημάτων έχουν διάρκεια ζωής πάνω από 150 ώρες.

6) Λειτουργία ηλεκτρονικού συστήματος. Για την καλή λειτουργία ενός ηλεκτρονικού συστήματος απαιτείται δύο ηλεκτρικές αντιστάσεις του Α και Β να έχουν τιμές μεταξύ 8,45 με 8,67 Ohm, και 1,56 με 1,60 Ohm, αντίστοιχα. Οι τιμές των ηλεκτρικών αντιστάσεων είναι τυχαίες μεταβλητές κανονικής κατανομής με μέσες τιμές 8,56 Ohm και 1,58 Ohm αντίστοιχα.

Είναι γνωστό ότι το 5% των αντιστάσεων τύπου Α έχει τιμή μικρότερη από 8,45 Ohm και από τις αντιστάσεις τύπου Β το 2% έχει τιμή πάνω από 1,60 Ohm.

(α) Να βρεθεί η πιθανότητα ότι το ηλεκτρονικό σύστημα δεν θα λειτουργήσει καλά.

(β) Ποια η πιθανότητα ότι σε ένα σύνολο 10 τέτοιων ηλεκτρονικών συστημάτων θα υπάρχουν περισσότερα από 2 ελαττωματικά;

Λύση

Συμβολίζω με X_A και X_B τις τυχαίες μεταβλητές για τις τιμές των ηλεκτρικών αντιστάσεων τύπου Α και Β αντίστοιχα. Είναι γνωστές οι αντίστοιχες μέσες τιμές μ_A=8,56 και μ_B=1,58. Από τις πληροφορίες του προβλήματος θα εκτιμήσουμε και τις τυπικές αποκλίσεις σ_A και σ_B, όπως

$$P(X_A < 8{,}45) = 0{,}05 \Rightarrow \Phi(\frac{8{,}45-8{,}56}{\sigma_A}) = \Phi(\frac{-0{,}11}{\sigma_A}) = 0{,}05 \Rightarrow \frac{-0{,}11}{\sigma_A} = -1{,}65$$

$$\Rightarrow \sigma_A = 0{,}067\ Ohm$$

$$P(X_B > 1{,}60) = 0{,}02 \Rightarrow P(X_B < 1{,}60) = 0{,}98 \Rightarrow \Phi(\frac{1{,}60-1{,}58}{\sigma_B}) = 0{,}98$$

$$\Rightarrow \Phi(\frac{0{,}02}{\sigma_B}) = 0{,}98 \Rightarrow \frac{0{,}02}{\sigma_B} = 2{,}05 \Rightarrow \sigma_B = 0{,}0098\ Ohm$$

(α) Γνωρίζοντας πλήρως την κανονική κατανομή των X_A και X_B, μπορώ να υπολογίσω τις πιθανότητες

$$P(8{,}45 \le X_A \le 8{,}67) = \Phi(\frac{8{,}67-8{,}56}{0{,}067}) - \Phi(\frac{8{,}45-8{,}56}{0{,}067}) = \Phi(\frac{0{,}11}{0{,}067}) - \Phi(\frac{-0{,}11}{0{,}067})$$

$$= 2\Phi(\frac{0{,}11}{0{,}067}) - 1 = 2\Phi(1{,}64) - 1 = 1{,}90 - 1 = 0{,}90$$

$$P(1{,}56 \le X_B \le 1{,}60) = \Phi(\frac{1{,}60-1{,}58}{0{,}0098}) - \Phi(\frac{1{,}56-1{,}58}{0{,}0098}) = \Phi(\frac{0{,}02}{0{,}0098}) - \Phi(\frac{-0{,}02}{0{,}0098})$$

$$= 2\Phi(\frac{0{,}02}{0{,}0098}) - 1 = 2\Phi(2{,}04) - 1 = 1{,}954 - 1 = 0{,}954$$

Έτσι, μπορώ να υπολογίσω την πιθανότητα το σύστημα να λειτουργήσει καλά,

$$P(8{,}45 \le X_A \le 8{,}67) \cdot P(1{,}56 \le X_B \le 1{,}60) = 0{,}90 \cdot 0{,}954 = 0{,}8586$$

Τέλος, η πιθανότητα να μην λειτουργήσει καλά είναι η πιθανότητα του συμπληρωματικού γεγονότος και είναι

p=1-0,8586=0,1414

(β) Εδώ θεωρώ ότι έχω n=10 δοκιμές Bernoulli με σταθερό p=0,1414 και αν Y είναι η τυχαία μεταβλητή που συμβολίζει πόσα στα δέκα συστήματα είναι ελαττωματικά, η Y είναι διωνυμικά κατανεμημένη, και μπορώ από τον τύπο της διωνυμικής να υπολογίσω την ζητούμενη πιθανότητα,

$$P(Y > 2) = 1 - P(Y \le 2) = 1 - \sum_{y=0}^{2} \binom{10}{y} 0{,}1414^{y} (1-0{,}1414)^{10-y}$$

Ασκήσεις προς λύση

Ομοιόμορφη κατανομή

1) Η συνάρτηση πυκνότητας πιθανότητας του καθαρού βάρους X σε κιλά ενός πακέτου χημικής βοτάνης (herbicide) είναι ομοιόμορφη για $49,74<x<50,25$.
(α) Προσδιορίστε την πιθανότητα ότι το πακέτο ζυγίζει περισσότερο από 50 κιλά.
(β) Ποια η ελάχιστη ποσότητα χημικής βοτάνης που περιέχεται στο 90% των πακέτων;
(γ) Προσδιορίστε την μέση τιμή και τυπική απόκλιση του καθαρού βάρους των πακέτων.

2) Η συνάρτηση πυκνότητας πιθανότητας του χρόνου που απαιτείται σε μία συναρμολόγηση είναι $f(x)=0,1$ για $30<x<40$ δευτερόλεπτα.
(α) Προσδιορίστε το ποσοστό των συναρμολογήσεων που απαιτούν περισσότερο από 35 δευτερόλεπτα για να ολοκληρωθούν.
(β) Ποιος χρόνος ξεπερνιέται από το 90% των συναρμολογήσεων;
(γ) Προσδιορίστε την μέση τιμή και τυπική απόκλιση του χρόνου συναρμολόγησης.

Εκθετική κατανομή

3) Υποθέστε ότι η τυχαία μεταβλητή έχει εκθετική κατανομή με μέση τιμή 10. Προσδιορίστε τα ακόλουθα.
(α) $P(X>10)$
(β) $P(X>20)$
(γ) $P(X>30)$
(δ) Να βρεθεί η τιμή της x έτσι ώστε $P(X<x)=0,95$.

4) Η πιθανότητα ότι μία λάμπα φωτός θα καεί είναι εκθετικά κατανεμημένη με μέση τιμή α=1000 ώρες. Βρέστε την πιθανότητα ότι η λάμπα θα συνεχίζει να φωτίζει μετά από 1200 ώρες.

5) Σε ένα μετρητή λήψης σημάτων, ο αριθμός σημάτων ακολουθεί εκθετική κατανομή με μέση συχνότητα δύο σήματα ανά λεπτό.
(α) Ποια είναι η πιθανότητα ότι σε διάστημα 30 δευτερολέπτων η συσκευή δεν μετρά κανένα σήμα;

(β) Ποια είναι η πιθανότητα ότι η πρώτη μέτρηση συμβαίνει σε λιγότερο από 10 δευτερόλεπτα;
(γ) Ποια η πιθανότητα ότι η πρώτη μέτρηση συμβαίνει μεταξύ 1 και 2 λεπτά μετά την έναρξη λειτουργίας του μετρητή;

6) Ο χρόνος μεταξύ δύο διαδοχικών ηλεκτρονικών μηνυμάτων στον υπολογιστή σας ακολουθεί εκθετική κατανομή με μέση τιμή δύο ώρες.
(α) Ποια η πιθανότητα ότι δεν θα πάρετε μήνυμα τις επόμενες δύο ώρες;
(β) Εάν δεν είχατε μήνυμα τις τελευταίες 6 ώρες, ποια η πιθανότητα ότι δεν θα έχετε μήνυμα για τις επόμενες δύο ώρες;
(γ) Ποια είναι η αναμενόμενη μέση διάρκεια μεταξύ του πέμπτου και έκτου ηλεκτρονικού μηνύματος που καταφθάνει στον υπολογιστή σας;

7) Ο χρόνος μεταξύ διαδοχικών αφίξεων ταξί σε ένα κυκλοφοριακό κόμβο ακολουθεί εκθετική κατανομή, με μέση διάρκεια 10 λεπτά.
(α) Ποια η πιθανότητα να περιμένετε περισσότερο από μία ώρα για ταξί;
(β) Ας πούμε ότι περιμένατε μία ώρα για ταξί και δεν ήρθε, ποια η πιθανότητα ότι κάποιο θα έρθει στα επόμενα 10 λεπτά;
(γ) Προσδιορίστε το x έτσι ώστε η πιθανότητα να περιμένετε πάνω από x λεπτά είναι 0,10.

8) Η απόσταση μεταξύ σοβαρών χαραμάδων σε ένα μεγάλο εθνικό δρόμο ακολουθεί εκθετική κατανομή με μέση τιμή 8 χιλιόμετρα.
(α) Ποια η πιθανότητα ότι δεν θα συναντήσετε χαραμάδες σε 15 χιλιόμετρα πάνω στον εθνικό δρόμο;
(β) Ποια η πιθανότητα ότι υπάρχουν δύο χαραμάδες σε απόσταση 15 χιλιομέτρων στον εθνικό δρόμο;
(γ) Ποια είναι η τυπική απόκλιση της απόστασης μεταξύ διαδοχικών χαραμάδων;

9) Ο αριθμός ψεγαδιών πάνω σε μία μαγνητική ταινία ακολουθεί Poisson κατανομή με μέση τιμή 0,2 ψεγάδια ανά μέτρο. Αν Χ παριστάνει την απόσταση μεταξύ δύο διαδοχικών ψεγαδιών,
(α) ποια η μέση τιμή της X;
(β) ποια η πιθανότητα ότι δεν υπάρχουν ψεγάδια στα επόμενα δέκα συνεχή μέτρα;
(γ) η απάντηση στην (β) ερώτηση αλλάζει αν τα δέκα μέτρα δεν είναι συνεχή;
(δ) Πόσα μέτρα ταινίας χρειάζονται να ελεγχθούν έτσι ώστε η πιθανότητα ότι τουλάχιστον ένα ψεγάδι θα βρεθεί είναι 0,90;

10) Η διάρκεια ζωής ενός ρυθμιστή ηλεκτρικής ισχύος σε ένα αυτοκίνητο ακολουθεί εκθετική κατανομή με μέση ζωή έξι ετών. Αγοράζετε ένα

αυτοκίνητο μεταχειρισμένο έξι ετών, με ρυθμιστή να λειτουργεί πολύ καλά, και σχεδιάζετε να κυκλοφορήσετε το αυτοκίνητο για έξι χρόνια.
(α) Πια η πιθανότητα ότι ρυθμιστής θα παρουσιάσει βλάβη στην διάρκεια λειτουργίας του αυτοκινήτου σας;
(β) Εάν ο ρυθμιστής χαλάσει τρία χρόνια μετά την απόκτηση του αυτοκινήτου, και αντικατασταθεί με έναν καινούργιο ίδιου τύπου, ποια η μέση διάρκεια μέχρι την επόμενη βλάβη;

11) Αν η τυχαία μεταβλητή X έχει πυκνότητα πιθανότητας

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda x}, & x \geq 0 \\ 0, \; x < 0 \end{cases}$$

να συγκριθεί η πιθανότητα $P(|X| \geq 2/\lambda)$ με το φράγμα της ανισότητας του Chebycheff:

$$P(|X-\mu| \geq t \cdot \sigma) \leq \frac{1}{t^2}.$$

(όπου μ η μέση τιμή της X και ισχύει μ=1/λ, και σ η τυπική απόκλιση της X και ισχύει σ=1/λ^2).

12) Η διάρκεια ζωής δύο μηχανών A και B είναι εκθετικά κατανεμημένη, με $E\{T_A\}=1/\lambda_A$ και $E\{T_B\}=1/\lambda_B$ αντίστοιχα. Να βρεθεί
(α) η πιθανότητα και οι δύο μηχανές να αστοχήσουν στο διάστημα (0, t)
(β) η πιθανότητα η μηχανή A να αστοχήσει προ της μηχανής B.

Κανονική κατανομή

13) Χρησιμοποίησε τους στατιστικούς πίνακες για να προσδιορίσεις τις ακόλουθες πιθανότητες για μια τυπική κανονική κατανομή Z.
(α) $P(Z<1.32)$
(β) $P(Z<3.0)$
(γ) $P(Z>-2.15)$
(δ) $P(-2.34<Z<1.76)$

14) Χρησιμοποίησε τους στατιστικούς πίνακες για να προσδιορίσεις τις τιμές z της τυπικής κανονικής κατανομής , έτσι ώστε,
(α) $P(Z<z)$=0.9
(β) $P(Z<z)$=0,5

(γ) $P(Z>z)=0,1$
(δ) $P(-1,24<Z<z)=0,8$

15) Υποθέτουμε ότι η τυχαία μεταβλητή X ακολουθεί κανονική κατανομή, $X \sim N(\mu=10\ ,\ \sigma=2)$. Προσδιορίστε τα ακόλουθα.
(α) $P(X<13)$
(β) $P(X>9)$
(γ) $P(6<X<14)$
(δ) $P(-2<X<8)$

16) Υποθέτουμε ότι η τυχαία μεταβλητή X ακολουθεί κανονική κατανομή, $X \sim N(\mu=10\ ,\ \sigma=2)$. Προσδιορίστε τις τιμές x έτσι ώστε
(α) $P(X>x)=0,5$
(β) $P(X>x)=0,95$
(γ) $P(x<X<10)=0,2$
(δ) $P(-x<X-10<x)=0,95$.

17) Μία τυχαία μεταβλητή X είναι κατανεμημένη κανονικά με μέση τιμή 100 και διακύμανση 10^2.
(α) Ποια είναι η πιθανότητα ότι η τιμή μίας τυχαία παρατήρησης δεν διαφέρει από το 100 περισσότερο από 15;
(β) Ποια είναι η τιμή του c έτσι $P(|X-100| \le c) = 99\%$;
(γ) Ποια είναι η πιθανότητα ότι μία τυχαία παρατήρηση δεν διαφέρει από το 80 περισσότερο από 15;

18) Σε ένα πληθυσμό ανδρών το βάρος ακολουθεί κανονική κατανομή με μέση τιμή 80 κιλά και τυπική απόκλιση 8 κιλά.
(α) Ποια είναι η πιθανότητα ότι από έναν άνδρα ο οποίος επιλέγεται τυχαία το βάρος του είναι

- λιγότερο από 55 κιλά
- περισσότερο από 100 κιλά
- περισσότερο από 10 κιλά από το μέσο βάρος;

(β) Ποιο είναι το βάρος το οποίο δεν ξεπερνιέται από το 90% του πληθυσμού;
(γ) Δώστε το διάστημα γύρω από την μέση τιμή μέσα στο οποίο βρίσκεται το 50 % του πληθυσμού.

19) Η αντοχή συμπίεσης X των δειγμάτων τσιμέντου ακολουθεί κανονική κατανομή με μέση τιμή 6000 κιλών ανά τετραγωνικό εκατοστό και τυπική απόκλιση 100 κιλά ανά τετραγωνικό εκατοστό.
(α) Ποια είναι η πιθανότητα ότι σε ένα δείγμα η αντοχή είναι λιγότερο από 6450 κιλά/τ.εκ.;

(β) Ποια είναι η πιθανότητα ότι σε ένα δείγμα η αντοχή είναι μεταξύ 5800 και 5950 κιλά/τ.εκ.;
(γ) Ποια αντοχή ξεπερνιέται από το 95% των δειγμάτων τσιμέντου;

20) Η αντοχή έντασης κάποιου χαρτιού ακολουθεί κανονική κατανομή με μέση τιμή 25 κιλά ανά τετραγωνικό εκατοστό και τυπική απόκλιση 2 κιλά ανά τετραγωνικό εκατοστό.
(α) Ποια είναι η πιθανότητα ότι η αντοχή ενός δείγματος χαρτιού είναι λιγότερο από 30 κιλά/ τ. εκ.;
(β) Εάν οι προδιαγραφές απαιτούν η αντοχή έντασης του χαρτιού να ξεπερνά τα 30 κιλά/τ.εκ., ποιο ποσοστό των δειγμάτων απορρίπτεται;
(γ) Κατά πόσο πρέπει να βελτιωθεί η τυπική απόκλιση της αντοχής του χαρτιού έτσι ώστε το ποσοστό των δειγμάτων που απορρίπτεται στην ερώτηση (β) να ελαττωθεί κατά 50%;

21) Ο όγκος ανθρακούχου ποτού X σε μία αυτοματοποιημένη διαδικασία γέμισης σε κουτιά ακολουθεί κανονική κατανομή με μέση τιμή 450 κ.ε. και μία τυπική απόκλιση 5 κ.ε..
(α) Ποιο ποσοστό από τα κουτιά περιέχει ποτό περισσότερο από 460 κ.ε.;
Εάν η μέση τιμή γέμισης μπορεί να ρυθμισθεί εύκολα, αλλά η τυπική απόκλιση παραμένει σταθερή,
(β) σε πια τιμή πρέπει να ανέβει η μέση τιμή γέμισης έτσι ώστε 99% από τα κουτιά να περιέχουν τουλάχιστον 450 κ.ε. του ανθρακούχου ποτού;
(γ) Εάν η τυπική απόκλιση βελτιωθεί σε 3 κ.ε, σε πια τιμή πρέπει να ανέβει η μέση τιμή γέμισης έτσι ώστε 99% από τα κουτιά να περιέχουν τουλάχιστον 450 κ.ε. του ανθρακούχου ποτού;

22) Το βάρος των αθλητικών παπουτσιών που κατασκευάζει μία εταιρεία για τους σπρίντερς είναι τυχαία μεταβλητή με κανονική κατανομή μέσης τιμής 160 γραμμάρια και τυπική απόκλιση 6 γραμμάρια.
(α) Ποια η πιθανότητα ότι ένα αθλητικό παπούτσι της εταιρείας αυτής ζυγίζει περισσότερο από 170 γραμμάρια;
(β) Κατά πόσο πρέπει η εταιρεία να βελτιώσει την διακύμανση του βάρους των παπουτσιών προκειμένου να δηλώσει ότι 95% των παπουτσιών της έχουν βάρος λιγότερο από 165 γραμμάρια;
(γ) Εάν η τυπική απόκλιση του βάρους παραμείνει στα 6 γραμμάρια, πόσο πρέπει να κατέβει η μέση τιμή του βάρους προκειμένου η εταιρεία να δηλώσει ότι το 95% των παπουτσιών της έχουν βάρος λιγότερο από 165 γραμμάρια;

23) Υποθέτουμε ότι η τυχαία μεταβλητή X ακολουθεί διωνυμική κατανομή με n=200 και p=0,4. Με κανονική κατανομή
(α) προσεγγίστε την πιθανότητα ότι η X είναι μικρότερη ή ίση από 70,

(β) προσεγγίστε την πιθανότητα ότι η *X* είναι μεγαλύτερη από 70 και μικρότερη από 90,
(γ) προσεγγίστε την πιθανότητα ότι η X είναι 70.

24) Ένα σύνθετο σύστημα αποτελείται από 100 εξαρτήματα λειτουργούντα ανεξάρτητα. Η πιθανότητα ότι ένα εξάρτημα αποτυγχάνει κατά την λειτουργία του συστήματος είναι 0,1. Να βρεθεί με κανονική προσέγγιση η πιθανότητα ότι το λιγότερο 85% των εξαρτημάτων δεν αποτυγχάνουν κατά την λειτουργία του συστήματος.

25) Ένα φορτίο περιέχει 1000 ηλεκτρικούς συνδετήρες. Ένα δείγμα από 25 συνδετήρες επιλέγεται τυχαία από το φορτίο (χωρίς επανάθεση). Υποθέστε ότι το φορτίο περιέχει 100 ελαττωματικούς συνδετήρες.
(α) Χρησιμοποιείστε διωνυμική προσέγγιση, για να προσδιορίστε την πιθανότητα ότι δεν υπάρχουν καθόλου ελαττωματικοί συνδετήρες στο δείγμα.
(β) Χρησιμοποιείστε κανονική προσέγγιση για να απαντήσετε στο ερώτημα (α). Είναι ικανοποιητικό το αποτέλεσμα;
(γ) Απαντήστε στα ερωτήματα (α) και (β) αν το μέγεθος του φορτίου είναι 500. Η κανονική προσέγγιση στην πιθανότητα ότι δεν υπάρχουν ελαττωματικοί συνδετήρες στο δείγμα είναι ικανοποιητική σε αυτήν την περίπτωση;

ERLANG και GAMMA κατανομή

26) Ο αριθμός κλήσεων σε ένα τηλεφωνικό κέντρο ακολουθεί Poisson κατανομή.
(α) Ποιο είναι το όνομα της κατανομής και των σχετικών παραμέτρων, που ακολουθεί ο χρόνος μέχρι την 10η τηλεφωνική κλήση;
(β) Ποιος είναι ο αναμενόμενος χρόνος μέχρι την 10η κλήση;
(γ) Ποια είναι η μέση τιμή του χρόνου μεταξύ της 9ης και 10ης κλήσης;
(δ) Εάν επιλέγουμε ξεχωριστά 10 τυχαία διαστήματα του ενός λεπτού, ποια η πιθανότητα ότι όλα τα διαστήματα μαζί περιέχουν πάνω από δυο τηλεφωνικές κλήσεις;

27) Ένα υγρό υλικό διερευνάται για τη μόλυνσή του ή προσβολή του με άλλα ξένα σωματίδια. Έχει διαπιστωθεί ότι ο αριθμός ξένων σωματιδίων που προσβάλουν το υγρό ακολουθεί Poisson κατανομή με μέσο όρο 0,01 σωματίδια ανά κιλό.
(α) Ποιος είναι ο μέσος αριθμός των κιλών υγρού υλικού που χρειάζεται να ελέγξουμε μέχρι να βρεθούν 15 σωματίδια;
(β) Ποια είναι η τυπική απόκλιση των κιλών υγρού υλικού που χρειάζεται να ελέγξουμε μέχρι να βρεθούν 15 σωματίδια;

28) Ο χρόνος μεταξύ διαδοχικών προβλημάτων σε μία διαδικασία παραγωγής ακολουθεί εκθετική κατανομή με μέση τιμή 30 ημέρες.
(α) Ποιος ο αναμενόμενος χρόνος μέχρι το τέταρτο πρόβλημα;
(β) Ποια είναι η πιθανότητα ότι ο χρόνος μέχρι να εμφανισθεί το τέταρτο πρόβλημα ξεπερνά τις 120 ημέρες;

29) Χρησιμοποιήστε τις ιδιότητες της Γάμα κατανομής για να υπολογίστε τα ακόλουθα.
(α) Γ(6)
(β) Γ(5/2)
(γ) Γ(9/2).

WEIBULL κατανομή

30) Υποθέστε ότι η X ακολουθεί Weibull κατανομή με β=0,2 και δ=100 ώρες.
(α) Προσδιορίστε την μέση τιμή και διακύμανση της X.
(β) Προσδιορίστε τις πιθανότητες $P(X>5000)$ και $P(X<10{,}000)$.

31) Η μέση ζωή μιας κυκλοφοριακής αντλίας σε ένα σύστημα ύδρευσης ακολουθεί Weibull κατανομή με παραμέτρους β=2 και δ=700 ώρες.
(α) Προσδιορίστε την μέση διάρκεια ζωής της αντλίας. Πόσες αντλίες αναμένετε να χρησιμοποιηθούν στο σύστημα ύδρευσης για 10,000 ώρες λειτουργίας του.
(β) Ποια είναι η πιθανότητα ότι η αντλία θα λειτουργήσει περισσότερο από την μέση διάρκειά της;

Προβλήματα προς λύση

1) Καθίζηση σε κολώνες. Ένα οικοδόμημα στηρίζεται κυρίως στις κολώνες Α, Β, και Γ. Οι καθιζήσεις του εδάφους στις κολώνες αυτές X_A, X_B, και X_Γ, είναι τυχαίες μεταβλητές κανονικής κατανομής, $X_A \sim N(\mu_A=3$ εκ., $\sigma_A=0{,}5$ εκ.), $X_B \sim N(\mu_B=4$ εκ., $\sigma_B=0{,}5$ εκ.), και $X_\Gamma \sim N(\mu_\Gamma=4{,}5$ εκ., $\sigma_\Gamma=0{,}7$ εκ.).
(α) Ποια η πιθανότητα ότι καμία καθίζηση δεν θα ξεπεράσει τα 4,5 εκατοστά;
(β) Ποια η πιθανότητα η μέγιστη καθίζηση να ξεπεράσει τα 4,5 εκατοστά;

2) Ποιοτικός έλεγχος καρφιών. Το μήκος X, των καρφιών που παράγονται από μια μηχανή, είναι γνωστό ότι κατανέμονται κανονικά $X \sim N(\mu=4, \sigma=1)$. Σ' έναν ποιοτικό έλεγχο απορρίπτονται όσα καρφιά έχουν μήκος λιγότερο από 3 πόντους. Εάν ένας καταναλωτής επιλέγει τυχαία ένα καρφί ποια η πιθανότητα να είναι μακρύτερο από 4,5 πόντους;

3) Έλεγχος ποιότητας σκυροδέματος. Μία εταιρεία προμηθεύει ένα κατασκευαστικό έργο με σκυρόδεμα του οποίου η αντοχή X είναι τυχαία μεταβλητή κανονικής κατανομής, $X \sim N(\mu=18, \sigma=2)$kips. Ο μηχανικός που επιβλέπει το έργο θεωρεί ότι το σκυρόδεμα πρέπει να έχει ελάχιστη αντοχή 14 kips. Γι' αυτό σε κάθε φορτίο που καταφθάνει στο εργοτάξιο παίρνει τυχαία 5 δοκίμια σκυροδέματος και δοκιμάζει την αντοχή τους. Για να γίνει αποδεκτό το κάθε φορτίο πρέπει τουλάχιστον 4 δοκίμια να είναι αποδεκτής ποιότητας.
(α) Ποια η πιθανότητα ένα φορτίο να επιστραφεί;
(β) Αν την ημέρα έρχονται στο εργοτάξιο 2 φορτία σκυροδέματος. Να προσδιορίσετε με κανονική προσέγγιση την πιθανότητα ότι σε ένα μήνα θα γίνουν αποδεκτά τουλάχιστον 30 φορτία.

4) Αντοχή τσιμέντου. Η αντοχή συμπίεσης τσιμέντου έχει μία μέση τιμή 60 N/mm^2 και μία τυπική απόκλιση 4 N/mm^2 και υποτίθεται ότι ακολουθεί κανονική κατανομή.

Ποια η πιθανότητα ότι σε ένα έλεγχο 8 δοκιμών αντοχής συμπίεσης τσιμέντου όλες οι τιμές θα βρεθούν στο διάστημα από 46 μέχρι 70 N/mm^2;

5) Αποστράγγιση βρόχινων νερών. *Ένας μηχανικός σχεδιάζει ένα σύστημα υπονόμων για μία πόλη, έτσι ώστε να υπάρχει η δυνατότητα να απορροφούνται ικανοποιητικά τα βρόχινα νερά. Από πληροφορίες που έχουν συλλεχθεί, ο αριθμός καταιγίδων κατά την διάρκεια ενός έτους είναι τυχαία μεταβλητή, και μπορεί να πάρει ισοπίθανα τιμές από 1 μέχρι 4. Ακόμη, η μέγιστη απορροή μετά από μία καταιγίδα είναι επίσης τυχαία μεταβλητή κανονικής κατανομής με μέση τιμή 5 κ.μ./δευτερ. και τυπική απόκλιση 1κ.μ./δευτερ.. Εάν ο μηχανικός αποφασίσει η δυνατότητα απορρόφησης του συστήματος των υπονόμων να είναι 7 κ.μ./δευτερ.,*
(α) ποια είναι η πιθανότητα πλημμύρας κατά την διάρκεια μιας καταιγίδας στην πόλη αυτή;
(β) ποια είναι η πιθανότητα πλημμύρας μέσα σε ένα χρόνο;

6) Μελέτη αντιπλημμυρικού φράγματος. Σε μία παραθαλάσσια περιοχή μελετάται η κατά σκευή ενός αντιπλημμυρικού φράγματος. Η συχνότητα των τρικυμιών, που συμβαίνουν στην περιοχή αυτή και προκαλούν ανύψωση της επιφάνειας της θάλασσας, ακολουθεί Poisson κατανομή με μέσο όρο 2 τρικυμίες ανά χρόνο. Σε κάθε τρικυμία που συμβαίνει η ανύψωση της επιφάνειας της θάλασσας H είναι τυχαία μεταβλητή εκθετικής κατανομής με μέση τιμή 4 μέτρα.

Το κόστος κατασκευής του φράγματος(σε ευρώ) εξαρτάται από ένα σταθερό κόστος A και το ύψος του φράγματος , όπως φαίνεται στον επόμενο τύπο,
K=A+20,000 h.
Εάν σε μία τρικυμία η ανύψωση της επιφάνειας της θάλασσας ξεπεράσει το φράγμα θα προκληθεί πλημμύρα και ο εργολάβος αναλαμβάνει τις ζημιές, οι οποίες ανέρχονται σε 120,000 *ευρώ*. Η συμφωνία αυτή ισχύει για 10 χρόνια μετά την κατασκευή του φράγματος.

Να προσδιορίσετε το βέλτιστο ύψος h_β του φράγματος.

7) Παραγωγή ράβδων. Μία βιομηχανία παράγει σιδερένιες ράβδους των οποίων το μήκος X είναι τυχαία μεταβλητή, η οποία ακολουθεί κανονική κατανομή, $X\sim N(\mu=5,\ \sigma^2=1)$. Μία ράβδος είναι ελαττωματική αν το μήκος της βρίσκεται εκτός του διαστήματος (4,8, 5,2).
(α) Ποιο είναι το ποσοστό των ελαττωματικών ράβδων της βιομηχανίας αυτής;
(β) Σένα δείγμα 20 ράβδων αρχίζουμε να ελέγχουμε τις ράβδους μία προς μία. Ποια είναι η πιθανότητα ότι 2 ελαττωματικοί ράβδοι θα βρεθούν ακριβώς στον έλεγχο της πέμπτης ράβδου;
(γ) Η ημερήσια ζήτηση των ράβδων είναι τυχαία μεταβλητή και ακολουθεί Poisson κατανομή με μέση τιμή λ ράβδους/ημέρα. Το κέρδος για κάθε καλή ή ελαττωματική ράβδο είναι v ή κ αντίστοιχα. Το ημερήσιο συνολικό κέρδος Y είναι προφανώς τυχαία μεταβλητή, να βρεθεί το ημερήσιο συνολικό αναμενόμενο κέρδος.

8) Αξιοπιστία ηλεκτρικής μηχανής. Ο χρόνος λειτουργίας μιας ηλεκτρικής μηχανής μέχρι την πρώτη βλάβη ακολουθεί εκθετική κατανομή με μέση τιμή 20 μήνες. Σύμφωνα με το πρόγραμμα ελέγχου η μηχανή επιθεωρείται κάθε 4 μήνες και σε περίπτωση βλάβης επισκευάζεται.
(α) Ποια είναι η πιθανότητα ότι η μηχανή θα υποστεί βλάβη μέχρι την επόμενη επιθεώρηση(σε 4 μήνες);
(β) Αν μέχρι την πρώτη επιθεώρηση η μηχανή δεν παρουσίασε βλάβη, ποια είναι η πιθανότητα ότι δεν θα παρουσιάσει βλάβη και μέχρι την δεύτερη επιθεώρηση;
(γ) Σ' ένα μεγάλο εργοστάσιο λειτουργούν 5 ανεξάρτητες μεταξύ τους μηχανές. Ποια είναι η πιθανότητα ότι το πολύ μια μηχανή θα παρουσιάσει βλάβη σε 4 μήνες;
(δ) Αν θέλουμε να περιορίσουμε σε 7% την πιθανότητα βλάβης μιας μηχανής μέχρι την επόμενη επιθεώρηση προσδιορίστε το διάστημα αυτό.

9) Περίοδος επιστροφής πλημμύρας. Η μέγιστη ετήσια πλημμύρα προκαλεί μια άνοδο της επιφανείας του νερού σ' έναν ποταμό σε ύψος *Η*. Υποθέτουμε ότι το ύψος *Η* είναι τυχαία μεταβλητή κανονικής κατανομής $H \sim N(\mu=60, \sigma^2=16^2)$.
(α) Προσδιορίστε το ύψος h_{20} που έχει περίοδο επιστροφής ίση με 20 χρόνια(δηλαδή η επιφάνεια του ποταμού ξεπερνά το ύψος h_{20} μια φορά στα 20 χρόνια);
(β) Ποια είναι η πιθανότητα ότι , μέσα στα επόμενα 10 χρόνια. το ύψος *Η* θα ξεπεράσει το h_{20} τουλάχιστον μια φορά;

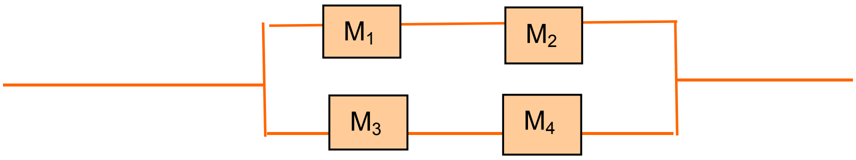
10) Τροφοδοσία συστήματος. Ο χρόνος καλής λειτουργίας, *X*, μιας μπαταρίας είναι τυχαία μεταβλητή και ακολουθεί κανονική κατανομή, $N(\mu=180$ ώρες, $\sigma^2=400$ ώρες$^2)$. Ένα σύστημα λειτουργεί σωστά αν τροφοδοτείται τουλάχιστον με 3 από τις 5 μπαταρίες του.
(α) Ποια η πιθανότητα ότι μια μπαταρία που τροφοδοτεί σωστά για 150 ώρες θα εξακολουθεί να τροφοδοτεί για τις επόμενες 50 ώρες;
(β) Ποια η πιθανότητα ότι το σύστημα θα λειτουργήσει σωστά τουλάχιστον 180 ώρες;
(γ) Ποιος ο ελάχιστος αριθμός μπαταριών που θα έπρεπε να διαθέτει το σύστημα, έτσι ώστε, η πιθανότητα αστοχίας του στο διάστημα [0,180 ώρες] να είναι μικρότερη του 0,4.

11) Αξιοπιστία παροχών ύδρευσης. Η ύδρευση μιας πόλης εξασφαλίζεται από τρεις ανεξάρτητες παροχές, Π_1, Π_2 και Π_3, όπου η κάθε παροχή δύναται να καλύψει το 50%, 60% και 70% των αναγκών της πόλης αντίστοιχα. Έστω ότι η διάρκεια των παροχών ακολουθεί εκθετική κατανομή με μέσες τιμές 3 χρόνια, 4 χρόνια και 5 χρόνια αντίστοιχα.
(α) Ποια η πιθανότητα ότι η πόλη θα έχει ικανοποιητική παροχή τον πρώτο χρόνο;
(β) Για τρία χρόνια η πόλη είχε ικανοποιητική παροχή, και όλες οι παροχές λειτουργούν καλά. Ποια η πιθανότητα ότι στα επόμενα 5 χρόνια θα υπάρξει ικανοποιητική παροχή νερού στην πόλη;

12) Αντλίες σε παράλληλη τοποθέτηση. Σένα μεγάλο αντλιοστάσιο αντλίες ίδιου τύπου σχεδιάζεται να τις τοποθετήσετε εν παραλλήλω. Η μέση διάρκεια λειτουργίας της κάθε αντλίας γνωρίζουμε από εμπειρία ότι είναι 10 χρόνια.

Ποιος ο ελάχιστος αριθμός των αντλιών που απαιτούνται έτσι ώστε το σύστημα θα αντλεί νερό συνεχώς για τρία χρόνια με πιθανότητα 0,95; (Υποθέστε εκθετική κατανομή)

13) Εξάρτηση μηχανών επεξεργασίας. Στο παρακάτω σύστημα επεξεργασίας ενός προιόντος ο χρόνος καλής λειτουργίας, *T*, των μηχανών M_1, M_2, M_3, M_4 είναι τυχαία μεταβλητή εκθετικής κατανομής με μέση τιμή α=50 ώρες. Οι μηχανές M_1 και M_2 αποτυγχάνουν ταυτόχρονα(όταν αποτυγχάνει η M_1 αμέσως αποτυγχάνει η M_2 και αντίστροφα) ενώ στους υπόλοιπους συνδυασμούς οι μηχανές είναι ανεξάρτητοι μεταξύ τους.
(α) Ποία η πιθανότητα ότι η λειτουργία του συστήματος δεν έχει σταματήσει πριν τις 80 ώρες;
(β) Να ευρεθεί η *σ.π.π.* $f_T(t)$ του χρόνου λειτουργίας του συστήματος.



Κατά πόσο αλλάζει η πιθανότητα της ερώτησης (α) αν όλες οι μηχανές λειτουργούν ανεξάρτητα;

14) Δομικές μηχανές σε εργοτάξιο. Μία δομική μηχανή αστοχεί με πιθανότητα 0,25 κατά την έναρξη της λειτουργίας της. Εάν η μηχανή δεν πάθει βλάβη στην αρχή, ο χρόνος καλής λειτουργίας της ακολουθεί την εξής συνάρτηση πυκνότητας πιθανότητας

$$f(x) = \begin{cases} 1{,}5 - \kappa x & \text{για } 0 < x < 1 \text{ έτος} \\ 0 & \text{για } x \geq 1 \text{ έτος} \end{cases}$$

(α) Να βρεθεί η συνάρτηση κατανομής πιθανότητας του χρόνου καλής λειτουργίας της μηχανής.
(β) Ποια η πιθανότητα ότι η μηχανή θα λειτουργήσει σωστά για έξι μήνες;
(γ) Ένα μεγάλο εργοτάξιο έχει 100 ίδιες δομικές μηχανές. Με κανονική προσέγγιση να υπολογιστεί η πιθανότητα ότι τουλάχιστον 20 μηχανές λειτουργούν σωστά μετά από 6 μήνες;

15) Παραθαλάσσια κατασκευή. Μία παραθαλάσσια κατασκευή έχει κατασκευαστεί έτσι ώστε να αντιμετωπίζει κύματα μέχρι 6 μέτρα πάνω από την μέση επιφάνεια της θάλασσας. Αν το ύψος των κυμάτων ξεπεράσει τα 6 μέτρα η κατασκευή αστοχεί .

Κάθε φορά που συμβαίνει τρικυμία στην περιοχή το ύψος των κυμάτων πάνω από την μέση επιφάνεια της θάλασσας, H, είναι τυχαίο μέγεθος και ακολουθεί κανονική κατανομή, $H \sim N(\mu_h = 5$ μέτρα, $\sigma_h = 1$μέτρο$)$. Ακόμη, είναι

γνωστό ότι στην περιοχή ο αριθμός τρικυμιών ακολουθεί Poisson κατανομή με μέση τιμή 2 τρικυμίες/έτος.
(α) Να υπολογιστεί η πιθανότητα ότι σε 5 τρικυμίες που θα συμβούν στην περιοχή, το πολύ σε δύο από αυτές το ύψος των κυμάτων θα ξεπεράσει τα 6 μέτρα.
(β) Να υπολογιστεί η πιθανότητα ότι η κατασκευή δεν θα αστοχήσει λόγω τρικυμίας μέσα στον επόμενο χρόνο.

16) Αντικατάσταση μηχανών. Η διάρκεια λειτουργίας T ηλεκτρικού εξαρτήματος είναι τυχαία μεταβλητή κατανεμημένη εκθετικά με μέση διάρκεια $E(T)=4$ έτη. Σένα μεγάλο σύστημα απατούνται 6 τέτοια εξαρτήματα. Όταν ένα εξάρτημα παρουσιάσει βλάβη, αντικαθίσταται από ένα καινούργιο.
(α) Ποια η πιθανότητα ότι δεν θα προκύψει ανάγκη αντικατάστασης εξαρτήματος μέσα σε ένα έτος;
(β) Ποια η πιθανότητα ότι θα προκύψει ανάγκη αντικατάστασης εξαρτήματος (ή εξαρτημάτων) για πρώτη φορά εντός του τρίτου έτους λειτουργίας του συστήματος;

17) Χωρητικότητα δρόμου. Ο αναμενόμενος ωριαίος όγκος κυκλοφορίας, X, για ένα προτεινόμενο δρόμο έχει σ.π.π. την ακόλουθη:

$$f(x)=\begin{cases} \dfrac{c}{500}x & \text{για} \quad 0<x\le 500 \\ -\dfrac{c}{200}x+3{,}5c & \text{για} \quad 500<x\le 700 \\ 0 & \text{για} \quad x>700 \end{cases}$$

Ο μηχανικός μπορεί να σχεδιάσει τον δρόμο έτσι ώστε η πραγματική χωρητικότητά του να είναι ίση με ένα από τα εξής:
- τη μέση τιμή της X, με πιθανότητα 0,25
- την πιθανότερη τιμή της X, με πιθανότητα 0,45
- την διάμεσο της X, με πιθανότητα 0,30.

(α) Να υπολογισθεί η χωρητικότητα του δρόμου σε κάθε μία από τις παραπάνω περιπτώσεις.
(β) Να υπολογισθεί η πιθανότητα ότι ο δρόμος που θα κατασκευασθεί θα ικανοποιεί τις κυκλοφοριακές ανάγκες.
(γ) Ποια η πιθανότητα της ερώτησης (β) αν η τυχαία μεταβλητή X ακολουθεί κανονική κατανομή $N(\mu=350, \sigma=40)$;

18) Μόλυνση υδάτων ποταμού. Η ημερήσια συγκέντρωση , X, μιας βλαβερής ουσίας στο νερό ενός ποταμού είναι τυχαία μεταβλητή με συνάρτηση

πυκνότητας πιθανότητας $f(x)=ce^{-cx}$. Μόλυνση του νερού προκύπτει όταν η συγκέντρωση ξεπεράσει τα 7 mg/1000*l*.
Αν η μέση ημερήσια συγκέντρωση είναι *E*(*X*)=1/c και η ημερήσια συγκέντρωση ξεπερνά τα 4 mg/1000*l* με πιθανότητα 0,20, *P*(*X*>4)=0,20,
(α) ποια είναι η πιθανότητα μόλυνσης σε μία συγκεκριμένη ημέρα;
(β) ποια είναι η περίοδος επιστροφής της μόλυνσης (μέσο χρονικό διάστημα μεταξύ δύο διαδοχικών μολύνσεων);
(γ) αν παρατηρούμε από σήμερα την συγκέντρωση βλαβερής ουσίας στον ποταμό, να υπολογιστεί η πιθανότητα ότι η δεύτερη μόλυνση θα συμβεί την δέκατη ημέρα.
Βοήθημα: $F(x)=1-e^{-cx}$

19) Αντοχή οδοστρώματος. Η διάρκεια αντοχής *X* ενός χιλιομέτρου οδοστρώματος είναι τυχαία μεταβλητή *εκθετικής* κατανομής, με μέση τιμή 5 χρόνια. Όταν το οδόστρωμα παρουσιάσει βλάβη επισκευάζεται αμέσως. Υποθέτουμε ακόμη ότι οι διάρκειες αντοχής των διαφορετικών χιλιομέτρων οδοστρώματος είναι στατιστικά ανεξάρτητες.
(α) Να υπολογισθεί η πιθανότητα ότι ένα χιλιόμετρο οδοστρώματος θα χρειαστεί επισκευές μέσα σε ένα χρόνο.
(β) Να υπολογισθεί η πιθανότητα ότι σε ένα τμήμα οδοστρώματος 4 χιλιομέτρων, μόνο 2 από αυτά θα χρειαστούν επισκευές μέσα σε 3 χρόνια.

20) Συντήρηση και αποθέματα εξαρτημάτων. Σε ένα ηλεκτρικό σύστημα η διάρκεια λειτουργίας Χ της ηλεκτρικής του λυχνίας είναι τυχαία μεταβλητή, *Εκθετικής κατανομής* με μέση τιμή 1700 ώρες. Όταν η λυχνία παρουσιάσει βλάβη αντικαθίσταται αμέσως με μία καινούργια και το σύστημα θεωρούμε ότι συνεχίζει την λειτουργία του χωρίς διακοπή.
(α) Δεδομένου ότι το σύστημα λειτουργεί 1500 ώρες με την ίδια λυχνία, ποια η πιθανότητα ότι δεν θα χρειαστεί αντικατάσταση της λυχνίας τις επόμενες 500 ώρες;
(β) Να υπολογιστεί η πιθανότητα ότι για να λειτουργήσει το σύστημα συνεχώς 17000 ώρες απαιτούνται 10 λυχνίες (θα υπάρξουν 10 βλάβες).
(γ) Ποιος ο ελάχιστος αριθμός διαθέσιμων λυχνιών έτσι ώστε να επαρκούν για την συνεχή λειτουργία του συστήματος 17000 ώρες, με πιθανότητα τουλάχιστον 90%;

21) Συσκευασία. Μία βιομηχανία πακετάρει ταμπλέτες σε μεταλλικά σωληνάρια, το κάθε ένα περιέχει 16 ταμπλέτες. Παρατηρήσεις δείχνουν ότι το πάχος της ταμπλέτας μεταβάλλεται γύρω από μία μέση τιμή 0,2 πόντους με τυπική απόκλιση 0,01 πόντους. Το μήκος του σωληναρίου μπορεί να έχει οιαδήποτε μέση τιμή αλλά έχει τυπική απόκλιση 0,03 πόντους. Πια μέση τιμή του μήκους του σωληναρίου θα εξασφάλιζε ότι 99% των σωληναρίων θα είχαν αρκετό χώρο για το περιεχόμενό τους; Με αυτή την μέση τιμή, πια είναι η πιθανότητα ότι ο κενός χώρος στο σωληνάριο ξεπερνά τους 0,25 πόντους;

22) Αριθμός κράτησης θέσεων. Μια αεροπορική εταιρεία κάνει 200 κρατήσεις για μία πτήση η οποία μπορεί να εξυπηρετήσει μόνο 185 ταξιδιώτες. Η πιθανότητα ότι ένας ταξιδιώτης που έχει κρατήσει θέση φθάνει στο αεροδρόμιο για την πτήση είναι 0,9. Υποτίθεται ότι οι ταξιδιώτες συμπεριφέρονται ανεξάρτητα.
(α) Με κανονική προσέγγιση να υπολογίστε την πιθανότητα ότι θα υπάρξουν θέσεις για όλους τους ταξιδιώτες που θα έρθουν για την πτήση αυτή.
(β) Με κανονική προσέγγιση να υπολογίστε την πιθανότητα ότι θα υπάρξουν κενές θέσεις στο αεροπλάνο.
(γ) Με κανονική προσέγγιση προσδιορίστε τον αριθμό κρατήσεων που πρέπει να κάνει η εταιρεία για την πτήση αυτή έτσι ώστε ο κάθε ταξιδιώτης που έρχεται για την πτήση να βρει θέση με πιθανότητα 0,95.

23) Αριθμός λειτουργούντων εκσκαφέων. Ένα μεγάλο κατασκευαστικό έργο ξεκινά με καινούργια χωματουργικά μηχανήματα. Η διάρκεια καλής λειτουργίας των χωματουργικών μηχανημάτων είναι τυχαία μεταβλητή κανονικής κατανομής. Το 30% των μηχανημάτων αυτών είναι παλαιάς τεχνολογίας με μέση διάρκεια λειτουργίας 15,000 ώρες και τυπική απόκλιση 2000 ώρες, ενώ τα υπόλοιπα είναι νέου τύπου με μέση διάρκεια λειτουργίας 25,000 ώρες και την ίδια τυπική απόκλιση. Ποιο ποσοστό από τα μηχανήματα συνεχίζει να λειτουργεί μετά από 20,000 ώρες εργασίας;

24) Μελέτη αποστράγγισης βρόχινων νερών. Μία πόλη έχει πρόβλημα με τα βρόχινα νερά, όταν συμβαίνει έντονη βροχόπτωση πλημμυρίζουν οι δρόμοι. Για τον λόγο αυτόν μελετάται ένα αποστραγγιστικό σύστημα με δυνατότητα απορρόφησης των βρόχινων νερών μέχρι 10 κυβικά μέτρα ανά δευτερόλεπτο. Από εμπειρία, είναι γνωστό ότι ο αριθμός καταιγίδων X στην πόλη αυτή είναι τυχαία μεταβλητή Poisson κατανομής με μέση συχνότητα 3 τον χρόνο. Ακόμη, Σε κάθε καταιγίδα η μέγιστη ποσότητα απορροής Y είναι στοχαστικό μέγεθος και ακολουθεί **λογάριθμο-κανονική κατανομή** με μέση τιμή 8 κυβικά μέτρα ανά δευτερόλεπτο και συντελεστή μεταβλητότητας 20%.
Κατασκευάζεται το αποστραγγιστικό έργο στη πόλη.
(α) Ποια η πιθανότητα στην επόμενη καταιγίδα να μην πλημμυρίσουν οι δρόμοι;
(β) Ποια η πιθανότητα να μην έχουμε πλημμύρα τον επόμενο χρόνο;
(γ) Ποια η πιθανότητα στα επόμενα 10 χρόνια να συμβούν το πολύ 2 πλημμύρες;

25) Βροχόπτωση και απορροή. Η ποσότητα ετήσιας βροχόπτωσης *X* σε μία περιοχή είναι τυχαίο μέγεθος κανονικής κατανομής με μέση τιμή 1000 χιλιοστά και απόκλιση 200 χιλιοστά. Η ετήσια απορροή των βρόχινων νερών *Y* στην περιοχή αυτή είναι συνάρτηση της ετήσιας βροχόπτωσης,

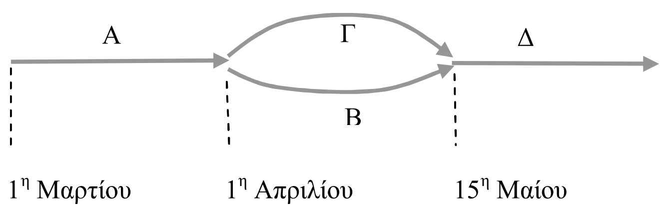
$$Y=100+0{,}6X.$$

(α) Προσδιορίστε την κατανομή της τυχαίας μεταβλητής Y.
(β) Ποια η πιθανότητα ότι μία χρονιά η ετήσια απορροή θα ξεπεράσει τα 900 χιλιοστά;
(γ) Ποια η αναμενόμενη τιμή της *Y* καθώς και η τυπική της απόκλιση;

26) Εκτροπή ποταμού. *Ένας ποταμός έχει ετήσια ροή* X, *η οποία είναι κανονικής κατανομής,* $X \sim N(\mu_X=400,\ \sigma_X=60)$. *Σε κάποιο σημείο ενώνεται με ένα μικρότερο ποτάμι του οποίου η ετήσια ροή* Y *είναι επίσης κανονικής κατανομής,* $Y \sim N(\mu_Y=200,\ \sigma_Y=40)$. *Οι μονάδες μέτρησης ετήσιας ροής αναφέρονται σε 1000 κυβικά μέτρα. Στην συνέχεια, σε κάποιο άλλο σημείο γίνεται μία τμηματική εκτροπή του ποταμού με ετήσια ροή* Z, *κανονικής κατανομής,* $Z \sim N(\mu_Z=150,\ \sigma_Z=40)$. *Υποθέστε ότι οι ετήσιες ροές νερού* X, Y, *και* Z *είναι ανεξάρτητες μεταξύ τους.*
(α) Προσδιορίστε την κατανομή της τελικής ετήσιας ροής R *του κεντρικού ποταμού μετά το σημείο εκτροπής.*
(β) Ποια η πιθανότητα ότι η ετήσια ροή R *θα ξεπεράσει τα 700,000 κυβικά μέτρα;*

27) Σχεδιασμός δραστηριοτήτων. *Οι δραστηριότητες σε ένα έργο φαίνονται στο Σχήμα. Μία δραστηριότητα δεν μπορεί να αρχίσει αν δεν έχουν τελειώσει όλες οι προηγούμενες, ακόμη δεν μπορούν να αρχίσουν νωρίτερα από την προγραμματιζόμενη ημερομηνία όπως φαίνεται στην αρχή της κάθε δραστηριότητας. Η διάρκεια* X *της κάθε δραστηριότητας είναι τυχαία μεταβλητή κανονικής κατανομής:*

$X_A \sim N(\mu_A=20$ *ημέρες*, $\sigma_A=4$ *ημέρες*),
$X_B \sim N(\mu_B=45$ ημέρες, $\sigma_B=6$ ημέρες),
$X_\Gamma \sim N(\mu_\Gamma=30$ ημέρες, $\sigma_\Gamma=5$ ημέρες),



(α) Να σχολιασθεί αν η υπόθεση ότι οι διάρκειες των δραστηριοτήτων ακολουθούν κανονική κατανομή είναι λογική.
(β) Εάν η δραστηριότητα Α αρχίσει την 1η Μαρτίου ,πια η πιθανότητα ότι οι δραστηριότητες Β και Γ θα αρχίσουν την 1η Απριλίου;
(γ) Εάν οι Β και Γ δεν καθυστερήσουν την έναρξή τους, πια η πιθανότητα ότι θα καθυστερήσει την έναρξή της η δραστηριότητα Δ;

28) Ημερήσια βάρδια. Μια μηχανή παραγωγής ρεύματος τροφοδοτεί ένα εργοστάσιο όλο το 24ώρο. Ο χρόνος καλής λειτουργίας *X* της μηχανής είναι τυχαία μεταβλητή εκθετικής κατανομής, $f(x)=0{,}02e^{-0{,}02x}$. Εάν η μηχανή πάθει βλάβη οιαδήποτε στιγμή μεταξύ τις 7 π.μ. και 7 μ.μ., ο υπεύθυνος μηχανικός την αντικαθιστά αμέσως με μία παρόμοια και η παροχή ρεύματος δεν διακόπτεται. Εάν η μηχανή πάθει βλάβη οιαδήποτε στιγμή μετά τις 7 μ.μ., σταματά η παροχή ρεύματος μέχρι το πρωί (7 π.μ.).
(α) Ποια η πιθανότητα ότι μία μηχανή παρέχει ρεύμα συνεχώς όλο το 24ωρο:
(β) Ποια η πιθανότητα ότι ο μηχανικός μία μέρα θα αντικαταστήσει δύο φορές τη μηχανή;
(γ) Ποιος ο μέσος χρόνος παροχής ρεύματος στο 24ωρο;

29) Επισκευές οδοστρώματος. Η διάρκεια αντοχής *X* ενός χιλιομέτρου οδοστρώματος είναι τυχαία μεταβλητή λογάριθμο-κανονικής κατανομής με μέση 4 χρόνια και συντελεστή μεταβλητότητα 60%. Υποθέτουμε ότι η αντοχή οδοστρώματος μεταξύ των χιλιομέτρων είναι στατιστικώς ανεξάρτητη.
(α) Ποια η πιθανότητα ότι ένα χιλιόμετρο δεν θα χρειασθεί επισκευή μέσα στα δύο πρώτα χρόνια;
(β) Ένα χιλιόμετρο οδοστρώματος δεν υπέστη ζημιές μέχρι τον τέταρτο χρόνο, ποια η πιθανότητα η διάρκεια αντοχής του να ξεπεράσει τα 6 χρόνια;
(γ) Ποια η πιθανότητα ότι για ένα μήκος οδοστρώματος 10 χιλιομέτρων τον πρώτο χρόνο δεν θα προκύψει ανάγκη επισκευής;
(δ Ποια η πιθανότητα ότι για ένα μήκος οδοστρώματος 10 χιλιομέτρων στα πρώτα 4 χρόνια θα προκύψει ανάγκη επισκευής μόνο σε 2 χιλιόμετρα;

30) Αξιοπιστία συστήματος παραγωγής. Κατά την επεξεργασία ενός προιόντος, το προιόν διέρχεται εν σειρά μέσα από τρεις τύπους μηχανών Α, Β, και Γ. Η παραγωγή σταματά εάν αστοχήσει τουλάχιστον μία από τρεις μηχανές.
Έστω οι χρόνοι διάρκειας ζωής των μηχανών είναι τυχαίες μεταβλητές εκθετικής κατανομής, με μέσες τιμές 130, 135 και 140 ώρες αντίστοιχα.
(α) Ποια η πιθανότητα ότι το σύστημα θα αστοχήσει εντός 100 ωρών από την έναρξη της λειτουργίας του;
(β) Ποια η μέση διάρκεια ζωής του συστήματος;

31) Εκφόρτωση πλοίων. Σένα λιμάνι υπάρχει μόνον ένας γερανός εκφόρτωσης πλοίων. Το κόστος των εγκαταστάσεων γερανών είναι μεγάλο. Έχει παρατηρηθεί ότι ο αριθμός φορτηγών πλοίων που φθάνουν στο λιμάνι είναι τυχαία μεταβλητή Poisson κατανομής με μέση συχνότητα 5 πλοία ανά 24ωρο. Σύμφωνα με τους διεθνείς κανονισμούς πρέπει το 80% των πλοίων να ξεφορτώνονται χωρίς καθυστέρηση.
(α) Ποια πρέπει να είναι η διάρκεια εκφόρτωσης ενός πλοίου;
(β) Ποια η πιθανότητα ότι από τα πέντε διαδοχικά πλοία που φθάνουν στο λιμάνι κανένα δεν θα περιμένει για εκφόρτωση;
(γ) Ποια η πιθανότητα ότι από 20 πλοία που φθάνουν στο λιμάνι το πολύ 4 θα περιμένουν για εκφόρτωση;

32) Προστατευτική κατασκευή. Για την κατασκευή των θεμελίων μιας γέφυρας στο λιμάνι πρέπει να προστατευθεί ο χώρος από την τυχών ανύψωση της επιφάνειας του νερού λόγω κυμάτων. Γι' αυτό πρέπει να εγκατασταθεί μία προστατευτική κατασκευή ύψους H, η οποία να προστατεύει τον χώρο θεμελίωσης με σιγουριά 98%. Είναι γνωστό ότι η μέγιστη μηνιαία ανύψωση της επιφάνειας του νερού X, λόγω κυμάτων, είναι τυχαία μεταβλητή κανονικής κατανομής, $X \sim N(\mu=3$μέτρα, $\sigma=0{,}5$μέτρα).
(α) Ποιο πρέπει να είναι το ύψος της προστατευτικής κατασκευής αν η θεμελίωση διαρκέσει ένα μήνα;
(β) Ποιο πρέπει να είναι το ύψος της προστατευτικής κατασκευής αν η θεμελίωση διαρκέσει 5 μήνες;
(γ) Ο εργολάβος έχει την δυνατότητα να συντομεύσει την διάρκεια της θεμελίωσης της γέφυρας από 5 μήνες σε 3 μήνες με επί πλέον επιβάρυνση 12000 ευρώ. Και αν το κόστος κατασκευής της προστατευτικής επιφάνειας ανέρχεται σε 25000 ευρώ ανά ένα μέτρο ύψους, τι θα προτείνατε στον εργολάβο;

33) Χρόνος αναμονής. Ο χρόνος μεταξύ διαδοχικών αφίξεων αεροπλάνων σε ένα αεροδρόμιο είναι τυχαία μεταβλητή εκθετικής κατανομής με μέση τιμή 25λεπτά. Το αεροδρόμιο διαθέτει μόνο ένα σημείο αποβίβασης επιβατών και αποσκευών, και ο χρόνος εξυπηρέτησης για ένα αεροπλάνο εκτιμάται ότι είναι 20 λεπτά.
(α) Ποιο ποσοστό αεροπλάνων που φθάνει στο αεροδρόμιο εξυπηρετείται αμέσως, χωρίς καθυστέρηση;
(β) Ποιος ο μέσος χρόνος αναμονής πριν αρχίσει η εκφόρτωση του αεροπλάνου;

34) Σχεδιασμός αναχωρήσεων λεωφορείων. Λεωφορεία αναχωρούν από το κέντρο της πόλης προς το αεροδρόμιο. Ο αριθμός ταξιδιωτών που φθάνουν στην αφετηρία των λεωφορείων με προορισμό το αεροδρόμιο μπορούμε να υποθέσουμε ότι είναι τυχαία μεταβλητή κανονικής κατανομής ,

$$X \sim N(\mu=120/\text{ώρα}, \sigma=20/\text{ώρα}).$$

(α) Εάν η χωρητικότητα των λεωφορείων είναι 45 θέσεων, ποια η συχνότητα αναχώρησης των λεωφορείων για αεροδρόμιο έτσι ώστε όλοι βρίσκουν θέση στο λεωφορείο με πιθανότητα 90%;
(β) Με την συχνότητα αναχώρησης που προσδιορίσατε στην (α) ερώτηση, ποια η πιθανότητα σε 8 αναχωρήσεις να σημειωθεί το πολύ σε μία συνωστισμός;

35) Χρόνος πρόσληψης προσωπικού. Η διάρκεια λειτουργίας μιας δομικής μηχανής είναι μεταβλητή εκθετικής κατανομής, με μέση τιμή 3000 ώρες. Όταν λειτουργεί η μηχανή το κόστος ανέρχεται σε 100 ευρώ την ώρα , ενώ το εισόδημα από την παραγωγή της μηχανής υπολογίζεται 200 ευρώ την ώρα. Το προσωπικό που χειρίζεται την μηχανή προσλαμβάνεται για 4000 ώρες και τα έξοδά του ανέρχονται σε 40 ευρώ την ώρα.
(α) Ποιο το αναμενόμενο κέρδος;
(β) Για ποιο χρόνο πρόσληψης του προσωπικού μεγιστοποιείται το κέρδος;

36) Χαμηλή ροή ποταμού σε ξερικές περιοχές. Τα ακόλουθα δεδομένα αφορούν την ετήσια ελάχιστη μέση ημερήσια ροή σε κυβικά μέτρα ανά δευτερόλεπτο, ενός ποταμού που διασχίζει μία ξερική και άγονη περιοχή. Τα στατιστικά δεδομένα προέρχονται από τα τελευταία 20 χρόνια.

2,5 2,8 1,1 3,9 1,5 1,6 2,8 3,1 1,2 1,8 1,7 2,6 2,1 2,4 3,9 1,4 4,8 2,6 1,8 1,2.

Υποθέτουμε ότι η ετήσια ελάχιστη μέση ημερήσια ροή ακολουθεί την **Weibull** με δύο παραμέτρους κατανομή. Να υπολογιστεί η πιθανότητα ότι τα επόμενα δύο χρόνια η ετήσια ελάχιστη ημερήσια μέση ροή θα είναι λιγότερο από 2 κυβικά νερό ανά δευτερόλεπτο.