Počítačová simulace logistických procesů II

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

› Jan Fábry

› 28.10.2017

Počítačová simulace logistických procesů II

Obsah předmětu

I.	Úvod, organizace, semestrální projekty, projekty Škoda
II.	Vysvětlení témat semestrálního projektu
III.	Analýza dat
IV.	Analýza dat
V.	Plant Simulation
VI.	Plant Simulation, pojmový model
VII.	Struktura simulačního modelu
VIII.	Tvorba simulačního modelu
IX.	**Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace**
X.	Simulační experimentování
XI.	Důsledky na reálný systém, Process Designer
XII.	Rozhraní (ProcessDesigner, MALAGA, TriCAD)

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Cíl přednášky

› Seznámit posluchače s problematikou náhodných čísel a náhodného pokusu.

› Uvést přehled běžných typů rozdělení pravděpodobnosti náhodných proměnných.

› Vysvětlit principy provádění testování hypotéz.

› Vysvětlit princip optimalizace ve spojení se simulačními modely.

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

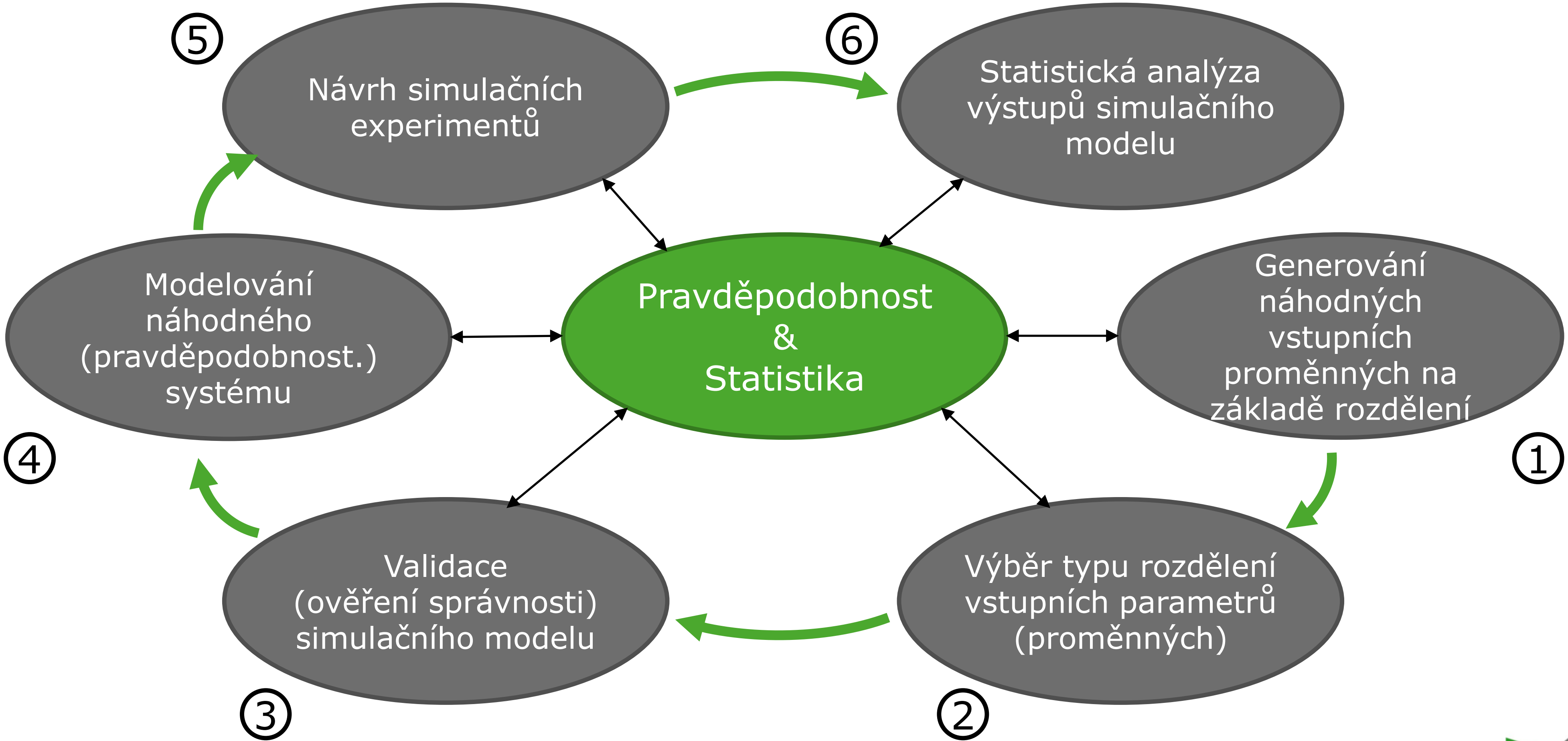
Obsah přednášky

- Náhodná proměnná (veličina), náhodný výběr, náhodný pokus.
- Charakteristiky náhodných proměnných.
- Typy rozdělení pravděpodobnosti náhodných proměnných.
- Testování hypotéz.
- Optimalizace.

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Náhodná proměnná, náhodný výběr, náhodný pokus

› Motivace – proč se zabývat stochastickými procesy a náhodnými veličinami?



9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Náhodná proměnná, náhodný výběr, náhodný pokus

- Příklady stochastických procesů, které jsou často modelovány
  - **Systém hromadné obsluhy** – vstupní parametry povahy náhodné proměnné:
    - Doba mezi vstupy zakázek.
    - Doba obsluhy zakázky.
    - …
  - **Zásobovací procesy** – vstupní parametry povahy náhodné proměnné:
    - Doba mezi objednávkou a vychystáním/dodáním materiálu.
    - Množství materiálu požadovaných při objednávce.
    - …
  - **Spolehlivost a údržba** – vstupní parametry povahy náhodné proměnné:
    - Doba mezi prostoji.
    - Trvání prostoje.
    - …

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Náhodná proměnná, náhodný výběr, náhodný pokus

Vymezení základních pojmů

- **Náhodný pokus** ≡ proces, jehož výsledek není s jistotou (dopředu) znám.

- **Náhodný jev** ≡ výsledek náhodného pokusu; požadujeme tyto vlastnosti:
  - **Hromadnost** (≡ dostatečná opakovatelnost).
  - **Stabilita** (≡ neměnnost pokusu).

- **Pravděpodobnost** ≡ číselné vyjádření míry možnosti nastoupení náhodného jevu.

- **Elementární jev ω** ≡ možný výsledek pokusu.

- **Základní prostor Ω** ≡ množina všech možných výsledků experimentu (elementárních jevů), tj. platí: $\forall \omega\colon \omega \in \Omega$ …speciální případy jsou jistý jev, resp. nemožný jev.

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Náhodná proměnná, náhodný výběr, náhodný pokus

Vymezení základních pojmů

- **Náhodný pokus** je proces, který při opakování může za stejných podmínek dávat rozdílné výsledky.
- **Náhodná proměnná** ≡ funkce, která každému výsledku náhodného pokusu přiřadí reálné číslo.
  - **Diskrétní náhodná proměnná** – může nabývat pouze konečný počet hodnot nebo hodnotu z nekonečné množiny jednoznačných hodnot.
    - Např.: trefa/minutí (0/1), počet čekajících zakázek {0, 1, 2, …}, počet vadných dílů stroje {0, 1, …, 1250 …celkový počet dílů stroje}.
  - **Spojitá náhodná proměnná** – může nabývat libovolnou hodnotu z určitého intervalu
    - Např.: doba mezi vstupy zakázek [s].

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Náhodná proměnná, náhodný výběr, náhodný pokus

- **Náhodný výběr** ~ při *n* opakováních náhodného pokusu získáme *n* hodnot náhodné proměnné; číslo *n* je **rozsah** náhodného výběru.
  - Tedy náhodný výběr je *n*-rozměrný náhodný vektor $X = (X_1, \dots, X_n)$.

- **Statistický soubor** s rozsahem *n* ≡ pozorovaná hodnota náhodného výběru.

- **Výběrový prostor** ≡ množina všech hodnot náhodného výběru, tj. množina všech statistických souborů.
- **Statistika** neboli **výběrová charakteristika** ≡ funkce náhodného výběru.

- **Empirická statistika** neboli **pozorovaná hodnota statistiky** ≡ hodnota statistiky na statistickém souboru.

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Náhodná proměnná, náhodný výběr, náhodný pokus

Náhodná proměnná
X
…abstraktní
Teoretická charakteristika
ϑ

~ náhodný výběr „několika" hodnot

Náhodný výběr
$(X_1, \ldots, X_n)$
…abstraktní
Výběrová charakteristika (statistika)
$T(X_1, \ldots, X_n)$

~ pozorování
~ pozorování

Statistický soubor
$(x_1, \ldots, x_n)$
Zpracování
…konkrétní
Empirická charakteristika
$t = T(x_1, \ldots, x_n)$

Příklad: výběrový průměr $\bar{X}$
jediná hodnota, nelze ji zjistit

aritmetický průměr $\bar{x}$
náhodná hodnota

Empirické charakteristiky se při opakovaných realizacích náhodného výběru náhodně mění.
Pro $n \to \infty$ se aritmetický průměr blíží neznámé střední hodnotě: $\bar{x} \to E(X)$

Jediné, s čím lze „pracovat" (co „existuje"), je statistický soubor. Náhodná proměnná a náhodný výběr jsou myšlenkové konstrukce…

Zdroj: Karpíšek, Z.: Statistická analýza – Přehledový učební text pro doktorské studium, 2008, Brno, FIS VUT v Brně

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Náhodná proměnná, náhodný výběr, náhodný pokus

- **Diskrétní náhodná proměnná (DNP)** – může nabývat pouze konečný počet hodnot nebo hodnotu z nekonečné množiny jednoznačných hodnot.
  - **Pravděpodobnostní funkce** popisuje pravděpodobnost, že DNP nabývá určité hodnoty:

$$p(x_i) = P(X = x_i)$$

  - Jde o funkci následujících vlastností:

$$0 \leq p(x_i) \leq 1 \qquad p(S) = 1 \; pro \; S \equiv všechny \; možnosti$$

  - **Kumulativní pravděpodobnostní funkce** popisuje pravděpodobnosti, že DNP nabývá hodnoty *nižší nebo rovné* hodnotě x_i

$$F(x) = P(X \leq x)$$

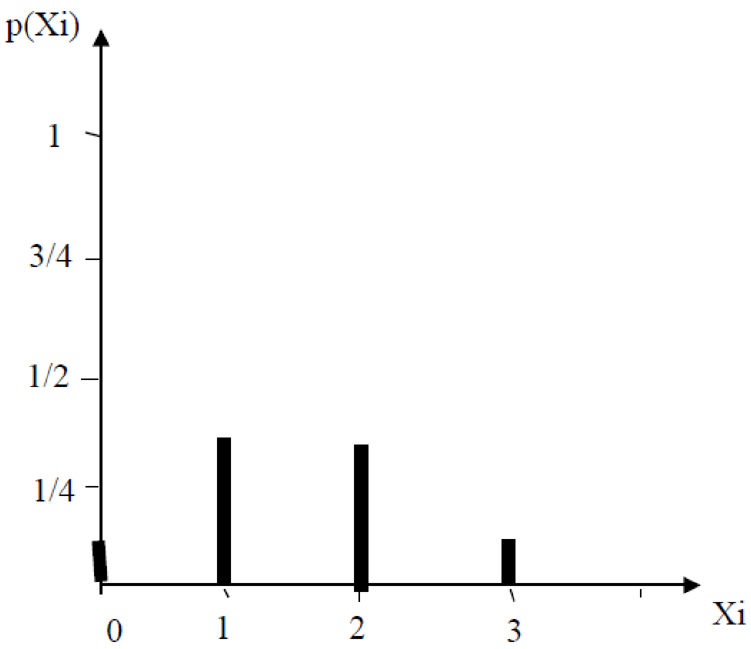
  - Jde o funkce následujících vlastností:

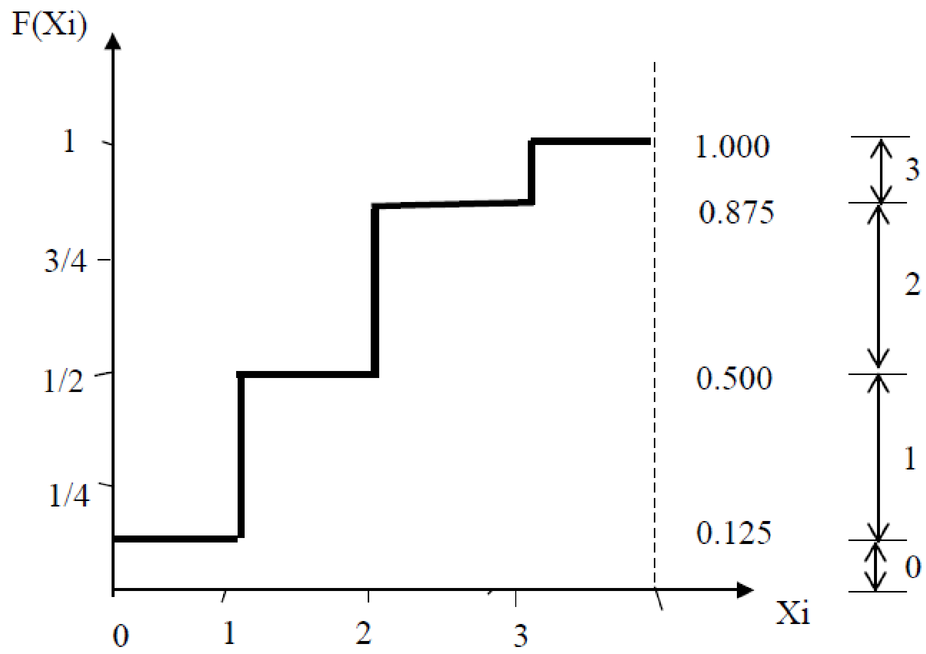
$$0 \leq F(x) \leq 1 \quad F(-\infty) = 0 \qquad F(\infty) = 1$$

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Náhodná proměnná, náhodný výběr, náhodný pokus

- **Diskrétní náhodná proměnná (DNP)** – může nabývat pouze konečný počet hodnot nebo hodnotu z nekonečné množiny jednoznačných hodnot.
  - **Pravděpodobnostní funkce** $p(X_i)$.
  - **Kumulativní pravděpodobnostní funkce** $F(X_i)$.





9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Náhodná proměnná, náhodný výběr, náhodný pokus

- **Spojitá náhodná proměnná (SNP)** – může nabývat hodnotu spadající do určitého intervalu hodnot, tj. hodnoty z konečné či nekonečné množiny hodnot.
  - **Hustota pravděpodobnosti** je funkce $f(x)$ popisující pravděpodobnost, že SNP nabývá určité hodnoty spadající do uzavřeného intervalu $\langle a; b\rangle$:

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

  - Jde o funkci následujících vlastností:

$$\mathrm{f(x)} \geq 0 \qquad \int_{-\infty}^{\infty} f(x)dx = 1$$

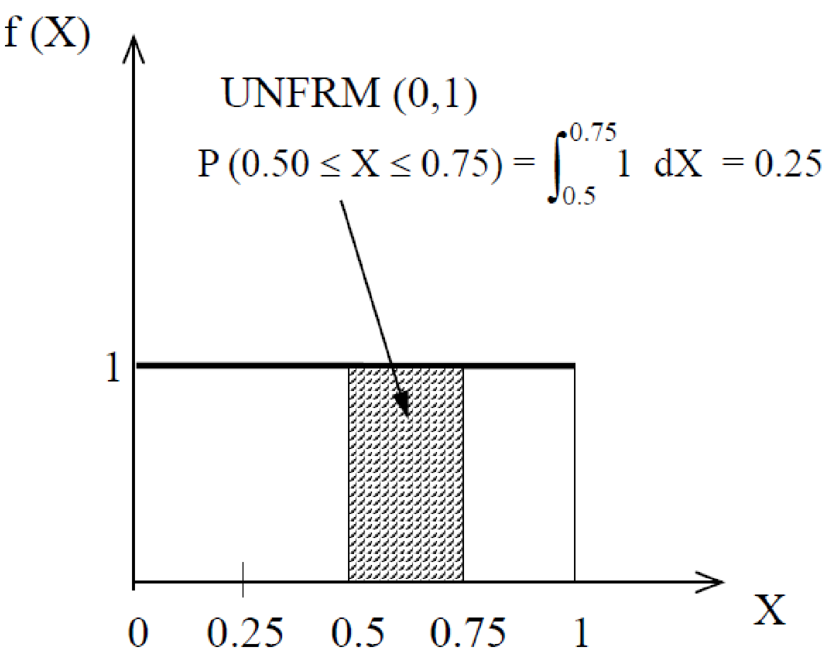
  - **Kumulativní distribuční funkce** popisuje pravděpodobnosti, že SNP nabývá hodnoty *nižší nebo rovné* hodnotě x:

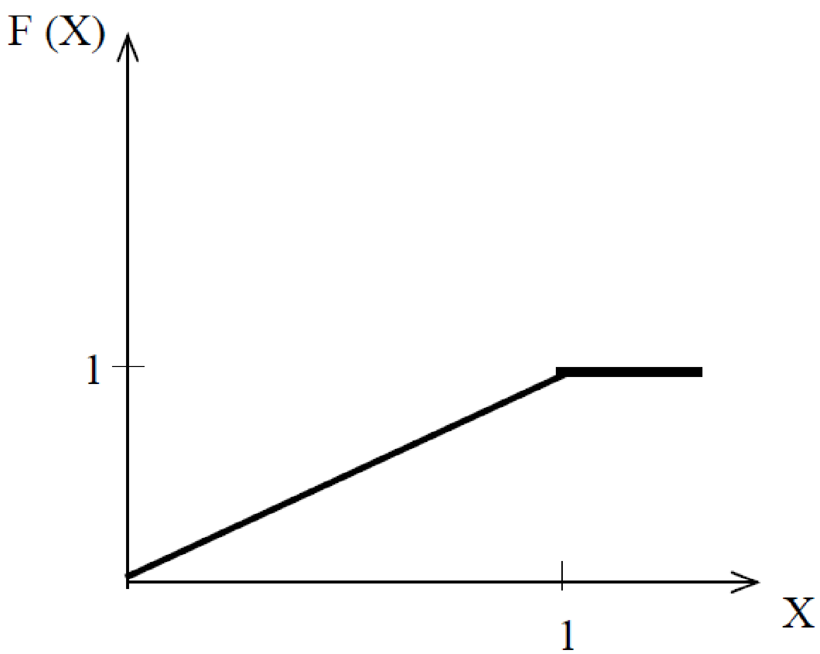
$$F(x) = \int_{-\infty}^{x} f(y)dy = P(X \leq x)$$

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Náhodná proměnná, náhodný výběr, náhodný pokus

- **Spojitá náhodná proměnná (SNP)** – může nabývat hodnotu spadající do určitého intervalu hodnot, tj. hodnoty z konečné či nekonečné množiny hodnot.
  - **Hustota pravděpodobnosti** $f(X)$.
  - **Kumulativní distribuční funkce** $F(X)$.





9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Charakteristiky náhodných proměnných

- **Charakteristiky** ≡ reálná čísla vyjadřující koncentrovaně důležité vlastnosti náhodného výběru (náhodné veličiny X), resp. statistického souboru.

- **Nejdůležitější charakteristiky**
  - Střední hodnota.
  - Rozptyl.

  Speciální případy tzv. obecného (centrálního) momentu.
  - P-Kvantily (100P%-Kvantily)
    - Medián,
    - Kvartil, percentil, centil, decil.
  - Modus.
  - Šikmost (asymetrie) a špičatost (exces).
  - Aritmetický průměr, vážený průměr, geometrický průměr, harmonický průměr.

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Charakteristiky náhodných proměnných

- **Střední hodnota** náhodné proměnné X: $E[X]$

$E[X] = \sum_{i=1}^{n} x_i \cdot p(x_i)$ ...DNP

$E[X] = \int_{-\infty}^{\infty} x \cdot f(x)dx$ ...SNP

- Obecně se jedná o funkci x:

$E[X^n] = \sum_{i=1}^{n} x_i^n \cdot p(x_i)$ ...DNP

$E[X^n] = \int_{-\infty}^{\infty} x^n \cdot f(x)dx$ ...SNP

- Dle definice je **střední hodnota** x^n označována jako n-tý moment náhodné veličiny x.

- Pokud $n = 1$, hovoříme o *prvním momentu* náhodné veličiny x, který právě nazýváme střední hodnota.

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Charakteristiky náhodných proměnných

- **Rozptyl (disperze)** je tzv. *druhý moment* náhodné veličin x proměnné: σ^2

  $\sigma^2 = Var[X] = E[(X - E[X])^2]$ …druhá mocnina kvůli stejnému zohlednění kladných i záporných hodnot výrazu (rozdílů) $X - E[X]$, tj. „vzdálenosti" od střední hodnoty.

- Odmocnina σ se nazývá **směrodatná odchylka** a popisuje, nakolik jsou hodnoty náhodné proměnné x „rozprostřeny" od střední hodnoty $E[X]$.

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Charakteristiky náhodných proměnných

- **P-Kvantil (100P%-Kvantil)** náhodné proměnné ≡ je hodnota $x_P$náhodné proměnné X, pro kterou platí $x_p = \inf\{x; F(x) \geq P\}$ …infimum množiny

  ~ hodnota rozdělující setříděný výběr prvků na dvě části, kde jedna obsahuje P % prvků, které jsou menší nebo rovny hodnotě tohoto P-kvantilu, a $(1-P)$ % prvků, které jsou větší nebo rovny hodnotě tohoto P-kvantilu.

- Mezi kvantily patří:
  - **Kvartily**
  - **Decily**
  - **Centily** (viz dále)

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Charakteristiky náhodných proměnných

- **Kvartily** ≡ kvantily, které dělí uspořádané prvky na čtyři části, přičemž každá obsahuje 25 % prvků. Kvartily jsou celkem tři:
  - **Dolní kvartil** ≡ odděluje čtvrtinu nejmenších prvků $\tilde{x}_{0.25}$
  - **Medián** ≡ rozděluje výběr prvků na dvě stejně početné části $\tilde{x}_{0.5}$

    Pokud seřadíme hodnoty náhodných pokusů (pozorování) vzestupně/sestupně, medián je hodnota uprostřed vzniklé řady. Je-li počet náhodných pokusů (pozorování) sudý, je medián (aritmetický) průměr dvou prostředních hodnot:

    $$\tilde{x}_{0.5} = \frac{X_{(n+1)/2} - X_{(n-1)/2}}{2}$$

    - **Horní kvartil** ≡ odděluje 75 % menších hodnot od zbylých 25 % vyšších $\tilde{x}_{0.75}$
- **Decily** ≡ kvantily, které dělí uspořádané prvky na deset stejně početných podmnožin
- **Centily** ≡ kvantily, které dělí uspořádané prvky na sto stejně početných podmnožin

  $\tilde{x}_{10}, \tilde{x}_{20}, \dots, \tilde{x}_{90}$ …první, druhý, … decil $\tilde{x}_{1}, \tilde{x}_{2}, \dots, \tilde{x}_{99}$ …první, druhý, … centil

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Charakteristiky náhodných proměnných

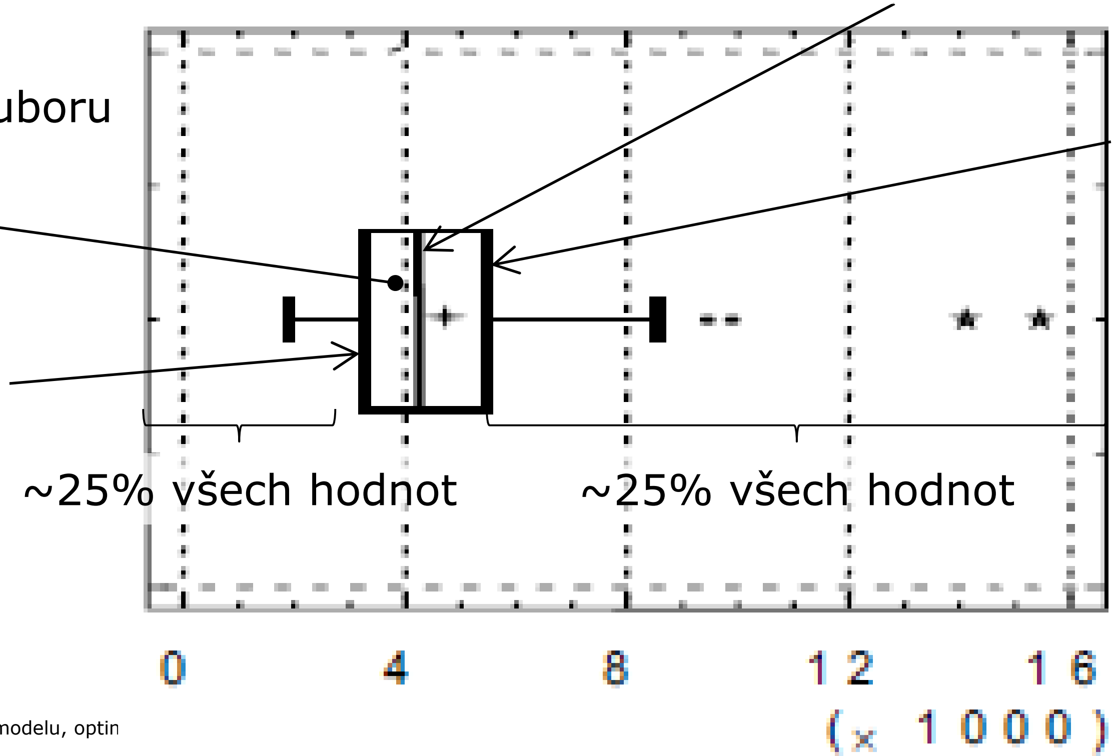
› **Modus** ≡ je lokální maximum na hustotě pravděpodobnosti náhodné proměnné X: $\hat{x}$

~ nejčastější hodnota náhodné proměnné X: $\hat{X}$

(hodnota, která se v rámci série náhodných pokusů (pozorování) vyskytuje nejčastěji)

› **Krabicový diagram:**

střední část statistického souboru (~50% všech hodhot)



9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Charakteristiky náhodných proměnných

- **Aritmetický průměr** ≡ první statistický moment m_1

$\bar{x} = \frac{\sum_{n}^{i=1} x_i}{n}$ …nejvěrohodnější odhad střední hodnoty $E(X)$ normálního rozdělení

- **Vážený průměr** je vhodné použít, nemají-li všechna pozorování (výsledky náhodného pokusu) stejnou důležitost (tj. stejnou „váhu"):

$\bar{x} = \frac{\sum_{i=1}^{n} w_i \cdot x_i}{\sum_{i=1}^{n} w_i}$ kde $\mathrm{X} = \{x_1, \dots, x_n\}$ jsou pozorování a $\mathrm{W} = \{w_1, \dots, w_n\}$ jejich váhy

  - **Geometrický průměr** je zvláštní případ váženého průměru, kde váhy všech prvků jsou shodné: $\overline{x_G} = \sqrt[n]{x_1 \cdot x_2 \cdots x_n} = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}$
  - **Harmonický průměr** je využíván pro znaky popisující rychlost děje apod.:

$$\overline{x_H} = \frac{1}{\frac{1}{n} \cdot \sum_{i=1}^{n} \frac{1}{x_i}}$$

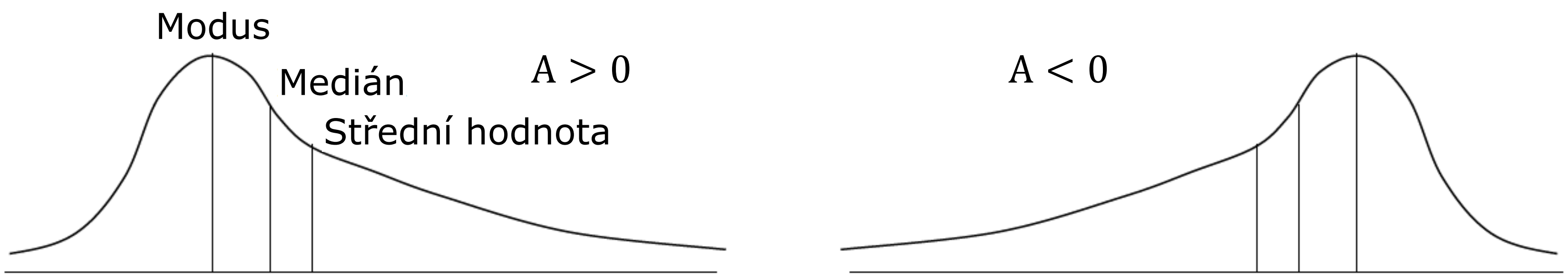
9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Charakteristiky náhodných proměnných

› **Šikmost** – měří, nakolik je rozdělení symetrické podle střední hodnoty náhodné proměnné

› **Koeficient šikmosti (asymetrie):**

$$A = \frac{\frac{1}{n}\cdot\sum_{i=1}^{n}(x_i-\bar{x})^3}{s^3}\,[-]$$



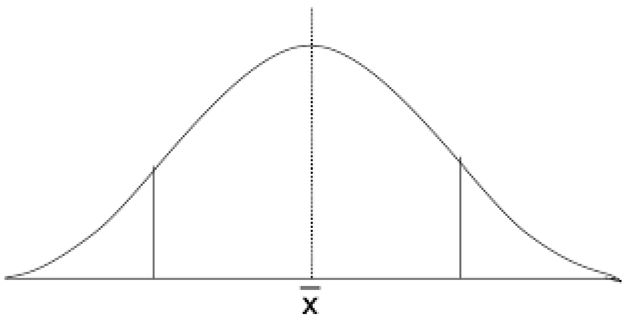
9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

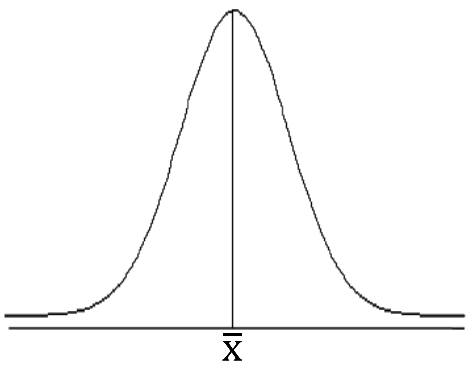
Charakteristiky náhodných proměnných

- **Špičatost** – měří, nakolik je graf hustoty pravděpodobnosti „špičatý"

- **Koeficient špičatosti (excesu):**

$$\mathrm{E} = \frac{\frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \bar{x})^4}{s^4} - 3[-]$$





9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Typy rozdělení pravděpodobnosti náhodných proměnných

- Je třeba rozlišit rozdělení pravděpodobnosti
  - **Diskrétních náhodných proměnných (DNP)**
    - Bernoulliho rozdělení
    - Binomické rozdělení
    - Geometrické rozdělení
    - Poissonovo rozdělení
  - **Spojitých náhodných proměnných (SNP)**
    - Rovnoměrné rozdělení
    - Exponenciální rozdělení
    - Erlangovo rozdělení
    - Normální rozdělení
    - Weibullovo rozdělení
    - Empirická rozdělení

Nejčastěji využívaná rozdělení v kontextu simulačních studií

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Typy rozdělení pravděpodobnosti diskrétních náhodných proměnných

- **Bernoulliho rozdělení** – souvisí s tzv. **Bernoulliho pokusem:** (viz dále)

  (Jacob Bernoulli, 1654-1705)

- **Využití:**
  - Počet náhodných jevů (výhra, opotřebení nástroje, …) v rámci provádění náhodných pokusů (zápas, výrobní operace, …) při známé pravděpodobnosti tohoto náhodného jevu (šance na výhru, pohotovost nástroje, …)

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Typy rozdělení pravděpodobnosti diskrétních náhodných proměnných

- **Bernoulliho rozdělení** – souvisí s tzv. **Bernoulliho pokusem:**

(Jacob Bernoulli, 1654-1705)

  - Uvažujme jediný j-tý hod mincí a výsledkům přiřaďme tyto hodnoty:

    $X_j = 1$ pokud je výsledkem j-tého hodu (např.) panna

    $X_j = 0$ pokud je výsledkem j-tého hodu orel

  - Pří prvním náhodném pokusu (hodu) platí pro pravděpodobnostní funkci

$$p_j(x_j) = p(x_j) = \begin{cases} p, & x_j = 1, j = 1, 2, \dots, n \\ 1 - p = q, & x_j = 0, j = 1, 2, \dots, n \\ 0 & \mathit{ostatní\ případy} \end{cases}$$

  - **Příklad:** n – počet pokusů, $p = 0{,}5$ – pravděpodobnost úspěchu každého pokusu,

    X – počet úspěšných pokusů

$$P(X = K) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$$ ...pravděpodobnost K úspěšných pokusů

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Typy rozdělení pravděpodobnosti diskrétních náhodných proměnných

- **Binomické rozdělení** ≡ zvláštní případ Bernoulliho rozdělení, kde N – počet pokusů $N > 1$ (viz předchozí případ)

- **Využití:**
  - Náhodný výběr více prvků (polotovarů, kusů pečiva,…), které vykazují s určitou pravděpodobnosti jistou vlastnost (vada, plíseň, …), kdy požadujeme zjistit pravděpodobnost výskytu zadaného počtu prvků ve výběrovém soboru

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Typy rozdělení pravděpodobnosti diskrétních náhodných proměnných

- **Binomické rozdělení** ≡ zvláštní případ Bernoulliho rozdělení, kde N – počet pokusů $N > 1$ (viz předchozí případ)
- **Příklad:** produkce vyrábí polotovary s pravděpodobností neshodného dílu 2%. Denně je kontrolován náhodný vzorek 50 polotovarů. Obsahuje-li více jak 2 neshodné díly, je produkce zastavena. Jaká ke pravděpodobnost zastavení produkce?
- **Řešení:**

$\mathrm{n} = 50$ – počet kontrolovaných dílů, $p = 0{,}02$ – pravděpodobnost neshodného dílu,

$X = 2$ – počet neshodných dílů, při kterém je zastavena produkce

$P(X > 2) = 1 - P(X \leq 2)$, kde

$$P(X \leq 2) = \binom{50}{2} \cdot 0{,}02^2 \cdot (1 - 0{,}02)^{50-2} = 0{,}92$$

Tedy $P(X > 2)$ = 0,08 …pravděpodobnost zastavení produkce je 8 %.

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

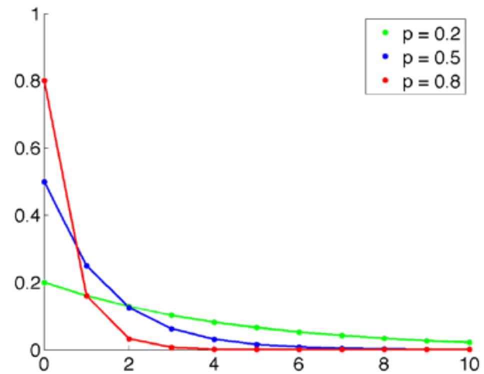
Typy rozdělení pravděpodobnosti diskrétních náhodných proměnných

- **Geometrické rozdělení** ≡ souvisí opět s Bernoulliho pokusem, kdy zde náhodná proměnná X označuje **počet pokusů nutných pro dosažení prvního úspěšného výsledku**. Pravděpodobnostní funkce X je dána:

$$p(x) = q^{x-1} \cdot p \qquad \text{kde } x = 1, 2, \dots$$

p – pravděpodobnost úspěchu

q – pradvěpodobnost neúspěchu $\mathrm{q} = 1 - \mathrm{p}$



- **Příklad:** 50% výrobků je při inspekci vyřazeno. Jaká je pravděpodobnost, že první vyhovující výrobek je třetí v pořadí na inspekci?
- **Řešení:**

$$q = 0{,}4 \qquad p = 0{,}6 \qquad \rightarrow \qquad P(3) = 0{,}4^{(3-1)} \cdot 0{,}6 = 0{,}096$$

Tedy asi v 10% případů je první vyhovujícím výrobkem ten, který je na inspekci zkoušen jako třetí v pořadí.

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Typy rozdělení pravděpodobnosti diskrétních náhodných proměnných

- **Poissonovo rozdělení** ≡ udává pravděpodobnost, že v určeném časovém intervalu nastane zadaný počet událostí

$$p(x) = \begin{cases} \frac{e^{-\lambda} \cdot \lambda^x}{x!}, & x \in \{0,1,\dots,t\} \\ 0 & ostatní\ příp. \end{cases}$$

kde $\lambda > 0$ – střední doba mezi příchody

- **Příklad:** Poruchy stroje vykazují co do počtu prostojů za daný časový interval Poissonovo rozdělení o střední hodnotě 2 poruchy za hodinu. Jaká ke pravděpodobnost 3 poruch v následující hodině?
- **Řešení:**

$$\mathrm{p}(3) = \frac{e^{-2} \cdot 2^3}{3!} \approx \frac{0{,}135 \cdot 8}{6} = 0{,}18$$

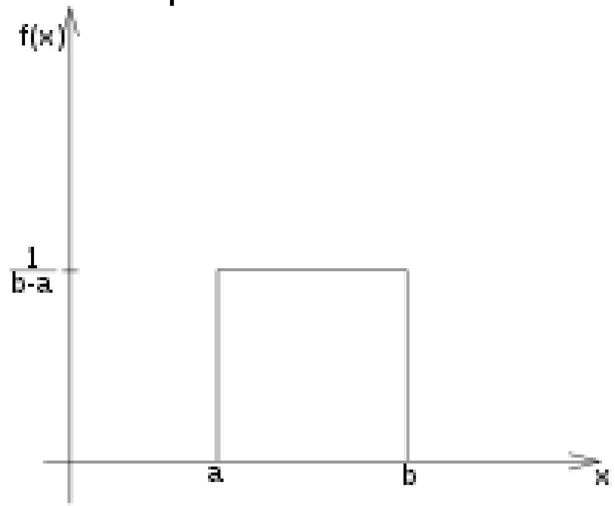
Jaká bude pravděpodobnost dvou a více prostojů během hodinového intervalu?

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Typy rozdělení pravděpodobnosti spojitých náhodných proměnných

- **Rovnoměrné rozdělení** ≡ spojitá náhodná proměnná X je rovnoměrně rozdělena v intervalu (a, b), pokud pro její hustotu pravděpodobnosti platí:

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0 & ostatní\ příp. \end{cases}$$



- Pro kumulativní distribuční funkci platí:

$$\mathrm{F}(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1 & x \geq b \end{cases}$$

- **Využití:**
  - Jakékoliv náhodné pokusy, které mají stejnou pravděpodobnost nastoupení všech možných jevů (všechny výsledky jsou stejně pravděpodobné)

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Typy rozdělení pravděpodobnosti spojitých náhodných proměnných

- **Rovnoměrné rozdělení** ≡ spojitá náhodná proměnná X je rovnoměrně rozdělena v intervalu (a, b), pokud pro její hustotu pravděpodobnosti platí:

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0 & ostatní\ příp. \end{cases}$$

- Pro kumulativní distribuční funkci platí:

$$\mathrm{F}(x) = \begin{Bmatrix} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1 & x \geq b \end{Bmatrix}$$

- **Příklad:** Autobusy přijíždí na zastávku v době od 6:55 do 17:55 každých 15 min. Cestující přichází každý den na zastávku mezi 7:00 a 7:30 (rovnoměrné rozdělení). Jaká je pravděpodobnost, že čeká více než 5 min?

- **Řešení:** $\mathrm{P}(\check{c}) = \mathrm{P}(5 < x < 15) + P(20 < x < 30) = F(15) - F(5) + F(30) - F(20) = \frac{15}{30} - \frac{5}{30} + 1 - \frac{20}{30} = \frac{2}{3}$

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

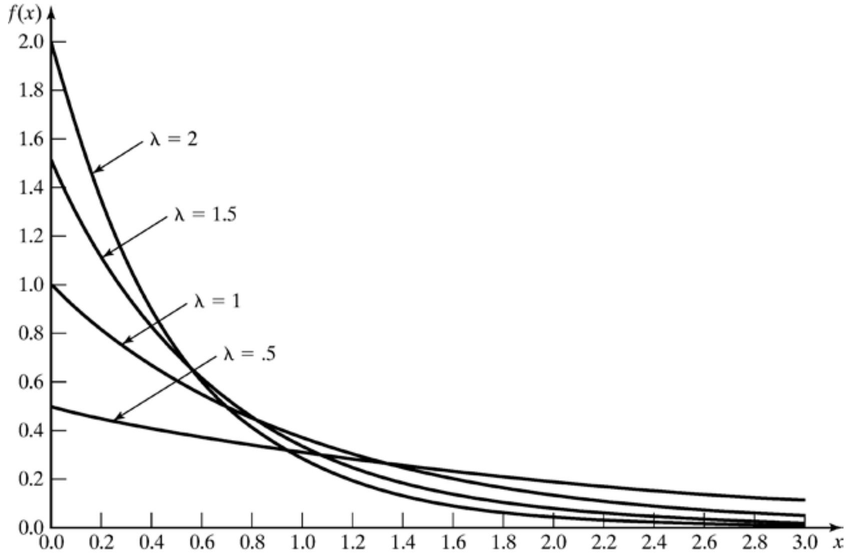
Typy rozdělení pravděpodobnosti spojitých náhodných proměnných

- **Exponenciální rozdělení** ≡ spojitá náhodná proměnná X je exponenciálně rozdělena, pokud pro její hustotu pravděpodobnosti platí:

$$f_x(u) = \lambda \cdot e^{-\lambda \cdot u} \qquad \text{kde } u \geq 0$$

- Pro kumulativní distribuční funkci platí:

$$\mathrm{F}(x) = \int_0^x f(y)dy = \int_0^x \lambda \cdot e^{-\lambda \cdot y} dy = 1 - e^{-\lambda \cdot y}$$



- **Využití:**
  - modelování doby mezi nezávislými událostmi, jako jsou příchody, nebo prostoje (poruchy), pracovní časy s vysokou mírou variability

- S časem se hodnota pravděpodobnosti nemění – rozdělení „nemá paměť":

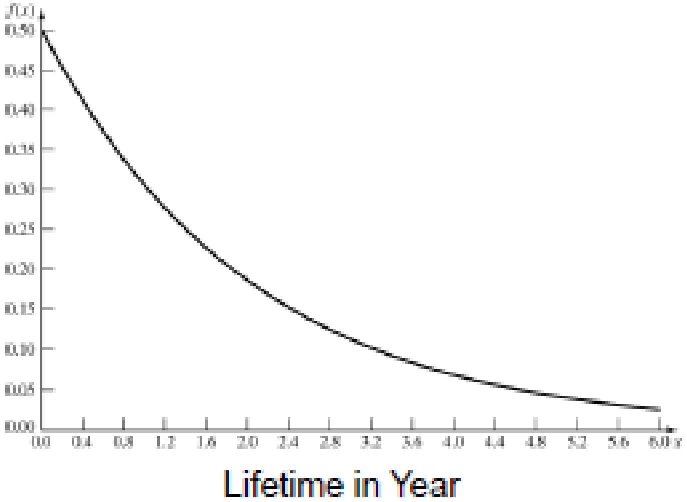
$$P\{X > s + t | X > t\} = P\{X > s\} \qquad \text{kde } t, s > 0$$

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Typy rozdělení pravděpodobnosti spojitých náhodných proměnných

› **Exponenciální rozdělení** ≡ (viz výše)

› **Příklad:** životnost žárovky je dána hodnotou X, která má exponenciální rozdělení pravděpodobnosti se střední hodnotou 2 roky. Graf její hustoty pravděpodobnosti vypadá takto:



Jaká je pravděpodobnost, že životnost žárovky bude mezi 2 a 3 roky?

› **Řešení:** $P(2 \leq X \leq 3) = \frac{1}{2} \cdot \int_2^3 e^{-\frac{x}{2}} dx = 0{,}14$

Jaká je pravděpodobnost, že žárovka bude sloužit další rok, pokud již je v provozu 2,5 roku?

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Typy rozdělení pravděpodobnosti spojitých náhodných proměnných

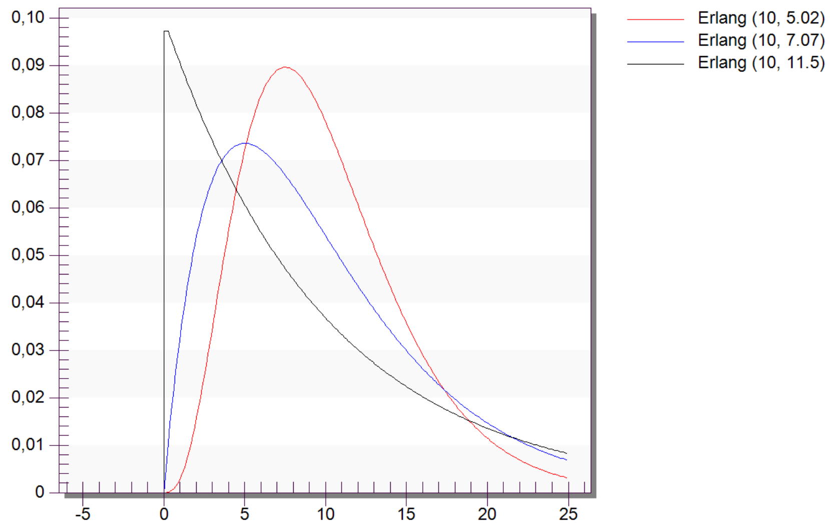
- **Erlangovo rozdělení** ≡ spojitá náhodná proměnná X má Erlangovo rozdělení pravděpodobnosti, pokud pro její hustotu pravděpodobnosti platí:

$$f(x;k;\lambda) = \frac{\lambda^k \cdot x^{k-1} \cdot e^{-\lambda \cdot x}}{(k-1)!}$$

pro $x, \lambda > 0$ k-tvar, λ-intenzita

- Pro kumulativní distribuční funkci platí:

$$\mathrm{F}(x;k;\lambda) = 1 - \sum_{n=0}^{k-1} e^{-\lambda \cdot x} \cdot (\lambda \cdot x)^n / n!$$



- Je rozdělení více nezávislých náhodných proměnných, které mají exponenciální rozdělení stejných parametrů.

- **Využití:**
  - modelování událostí, které nastávají vzájemně nezávisle se stanovenou četností. Doba mezi událostmi je modelována Erlangovým rozdělením.

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Typy rozdělení pravděpodobnosti spojitých náhodných proměnných

- **Erlangovo rozdělení** ≡ (viz výše)

- **Příklad:** Svařovací robot má dvě hlavice. V případě poruchy první hlavice přepne proud do druhé. Živostnost hlavic je exponenciálně rozdělena, průměrná hodnota je 1000 hod. Jaká je pravděpodobnost, že robot bude i po 90 dnech svařovat?

- **Řešení:** jde o výpočet tzv. spolehlivosti $\mathrm{R}(x)$, kde $\mathrm{R}(x) = 1 - F(x)$, kde $\mathrm{F}(x)$ je pravděpodobnost poruchy. V tomto případě $\mathrm{k} = 2$ a $\lambda = 1/1000$.
  $\mathrm{F}(90 \cdot 24) = 0{,}636$, tedy
  s pravděpodobností 36 % bude systém i po 90 dnech stále pracovat.

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Typy rozdělení pravděpodobnosti spojitých náhodných proměnných

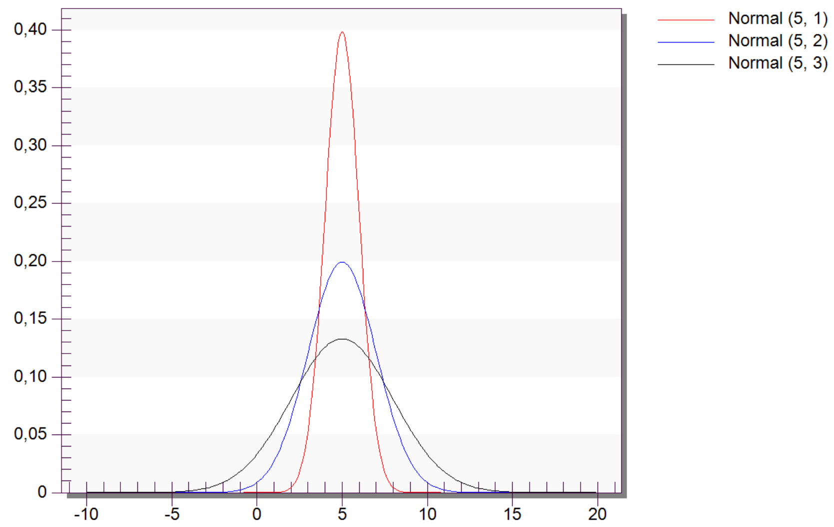
- **Normální rozdělení** ≡ spojitá náhodná proměnná X má normální rozdělení pravděpodobnosti, pokud pro její hustotu pravděpodobnosti platí:

$$f_x(u) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-(u-\mu)^2/2 \cdot \sigma^2}$$

kde μ-střední hodnota σ-standardní odchylka

- Pro kumulativní distribuční funkci platí:

$$F(X) = \int_{-\infty}^{Z} \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{t^2}{2}} dt \cdot \left(\frac{x-\mu}{\sigma}\right)$$



- **Využití:**
  - Stanovení doby operací automatických zařízení
  - Stanovení trvání stabilních procesů

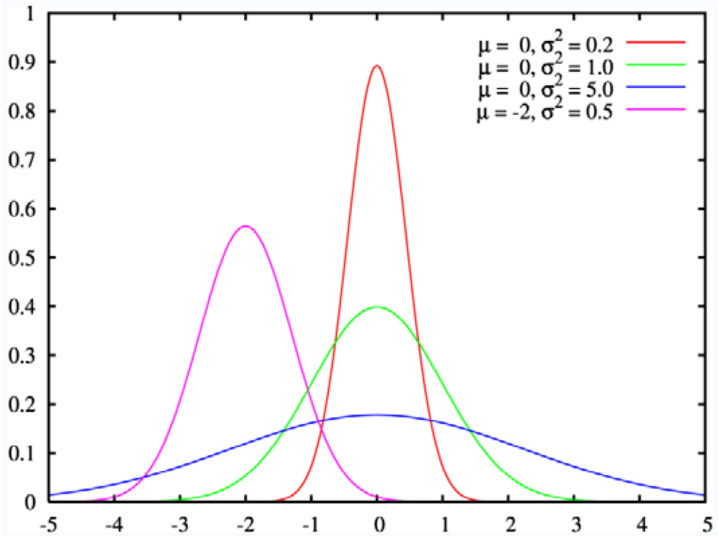
9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Typy rozdělení pravděpodobnosti spojitých náhodných proměnných

- **Normální rozdělení** ≡ spojitá náhodná proměnná X má normální rozdělení pravděpodobnosti, pokud pro její hustotu pravděpodobnosti platí:

$$f_x(u) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-(u-\mu)^2/2 \cdot \sigma^2}$$

kde μ-střední hodnota σ-standardní odchylka



- Pro kumulativní distribuční funkci platí:

$$F(X) = \int_{-\infty}^{z} \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{t^2}{2}} dt \cdot \left(\frac{x-\mu}{\sigma}\right)$$

- **Příklad:** čas potřebný k naložení lodi X má normální rozdělení o střední hodnotě $\mu = 12\ hod$ a standardní rozptylu $\sigma^2 = 4$. Jaká je pravděpodobnost, že loď bude naložena za méně jak 10 hodin?

- **Řešení:** $\mathrm{F}(10) = \int_{-\infty}^{z} \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{t^2}{2}} dt \cdot \left(\frac{10-12}{2}\right) = 0{,}1587$

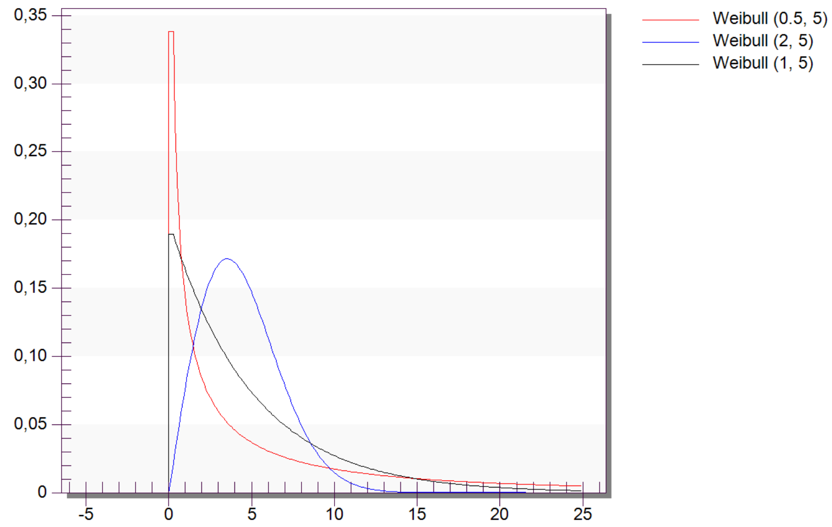
9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Typy rozdělení pravděpodobnosti spojitých náhodných proměnných

- **Weibullovo rozdělení** ≡ spojitá náhodná proměnná X má Weibullovo rozdělení pravděpodobnosti, pokud pro její hustotu pravděpodobnosti platí:

$$\mathrm{f}(x) = \frac{\beta}{\delta} \cdot \left(\frac{x}{\delta}\right)^{\beta-1} exp\left[-\left(\frac{x}{\delta}\right)^{\beta}\right]$$

kde $\mathrm{x} > 0$, δ-parametr rozsahu, β-parametr tvaru



- Pro kumulativní distribuční funkci platí:

$$F(X) = 1 - e^{-\left(\frac{x}{\delta}\right)^{\beta}}$$

- Pro $\beta = 1$ se jedná o exponenciální rozdělení
- **Využití:**
  - modelování doby do poruchy
  - dle parametrů lze modelovat systém, kde s časem počet prostojů roste $(\beta < 1)$, klesá $(\beta > 1)$, je konstantní $(\beta = 1)$

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Typy rozdělení pravděpodobnosti spojitých náhodných proměnných

- **Weibullovo rozdělení** ≡ (viz výše)

- **Příklad:** čas do poruchy elektrického zařízení vykazuje Weibullovo rozdělení pravděpodobnosti s parametry $\beta = 1/3$ a $\delta = 200\ hod$. Jaká je střední doba do poruchy zařízení? Jaká je pravděpodobnost, že k poruše dojde dříve než za 2000 hodin?

- **Řešení:**

$$E(X) = 200\neg(3+1) = 200 \cdot (3!) = 1200\ hod$$

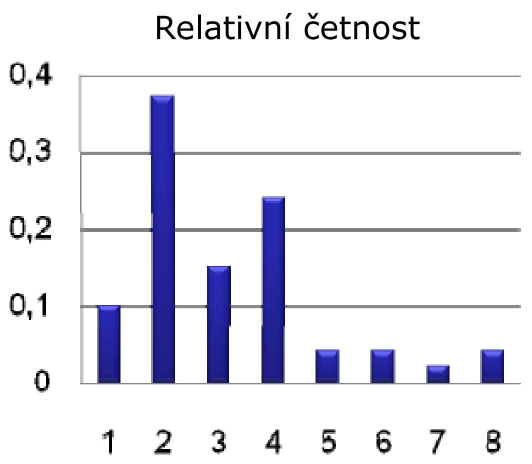
$$F(2000) = 1 - \exp\left[-\left(\frac{2000}{200}\right)^{\frac{1}{3}}\right] = 0{,}884$$

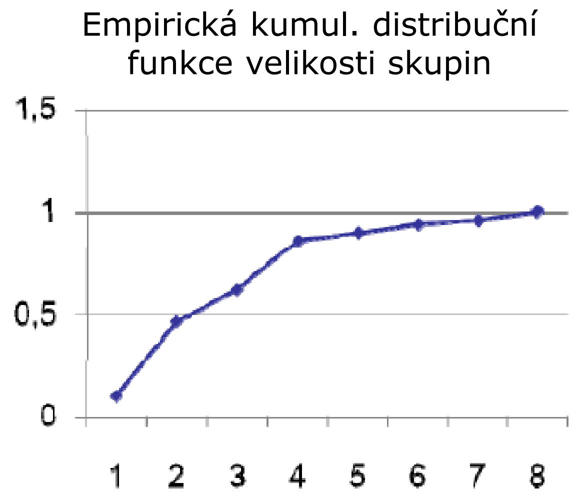
9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Typy rozdělení pravděpodobnosti spojitých náhodných proměnných

- **Empirické rozdělení** ≡ rozdělení, jehož parametry nejsou známy a místo nich jsou využity pozorované hodnoty (výsledky náhodného pokusu) v rámci náhodného výběru.
- **Využití:**
  - pokud není možné nebo není nutné stanovit náhodnou veličinu parametrického rozdělení pravděpodobnosti.

Velikost skupiny	Četnost	Relativní četnost	Kumul. rel. četnost
1	30	0,10	0,10
2	110	0,37	0,47
3	45	0,15	0,62
4	71	0,24	0,86
5	12	0,04	0,90
6	13	0,04	0,94
7	7	0,02	0,96
8	12	0,04	1,00





9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Testování hypotéz

- Pomocí tzv. **statistického testování** ověřujeme platnost tzv. **nulové hypotézy** H_0, kterou postavíme proti tzv. **alternativní hypotéze** H_A. Z pohledu výrokové logiky jsou obě hypotézy výroky s opačnou hodnotou pravdivosti, tj. vždy platí buď H_0 nebo platí H_A.

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Testování hypotéz

› **Postup testování statistických hypotéz:**

1. Formulace nulové hypotézy H_0 a alternativní hypotézy H_A
2. Volba hladiny významnosti α
3. Volba testovací statistiky (postupu, podle kterého test provedeme) – např. t
4. Určení kritického oboru testovací statistiky
5. Vyčíslení testovací statistiky a jejích kvantilů
6. Rozhodnutí, zda
   a) Zamítnout nulovou hypotézu H_0 a přijmout alternativní hypotézu H_A - jestliže testovací statistika padne do kritického oboru
   b) Nezamítnout nulovou hypotézu H_0 - testovací statistika nepadne do kritického oboru

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Testování hypotéz

› **Postup testování statistických hypotéz:**

1. Formulace nulové hypotézy H_0 a alternativní hypotézy H_A

› Hypotézy mohou být
  › **Parametrické** ...tvrzení o parametrech pozorované náhodné veličiny X
  › **Neparametrické** ...tvrzení o kvalitativních vlastnostech náhodné veličiny X

› V případě parametrické nulová hypotéza H_0 může být tato hypotéza:
  › **Jednoduchá** ...uvažujeme jedinou hypotetickou hodnotu $\vartheta = \vartheta_0$
  › **Složená** ...uvažujeme více hypotetických hodnot, např. $\vartheta \neq \vartheta_0$

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Testování hypotéz

› **Postup testování statistických hypotéz:**

1. Formulace nulové hypotézy H_0 a alternativní hypotézy H_A
2. Volba hladiny významnosti

› **Hladina významnosti α** vyjadřuje pravděpodobnost **chyby I. druhu** – zamítáme nulovou hypotézu H_0, přestože platí, neboť testovací statistika nabude hodnotu z **tzv. kritického oboru W_α** (což je kritérium pro rozhodnutí o ne/platnosti H_0)

› Hladinu významnosti $\alpha \in (0; 1)$ volíme obvykle $\alpha = 0{,}05$, nebo $\alpha = 0{,}01$

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Testování hypotéz

› **Postup testování statistických hypotéz:**

1. Formulace nulové hypotézy H_0 a alternativní hypotézy H_A
2. Volba hladiny významnosti
3. Volba testovací statistiky (postupu, podle kterého test provedeme) – např. t

› Lze vycházet ze statistických tabulek a vztahů

› Volba závisí také na typu prováděného testu – dle posuzované hypotézy (viz dále)

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Testování hypotéz

- **Postup testování statistických hypotéz:**

1. Formulace nulové hypotézy H_0 a alternativní hypotézy H_A
2. Volba hladiny významnosti
3. Volba testovací statistiky (postupu, podle kterého test provedeme) – např. t
4. Určení kritického oboru testovací statistiky

- Obor hodnot testovací statistiky (testového kritéria) se za předpokladu platnosti nulové hypotézy $H_0: \vartheta = \vartheta_0$ rozdělí na dvě disjunktní množiny
  - Kritický obor W_0
  - Doplněk kritického oboru $\overline{W_0}$

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Testování hypotéz

› **Postup testování statistických hypotéz:**

1. Formulace nulové hypotézy H_0 a alternativní hypotézy H_A
2. Volba hladiny významnosti
3. Volba testovací statistiky (postupu, podle kterého test provedeme) – např. t
4. Určení kritického oboru testovací statistiky
5. Vyčíslení testovací statistiky a jejích kvantilů

› Lze vycházet ze statistických tabulek a vztahů

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Testování hypotéz

› **Postup testování statistických hypotéz:**

1. Formulace nulové hypotézy H_0 a alternativní hypotézy H_A
2. Volba hladiny významnosti
3. Volba testovací statistiky (postupu, podle kterého test provedeme) – např. t
4. Určení kritického oboru testovací statistiky
5. Vyčíslení testovací statistiky a jejích kvantilů
6. Rozhodnutí, zda
   a) Zamítnout nulovou hypotézu H_0 a přijmout alternativní hypotézu H_A - jestliže testovací statistika padne do kritického oboru $\quad t \in W_0$
   b) Nezamítnout nulovou hypotézu H_0 - testovací statistika nepadne do kritického oboru

$$t \in \overline{W}_0$$

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Testování hypotéz

- Výsledek testování statistických hypotéz:
  - **Zamítnutí nulové hypotézy H_0 neznamená, že H_0 neplatí!** Zamítnutí říká, že nulovou hypotézu nepředpokládáme a naopak předpokládáme platnost alternativní hypotézy H_A.
  - **Nezamítnutí nulové hypotézy H_0 neznamená, že H_0 platí!** Výsledek testu pouze neukázal dostatečnou neshodu mezi testovanou skutečností a hypotézou.

- Dvě možné chyby při testování statistických hypotéz:
  - **Chyba I. druhu:** Testovací statistika padne mimo obor pro přijetí H_0 - H_0 tedy zamítáme, avšak ve skutečnosti H_0 platí. Pravděpodobnost této chyby je α a nazývá se **hladina významnosti**.
  - **Chyba II. Druhu:** Testovací statistika padne do oboru pro přijetí H_0 - H_0 tedy nezamítáme, avšak ve skutečnosti platí alternativní hypotéza H_A. Pravděpodobnost této chyby je β. Platí přitom $\alpha + \beta = 1$.

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Testování hypotéz

› Dvě možné chyby při testování statistických hypotéz z pohledu modelování a simulací:

Terminologie statistiky	Terminologie modelování	Související riziko
Typ I: zamítnutí H_0, přestože H_0 platí	Zamítnutí validního modelu	α
Typ II: nezamítnutí H_0, přestože platí H_A	Použití nevalidního modelu	β

› Chybě I druhu lze předejít volbou nízké hladiny významnosti α

› Chybě II druhu lze předejít určením kritické hodnoty rozdílu (skutečnost-model) a provedením „dostatečného“ množství pozorování (simulačních experimentů) n

› Protože $\alpha + \beta = 1$, snížením rizika α se zvyšuje riziko β a naopak…

› **Chyba III. druhu** – řešení nesprávného problému

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

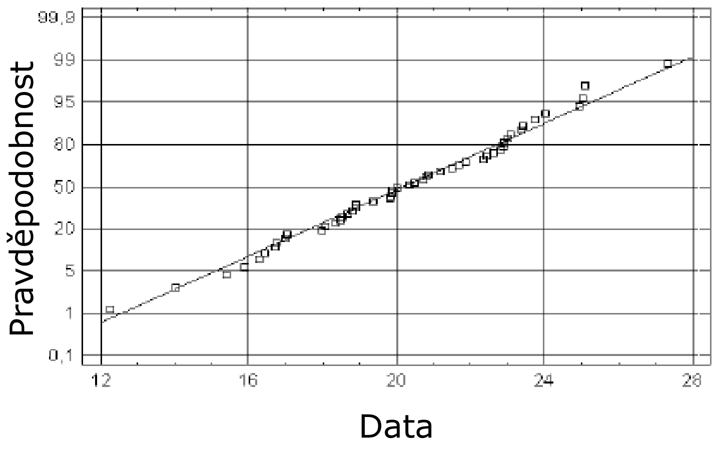
Testování hypotéz

- Nejběžnější statistické testy hypotéz o normálním rozdělení:
  - **t-test** neboli **Studentův test pro jeden výběr**
    - H_0: $\mu = \mu_0$ při neznámém rozptylu σ^2
  - **Pearsonův test**
    - H_0: $\sigma^2 = \sigma_0^2$
  - **t-test** neboli **Studentův test pro dva výběry**
    - H_0: $\mu(X - Y) = 0$ pro dvojice hodnot náhodných výběrů (X, Y)
  - **t-test** neboli **Studentův test pro dva výběry při stejných rozptylech**
    - H: $\mu(X) - \mu(Y) = \mu_0$ při neznámých rozptylech $\sigma^2(X) = \sigma^2(Y)$
  - **t-test** neboli **Studentův test pro dva výběry při různých rozptylech**
    - H: $\mu(X) - \mu(Y) = \mu_0$ při neznámých rozptylech $\sigma^2(X) \neq \sigma^2(Y)$
  - **F-test** neboli **Fischerův test**
    - H_0: $\sigma^2(X) = \sigma^2(Y)$

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Testování hypotéz

- Nejběžnější statistické testy hypotéz o rozdělení (testy dobré shody):
  - **Grafická metoda pravděpodobnostního papíru**
    - Graf libovolné distribuční funkce
    - Je-li statistický soubor realizací náhodného výběru ze základního souboru s rozdělením pravděpodobnosti pro daný pravděpodobnostní papír, leží body na přímce



  - **Test chí-kvadrát (Pearsonův test) o rozdělení**
    - Test, zda pozorovaná náhodná veličina X má distribuční funkci $F(X)$

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Testování hypotéz – nejběžnější statistické testy hypotéz v kontextu modelování a simulací

› **Turingův test** – popsán Alanem Turingem r. 1950 – použit v důsledku nemožnosti aplikace klasických statistických testů – namísto nich využit lidský úsudek, např.:

1. Přeložíme 10 výstupů systému (výstupní hodnoty KPI) – 5 z nich jsou z reálného systému a 5 jsou ze simulačního modelu
2. Pokud nezávislý posuzovatel identifikuje správně „podstatný" počet výstupů ze simulačního modelu, konzultace na vylepšení modelu
3. Pokud nezávislý posuzovatel naopak není schopen rozeznat výstupy z reálného systému a modelu, předpokládejme, že simulační model je validní.

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Optimalizace

- Simulační modely popisují většinou složité stochastické systémy z dynamického hlediska. Vyžadující optimalizaci vzhledem k vytížení výrobních zdrojů, nákladů, doby výroby apod.
- Optimalizační metody jsou aplikovány na jednodušší systémy, resp. za použití zjednodušených předpokladů.
- **Cílem optimalizace je nalezení kombinace parametrů zajišťující maximální/minimální hodnotu sledovaných kritérií (kriteriální funkce).**
- Kombinace simulačních modelů a optimalizačních metod ~ synergický efekt.
- V kontextu simulací a simulačních studií lze optimalizaci provádět dvojím způsobem
  - **Návrh a realizace optimalizačních experimentů.**
  - **Využití heuristických optimalizačních algoritmů.**

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Optimalizace

- **Návrh a realizace optimalizačních experimentů** – v těchto krocích:
  1. Postupná změna hodnot všech parametrů modelu – obvykle mnoho variant.
  2. Provedení simulačních experimentů se všemi variantami.
  3. Výběr nejlepšího řešení vzhledem k hodnotám sledovaných kritérií pro jednotlivé varianty.
  - V případě **malého počtu variant** lze ověřit všechny.
  - V případě **velkého počtu variant**:
    - Hledáme suboptimální (dostatečně dobré) řešení.
    - Převádíme vícerozměrnou optimalizaci o N parametrech na sled N jednorozměrných optimalizací:
      - Fixace hodnot $N-1$ parametrů, měněn pouze parametr X_1 - nalezeno suboptimum $\widetilde{X_1}$.
      - Postupně s nalezeými dalšími suboptimy získáme řešení s parametry $\widetilde{X_1}, \widetilde{X_2}, \dots, \widetilde{X_N}$.

9. přednáška – Stochastické procesy a jejich zohlednění v modelu, optimalizace na bázi simulace

Otázky z dané problematiky

- Co je to **náhodné číslo** a co je **náhodný pokus**? – jak tyto pojmy souvisí s problematikou diskrétních simulačních modelů?
- Jak chápete rozdíl mezi **náhodnou proměnnou**, **náhodným výběrem** a **statistickým souborem**?
- Jaké jsou **typy náhodných proměnných**? Co popisují jejich **rozdělení pravděpodobnosti**?
- Jaké znáte **typy rozdělení pravděpodobnosti**?
- Proč se provádí (v kontextu vytváření a práce se simulačními modely) **testování statistických hypotéz**?
- Jaký je **princip testu hypotézy**?
- Jaké jsou **možné chyby při testování statistických hypotéz**? Jak je lze omezit?
- Jaký je **princip optimalizace**?

Děkuji.