PROCESOS DE DIFUSION

Tesis de Licenciatura

Presenta:

Jesús Antonio Astorga Moreno

Director de Tesis:

Dr. Eugenio Pacelli Balanzario Gutiérrez

Indice

Intodución 1

Capítulo 1

El Teorema de DeMoivre-Laplace

§1. Introducción 4

§2. Deducción Heurística 5

§3. El Teorema de DeMoivre-Laplace con Término de Error 8

Capítulo 2

Ecuaciones de Kolmogorov

§1. Introducción 22

§2. Movimiento Browniano 24

§3. Movimiento Browniano con Arrastre 26

§4. Ecuaciones de Kolmogorov 29

Capítulo 3

Ejemplo de Procesos de Difusión

§1. Difusión de un Contaminante 38

§2. Proceso de Ornstein-Uhlenbeck 42

Glosario 47

Conclusiones 48

Bibliografía 49

Introducción

Esta tesis es una primera introducción básica al tema de los procesos de difusión. El objetivo es el de poner en manos de los participantes en la redacción de la misma, algunas de las herramientas sobre procesos de difusión que les permitan abordar problemas prácticos tanto en problemas de explotación óptima de recursos naturales así como en problemas que se presentan en los mercados financieros. Después de hacer una descripción del contenido de la tesis, indicaremos la conexión entre los procesos de difusión con algunos problemas de explotación de recursos naturales y con los problemas financieros.

En el primer capítulo de la tesis se estudia el teorema de DeMoivre-Laplace, primero desde un punto de vista intuitivo y después mediante una prueba rigurosa del mismo. La deducción intuitiva del teorema de DeMoivre-Laplace que se presenta en este trabajo, y que se debe a M.A. Proschan [7], se puede presentar en los cursos introductorios de estadística más básicos. Sin embargo, debe notarse que Proschan no nos ofrece una demostración matematicamente correcta del teorema de DeMoivre-Laplace. Para una demostración rigurosa del teorema de DeMoivre-Laplace se consultó el texto de E. Lesigne [5]. En este trabajo, presentamos el teorema de DeMoivre-Laplace con una estimación de la rapidez de convergencia hacia la distribución normal. Puesto que Lesigne no obtiene esta estimación de la rapidez de convergencia, entonces los cálculos reportados en este trabajo sirven para complementar el trabajo de Lesigne [5].

En el segundo capítulo se aplica el teorema de DeMoivre-Laplace para obtener la ecuación de difusión más simple. Se trata de la ecuación de conducción de calor. También se obtiene a partir del teorema de DeMoivre-Laplace la ecuación de difusión con coeficiente de arrastre. Cuando el coeficiente de arrastre es igual a cero, la ecuación de difusión con coeficiente de arrastre se reduce a la ecuación de calor. La deducción de ambas ecuaciones se logra al considerar las caminatas aleatorias de partículas suspendidas en un medio líquido. En el capítulo XIV del libro de W. Feller [3] se puede encontrar una deducción de la ecuación de difusión con coeficiente de arrastre similar a la deducción que se presenta en esta tesis. Sin embargo Feller da pocos detalles de los cálculos que llevan a la ecuación de difusión con coeficiente de arrastre y en esta tesis se trata de ser más explícitos en la deducción de esta ecuación. Como segundo objetivo del segundo capítulo se deducen las dos ecuaciones de Kolmogorov que todo proceso de difusión debe satisfacer. La ecuación de calor y la ecuación de difusión con coeficiente de arrastre son dos casos particulares de las ecuaciones de Kolmogorov.

El objeto del tercer capítulo es dar dos ejemplos concretos de procesos de difusión. En el primer ejemplo se estudia la difusión de un contaminante en un medio líquido, para lo cual, la ecuación de calor resulta apropiada. En el segundo ejemplo se considera el proceso de Ornstein-Uhlenbeck. Para estudiar el proceso de Ornstein-Uhlenbeck se utiliza la segunda de las ecuaciones de Kolmogorov. Para los ejemplos de este capítulo se resuelven las ecuaciones de Kolmogorov; por el método de separación de variables en el primer caso (ver [2]) y por el método de las características en el segundo caso (ver [4]).

Se puede modelar la difusión de algunas poblaciones de animales mediante las ecuaciones de difusión. Al incorporar los fenómenos difusivos en los modelos de explotación óptima de algunos recursos naturales (como las pesquerías) se podrían obtener políticas de explotación mejoradas de los recursos naturales. Ver la obra de C. Clark [1] para un tratamiento de este tema.

Finalmente, el proceso de Ornstein-Uhlenbeck se ha usado para modelar, por ejemplo, la variación de las tasas de interés en los mercados financieros. De mayor importancia que el haber obtenido en esta tesis la función de densidad condicional del proceso de Ornstein-Uhlenbeck, es el haber hecho conciencia en los participantes de esta tesis que un proceso de difusión está sujeto a ser modelado de tal forma que se ajuste al fenómeno considerado en la investigación en turno. Para una introducción al tema de los procesos de difusión en los modelos financieros se puede consultar la obra de S.E. Shreve [9].

Capítulo 1

El Teorema de DeMoivre-Laplace

§1. Introducción.

Este capítulo está dedicado al teorema de DeMoivre-Laplace el cual apareció por primera vez en la segunda edición del libro *"The doctrine of chances"* escrito por el matemático francés Abraham DeMoivre y publicado en 1718. El teorema de DeMoivre-Laplace es quizá la versión más simple del Teorema Central del Límite y se puede considerar como el segundo teorema límite básico en teoría de probabilidades, sólo después de la ley de los grandes números de Bernoulli. En 1810 Pierre-Simon Laplace extendió el teorema de DeMoivre a valores arbitrarios de p distintos de 0 y 1, ya que DeMoivre sólo probó el caso especial $p = \frac{1}{2}$. Esta generalización representa el trabajo más importante en teoría de probabilidades de Laplace.

El Teorema de DeMoivre-Laplace proporciona una buena aproximación a la distribución binomial mediante la distribución normal. Esta aproximación normal a la distribución binomial tiene un gran valor teórico y práctico ya que desempeñó un papel importante en el desarrollo de la teoría de probabilidad, al conducir al primer Teorema Central del Límite. El Teorema Central del Límite no es un único teorema, sino que consiste en un conjunto de resultados acerca del comportamiento de la suma de variables aleatorias. El término "central" debido a Polya (1920), significa fundamental o de importancia central. Esto describe el rol que cumple este teorema tanto en la teoría como en la práctica. Su importancia radica en que el Teorema Central del Límite devela las razones por la cuales en muchos campos de aplicación se encuentran distribuciones normales. Dentro de la historia del Teorema Central del Límite, el trabajo de Laplace ocupa un lugar fundamental. A pesar de que él nunca enunció formalmente este resultado ni procedió con rigor en sus demostraciones, su trabajo en la teoría de la probabilidad es fundamental.

§2. Deducción Heurística.

Tanto DeMoivre como Laplace probaron que la distribución binomial con un número muy grande de ensayos se acerca a la distribución de probabilidad normal. En lo siguiente procederemos de manera heurística para obtener dicha aproximación. Este proceder de manera heurística constituye una introducción al teorema que nos compete y en el apartado §3 damos una demostración rigurosa del Teorema de DeMoivre-Laplace. El término "heurística" se debe a que estamos aproximando la distribución de una variable aleatoria discreta con la de una variable continua, pero en esta aproximación algunos pasos necesitan una justificación más adecuada de la que nosotros presentamos.

Consideremos el experimento aleatorio que consiste de n ensayos de Bernoulli repetidos, donde p es la probabilidad de éxito y $q = 1 - p$ la probabilidad de fracaso en cada ensayo individual. El término "repetido" se usa para indicar que la probabilidad de éxito es la misma ensayo tras ensayo. Sea X_n la variable aleatoria que representa el número de éxitos en los n ensayos. Notemos que dicha variable también se puede considerar como la suma de n variables aleatorias independientes con distribución Bernoulli.

Se plantea la siguiente pregunta, ¿qué podemos decir del comportamiento de dicha variable X_n cuando n tiende a infinito? La respuesta es sencilla, ya que cada uno de los ensayos tiene probabilidad p de éxito y q de fracaso, y además existen $\binom{n}{x}$ experimentos en donde se registran n éxitos y $n-x$ fracasos, entonces tenemos que

$$\mathbf{P}\{X_n = x\} = \mathbf{P}\{x \text{ éxitos y } n-x \text{ fracasos}\} = \binom{n}{x} p^x q^{n-x}. \qquad (\mathbf{1})$$

Encontramos que la variable aleatoria X_n tiene distribución binomial. Mediante $f_n(x)$ denotamos la función de masa de probabilidad de la variable X_n de modo que $f_n(x) = \mathbf{P}\{X_n = x\}$. Así entonces $f_n(x)$ es la distribución binomial descrita por la ecuación (1). Ahora bien, para cada $x = 1, \cdots, n-1$, tenemos que

$$\frac{f_n(x)}{f_n(x+1)} = \frac{\binom{n}{x} p^x q^{n-x}}{\binom{n}{x+1} p^{x+1} q^{n-x-1}} = \frac{q(x+1)}{p(n-x)}.$$

De aquí tenemos que

$$\frac{f_n(x+1)}{f_n(x)} = \frac{(n-x)p}{(x+1)q} = 1 - \frac{x - np + q}{q(x+1)}.$$

Lo que queremos saber ahora es cuándo $f_n(x)$ crece monótonamente y cuándo $f_n(x)$ decrece monótonamente. Para ésto nótese que

$$\frac{f_n(x+1)}{f_n(x)} > 1 \qquad \text{si} \qquad x < np - q$$

y

$$\frac{f_n(x+1)}{f_n(x)} < 1 \qquad \text{si} \qquad x > np - q$$

mientras que la igualdad se da cuando $x = np - q$. De este análisis podemos inferir que $f_n(x)$ tiene un máximo en el punto $x = np - q$, ya que para valores menores que éste la función crece y lo opuesto para valores mayores. De esta manera se puede conjeturar que la distribución binomial tiene cierta forma de "campana". Lo interesante a saber es cuál de todas las densidades continuas en forma de campana corresponde a la distribución binomial. Para ésto, sea z tal que $x = np + z\sqrt{npq}$ es un entero. Entonces

$$\frac{f_n(x+1)}{f_n(x)} = \frac{n-x}{x+1}\Big(\frac{p}{q}\Big) = \frac{n-np-z\sqrt{npq}}{np+z\sqrt{npq}+1}\Big(\frac{p}{q}\Big) = \frac{1-p-z\sqrt{pq/n}}{p+z\sqrt{pq/n}+1/n}\Big(\frac{p}{q}\Big).$$

Puesto que $1 = p + q$ entonces

$$\frac{f_n(x+1)}{f_n(x)} = \frac{q-z\sqrt{pq/n}}{p+z\sqrt{pq/n}+1/n}\Big(\frac{p}{q}\Big) = \frac{1-z\sqrt{p/qn}}{1+z\sqrt{q/pn}+1/pn}.$$

Tomando $\Delta = \frac{1}{\sqrt{npq}}$ vemos que

$$\frac{f_n(x+1)}{f_n(x)} = \frac{1-zp\Delta}{1+zq\Delta+q\Delta^2}. \qquad (\mathbf{2})$$

Escribiendo ahora $Z_n = \frac{X_n - np}{\sqrt{npq}}$ se obtiene que

$$f_n(x+1) = \mathbf{P}\big\{X_n = x+1\big\} = \mathbf{P}\big\{X_n = np + z\sqrt{npq}+1\big\} = \mathbf{P}\big\{Z_n = z+\Delta\big\}.$$

De manera similar vemos que $f_n(x) = \mathbf{P}\big\{Z_n = z\big\}$. Todo lo anterior nos lleva a

$$\frac{f_n(x+1)}{f_n(x)} = \frac{\mathbf{P}\big\{Z_n = z+\Delta\big\}}{\mathbf{P}\big\{Z_n = z\big\}} = \frac{1-zp\Delta}{1+zq\Delta+q\Delta^2}.$$

Nuestro siguiente paso es aproximar la distribución de Z_n usando una función de densidad continua $g(z)$. Para esto queremos escoger $g(z)$ de tal forma que la razón

$$\frac{g(z+\Delta)}{g(z)}$$

sea muy cercana a

$$\frac{\mathbf{P}\big\{Z_n = z+\Delta\big\}}{\mathbf{P}\big\{Z_n = z\big\}}.$$

Dicho de otra manera, queremos que

$$\frac{g(z+\Delta)}{g(z)} \simeq \frac{1-zp\Delta}{1+zq\Delta+q\Delta^2}.$$

Tomando logaritmo y dividiendo entre Δ tenemos

$$\frac{\log\big(g(z+\Delta)\big) - \log\big(g(z)\big)}{\Delta} \;=\; \frac{1}{\Delta}\log(1 - zp\Delta) - \frac{1}{\Delta}\log(1 + zq\Delta + q\Delta^2).$$

Ahora tomamos el límite cuando $\Delta \to 0$ y obtenemos

$$\frac{d}{dz}\log\big(g(z)\big) \;=\; -zp - zq \;=\; -z.$$

Por el teorema fundamental del cálculo sabemos que existe una constante c tal que

$$\log\big(g(z)\big) \;=\; -\frac{z^2}{2} + \log c.$$

Aplicando la función exponencial, obtenemos

$$g(z) \;=\; ce^{-\frac{1}{2}x^2}.$$

Ya que $g(z)$ es una función de densidad, entonces $\int_{-\infty}^{\infty} g(z)\,dz = 1$. Por lo tanto $c = \frac{1}{\sqrt{2\pi}}$.

§3. El Teorema de DeMoivre-Laplace con Término de Error.

El objetivo de este apartado es demostrar el siguiente teorema. Se trata de una versión del teorema de DeMoivre-Laplace en términos de la esperanza matemática de una función f. El caso clásico se obtiene cuando f es la función indicadora de un intervalo $[a, b]$, es decir, cuando f es de la forma

$$f(x) \;=\; \begin{cases} 1, & \text{si } x \in [a,b], \\ 0, & \text{si } x \notin [a,b]. \end{cases}$$

También se ha estimado la razón de convergencia hacia la distribución normal.

Teorema. (DeMoivre-Laplace). *Sea* $0 < p < 1$. *Sea* $q = 1 - p$. *Sea* $\{X_j\}_{j=1}^{\infty}$ *una sucesión de variables aleatorias independientes e idénticamentes distribuidas tales que*

$$\mathbf{P}\{X_1 = 1\} = p \qquad \text{y además} \qquad \mathbf{P}\{X_1 = 0\} = q.$$

Sea $\xi_n = X_1 + \cdots + X_n$. *Sea* $f\colon \mathbf{R} \to \mathbf{R}$ *una función acotada, Lipschitz continua y que se anula fuera de un intervalo cerrado. Entonces*

$$\mathbf{E}\left[f\left(\frac{\xi_n - np}{\sqrt{npq}}\right)\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-\frac{1}{2}x^2}\,dx + O\left[\frac{(\log n)^{\frac{3}{2}}}{\sqrt{n}}\right].$$

Como preámbulo a la demostración del teorema se probaran 5 lemas. El primero de ellos se conoce como la fórmula de Stirling y nos permite aproximar factoriales de números naturales grandes. Antes de abordar la demostración del primer lema, veremos, mediante un cálculo informal, su importancia e interés. Se observará que la fórmula de Stirling involucra la expresión $\sqrt{2\pi}$ que aparece en la distribución normal.

Si n es un entero muy grande, tenemos que

$$\log n! = \sum_{k=1}^{n} \log k \simeq \int_{1}^{n} \log t\,dt = n\log n - n.$$

Esto sugiere que $n! \simeq n^n e^{-n}$. Una comparación con el Lema 1 hará evidente que esta aproximación no es del todo correcta. La versión correcta de esta aproximación es conocida como la fórmula de Stirling, en honor al matemático escocés del mismo apellido. Es interesante hacer notar que algunas referencias citan que DeMoivre fue quien la publicó por primera vez en 1730 en su libro *Miscellanea Analytica*.

Lema 1 (Stirling). *Para $n \to \infty$, se cumple que*

$$n! \;=\; \sqrt{2\pi n}\,\Big(\frac{n}{e}\Big)^n \Big\{1 + O\Big(\frac{1}{n}\Big)\Big\}.$$

Demostración. Sea

$$S_n \;=\; \log(n!) - n\big(\log(n) - 1\big) \;=\; \sum_{k=1}^{n} \log(k) - n\big(\log(n) - 1\big).$$

Cuando $n \geq 2$ tenemos que

$$S_n - S_{n-1} \;=\; \log(n) - n\big(\log(n) - 1\big) + (n-1)\big(\log(n-1) - 1\big) \;=\; (n-1)\log\Big(1 - \frac{1}{n}\Big) + 1.$$

Utilizando la serie de Taylor para $\log(1-x)$ con $|x| < 1$ encontramos que

$$S_n - S_{n-1} \;=\; 1 - (n-1)\sum_{j=1}^{\infty} \frac{1}{j}\Big(\frac{1}{n}\Big)^j \;=\; \sum_{j=1}^{\infty} \frac{1}{j(j+1)n^j}.$$

Nótese que

$$0 \;\leq\; \sum_{j=2}^{\infty} \frac{1}{j(j+1)n^j} \;\leq\; \frac{1}{n^2}\sum_{j=2}^{\infty} \frac{1}{j(j+1)} \;=\; \frac{1}{n^2}.$$

Por lo tanto tenemos que

$$S_n - S_{n-1} \;=\; \frac{1}{2n} + E_n \qquad \text{en donde} \qquad E_n = O\Big(\frac{1}{n^2}\Big).$$

Dado este resultado para el S_n encontramos que

$$S_n \;=\; \sum_{j=2}^{n} (S_j - S_{j-1}) + S_1 \;=\; \frac{1}{2}\sum_{j=2}^{n} \frac{1}{j} + \sum_{j=1}^{n} E_j$$

en donde hemos puesto $E_1 = S_1$. Por lo tanto

$$S_n \;=\; \frac{1}{2}\log(n) + \frac{1}{2}\gamma + O\Big(\frac{1}{n}\Big) + \sum_{j=1}^{n} E_j$$

en donde γ es la constante de Euler. Ahora escribimos

$$\omega \;=\; \frac{1}{2}\gamma + \sum_{j=1}^{\infty} E_j.$$

En vista de que

$$\sum_{j=n+1}^{\infty} E_j \;\ll\; \int_{n}^{\infty} \frac{dt}{t^2} \;=\; \frac{1}{n}$$

entonces podemos concluir que

$$S_n \;=\; \frac{1}{2}\log(n) + \omega + O\Big(\frac{1}{n}\Big).$$

Del análisis precedente se deduce que

$$\log(n!) \;=\; n\big(\log(n) - 1\big) + \frac{1}{2}\log(n) + \omega + O\Big(\frac{1}{n}\Big).$$

Al aplicar la función exponencial, y en vista de que $e^{O(\frac{1}{n})} = 1 + O(\frac{1}{n})$, obtenemos

$$n! \;=\; \Big(\frac{n}{e}\Big)^n \sqrt{n}\, e^{\omega}\Big\{1 + O\Big(\frac{1}{n}\Big)\Big\}.$$

Para encontrar el valor de e^{ω} consideremos la siguiente integral

$$I_k \;=\; \int_0^{\frac{\pi}{2}} \big\{\operatorname{sen}(x)\big\}^k\,dx \;=\; (k-1)\int_0^{\frac{\pi}{2}} \big\{\operatorname{sen}(x)\big\}^{k-2}\big\{\cos(x)\big\}^2\,dx.$$

Al simplificar encontramos que $I_k = \frac{k-1}{k} I_{k-2}$. Así obtenemos las expresiones

$$I_{2n} \;=\; \frac{\prod_{k=1}^{n}(2k-1)}{\prod_{k=1}^{n} 2k}\frac{\pi}{2} \qquad \text{y} \qquad I_{2n+1} \;=\; \frac{\prod_{k=1}^{n} 2k}{\prod_{k=1}^{n}(2k+1)}.$$

Haciendo el cociente de I_{2n} con I_{2n-1} y de este último con I_{2n+1} llegamos a los siguientes resultados

$$\frac{I_{2n}}{I_{2n+1}} \;=\; \frac{\big(\prod_{k=1}^{n}(2k-1)\big)^2(2n+1)}{\big(\prod_{k=1}^{n} 2k\big)^2}\frac{\pi}{2} \qquad \text{y} \qquad \frac{I_{2n-1}}{I_{2n+1}} \;=\; \frac{2n+1}{2n}. \tag{3}$$

Nótese que para cada $x \in (0, \frac{\pi}{2})$ se cumple que $\operatorname{sen}(x) < 1$. Por lo tanto, para cada n natural se cumple que

$$(\operatorname{sen} x)^{2n-1} \; > \; (\operatorname{sen} x)^{2n} \; > \; (\operatorname{sen} x)^{2n+1}.$$

Integrando de 0 a $\frac{\pi}{2}$ las desigualdades anteriores vemos que $I_{2n-1} > I_{2n} > I_{2n+1}$. Por lo tanto

$$\frac{I_{2n-1}}{I_{2n+1}} > \frac{I_{2n}}{I_{2n+1}} > 1. \qquad (\mathbf{4})$$

Por la segunda identidad en (3) tenemos que $\frac{I_{2n}}{I_{2n+1}} \to 1$ cuando $n \to \infty$. Las desigualdades (4) ahora implican que

$$\lim_{n\to\infty} \frac{I_{2n}}{I_{2n+1}} \; = \; 1.$$

Esto es lo mismo que

$$\frac{\pi}{2} \; = \; \lim_{n\to\infty} \frac{\left(\prod_{k=1}^{n} 2k\right)^2}{\left(\prod_{k=1}^{n}(2k-1)\right)^2 2n}.$$

Para obtener un resultado más compacto se multiplicará la expresión por el número $\left(\prod_{k=1}^{n} 2k\right)^2$ tanto en el numerador como en el denominador. Obtenemos

$$\pi \; = \; \lim_{n\to\infty} \frac{\left(\prod_{k=1}^{n} 2k\right)^4}{\left(\prod_{k=1}^{n} k\right)^2 n} \; = \; \frac{2^{4n}(n!)^4}{(2n!)^2 n}.$$

La expresión obtenida es conocida como la fórmula del producto de Wallis para π en honor al matemático inglés de mismo apellido. Ya que tenemos una expresión para el factorial la sustituimos en la relación anterior y obtenemos

$$\pi \; = \; \lim_{n\to\infty} \frac{2^{4n} n^{4n+2} e^{4\omega-4n}}{2^{4n+1} n^{4n+2} e^{2\omega-4n}} \; = \; \frac{e^{2\omega}}{2}.$$

Se obteniene $e^{\omega} = \sqrt{2\pi}$, con lo que queda demostrado el lema. ∎

Lema 2. *Sea $a > 1$. Sea $0 < p < 1$ y sea $q = 1 - p$. Para cada k tal que $|k - np| < a\sqrt{n}$, se cumple que*

$$\binom{n}{k} p^k q^{n-k} \;=\; \frac{1}{\sqrt{2\pi pqn}} \exp\left\{ \frac{-(k-np)^2}{2npq} \right\} \left[1 + O\Big(\frac{a^3}{\sqrt{n}} \Big) \right].$$

Demostración. Por el lema anterior

$$\frac{n! p^k q^{n-k}}{k!(n-k)!} \;=\; \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{k(n-k)}} \left(\frac{qn}{n-k} \right)^{n-k} \left(\frac{pn}{k} \right)^k \frac{(1+\epsilon_n)}{(1+\epsilon_k)(1+\epsilon_{n-k})}$$

donde $\epsilon_n = O(\frac{1}{n})$, $\epsilon_k = O(\frac{1}{k})$ y $\epsilon_{n-k} = O(\frac{1}{n-k})$. Puesto que $|k - np| < a\sqrt{n}$ y $p + q = 1$ entonces

$$\frac{1}{a\sqrt{n} + np} \;<\; \frac{1}{k} \;<\; \frac{1}{np - a\sqrt{n}} \qquad \text{y} \qquad \frac{1}{nq + a\sqrt{n}} \;<\; \frac{1}{n-k} \;<\; \frac{1}{nq - a\sqrt{n}}.$$

Multiplicando ambas ecuaciones

$$\frac{1}{(nq + \sqrt{n})(np + \sqrt{n})} \;<\; \frac{1}{k(n-k)} \;<\; \frac{1}{(nq - \sqrt{n})(np - \sqrt{n})}.$$

De lo anterior se deduce

$$\frac{1}{k(n-k)} \;<\; \frac{1}{nq\big(1 - \frac{1}{q\sqrt{n}}\big) np\big(1 - \frac{1}{p\sqrt{n}}\big)} \;\leq\; \frac{1}{n^2 pq} \Big\{ 1 + O\Big(\frac{1}{\sqrt{n}} \Big) \Big\}.$$

De forma similar se obtiene que

$$\frac{1}{k(n-k)} \;\geq\; \frac{1}{n^2 pq} \Big\{ 1 + O\Big(\frac{1}{\sqrt{n}} \Big) \Big\}.$$

Resumiendo lo anterior obtenemos

$$\frac{n}{k(n-k)} \;=\; \frac{1}{npq} \Big\{ 1 + O\Big(\frac{1}{\sqrt{n}} \Big) \Big\}.$$

En vista de la serie de Taylor de la función raíz cuadrada, también se cumple que

$$\sqrt{\frac{n}{k(n-k)}} \;=\; \sqrt{\frac{1}{npq}} \Big\{ 1 + O\Big(\frac{1}{\sqrt{n}} \Big) \Big\}. \tag{5}$$

Por otro lado, se tiene que

$$\log\left[\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}\right] \;=\; k\log\left(1-\frac{k-np}{k}\right)+(n-k)\log\left(1+\frac{k-np}{n-k}\right).$$

Usando la serie de Taylor de $\log(1-x)$ donde $|x|<1$,

$$\log(1-x)+x+\frac{x^2}{2} \;=\; -\sum_{k=3}^{\infty}\frac{x^k}{k} \;<\; \sum_{k=3}^{\infty}x^k \;=\; \frac{x^3}{1-x}.$$

Ahora bien, si $|x|\leq\frac{1}{2}$, entonces

$$\log(1-x)+x+\frac{x^2}{2} \;\leq\; 2x^3 \qquad \text{y} \qquad \log(1-x) \;=\; -x-\frac{x^2}{2}+O(x^3).$$

Análogamente tenemos que $\log(1+x)=x-\frac{x^2}{2}+O(x^3)$.

Ya que $np-a\sqrt{n}<k<np+a\sqrt{n}$ entonces

$$\left|\frac{k-np}{k}\right| \;\leq\; \frac{np-k}{np-a\sqrt{n}} \;\leq\; \frac{a\sqrt{n}}{np-a\sqrt{n}} \;\ll\; \frac{a}{\sqrt{n}}.$$

Puesto que $nq-a\sqrt{n}<n-k<nq+a\sqrt{n}$ entonces

$$\left|\frac{k-np}{n-k}\right| \;\leq\; \frac{a\sqrt{n}}{nq-a\sqrt{n}} \;\ll\; \frac{a}{\sqrt{n}}.$$

Resulta entonces que

$$k\log\left(1-\frac{k-np}{k}\right)+(k-np)+\frac{(k-np)^2}{2k} \;\ll\; \frac{a^3}{\sqrt{n}},$$

$$(n-k)\log\left(1+\frac{k-np}{n-k}\right)-(k-np)+\frac{(k-np)^2}{2(n-k)} \;\ll\; \frac{a^3}{\sqrt{n}}.$$

Sumando ambas ecuaciones

$$k\log\left(\frac{np}{k}\right)+(n-k)\log\left(\frac{nq}{n-k}\right) \;=\; -\frac{(k-np)^2}{2}\Big[\frac{n}{k(n-k)}\Big]+O\Big(\frac{a^3}{\sqrt{n}}\Big).$$

En vista de la relación (5) tenemos

$$\begin{aligned} k\log\Big(\frac{np}{k}\Big)+(n-k)\log\Big(\frac{nq}{n-k}\Big) &= -\frac{(k-np)^2}{2npq}\Big[1+O\Big(\frac{1}{\sqrt{n}}\Big)\Big]+O\Big(\frac{a^3}{\sqrt{n}}\Big) \\ &= -\frac{(k-np)^2}{2npq}+O\Big(\frac{a^3}{\sqrt{n}}\Big). \end{aligned}$$

Tomando exponencial

$$\Big(\frac{np}{k}\Big)^k\Big(\frac{nq}{n-k}\Big)^{n-k} = e^{\frac{-(k-np)^2}{2npq}}\Big[1+O\Big(\frac{a^3}{\sqrt{n}}\Big)\Big]. \qquad (\mathbf{6})$$

Por otro lado existen constantes c_1, c_2, c_3 tales que

$$\epsilon_n<\frac{c_1}{n}, \qquad \epsilon_k<\frac{c_2}{k}, \qquad \epsilon_{n-k}<\frac{c_3}{n-k}.$$

También existen constantes c_4, c_5 tales que

$$\frac{1}{k}<\frac{c_4}{n} \qquad \text{y} \qquad \frac{1}{n-k}<\frac{c_5}{n}.$$

Se obtiene entonces que

$$\epsilon_k = O\Big(\frac{1}{n}\Big), \qquad \text{y} \qquad \epsilon_{n-k} = O\Big(\frac{1}{n}\Big).$$

Con los cálculos anteriores vemos que

$$\frac{1+\epsilon_n}{(1+\epsilon_k)(1+\epsilon_{n-k})} = 1+O\Big(\frac{1}{n}\Big). \qquad (\mathbf{7})$$

Sustituyendo (5), (6) y (7) en nuestra ecuación inicial

$$\begin{aligned} \binom{n}{k}p^kq^{n-k} &= \frac{1}{\sqrt{2\pi npq}}e^{\frac{-(k-np)^2}{2npq}}\Big[1+O\Big(\frac{1}{\sqrt{n}}\Big)\Big]\Big[1+O\Big(\frac{a^3}{\sqrt{n}}\Big)\Big]\Big[1+O\Big(\frac{1}{n}\Big)\Big] \\ &= \frac{1}{\sqrt{2\pi npq}}e^{\frac{-(k-np)^2}{2npq}}\Big[1+O\Big(\frac{a^3}{\sqrt{n}}\Big)\Big] \end{aligned}$$

con lo que queda terminada la demostración. Notemos que la constante implicada en el símbolo O no depende de k pero sí de p. ∎

Lema 3. *Sea $f : \mathbf{R} \to \mathbf{R}$ una función continua en el intervalo $[a, b]$ y que se anula fuera de $[a, b]$. Suponga que existe una constante $A > 0$ tal que $|f(x) - f(y)| < A|x - y|$ para cada x, y tales que $[x, y] \subset [a, b]$. Sea $M = \sup\{|f(x)| \,:\, a \le x \le b\}$. Para cada $t \in \mathbf{R}$ se cumple que*

$$h \sum_{k \in \mathbf{Z}} f(t + kh) \;=\; \int_a^b f(x) + O(h(M + b - a)).$$

Demostración. Puesto que

$$h \sum_{k \in \mathbf{Z}} f(t + kh) - h \int_{-\infty}^{\infty} f(t + xh)\, dx \;\le\; h \sum_{k \in \mathbf{Z}} \int_k^{k+1} |f(t + kh) - f(t + xh)|\, dx$$

$$\le\; h(4M + h\,\mathrm{Card}\{k \,:\, a \le t + kh \le b\})$$

entonces

$$h \sum_{k \in \mathbf{Z}} f(t + kh) - h \int_{-\infty}^{\infty} f(t + xh)\, dx \;\le\; 4Mh + h(b - a). \tag{8}$$

Como la función se anula fuera del intervalo $[a, b]$ y haciendo el cambio de variable $u = t + xh$ en la integral en (8) obtenemos el resultado. ∎

Para la prueba del último lema, conocido como la Estimación para las Grandes Desviaciones se hará uso de un lema conocido como la Desigualdad de Markov el cual se probará a continuación.

Lema 4 (A. Markov). *Sea $X : \Omega \to \mathbf{N}_0$ una variable aleatoria con valores en el conjunto de enteros no negativos $\mathbf{N}_0$. Sea a un número real positivo. Entonces*

$$\mathbf{P}\big\{X \ge a\big\} \;\le\; \frac{1}{a}\mathbf{E}(X).$$

Demostración. Si $p_j = \mathbf{P}\big\{X = j\big\}$, entonces

$$\mathbf{P}\big\{X \geq a\big\} \;=\; \sum_{a \leq j} p_j \;\leq\; \sum_{a \leq j} \frac{j}{a} p_j \;\leq\; \frac{1}{a}\mathbf{E}(X).$$

Lema 5. *Sea $0 < p < 1$ y sea $q = 1 - p$. Sea $0 < \epsilon < q$. Sea*

$$h_+(\epsilon) \;=\; (p+\epsilon)\log\frac{p+\epsilon}{p} + (q-\epsilon)\log\frac{q-\epsilon}{q}.$$

Sea $h_-(\epsilon) = h_+(-\epsilon)$. Sea $\{X_j\}_{j=1}^{\infty}$ una sucesión de variables aleatorias de tipo Bernoulli, independientes e idénticamente distribuidas, con p como probabilidad de éxito. Sea $\xi_n = X_1 + \cdots + X_n$. Entonces

(a) $\mathbf{P}\Big\{\frac{1}{n}\xi_n \;\geq\; p+\epsilon\Big\} \;\leq\; e^{-nh_+(\epsilon)}$.

(b) $\mathbf{P}\Big\{\frac{1}{n}\xi_n \;\leq\; p-\epsilon\Big\} \;\leq\; e^{-nh_-(\epsilon)}$.

(c) $\mathbf{P}\Big\{\Big|\frac{1}{n}\xi_n - p\Big| \;\geq\; \epsilon\Big\} \;\leq\; e^{-nh_+(\epsilon)} + e^{-nh_-(\epsilon)}$.

Demostración. Ya que ξ_n es la suma de n variables de tipo Bernoulli, entonces ξ_n tiene distribución binomial. Para demostrar (a) sea $t > 0$. Entonces

$$\xi_n \;\geq\; np + n\epsilon \qquad \text{si y sólo si} \qquad e^{t(\xi_n - np - n\epsilon)} \;\geq\; 1.$$

Utilizando el lema anterior vemos que

$$\mathbf{P}\Big\{\frac{1}{n}\xi_n \;\geq\; p+\epsilon\Big\} \;=\; \mathbf{P}\Big\{e^{t(\xi_n - np - n\epsilon)} \;\geq\; 1\Big\} \;\leq\; \mathbf{E}\Big[e^{t(\xi_n - np - n\epsilon)}\Big].$$

Puesto que ξ_n tiene una distribución binomial, entonces

$$\begin{aligned} e^{-nt(p+\epsilon)}\mathbf{E}\left[e^{t\xi_n}\right] \;&=\; e^{-nt(p+\epsilon)} \sum_{k=0}^{n} e^{tk}\binom{n}{k} p^k q^{n-k} \\ &=\; e^{-nt(p+\epsilon)}(q+pe^t)^n \\ &=\; \exp\big\{-ng(t)\big\} \end{aligned}$$

en donde $g(t) = t(p+\epsilon) - \log(q+pe^t)$. Por lo tanto

$$\mathbf{P}\Big\{\frac{1}{n}\xi_n \ \geq\ p+\epsilon\Big\} \ \leq\ e^{-ng(t)}$$

se cumple para cada $t > 0$ y en particular se cumple para $t = \alpha$, en donde

$$\alpha \ =\ \log\Big[\frac{q(p+\epsilon)}{p(q-\epsilon)}\Big].$$

La condición $\epsilon > 0$ implica que $\alpha > 0$. Ahora sólo basta observar que $g(\alpha) = h_+(\epsilon)$. Es interesante hacer notar que $g(t) \leq g(\alpha)$ se cumple para cada $t > 0$ y por lo tanto la desigualdad (a) del lema es óptima.

Para demostrar (b) definamos la variable aleatoria $\tilde{\xi}_n \ =\ n - \xi_n$. Definimos también $\tilde{p} = q$ y $\tilde{q} = p$. Entonces

$$\mathbf{P}\Big\{\frac{\xi_n}{n} \ \leq\ p-\epsilon\Big\} \ =\ \mathbf{P}\Big\{\frac{\tilde{\xi}_n}{n} \ \geq\ \tilde{p}+\epsilon\Big\} \ \leq\ e^{-n\tilde{h}_+(\epsilon)}$$

en donde

$$\begin{aligned}
\tilde{h}_+(\epsilon) \ &=\ (\tilde{p}+\epsilon)\log\frac{\tilde{p}+\epsilon}{\tilde{p}} + (\tilde{q}-\epsilon)\log\frac{\tilde{q}-\epsilon}{\tilde{q}} \\
&=\ (q+\epsilon)\log\frac{q+\epsilon}{q} + (p-\epsilon)\log\frac{p-\epsilon}{p} \\
&=\ h_-(\epsilon).
\end{aligned}$$

Para probar (c) nótese que por los incisos anteriores se cumple

$$\begin{aligned}
\mathbf{P}\Big\{\Big|\frac{1}{n}\xi_n - p\Big| \geq \epsilon\Big\} &= \mathbf{P}\Big\{\Big(\epsilon + p \leq \frac{\xi_n}{n}\Big) \cup \Big(\frac{\xi_n}{n} \leq p-\epsilon\Big)\Big\} \\
&= \mathbf{P}\Big\{\Big(\epsilon + p \leq \frac{\xi_n}{n}\Big)\Big\} + \mathbf{P}\Big\{\Big(\frac{\xi_n}{n} \leq p-\epsilon\Big)\Big\} \\
&\leq\ e^{-nh_+(\epsilon)} + e^{-nh_-(\epsilon)}. \ \blacksquare
\end{aligned}$$

Para demostrar el Teorema de DeMoivre-Laplace sea $K > 0$ y consideremos

$$\sigma_1 + \sigma_2 \;=\; \sum_{k=0}^{n} f\Big(\frac{k-np}{\sqrt{npq}}\Big)\,\mathbf{P}\big\{\xi_n = k\big\}$$

donde σ_1 es la suma sobre aquellos índices k tales que cumple $|k-np| < K\sqrt{n}$ y la suma σ_2 es sobre los k tales que $|k-np| \geq K\sqrt{n}$. Como $\epsilon \to 0$ entonces

$$\log\Big(1 \pm \frac{\epsilon}{p}\Big) \mp \frac{\epsilon}{p} + \frac{1}{2}\Big(\frac{\epsilon}{p}\Big)^2 \;\ll\; \epsilon^3.$$

Por lo tanto

$$(p+\epsilon)\log\Big(1+\frac{\epsilon}{p}\Big) - \frac{(p+\epsilon)\epsilon}{p} + \frac{(p+\epsilon)\epsilon^2}{2p^2} \;\ll\; \epsilon^3,$$

$$(q-\epsilon)\log\Big(1-\frac{\epsilon}{q}\Big) + \frac{(q-\epsilon)\epsilon}{q} + \frac{(q-\epsilon)\epsilon^2}{2q^2} \;\ll\; \epsilon^3.$$

Sumando ambas ecuaciones vemos que

$$h_+(\epsilon) \;=\; \frac{\epsilon^2}{2pq} + O(\epsilon^3).$$

Con el mismo procedimiento se puede ver que $h_-(\epsilon) = \frac{\epsilon^2}{2pq} + O(\epsilon^3)$. Por el Lema 5 con $\epsilon = K\sqrt{\frac{2pq}{n}}$ se obtiene que $h_\pm(\epsilon) = O(\frac{K^2}{n})$. Por lo tanto

$$\sigma_2 \;\ll \sum_{|k-np|\geq K\sqrt{n}} \mathbf{P}\big\{\xi_n = k\big\} \;=\; \mathbf{P}\Big\{\Big|\frac{1}{n}\xi_n - p\Big| \;\geq\; \frac{K}{\sqrt{n}}\Big\} \;\ll\; e^{-K^2}.$$

Ya que σ_1 es una suma sobre los k tales que $|k-np| < K\sqrt{n}$, entonces es posible aplicar el Lema 2 y así obtener

$$\sigma_1 \;=\; \Big[1 + O\Big(\frac{K^3}{\sqrt{n}}\Big)\Big] \sum_k f\Big(\frac{k-np}{\sqrt{npq}}\Big)\,\frac{e^{-\frac{(k-np)^2}{2\pi npq}}}{\sqrt{2\pi npq}}.$$

Con $h = \frac{1}{\sqrt{npq}}$ y $t = -\sqrt{\frac{np}{q}}$ el lema 3 implica

$$\begin{aligned}
\sum_k f\Big(\frac{k-np}{\sqrt{npq}}\Big)\,\frac{e^{-\frac{(k-np)^2}{2\pi npq}}}{\sqrt{npq}} \;&=\; \frac{1}{\sqrt{npq}}\sum_k f\Big(\frac{k}{\sqrt{npq}} - \sqrt{\frac{np}{q}}\Big)e^{-\frac{1}{2\pi}\left(\frac{k}{\sqrt{npq}}-\sqrt{\frac{np}{q}}\right)} \\
&=\; h\sum_k f(t+kh)e^{-\frac{(t+kh)^2}{2\pi}} \\
&=\; \int_{-\frac{K}{\sqrt{pq}}}^{\frac{K}{\sqrt{pq}}} f(x)\,e^{-\frac{x^2}{2}}\,dx + O\Big(\frac{K}{\sqrt{n}}\Big).
\end{aligned}$$

Sustituyendo en la ecuación para σ_1

$$\sigma_1 \;=\; \Big[1+O\Big(\frac{K^3}{\sqrt{n}}\Big)\Big]\Big[\frac{1}{\sqrt{2\pi}}\int_{-\frac{K}{\sqrt{pq}}}^{\frac{K}{\sqrt{pq}}} f(x)e^{-\frac{x^2}{2}}\,dx + O\Big(\frac{K}{\sqrt{n}}\Big)\Big].$$

Por lo tanto

$$\begin{aligned}
\mathbf{E}\Big[f\Big(\frac{\xi_n-np}{\sqrt{npq}}\Big)\Big] \;&=\; \sum_k f\Big(\frac{k-np}{\sqrt{npq}}\Big)\mathbf{P}\{\xi_n=k\} \\
&=\; \Big[1+O\Big(\frac{K^3}{\sqrt{n}}\Big)\Big]\Big[\frac{1}{\sqrt{2\pi}}\int_{-\frac{K}{\sqrt{pq}}}^{\frac{K}{\sqrt{pq}}} f(x)\,e^{-\frac{x^2}{2}}\,dx + O\Big(\frac{K}{\sqrt{n}}\Big)\Big] + O(e^{-K^2}) \\
&=\; \frac{1}{\sqrt{2\pi}}\int_{-\frac{K}{\sqrt{pq}}}^{\frac{K}{\sqrt{pq}}} f(x)\,e^{-\frac{x^2}{2}}\,dx + O(e^{-K^2}) + O\Big(\frac{K}{\sqrt{n}}\Big).
\end{aligned}$$

Tomando $K^2 = \log n$ obtenemos

$$\int_{\frac{K}{\sqrt{pq}}}^{\infty} f(x)\,e^{-\frac{x^2}{2}}\,dx \;\ll\; \int_K^{\infty} x\,e^{-\frac{x^2}{2}}\,dx \;=\; e^{-\frac{K^2}{2}} \;=\; \frac{1}{\sqrt{n}}.$$

Ahora podemos concluir que

$$\begin{aligned}
\mathbf{E}\Big[f\Big(\frac{\xi_n - np}{\sqrt{npq}}\Big)\Big] &= \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} f(x)\, e^{-\frac{x^2}{2}}\, dx + O\Big(\frac{1}{n}\Big) + O\Big(\frac{(\log n)^{\frac{3}{2}}}{\sqrt{n}}\Big) \\
&= \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} f(x)\, e^{-\frac{x^2}{2}}\, dx + O\Big(\frac{(\log n)^{\frac{3}{2}}}{\sqrt{n}}\Big)
\end{aligned}$$

con lo que queda demostrado el teorema.

Capítulo 2

Ecuaciones de Kolmogorov

§1. Introducción.

Muchos fenómenos que ocurren en algunas ramas de la ciencia, como por ejemplo la física, la biología y las ciencias sociales, se estudian no solamente como fenómenos aleatorios sino también como variables aleatorias que dependen del tiempo. El estudio de este tipo de variables aleatorias que dependen del tiempo ha ido en aumento dando lugar a una parte importante de la teoría de probabilidades conocida como *procesos estocásticos*. El primer paso para incorporar lo estocástico a un modelo matemático de un fenómeno dinámico es determinar qué tipo de variable aleatoria se puede observar. En el siguiente paso se busca determinar cómo evoluciona la distribución de esta variable aleatoria a través del tiempo.

Como primer ejemplo consideremos un experimento tan simple como lanzar una moneda. Para cada número natural n, hacemos que X_n represente al enésimo lanzamiento. Entonces X_n es una variable aleatoria de tipo Bernoulli y $X_n = 0$ denota que en el enésimo lanzamiento se observó un "águila". Si $X_n = 1$ entonces sabemos que en el enésimo lanzamiento se observó "sol". En nuestro modelo matemático el enésimo lanzamiento ocurre en la enésima unidad de tiempo. La ley de evolución en este proceso estocástico es la más simple, ya que la distribución de las variables X_n no depende de n. En particular las variables X_n son independientes.

El conjunto de valores de una variable aleatoria X_n en un proceso estocástico se conoce como *espacio estado*. El espacio de estado es discreto si contiene un número de puntos a lo más numerable, de otra forma es continuo.

Volviendo al ejemplo anterior sobre la sucesión de lanzamientos de una moneda, el espacio de estados es el conjunto $\{0, 1\}$. Por lo tanto, se trata de un espacio de estado discreto.

El matemático ruso A. A. Markov estudió ciertos procesos estocásticos en los cuales la evolución en el tiempo tiene la característica de que el futuro depende solamente del estado presente del sistema, pero no de toda la historia previa al estado presente del sistema en cuestión. Este tipo de procesos reciben el nombre de *procesos de Markov*. Cuando el espacio de estado de un proceso de Markov es discreto, entonces se le conoce como *cadena de Markov*. Dadas sus múltiples aplicaciones, las cadenas de Markov son los ejemplos más conocidos de procesos estocásticos.

En este capítulo son de nuestro interés ciertos procesos estocásticos conocidos como *procesos de difusión*. Hay muchos ejemplos de procesos de difusión. En el capítulo 3 damos dos ejemplos interesantes de procesos de difusión con espacio de estado continuo. Los ejemplos del capítulo 3 se obtienen a partir de las ecuaciones de Kolmogorov que se van a deducir en el apartado §4 de este capítulo. En el apartado §2 de este capítulo se aplica el teorema de DeMoivre-Laplace para obtener los casos más sencillos de las ecuaciones de Kolmogorov. Si bien el apartado §2 no es necesario para la deducción de las ecuaciones de Kolmogorov, la discusión del apartado §2 ayuda bastante a formarse una intuición sobre los procesos de difusión y sobre las ecuaciones de Kolmogorov.

§2. Movimiento Browniano.

En 1828 el botánico inglés Robert Brown observó el movimiento de una partícula de polen suspendida en agua llamando su atención el movimiento irregular que se podía observar. Fue en 1905 cuando Albert Einstein explicó que dicha partícula se ve bombardeada en todas direcciones por moléculas de agua y estos impulsos la hacen moverse aleatoriamente. La explicación matemática para este fenómeno, como un proceso estocástico, fue dada por N. Wiener en 1931. En este apartado se presenta un modelo muy simple para el movimiento del polen observado por Brown. Éste es un ejemplo de proceso estocástico de Markov con un espacio de estados continuo que se conoce como proceso de Wiener o *Movimiento Browniano*.

Una caminata aleatoria es un proceso estocástico discreto que es análogo al movimiento browniano. Se puede pensar que el movimiento browniano se obtiene al tomar el límite de una caminata aleatoria, cuando el tamaño de los pasos de la caminata se hacen cada vez más pequeños. Se utilizará el teorema de DeMoivre-Laplace para tomar el límite anterior y así poder pasar de un estado discreto a uno continuo. Después utilizaremos este modelo para obtener la ecuación diferencial parcial que rige a un proceso de difusión.

Consideremos una partícula que realiza una caminata aleatoria simétrica unidimensional en el intervalo $[0, T]$ y con magnitud de los pasos Δ. La partícula se puede mover hacia la derecha o izquierda. Sea n un entero positivo. En cada uno de los n instantes de tiempo $j\delta$, con $\delta = \frac{T}{n}$ y $0 < j \leq n$ se observa el valor de una variable aleatoria ξ_j. Dicha variable determina en qué dirección (derecha o izquierda) se mueve la partícula en la unidad de tiempo j. Puesto que se trata de una caminata aleatoria simétrica, entonces

$$\mathbf{P}\{\xi_j = \Delta\} \;=\; \frac{1}{2} \;=\; \mathbf{P}\{\xi_j = -\Delta\}.$$

La posición en que se encuentra la partícula al tiempo k está dada por la expresión

$$w_k \;=\; \sum_{j=1}^{k} \xi_j.$$

Para poder aplicar el teorema de DeMoivre-Laplace escribimos

$$X_j \;=\; \frac{\xi_j + \Delta}{2\Delta} \qquad \text{y} \qquad S_k = \sum_{j=1}^{k} X_j.$$

Tomando $p = q = \frac{1}{2}$ tenemos

$$\frac{S_k - kp}{\sqrt{kpq}} \;=\; \frac{2}{\sqrt{k}}\Big(S_k - \frac{k}{2}\Big) \;=\; \frac{2}{\sqrt{k}}\Big\{\frac{1}{2\Delta}\sum_{j=1}^{k}\xi_j\Big\} \;=\; \frac{w_k}{\Delta\sqrt{k}}.$$

Dado un valor fijo de t que satisface $0 < t \leq T$ seleccionamos un entero k tal que $k \simeq \frac{tn}{T}$. Se obteniene entonces que

$$k\delta \;=\; \frac{kT}{n} \;\simeq\; t.$$

Esta aproximación justifica que de aquí en adelante escribamos w_t en lugar de w_k. Como queremos hacer que $n \to \infty$ para que las observaciones de las variables aleatorias ξ_j sean cada vez más frecuentes, entonces debemos eliminar el parámetro n. Hacemos que $\Delta = \sigma\sqrt{\delta}$, donde $\sigma > 0$ es un parámetro que nos permite controlar la longitud de los pasos. Un valor grande de σ nos dice que la partícula se desplaza con grandes pasos y lo contrario cuando es pequeño. Por esto σ es nombrado *coeficiente de difusión.* Así entonces

$$\frac{w_k}{\Delta\sqrt{k}} \;\simeq\; \frac{w_t}{\sigma\sqrt{t}}.$$

Ahora ya podemos aplicar el teorema de DeMoivre-Laplace para ver que

$$\mathbf{P}\big\{w_t \le x\big\} \;=\; \mathbf{P}\Big\{\frac{w_t}{\sigma\sqrt{t}} \le \frac{x}{\sigma\sqrt{t}}\Big\} \;=\; \mathbf{P}\Big\{\frac{S_k - kp}{\sqrt{kpq}} \le \frac{x}{\sigma\sqrt{t}}\Big\} \;\simeq\; \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{x}{\sigma\sqrt{t}}} e^{-\frac{1}{2}y^2}\, dy.$$

Haciendo el cambio de variable $u = \sigma\sqrt{t}y$ y escribiendo $=$ en lugar de $\simeq$ tenemos

$$\mathbf{P}\big\{w_t \le x\big\} \;=\; \frac{1}{\sigma\sqrt{2\pi t}} \int\limits_{-\infty}^{x} e^{-\frac{y^2}{2t\sigma^2}}\, dy.$$

Hemos obtenido la función de densidad $f(x,t) = \frac{1}{\sigma\sqrt{2\pi t}} e^{-\frac{x^2}{2t\sigma^2}}$. Tomando la derivada parcial de f con respecto a t y la segunda parcial de f con respecto a x vemos que dicha función f satisface la ecuación diferencial parcial

$$\frac{\partial f}{\partial t} \;=\; \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2}.$$

Ésta se conoce como ecuación de difusión o también como ecuación de calor.

§3. Movimiento Browniano con Arrastre.

Consideremos ahora una caminata aleatoria no simétrica. En este caso tenemos

$$\mathbf{P}\big\{\xi_j = \Delta\big\} \;=\; p \quad \text{y} \quad \mathbf{P}\big\{\xi_j = -\Delta\big\} \;=\; q \quad \text{con} \quad 0 < p < 1 \quad \text{y} \quad p + q = 1.$$

Se cumple entonces que

$$S_k - kp \;=\; \frac{1}{2\Delta}\sum_{j=1}^{k} \xi_j + k\big(\frac{1}{2} - p\big) \;=\; \frac{w_k}{2\Delta} + k\big(\frac{1}{2} - p\big).$$

Calculemos el valor esperado de ξ_j

$$\mathbf{E}(\xi_i) \;=\; \Delta\,\mathbf{P}\big\{\xi_j = \Delta\big\} - \Delta\,\mathbf{P}\big\{\xi_j = -\Delta\big\} \;=\; \Delta(p - q).$$

Ahora calculemos la varianza de ξ_j

$$\begin{aligned}
\mathbf{V}(\xi_i) &= p\Delta^2\big(1-(p-q)\big)^2 + q\Delta^2\big(1+(p-q)\big)^2 \\
&= p\Delta^2(2q)^2 + q\Delta^2\big(2p\big)^2 \\
&= 4\Delta^2 pq(q+p) \\
&= 4\Delta^2 pq.
\end{aligned}$$

Como en el caso simétrico, consideramos un valor t tal que $0 \leq t \leq T$. Sea k tal que $k \simeq \frac{tn}{T}$. Entonces

$$k\delta = \frac{kT}{n} \simeq t.$$

También como antes, escribimos w_t en lugar de w_k. Se cumple que

$$\mathbf{E}(w_t) = k\mathbf{E}(\xi_j) = k\Delta(p-q) \simeq t\,\frac{p-q}{\delta}\,\Delta.$$

Ya que las variables ξ_j son independientes, entonces

$$\mathbf{V}(w_t) = k\mathbf{V}(\xi_j) = 4\frac{t}{\delta}\,pq\,\Delta^2.$$

Ahora vamos a suponer que existen dos números μ y σ^2 tales que

$$\mu = \lim_{\Delta\to 0} \frac{p-q}{\delta}\,\Delta \qquad \text{y} \qquad \sigma^2 = \lim_{\Delta\to 0} \frac{4}{\delta}\,pq\,\Delta^2.$$

Esto se satisface tomando

$$\Delta = \sigma\sqrt{\delta}, \qquad p = \frac{1}{2} + \frac{\mu\sqrt{\delta}}{2\sigma} \qquad \text{y} \qquad q = \frac{1}{2} - \frac{\mu\sqrt{\delta}}{2\sigma}.$$

El parámetro μ es conocido como el *coeficiente de arrastre* y nos indica la posición promedio de la partícula en cada instante de tiempo. El parámetro σ^2 es conocido como el *coeficiente de difusión* y describe qué tanto varía dicho movimiento.

Con los valores anteriores y escribiendo w_t en lugar de w_k,

$$\frac{w_k}{2\Delta} + k\big(\frac{1}{2} - p\big) \ = \ \frac{w_t}{2\sigma\sqrt{\delta}} - k\frac{\mu\sqrt{\delta}}{2\sigma} \ = \ \frac{w_t}{2\sigma\sqrt{\delta}} - \frac{t\mu}{2\sigma\sqrt{\delta}}.$$

Por otro lado,

$$\frac{1}{\sqrt{kpq}} \ = \ 2\sqrt{\frac{\delta}{t}} + O(\delta).$$

Por lo tanto

$$\lim_{\delta\to 0} \frac{S_k - kp}{\sqrt{kpq}} \ = \ \frac{w_t - t\mu}{\sigma\sqrt{t}} \ = \ \frac{aw_t - bt}{\sqrt{t}}$$

donde $a = \frac{1}{\sigma}$ y $b = \frac{\mu}{\sigma}$. Aplicando el teorema de DeMoivre-Laplace,

$$\mathbf{P}\big\{w_t \leq x\big\} \ = \ \mathbf{P}\Big\{\frac{aw_t - bt}{\sqrt{t}} \leq \frac{ax - bt}{\sqrt{t}}\Big\} \ = \ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{ax-bt}{\sqrt{t}}} e^{-\frac{1}{2}y^2}\, dy.$$

Al estudiar la difusión de un contaminante en algún líquido, la expresión anterior nos indica el pocentaje de moléculas del contaminante cuya posición w_t al tiempo t es menor o igual a un número x.

Haciendo el cambio de variable $u = \frac{\sqrt{t}y + bt}{a}$ la integral en la última expresión nos queda es igual a

$$\mathbf{P}\big\{w_t \leq x\big\} \ = \ \frac{a}{\sqrt{2\pi t}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{au-bt}{\sqrt{t}}\right)^2}\, du$$

siendo el integrando la función de densidad $f(x,t)$ requerida. Ahora calculamos la derivada de f con respecto a t,

$$\frac{\partial f}{\partial t} \ = \ \frac{-ae^{\frac{-(ax-bt)^2}{2t}}}{\sqrt{2\pi t}} \left(\frac{1}{2t} - \frac{a^2x^2}{2t^2} + \frac{b^2}{2}\right).$$

Tomando la primera y segunda derivada parcial con respecto a x, se obtiene

$$\frac{\partial f}{\partial x} = \frac{-a^2 e^{\frac{-(ax-bt)^2}{2t}}}{\sqrt{2\pi t}} \left(\frac{ax}{t} - b\right),$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{-a^3 e^{\frac{-(ax-bt)^2}{2t}}}{\sqrt{2\pi t}} \left(\frac{1}{t} - \frac{a^2x^2}{t^2} + \frac{2axb}{t} - b^2\right).$$

Al multiplicar la parcial con respecto a x por $-\frac{b}{a}$ y la segunda parcial por $\frac{1}{2a^2}$ y sumando, se obtiene que

$$\frac{1}{2a^2}\frac{\partial^2 f}{\partial x^2} - \frac{b}{a}\frac{\partial f}{\partial x} = \frac{\partial f}{\partial t}.$$

Como $\frac{b}{a} = \mu$ y $\frac{1}{2a^2} = \frac{\sigma^2}{2}$ entonces vemos que f satisface la ecuación diferencial parcial

$$\frac{\partial f}{\partial t} = \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2} - \mu\frac{\partial f}{\partial x}$$

conocida como ecuación de difusión con coeficiente de arrastre μ.

§4. Ecuaciones de Kolmogorov.

En este apartado vamos a obtener las ecuaciones de Kolmogorov. Será conveniente dar una definición precisa de un proceso estocástico de tipo Markov.

Un proceso estocástico $\xi(t)$, con $t \geq 0$, se llama un proceso de Markov con espacio de estado $\mathbf{R}$, si se cumple que para para cada colección de tiempos $0 \leq t_0 < t_1 < \cdots < t_n < \infty$ y cada $x \in \mathbf{R}$, se tiene

$$\mathbf{P}\big\{\xi(t_n) \leq x \,\big|\, \xi(t_0), \cdots, \xi(t_{n-1})\big\} = \mathbf{P}\big\{\xi(t_n) \leq x \,\big|\, \xi(t_{n-1})\big\}.$$

Es decir, la distribución de $\xi(t_n)$ depende del valor de $\xi(t_{n-1})$ pero no de los valores de $\xi(t_0), \cdots, \xi(t_{n-2})$.

Sea $s < t$. La función

$$F(s, x \,|\, t, y) \;=\; \mathbf{P}\big\{\xi(t) \leq y \;\big|\; \xi(s) = x\big\}$$

se llama *función de distribución condicional* del proceso $\xi(t)$. La función

$$p(s, x \,|\, t, y) \;=\; \frac{\partial}{\partial y} F(s, x \,|\, t, y)$$

se llama *función de densidad condicional* de $\xi(t)$. En particular se cumple que

$$\mathbf{P}\big\{a \leq \xi(t) \leq b \;\big|\; \xi(s) = x\big\} \;=\; \int\limits_a^b p(s, x \,|\, t, y)\, dy.$$

para cualesquiera números reales $a < b$.

Por ejemplo, para el movimiento Browniano la función de densidad condicional está dada por

$$p(s, x \,|\, t, y) \;=\; \frac{1}{\sigma\sqrt{2\pi(t-s)}} \exp\Big\{-\frac{(y-x)^2}{2\sigma^2(t-s)}\Big\}.$$

Nótese que en este caso la función de densidad condicional depende sólo de la diferencia $t - s$. Por esta razón podemos escribir la función de densidad condicional del movimiento Browniano como $p(x \,|\, t-s, y)$ en donde

$$p(x \,|\, t, y) \;=\; \frac{1}{\sigma\sqrt{2\pi t}} \exp\Big\{-\frac{(y-x)^2}{2\sigma^2 t}\Big\}.$$

Supongamos que en un proceso de tipo Markov tenemos tres eventos A, B_j, C que ocurren en los tiempos $s < u < t$ respectivamente. La propiedad de Markov implica que

$$\mathbf{P}\big\{C \;\big|\; B_j A\big\} \;=\; \mathbf{P}\big\{C \;\big|\; B_j\big\}.$$

Por lo tanto
$$\frac{\mathbf{P}\{CB_j\}}{\mathbf{P}\{B_j\}} \;=\; \frac{\mathbf{P}\{CB_jA\}}{\mathbf{P}\{B_jA\}}.$$
Se cumple entonces que

$$\mathbf{P}\{AC\} \;=\; \sum_j \frac{\mathbf{P}\{CB_j\}}{\mathbf{P}\{B_j\}}\mathbf{P}\{B_jA\}$$

en donde $\{B_j\}$ es una partición del espacio muestral. Al dividir entre $\mathbf{P}\{A\}$ se obtiene la siguiente fórmula de probabilidad total

$$\mathbf{P}\{C \mid A\} \;=\; \sum_j \mathbf{P}\{C \mid B_j\}\mathbf{P}\{B_j \mid A\}.$$

Si ahora ponemos (identificando j con z)

$$A \;=\; \{\xi(s) = x\}, \qquad B_j \;=\; \{z \le \xi(u) < z + dz\}, \qquad C \;=\; \{y \le \xi(t) < y + dy\}$$

entonces
$$\mathbf{P}\{C \mid A\} \;=\; p(s, x \,|\, t, y)\, dy,$$
$$\mathbf{P}\{C \mid B_j\} \;=\; p(u, z \,|\, t, y)\, dy,$$
$$\mathbf{P}\{B_j \mid A\} \;=\; p(s, x \,|\, u, z)\, dz.$$
Por la fórmula de probabilidad total, tenemos que

$$p(s, x \,|\, t, y) \;=\; \int_{-\infty}^{\infty} p(s, x \,|\, u, z)\, p(u, z \,|\, t, y)\, dz.$$

Esta última expresión se conoce como la fórmula de *Chapman-Kolmogorov.*

Se dice que un proceso de tipo Markov $\xi(s)$ con función de densidad condicional $p(s, x \,|\, t, y)$ es un *proceso de difusión* si, dado $\epsilon > 0$ y dado un intervalo de longitud finita $I = [a, b]$, las tres propiedades

$$\int_{|y-x|>\epsilon} p(s, x \,|\, s+h, y)\, dy \;=\; o(h), \tag{\textbf{1}}$$

$$\int_{|y-x|\le\epsilon} (y-x)\, p(s, x \,|\, s+h, y)\, dy \;=\; a(s, x)h + o(h), \tag{\textbf{2}}$$

$$\int_{|y-x|\le\epsilon} (y-x)^2\, p(s, x \,|\, s+h, y)\, dy \;=\; b^2(s, x)h + o(h) \tag{\textbf{3}}$$

se satisfacen de manera uniforme para cada $s \in I$.

Queremos ahora entender mejor el significado de las funciones $a(s, x)$ y $b^2(s, x)$. Para esto, supongamos que en lugar de (1) se cumple que

$$\lim_{h\to 0} \frac{1}{h} \int_{|y-x|>\epsilon} p(s, x \,|\, s+h, y)\, dy \;=\; 0. \tag{\textbf{4}}$$

Es claro que (4) implica (1). Ahora bien, si se cumple (4) entonces (2) y (3) se pueden escribir como

$$\lim_{h\to 0} \frac{1}{h} \int_{-\infty}^{\infty} (y-x)\, p(s, x \,|\, s+h, y)\, dy \;=\; a(s, x),$$

$$\lim_{h\to 0} \frac{1}{h} \int_{-\infty}^{\infty} (y-x)^2\, p(s, x \,|\, s+h, y)\, dy \;=\; b^2(s, x).$$

Por otro lado tenemos que

$$\int_{-\infty}^{\infty} (y-x)\, p(s, x \,|\, s+h, y)\, dy \;=\; \mathbf{E}\big[\xi(s+h) - \xi(s) \,\big|\, \xi(s) = x\big],$$

$$\int_{-\infty}^{\infty} (y-x)^2\, p(s, x \,|\, s+h, y)\, dy \;=\; \mathbf{V}\big[\xi(s+h) - \xi(s) \,\big|\, \xi(s) = x\big].$$

Por lo tanto

$$a(s,x) \;=\; \lim_{h\to 0}\frac{1}{h}\mathbf{E}\big[\xi(s+h)-\xi(s)\,\big|\,\xi(s)=x\big],$$

$$b^2(s,x) \;=\; \lim_{h\to 0}\frac{1}{h}\mathbf{V}\big[\xi(s+h)-\xi(s)\,\big|\,\xi(s)=x\big].$$

A continuación vamos a deducir la primera de las ecuaciones de Kolmogorov.

Teorema 1. *Suponga que la función $p(s,x\,|\,t,y)$ satisface las propiedades (1), (2), (3). Suponga además que las derivadas*

$$\frac{\partial p}{\partial s}, \qquad \frac{\partial p}{\partial x}, \qquad \frac{\partial^2 p}{\partial x^2}$$

son uniformemente continuas con respecto a x para $y \in [y_0, y_1]$. Se cumple entonces que

$$-\frac{\partial p}{\partial s} \;=\; a(s,x)\frac{\partial p}{\partial x} + \frac{1}{2}b^2(s,x)\frac{\partial^2 p}{\partial x^2}.$$

Demostración. Sea $\varphi(x)$ una función que se anula fuera de un intervalo de longitud finita. Sea

$$\Phi(s,x) \;=\; \int_{-\infty}^{\infty} \varphi(y)\,p(s,x\,|\,t,y)\,dy.$$

Suponga que $t_0 < s < u < t$. Por la fórmula de Chapman-Kolmogorov

$$\begin{aligned}
\Phi(s,x) \;&=\; \int_{-\infty}^{\infty} \varphi(y) \int_{-\infty}^{\infty} p(s,x\,|\,u,z)\,p(u,z\,|\,t,y)\,dz\,dy \\
&=\; \int_{-\infty}^{\infty} p(s,x\,|\,u,z) \int_{-\infty}^{\infty} \varphi(y)\,p(u,z\,|\,t,y)\,dy\,dz \\
&=\; \int_{-\infty}^{\infty} \Phi(u,z)\,p(s,x\,|\,u,z)\,dz.
\end{aligned}$$

Es claro que la función Φ tiene derivadas continuas Φ_s, Φ_x y Φ_{xx}. Para un número fijo u tenemos que

$$\Phi(u,z) - \Phi(u,x) \;=\; \Phi_x(u,x)(z-x) + \frac{1}{2}\Phi_{xx}(u,x)(z-x)^2 + O(z-x)^3.$$

La propiedad (1) implica que

$$\begin{aligned}\Phi(s,x) - \Phi(u,x) \;&=\; \int_{-\infty}^{\infty} \Big\{\Phi(u,z) - \Phi(u,x)\Big\}\, p(s,x\,|\,u,z)\,dz \\ &=\; \int_{|z-x|\le\epsilon} \Big\{\Phi(u,z) - \Phi(u,x)\Big\}\, p(s,x\,|\,u,z)\,dz + o(h)\end{aligned}$$

en donde hemos escrito $h = u - s$. Escribimos además

$$\Phi_x = \Phi_x(u,x), \qquad \Phi_{xx} = \Phi_{xx}(u,x), \qquad a = a(s,x) \qquad \text{y} \qquad b^2 = b^2(s,x).$$

La propiedad (3) implica que

$$\int_{|z-x|\le\epsilon} (z-x)^3\, p(s,x\,|\,u,z)\,dz \;\ll\; h\epsilon.$$

Por lo tanto

$$\begin{aligned}\Phi(s,x) - \Phi(u,x) \;&=\; \Phi_x \int_{|z-x|\le\epsilon} (z-x)\, p(s,x\,|\,u,z)\,dz \\ &\quad + \frac{\Phi_{xx}}{2} \int_{|z-x|\le\epsilon} (z-x)^2\, p(s,x\,|\,u,z)\,dz + o(h) + O(\epsilon h).\end{aligned}$$

Las propiedades (2) y (3) implican que

$$\Phi(s,x) - \Phi(u,x) \;=\; a\Phi_x h + \frac{b^2}{2}\Phi_{xx} h + o(h) + O(\epsilon h).$$

Al dividir entre $h = s - u$ tenemos

$$-\frac{\Phi(s,x) - \Phi(u,x)}{s-u} \;=\; a\Phi_x + \frac{b^2}{2}\Phi_{xx} + o(1) + O(\epsilon).$$

Haciendo que $s \to u$ obtenemos

$$-\Phi_s(u,x) \;=\; a\Phi_x(u,x) + \frac{b^2}{2}\Phi_{xx}(u,x) + O(\epsilon).$$

Ya que ϵ es positivo pero arbitrario, entonce el término $O(\epsilon)$ se puede omitir. Así se obtiene que

$$0 \;=\; \int_{-\infty}^{\infty} \varphi(y)\Big\{p_s + ap_x + \frac{b^2}{2}p_{xx}\Big\}\,dy.$$

Puesto que la función $\varphi(x)$ es arbitraria, entonces

$$-p_s(u,x) \;=\; ap_x(u,x) + \frac{b^2}{2}p_{xx}(u,x).$$

Esto termina la prueba. ∎

A continuación vamos a deducir la segunda de las ecuaciones de Kolmogorov.

Teorema 2. *Suponga que la función $p(s,x\,|\,t,y)$ satisface las propiedades (1), (2), (3). Suponga además que las derivadas*

$$\frac{\partial}{\partial t}p(s,x\,|\,t,y), \qquad \frac{\partial}{\partial y}\big[a(t,y)\,p(s,x\,|\,t,y)\big], \qquad \frac{\partial^2}{\partial y^2}\big[b^2(t,y)\,p(s,x\,|\,t,y)\big]$$

son uniformemente continuas. Se cumple entonces que

$$\frac{\partial}{\partial t}p(s,x\,|\,t,y) \;=\; -\frac{\partial}{\partial y}\big[a(t,y)\,p(s,x\,|\,t,y)\big] + \frac{1}{2}\frac{\partial^2}{\partial y^2}\big[b^2(t,y)\,p(s,x\,|\,t,y)\big].$$

Demostración. Sea $\varphi(x)$ una función con dos derivadas continuas y que se anula fuera de un intervalo finito, de modo que

$$\varphi(y) \;=\; \varphi(x) + (y-x)\varphi'(x) + \frac{1}{2}(y-x)^2\varphi''(x) + O(y-x)^3.$$

Se cumple entonces que

$$\lim_{h\to 0}\frac{1}{h}\Big\{\int\limits_{-\infty}^{\infty}\varphi(y)\,p(t,x\,|\,t+h,y)\,dy-\varphi(s)\Big\} = a(t,x)\,\varphi'(x)+\frac{1}{2}b^2(t,x)\,\varphi''(x). \qquad \mathbf{(5)}$$

Para ver que esta fórmula se cumple nótese que las propiedades (1) y (2) implican

$$\int\limits_{-\infty}^{\infty}(y-x)\,p(t,x\,|\,t+h,y)\,dy = \int\limits_{|y-x|\le\epsilon}(y-x)p(t,x\,|\,t+h,y)\,dy+o(h) = a(t,x)\,h+o(h)$$

para cada $\epsilon > 0$. Las propiedades (1) y (3) implican que

$$\int\limits_{-\infty}^{\infty}(y-x)^2\,p(t,x\,|\,t+h,y)\,dy \;=\; a(t,x)\,h+o(h).$$

Suponga que $\varphi(x)$ se anula fuera del intervalo $[a,b]$. Las propiedades (1) y (3) implican que

$$\int\limits_{a}^{b}(y-x)^3\,p(t,x\,|\,t+h,y)\,dy \;\ll\; \epsilon h+o(h).$$

Ahora es claro que (5) se verifica.

Sea $A=\dfrac{\partial}{\partial t}\int_{-\infty}^{\infty}\varphi(y)\,p(s,x\,|\,t,y)\,dy$. Entonces

$$A \;=\; \lim_{h\to 0}\frac{1}{h}\Big\{\int\limits_{-\infty}^{\infty}\varphi(y)\,p(s,x\,|\,t+h,y)\,dy-\int\limits_{-\infty}^{\infty}\varphi(z)\,p(s,x\,|\,t,z)\,dz\Big\}.$$

En vista de las ecuaciones de Chapman-Kolmogorov

$$A \;=\; \lim_{h\to 0}\frac{1}{h}\Big\{\int\limits_{-\infty}^{\infty}\varphi(y)\int\limits_{-\infty}^{\infty}p(s,x\,|\,t,z)\,p(t,z\,|\,t+h,y)\,dz\,dy-\int\limits_{-\infty}^{\infty}\varphi(z)\,p(s,x\,|\,t,z)\,dz\Big\}.$$

Al intercambiar el orden de las integrales

$$A \;=\; \int\limits_{-\infty}^{\infty}p(s,x\,|\,t,z)\lim_{h\to 0}\frac{1}{h}\Big\{\int\limits_{-\infty}^{\infty}\varphi(y)\,p(t,z\,|\,t+h,y)\,dy-\varphi(z)\Big\}\,dz.$$

Ya que (5) se cumple, entonces

$$\int_{-\infty}^{\infty} \varphi(y)\,\frac{\partial}{\partial t} p(s,x\,|\,t,y)\,dy \;=\; \int_{-\infty}^{\infty} p(s,x\,|\,t,z)\Big\{a(t,z)\,\varphi'(z) + \frac{1}{2} b^2(t,z)\,\varphi''(z)\Big\}\,dz.$$

Integrando por partes vemos que

$$\int_{-\infty}^{\infty} p(s,x\,|\,t,y)\,a(t,z)\,\varphi'(y)\,dy \;=\; -\int_{-\infty}^{\infty} \varphi(y)\,\frac{\partial}{\partial y}\Big[a(t,z)\,p(s,x\,|\,t,y)\Big]\,dy.$$

Se cumple también que

$$\int_{-\infty}^{\infty} p(s,x\,|\,t,y)\,b^2(t,z)\,\varphi''(y)\,dy \;=\; \int_{-\infty}^{\infty} \varphi(y)\,\frac{\partial^2}{\partial y^2}\Big[b^2(t,z)\,p(s,x\,|\,t,y)\Big]\,dy.$$

Resumimos lo anterior escribiendo

$$0 \;=\; \int_{-\infty}^{\infty} \varphi(y)\Big[\frac{\partial p}{\partial t} + \frac{\partial}{\partial y}(ap) - \frac{1}{2}\,\frac{\partial^2}{\partial y^2}(b^2 p)\Big]\,dy.$$

El teorema se cumple ya que $\varphi(x)$ es una función arbitraria. ∎

Capítulo 3

Ejemplos de Procesos de Difusión

§1. Difusión de un Contaminante.

Consideremos la difusión de un contaminante en un medio ambiente como el agua. Supongamos que tenemos un recipiente de longitud igual a L. Supongamos también que la altura y anchura del recipiente son muy pequeños en comparación con la longitud. Entonces nuestro problema se reduce a un problema difusión en una dimensión espacial.

Pensemos en una molécula de contaminante. Esta molécula es bombardeada por las moléculas del medio haciendo que en cada instante se mueva en una dirección aleatoria emulando un movimiento browniano. Si $u = u(x,t)$ es la densidad del contaminante en el punto x al tiempo t entonces u satisface la ecuación de difusión que consideramos en el capítulo anterior. Además, en el problema que nos ocupa tenemos un recipiente con una capacidad finita de contener agua. En particular, no hay flujos de agua ni de contaminate desde ni hacia el recipiente. Por lo tanto en nuestro problema se tienen las siguientes cuatro condiciones

$$u_t(x,t) = kf_{xx}(x,t), \qquad (\mathbf{1})$$

$$u(x,0) = g(x), \qquad (\mathbf{2})$$

$$u_x(0,t) = 0, \qquad (\mathbf{3})$$

$$u_x(L,t) = 0. \qquad (\mathbf{4})$$

Las condiciones (3) y (4) expresan el hecho de que no hay flujo en las fronteras del recipiente, mientras la condición (2) nos dice que la densidad del contaminante en el punto x al tiempo $t = 0$ es igual a $g(x)$. Notemos que la cantidad total de contaminante en el recipiente está dada por $\int_0^L g(x)\,dx$.

Ahora vamos a emplear el método de separación de variables para obtener una solución que satisfaga (3) y (4). Para esto proponemos una función de la forma $u(x,t) = X(x)T(t)$, donde X y T deben ser no triviales. Si $u = X(x)T(t)$ satisface (1) entonces tenemos

$$X(x)T'(t) \;=\; kX''(x)T(t).$$

Por la condición de no trivialidad podemos dividir entre $kX(x)T(t)$ y así vemos que

$$\frac{T'(t)}{kT(t)} \;=\; \frac{X''(x)}{X(x)}.$$

De aquí se obteniene que las dos funciones $T(t)$ y $X(x)$ tienen una valor constante $-\lambda$ en común. El parámetro λ se conoce como constante de separación. Si $u = X(x)T(t)$ satisface las condiciones de frontera (3) y (4), entonces tenemos que para todo $t > 0$ se cumple que $X'(0)T(t) = 0$. Ya que T no es identicamente igual a cero entonces se sigue que $X'(0) = 0$. De igual manera vemos que $X'(L) = 0$. Por lo tanto entonces X y T satisfacen los siguientes problemas homogéneos

$$\frac{d^2X(x)}{dx^2} + \lambda X(x) \;=\; 0, \qquad X'(0) = 0, \quad X'(L) = 0, \tag{5}$$

$$\frac{dT(t)}{dt} + \lambda k T(t) \;=\; 0. \tag{6}$$

Puesto que el problema (6) no tiene condiciones de frontera entonces tiene soluciones no triviales para cualquier valor de λ. Por otro lado, el problema de Sturm-Louville (5) tiene soluciones no triviales solamente cuando λ toma ciertos valores.

Si $\lambda = 0$ entonces la ecuación (6) toma la forma $X''(x) = 0$, cuya solución general es

$$X(x) \;=\; Ax + B \tag{7}$$

donde A y B son constantes. Tomando derivadas en (7) y aplicando las condiciones de frontera obtenemos que $X_0(x) = \frac{1}{2}$ resuelve el problema (5).

Sea $\lambda > 0$. Entonces podemos escribir $\lambda = \alpha^2$ con $\alpha > 0$. La solución general de (5) está dada por

$$X(x) \;=\; c_1 \cos(\alpha x) + c_2 \operatorname{sen}(\alpha x). \tag{8}$$

Tomando derivada en (8) y como α es positivo y además $X'(0) = 0$, entonces se obtiene que $c_2 = 0$. La otra condición implica $c_1 \alpha \sin \alpha L = 0$. Al ser $X(x)$ no trivial α debe ser una raíz de la ecuación $\operatorname{sen}(\alpha L) = 0$. Por lo tanto

$$\alpha \;=\; \frac{n\pi}{L} \qquad \text{donde} \qquad n \geq 1.$$

En resumen, cuando $\lambda > 0$ entonces

$$X_n(x) \;=\; c_1 \cos\left(\frac{n\pi x}{L}\right)$$

es solución de (5) para cualquier constante c_1.

Sea $\lambda < 0$. Entonces $\lambda = -\alpha^2$ con $\alpha > 0$. La solución de la ecuación en (5) está dada por

$$X(x) \;=\; c_1 e^{\alpha x} - c_2 e^{-\alpha x}. \tag{9}$$

Derivando (9) y aplicando la condición $X'(0) = 0$ se obtiene $c_2 = c_1$. Aplicando la otra condición de frontera

$$X(L) \;=\; c_1 \alpha (e^{\alpha L} + e^{-\alpha L}) \;=\; 2c_1 \alpha \cosh(\alpha L).$$

Como α y $\cosh \alpha L$ son distintos de cero, entonces $c_1 = 0$. En éste caso sólo se tiene la solución trivial.

Consideremos ahora la ecuación (6). Cuando $\lambda = 0$, tenemos a $T_0(t) = 1$ como solución. Cuando $\lambda = \left(\frac{n\pi}{L}\right)^2$ entonces

$$T_n(t) = e^{-\frac{(n\pi)^2 k}{L^2} t}$$

es solución de (6).

Cada uno de los productos ($n \geq 1$)

$$u_0(x,t) = X_0(x) T_0(t) = \frac{1}{2}, \tag{10}$$

$$u_n(x,t) = X_n(x) T_n(t) = \cos\left(\frac{n\pi x}{L}\right) e^{-\frac{n^2\pi^2 k}{L^2} t} \tag{11}$$

satisfacen las condiciones (3) y (4) y son solución a nuestro problema de difusión. Aplicando el principio de superposición, obtenemos que la serie

$$u(x,t) = \frac{c_0}{2} + \sum_{n=1}^{\infty} c_n e^{-\frac{(n\pi)^2 k}{L^2} t} \cos\left(\frac{n\pi x}{L}\right) \tag{12}$$

también es solución de dicho problema. Ya que la condición inicial (2) tiene lugar, entonces

$$g(x) = \frac{c_0}{2} + \sum_{n=1}^{\infty} c_n \cos\left(\frac{n\pi x}{L}\right).$$

De aquí se obtiene el valor de los coeficientes c_n. En efecto, c_n está dado por

$$c_n = \frac{2}{L} \int_0^L g(x) \cos\left(\frac{n\pi x}{L}\right) dx. \tag{13}$$

La solución de nuestro problema de difusión está dado por la serie (12) junto con la fórmula (13) para los coeficientes c_n.

§2. Proceso de Ornstein-Uhlenbeck.

En este apartado vamos a aplicar las ecuaciones de Kolmogorov para obtener la función de densidad condicional de un proceso de difusión sujeto a una fuerza que lo hace moverse hacia el origen como estado de equilibrio. Así entonces vamos a pedir que

$$\lim_{h\to 0}\frac{1}{h}\mathbf{E}\big[\xi(t+h)-\xi(t)\,\big|\,\xi(t)=y\big] \;=\; -\alpha y$$

en donde α es una constante. Esta relación nos dice que si al tiempo t el valor del proceso es $\xi(t)=y$, entonces hay una tendencia a moverse hacia el origen, es decir hacia el estado $\xi=0$. Por lo tanto podemos aplicar la segunda de las ecuaciones de Kolmogorov con

$$a(t,y) \;=\; -\alpha y \qquad\qquad \text{y} \qquad\qquad b^2(t,y) \;=\; 2\beta$$

en donde β es una constante.

Consideremos entonces la ecuación

$$\frac{\partial p}{\partial t} \;=\; \alpha\frac{\partial}{\partial y}(yp)+\beta\frac{\partial^2 p}{\partial y^2} \tag{14}$$

con las tres condiciones

$$\lim_{|y|\to\infty} yp(x\,|\,t,y) \;=\; 0, \qquad \lim_{|y|\to\infty}\frac{\partial}{\partial y}p(x\,|\,t,y) \;=\; 0, \qquad \lim_{t\to 0}p(x\,|\,t,y) \;=\; \delta(y-x)$$

en donde $\delta(x)$ es la delta de Dirac. Para resolver (14) sea

$$\phi(\theta) \;=\; \int_{-\infty}^{\infty} e^{i\theta y}p(x_0\,|\,t,y)\,dy.$$

Nótese primero que

$$\begin{aligned}
\int_{-\infty}^{\infty} e^{i\theta y}\frac{\partial}{\partial y}(yp)\,dy \;&=\; e^{i\theta y}yp\Big|_{-\infty}^{\infty} - i\theta\int_{-\infty}^{\infty} e^{i\theta y}yp\,dy \\
&=\; -\theta\frac{\partial}{\partial\theta}\int_{-\infty}^{\infty} e^{i\theta y}p\,dy \\
&=\; -\theta\frac{\partial\phi}{\partial\theta}.
\end{aligned}$$

Por otro lado tenemos que

$$\int_{-\infty}^{\infty} e^{i\theta y}\frac{\partial p}{\partial y}\,dy \;=\; e^{i\theta y}p\Big|_{-\infty}^{\infty} - i\theta\int_{-\infty}^{\infty} e^{i\theta y}p\,dy \;=\; -i\theta\phi(\theta).$$

De forma similar se obtiene

$$\int_{-\infty}^{\infty} e^{i\theta y}\frac{\partial^2 p}{\partial y^2}\,dy \;=\; e^{i\theta y}\frac{\partial p}{\partial y}\Big|_{-\infty}^{\infty} - i\theta\int_{-\infty}^{\infty} e^{i\theta y}\frac{\partial p}{\partial y}\,dy \;=\; -\theta^2\phi(\theta).$$

De la ecuación (14) ahora tenemos

$$A\frac{\partial\phi}{\partial t} + B\frac{\partial\phi}{\partial\theta} \;=\; C \quad\text{ en donde }\quad A = 1,\quad B = \alpha\theta,\quad C = -\beta\theta^2\phi. \qquad \mathbf{(15)}$$

Esta última es una ecuación diferencial parcial cuasilineal. Supongamos ahora que la ecuación

$$F(t,\theta,\phi) \;=\; c \qquad \mathbf{(16)}$$

define una solución de (15) en el sentido de que (16) determina a ϕ como función de t y θ de tal manera que (15) se satisface. Se cumple entonces que

$$\frac{\partial F}{\partial t} + \frac{\partial F}{\partial\phi}\frac{\partial\phi}{\partial t} \;=\; 0 \qquad \text{y} \qquad \frac{\partial F}{\partial\theta} + \frac{\partial F}{\partial\phi}\frac{\partial\phi}{\partial\theta} \;=\; 0.$$

Por lo tanto

$$\frac{\partial\phi}{\partial t} = -\frac{F_t}{F_\phi} \qquad \text{y} \qquad \frac{\partial\phi}{\partial\theta} = -\frac{F_\theta}{F_\phi}.$$

Al sustituir estas expresiones en (15) vemos que

$$AF_t + BF_\theta + CF_\phi = 0. \tag{\textbf{17}}$$

Esta última ecuación se puede escribir como

$$(A\mathbf{i} + B\mathbf{j} + C\mathbf{k}) \cdot \nabla F = 0.$$

Puesto que ∇F es un vector normal a la superficie $F = c$, entonces la ecuación (17) afirma que el vector $A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$ es perpendicular a la normal de la superficie $F = c$ y por lo tanto $A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$ yace en el plano tangente. Por lo tanto, en cualquier punto de la superficie $F = c$, el vector $A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$ es tangente a una curva contenida en $F = c$.

Se desprende de nuestra discusión que la ecuación (15) determina un campo vectorial $\mathbf{V} = A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$ en alguna región de $\mathbf{R}^3$. Si una partícula se mueve a partir de un punto inicial de tal manera que su dirección siempre coincide con $\mathbf{V}$, entonces se genera una curva. La curva así generada se llama *curva característica* de la ecuación (15).

Es fácil ver que cualquier superficie construida "pegando" un conjunto de curvas características tendrá la propiedad de que cada plano tangente contiene al vector $A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$.

Sea $\mathbf{r}$ el vector posición de un punto en una curva característica y sea s la longitud de arco de la curva. Entonces el vector tangente unitario está dado por

$$\frac{d\mathbf{r}}{ds} = \frac{dt}{ds}\mathbf{i} + \frac{d\theta}{ds}\mathbf{j} + \frac{d\phi}{ds}\mathbf{k}.$$

La condición de que este vector tenga la misma dirección que $A\mathbf{i}+B\mathbf{j}+C\mathbf{k}$ implica que existe un número μ tal que

$$A=\mu\frac{dt}{ds}, \qquad B=\mu\frac{d\theta}{ds}, \qquad C=\mu\frac{d\phi}{ds}.$$

Se obtiene entonces que

$$\frac{d\theta}{dt} \;=\; \frac{B}{A} \;=\; \alpha\theta \qquad \text{y} \qquad \frac{d\phi}{d\theta} \;=\; \frac{C}{B} \;=\; -\frac{\beta}{\alpha}\theta\phi.$$

Por lo tanto

$$\theta\, e^{-\alpha t} \;=\; c_1 \qquad \text{y} \qquad \phi e^{\frac{\beta}{2\alpha}\theta^2} \;=\; c_2.$$

Estas dos ecuaciones representan dos familias de superficies, tales que la intersección de una superficie en una familia con una superficie en la otra familia resulta en una curva característica. Si consideramos ahora una familia uniparamétrica de tales curvas características, entonces vemos que entre c_1 y c_2 existe una relación funcional. Por lo tanto, existe una función ψ tal que

$$\phi e^{\frac{\beta}{2\alpha}\theta^2} \;=\; \psi(\theta\, e^{-\alpha t}).$$

Ya que $p(x\,|t,y)\to\delta(y-x)$ cuando $t\to 0$, entonces

$$e^{i\theta x} \;=\; \lim_{t\to 0}\phi(\theta) \;=\; \psi(\theta)\, e^{-\frac{\beta}{2\alpha}\theta^2}.$$

Se deduce entonces que

$$\phi(\theta) \;=\; \exp\Big\{ i\theta x\, e^{-\alpha t} - \frac{\beta}{2\alpha}\theta^2(1-e^{-2\alpha t})\Big\}.$$

Se reconoce esta como la transformada de Fourier de la densidad

$$p(x\,|t,y) \;=\; \frac{1}{\sigma(t)\sqrt{2\pi}}\exp\Big\{ -\frac{1}{2\sigma^2(t)}\big(y-m(t)\big)^2\Big\}$$

en donde

$$m(t) \;=\; x\,e^{-\alpha t} \qquad\qquad \text{y} \qquad\qquad \sigma^2(t) \;=\; \frac{\beta}{\alpha}(1 - e^{-2\alpha t}).$$

La función obtenida $p(x\,|t, y)$ caracteriza completamente al proceso de Ornstein-Uhlenbeck.

Glosario

Lipschitz. Una función $f : [a, b] \to \mathbf{R}$ es de tipo Lipschitz, si existen dos constantes positivas M y α tales que

$$|f(x) - f(y)| \ \leq \ M|x - y|^{\alpha}$$

se cumple siempre que $a \leq x < y \leq b$.

Notación O. Sea $a \in \mathbf{R} \cup \big\{ \pm \infty \big\}$. La notación $f(x) = O\big(g(x)\big)$ cuando $x \to a$ significa lo siguiente.

(a) Que $g(x) > 0$.

(b) Que existen dos constantes $A > 0$ y $B > 0$ tales que $|f(x)| \leq Ag(x)$ se cumple para todo x tal que $|x - a| \leq B$.

Por lo general se omite escribir "$x \to a$" ya que del contexto es claro hacia donde tiende x. Se dice que O es el símbolo de Landau (Edmundo). La constante A se conoce como la constante implicada en la notación de Landau.

Notación o. Escribimos $f(x) = o\big(g(x)\big)$ cuando $x \to a$ siempre que $f(x)/g(x) \to 0$ cuando $x \to a$. Se dice que f es o chica de g. Esta Notación también se debe a Edmundo Landau.

Notación $\ll$. Escribimos $f \ll g$ en lugar de escribir $f = O(g)$. Esta notación se debe a I.M. Vinogradov.

Debe notarse que algunos autores escriben $f \ll g$ en lugar de escribir $f = o(g)$. Sin embargo este no es el sentido de la notación de Vinogradov. Los símbolos "o" y "$\ll$" no son sinónimos. Los símbolos "O" y "$\ll$" sí son sinónimos.

Conclusiones

Gracias al trabajo de redacción de esta tesis, los participantes en la redacción de la misma aprendieron sobre los procesos de difusión. Este conocimiento adquirido sobre procesos de difusión les permitirá abordar en el futuro distintos problemas sobre la explotación óptima de recursos naturales así como también problemas de tipo financiero. En particular, se aprendió el significado de las dos ecuaciones de Kolmogorov.

Como tema de trabajo posterior queda el abordar los procesos de difusión desde el punto de vista de la teoría de integración de K. Ito. Ver [9].

Bibliografía

[1] **Clark, C.W.** *Mathematical bioeconomics: the optimal management of renewable resources.* Wiley-Interscience, 1976.

[2] **Churchill, Ruel V.** *Fourier Series and Boundary Value Problems.* McGraw-Hill Book Company, Inc., New York and London, 1941.

[3] **Feller, W.** *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons, Inc., 1968.

[4] **Hildebrand, F. B.** *Advanced Calculus for Engineers.* Prentice-Hall, Inc., New York, N. Y., 1949.

[5] **Lesigne, Emmanuel** *Heads or tails. An introduction to limit theorems in probability.* Student Mathematical Library, 28. American Mathematical Society, Providence, RI, 2005.

[6] **Prabhu, N. U.** *Stochastic processes. Basic theory and its applications.* The Macmillan Co., New York; Collier-Macmillan Ltd., London 1965.

[7] **Proschan, Michael A.** *The normal approximation to the binomial.* Amer. Statist. 62 (2008), no. 1, 62-63.

[8] **Rozanov, Y. A.** *Introduction to random processes.* Springer Series in Soviet Mathematics. Springer-Verlag, Berlin, 1987.

[9] **Shreve, S.E.** *Stochastic calculus for finance. II. Continuous-time models.* Springer Finance. Springer-Verlag, 2004.

[**10**] **Stirzaker, D.** *Stochastic processes and models.* Oxford University Press, 2005.