SAHUARUS
Revista Electrónica de Matemáticas,
Departamento de Matemáticas, Universidad de Sonora,
Vol. 3, No. 1. Octubre, 2018, pp. 1–17.

Análisis de la difusión de Feller vía el cálculo estocástico

Dorilian García Cerino†
Aroldo Pérez Pérez‡

Universidad Juárez Autónoma de Tabasco
e-mail: †w18dori@hotmail.com, ‡aroldopz2@gmail.com

Resumen

En este artículo de difusión describiremos un crecimiento poblacional continuo en el que los individuos se multiplican y son estadísticamente independientes entre sí con una tasa de crecimiento relativa o per cápita constante. En 1951 William Feller realizó el análisis encontrando la densidad de transición del proceso resolviendo la ecuación de Fokker-Plank. En este trabajo se desarrolla el tratamiento dado en el libro de Fima C. Klebaner titulado "Introduction to stochastic calculus with applications", en el cual la teoría del cálculo estocástico es usada para obtener información del proceso directamente de su ecuación diferencial estocástica correspondiente.

1 Introducción

Después de las definiciones de la derivada de una función y de la integral de Riemann, una enorme gama de problemas de las ciencias naturales, sociales y biológicas quedaron bajo el dominio de la teoría de funciones de una variable real. Las principales componentes de esta teoría son el uso de la diferenciación para describir razones de cambio, el uso de la integración para pasar al límite en la aproximación por sumas y el teorema fundamental del cálculo, el cual relaciona ambos conceptos y hace más accesible los cómputos de estos límites de sumas (integrales). Esta teoría dió lugar al concepto de ecuaciones diferenciales ordinarias, y la aplicación de estas ecuaciones para modelar fenómenos del mundo real revela el enorme potencial del cálculo.

El cálculo estocástico surge de la necesidad de dar sentido a ecuaciones diferenciales ordinarias que, con la finalidad de obtener modelos matemáticos más realistas del fenómeno en cuestión, involucran procesos estocásticos. A manera de ilustración consideremos el siguiente modelo sencillo de crecimiento de población

$$dX_t = \alpha(t)X_t dt, \qquad X_0 = x_0,$$

donde X_t es el tamaño de la población al tiempo t y $\alpha(t)$ es la tasa relativa de crecimiento al tiempo t. Podría ocurrir que dicha tasa de crecimiento no esté completamente determinada debido a su dependencia de algunos fenómenos ambientales aleatorios, de manera que tendremos

$$\alpha(t) = \mu(t) + R_t,$$

donde la función $\mu(t)$ es determinista y se desconoce el comportamiento exacto del “ruido”, R_t. Formalmente, el proceso R_t puede verse como la derivada de un proceso estocástico W_t, es decir, $R_t = \frac{dW_t}{dt}$, y así, formalmente, obtenemos la ecuación diferencial estocástica

$$dX_t = \mu(t)X_t dt + X_t dW_t.$$

Sin embargo, la mayoría de los procesos estocásticos importantes, tales como el proceso de Wiener, no son diferenciables (véase por ejemplo, [11] pp. 110-111). Por tal motivo, el cálculo estocástico toma el camino opuesto al del cálculo ordinario; primero se define la integral estocástica

$$Y_t = \int_0^t X_s dW_s$$

de cierto proceso estocástico $\{X_t\}_{t\geq 0}$ con respecto a cierto proceso estocástico $\{W_t\}_{t\geq 0}$ y luego se le da sentido a la diferencial estocástica a través del “teorema” fundamental del cálculo. Este “teorema” es en realidad una definición en el cálculo estocástico, debido a que la diferencial no tiene más significado que el que se le asigna mediante una integral estocástica, esto es, si $\{X_t\}_{t\geq 0}$ es un proceso estocástico que satisface la ecuación integral estocástica

$$X_t = X_0 + \int_0^t \mu(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dW_s,$$

donde $\mu(t, x)$ y $\sigma(t, x)$ son funciones dadas, se escribe

$$dX_t = \mu(t, X_t)\, dt + \sigma(t, X_t)\, dW_t. \tag{1}$$

Existen muchos textos que consideran el caso donde $\{W_t\}_{t\geq 0}$ es una semimartingala (para su definición, véase por ejemplo, [12] p. 211), sin embargo, en este artículo de difusión nos restringimos solo al caso en que $\{W_t\}_{t\geq 0}$ es un proceso de Wiener.

Las ecuaciones diferenciales estocásticas (1) son el análogo probabilista de las ecuaciones diferenciales deterministas, y como se ejemplificó arriba, aparecen cuando se presenta aleatoriedad en los coeficientes de las ecuaciones deterministas. Hoy en día, la teoría de las ecuaciones diferenciales estocásticas es un tema bastante desarrollado de la teoría de los procesos estocásticos, con aplicaciones muy importantes en varios campos del conocimiento tales como la economía, ingeniería, física, medicina y biología. En particular, en la biología el cálculo estocástico se usa para modelar los efectos de la variabilidad estocástica en la reproducción y evolución genética de poblaciones (véase por ejemplo, [4] y [12]). En este artículo abordaremos la **difusión de ramificación de Feller**, modelo para el crecimiento de cierto tipo de poblaciones, que como veremos tiene un crecimiento medio exponencial; y encontraremos además su probabilidad de extinción. Este tipo de procesos, aparte de ser de interés por sí mismos, surgen también como límite de sucesiones de procesos de ramificación simples cuando el tamaño de la población crece a infinito y la escala de tiempo se comprime de una manera específica (véase por ejemplo, [2] p. 262).

2 Algunos conceptos y resultados sobre funciones y variables aleatorias

Recordemos que una partición $\mathcal{P}$ de $[a,b]$ es un conjunto finito de puntos de $[a,b]$, digamos, $\mathcal{P}=\{x_0,x_1,\ldots,x_n\}$ tal que $a=x_0<x_1<\cdots<x_n=b$. Al número $|\mathcal{P}|=\max_{1\leq i\leq n}(x_i-x_{i-1})$ le llamaremos la **malla** de la partición $\mathcal{P}$.

Sean $f,F:[a,b]\to\mathbb{R}$ funciones acotadas. Para una partición $\mathcal{P}=\{x_0,x_1,\ldots,x_n\}$ de $[a,b]$ y una elección de puntos $y_i\in[x_{i-1},x_i]$, $i=1,2,...n$, se define una **suma de Riemann-Stieltjes** de f con respecto a F por

$$S(f,F,\mathcal{P}):=\sum_{i=1}^{n}f(y_i)[F(x_i)-F(x_{i-1})].$$

Definición 2.1 *Se dice que f es **Riemann-Stieltjes integrable** respecto de F en $[a,b]$ si existe un número real, denotado por $\int_a^b fdF$ o por $\int_a^b f(x)dF(x)$, con la propiedad de que, para cada $\epsilon>0$, existe $\delta>0$ tal que $|S(f,F,\mathcal{P})-\int_a^b fdF|<\epsilon$ para cualquier partición $\mathcal{P}=\{x_0,x_1,\ldots,x_n\}$ de $[a,b]$ con $|\mathcal{P}|<\delta$ y para cualquier elección de puntos $y_i\in[x_{i-1},x_i]$, $i=1,2,...n$. Si existe, a la cantidad $\int_a^b fdF$ se le llama la **integral de Riemann- Stieltjes** de f con respecto a F en $[a,b]$, y en este caso se escribe*

$$\int_a^b fdF=\lim_{|\mathcal{P}|\to 0}S(f,F,\mathcal{P}).$$

Observemos que en el caso particular en que $F(x)=x$ se tiene la definición de que una función sea Riemann integrable.

El siguiente teorema nos brinda una condición suficiente para reducir una integral de Riemann-Stieltjes a una integral de Riemann. Su demostración puede verse en [5] p. 232.

Teorema 2.2 *Sean $f,F:[a,b]\to\mathbb{R}$ funciones acotadas. Si F' existe y es Riemann integrable en $[a,b]$, entonces f es Riemann-Stieltjes integrable con respecto a F en $[a,b]$ si, y sólo si, fF' es Riemann integrable en $[a,b]$. En cuyo caso,*

$$\int_a^b f(x)dF(x)=\int_a^b f(x)F'(x)dx.$$

Los siguientes tres conceptos serán requeridos como condiciones sobre los coeficientes en los resultados concernientes a la existencia y unicidad de soluciones de ecuaciones diferenciales estocásticas (véase los teoremas 4.9 y 4.10).

Definición 2.3 *Una función $f:\mathbb{R}\to\mathbb{R}$ satisface la **condición de Hölder** de orden $\alpha\in(0,1]$ si existe $K>0$ tal que para todo $x,y\in\mathbb{R}$*

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

*Cuando f satisface la condición de Hölder de orden $\alpha=1$, se dice que f satisface la **condición de Lipschitz** o que es una **función de Lipschitz**.*

Definición 2.4 *Una función $f : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ es **localmente Lipschitz** en x uniformemente en t, si para cada $N > 0$ existe $K(N) > 0$ tal que para todo $|x|, |y| \leq N$ y todo $t \geq 0$*

$$|f(t,x) - f(t,y)| \leq K|x-y|.$$

Definición 2.5 *Una función $f : \mathbb{R} \to \mathbb{R}$ satisface la **condición de crecimiento lineal** si existe $K > 0$ tal que para todo $x \in \mathbb{R}$*

$$|f(x)| \leq K(|x|+1).$$

La prueba de la desigualdad que exponemos a continuación puede verse en [13] p. 188.

Teorema 2.6 *(**Desigualdad de Gronwall**) Sean $f, g : [0,T] \to [0,\infty)$ funciones y $\alpha > 0$ tales que*

$$f(t) \leq g(t) + \alpha \int_0^t f(s)ds$$

para todo $t \in [0,T]$. Entonces para $t \in [0,T]$

$$f(t) \leq g(t) + \alpha \int_0^t g(s)e^{\alpha(t-s)}ds.$$

Dado un espacio de probabilidad $(\Omega, \mathcal{F}, P)$, diremos que una propiedad es válida **casi seguramente** (c.s.), cuando exista $A \in \mathcal{F}$ tal que dicha propiedad se valga en $\Omega \backslash A$ y $P(A) = 0$. De esta manera, diremos que una sucesión de variables aleatorias X_n **converge casi seguramente** a una variable aleatoria X si existe $A \in \mathcal{F}$ con $P(A) = 0$ tal que $X_n(\omega) \to X(\omega)$ para todo $\omega \in \Omega \backslash A$. En este caso escribiremos $X_n \to X$ c.s.

Otros tipos de convergencia de sucesiones de variables aleatorias son dadas en la siguiente definición, donde F_X denota a la función de distribución de la variable aleatoria X.

Definición 2.7 *Sean $X, X_1, X_2, \ldots$ variables aleatorias.*

*a) X_n **converge en probabilidad** a X, denotado por $X_n \xrightarrow{P} X$, si para cualquier $\epsilon > 0$,*

$$P(|X_n - X| > \epsilon) \to 0 \quad \text{cuando} \quad n \to \infty.$$

*b) X_n **converge en distribución** a X, denotado por $X_n \xrightarrow{d} X$, si para cualquier punto de continuidad x de F_X,*

$$F_{X_n}(x) \to F_X(x) \quad \text{cuando} \quad n \to \infty.$$

*c) X_n **converge en** L^r a X, $r \geq 1$, denotado por $X_n \xrightarrow{r} X$, si*

$$E(|X_n - X|^r) \to 0 \quad \text{cuando} \quad n \to \infty.$$

Es conocido (véase por ejemplo, las secciones 2.4, 2.5, 2.8 y 7.1 de [1]) que $X_n \xrightarrow{r} X \Rightarrow X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$. También, $X_n \to X$ c.s. $\Rightarrow X_n \xrightarrow{P} X$.

Tres importantes resultados bien conocidos en la teoría de integración de Lebesgue son los siguientes. La prueba de cada uno de estos puede verse en la sección 1.6 de [1].

Teorema 2.8 *Sean $Y, X, X_1, X_2, \ldots$ variables aleatorias. Se tiene que*

*(**Teorema de convergencia monótona**) Si $Y \leq X_n \leq X_{n+1}$ c.s. para todo $n = 1, 2, \ldots$, $E|Y| < \infty$ y $X_n \to X$ c.s., entonces*

$$\lim_{n\to\infty} EX_n = EX.$$

*(**Lema de Fatou**) Si $X_n \geq Y$ c.s. para todo $n = 1, 2, \ldots$ y $EY > -\infty$, entonces*

$$\liminf_{n\to\infty} EX_n \geq E\left(\liminf_{n\to\infty} X_n\right).$$

*(**Teorema convergencia dominada**) Si $|X_n| \leq Y$ c.s. para todo $n = 1, 2, \ldots$, $E|Y| < \infty$ y $X_n \to X$ c.s., entonces EX existe y*

$$\lim_{n\to\infty} EX_n = EX.$$

3 Algunos preliminares sobre procesos estocásticos

Un **proceso estocástico a tiempo continuo** es una colección $X = \{X_t\}_{t\geq 0}$ de variables aleatorias definidas sobre un espacio de probabilidad $(\Omega, \mathcal{F}, P)$, las cuales toman valores en un espacio medible, llamado **espacio de estados**. Para nuestros propósitos el espacio de estados es $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, donde $\mathcal{B}(\mathbb{R})$ es la σ-álgebra de Borel en $\mathbb{R}$. El índice $t \in [0, \infty)$ es interpretado como el tiempo. Para cada $\omega \in \Omega$ fijo, la función $t \mapsto X_t(\omega)$, que mapea $[0, \infty)$ en $\mathbb{R}$, es una **trayectoria** o **realización** del proceso X. Si las trayectorias son continuas, se dice que el proceso estocástico X es continuo.

Necesitaremos, más adelante, la siguiente versión simplificada del teorema de Fubini (para una versión más general véase [1] pp. 105-106).

Teorema 3.1 *(**Teorema de Fubini**) Sea $\{X_t\}_{t\geq 0}$ un proceso estocástico continuo. Si I es un subintervalo de $[0, \infty)$ tal que*

$$\int_I E\,|X_t|\,dt < \infty,$$

entonces las trayectorias de X son integrables y

$$\int_I EX_t dt = E\int_I X_t dt.$$

Dado un espacio de probabilidad $(\Omega, \mathcal{F}, P)$, a una colección $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$ de sub-σ-álgebras de $\mathcal{F}$ se le llama una **filtración**, si $\mathcal{F}_s \subset \mathcal{F}_t$ para $s \leq t$. En este caso a $(\Omega, \mathcal{F}, \mathbb{F}, P)$ se le llama **espacio de probabilidad filtrado**.

Definición 3.2 *Un proceso estocástico $X = \{X_t\}_{t\geq 0}$ es **adaptado** a la filtración $\mathbb{F}$ si, para cada $t \geq 0$, X_t es $\mathcal{F}_t$-medible.*

La filtración más pequeña respecto a la cual un proceso $X = \{X_t\}_{t\geq 0}$ es adaptado es la generada por sí mismo, $\{\mathcal{F}_t^X\}_{t\geq 0}$, dada por $\mathcal{F}_t^X = \sigma(X_s : s \leq t)$. Esta filtración es conocida como la **filtración canónica** de X y se interpreta como la información generada por el proceso X.

Definición 3.3 *Sean $X = \{X_t\}_{t\geq 0}$ y $Y = \{Y_t\}_{t\geq 0}$ procesos estocásticos.*

*a) La **variación cuadrática** de X sobre $[0, t]$ se define como el límite en probabilidad (cuando exista)*

$$[X, X]_t = \lim \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})^2,$$

donde el límite es tomado sobre las particiones $\mathcal{P} = \{t_0, t_1, ..., t_n\}$ de $[0, t]$ con $|\mathcal{P}| \to 0$.

*b) La **covariación cuadrática** de X y Y sobre $[0, t]$ se define como*

$$[X, Y]_t = \frac{1}{2}([X + Y, X + Y]_t - [X, X]_t - [Y, Y]_t).$$

Definición 3.4 *Una variable aleatoria $\tau : \Omega \to [0, \infty]$ es un **tiempo de paro** respecto a la filtración $\{\mathcal{F}_t\}_{t\geq 0}$ si, para cada $t \geq 0$, $\{\tau \leq t\} \in \mathcal{F}_t$.*

La demostración del siguiente resultado puede verse en [3] p. 44.

Teorema 3.5 *Si X es un proceso estocástico continuo, adaptado a una filtración $\mathbb{F}$ y $A \in \mathcal{B}(\mathbb{R})$, entonces la variable aleatoria*

$$\tau_A = \inf \{t > 0 : X_t \in A\},$$

donde convenimos que $\inf \emptyset = \infty$, es un tiempo de paro.

Definición 3.6 *Sea τ un tiempo de paro respecto a la filtración $\{\mathcal{F}_t\}_{t\geq 0}$. Los eventos observados antes o en el tiempo τ son descritos por la σ-álgebra*

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ para todo } t \geq 0\}.$$

Definición 3.7 *Sean τ un tiempo de paro y $X = \{X_t\}_{t\geq 0}$ un proceso estocástico.*

a) Se define la variable aleatoria X_τ sobre el evento $\{\tau < \infty\}$ como $X_\tau(\omega) = X_{\tau(\omega)}(\omega)$.

*b) El **proceso parado** en τ, $X^\tau = \{X_{t\wedge\tau}\}_{t\geq 0}$, se define por $X_{t\wedge\tau}(\omega) = X_{t\wedge\tau(\omega)}(\omega)$.*

Observemos que el proceso parado está bien definido puesto que $t \wedge \tau := \min\{t, \tau\} < \infty$.

Definición 3.8 *Un proceso estocástico $X = \{X_t\}_{t\geq 0}$ adaptado a una filtración $\{\mathcal{F}_t\}_{t\geq 0}$ con $E|X_t| < \infty$ para todo $t \geq 0$ es una **martingala** si, para cualesquiera $0 \leq s < t$,*

$$E(X_t|\mathcal{F}_s) = X_s \ \ c.s.$$

El siguiente resultado nos proporciona una condición suficiente para que una martingala converja. Su demostración puede ser consultada en [12] p. 187.

Teorema 3.9 *Si $X = \{X_t\}_{t\geq 0}$ es una **martingala cuadrado integrable**, es decir, $\sup_{t\geq 0} E(X_t^2) < \infty$, entonces $X_\infty = \lim_{t\to\infty} X_t$ existe c.s. y $E|X_\infty| < \infty$.*

Definición 3.10 *Un proceso estocástico $X = \{X_t\}_{t\geq 0}$ adaptado a una filtración $\mathbb{F}$ es una **martingala local** si existe una sucesión creciente de tiempos de paro τ_n, llamada **sucesión localizante**, tal que $\tau_n \uparrow \infty$ cuando $n \to \infty$ y para cada n, $X^{\tau_n} = \{X_{t\wedge\tau_n}\}_{t\geq 0}$ es una martingala.*

Es fácil ver que cualquier martingala es una martingala local. El teorema que exponemos enseguida nos da una condición suficiente para que el recíproco de este hecho se cumpla. Su prueba puede verse en [12] p. 202.

Teorema 3.11 *Sea X una martingala local con $X_0 = 0$. Si $E[X,X]_t < \infty$ para todo $t \geq 0$, entonces X es una martingala. Si $\sup_{t\geq 0} E[X,X]_t < \infty$, entonces X es una martingala cuadrado integrable.*

Definición 3.12 *Sea $X = \{X_t\}_{t\geq 0}$ un proceso estocástico. Si para todo $y \in \mathbb{R}$ y para cualesquiera $0 \leq s \leq t$ se satisface*

$$P(X_t \leq y|\mathcal{F}_s^X) = P(X_t \leq y|X_s) \quad c.s., \tag{2}$$

*se dice que X es un **proceso de Markov**.*

La condición (2), conocida como la **propiedad de Markov**, nos dice intuitivamente que el comportamiento de X en el futuro depende únicamente de su estado presente y no de su comportamiento en el pasado.

Definición 3.13 *Un proceso de Markov $X = \{X_t\}_{t\geq 0}$ es un **proceso de Markov fuerte** si, para cualquier tiempo de paro finito τ respecto a $\{\mathcal{F}_t^X\}_{t\geq 0}$ y para cualesquiera $y \in \mathbb{R}$ y $t \geq 0$ se cumple*

$$P(X_{\tau+t} \leq y|\mathcal{F}_\tau) = P(X_{\tau+t} \leq y|X_\tau) \quad c.s.$$

Definición 3.14 *Un proceso estocástico* $W = \{W_t\}_{t\geq 0}$ *es un* ***proceso de Wiener*** *o* ***movimiento browniano*** *si*

a) $W_0 = 0$ *c.s.*

b) Tiene incrementos independientes, esto es, para cualesquiera $0 \leq t_1 < t_2 < \cdots < t_n$*, las variables aleatorias* W_{t_1}*,* $W_{t_2} - W_{t_1}$*,...,*$W_{t_n} - W_{t_{n-1}}$ *son independientes.*

c) Posee incrementos gaussianos, es decir, para cualesquiera $0 \leq s < t$*,* $W_t - W_s \sim N(0, t-s)$*.*

d) Tiene trayectorias continuas c.s.

4 Algunos preliminares del cálculo estocástico

Denotaremos por 1_A a la función indicadora de un conjunto A, es decir, $1_A(x) = 1$ si $x \in A$ y $1_A(x) = 0$ si $x \notin A$.

Definición 4.1 *Sea* W *un proceso de Wiener. Un proceso estocástico* $X = \{X_t\}_{0\leq t\leq T}$ *es un* ***proceso simple adaptado*** *si existen tiempos* $0 = t_0 < t_1 < \cdots < t_n = T$ *y variables aleatorias* $\xi_0, \xi_1, \ldots, \xi_{n-1}$ *con* ξ_0 *constante,* ξ_i $\mathcal{F}_{t_i}^W$*-medible y* $E\left(\xi_i^2\right) < \infty$*,* $i = 0, 1, \ldots, n-1$*; tales que*

$$X_t = \xi_0 1_{\{0\}}(t) + \sum_{i=0}^{n-1} \xi_i 1_{(t_i, t_{i+1}]}(t). \tag{3}$$

Definición 4.2 *La* ***integral de Itô*** *del proceso simple* X *dado por (3) con respecto al proceso de Wiener* W *se define como*

$$\int_0^T X_t dW_t = \sum_{i=1}^{n} \xi_{i-1}(W_{t_i} - W_{t_{i-1}}).$$

Sea X^n una sucesión de procesos simples adaptados convergentes en probabilidad a un proceso X. Entonces bajo ciertas condiciones, la sucesión de sus integrales $\int_0^T X_t^n dW_t$ también converge en probabilidad a una variable aleatoria J, la cual es tomada como la integral de Itô $\int_0^T X_t dW_t$.

Teorema 4.3 *Si* X *es un proceso continuo y adaptado, entonces la integral de Itô* $\int_0^T X_t dW_t$ *está bien definida. Más aún,* $Y_t = \int_0^t X_s dW_s$ *está bien definida para todo* $t \leq T$*. Además, el proceso* $Y = \{Y_t\}_{0\leq t\leq T}$ *es una martingala local continua con variación cuadrática*

$$[Y, Y]_t = \int_0^t X_s^2 ds.$$

La prueba del teorema 4.3 puede ser consultada en [13] pp. 70-73 y en [6] pp. 77-79.

Definición 4.4 *Un proceso estocástico X es un* ***proceso de Itô*** *si*

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s, \quad 0 \le t \le T, \tag{4}$$

donde X_0 es $\mathcal{F}_0^W$-medible, los procesos μ y σ son adaptados a $\mathbb{F}^W$ con $\int_0^T |\mu_s|\, ds < \infty$ y $\int_0^T \sigma_s^2 ds < \infty$. Si el proceso X satisface (4), se dice que tiene ***diferencial estocástica***

$$dX_t = \mu_t dt + \sigma_t dW_t, \quad 0 \le t \le T.$$

Definición 4.5 *Si X es un proceso de Itô que satisface (4) y Y es un proceso adaptado que satisface $\int_0^t |Y_s \mu_s|\, ds < \infty$ y $\int_0^t Y_s^2 \sigma_s^2 ds < \infty$, entonces $\int_0^t Y_s dX_s$ se define por*

$$\int_0^t Y_s dX_s = \int_0^t Y_s \mu_s ds + \int_0^t Y_s \sigma_s dW_s. \tag{5}$$

Las demostraciones de los siguientes dos teoremas pueden ser consultadas, por ejemplo, en la sección 7 del capítulo II de [15].

Teorema 4.6 ***(Fórmula de Itô)*** *Sean X un proceso de Itô que satisface (4) y $f : \mathbb{R} \to \mathbb{R}$ una función dos veces continuamente diferenciable. Entonces $f(X)$ es también un proceso de Itô con la representación*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) dX_s + \frac{1}{2}\int_0^t f''(X_s)\sigma_s^2 ds. \tag{6}$$

La representación diferencial de (6) es

$$df(X_t) = \left(f'(X_t)\mu_t + \frac{1}{2} f''(X_t)\sigma_t^2 \right) dt + f'(X_t)\sigma_t dW_t.$$

Sean X y Y procesos de Itô que satisfacen

$$dX_t = \mu_t^X dt + \sigma_t^X dW_t \tag{7}$$

$$dY_t = \mu_t^Y dt + \sigma_t^Y dW_t. \tag{8}$$

Teorema 4.7 ***(Fórmula de integración por partes)*** *Sean X y Y procesos de Itô que satisfacen (7) y (8), entonces,*

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \int_0^t \sigma_s^X \sigma_s^Y ds. \tag{9}$$

La representación diferencial de (9) es

$$d\left(X_t Y_t\right) = X_t dY_t + Y_t dX_t + \sigma_t^X \sigma_t^Y dt.$$

Una **ecuación diferencial estocástica** conducida por el proceso de Wiener, W, tiene la forma

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \tag{10}$$

donde las funciones $\mu(t,x)$ y $\sigma(t,x)$, llamadas **deriva** y **coeficiente de difusión** respectivamente, son conocidas y el proceso estocástico X es desconocido.

Definición 4.8 *Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad, y sea W un proceso de Wiener sobre $(\Omega, \mathcal{F}, P)$. Un proceso estocástico X adaptado a la filtracción $\{\mathcal{F}_t^W\}_{t\geq 0}$ es una* ***solución fuerte*** *de la ecuación diferencial estocástica (10) si para todo $t > 0$ las integrales $\int_0^t \mu(s, X_s)ds$ y $\int_0^t \sigma(s, X_s)dW_s$ existen, y*

$$X_t = X_0 + \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s.$$

A una solución fuerte X de (10) se le conoce también como **difusión de Itô**.

Así como en el caso de ecuaciones diferenciales deterministas, en el caso estocástico también existen teoremas para la existencia y unicidad de soluciones fuertes de ecuaciones diferenciales estocásticas. La demostración del siguiente teorema de existencia y unicidad puede ser consultada en [8] pp. 104-107.

Teorema 4.9 ***(Existencia y unicidad)*** *Si $\mu(t,x)$ y $\sigma(t,x)$ son localmente de Lipschitz en x uniformemente en t, y satisfacen la condición de crecimiento lineal en x, entonces existe una única solución fuerte X de la ecuación diferencial estocástica (10). Además, X es continuo.*

Se conoce también que las difusiones son procesos de Markov fuertes (véase por ejemplo, [14] pp. 20-21).

El siguiente resultado, específico para ecuaciones diferenciales estocásticas de dimensión uno, requiere condiciones menos estrictas para garantizar la existencia y unicidad de ecuaciones diferenciales estocásticas, el cual exponemos en su versión más simple, esto es, para ecuaciones diferenciales estocásticas con coeficientes homogéneos en el tiempo. Su demostración puede verse en [16] p. 265.

Teorema 4.10 ***(Yamada-Watanabe)*** *Si $\mu(x)$ satisface la condición de Lipschitz y $\sigma(x)$ satisface una condición de Hölder de orden $\alpha \geq 1/2$, entonces existe una única solución fuerte para la ecuación*

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t. \tag{11}$$

Otro concepto de solución (que no trabajaremos aquí) de una ecuación diferencial estocástica es el de **solución débil**, el cual permite dar significado a las ecuaciones diferenciales estocásticas que no tienen soluciones fuertes. Las soluciones débiles son soluciones en distribución, pueden ser definidas sobre algún espacio de probabilidad filtrado y existir bajo condiciones más débiles sobre los coeficientes de la ecuación (véase por ejemplo, la sección 5.3 de [11]).

Si las derivadas parciales $\frac{\partial}{\partial x}\mu(t,x)$, $\frac{\partial}{\partial x}\sigma(t,x)$ y $\frac{\partial^2}{\partial x^2}\sigma(t,x)$ existen, es conocido (véase [9] p. 102) que si la probabilidad de transición $P(s,x,t,y) = P\left(X_t \leq y | X_s = x\right)$ de la difusión X tiene una densidad de transición $p(s,x,t,y)$, es decir, $P\left(X_t \leq y | X_s = x\right) = \int_{-\infty}^{y} p(s,x,t,z)dz$, tal que $\frac{\partial}{\partial t}p(s,x,t,y)$, $\frac{\partial}{\partial y}p(s,x,t,y)$ y $\frac{\partial^2}{\partial y^2}p(s,x,t,y)$ existen, entonces para $t > s$, $p(s,x,t,y)$ satisface la ecuación diferencial parcial

$$\frac{\partial}{\partial t}p(s,x,t,y) = -\frac{\partial}{\partial y}\left(\mu(t,y)p(s,x,t,y)\right) + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left(\sigma^2(t,y)p(s,x,t,y)\right). \tag{12}$$

A esta ecuación se le conoce como la **ecuación hacia adelante de Kolmogorov** o la **ecuación de Fokker-Plank**.

Recíprocamente, si las condiciones de existencia y unicidad de soluciones de la ecuación (10) se cumplen y $p(s,x,t,y)$ como una función en t y y, satisface la ecuación (12), entonces $p(s,x,t,y)$ es la densidad de transición de la solución X de (10) (véase [11] p. 369).

Nos restringiremos ahora a ecuaciones diferenciales estocásticas de la forma (11). En este caso la probabilidad de transición $P(s,x,t,y) = P(0,x,t-s,y)$ depende sólo de $t-s$ (véase el teorema 6.10 de [12]). Por lo que en este caso, la probabilidad de transición es denotada por

$$P(t,x,y) = P(s,x,t+s,y) = P(0,x,t,y) = P\left(X_t \leq y | X_0 = x\right),$$

y su densidad, cuando existe, es denotada por $p(t,x,y)$. La ecuación de Fokker-Plank (12) queda expresada como

$$\frac{\partial}{\partial t}p(t,x,y) = -\frac{\partial}{\partial y}\left(\mu(y)p(t,x,y)\right) + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left(\sigma^2(y)p(t,x,y)\right).$$

Consideremos una difusión X que satisface la ecuación diferencial estocástica con coeficientes homogéneos en el tiempo (11). Sean τ el primer tiempo de salida de X de un intervalo (a,b), esto es,

$$\tau := \inf\left\{t > 0 : X_t \notin (a,b)\right\},$$

y τ_x el primer tiempo en el cual la difusión X alcanza el valor $x \in [a,b]$, es decir,

$$\tau_x := \inf\left\{t > 0 : X_t = x\right\}. \tag{13}$$

Del teorema 3.5 se observa que tanto τ como τ_x son tiempos de paro. Más aún, por la continuidad de X, es claro que $\tau = \min\left\{\tau_a, \tau_b\right\}$.

El siguiente teorema (cuya demostración puede ser consultada en [12] p. 163) nos da una fórmula para calcular la probabilidad de que el proceso alcance b antes de alcanzar a, $P_x\left(\tau_b < \tau_a\right) \equiv P\left(\tau_b < \tau_a | X_0 = x\right)$.

Teorema 4.11 *Si $\sigma(x) > 0$ es continua en $[a,b]$ y $X_0 = x$, $a < x < b$, entonces*

$$P_x(\tau_b < \tau_a) = \frac{S(x) - S(a)}{S(b) - S(a)}, \tag{14}$$

donde

$$S(x) = \int_{x_1}^{x} \exp\left(-\int_{x_0}^{u} \frac{2\mu(y)}{\sigma^2(y)} dy\right) du, \tag{15}$$

con x_0 y x_1 siendo constantes indeterminadas.

A la función S se le conoce como **función de escala**, y es una solución de la ecuación diferencial ordinaria (véase [12] p. 162)

$$\frac{1}{2}\sigma^2(x)S''(x) + \mu(x)S'(x) = 0.$$

5 Difusión de ramificación de Feller

Deseamos, en esta sección, describir un crecimiento poblacional continuo en el que los individuos se multiplican y son estadísticamente independientes entre sí con una tasa de crecimiento relativa o per cápita constante; en este caso, el desplazamiento medio infinitesimal y la varianza son necesariamente proporcionales al tamaño instantáneo de la población. De esta manera, si X_t es una variable aleatoria que representa al tamaño de la población en el instante $t \geq 0$, entonces el proceso $X = \{X_t\}_{t\geq 0}$ será un proceso de Markov para el cual existirán constantes α y σ tales que

$$\lim_{t\to 0} t^{-1} E\left(X_t - x | X_0 = x\right) = \alpha x$$

y

$$\lim_{t\to 0} t^{-1} E\left[(X_t - x)^2 | X_0 = x\right] = \sigma^2 x.$$

Así, por el teorema de Yamada-Watanabe (teorema 4.10), el proceso es la solución única de la ecuación diferencial estocástica

$$dX_t = \alpha X_t dt + \sigma\sqrt{X_t} dW_t. \tag{16}$$

La ecuación de Fokker-Plank correspondiente es

$$\frac{\partial}{\partial t} p(t,x,y) = -\alpha \frac{\partial}{\partial y}\left(y p(t,x,y)\right) + \frac{\sigma^2}{2}\frac{\partial^2}{\partial y^2}\left(y p(t,x,y)\right). \tag{17}$$

En 1951 el matemático William Feller [7] realizó el análisis de este proceso de difusión X resolviendo la ecuación diferencial parcial (17). En el presente artículo desarrollaremos el tratamiento realizado en [12] (sección 13.1) mediante la aplicación de la teoría del cálculo estocástico para obtener información de dicho proceso directamente de la ecuación diferencial estocástica (16).

Cabe destacar que aunque el crecimiento de una población es, estrictamente hablando, un proceso discontinuo, para **poblaciones grandes** el modelo continuo es una aproximación muy razonable, siempre y cuando se use una escala de tiempo apropiada (véase por ejemplo, la sección 6 del capítulo VI de [2], o la sección 4 de [10]).

El primer resultado muestra, como es de esperarse, que la solución de (16) es no negativa y que una vez que toma el valor 0, permanece por siempre en 0.

Teorema 5.1 *$P(X_t \geq 0$ para todo $t \geq 0) = 1$. Además, si $\tau_0 = \inf\{t > 0 : X_t = 0\}$, entonces $X_t = 0$ para todo $t \geq \tau_0$.*

Prueba. Consideremos primero a (16) con $X_0 = 0$. Así, es claro que $X \equiv 0$ es una solución, y como la solución es única, entonces en este caso $X \equiv 0$ es la solución. Consideremos ahora a (16) con $X_0 > 0$ y al tiempo aleatorio $\tau_0 = \inf\{t > 0 : X_t = 0\}$. Notemos que por el teorema 3.5, τ_0 es un tiempo de paro con $X_{\tau_0} = 0$. De este modo, como X es un proceso de Markov fuerte, entonces X_t para $t > \tau_0$ es la solución a (16) con $X_{\tau_0} = 0$, y así, $X_t = 0$ para todo $t > \tau_0$. Por lo tanto, si $X_0 > 0$, entonces $X_t > 0$ para todo $t < \tau_0$ y $X_t = 0$ para todo $t \geq \tau_0$. ■

El siguiente resultado describe el crecimiento medio de la población.

Teorema 5.2 *Sean $X_0 > 0$ y $\alpha > 0$. Entonces $EX_t = X_0 e^{\alpha t}$ y $\{X_t e^{-\alpha t}\}_{t\geq 0}$ es una martingala no negativa que converge casi seguramente cuando $t \to \infty$ a una variable aleatoria Z con $E|Z| < \infty$.*

Prueba. Probaremos primero que $EX_t = X_0 e^{\alpha t}$. Para ello veamos que X_t es integrable. Por el teorema 4.3 tenemos que $\{\int_0^t \sqrt{X_s} dW_s\}_{t\geq 0}$ es una martingala local. Así, podemos elegir una sucesión localizante τ_n tal que el proceso parado en τ_n, $\{\int_0^{t\wedge\tau_n} \sqrt{X_s} dW_s\}_{t\geq 0}$, es una martingala. Luego, escribiendo

$$X_{t\wedge\tau_n} = X_0 + \alpha \int_0^{t\wedge\tau_n} X_s ds + \sigma \int_0^{t\wedge\tau_n} \sqrt{X_s} dW_s,$$

y tomando esperanzas obtenemos

$$EX_{t\wedge\tau_n} = X_0 + \alpha E \int_0^{t\wedge\tau_n} X_s ds,$$

pues debido a que el proceso $\{\int_0^{t\wedge\tau_n} \sqrt{X_s} dW_s\}_{t\geq 0}$ es una martingala, $0 = \sigma E \int_0^{0\wedge\tau_n} \sqrt{X_s} dW_s = \sigma E \int_0^{t\wedge\tau_n} \sqrt{X_s} dW_s$ para todo $t \geq 0$. Como $t \wedge \tau_n \uparrow t$ cuando $n \uparrow \infty$ y dado que X es no negativo, $0 \leq \int_0^{t\wedge\tau_n} X_s ds \uparrow \int_0^t X_s ds$ cuando $n \uparrow \infty$. Luego, por el teorema de convergencia monótona,

$$E \int_0^{t\wedge\tau_n} X_s ds \to E \int_0^t X_s ds$$

cuando $n \uparrow \infty$. Por consiguiente,

$$EX_{t\wedge\tau_n} \to X_0 + \alpha \int_0^t EX_s ds$$

cuando $n \uparrow \infty$, donde hemos usado que, por el teorema de Fubini, $E\int_0^t X_s ds = \int_0^t EX_s ds$. Debido a que X es continuo, $\lim_{n\uparrow\infty} X_{t\wedge\tau_n} = X_t$, y por el lema de Fatou,

$$EX_t = E\lim_{n\uparrow\infty} X_{t\wedge\tau_n} = E\liminf_{n\uparrow\infty} X_{t\wedge\tau_n} \leq \liminf_{n\uparrow\infty} EX_{t\wedge\tau_n} = \lim_{n\uparrow\infty} EX_{t\wedge\tau_n} = X_0 + \alpha \int_0^t EX_s ds.$$

Así, usando la desigualdad de Gronwall (teorema 2.6) obtenemos

$$EX_t \leq X_0 + \alpha \int_0^t X_0 e^{\alpha(t-s)} ds = X_0 e^{\alpha t}.$$

Tenemos así que $EX_t \leq X_0 e^{\alpha t} < \infty$ para todo $t \geq 0$, lo cual prueba que X_t es integrable.

Veamos ahora que la martingala local $\{\int_0^t \sqrt{X_s} dW_s\}_{t\geq 0}$ es en realidad una martingala. Notemos que por el teorema 4.3 y el teorema de Fubini

$$\begin{aligned} E\left[\int_0^\cdot \sqrt{X_s} dW_s, \int_0^\cdot \sqrt{X_s} dW_s\right]_t &= E\int_0^t X_s ds = \int_0^t EX_s ds \\ &\leq X_0 \int_0^t e^{\alpha s} ds = (X_0/\alpha)(e^{\alpha t} - 1) < \infty \end{aligned}$$

para todo $t \geq 0$. De este modo, por el teorema 3.11, $\{\int_0^t \sqrt{X_s} dW_s\}_{t\geq 0}$ es una martingala. En particular, $0 = E\int_0^0 \sqrt{X_s} dW_s = E\int_0^t \sqrt{X_s} dW_s$, y entonces, tomando la esperanza en cada lado de la expresión integral de (16) resulta que $EX_t = X_0 + E\int_0^t \alpha X_s ds$, de lo cual se obtiene la ecuación diferencial ordinaria $\frac{dEX_t}{dt} = \alpha EX_t$ con condición inicial $EX_0 = X_0$. Por lo tanto $EX_t = X_0 e^{\alpha t}$.

Para probar que $\{X_t e^{-\alpha t}\}_{t\geq 0}$ es una martingala no negativa que converge c.s. a una variable aleatoria integrable, observemos que tanto $\{X_t\}_{t\geq 0}$ como $\{e^{-\alpha t}\}_{t\geq 0}$ son procesos de Itô para los cuales la fórmula de integración por partes (teorema 4.7) da

$$X_t e^{-\alpha t} = X_0 + \int_0^t X_s de^{-\alpha s} + \int_0^t e^{-\alpha s} dX_s.$$

Pero, de la ecuación (5), tenemos que $\int_0^t e^{-\alpha s} dX_s := \int_0^t e^{-\alpha s}\alpha X_s ds + \int_0^t e^{-\alpha s}\sigma\sqrt{X_s} dW_s$, y aplicando el teorema 2.2 obtenemos que $\int_0^t X_s de^{-\alpha s} = \int_0^t -\alpha X_s e^{-\alpha s} ds$. De manera que

$$Y_t := X_t e^{-\alpha t} = X_0 + \sigma \int_0^t e^{-\alpha s}\sqrt{X_s} dW_s,$$

y así, nuevamente por el teorema 4.3, $\{Y_t\}_{t\geq 0}$ es una martingala local tal que

$$\begin{aligned} E[Y,Y]_t &= E\left[\sigma\int_0^{\cdot} e^{-\alpha s}\sqrt{X_s}dW_s, \sigma\int_0^{\cdot} e^{-\alpha s}\sqrt{X_s}dW_s\right]_t = E\int_0^t \sigma^2 e^{-2\alpha s}X_s ds \\ &= \int_0^t \sigma^2 e^{-2\alpha s} EX_s ds = \int_0^t \sigma^2 e^{-2\alpha s}X_0 e^{\alpha s} ds = X_0\sigma^2\int_0^t e^{-\alpha s}ds \\ &= \left(X_0\frac{\sigma^2}{\alpha}\right)\left(1-e^{-\alpha t}\right). \end{aligned}$$

Consecuentemente, $\sup_{t\geq 0} E\left[Y,Y\right]_t = X_0\sigma^2\int_0^\infty e^{-\alpha s}ds = X_0\frac{\sigma^2}{\alpha} < \infty$ pues $\alpha > 0$. Por lo tanto del teorema 3.11, $\{Y_t\}_{t\geq 0}$ es una martingala cuadrado integrable, y así del teorema 3.9 tenemos que existe una variable aleatoria Z con $E|Z| < \infty$ tal que

$$\lim_{t\to\infty} X_t e^{-\alpha t} = Z \quad \text{c.s.}$$

■

El siguiente resultado nos da la probabilidad de extinción de la población cuando $X_0 = x > 0$.

Teorema 5.3 *Sea $X_0 = x > 0$. Entonces la probabilidad de extinción es $e^{(-2\alpha/\sigma^2)x}$ si $\alpha > 0$ y 1 si $\alpha \leq 0$.*

Prueba. Consideremos los tiempos de paro τ_0 y τ_b con $b > 0$ definidos por (13). De la ecuación (14) obtenemos

$$P_x(\tau_0 < \tau_b) = 1 - \frac{S(x)-S(0)}{S(b)-S(0)} = \frac{S(b)-S(x)}{S(b)-S(0)},$$

donde $S(x)$ está dada por (15). En este caso particular resulta

$$S(x) = \int_{x_1}^x \exp\left(-\int_{x_0}^u \frac{2\alpha}{\sigma^2}dy\right)du = -\frac{1}{c}e^{cx_0}\left(e^{-cx}-e^{-cx_1}\right),$$

donde $c = \frac{2\alpha}{\sigma^2}$ y $x_0, x_1 \geq 0$. Así,

$$P_x(\tau_0 < \tau_b) = \frac{e^{-cx}-e^{-cb}}{1-e^{-cb}}.$$

Observemos ahora que $\lim_{b\to\infty} P_x(\tau_0 < \tau_b) = e^{-cx}$ si $\alpha > 0$ ya que $e^{-cb} \to 0$ cuando $b \to \infty$, y $\lim_{b\to\infty} P_x(\tau_0 < \tau_b) = 1$ si $\alpha \leq 0$ pues por la regla de L' Hopital

$$\lim_{b\to\infty}\frac{e^{-cx}-e^{-cb}}{1-e^{-cb}} = \lim_{b\to\infty}\frac{\frac{d}{db}\left(e^{-cx}-e^{-cb}\right)}{\frac{d}{db}\left(1-e^{-cb}\right)} = \lim_{b\to\infty}\frac{ce^{-cb}}{ce^{-cb}} = 1.$$

Por lo tanto, la probabilidad de extinción es

$$P_x(\tau_0 < \tau_\infty) = \lim_{b\to\infty} P_x(\tau_0 < \tau_b) = \begin{cases} e^{\left(-\frac{2\alpha}{\sigma^2}\right)x}, & \text{si } \alpha > 0 \\ 1, & \text{si } \alpha \leq 0, \end{cases}$$

donde $\tau_\infty = \lim_{b\to\infty} \tau_b$ es el tiempo de explosión, el cual es infinito si la explosión no ocurre. La explosión ocurre si el proceso X toma el valor ∞ en un tiempo τ_∞ finito. ■

Referencias

[1] R. B. Ash, Probability and measure theory, 2nd. Edition, Academic Press, USA (2000).

[2] K. B. Athreya and P. E. Ney, Branching processes, Springer-Verlag, Germany (1972).

[3] T. Bojdecki, Teoría general de procesos e integración estocástica, Aportaciones Matemáticas, Serie textos 6, Soc. Mat. Mex. (1995).

[4] V. Capasso and D. Bakstein, An introduction to continuous-time stochastic processes: theory, models, and applications to finance, biology, and medicine, Birkhäuser, USA (2005).

[5] N. L. Carothers, Real analysis, Cambridge University Press, USA (2000).

[6] K. L. Chung and R. J. Williams, Introduction to stochastic integration, 2dn. Edition, Birkhäuser, USA (1990).

[7] W. Feller, Diffusion processes in genetics, Second Berkeley Symposium on Math. Stat. and Probab., 227-246 (1951).

[8] A. Friedman, Stochastic differential equations and applications, Volume 1, Academic Press, USA (1975).

[9] I. I. Gihman and A. V. Skorohod, Stochastic differential equations, Springer-Verlag, USA (1972).

[10] L. G. Gorostiza, Ramificación y superprocesos, Aportaciones Matemáticas, Comunicaciones, Vol. 39, 79-105 (2008).

[11] I. Karatzas and S. E. Shreve, Brownian motion and stochastic calculus, 2nd. Edition, Springer-Verlag, USA (1998).

[12] F. C. Klebaner, Introduction to stochastic calculus with applications, 2nd. Edition, Imperial College Press, Singapore (2005).

[13] H. H. Kuo, Introduction to stochastic integration, Springer-Verlag, USA (2006).

[14] R. G. Pinsky, Positive harmonic functions and diffusion, Cambridge University Press, USA (1995).

[15] P. E. Protter, Stochastic integration and differential equations: version 2.1 (stochastic modeling and applied probability), 2nd. Edition, Springer, Germany (2010).

[16] L. C. G. Rogers and D. Williams, Diffusions, Markov processes and martingales, Volume 2: Itô calculus, 2nd. Edition, Cambridge University Press, USA (2000).