Universidad de la República
Facultad de Ciencias Sociales
DEPARTAMENTO DE ECONOMIA

Notas Docentes

Aplicación del Cálculo Estocástico al Análisis de la Estructura Temporal de las Tasas de Interés

Serafín Frache y Gabriel Katz

Nota Docente No. 13

Indice

I. Introducción

La estructura temporal de las tasas de interés es la relación entre el plazo de instrumentos financieros de determinada calidad crediticia y sus rendimientos o precios respectivos. Contar con una estimación precisa de la estructura temporal de las tasas de interés permite determinar los precios de instrumentos financieros dependientes de tasas de interés tales como bonos del tesoro, letras de tesorería, obligaciones negociables, acciones, certificados de depósitos, etc. Incluso los precios de instrumentos más complejos, tales como *swaps* de tasas de interés, futuros, opciones y otros derivados[1] de tasas de interés, pueden computarse a partir de la estructura temporal.

A fin de determinar los precios de estos instrumentos, es necesario modelar la incertidumbre respecto a la evolución futura de la estructura temporal de las tasas de interés. Aún cuando los flujos de fondos generados por los instrumentos financieros sean conocidos, las tasas de interés futuras dependerán del estado futuro, incierto, de la economía. La literatura financiera moderna recurre a procesos estocásticos para modelar dicha incertidumbre. Este trabajo introduce someramente y de manera relativamente informal los procesos estocásticos y las herramientas matemáticas utilizadas en la literatura financiera a fin de especificar la evolución intertemporal de la estructura temporal de las tasas de interés. Omitiremos aquellos detalles técnicos que no sean imprescindibles para una comprensión razonable de estas herramientas, centrándonos en los procesos y resultados más relevantes. El lector interesado puede remitirse a Lamberton y Lapeyre (1996), Shreve (1997) y Mikosch (1999) para un examen más riguroso y detallado. Si bien a efectos de facilitar la exposición consideraremos básicamente procesos unidimensionales, la extensión de las definiciones al caso multidimensional resulta inmediata.

[1] Un derivado es cualquier instrumento financiero cuyo precio está directamente vinculado al valor de un activo subyacente: acciones, canastas de acciones, índices bursátiles, *commodities*, otro derivado, etc.

II. Espacios de probabilidad

Para la representación de eventos inciertos, se define un **espacio de probabilidad** $(\Omega, \mathcal{F}, \mathcal{P})$.

El primer elemento del espacio de probabilidad, es un conjunto arbitrario designado como Ω, se denomina el **espacio muestral**, y puede concebirse como el conjunto de todos los posibles resultados, ω, de un experimento aleatorio o el conjunto de todos los posibles resultados de un "estado del mundo". Un **evento** es un subconjunto del espacio muestral, un conjunto de los resultados posibles.

Una **σ -álgebra** $\mathcal{F}$ es una colección de eventos "interesantes", un conjunto de todos los eventos relevantes en una u otra situación, a los que le asignaremos cierta probabilidad. Intuitivamente, $\mathcal{F}$ puede interpretarse como un conjunto de información.[2] Más rigurosamente, sea Ω un conjunto no vacío, una σ –álgebra es una colección de subconjuntos de Ω que cumple las siguientes tres propiedades:

(i) $\varnothing \in \mathcal{F}$, donde $\varnothing$ representa el conjunto vacío

(ii) si $A \in \mathcal{F}$, entonces su complemento $A^c \in \mathcal{F}$

(iii) si $A_1, A_2, A_3, \ldots \in \mathcal{F}$, entonces $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Dos casos extremos de una σ – álgebra son $\mathcal{F}_0 = \{\varnothing, \Omega\}$ y $\mathcal{F}_\infty$. Esta última contiene todos los posibles subconjuntos de Ω.

$\mathcal{G}$ es una **sub - σ – álgebra** de $\mathcal{F}$ si $\mathcal{G} \subseteq \mathcal{F}$ y $\mathcal{G}$ es una **σ – álgebra**.

Una **filtración** $\{\mathcal{F}_t\}_{t\geq 0}$ es una familia creciente de σ – álgebras incluidas en $\mathcal{F}$ tal que cada σ-álgebra contiene todos los conjuntos contenidos en la σ – álgebra anterior, por lo que $\mathcal{F}_s \subseteq \mathcal{F}_t$ para todo $s \leq t$. Es decir, una filtración puede ser concebida como una corriente de información siempre creciente, se asume que los agentes "no olvidan".

[2] Generalmente, la literatura financiera suele modelar la información disponible mediante σ - álgebras. No obstante, Dubra y Echenique (2001) muestran que esta definición de $\mathcal{F}$ no es enteramente válida, ya que el teorema de Blackwell (1951) no se cumple si la infomación es modelada como σ - álgebras en la manera en que habitualmente lo hacen los modelos financieros.

El tercer componente del espacio de probabilidad es una **medida de probabilidad** $\mathcal{P}$. $\mathcal{P}$ es una función $\mathcal{P}:\mathcal{F}\rightarrow[0,1]$ que asigna a cada evento $A\in\mathcal{F}$ un número $\mathcal{P}(A)\in[0,1]$. Formalmente, $\mathcal{P}$ satisface las siguientes propiedades:

(i) $\mathcal{P}(\varnothing)=0$

(ii) $\mathcal{P}(\Omega)=1$

(iii) si $A_1, A_2,\ldots.$ es una sucesión de conjuntos disjuntos en $\mathcal{F}$, entonces:

$$\mathcal{P}\left(\bigcup_{k=1}^{\infty}A_k\right)=\sum_{k=1}^{\infty}\mathcal{P}(A_k).$$

Dos medidas de probabilidad $\mathcal{P}$ y $\mathcal{Q}$ definidas en el mismo espacio muestral y $\sigma-$ álgebra $(\Omega,\mathcal{F})$ son **equivalentes** si ambas asignan probabilidad cero a exactamente los mismos eventos, es decir, si $\mathcal{P}(A)=0\Leftrightarrow\mathcal{Q}(A)=0$.[3] Intuitivamente, esto quiere decir que los eventos que no pueden ocurrir de acuerdo a $\mathcal{P}$ no pueden hacerse posibles simplemente cambiando la medida de probabilidad a $\mathcal{Q}$. Del mismo modo, eventos que pueden ocurrir bajo $\mathcal{P}$ no pueden tornarse imposibles mediante un cambio en la medida de probabilidad.

III. Variables aleatorias, procesos estocásticos en tiempo continuo y martingalas

Una **variable aleatoria** x es una función definida sobre el espacio muestral Ω que asigna a cada $\omega\in\Omega$ un único número real, $x(\omega)$. $X=(x_1,x_2,\ldots x_n)$ es un **vector aleatorio** n – dimensional si sus componentes, $x_1,x_2,\ldots x_n$ son variables aleatorias.

Un **proceso estocástico en tiempo continuo**, $\{x_t\}_{t\geq 0}$, es una colección de variables aleatorias definidas en Ω, donde x_t representa el valor del objeto descrito por el proceso en el momento t. Una **trayectoria muestral** es una posible realización de la evolución del proceso en el tiempo; puede concebirse como el equivalente al concepto de serie de tiempo para las realizaciones de procesos estocásticos en tiempo discreto. El espacio muestral Ω es el conjunto de todas las trayectorias muestrales.

[3] Es decir, $\mathcal{P}$ y $\mathcal{Q}$ son cada una absolutamente continua respecto a la otra.

Los agentes en los mercados financieros pueden operar virtualmente en cualquier momento del tiempo. En realidad, debido a consideraciones prácticas (los mercados no están “abiertos” todo el tiempo) y a la existencia de costos de transacción, ningún inversor realizará transacciones continuamente. No obstante, dado un número suficientemente grande de agentes, habrá transacciones prácticamente en todo momento, por lo que los precios de los instrumentos y las tasas de interés variarán casi continuamente. Por esto, los modelos dinámicos de la ETTI suelen adoptar el supuesto de que las transacciones financieras ocurren continuamente, recurriendo entonces a procesos estocásticos en tiempo continuo para dar cuenta de la evolución de las variables de estado.

Asumiremos que todas las variables aleatorias x_t toman valores en el mismo conjunto S, denominado el **espacio de estados** del proceso. Más rigurosamente, S es el conjunto más pequeño con la propiedad de que $\mathcal{P}\{x_t \in S\} = 1 \ \forall t$. Si $S \subseteq \mathbb{R}$, estaremos ante un proceso unidimensional valuado en los reales. Si S es un subconjunto de $\mathbb{R}^n$, se tratará de un proceso n – dimensional valuado en los reales, que puede concebirse como un vector de n procesos unidimensionales valuados en los reales. En tanto consideremos medidas de probabilidad equivalentes, el espacio de estados no se verá afectado por cambios en la medida de probabilidad.

A medida que transcurre el tiempo, podemos observar la evolución del objeto descrito por el proceso estocástico. En cualquier momento del tiempo t', los valores previos $\{x_t\}_{t \in [0,t')}$, donde $x_t \in S$, serán conocidos. Estos valores constituyen la historia del proceso hasta el momento t'; los valores futuros son estocásticos.

Conforme pasa el tiempo, además, será posible revisar nuestras expectativas acerca de los valores futuros del proceso o, más precisamente, revisar la distribución de probabilidad que atribuimos al valor del proceso $\{x_t\}_{t \geq 0}$ en cualquier punto futuro del tiempo. Supongamos que estamos situados en el momento t, y consideremos el valor del proceso en algún momento futuro $u > t$. La distribución del valor de x_u está caracterizada por las probabilidades $\mathcal{P}(x_u \in A)$ para subconjuntos A medibles del espacio de estados S. Si para todo $t, u \in \mathbb{R}_+$, con $t < u$, y todo $A \subseteq S$ se tiene que

$$\mathcal{P}\left[x_u \in A / \{x_s\}_{s \in [0,t]}\right] = \mathcal{P}\left(x_u \in A / x_t\right),$$

entonces $\{x_t\}_{t\geq 0}$ se denomina un **proceso de Markov**. Esta condición implica que, dada la información disponible en el presente, el futuro es independiente del pasado; la historia no contiene otra información respecto a los valores futuros más que la que puede extraerse del valor presente.

Los procesos de Markov son ampliamente utilizados en los modelos financieros para describir la evolución de los precios de los activos, ya que la propiedad de Markov es consistente con la denominada "forma débil" de eficiencia del mercado, que establece que no pueden obtenerse retornos extraordinarios mediante la utilización de información acerca de la evolución histórica del precio de un activo.[4] Si pudieran obtenerse retornos extraordinarios de esta manera, todos los inversores tratarían de aprovecharse, de modo que los precios cambiarían inmediatamente y se situarían en un nivel en el que no podrían lograrse beneficios extraordinarios. Por este motivo es razonable modelar los precios mediante procesos de Markov, asumiendo que la información pública es rápidamente incorporada al precio de los instrumentos, por lo que el comportamiento pasado de los mismos no tiene valor predictivo alguno. Además, los modelos basados en procesos de Markov son por lo general analíticamente más tratables que los modelos que no los utilizan. La mayor parte de los procesos no-markovianos pueden ser transformados en procesos de Markov mediante una técnica denominada "expansión de estados" [Cox y Miller (1965, pp. 16 – 18)].

Un proceso estocástico es una **martingala** si, en cualquier punto del tiempo, la variación esperada en el valor del proceso a lo largo de cualquier período futuro dado es igual a cero. Puesto que la esperanza depende de una cierta medida de probabilidad, el concepto de martingala debe concebirse en relación a la medida de probabilidad aplicada. Más formalmente, un proceso estocástico $\{x_t\}_{t\geq 0}$ es una martingala bajo $\mathcal{P}$ (o $\mathcal{P}$- martingala) si para todo $t \geq 0$ se tiene que:

$$E_t^{\mathcal{P}}\left[x_u\right] = x_t \text{ para todo } u \geq t$$

donde $E_t^{\mathcal{P}}$ denota la esperanza computada bajo la medida de probabilidad $\mathcal{P}$, dada la información disponible en el momento t, $\mathcal{F}_t$; es decir, dada la historia del proceso hasta (e

[4] Esta propiedad no entra en conflicto con el hecho de que la información histórica es a menudo utilizada para identificar algunas características específicas del proceso; por ejemplo, para estimar medias y varianzas.

incluyendo) el momento t. Esto implica que x_t es la mejor predicción de x_u dado $\mathcal{F}_t$. Deben cumplirse, además, dos condiciones técnicas adicionales:

- $E\left|x_t\right| < \infty$ para todo t
- $\left\{x_t\right\}_{t\geq 0}$ es un proceso **adaptado** a $\left\{\mathcal{F}_t\right\}_{t\geq 0}$, lo que significa que para todo t, x_t es medible en $\mathcal{F}_t$.

IV. Procesos de Wiener

Prácticamente todos los procesos estocásticos aplicados en los modelos dinámicos de la ETTI se basan en una clase particular de procesos, denominados **procesos de Wiener** o **movimientos Brownianos estándar**. A partir del trabajo de Bachelier (1900), quien representó los movimientos en el precio de las acciones mediante procesos de Wiener, este tipo de procesos ha sido ampliamente utilizado en la literatura financiera.

Un proceso estocástico unidimensional $\left\{w_t\right\}_{t\geq 0}$ es un movimiento Browniano estándar si satisface las siguientes condiciones:

(i) $w_0 = 0$

(ii) para todo $t, t' \geq 0$ con $t < t' : w_{t'} - w_t \sim N\left(0, t'-t\right)$ (incrementos normalmente distribuidos)

(iii) para todo $0 \leq t_0 < t_1 ... < t_n$, las variables aleatorias $w_{t_1} - w_{t_0}, ..., w_{t_n} - w_{t_{n-1}}$ son mutuamente independientes (incrementos independientes).

De (ii) y (iii) se sigue que para todo $t > 0$, $w_t \sim N(0,t)$ y $Cov\left[w_s, w_t\right] = \min(s,t)$ [5] para todo $0 \leq s < t$.

(iv) $\left\{w_t\right\}_{t\geq 0}$ tiene trayectorias muestrales continuas con probabilidad 1.

---

[5] Puesto que los incrementos $w_s - w_0 = w_s$ y $w_t - w_s$ son independientes para $t > s$, su covarianza: $Cov\left[w_s, w_t\right] = E\left[\left(\left(w_t - w_s\right) + w_s\right) w_s\right] = E\left[\left(w_t - w_s\right) w_s\right] + E\left[w_s^{\,2}\right] = E\left[w_t - w_s\right] E\left[w_s\right] + s = 0 + s = s$ para $0 \leq s < t$.

Un proceso de Wiener está definido en relación a una medida de probabilidad $\mathcal{P}$ bajo la cual los incrementos tienen las propiedades arriba mencionadas. Por ejemplo para todo $t < t'$ y para todo $h \in \mathbb{R}$, tenemos que:

$$\mathcal{P}\left(\frac{w_{t'} - w_t}{\sqrt{t'-t}} < h \right) = \phi(h) \equiv \int_{-\infty}^{h} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du \ ,$$

donde $\phi(\cdot)$ denota la función de distribución acumulada de una normal estándar. Nótese que un movimiento Browniano estándar es un proceso de Markov, puesto que el incremento desde hoy a cualquier punto futuro en el tiempo es independiente de la historia del proceso. Un movimiento Browniano es asimismo una martingala, ya que el cambio esperado en el valor del proceso es cero.

Puede demostrarse [véase Mikosch (1999, sección 1.3)] que las trayectorias muestrales de un proceso de Wiener no son diferenciables en ningún punto. Otra indicación de la irregularidad de las trayectorias muestrales Brownianas es que éstas no tienen una variación acotada en ningún intervalo finito $[0,T]$:

$$\sup_{\tau} \sum_{i=1}^{n} \left| w_{t_i}(\omega) - w_{t_{i-1}}(\omega) \right| = \infty$$

donde el supremo se toma a lo largo de todas las particiones posibles $\tau : 0 = t_0 < t_1 < .... < t_n = T$ de $[0,T]$.

La variación no acotada y la no diferenciabilidad de las trayectorias muestrales Brownianas son las principales causas de las dificultades que enfrenta la aplicación de los métodos de integración clásicos a dichas trayectorias [véase Mikosch (1999, secciones 2.1 y 2.2)], resultando necesario apelar al cálculo estocástico.

Sea $\{w_t\}_{t\geq 0}$ un movimiento Browniano unidimensional, y defínase un nuevo proceso estocástico $\{x_t\}_{t\geq 0}$ dado por:

$$x_t = x_0 + \mu t + \sigma w_t, \ t \geq 0 \ ,$$

donde x_0, μ y σ son constantes; la constante x_0 es el valor inicial del proceso. A partir de las propiedades ya mencionadas del movimiento Browniano estándar se sigue que, visto desde el momento 0, x_t se distribuye normalmente con media $\mu\, t$ y varianza $\sigma^2 t$. El cambio en el valor del proceso $\{x_t\}_{t\geq 0}$ entre dos puntos arbitrarios en el tiempo t y t', donde $t < t'$, viene dado por:

$$x_{t'} - x_t = \mu(t' - t) + \sigma(w_{t'} - w_t)$$

El cambio a lo largo de un intervalo infinitesimalmente corto $[t, t + \Delta t]$, con $\Delta t \to 0$, suele escribirse como

$$dx_t = \mu dt + \sigma dw_t \tag{1}$$

donde dw_t puede ser interpretado como una variable aleatoria con distribución $N(0, dt)$.[6] El proceso $\{x_t\}_{t\geq 0}$ se denomina un **proceso de Wiener generalizado**. El parámetro μ refleja el cambio esperado en el proceso por unidad de tiempo, y se denomina **componente tendencial**. El parámetro σ refleja la incertidumbre acerca de los valores futuros del proceso, y se denomina la **volatilidad** del proceso. Un proceso de Wiener generalizado comparte varias de las propiedades características de un movimiento Browniano estándar: al igual que aquél, es un proceso de Markov, y sus trayectorias son continuas pero no son diferenciables en ningún punto. No obstante, un proceso de Wiener generalizado no es una martingala a menos que $\mu = 0$.

Si los parámetros μ y σ varían en el tiempo en forma determinística, $\{x_t\}_{t\geq 0}$ es un proceso de Wiener generalizado no homogéneo respecto al tiempo, y puede ser expresado en términos diferenciales como

$$dx_t = \mu_t dt + \sigma_t dw_t \quad (2).$$

[6] Para dar a la expresión (1) un significado matemático preciso, debe ser interpretada como el límite de la expresión $x_{t+\Delta t} - x_t = \mu\Delta t + \sigma(w_{t+\Delta t} - w_t)$ para $\Delta t \to 0$.

A lo largo de un intervalo pequeño de tiempo $[t, t+\Delta t]$, el cambio esperado en el proceso será aproximadamente igual a $\mu_t \Delta t$ y la varianza del cambio será aproximadamente $\sigma_t^2 \Delta t$. Más precisamente, el incremento a lo largo de cualquier intervalo $[t, t']$ vendrá dado por:

$$x_{t'} - x_t = \int_t^{t'} \mu_u du + \int_t^{t'} \sigma_u dw_u$$

donde la última integral es una **integral estocástica** o **Integral de Ito**. La consideración exhaustiva de las propiedades y características de las integrales estocásticas excede los objetivos de este trabajo; un análisis detallado y riguroso puede encontrarse en Lamberton y Lapeyre (1996, sección 3.4) y Shreve (1997, cap. 14). No obstante, vale la pena mencionar aquí algunas de las propiedades de la Integral de Ito: [7]

(i) $E_t\left[\int_t^{t'} C_u dw_u\right] = 0$ para un proceso integrando C adaptado a $\{w_t\}_{t\geq 0}$ que cumple que $\int_0^T EC_u^2 du$ es finita [véase Mikosch (1999, sección 2.2.3)].

(ii) $E_t\left[\int_t^{t'} C_u dw_u\right]^2 = \int_t^{t'} E_t\left[C_u^{\,2}\right] du$ (propiedad de "isometría").

(iii) Si $C(u)$ es una función determinística del tiempo, $\int_t^{t'} C(u) dw_u$ tiene una distribución normal con media cero y varianza $\int_t^{t'} C(u)^2 du$.

V. Procesos de difusión

Un **proceso de difusión** unidimensional es un proceso estocástico $\{x_t\}_{t\geq 0}$ para el que el cambio a lo largo de un intervalo de tiempo infinitesimal, $[t, t+dt]$, puede representarse como

$$dx_t = \mu(x_t, t)dt + \sigma(x_t, t)dw_t \tag{3}$$

donde $\{w_t\}_{t\geq 0}$ es un movimiento Browniano estándar, pero el componente tendencial μ y la volatilidad σ son ahora funciones del tiempo y del valor actual del proceso. Esta expresión generaliza (1), donde μ y σ se asumían constantes, y (2), donde eran funciones del tiempo únicamente. Una ecuación como (3), en la que el proceso estocástico entra en ambos lados de la igualdad, se denomina una **ecuación diferencial estocástica**.[8] Por ende, un proceso de difusión es una solución a una ecuación diferencial estocástica. En rigor, $\mu(x,t)$ y $\sigma(x,t)$ deben satisfacer las condiciones de crecimiento y de Lipschitz: deben existir constantes c y d tal que para todo $(x,t)\in\mathbb{R}$,

(i) $$\mu(x,t)^2+\sigma(x,t)^2\leq c\left(1+|x|\right)$$

(ii) $$|\mu(x,t)-\mu(y,t)|+|\sigma(x,t)-\sigma(y,t)|\leq d\,|x-y|$$

La condición de crecimiento (i) impide un comportamiento explosivo de x_t y asegura que existe una solución de (3); la condición de Lipschitz (ii) establece que $\mu(x,t)$ y $\sigma\left(x,t\right)$ no varían más rápidamente que $\{x_t\}_{t\geq 0}$, lo que asegura que la solución es única. La solución $\{x_t\}_{t\geq 0}$ así obtenida es una solución "fuerte", lo que significa que cualquier otro proceso de difusión que resuelva (3) es igual a $\{x_t\}_{t\geq 0}$ en casi cualquier punto.[9]

Si tanto μ como σ son independientes del tiempo, se dice que la difusión es homogénea respecto al tiempo. En este caso, la distribución del valor futuro del proceso dependerá del valor actual y de "cuán lejos en el futuro estemos viendo", no del momento particular en el que estemos situados. A modo de ejemplo, la distribución de $x_{t+\delta}$ dado $x_t=x$ dependerá únicamente de x y de δ, pero no de t. Esto no se cumple en el caso de una difusión no homogénea respecto al tiempo, en cuyo caso la distribución dependerá también de t.

[7] Una demostración de estos resultados puede encontrarse en Mikosch (1999, Apéndice A4) y en Munk (2002, sección 3.6.1).

[8] Intuitivamente, una ecuación diferencial estocástica puede ser concebida como una ecuación diferencial ordinaria o determinística que se ve perturbada por el arribo de nueva información, modelado por el movimiento Browniano estándar. Una caracterización minuciosa de las ecuaciones diferenciales estocásticas puede encontrarse en Mikosch (1999, cap. 3).

[9] Para una definición más precisa del concepto de "solución fuerte" de una ecuación diferencial estocástica, véase Mikosch (1999, sección 3.2).

Como mencionamos anteriormente, dw_t presenta una distribución $N\left(0,dt\right)$, de modo que la media y la varianza del cambio en el proceso a lo largo de un intervalo infinitesimal $\left[t,t+dt\right]$ vienen dadas por

$$E_t\left[dx_t\right] = \mu(x_t,t)dt$$
$$Var_t\left[dx_t\right] = \sigma(x_t,t)^2 dt$$

donde $E_t[\ ]$ y $Var_t[\ ]$ denotan la media y la varianza, respectivamente, condicional a la información disponible en el momento t. Más rigurosamente, el cambio en el proceso de difusión a lo largo del intervalo $\left[t,t'\right]$ es:

$$x_{t'} - x_t = \int_t^{t'} \mu(x_u,u)du + \int_t^{t'} \sigma(x_u,u)dw_u \ ,$$

donde la primer integral es una Integral de Riemann, en tanto que la segunda es una integral estocástica o de Ito. El componente tendencial y la varianza vienen en realidad dados por los límites:

$$\mu(x_t,t) = \lim_{\Delta\to 0}\frac{E_t\left[x_{t+\Delta t}-x_t\right]}{\Delta t}$$
$$\sigma(x_t,t)^2 = \lim_{\Delta\to 0}\frac{Var_t\left[x_{t+\Delta t}-x_t\right]}{\Delta t}$$

Un proceso de difusión es un proceso de Markov, puesto que se desprende de (3) que tanto el componente tendencial como la volatilidad dependen del valor actual del proceso, no de sus valores anteriores. Una difusión no es una martingala a menos que $\mu(x_t,t)$ sea igual a cero para todo x_t y t. Las trayectorias muestrales de un proceso de difusión serán continuas, pero no diferenciables en ningún punto; el espacio de estados S y la distribución de los valores futuros dependerán de las funciones μ y σ.

Algunos de los procesos de difusión más frecuentemente utilizados en la modelización de la estructura temporal son los movimientos Brownianos geométricos, los procesos de Ornstein – Uhlenbeck y los procesos de raíz cuadrada.

Un proceso estocástico unidimensional $\{x_t\}_{t\geq 0}$ es un **movimiento Browniano geométrico** si constituye una solución a la ecuación diferencial estocástica

$$dx_t = \mu x_t \, dt + \sigma x_t \, dw_t \tag{4}$$

donde μ, σ son constantes. El valor inicial del proceso se asume positivo, $x_0 > 0$. En definitiva, un movimiento Browniano geométrico es un proceso de difusión particular que se obtiene a partir de (3) insertando $\mu(x_t, t) = \mu x_t$ y $\sigma(x_t, t) = \sigma x_t$.
La expresión (4) puede escribirse como

$$\frac{dx_t}{x_t} = \mu dt + \sigma dz$$

que es la variación relativa (porcentual) en el valor del proceso a lo largo del intervalo de tiempo infinitesimal $[t, t+dt]$. Si x_t es el precio de un instrumento financiero, $\frac{dx_t}{x_t}$ será entonces la tasa de retorno instantánea del instrumento. La constante μ será en ese caso la tasa de retorno esperada por período, en tanto que σ será el desvío estándar de la tasa de retorno.

El precio de un instrumento financiero con una tasa esperada de retorno constante y una volatilidad relativa σ también constante seguirá un movimiento Browniano geométrico; el famoso modelo de determinación del precio de las opciones de Black – Scholes – Merton asume que el precio de la acción sobre la que se suscribe la opción sigue precisamente un movimiento Browniano geométrico. El modelo de Merton (1970), uno de los primeros modelos dinámicos de la estructura temporal, recurre asimismo a un movimiento Browniano geométrico a fin de representa la dinámica de la tasa de interés.

Por su parte, un proceso estocástico unidimensional $\{x_t\}_{t\geq 0}$ es un **proceso de Ornstein – Uhlenbeck** si su dinámica es de la forma

$$dx_t = [\varphi - \kappa x_t] dt + \sigma dw_t \tag{5}$$

donde φ, κ y σ son constantes, con $\kappa > 0$. Alternativamente, puede expresarse como:

$$dx_t = \kappa\left[\theta - x_t\right]dt + \sigma dw_t \qquad (6)$$

donde

$$\theta = \frac{\varphi}{\kappa}.$$

Un proceso de Ornstein – Uhlenbeck exhibe reversión en media: el componente tendencial es positivo cuando $x_t < \theta$ y negativo cuando $x_t > \theta$. Por consiguiente, el proceso tiende siempre hacia su nivel "normal" o de largo plazo, θ. El parámetro κ controla el tamaño o la velocidad del ajuste hacia el nivel de largo plazo, y se denomina el parámetro de reversión en media o velocidad del ajuste.

La ecuación (6) es la versión en tiempo continuo de un proceso autorregresivo de orden uno. En concreto, (6) es el límite, cuando $\Delta t \to 0$, del siguiente proceso AR(1):

$$x_t - x_{t-1} = \theta\left(1 - e^{-\kappa}\right) + \left(e^{-\kappa} - 1\right)x_{t-1} + \varepsilon_t,$$

donde $\varepsilon_t \sim N(0, \sigma_\varepsilon{}^2)$, con

$$\sigma_\varepsilon{}^2 = \frac{\sigma^2}{\kappa}\left(1 - e^{-2\kappa}\right)$$

[véase Dixit y Pindyck (1994, cap. 3, sección 3.B)].

Resulta evidente que un bajo valor del parámetro de reversión en media κ implica un alto grado de autocorrelación. Vasicek (1977) fue uno de los primeros autores en utilizar procesos de Ornstein – Uhlbenbeck para modelizar la estructura temporal de las tasas de interés.

Finalmente, un proceso estocástico unidimensional $\{x_t\}_{t \geq 0}$ es un **proceso de raíz cuadrada** si su dinámica es de la forma

$$dx_t = \left[\varphi - \kappa x_t\right]dt + \beta\sqrt{x_t}\ dwt = \kappa\left[\theta\text{-}x_t\right]dt + \beta\sqrt{x_t}\ dw_t\,,$$

donde φ, θ, β son constantes positivas y se asume que el valor inicial del proceso, x_0, es positivo, de modo de poder aplicar la raíz cuadrada. La única diferencia con la dinámica de los procesos de Ornstein – Uhlenbeck radica en el término $\sqrt{x_t}$ en la volatilidad y en que la varianza, dada por $\beta^2 x_t$, es en este caso proporcional al nivel del proceso. Al igual que en el caso de los procesos de Ornstein – Uhlenbeck, los procesos de raíz cuadrada exhiben reversión en media, si bien en este caso el proceso no puede tomar valores negativos. Para ver esto, nótese que si el valor del proceso fuera cero, entonces el componente tendencial sería negativo y la volatilidad igual a cero, por lo que el valor del proceso se tornará inmediatamente positivo. Puede demostrarse que si $2\varphi \geq \beta^2$, el componente tendencial positivo para valores bajos del proceso será tan grande en relación a la volatilidad que éste permanecería estrictamente positivo. Por consiguiente, el espacio de estados de un proceso de raíz cuadrada será $S = [0,\infty)$ o $S = (0,\infty)$.

Los procesos de raíz cuadrada son utilizados en el modelo de Cox, Ingersoll y Ross (1985) y en los modelos dinámicos derivados de aquél (véase Lund, 1993, cap. 4).

VI. Procesos de Ito

Es posible definir procesos más generales que aquellos comprendidos dentro de los procesos de difusión. Un proceso estocástico unidimensional se denomina un **proceso de Ito** si los incrementos locales son de la forma:

$$dx_t = \mu_t dt + \sigma_t dw_t,$$

donde μ y σ son ellos mismos procesos estocásticos. Obsérvese que un proceso de difusión es un caso particular de un proceso de Ito en el que los valores del componente tendencial y la volatilidad están dados por t y x_t. En el caso de un proceso de Ito genérico, μ y σ pueden depender también de los valores pasados del proceso, por lo que los procesos de Ito no son generalmente markovianos. Tampoco son martingalas, a menos que μ_t sea idénticamente igual a cero y σ_t satisfaga algunas condiciones técnicas (véase Munk, 2002, sección 3.5).

VII. Lema de Ito

Los modelos dinámicos de la ETTI asumen ciertos procesos estocásticos para las variables de estado o factores. La tasa de interés instantánea, los precios de los bonos y demás variables de interés son funciones de estos factores. Para determinar la dinámica de estas otras variables, será necesario recurrir al **Lema de Ito**, que es básicamente la versión estocástica de la regla de la cadena del cálculo ordinario. En esencia, establece que una función "bien comportada" f de un proceso de Ito sigue a su vez un proceso de Ito. Es decir, el lema preserva la propiedad de Ito aún cuando f no sea lineal.

Sea $\{x_t\}_{t\geq 0}$ un proceso de Ito unidimensional valuado en los reales cuya dinámica viene dada por

$$dx_t = \mu_t dt + \sigma_t dw_t ,$$

donde μ y σ son procesos valuados en los reales y w es un movimiento Browniano estándar unidimensional.

Sea $f(x_t, t)$ una función valuada en los reales continuamente diferenciable dos veces respecto a x y una vez respecto a t. Entonces el proceso y definido por $y_t = f(x_t, t)$ es también un proceso de Ito, y su dinámica viene dada por:

$$dy_t = \left[\frac{\partial f(x_t, t)}{\partial t} + \frac{\partial f(x_t, t)}{\partial x} \mu_t + \frac{1}{2} \frac{\partial^2 f(x_t, t)}{\partial x^2} \sigma_t^{\ 2} \right] dt + \frac{\partial f(x_t, t)}{\partial x} \sigma_t dw_t \qquad (7).$$

La demostración del Lema de Ito se basa en la expansión de Taylor de $f(x_t, t)$ combinada con límites apropiados. El lector interesado puede consultar Shreve (1997, cap. 15).

La expresión (7) puede también escribirse como:

$$dy_t = \frac{\partial f(x_t, t)}{\partial t} dt + \frac{\partial f(x_t, t)}{\partial x} dx_t + \frac{1}{2} \frac{\partial^2 f(x_t, t)}{\partial x^2} \left(dx_t \right)^2 .$$

En el cálculo de $\left(dx_t \right)^2$ debe aplicarse que

$$\left(dt \right)^2 = dt \ dw_t = 0$$

y

$$\left(dw_t \right)^2 = dt ,$$

por lo que

$$\left(dx_t\right)^2 = \left(\mu_t dt + \sigma_t dw_t\right)^2 = \mu_t^2 \left(dt\right)^2 + 2\mu_t\sigma_t dt + \sigma_t^2 \left(dw_t\right)^2 = \sigma_t^2 dt .$$

La intuición detrás de estos resultados es la siguiente: cuando dt está cercano a cero, $\left(dt\right)^2$ es sensiblemente menor que dt y puede por ende ser ignorado. Puesto que $dw_t \sim N(0,dt)$, se obtiene que

$$E\left[dt\ dw_t\right] = dt\ E\left[dw_t\right] = 0 ,$$

y

$$Var\left[dt\ dw_t\right] = \left(dt\right)^2 Var\left[dw_t\right] = \left(dt\right)^3 ,$$

que también resulta muy pequeño en comparación con dt y puede ser entonces ignorado. Finalmente, se tiene que

$$E\left[\left(dw_t\right)^2\right] = Var\left[dw_t\right] - \left(E\left[dw_t\right]\right)^2 = dt ,$$

y puede mostrarse[10] que

$$Var\left[\left(dw_t\right)^2\right] = 2\left(dt\right)^2 .$$

Para dt próximo a cero, la varianza es sustancialmente menor que la media, por lo que $\left(dw_t\right)^2$ puede aproximarse por su media dt.

En comparación con la regla de la cadena del cálculo ordinario, la ecuación (7) tiene un término "extra". Supóngase, por simplicidad, que $\mu_t = 0$ y que $\frac{\partial f(x_t,t)}{\partial t} = 0$. En ese caso, $E\left[dx_t\right] = 0$, pero $E\left[dy_t\right] \neq 0$: si f es una función convexa de x, esto es, si $\frac{\partial^2 f(x_t,t)}{\partial x^2} > 0$, $E\left[dy_t\right]$ será positiva, mientras que si f es una función cóncava de x (vale decir, si $\frac{\partial^2 f\left(x_t,t\right)}{\partial x^2} < 0$),

---

[10] Este resultado se basa en el cálculo de

$$Var\left[\left(w_{t+\Delta t} - w_t\right)^2\right] = E\left[\left(w_{t+\Delta t} - w_t\right)^4\right] - \left(E\left[\left(w_{t+\Delta t} - w_t\right)^2\right]\right)^2 = 3\left(\Delta t\right)^2 - \left(\Delta t\right)^2 = 2\left(\Delta t\right)^2$$

y el pasaje al límite.

$E\left[dy_t\right]$ será negativa. Puesto que $\left\{x_t\right\}_{t\geq 0}$ es un proceso de Ito, dx_t se comporta como $\sqrt{dt}$ y $\left(dx_t\right)^2$ como dt, por lo que el efecto de la convexidad o la concavidad es de orden dt y no puede ignorarse al escribir el diferencial de y_t. El término adicional en (7) captura precisamente este efecto.

La ecuación (7) puede extenderse inmediatamente al caso de procesos multidimensionales, asumiendo que f es función de un vector nx1 X_t y de t: $f\left(X_t, t\right)$. μ_t y dW_t serán ahora vectores de nx1 y σ_t será una matriz de nxn. La versión multivariada del Lema de Ito puede escribirse entonces como:

$$dY_t = \left(\frac{\partial f(X,t)}{\partial t} + \mu_t^T \frac{\partial f(X,t)}{\partial X} + \frac{1}{2} tr\left[\sigma_t \sigma_t^T \frac{\partial^2 f(X,t)}{\partial X^2} \right] \right) dt + \frac{\partial f(X,t)}{\partial X}^T \sigma_t \; dW_t$$

donde

$\dfrac{\partial f(X,t)}{\partial X}$ es el vector de derivadas parciales de primer orden, $\begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{pmatrix}$,

$\dfrac{\partial^2 f(X,t)}{\partial X^2}$ es la matriz nxn de las derivadas de segundo orden, $\begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$, y

$tr\left[M\right]$ denota la traza de la matriz M.

VIII. Procesos de salto y procesos de difusión con saltos

Hasta ahora hemos considerado únicamente procesos de difusión, es decir, procesos estocásticos cuyas trayectorias son continuas con probabilidad uno. Dicha propiedad podría parecer a priori cuestionable en el caso de los precios de los instrumentos financieros: una simple observación revelaría que en éstos experimentan cambios abruptos en la práctica.[11] En particular, en el caso de instrumentos financieros sujetos a riesgo de crédito, la ocurrencia de “eventos crediticios” tales como renegociaciones o reestructuraciones de deuda, cambios en la calificación crediticia de los títulos o *defaults* da lugar a cambios abruptos en los precios. Por consiguiente, algunos modelos utilizados para valorar instrumentos financieros recurren a procesos de salto (generalmente procesos de Poisson o de Cox) o procesos de difusión con saltos para la valoración de instrumentos financieros.

Denótese por $T_1, T_2, ...$ los tiempos de ocurrencia de algún evento físico. La sucesión $(T_i)_{i\geq 1}$ es un **proceso (homogéneo) de Poisson** con intensidad h si los lapsos entre eventos, $T_{i+1} - T_i$, son independientes y exponencialmente distribuidos con parámetro h. De manera equivalente, si

$$x_t = \sum_i 1_{\{T_i \leq t\}}$$

cuenta el número de ocurrencias del evento en el intervalo $[0,t]$, $\{x_t\}_{t\geq 0}$ es un proceso homogéneo de Poisson con intensidad h si los incrementos $x_t - x_s$ son independientes y tienen una distribución Poisson con parámetro $h(t-s)$ para $s < t$:

$$\mathcal{P}\left[x_t - x_s = k\right] = \frac{1}{k!} h^k \left(t-s\right)^k e^{-h(t-s)}.$$

La mayor parte de los modelos de valoración de instrumentos sujetos a riesgo de *default* suponen que el momento en que ocurre el *default* es el primer salto en un proceso de Poisson con intensidad h (Giesecke, 2002). En ese caso, $\tau = T_1$ se distribuye exponencialmente con parámetro h, y la probabilidad de *default* viene dada por

$$F(T) = \mathcal{P}\left[\tau \leq T\right] = 1 - e^{-hT}$$

y la intensidad del *default* viene dada por

[11] Una pregunta interesante a este respecto es si las discontinuidades observadas son el resultado del caracter discreto del muestreo de los datos, o si constituyen evidencia genuina de que es erróneo representar la dinámica de las variables de estado y, en consecuencia, de las variables financieras relevantes, mediante procesos de difusión. El lector interesado puede remitirse a Bertsimas et al. (2000) y Aït – Sahalia (2002) para una discusión pormenorizada de esta cuestión técnica.

$$\lim_{j\to 0}\frac{1}{j}\mathcal{P}\left[\tau \in \left(t, t+j\right] \middle| \tau > t\right] = h$$

que puede también expresarse como

$$\mathcal{P}\left[\tau \in \left[t, t+dt\right) \middle| \tau > t\right] \approx hdt \,.$$

No obstante, en general los modelos de riesgo de *default* asumen que la intensidad del *default* varía a lo largo del tiempo, frecuentemente en función de las variaciones en el nivel de las tasas de interés.[12]

El proceso $\{x_t\}_{t\geq 0}$ es un **proceso no homogéneo de Poisson** con una intensidad determinística h_t si los incrementos $x_t - x_s$ son independientes y para $s < t$ se tiene que

$$\mathcal{P}\left[x_t - x_s = k\right] = \frac{1}{k!}\left(\int_s^t h_u du\right)^k e^{-\int_s^t h_u du} \,.$$

En este caso, la probabilidad de *default* viene dada por

$$\mathcal{P}\left[\tau \leq T\right] = 1 - \mathcal{P}\left[x_T = 0\right] = 1 - e^{-\int_0^T h_u du} \,.$$

Un **proceso de Cox** $\{x_t\}_{t\geq 0}$ con intensidad $\{h_t\}_{t\geq 0}$ es la generalización de un proceso no homogéneo de Poisson en el que la intensidad es aleatoria, pero adoptando la restricción de que, condicional a las realizaciones de h, $\{x_t\}_{t\geq 0}$ es un proceso no homogéneo de Poisson. Por este motivo, los procesos de Cox se denominan también "procesos condicionales de Poisson" o procesos de Poisson "doblemente estocásticos". La probabilidad condicional de *default* en ese caso es

[12] Una excepción la constituye el modelo de Jarrow y Turnbull (1995), en el que se asume que la intensidad del *default* permanece constante.

$$\mathcal{P}\left[\tau \leq T \left| (h_t)_{0\leq t\leq T} \right.\right] = 1 - \mathcal{P}\left[x_T = 0 \left| (h_t)_{0\leq t\leq T} \right.\right] = 1 - e^{-\int_0^T h_u du}$$

y a partir de la propiedad de las esperanzas iteradas se obtiene

$$\mathcal{P}\left[\tau \leq T\right] = E\left[\mathcal{Q}\left[\tau \leq T \left| (h_t)_{0\leq t\leq T} \right.\right]\right] = 1 - E\left[e^{-\int_0^T h_u du}\right].$$

Finalmente, es posible combinar los procesos de salto analizados en esta sección y los procesos de difusión reseñados anteriormente. Un **proceso de difusión con saltos** (*jump – difussion process*) es un proceso estocástico $\{x_t\}_{t\geq 0}$ que soluciona

$$dx_t = \mu(x_{t\_})dt + \sigma(x_{t\_})dw_t + dJ_t \tag{8}$$

donde J es un proceso puro de salto y $x_{t\_} = \lim_{s\uparrow t} x_s$.

El proceso de salto J puede ser "activado" de dos maneras posibles: los saltos pueden ser causados por un proceso de Poisson con intensidad h, o pueden ocurrir en momentos determinísticos del tiempo. La extensión de (8) para múltiples procesos de salto es inmediata, e implica solamente sumar diferentes procesos de salto en la fórmula anterior.

El Lema de Ito puede ser generalizado al caso de procesos de salto y procesos difusión con saltos. El lector interesado puede remitirse a Dixit y Pyndick (1994, cap. 3, sección 6).

Bibliografía

Äit – Sahalia, Y. (2002): "Maximum-Likelihood Estimation of Discretely-Sampled Diffusions: A Closed-Form Approximation Approach", Econometrica, 2002, 70, 223-262.

Bachelier, L. (1900): "Théorie de la spéculation", Gauthier – Villars, Paris. También en Annales Scientifiques de l'Ecole Normale Superieure, III -17, 21-86.

Bertsimas, D., Lauprete, G.J., Samarov, A. (2000): "Shortfall as a risk measure: properties properties, optimization and applications". Working paper, Sloan School of Management, MIT, Cambridge.

Cox, J., Ingersoll, J. y S. Ross (1985): "A Theory of the Term Structure of Interest Rates", Econometrica, 53(2), 385 – 407.

Cox, D. y H. Miller (1965): "The Theory of Stochastic Processes", Methuen, Londres.

Dixit, A. y R. Pindyck (1994): "Investment under Uncertainty", Princeton University Press, Nueva Jersey.

Dubra, J. y F. Echenique (2001): "Measurability Is Not About Information", Cowles Foundation Discussion Paper N° 1296, Cowles Foundation for Research in Economics, Yale University.

Giesecke, K. (2002): "Credit Risk Modeling and Valuation: and Introduction", Working Paper, Humboldt – Universität zu Berlin.

Jarrow, R. y S. Turnbull (1995): "Pricing Options on Financial Securities Subject to Default Risk", The Journal of Finance, 50, pp. 53 – 86.

Lamberton, D. y B. Lapeyre (1996): "Introduction to Stochastic Calculus Applied to Finance", Chapman & Hall, Londres.

Merton, R. (1970): "The Dynamic General Equilibrium Model of the Asset Market and its Application to the Pricing of the Capital Structure of the Firm", Working Paper, Sloan School of Management, M.I.T. Reprinted as chapter 11 in Merton (1992): "Continuous-Time Finance, Revised Edition", Basil Blackwell, Cambridge, Mass.

Mikosch, T. (1999): "Elementary Stochastic Calculus with Finance in View", Advanced Series on Statistical Science & Applied Probability, World Scientific Publishing Co.

Munk, C. (2002): "Fixed Income Analysis – Securities, Pricing and Risk Management in the Markets for Interest Rates", Department of Accounting, Finance & Law, University of Southern Denmark.

Shreve, S. (1997): "Stochastic Calculus and Finance", Versión Preliminar, no publicada.

Vasicek, O. (1977): "An Equilibrium Characterization of the Term Structure", Journal of Financial Economics, 5, pp. 177 – 188.