

INSTITUTO TECNOLÓGICO Y DE ESTUDIOS
SUPERIORES DE MONTERREY

CAMPUS CIUDAD DE MÉXICO



MODELOS DE TIEMPO CONTINUO PARA TASAS
DE INTERÉS SPOT EN MÉXICO

DOCTORADO EN CIENCIAS FINANCIERAS

TESIS PRESENTADA POR
ELIZABETH ORTEGA BENÍTEZ

ASESOR

DR. JOSÉ ANTONIO NUÑEZ MORA

LECTORES

DR. ARTURO LORENZO VALDÉS

DR. ANTONIO RUÍZ PORRAS

DRA. BÁRBARA TREJO BECERRIL

JUNIO 2008

Agradecimientos...

Quiero agradecer el apoyo que me brindaron mis padres durante toda mi vida y en especial durante el posgrado, porque me entendieron y me ayudaron a que todo estuviera no bien, sino excelente. También quiero agradecer a quien es ahora mi esposo porque me dio su apoyo incondicional, porque es una parte esencial en mi vida y ha estado ha mi lado

Además quiero dedicar este trabajo a mis dos hermanos. A ti Angy porque siempre me das esa alegría que tienes en tu rostro y que me hace sentir bien, espero algún día pueda compensar todo lo que me has dado. Y a ti Alfredín porque fuiste, eres y serás el motor de mi vida, y aunque ya no estés con nosotros siempre estarás en mi corazón.

A todos los quiero.

RESUMEN

Uno de los factores más importantes en una economía son las tasas de interés, en las últimas décadas se han desarrollado diversas metodologías de estimación para las tasas de interés a corto plazo. En este trabajo se comparan los modelos paramétricos de tiempo continuo de la tasa de interés spot con un modelo no paramétrico, aplicado a la tasa cetes del mercado secundario en México. Se consideran a los modelos paramétricos de Vasicek, CIR, Brennan y Schwartz, Chan, entre otros. El trabajo se divide en dos partes esencialmente, la primera es obtener y comparar las densidades marginales que se generan con una familia de modelos paramétricos y el modelo no-paramétrico. La segunda parte es hacer un análisis análogo cuando los modelos paramétricos de tiempo continuo tienen asociado un proceso de Poisson.

ÍNDICE

| | |
|--|----|
| INTRODUCCIÓN..... | 6 |
| 1. DEFINICIÓN DEL PROBLEMA DE INVESTIGACIÓN..... | 8 |
| 2. OBJETIVOS..... | 9 |
| 3. REVISIÓN LITERARIA..... | 10 |
| 4. JUSTIFICACIÓN..... | 14 |
| 5. METODOLOGÍA..... | 19 |
| 5.1 FUNCIONES KERNEL..... | 20 |
| 5.2 ECUACIONES FORWARD Y BACKWARD DE KOLMOGOROV..... | 24 |
| 5.3 SUPUESTOS DEL MODELO..... | 27 |
| 5.4 HIPÓTESIS NULA, HIPÓTESIS ALTERNATIVA Y PRUEBA ESTADÍSTICA..... | 30 |
| 5.5 DENSIDAD MARGINAL PARA EL MODELO NO-PARAMÉTRICO..... | 36 |
| 5.6 DENSIDAD MARGINAL PARA LOS MODELOS PARAMÉTRICOS..... | 38 |
| 5.7 DENSIDAD DE TRANSICIÓN PARA LOS MODELOS PARAMÉTRICOS CON SALTOS..... | 39 |
| 6. MODELO NO-PARAMÉTRICO..... | 42 |
| 7. MODELOS PARAMÉTRICOS..... | 46 |

| | |
|--|-----------|
| 8. PARTE I. ANÁLISIS DE LOS MODELOS DE TIEMPO CONTINUO SIN SALTOS DE POISSON..... | 52 |
| 8.1 ANÁLISIS MATEMÁTICO SIN SALTOS DE POISSON..... | 52 |
| 8.2 RESULTADOS..... | 57 |
| | |
| 9. PARTE II. ANÁLISIS DE LOS MODELOS DE TIEMPO CONTINUO CON SALTOS DE POISSON. | 59 |
| 9.1 ANÁLISIS MATEMÁTICO..... | 60 |
| 9.2 RESULTADOS..... | 72 |
| | |
| 10. CONCLUSIONES..... | 73 |
| | |
| 11. BIBLIOGRAFÍA..... | 75 |
| | |
| APÉNDICE 1..... | 82 |

Introducción.

Para la economía de cualquier país una parte importante la constituye su mercado financiero. Es a través del mercado financiero que las instituciones bancarias hacen circular la moneda local y regulan la inflación, se financian actividades del estado y del sector privado; se canalizan los pagos del sistema bancario, el comercio y la industria. Además, los mercados financieros entregan una gran cantidad de información, que es considerada por las personas y empresas a la hora de tomar sus decisiones, para lo cual se apoyan en una serie de indicadores y precios de mercado.

Uno de los precios de mercado más importantes en una economía son las tasas de interés, pues no sólo permiten conocer la relación intertemporal del valor de los bienes y el consumo, sino que además cuentan con una cierta noción del valor del dinero en el tiempo, del valor presente y futuro de los activos, contienen información acerca de la actividad económica, de las perspectivas de crecimiento y de las expectativas de inflación futura.

La relación entre la tasa de interés y fechas futuras en el tiempo se conoce como estructura de tasas de interés, y se puede representar mediante una curva de tasas de descuento o una curva de tasas internas de retorno asociadas a algún instrumento o una curva de tasas forward, etc. Ya que conocer el valor de la tasa de interés a diversos plazos futuros resulta fundamental a la hora de valorizar activos, oportunidades de consumo y proyectos de inversión. La estimación de la tasa de interés se reduce a encontrar aquel método que mejor la ajuste a los datos observados y, en el proceso, cuente con una interpretación económica sólida.

En las últimas décadas, varios autores han desarrollado metodologías de estimación de la estructura de tasas de interés, destacando los modelos de Nelson y Siegel (1987) y Svensson (1994). Estos modelos presuponen la existencia y estabilidad en los datos, pero generalmente la información disponible no es suficiente o no posee la calidad adecuada para estimar una tasa de interés.

Así, se hace necesaria una metodología que sea capaz no sólo de realizar una estimación confiable de la forma en que evoluciona la estructura de tasas de interés en el tiempo, sino que, además, pueda extraer la información necesaria para ello, a partir de las observaciones del mercado.

De esta manera, los modelos de tasas de interés dinámicos comienzan a resultar atractivos, pues su objetivo, además de ajustarse a los datos observados en un instante determinado del tiempo, es mantener una consistencia intertemporal de sus estimaciones de tasas. En esta línea de investigación se encuentran el trabajo de Merton (1973) y el modelo de Vasicek (1977).

Este trabajo surge de la idea de dar una aportación en el área de estimación de las tasas de interés en México, se pretende probar, validar o rechazar la aplicación de los modelos de tasas de interés en tiempo continuo en el mercado secundario de México, a través de una comparación con una serie de datos observados.

Este trabajo se compone de diez capítulos, en el primero y segundo se muestra la definición del problema y los objetivos de este trabajo. En el tercero, se hace una revisión literaria de temas relacionados que se han hecho recientemente. El cuarto contiene una justificación de porque se puede llevar a cabo esta metodología. El quinto define algunos conceptos matemáticos, así como los supuestos que se tomarán para este trabajo. El sexto y séptimo detallan los modelos que son utilizados y algunas de sus características. Los capítulos ocho y nueve hacen el análisis matemático y muestran los resultados para los modelos sin y con saltos de Poisson, respectivamente. Finalmente el capítulo diez muestra las conclusiones de este trabajo.

1. Definición del Problema de Investigación.

En este trabajo se pretende comparar los modelos paramétricos de tiempo continuo de la tasa de interés spot con un modelo no paramétrico, aplicado a la tasa cetes del mercado secundario en México. El trabajo se divide en dos partes esencialmente, la primera es obtener y comparar las densidades marginales que se generan con una familia de modelos paramétricos y el no-paramétrico, esta parte implica el desarrollo de la metodología a utilizar. La segunda parte es hacer el análisis cuando los modelos paramétricos de tiempo continuo tienen asociado un proceso de Poisson, para este caso se obtienen las funciones de densidad de probabilidad para cada uno de los modelos paramétricos de la familia de tasa de interés.

Finalmente para ambos casos, y utilizando nuestra serie de datos se obtienen los resultados, con los cuales se prueba que los modelos de tiempo continuo usados no reflejan el comportamiento de la tasa corta para el caso de México, y también se prueba que un modelo de tiempo continuo con saltos de Poisson puede ser una representación aceptable de los datos.

2. Objetivos

En la literatura existen diferentes modelos de tiempo continuo para tasas de interés. En este trabajo se pretende discernir entre los modelos de tiempo continuo con un modelo no-paramétrico. La forma de hacerlo posible es comparando los parámetros de densidad implícitos con la densidad no-paramétrica estimada. El origen principal de rechazo de los modelos existentes es la extraña no-linealidad en el drift por Aït-Sahalia (1996). Alrededor de su media, donde el drift es casi cero, la tasa spot se comporta como una caminata aleatoria. El drift entonces tiene reversión a la media cuando se va alejando de la media. La volatilidad es más alta cuando se va alejando de la media.

También se pretende hacer el mismo análisis pero considerando los modelos de tiempo continuo con saltos de Poisson y de igual forma se comparan los parámetros de la función de probabilidad de transición implícitos. En este caso se provee una aproximación cerrada de la función de densidad de probabilidad de transición de procesos multivariados con difusión de saltos para cada uno de los modelos paramétricos de tiempo continuo de la tasa de interés.

3. Revisión Literaria

Principalmente se revisaron artículos de Yacine Aït-Sahalia (1996) quien publica “Testing continuous-Time Models of the Spot Interest Rate”, en donde rechaza que algunos de los modelos paramétricos en tiempo continuo tienen la misma densidad que la densidad estimada de una serie de datos observados. Otro artículo del mismo autor (1996) es “NonParametric Pricing of Interest Rate Derivative Securities” en donde utiliza procesos que cumplan con la propiedad markoviana, lo cual reduce la elección de modelos bien especificados, y compara a través de simulaciones las estimaciones de las diferentes densidades de transición de los modelos paramétricos de tiempo continuo con saltos de Poisson con la estimación de la densidad de transición de una serie de datos observados.

Otro artículo sobre tasas de interés es “A Theory of the Term Structure of Interest Rates” de John C. Cox, Jonathan E. Ingersoll y Ross (1985), el cual usa un activo en un estado de equilibrio general que valúa el modelo para estudiar la estructura de tasas de interés. En este modelo los factores como, pagos anticipados, aversión al riesgo, alternativas de la inversión, y preferencias sobre el cronometraje de consumo, son utilizados en la determinación de precios. Muchos de los factores tradicionalmente mencionados como la influencia en la estructura de plazos son incluidos en un camino que es totalmente compatible con el maximizar el comportamiento y expectativas racionales.

Recientes artículos en la literatura de estructura de plazos han modelado cambios de regímenes discretos en el proceso de tasa de interés spot como Hamilton (1988), y Naik y Lee (1993). Estos modelos típicamente suponen que un proceso de tasa de interés puede cambiar aleatoriamente entre uno o más regímenes, por ejemplo, de un régimen con media-baja normal a uno con media-alta. Ellos tienen la misma función para la media o varianza pero diferentes valores en los parámetros para cada régimen.

La convergencia a aproximaciones discretas en modelos de tiempo continuo han sido estudiadas en la literatura financiera por He (1990), Nelson (1990), y Nelson y Ramaswamy (1990). La aproximación puede ser únicamente exacta en pocos casos, típicamente cuando su función de verosimilitud es conocida en su forma cerrada.

Stanton (1997) utiliza una técnica no paramétrica alternativa para estimar los procesos continuos de la difusión del tiempo, que se observan en los intervalos discretos. Stanton aplica la metodología a tres y seis meses la Treasury Bill, para la estimación de las funciones drift y de difusión de la tasa corta. En la difusión estimada hay una evidencia fuerte de no linealidad en la función del drift.

Jiang (1998) desarrolla otro modelo no paramétrico de la estructura de plazos, permite una máxima flexibilidad en los datos, el cual se basa solamente sobre un proceso de la tasa corta que admite solamente tasas de interés positivas y una prima al riesgo de mercado que imposibilita oportunidades de arbitraje. La densidad marginal de la tasa de interés a corto plazo, tanto como la trayectoria histórica de la estructura de plazos, se utiliza para permitir una estimación robusta de la estructura de plazo. Se estima el modelo usando bonos gubernamentales de Estados Unidos, sus resultados sugieren que la mayoría de los modelos tradicionales de la tasa de interés spot no están bien especificados y que el modelo no paramétrico genera diferentes estructuras de plazo y precios de mercado de la tasa de interés.

Matt Pritsker (1997), escribió un artículo llamado “NonParametric Density Estimation and Tests of Continuous Time Interest Rate Models”, donde usa el modelo de Vasicek (1977) para estudiar el funcionamiento de las estimaciones de densidad de la distribución ergódica de la tasa de interés instantánea. Este modelo permite analizar el MISE¹ del estimador Kernel como una función de persistencia,

¹ *Mean Integrated Squared Error* (MISE) es una medida de distancia, es la diferencia al cuadrado esperada entre la densidad estimada y la densidad teórica sobre un intervalo, y está dada por:

la varianza de la distribución ergódica, la frecuencia de la muestra, y la amplitud del ancho de banda del Kernel. El resultado principal del artículo es que la persistencia tiene un impacto importante sobre la selección del ancho de banda óptima y sobre una muestra finita. A menudo, la prueba se rechaza usando valores asintóticos críticos.

Pearson (1994), propone un método empírico que utiliza la densidad condicional de las variables de estado y prueba un modelo de estructura de plazos, usando datos de bonos cuponados descontados. El método es aplicado a una extensión del modelo de CIR (1985) de dos factores. Muestran que utilizando este tipo de datos el modelo de CIR es rechazado con una prueba de máxima verosimilitud.

Otro artículo que aporta herramientas importantes es “Nonparametric Specification Testing for Continuous-Time Models with Application to Spot Interest Rates” de Yongmiao Hong y Haitao Li (2002). Se proponen dos pruebas basadas en densidades de transición para modelos continuos no paramétricos. En contraste con la densidad marginal, la densidad de transición puede capturar la dinámica completa de un proceso de difusión, y en particular, una función de densidad de transición puede distinguir procesos con la misma densidad marginal, aunque dichos procesos cuenten con diferentes densidades de transición. El factor difusión de los modelos pueden ser aplicadas a una amplia clase de modelos dinámicos económicos, como modelos de series de tiempo discretos, modelos de difusión no homogéneos de tiempo, modelos de volatilidad estocásticos, los modelos de salto-difusión, y el término de multifactores en la estructura de los modelos. Esta metodología fue aplicada a tasas de intereses de Euro-dólar diarias, donde se rechazan algunos modelos de tasa instantánea, incluyendo aquellos con flujos no lineales que algunas pruebas existentes no pueden rechazar después de la corrección de distorsiones de la

$$MISE = E \int_u [\hat{\pi}(u) - \pi(u)]^2 du$$

donde $\pi(u)$ es la función de densidad teórica, y $\hat{\pi}(u)$ es la función de densidad estimada.

muestra. Se muestra que los métodos no paramétricos son un instrumento confiable y poderoso, para analizar datos financieros.

Un número de métodos econométricos han ido desarrollados para estimar los parámetros del proceso sin requerir que las observaciones sean continuas. Algunos métodos están basados en simulaciones como Duffie y Singleton (1993); Guriéroux, Monfort y Renault (1993); Gallant y Tauchen (1992); Honoré (1997) hizo aplicaciones; al igual que Andersen y Lund (1997). Otros estudios con el Método Generalizado de Momentos fueron hechos por Hansen y Scheinkman (1991); Bibby y Sorensen (1995). Duffie y Glynn (1997) utilizaron un proceso de muestras aleatorias para generar momentos condicionales.

La difusión de saltos ha sido una herramienta muy útil para describir fenómenos económicos, se han utilizado métodos probabilístico tales como estimación de máxima verosimilitud, métodos de estimaciones Bayesianas; sin embargo, estos métodos no han podido ser usados en la práctica debido a la falta de funciones de probabilidad con formas cerradas. Debido a estas complicaciones se han propuesto métodos alternativos de estimación, por ejemplo, Gouriéroux y Monfort (1996) propusieron una estimación basada en simulaciones; Hansen y Scheinkman (1995) una estimación generalizada en el método de los momentos; Aït-Sahalia (1996) una estimación no-paramétrica; Singleton (1996) un cálculo numérico o aproximación analítica de Fourier de las funciones características de la densidad de transición.

Otros estudios se han desarrollado para calcular la función de transición, entre ellos el de Dacunha-Castelle y Florens-Zmirou (1986), ellos calcularon expresiones que involucran funcionales al movimiento Browniano, y lograron hacer las simulaciones. Pero aún así, éstos no producían una forma cerrada para poder maximizar los parámetros del proceso.

4. Justificación.

La teoría financiera de tiempo continuo ha desarrollado herramientas extensivas para valorar los títulos de derivados cuando los subyacentes de los activos comerciables o los factores no comerciables siguen una ecuación diferencial estocástica. Sin embargo, no existe una ecuación diferencial estocástica específica apropiada. Por ejemplo, muchos procesos diferentes han propuesto capturar la dinámica de la tasa spot instantánea. Estos modelos generalmente son exclusivos entre sí cuando son usados para valorar títulos de derivados, es decir, generan precios significativamente diferentes. Este trabajo pretende desarrollar una metodología simple para probar su especificación y evaluar varios modelos para la tasa de interés spot corta.

El riesgo de las tasas de interés a corto plazo es uno de los temas principales e importantes en las valuaciones en los mercados financieros. Varios modelos han sido propuestos para explicar su comportamiento como cualquier tema en finanzas. Muchos de los modelos más populares actualmente son usados por investigadores académicos y practicantes, éstos han sido desarrollados en escenarios de tiempo continuo, los cuales proveen una buena aproximación del comportamiento dinámico de la tasa a corto plazo.

Las funciones de difusión y de forma general los procesos Markovianos en tiempo continuo son especificados generalmente en la economía y finanzas por su evolución sobre instantes infinitesimales; es decir, la ecuación diferencial estocástica es seguida por un vector de estado. Sin embargo, como la mayoría de las estrategias dependen de muestras de datos discretos, se necesita ser capaz de inferir las implicaciones de la evolución del tiempo infinitesimal de los procesos para intervalos de tiempo finito, como es el caso de nuestra muestra. Desafortunadamente, la función de transición casi siempre es desconocida.

Un modelo de tiempo continuo en finanzas típicamente se apoya en uno o más procesos de difusión estacionarios $\{X_t, t \geq 0\}$, con una dinámica representada por una ecuación diferencial estocástica de Itô:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t \quad (4.1)$$

Donde X_t es la variable de estado y $\{W_t, t \geq 0\}$ es un movimiento browniano estándar. Las funciones $\mu(X_t)$ y $\sigma^2(X_t)$ son respectivamente, el drift y la función de difusión del proceso, éstas veces se encuentran parametrizadas.

$$\mu(X) = \mu(X, \theta) \quad \text{y} \quad \sigma^2(X) = \sigma^2(X, \theta), \quad \text{donde} \quad \theta \in \Theta \subset R^K \quad (4.2)$$

La elección del drift paramétrico y la familia de difusión puede hacerse de manera arbitraria. Por ejemplo, varios modelos de tasas de interés especifican un drift con reversión a la media lineal, digamos $\mu(X, \theta) = \beta(\alpha - X)$, donde la tasa spot se mueve en la dirección de su media α a una velocidad β . La especificación de la función de difusión, sin embargo, tiende a variar ampliamente.

En este trabajo, se busca una caracterización del proceso

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t \quad (4.1)$$

el cual depende de cantidades que pueden ser estimadas bajo los datos observados. Cualquier parametrización de $\mu(X_t)$ y $\sigma^2(X_t)$ corresponden a una parametrización de la densidad marginal. La idea básica es usar el mapeo entre el drift y la difusión contra la densidad marginal para probar la especificación del modelo.

Existe una lista exhaustiva de modelos propuestos de la dinámica de tiempo continuo para la tasa corta, para discriminar entre estos modelos, nos enfocamos en

la siguiente pregunta: ¿Cuándo un modelo paramétrico es un modelo adecuado para describir los movimientos de la tasa de interés corta? Para responder a esta pregunta, vamos a utilizar la hipótesis nula, la cual dado un modelo paramétrico, establecerá por lo tanto que existen valores paramétricos para los cuales el modelo paramétrico considerado es un representación aceptable de los datos. La hipótesis alternativa, por el contrario, afirma que ningún valor paramétrico es capaz de reproducir las funciones exactas de μ y σ^2 .

La Tabla 4.1 muestra la lista de modelos de tasa de interés considerados en este trabajo:

| Modelo Paramétrico | $\mu(X, \theta)$ | $\sigma(X, \theta)$ |
|------------------------------|---|---|
| Vasicek (1977) | $\alpha_0 + \alpha_1 X$ | β_0 |
| Cox, Ingersoll y Ross (1985) | $\alpha_0 + \alpha_1 X$ | $\beta_1 \sqrt{X}$ |
| CIR VR | - | $\beta_1 X^{3/2}$ |
| Brennan & Schwartz (1982) | $\alpha_0 + \alpha_1 X$ | $\beta_2 X$ |
| * Brennan & Schwartz | $\alpha_0 + \alpha_1 X$ | $\beta_1 X^{3/2}$ |
| Chan (1992) | $\alpha_0 + \alpha_1 X$ | $\beta_2 X^{\beta_3}$ |
| General Drift. | $\alpha_0 + \alpha_1 x X + \alpha_2 X^2 + \alpha_3 / X$ | $\beta_0 + \beta_1 X + \beta_2 X^{\beta_3}$ |
| Merton | α_0 | β_0 |
| Dothan | - | $\beta_1 X$ |
| GBM | $\alpha_1 X$ | $\beta_1 X$ |
| CEV | $\alpha_1 X$ | $\beta_2 X^\gamma$ |

Tabla 4.1

Es deseable ser capaz de probar la especificación de los procesos de la tasa spot actual sin hacer uso de las observaciones sobre la colección de derivados. La prueba estadística utilizada se basa en la siguiente idea: una propiedad esencial del proceso de difusión es que éste es caracterizado por sus primeros dos momentos de tiempo continuo, las funciones de μ y σ^2 . Entonces cada modelo paramétrico para las funciones de μ y σ^2 , tales como CIR, implicaran una familia de funciones de densidad de las observaciones de las tasas de interés.

En otras palabras, la prueba estadística consiste en la comparación de la densidad implícita en el modelo paramétrico y el estimador no paramétrico, incluso si el modelo paramétrico no está bien especificado. Así, la densidad de los estimadores paramétricos y no-paramétricos estarán próximos si el modelo paramétrico está bien especificado. Aquí es cuando se tiene que introducir una medida de comparación, a la cual denotaremos como M la cual medirá la distancia entre las dos densidades estimadas y estará dada por la siguiente ecuación:

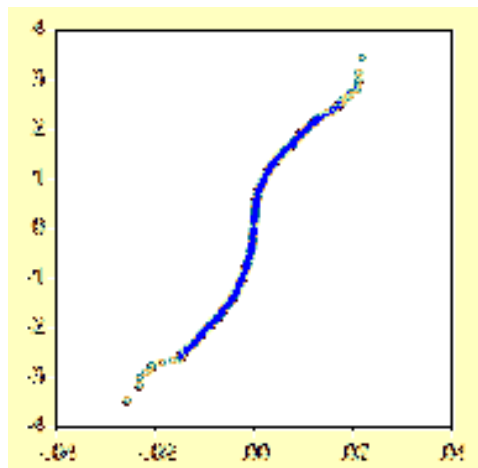
$$M \equiv \min_{\theta \in \Theta} \int_x^{\bar{x}} (\pi(X, \theta) - \pi_0(X))^2 \pi_0(X) du \quad (4.3)$$

La prueba estadística se rechaza cuando el valor de M es suficientemente grande. La estacionariedad de la densidad marginal del proceso no muestra toda la información disponible en los datos. Para la prueba, un rechazo sobre la densidad marginal es cuando M es muy grande, lo que invalida al vector paramétrico de la función del drift y difusión.

La segunda parte del análisis se basa en agregar un proceso de difusión de saltos a los modelos descritos en la Tabla 1.1. ¿Por qué utilizar saltos? Comenzando con un modelo más general, un proceso de difusión con saltos puede aproximar un mejor generador de datos, los modelos de saltos pueden tener un gran impacto sobre

periodos cortos de tiempo. Modelos explícitos de saltos mejoran nuestra interpretación para fenómenos llamados “eventos raros”.

Generalmente la función de distribución para estos casos no es normal, la cresta de la distribución empírica con estos eventos es superior a la de la distribución normal, lo que implica que las colas de la distribución empírica son más pesadas, es decir, se tiene una mayor probabilidad de que ocurran valores extremos en comparación con la distribución normal. La Gráfica 4.1 muestra la gráfica cuantil-cuantil de los rendimientos de las tasas, la cuál claramente que no es una línea recta, lo que refleja una presencia de colas pesadas.



Gráfica 4.1

Para estos casos la función que se pretende obtener en cada uno de los modelos paramétricos de tasas de interés es la función de probabilidad de transición, a la que se le aplicará el estadístico sugerido por Ait Sahalia (1996) siguiente:

$$M \equiv \min_{\theta \in \Theta} E[p(X, \theta) - p_0(X)]^2 \quad (4.4)$$

en este caso, se medirá la distancia entre las dos funciones de probabilidad.

5. Metodología

Se pretende probar empíricamente varios modelos de difusión de un factor para la tasa cetes del mercado secundario en México. Entonces identificamos los orígenes del rechazo de los modelos existentes. La linealidad del drift impuesta en la literatura aparece como el principal origen de la no especificación.

Para hacer esta prueba se utiliza el Método Generalizado de los Momentos para el modelo no paramétrico, y mediante los dos primeros momentos y utilizando el kernel Gaussiano se puede obtener la aproximación de la función de densidad marginal. La ventaja de usar esta técnica es que la distribución de los cambios de la tasa de interés no necesariamente tiene que ser normal.

Para los modelos paramétricos se obtiene la función de densidad marginal y la función de densidad de probabilidad de transición, la primera se utiliza para los modelos paramétricos que no consideran saltos de Poisson, y la segunda se considera cuando se hace el análisis con modelos paramétricos con saltos de Poisson. Para obtener la función de densidad marginal y la función de densidad de probabilidad de transición se hace uso de las ecuaciones *forward* y *backward* de Kolmogorov.

En este capítulo primero se especifican algunos tipos de Kernels que existen en la literatura y que pueden ser utilizados en el modelo no-paramétrico, y se explica por qué se eligió el kernel Gaussiano. Además se hace un esbozo de las ecuaciones de *forward* y *backward* Kolmogorov, en el apéndice se muestra sólo una de las soluciones de éstas ecuaciones. Se dan los supuestos que se tienen y se deben cumplir para esta metodología, así como se detalla la prueba estadística que se utiliza para la comparación de las funciones de densidad marginal, o en su caso, función de densidad de probabilidad, para los modelos paramétricos y el modelo no paramétrico, respectivamente. Se muestra la función de densidad marginal que se

utiliza para el modelo no-paramétrico, así como la función de densidad marginal para los modelos paramétricos sin saltos de Poisson. Y finalmente se muestra la función de densidad de probabilidad de transición para los modelos paramétricos con saltos de Poisson.

En la siguiente sección “Funciones Kernel” se explica que significa una función y se dan algunas funciones de la literatura, además de las propiedades que deben cumplir dichas funciones.

5.1 Funciones Kernel

La función de densidad de probabilidad es un concepto fundamental en estadística. Se considera cualquier variable X_i que tiene una función de densidad de probabilidad f . La función f da una descripción natural de la distribución de X_i . Si se tiene un conjunto de datos observados con una función de densidad de probabilidad desconocida, entonces se puede estimar una función de densidad.

Se asume que los datos son descritos por una familia de parámetros, por ejemplo la distribución normal tiene como parámetros a la media y la varianza. La densidad estimada es la construcción de un estimador de la función de densidad de los datos observados. La distribución estimada más simple es el histograma, éste es sensible a la elección de cambios, pero no es continua.

La suavidad en las funciones se hace poniendo menor peso en observaciones que están más alejadas desde el punto que están siendo valuadas. Más técnicamente, la densidad kernel estimada de una serie X en un punto x es estimada por:

$$f(x) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x - X_i}{h}\right) \quad (5.1)$$

Donde N es el número de observaciones, h es el ancho de banda (o parámetro de suavizamiento) y $K(\cdot)$ es la función kernel ponderadora cuya integral es uno. El ancho de banda controla la suavidad de la densidad estimada, a menor ancho de banda, el estimador es más suave. Existen diferentes formas de calcular el ancho de banda, uno de ellos es

$$h = 0.9kN^{-1/5} \min\left(s, \frac{R}{1.34}\right) \quad (5.2)$$

Donde N es el número de observaciones, s es la desviación estándar, y R es el rango intercuantil de la serie.

Aït Sahalia (1996) sugiere el siguiente ancho de banda²:

$$h_n = c_n n^{-1/4.5} \quad (5.3)$$

donde $c_n = c$ veces la desviación estándar de la serie de tiempo de la tasa de interés spot dividida por $\ln(n)$. Donde c es elegido minimizando la media integrada del error cuadrado del estimador. Utilizando este ancho de banda se obtiene un intervalo de confianza asintótico³.

En la literatura existen diferentes tipos de *kernel*, éstas son funciones ponderadoras que determinan la forma de los choques. Algunas funciones son:

Normal (Gaussiano):

$$f(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right)$$

Epanechnikov

$$f(u) = \frac{3}{4}(1-u^2)I(|u| \leq 1)$$

² Aït Sahalia (1996). "Testing Continuous-Time Models of the Spot Interest Rate".

³ Härdle Wolfgang, Müller Marlene.(2004) "Nonparametric and Semiparametric Models".

| | |
|------------|---|
| Triangular | $f(u) = \frac{1}{2}(1- u)I(u \leq 1)$ |
| Uniforme | $f(u) = \frac{1}{2}I(u \leq 1)$ |
| Coseno | $f(u) = \frac{\pi}{4} \cos\left(\frac{\pi}{2}u\right)I(u \leq 1)$ |
| Cuártica | $f(u) = \frac{15}{16}(1-u^2)^2 I(u \leq 1)$ |

Donde I es la función indicadora, es decir,

$$I_S(x) = \begin{cases} 1 & \text{si } x \in S. \\ 0 & \text{si } x \notin S. \end{cases} \quad (5.4)$$

Donde S es un subconjunto medible contenido en \mathfrak{R} , y se cumple que la integral de ésta función es la medida del conjunto S ,

$$\int I_S(x) = \mu(S) \quad (5.5)$$

Conociendo éstas funciones de densidad del kernel, se hizo el análisis y se obtuvo que el ancho de banda no variaba tanto de un kernel a otro, pero la gráfica de la función de densidad kernel más suave y continua fue con el kernel Gaussiano, por esta razón se optó por considerar dicha función.

La función de densidad kernel Gaussiano está dada por:

$$f(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right) \quad (5.6)$$

Donde $u = \frac{x - x_i}{h}$, h es el ancho de banda y x_i son los valores de la variable independiente en los datos, y x el valor de la variable independiente para el cual se busca un estimador.

La función de densidad kernel cumple las siguientes propiedades:

1. Las observaciones x_1, x_2, \dots, x_n son independientes e idénticamente distribuidas.
2. La función kernel $K(\cdot)$ es una función simétrica alrededor de cero y satisface:
 - a. $\int K(u)du = 1$;
 - b. $\int u^2 K(u)du \neq 0$;
 - c. $\int K^2(u)du < \infty$.
3. Las derivadas de segundo orden de f son continuas y acotadas en alguna vecindad de x .
4. $h = h_n \rightarrow 0$ cuando $n \rightarrow \infty$.
5. $nh_n \rightarrow \infty$ cuando $n \rightarrow \infty$.
6. Sea $K(\cdot)$ el kernel simétrico tal que

$$\int u^j K(u)du = \begin{cases} 1 & \text{si } j = 0 \\ 0 & \text{si } j = 1, \dots, r-1, \quad r \geq 2 \\ < \infty & \text{si } j = r \end{cases}$$

7. Las derivadas de orden r de la función de densidad son continuas.
8. Sea $K(\cdot)$ una función de la clase de todas las medidas acotadas de Borel en valores reales, $K(u)$ tal que:

- a. $\int K(u)du = 1$,
- b. $\int |K(u)|du < \infty$,
- c. $|u| \cdot |K(u)| \rightarrow 0$ cuando $|u| \rightarrow \infty$,

- d. $\sup |K(u)| < \infty$,
 - e. $\int K^2(u)du < \infty$.
9. $f(\cdot)$ es continua en cualquier punto x_0 y $\int |f(u)|du < \infty$.
 10. $\sum_{n=1}^{\infty} \exp(-\alpha nh) < \infty$ para todo $\alpha > 0$.
 11. $nh(\log n)^{-1} \rightarrow \infty$ cuando $n \rightarrow \infty$.
 12. La función característica $\phi(t)$ de $K(\cdot)$ es absolutamente integrable.
 13. $f(\cdot)$ es uniformemente continua en \mathfrak{R}^1 .
 14. $nh^2 \rightarrow \infty$ cuando $n \rightarrow \infty$.
 15. $(nh)^{1/2} h^2 \rightarrow 0$ cuando $n \rightarrow \infty$.
 16. $\{x_i, i = 0, \pm 1, \dots\}$ es un proceso estocástico de valores reales estrictamente estacionario.

Considerando estas propiedades y verificando que se satisfaga cada una de ellas, se puede proseguir al análisis del modelo no paramétrico.

A continuación se explica acerca del origen de las ecuaciones de *forward* y *backward* de Kolmogorov, y se da un esbozo de una de las metodologías para resolver dichas ecuaciones.

5.2 Ecuaciones Forward y Backward de Kolmogorov.

Muchos de los problemas de la ciencia y la ingeniería pueden ser formulados con procesos Markovianos y tienen asociado la ecuación de “*Fokker-Planck o forward Kolmogorov*”, ésta gobierna la evolución de la función de densidad de probabilidad de transición.

Se puede llegar a la solución de esta ecuación por varios métodos. Una solución es con funciones de densidad de probabilidad exactas, únicamente existe para los sistemas más simples, como sistemas escalares, sistemas lineales, y algunas soluciones estacionarias para sistemas conservativos.

Las soluciones analíticas para problemas pasajeros están limitados a sistemas escalares. Estas limitaciones han dado alcance a un gran número de aproximaciones o métodos analíticos locales tales como linealización equivalente, métodos de perturbación, promedios estocásticos, etc. Los métodos de simulación Monte Carlo son aplicables a sistemas más generales, pero requiere demasiados recursos computacionales, especialmente para cálculos exactos de probabilidades para las colas de una distribución, tal como son a veces requeridos en determinar probabilidades de incumplimiento de sistemas de ingeniería. Otras soluciones para las ecuaciones de Kolmogorov, *forward* y *backward*, han sido encontradas usando métodos Galerkin, métodos de elementos finitos o métodos de diferencias finitas.

Muchos sistemas estocásticos pueden ser modelados con un vector de ecuaciones diferenciales estocásticas de Itô de dimensión n , de la forma

$$dX = m(X)dt + G(X)W_t \quad (5.7)$$

Donde $X \in \mathfrak{R}^n$ es un proceso estocástico. $m(X)$ y $G(X)$ son el vector del drift y la matriz de difusión, respectivamente. W_t es un vector de dimensión m de ruido blanco no correlacionados. El vector de ruido blanco es completamente definido por el primero y segundo momento de sus componentes, dados por:

$$\begin{aligned} E[W_i(t)] &= E[W_j(t)] = 0, & E[W_i(t_1)W_j(t_2)] &= 2D_i\delta(t_2 - t_1), \\ E[W_i(t)W_j(t)] &= 0, \end{aligned} \quad (5.8)$$

Donde D_i es la densidad espectral de la i -ésima excitación y $\delta(\cdot)$ es la función delta de Dirac. El sistema mencionado forma un vector de procesos Markovianos en \mathfrak{R}^n , el comportamiento de los cuales es completamente determinado por la función de densidad de probabilidad de transición $p(X, t | X_0)$. La función de densidad de probabilidad de transición es proporcional a la probabilidad de comenzar en un elemento diferencial $(X, X + dX)$ del plano en el tiempo t , habiendo comenzado en X_0 en el tiempo cero y satisfaciendo ambas ecuaciones de Kolmogorov, *forward* y *backward*.

La ecuación de *Fokker-Planck* asociada a la ecuación (5.7) puede ser derivada como:

$$dX_t = m(X)dt + G(X)dB_t \quad (5.9)$$

Donde B_t es un proceso de Wiener. Así para el sistema la ecuación de *Fokker-Planck* está dada por

$$\frac{\partial p}{\partial t} = -\sum_{i=1}^n \frac{\partial [m_i(X)p]}{\partial x_i} + \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 [h_{ij}(X)p]}{\partial x_i \partial x_j} \quad (5.10)$$

Donde h_{ij} es el i -ésimo y j -ésimo elementos de la matriz $H(X) = G(X)DG^T(X)$ y $p(X, t | X_0)$ es la función de densidad de probabilidad de transición de X_t con la condición de normalización

$$\int_{\mathfrak{R}^n} p(X, t | X_0) dX = 1 \quad (5.11)$$

Sujeto a la condición inicial

$$\lim_{t \rightarrow 0} p(X, t | X_0) = \delta(X - X_0) \quad (5.12)$$

En otras palabras, la función de probabilidad de transición $p(\cdot)$ es solución de las ecuaciones *forward* y *backward* de Kolmogorov, esta demostración se puede ver en el Apéndice 1.

Las ecuaciones de *forward* y *backward* de Kolmogorov son aplicables a procesos de difusión multivariado. En particular, para los procesos de difusión con saltos la diferencia en el argumento del salto entre las ecuaciones de *forward* y *backward* de Kolmogorov es que en el caso de la ecuación *backward* en el término de salto se usa un salto posterior, y en la ecuación *forward* se usa un salto anterior.

Existen varias aplicaciones de este tipo de ecuaciones, entre ellas se encuentran en áreas como física e ingeniería, como por ejemplo la difusión de concentración de un fluido en un líquido o gas, de acuerdo a la ley de Fick; aplicaciones para flujos de concentración o difusión de calor en un medio conductivo de acuerdo a la ley de Fourier para flujos de calor.

Hasta el momento, se ha mostrado en las dos secciones previas algunos fundamentos matemáticos de por qué utilizar una función de densidad kernel para nuestro modelo no paramétrico, además se mostró que tanto la función de densidad marginal como la función de densidad de transición son soluciones de las ecuaciones de *forward* y *backward* de Kolmogorov. A continuación se muestran algunos supuestos generales para poder continuar con la metodología.

5.3 Supuestos del Modelo

Para el correcto desarrollo de la metodología, se consideran los siguientes supuestos:

Supuesto 1. El espacio de parámetros $\Theta \subset \mathfrak{R}^K$ es compacto. Es decir, en una vecindad de los parámetros reales θ_0 , $\pi(x, \theta)$ es dos veces continuamente diferenciable en θ , $E \left[\left(\frac{\partial \pi(x, \theta)}{\partial \theta} \right) \left(\frac{\partial \pi(x, \theta)}{\partial \theta'} \right)^T \right]$ tiene rango completo, y $\frac{\partial^2 \pi(x, \theta)}{\partial \theta_i \partial \theta_j}$ está acotada en valor absoluto para todo $\theta \in \Theta$.

Supuesto 2. Para cada $\theta \in \Theta$;

- i) Las funciones de drift y el proceso de difusión son seis veces continuamente diferenciables en x sobre $D = (0, \infty)$, y $\sigma^2 > 0$ sobre $(0, \infty)$.
- ii) La integral de $m(v, \theta) \equiv \frac{1}{\sigma^2(v, \theta)} \exp \left\{ - \int_v^{\bar{x}} \frac{2\mu(u, \theta)}{\sigma^2(u, \theta)} du \right\}$, la velocidad de la medida M , convergen a ambas cotas de D .
- iii) La integral de $s(v, \theta) \equiv \exp \left\{ \int_v^{\bar{x}} \frac{2\mu(u, \theta)}{\sigma^2(u, \theta)} du \right\}$, la medida de escala S , divergen a ambas cotas de D .

El inciso i) implica la condición local de Lipschitz y propiedades de crecimiento (sobre el espacio compacto D) por el teorema del valor medio. La condición local de Lipschitz es suficiente para asegurar la solución única. Los incisos i) y ii) garantizan que la ecuación diferencial estocástica $dX_t = \mu(X_t)dt + \sigma(X_t)dW_t$ admite una única solución, la cual es un proceso de Markov con densidad de transición en tiempo homogéneo y una densidad marginal estacionaria. El inciso iii) garantiza que comenzando desde cualquier punto interior del espacio de estado $D = (0, \infty)$, las cotas $\underline{x} = 0$ y $\bar{x} = +\infty$ no pueden alcanzarse en un tiempo esperado finito.

Supuesto 3. La sucesión de datos observados es una sucesión β estrictamente estacionaria que satisface: $k^\delta \beta(k) \xrightarrow{k \rightarrow +\infty} 0$ para algún valor fijo $\delta > 1$. Una condición sobre la función del drift y la difusión suficiente para asegurar que las observaciones discretas deben satisfacer este supuesto, se debe asumir:

$$i) \lim_{x \rightarrow 0 \text{ ó } x \rightarrow \infty} \sigma(x, \theta) \pi(x, \theta) = 0$$

$$ii) \lim_{x \rightarrow 0 \text{ ó } x \rightarrow \infty} \left| \frac{\sigma(x, \theta)}{2\mu(x, \theta) - \sigma(x, \theta) \frac{\partial \sigma(x, \theta)}{\partial x}} \right| < \infty$$

Supuesto 4. Se cumple;

- i) La densidad real π_0 es positiva sobre el interior de D y ambos π_0 y π_0^2 están en $L^2(D)$. La densidad conjunta p_0 es seis veces continuamente diferenciable.
- ii) La variable aleatoria inicial se distribuye como π_0 .
- iii) Las funciones reales de drift y de difusión satisfacen los supuestos 1 y 3.

Supuesto 5. La integral de la función kernel $K(\cdot)$ es igual a uno y su orden es $r = 2$. $K(\cdot)$ está acotada, es dos veces continuamente diferenciable sobre \mathfrak{R} . $K(\cdot)$ y sus derivadas están en $L^2(D)$.

El orden de un kernel $K(\cdot)$ esta definida como el entero r tal que

$$\int_{-\infty}^{+\infty} x^i K(x) dx = 0 \quad i = 1, \dots, r-1, \text{ y } \int_{-\infty}^{+\infty} x^r K(x) dx \neq 0, \int_{-\infty}^{+\infty} |x|^r |K(x)| dx < \infty.$$

Supuesto 6. Para estimar la densidad marginal, el ancho de banda es elegido de tal

$$\text{forma que } \lim_{n \rightarrow \infty} nh_n = \infty \text{ y } \lim_{n \rightarrow \infty} nh_n^{4.5} = 0.$$

Este supuesto determina que tan rápido el ancho de banda h_n debería ir a cero según incrementa el tamaño de la muestra. Los resultados empíricos son obtenidos con el kernel Gaussiano $K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right)$ de orden $r = 2$. De forma general, para una densidad de dimensión d y un kernel de orden r , $\lim_{n \rightarrow \infty} nh_n^d = \infty$ y $\lim_{n \rightarrow \infty} nh_n^{2r + \frac{d}{2}} = 0$. La calidad del estimador de densidad es ampliamente reconocida para determinar la elección del ancho de banda. El ancho de banda es calculado tal que $h_n = c_n n^{-1/4.5}$ donde $c_n = c$ veces la desviación estándar de la serie de tiempo de la tasa de interés spot dividida por $\ln(n)$. Donde c es elegido minimizando la media integrada del error cuadrado del estimador. Los resultados son cualitativamente robustos a la elección del kernel tanto como a los pequeños cambios del parámetro del ancho de banda alrededor del valor óptimo.

Dados dichos supuestos, ahora se va a mostrar cuál es la prueba estadística que se llevará a cabo en el trabajo, de igual forma se tratará de explicar el análisis para haber llegado a dicha prueba y los supuestos que se deben considerar para el proceso estocástico.

5.4 Hipótesis Nula, Hipótesis Alternativa y Prueba Estadística.

El principal objetivo de la prueba estadística es sobre los modelos de la tasa de interés corta que se clasifican dentro de los procesos Markovianos de difusión univariada los cuales son estrictamente estacionarios. Las propiedades que se consideran son las siguientes:

- i) La propiedad Markoviana es fundamental, ya que cuando un modelo no tiene el comportamiento dinámico Markoviano, es posible redefinir el espacio de estado. Con la propiedad de Markov, las transiciones entre los estados, sólo puede producirse con el estado anterior o posterior (en nuestro caso, para tasas de interés, con el estado anterior).
- ii) La función de difusión es univariada. Esto se pide para que el proceso X_t sea un caso reducible.
- iii) Estrictamente Estacionaria significa que el proceso de distribución conjunta satisface para cada valor m , cada $0 < t_1 < \dots < t_m < \infty$, y cada x_1, \dots, x_m en el espacio de estado:

$$f(X_{t_1} \leq x_1, \dots, X_{t_m} \leq x_m) = f(X_{t_1+\Delta} \leq x_1, \dots, X_{t_m+\Delta} \leq x_m)$$

En particular, aplicado a $m = 1$, la densidad del proceso puede depender de Δ pero no del tiempo. Así el proceso admite densidad marginal estacionaria. Esta propiedad se le pide a las funciones $\mu(X_t)_t$ y $\sigma^2(X_t)$ que corresponden a las funciones del drift y el proceso de difusión del modelo de la tasa de interés. Esto es debido a que las funciones pueden variar a través del tiempo y pueden ser redefinidas continuamente, y esta prueba no se basa en describir series de tiempo.

- iv) Se restringe a los modelos de la tasa de interés spot, es decir, no tomar valores de cero e infinito.
- v) La teoría asintótica clásica es aplicable a los datos discretos generados por su proceso de tiempo continuo, es decir, que los datos observados puedan ser mezclados lo suficientemente rápido con la tasa.

Bajo estas propiedades sobre el modelo de tasa de interés corta, existe una solución única para la ecuación (4.1), dadas las funciones de drift y el proceso de difusión

$(\mu_0(\cdot), \sigma_0^2(\cdot))$. Se considera una parametrización específica, la cual es una familia conjunta de parámetros:

$$P \equiv \{(\mu(\cdot, \theta), \sigma^2(\cdot, \theta)) \mid \theta \in \Theta\} \quad (5.13)$$

Donde Θ es un subconjunto compacto de \mathfrak{R}^K . Por ejemplo para el modelo de CIR, se tiene $P \equiv \{(\mu(X, \theta), \sigma^2(X, \theta)) = ((\alpha_0 + \alpha_1 X), \beta_1^2 X) \mid \theta \equiv (\alpha_0, \alpha_1, \beta_1^2) \in \Theta\}$. Donde nuestro objetivo es encontrar valores de los parámetros en Θ de tal forma que el modelo paramétrico tenga una representación aceptable del proceso, es decir, existan funciones $(\mu_0(\cdot), \sigma_0^2(\cdot))$ que encajen en el espacio paramétrico P . Así nuestra hipótesis nula y alternativa son:

$$\begin{aligned} H_0 : & \exists \theta_0 \in \Theta \mid \mu(\cdot, \theta_0) = \mu_0(\cdot) \quad y \quad \sigma^2(\cdot, \theta_0) = \sigma_0^2(\cdot) \\ H_1 : & (\mu_0(\cdot), \sigma_0^2(\cdot)) \notin P \end{aligned} \quad (5.14)$$

Así, la hipótesis nula establecerá que existen valores paramétricos para los cuales el modelo considerado es una representación aceptable de los datos. En cambio, la hipótesis alternativa afirma que ningún valor paramétrico es capaz de reproducir las funciones exactas de μ y σ^2 .

Pero es difícil de estimar cada parámetro de (5.14), muchas veces las aproximaciones son válidas cuando el horizonte de tiempo es muy pequeño, para datos intradía, entonces, se busca una caracterización del proceso:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t \quad (5.15)$$

Y que sea equivalente a:

$$\mu(X) = \mu(X, \theta) \quad \text{and} \quad \sigma^2(X) = \sigma^2(X, \theta), \quad \text{donde} \quad \theta \in \Theta \subset \mathfrak{R}^K \quad (5.16)$$

Las cuales dependen de cantidades que pueden actualmente ser estimados bajo las restricciones impuestas por los datos discretos, en nuestro caso datos diarios. Bajo los supuestos hechos, las funciones del proceso $\mu(X, \theta)$ y $\sigma^2(X, \theta)$ serán caracterizadas únicamente por sus densidad marginal y densidad de transición sobre sus intervalos de tiempo discreto.

Cualquier parametrización de P de μ y σ^2 corresponden a una parametrización de la densidad marginal:

$$\Pi \equiv \left\{ \pi(\cdot, \theta), p(\cdot, \cdot, | \cdot, \cdot; \theta) \mid (\mu(\cdot, \theta), \sigma^2(\cdot, \theta)) \in P, \theta \in \Theta \right\} \quad (5.17)$$

Donde $\pi(x, \theta)$ es la densidad marginal en x y es la densidad condicional de que $X_s = y$ dado $X_t = x$.

Se define el espacio Π_M de funciones de densidad correspondiente al par $(\mu(\cdot, \cdot), \sigma^2(\cdot, \cdot))$ en P , es decir,

$$\Pi_M \equiv \left\{ \pi(\cdot, \theta) \mid (\mu(\cdot, \theta), \sigma^2(\cdot, \theta)) \in P, \theta \in \Theta \right\} \quad (5.18)$$

Ahora la nueva hipótesis nula asegura que existe una función de densidad marginal correspondiente al par μ y σ^2 y la hipótesis alternativa que la densidad marginal no pertenece al espacio formado por μ y σ^2 , es decir:

$$\begin{aligned} H_{M0} : & \exists \theta_0 \in \Theta \mid \pi(\cdot, \theta_0) = \pi_0(\cdot) \\ H_{M1} : & \pi_0(\cdot) \notin \Pi_M \end{aligned} \quad (5.19)$$

Es necesario que H_{M_0} se cumpla para que se cumpla H_0 . Si la función de densidad $\pi_0(\cdot)$ fuera conocida, se puede verificar fácilmente la conjetura con la densidad del modelo paramétrico. Pero cuando $\pi_0(\cdot)$ es desconocida, se puede estimar, y así hacer la una prueba estadística para probar la hipótesis nula.

Se debe de proponer un estimador no paramétrico, es decir, que no dependa de ningún parámetro de $\mu(X, \theta)$ y $\sigma^2(X, \theta)$, el estimador deberá converger a la función de densidad real con o sin el modelo paramétrico. Luego se considera un estimador paramétrico de la densidad implícita $\pi(x, \theta_0)$, este converge a la densidad real si y sólo si el modelo está correctamente especificado.

Por lo tanto la densidad del modelo paramétrico y la del modelo no paramétrico deben de estar muy cercanas si el modelo paramétrico está bien definido, y muy lejanas en cualquier otro caso. Así se introduce una medida de comparación, a la cual denotaremos como M la cual medirá la distancia entre las dos densidades estimadas, y provee un estadístico natural para probar la hipótesis nula de la especificación paramétrica correcta y estará dada por la siguiente ecuación:

$$\begin{aligned}
 M &\equiv \min_{\theta \in \Theta} \int_{\bar{x}}^{\bar{s}} (\pi(u, \theta) - \pi_0(u))^2 \pi_0(u) du & (5.20) \\
 &= \min_{\theta \in \Theta} E[(\pi(X, \theta) - \pi_0(X))^2]
 \end{aligned}$$

La prueba estadística se rechaza cuando el valor de M es muy grande. La distancia M pondera la diferencia entre los estimadores paramétrico y no paramétrico de acuerdo a su relevancia, pondera con mayor peso a los valores del proceso que más aparecen, y cuando la densidad es muy baja la diferencia tiene menor ponderación. Bajo la hipótesis nula M tiene un valor pequeño, mientras que para la hipótesis alternativa M debe tener un valor grande.

Se propone como prueba estadística a

$$\hat{M} \equiv n \cdot h_n \cdot \min_{\theta \in \Theta} \frac{1}{n} \sum_{i=1}^n (\pi(x, \theta) - \hat{\pi}_0(x_i))^2 \quad (5.21)$$

Donde $\hat{\pi}_0(\cdot)$ es el estimador no paramétrico de la función de densidad, n es el número de observaciones y h_n es el ancho de banda. Así es como se van hacer las comparaciones entre las funciones de densidad marginal.

Bajo la hipótesis H_0 , el parámetro estimado $\hat{\theta}_M$ tiene la distribución

$$n^{1/2} \{\hat{\theta}_M - \theta_0\} \xrightarrow{d} N(0, \Omega_M) \quad (5.22)$$

Donde Ω_M está dado por

$$\begin{aligned} \Omega_M \equiv & \text{var}(\varphi[F](x_t)) \\ & + \sum_{k=i}^{+\infty} \{\text{cov}(\varphi[F](x_t), \varphi[F](x_{t+k})) + \text{cov}(\varphi[F](x_{t+k}), \varphi[F](x_t))\} \end{aligned} \quad (5.23)$$

Donde $F_0(\cdot)$ denota la función de densidad acumulada asociada con la densidad $\pi_0(\cdot)$; $\varphi(F_0)$ es la derivada del funcional $\Theta_M(\cdot)$ que está dada por:

$$\Theta_M[F] \equiv \arg \min_{\theta \in \Theta} \int_{\underline{x}}^{\bar{x}} (\pi(x, \theta) - \pi_0(x))^2 \pi_0(x) dx. \quad (5.24)$$

La prueba estadística \hat{M} está distribuida como:

$$h_n^{-1/2} \{\hat{M} - E_M\} \xrightarrow{d} N(0, V_M) \quad (5.25)$$

Donde \hat{E}_M y \hat{V}_M que están dados por:

$$\begin{aligned}\hat{E}_M &\equiv \left(\int_{-\infty}^{+\infty} K^2(x) dx \right) \left(\frac{1}{n} \sum_{i=1}^n \hat{\pi}_0(r_i) \right), \\ \hat{V}_M &\equiv \left(\int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{+\infty} K(u)K(u+x) du \right\}^2 dx \right) \left(\frac{1}{n} \sum_{i=1}^n \hat{\pi}_0^3(r_i) \right)\end{aligned}\quad (5.26)$$

La prueba de la hipótesis nula a un nivel de confianza α , la región crítica será:

$$\text{rechazo } H_0 \text{ cuando } \hat{M} \geq \hat{c}(\alpha) \equiv \hat{E}_M + h_n^{1/2} \cdot z_{1-\alpha} \cdot \{\hat{V}_M\}^{1/2} \quad (5.27)$$

De esta forma se obtendrán los valores de M , si M es muy grande quiere decir que el modelo utilizado no refleja el comportamiento de los datos observados, y viceversa. En cualquier forma se busca el valor mínimo de M .

Ya se ha explicado como se va a considerar la prueba estadística, ahora se va a comenzar a explicar como se va a relacionar la función de densidad kernel elegida con la función de densidad marginal para el modelo no paramétrico.

5.5 Densidad Marginal para el modelo no-paramétrico.

Los datos utilizados para formar los estimadores de densidad consisten en observaciones discretas diarias de la tasa de interés corta. Estos datos son usados para estimar la densidad continua del proceso sin remplazar el modelo de tiempo continuo. Las propiedades asintóticas de los estimadores son derivadas para un periodo dado, es decir, que la tasa de interés es observada sobre periodo de tiempo cada vez más largo.

Para obtener la función de densidad marginal para el modelo no paramétrico se utilizó la siguiente función:

$$\hat{\pi}_0(u) \equiv \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n} K\left(\frac{u-r_i}{h_n}\right) \quad (5.28)$$

Donde $K(\cdot)$ es la función *kernel* gaussiano y h_n el ancho de banda. Además se calculan los siguientes los estimadores \hat{E}_M y \hat{V}_M que están dados por:

$$\begin{aligned} \hat{E}_M &\equiv \left(\int_{-\infty}^{+\infty} K^2(x) dx \right) \left(\frac{1}{n} \sum_{i=1}^n \hat{\pi}_0(r_i) \right), \\ \hat{V}_M &\equiv \left(\int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{+\infty} K(u)K(u+x) du \right\}^2 dx \right) \left(\frac{1}{n} \sum_{i=1}^n \hat{\pi}_0^3(r_i) \right) \end{aligned} \quad (5.29)$$

Cualquier otro estimadores consistentes de E_M y V_M pueden ser usados. Siendo $z_{1-\alpha} = 1.64$ se obtiene el 95 por ciento el nivel de confianza. Para la función de kernel Gaussiano está dado por:

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right) \quad (5.30)$$

Las dos constantes de esta función kernel cumplen:

- $\int_{-\infty}^{+\infty} K^2(x) dx = \frac{1}{2\sqrt{\pi}}$
- $\int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} K(u)K(u+x) du \right)^2 dx = \frac{1}{2\sqrt{2\pi}}$

$$(5.31)$$

Y el ancho de banda está dado por $h_n = c_n n^{-1/4.5}$ donde $c_n = c$ veces la desviación estándar de la serie de tiempo de la tasa de interés spot dividida por $\ln(n)$. Donde c es elegido minimizando la media integrada del error cuadrado del estimador.

Falta desarrollar la metodología que se va a tomar en cuenta con los modelos paramétricos, en la siguiente sección se va a mostrar de forma general la forma del proceso estocástico que se considera y la forma general de la función de densidad marginal para este tipo de modelos.

5.6 Densidad Marginal para los modelos paramétricos

Se consideran los modelos de tasa de interés con funciones de difusión univariados, estrictamente estacionarios y que cuenten con la propiedad Markoviana. En general estos modelos siguen el siguiente proceso:

$$dX_t = \mu(X_t, \theta)dt + \sigma(X_t, \theta)dW_t \quad (5.32)$$

Donde $\{W_t, t \geq 0\}$ es un movimiento browniano estándar. Las funciones $\mu(X_t, \theta)$ y $\sigma^2(X_t, \theta)$ son respectivamente, la función del drift y difusión del proceso, donde $\theta \in \Theta \subset \mathfrak{R}^K$. Para cada parametrización específica del proceso, se puede trabajar con la familia paramétrica conjunta:

$$P \equiv \{(\mu(\cdot, \theta), \sigma^2(\cdot, \theta)) \mid \theta \in \Theta\} \quad (5.33)$$

Donde Θ es un subconjunto compacto de \mathfrak{R}^K . Cualquier parametrización de P de μ y σ^2 corresponden a una parametrización de la densidad marginal:

$$\Pi \equiv \left\{ \pi(\cdot, \theta), p(\cdot, \cdot | \cdot, \cdot; \theta) \mid (\mu(\cdot, \theta), \sigma^2(\cdot, \theta)) \in P, \theta \in \Theta \right\} \quad (5.34)$$

Donde $\pi(x, \theta)$ es la densidad marginal en x y es la densidad condicional de que $X_s = y$ dado $X_t = x$. La densidad marginal para los modelos paramétricos correspondiente a $(\mu(\cdot, \cdot), \sigma^2(\cdot, \cdot))$ está dada por:

$$\pi(x, \theta) = \frac{\varepsilon(\theta)}{\sigma^2(x, \theta)} \exp \left\{ \int_{x_0}^x \frac{2 \cdot \mu(u, \theta)}{\sigma^2(u, \theta)} du \right\} \quad (5.35)$$

Donde $\varepsilon(\theta)$ sirve para asegurar que la integral de la función de densidad es 1. La ecuación (5.39) se obtiene de “*forward Kolmogorov equation*” (la demostración se hace en la sección 8.1).

Después de realizar lo anterior, la segunda parte del trabajo es desarrollar la metodología para modelos paramétricos con saltos de Poisson, a continuación se muestra el modelo general del proceso estocástico que se pretende utilizar, y debido a la discretización del proceso, para estos casos se determina la función de probabilidad de transición.

5.7 Densidad de Transición para los modelos paramétricos con Saltos

Para la segunda parte del análisis, se considera un proceso con difusión de saltos X_t definido sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, (\mathfrak{F}_t)_{t \in [0, T]}, \mathcal{P})$ con filtración $\{\mathfrak{F}_t\}$ que satisface las condiciones usuales,

$$dX_t = \mu(X_t, \theta)dt + \sigma(X_t, \theta)dW_t + J_t dN_t \quad (5.36)$$

Donde

- $\{W_t, t \geq 0\}$ es un movimiento browniano estándar;

- $\theta \in \mathfrak{R}^p$ es un vector finito de parámetros que serán estimados;
- $\mu(X_t, \theta): \mathfrak{R}^n \rightarrow \mathfrak{R}^n$ es la función del drift;
- $\sigma^2(X_t, \theta): \mathfrak{R}^n \rightarrow \mathfrak{R}^{n \times d}$ es la función de difusión del proceso;
- El proceso de salto puro N tiene una intensidad estocástica de $\lambda(X_t, \theta)$ y tamaño de salto de 1, y
- El salto de tamaño J_t es independiente de \mathfrak{F}_{t-} y tiene una densidad de probabilidad $\nu(\cdot, \theta): \mathfrak{R}^n \rightarrow \mathfrak{R}$. En nuestro caso $\nu(\cdot, \theta)$ se supone normal.

Para los modelos con saltos de Poisson denotaremos la densidad condicional de $X_{t+\Delta} = y$ dado $X_t = x$, por $p(\Delta, y|x, \theta)$ y también se le conoce como la densidad de probabilidad de transición.

Sea X_t un proceso de Markov en tiempo continuo, es decir, un proceso de difusión con saltos en el espacio de estado Θ . La función de distribución de probabilidad de transición está dada por

$$p(\Delta, y|x, \theta) = \Pr\{X_t < x | X_{t_0} = x_0\} \quad (5.37)$$

Donde $t > t_0$, $X_t < x$ significa $X_{it} < x_i$ para $i = 1:n$.

Se considera una aproximación a $p(\Delta, y|x, \theta)$ usando las ecuaciones *forward* y *backward* de Kolmogorov.

$$p(\Delta, y|x) = \Delta^{-n/2} \exp\left[-\frac{C^{(-1)}(x, y)}{\Delta}\right] \sum_{k=0}^{\infty} C^{(k)}(x, y) \Delta^k + \sum_{k=0}^{\infty} D^{(k)}(x, y) \Delta^k \quad (5.38)$$

Para alguna función $C^{(k)}(x, y)$ y $D^{(k)}(x, y)$ para ser determinada. (la demostración se hace en la sección 9.1)

6. Modelo No-Paramétrico.

En este capítulo se muestran algunas de las propiedades que adquieren los estimadores de los datos observados, y se especifica la función de densidad marginal que se considera para el desarrollo de la metodología.

Los datos utilizados para formar los estimadores de la densidad marginal consisten en observaciones discretas (diarias) de la tasa de interés corta. Los parámetros de este proceso en tiempo discreto se estiman usando la técnica generalizada del método de los momentos. Así, los datos son usados para estimar la densidad continua del proceso. Dado que el tamaño del periodo de tiempo es grande, los estimadores poseen diversas propiedades estadísticas deseables conocidas como propiedades asintóticas. Entre ellas se encuentran:

- a) Insesgamiento asintótico. $\hat{\theta}$ es un estimador asintóticamente insesgado de θ , su valor esperado, o media se aproxima al verdadero valor a medida que el número de la muestra se hace cada vez más grande.
- b) Consistencia. $\hat{\theta}$ es un estimador consistente de θ si se puede decir que $\hat{\theta}$ es un estimador confiable, en otras palabras, si la probabilidad de que el valor absoluto de la diferencia entre $\hat{\theta}$ y θ sea menor que δ , $\delta > 0$, y se aproxima a la unidad, matemáticamente:

$$\Pr\{|\hat{\theta} - \theta| < \delta\} = 1$$

- c) Eficiencia asintótica. Sea $\hat{\theta}$ un estimador de θ . La varianza de la distribución asintótica de $\hat{\theta}$ se denomina varianza asintótica de $\hat{\theta}$. Si $\hat{\theta}$ es consistente y su varianza asintótica es menor que la varianza asintótica de todos los demás estimadores consistentes de θ , $\hat{\theta}$ es llamado asintóticamente eficiente.

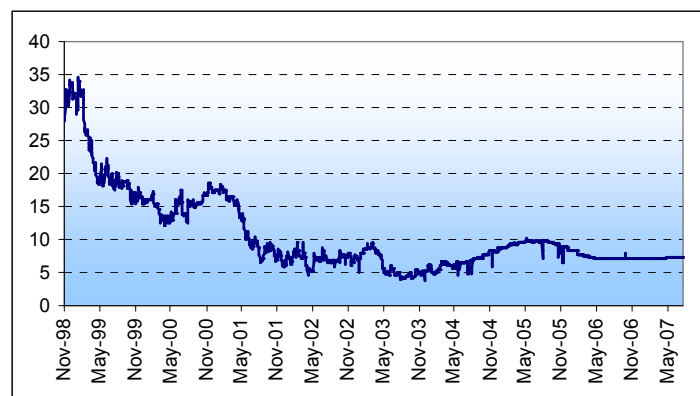
- d) Normalidad asintótica. Se dice que un estimador $\hat{\theta}$ está distribuido normal asintóticamente si su distribución muestral tiende a aproximarse a la distribución normal a medida que el tamaño de la muestra aumenta indefinidamente.
- e) Las funciones kernel son simétricas y acotadas, y satisfacen la condición de Lipschitz.
- f) El proceso Markoviano estacionario X_t es β -mezclado con una tasa de decaimiento exponencial $\beta(n) = O(\exp(-\lambda n))$ para algún $\lambda > 0$.
- g) La función $p(y|x, \theta)$ tiene derivadas continuas parciales de orden cuatro con respecto a x y y sobre el conjunto Θ . La densidad invariante $\pi(x)$ tiene una segunda derivada continua. Más aún, $\pi(x) > 0$ y $p(y|x, \theta) > 0$ para todo $(x, y) \in \Theta$.

El modelo de tiempo discreto tiene la ventaja de permitir cambios de tasas de interés en la varianza que dependan directamente sobre el nivel de la tasa de interés en una forma consistente con el modelo de tiempo continuo. Además el modelo captura la volatilidad de cambios en el riesgo de tasa, se sabe que la volatilidad juega un papel central en dos de los más importantes aplicaciones de los modelos de estructura de plazos: valuar contingencias y coberturas de riesgo de tasas de interés. Por ejemplo, la volatilidad de tasas de interés es un determinante fundamental del valor de las opciones de tasas de interés. Además, estrategias de coberturas óptimas para inversionistas adversos al riesgo pueden ser muy sensibles a cambios en la volatilidad esperada de las tasas de interés.

Se considera la tasa cetes del mercado secundario en México en el periodo del 4 de noviembre de 1988 al 23 de julio de 2007. Así se tiene una muestra de $n = 2,196$ observaciones. Se aplicó la prueba de no estacionariedad de Dickey-Fuller, y la

presencia de una raíz unitaria fue rechazada, la serie de los rendimientos de los datos es estacionaria. Un proceso estocástico es estacionario si su media y su varianza son constantes en el tiempo y el valor de la covarianza entre dos periodos depende solamente de la distancia o rezago entre estos dos periodos de tiempo y no del tiempo en el cual se ha calculado la covarianza.

Los datos se encuentran en la gráfica 6.1.



Gráfica 6.1

Para obtener la función de densidad marginal para el modelo no paramétrico se utiliza la siguiente función:

$$\hat{\pi}_0(u) \equiv \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n} K\left(\frac{u - r_i}{h_n}\right) \quad (6.1)$$

Donde $K(\cdot)$ es la función Kernel Gaussiano está dado por:

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right) \quad (6.2)$$

y h_n es el ancho de banda está dado por $h_n = c_n n^{-1/4.5}$ donde $c_n = c$ veces la desviación estándar de la serie de tiempo de la tasa de interés spot dividida por

$\ln(n)$. Donde c es elegido minimizando la media integrada del error cuadrado del estimador.

7. Modelos Paramétricos

Como se ha dicho en la literatura existen diversos modelos de tasas de interés, los cuales son caracterizados por su drift y su función de difusión. Un ejemplo, se encuentra los modelos con reversión a la media, $\mu(X, \theta) = \beta(\gamma - X)$, donde la tasa spot se mueve en dirección a su media γ a una velocidad β . En este capítulo se especifican los modelos paramétricos que se van a considerar en la metodología.

Modelo de Vasicek

O. A. Vasicek publica el artículo “An Equilibrium Characterization of the Term Structure” en el que describe la dinámica estocástica de una tasa de interés instantánea que presenta reversión a la media. Sea $(W_t)_{t \geq 0}$ un movimiento Browniano definido sobre un espacio fijo de probabilidad equipado con su filtración aumentada $(\Omega, \mathfrak{F}, (\mathfrak{F}_t)_{t \in [0, T]}, \mathcal{P})$, el comportamiento de la tasa corta es:

$$dX_t = (\alpha_0 + \alpha_1 X)dt + \beta_0 dW_t \quad (7.1)$$

Donde $\alpha_0, \alpha_1, \beta_0 \in \mathfrak{R}$. X_t es obligada a moverse en promedio, hacia un nivel de largo plazo α_0 / α_1 a una velocidad $-\alpha_1$.

Modelo de Cox, Ingersoll y Ross

Se considera un movimiento Browniano $(W_t)_{t \geq 0}$ definido sobre un espacio de probabilidad equipado con su filtración aumentada $(\Omega, \mathfrak{F}, (\mathfrak{F}_t)_{t \in [0, T]}, \mathcal{P})$, el comportamiento de la tasa corta es:

$$dX_t = (\alpha_0 + \alpha_1 X) \cdot dt + \beta_1 \sqrt{X} \cdot dW_t \quad (7.2)$$

Donde $\alpha_0, \alpha_1, \beta_1 \in \mathfrak{R}$. Al considerar \sqrt{X} en el término estocástico, el proceso de la tasa corta deja de tener una distribución normal, ahora corresponde a una χ^2 no central. Al igual que el modelo de Vasicek, se presenta reversión a la media y la varianza es proporcional a $\beta_1^2 X$ por unidad de tiempo, es decir que conforme la tasa de interés corta aumenta, la desviación estándar aumenta, lo que impide la existencia de tasas de interés negativas. Este modelo también es conocido como proceso de raíz cuadrada y se introdujo en 1985.

Otro modelo que CIR introdujo es:

$$dX_t = \beta_1 X^{3/2} \cdot dW_t \quad (7.3)$$

Este modelo se introdujo con un estudio de tasas variables, “*variable rates (VR)*”.

Modelo de Brennan y Schwartz

Brennan y Schwartz (1980) desarrollaron el siguiente modelo, derivando un modelo numérico para valuar bonos convertibles:

$$dX_t = (\alpha_0 + \alpha_1 X)dt + \beta_2 X dW_t \quad (7.4)$$

Donde $\alpha_0, \alpha_1, \beta_2 \in \mathfrak{R}$. Otro modelo que Brennan y Schwartz introdujeron a la literatura es el siguiente

$$dX_t = (\alpha_0 + \alpha_1 X)dt + \beta_2 X^{3/2} dW_t \quad (7.5)$$

Modelo de Chan

K. C. Chan (1992) propuso el siguiente modelo en su artículo “An Empirical Comparison of Alternative Models of the Short-Term Interest rate”. Chan utilizó el Método Generalizado de Momentos para mostrar que la dinámica de la tasa de interés a corto plazo permite cambios muy sensibles en la volatilidad de misma tasa de interés. Su modelo propuesto fue:

$$dX_t = (\alpha_0 + \alpha_1 X)dt + \beta_2 X^{\beta_3} dW_t \quad (7.6)$$

Donde $\alpha_0, \alpha_1, \beta_2 \in \mathfrak{R}$. Esta dinámica implica que los cambios condicionales en la media y la varianza en la tasa de interés dependen del nivel de X . Encontró que para los modelos donde $\beta_3 \geq 1$ se captura mejor la dinámica de la tasa de interés a corto plazo que cuando $\beta_3 < 1$.

Modelo de Drift General Drift, (Constant Elasticity Volatility).

Se considera un movimiento Browniano $(W_t)_{t \geq 0}$ definido sobre un espacio de probabilidad equipado con su filtración aumentada $(\Omega, \mathfrak{F}, (\mathfrak{F}_t)_{t \in [0, T]}, \mathcal{P})$, el comportamiento de la tasa corta para el modelo de media general:

$$dX_t = \left(\alpha_0 + \alpha_1 X + \alpha_2 X^2 + \frac{\alpha_3}{X} \right) \cdot dt + \left(\beta_0 + \beta_1 X + \beta_2 X^{\beta_3} \right) \cdot dW_t \quad (7.7)$$

Donde $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \beta_0, \beta_1, \beta_2 \in \mathfrak{R}$. Un caso particular es un proceso de elasticidad en la varianza constante, es introducido por Cox (1985):

$$dX_t = \alpha_1 X dt + \beta_1 X^\gamma dW_t \quad (7.8)$$

Modelo de Merton

Robert C. Merton (1973) en su artículo “Theory of Rational Option Pricing”, propuso uno de los primeros modelos para explicar la dinámica estocástica de la tasa de interés instantánea, o tasa corta. Se considera un movimiento Browniano $(W_t)_{t \geq 0}$ definido sobre un espacio de probabilidad equipado con su filtración aumentada $(\Omega, \mathfrak{F}, (\mathfrak{F}_t)_{t \in [0, T]}, \mathcal{P})$, el comportamiento de la tasa corta es:

$$dX_t = \alpha_0 dt + \beta_0 dW_t \quad (7.9)$$

Donde $\alpha_0, \beta_0 \in \mathfrak{R}$ y X_t se distribuye normal. Este modelo tiene varias limitaciones como: i) existe una probabilidad positiva de que la tasa corta tome valores negativos; ii) no presenta reversión a la media, es decir, no existe un mecanismo que

obligue a la tasa corta a regresar a un nivel de largo plazo conforme el tiempo transcurre; iii) la esperanza y la varianza condicionales de la tasa corta crecen si límite al transcurrir el tiempo; y iv) la curva de rendimiento y la tasa *forward* decrecen sin cota conforme el tiempo aumenta.

Modelo de Dothan

Dothan (1978) utiliza este modelo en la valuación de bonos, en su artículo “On the Term Structure of Interest Rates”, además Brennan y Schwartz también lo usan en el desarrollo de modelos numéricos para bonos. Dothan toma el supuesto de Vasicek de que el mercado sea libre de riesgo, pero modifica el supuesto sobre el proceso que sigue la tasa de Itnez spot, de tal forma:

$$dX_t = \beta_1 X dW_t \quad (7.10)$$

Donde $\beta_1 \in \Re$ y X_t se distribuye normal.

Dotan utiliza fundamentos macroeconómicos de maximización de utilidad bajo ciertas preferencias y de argumentos de arbitraje.

Modelo de GBM

Este es el proceso del Modelo Geométrico Browniano simple, Geometric Brownian Motion (GBM), de Black y Scholes. Sea $(W_t)_{t \geq 0}$ un movimiento Browniano definido sobre un espacio de probabilidad equipado con su filtración aumentada $(\Omega, \mathfrak{F}, (\mathfrak{F}_t)_{t \in [0, T]}, \mathcal{P})$, el comportamiento de la tasa corta es:

$$dX_t = \alpha_1 X dt + \beta_1 X dW_t \quad (7.11)$$

Donde $\alpha_1, \beta_1 \in \mathfrak{R}$.

En la Tabla 7.1 se muestra la lista de modelos de tasa de interés que son considerados para este trabajo:

| Modelo Paramétrico | $\mu(X, \theta)$ | $\sigma(X, \theta)$ |
|---------------------------------|---|---|
| Vasicek (1977) | $\alpha_0 + \alpha_1 X$ | β_0 |
| Cox, Ingersoll y Ross (1985) | $\alpha_0 + \alpha_1 X$ | $\beta_1 \sqrt{X}$ |
| CIR VR | - | $\beta_1 X^{3/2}$ |
| Brennan & Schwartz (1982) | $\alpha_0 + \alpha_1 X$ | $\beta_2 X$ |
| * Brennan & Schwartz | $\alpha_0 + \alpha_1 X$ | $\beta_1 X^{3/2}$ |
| Chan (1992) | $\alpha_0 + \alpha_1 X$ | $\beta_2 X^{\beta_3}$ |
| General Drift. | $\alpha_0 + \alpha_1 x X + \alpha_2 X^2 + \alpha_3 / X$ | $\beta_0 + \beta_1 X + \beta_2 X^{\beta_3}$ |
| Merton | α_0 | β_0 |
| Dothan | - | $\beta_1 X$ |
| GBM | $\alpha_1 X$ | $\beta_1 X$ |
| CEV | $\alpha_1 X$ | $\beta_2 X^\gamma$ |

Tabla 7.1

8. Parte I. Análisis de los Modelos de Tiempo Continuo sin saltos de Poisson.

En esta parte del análisis se obtienen las funciones de densidad marginal para cada uno de los modelos de tiempo continuo a través de la función de “*forward Kolmogorov*” y se aplica la metodología para discernir con el modelo no-paramétrico.

8.1 Análisis Matemático sin saltos de Poisson.

Se considera el siguiente proceso X_t definido sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ con filtración $\{\mathfrak{F}_t\}$ que satisfacen las condiciones usuales:

$$dX_t = \mu(X_t, \theta)dt + \sigma(X_t, \theta)dW_t \quad (8.1)$$

Donde

- $\{W_t, t \geq 0\}$ es un movimiento browniano estándar;
- $\theta \in \mathfrak{R}^p$ es un vector finito de parámetros que serán estimados;
- $\mu(X_t, \theta): \mathfrak{R}^n \rightarrow \mathfrak{R}^n$ es la función del drift, o bien el promedio de crecimiento del proceso estocástico X_t , y
- $\sigma^2(X_t, \theta): \mathfrak{R}^n \rightarrow \mathfrak{R}^{n \times d}$ es la función de difusión del proceso, o bien, $\sigma(X_t, \theta)$ es la varianza del proceso estocástico X_t .

Para los modelos paramétricos como se había mencionado la densidad de marginal viene dada por:

$$\pi(x, \theta) = \frac{\varepsilon(\theta)}{\sigma^2(x, \theta)} \exp \left\{ \int_{x_0}^x \frac{2 \cdot \mu(u, \theta)}{\sigma^2(u, \theta)} du \right\} \quad (8.2)$$

Donde $\varepsilon(\theta)$ sirve para asegurar que la integral de la función de densidad es 1. Para obtener la ecuación (8.2), se parte de la ecuación de *forward-Kolmogorov*. A continuación se muestra la demostración.

Sea $\psi(x)$ la distribución de densidad estacionaria de la variable X_t :

$$\psi(y) = \int \psi(x)p(t,x,y)dx \quad \text{para todo } t > 0. \quad (8.3)$$

la ecuación *forward-Kolmogorov*, que está definida por:

$$\frac{\partial \varphi(t,x)}{\partial t} = \frac{1}{2} \frac{\partial^2}{\partial x^2} [\sigma^2(x)\varphi(t,x)] - \frac{\partial}{\partial x} [\mu(x)\varphi(t,x)] \quad (8.4)$$

La distribución empírica observada es solución de la estacionariedad formada de la ecuación *forward-Kolmogorov*, así, la función de distribución $\psi(y)$ cumple:

$$0 = \frac{1}{2} \frac{\partial^2}{\partial y^2} [\sigma^2(y)\psi(y)] - \frac{\partial}{\partial y} [\mu(y)\psi(y)] \quad (8.5)$$

Integrando la ecuación (8.5) se tiene:

$$\frac{d}{dy} \left[\frac{\sigma^2(y)}{2} \psi(y) \right] - \mu(y)\psi(y) = \frac{1}{2} C_1 \quad (8.6)$$

Donde C_1 es la constante de integración, despejando el segundo término se tiene:

$$\frac{d}{dy} [\sigma^2(y)\psi(t,y)] = 2\mu(y)\psi(t,y) + C_1 \quad (8.7)$$

Por otro lado, sea el factor de multiplicación $s(y)$, y considerando que se debe de cumplir la siguiente expresión:

$$\begin{aligned} \frac{\partial}{\partial y} [s(y)\sigma^2(y)\psi(t,y)] &= C_1 s(y) \\ \Leftrightarrow \frac{1}{s(y)} \left[s(y) \frac{\partial [\sigma^2(y)\psi(t,y)]}{\partial y} + \sigma^2(y)\psi(t,y) \frac{\partial [s(y)]}{\partial y} \right] &= C_1 \end{aligned} \quad (8.8)$$

Para obtener la ecuación (8.6) a partir de la ecuación (8.7), se utiliza la ecuación (8.8) y se debe cumplir:

$$\begin{aligned} \frac{\sigma^2(y)\psi(y)}{s(y)} \frac{\partial [s(y)]}{\partial y} &= -2\mu(y)\psi(y) \\ \Leftrightarrow \frac{1}{s(y)} \frac{\partial [s(y)]}{\partial y} &= -\frac{2\mu(y)\psi(y)}{\sigma^2(y)\psi(y)} \\ \Leftrightarrow \frac{\partial}{\partial y} \ln[s(y)] &= -\frac{2\mu(y)}{\sigma^2(y)} \\ \Leftrightarrow s(y) &= \exp \left\{ -\int \frac{2\mu(y)}{\sigma^2(y)} \right\} \end{aligned} \quad (8.9)$$

La integral $S(x) = \int s(\varepsilon) d\varepsilon$, es llamada función de escala y la velocidad de la

densidad es $m(x) = \frac{1}{\sigma^2(x)s(x)}$. Si la densidad es estacionaria, se tiene:

$$\psi(x) = m(x)[C_1 s(x) + C_2] \quad (8.10)$$

Donde $s(y) = \exp\left\{-\int \frac{2\mu(y)}{\sigma^2(y)}\right\}$, y las constantes están determinadas para garantizar

las restricciones $\psi(x) \geq 0$ en el intervalo (l, r) , y $\int_l^r \psi(\xi) d\xi = 1$; si $C_1 = 0$ se garantiza que $\psi(x) > 0$ y C_2 sirve para asegurar que:

$$\int_l^r \psi(\xi) d\xi = 1.$$

Si se sustituyen los valores de $m(x)$ y $s(x)$ en (10) se tiene:

$$\begin{aligned} \psi(x) &= \frac{1}{\sigma^2(x)s(x)} [C_1 s(x) + C_2] \\ &= \frac{C_1}{\sigma^2(x)} + \frac{C_2}{\sigma^2(x)} [s(x)]^{-1} \\ &= \frac{C_2}{\sigma^2(x)} \exp\left\{\int \frac{2\mu(y)}{\sigma^2(y)} dy\right\} \end{aligned} \quad (8.11)$$

Finalmente se obtiene la ecuación (3) haciendo el cambio de ψ por π :

$$\pi(x, \theta) = \frac{\varepsilon(\theta)}{\sigma^2(x, \theta)} \exp\left\{\int_{x_0}^x \frac{2 \cdot \mu(u, \theta)}{\sigma^2(u, \theta)} du\right\} \quad (8.12)$$

Haciendo los cálculos correspondientes, se encuentran las siguientes funciones de densidad marginal para cada uno de los modelos de tasa de interés que se obtuvieron. Los resultados se presentan en la tabla 8.1 (cabe mencionar que sólo se expresan los resultados, sin el desarrollo para llegar a estos resultados).

| Modelo Paramétrico | Función de densidad marginal $\pi(x, \theta)$ |
|------------------------------------|--|
| Vasicek (1977) | $\frac{\varepsilon(\theta)}{\beta_0} \exp\left\{\frac{2\alpha_0}{\beta_0} x + \frac{\alpha_1}{\beta_0} x^2\right\}$ |
| CIR -Cox, Ingersoll y Ross- (1985) | $\frac{\varepsilon(\theta)}{\beta_2 \cdot x^{1/2}} \cdot \exp\left\{\frac{4\alpha_0}{\beta_1} x^{1/2} + \frac{4\alpha_1}{3\beta_1} x^{3/2}\right\}$ |
| * CIR VR | $\frac{\varepsilon(\theta)}{\beta} x^{-3/2}$ |
| Brennan & Schwartz (1982) | $\frac{\varepsilon(\theta)}{\beta_2 \cdot x_0^{2\alpha_1/\beta_2}} \cdot x^{\left(\frac{2\alpha_1-2}{\beta_2}\right)} \cdot \exp\left\{\frac{2\alpha_0}{\beta_2 \cdot x_0} - \frac{2\alpha_0}{\beta_2 \cdot x}\right\}$ |
| * Brennan & Schwartz | $\frac{\varepsilon(\theta)}{\beta_1 \cdot x^{3/2}} \cdot \exp\left\{\frac{4\alpha_1}{\beta_1} x^{1/2} - \frac{4\alpha_0}{\beta_1} x^{-1/2}\right\}$ |
| Chan (1992) | $\frac{\varepsilon(\theta)}{\beta_2 \cdot x^{\beta_3}} \cdot \exp\left\{\frac{2}{\beta_2} \cdot x^{-\beta_3} \cdot \left(\frac{\alpha_0 \cdot x}{1-\beta_3} + \frac{\alpha_1 \cdot x^2}{2-\beta_2}\right)\right\}$ |
| General Drift, CEV. | $\frac{\varepsilon(\theta)}{\beta_2 \cdot x^{\beta_3}} \cdot \exp\left\{\frac{2}{\beta_2} \cdot x^{-\beta_3} \cdot \left(\frac{\alpha_0 \cdot x}{1-\beta_3} + \frac{\alpha_1 \cdot x^2}{2-\beta_3} + \frac{\alpha_2 \cdot x^3}{3-\beta_3} - \frac{\alpha_3}{\beta_3}\right)\right\}$ |
| * Merton | $\frac{\varepsilon(\theta)}{\beta_0} \exp\left\{\frac{2\alpha_0}{\beta_0} x\right\}$ |
| * Dothan | $\frac{\varepsilon(\theta)}{\beta_1} x^{-1}$ |
| * GBM | $\frac{\varepsilon(\theta)}{\beta_1} x^{-1} \cdot \exp\left\{\frac{2\alpha_1}{\beta_1} x\right\}$ |
| * CEV | $\frac{\varepsilon(\theta)}{\beta_1} x^{-\gamma} \cdot \exp\left\{\frac{2\alpha_1}{\beta_1 \cdot (2-\gamma)} x\right\}$ |

Tabla 8.1

Después se calcula el valor de $\varepsilon(\theta)$ para cada uno de los modelos, este valor depende de los parámetros $\theta = (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \beta_0, \beta_1, \beta_2, \beta_3)$ según el modelo.

8.2 Resultados.

Considerando los datos descritos anteriormente, y las funciones de densidad marginal de la sección anterior se han creado las funciones para cada modelo paramétrico en MATLAB. Además los cálculos que se han hecho en Excel. Estas funciones también hacen la minimización para encontrar el valor M de la prueba estadística. Los resultados que se obtuvieron para el modelo no-paramétrico se encuentran en la Tabla (8.2).

| | |
|---------------------------------|-----------|
| Observaciones (n) | 2,196 |
| \hat{E}_M | 11.65 |
| \hat{V}_M | 82,970.00 |
| Nivel de Confianza (95%) | 1.64485 |
| Ancho de Banda | 0.00002 |
| c(alfa)= | 11.65030 |

Tabla 8.2

A un nivel de confianza del 95 por ciento, el valor con el que se tiene que comparar el estadístico M es 11.65030, es decir si el estadístico M está por debajo de este valor, no se rechaza la hipótesis nula. En la tabla (8.3) se muestran los resultados para los modelos paramétricos:

| | VASICEK | CIR | CIR VR | Brennan & Schwartz (1982) | Brennan & Schwartz | Chan |
|----------------------|-------------|-------------|-------------|---------------------------|--------------------|-----------|
| Mínimo | 2,565.40747 | 2,564.47988 | 2,564.19831 | 645.43314 | 2,565.87205 | 612.37889 |
| Estadístico M | 74.45160 | 74.42469 | 74.41651 | 18.73135 | 74.46509 | 17.77207 |
| Resultado | RECHAZO | RECHAZO | RECHAZO | RECHAZO | RECHAZO | RECHAZO |

| | General Drift CEV | MERTON | Dothan | GBM | CEV |
|----------------------|-------------------|-------------|-----------|-------------|-------------|
| Mínimo | 2,565.43618 | 2,225.27921 | 908.92776 | 2,565.87224 | 2,565.92205 |
| Estadístico M | 74.45244 | 64.58062 | 26.37832 | 74.46509 | 74.46654 |
| Resultado | RECHAZO | RECHAZO | RECHAZO | RECHAZO | RECHAZO |

Tabla 8.3

Estos son los resultados que se han obtenido y como se observa se ha rechazado la hipótesis nula en cada uno de ellos, es decir no existen evidencia de que existan

valores paramétricos para los cuales los modelos anteriores sean una representación aceptable de los datos.⁴ En el Apéndice 2 se muestra la tabla de los coeficientes obtenidos para estos modelos.

A manera de comparación, se hizo lo mismo para otros datos como cetes 7días, cetes 1M, cetes 3M, cetes 6M, cetes 9M, cetes1Y. Para estos últimos casos, se observa de igual forma que se ha rechazado la hipótesis nula en cada uno de éstos, es decir no existen valores paramétricos para los cuales los modelos anteriores sean una representación aceptable de los datos.

Esto nos lleva a la conclusión de que para el caso de las tasas de interés de México como cetes a diferentes plazos, los modelos paramétricos considerados no son una buena representación de los datos.

⁴ Estos resultados fueron publicados en mayo de 2007 y aceptada en agosto de 2007, por José Antonio Núñez, José Luis de la Cruz y Elizabeth Ortega, en el artículo “Continuous Time Models of Interest Rate: Testing the Mexican Data (1998-2006).”, *Economics Bulletin*, Vol.7.

9. Parte II. Análisis de los Modelos de Tiempo Continuo con saltos de Poisson.

En el capítulo anterior, se hizo la comparación de los modelos paramétricos con el modelo no-paramétrico, del cual se dedujo que no existen valores paramétricos para las funciones del drift y el proceso de difusión que sean una buena representación de los datos observados. En este capítulo se pretende abordar el problema usando los modelos paramétricos que tengan asociado un proceso de difusión de saltos, con lo que se pretende que el modelo de la tasa de interés sea más preciso en el comportamiento de la tasa de interés corta observada.

Los procesos con difusión de saltos son una herramienta muy útil para varios modelos de fenómenos económicos tales como crisis en monedas, crisis en mercados financieros, incumplimientos, etc.

Cuando un proceso con saltos es incluido, la situación es más complicada. Los procesos con saltos son discontinuos, los saltos pueden incluso disparar la colas de las distribuciones lo que es más difícil de modelar, la teoría de probabilidad ayuda a este tipo de fenómenos.

En esta parte del análisis se obtienen las funciones de densidad marginal para cada uno de los modelos de tiempo continuo con saltos de Poisson a través de la ecuación “*forward Kolmogorov*” y se aplica la metodología para discernir con el modelo no-paramétrico.

9.1 Análisis Matemático.

Se considera un proceso multivariado de difusión de saltos X_t definido sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ con filtración $\{\mathfrak{F}_t\}$ que satisfacen las condiciones usuales:

$$dX_t = \mu(X_t, \theta)dt + \sigma(X_t, \theta)dW_t + J_t dN_t \quad (9.1)$$

Donde

- $\{W_t, t \geq 0\}$ es un movimiento browniano estándar;
- $\theta \in \mathfrak{R}^p$ es un vector finito de parámetros que serán estimados;
- $\mu(X_t, \theta): \mathfrak{R}^n \rightarrow \mathfrak{R}^n$ es la función del drift;
- $\sigma^2(X_t, \theta): \mathfrak{R}^n \rightarrow \mathfrak{R}^{n \times d}$ es la función de difusión del proceso;
- El proceso de salto puro N tiene una intensidad estocástica de $\lambda(X_t, \theta)$ y tamaño de salto 1, y
- El salto de tamaño J_t es independiente de \mathfrak{F}_{t-} y tiene una densidad de probabilidad $\nu(\cdot, \theta): \mathfrak{R}^n \rightarrow \mathfrak{R}$. En nuestro caso $\nu(\cdot, \theta)$ se supone normal.

Definición: La densidad de probabilidad de transición $p(\Delta, y|x, \theta)$, cuando existe, es la densidad condicional de $X_{t+\Delta} = y \in \mathfrak{R}^n$ dado $X_t = x \in \mathfrak{R}^n$.

Sea X_t un proceso Markoviano en tiempo continuo con una difusión con saltos, sobre el espacio de estado Θ . La función de distribución de probabilidad de transición está dada por:

$$p(\Delta, y|x, \theta) = \Pr\{Y_t < y | X_{t_0} = x\} \quad (9.2)$$

Donde $t > t_0$, $Y_t < y$ significa $Y_u < y_i$ para $i = 1:n$.

En este caso de difusión con saltos, cualquier parametrización P corresponde a una parametrización de la densidad marginal y de la densidad de transición:

$$\left\{ (\pi(\cdot, \theta), p(\cdot | \cdot, \theta)) \mid (\mu(\cdot, \theta), \sigma^2(\cdot, \theta), \lambda(\cdot, \theta), \nu(\cdot, \theta)) \in P, \theta \in \Theta \right\} \quad (9.3)$$

En este caso la hipótesis nula y la alternativa son de la forma:

$$\begin{aligned} H_0 : & \quad p(y|x, \theta) = p(y|x) \\ H_1 : & \quad p(y|x, \theta) \neq p(y|x) \end{aligned} \quad (9.4)$$

Se prueba la hipótesis nula para el modelo (9.1) contra la familia de modelos alternativos, con modelos Markovianos con la variable de estado X_t . La función de transición de una variable no está bien especificada por la desviación del marco Markoviano, la prueba entonces se convierte a una prueba de la adecuación por el modelo paramétrico como una proyección dentro del espacio de funciones de transición Markovianas de una única variable.

Para hacer tal prueba, se deben obtener dos estimadores de p . Uno es un estimador no paramétrico de la función de densidad de transición $p(y|x)$, el segundo es una expansión de forma cerrada de $p(y|x, \theta)$ que corresponde al modelo paramétrico asumido bajo la hipótesis nula.

Los supuestos para este modelo son los siguientes:

Supuesto 1. La matriz de varianza $V(x)$ es definida positiva para todo x en el dominio del proceso X .

Supuesto 2. La ecuación diferencial estocástica (9.1) tiene una única solución. La densidad de transición $p(\Delta, y|x)$ existe para todo x, y en el dominio de X y todo $\Delta > 0$. $p(\Delta, y|x)$ es continuamente diferenciable con respecto a Δ , y dos veces continuamente diferenciable con respecto a x y y .

Supuesto 3. La densidad de transición $p(\Delta, y|x)$ es determinada por las ecuaciones backward y forward Kolmogorov.

Supuesto 4. $\nu(\cdot)$, $\mu(\cdot)$, $\sigma(\cdot)$ y $\lambda(\cdot)$ son infinitamente diferenciables casi en cualquier sitio del dominio de X .

Si la intensidad del salto es constante, se cumple el siguiente lema:

Lema 1. Si la densidad de transición $p(\Delta, y|x)$ existe para toda x y y en el dominio de X y todo $\Delta > 0$; $p(\Delta, y|x)$ es continuamente diferenciable con respecto a Δ , y dos veces continuamente diferenciable con respecto a x y y . Entonces, la ecuación diferencial (9.1) tiene una única solución, bajo las siguientes condiciones:

- La intensidad del salto $\lambda(\cdot)$ es constante;
- $\mu(\cdot)$ y $\sigma(\cdot)$ son infinitamente diferenciables continuamente con derivadas acotadas;
- Los eigenvalores de $V(\cdot)$ están acotados inferiormente por una constante positiva, donde $V(\cdot)$ es la matriz de varianza definida positiva; y,
- La distribución de saltos de J_t tiene momentos de cualquier orden.

Proposición 1. Bajo el supuesto 2, la densidad de transición satisface las ecuaciones de “*backward and forward-Kolmogorov equations*”. En particular, “*forward-Kolmogorov equation*”:

$$\begin{aligned} \frac{\partial p(\Delta, y|x)}{\partial t} = & \frac{1}{2} \frac{\partial^2}{\partial y^2} [\sigma^2(y)p(\Delta, y|x)] - \frac{\partial}{\partial t} [\mu(y)p(\Delta, y|x)] \\ & + \int [\lambda(y-c)p(\Delta, y-c|x) - \lambda(y)p(\Delta, y|x)] \nu(c) dc \end{aligned} \quad (9.5)$$

Luego, $\nu(\cdot)$, $\mu(\cdot)$, $\sigma(\cdot)$ y $\lambda(\cdot)$ son infinitamente diferenciables al menos en cualquier sitio del dominio de X .

Para los modelos Paramétricos con saltos de Poisson, la densidad de probabilidad de transición viene dada por:

$$p(\Delta, y|x) = \Delta^{-n/2} \exp \left[-\frac{C^{(-1)}(x, y)}{\Delta} \right] \sum_{k=0}^{\infty} C^{(k)}(x, y) \Delta^k + \sum_{k=1}^{\infty} D^{(k)}(x, y) \Delta^k \quad (9.6)$$

Para alguna función $C^{(k)}(x, y)$ y $D^{(k)}(x, y)$ que serán determinadas, dependiendo el modelo paramétrico. La ecuación (9.3) es solución de “*backward and forward Kolmogorov equations*”, así sustituyendo (9.3) en éstas dos ecuaciones y haciendo sus coeficientes igual a cero, se puede resolver para $C^{(k)}(x, y)$ y $D^{(k)}(x, y)$. Las dos ecuaciones son necesarias, ya que la ecuación *backward* es una ecuación diferencial con respecto a x y su solución, el coeficiente $C^{(0)}(x, y)$ en particular, es únicamente determinada por una función multiplicativa de y . Similarmente, la solución de la ecuación *forward* es determinada por una función multiplicativa de x .

Considerando ambas ecuaciones, “*backward and forward Kolmogorov equations*”, y las dos siguientes condiciones se garantiza que la densidad converge a una función delta de Dirac en $y = x$ cuando $\Delta \rightarrow 0$, la cual es una propiedad de la función de densidad,

- 1) $C^{(-1)}(x, y) = 0$ si y sólo si $y = x$.
- 2) $C^{(0)}(x, x) = (2\pi)^{-n/2} [\det V(x)]^{-1/2}$.

Cuando $\Delta \rightarrow 0$, la condición (1) garantiza la aproximación en la densidad en los picos de x y la condición (2) sirve para asegurar que la integral de la densidad es 1 con respecto a y cuando Δ se hace pequeño.

Se puede apreciar de la ecuación (9.3) que de la primer parte de la ecuación,

$$\Delta^{-n/2} \exp\left[-\frac{C^{(-1)}(x, y)}{\Delta}\right] \sum_{k=0}^{\infty} C^{(k)}(x, y) \Delta^k, \text{ se captura el comportamiento de } p(\Delta, y|x)$$

en y cerca de x , y el término $\sum_{k=1}^{\infty} D^{(k)}(x, y) \Delta^k$ captura el comportamiento de la cola

de $p(\Delta, y|x)$. Se comienza con $k=1$ debido a que con procesos de Poisson, $p(\Delta, y|x)$ tiene una cola de orden $O(\Delta)$.

Esta relación se puede ver de forma más clara haciendo la comparación de la ecuación (9.3) con la siguiente ecuación. Suponga que $D_{t,\Delta}$ denota el evento de que no existen saltos en el intervalo $[t, t + \Delta]$, luego

$$p(\Delta, y|x) = \Pr\{D_{t,\Delta}|X_t = x\} \cdot pdf\{X_{t+\Delta} = y|X_t = x, D_{t,\Delta}\} + \Pr\{D_{t,\Delta}^c|X_t = x\} \cdot pdf\{X_{t+\Delta} = y|X_t = x, D_{t,\Delta}^c\} \quad (9.7)$$

Donde $pdf\{X_{t+\Delta} = y|X_t = x, D_{t,\Delta}\}$ denota la función de densidad de transición para un a difusión X . El salto de Poisson llega cuando ocurre el evento $D_{t,\Delta}^c$, es decir el segundo término de la ecuación anterior.

El Teorema 1 y el Teorema 2, dan un conjunto de restricciones sobre $C^{(k)}(x,y)$ y $D^{(k)}(x,y)$ para las ecuaciones “*backward*” y “*forward*” de Kolmogorov, junto con las condiciones 1 y 2 son usados para resolver la aproximación de la densidad de transición (ecuación 9.3).

Teorema 1. La ecuación “backward” impone las siguientes restricciones, para el caso univariado

$$\text{i. } C^{(-1)}(x, y) = \frac{1}{2} \left[\int_x^y \sigma(s)^{-1} ds \right]^2 \quad (9.8)$$

$$\text{ii. } C^{(0)}(x, y) = \frac{1}{\sqrt{2\pi}\sigma(y)} \exp \left[\int_x^y \left(\frac{\mu(s)}{\sigma^2(s)} - \frac{\sigma'(s)}{2\sigma(s)} \right) ds \right]$$

iii. Sea

$$C^{(k+1)}(x, y) = - \left[\int_x^y \sigma^{-1}(s) ds \right]^{-(k+1)} \times \\ \int_x^y \left[\exp \left(\int_s^x \left(\frac{\sigma'(u)}{2\sigma(u)} - \frac{\mu(u)}{\sigma^2(u)} \right) du \right) \cdot \sigma^{-1}(s) \cdot \left(\int_s^y \sigma^{-1}(u) du \right)^k \cdot [\lambda(s) - L^B] C^{(k)}(s, y) \right] ds \\ \text{para } k \geq 0$$

Donde

$$L^B(\vartheta(\cdot, \cdot)) = \sum_{i=1}^n \mu_i(x) \frac{\partial}{\partial x_i} \vartheta(\cdot, \cdot) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n v_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} \vartheta(\cdot, \cdot)$$

$$\text{iv. } D^{(1)}(x, y) = \lambda(x) \nu(y - x)$$

v. Sea

$$D^{(k+1)}(x, y) = \frac{1}{1+k} \times \\ \left[A^B D^{(k)}(x, y) + \sqrt{2\pi} \lambda(x) \sum_{r=0}^k \frac{M_{2r}^1}{(2r)!} \frac{\partial^{2r}}{\partial w^{2r}} g_{k-r}(x, y, w) \Big|_{w=0} \right] \text{ para } k > 0$$

Donde

$$g_k(x, y, w) \equiv C^{(k)}(w_B^{-1}(w), y) \cdot \nu(w_B^{-1}(w) - x) \cdot \sigma(w_B^{-1}(w)), \quad y \\ M_{2r}^1 \equiv \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{s^2}{2}\right) s^{2r} ds \quad y \quad w_B(x, y) = \int_y^x \sigma(s)^{-1} ds$$

Teorema 2. La ecuación “forward” impone las siguientes restricciones, para el caso univariado

$$i. \quad C^{(-1)}(x, y) = \frac{1}{2} \left[\int_x^y \sigma(s)^{-1} ds \right]^2 \quad (9.9)$$

$$ii. \quad C^{(0)}(x, y) = \frac{1}{\sqrt{2\pi}\sigma(y)} \exp \left[\int_x^y \left(\frac{\mu(s)}{\sigma^2(s)} - \frac{3\sigma'(s)}{2\sigma(s)} \right) ds \right]$$

iii. Sea

$$C^{(k+1)}(x, y) = - \left[\int_x^y \sigma^{-1}(s) ds \right]^{-(k+1)} \times \\ \int_x^y \left[\exp \left(\int_s^x \left(\frac{3\sigma'(u)}{2\sigma(u)} - \frac{\mu(u)}{\sigma^2(u)} \right) du \right) \cdot \sigma^{-1}(s) \cdot \left(\int_s^y \sigma^{-1}(u) du \right)^k \cdot [\lambda(s) - L^F] C^{(k)}(s, y) \right] ds \\ \text{para } k \geq 0$$

Donde

$$L^F(\mathcal{G}(\cdot, \cdot)) = - \sum_{i=1}^n \frac{\partial}{\partial x_i} [\mu_i(x) \mathcal{G}(\cdot, \cdot)] + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} [v_{ij}(x) \mathcal{G}(\cdot, \cdot)]$$

$$iv. \quad D^{(1)}(x, y) = \lambda(x) \nu(y - x)$$

v. Sea

$$D^{(k+1)}(x, y) = \frac{1}{1+k} \times \\ \left[A^B D^{(k)}(x, y) + \sqrt{2\pi} \sum_{r=0}^k \frac{M_{2r}^1}{(2r)!} \frac{\partial^{2r}}{\partial w^{2r}} [\lambda(w_F^{-1}(w)) h_{k-r}(x, y, w)]_{w=0} \right] \\ \text{para } k > 0.$$

Donde

$$h_k(x, y, w) \equiv C^{(k)}(x, w_F^{-1}(w)) \cdot \nu(y - w_F^{-1}(w)) \cdot \sigma(w_F^{-1}(w)), \quad y$$

$$M_{2r}^1 \equiv \frac{1}{\sqrt{2\pi}} \int_{\mathfrak{R}} \exp\left(-\frac{s^2}{2}\right) s^{2r} ds \quad y \quad w_F(x, y) = \int_x^y \sigma(s)^{-1} ds$$

Por ejemplo, se considera el siguiente proceso, un movimiento browniano con drift:

$$dX_t = \mu dt + \sigma dW_t \quad (9.10)$$

La verdadera densidad de transición es de una normal con media $\mu\Delta$ y varianza $\sigma^2\Delta$, es decir:

$$\begin{aligned} p(\Delta, y|x) &= \frac{1}{\sqrt{2\pi\sigma^2\Delta}} \exp\left[-\frac{(y-x-\mu\Delta)^2}{2\sigma^2\Delta}\right] \\ &= \frac{1}{\sqrt{2\pi\sigma^2\Delta}} \exp\left[-\frac{(y-x)^2}{2\sigma^2\Delta} + \frac{\mu}{\sigma^2}(y-x)\right] \exp\left[-\frac{\mu}{2\sigma^2}\Delta\right] \end{aligned} \quad (9.11)$$

Por el Teorema 1 y 2, para el caso univariado se tiene,

$$C^{(-1)}(x, y) = \frac{1}{2\sigma^2} \left[\int_x^y ds \right]^2 = \frac{1}{2\sigma^2} (y-x)^2 \quad (9.12)$$

$$C^{(0)}(x, y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{\mu}{\sigma^2} \int_x^y ds\right] = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{\mu}{\sigma^2} (y-x)\right]$$

$$\begin{aligned} C^{(1)}(x, y) &= -\frac{1}{(y-x)} \int_x^y \exp\left(-\frac{\mu}{\sigma^2}(x-s)\right) \cdot L^B C^0(x, y) ds \\ &= -C^{(0)}(x, y) \frac{\mu^2}{2\sigma^2} \end{aligned}$$

$$D^{(k)}(x, y) = 0 \quad \text{para } k > 0.$$

La aproximación a la densidad $p(\Delta, y|x)$ es por lo tanto:

$$p(\Delta, y|x) = \frac{1}{\sqrt{2\pi\sigma^2\Delta}} \exp\left[-\frac{(y-x)^2}{2\sigma^2\Delta} + \frac{\mu}{\sigma^2}(y-x)\right] \cdot \left(1 - \frac{\mu^2}{2\sigma^2}\Delta\right) \quad (9.13)$$

Utilizando la expansión de Taylor para el término $1 - \frac{\mu^2}{2\sigma^2} \Delta$ se reemplaza con $\exp\left(-\frac{\mu^2}{2\sigma^2} \Delta\right)$ y se obtiene la expresión requerida.

El cálculo de $C^{(-1)}(x, y)$ en casos multivariados no siempre tiene soluciones explícitas. Se puede encontrar una solución explícita cuando los procesos son univariados, o bien, son casos reducibles⁵. Los modelos considerados son casos reducibles, por lo que $C^{(-1)}(x, y)$ se puede estimar fácilmente.

Sin embargo, la metodología se puede aplicar para casos irreducibles, ya que existe un método irreducible que convierte un caso irreducible a uno reducible reemplazando a $C^{(-1)}(x, y)$ con sucesiones de expresiones exactas.

Retomando los modelos paramétricos de tasas de interés y haciendo los cálculos correspondientes se obtuvieron las funciones de densidad de transición para cada uno de éstos, se encuentran en la tabla 9.1. Además se supone que el tamaño del salto J_t es normal e i.i.d. con media μ_s y varianza σ_s^2 . (Se omitieron los pasos de cada cálculo, sólo se muestra el resultado en la siguiente tabla)

⁵ Un proceso de difusión X_t es llamado reducible a una difusión unitaria si y sólo si existe una transformación uno a uno del proceso de difusión X_t en un proceso de difusión Y_t , el cual tiene una matriz identidad de difusión σ_Y . Esto es, existe una función invertible $\gamma(X_t)$, infinitamente diferenciable en X_t sobre $D = (0, \infty)$, tal que $Y_t = \gamma(X_t)$ satisfaga la ecuación diferencial estocástica $dY_t = \mu_Y(Y_t)dt + dW_t$ sobre el dominio de Y_t .

| Modelo Paramétrico | Función de densidad de transición $p(\Delta, y x)$ |
|------------------------------------|---|
| Vasicek (1977) | $= \frac{1}{\sqrt{2\pi\beta_0^2\Delta}} e^{\left\{ \frac{y-x}{\beta_0^2} \left[\frac{1}{2}(y-x) + \alpha_0 + \frac{1}{2}\alpha_1(y+x) \right] - \left(\frac{\alpha_0 + \alpha_1 y}{2\beta_0^2} + \lambda \right) \Delta \right\}}$ $+ \frac{\lambda\Delta}{\sqrt{2\pi\sigma_s}} e^{-\frac{1}{2} \left(\frac{y-x-\mu_s}{\sigma_s} \right)^2}$ |
| CIR -Cox, Ingersoll y Ross- (1985) | $= \frac{1}{\sqrt{2\pi\Delta\beta_1}} \left(\frac{y}{x} \right)^{-\frac{\alpha_0}{\beta_1^2}} e^{\left\{ \frac{-2}{2\beta_1^2\Delta} (\sqrt{y}-\sqrt{x})^2 + (y-x) \left(\frac{\alpha_1}{2\beta_1^2} - \frac{3}{4} \right) - \left(\frac{\alpha_0 + \alpha_1 y}{2\beta_1^2 y} + \lambda \right) \Delta \right\}}$ $+ \frac{\lambda\Delta}{\sqrt{2\pi\sigma_s}} e^{-\frac{1}{2} \left(\frac{y-x-\mu_s}{\sigma_s} \right)^2}$ |
| * CIR VR | $= \frac{1}{\sqrt{2\pi\Delta x^3 \beta_1}} \left(\frac{y}{x} \right)^{-\frac{9}{4}} e^{\left\{ \frac{1}{\beta_1^2\Delta} \left(\frac{1}{\sqrt{y}} - \frac{1}{\sqrt{x}} \right)^2 + \lambda\Delta \right\}}$ $+ \frac{\lambda\Delta}{\sqrt{2\pi\sigma_s}} e^{-\frac{1}{2} \left(\frac{y-x-\mu_s}{\sigma_s} \right)^2}$ |
| Brennan & Schwartz (1982) | $= \frac{1}{\sqrt{2\pi\Delta\beta_2 x^2}} \left(\frac{y}{x} \right)^{\frac{3}{2}} e^{\left\{ \frac{\alpha_0}{3\beta_2^2} \left[\frac{1}{x^3} - \frac{1}{y^3} \right] + \frac{\alpha_1}{2\beta_2^2} \left[\frac{1}{x^2} - \frac{1}{y^2} \right] - \frac{1}{2\Delta\beta_2^2} \left[\frac{1}{x} - \frac{1}{y} \right] - \left(\frac{\alpha_0 + \alpha_1 y}{2\beta_2^2} + \lambda \right) \Delta \right\}}$ $+ \frac{\lambda\Delta}{\sqrt{2\pi\sigma_s}} e^{-\frac{1}{2} \left(\frac{y-x-\mu_s}{\sigma_s} \right)^2}$ |
| Chan (1992) | $= \frac{1}{\sqrt{2\pi\Delta\beta_2 x^{\beta_3-1}}} \left(\frac{y}{x} \right)^{-\frac{3}{2}\beta_3} e^{\left\{ \frac{x^{\beta_3-1} - y^{\beta_3-1}}{2\beta_2^2(1-\beta_3)\Delta} + \frac{\alpha_0(y^{1-2\beta_3} - x^{1-2\beta_3})}{\beta_2^2(1-2\beta_3)} + \frac{\alpha_1(y^{2\beta_3} - x^{2\beta_3})}{2\beta_2^2\beta_3} - \left(\frac{\alpha_0 + \alpha_1 y}{2\beta_2^2 y^{2\beta_3}} + \lambda \right) \Delta \right\}}$ $+ \frac{\lambda\Delta}{\sqrt{2\pi\sigma_s}} e^{-\frac{1}{2} \left(\frac{y-x-\mu_s}{\sigma_s} \right)^2}$ |

Tabla 9.1

| Modelo Paramétrico | Función de densidad de transición $p(\Delta, y x)$ |
|--------------------|---|
| * Merton | $= \frac{1}{\sqrt{2\pi\Delta\beta_0}} e^{\left\{-\frac{1}{2\beta_0^2\Delta}(y-x)^2 + \frac{\alpha_0}{\beta_0^2}(y-x) - \left(\frac{\alpha_0^2}{2\beta_0^2} + \lambda\right)\Delta\right\}}$ $+ \frac{\lambda\Delta}{\sqrt{2\pi\sigma_s}} e^{-\frac{1}{2}\left(\frac{y-x-\mu_s}{\sigma_s}\right)^2}$ |
| * Dothan | $= \frac{1}{\sqrt{2\pi\Delta\beta_1 x}} \left(\frac{y}{x}\right)^{-\frac{3}{2}} e^{\left\{-\frac{1}{2\beta_1^2\Delta}\left(\ln\frac{y}{x}\right)^2 - \lambda\Delta\right\}}$ $+ \frac{\lambda\Delta}{\sqrt{2\pi\sigma_s}} e^{-\frac{1}{2}\left(\frac{y-x-\mu_s}{\sigma_s}\right)^2}$ |
| * GBM | $= \frac{1}{\sqrt{2\pi\Delta\beta_1 x}} \left(\frac{y}{x}\right)^{\frac{\alpha_1}{\beta_1^2} - \frac{3}{2}} e^{\left\{\frac{1}{2\beta_1^2\Delta}\left(\ln\frac{y}{x}\right)^2 - \left(\frac{\alpha_1^2}{\beta_1^2} + \lambda\right)\Delta\right\}}$ $+ \frac{\lambda\Delta}{\sqrt{2\pi\sigma_s}} e^{-\frac{1}{2}\left(\frac{y-x-\mu_s}{\sigma_s}\right)^2}$ |
| * CEV | $= \frac{1}{\sqrt{2\pi\Delta\beta_1 x^\gamma}} \left(\frac{y}{x}\right)^{-\frac{3}{2}\gamma} e^{\left\{\frac{(y^{1-\gamma} - x^{1-\gamma})^2}{2\beta_1^2(1-\gamma)^2\Delta} + \frac{\alpha_1}{2\beta_1^2\gamma}(y^{2\gamma} - x^{2\gamma}) - \left(\frac{\alpha_1^2}{2\beta_1^2 x^{2\gamma-1}} + \lambda\right)\Delta\right\}}$ $+ \frac{\lambda\Delta}{\sqrt{2\pi\sigma_s}} e^{-\frac{1}{2}\left(\frac{y-x-\mu_s}{\sigma_s}\right)^2}$ |

Tabla 9.1
Continuación

El siguiente paso es obtener los resultados.

9.2 Resultados

Ya que se obtuvieron los resultados de la Tabla 9.1, se programaron éstas funciones en MATLAB y Excel, también se hace la minimización para encontrar el valor de la prueba estadística M . Los resultados para el modelo no-paramétrico se encuentran en la Tabla 9.2.

| | |
|---------------------------------|------------|
| Observaciones (n) | 2,196 |
| \hat{E}_M | 28.34 |
| \hat{V}_M | 590,613.48 |
| Nivel de Confianza (95%) | 1.64485 |
| Ancho de Banda | 0.00240 |
| c(alfa)= | 28.33970 |

Tabla 9.2

A un nivel de confianza del 95 por ciento, el valor con el que se tiene que comparar el estadístico M es 28.33970, es decir si el estadístico M está por debajo de este valor, no se rechaza la hipótesis nula. En la tabla 9.3 se muestran los resultados para los modelos paramétricos

| | VASICEK | Brennan & Schwartz | Chan | MERTON |
|----------------------|----------------------|--------------------|--------------|--------------|
| Mínimo | 3.98576 | 12,011.73826 | 12,014.40247 | 10,332.47595 |
| Estadístico M | 20.83483 | 62,789.21253 | 62,803.13917 | 54,011.16923 |
| Resultado | No se puede Rechazar | RECHAZO | RECHAZO | RECHAZO |

| | CIR VR | Dothan | GBM | CEV |
|----------------------|--------------|--------------|--------------|--------------|
| Mínimo | 5,495.03913 | 1,965.32643 | 9,052.06187 | 11,961.09427 |
| Estadístico M | 28,724.33381 | 10,273.39225 | 47,318.03374 | 62,524.48008 |
| Resultado | RECHAZO | RECHAZO | RECHAZO | RECHAZO |

Tabla 9.3

Estos son los resultados que se han obtenido y como se observa se ha rechazado la hipótesis nula en cada uno de ellos, excepto para el modelo Vasicek, es decir existen valores paramétricos para los cuales el modelo de Vasicek es una representación aceptable de los datos. En el Apéndice 2 se muestra la tabla de los coeficientes obtenidos para estos modelos.

10. Conclusiones.

El primer objetivo del trabajo era discernir entre modelos paramétricos de tiempo continuo y un modelo no paramétrico, después del análisis con los resultados obtenidos, se observa que los modelos de tiempo continuo usados no reflejan el comportamiento de la tasa corta para el caso de México.

A pesar de haber considerado la mayor parte de los modelos de tiempo continuo y de tratar de generalizarlos, no se pudo obtener parámetros tales que dieran una buena representación de los datos. A pesar de los ajustes que se hicieron a los datos, como considerar diferentes curvas de tasas de interés, considerar diferentes ventanas de datos, los resultados fueron similares, es decir, los modelos considerados fueron rechazados.

Sin embargo, cuando se compararon los modelos paramétricos de tiempo continuo con saltos de Poisson y el modelo no paramétrico, se obtuvo que el único modelo de tasa de interés corta con una representación aceptable de los datos para el caso de México, fue el modelo de Vasicek con saltos de Poisson. Esto se debe a que la distribución de los rendimientos no corresponde a una distribución normal sino que esta distribución tiene colas pesadas.

Con este trabajo se propone la utilización de un modelo de tiempo continuo para que las instituciones bancarias lo utilicen en la proyección de la tasa de interés corta para el mercado secundario mexicano. Aún cuando las instituciones bancarias en México están reguladas por la Secretaría de Hacienda y Crédito Público y supervisadas por el Banco de México, éste último regula el sistema de pagos. Las instituciones financieras ofrecen o cobran una tasa de interés a los ahorradores o a los que piden un préstamo, por esta razón es importante conocer la evolución de la tasa de interés.

Se recomienda aplicar esta misma metodología al tipo de cambio peso-dólar, es decir, suponer que la divisa tiene asociado un proceso de difusión estacionario $\{X_t, t \geq 0\}$, con una dinámica representada por una ecuación diferencial estocástica de Itô:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t$$

O de la forma:

$$dX_t = \mu(X_t, \theta)dt + \sigma(X_t, \theta)dW_t$$

Luego aplicar las metodologías expuestas en este trabajo para discernir entre modelos paramétricos de tiempo continuo con y sin saltos de Poisson y el modelo no paramétrico. Así al obtener los parámetros óptimos se puede analizar el comportamiento del tipo de cambio a través del tiempo.

11. Bibliografía.

- Aït-Sahalia, Y., (1992), “The Delta and Bootstrap Methods for Nonlinear Kernel Functional”. *Econometrica*.
- Aït-Sahalia, (1996). “Testing Continuous-Time Models of the Spot Interest Rate”.
- Aït-Sahalia, (1996). “Nonparametric Pricing of Interest Rate Derivative Securities”. *Econometrica* 64.
- Aït-Sahalia, (2002). “Maximum Likelihood Estimation of Discretely Sampled Diffusions”. *Econometrica* 70.
- Aït-Sahalia, (2006). “Nonparametric Transition-based Test for Jump-Diffusions”. Princeton University.
- Andersen, T. y J. Lund. (1997) “Estimating Continuous-time Stochastic Volatility Models of the Short-Term Interest Rate”. *Journal of Econometrics*.
- Bibby, B. y M. Sorensen. (1995) “Martingale Estimation Functions for Discretely Observed Diffusion Processes”. *Bernoulli* 1.
- Bickel, P., Rosenblatt, (1973). “On some global Measures of the deviations of Density Function Estimates”. *The Annals of Statistics*, Vol. 1.
- Brandt, Michael y Pedro Santa-Clara, (2002). “Simulated Likelihood Estimation of Diffusion with an Application to Exchange Rate Dynamics in Incomplete Markets”. *Journal of Financial Economics*, Vol. 63.

- Briand Philippe, (2006). “Differentiability of Backward Stochastic Differential Equations in Hilbert Spaces with Monotone Generators”.
- Bühler, Wolfgang, (1997). “An Empirical Comparison of Alternative Models for Valuing Interest Rate Options” Working paper, University Mannheim.
- Chan, K.C., Karolyi Andrew (1992). “An Empirical Comparison of Alternative models of the Short-Term Interest Rate”. Journal of Finance.
- Cox, Ingersoll y Ross (1985). “A Theory of the Term Structure of Interest Rates”. Journal of Finance.
- Dacunha Castelle y Florens-Zmirou. (1986). “Estimation of the Coefficients of a Diffusion from Discrete Observations”.
- Das, S. (1999). “Mean Rate Shifts and Alternative Models of the Interest Rate: Theory and Evidence”.
- Donald W. K. Andrews, (1997). “A Conditional Kolmogorov Test”. *Econometrica*, Vol 65, No. 5.
- Dothan, Uri. (1978) “On the Term Structure of interest Rates”. *Journal of Financial Economics*, Vol. 6.
- Duffie, D. y P. Glynn. (1997) “Estimation of continuous-Time Markov Processes Sampled at Random Time Intervals”. Working Paper, Stanford University.
- Duffie, D. y K. Singleton. (1993) “Simulated Moments Estimation of Markov Models of asset Prices”.

- Floyd B. Hanson, (2006). “Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis and Computation”. University of Illinois, Chicago.
- Gallant, R. y G. Tauchen. (1992). “Which Moments to match?”
- Gardiner, C.W., (1985). “Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences”. Springer.
- Gouriéroux, A. Monfort. (1993). “Simulation Based Inference: A Survey with special Reference to Panel Data Models”. Journal of Econometrics.
- Gouriéroux, A. Monfort y E. Renault. (1996). “Indirect Inference”. Journal of Applied Econometrics.
- Hamilton, J. D., (1988). “Rational Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates”. Journal of Economic Dynamics and Control.
- Hansen, L. P. y J. Scheinkman. (1991). “Back to the Future: Generating Moment Implications for Continuous Time Markov Processes”. University of Chicago.
- Härdle, W., Müller M., (2004). “Nonparametric and Semi parametric Models”. Springer
- Härdle, W., (1990). “Applied Nonparametric Regression”, Cambridge University Press.
- He, H. (1990) “Convergence from Discrete to Continuous-Time Contingent Claims Prices”. Review of Financial Studies.

- Hong, Y., (2005). “Nonparametric Specification Testing for Continuous-time models with Applications to term Structure of Interest Rates”. *Review of Financial Studies*, 18.
- Honoré, Peter. (1997). “Maximum-Likelihood Estimation of Non-Linear Continuous-Time Term Structure Models”. Working Paper, Aarhus University.
- Jensen, Bjarke y Rolf Poulsen, (2002). “Transition Densities of Diffusion Processes: numerical Comparison of Approximation Techniques”. *Journal of Derivatives*, Vol. 9.
- Jiang, G. J. (1998) “Nonparametric Modeling of U.S. Interest Rate Term Structure Dynamics and Implications on the Price of Derivative Securities”. *Journal of Financial and Quantitative Analysis*.
- Jilian Yu, (2003). “Closed Form likelihood Estimation of Jump-Diffusions with an Application to the Realignment Risk”. *Journal of Finance*.
- Johnson, E., Wojtkiewicz, S., (1997). “Finite element and finite difference solutions to the transient Fokker-Planck equation”.
- Karlin, S., Taylor,H, (1981). “A second course in Stochastic Processes”. Academic Press, INC.
- Kumar, Pankaj y Narayanan, (2006). “Solution of Fokker-Planck equation by finite element and finite difference methods for nonlinear systems”. *Sadhana* Vol. 31.

- Langtangeh, (1991). "A general numerical solution method for Fokker-Planck equations with applications to structural reliability". Probabilistic Engineering Mechanics 6.
- Lo, A.W., (1988). "Maximum Likelihood Estimation of Generalized Ito Processes with Discretely Sampled Data". Econometric Theory.
- Mazumdar, Ravi. Guillermin, Fabrice, (1996). "Forward Equations for Reflected Diffusions with jumps" Applied Mathematics and Optimization.
- Merton R. C., (1973). "A dynamic General Equilibrium Model of the Asset Market and its Application to the Pricing of the Capital Structure of the Firm" Working Paper, Cambridge.
- Naik, v. y M. Lee, (1993) "The Yield Curve and Bond Option Prices with Discrete Shifts in Economic Regimes". Working paper. University of British Columbia.
- Nelson D. B. (1990). "ARCH Models as Diffusion Approximations". Journal of Econometrics.
- Nelson D. B. (1991). "Conditional Heteroskedasticity in Asset Returns: A New Approach". Econometrica.
- Nelson D.B. y K. Ramaswamy. (1990). "Single Binomial Proceses as Diffusion Approximations in Financial Models". Review of Financial Studies.
- Núñez, José Antonio, José Luis de la Cruz, y Elizabeth Ortega, (2007). "Continuous Time Models of Interest Rate: Testing the Mexican Data (1998-2006)." *Economics Bulletin*, Vol. 7, No. 11 pp. 1-9

- Pagan y Ullah, (1999). “Nonparametric Econometrics”. Cambridge University.
- Pagan, A, (1990). “Alternative Models for Conditional Stock Volatility”, Journal of Econometrics.
- Pearson, N. D. (1994) “Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll, and Ross Model”. Journal of Finance.
- Pritsker, Matt. (1997). “Nonparametric Density Estimation and Tests of Continuous Time Interest Rate Models”. Review of Financial Studies
- Silverman B. W., (1986). “Density Estimation for Statistics and Data Analysis”. Chapman and Hall.
- Singleton, K. (1996). “Estimation of Affine Asset Pricing Models Using the Empirical Characteristic Function”. Journal of Econometrics.
- Stanton, Richard. (1997). “A Nonparametric model of term Structure Dynamics and the Market Price of Interest Rate Risk”. Working Paper.
- Vasicek, Oldrich, (1977). “An Equilibrium characterization of the Term Structure”. Journal of Financial Economics.
- Venegas, F. “Riesgos Financieros y Económicos”. Thompson.
- Yongmiao Hong. y Haitao Li, (2002). “Nonparametric Specification Testing for Continuous-Time Models with Application to Spot Interest Rates”

Apéndice 1.

Demostración de que la función de probabilidad de transición $p(\cdot)$ es solución de las ecuaciones *forward* y *backward* de Kolmogorov.

Sea $p_{i,j}$ la función de densidad en puntos discretos relativas a puntos candidatos y viene dada por

$$p_{i,j} = p(x + i\Delta x, y + j\Delta y) \quad (\text{A1.1})$$

La función de densidad en $p_{0,1} = p(x, y + \Delta y)$ y $p_{0,-1} = p(x, y - \Delta y)$ pueden ser expresados en series de Taylor,

$$p_{0,1} = p_{0,0} + p'_{0,0} \Delta y + \frac{1}{2} p''_{0,0} \Delta y^2 + O(\Delta y^3) \quad (\text{A1.2})$$

$$p_{0,-1} = p_{0,0} - p'_{0,0} \Delta y + \frac{1}{2} p''_{0,0} \Delta y^2 + O(\Delta y^3)$$

Resolviendo los términos derivados como puntos candidatos da

$$p'_{0,0} = \frac{\partial p(x, y)}{\partial y} = \frac{p_{0,1} - p_{0,-1}}{2\Delta y} + O(\Delta y^3) \quad (\text{A1.3})$$

$$p''_{0,0} = \frac{\partial^2 p(x, y)}{\partial y^2} = \frac{p_{0,1} - 2p_{0,0} + p_{0,-1}}{\Delta y^2} + O(\Delta y^3)$$

Obteniendo además $\frac{\partial p}{\partial x}$ y sustituyéndola en la ecuación de *Fokker Planck*, se formula una ecuación para cada uno de los N nodos, resultando en un sistema

$$\dot{p} + Kp = 0 \quad (\text{A1.4})$$

Esta ecuación puede ser discretizada en el tiempo usando método de Crank-Nicholson, así se tiene que la función de probabilidad de transición $p(\cdot)$ es solución de las ecuaciones de Kolmogorov, forward y backward.

Apéndice 2.

A continuación se muestra en la siguiente Tabla A2.1 los coeficientes obtenidos en los modelos de tiempo continuo sin considerar saltos de Poisson.

| | alfa_0 | alfa_1 | alfa_2 | alfa_3 | beta_0 | beta_1 | beta_2 | beta_3 | gamma_1 |
|---------------------------|-----------|-------------|----------|----------|----------|------------|-----------|----------|----------|
| Vasicek | -0.058721 | 0.086430 | - | - | 0.245428 | - | - | - | - |
| CIR | -0.531200 | 0.077600 | - | - | - | 2.128600 | - | - | - |
| CIR VR | - | - | - | - | - | 0.040000 | - | - | - |
| Brennan & Schwartz (1982) | 18.141384 | -161.243903 | - | - | - | - | 69.175130 | - | - |
| Brennan & Schwartz | 5.490172 | -76.939313 | - | - | - | 49.904493 | - | - | - |
| Chan | 0.035182 | 0.125418 | - | - | - | - | 2.034988 | 2.619099 | - |
| General Drift | 0.035182 | 0.125418 | 2.034988 | 2.619099 | - | - | 2.000000 | 1.330000 | - |
| Merton | -0.203503 | - | - | - | 0.032086 | - | - | - | - |
| Dothan | - | - | - | - | - | 0.000015 | - | - | - |
| GBM | - | -7.249043 | - | - | - | 1.434952 | - | - | - |
| CEV | - | 0.397323 | - | - | - | 18,966.985 | - | - | 3.294442 |

Tabla A2.1

A continuación se muestra en la siguiente Tabla A2.2 los coeficientes obtenidos en los modelos de tiempo continuo que tienen asociado un proceso de Poisson.

| | alfa_0 | alfa_1 | beta_0 | beta_1 | beta_2 | beta_3 | gamma_1 |
|--------------------|-----------|------------|----------|-----------|----------|----------|----------|
| Vasicek | -0.004298 | 0.509472 | 0.026494 | - | - | - | - |
| CIR VR | - | - | - | 0.191776 | - | - | - |
| Brennan & Schwartz | 4.892013 | -18.412410 | - | 3.899215 | - | - | - |
| Chan | 0.486253 | 1.123600 | - | - | 1.985233 | 0.800365 | - |
| Merton | 0.5203191 | - | 1.985946 | - | - | - | - |
| Dothan | - | - | - | 0.231955 | - | - | - |
| GBM | - | 5.105401 | - | 67.207303 | - | - | - |
| CEV | - | 0.977421 | - | 16.432580 | - | - | 4.036958 |

Tabla A2.2