Índice

1. Introducción **1**

2. Descripción de la Base de Datos **3**

3. Reconstrucción del Libro de Posturas **7**

4. Características Estadísticas del Libro de Posturas **13**
4.1. Precios 13
4.2. Volúmenes 17
4.3. Cancelaciones 20

5. Desarrollo Matemático **23**
5.1. Procesos Estocásticos y Tiempos de Transición 23
5.2. La Transformación de Laplace Inversa 27

6. Modelo del Libro de Posturas Limitado **32**
6.1. Modelo 32
6.2. Dirección de los Movimientos en el Precio 34
6.3. Retos de la Transformación Inversa de Laplace 36
6.4. Estimación de los Parámetros 36
6.5. Resultados 40

7. Conclusiones **44**

Índice de tablas **46**

Índice de figuras **47**

Bibliografía **48**

1. Introducción

Los mercados de valores permiten el intercambio de activos con relativa facilidad. La mayoría de los estudios del mercado tratan de explicar su comportamiento a través de procesos autorregresivos con el objetivo de establecer la relación del valor actual de una serie de tiempo con los valores anteriores; sin embargo, la mayoría de estos estudios ignoran el mecanismo básico detrás de estos mercados: la lógica racional de oferta y demanda entre compradores y vendedores.

Los mercados se pueden clasificar en dos tipos:

- Regidos por formadores de mercado (Quote-driven market)
- Regidos por órdenes (Order-driven market)

En los mercados regidos por formadores de mercado, éstos centralizan las órdenes proveyendo liquidez al fijar una postura de compra y una de venta; un ejemplo de esto es el especialista en el NYSE, el cual es el punto de contacto entre compradores y vendedores. El especialista tiene como misión mantener la continuidad del mercado, mantener a los inversionistas informados, comprometer capital cuando es necesario y asegurar equidad de oportunidades entre todos los clientes [1]. En los mercados regidos por órdenes, las posturas de todos los participantes se agregan en lo que se conoce como el libro de posturas. Los mercados regidos por órdenes tienen la ventaja de dar transparencia a todos los participantes, ya que éstos conocen las características de cada una de las posturas. De igual forma, existen algunos mercados híbridos donde los formadores de mercado interactúan en el libro de posturas; la Bolsa Mexicana de Valores y NASDAQ son algunos ejemplos.

La microestructura del mercado se define como el estudio del proceso y los resultados del intercambio de activos bajo un conjunto especifico de reglas [2]. El estudio de la microestructura del mercado es de utilidad ya que permite identificar patrones dentro del mismo. Estos patrones pueden ser utilizados para la predicción de los movimientos del mercado y para la identificación de comportamientos fraudulentos entre los participantes del mercado como se ha hecho en otros trabajos [3]. Esta información es de gran interés para los reguladores, ya que permite identificar casos de manipulación de precios.

Con el objetivo de estudiar el comportamiento de la Bolsa Mexicana de Valores, se obtuvo una base de datos con todas las posturas del libro de órdenes

en un periodo determinado. El trabajo se divide en dos partes; en la primera, se describirá la base datos sujeta al estudio y se explicará la forma en que se reconstruyó el libro de posturas para llevar a cabo el análisis. Los resultados obtenidos tras la reconstrucción se compararán con los estudios de otros mercados similares. En la segunda, se describirá un modelo del libro de posturas y un método basado en la transformación inversa de Laplace para llevar a cabo la predicción del movimiento en el precio de una emisora. Se considera que esta es la primera vez que se realiza un análisis como éste con los datos de la BMV.

2. Descripción de la Base de Datos

La Bolsa Mexicana de Valores proporcionó para este trabajo una base de datos con las posturas en todas las emisoras entre el 16 de noviembre de 2010 y el 18 de febrero de 2011. La base de datos se compone de dos tablas:

- Posturas Iniciales
- Registro de Posturas

La tabla de registro tiene 25,491,702 renglones; mientras que la tabla de posturas iniciales tiene 275,005. En el registro se encuentran todas las posturas que se reciben a lo largo del día. Las tablas tienen los siguientes campos:

id Número de identificación

folio Número asignado por la BMV a la postura
Según el reglamento de la BMV; las órdenes que, en su caso, tengan modificaciones, perderán el folio de recepción que en un inicio les haya correspondido y se les asignará uno nuevo. No perderán su folio aquellas órdenes que sean modificadas únicamente para disminuir su volumen.

fecha_vig/vigencia Fecha de vigencia de la postura

casabolsa Casa de Bolsa

tipo_mov Tipo de Movimiento
Posibles Valores

CO	Compra
VE	Venta
MO	Modificación
CA	Cancelación
AH	Reconocimiento de Hecho
MH	Modificación de Hecho
CH	Cancelación de Hecho

tipo_op Tipo de Operación
Posibles Valores

CO	Contado

PI	Pico

Son operaciones que implican la compra/venta de una cantidad de acciones inferior a un lote (para la mayoría de la emisoras son 100 acciones).

HC	Operación a Precio de Cierre
DC	Operaciones al Precio de Cierre Modalidad Después del Cierre
OR	Operaciones de Registro
OP	Oferta Pública
SR	Suscripción Recíproca
SA	Sobre Asignación

tipo_ord Tipo de Orden

Posibles Valores

LP	Limitada
MC	Mercado
MA	Mejor Postura Limitada Activa

Aquella que sigue el mejor precio límite visible en su mismo sentido y su precio buscará cerrar posturas en sentido contrario que se ubiquen dentro de su precio de protección.

MP	Mejor Postura Limitada Pasiva

Aquella que sigue el mejor precio límite visible en su mismo sentido. Aún y cuando existan posturas en sentido contrario dentro de su precio de protección la postura no buscará cerrarlas.

VO	Postura Limitada con Volumen Oculto
LO	Mejor Postura Limitada Activa con Volumen Oculto
HC	Operación a Precio de Cierre
DC	Operaciones al Precio de Cierre Modalidad Después del Cierre
PQ	Paquete

tipo_val Tipo de Valor

emisora Emisora

serie Serie

precio Precio

volumen Volumen

La tabla de posturas iniciales utiliza además estos campos:

fecha Fecha
Este campo se utiliza para distinguir las órdenes que están activas al principio del día, actualiza las posiciones activas para inicializar el libro de órdenes.

orig_timestamp Fecha original de la postura expresada como la cantidad de milésimas de segundo transcurridas desde la medianoche del 1 de enero de 1970

La tabla de registro utiliza estos campos:

timestamp Fecha de la postura expresada como la cantidad de milésimas de segundo transcurridas desde la medianoche del 1 de enero de 1970

folio_anterior Folio Anterior
Se utiliza para localizar el folio de una postura que se va a modificar o cancelar.

fecha_folio_ant Fecha de Folio Anterior

Tabla 1: Posturas Iniciales

id	fecha	orig_timestamp	folio	fecha_vig	tipo_mov	casabolsa	tipo_op	tipo_ord	tipo_val	emisora	serie	precio	volumen
247	24/11/2010	1290011934040	318480	10/12/2010	CO	1277	CO	LP	1	LAB	B	25.95	2700
292	23/11/2010	1290189919100	368726	26/11/2010	VE	1392	CO	LP	1	CHDRAUI	B	39.8	100
519	23/12/2010	1292863358080	242642	24/12/2010	VE	1277	CO	LP	1	TELMEX	L	10.18	25000
651	27/01/2011	1295881813280	93317	28/01/2011	VE	1664	CO	LP	1	ASUR	B	72	100
848	15/02/2011	1295539042060	268023	18/02/2011	CO	1392	CO	LP	1	INCARSO	B-1	12.01	1000
...	...	...	...	...	...	...	...	...	...	...	...	...	...

Tabla 2: Registro de Posturas

id	timestamp	folio	vigencia	folio_anterior	fecha_folio_ant	tipo_mov	casa_bolsa	tipo_op	tipo_ord	tipo_val	emisora	serie	precio	volumen
48	1289916001700	19		0		VE	1369	CO	LP	1	ARA	*	39.6	800
329	1290785630710	124082	26/11/2010	0		CO	1369	CO	LP	1	AMX	L	35.69	12000
519	1291907916550	16032		0		AH	1305	CO	VO	1	GNP	*	58.1	1100
652	1296246310460	911020		0		VE	1369	PI	MC	1	ASUR	B	64.05	76
998	1298062094720	566388	18/02/2011	566388	18/02/2011	MO	1288	CO	MA	1	BIMBO	A	97.51	300
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

3. Reconstrucción del Libro de Posturas

Para realizar los diversos análisis fue necesario reconstruir el libro de posturas. El libro de posturas registra todas las posturas para cada una de las emisoras. Con los datos proporcionados se reconstruyó el libro de posturas de la BMV. Uno de los principales retos que se identificaron fueron las modificaciones y cancelaciones. A lo largo del día existen un gran número de modificaciones y cancelaciones; aproximadamente el 11 % de los registros en la base de datos son modificaciones y el 35 % son cancelaciones. Por lo tanto, es indispensable identificar las posturas activas en todo momento. En otros trabajos esto no ha presentado un reto ya que se ha contado con una representación del libro que facilita el análisis. Trabajos anteriores han realizado el análisis con lo que se conoce como las "puntas". Las puntas son las mejores posturas de compra y venta en el libro de posturas. El análisis que se realiza en este trabajo es más completo ya que cuenta con la información de todas las posturas en cada instante de tiempo.

Otro reto que se identificó fue el de las posturas de tipo VO (Volumen Oculto). La base de datos que se utilizó no cuenta con la información del volumen que se encuentra oculto en este tipo de órdenes. Se optó por ignorar esta información ya que las posturas de tipo VO comprenden sólo el 0.55 % del total, mientras que las de tipo LP (Limitadas), las más comunes, representan el 93.81 %; estas últimas sí cuentan con la información del volumen. Como se muestra más adelante, ignorar las posturas de tipo VO no tuvo un gran impacto ya que los hechos reportados por el algoritmo son similares a los reportados por la BMV. Las operaciones del tipo PI (Picos) representan el 0.61 % de la muestra y de acuerdo al reglamento de la BMV no se incluyen en el libro de posturas, por lo que para el análisis también se ignoraron. Adicionalmente el algoritmo respeta el horario establecido en el reglamento de la bolsa; por lo tanto, antes de las 8:30am no se concreta ningún hecho. En la primera fase, el algoritmo carga las posturas que quedaron abiertas en días anteriores de la tabla de Posturas Iniciales y en la segunda se procesa el registro.

El algoritmo se implementó en Java y se desarrolló una interfaz gráfica. Como se puede ver en la Figura 1, la interfaz permite analizar individualmente cada entrada del registro de una emisora para un día. La interfaz cuenta con 5 ventanas:

- Corros de Venta
- Corros de Compra

- Hechos Reportados por la BMV
- Hechos Reportados por el Algoritmo
- Registro

Con el botón de "Siguiente" se van procesando una a una las entradas del registro. En el caso de una postura de compra, se trata de buscar una postura en los corros de venta para perfeccionarla; si no existe, ésta se registra en los corros de compra, los cuales están ordenados de mayor a menor precio. El mismo procedimiento se realiza para las posturas de venta, de tal forma, que la postura que se encuentra en la parte inferior de los corros de venta y la que se encuentra en la parte más alta de los corros de compra se conocen como las puntas. Una función adicional de la interfaz es la cancelación de posturas de compra o de venta. Las posturas se cancelan al seleccionarlas y oprimir el botón correspondiente. Esta función se implementó con el objeto de poder depurar errores en el funcionamiento del algoritmo.

Como se mencionó antes, los hechos se registran con la clave AH; éstos se mantienen en la ventana de "Hechos Registrados por la BMV". Cuando el algoritmo desarrollado perfecciona un hecho, éste se registra en la ventana de "Hechos Reportados por el Algoritmo". Para las cancelaciones y las modificaciones, es necesario localizar el folio correspondiente en los corros de compra o venta para realizar lo que a cada una corresponde. Con el objetivo de poder analizar el registro con mayor rapidez, se implementó el botón de "Play", el cual permite avanzar 100 entradas del registro.

Figura 1: Captura de pantalla de la Interfaz desarrollada en Java

Para medir la efectividad del algoritmo se compararon los hechos reportados por la BMV con los hechos arrojados por el algoritmo y se ignoraron acciones de muy baja bursatilidad. Por ejemplo, se analizó la acción de KUO serie B. Ésta tiene en promedio alrededor de 28 posturas diarias en la base de datos comparado con la media de 2,550 posturas diarias para las demás emisoras; y se observó que para la muestra de 68 días, el algoritmo solamente reprodujo el 50.3 % del volumen operado según la BMV. Al realizar un análisis más profundo del volumen operado de esta acción durante el plazo mencionado, se encontró que la baja efectividad del algoritmo se debe a una serie de posturas de volumen oculto que se registraron el 29 de diciembre de 2010. Si se excluye esta fecha, el algoritmo reporta un volumen de 5,607,600 acciones mientras que la BMV reporta 5,972,500, es decir, el algoritmo tiene una efectividad del 93.9 %. Para AMX serie L, la acción con mayor volumen en la BMV con un promedio 44,396 posturas por día, el algoritmo reproduce el 89.7 % de las operaciones.

No sólo es importante analizar el volumen operado, sino también la evolución del precio de una acción. Para realizar el análisis de la tendencia del precio, se utilizó el precio promedio ponderado por volumen, VWAP (Volume-weighted average price). Para la acción de KUO el VWAP del 28 de enero de 2011 registrado por la BMV es de 21.72491, el registrado por el algoritmo es de 21.72289. Por ejemplo, el 31 de diciembre de 2010 se encontró una diferencia entre el volumen reportado por la bolsa y el del algoritmo mayor al 50 %; sin embargo, el VWAP de la bolsa y del algoritmo son similares, 20.35595 y 20.27741 respectivamente. En la Figura 2 y la Figura 3 se muestra de forma gráfica la evolución del precio de KUO a lo largo de los días mencionados antes.

Para el caso de AMX, el 28 de diciembre el algoritmo tuvo solamente un 69.0 % de efectividad para replicar el volumen, sin embargo, el VWAP de la bolsa y el algoritmo difieren en menos de 0.08 %. El 20 de enero de 2011, día de mayor efectividad en términos del volumen, la bolsa reporta un VWAP de 35.28202 mientras que el algoritmo 35.26883. En la Figura 4 y la Figura 5 se encuentran unas gráficas con los precios de AMX a lo largo de esos días.

Figura 2: Reconstrucción: KUO diciembre 31

Figura 3: Reconstrucción: KUO enero 28

Figura 4: Reconstrucción: AMX enero 20

Figura 5: Reconstrucción: AMX diciembre 28

4. Características Estadísticas del Libro de Posturas

Los mercados financieros son inherentemente difíciles de analizar; sin embargo, es posible identificar algunos patrones en su comportamiento. Estudios anteriores han identificado patrones para el desempeño de precios, volúmenes y cancelaciones en el libro de posturas. A continuación, se compararán resultados obtenidos en otros mercados con una muestra de acciones de la Bolsa Mexicana de Valores. Para la muestra se eligieron emisoras de alta bursatilidad (AMX), media (BIMBO, COMERCI y GRUMA) y baja (HERDEZ y KUO). Este análisis podría extenderse fácilmente a otras acciones.

4.1. Precios

Para los precios, se ha identificado que la distribución de sus posturas sigue una ley de potencias [4], es decir, que las frecuencias decrecen exponencialmente al aumentar la variable aleatoria. Es importante mencionar que esto no se ha probado formalmente mediante métodos estadísticos.

Si se define a $b(t)$ como el precio de la mejor postura de compra ("bid") y $a(t)$ como el precio de la mejor postura de venta ("ask"), el precio de una nueva postura de compra o venta se puede expresar como $b(t) - \Delta$ ó $a(t) + \Delta$ respectivamente, donde la Δ está expresada en pujas o ticks. Una puja es el importe mínimo en que puede variar el precio unitario de cada título. Cabe mencionar que Δ puede tomar valores negativos. Si por ejemplo, con $\Delta < 0$, se registra una postura con el precio $b(t) - \Delta \geq a(t)$ entonces se concretará un hecho, mientras que si $b(t) - \Delta < a(t)$ entonces el precio $b(t+1) = b(t) - \Delta$. Si se excluyen las posturas donde $\Delta \leq 0$, se puede decir que los precios siguen una distribución definida por la siguiente función de densidad:

$$P(\Delta) \propto \frac{A}{(\Delta + \lambda)^{\alpha+1}} \tag{4.1}$$

Con el objetivo de analizar si el comportamiento de los precios de la Bolsa Mexicana sigue los patrones antes mencionados, se estimaron los parámetros del modelo utilizando el método de máxima verosimilitud para distintas acciones, excluyendo todas las modificaciones, las cuales se podrían en algunos casos considerar como una nueva postura. Otros estudios han encontrado que el parámetro α es "similar" entre las distintas emisoras [5] de un mercado en particular, sin embargo, como se puede ver en las Tablas 3 y 4, las acciones

tienen parámetros distintos. La distribución ajustada es conocida como Lomax o Pareto Tipo II con $\mu = 0$. La media de esta distribución está definida cuando $\alpha > 1$; en este análisis existen algunas emisoras para las que no se podría definir la media. En otros estudios se destaca que el parámetro α es similar para las posturas de compra y venta de una misma emisora [4]. En las acciones utilizadas en este trabajo el parámetro es similar en la mayoría de los casos; sin embargo, en el caso de KUO el parámetro α para las posturas de venta es 50 % mayor que el de las de compra.

Tabla 3: Distribución de Precios de las Posturas de Compra

	λ	α	#posturas	$\Delta > 0$
AMX	7.4441	1.7098	629,302	454,039
BIMBO	14.7809	0.8242	37,021	9,042
COMERCI	7.6756	1.2256	20,955	4,201
GRUMA	58.6601	4.4624	49,895	27,845
HERDEZ	21.1333	2.3332	6,811	1,773
KUO	11.3803	0.8556	495	93

Tabla 4: Distribución de Precios de las Posturas de Venta

	λ	α	#posturas	$\Delta > 0$
AMX	6.2232	1.5022	767,215	555,595
BIMBO	31.2291	1.2005	36,671	11,400
COMERCI	11.0780	1.6555	28,893	6,584
GRUMA	63.2991	4.1402	49,026	30,012
HERDEZ	21.1781	2.3215	5,493	1,263
KUO	42.1747	1.3034	319	19

En la Figura 6 se muestra el ajuste del modelo propuesto anteriormente contra los datos empíricos de la base de datos. Adicionalmente en la Figura 7 se pueden ver los QQ-Plots hasta el percentil 99. Para AMX y COMERCI al ser emisoras con mayor liquidez, las órdenes llegan dentro de un rango más compacto. Como puede notarse, con excepción de AMX, el modelo es adecuado para los precios en la BMV. La distribución Lomax se concentra cerca del cero, por lo que la mayoría de las observaciones al menos en el caso de COMERCI y KUO parecen seguir esta distribución. Los punto lejanos se desvían, sin embargo, no parece haber un patrón claro, por lo que podría deberse a los reducidos tamaños de las muestras.

(a) AMX compra

(b) AMX venta

(c) COMERCI compra

(d) COMERCI venta

(e) KUO compra

(f) KUO venta

Figura 6: Ajuste del Modelo de Distribución de Precios de las Posturas

(a) AMX compra

(b) AMX venta

(c) COMERCI compra

(d) COMERCI venta

(e) KUO compra

(f) KUO venta

Figura 7: QQ-Plot Distribución de Precios de las Posturas

4.2. Volúmenes

En otro trabajo para un mercado diferente [6] se demostró que los volúmenes también se pueden representar con una distribución Lomax. Se realizó un estudio similar al de precios para la distribución del volumen de las posturas en la Bolsa Mexicana de Valores. Como se puede ver en la Tabla 5, en la mayoría de los casos los parámetros para una misma emisora son "similares" para las posturas de compra y venta. Para 4 de las 6 emisoras estudiadas se puede establecer la media, ya que $\alpha > 1$. En el caso de COMERCI no se puede definir una media con el modelo ajustado para las posturas de compra ni para las de venta. En el caso de KUO no se puede definir la media para las posturas de compra; para HERDEZ no está definida para las posturas de venta.

Tabla 5: Distribución del Volumen de las Posturas

	Compra		**Venta**	
	λ	α	λ	α
AMX	115,318.20	7.2939	112,105.20	7.4092
BIMBO	689.36	1.4184	550.10	1.3052
COMERCI	863.75	0.8432	648.47	0.8336
GRUMA	6,135.30	2.6177	9,413.25	3.1836
HERDEZ	4,064.11	1.2659	2,316.45	0.8967
KUO	975.37	0.5320	24,193.91	1.9253

Por ejemplo, con el modelo ajustado, la media del volumen de las posturas de compra de AMX es de 18,322.35 títulos, mientras que el de venta es de 17,491.36. La media del volumen de compra para BIMBO es de 1,647.74 mientras que la media las posturas de venta es de 1,802.22.

En el caso de los volúmenes el ajuste para AMX es mejor que el observado en los precios; esto no es así para COMERCI y KUO. En la Figura 8 se muestran algunas gráficas de la distribución, mientras que en la Figura 9 se muestran los QQ-Plots. Es interesante observar la preferencia de los participantes por volúmenes múltiplos de 1,000.

(a) AMX compra

(b) AMX venta

(c) COMERCI compra

(d) COMERCI venta

(e) KUO compra

(f) KUO venta

Figura 8: Ajuste del Modelo de Distribución del Volumen de las Posturas

(a) AMX compra

(b) AMX venta

(c) COMERCI compra

(d) COMERCI venta

(e) KUO compra

(f) KUO venta

Figura 9: QQ-Plot Distribución del Volumen de las Posturas

4.3. Cancelaciones

La literatura en el caso de las cancelaciones es más limitada; sin embargo, se intentará también modelar la distribución de la distancia Δ de las cancelaciones del mejor precio de compra $b(t)$ y de venta $a(t)$ mediante una distribución Lomax. Se debe considerar, como se mostrará en las distribuciones empíricas, que un gran número de las cancelaciones suceden con $\Delta = 0$; es decir, al mejor precio de compra y venta. La distribución Lomax asume que $F(0) = 0$, lo cual no sucede en este análisis. Para evitar esto, sólo se considerarán las posturas donde $\Delta > 0$, y se mencionará por separado la proporción de las cancelaciones donde $\Delta = 0$.

La media de cada una de las emisoras en la muestra elegida es menor que 35; por lo tanto, las cancelaciones se concentran cerca de los mejores precios de compra y venta respectivamente. Este resultado era de esperarse ya que se presume que los inversionistas que colocan posturas lejos de las puntas tienen mayor paciencia, es decir están dispuestos a esperar más tiempo para que su postura se concrete. Es importante considerar que en las emisoras estudiadas, en el peor de los casos, cerca del 15 % de las cancelaciones suceden al mejor precio de compra o venta. En el caso de HERDEZ para las posturas de venta, más del 50 % de las cancelaciones se realizan en la punta, como puede verse en la Tabla 6.

Tabla 6: Distribución de las Cancelaciones de las Posturas

	Compra			**Venta**		
	λ	α	$\Delta = 0$	λ	α	$\Delta = 0$
AMX	7.6567	1.7677	0.1570 %	6.2821	1.5505	0.1575 %
BIMBO	3.9398	0.6827	0.4175 %	14.4289	1.1085	0.3817 %
COMERCI	5.2260	1.3656	0.4670 %	6.8606	1.7948	0.5354 %
GRUMA	88.3815	6.9082	0.2002 %	91.1204	6.4492	0.1461 %
HERDEZ	15.5851	1.7531	0.4269 %	20.0407	1.9113	0.5392 %
KUO	26.0058	1.7855	0.4031 %	22.8162	1.7143	0.6486 %

Como se puede ver en la Figura 10 y en los QQ-Plots de la Figura 11, la distribución Lomax no es la más adecuada, ya que el grueso de las cancelaciones se concentra en el intervalo donde $\Delta < 20$. Se puede afirmar que las cancelaciones para la muestra de acciones considerada de la Bolsa Mexicana no siguen una ley de potencias.

(a) AMX compra

(b) AMX venta

(c) COMERCI compra

(d) COMERCI venta

(e) KUO compra

(f) KUO venta

Figura 10: Ajuste del Modelo de Distribución de las Cancelaciones de las Posturas

(a) AMX compra

(b) AMX venta

(c) COMERCI compra

(d) COMERCI venta

(e) KUO compra

(f) KUO venta

Figura 11: QQ-Plot Distribución de las Cancelaciones de las Posturas

5. Desarrollo Matemático

Se considera que la base matemática necesaria para el modelo que se desarrollará en el próximo capítulo puede ser valiosa para el lector, por lo que a continuación se presenta en forma exhaustiva. Primero, se describen los procesos estocásticos de nacimiento y muerte; luego, se encuentra la distribución de los tiempos de transición mediante la transformación de Laplace y fracciones infinitas.

5.1. Procesos Estocásticos y Tiempos de Transición

En esta sección se define formalmente el concepto de cadena de Markov homogénea y un caso particular de éstos; los procesos de nacimiento y muerte. Un proceso estocástico $\{X(t); t \geq 0\}$ con un conjunto de estados $\mathcal{S} \subset I$ se dice que es una cadena de Markov homogénea si para toda $n \in I$, estados $i_0, i_1, \ldots, i_n, i, j$ y tiempos $0 \leq \tau_0 < \tau_1 < \cdots < \tau_n < s$

$$\begin{aligned} &P\{X(t+s) = j | X(\tau_0) = i_0 \ldots X(\tau_n)\tau_n) = i_n, X(s) = i\} \\ &= P\{X(t+s) = j | X(s) = i\} = P\{X(t) = j | X(0) = i\} \end{aligned} \tag{5.1}$$

Un proceso estocástico se conoce como proceso de nacimiento y muerte con un espacio de estados $\mathcal{S}$, si cumple con los siguientes axiomas:

1. $\mathcal{S} \subset I = \{0, 1, 2, \ldots\}$
2. Es una cadena Markov homogénea
3. Existen las constantes no negativas λ_j y μ_j, con $j = 0, 1, \ldots$ con $\mu_0 = 0$, tal que para $s > 0$ y $t \geq 0$, se cumple:

$$P\{X(t+s) = j+1 | X(t) = j\} = \lambda_j s + o(s) \quad j \in I \tag{5.2a}$$
$$P\{X(t+s) = j-1 | X(t) = j\} = \mu_j s + o(s) \quad j \in I \tag{5.2b}$$
$$P\{X(t+s) = j | X(t) = j\} = 1 - (\lambda_j + \mu_j)s + o(s) \quad j \in I \tag{5.2c}$$

$X(t)$ se conoce como el tamaño de la población al tiempo t. A λ_j se le conoce como la tasa de nacimiento y μ_j como la tasa de muerte. Una medida interesante de los procesos de nacimiento y muerte es el tiempo del primer paso del estado i al estado j, que se denota como $\sigma_{i,j}$. Para calcular la distribución de esta variable aleatoria es conveniente utilizar la transformación de Laplace bilateral, la cual se define de la siguiente forma:

$$\hat{f}(s) = \mathcal{L}\{f(t)\} = \int_{-\infty}^{\infty} e^{-st} f(t)\, dt \tag{5.3}$$

De aquí en adelante se referirá a la transformación bilateral de Laplace simplemente como la transformación de Laplace.

El tiempo $\sigma_{i,j}$ se puede expresar de la siguiente forma:

$$\sigma_{i,j} = \sigma_{i,i+1} + \sigma_{i+1,i+2} + \ldots + \sigma_{j-1,j} \tag{5.4}$$

donde $\sigma_{k,k+1}$ con $k = 1, 2, \ldots, j-1$ son independientes, lo cual permite utilizar la siguiente propiedad de la transformación de Laplace:

$$\mathcal{L}\{f_{X+Y}(t)\} = \hat{f}_{X+Y}(s) = \mathrm{E}[e^{-s(X+Y)}] \tag{5.5}$$

$$\mathrm{E}[e^{-s(X+Y)}] = \mathrm{E}[e^{-sX}]\,\mathrm{E}[e^{-sY}] = \hat{f}_X(s)\hat{f}_Y(s) \tag{5.6}$$

La transformación de la función de densidad de $\sigma_{i,j}$ se puede expresar como:

$$\mathcal{L}\{\sigma_{i,j}(t)\} = \hat{f}_{i,j}(s) = \prod_{k=i}^{j-1} \hat{f}_{k,k+1}(s) \tag{5.7}$$

haciendo deseable encontrar una transformación $\mathcal{L}^{-1}$ tal que:

$$\mathcal{L}^{-1}\{\hat{f}(s)\} = f(t) \tag{5.8}$$

Primero es necesario encontrar una expresión para $\hat{f}_{i,i-1}$. Ahora, si se condiciona $\sigma_{i,i-1}$ a la primera transición, donde S_1 es el tiempo para el primer evento, se puede expresar como:

$$\sigma_{i,i-1} = \begin{cases} S_1, & \text{si } X(S_1) = i-1 \\ S_1 + \sigma_{i+1,i-1}, & \text{si } X(S_1) = i+1 \end{cases} \tag{5.9}$$

$$\sigma_{i+1,i-1} = \sigma_{i+1,i} + \sigma_{i,i-1} \tag{5.10}$$

$$\begin{aligned} F_{i,i-1}(t) = P\{\sigma_{i+1,i-1} \le t\} &= P\{\sigma_{i,i-1} \le t | X(S_1) = i-1\} P\{X(S_1) = i-1\} + \\ &\quad P\{\sigma_{i,i-1} \le t | X(S_1) = i+1\} P\{X(S_1) = i+1\} \\ &= P\{S_1 \le t\} P\{X(S_1) = i-1\} + \\ &\quad P\{S_1 + \sigma_{i+1,i-1} \le t\} P\{X(S_1) = i+1\} \\ &= \tfrac{\mu_i}{\lambda_i+\mu_i} P\{S_1 \le t\} + \tfrac{\lambda_i}{\lambda_i+\mu_i} P\{S_1 + \sigma_{i+1,i-1} \le t\} \\ &= \tfrac{\mu_i}{\lambda_i+\mu_i} F_{S_1}(t) + \tfrac{\lambda_i}{\lambda_i+\mu_i} F_{S_1+\sigma_{i+1,i-1}}(t) \end{aligned}$$

$$\begin{aligned}\tfrac{d}{dt}F_{i,i-1}(t) &= f_{i,i-1}(t) &= \tfrac{\mu_i}{\lambda_i+\mu_i}\tfrac{d}{dt}F_{S_1}(t) + \tfrac{\lambda_i}{\lambda_i+\mu_i}\tfrac{d}{dt}F_{S_1+\sigma_{i+1,i-1}}(t)\\ &&= \tfrac{\mu_i}{\lambda_i+\mu_i}f_{S_1}(t) + \tfrac{\lambda_i}{\lambda_i+\mu_i}f_{S_1+\sigma_{i+1,i-1}}(t)\end{aligned}$$

Ahora S_1 tiene una distribución exponencial con media $\frac{1}{\lambda_i+\mu_i}$ por lo que:

$$\mathcal{L}\{f_{S_1}(t)\} = \int_0^\infty e^{-st} f_{S_1}(t)\,dt = (\lambda_i+\mu_i)\int_0^\infty e^{-(\lambda_i+\mu_i+s)t}\,dt = \frac{\lambda_i+\mu_i}{\lambda_i+\mu_i+s}$$

$$\begin{aligned}\mathcal{L}\{f_{i,i-1}(t)\} &= \hat{f}_{i,i-1}(s) &= \mathcal{L}\{\tfrac{\mu_i}{\lambda_i+\mu_i}f_{S_1}(t) + \tfrac{\lambda_i}{\lambda_i+\mu_i}f_{S_1+\sigma_{i+1,i-1}}(t)\}\\ &&= \tfrac{\mu_i}{\lambda_i+\mu_i}\hat{f}_{S_1}(s) + \tfrac{\lambda_i}{\lambda_i+\mu_i}f_{S_1+\sigma_{i+1,i-1}}(s)\\ &&= \tfrac{\mu_i}{\lambda_i+\mu_i}\tfrac{\lambda_i+\mu_i}{\lambda_i+\mu_i+s} + \tfrac{\lambda_i}{\lambda_i+\mu_i}\hat{f}_{S_1}(s)\hat{f}_{\sigma_{i+1,i}}(s)\hat{f}_{\sigma_{i,i-1}}(s)\\ &&= \tfrac{\mu_i}{\lambda_i+\mu_i}\tfrac{\lambda_i+\mu_i}{\lambda_i+\mu_i+s} + \tfrac{\lambda_i}{\lambda_i+\mu_i}\tfrac{\lambda_i+\mu_i}{\lambda_i+\mu_i+s}\hat{f}_{i+1,i}(s)\hat{f}_{i,i-1}(s)\end{aligned}$$

$$\hat{f}_{i,i-1}(s) = \frac{\mu_i}{\lambda_i+\mu_i+s-\lambda_i\hat{f}_{i+1,i}(s)} \tag{5.11}$$

Otro concepto que es necesario introducir es lo que se conoce como fracciones continuas. Las cuales son expresiones de la forma:

$$\cfrac{a_1}{b_1+\cfrac{a_2}{b_2+\cfrac{a_3}{b_3+\ddots}}}$$

Una fracción continua infinita se define con las sucesiones $[\{a_n\}_1^\infty, \{b_n\}_1^\infty, \{w_n\}_1^\infty]$ donde t_k se define como:

$$t_k :\rightarrow \frac{a_k}{b_k+u}$$

donde:

$$w_n := t_1 \circ t_2 \circ \cdots \circ t_n(0)$$

por lo que se puede denotar a la fracción continua $[\{a_n\}_1^\infty, \{b_n\}_1^\infty, \{w_n\}_1^\infty]$ de la siguiente forma:

$$\Phi_{k=i}^\infty \frac{a_n}{b_n}$$

A w_n se le conoce como el n-ésimo aproximante y de manera general se puede calcular de la siguiente forma:

$$w_n = \frac{P_n}{Q_n}$$

donde $P_0 = 0, P_1 = a_1, Q_0 = 1, Q_1 = b_1$ con:

$$P_n = b_n P_{n-1} + a_n P_{n-2}$$

$$Q_n = b_n Q_{n-1} + a_n Q_{n-2}$$

Ahora iterando en la expresión que se obtuvo previamente en 5.11 para $\hat{f}_{i,i-1}(s)$, se obtiene:

$$\hat{f}_{i,i-1}(s) = \cfrac{\mu_i}{\lambda_i + \mu_i + s - \lambda_i \cfrac{\mu_{i+1}}{\lambda_{i+1} + \mu_{i+1} + s - \lambda_{i+1} \cfrac{\mu_{i+2}}{\lambda_{i+2} + \mu_{i+2} + s + \ddots}}}$$

$$\hat{f}_{i,i-1}(s) = -\frac{1}{\lambda_{i-1}} \cfrac{-\lambda_{i-1}\mu_i}{\lambda_i + \mu_i + s + \cfrac{-\lambda_i \mu_{i+1}}{\lambda_{i+1} + \mu_{i+1} + s + \cfrac{-\lambda_{i+1}\mu_{i+2}}{\lambda_{i+2} + \mu_{i+2} + s + \ddots}}}$$

$$\hat{f}_{i,i-1}(s) = -\frac{1}{\lambda_{i-1}} \Phi_{k=i}^{\infty} \frac{-\lambda_{k-1}\mu_k}{\lambda_k + \mu_k + s} \tag{5.12}$$

Debido a las propiedades de la función de densidad f, se puede probar [7] que $\hat{f}$ converge y está acotada entre la sucesión creciente de aproximantes pares w_{2n} y la sucesión decreciente de aproximantes impares w_{2n+1}:

$$w_{2n}(s) < \hat{f}(s) < w_{2n+1}(s) \text{ para toda } n$$

Más adelante será de interés para el análisis encontrar la distribución del tiempo del primer paso del estado i al estado 0, definido como σ_i, cuya transformada de Laplace se calcula por 5.7 de esta forma:

$$\hat{f}_i(s) = \prod_{k=1}^{i} \hat{f}_{k,k-1}(s) \tag{5.13}$$

por lo que sustituyendo de 5.12 y asumiendo que $\lambda_i = \lambda$ para toda i:

$$\hat{f}_i(s) = \left(-\frac{1}{\lambda}\right)^i \left(\prod_{k=1}^{i} \Phi_{k=i}^{\infty} \frac{-\lambda\mu_k}{\lambda + \mu_k + s}\right) \tag{5.14}$$

5.2. La Transformación de Laplace Inversa

El reto que aún queda es encontrar una forma de calcular numéricamente $\mathcal{L}^{-1}$, con el objetivo de poder calcular la distribución del tiempo del primer paso del estado i al estado 0. Esta distribución se utilizará más adelante para calcular la probabilidad de que en el libro de posturas el siguiente movimiento aumente el precio medio. $\mathcal{L}^{-1}$ está definida por la siguiente integral de variable compleja, que se puede evaluar a lo largo de la línea vertical $\gamma = a$, de la siguiente manera:

$$\begin{aligned}
f(t) = \mathcal{L}^{-1}\{\hat{f}(s)\} &= \frac{1}{2\pi i} \lim_{T\to\infty} \int_{a-iT}^{a+iT} e^{st}\hat{f}(s)\,ds \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{(a+iu)t}\hat{f}(a+iu)\,du \\
&= \frac{e^{at}}{2\pi} \int_{-\infty}^{\infty} (\cos ut + i \sin ut)\hat{f}(a+iu)\,du \\
&= \frac{e^{at}}{2\pi} \int_{-\infty}^{\infty} \Big[Re(\hat{f}(a+iu)) \cos ut - Im(\hat{f}(a+iu)) \sin ut\Big]\,du \\
&= \frac{2e^{at}}{\pi} \int_{0}^{\infty} Re(\hat{f}(a+iu)) \cos ut\,du
\end{aligned}$$

Para evaluar esta integral se utiliza un método [8] que se basa en la expansión de series de Fourier. Si se considera una función $h(t)$ tal que $h(t) = 0$ para $t < 0$ la cual se dividirá en segmentos de tamaño T, al reflejar cada segmento, se puede construir un conjunto de funciones pares $g_n(t)$, tal que (Ver Figura 12):

$$g_n(t) = \begin{cases} h(t), & \text{si } nT \le t \le (n+1)T \\ h(2nT - t), & \text{si } (n-1)T \le t \le nT \end{cases}$$

Estas funciones se pueden expresar en el intervalo $(-T, T)$. De tal forma que para $n = 0, 2, 4, \ldots$:

$$g_n(t) = \begin{cases} h(nT + t), & \text{si } 0 \leq t \leq T \\ h(nT - t), & \text{si } -T \leq t \leq 0 \end{cases}$$

y para $n = 1, 3, 5, \ldots$

$$g_n(t) = \begin{cases} h((n+1)T - t), & \text{si } 0 \leq t \leq T \\ h((n+1)T + t), & \text{si } -T \leq t \leq 0 \end{cases}$$

La elección de T es de gran importancia, en 6.3 se profundiza en esto.

La expansión en series de Fourier de $g_n(t)$ es la siguiente:

$$g_n(t) = \frac{a_{n,0}}{2} + \sum_{k=1}^{\infty} a_{n,k} \cos\left(\frac{k\pi t}{T}\right)$$

donde $a_{a,k}$ está definido como:

$$a_{n,k} = \begin{cases} \dfrac{2}{T} \displaystyle\int_0^T h(nT + x) \cos\left(\frac{k\pi}{T} x\right) dx, & \text{si } n = 0, 2, 4, \ldots \\ \dfrac{2}{T} \displaystyle\int_0^T h((n+1)T - x) \cos\left(\frac{k\pi}{T} x\right) dx, & \text{si } n = 1, 3, 5, \ldots \end{cases}$$

Esto se puede expresar de forma compacta como:

$$a_{n,k} = \frac{2}{T} \int_{nT}^{(n+1)T} h(t) \cos\left(\frac{k\pi}{T} t\right) dt$$

Ahora sumando las funciones $g_n(t)$:

$$\sum_{n=0}^{\infty} g_n(t) = \frac{2}{T} \left[\frac{A(w_0)}{2} + \sum_{k=1}^{\infty} A(w_k) \cos\left(\frac{k\pi}{T} t\right) \right]$$

donde:

$$A(w_k) = \int_0^{\infty} h(t) \cos\left(\frac{k\pi}{T} t\right) dt$$

Ahora si se introduce un factor de atenuación tal que:

$$h(t) = e^{-at} f(t)$$

Figura 12: Funciones $g_n(t)$

de tal forma que $A(w_k)$ es la transformación de Laplace de una función en los números reales.

De esta forma, multiplicando ambos lados por el factor de atenuación:

$$\sum_{n=0}^{\infty} e^{at} g_n(t) = \frac{2e^{at}}{T}\left[\frac{1}{2}\{\hat{f}(a)\} + \sum_{k=1}^{\infty} Re\left(\hat{f}\left(a + \frac{k\pi i}{T}\right)\right)\cos\left(\frac{k\pi}{T}t\right)\right]$$

por lo que de la definición de $g_n(t)$:

$$\sum_{n=0}^{\infty} e^{at} g_n(t) = \sum_{n=0}^{\infty} e^{at} h(2nT + t) + \sum_{n=1}^{\infty} e^{at} h(2nT - t)$$

de tal forma que sustituyendo $h(t) = e^{-at} f(t)$:

$$\sum_{n=0}^{\infty} e^{at} g_n(t) = f(t) + \text{error}$$

Con esto se puede deducir el resultado:

$$f(t) \approx \frac{2e^{at}}{T}\left[\frac{1}{2}\{\hat{f}(a)\} + \sum_{k=1}^{\infty} Re\left(\hat{f}\left(a + \frac{k\pi i}{T}\right)\right)\cos\left(\frac{k\pi}{T}t\right)\right] \tag{5.15}$$

Esto es similar al resultado que se obtiene si se utilizara la regla del trapecio para evaluar la integral de la transformación inversa, la cual tiene como resultado la siguiente expresión:

$$f(t) \approx f_h(t) = \frac{he^{at}}{\pi} Re(\hat{f}(a)) + \frac{2he^{at}}{\pi}\sum_{k=1}^{\infty} Re(\hat{f}(a + ikh))\cos kht$$

para acelerar la convergencia de la serie [9], se toma $h = \pi/2t$ y $a = A/2t$, se llega a la siguiente serie, la cual es casi alternante:

$$f_h(t) = \frac{e^{A/2}}{2t} Re\left(\hat{f}\left(\frac{A}{2t}\right)\right) + \frac{e^{A/2}}{t}\sum_{k=1}^{\infty}(-1)^k Re\left(\hat{f}\left(\frac{A + 2k\pi i}{2t}\right)\right)$$

esta aproximación tiene un error de discretización acotado [9].

Debido a que la serie es casi alternante (la serie sería alternante si $Re\left(\hat{f}\left(\frac{A+2k\pi i}{2t}\right)\right)$ tuviera el mismo signo para toda k) se pueden utilizar métodos para acelerar la convergencia, tal como la transformación de Euler. Al utilizar este método se obtiene la siguiente aproximación:

$$E(m,n,t) = \sum_{k=0}^{m} \binom{m}{k} 2^{-m} s_{n+k}(t)$$

donde:

$$s_n = \frac{e^{A/2}}{2t} Re\left(\hat{f}\left(\frac{A}{2t}\right)\right) + \frac{e^{A/2}}{t} \sum_{k=1}^{n} (-1)^k Re\left(\hat{f}\left(\frac{A+2k\pi i}{2t}\right)\right)$$

Se sugiere utilizar $m = 11$ y $n = 15$. Para probar la ventaja de la aceleración se encontró la transformación de Laplace de la función de densidad de una variable aleatoria con distribución exponencial. La aceleración permite encontrar una aproximación muy cercana sumando 26 términos; mientras que cuando no se recurre a este método, el error de aproximación es aún mayor utilizando más de 7,000 términos como se puede ver en la Tabla 7. Es importante mencionar que las funciones deben de tener ciertas propiedades para que la transformación de Euler sea útil.

Tabla 7: Transformación de Euler

$$f(x) = \tfrac{1}{2} e^{-\frac{1}{2}x}$$

$f(2)$	0.183940
$f(4)$	0.067668

	Normal	**Euler**
$f(2) = \mathcal{L}^{-1}\left\{\hat{f}(s)\right\}$	0.183990	0.183940
Iteraciones	7,152	26

	Normal	**Euler**
$f(4) = \mathcal{L}^{-1}\left\{\hat{f}(s)\right\}$	0.067718	0.067668
Iteraciones	7,494	26

6. Modelo del Libro de Posturas Limitado

Para modelar la dinámica del libro de posturas limitado se considera el modelo propuesto en [10]. Este modelo tiene varias bondades que permiten la utilización de las herramientas matemáticas presentadas en el capitulo anterior para realizar el cálculo de las probabilidades de algunos eventos. El modelo también tiene ciertas limitaciones. En primer lugar considera que las llegadas de las posturas siguen un proceso Poisson, es decir, que son completamente aleatorias. También considera posturas con un volumen estandarizado o único.

6.1. Modelo

Si se considera un mercado donde las posturas se pueden colocar en una matriz de precios $\{1, \ldots, n\}$ múltiplos de una puja. n se elige de forma que es suficientemente grande para que el precio de todas las posturas estén dentro de este rango. El estado del libro se puede representar mediante el proceso $X(t) = \{X_1(t), \ldots, X_n(t)\}$ donde $|X_p(t)|$ es el número de posturas a un precio p. Si $X_p(t) < 0$ se dice que hay $-X_p(t)$ posturas de compra, cuando $X_p > 0$ existen X_p posturas de venta.

El precio de la mejor postura de venta $p_A(t)$ se define como:

$$p_A(t) = \text{ínf}\{p = 1, \ldots, n, X_p(t) > 0\}$$

de igual forma el precio de la mejor postura de venta se define como:

$$p_B(t) = \sup\{p = 1, \ldots, n, X_p(t) < 0\}$$

También se pueden definir las siguientes cantidades, el precio medio $p_M(t)$ y la diferencia o "spread" $p_S(t)$ entre la mejor postura de compra y venta respectivamente:

$$p_M(t) = \frac{p_B(t) + p_A(t)}{2}$$
$$p_S(t) = p_A(t) - p_B(t)$$

El libro de posturas limitado se actualiza al llegar nuevas posturas. Para un estado $x \in \mathbb{Z}^n$ con $1 \leq p \leq n$ se define:

$$x^{p\pm 1} = x \pm (0, \ldots, 1, \ldots, 0)$$

donde el 1 se encuentra en la posición p del vector anterior. De esta forma:

- Una postura de compra a un precio $p < p_A(t)$ incrementa la cantidad al nivel p: $x \to x^{p-1}$
- Una postura de venta a un precio $p > p_B(t)$ incrementa la cantidad al nivel p: $x \to x^{p+1}$
- Una postura de compra a mercado disminuye la cantidad al mejor precio de venta: $x \to x^{p_A(t)-1}$
- Una postura de venta a mercado disminuye la cantidad al mejor compra de venta: $x \to x^{p_B(t)+1}$
- Una cancelación de una postura activa de compra a un precio $p < p_A(t)$ disminuye la cantidad al nivel p: $x \to x^{p+1}$
- Una cancelación de una postura activa de venta a un precio $p > p_B(t)$ disminuye la cantidad al nivel p: $x \to x^{p-1}$

Los eventos mencionados arriba se modelan con procesos de Poisson independientes, de tal forma que para $i \geq 0$:

- Las posturas limitadas de compra (venta) arriban a una distancia de i pujas de la mejor postura contraria de acuerdo a distribuciones exponenciales independientes con tasa $\lambda(i)$
- La llegada de las posturas a mercado siguen distribuciones exponenciales independientes con tasa μ
- La cancelación de posturas a una distancia de i pujas de la mejor postura contraria sucede de acuerdo a una distribución exponencial donde la tasa es proporcional al número de órdenes activas, es decir, si a cierto nivel existe un número x de posturas activas, la tasa de cancelación es $\theta(i)x$
- Los eventos mencionados anteriormente son independientes

De esta forma se puede decir que el proceso X es una cadena de Markov en tiempo continuo, con el espacio de estados $\mathbb{Z}^n$ con las siguientes tasas de transición:

- $x \to x^{p-1}$ con tasa $\lambda(p_A(t) - p)$ para $p < p_A(t)$

- $x \to x^{p+1}$ con tasa $\lambda(p - p_B(t))$ para $p > p_B(t)$
- $x \to x^{p_B(t)+1}$ con tasa μ
- $x \to x^{p_A(t)-1}$ con tasa μ
- $x \to x^{p+1}$ con tasa $\theta(p_A(t) - p)|x_p|$ para $p < p_A(t)$
- $x \to x^{p-1}$ con tasa $\theta(p - p_B(t))|x_p|$ para $p > p_B(t)$

6.2. Dirección de los Movimientos en el Precio

Se desea calcular la probabilidad de que en el siguiente movimiento el precio medio $p_M(t)$ aumente. El precio medio se altera en el tiempo del primer paso de las colas al mejor precio de compra o venta a cero. También este precio se puede alterar si el "spread" $p_S(t) > 1$ y llega una postura dentro de este "spread". Se define a $X_A = X_{p_A(\cdot)}(\cdot)$ y $X_B = X_{p_B(\cdot)}(\cdot)$ y $\tau = \inf\{t \geq 0, p_M(t) \neq p_M(0)\}$.

Dada una configuración inicial del libro de posturas y un "spread" inicial S, la probabilidad de que en el siguiente movimiento el precio medio aumente se puede expresar de la siguiente forma:

$$P\{p_M(\tau) > p_M(0)|X_A(0) = a, X_B(0) = b, p_S(0) = S\}$$

Los procesos X_A y X_B son procesos de nacimiento y muerte con las siguientes tasas de transición:

$$n \to n + 1 \text{ con una tasa } \lambda(S)$$

$$n \to n - 1 \text{ con una tasa } \mu + n\theta(S)$$

En particular cuando $S = 1$ el precio medio $p_M(t)$ se altera cuando X_A ó X_B alcanzan el estado 0. Ahora, si se denota σ_A y σ_B como el tiempo del primer paso de los procesos X_A y X_B a cero, la distribución de τ está dada por el mínimo de los tiempos del primer paso a cero de los procesos X_A y X_B. La probabilidad de interés se puede expresar como $P\{\sigma_A < \sigma_B|X_A(0) = a, X_B(0) = b, p_S(0) = S\}$. La transformación de Laplace de la distribución condicional de $\sigma_A - \sigma_B$ está dada por:

$$\hat{f}_a^1(s)\hat{f}_b^1(-s)$$

donde por 5.14:

$$\hat{f}_j^S(s) = \left(-\frac{1}{\lambda(S)}\right)^j \left(\prod_{i=1}^{j} \Phi_{k=i}^{\infty} \frac{-\lambda(S)(\mu + k\theta(S))}{\lambda(S) + \mu + k\theta(S) + s}\right)$$

Por lo que la probabilidad buscada se obtiene al evaluar la función de distribución acumulada en 0. Para hacer esto es necesario encontrar la transformación inversa de $\hat{f}_a^1(s)\hat{f}_b^1(-s)$ e integrar esta función.

En el caso de que $S > 1$, es necesario considerar que si se recibe una postura dentro del spread el precio medio $p_M(t)$ también se alterará. Por lo que si se denota σ_A^i y σ_B^i como el tiempo en el que llega la primera orden a i pujas del mejor precio de venta o compra respectivamente. El tiempo en el que cambia por primera vez $p_M(t)$ está dado por el mínimo de los tiempos $\sigma_A, \sigma_B, \sigma_A^i$ y σ_B^i con $i = 1, \ldots, S-1$. Es importante mencionar que σ_A^i y σ_B^i se distribuyen de manera exponencial con tasa $\lambda(i)$. Para que $p_M(t)$ aumente es necesario que arribe una orden de compra a $S-1$ pujas del mejor precio de venta o se agoten las posturas al mejor precio de venta antes de que llegue una postura de venta a $S-1$ pujas de la mejor postura de compra o se agoten las posturas al mejor precio de compra.

En este caso la probabilidad buscada está dada por $P\{\sigma_A \wedge \sigma_B^1 \wedge \cdots \wedge \sigma_B^{S-1} < \sigma_B \wedge \sigma_A^1 \wedge \cdots \wedge \sigma_A^{S-1} | X_A(0) = a, X_B(0) = b, p_S(0) = S\}$. El mínimo de un conjunto de variables aleatorias que se distribuyen de forma exponencial con parámetros $\lambda_1, \ldots, \lambda_n$ tiene una distribución exponencial con parámetro $\lambda_1 + \ldots + \lambda_n$. Por lo que se puede expresar la probabilidad de interés de la siguiente forma $P\{\sigma_A \wedge \sigma_B^{\Sigma} < \sigma_B \wedge \sigma_A^{\Sigma} | X_A(0) = a, X_B(0) = b, p_S(0) = S\}$. Por lo que se puede probar, que la transformación de Laplace de la función de densidad de $\sigma_A \wedge \sigma_B^{\Sigma} - \sigma_B \wedge \sigma_A^{\Sigma}$ está dada por [10]:

$$\left(\hat{f}_a^S(\Lambda + s) + \frac{\Lambda}{\Lambda + s}(1 - \hat{f}_a^S(\Lambda + s))\right)\left(\hat{f}_b^S(\Lambda + s) + \frac{\Lambda}{\Lambda + s}(1 - \hat{f}_b^S(\Lambda + s))\right)$$

donde $\Lambda = \sum_{i=1}^{S-1} \lambda(i)$. De igual forma es necesario encontrar la transformación, integrar la función y evaluar en 0.

6.3. Retos de la Transformación Inversa de Laplace

Los métodos para invertir la transformación de Laplace discutidos en la sección anterior están diseñados para la transformación unilateral de Laplace, es decir definida en los reales positivos. La distribución de $\sigma_A \wedge \sigma_B^{\Sigma} - \sigma_B \wedge \sigma_A^{\Sigma}$ están definidas en todos los reales, por lo que es necesario utilizar algunas herramientas para poder encontrar la transformación inversa.

Es importante tener presente que el método utilizado para encontrar la transformación inversa presentado en este trabajo está basado en la expansión en series de Fourier, el cual utiliza funciones pares. Además es importante notar que el método permite encontrar la transformación inversa para $t \in (0, T)$. Las funciones de densidad estudiadas no son pares y están definidas para $t < 0$ por lo que para resolver este problema se define la función $g(t) = f(t - a)$, donde a es tal que:

$$\lim_{t \to 0^-} g(t) = 0$$

En este caso se realizaron diversas simulaciones de los procesos X_A y X_B para evaluar la distribución de τ, la cual disminuye exponencialmente de ambos lados a partir de un cierto máximo. En la Figura 13 se muestra el resultado que se obtiene de las simulaciones y el que se obtiene al utilizar el método de la transformación inversa de Laplace.

Si se define a $u(t)$ como la función de escalón unitario, es decir:

$$u(t) = \begin{cases} 0, & \text{si } x < 0 \\ 1, & \text{si } x \geq 0 \end{cases}$$

debido a las propiedades de la transformación de Laplace:

$$\mathcal{L}\left\{g(t)u(t-a)\right\} = \mathcal{L}\left\{f(t-a)u(t-a)\right\} = e^{-as}\hat{f}(s)$$

por lo que de esta forma se puede desplazar la función de densidad de forma que $\lim_{t \to 0^-} g(t) = 0$, como se puede apreciar en la Figura 14.

También es importante utilizar el parámetro correcto T para encontrar la distribución correcta. Como se puede ver en la Figura 15, la función de densidad es radicalmente distinta cuando T no es lo suficientemente grande.

6.4. Estimación de los Parámetros

Debido a que el modelo considera posturas con un volumen estandarizado o único, es necesario encontrar el promedio de las posturas limitadas S_l, posturas

Figura 13: Función de densidad de τ

Figura 14: Función de densidad de τ

Figura 15: Función de densidad de τ

a mercado S_m y de las cancelaciones S_c. Con esto se estima la tasa de llegada de las ordenes limitadas λ a una distancia de i pujas, $1 \leq i \leq 5$

$$\hat{\lambda}(i) = \frac{N_l(i)}{T_*}$$

donde $N_l(i)$ es el número de posturas que llegaron a una distancia i de la mejor postura contraria y T_* es el tiempo total en el que el mercado estuvo abierto expresado en minutos. Para $i > 5$ esto se extrapola siguiendo la siguiente ley de potencias:

$$\hat{\lambda}(i) = \frac{k}{i^\alpha}$$

La tasa de llegada de las posturas a mercado μ se estima de la siguiente forma:

$$\hat{\mu} = \frac{N_m}{T_*}\frac{S_m}{S_l}$$

N_m es el número de posturas de mercado; en este caso se consideran como posturas de mercado todas aquellas que superan la mejor postura contraria.

En trabajos anteriores no se cuenta con información tan detallada del libro de posturas, por lo que no se han considerado las modificaciones que en este caso forman una parte muy importante de la muestra. El promedio del volumen de las modificaciones para BIMBO es casi 10 veces mayor al de las nuevas posturas. Esto se puede deber a que los participantes del mercado que están interesados en comprar o vender un volumen mayor seguramente lo harán a través de posturas limitadas, ya que son más sensibles al precio. Las posturas limitadas garantizan un precio, mientras que a través de una postura a mercado se garantiza la ejecución pero no el precio. Los participantes con posturas de gran volumen se verán obligados a adaptar el precio de sus posturas a través de modificaciones para lograr la ejecución de las mismas. Existen tres alternativas:

- Considerar las modificaciones como una cancelación y una nueva postura
- Ignorar las modificaciones
- Modelar las modificaciones como otro proceso

La primera alternativa es la que más se acerca a estudios anteriores ya que estos no cuentan con la noción de las modificaciones, al tener la información de las puntas del libro de posturas sólo notan la cancelación de una postura y la aparición de una nueva a otro nivel. No se considera la diferencia en el volumen de las nuevas posturas y de las modificaciones. Por ejemplo, en la acción de BIMBO el volumen de las nuevas posturas limitadas de compra es de 2,005.23. Mientras que si se toman en conjunto las nuevas posturas y las modificaciones el promedio del volumen es de 10,518.09. Esto se debe a que el promedio del volumen de las modificaciones de compra es de 27,237.23. La segunda alternativa no es viable ya que hay un gran número de modificaciones, al ignorarlas se descartaría información importante en el modelo. La tercera alternativa le agrega un poco de complejidad al modelo, ya que es necesario modelar de alguna forma la llegada de las modificaciones y el comportamiento de ellas. En este caso se optará por la primera alternativa.

Finalmente para calcular los parámetros de cancelación $\theta(i)$, es necesario calcular el libro promedio. En la Figura 16 se encuentra el libro promedio para BIMBO. Bajo los supuestos del modelo, el número de cancelaciones depende del número de órdenes activas. Si se define Q_i^A y Q_i^B como el número promedio de órdenes activas a i pujas de la mejor postura de venta y compra respectivamente, $\theta(i)$ para $1 \leq i \leq 5$ se pueden calcular de la siguiente forma, donde

Q_i es el promedio de Q_i^A y Q_i^B:

$$\hat{\theta}(i) = \frac{N_c(i)S_c}{Q_i T_*}$$

$N_c(i)$ es el número de posturas de cancelación a una distancia i de la mejor postura contraria. Para $i > 5$, $\hat{\theta}(i) = \hat{\theta}(5)$.

(a) BIMBO compra promedio **(b)** BIMBO venta promedio

Figura 16: Libro Promedio

Al realizar los cálculos antes descritos para la emisora BIMBO se obtienen los siguientes parámetros:

Tabla 8: Parámetros del Modelo

	1	2	3	4	5
$\lambda(i)$	0.1669	0.0732	0.0655	0.0601	0.0731
$\theta(i)$	0.6480	0.9954	1.3240	1.5714	0.7582
μ	0.1438				

6.5. Resultados

Para entender los resultados, en primer lugar, es necesario obtener las probabilidades empíricas para compararlas con las obtenidas por el modelo. En este caso, se eligió la acción de BIMBO, en particular los casos donde $S = 1$, es decir, la distancia de la mejor postura de compra a la mejor postura de venta es de 1 puja. Para el caso de BIMBO, existen 21,256 eventos donde $S = 1$ de los 173,616 eventos que sucedieron en los 68 días de estudio.

Tabla 9: Ejemplo: Posturas BIMBO

id	timestamp	$p_B(t)$	$X_{p_B}(t)$	$p_A(t)$	$X_{p_A}(t)$	S	$p_M(t)$
151	1289919878430	97.79	600	97.8	2100	1	97.7950
152	1289919878450	97.79	600	97.8	2100	1	97.7950
153	1289919882930	97.79	600	97.8	1500	1	97.7950
154	1289919882940	97.79	600	97.8	1500	1	97.7950
155	1289919883070	97.75	39000	97.79	900	4	97.7700

Para analizar los resultados, se considerará la primera postura donde $S = 1$, en el ejemplo de la Tabla 9 esto sucede en el renglón 151, donde $p_M(t)$ es 97.7950 y el volumen de la mejor postura de venta excede el de la mejor postura de compra. Los siguientes eventos modifican posturas distintas a las puntas o el volumen en las puntas, hasta el evento 155 donde se modifica $p_M(t)$.

Para la acción de BIMBO existen 2,196 instancias como esta; para poder realizar el análisis se encontró la media del volumen del primer evento de cada instancia donde $S = 1$, la cual es de 2,707.45. El volumen de todos los eventos se dividió entre la media y se tomó el número entero más cercano mayor a este cociente para realizar los cálculos. Para el análisis se consideran las posturas donde el volumen normalizado es menor o igual que 5. De las 2,196 instancias mencionadas antes 2,043 cumplen esto; éstas se muestran en la Tabla 10.

Tabla 10: Número de Posturas

	venta				
compra	1	2	3	4	5
1	1,217	253	116	59	5
2	193	23	9	4	1
3	62	3	4	1	1
4	39	4	0	0	0
5	15	1	0	0	0

En la Tabla 11 se encuentran los resultados que se obtuvieron. En los casos donde existen el mismo número de posturas de compra que de venta los resultados no son muy claros, ya que sólo en el 39.93 % y 34.78 % para 1 y 2 posturas respectivamente aumenta el precio; se esperaría un número cercano al 50 %. Sin embargo, los resultados sí reflejan algo que se esperaría, en los casos donde el número de posturas de compra es mayor que el número de posturas de venta la probabilidad de un aumento de precio es mayor; al aumentar esta

diferencia aumenta también la probabilidad de que el precio suba. Muchas de las combinaciones son poco concluyentes, ya que la muestra es muy pequeña.

Tabla 11: Resultados Empíricos

	venta				
compra	1	2	3	4	5
1	39.93 %	11.86 %	9.48 %	11.86 %	2.63 %
2	60.62 %	34.78 %	11.11 %	50.00 %	0.00 %
3	67.74 %	33.33 %	75.00 %	0.00 %	100.00 %
4	76.92 %	25.00 %	-	-	-
5	93.33 %	0.00 %	-	-	-

Los resultados obtenidos al aplicar el método de la transformación de Laplace siguen una tendencia similar como se puede ver en la Tabla 12. En este caso cuando el número de órdenes de compra es el mismo al de las de venta, la probabilidad de un aumento en $p_M(t)$ es cercana al 50 %. Cuando el número de posturas de compra excede al de venta existe una mayor probabilidad de que el precio suba para todos los casos; también al aumentar la diferencia entre las posturas de compra y venta aumenta la probabilidad de un aumento en el precio. Un resultado interesante es que se esperaría que $P\{p_M(T) > p_M(0)|X_A(0) = a, X_B(0) = b, p_S(0) = 1\} = 1 - P\{p_M(T) > p_M(0)|X_A(0) = b, X_B(0) = a, p_S(0) = 1\}$; sin embargo, en este caso $P\{p_M(T) > p_M(0)|X_A(0) = 1, X_B(0) = 2, p_S(0) = 1\} = 69.31\,\%$ mientras que $P\{p_M(T) > p_M(0)|X_A(0) = 2, X_B(0) = 1, p_S(0) = 1\} = 35.32\,\%$. Esto se puede deber a la precisión de algunos de los cálculos que se realizaron, en particular las integraciones numéricas; esto se podría mejorar al alterar los parámetros.

Tabla 12: Resultados Transformación de Laplace

	venta				
compra	1	2	3	4	5
1	53.49 %	35.32 %	26.21 %	20.82 %	17.27 %
2	69.31 %	52.28 %	41.79 %	34.75 %	29.73 %
3	77.16 %	62.26 %	52.02 %	44.60 %	39.01 %
4	81.81 %	68.80 %	59.22 %	51.91 %	46.17 %
5	84.89 %	73.42 %	64.55 %	57.52 %	51.85 %

Con los mismos parámetros que se utilizaron para el método de Laplace se realizaron simulaciones. Para cada una de las combinaciones se realizaron

10,000 simulaciones en el paquete estadístico R. Para las simulaciones en los casos donde el número de posturas de compra y venta son iguales, el porcentaje de las ocasiones donde el precio sube es el más cercano al 50 %. También en este método se cumplen las premisas que se establecieron en los otros dos métodos en cuanto al aumento de la probabilidad de que el precio suba dada la diferencia entre el número de posturas de compra y venta. En este caso se podría decir que $P\{p_M(T) > p_M(0)|X_A(0) = a, X_B(0) = b, p_S(0) = 1\} = 1 - P\{p_M(T) > p_M(0)|X_A(0) = b, X_B(0) = a, p_S(0) = 1\}$ sí se cumple. Este método resultó ser bastante más rápido que el método de la transformación inversa de Laplace, sin embargo, probablemente el método de Laplace se podría mejorar considerablemente si se optimizaran algunos de los algoritmos utilizados.

Tabla 13: Resultados Simulación

	venta				
compra	1	2	3	4	5
1	49.85 %	33.23 %	25.10 %	19.71 %	16.29 %
2	66.78 %	50.26 %	40.35 %	33.20 %	28.28 %
3	75.55 %	60.08 %	50.41 %	42.41 %	36.89 %
4	79.55 %	66.27 %	57.13 %	49.40 %	43.76 %
5	83.90 %	71.98 %	63.65 %	55.75 %	50.59 %

7. Conclusiones

En el trabajo se reconstruyó el libro de posturas de la Bolsa Mexicana de Valores, el algoritmo logró reproducir con gran eficacia los hechos reportados por la BMV, medidos desde el punto de vista del volumen y del VWAP.

Se compararon los resultados obtenidos con los análisis que se han realizado para otros mercados. Para algunas de las variables los resultados fueron similares a los encontrados en otros mercados; por ejemplo los precios siguen una ley de potencias. También se puede decir que el volumen de las órdenes sigue la ley de potencias, sin embargo, las cancelaciones no tienen un comportamiento similar. La mayoría de las cancelaciones suceden cerca de las puntas.

En el trabajo se describió un modelo del libro de posturas simplificado, el cual clasifica las órdenes en: órdenes limitadas, de mercado y cancelaciones, a las cuales les asigna una distinta tasa de llegada. El modelo tiene la bondad de permitir el cálculo de algunos resultados a través de la transformación inversa de Laplace, la cual facilita los mismos debido a sus propiedades.

El trabajo se concentró en describir el comportamiento de los movimientos del precio. Para realizar los cálculos necesarios se utilizó el método de Laplace y una expresión de la probabilidad buscada en términos de fracciones continuas. Para realizar los cálculos fue necesario estimar los parámetros. Para poder comparar los resultados obtenidos a través del método de Laplace, se realizaron los cálculos de las frecuencias empíricas y a través de simulación.

En particular se calculó la probabilidad de un aumento en el precio cuando el "spread" es de una puja; las frecuencias empíricas muestran algunas diferencias con los resultados obtenidos a través de la transformación de Laplace y simulación, sin embargo los resultados eran consistentes con los que se esperaban; al aumentar la diferencia entre las posturas de compra y venta, aumenta la probabilidad de que aumente el precio. El método de la transformación de Laplace resultó ser significativamente más lento que las simulaciones.

En trabajos futuros se podría probar el método de la transformación inversa en casos donde el "spread" es mayor que 1, para estos casos podría ser que este método fuera más veloz que la simulación. También se deja para un trabajo posterior la implementación de estrategias para explotar los resultados encontrados en el trabajo. Para implementar estas estrategias es necesario optimizar los algoritmos utilizados en este trabajo. El algoritmo de la trans-

formación de Laplace utilizada en el trabajo, utiliza los mismos resultados en varias ocasiones, por lo que estos se podrían almacenar para evitar calcularlos en cada ocasión.

El modelo utilizado también puede ser perfeccionado; en particular se podría eliminar el supuesto de órdenes de un tamaño unitario o estándar. También se podrían utilizar distintos parámetros para las órdenes de compra y venta.

Índice de tablas

1. Posturas Iniciales 6
2. Registro de Posturas 6
3. Distribución de Precios de las Posturas de Compra 14
4. Distribución de Precios de las Posturas de Venta 14
5. Distribución del Volumen de las Posturas 17
6. Distribución de las Cancelaciones de las Posturas 20
7. Transformación de Euler 31
8. Parámetros del Modelo 40
9. Ejemplo: Posturas BIMBO 41
10. Número de Posturas 41
11. Resultados Empíricos 42
12. Resultados Transformación de Laplace 42
13. Resultados Simulación 43

Índice de figuras

1. Captura de pantalla de la Interfaz desarrollada en Java 9
2. Reconstrucción: KUO diciembre 31 11
3. Reconstrucción: KUO enero 28 11
4. Reconstrucción: AMX enero 20 12
5. Reconstrucción: AMX diciembre 28 12
6. Ajuste del Modelo de Distribución de Precios de las Posturas . 15
7. QQ-Plot Distribución de Precios de las Posturas 16
8. Ajuste del Modelo de Distribución del Volumen de las Posturas 18
9. QQ-Plot Distribución del Volumen de las Posturas 19
10. Ajuste del Modelo de Distribución de las Cancelaciones de las Posturas 21
11. QQ-Plot Distribución de las Cancelaciones de las Posturas . . . 22
12. Funciones $g_n(t)$ 29
13. Función de densidad de τ 37
14. Función de densidad de τ 37
15. Función de densidad de τ 38
16. Libro Promedio 40

Bibliografía

[1] *Inside the NYSE: The Specialist*, `http://www.nyse.com/pdfs/specialistmagarticle.pdf`.

[2] Maureen O'Hara, *Market microstructure theory*, Blackwell Business, John Wiley & Sons, 1995.

[3] Rasika Maduranga Withanawasam, *Modelling trade-based manipulation strategies in limit-order markets*, Ph.D. thesis, University of Otago, 2013.

[4] Jean-Philippe Bouchaud, Marc Mézard, and Marc Potters, *Statistical properties of stock order books: empirical results and models*, Quantitative Finance **2** (2002), no. 4, 251–256.

[5] I. Zovko and J.D. Farmer, *The power of patience: a behavioural regularity in limit-order placement*, Quantitative Finance **2** (2002), no. 5, 387–392.

[6] Jean-Philippe Bouchaud, J. Doyne Farmer, and Fabrizio Lillo, *How markets slowly digest changes in supply and demand*, Handbook of Financial Markets: Dynamics and Evolution (Thorsten Hens and Klaus Reiner Schenk-Hoppé, eds.), North-Holland, San Diego, 2009, pp. 57 – 160.

[7] J. Abate and W. Whitt, *Computing laplace transforms for numerical inversion via continued fractions*, INFORMS Journal on Computing **11** (1999), no. 4, 394–405.

[8] H. Dubner and J. Abate, *Numerical inversion of laplace transforms by relating them to the finite fourier cosine transform*, Journal of the ACM (JACM) **15** (1968), no. 1, 115–123.

[9] J. Abate and W. Whitt, *Numerical inversion of laplace transforms of probability distributions*, ORSA Journal on Computing **7** (1995), no. 1, 36–43.

[10] Rama Cont, Sasha Stoikov, and Rishi Talreja, *A stochastic model for order book dynamics*, Operations Research **58** (May/June 2010), no. 3, 549–563.