



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

EL CAR DE LA DEUDA PÚBLICA DE MÉXICO:
MODELOS DE PREDICCIÓN DE TASAS DE
INTERÉS Y TIPO DE CAMBIO PARA SU
ESTIMACIÓN

TESIS

QUE PARA OBTENER EL TÍTULO DE:
ACTUARIO

PRESENTA:

RODRIGO GARCÍA CASTILLO

DIRECTOR DEL TRABAJO:
DR. RICARDO RAMÍREZ ALDANA

CIUDAD DE MÉXICO 2019





Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Hoja de Datos del Jurado

1. Datos del alumno

García
Castillo
Rodrigo
55 71 56 04 68
Universidad Nacional Autónoma de México
Facultad de Ciencias
Actuaría
311046436

2. Datos del tutor

Dr.
Ricardo
Ramírez
Aldana

3. Datos del sinodal 1

Dra.
Ruth Selene
Fuentes
García

4. Datos del sinodal 2

Dra.
Ana
Meda
Guardiola

5. Datos del sinodal 3

Act.
Edgar
Díaz
Ordoñez

6. Datos del sinodal 4

Act.
Fernando
Díaz
López

7. Datos de la tesis.

El CaR de la Deuda Pública de México: Modelos de predicción de tasas de interés y tipo de cambio para su estimación
133 p.
2019

Agradecimientos

Este texto esta dedicado a mis padres, Beatriz y Manuel, a mi hermano, Memo, quienes me han apoyado durante toda la vida para salir adelante y ser una persona de bien.

A la Universidad Nacional Autónoma de México que no solo me ha brindado una educación de excelencia y una formación humana, sino que me ha hecho conocer a un gran número de personas increíbles.

A la Unidad de Crédito Público, de la SHCP, que durante mi estancia me brindaron su amistad, además de inspirar la realización de este trabajo y mostrar su apoyo con para que este trabajo fuera posible.

A mis sinodales, la Dra. Ana Meda Guardiola, Dra. Ruth Selene Fuentes García, Act. Edgar Díaz Ordoñez y Act. Fernando Díaz López quienes me apoyaron con sus valiosos comentarios y observaciones que me permitieron concluir este proyecto.

Finalmente agradezco infinitamente al Dr. Ricardo Ramírez Aldana quien mostró una gran disposición apoyándome aportando su conocimiento y experiencia dirigiendo este trabajo. Sin su ayuda no hubiera sido posible.

Índice general

Índice general	3
Índice de tablas	6
Índice de figuras	7
1. La Deuda Pública	14
2. El Costo en Riesgo	19
2.1. Elementos necesarios para la estimación	20
2.2. Estructura de la Deuda Pública del Gobierno Federal	21
2.3. Características de la Deuda Pública futura de México	23
3. Conceptos básicos y supuestos en las tasas de interés	25
3.1. Condición de no arbitraje	29
3.2. Modelos de tasas cortas de un factor	32
4. Modelo de Cox-Ingersoll-Ross	35
4.1. Dinámica en la medida \mathbb{P} y estimación histórica	38
4.2. Estimación por mínimos cuadrados condicionales	40
5. Series de Tiempo	43
5.1. Conceptos básicos de las series de tiempo	43
5.1.1. Transformaciones de las series	45

<i>ÍNDICE GENERAL</i>	5
5.1.2. Ruido blanco y series de tiempo univariadas	46
5.2. Representaciones de series de tiempo	47
5.2.1. Procesos causales	47
5.2.2. Procesos invertibles	49
5.2.3. Modelos MA(q)	50
5.2.4. Modelos AR(p)	52
5.2.5. Modelos ARMA(p,q)	55
6. Identificación del modelo y pronósticos	56
6.0.1. Función de autocorrelaciones parciales	57
6.0.2. Ajuste del modelo	58
6.1. Técnicas de pronósticos	60
6.1.1. Algoritmos recursivos	63
6.1.2. Pronósticos en Modelos ARMA	68
6.2. Estimación de parámetros	70
6.2.1. Validación del modelo	72
6.2.2. Modelos ARIMA y SARIMA	74
7. Estimación del CaR	77
7.1. CaR de Tipo de Cambio	78
7.2. Ajuste ARIMA a los tipos de cambio	81
7.2.1. Código en R para ajuste ARIMA a la serie de tiempo USD/MXN	89
7.2.2. Código en Matlab para CaR de tipo de cambio	93
7.3. CaR de Tasas de interés	95
7.3.1. Código para CaR de tasas de interés	100
7.3.2. Código para hallar parámetros en modelo CIR	101
8. Conclusiones	104

9. Apéndice	107
--------------------	------------

10. Bibliografía	130
-------------------------	------------

Índice de cuadros

3.1. Modelos clásicos de tasas instantáneas	34
---	----

Índice de figuras

1.1.	Informes sobre la Situación Económica, las Finanzas Públicas y la Deuda Pública 2017 . . .	16
1.2.	Informes sobre la Situación Económica, las Finanzas Públicas y la Deuda Pública 2017 . . .	16
1.3.	Porcentaje de deuda respecto al PIB en 2014. Fuente: Casar J. (2017). Deuda Pública y Política Fiscal en México. Revista de Economía Mexicana. Núm. 2.	17
1.4.	Indicadores de endeudamiento en México. Fuente: Casar J. (2017). Deuda Pública y Política Fiscal en México. Revista de Economía Mexicana. Núm. 2.	17
2.1.	Fuente: Plan Anual de Financiamiento 2018	22
2.2.	Fuente: Plan Anual de Financiamiento 2018	23
3.1.	Ejemplo de tasa cupón cero. Fuente: Brigo D., Mercurio F. (2006). <i>Interest Rate Models-Theory and Practice</i> . Springer, 2a edición.	28
4.1.	Simulación de la dinámica de una tasa de acuerdo al modelo de Cox-Ingersoll-Ross.	36
5.1.	Simulación de tamaño 300 de un ruido blanco.	47
6.1.	ACF de la simulación de dos procesos ARMA(1,1)	66
7.1.	Fragmento de la matriz de tipo de cambio. De izquierda a derecha: Dólar Americano, Euro, Yen, Dólar Canadiense, Libra	79
7.2.	Intereses por pagar en unidades de moneda extranjera. De izquierda a derecha: USD, EUR, JPY, CAD, GBP	80
7.3.	Tipos de cambio vistos como series de tiempo	82
7.4.	Tipo de cambio USD/MXN	82
7.5.	Regresión lineal al tipo de cambio USD/MXN	83

<i>ÍNDICE DE FIGURAS</i>	9
7.6. Resumen de regresión lineal al tipo de cambio USD/MXN	83
7.7. Regresión cuadrática al tipo de cambio USD/MXN	83
7.8. Resumen de regresión cuadrática al tipo de cambio USD/MXN	84
7.9. De arriba a abajo: Serie diferenciada una vez, serie diferenciada dos veces.	84
7.10. De arriba a abajo: Serie diferenciada en la parte estacional, Serie diferenciada en la parte estacional y luego una vez más.	85
7.11. Prueba de Dickey-Fuller. De arriba a abajo: Serie diferenciada en la parte estacional, Serie diferenciada en la parte estacional y luego una vez más.	85
7.12. Sucesión de autocorrelaciones muestrales de la serie DiffIag2USD.	85
7.13. Sucesión de autocorrelaciones parciales muestrales de la serie DiffIag2USD.	86
7.14. Ejemplo de distintos ajustes ARIMA la serie DiffIag2USD.	86
7.15. Residuales del ajuste ARIMA(2,2,1) a la serie DiffIag2USD	87
7.16. ACF y el estadístico de Ljung-Box para los distintos lags de los residuales	87
7.17. Prueba Kolmogorov-Smirnov para una distribución normal. La primera con la media muestral y la segunda con media cero.	88
7.18. Pronostico de un ajuste ARIMA(2,2,1) a la serie de tipos de cambio USD/MXN	88
7.19. Pronostico de un ajuste SARIMA a la serie de tipos de cambio USD/MXN	89
7.20. Matriz ARIMA	93
7.21. Histograma de la variable aleatoria que mide el costo financiero generado por la deuda pública en moneda extranjera.	95
7.22. Resultados del código para el CaR en moneda extranjera.	95
7.23. Matrices de colocación de Bonos. De arriba a abajo: Bonde D, Bonos a Tasa Fija, CETEs, Udibonos	97
7.24. Calendario de cupones de Bonos a tasa fija para un periodo de 13 meses. De izquierda a derecha: 3 años, 5 años, 10 años, 20 años y 30 años.	98
7.25. Pronóstico de modelo ARIMA(2,2,2) a la serie de tiempo generada por los valores de la udi.	99
7.26. Histograma de la variable aleatoria que mide el costo financiero generado por la deuda pública en moneda local.	103
7.27. Resultados del código para el CaR en moneda local.	103

9.1. Serie de tiempo generada por el tipo de cambio EUR/MXN.	112
9.2. Serie diferenciada dos veces del tipo de cambio EUR/MXN.	112
9.3. Prueba de Dickey-Fuller para determinar la existencia de raíces unitarias a la serie diferenciada del tipo de cambio EUR/MXN.	112
9.4. ACF y PACF de la serie de tiempo diferenciada del tipo de cambio EUR/MXN para determinar modelo.	113
9.5. Ajuste de modelo ARIMA a la serie de tiempo del tipo de cambio EUR/MXN.	113
9.6. ACF y estadísticos de Ljung-Box para los distintos lags de los residuales del modelo ajustado.	113
9.7. Serie generada de los residuales e histograma con densidad de una distribución normal ajustada.	114
9.8. Distribución empírica generada por los residuales contra distribución teórica normal ajustada.	114
9.9. Pruebas de Kolmogorov-Smirnov, Lilliefors y Anderson-Darling para verificar normalidad en residuales.	114
9.10. Serie de tiempo generada por el tipo de cambio JPY/MXN.	115
9.11. Serie diferenciada dos veces del tipo de cambio JPY/MXN.	115
9.12. Prueba de Dickey-Fuller para determinar la existencia de raíces unitarias a la serie diferenciada del tipo de cambio JPY/MXN.	115
9.13. ACF y PACF de la serie de tiempo diferenciada del tipo de cambio JPY/MXN para determinar modelo.	116
9.14. Ajuste de modelo ARIMA a la serie de tiempo del tipo de cambio JPY/MXN.	116
9.15. ACF y estadísticos de Ljung-Box para los distintos lags de los residuales del modelo ajustado.	117
9.16. Serie generada de los residuales e histograma con densidad de una distribución normal ajustada.	117
9.17. Distribución empírica generada por los residuales contra distribución teórica normal ajustada.	118
9.18. Pruebas de Kolmogorov-Smirnov, Lilliefors y Anderson-Darling para verificar normalidad en residuales.	118
9.19. Serie de tiempo generada por el tipo de cambio CAD/MXN.	118
9.20. Serie diferenciada dos veces del tipo de cambio CAD/MXN.	119
9.21. Prueba de Dickey-Fuller para determinar la existencia de raíces unitarias a la serie diferenciada del tipo de cambio CAD/MXN.	119
9.22. ACF y PACF de la serie de tiempo diferenciada del tipo de cambio CAD/MXN para determinar modelo.	119

9.23. Ajuste de modelo ARIMA a la serie de tiempo del tipo de cambio CAD/MXN. 120

9.24. ACF y estadísticos de Ljung-Box para los distintos lags de los residuales del modelo ajustado. 120

9.25. Serie generada de los residuales e histograma con densidad de una distribución normal ajustada. 121

9.26. Distribución empírica generada por los residuales contra distribución teórica normal ajustada. 121

9.27. Pruebas de Kolmogorov-Smirnov, Lilliefors y Anderson-Darling para verificar normalidad en residuales. 121

9.28. Serie de tiempo generada por el tipo de cambio GBP/MXN. 122

9.29. Serie diferenciada dos veces del tipo de cambio GBP/MXN. 122

9.30. Prueba de Dickey-Fuller para determinar la existencia de raíces unitarias a la serie diferenciada del tipo de cambio GBP/MXN. 122

9.31. ACF y PACF de la serie de tiempo diferenciada del tipo de cambio GBP/MXN para determinar modelo. 123

9.32. Ajuste de modelo ARIMA a la serie de tiempo del tipo de cambio GBP/MXN. 123

9.33. ACF y estadísticos de Ljung-Box para los distintos lags de los residuales del modelo ajustado. 124

9.34. Serie generada de los residuales e histograma con densidad de una distribución normal ajustada. 124

9.35. Distribución empírica generada por los residuales contra distribución teórica normal ajustada. 125

9.36. Pruebas de Kolmogorov-Smirnov, Lilliefors y Anderson-Darling para verificar normalidad en residuales. 125

9.37. Serie de tiempo generada por el precio de la UDI. 125

9.38. Serie diferenciada dos veces del precio de la UDI. 126

9.39. Prueba de Dickey-Fuller para determinar la existencia de raíces unitarias a la serie diferenciada del precio de la UDI. 126

9.40. ACF y PACF de la serie de tiempo diferenciada del precio de la UDI para determinar modelo. 126

9.41. Ajuste de modelo ARIMA a la serie de tiempo del precio de la UDI. 127

9.42. ACF y estadísticos de Ljung-Box para los distintos lags de los residuales del modelo ajustado. 127

9.43. Serie generada de los residuales e histograma con densidad de una distribución normal ajustada. 127

9.44. Distribución empírica generada por los residuales contra distribución teórica normal ajustada. 128

9.45. Pruebas de Kolmogorov-Smirnov, Lilliefors y Anderson-Darling para verificar normalidad en residuales. 128

9.46. Resumen de los modelos ajustados a las series de tiempo.	129
--	-----

Introducción

Los objetivos del trabajo son mostrar la teoría que respalda el ajuste y cálculo de la estimación del Costo en Riesgo, CaR, por sus siglas en inglés *Cost at Risk* y así mismo difundir a la sociedad el concepto de costo en riesgo, como una medida del costo financiero que enfrenta una entidad (en el caso de este trabajo será sobre la Deuda Pública de México) y que sirve para la toma de decisiones que permitan mitigar este riesgo. Se plantea realizar la estimación del Costo en Riesgo de la Deuda Pública del México para el corto plazo (1 año).

El enfoque que tendrá el trabajo es teórico-práctico ya que primero se presentarán los modelos teóricos que sustentan la estimación del CaR y luego se enfocará en la parte de la estimación detallando los inputs que los códigos requieren. Se presentará la metodología que implementa la Secretaría de Hacienda y Crédito Público, SHCP, para la estimación de CaR, se compone de dos modelos individuales paramétricos, el primero para la estimación de tasas de interés a corto plazo (*Modelo de Cox-Ingersoll-Ross*), mientras que el segundo servirá para la estimación de los tipos de cambio en los que emite deuda el país (*Modelos ARIMA*). Existe procedimiento no paramétrico para la estimación, sin embargo este trabajo no se enfocará en ella debido a que el proceso se realiza de manera análoga¹.

Cabe señalar que los códigos para la estimación del costo están basados en los desarrollados por el Banco de Canadá, abiertos al público para su consulta.²

Finalmente se hará una comparación de los modelos y las estimaciones resultantes a partir de éstos.

¹Con excepción de las simulaciones que se enfoca en distribuir los choques históricos de manera uniforme para los pronósticos de los tipos de cambio y tasa de interés.

²La metodología se encuentran en la siguiente ruta <https://www.bankofcanada.ca/markets/government-securities-auctions/>

Capítulo 1

La Deuda Pública

Se comenzará este capítulo dando una breve explicación acerca de la deuda pública, cómo esta compuesta en el caso de México, y algunas cifras de manera general.

¿Qué es la deuda pública? Una manera de definirla es que son todos los pasivos a cargo del sector público de una nación. Una definición más precisa de deuda pública que brinda el Banco Mundial dice: "*Son todas las obligaciones insolutas del sector público, contraídas en forma directa o a través de sus agentes financieros*". [31].

Por lo tanto, se puede decir que la deuda constituye una herramienta de financiación de los gobiernos que le permite hacer frente a sus gastos y poder así cumplir con las obligaciones que le competen. Es conveniente señalar que el nivel de deuda también puede ser utilizado para ejercer política fiscal (una mayor recaudación fiscal puede implicar que el gobierno requiera disminuir su nivel de deuda, así como una política fiscal más laxa puede implicar aumentar el nivel de endeudamiento soberano), así como para financiar inversiones necesarias para mejorar la provisión de bienes públicos, y con ello fomentar el desarrollo económico. Es así que si el nivel de deuda logra estimular el crecimiento económico al corto plazo aunado con una política fiscal adecuada, entonces puede ser positiva para una nación. Sin embargo, un alto nivel de deuda puede traer consecuencias adversas para la economía porque aumenta el riesgo potencial de una crisis soberana, lo que conlleva a un aumento desmedido en impuestos o un recorte en el gasto público que puede repercutir en el crecimiento económico de una nación. Es por ello que la cuestión más importante es saber cuál es el nivel óptimo de deuda que una nación requiere.

La deuda pública se puede dividir en deuda bruta y deuda neta. La deuda neta es la diferencia entre deuda bruta y los activos financieros del país.

La deuda pública de México puede clasificarse de diversas formas[27] :

1. *Por su origen:*

- a) *Deuda interna* son los financiamientos obtenidos en el mercado doméstico, mediante la colocación de valores gubernamentales y de créditos directos con otras instituciones, cuyas características principales son el ser pagaderos dentro del país y en pesos mexicanos.
- b) *Deuda externa* se considera a los créditos contratados por el sector público con

entidades financieras del exterior y pagaderos en el extranjero en moneda diferente a la moneda nacional.

2. *Por su plazo:*

- a) *Corto plazo* son préstamos que se obtienen a plazo menor de un año.
- b) *Largo plazo* son préstamos que se contratan a un plazo de un año o más.

3. *Por fuente de financiamiento.* Se clasifica según la naturaleza de los acreedores financieros. Para la deuda externa: mercado de capitales, organismos financieros internacionales, mercado bancario, comercio exterior, deuda reestructurada y otros pasivos. Para la deuda interna: valores gubernamentales, banca comercial, Sistema de Ahorro para el Retiro (SAR), Sociedades de Inversión Especializadas en Fondos para el Retiro (Siefores), entre los más importantes.

4. *Por instrumento.* Según las características jurídicas de los instrumentos que describen la relación que existe entre el acreedor y el deudor. Los instrumentos que actualmente están vigentes son los pagarés, bonos, valores gubernamentales y los contratos o líneas de crédito.

Existen 3 indicadores de la deuda pública de México: la deuda del Gobierno Federal, la deuda del Sector Público y el Saldo Histórico de los Requerimientos Financieros del Sector Público (*SHRFSP*). A continuación se describe cada uno de estos indicadores.

La deuda del Gobierno Federal son, como su nombre los indica, todas las obligaciones contratadas por las dependencias del Poder Ejecutivo Federal, los poderes legislativo y judicial y sus órganos desconcentrados y comprende los siguientes conceptos: Préstamos bancarios, emisiones de valores gubernamentales, deuda con organismos financieros internacionales, bonos del ISSSTE y cuentas relacionadas a la seguridad social, bonos de pensión de PEMEX y CFE, entre otros.

La deuda del Sector Público está integrada por la deuda del Gobierno Federal más las obligaciones de las entidades de control presupuestario directo, de las Empresas Productivas del Estado (PEMEX y CFE) y sus subsidiarias, y de la banca de desarrollo (Banco Nacional de Obras y Servicios Públicos (BANOBRAS), Sociedad Hipotecaria Federal (SHF), Banco Nacional de Comercio Exterior (BANCOMEXT), Nacional Financiera (NAFIN), Financiera Nacional de Desarrollo Agropecuario, Rural, Forestal y Pesquero (FND), Banco del Ahorro Nacional y Servicios Financieros (BANSEFI)).

Por último, el Saldo Histórico de los Requerimientos Financieros del Sector Público, (*SHRFSP*), es la medida más amplia de la deuda pública, al incluir todos los instrumentos de política pública que pudieran implicar endeudamiento a cargo del Sector Público. Entonces el *SHRFSP* está integrado por la deuda del Sector Público Presupuestario más las obligaciones que, por diversos ordenamientos jurídicos y prácticas presupuestarias, no se incluyen en la deuda pública. Por lo tanto, es así que el *SHRFSP* se compone principalmente por la deuda del Sector Público Presupuestario más los pasivos del Instituto para la Protección al Ahorro Bancario (IPAB), obligaciones derivadas del Programa de Apoyo a Deudores de la Banca, Fondo Nacional de Infraestructura (FONADIN), Banca de Desarrollo, Fondos de Fomento, entre otros.

De acuerdo a los Informes sobre la Situación Económica, las Finanzas Públicas y la Deuda Pública en el caso de México, la deuda neta del Gobierno Federal en millones de

pesos al cierre del cuarto trimestre del 2017 fue de aproximadamente 7.507 billones de pesos (millones de millones de pesos), mientras que el saldo de la deuda interna neta del Gobierno Federal ascendió a 5 billones 714 mil 288 millones de pesos que representa el 76,1 % del portafolio de la deuda. La imagen 1.1 muestra los saldos de la deuda del Gobierno Federal, así como el monto de la deuda del Sector Público y el SHRFSP que a fines del 2017 ascienden a más de 10.088 y 10.031 billones de pesos respectivamente.

SALDOS DE LA DEUDA NETA (Millones de pesos)						
Concepto	2016 diciembre			2017 diciembre		
	Total	Interna	Externa	Total	Interna	Externa
Gobierno Federal	7,193,008.9	5,396,301.4	1,796,707.5	7,507,461.6	5,714,288.0	1,793,173.6
Sin Bonos de Pensión	6,894,289.0	5,097,581.5	1,796,707.5	7,210,254.0	5,417,080.4	1,793,173.6
Bonos de Pensión Pemex ^{1/}	137,639.7	137,639.7	0.0	136,127.4	136,127.4	0.0
Bonos de Pensión CFE ^{2/}	161,080.2	161,080.2	0.0	161,080.2	161,080.2	0.0
Sector Público	9,693,217.5	6,009,403.1	3,683,814.4	10,088,979.7	6,283,058.4	3,805,921.3
SHRFSP	9,797,439.6	6,217,513.5	3,579,926.1	10,031,652.7	6,326,044.5	3,705,608.2

Figura 1.1: Informes sobre la Situación Económica, las Finanzas Públicas y la Deuda Pública 2017

Como se puede ver en la imagen 1.2 la participación de la deuda interna en la composición de la cartera de deuda pública ha venido aumentando con respecto a la deuda en moneda extranjera pasando de un 55,3 % en el año 2000 al 76,1 % en el año 2017, mostrando claramente la respuesta ante el riesgo que conlleva la volatilidad de los tipos de cambio emitiendo en mayor proporción instrumentos denominados en pesos mexicanos.

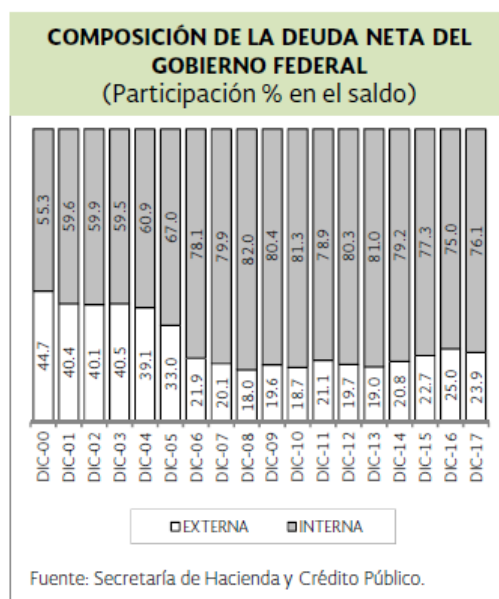


Figura 1.2: Informes sobre la Situación Económica, las Finanzas Públicas y la Deuda Pública 2017

Para el año 2014, la Organización para la Cooperación y el Desarrollo Económico, *OCDE*, realizó comparaciones internacionales de la deuda con respecto al PIB de los países miembros, ver figura 1.3, el promedio de la deuda del Gobierno Central en los países de la organización era de 88,1 % del PIB. La mediana de dicha deuda, a su vez, era de 79,5 %. Además 26 de 32 países miembros tenían en ese año un coeficiente de endeudamiento mayor al de México y en 12 de ellos la deuda superaba 100 % del PIB. El coeficiente alcanza un máximo de 246 % en Japón. El nivel de deuda de México a finales del 2016 corresponde a el 50 % del Producto Interno Bruto, que comparado con con otras naciones, desarrolladas o

emergentes, pertenecientes a la OCDE, no es un nivel relativamente grande.

Coefficiente de endeudamiento a PIB, países seleccionados, 2014

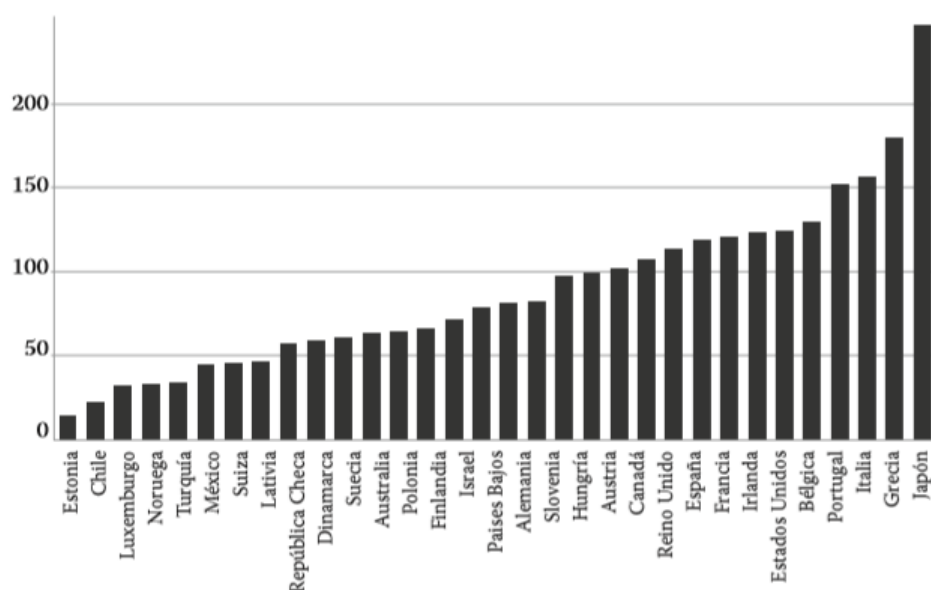


Figura 1.3: Porcentaje de deuda respecto al PIB en 2014. Fuente: Casar J. (2017). Deuda Pública y Política Fiscal en México. Revista de Economía Mexicana. Núm. 2.

Indicadores de endeudamiento. México, 1996 - 2016

	1996	2000	2006	2012	2013	2014	2015	2016(P)
Deuda Bruta Sector Público (% del PIB)	29.2	22.6	21.4	33.9	36.6	40.1	44.7	50.0
Deuda Bruta/Ingresos Tributarios (veces)	4.4	2.6	2.6	4.2	4.0	4.0	3.6	3.7
RFSP (% del PIB)	n/d	n/d	n/d	3.8	3.7	4.6	4.1	3.0
SHRFSP (% del PIB)	n/d	31.7	29.8	37.7	40.4	43.2	47.6	50.5
SHRFSP/Ingresos Tributarios (veces)	n/d	3.5	3.5	4.5	4.2	4.1	3.6	3.7
Costo Financiero (% PIB)	3.6	3.2	2.4	2.0	1.9	2.0	2.3	2.5

Figura 1.4: Indicadores de endeudamiento en México. Fuente: Casar J. (2017). Deuda Pública y Política Fiscal en México. Revista de Economía Mexicana. Núm. 2.

La figura 1.4, muestra la deuda bruta de México como porcentaje del PIB durante el periodo 1996-2016, además del SHRFSP durante el mismo periodo. Como se puede apreciar, estas dos medidas de la deuda han crecido durante los últimos 20 años, la deuda bruta creció 20.8 puntos porcentuales respecto al PIB en 20 años, mientras que el SHRFSP creció 18,8 %. Otro indicador que muestra la figura 1.4, es el coeficiente de la capacidad de pago que tiene el país de la deuda, este es el monto de la deuda (o en su caso el SHRFSP) dividido entre los ingresos tributarios (percepciones que obtiene el Gobierno Federal por las imposiciones fiscales). Este coeficiente muestra que a pesar de el aumento de la deuda pública en los últimos años, la capacidad de pago ha mejorado a partir del año 2012. En 2016, la capacidad de pago de la deuda pública por parte del Estado es mucho más robusta que la que tenía en 2012, a pesar de que el saldo de la deuda es considerablemente mayor.

En resumen la deuda es una herramienta financiera del gobierno para repartir sus cos-

tos en el tiempo. Si bien la deuda puede ser utilizada para ejercer la política fiscal, así como financiar inversiones necesarias para mejorar la provisión de bienes públicos, y con ello fomentar el desarrollo económico, si se emite sin medida puede dañar al crecimiento económico. Es posible realizar un análisis mucho más profundo acerca de esta, de cual es una mejor medida de la deuda, de la interpretación que se le debe dar, de sus causas y consecuencias, sin embargo dichas cuestiones deben de ser tratadas en un trabajo dedicado cien por ciento a su estudio, por lo que no se profundizará mas allá este tema.

Es importante señalar que para fines de este documento se trabajará solo con la Deuda del Gobierno Federal, especialmente con la emisión de valores gubernamentales y las obligaciones en moneda extranjera.

Capítulo 2

El Costo en Riesgo

La historia de los años más recientes (en particular la crisis del 2008) nos ha enseñado que tanto países emergentes como desarrollados deben contar con políticas fiscal, monetaria, cambiaria y de deuda soberana capaces de hacer frente a cualquier escenario incluso a los más adversos. Es por ello que las naciones deben contar con un portafolio de deuda que bajo supuestos escenarios a través del tiempo sea sostenible con el fin de mantener un costo financiero bajo y estable. Para poder mantener un control sostenible en la deuda pública, se debe disponer de diversos elementos que permitan evaluar y medir el valor de la deuda y su costo financiero.

En la actualidad, ante la apertura que han logrado los mercados financieros, es inminente que los cambios en las variables de mercado, como tasas de interés, tipo de cambio de moneda o spreads de crédito, representan un riesgo de pérdida en los valores de los instrumentos o portafolios financieros. Este tipo de riesgo se le conoce como *riesgo de mercado*. En el caso del mercado de deuda soberana el riesgo que conllevan los cambios en las variables de mercado son el aumento en el pago de intereses de la deuda pública (*costo financiero*).

La volatilidad de estas variables obliga a contar con herramientas que permitan la detección de estos riesgos y poder así mitigarlos, acotando el costo financiero. Un adecuado manejo de los riesgos que conlleva la deuda pública reduce el riesgo de que los cambios en las variables de mercado eleven el costo financiero de la deuda, así mismo, minimiza el *riesgo soberano* (probabilidad de incumplimiento de una entidad gubernamental con sus obligaciones) y la posibilidad de ajustes fiscales que serían necesarios para hacer sostenible la deuda.

Una de las herramientas que permitirá manejar y medir el riesgo asociadas a las variables de mercado antes mencionadas, es el Costo en Riesgo (CaR, por si siglas en inglés *Cost at Risk*), que es estimado por la Secretaría de Hacienda y Crédito Público (SHCP). El Costo en Riesgo puede definirse como el costo financiero máximo que podría enfrentar el país a un nivel de confianza predeterminado, $(1 - \alpha) \%$, dado su esquema de endeudamiento y las características de su portafolio (composición de plazos, monedas, tipos de tasa, etc.). Es así que si Λ es una variable aleatoria que indica el costo financiero, entonces el CaR a un nivel de confianza $(1 - \alpha) \%$, con $\alpha \in (0, 1)$, está dado por la siguiente definición:

$$\begin{aligned} CaR_{(1-\alpha)} &= \inf\{\lambda \in \mathbb{R} : \mathbb{P}[\Lambda > \lambda] < \alpha\} \\ &= \inf\{\lambda \in \mathbb{R} : \mathbb{P}[\Lambda \leq \lambda] \geq 1 - \alpha\} \end{aligned} \quad (2.1)$$

El objetivo básico de estos modelos es generar distribuciones de probabilidad del costo financiero de la deuda pública en el mediano y largo plazo con base en distintas estrategias de emisión. Estos modelos resultan similares a los modelos más comúnmente utilizados de Valor en Riesgo (VaR, por sus siglas en inglés *Value at Risk*) en el sector financiero. Es conveniente precisar que en los modelos VaR se estiman distribuciones de probabilidad de los valores de mercado de distintos instrumentos ante diferentes trayectorias en las variables financieras. Mientras tanto, para un administrador soberano de deuda cuya participación en los mercados no es oportunista, el foco de atención tiene que ver con la volatilidad de los flujos asociados con el servicio de la deuda pública y no con las fluctuaciones en el valor de la deuda, en tanto que los instrumentos emitidos generalmente llegan a su vencimiento.

Así, mientras que los modelos VaR se utilizan para estimar la pérdida potencial del portafolio a cierto nivel estadístico de confianza, el CaR se utiliza para estimar el **costo financiero** máximo de la deuda a determinado nivel estadístico de confianza.

De la misma manera se define el *CaR relativo* como la desviación porcentual entre el CaR y el valor esperado del costo financiero.

$$CaR_{relativo} = \frac{CaR_{(1-\alpha)}}{\mathbb{E}[\Lambda]} \quad (2.2)$$

Además se define el *CaR extremo* como la desviación porcentual entre un costo superior extremo y el costo medio esperado de la distribución completa, el costo superior es estimado como el valor medio del costo financiero dentro del α % de la cola derecha.

$$CaR_{extremo} = \frac{\mathbb{E}[\Lambda | \Lambda > CaR_{(1-\alpha)}]}{\mathbb{E}[\Lambda]} \quad (2.3)$$

La metodología CaR se desarrolló en distintas oficinas de administración de deuda en Europa, principalmente en Dinamarca y Suecia, ante la necesidad de tener mayores elementos para analizar el intercambio entre costos y riesgos de la deuda pública.

El costo en riesgo permite estimar el costo financiero de la deuda pública con base en escenarios sobre la evolución futura de las tasas de interés o el tipo de cambio ante diferentes estrategias de financiamiento. Para cada estrategia de financiamiento es posible obtener una distribución de probabilidad del costo financiero de la deuda pública originada por estos escenarios.

2.1. Elementos necesarios para la estimación

El costo en riesgo mide el riesgo al que se encuentra sujeto el costo financiero de la deuda pública para diversas estrategias de financiamiento dado un perfil de endeudamiento. El

análisis se basa en la valuación de un conjunto de escenarios utilizando técnicas de simulación estocástica, es decir, se simulan distribuciones de probabilidad del costo financiero para el corto plazo de la deuda pública del Gobierno Federal. La estimación del CaR se basa en los siguientes tres elementos:

1. Las características actuales del portafolio de la deuda
2. Las características de la deuda futura
3. Un modelo estadístico (o econométrico) para pronosticar las curvas de tasa de interés o tipo de cambio

Las características del portafolio actual de la deuda, hace referencia a la constitución de los pasivos del Gobierno Federal, algunas de estas pueden ser la denominación en la que se encuentra emitida la deuda (extranjera o local), el tipo de tasa que pagan que puede ser fija o variable, los distintos plazos emitidos, si se encuentran indexados a inflación o en valor nominal, si son instrumentos que pagan cupón o instrumentos *zero coupon*, etc.

Las características de la estrategia futura de endeudamiento implica conocer cómo el país está planeando financiar sus déficits fiscales. Sobre esto Díaz de León, Gutierrez y Vela (2016) mencionan: "Si bien el cálculo del CaR implica la suposición de una estrategia de endeudamiento también es una herramienta que permite elegir entre distintas estrategias de colocación de valores." [14]. Esto permite que se pueda plantear diferentes estrategias de colocación de deuda y valorar cada una de ellas para poder elegir aquella que se adecue a las necesidades del Gobierno Federal.

Puntualizando lo anterior, para el pronóstico de tipos de cambio se utilizarán modelos *ARIMA*, por sus siglas en inglés *Autoregressive Integrated Moving Average*, mientras que para los modelos de tasa de interés se usará el modelo de *Cox-Ingersoll-Ross*, *CIR*. El CaR denominado en moneda local está asociado a los movimientos de las tasas de interés, mientras que el costo en riesgo denominado en moneda extranjera se encuentra asociado a los tipos de cambio en el cual el Gobierno Federal ha emitido deuda.

La metodología para la estimación del modelo *CIR* se profundizará en capítulo 4, mientras que el ajuste y estimación de los modelos *ARIMAs* se verán en los capítulos 5 y 6.

2.2. Estructura de la Deuda Pública del Gobierno Federal

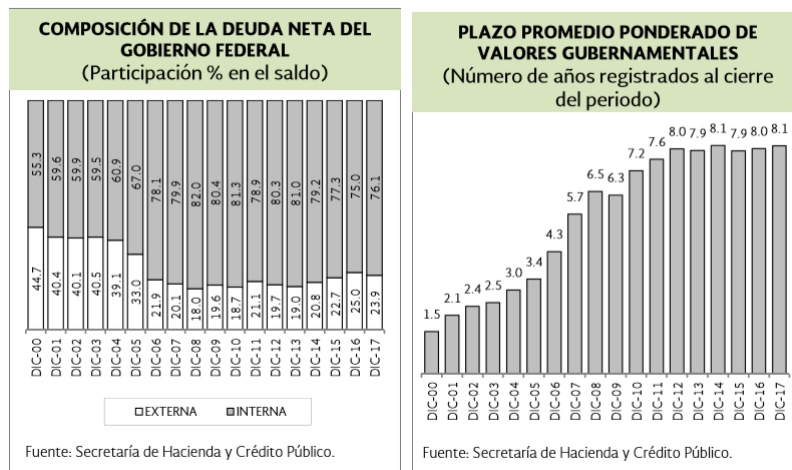
Además de las ya mencionadas características de los instrumentos que componen el portafolio de deuda del Gobierno Federal, es conveniente mencionar que este portafolio se encuentra bien diversificado con el objetivo de garantizar el acceso permanente a los mercados de deuda y es primordial conocer su composición para poder así llevar a cabo una estimación correcta del costo financiero. El portafolio de instrumentos del Gobierno Federal está principalmente compuesto por los siguientes instrumentos: Cetes, Bonos D, Bonos M, Udibonos, que se encuentran en moneda local, además de bonos colocados en el mercado internacional en moneda extranjera. La deuda interna representa, al cuarto trimestre del 2017, el 76,1 % de la deuda del Gobierno Federal, mientras que la deuda externa representa el 23,9 %.

Los Certificados de la Tesorería de la Federación mejor conocidos como Cetes son instrumentos cupón cero que vencen al corto plazo, es decir, tienen periodo de maduración menor a un año. Tienen una madurez de 28, 91, 182 y 364 días. Se venden a descuento y ofrecen una tasa fija. Los Cetes representan el 13,2 % del portafolio de la deuda interna. Los Cetes a 28, 91 y a 182 días se colocan de manera semanal, mientras que los Cetes a 364 días son colocados una vez cada cuatro semanas.

Los Bondes D son valores gubernamentales con tasa variable, plazo de vencimiento a 5 años y con valor nominal de 100 pesos. La tasa de rendimiento de estos instrumentos está indexada a la tasa de fondeo bancaria diaria, llamada también en México *Tasa de Interés Interbancaria de Equilibrio*, TIIE. Este instrumento paga cupones cada 28 días y tienen una colocación cada 2 semanas. Los Bondes D representan el 8,8 % del portafolio de deuda interna del Gobierno Federal.

Los Bonos M son instrumentos con una tasa de rendimiento fija y plazo de vencimientos de 3, 5, 10, 20 y 30 años. Estos valores tienen valor nominal de 100 pesos y pagan cupones cada 182 días. Los Bonos de tasa fija dominan el portafolio de deuda interna con una representación del 51,6 %. En cuanto a su colocación aquellos que tienen plazo de 3 y 5 años son colocados cada 4 semanas, mientras que los bonos a plazos de 10, 20 y 30 años son colocados una vez cada 6 semanas.

Los Udibonos son valores que se encuentran indexados a inflación, es decir, pagan una tasa real fija. Estos instrumentos pagan cupones cada 182 días y tienen un valor nominal de 100 udis (unidades de inversión). Estos instrumentos tienen plazo de vencimiento de 3, 10 y 30 años. Es de destacar que en el portafolio de la deuda interna aún siguen en vigencia instrumentos emitidos antes de 2008 con un plazo de vencimiento de 20 años. Los Udibonos se colocan una vez cada 4 semanas y representan el 26,4 % del portafolio de deuda interna del Gobierno Federal.

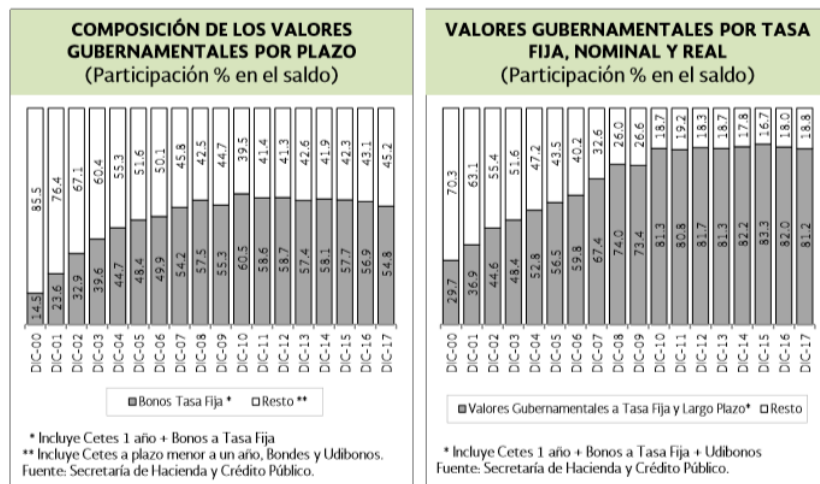


(a) Deuda Interna vs Deuda Externa (b) Plazo promedio de los valores

Figura 2.1: Fuente: Plan Anual de Financiamiento 2018

En la figura 2.1 (a) se muestran el porcentaje que representan los bonos a tasa fija en la cartera de la deuda interna. Esta gráfica muestra como la participación de la deuda denominada en moneda extranjera ha disminuido con el transcurrir de los años, indicando la preferencia del Gobierno Federal por así reducir el riesgo de la volatilidad de los tipos de cambio. La imagen 2.1 (b) indica una reducción en el plazo promedio de la cartera de la deuda, reflejando la implementación de estrategias de colocación que disminuyen la

exposición al riesgo.



(a) Bonos a tasa fija y de largo plazo vs. (b) Instrumentos que pagan tasa fija vs. resto de instrumentos que pagan a tasa flotante

Figura 2.2: Fuente: Plan Anual de Financiamiento 2018

En cuanto a la composición de la deuda interna del Gobierno Federal, la gráfica 2.2 (a) muestra como la mayor parte se encuentra contratada en instrumentos a tasa fija nominal denominada en pesos, por lo que la participación en el total de estos valores se mantuvo en niveles similares al pasar de 56,9% al cierre del año 2016 a 54,8% al cuarto trimestre de 2017. Además, esta gráfica muestra como los bonos a plazo largo y de tasa fija han cobrado importancia en el portafolio de deuda, reflejando la preferencia por aumentar el plazo promedio de vencimiento y poder así disminuir reducir la exposición a las tasas de interés. La imagen 2.2 (b) muestra como los instrumentos que pagan tasa fija han adquirido mayor importancia en el portafolio de la deuda, reflejando la implementación de estrategias que disminuyen la exposición al riesgo.

2.3. Características de la Deuda Pública futura de México

El Gobierno Federal cuenta con diversas formas de cubrir sus necesidades de financiamiento contando con múltiples instrumentos de deuda con los cuales capta recursos. La estrategia de elección dependerá de el comportamiento del mercado. Cada año el Gobierno Federal diseña diversas estrategias de colocación de valores, cada una de estas supone variantes en la composición del portafolio de deuda. La elección de la estrategia adecuada con base en en múltiples criterios y el marco de administración de riesgos.

De acuerdo al Plan Anual de Financiamiento 2018[28], PAF 2018, la política de endeudamiento interno se llevó a cabo buscando cubrir las necesidades de financiamiento del Gobierno Federal a costos reducidos, considerando un horizonte de largo plazo y un bajo nivel de riesgo, manteniendo una estrategia proactiva y flexible. De acuerdo al PAF 2018, el Gobierno Federal busca implementar una estrategia que incorpore las siguientes características:

- Cubrir las necesidades del Gobierno Federal a costos reducidos considerando un horizonte de largo plazo.

- Recurrir principalmente al mercado local de deuda.
- Analizar las condiciones de los mercados externos y de ser favorables poder acceder a los mismos para mantener una diversificación del portafolio de la deuda.
- Orientar las acciones en el manejo de la deuda a mejorar la eficiencia del portafolio de deuda del Gobierno Federal.

El escenario base de las necesidades de financiamiento del Gobierno Federal se basa en el Escenario Macroeconómico que se prevé en los Criterios Generales de Política Económica y en la Ley de Ingresos y el Presupuesto de Egresos de la Federación. Una vez que se cuenta con el escenario base se definen los límites de endeudamiento tanto interno como externo para poder así plantear diversos programas de emisión de deuda, los cuales son valuados y se estima una distribución de probabilidad de los montos a captar. La estrategia final seleccionada se encuentra basada en los objetivos de la estrategia de endeudamiento considerando el Costo en Riesgo de las tasa de interés para la deuda interna y el Costo en Riesgo de los tipo de cambio para deuda externa.

Capítulo 3

Conceptos básicos y supuestos en las tasas de interés

Una vez que se tiene el conocimiento de la estructura del portafolio de la deuda y las características de endeudamiento futuro, se deben calcular los pronósticos de las trayectorias de las curvas de rendimiento y los tipos de cambio para poder estimar el costo financiero a corto plazo.

La estimación del CaR para las tasas de interés para la deuda denominada en moneda local que realiza la SHCP se basa en el modelo de estructura de tasas de Cox-Ingersoll-Ross, CIR. Antes de describir de forma detallada el modelo, así como la estimación de los parámetros correspondientes para la estimación del costo financiero en pesos, es necesario dar los fundamentos básicos para el estudio de los modelos de tasas de interés, en particular para comprender mejor la dinámica del modelo CIR.

La primera definición a considerar es la de cuenta bancaria, que representa una inversión (localmente) libre de riesgo donde los beneficios se acumulan continuamente de acuerdo a la tasa libre de riesgo que prevalece en el mercado en cada instante [8].

Definición 3.1 *Definimos $B(t)$ como el valor de una cuenta en el banco al tiempo t . Se asume que $B(0) = 1$ y que la el valor de la cuenta evoluciona de acuerdo a la ecuación diferencial:*

$$dB(t) = r(t)B(t)dt \quad (3.1)$$

Donde $r(t)$ es una función positiva del tiempo. Como consecuencia:

$$B(t) = e^{\int_0^t r(s)ds} \quad (3.2)$$

La definición muestra que una unidad invertida al tiempo 0 rinde el valor en (3.2) al tiempo t . La tasa r es conocida como la *tasa spot instantánea* o también conocida como la *tasa corta*. En muchas ocasiones, la cuenta bancaria funge como numerario (por lo pronto se entenderá por numerario como una referencia) que sirve para normalizar los precios de los activos disponibles.

Definición 3.2 El factor de descuento $D(t, T)$ entre los tiempos t y T , $t < T$, es el monto al tiempo t que es equivalente a una unidad de moneda pagadera al tiempo T , y está dado por:

$$D(t, T) = \frac{B(t)}{B(T)} = e^{-\int_t^T r(s)ds} \quad (3.3)$$

Definición 3.3 Un bono Cupón Cero con periodo de maduración T , es un contrato que le garantiza al poseedor el pago de una unidad de moneda al tiempo T , sin pagos intermedios. El valor de este contrato al tiempo t , donde $t < T$, es denotado por $P(t, T)$ y además por definición $P(T, T) = 1$ para toda T .

Si r es determinista entonces el valor de D es determinista también y por tanto $D(t, T) = P(t, T)$. Sin embargo, si r es estocástica, entonces $D(t, T)$ es estocástico al tiempo t , y depende de la evolución de r entre t y T . Por tanto el valor de $D(t, T)$ y $P(t, T)$ se encuentran muy relacionadas y $P(t, T)$ puede ser vista como la *esperanza* de la variable aleatoria $D(t, T)$ bajo una medida de probabilidad en particular.

Definición 3.4 Fecha de convención. Se denota como $\tau(t, T)$ a la medida elegida entre los tiempos t y T . A $\tau(t, T)$ se le puede interpretar como la diferencia entre los tiempos t y T .

Si D_t y D_T representan dos fechas tal que $D_t < D_T$, asumiendo que el tiempo en el modelo es discreto, y además es medido en días, entonces algunos ejemplos de fechas de convención son:

1. *Actual 360* cuya definición es $\tau(t, T) = \frac{D_T - D_t}{360}$, es decir, todos los meses se computan por los días reales que tienen y los años como si tuvieran 360 días.
2. *Actual 365* cuya definición es $\tau(t, T) = \frac{D_T - D_t}{365}$.
3. *Business 252* cuya definición es $\tau(t, T) = \frac{D_T - D_t}{252}$.

Definición 3.5 La tasa de interés simple que prevalece del tiempo t al tiempo de maduración T , $L(t, T)$, es la tasa constante a la cual una inversión de $P(t, T)$ unidades de alguna moneda al tiempo t rinde una unidad en moneda al tiempo T .

Por lo tanto se representa de la siguiente forma:

$$L(t, T) := -\frac{1 - P(t, T)}{\tau(t, T)P(t, T)} \quad (3.4)$$

La tasa $L(t, T)$ satisface entonces:

$$P(t, T)(1 + \tau(t, T)L(t, T)) = 1 \quad (3.5)$$

Definición 3.6 La tasa de interés compuesta que compone k veces al año que prevalece del tiempo t al tiempo de maduración T , $Y^k(t, T)$, es la tasa constante a la cual una inversión de $P(t, T)$ unidades de alguna moneda al tiempo t rinde una unidad en moneda al tiempo T siendo reinvertido el monto generado k veces cada año.

Por lo tanto se representa de la siguiente forma:

$$Y^k(t, T) := \frac{k}{P(t, T)^{\frac{1}{k\tau(t, T)}}} - k \quad (3.6)$$

La tasa $Y^k(t, T)$ satisface entonces:

$$P(t, T) \left(1 + \frac{Y^k(t, T)}{k}\right)^{k\tau(t, T)} = 1 \quad (3.7)$$

Una propiedad fundamental de esta definición es la existencia de una tasa que compone de manera continua que se puede obtener a partir de llevar al límite el número de veces que compone la tasa al año. De hecho, de la ecuación 3.7, podemos encontrar la tasa continua de la siguiente manera:

$$\begin{aligned} 1 &= \lim_{k \rightarrow \infty} P(t, T) \left(1 + \frac{Y^k(t, T)}{k}\right)^{k\tau(t, T)} \\ &= P(t, T) \lim_{k \rightarrow \infty} \left(1 + \frac{Y^k(t, T)}{k}\right)^{k\tau(t, T)} \\ &= P(t, T) e^{Y^k \tau(t, T)} \end{aligned} \quad (3.8)$$

Despejando a la tasa se puede encontrar a la tasa que compone continuamente, $Y^k = -\frac{\ln P(t, T)}{\tau(t, T)}$, lo que nos lleva a la siguiente definición:

Definición 3.7 La tasa de interés continua que prevalece del tiempo t al tiempo de maduración T , $R(t, T)$, es la tasa constante a la cual una inversión de $P(t, T)$ unidades de alguna moneda al tiempo t se compone continuamente hasta rendir una unidad en moneda al tiempo T .

Por lo tanto se representa de la siguiente forma:

$$R(t, T) := -\frac{\ln P(t, T)}{\tau(t, T)} \quad (3.9)$$

La tasa $R(t, T)$ satisface entonces:

$$e^{R(t, T)\tau(t, T)} P(t, T) = 1 \quad (3.10)$$

Definición 3.8 La curva cupón cero (también llamada curva de rendimientos o estructura de plazo de tasa de interés) al tiempo t (en años) es la gráfica de la función:

$$T \rightarrow \begin{cases} L(t, T) & t < T \leq t + 1 \\ Y(t, T) & T > t + 1 \end{cases} \quad (3.11)$$

La curva cupón cero al tiempo t es un gráfico de la tasa de interés simple para todos los periodos de maduración T hasta un año y es el gráfico de la tasa de interés compuesta para todos los periodos de maduración T mayores a un año. En la figura 3.1 se muestra como luce una tasa cupón cero, en particular la del EURO mercado del día 13 de febrero de 2001.

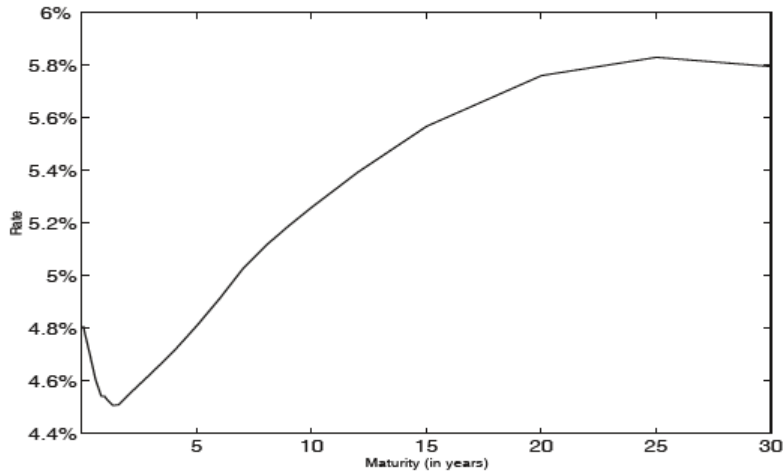


Figura 3.1: Ejemplo de tasa cupón cero. Fuente: Brigo D., Mercurio F. (2006). *Interest Rate Models- Theory and Practice*. Springer, 2a edición.

Definición 3.9 La tasa de interés forward simple es el valor de la tasa fija en un contrato FRA (Forward Rate Agreement) con periodos de expiración T y maduración S , $T < S$ que hacen que el contrato sea justo al tiempo t y se define de la siguiente manera:

$$F(t; T, S) := \frac{1}{\tau(T, S)} \left(\frac{P(t, T)}{P(t, S)} - 1 \right) \quad (3.12)$$

De la misma manera que en la tasa de interés que compone k veces al año, se puede encontrar una tasa continua, que se puede obtener cuando el periodo de maduración se contrae hacia su periodo de expiración, entonces se puede hablar de la *tasa forward instantanea* y se define de la siguiente manera:

$$f(t, T) := \lim_{S \rightarrow T^+} F(t; T, S) = - \frac{d \ln P(t, T)}{dT} \quad \forall t < T. \quad (3.13)$$

De hecho la tasa a corto plazo puede ser vista como un caso particular de la tasa forward, es decir $r(t) = f(t, t)$.

Cuando se trata de valorar productos en los cuales intervienen tasas de interés es necesario abandonar la suposición de tasas deterministas y empezar a valorar el comportamiento de r a través del tiempo como un proceso estocástico, lo que conlleva a que 3.2 y 3.3 se

transformen en procesos estocásticos también. Se asume que la tasa de interés corta $\{r(t)\}_{t \geq 0}$ es un proceso de Ito¹, es decir satisface:

$$dr(t) = \mu(t, r(t)) + \sigma(t, r(t))dW(t) \quad \forall t \in [0, T] \quad (3.14)$$

O bien en su forma integral:

$$r(t) = r(0) + \int_0^t \mu(s, r(s))ds + \int_0^t \sigma(s, r(s))dW(s) \quad (3.15)$$

Donde W es un proceso de Wiener², además μ y σ son funciones de $[0, T] \times \mathbb{R}$ en \mathbb{R} y son conocidas como el coeficiente de tendencia, (*drift*), y coeficiente de difusión respectivamente, así mismo la parte derecha de la ecuación 3.15 corresponde a una integral de Ito, cuya definición se encuentra en el apéndice. La suposición de una tasa estocástica que satisface las condiciones anteriores, implican que la cuenta bancaria y el factor de descuento, 3.2 y 3.3 respectivamente, se encuentren dadas por las siguientes expresiones:

$$B(t) = \exp\left(\int_0^t r(0)ds + \int_0^t \int_0^t \mu(s, r(s))dr_s ds + \int_0^t \int_0^t \sigma(s, r(s))dsdW(s)\right) \quad (3.16)$$

$$D(t, T) = \exp\left(-\int_t^T r(0)ds - \int_t^T \int_0^t \mu(s, r(s))dr_s ds - \int_t^T \int_0^t \sigma(s, r(s))dsdW(s)\right) \quad (3.17)$$

3.1. Condición de no arbitraje

Una suposición fundamental económica en este trabajo sobre las tasas de interés, es la presentada por Black-Scholes [7], que se refiere a la ausencia de oportunidades de arbitraje en el mercado financiero sobre el que se trabaja. Formalmente hablando, la ausencia de arbitraje es equivalente a la imposibilidad de invertir cero y recibir en un futuro, un monto positivo con probabilidad positiva. Otra manera de verlo es que teniendo dos portafolios que tienen el mismo *payoff* a una misma fecha dada, deben necesariamente tener el mismo precio al día de hoy. Por la construcción de un portafolio adecuado que replica los retornos instantáneos de una inversión de mínimo riesgo es que Black-Scholes pudo concluir que el retorno instantáneo del portafolio fue de hecho la tasa libre de riesgo instantánea.

En esta sección se muestran los resultados principales de la suposición de ausencia de arbitraje, las principales herramientas que se usarán son el lema de Ito, el teorema de Girsanov, aunado a las medidas de probabilidad equivalentes y la derivada de Radon-Nikodym, estos conceptos se encuentran explicados más detalladamente en el apéndice.

¹Esta definición se encuentra más detalladamente en el apéndice.

²Para mayor detalle de esta definición véase el apéndice

Consideraremos el espacio de probabilidad filtrado $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathbb{P})^3$ y T el horizonte del tiempo, $T > 0$. En la economía estudiada, $k + 1$ activos que no pagan dividendos son comercializados de manera continua desde el tiempo 0 hasta el tiempo T . Se asume que los precios de los activos son modelados por el vector $S(t) = (S_t^0, S_t^1, \dots, S_t^k)$ que es un proceso estocástico a tiempo continuo, \mathcal{F}_t -adaptado⁴, cuyas componentes son positivas. Además S^0 representa una cuenta en el banco.

Definición 3.10 Una estrategia de trading es un proceso a tiempo continuo $(k+1)$ -dimensional $\phi(t) = (\phi_t^0, \phi_t^1, \dots, \phi_t^k)$, cuyas componentes son localmente acotadas y predecibles. El valor del proceso asociado a la estrategia ϕ es definido por:

$$V_t(\phi) = \phi(t)S(t) = \sum_{j=0}^k \phi_t^j S_t^j, \quad 0 \leq t \leq T \quad (3.18)$$

Y la ganancia del proceso asociado a la estrategia ϕ es definido por:

$$G_t(\phi) = \int_0^t \phi(u) dS(u) = \sum_{j=0}^k \int_0^t \phi_u^j dS_u^j, \quad 0 \leq t \leq T \quad (3.19)$$

La k -ésima componente, ϕ_t^k , de la estrategia ϕ es interpretada como el número de unidades del activo k que se encuentran en posesión del inversor al tiempo t . Además $V_t(\phi)$ y $G_t(\phi)$ se interpretan respectivamente como el precio de mercado del portafolio ϕ_t y las ganancias acumuladas por el inversionista hasta el tiempo t adoptando la estrategia ϕ_t .

Definición 3.11 Se dice que una estrategia de trading es autofinanciada si $V_t(\phi) \geq 0$ y además:

$$V_t(\phi) = V_0(\phi) + G_t(\phi), \quad 0 \leq t \leq T \quad (3.20)$$

La interpretación es que estrategia es autofinanciada si su valor cambia solamente debido a los cambios en los precios de los activos

Un resultado muy importante que mostraron Harrison-Kreps[18] y Harrison-Pliska [19] [20] es la relación entre el concepto de ausencia de arbitraje y la propiedad de la existencia de una medida de probabilidad equivalente, mejor conocida como la medida de riesgo neutral, cuya definición se da a continuación:

Definición 3.12 Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Una medida de riesgo neutral, \mathbb{Q} , es una medida de probabilidad en el espacio (Ω, \mathcal{F}) tal que:

- \mathbb{P} y \mathbb{Q} son medida equivalentes, es decir que $\mathbb{Q}(A) = 0 \Leftrightarrow \mathbb{P}(A) = 0 \quad \forall A \in \mathcal{F}$.
- La derivada de Radon-Nikodim⁵, $\frac{d\mathbb{Q}}{d\mathbb{P}}$, pertenece a $L^2(\Omega, \mathcal{F}, \mathbb{P})$ (es decir es cuadrado integrable con respecto a \mathbb{P}).

³Donde \mathcal{F}_t es la filtración generada por $\{r(t)\}$ hasta el tiempo t

⁴Para mayor detalle de esta definición véase el apéndice

⁵Para mayor detalle de esta definición véase el apéndice.

c) El proceso del precio del activo descontado, $\{D(0, t)S(t)\}_{t \geq 0}$, es una $(\mathcal{F}_t, \mathbb{Q})$ -martingala (es decir $\mathbb{E}[D(0, t)S_t^j | \mathcal{F}_u] = D(0, u)S_u^j \quad \forall j = 0, \dots, k$ y $\forall 0 \leq u \leq t \leq T$ con \mathbb{E} la esperanza bajo \mathbb{Q}).

Es entonces que Harrison y Pilska probaron que la existencia de una medida de probabilidad equivalente implica la ausencia de oportunidades de arbitraje.

Definición 3.13 Una oportunidad de arbitraje es definida, en términos matemáticos, como una estrategia autofinanciada ϕ tal que $V_0(\phi) = 0$, y además $\mathbb{P}[V_T(\phi) > 0] > 0$.

Definición 3.14 Una reclamación contingente, o pago futuro, es una variable aleatoria cuadrado integrable y positiva en $(\Omega, \mathcal{F}, \mathbb{P})$. Se dice que una reclamación contingente, H , es alcanzable al tiempo T si existe una estrategia autofinanciada, ϕ , tal que $V_T(\phi) = H(T)$

La siguiente proposición fue probada por Harrison y Pilska en 1981, proporciona la caracterización matemática de un precio único no arbitrable con cualquier reclamación contingente alcanzable.

Proposición 3.1 Asuma que existe una medida de probabilidad equivalente, \mathbb{Q} , y sea H una reclamación contingente alcanzable al tiempo T . Entonces para toda t , $0 \leq t \leq T$, **existe un único precio $\pi(t)$** asociado a $H(T)$, es decir:

$$\pi(t) = \mathbb{E}[D(t, T)H(T) | \mathcal{F}_t] \quad (3.21)$$

Definición 3.15 Un mercado financiero es completo si y solo si toda reclamación contingente es alcanzable.

Cuando el conjunto de todas las medidas de probabilidad equivalentes es no vacío, entonces es posible obtener un único precio no arbitrable asociado a cualquier reclamación contingente. Este precio está dado por la esperanza del *payoff* descontado bajo la medida \mathbb{Q} que es equivalente a \mathbb{P} . Este resultado generaliza el de Black-Scholes para la valoración de reclamos contingentes.

Es así que finalmente Harrison y Pilska [20] demostraron el siguiente resultado:

- a. Un mercado financiero es libre de arbitraje si y solo si existe una medida de probabilidad equivalente.
- b. Un mercado financiero es completo si la medida equivalente es única.
- c. La existencia de una única medida de probabilidad equivalente no solo hace al mercado libre de arbitraje sino que además le asocia un único precio a toda reclamación contingente.

Se le denomina *numerario* a todo activo, positivo, que no paga dividendos. Un numerario es un activo de referencia que es elegido para normalizar los otros precios de activos. Dado un numerario Z los precios relativos de un activo S están dados por $\frac{S^j}{Z}$, $j = 0, \dots, k$.

En la definición de medida de probabilidad equivalente vista hace un momento se ha asumido como numerario a la cuenta en el banco S^0 . Sin embargo, esta es una de muchas posibilidades, de hecho, existen numerarios más convenientes que se pueden elegir para la valuación de reclamaciones contingentes. En 1995, German et. al. [16] proporcionó una herramienta para la valuación de derivados y es, de hecho, una generalización de la proposición 3.1.

Proposición 3.2 *Se asume que existe un numerario N y una medida de probabilidad \mathbb{Q}^N equivalente a la medida \mathbb{P} tal que el precio de cualquier activo, X , relativo a N es una \mathbb{Q}^N -martingala⁶, es decir:*

$$\mathbb{E}_{\mathbb{Q}^N} \left[\frac{X(T)}{N(T)} \middle| \mathcal{F}_t \right] = \frac{X(t)}{N(t)}$$

Además si se asume que existe un numerario Z y una medida de probabilidad \mathbb{Q}^Z equivalente a la medida \mathbb{P} tal que el precio de cualquier reclamación contingente alcanzable, Y , relativo a Z es una \mathbb{Q}^Z -martingala, es decir:

$$\mathbb{E}_{\mathbb{Q}^Z} \left[\frac{Y(T)}{Z(T)} \middle| \mathcal{F}_t \right] = \frac{Y(t)}{Z(t)}$$

Entonces la derivada de Radon-Nikodym que determina a la medida \mathbb{Q}^Z está dada por:

$$\frac{d\mathbb{Q}^Z}{d\mathbb{Q}^N} = \frac{Z(T)N(0)}{Z(0)N(T)} \quad (3.22)$$

Esta proposición nos otorga la posibilidad de hacer un cambio de numerario lo cual abre la posibilidad de trabajar con diferentes medidas de probabilidad que pueden resultar muy útiles en la valoración de derivados ya que los procesos pueden ser convertidos en martingalas con respecto a la nueva medida elegida. La construcción de una nueva medida de probabilidad se verá a continuación.

3.2. Modelos de tasas cortas de un factor

La teoría de modelos de tasas de interés fue desarrollada originalmente basada en la suposición de una dinámica específica unidimensional para los procesos de las tasa de interés de corto plazo. Modelar las dinámicas de las tasa de interés se convirtió en algo conveniente ya que las cantidades son claramente definidas como la esperanza de un proceso bajo las suposiciones de no arbitraje. De hecho la existencia de una medida de riesgo neutral implica que el precio libre de arbitraje, al tiempo t , de una reclamación contingente con *payoff*, $H(T)$ al tiempo T se encuentra dado por la ecuación 3.21.

⁶ \mathbb{Q}^N y \mathbb{Q}^Z son medidas de probabilidad equivalentes que dependen del numerario elegido.

La construcción de un portafolio adecuado de riesgo mínimo, tal como lo hiciera Black-Scholes, conlleva a la existencia de un proceso estocástico que solamente depende del tiempo presente, t , y de la tasa corta instantánea y no de las maduraciones de los pagos futuros que constituyen el portafolio. Este proceso es conocido como el **precio de mercado del riesgo**, que define un cambio de medida de acuerdo al teorema de Girsanov, de la medida del mundo real, \mathbb{P} , a una medida de riesgo neutral, \mathbb{Q} . Para precisar, se asume que la tasa corta instantánea se desenvuelve de acuerdo a la medida del mundo real, \mathbb{P} , de acuerdo a la siguiente ecuación diferencial estocástica:

$$dr(t) = \mu(t, r(t))dt + \sigma(t, r(t))dW(t) \quad (3.23)$$

Donde μ y σ son funciones definidas y W es un proceso de Wiener, es decir $\{r(t)\}_{t \geq 0}$ sigue un proceso de Ito. Es posible demostrar la existencia de un proceso estocástico $\{\lambda(t)\}_{t \geq 0}$ tal que si:

$$dP(t, T) = \mu^*(t, r(t))dt + \sigma^*(t, r(t))dW(t) \quad (3.24)$$

entonces:

$$\frac{\mu^*(t, r(t)) - r(t)P(t, T)}{\sigma^*(t, r(t))} = \lambda(t) \quad \forall \text{ fecha de maduración } T. \quad (3.25)$$

Este proceso estocástico puede depender de r pero no de T . Si además $\{\lambda(t)\}_{t \geq 0}$ es un proceso acotado, entonces el proceso $\{M(t)\}_{t \geq 0}$ definido por:

$$M(t) = \exp\left(-\frac{1}{2} \int_0^t \lambda^2(s)ds - \int_0^t \lambda(s)dW(s)\right) \quad (3.26)$$

es una \mathbb{P} -martingala⁷ y además podemos definir a partir de este proceso una nueva medida de probabilidad \mathbb{Q} equivalente a \mathbb{P} definida por:

$$\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}[I_A M(T)] = \int_A M(T)dP \quad (3.27)$$

En consecuencia la dinámica del proceso $\{r(t)\}_{t \geq 0}$ bajo \mathbb{Q} está dada por la siguiente ecuación:

$$dr(t) = [\mu(t, r(t)) - \lambda(t)\sigma(t, r(t))]dt + \sigma(t, r(t))d\hat{W}(t) \quad (3.28)$$

donde $\hat{W}(t) = W(t) + \int_0^t \lambda(s)ds$ es un movimiento Browniano bajo \mathbb{Q} .

Cabe señalar que a la variable aleatoria $\{M(t)\}_{t \geq 0}$ se le conoce como la derivada de Radon-Nikodym y su notación más común es:

⁷La demostración se encuentra en Steele, Stochastic Calculus and Financial Applications, p.p 222-224. [30]

$$\left. \frac{dQ}{dP} \right|_{\mathcal{F}_t} \quad \text{donde } \mathcal{F}_t \text{ es la filtración generada por } r \text{ hasta el tiempo } t. \quad (3.29)$$

La ecuación 3.24 expresa la dinámica del precio de un bono en términos de la tasa corta $r(t)$. Ahora, se recuerda que $r(t)$ es la tasa de retorno instantáneo de una inversión libre de riesgo, por lo que la expresión $\mu - r$, representa la diferencia entre los retornos, es decir, muestra cuan mejor es la inversión con respecto a poner el dinero en una cuenta bancaria. Cuando esta cantidad es dividida entre σ , lo que se está haciendo es dividirlo entre la cantidad de riesgo al que estamos sujetos. Es por esta razón por la cual al proceso $\{\lambda(t)\}_{t \geq 0}$ se le conoce como el **precio de mercado del riesgo**. Una interpretación alternativa es que $\lambda(t)$ es el *exceso de retorno con respecto a una inversión libre de riesgo por unidad de riesgo*.

Brigo y Mercurio[9] comentan con respecto a los cambios de medidas de probabilidad, que si solamente se está interesado en la valoración de instrumentos de tasas de interés, se puede modelar la dinámica de las tasas bajo la medida de riesgo neutral, \mathbb{Q} , por lo que λ estará implícita en la dinámica de la tasa de acuerdo a la medida de nuestra elección. Señalan que es posible *moverse* en el mundo del riesgo neutral, \mathbb{Q} , sin importar la dinámica de las tasas bajo la medida \mathbb{P} y que si solamente se quiere valorar derivados, la medida \mathbb{P} no es necesaria, lo más importante es el valor de los parámetros de los modelos bajo a medida \mathbb{Q} [8].

Los primeros modelos de tasas spot instantáneas que fueron introducidos en la literatura financiera fueron homogéneos en el tiempo, lo que significa que la dinámica de la tasa depende solamente de coeficientes constantes. Alguno de los modelos más famosos de un factor son, por orden cronológico, Vasicek [33], Dothan [13], Cox-Ingersoll-Ross [12]. En particular estos modelos son muy utilizados por su facilidad para poder modelar los precios de los bonos, de hecho una vez que se conoce la distribución de la tasa corta, cualquier *payoff* puede ser estimado por medio de la ecuación 3.21.

El siguiente cuadro muestra la dinámica de los modelos ya mencionados e indica además si la tasas generadas por el modelo son positivas. En la última columna se muestra la distribución de las tasa que respectivamente son Normal (N), Chi-cuadrada no central (NC_{χ^2}) y Lognormal (LN).

Modelo	Dinámica	$r > 0$	$r \sim$
Vasicek	$dr_t = k[\theta - r_t]dt + \sigma dW_t$	No	N
Cox-Ingersoll-Ross	$dr_t = k[\theta - r_t]dt + \sigma \sqrt{r_t} dW_t$	Sí	NC_{χ^2}
Dothan	$ar_t dt + \sigma r_t dW_t$	Sí	LN

Cuadro 3.1: Modelos clásicos de tasas instantáneas

Capítulo 4

Modelo de Cox-Ingersoll-Ross

En 1985 John Cox, Jonathan Ingersoll y Stephen Ross presentaron su trabajo, *A Theory of the Term Structure of Interest Rates*[12], en el cual proponen una dinámica para la evolución de las tasas de interés que se utiliza para la estimación de precios de activos. El modelo pertenece a la clase de modelos afines, cuya principal característica es que describen la estructura de las tasas de interés como una función lineal de varios factores estocásticos. Este es ya un modelo clásico en el mundo financiero, acerca de la dinámica del comportamiento de una tasa de interés y ha sido un marco de referencia por la manejabilidad analítica y por el hecho de que las tasas resultan ser siempre positivas.

El modelo que sigue la tasa $\{r(t)\}_{t \geq 0}$ bajo la medida \mathbb{Q} está dado por la siguiente ecuación diferencial estocástica:

$$dr(t) = k[\theta - r(t)]dt + \sigma\sqrt{r(t)}dW(t), \quad r(0) = r_0 \quad (4.1)$$

Donde los parámetros k, θ, σ son constantes positivas, W_t es un proceso de Wiener y el valor r_0 es no negativo. El proceso corresponde a un autorregresivo continuo de primer orden, que se definirá en el siguiente capítulo, de tal manera que la tasa de interés tiende al valor de largo plazo θ , conocido también como el parámetro de tendencia (*deriva* o *drift* en inglés), cuando t tiende a infinito. El valor del parámetro k determina la velocidad de ajuste a θ y σ determina el parámetro volatilidad.

La condición $2k\theta \geq \sigma^2$ es impuesta para asegurar que el origen se vuelva inaccesible y poder así garantizar que la tasa $r(t)$ sea siempre no negativa.

El comportamiento de la tasa de interés que sigue este modelo siguen empíricamente las siguientes propiedades:

- No se producen tasas negativas.
- Si la tasa alcanza el 0, esta subsecuentemente puede llegar a ser positiva.
- La varianza de las tasas de interés aumenta cuando la tasa por sí misma aumenta.
- Existe una distribución estacionaria de la tasa de interés, es decir la distribución de la tasa se aproxima a una distribución Gamma .

A continuación se presenta la simulación de una tasa que sigue este modelo y que inicia en el origen.

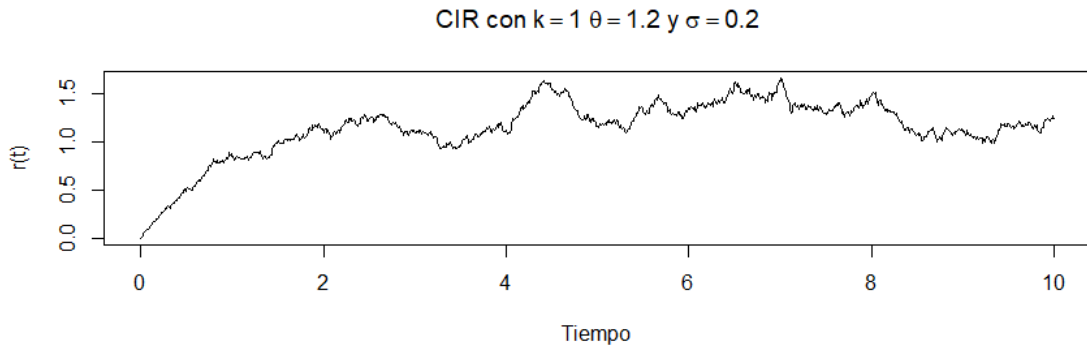


Figura 4.1: Simulación de la dinámica de una tasa de acuerdo al modelo de Cox-Ingersoll-Ross.

La función de densidad de probabilidad de la tasa de interés al tiempo s , condicionada a su valor en el tiempo presente, t donde $s > t$, se encuentra dada por:

$$f_{r(s)|r(t)}(r_s; r_t) = \frac{d}{dr_s} \mathbb{P}[r(s) \leq r_s | r(t) = r_t] = c e^{-u-v} \left(\frac{v}{u}\right)^{\frac{q}{2}} I_q(2\sqrt{uv}) \quad (4.2)$$

Donde:

$$\begin{aligned} c &= \frac{2k}{\sigma^2(1 - e^{-k(s-t)})} \\ u &= cr_t e^{-k(s-t)} \\ v &= cr_s \\ q &= \frac{2k\theta}{\sigma^2} - 1 \end{aligned}$$

Y además I_q es la función de Bessel modificada del primer tipo de orden q , que esta dada por:

$$I_q(z) = \sum_{j=0}^{\infty} \frac{(z/2)^{q+2j}}{(j!) \Gamma(q+j+1)}$$

Kladívko [22] menciona que a veces se utiliza la siguiente transformación $r(k) = 2cr(s)$ lo que implica que la función de densidad de esta variable sea la siguiente:

$$f_{r(k)|r(t)}(r_k; r_t) = \frac{1}{2c} f_{r(s)|r(t)}(r_s; r_t) \quad (4.3)$$

La función de distribución resultante es una Xi-cuadrada no central¹, $\chi^2_{(2q+2, 2u)}$, con $2q+2$ grados de libertad y parámetro de no centralidad $2u$.

La media y la varianza de $r(s)$ condicionado a su valor en el pasado al tiempo t , es decir, condicionado a la filtración $\{\mathcal{F}_t\}_{0 \leq t < s}$, se encuentran dadas por las siguientes ecuaciones²:

¹Para mayor detalle de esta definición véase el apéndice

²La demostración de la media y varianza se pueden encontrar en Salinas L. (2011). *Modelo de tasas de Interés Vasicek y algunas generalizaciones*. (Tesis de maestría). IPN, México. pp. 36-38

$$\mathbb{E}[r(s)|\mathcal{F}_t] = r(s)e^{-k(s-t)} + \theta(1 - e^{-k(s-t)}) \quad (4.4)$$

$$\text{Var}(r(s)|\mathcal{F}_t) = r(s)\left(\frac{\sigma^2}{k}\right)\left(e^{-k(s-t)} - e^{-2k(s-t)}\right) + \theta\left(\frac{\sigma^2}{2k}\right)\left(1 - e^{-k(s-t)}\right)^2 \quad (4.5)$$

La esperanza y varianza tienden a θ y a cero, respectivamente, a medida que k (parámetro de velocidad de ajuste) se aproxima a infinito. Por otro lado cuando k tiende a cero entonces la esperanza se aproximará a la tasa de interés presente, mientras que la varianza se aproximará al valor $\sigma^2 r(t)(s-t)$. El siguiente resumen ayuda de manera visual mostrando convergencias mencionadas.

$$\begin{array}{ll} \mathbb{E}_t \xrightarrow[k \rightarrow \infty]{} \theta & \text{Var}_t \xrightarrow[k \rightarrow \infty]{} 0 \\ \mathbb{E}_t \xrightarrow[k \rightarrow 0]{} r(t) & \text{Var}_t \xrightarrow[k \rightarrow 0]{} \sigma^2 r(t)(s-t) \end{array}$$

Donde \mathbb{E}_t y Var_t son la media y varianza condicionadas del proceso $r(t)$ al tiempo t .

La propiedad fundamental de este modelo es que si el proceso muestra reversión a la media, es decir, si $k, \sigma > 0$, entonces a medida que el tiempo en el cual se está valorando, s , crece, la distribución de la tasa se aproxima a una distribución Gamma, y por tanto, la función de la distribución estacionaria se encuentra dada por:

$$f_{r(\infty)|r(t)}(r_\infty; r_t) = \frac{w^\nu}{\Gamma(\nu)} r_t^{\nu-1} e^{-wr} \quad (4.6)$$

Donde $w = \frac{2k}{\sigma^2}$ y $\nu = \frac{2k\theta}{\sigma^2}$. La media y varianza de la distribución estacionaria son respectivamente θ y $\frac{\sigma^2\theta}{2k}$.

Cox, Ingersoll y Ross mostraron que el precio del bono cupón cero con fecha de madurez T es:

$$P(t, T) = A(t, T)e^{-B(t, T)r(t)} \quad (4.7)$$

Donde:

$$\begin{aligned} h &= \sqrt{k^2 + 2\sigma^2} \\ A(t, T) &= \left[\frac{2h(\exp[(k+h)(T-t)/2])}{2h + (k+h)(\exp[(T-t)h] - 1)} \right]^{\frac{2k\theta}{\sigma^2}} \\ B(t, T) &= \frac{2(\exp[(T-t)h] - 1)}{2h + (k+h)(\exp[(T-t)h] - 1)} \end{aligned}$$

Bajo la medida del riesgo neutral, \mathbb{Q} , la dinámica del precio del bono puede ser obtenida aplicando el Lema de Ito³ a la ecuación 4.7:

$$dP(t, T) = r(t)P(t, T)dt - B(t, T)P(t, T)\sigma\sqrt{r(t)}dW(t) \quad (4.8)$$

Despejando r de la ecuación 4.7 para dejar la expresión anterior en términos del precio del bono, se tiene:

$$dP(t, T) = \frac{1}{B(t, T)} \ln \left[\frac{A(t, T)}{P(t, T)} \right] P(t, T)dt - \sigma P(t, T) \sqrt{B(t, T) \ln \left[\frac{A(t, T)}{P(t, T)} \right]} dW(T) \quad (4.9)$$

El precio del bono, ecuación 4.7, de acuerdo a Cox-Ingersoll-Ross es una función convexa decreciente de la tasa de interés y una función creciente (decreciente) de la fecha de maduración T . Mientras que los parámetros del proceso de la tasa, $\{r(t)\}_{t \geq 0}$, tienen los siguientes efectos. El precio del bono es una función convexa decreciente del nivel medio, θ , de la tasa de interés y creciente cóncava (decreciente convexa) del parámetro de velocidad de ajuste, k , si la tasa de interés es mayor (menor) que θ . El precio del bono es una función cóncava decreciente de la volatilidad σ .

Una vez que se conoce la dinámica del modelo de tasas de Cox-Ingersoll-Ross de manera teórica, así como sus propiedades básicas, el siguiente paso para poder realizar la estimación del Costo en Riesgo de la deuda en moneda local, es conocer la manera en que los parámetros del modelo CIR son estimados. A continuación se detalla el proceso de estimación de parámetros que utiliza la SHCP, que es la técnica propuesta por Overbeck y Rydén.

4.1. Dinámica en la medida \mathbb{P} y estimación histórica

Antes de iniciar con el proceso de estimación, es importante mencionar que a pesar de contar con un modelo manejable de tasas de interés en la medida \mathbb{Q} , es deseable que la dinámica de la tasa sea también un modelo manejable bajo la medida del mundo real \mathbb{P} .

Es claro que es deseable que el modelo debe ser manejable bajo la medida \mathbb{Q} , ya que bajo esta medida la esperanza de las reclamaciones contingentes se vuelven relativamente sencillas de calcular. Sin embargo, contestar por qué se desea que también la dinámica de la tasa bajo \mathbb{P} sea manejable requiere otra explicación. Suponga que se cuenta con un conjunto de observaciones diarias observadas $r_0, r_1, r_2, \dots, r_n$ que son asumidas como una realización del proceso $r(t)$ y quisiéramos incorporar la información a nuestro modelo. Entonces se puede estimar los parámetros del modelo con la serie de datos observada con la que se cuenta. Sin embargo, los datos con los que se cuentan provienen del mundo real, y sus propiedades estadísticas caracterizan la distribución del proceso de tasa de interés, $r(t)$, bajo la medida del mundo real \mathbb{P} . *Por lo tanto, lo que se estima (parámetros) de las observaciones históricas es la dinámica del proceso bajo \mathbb{P} .*

³Esta definición se encuentra más detalladamente en el apéndice.

Por otro lado, los *precios* son calculados a través de esperanzas bajo la medida del riesgo neutral, \mathbb{Q} . Cuando se observan precios, lo que se observan son esperanzas bajo la medida \mathbb{Q} . Por lo tanto, *para estimar (pronosticar) las tasas es necesario estimar los parámetros k , σ y θ bajo la dinámica de \mathbb{Q} .*

Entonces si contamos con la información histórica de la tasa, se necesita modelar la dinámica del modelo bajo la medida \mathbb{P} , Brigo-Mercurio recomiendan adoptar el siguiente modelo bajo esta medida:

$$dr(t) = [k\theta - (k + \lambda\sigma)r(t)]dt + \sigma\sqrt{r(t)}dW(t), \quad r(0) = r_0 \quad (4.10)$$

Donde λ es un nuevo parámetro, es, de hecho, el *precio de mercado del riesgo* mencionado en el capítulo anterior. Note que para el caso $\lambda = 0$ las dos dinámicas son iguales. Ya que se requiere que la naturaleza de la dinámica debe ser la misma bajo las dos medidas, se impone un cambio de medida Girsanov, de la misma forma que en el capítulo anterior, para pasar de 4.1 a 4.10 de la siguiente manera:

$$\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_t} = \exp\left(-\frac{1}{2} \int_0^t \lambda^2(s)r(s)ds + \int_0^t \lambda(s)\sqrt{r(s)}dW(s)\right) \quad (4.11)$$

Otra manera de verlo, es que se asume que el proceso del precio de mercado del riesgo, $\{\lambda(t)\}_{t \geq 0}$, tiene la forma particular:

$$\lambda(t) = \lambda\sqrt{r(t)} \quad (4.12)$$

Bajo esta elección se obtiene un proceso que es manejable bajo ambas medidas, en particular para la medida \mathbb{P} será de gran utilidad para la estimación de parámetros.

Cabe señalar que para la estimación de parámetros se utiliza un cambio de variable en la ecuación 4.10, con el fin de simplificar el modelo para poder realizar una estimación adecuada de los parámetros.

$$dr(t) = [a + br(t)]dt + \sigma\sqrt{r(t)}dW(t) \quad (4.13)$$

Donde $a = k\theta > 0$, $b = -(k + \lambda\sigma) < 0$ y $\sigma > 0$.

En 1997, Ludger Overbeck y Tobias Rydén presentaron su trabajo *Estimation in the Cox-Ingersoll-Ross model* [24] en cual proponen una metodología para la estimación de los parámetros del modelo CIR, para observaciones de las tasas en puntos equidistantes en el tiempo $t(k) = k\Delta t$, para $k = 0, 1, \dots, n$.

La idea para la estimación de parámetros que proponen Overbeck y Rydén está basada en la discretización del modelo para luego poder utilizar la estimación por mínimos cuadrados condicionales para los parámetros a y b . Mientras que para σ^2 se utiliza la estimación por máxima verosimilitud, ya que el termino de la varianza es el mismo bajo ambas medidas, es

decir igual en las ecuaciones 4.1 y 4.10. Overbeck-Rydén mencionan que dos propiedades atractivas de los estimadores por mínimos cuadrados condicionales son:

- a. Estimadores consistentes, asintóticamente normales, fáciles de calcular.
- b. Los estimadores de a y b son robustos en contra de la especificación incorrecta de σ .

4.2. Estimación por mínimos cuadrados condicionales

Sea $\phi = (a, b, \sigma)$ los parámetros del modelo y sea además el dominio de ϕ el espacio $\Phi = (0, \infty) \times (-\infty, 0) \times (0, \infty)$. De acuerdo a Revuz y Yor [25] la ecuación 4.13 tiene una única solución para cada $\phi \in \Phi$ con condición inicial $r(0) = r_0$. La estimación por mínimos cuadrados condicionales, la cual es una aproximación general para la estimación de los parámetros implicados en la función de la esperanza condicional de un proceso estocástico, fue dada a través de una técnica por Klimko y Nelson en 1978. El estimador por mínimos cuadrados condicionales, $\hat{\phi}_n$, busca minimizar la siguiente suma:

$$\sum_{k=1}^n (r_k - \mathbb{E}[r(k)|\mathcal{F}_{k-1}])^2 \quad (4.14)$$

Donde \mathcal{F}_{k-1} es la filtración generada por las observaciones r_1, \dots, r_{k-1} .

Como ya se había visto la media condicionada del proceso dada por 4.4, cambia de acuerdo a la nueva notación adoptada 4.13 para quedar de la siguiente manera:

$$\mathbb{E}[r(k)|\mathcal{F}_{k-1}] = \mathbb{E}[r(k)|r(k-1) = r_{k-1}] = \gamma_0 + \gamma_1 r_{k-1} \quad (4.15)$$

Donde $\gamma_0 = \frac{a}{b}(e^{b\Delta t} - 1)$ y $\gamma_1 = e^{b\Delta t}$.

Así los únicos parámetros que aparecen en la media condicionada son a y b y por tanto podemos ver a la suma dada por 4.14 como una función que depende de a y b , es decir $\varphi(n; a, b) = \sum_{k=1}^n (r_k - \mathbb{E}[r(k)|\mathcal{F}_{k-1}])^2$.

A partir de la ecuación 4.15 discretizan el modelo, además asumen que el valor esperado será el valor observado y de esta manera proponen que el modelo sea de la siguiente manera:

$$r(k) = \gamma_0 + \gamma_1 r(k-1) + \epsilon(k) \quad (4.16)$$

Donde $\{\epsilon(t)\}_{t \geq 0}$ es una martingala con respecto a $\{\mathcal{F}_t\}$, es decir ϵ_t es \mathcal{F}_t -medible y además $\mathbb{E}[\epsilon(t)|\mathcal{F}_{t-1}] = 0$. Lo que se pretende es hacer una regresión de este proceso

⁴Ya que el proceso $r(t)$ es un proceso de Markov es posible cambiar a \mathcal{F}_{k-1} por $r(k-1) = r_{k-1}$

⁵Para precisar la notación r_k hace referencia al valor de la observación k , es decir, es un valor real, mientras que $r(k)$ es el proceso $\{r\}$ evaluado al tiempo k , es decir es una variable aleatoria.

autoregresivo de primer orden, es decir, se asume que la variable $r(k)$ es respuesta de $r(k-1)$.

El estimador se puede encontrar al minimizar la ecuación 4.14. Para ello se deriva a $\varphi(n; a, b)$ con respecto a cada parámetro, se iguala a cero cada expresión, se encuentra un sistema de dos ecuaciones y se resuelve para cada parámetro. Los estimadores resultantes presentados por Overbeck y Rydén fueron los siguientes:

$$\hat{b}_n = \frac{1}{\Delta t} \ln \left[\frac{n^{-1} \sum_{k=1}^n (r_k - \bar{r})(r_{k-1} - \bar{r}')}{n^{-1} \sum_{k=1}^n (r_{k-1} - \bar{r}')^2} \right] = \frac{1}{\Delta t} \ln [\hat{\beta}_1]^6 \quad (4.17)$$

Y

$$\hat{a}_n = \left(\frac{\bar{r} - e^{\hat{b}_n \Delta t} \bar{r}'}{e^{\hat{b}_n \Delta t} - 1} \right) \hat{b}_n \quad (4.18)$$

Donde $\bar{r} = \frac{1}{n} \sum_{k=1}^n r_k$ y $\bar{r}' = \frac{1}{n} \sum_{k=1}^n r_{k-1}$.

De la ecuación 4.17 despejando a $\hat{\beta}_1$, se tiene que es igual a $e^{\hat{b}_n \Delta t}$. Además recordando que en una regresión lineal se cumple $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$, es que podemos reexpresar la ecuación 4.18:

$$\hat{a}_n = \left(\frac{\hat{b}_n}{\hat{\beta}_1 - 1} \right) \hat{\beta}_0^7 \quad (4.19)$$

Entonces se puede encontrar a los estimadores de γ_0 y γ_1 , simplemente sustituyendo a los parámetros a y b por sus respectivos estimadores \hat{a}_n y \hat{b}_n de la siguiente manera:

$$\hat{\gamma}_0 = \frac{\hat{a}_n}{\hat{b}_n} (e^{\hat{b}_n \Delta t} - 1) = \frac{\hat{\beta}_0}{\hat{\beta}_1 - 1} (\hat{\beta}_1 - 1) = \hat{\beta}_0 \quad (4.20)$$

$$\hat{\gamma}_1 = e^{\hat{b}_n \Delta t} = \hat{\beta}_1 \quad (4.21)$$

Es así que los estimadores γ_0 y γ_1 resultan entonces ser $\hat{\beta}_0$ y $\hat{\beta}_1$ (estimadores por mínimos cuadrados de una regresión lineal) respectivamente, por tanto este procedimiento tiene como objeto encontrar los estimadores \hat{a}_n y \hat{b}_n que servirán más adelante para el estimador de σ^2 .

Para el estimador de σ^2 se utilizará la varianza condicional dada por la ecuación 4.5, además, así como ocurre con la media condicional, la expresión cambia de acuerdo a la nueva notación adoptada, por lo que la varianza ahora se encuentra expresada de la siguiente manera:

⁶ $\hat{\beta}_1$ es el estimador por mínimos cuadrados de una regresión lineal.

⁷ $\hat{\beta}_0$ es el estimador por mínimos cuadrados de una regresión lineal.

$$\text{Var}(r(k)|\mathcal{F}_{k-1}) = \text{Var}(r(s)|r(k-1) = r_{k-1}) = \sigma^2(\eta_0 + \eta_1 r_{k-1}) \quad (4.22)$$

Donde:

$$\eta_0 = \frac{a}{2b^2}(e^{b\Delta t} - 1)^2 \quad (4.23)$$

$$\eta_1 = \frac{1}{b}e^{b\Delta t}(e^{b\Delta t} - 1) \quad (4.24)$$

Los estimadores para η_0 y η_1 se obtienen sustituyendo a los parámetros a y b por sus respectivos estimadores vistos anteriormente. Existen diversas formas de encontrar el estimador de σ^2 , inspirada por una regresión lineal, se puede considerar es siguiente estimador, propuesto por Carroll y Rupert (1988) al que llamaron *método de pseudo-verosimilitud*:

$$\tilde{\sigma}_n^2 = \frac{1}{n} \sum_{k=1}^n \frac{(r_k - \mathbb{E}[r(k)|r(k-1) = r_{k-1}])^2}{\hat{\eta}_0 + \hat{\eta}_1 r_{k-1}} \quad (4.25)$$

Otra manera de estimación es considerar la ecuación 4.22 como una regresión con $(r_k - \mathbb{E}[r_k|r_{k-1}])^2$ como respuesta de $\eta_0 + \eta_1 r_{k-1}$, este método fue propuesto una vez más por Carroll y Rupert [10] y lo llamaron *mínimos cuadrados no ponderados basados en residuales cuadrados*, y se puede obtener de la siguiente manera:

$$\hat{\sigma}_n^2 = \frac{\sum_{k=1}^n (\hat{\eta}_0 + \hat{\eta}_1 r_{k-1})(r_k - \mathbb{E}[r(k)|r(k-1) = r_{k-1}])^2}{\sum_{k=1}^n (\hat{\eta}_0 + \hat{\eta}_1 r_{k-1})^2} \quad (4.26)$$

Es de suma importancia señalar que una vez estimados los parámetros, para realizar predicciones de las tasas de interés se debe utilizar la ecuación 4.13.

Para hallar esta ecuación se realizó un cambio de variable para adaptar la dinámica de 4.10, pero también puede ser obtenida a partir de la ecuación 4.1, por lo que con esta particular forma de expresión es posible realizar tanto la estimación de parámetros, como la estimación, o mejor dicho la predicción, de la tasa en el futuro. Cada individuo es libre de elegir la dinámica en particular para poder estimar parámetros o predecir, pero la forma ya descrita en la ecuación 4.13 es la más conveniente ya que es capaz de hacer un *match* entre los mundos, la medida del mundo real, \mathbb{P} , y la medida del mundo del riesgo neutral, \mathbb{Q} .

Con estos elementos ya es posible hacer una estimación del Costo en Riesgo para la deuda en moneda local, es decir, para la deuda que está referenciada a las tasas de interés, precisando más, a los bonos que pagan a tasas flotante. La manera de estimar el CaR, tanto el de tipo de cambio como el de tasa de interés, se verá hasta el capítulo siete. En los siguientes dos capítulos se explicará de breve manera los modelos *ARIMA*, al igual que en el modelo *CIR*, se verá un contexto general del modelo así como la manera de estimar los parámetros correspondientes, que servirán para la estimación del CaR de los instrumentos que pagan en moneda extranjera.

Capítulo 5

Series de Tiempo

En este capítulo se discutirá brevemente acerca de la teoría básica de las series de tiempo, así como la estimación de parámetros en los modelos con el fin de mostrar la manera en que se estimará el Costo en Riesgo en moneda extranjera de la deuda pública.

El análisis de las series de tiempo lineales (univariadas) provee un marco de referencia para el estudio de la dinámica de la colección de variables aleatorias en el tiempo. Se definirán los conceptos básicos de estacionariedad, causalidad, invertibilidad, función de autocorrelación. Los modelos que se mencionarán son los modelos autorregresivos (AR), de medias móviles (MA) y autorregresivos de medias móviles (ARMA) para poder así presentar los modelos autorregresivos de medias móviles integrados (ARIMA).

Por último se definirán los estimadores de los parámetros y la manera de realizar pronósticos sobre los valores que tomen las series.

El enfoque de las series de tiempo como un proceso estocástico, fue desarrollado por Box-Jenkins en la década de 1970. Para el estudio de esta metodología se debe trabajar con aquellos procesos que sean estacionarios, que de manera ambigua, significa que las propiedades estadísticas del proceso no cambian a través del tiempo.

5.1. Conceptos básicos de las series de tiempo

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Un proceso estocástico $\{X(t)\}_{t \in T}$, es una colección de variables aleatorias, donde T es el espacio parametral, que es a lo más numerable.

Definición 5.1 *Una serie de tiempo es una sucesión de observaciones (realizaciones) correspondientes a un proceso estocástico tomados en intervalos regulares y de duración constante. La notación que se usará, para una mayor comprensión, será X_s para el valor al tiempo s del proceso $\{X(s)\}$.*

Definición 5.2 *Un proceso estocástico $\{X(t)\}_{t \in T}$ es fuertemente estacionario si la distribución conjunta de $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ es igual a la distribución conjunta de $(X_{t_1+k}, X_{t_2+k}, \dots, X_{t_n+k})$ para todo $k \in \mathbb{R}$ tal que t_1, t_2, \dots, t_n y $t_1 + k, t_2 + k, \dots, t_n + k$ pertenecen a T .*

Es decir:

$$F_{X_{t_1}, X_{t_2}, \dots, X_{t_n}}(a_1, a_2, \dots, a_n) = F_{X_{t_1+k}, X_{t_2+k}, \dots, X_{t_n+k}}(a_1, a_2, \dots, a_n) \quad (5.1)$$

Esta es una muy fuerte condición, que es difícil de verificar empíricamente. A partir de este momento se entenderá por un proceso estacionario, a un proceso que es estacionario solamente en el sentido débil, cuya definición se muestra a continuación:

Definición 5.3 Se dice que un proceso estocástico $\{X(t)\}_{t \in T}$ es **débilmente estacionario** si todos los momentos de primer y segundo orden de $X_{t_1}, X_{t_2}, \dots, X_{t_n}$ existen y son iguales a los momentos de $X_{t_1+k}, X_{t_2+k}, \dots, X_{t_n+k}$ para todo $k \in \mathbb{R}$ tal que t_1, t_2, \dots, t_n y $t_1+k, t_2+k, \dots, t_n+k$ pertenecen a T .

Esto equivale a decir que la media y covarianza es invariante en el tiempo, esto es, $\mathbb{E}[X_{t_1}] = \mathbb{E}[X_{t_2}]$ y además $Cov(X_{t_1}, X_{t_2}) = Cov(X_{t_1+k}, X_{t_2+k})$, donde $t_1 \neq t_2$ y $k \in \mathbb{R}$.

Definición 5.4 Se dice que el proceso $\{X(t)\}_{t \in T}$ es de segundo orden si es cuadrado integrable, es decir, si $\mathbb{E}[X_t^2] < \infty \forall t \in T$.

Definición 5.5 Sea $\{X(t)\}_{t \in T}$ un proceso de segundo orden, se define la función de autocovarianza entre X_{t_1} y X_{t_2} como:

$$\gamma_X(t_1, t_2) = \mathbb{E}[(X_{t_1} - \mu_{t_1})(X_{t_2} - \mu_{t_2})] \quad (5.2)$$

Definición 5.6 Sea $\{X(t)\}_{t \in T}$ un proceso de segundo orden, se define la función de autocorrelación entre X_{t_1} y X_{t_2} como:

$$\rho_X(t_1, t_2) = \frac{\gamma_X(t_1, t_2)}{\sqrt{\sigma_{t_1}^2} \sqrt{\sigma_{t_2}^2}} \quad (5.3)$$

Si se supone que los procesos son estacionarios y de segundo orden, entonces los procesos son débilmente estacionarios. Esta suposición implica que en todos los procesos la media, μ , y varianza, σ^2 , son constantes, a partir de esto, la función de autocorrelación y autocovarianza solo depende de la diferencia de entre los tiempos $|t_2 - t_1|$, conocido como *lago* retraso y denotado como τ , por lo que podemos reescribirlas de la siguiente manera:

Definición 5.7 Sea $\{X(t)\}_{t \in T}$ un proceso estacionario y de segundo orden. Se define la sucesión de autocovarianzas teórica, $\{\gamma(\tau)\}_\tau$, como la sucesión generada por las covarianzas entre X_t y $X_{t+\tau} \forall \tau \in \mathbb{Z}$, es decir:

$$\gamma_X(\tau) = Cov(X_t, X_{t+\tau}) = \mathbb{E}[(X_t - \mu)(X_{t+\tau} - \mu)] \quad (5.4)$$

Definición 5.8 Sea $\{X(t)\}_{t \in T}$ un proceso estacionario y de segundo orden. Se define la sucesión de autocorrelaciones teórica, ACF por sus siglas en inglés, $\{\rho(\tau)\}_\tau$, como la sucesión generada por las correlaciones entre X_t y $X_{t+\tau} \forall \tau \in \mathbb{Z}$, es decir:

$$\rho_X(\tau) = \frac{\gamma_X(\tau)}{\sqrt{\sigma^2} \sqrt{\sigma^2}} = \frac{\gamma_X(\tau)}{\gamma_X(0)} \quad (5.5)$$

La última igualdad se obtiene a partir de que $\gamma_X(0) = Cov(X_t, X_{t+0}) = Var(X_t) = \sigma^2$. A partir de este momento se asumirá que todos los procesos que se verán son estacionarios y de segundo orden. También a partir de este momento se omite el subíndice en las sucesiones autocovarianzas y autocorrelaciones, teniendo en cuenta que ya se sabe de que proceso se está haciendo referencia.

Es fácil de ver que para un proceso estacionario, la función de autocovarianzas, $\gamma(\tau)$, y la función de autocorrelaciones, $\rho(\tau)$, satisfacen las siguientes propiedades:

- a. $\gamma(0) = Var(X_t)$; y demás $\rho(0) = 1$.
- b. $|\gamma(\tau)| \leq \gamma(0)$; y demás $|\rho(\tau)| \leq 1$.
- c. $\gamma(\tau) = \gamma(-\tau)$; y demás $\rho(\tau) = \rho(-\tau) \forall k \in \mathbb{Z}$.

5.1.1. Transformaciones de las series

A partir de este momento, se requiere que las series a estudiar sean estacionarias con el fin de poderle asignar a los procesos algún modelo ARIMA, el cual se verá en las siguientes secciones. Antes de iniciar con su estudio de estos modelos, se mostrarán las maneras más comunes para transformar una serie de tal manera que la serie transformada sea estacionaria.

Una serie de tiempo puede ser vista como una función que depende de una *tendencia*, T_t , una *estacionalidad*, E_t , y de la *variabilidad*, ϵ_t . Una serie de tiempo entonces se puede descomponer en una de las siguientes formas:

- a. Esquema aditivo: $X_t = T_t + E_t + \epsilon_t$.
- b. Esquema multiplicativo: $X_t = T_t * E_t * \epsilon_t$

Se recomienda que antes de empezar con métodos que ayuden a la tendencia y estacionalidad de una serie, se empiece estabilizando la varianza de la misma. Para estabilizar una serie se recomiendan utilizar las *transformaciones de Box-Cox*, que son una serie de transformaciones que se le pueden aplicar a una serie, $\{X_t\}$, dadas por:

$$f_\lambda(X_t) = \begin{cases} \lambda^{-1}(X_t^\lambda - 1) & \text{si } X_t \geq 0, \lambda \neq 0. \\ \ln(X_t) & \text{si } \lambda = 0, \lambda > 0. \end{cases} \quad (5.6)$$

Para eliminar tendencia se puede hacer una regresión (lineal, polinómica, exponencial) o utilizar el método de medias móviles, y los residuos que se generan entre la serie y el ajuste, $\hat{\epsilon}_t = X_t - \hat{X}_t$, debería de ser una serie estacionaria. Otro método muy utilizado es el uso de los operadores de diferencias ∇ . Primeramente definimos al operador de retraso B , tal que:

1. $BX_t = X_{t-1}$.

2. $\nabla X_t = X_t - X_{t-1}$.
3. $\nabla^k X_t = (\nabla \cdots \nabla) X_t$ (se aplica k veces el operador).

Para $k \geq 1$ se tiene que $B^k X_t = X_{t-k}$ y además a partir de la definición de ∇ es fácil obtener el siguiente resultado:

$$\nabla^k X_t = (1 - B)^k X_t = \sum_{j=0}^k \binom{k}{j} (-B)^j X_t$$

En general, al diferenciar una serie k veces, es decir aplicar el operador ∇^k , lo que se está realizando es extraer una tendencia de un polinomio de orden k . Entonces si se tiene una serie que presenta tendencia, es decir, una serie de la forma $y_t = T_t + x_t$ de tal manera que x_t es una serie estacionaria, entonces al aplicar, un cierto número de veces, el operador ∇ la serie resultante debería de ser estacionaria. Se recomienda que el número de veces que una serie sea diferenciada no sea muy grande (más de 2 veces) ya que se puede perder información y además resulta difícil regresar la serie a su escala original.

En cuanto a la estacionalidad al igual que como en el caso de la tendencia se puede utilizar el método de Holt (doble exponencial)¹, y los residuos generados de la serie y el ajuste deberían de ser una serie estacionaria.

Otra manera de eliminar la estacionalidad de las series de tiempo es utilizar el operador de diferencias en la estacionalidad.

Sea $s = 2, 3, \dots$ el número de períodos distintos que componen una serie de tiempo (por ejemplo, estos pueden ser 12 si se trata de una serie mensual o 3 si son series cuatrimestrales) entonces se define el operador de diferencias para la estacionalidad, ∇_s , como sigue:

$$\nabla_s X_t = X_t - X_{t-s}$$

Al aplicar este operador a una serie que presenta estacionalidad cada s períodos, al ser las componentes estacionales iguales en cada estación, esto es $E_t = E_{t+ks} \forall k \in \mathbb{Z}$, la componente de estacionalidad es eliminada y por tanto, si no presenta tendencia o varianza inconsistente, la serie transformada es estacionaria.

5.1.2. Ruido blanco y series de tiempo univariadas

Un *ruido blanco* es una sucesión de variables aleatorias $\{\epsilon(t)\}_{t>0}$ no correlacionadas de una distribución dada, tales que $\mathbb{E}[\epsilon(t)] = 0$ y $Var(\epsilon(t)) = \sigma_\epsilon^2 \forall t > 0$.

Inmediatamente de la definición se tiene que la función de autocovarinzas del ruido blanco se encuentra dada de la siguiente manera:

¹Ver Brockwell, capítulo 9.

$$\gamma_{\epsilon}(\tau) = \begin{cases} \sigma_{\epsilon}^2 & \text{si } \tau = 0 \\ 0 & \text{si } \tau \neq 0 \end{cases} \quad (5.7)$$

De esta forma la sucesión de autocorrelaciones del ruido blanco es:

$$\rho_{\epsilon}(\tau) = \begin{cases} 1 & \text{si } \tau = 0 \\ 0 & \text{si } \tau \neq 0 \end{cases} \quad (5.8)$$

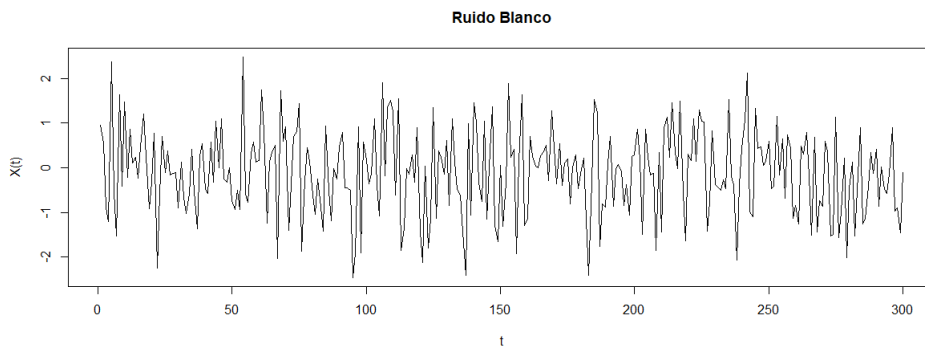


Figura 5.1: Simulación de tamaño 300 de un ruido blanco.

En la imagen 5.1 se puede visualizar el comportamiento de un ruido blanco Gaussiano. Un ruido blanco gaussiano es aquel cuyas componentes presentan distribución Normal con media cero y varianza uno.

Como nota, en algunas ocasiones se requerirá que el proceso de ruido blanco sea un conjunto de variables aleatorias independientes e idénticamente distribuidas con media 0 y varianza constante que es un tipo particular de ruido blanco.

5.2. Representaciones de series de tiempo

En la presente sección se presentarán los modelos autorregresivos y de medias móviles, así como sus principales características. Antes de iniciar con su estudio, se debe hacer mención de los procesos causales e invertibles, ya que algunas propiedades de los procesos ARMA se basan en la suposición de causalidad o invertibilidad del proceso.

5.2.1. Procesos causales

Una propiedad deseable de un proceso es que sea un proceso causal. Un proceso $\{X_t\}_{t \in T}$ es causal si existe una sucesión $\{\psi_j\}$ de tal manera que $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ y además X_t se puede expresar como una combinación lineal del proceso $\{\epsilon_t\}$. Es decir si:

$$X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} \quad \forall t \in T. \quad (5.9)$$

La siguiente proposición es presentada y demostrada por Brockwell [34] en su capítulo 2.

Proposición 5.1 *Sea $\{Y_t\}$ un proceso estacionario con función de autocovarianzas $\{\gamma_Y(\tau)\}_{\tau \in \mathbb{N}}$, asumiendo que $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ entonces la serie:*

$$\psi(B)Y_t = \sum_{j=-\infty}^{\infty} \psi_j Y_{t-j} \quad (5.10)$$

converge absolutamente y en media cuadrática² al mismo límite c.s. El proceso $X_t \equiv \psi(B)Y_t$ es un proceso estacionario de segundo orden con sucesión de autocovarianzas:

$$\gamma_X(\tau) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \gamma_Y(\tau - j + k) \quad (5.11)$$

Esta proposición indica que si $\{Y_t\}$ es un proceso estacionario, entonces el proceso $X_t \equiv \sum_{j=-\infty}^{\infty} \psi_j X_{t-j}$ es un proceso estacionario y de segundo orden, además de contar con una expresión para la sucesión de autocovarianzas. De hecho un proceso causal de la forma $\sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$ es un caso particular con $\psi_j = 0$ para $j \leq 0$, de esta manera si se logra identificar que un proceso ARMA(p,q) es causal, entonces es estacionario, porque el ruido blanco es estacionario y es posible obtener la sucesión de autocovarianzas según la proposición.

Teorema 5.1 *Sea $\{X_t\}$ un proceso ARMA(p,q) para el cual $\phi(\cdot)$ y $\theta(\cdot)$ no tienen ceros en común. Entonces $\{X_t\}$ es causal $\Leftrightarrow \phi(z) \neq 0 \forall z \in \mathbb{C}$ tal que $\|z\| \leq 1$. Los coeficientes $\{\psi_j\}$ quedan determinados por:*

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\phi(z)} \quad \text{con } \|z\| \leq 1. \quad (5.12)$$

Este teorema indica que un proceso es causal si y sólo si las raíces del polinomio autorregresivo se encuentran fuera del círculo unitario.³

De la ecuación anterior se tiene que $\psi(z)\phi(z)$ es un nuevo polinomio de la forma $\sum_{j=0}^{\infty} c_j z^j$, por tanto para obtener los coeficientes del polinomio se tiene que:

²Se entiende por convergente en media cuadrática $\lim_{n \rightarrow \infty} \mathbb{E}[(X_t - \sum_{j=-n}^n \psi_j X_{t-j})^2] = 0$

³Se refiere a la norma de la raíz ya que esta puede ser compleja.

$$\begin{aligned}
\psi(z)\phi(z) &= (\psi_0 + \psi_1z + \psi_2z^2 + \dots)(1 - \phi_1z - \phi_2z^2 - \dots - \phi_pz^p) \\
&= \psi_0 + (\psi_1 - \phi_1\psi_0)z + (\psi_2 - \psi_1\phi_1 - \psi_0\phi_2)z^2 + \dots \\
&= \sum_{j=0}^{\infty} c_j z^j
\end{aligned} \tag{5.13}$$

De esta manera el coeficiente asociado al nuevo polinomio es:

$$c_j = - \sum_{k=0}^j \phi_k \psi_{j-k} = -\phi_0 \psi_j - \sum_{k=1}^j \phi_k \psi_{j-k} = \theta_j \tag{5.14}$$

Por convención $c_0 = \theta_0 = \phi_0 = 1$. Además si $j \geq p$ y $j \geq q + 1$ entonces $\theta_j = 0$. Entonces se obtiene el siguiente resultado:

$$\begin{aligned}
\psi_j - \sum_{k=1}^j \phi_k \psi_{j-k} &= \theta_j \quad 1 \leq j < \max\{p, q + 1\} \\
\psi_j - \sum_{k=1}^q \phi_k \psi_{j-k} &= 0 \quad j \geq \max\{p, q + 1\}
\end{aligned} \tag{5.15}$$

Se pueden utilizar el sistema anterior para poder encontrar la sucesión para encontrar los pesos de $\{\psi_j\}$, con las condiciones iniciales $\theta_0 = 1 = \phi_0$.

5.2.2. Procesos invertibles

Un proceso $\{X_t\}_{t \in T}$ es invertible si existe una sucesión $\{\pi_j\}$ de tal manera que $\sum_{j=-\infty}^{\infty} |\pi_j| < \infty$ y además el proceso ϵ_t se puede expresar como una combinación lineal del proceso $\{x_t\}$. Es decir si:

$$\epsilon_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} \quad \forall t \in T. \tag{5.16}$$

Esta definición indica que es posible expresar el ruido blanco en términos de un proceso autorregresivo infinito.

Teorema 5.2 Sea $\{X_t\}$ un proceso ARMA(p, q) para el cual $\phi(\cdot)$ y $\theta(\cdot)$ no tienen ceros en común. Entonces $\{X_t\}$ es un proceso invertible $\Leftrightarrow \phi(z) \neq 0 \forall z \in \mathbb{C}$ tal que $\|z\| \leq 1$. Los coeficientes $\{\psi_j\}$ quedan determinados por:

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\phi(z)}{\theta(z)} \quad \text{con } \|z\| \leq 1. \tag{5.17}$$

De la ecuación anterior se tiene que $\pi(z)\theta(z)$ es un nuevo polinomio de la forma $\sum_{j=0}^{\infty} c_j z^j$, por tanto para obtener los coeficientes del polinomio se tiene que:

$$\begin{aligned}\pi(z)\theta(z) &= (\pi_0 + \pi_1 z + \pi_2 z^2 + \cdots)(1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q) \\ &= \pi_0 + (\pi_1 + \theta_1 \pi_0)z + (\pi_2 + \pi_1 \theta_1 + \pi_0 \theta_2)z^2 + \cdots \\ &= \sum_{j=0}^{\infty} c_j z^j\end{aligned}\tag{5.18}$$

De esta manera el coeficiente asociado al nuevo polinomio es:

$$c_j = \sum_{k=0}^j \theta_k \pi_{j-k} = \theta_0 \pi_j + \sum_{k=1}^j \theta_k \pi_{j-k} = \phi_j\tag{5.19}$$

Por convención $c_0 = \theta_0 = -\phi_0 = 1$. Además si $j \geq p$ y $j \geq q + 1$ entonces $\theta_j = 0$. Entonces se obtiene el siguiente resultado:

$$\begin{aligned}\pi_j + \sum_{k=1}^j \theta_k \pi_{j-k} &= \phi_j \quad 1 \leq j < \text{máx} \{p, q + 1\} \\ \pi_j + \sum_{k=1}^q \theta_k \pi_{j-k} &= 0 \quad j \geq \text{máx} \{p, q + 1\}\end{aligned}\tag{5.20}$$

Se pueden utilizar el sistema anterior para poder encontrar la sucesión para encontrar los pesos de $\{\pi_j\}$, con las condiciones iniciales $\theta_0 = 1 = -\phi_0 = \pi_0$.

5.2.3. Modelos MA(q)

Sea $\{X_t\}_{t>0}$ un proceso estocástico. Se dice que el proceso es de medias móviles (MA por sus siglas en inglés) de orden q si el proceso se puede escribir de la siguiente manera:

$$X_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} = \mu + \sum_{j=0}^q \theta_j \epsilon_{t-j}\tag{5.21}$$

Donde $\theta_0 = 1$, $\mu, \theta_j \in \mathbb{R} \forall j \in \mathbb{Z}$ y además $\{\epsilon_t\}_{t>0}$ es ruido blanco. A los coeficientes θ_i se les conoce como los pesos de la serie $\{X_t\}_{t>0}$

La esperanza de este proceso se puede encontrar de la siguiente manera:

$$\begin{aligned}
\mathbb{E}[X_t] &= \mathbb{E}[\mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}] \\
&= \mathbb{E}\left[\mu + \sum_{j=0}^q \theta_j \epsilon_{t-j}\right] = \mathbb{E}[\mu] + \mathbb{E}\left[\sum_{j=0}^q \theta_j \epsilon_{t-j}\right] \\
&= \mu + 0 = \mu
\end{aligned} \tag{5.22}$$

Esto es posible debido a que el ruido blanco, $\{\epsilon_t\}$, tiene media cero. De la misma manera para calcular la varianza de la serie, se tiene:

$$\begin{aligned}
\text{Var}(x_t) &= \text{Var}(\mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}) \\
&= \text{Var}\left(\mu + \sum_{j=0}^q \theta_j \epsilon_{t-j}\right) \\
&= \text{Var}(\mu) + \text{Var}\left(\sum_{j=0}^q \theta_j \epsilon_{t-j}\right) \\
&= 0 + \sigma_\epsilon^2 [1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2] \\
&= \sigma_\epsilon^2 \sum_{j=0}^q \theta_j^2
\end{aligned} \tag{5.23}$$

Donde σ_ϵ^2 es la varianza del ruido blanco y además $\theta_0 = 1$. Ahora, se calculará la sucesión de autocovarianzas teóricas, $\{\gamma_X(\tau)\}$, para el lag τ :

$$\begin{aligned}
\gamma_X(\tau) &= \text{Cov}(X_t, X_{t+\tau}) = \mathbb{E}[(X_t - \mu)(X_{t+\tau} - \mu)] \\
&= \mathbb{E}\left[\left(\sum_{i=0}^q \theta_i \epsilon_{t-i}\right)\left(\sum_{j=0}^q \theta_j \epsilon_{t+\tau-j}\right)\right] \\
&= \mathbb{E}\left[\sum_{i,j=0}^q \theta_i \epsilon_{t-i} \theta_j \epsilon_{t+\tau-j}\right] \\
&= \sum_{j=0}^q \theta_j \theta_{j+\tau} [\mathbb{E}[\epsilon_{t-j} \epsilon_{t+\tau-j}] - \mathbb{E}[\epsilon_{t-j}] \mathbb{E}[\epsilon_{t+\tau-j}]] \\
&= \sum_{j=0}^q \theta_j \theta_{j+\tau} \gamma_\epsilon(\tau)
\end{aligned} \tag{5.24}$$

Por lo tanto, dado que $\gamma_\epsilon(\tau) = \sigma_\epsilon^2$ si $\tau = 0$ y como debe ocurrir que $\tau + j \leq q$ y como $j \geq 0$, entonces $0 \leq j \leq q - \tau$, es decir la sucesión de autocovarianzas teóricas es la siguiente:

$$\gamma_X(\tau) = \begin{cases} \sigma_\epsilon^2 \left(\sum_{j=0}^{q-\tau} \theta_j \theta_{j+\tau} \right) & \text{si } 0 \leq \tau \leq q \\ 0 & \text{si } \tau > q. \end{cases} \tag{5.25}$$

Y por tanto, la sucesión de autocorrelaciones teórica (ACF) está dada por:

$$\rho_X(\tau) = \begin{cases} \frac{\left(\sum_{j=0}^{q-\tau} \theta_j \theta_{j+\tau} \right)}{\sum_{j=0}^q \theta_j^2} & \text{si } 0 \leq \tau \leq q \\ 0 & \text{si } \tau > q. \end{cases} \quad (5.26)$$

Así que si se pudieran obtener la sucesión de autocorrelaciones y a partir de algún momento p , las $\gamma_X(p)$ son cero entonces se puede concluir que el modelo corresponde a un MA(p).

A partir de estos resultados, se concluye que un proceso de medias móviles de orden q , MA(q), es un proceso estacionario ya que su esperanza (\mathbb{E}), segundo momento (\mathbb{E}^2)⁴ y sucesión de autocovarianzas, γ , son finitas y no dependen del tiempo, t , solamente del lag τ .

5.2.4. Modelos AR(p)

Sea $\{X_t\}_{t>0}$ un proceso. Se dice que el proceso es autoregresivo (AR por sus siglas en inglés) de orden p si el proceso se puede escribir de la siguiente manera:

$$X_t = \mu + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t = \mu + \sum_{j=1}^p \phi_j X_{t-j} + \epsilon_t \quad (5.27)$$

Donde $\mu, \phi_j \in \mathbb{R} \forall j \in \mathbb{Z}$ y además $\{\epsilon_t\}_{t>0}$ es ruido blanco.

A partir de este momento se tomara el proceso centrado $\{Y_t\}_{t>0} = \{X_t - \mu\}_{t>0}$, es decir, el nuevo proceso tiene la forma:

$$Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \epsilon_t = \sum_{j=1}^p \phi_j Y_{t-j} + \epsilon_t \quad (5.28)$$

La esperanza del proceso centrado $\{Y_t\}$ se puede obtener de la siguiente manera:

$$\begin{aligned} \mathbb{E}[Y_t] &= \mathbb{E} \left[\sum_{j=1}^p \phi_j Y_{t-j} + \epsilon_t \right] \\ &= \mathbb{E}[\phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \epsilon_t] \\ &= 1 + \phi_1 \mathbb{E}[Y_{t-1}] + \cdots + \phi_p \mathbb{E}[Y_{t-p}] \end{aligned} \quad (5.29)$$

Entonces:

⁴Esta puede ser hallada al despejarla de la definición de varianza y covarianza

$$\begin{aligned}\mathbb{E}[Y_t] - \phi_1\mathbb{E}[Y_{t-1}] - \cdots - \phi_p\mathbb{E}[Y_{t-p}] &= 1 \\ \mathbb{E}[Y_t](1 - \phi_1 - \cdots - \phi_p) &= 1\end{aligned}\quad (5.30)$$

Esta última expresión es posible realizarla debido a que el proceso $\{y_t\}$ es un proceso estacionario, de acuerdo a la suposición que se hizo en la sección pasada. Y por tanto la esperanza de Y_t es:

$$\mathbb{E}[Y_t] = \frac{1}{(1 - \phi_1 - \cdots - \phi_p)} \quad (5.31)$$

De la misma manera, es posible encontrar la varianza del proceso y es posible obtener:

$$\text{Var}(Y_t) = \frac{1 + \sigma_\epsilon^2}{(1 - \phi_1^2 - \cdots - \phi_p^2)} \quad (5.32)$$

Para encontrar la función de autocovarianzas teórica del proceso es necesario suponer que el proceso es causal, además se debe multiplicar por Y_{t-k} ambos lados de la igualdad 5.28 para así obtener:

$$Y_t Y_{t-k} = \phi_1 Y_{t-1} Y_{t-k} + \cdots + \phi_p Y_{t-p} Y_{t-k} + \epsilon_t Y_{t-k} \quad (5.33)$$

Tomando la esperanza se tiene:

$$\mathbb{E}[Y_t Y_{t-k}] = \phi_1 \mathbb{E}[Y_{t-1} Y_{t-k}] + \cdots + \phi_p \mathbb{E}[Y_{t-p} Y_{t-k}] + \mathbb{E}[\epsilon_t Y_{t-k}] \quad (5.34)$$

De la suposición de causalidad se tiene $\mathbb{E}[Y_t] = 0$ para toda $t > 0$. De esta manera podemos realizar lo siguiente:

$$\begin{aligned}\mathbb{E}[Y_t Y_{t-k}] &= \mathbb{E}[Y_t Y_{t-k}] - \mathbb{E}[Y_t] \mathbb{E}[Y_{t-k}] \\ &= \text{Cov}(Y_t Y_{t-k}) \\ &= \gamma_Y(k)\end{aligned}$$

De esta forma reexpresando y despejando la ecuación 5.34 es que se obtienen la siguiente expresión:

$$\gamma_Y(k) - \phi_1 \gamma_Y(k-1) - \cdots - \phi_p \gamma_Y(k-p) = \mathbb{E}[\epsilon_t Y_{t-k}] \quad (5.35)$$

Es posible demostrar, de nuevo a partir de la definición de causalidad que $\mathbb{E}[\epsilon_t Y_{t-k}] = 0$ para toda $1 < k < p$. De esta manera se obtienen el siguiente sistema de ecuaciones:

$$\begin{aligned}
\gamma(1) &= \phi_1\gamma(0) + \phi_2\gamma(1) + \cdots + \phi_p\gamma(p-1) \\
\gamma(2) &= \phi_1\gamma(1) + \phi_2\gamma(2) + \cdots + \phi_p\gamma(p-2) \\
\gamma(3) &= \phi_1\gamma(2) + \phi_2\gamma(1) + \cdots + \phi_p\gamma(p-3) \\
&\vdots \\
\gamma(p) &= \phi_1\gamma(p-1) + \phi_2\gamma(p-2) + \cdots + \phi_p\gamma(0)
\end{aligned} \tag{5.36}$$

Y dividiendo entre $\gamma(0)$ se obtienen la ecuaciones de Yule-Walker:

$$\begin{aligned}
\rho(1) &= \phi_1 + \phi_2\rho(1) + \cdots + \phi_p\rho(p-1) \\
\rho(2) &= \phi_1\rho(1) + \phi_2\rho(2) + \cdots + \phi_p\rho(p-2) \\
\rho(3) &= \phi_1\rho(2) + \phi_2\rho(1) + \cdots + \phi_p\rho(p-3) \\
&\vdots \\
\rho(p) &= \phi_1\rho(p-1) + \phi_2\rho(p-2) + \cdots + \phi_p
\end{aligned} \tag{5.37}$$

En general, las ecuaciones de Yule-Walker para un proceso autorregresivo de orden p están dadas por:

$$\gamma(j) = - \sum_{k=1}^p \phi_k \gamma(j-k) \quad \text{para } j = 1, \dots, p \tag{5.38}$$

Mientras que para el caso $j = 0$ la ecuación se encuentra dada por:

$$\gamma(0) = - \sum_{k=1}^p \phi_k \gamma(k) + \sigma_\epsilon^2 \tag{5.39}$$

Además el sistema 5.36 se puede expresar en su forma matricial de la siguiente manera:

$$\begin{bmatrix}
\gamma(0) & \gamma(1) & \gamma(2) & \cdots & \gamma(p-1) \\
\gamma(1) & \gamma(0) & \gamma(1) & \cdots & \gamma(p-2) \\
\gamma(2) & \gamma(1) & \gamma(0) & \cdots & \gamma(p-3) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\gamma(p-1) & \gamma(p-2) & \gamma(p-3) & \cdots & \gamma(0)
\end{bmatrix}
\begin{bmatrix}
\phi_1 \\
\phi_2 \\
\phi_3 \\
\vdots \\
\phi_p
\end{bmatrix}
=
\begin{bmatrix}
\gamma(1) \\
\gamma(2) \\
\gamma(3) \\
\vdots \\
\gamma(p)
\end{bmatrix} \tag{5.40}$$

Finalmente, las ecuaciones de Yule-Walker pueden escribirse, de la misma manera que el sistema anterior, en términos de autocorrelaciones de la siguiente forma:

$$\begin{bmatrix}
1 & \rho(1) & \rho(2) & \cdots & \rho(p-1) \\
\rho(1) & 1 & \rho(1) & \cdots & \rho(p-2) \\
\rho(2) & \rho(1) & 1 & \cdots & \rho(p-3) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\rho(p-1) & \rho(p-2) & \rho(p-3) & \cdots & 1
\end{bmatrix}
\begin{bmatrix}
\phi_1 \\
\phi_2 \\
\phi_3 \\
\vdots \\
\phi_p
\end{bmatrix}
=
\begin{bmatrix}
\rho(1) \\
\rho(2) \\
\rho(3) \\
\vdots \\
\rho(p)
\end{bmatrix} \tag{5.41}$$

Para encontrar los parámetros del proceso basta resolver el sistema de ecuaciones anterior.

5.2.5. Modelos ARMA(p,q)

Se dice que un proceso $\{X_t\}_{t \in T}$ es un ARMA(p, q) si el proceso es estacionario para cada $t \in T$ y además puede escribirse como:

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} \quad (5.42)$$

Donde $\phi_i, \theta_j \in \mathbb{R} \forall i, j \in T$ y además $\{\epsilon(t)\}_{t > 0}$ es un ruido blanco. Un proceso ARMA dado por la ecuación 5.42 también puede ser expresado como:

$$\phi(B)X_t = \theta(B)e_t \quad (5.43)$$

Donde $\phi(B)$ y $\theta(B)$ representan a los polinomios asociados a los procesos $\{X_t\}$ y $\{\epsilon_t\}$. Los polinomios asociados al proceso se definen de la siguiente manera:

$$\begin{aligned} \phi(Z) &= 1 - \phi_1 Z - \phi_2 Z^2 - \dots - \phi_p Z^p & Z \in \mathbb{C} \\ \theta(Z) &= 1 + \theta_1 Z + \theta_2 Z^2 + \dots + \theta_p Z^p & Z \in \mathbb{C} \end{aligned}$$

A $\phi(Z)$ se le conoce como el polinomio autorregresivo de orden p , mientras que a $\theta(Z)$ se le llama el polinomio de promedios móviles de orden q . Retomando al operador de retraso B , cuyas propiedades se mencionaron al principio del capítulo, $BX_t = X_{t-1}$ y $B^j X_t = X_{t-j}$, es posible mostrar que:

$$\begin{aligned} \phi(B)X_t &= (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)X_t \\ &= X_t - \phi_1 B X_t - \phi_2 B^2 X_t - \dots - \phi_p B^p X_t \\ &= X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} \end{aligned}$$

De la misma manera $\theta(B)\epsilon_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q}$. Un proceso ARMA(p,q) puede ser visto como una generalización de los modelo AR(p) y MA(q).

Capítulo 6

Identificación del modelo y pronósticos

Si se cuenta con una serie de datos, y se desea ajustar un modelo ARMA(p,q), evidentemente no es posible hacerlo con la sucesión de autocovarianzas y autocorrelaciones teóricas, así que se procede a estimarlas a partir de la muestra con la que se cuenta. Se recomienda ampliamente que el tamaño de la muestra sea mayor a 50 observaciones, para que las estimaciones resulten ser confiables.

La sucesión de autocovarianzas muestrales de una serie de datos $\{X_t\}$ ¹ está dado por:

$$\hat{\gamma}_X(\tau) = \begin{cases} \frac{1}{n} \sum_{t=1}^{n-\tau} (x_t - \bar{x}_n)(x_{t+\tau} - \bar{x}_n) & \text{si } 0 \leq \tau \leq n-1 \\ -\hat{\gamma}_X(\tau) & \text{si } n-1 \leq \tau \leq 0 \end{cases} \quad (6.1)$$

Mientras que la sucesión de autocorrelaciones (ACF) muestrales está dada por:

$$\hat{\rho}_X(\tau) = \frac{\hat{\gamma}_X(\tau)}{\hat{\gamma}_X(0)} \quad (6.2)$$

El estimador $\hat{\gamma}_X(\tau)$ presenta un sesgo, sin embargo su distribución cuando el número de observaciones tiende a infinito, tiene media $\gamma_X(\tau)$. La sucesión de autocovarianzas muestrales puede representarse de forma matricial, $\hat{\Gamma}_X$, de la siguiente manera:

$$\hat{\Gamma}_X(\tau) = \begin{bmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) & \cdots & \hat{\gamma}(n-1) \\ \hat{\gamma}(-1) & \hat{\gamma}(0) & \cdots & \hat{\gamma}(n-2) \\ \hat{\gamma}(-2) & \hat{\gamma}(-1) & \cdots & \hat{\gamma}(n-3) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\gamma}(n-1) & \hat{\gamma}(n-2) & \cdots & \hat{\gamma}(0) \end{bmatrix} \quad (6.3)$$

Sin embargo ya que se asume que el proceso $\{X_t\}$ es un proceso estacionario, entonces

¹Notese que se entiende por X_t al proceso $\{X(t)\}_{t \in T}$ al tiempo t , mientras que x_t es una realización del proceso al tiempo t

$\gamma_X(-\tau) = \gamma_X(\tau)$ para toda τ , y por tanto, la representación de la sucesión de autocovarianzas muestrales queda de la siguiente manera:

$$\Gamma_X(\tau) = \begin{bmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) & \cdots & \hat{\gamma}(n-1) \\ \hat{\gamma}(1) & \hat{\gamma}(0) & \cdots & \hat{\gamma}(n-2) \\ \hat{\gamma}(2) & \hat{\gamma}(1) & \cdots & \hat{\gamma}(n-3) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\gamma}(n-1) & \hat{\gamma}(n-2) & \cdots & \hat{\gamma}(0) \end{bmatrix} \quad (6.4)$$

De esta manera también es posible representar la sucesión de autocorrelaciones muestrales, de la siguiente forma:

$$P_X(\tau) = \begin{bmatrix} 1 & \hat{\rho}(1) & \cdots & \hat{\rho}(n-1) \\ \hat{\rho}(1) & 1 & \cdots & \hat{\rho}(n-2) \\ \hat{\rho}(2) & \hat{\rho}(1) & \cdots & \hat{\rho}(n-3) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\rho}(n-1) & \hat{\rho}(n-2) & \cdots & 1 \end{bmatrix} \quad (6.5)$$

Anteriormente se ha mencionado que si en un proceso MA a partir de algún retraso $\tau \geq q$ la sucesión de autocorrelaciones (ACF) es cero, entonces se está hablando de un proceso MA de orden q . Por tanto, si se tiene una muestra y se desea ajustar un modelo de medias móviles, entonces basta con identificar a partir de que momento ($\tau \geq q$) las autocorrelaciones muestrales son casi cero. Sin embargo, cuando se trata de ajustar un modelo AR o ARMA no es posible determinar el orden del modelo mediante las ACF, por ello existen dos aproximaciones para tratar de determinar el orden de la parte autorregresiva, la primera son las autocorrelaciones parciales (PACF), mientras que la segunda utiliza algunos criterios de información, los cuales se mencionarán más adelante.

6.0.1. Función de autocorrelaciones parciales

La función de autocorrelaciones parciales de una serie de tiempo es una herramienta muy útil para determinar el orden de la parte autorregresiva de un proceso. Una autocorrelación parcial mide el grado de asociación entre 2 variables, X e Y , sin considerar el efecto de una variable o un conjunto de variables, Z . Formalmente, es la correlación entre los residuos, $\{e_X\}$, de la regresión de la variable X con Z y los residuos, $\{e_Y\}$, de Y con Z .

Algebraicamente, la autocorrelación entre X e Y sin considerar los efectos de Z se encuentran determinados de la siguiente forma:

$$\rho_{XY.Z} = \frac{\rho(X, Y) - \rho(X, Z)\rho(Y, Z)}{\sqrt{1 - \rho^2(X, Z)}\sqrt{1 - \rho^2(Y, Z)}} \quad (6.6)$$

La autocorrelación parcial se puede generalizar para retrasos cualesquiera de la siguiente manera:

Sea $\{X_t\}_{t \in T}$ un proceso, donde $T = \{1, 2, 3, \dots\}$, tomar la regresión de X_1 y X_{k+1} con respecto a los valores intermedios, es decir con respecto de $\{X_2, X_3, \dots, X_k\}$ de tal manera que minimiza el error cuadrático medio :

$$\begin{aligned}\hat{X}_1 &= \alpha_1 + \alpha_2 X_2 + \alpha_3 X_3 + \dots + \alpha_k X_k \\ \hat{X}_{k+1} &= \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k\end{aligned}$$

La regresión no tiene intercepto ya que la esperanza del proceso $\{X_t\}_{t \in T}$ es cero. De esta manera los residuales $\{e_X\}$ y $\{e_Y\}$ quedan determinados de la siguiente manera:

$$\begin{aligned}e(1) &= X_1 - [\hat{\alpha}_1 + \hat{\alpha}_2 X_2 + \hat{\alpha}_3 X_3 + \dots + \hat{\alpha}_k X_k] \\ e(k+1) &= X_{k+1} - [\hat{\beta}_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 + \dots + \hat{\beta}_k X_k]\end{aligned}$$

De esta manera la autocorrelación parcial en k pasos (es decir $\tau = k$) corresponde a $PACF(k) = \rho(e(1), e(k+1))$, mientras que la autocorrelación parcial a un paso es $PACF(1) = \rho(e(1), e(2))$.

Definición 6.1 Sea $\{X_t\}_{t \in T}$ un proceso estacionario, entonces se define la sucesión de autocorrelaciones parciales (PACF), $\{\phi_{\tau\tau}\}$ para $\tau = 1, 2, \dots$ de la siguiente forma:

$$\{\phi_{\tau\tau}\} = \begin{cases} \rho(X_t, X_{t+1}) & \text{si } \tau = 1 \\ \rho(e(\tau+1), e(1)) & \text{si } \tau \geq 2 \end{cases} \quad (6.7)$$

Brockwell [9] p.p. 95 y Tsay [32] p.p. 36 mencionan que en un proceso $AR(p)$ causal, debido a que el valor de la PACF se vuelven cero (o casi cero en el caso de la PACF muestral) a partir del orden p , entonces los términos de la sucesión de autocorrelaciones muestrales sirve entonces para determinar el valor del orden p de la parte autorregresiva de un proceso. Además mencionan que $\hat{\phi}_{k,k}$ converge a ϕ_k a medida que el tamaño de la muestra tiende a infinito. En la sección 6.0.3 se mencionará de qué manera se debe elegir el orden p de manera más formal.

La siguiente tabla muestra las características que se pueden observar cuando se grafican las ACF y PACF teóricas de una serie de tiempo:

Proceso	ACF	PACF
$AR(p)$	Se presenta decaimiento exponencial.	Se vuelven cero después del lag p .
$MA(q)$	Se vuelven cero después lag q .	Se presenta decaimiento exponencial.
$ARMA(p, q)$	Se vuelven cero después del lag $(q - p)$.	Se vuelven cero después del lag $(p - q)$.

6.0.2. Ajuste del modelo

Para poder implementar un modelo ARMA a una serie de datos, es necesario, que la serie en cuestión sea estacionaria ya que si se le está ajustando un modelo, entonces los

datos deberían de ser una realización del modelo ajustado. Por tanto, el primer paso para poder ajustar un modelo ARMA adecuado, es verificar la estacionariedad en la serie. De no ser así, entonces se debe identificar la causa del por qué los datos corresponden a una serie estacionaria, esto puede ocurrir debido a las siguientes causas:

- a. Variabilidad no constante en los datos.
- b. Tendencia en los datos.
- c. Estacionalidad.

Recordando lo visto al principio del capítulo, para eliminar variabilidad se puede hacer uso de las transformaciones de Box-Cox. Para la tendencia de datos es posible identificarla visualmente graficando los datos que se tienen para poder después realizar una regresión (polinómica, exponencial, logarítmica, etc.) y sustraer los valores ajustados de la serie. Otra manera de eliminar tendencia es mediante la diferenciación de las series, esto debido a que, el orden k de diferenciación está asociado a la eliminación de la tendencia de un polinomio de orden k por lo que la serie diferenciada ya no debería presentar tendencia.

Finalmente, para estacionalidad es posible su identificación graficando los ACF ya que una lenta decaída y un comportamiento periódico en las autocorrelaciones, respectivamente, sería indicativo de estacionalidad. En caso de presentarla se puede eliminar mediante la diferenciación estacional mediante el operador ∇_s , donde s es el orden de la ciclicidad de la serie.

Cuando se tiene una serie que no es estacionaria una técnica utilizada para volverla estacionaria es la diferenciación de la serie. El número de diferencias por aplicar coincide con el número de raíces unitarias que presente la serie. El número de raíces unitarias se determina de manera formal con la prueba Dickey-Fuller aumentada (DFA).

Antes de detallar la prueba de Dickey-Fuller aumentada, se mostrará como se obtiene la prueba de Dickey-Fuller simple (DF) que se muestra en Wei [34].

Sea $\{X_t\}$ un proceso $AR(1)$ causal, es decir $X_t = \phi_1 X_{t-1} + \epsilon_t$ donde $|\phi_1| < 1$. Suponga que le aplicamos el operador ∇ al proceso, si una serie presenta una raíz unitaria entonces la serie no es estacionaria, por el contrario si la serie presenta raíces dentro del círculo unitario, entonces la serie diferenciada es una serie estacionaria. La prueba de Dickey-Fuller simple plantea lo siguiente:

$H_0 : \phi_1 = 1$ vs $H_1 : |\phi_1| < 1$ por tanto lo tanto:

$$\begin{aligned}
 \nabla X_t &= X_t - X_{t-1} \\
 &= \phi_1 X_{t-1} + \epsilon_t - X_{t-1} \\
 &= (\phi_1 - 1)X_{t-1} + \epsilon_t \\
 &= \gamma X_{t-1} + \epsilon_t
 \end{aligned} \tag{6.8}$$

Donde $\gamma = \phi_1 - 1$. Haciendo este cambio de variable, la prueba de hipótesis queda como sigue: $H_0 : \gamma = 0$ vs $H_1 : \gamma \neq 0$.

Por lo tanto, si en una serie de datos se quiere verificar la estacionariedad de la serie, el estadístico de prueba es :

$$\hat{d}f = \frac{\hat{\gamma}}{\sqrt{\text{Var}(\hat{\gamma})}} \quad \text{Donde } \hat{\gamma} = \hat{\phi}_1 - 1 \quad (6.9)$$

La prueba Dickey-Fuller aumentada es una generalización de la prueba Dickey-Fuller simple, pero ahora aplicada a los modelos $AR(p)$. Siguiendo el mismo proceso el estadístico se convierte en :

$$\gamma = \phi_1 + \phi_2 + \phi_3 + \dots - 1$$

El estadístico de prueba es el mismo que para la prueba DF, es decir, el que se utilizó para probar la existencia de una raíz unitaria en la serie. Si se rechaza H_0 se concluye que la serie no presenta raíz unitaria y por lo tanto es estacionaria.

Otra manera de hallar el orden de p y q es eligiendo aquellos que minimicen el criterio de información de Akaike Corregido (AICC). El AICC utiliza a los estimadores de θ y ϕ vía Máxima Verosimilitud. Este criterio de información se encuentra dado por:

$$AICC = -2 \ln \{\mathbb{L}(\zeta)\} + \frac{2(p + q + 1)n}{n - p - q - q} \quad (6.10)$$

Donde $\mathbb{L}(\zeta)$ es la función de verosimilitud de $\zeta = (\mu, \sigma_\epsilon^2, \phi, \theta)$ la cual se detallará más adelante. La idea es establecer un rango en los que puedan encontrarse tanto p como q , estimar vía máxima verosimilitud (la estimación se verá en la siguiente sección) los parámetros θ_i , σ_j y σ_ϵ^2 (donde $i \in \{1, \dots, p\}$ y $j \in \{1, \dots, q\}$) de los posibles modelos ARMA y finalmente elegir aquel modelo con el par p y q que minimicen el AICC.

La causa de por qué se debe minimizar este criterio es debido a que el tanto el AIC como el AICC son medias que indican la calidad del ajuste realizando que penaliza el número de parámetros, por lo que se deber elegir aquel modelo que contenga estas medidas lo más pequeña posible.

Existen otros modelos que también pueden ser utilizados como el Criterio de Información Bayesiana (BIC), cuya idea también es elegir aquellos pares de p y q que minimicen este criterio de información

6.1. Técnicas de pronósticos

En este capítulo lo que se pretende encontrar es un predictor de $X_{k+\tau}$ dado un conjunto de observaciones, $\{x_1, x_2, \dots, x_k\}$ de una serie estacionaria, de tal manera que se produzca el menor error posible, lo que conlleva a que se trate de minimizar el error cuadrático medio del pronóstico.

Definición 6.2 Sea $\{X_t\}$ un proceso estacionario, de segundo orden con media μ y sucesión de autocovarianzas $\{\gamma(\tau)\}$. Se define el mejor predictor lineal de $X_{k+\tau}$ en términos de $\{X_1, X_2, \dots, X_k\}$, como:

$$P_k X_{k+\tau} := a_0 + a_1 X_1 + \dots + a_k X_k \quad (6.11)$$

Que es la combinación lineal que minimiza el error cuadrático medio.

Dicho predictor es de hecho la proyección ortogonal de $X_{k+\tau}$ sobre el plano generado por $\{X_1, X_2, \dots, X_k\}$, de allí la notación $P_k X_{k+\tau}$. Al minimizar el error cuadrático medio lo que ocurre es que se minimiza la siguiente función:

$$S(a_0, \dots, a_k) = \mathbb{E}[(X_{k+\tau} - P_k X_{k+\tau})^2] \quad (6.12)$$

Como S es una función cuadrática y se encuentra acotada inferiormente por cero, está claro que existe al menos un valor (a_0, \dots, a_k) que minimizan a S y mínimo (a_0, \dots, a_k) satisface las ecuaciones:

$$\frac{d}{da_j} S(a_0, \dots, a_k) = 0 \quad j = 0, \dots, k$$

Es decir, para estimar a a_0 lo que se debe hacer es diferenciar esta función con respecto a a_0 e igualarla a cero de la siguiente manera :

$$\begin{aligned} \frac{d}{da_0} \mathbb{E}[(X_{k+\tau} - P_k X_{k+\tau})^2] &= \mathbb{E}\left[\frac{d}{da_0} (X_{k+\tau} - P_k X_{k+\tau})^2\right] \\ &= -2\mathbb{E}[(X_{k+\tau} - P_k X_{k+\tau}) * 1] \\ &= -2\mathbb{E}[X_{k+\tau} - (a_0 + a_1 X_1 + \dots + a_k X_k) * 1] \\ &= 0 \end{aligned} \quad (6.13)$$

Por lo tanto lo que se debe resolver es:

$$\begin{aligned} 0 &= \mathbb{E}[X_{k+\tau} - (a_0 + a_1 X_1 + \dots + a_k X_k)] \\ &= \mathbb{E}[X_{k+\tau}] - a_0 - \sum_{i=1}^k a_i \mathbb{E}[X_i] \end{aligned} \quad (6.14)$$

Es decir $\mathbb{E}[X_{k+\tau}] = a_0 + \sum_{i=1}^k a_i \mathbb{E}[X_i]$. Y por lo tanto, teniendo en cuenta que $\mathbb{E}[X_j] = \mu$ para toda $j \in T$, se tiene que:

$$a_0 = \mu \left(1 - \sum_{i=1}^k a_i\right) \quad (6.15)$$

Además por otra parte, si se quiere estimar a a_j se debe realizar el procedimiento análogamente, solamente se debe derivar con respecto a a_j , es decir:

$$\begin{aligned} \frac{d}{da_j} \mathbb{E}[(X_{k+\tau} - P_k X_{k+\tau})^2] &= \mathbb{E}\left[\frac{d}{da_j}(X_{k+\tau} - P_k X_{k+\tau})^2\right] \\ &= -2\mathbb{E}[(X_{k+\tau} - P_k X_{k+\tau}) * X_j] \\ &= -2\mathbb{E}[(X_{k+\tau} - a_0 + a_1 X_1 + \dots + a_k X_k) * X_j] \\ &= 0 \end{aligned} \quad (6.16)$$

Es decir se debe resolver:

$$0 = \mathbb{E}[(X_{k+\tau} - a_0 + a_1 X_1 + \dots + a_k X_k) * X_j] \quad (6.17)$$

Por lo tanto tomando la esperanza en la ecuación anterior se tiene:

$$\mathbb{E}[X_{k+\tau} X_j] = a_0 \mathbb{E}[X_j] + \sum_{i=1}^k a_i \mathbb{E}[X_i X_j] \quad (6.18)$$

Debido a la ortogonalidad de $X_{k+1} - \hat{X}_{k+1}$ con respecto a cualquier subconjunto de $\{X_0 = 1, X_1, \dots, X_k\}$ en particular con respecto a X_0 implica que es posible sustraer a μ de todas las variables ($P_k X_{k+\tau}$ y X_i), por lo que se puede asumir que $\mu = 0$ y recordando que $\gamma(|t - k|) = \mathbb{E}[X_t X_k] - \mathbb{E}[X_t] \mathbb{E}[X_k]$, en particular para un proceso estacionario y bajo lo comentado $\gamma(|t - k|) = \mathbb{E}[X_t X_k] - \mu^2 = \mathbb{E}[X_t X_k]$ entonces la ecuación anterior también se puede ver como:

$$\gamma(|k + \tau - j|) = a_0 \mu + \sum_{i=1}^k a_i \gamma(|i - j|) = \sum_{i=1}^k a_i \gamma(|i - j|) \quad (6.19)$$

La ecuación 6.19 se puede escribir de manera matricial de la siguiente manera:

$$\bar{\gamma}(\tau) = \Gamma_k A_k \quad (6.20)$$

Donde:

$$\begin{aligned} \bar{\gamma}(\tau) &= [\gamma(\tau), \gamma(\tau + 1), \dots, \gamma(\tau + k - 1)]^T \\ \Gamma_k &= \{\gamma(|i - j|)\}_{i,j=1}^k \\ A_k &= (a_1, a_2, \dots, a_k)^T \end{aligned} \quad (6.21)$$

Al sistema anterior se le conoce como las ecuaciones de predicción a τ pasos. El mejor predictor lineal de $X_{k+\tau}$ es el dado por la ecuación 6.11 el cual cumple que sus coeficientes asociados dados en la ecuación anterior, A_k , resuelve las ecuaciones de predicción.

El valor esperado del error de la predicción es:

$$\begin{aligned}\mathbb{E}[X_{k+\tau} - \hat{X}_{k+\tau}] &= \mathbb{E}[X_{k+\tau} - P_k X_{k+\tau}] \\ &= \mathbb{E}[X_{k+\tau}] - a_0 - \sum_{i=1}^k a_i \mathbb{E}[X_i] \\ &= 0\end{aligned}\quad (6.22)$$

Por otro lado el error cuadrático medio de la predicción es:

$$\begin{aligned}\mathbb{E}[(X_{k+\tau} - \hat{X}_{k+\tau})^2] &= \mathbb{E}[(X_{k+\tau} - \hat{X}_{k+\tau})(X_{k+\tau} - \hat{X}_{k+\tau})] \\ &= \mathbb{E}[X_{k+\tau}(X_{k+\tau} - \hat{X}_{k+\tau})]\end{aligned}\quad (6.23)$$

Esta última expresión es posible ya que $\hat{X}_{k+\tau}$ es ortogonal a $(X_{k+\tau} - \hat{X}_{k+\tau})$ y por lo tanto su producto es cero.

De las ecuaciones 6.11 y 6.15 se puede reexpresar la definición de mejor predictor lineal de la siguiente manera:

$$P_k X_{k+\tau} = \mu + \sum_{i=1}^k a_i (X_i - \mu) \quad (6.24)$$

Donde las a_i satisfacen la ecuación 6.20.

Brockwell señala que si $\{Y_t\}$ es un proceso estacionario y $\{X_t\}$ es un proceso con media cero tal que se encuentra determinado de la siguiente manera $X_t = Y_t - \mu$, entonces el mejor predictor lineal de Y_t es:

$$P_k Y_{k+\tau} = P_k X_{k+\tau} + \mu \quad (6.25)$$

6.1.1. Algoritmos recursivos

Sean Y y X_1, X_2, \dots, X_k variables aleatorias cualesquiera de segundo orden, es decir con segundo momento finito, con esperanza dadas por $\mathbb{E}[Y] = \mu_Y$, $\mathbb{E}[X_j] = \mu_j$ y covarianzas $Cov(Y, Y)$, $Cov(Y, X_j)$ y $Cov(X_j, X_i)$ todas conocidas.

Sea X el vector dado por $X = (X_1, \dots, X_k)^T$, el vector de medias $\mu_X = (\mu_1, \dots, \mu_k)^T$, el vector de autocovarianzas:

$$\gamma = Cov(Y, X) = (Cov(Y, X_1), \dots, Cov(Y, X_k))^T$$

y la matriz de covarianzas:

$$\Gamma = \text{Cov}(X, X) = [\text{Cov}(X_i, X_j)]_{i,j=1}^k$$

Entonces de la misma manera en que se encontró a $P_k X_{k+\tau}$ anteriormente, se puede encontrar el mejor predictor lineal de Y en términos de $\{1, X_1, \dots, X_k\}$, denotado por $P[Y|X]$ y esta dado por:

$$P[Y|X] = \mu_Y + a^T(X - \mu_X) \quad (6.26)$$

Donde $a = (a_1, \dots, a_k)^T$ es solución de:

$$\Gamma a = \gamma$$

El error cuadrático medio de este predictor esta dado por:

$$\mathbb{E}[(Y - P[Y|X])^2] = \text{Var}(Y) - a^T \gamma \quad (6.27)$$

Un caso muy particular es que si el proceso $\{X_t\}$ es un $AR(P)$ de hecho el vector $A_k = (a_1, a_2, \dots, a_k)^T$ es en realidad el vector compuesto por la autocorrelaciones parciales del proceso, es decir $A_k = \Phi_k = (\phi_{n,1}, \phi_{n,2}, \dots, \phi_{n,k})^T$, por lo que la ecuación 6.21 queda de la siguiente manera:

$$\bar{\gamma}(\tau) = \Gamma_k \Phi_k \quad (6.28)$$

Si Γ_k es no singular (es decir si existe su matriz inversa) entonces los elemento de Φ_k son únicos y además se encuentran dados por:

$$\Phi_k = \Gamma_k^{-1} \bar{\gamma}(\tau) \quad (6.29)$$

El sistema de ecuaciones 6.28 también se puede expresar en términos de correlaciones de la siguiente manera:

$$\bar{\rho}_k = P_k \bar{\Phi}_k \quad (6.30)$$

Donde:

$$\begin{aligned} \bar{\rho}_k &= [\rho(1), \rho(2), \dots, \rho(k)]^T \\ P_k &= \{\rho(|i-j|)\}_{i,j=1}^k \\ \Phi_k &= (\phi_{n,1}, \phi_{n,2}, \dots, \phi_{n,k})^T \end{aligned} \quad (6.31)$$

Las ecuaciones de predicción dadas por 6.28 o bien por 6.30 son muy similares a la ecuaciones de Yule- Walker, 5.40 y 5.41, vistas anteriormente por lo que un método que resuelve las ecuaciones de predicción sirve también para resolver la ecuaciones de Yule-Walker y viceversa. Para la resolución de estos sistemas existen métodos numéricos de gran utilidad, uno de ellos es el algoritmo de Durbin-Levinson que permite resolver de forma recursiva el sistema, lo que conlleva a poder realizar predicciones y obtener los estimadores vía Yule-Walker e incluso obtener las autocorrelaciones parciales (PACF).

El algoritmo de Durbin-Levinson se basa en obtener recursivamente para $n \geq 1$:

$$\phi_{nn} = \left[\gamma(n) - \sum_{j=1}^{n-1} \phi_{n-1,j} \gamma(n-j) \right] V_{n-1}^{-1} \quad (6.32)$$

Donde:

$$\begin{aligned} V_n &= V_{n-1} [1 - \phi_{n,n}^2] \\ \phi_{1,1} &= \frac{\gamma(1)}{\gamma(0)} = \rho(1) \\ V_0 &= \gamma(0) \end{aligned}$$

Resultado: Si $\{x_t\}$ es un proceso $AR(p)$ con media cero y causal, es decir es de la forma $X_t - \sum_{j=1}^p \phi_j x_{t-j} = \epsilon_t$, donde $\{\epsilon_t\}$ es ruido blanco, y si además $\Phi_k = P_k^{-1} \bar{\rho}_k$ con $k > p$ entonces:

$$\sqrt{n}(\hat{\Phi}_k - \Phi_k) \xrightarrow{n \rightarrow \infty} N(\underline{0}, \sigma_\epsilon^2 \Gamma_k^{-1}) \quad (6.33)$$

Y por lo tanto para $k > p$:

$$\sqrt{n}(\hat{\phi}_{k,k} - \phi_{k,k}) \xrightarrow{n \rightarrow \infty} N(0, 1) \quad (6.34)$$

De aquí que el intervalo de confianza al $(1 - \alpha)\%$ para $\phi_{k,k}$ se obtiene usando:

$$\mathbb{P}[-Z_{1-\frac{\alpha}{2}} < \sqrt{n}(\hat{\phi}_{k,k} - \phi_{k,k}) < Z_{1-\frac{\alpha}{2}}] \quad (6.35)$$

Donde $Z_{1-\frac{\alpha}{2}}$, es el cuántil de una distribución Normal estándar. Entonces el intervalo de confianza al $(1 - \alpha)\%$ para $\phi_{k,k}$ es:

$$(\phi_{kk} - Z_{1-\frac{\alpha}{2}}/\sqrt{n}, \phi_{kk} + Z_{1-\frac{\alpha}{2}}/\sqrt{n}) \quad (6.36)$$

Entonces de debe elegir aquel orden, " p ", a partir de cual no se rechaza la hipótesis nula de que los términos ϕ_{kk} relacionados con las autocorrelaciones parciales son cero. Es decir se elige al orden p como:

$$p = \min\{n \in \mathbb{N} | \forall k > n \quad |\hat{\phi}_{kk}| < Z_{1-\frac{\alpha}{2}} / \sqrt{n}\} \quad (6.37)$$

En resumen, mediante la sucesión de autocorrelaciones muestrales (ACF) es que se decide el orden q de la parte promedios móviles, MA, mientras que con las autocorrelaciones parciales muestrales (PACF) se decide el orden, p , de la parte autorregresiva del proceso $ARMA(p, q)$. Para ilustrar, en la figura 6.1 se muestra la simulación de 2 procesos $ARMA(1, 1)$, con sus respectivas ACF, el primero con ϕ_1 y θ_1

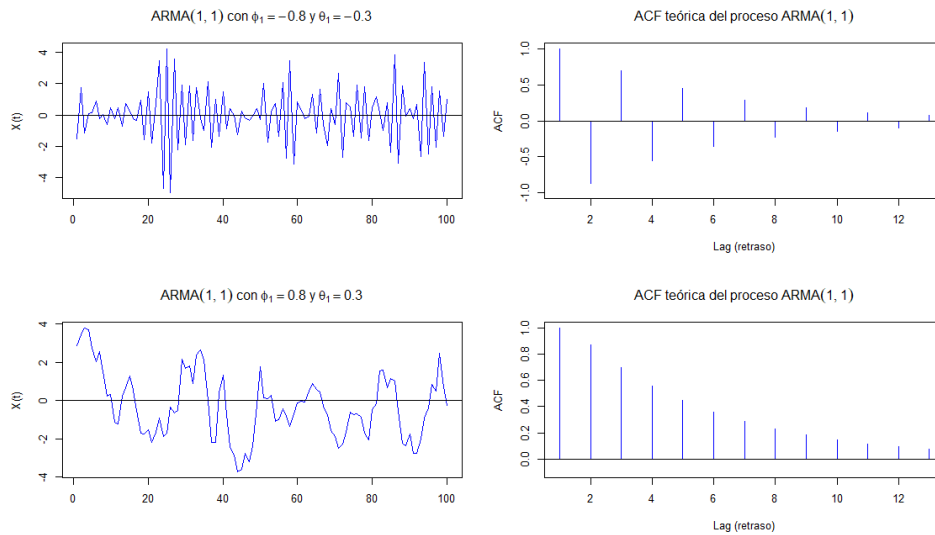


Figura 6.1: ACF de la simulación de dos procesos $ARMA(1,1)$

Otro algoritmo recursivo, que es aplicable a todos los procesos cuadrado integrables, sin importar si son o no estacionarias se muestra a continuación.

Sea $\{X_t\}$ un proceso con media cero y $\mathbb{E}[X_t^2] < \infty$ para toda t . La suposición $\mathbb{E}[X_t] = 0$ implica $Cov(X_i X_j) = \mathbb{E}[X_i X_j] = \kappa(i, j)$.

Es conveniente introducir a continuación la siguiente notación del mejor predictor lineal a un paso:

$$\hat{X}_n = \begin{cases} 0 & \text{si } n=1, \\ P_{n-1} X_n & \text{si } n=2,3,.. \end{cases} \quad (6.38)$$

y

$$v_n = \mathbb{E}[(X_{n+1} - P_n X_{n+1})^2] \quad (6.39)$$

Además se introduce el concepto de *innovaciones* o errores de predicción a un paso como:

$$U_n = X_n - \hat{X}_n \quad (6.40)$$

Las ecuaciones anteriores en términos de $\mathbf{U} = (U_1, \dots, U_n)^T$ y $\mathbf{X} = (X_1, \dots, X_n)^T$ pueden ser escritas de la siguiente manera:

$$\mathbf{U}_n = A_n \mathbf{X}_n \quad (6.41)$$

Donde A_n están dadas por:

$$A_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{11} & 1 & 0 & \cdots & 0 \\ a_{22} & a_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ a_{n-1,n-1} & a_{n-1,n-2} & a_{n-1,n-3} & \cdots & 1 \end{bmatrix} \quad (6.42)$$

Si el proceso X_t es un proceso estacionario, entonces los términos $a_{ij} = a_j$ correspondientes a la ecuación 6.21 con $\tau = 1$. La matriz inversa de A_n denotada por \mathbb{C}_n se encuentra dada por:

$$\mathbb{C}_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \theta_{11} & 1 & 0 & \cdots & 0 \\ \theta_{22} & \theta_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \theta_{n-1,n-1} & \theta_{n-1,n-2} & \theta_{n-1,n-3} & \cdots & 1 \end{bmatrix} \quad (6.43)$$

Entonces el vector de predictores a un paso $\hat{\mathbf{X}}_n := (X_1, P_1 X_2, \dots, P_{n-1} X_n)^T$ se puede expresar como:

$$\begin{aligned} \hat{\mathbf{X}}_n &= \mathbf{X}_n - \mathbf{U}_n \\ &= \mathbb{C}_n \mathbf{U}_n - \mathbf{U}_n \\ &= \Theta_n (\mathbf{X}_n - \hat{\mathbf{X}}_n) \end{aligned} \quad (6.44)$$

donde

$$\Theta_n = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ \theta_{11} & 0 & 0 & \cdots & 0 \\ \theta_{22} & \theta_{21} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \theta_{n-1,n-1} & \theta_{n-1,n-2} & \theta_{n-1,n-3} & \cdots & 0 \end{bmatrix} \quad (6.45)$$

Y \mathbf{X}_n satisface por sí misma $\mathbf{X}_n = \mathbb{C}_n (\mathbf{X}_n - \hat{\mathbf{X}}_n)$.

La ecuación 6.44 puede escribirse de la siguiente manera:

$$\hat{X}_{n+1} = \begin{cases} 0 & \text{si } n=0, \\ \sum_{j=1}^n \theta_{nj}(\hat{X}_{n+1-j} - X_{n+1-j}) & \text{si } n=1,2,\dots \end{cases} \quad (6.46)$$

A partir de esta ecuación es posible calcular los predictores a un paso $\hat{X}_1, \hat{X}_2, \dots$ de manera recursiva una vez que los coeficientes θ_{ij} han sido calculados.

De acuerdo a Brockwell [9] los coeficientes $\theta_{n1}, \dots, \theta_{nn}$ pueden ser calculados recursivamente de la siguiente manera:

$$\theta_{n,n-k} = v_k^{-1} \left(\kappa(n+1, k+1) - \sum_{j=0}^{k-1} \theta_{k,k-j} \theta_{n,n-j} v_j \right) \quad 0 \leq k < n \quad (6.47)$$

Donde

$$v_n = \kappa(n+1, n+1) - \sum_{j=0}^{n-1} \theta_{n,n-j}^2 v_j \quad (6.48)$$

$$v_0 = \kappa(1, 1)$$

Los algoritmos vistos en esta sección sirven exclusivamente para procesos cuadrado integrables. El algoritmo de innovación se aplica, generalmente, a los procesos $MA(q)$ mientras que el algoritmo de Durbin-Levinson se aplica a los procesos $AR(p)$. Sin embargo, para realizar pronósticos a modelos ARMA es necesario estudiar otros métodos, que se tratarán de presentar en la siguiente sección.

6.1.2. Pronósticos en Modelos ARMA

Para el modelo ARMA causal:

$$\phi(B)X_t = \theta(B)\epsilon_t,$$

Brockwell [9] menciona un método que propuesto por Ansley [1], que se basa en la aplicación del algoritmo de innovación, no al proceso $\{X_t\}$ en sí, sino una transformación de este, denotado por $\{W_t\}$, dada por:

$$W_t = \begin{cases} \sigma^{-1} X_t & \text{para } 1 \leq t \leq m \\ \sigma^{-1} \phi(B) X_t & \text{para } t > m \end{cases} \quad (6.49)$$

Donde $m = \max(p, q)$ y σ^{-1} es el inverso de la desviación estándar asociada al ruido blanco del modelo.

Por conveniencia se asume, $\theta_0 = 1$ y $\theta_j = 0$ para $j > q$. Además se asume, sin pérdida de generalidad, que $p, q \geq 1$.

La función de autocovarianzas γ_t de $\{X_t\}_t$ se calcula de la misma manera que en las secciones anteriores, mientras que las autocovarianzas $\kappa(i, j) = \mathbb{E}[W_i, W_j]$ se pueden encontrar de la siguiente manera:

$$\kappa(i, j) = \begin{cases} \sigma^{-2}\gamma(i-j) & \text{si } 1 \leq i, j \leq m. \\ \sigma^{-2}\left[\gamma(i-j) - \sum_{r=1}^p \phi_r \gamma(r - |i-j|)\right] & \text{si } \min(i, j) \leq m \leq \max(i, j) \leq 2m \\ \sum_{r=1}^q \theta_r \theta_{r+|i-j|} & \text{si } \min(i, j) > m \\ 0 & \text{En otro caso.} \end{cases} \quad (6.50)$$

Si se le aplica el algoritmo de innovación al proceso $\{W_t\}$ se obtiene:

$$\hat{W}_t = \begin{cases} \sum_{j=1}^n \theta_{nj}(W_{n+1-j} - \hat{W}_{n+1-j}) & 1 \leq n < m \\ \sum_{j=1}^q \theta_{nj}(W_{n+1-j} - \hat{W}_{n+1-j}) & n \geq m \end{cases} \quad (6.51)$$

Donde θ_{nj} puede ser encontrada por medio del algoritmo de innovación con κ definida como en la ecuación 6.50.

Brockwell también señala que la ecuación 6.49 permite a X_n , $n \geq 1$, ser escrita como una combinación lineal de W_j , donde $1 \geq j \geq n$, inversamente cada W_n , $n \geq 1$, puede ser escrita como una combinación lineal de X_j , donde $1 \geq j \geq n$. Esto significa que el mejor predictor lineal de cualquier variable aleatoria Y en términos de $\{1, X_1, X_2, \dots, X_n\}$ es el mismo predictor de Y en términos de $\{1, W_1, W_2, \dots, W_n\}$. Se denota a este predictor como $P_n Y$. En particular los predictores a un paso de X_n y W_n son respectivamente:

$$\hat{X}_{n+1} = P_n X_{n+1}$$

Y

$$\hat{W}_{n+1} = P_n W_{n+1}$$

Usando la linealidad de P_n y la ecuación 6.49 se obtiene lo siguiente:

$$\hat{W}_t = \begin{cases} \sigma^{-1} \hat{X}_t & \text{para } 1 \leq t \leq m \\ \sigma^{-1} \left[\hat{X}_t - \theta_1 X_{t-1} - \dots - \theta_p X_{t-p} \right] & \text{para } t > m \end{cases} \quad (6.52)$$

De esta de ecuación, realizando el despeje correspondientes se obtiene:

$$X_t - \hat{X}_t = \sigma \left[W_t - \hat{W}_t \right] \quad \forall t \geq 1. \quad (6.53)$$

Por lo tanto $(W_t - \hat{W}_t) = \sigma^{-1}(X_t - \hat{X}_t)$. Brockwell indica que si se reemplaza esta expresión en la ecuación 6.50 y después en 6.51 finalmente se obtiene la siguiente expresión:

$$\hat{X}_{t+1} = \begin{cases} \sum_{j=1}^n \theta_{nj} (X_{n+1-j} - \hat{X}_{n+1-j}) & \text{si } 1 \geq n < m \\ \phi_1 X_n + \dots + \phi_p X_{n+1-p} + \sum_{j=1}^q (X_{n+1-j} - \hat{X}_{n+1-j}) & \text{si } n \geq m. \end{cases} \quad (6.54)$$

Y

$$\mathbb{E}[(X_{n+1} - \hat{X}_{n+1})^2] = \sigma^2 \mathbb{E}[(W_{n+1} - \hat{W}_{n+1})^2] = \sigma^2 r_n \quad (6.55)$$

Donde θ_{nj} y r_n son encontradas mediante el algoritmo de innovación con κ definida en 6.50. La ecuación 6.54 determina los predictores a un paso, $\hat{X}_2, \hat{X}_3, \dots$, de manera recursiva.

6.2. Estimación de parámetros

Existen diversas maneras de obtener los estimadores de los parámetros del modelo ARMA, en el caso de este trabajo, debido a sus sencillas propiedades y simplicidad, se utilizará el método de máxima verosimilitud condicional.

La etapa de estimación supone que ya se ha identificado el modelo ARMA(p,q) asociado a la serie observada.

Se asume que que el ruido blanco es gaussiano, es decir $\epsilon_t \sim N(0, \sigma_\epsilon^2)$. Entonces, sea $\{Y_t\} = \{X_t - \mu\}$ un proceso ARMA(p,q) centrado con $\{\epsilon_t\}$ ruido blanco gaussiano:

$$Y_t = \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} \quad (6.56)$$

La función de verosimilitud del conjunto de observaciones $y = \{y_1, y_2, \dots, y_n\}$ para un proceso (cualquiera) se encuentra dada por:

$$\mathbb{L}(y|\mathcal{S}) = f_{Y_1, Y_2, \dots, Y_n}(y_1, y_2, \dots, y_n; \mathcal{S}) \quad (6.57)$$

Donde $\mathcal{S} = (\mu, \sigma_\epsilon^2, \phi, \theta)$, además $\phi = \{\phi_1, \dots, \phi_p\}$ y $\theta = \{\theta_1, \dots, \theta_q\}$. Esta ecuación representa la probabilidad de que las variables aleatorias $\{Y_1, Y_2, \dots, Y_n\}$ tomen los respectivos valores que se obtuvieron de las observaciones, es decir, que tomen los siguientes valores respectivamente y_1, y_2, \dots, y_n . Por tanto, el objetivo de esta idea es maximizar esta función, para ello se reexpresa la función condicionando sobre sus valores tomados en el pasado. Primero de utiliza la definición de densidad condicional² sobre la densidad conjunta de las Y_i como se muestra a continuación:

²Ver apéndice.

$$\begin{aligned}
f_{Y_1, Y_2, \dots, Y_n}(y_1, y_2, \dots, y_n; \mathcal{S}) &= f_{Y_n|Y_{n-1}}(y_n|y_{n-1}; \mathcal{S}) f_{Y_1, Y_2, \dots, Y_{n-1}}(y_1, y_2, \dots, y_{n-1}; \mathcal{S}) \\
&= \left[f_{Y_n|Y_{n-1}}(y_n|y_{n-1}; \mathcal{S}) \right] \left[\prod_{j=2}^n f_{Y_j|Y_{j-1}}(y_j|y_{j-1}; \mathcal{S}) \right] \quad (6.58)
\end{aligned}$$

Entonces si a esta función se le aplica el logaritmo se obtiene la función de log-verosimilitud dada por:

$$\ln \{\mathbb{L}(\mathcal{S})\} = \log \left[f_{Y_1}(y_1|\mathcal{S}) \right] + \sum_{j=2}^n \log \left[f_{Y_j|Y_{j-1}}(y_j|y_{j-1}; \mathcal{S}) \right] \quad (6.59)$$

Es posible describir la ecuación 6.56 de la siguiente manera:

$$\epsilon_t = \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} + Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} \quad (6.60)$$

Sea $Y = (Y_1, \dots, Y_n)^T$, se sume como condición inicial a $Y^* = (Y_{1-p}, \dots, Y_{-1}, Y_0)^T$ y $\epsilon^* = (\epsilon_{1-q}, \dots, \epsilon_{-1}, \epsilon_0)^T$ conocidas.

Retomando al ruido blanco, la función de densidad conjunta de $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_n)^T$, se puede obtener aplicando la ecuación de verosimilitud 6.58 al ruido blanco 6.60, que de acuerdo a lo asumido tiene una distribución Normal con media cero y varianza σ_ϵ^2 ³. Por lo tanto está se encuentra dada por:

$$\begin{aligned}
\mathbb{L}(\epsilon|\mathcal{S}) &= \left[(2\pi\sigma_\epsilon^2)^{-1/2} \exp\left[-\frac{\epsilon_t^2}{2\sigma_\epsilon^2}\right] \right] \left[\prod_{j=2}^n (2\pi\sigma_\epsilon^2)^{-1/2} \exp\left[-\frac{\epsilon_j^2}{2\sigma_\epsilon^2}\right] \right] \\
&= (2\pi\sigma_\epsilon^2)^{-n/2} \exp\left[-\frac{1}{2\sigma_\epsilon^2} \sum_{t=1}^n \epsilon_t^2\right] \quad (6.61)
\end{aligned}$$

Y por lo tanto la función de log-verosimilitud para este caso particular queda de la siguiente manera:

$$\ln \{\mathbb{L}(\epsilon|\mathcal{S})\} = -\frac{n}{2} \log[2\pi\sigma_\epsilon^2] - \sum_{j=1}^n \frac{\epsilon_j^2}{2\sigma_\epsilon^2} \quad (6.62)$$

Una vez que se tiene esta expresión lo que se debe hacer es derivar esta expresión con respecto al parámetro del que se desea encontrar el estimador, para después igualar a cero la derivada hallada y despejar al estimador.

³Su función de densidad es $f(\epsilon_t) = (2\pi\sigma_\epsilon^2)^{-1/2} \exp\left[-\frac{\epsilon_t^2}{2\sigma_\epsilon^2}\right]$

Los estimadores encontrados $\hat{\phi}$, $\hat{\theta}$ y $\hat{\mu}$ son tales que maximizan a la función 6.62 y son conocidos como los estimadores de máxima verosimilitud condicionales. Como $\mathbb{L}(\epsilon|\mathcal{S})$ involucra datos solamente a través de $\sum_{j=1}^n \epsilon_n^2$, estos estimadores resultan ser los mismos que los estimadores por mínimos cuadrados condicionales que se obtienen al minimizar la suma de residuos cuadrados $\sum_{j=1}^n \epsilon_n^2$, la cual no contiene al parámetro σ_ϵ^2 .

Existen diversas alternativas para elegir las condiciones iniciales Y^* y ϵ^* . De acuerdo a Wei [34] si se asume que $\{x_t\}$ es un proceso estacionario y $\{\epsilon_t\}$ es un ruido blanco gaussiano, se puede remplazar a x_t por la media muestral \bar{x} y a ϵ_t por su valor esperado de cero.

Para el modelo 6.56 se podría asumir que $\epsilon_p = \epsilon_{p-1} = \epsilon_{p-2} = \dots = \epsilon_{p-q+1} = 0$ y calcular ϵ_t para $t \geq (p+1)$ utilizando la ecuación 6.60. Bajo estas condiciones sobre el error y con un proceso similar utilizado anteriormente la log-verosimilitud condicionada estará dada por

$$\ln \{\mathbb{L}(\mathcal{S})\} = -\frac{n}{2} \log[2\pi\sigma_\epsilon^2] - \sum_{j=p+1}^n \frac{\epsilon_n^2}{2\sigma_\epsilon^2} \quad (6.63)$$

Esta forma es la forma en que muchos programas computacionales realizan el cálculo.

Luego de obtener los estimadores $\hat{\phi}$, $\hat{\theta}$ y $\hat{\mu}$, se puede obtener el estimador $\hat{\sigma}_\epsilon^2$ diferenciando a 6.63 con respecto a σ_ϵ^2 e igualando a cero, para finalmente obtener:

$$\hat{\sigma}_\epsilon^2 = \frac{\sum_{j=p+1}^n \epsilon_n^2}{df} \quad (6.64)$$

Donde df son los grados de libertad, que es igual al número de términos utilizados en la suma de residuos menos el número de parámetros estimados.

6.2.1. Validación del modelo

Por último, una vez ajustado el modelo, hay que proceder a la validación del mismo. Se debe validar los supuestos acerca del comportamiento de los residuales. Para la estimación bajo máxima verosimilitud, se asumió que el proceso $\{\epsilon_t\}$ era un ruido blanco gaussiano, es decir tiene una distribución normal con media cero y varianza constante.

Sea \hat{X}_t el valor estimador vía máxima verosimilitud basado en las observaciones al tiempo $t-1$, entonces un estimador de ϵ_t es $\hat{\epsilon}_t = X_t - \hat{X}_t$. Este se puede estandarizar, utilizando el algoritmo de innovación[23] pues a partir de este, se obtiene el error cuadrático medio al tiempo i , $r_i = \mathbb{E}[(X_t - \hat{X}_t)^2]$ y entonces se puede definir un segundo tipo de residuo se la siguiente manera:

$$\hat{W}_t = \frac{X_t - \hat{X}_t}{\sqrt{r_t}} \quad (6.65)$$

Entonces el proceso $\{\hat{W}_t\} \sim N(0, \sigma_\epsilon^2)$, es decir se comporta como ruido blanco. Con el fin de facilitar las validaciones de puede definir a los residuales re-escalados:

$$\hat{R}_t = \frac{\hat{W}_t}{\hat{\sigma}_t} \quad (6.66)$$

$$\text{Donde } \hat{\sigma}_t = \sqrt{\sum_{t=1}^n W_t^2 / n}.$$

De manera visual se puede identificar si los residuos ajustan a una distribución normal si se realiza un QQ-plot. Esta gráfica muestra los cuantiles del conjunto de datos contra los cuantiles teóricos de una distribución normal, si los puntos se apegan a la recta identidad, entonces es señal que los residuos se asemejan a los residuos teóricos.

Para validar que la media de los residuos, ya sea del proceso $\{\hat{W}_t\}$ o $\{\hat{X}_t\}$, es cero, entonces es posible realizar una prueba de hipótesis $H_0 : \mu = 0$ vs $H_1 : \mu \neq 0$, tomando estadístico de prueba T , definido como $T = \frac{\hat{\mu}}{\sqrt{\hat{\sigma}^2}}$. Donde $\hat{\mu}$ y σ^2 son la media y varianza muestrales de los residuos \hat{W}_t o \hat{X}_t . El estadístico T tiene una distribución t-student con $(n - 1)$ grados de libertad.

Para validar el supuesto de normalidad se pueden realizar distintas pruebas de bondad de ajuste, entre las más famosas están prueba de Kolmogorov-Smirnov o Anderson-Darling.

La primer prueba se basa en la función de distribución empírica de la muestra observada. Se toma como estadístico de prueba a:

$$D_n = \sup_{x \in \mathbb{R}} \{|F_n(x) - F_*(x)|\} \quad (6.67)$$

Donde F_n es la función empírica de la muestra mientras que F_* es la función teórica a ajustar (para este caso normal), posteriormente mediante el uso de tablas o en su caso simulando la distribución se encuentra el cuantil al $(1 - \alpha)\%$ y se compara contra el estadístico. La prueba de hipótesis es $H_0 : F_n = F_*$ vs $H_1 : F_n \neq F_*$, es decir, se valida que la muestra proviene efectivamente de la distribución ajustar. Se decide no rechazar H_0 si el estadístico es en términos absolutos más chico que el cuantil de la distribución de Kolmogorov-Smirnov bajo H_0 .

La prueba de Anderson-Darling es otra opción como prueba de bondad de ajuste. El estadístico de prueba para esta prueba es:

$$A_n^2 = -n - \sum_{i=1}^n \left(\frac{2i-1}{n} \right) \left(\ln(F(x_{(i)})) + \ln(1 - F(x_{(n-1+1)})) \right) \quad (6.68)$$

La prueba de hipótesis es la misma que en el caso de la prueba de Kolmogorov, la regla de decisión es rechazar H_0 si $A_n^2 > w_{1-\alpha}$, donde $w_{1-\alpha}$ es el cuantil asociado a la distribución Anderson-Darling bajo H_0 .

Finalmente, para ver si los errores se encuentran no correlacionados, hasta en un número k de observaciones, se utiliza la prueba de Ljung-Box, mediante la prueba $H_0 : \rho(0) =$

$\rho(1) = \dots = \rho(k) = 0$ vs $H_1 : \exists \rho(i)$ tal que $\rho(i) \neq 0$ p.a. $i \in \{1, \dots, n\}$. El estadístico de prueba esta dado por:

$$Q(k) = n(n+2) \sum_{j=1}^k \frac{\hat{\rho}(j)_\epsilon^2}{n-j} \quad (6.69)$$

Donde $\hat{\rho}_\epsilon^2$ es la correlación entre los residuos y n es el número de observaciones. El estadístico Q tiene una distribución ji-cuadrada con $k - p - q$ grados de libertad, por lo que H_0 no se rechaza si el estadístico es menor que el cuantil $(1 - \alpha)\%$ de la distribución ji-cuadrada con $k - p - q$ grados de libertad.

En la práctica, la elección de k para la prueba estadística, puede afectar el desempeño del estadístico $Q(k)$, por lo que de acuerdo a Tsay[32] se utilizan diversos valores de k , más aún, menciona que diversos estudios sugieren como elección a $k \approx \ln(n)$ proporcionan el mejor desempeño.

6.2.2. Modelos ARIMA y SARIMA

En esta sección se examina el problema de ajustar un modelo adecuado a un conjunto de datos que no necesariamente provienen de una serie de tiempo estacionaria. Cuando los datos presentan estacionariedad y su función de autocovarianzas decae rápidamente en general se puede ajustar un modelo ARMA. Sin embargo, si presentan desviaciones cada cierto número de periodos es posible adaptar un modelo a los datos observados mediante diferenciación.

Una generalización de los modelos $ARMA(p, q)$, como fue mencionado al principio del capítulo anterior, son modelos autorregresivos e integrados de promedios móviles o modelos $ARIMA(p, d, q)$. Este es un modelo para procesos no estacionarios, pero que luego de aplicar diferencias en la serie logra comportarse como una serie estacionaria.

Definición 6.3 *Un proceso $\{X_t\}$ es un proceso autorregresivo integrado y de medias móviles, $ARIMA(p, d, q)$ si puede expresarse de la siguiente manera:*

$$\phi(B)\nabla^d X_t = \phi(B)(1 - B)^d X_t = \theta(B)\epsilon_t \quad (6.70)$$

Donde $\{\epsilon_t\}$ es un ruido blanco. Cabe notar, que la expresión $\phi(B)\nabla^d X_t = \theta(B)\epsilon_t$ indica que el proceso que se obtiene de diferenciar d veces al proceso $\{x_t\}$, $\{Z_t\} = \{\nabla^d X_t\}$, se puede modelar como un modelo $ARMA(p, q)$.

Retomando lo anteriormente comentado, el orden de diferenciación se puede obtener a partir de la sucesión de autocorrelaciones muestrales (ACF), teniendo en cuenta que un decaimiento rápido es indicativo de un proceso estacionario.

Wei [34], sugiere una serie de pasos para la identificación del modelo, dada una serie de tiempo.

1. Graficar la serie de tiempo y elegir la transformaciones más adecuadas. Es a través de un primer escaneo visualizando el comportamiento de serie de tiempo que se puede apreciar si el proceso presenta tendencia, estacionalidad, varianza inconsistente o outliers. Debido a que muchos métodos para estabilizar la varianza requieren valores no negativos en la serie, y la diferenciación provoca estos, se recomienda primero estabilizar la varianza antes de diferenciar para eliminar tendencias.
2. Calcular y analizar los ACF y PACF de la serie original con el fin de identificar si se requiere algún orden de diferenciación. Si la ACF muestral decae lentamente y además la PACF se interrumpe después del retraso 1, es indicativo que la serie necesita ser diferenciada. También puede ser utilizada la prueba de Dickey-Fuller o Dickey-Fuller Aumentada para saber si se requiere diferenciar una vez más.
3. Calcular y analizar los ACF y PACF de la serie transformada o diferenciada para identificar los ordenes de p y q . Usualmente el orden de p y q son menores a tres. Para la asignación de un modelo ARIMA razonable se recomienda un mínimo de 50 observaciones.

Antes de concluir este capítulo se debe mencionar que existen otros modelos más generales de series de tiempo que son los modelos estacionales autorregresivos integrados de medias móviles, mejor conocido como modelos *SARIMA*. Este tipo de modelos contemplan la existencia de componente estacional en la serie de tiempo.

Definición 6.4 Sea $\{X_t\}$ un proceso, entonces se dice que $\{X_t\}$ es un modelo *SARIMA*(p, d, q) \times (P, D, Q) $_s$ con periodos s si la serie diferenciada $Y_t = (\nabla_s)^D(\nabla^d X_t) = (1 - B^s)^D(1 - B)^d X_t$ es un proceso ARMA de la forma:

$$\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)\epsilon_t \quad (6.71)$$

Donde s representa el número de periodos de la serie, es decir si la serie es mensual $s = 12$ o si la serie es semestral entonces $s = 2$. Además $\phi(B)$ y $\theta(B)$ están definidos como antes, mientras que $\Phi(B^s)$ y $\Theta(B^s)$ se definen a continuación :

$$\begin{aligned} \Phi(B^s) &= 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps} \\ \Theta(B^s) &= 1 + \Theta_1 B^s - \Theta_2 B^{2s} + \dots + \Theta_Q B^{Qs} \end{aligned} \quad (6.72)$$

Otra forma de ver un proceso *SARIMA* en términos del proceso original X_t , es si satisface:

$$\phi(B)\Phi(B^s)(\nabla_s)^D(\nabla^d X_t) = \theta(B)\Theta(B^s)\epsilon_t \quad (6.73)$$

En esta última sección se menciona de manera breve estas dos generalizaciones de los modelo ARMA, sin embargo profundizar en estos modelos implicaría desviarse del objeto de este trabajo, que es el estudio y estimación del CaR como medida del costo financiero del país. Es por ello que, a pesar del gran interés que generan estos modelos, ARIMA y SARIMA y otros modelos con los cuales es posible describir y predecir fenómenos financieros como los modelos ARCH y GARCH, que suponen heterocedasticidad en la volatilidad, que solamente

se tocaron de manera superficial o en su caso solo se hicieron mención, pero no por ello son toda esta gama de modelos menos importantes.

A continuación se mostrará lo visto acerca de estos modelos de tasas de interés y series de tiempo implementado para la estimación del CaR en la deuda pública, además de un modelo no paramétrico con el que se comparará y se realizará un análisis de la información generada, así como de las ventajas y desventajas de los modelos presentados.

Capítulo 7

Estimación del CaR

La estimación del costo en riesgo se puede realizar para un determinado número de periodos futuros. Debido a que la información con la que se cuenta es hasta noviembre de 2018 y se desea realizar la estimación para el año 2019, se realizará una estimación de 13¹ periodos (diciembre de 2018 y 12 meses para 2019), lo cual implica que se considerará información mensual para los cálculos a realizar.

El primer costo en riesgo a estimar es el generado por moneda extranjera, de manera muy general, los pasos a realizar son los siguientes:

1. Se calcula la matriz del costo en riesgo que implican los bonos a tasa fija en moneda extranjera.²
2. Se ajusta un modelo ARMA³ a la serie de tiempo generada por cada tipo de cambio (dólar americano, euro, yen, dólar canadiense y libra).
3. Se realizan simulaciones del costo financiero ajustado por la predicción del tipo de cambio para cada mes.
4. Con los pronósticos se realizan las proyecciones del costo financiero en moneda extranjera.
5. Se estima el costo financiero de la deuda a emitirse de todos los instrumentos.⁴
6. Se calcula el CaR al 95 % del costo total por moneda extranjera para el año 2019.

Para su estimación en moneda local, los pasos a realizar son los siguientes:

¹Cabe mencionar, que si la estimación se requiere para cualquier otro número de periodos, únicamente se deberá cambiar las dimensiones de las matrices que se requieren para correr el código (solo se deben cambiar las matrices de calendarios) y el valor de la variable pred (variable que contiene el número de periodos a pronosticar), y de esta forma del código funcionará sin problema. Más adelante se describirán las matrices (inputs) que son necesarios para poder correr el código para la estimación.

²Este dato se debe obtener de la proyección de intereses proporcionado por la SHCP, debido a que esa información no esta disponible se utiliza la información proporcionada de años anteriores

³En caso de ajustar un ARIMA(p,d,q) se toma el modelo ARMA(p,q) de la serie diferenciada d veces, es por eso que se dice que se toma un ARMA

⁴La deuda a emitirse se debe obtener de la programación de deuda la SHCP, debido a que esa información no esta disponible se utiliza la información del resultado del año anterior proporcionada por el Banco de México. El resultado de las subastas realizadas se pueden encontrar en <http://www.banxico.org.mx/SieInternet/>

1. Se calculan las estimaciones de los parámetros del modelo CIR para cada instrumento que se consideró como formador de la deuda pública en pesos (Cetes, Bonos a Tasa Fija, Udibonos).
2. Se recrean las tasas de interés de acuerdo al modelo de CIR, para recrear posibles escenarios del costo financiero de la deuda.
3. Se estima el costo financiero de la deuda emitida (solo es necesaria la de Udibonos debido a que es a tasa variable y ya se conoce el costo financiero de los demás instrumentos) proyectando el valor de la deuda con las tasas pronosticadas.
4. De la misma manera que en el caso anterior, se estima el costo financiero de la deuda a emitirse de todos los instrumentos.⁵
5. Se estima el costo financiero total.
6. Se calcula el CaR al 95 % del costo total para la deuda interna para el año 2019 .

En las siguientes dos secciones se describirá, de manera más detallada, los pasos completos a seguir para poder realizar las estimación del costos en riesgo. Así mismo, se presentarán los códigos realizados en Matlab, version 2018, utilizados para la estimación en los cuales se puede apreciar de mejor manera la implementación de los modelos tratados durante este trabajo.

7.1. CaR de Tipo de Cambio

En esta sección se describirá paso a paso la manera en que el costo en riesgo del tipo de cambio, es decir aquel que mide el costo financiero de la deuda externa, es estimado.

Como se ha mencionado antes, para su estimación se ajustará un modelo ARIMA a cada serie de tipo de cambio, con el fin de realizar un pronósticos de la paridad cambiaria. Además se utilizará simulación Monte Carlo⁶ de 10,000 escenarios de tipos de cambio para poder proporcionar el valor del CaR a un nivel $(1 - \alpha)$ % de confianza.

Los primero insumos que se utilizarán son los tipos de cambio, debido a que la estimación del costo financiero que realiza la SHCP es mensual, entonces se utilizarán promedios mensuales (ya que los tipo de cambio se obtienen diariamente) del tipo de cambio de los últimos K -meses (en este caso se tomaran los últimos 92 meses, i.e $K = 92$).

Los tipos de cambio fueron descargados de un proveedor de índices y precios⁷ de los últimos K -meses y posteriormente se realizó el cálculo del promedio mensual y se acomodaron en una matriz., ver figura 7.1.

De acuerdo a la proyección mensual del costo financiero de la deuda emita en moneda extranjera, los tipos de cambio a utilizar serán: el dólar norteamericano (USD), el euro (EUR), el yen japonés (JPY), el dólar canadiense (CAD) y la libra británica (GBP).

⁵La deuda a emitirse se debe obtener de la programación de deuda la SHCP, debido a que esa información no esta disponible se utiliza la información del resultado del año anterior proporcionada por el Banco de México. El resultado de las subastas realizadas se pueden encontrar en <http://www.banxico.org.mx/SieInternet/>

⁶Técnica de simulación numérica que sirve para generar variables aleatorias y evaluar la incertidumbre en sistemas complejos.[17]

⁷Investing.com, sitio web: <https://mx.investing.com/> .

	USD	EUR	JPY	CAD	GBP
1	13.70422	18.57794	0.1767205	13.36274	21.65695
2	13.76591	18.11882	0.1768659	13.45237	21.46867
3	13.42325	17.31428	0.1743477	13.23509	20.81197
4	12.78741	16.91761	0.1629476	12.81922	20.20425
5	12.74233	16.82852	0.1544659	12.82301	20.15946
6	13.05429	17.17608	0.1604548	13.14186	20.88332
7	13.62500	17.43440	0.1708348	13.49285	21.67937
8	13.92825	17.45940	0.1756024	13.54414	21.64944
9	13.36353	16.41835	0.1691477	13.16756	20.82903
10	13.19227	16.35399	0.1676391	13.27947	20.72764
11	12.94606	16.64530	0.1655350	13.22116	20.84795
12	12.89802	16.72002	0.1632739	13.06745	20.72771
13	13.07296	16.77472	0.1614568	13.11008	20.86657
14	12.86858	16.88224	0.1535452	12.99685	20.76057
15	12.70570	16.88222	0.1428522	12.80309	20.29672

Figura 7.1: Fragmento de la matriz de tipo de cambio. De izquierda a derecha: Dólar Americano, Euro, Yen, Dólar Canadiense, Libra

El siguiente paso fue ajustar un modelo a $ARIMA(p, d, q)$ a la series mensuales de los tipos de cambio. La manera en que se realizó el ajuste se mostrará en la siguiente sección, antes de presentar el código. El modelo se ajusta a cada tipo de cambio de manera individual, ya que hacerlo conjuntamente implica un modelo diferente que no se realiza en la SHCP para su estimación, como los modelos VAR, SVAR o SVARMA cuyo alcance se encuentra fuera de este trabajo[34].

Una vez que se ha ajustado el modelo a cada serie de tiempo, lo siguiente es acomodar los parámetros en una matriz. Esta contendrá en su primera fila los parámetros μ de la distintas series, la segunda fila los parámetros ϕ_1 , mientras que la tercer fila aquellos correspondientes a ϕ_2 y así sucesivamente con los parámetros autorregresivos. Una vez que se han colocado en el modelo los parámetros ϕ , se colocan los parámetros correspondientes a la parte de medias móviles, $\{\theta\}_{j=1}^q$, de cada serie de tiempo y en la última fila la desviación estándar de los errores asociados al ruido blanco.

A este segundo input se le llamará la matriz ARIMA, la cual de acuerdo a la descripción realizada en el párrafo anterior tiene la siguiente estructura:

$$\begin{bmatrix}
 \mu_1 & \mu_2 & \mu_3 & \mu_4 & \mu_5 \\
 \phi_{1,1} & \phi_{2,1} & \phi_{3,1} & \phi_{4,1} & \phi_{5,1} \\
 \phi_{1,2} & \phi_{2,2} & \phi_{3,2} & \phi_{4,2} & \phi_{5,2} \\
 \vdots & \vdots & \vdots & \vdots & \vdots \\
 \phi_{1,p} & \phi_{2,p} & \phi_{3,p} & \phi_{4,p} & \phi_{5,p} \\
 \theta_{1,1} & \theta_{2,1} & \theta_{3,1} & \theta_{4,1} & \theta_{5,1} \\
 \theta_{1,2} & \theta_{2,2} & \theta_{3,2} & \theta_{4,2} & \theta_{5,2} \\
 \vdots & \vdots & \vdots & \vdots & \vdots \\
 \theta_{1,q} & \theta_{2,q} & \theta_{3,q} & \theta_{4,q} & \theta_{5,q} \\
 \sigma_1 & \sigma_2 & \sigma_3 & \sigma_4 & \sigma_5
 \end{bmatrix} \quad (7.1)$$

Donde $\phi_{i,j}$ se refiere al j -ésimo parámetro de la parte autorregresiva de la serie i . De la misma manera para el conjunto de parámetros de medias móviles $\theta_{i,j}$. En el caso de este ejercicio, se consideran solo las 5 series de tiempo que contempla el modelo (USD, EUR, JPY, CAD, GBP), el orden p y q que se utilizarán en la matriz será el máximo de los ordenes individuales de las p_i y q_i , es decir, $p = \max [p_i]$ y $q = \max [q_i]$.

Cabe señalar que debido a que las distintas series no tienen los mismos ordenes y tomando en cuenta la elección de p y q en la matriz, se consideraran cero los parámetros dentro de la matriz, de aquellas series que tengan orden mas pequeño.

El tercer input (en caso de ser necesario) es la matriz de tipo de cambio diferenciada d veces de acuerdo al modelo ARIMA(p,d,q) ajustado, ya que en el código que estima el CaR, los pronósticos se realizan de acuerdo a la parte autorregresiva y medias móviles de una serie por lo que debe recibir la serie ya diferenciada.

El último input, es una matriz la cual contiene el costo por interés de cada mes, de los instrumentos a tasa fija en moneda extranjera, ver imagen 7.2⁸, este simplemente son los intereses de cada emisión que aún se encuentran en activo que pagan cupón durante el año en curso (en este caso que pagan cupón durante el año 2019) en moneda extranjera, por lo que se debe pronosticar el tipo de cambio para cada mes con el fin de poder valorizar la deuda en pesos mexicanos.

calen	1	2	3	4	5
1	0	0	0	0	0
2	0	0	0	0	0
3	4.6973e+08	0	2.9696e+08	0	0
4	9.6241e+07	2.6731e+07	5.1763e+08	0	3.2291e+07
5	4.9243e+08	3.2223e+05	6.5902e+06	70.32	5.6467e+07
6	3.395e+08	1.0604e+08	1.1507e+09	0	0
7	1.5004e+08	1.8801e+07	5.0989e+06	0	0
8	1.6513e+08	3.1528e+07	2.2239e+09	0	0
9	5.1698e+08	2.746e+07	4.9716e+08	0	0
10	9.4792e+07	4.2416e+06	5.1763e+08	0	0
11	4.6393e+08	3.0018e+05	6.6994e+06	56.25	0
12	4.256e+08	73996	1.148e+09	0	0
13	1.5052e+08	0	2.5879e+06	0	0
14	1.6074e+08	8.0269e+06	1.9005e+09	0	0

Figura 7.2: Intereses por pagar en unidades de moneda extranjera. De izquierda a derecha: USD, EUR, JPY, CAD, GBP

Una vez que se cuenta con los inputs iniciales, el algoritmo a seguir para poder realizar la estimación del CaR es el siguiente:

1. Crear una matriz para los tipos de cambio futuros (matriz $tcfd$), la cual contendrá los pronósticos de las tasas de conversión que se obtuvieron a partir del ajuste ARIMA realizado previamente, para los meses necesarios para la estimación, en este caso para 13 meses (diciembre de 2018 y 2019).

⁸La deuda a emitirse se debe obtener de la programación de deuda la SHCP, debido a que esa información no esta disponible se utiliza la información del resultado del año anterior proporcionada por el Banco de México. El resultado de las subastas realizadas se pueden encontrar en <http://www.banxico.org.mx/SieInternet/>

2. Se realizan los pronósticos de la serie diferenciada de acuerdo al modelo ARIMA ajustado, esto debido a que por la manera en que se realiza el pronóstico, solo se realiza a la parte autorregresiva y de medias móviles de la serie.
3. Para que el pronóstico sea sobre la serie original, las series que contienen los pronósticos de las series diferenciadas (matriz $tcfd$) deben ser regresadas a sus valores originales, para ello se utilizó una función que utiliza la serie diferenciada y un valor inicial para que pueda reconstruir la serie original y a su vez la parte pronosticada como parte de serie original. El código de la función para regresar a la serie original se presentará en la siguiente sección.
4. Crear una matriz de ceros, la cual contendrá el ruido blanco para los pronósticos. Los errores asociados al ruido blanco son calculados como la simulación de una variable aleatoria normal (0,1) multiplicada por la desviación estándar de acuerdo al ajuste ARIMA realizado anteriormente.
5. Se realiza un pronóstico a 13 meses de los tipos de cambio de acuerdo al modelo ARIMA que se determinó para cada serie, para el cual se utiliza la matriz con los parámetros del modelo y la matriz de errores. Los pronósticos son almacenados en la matriz de ceros creada previamente, matriz $tcfd$.
6. Con las tasa pronosticadas se realiza una estimación del costo financiero de la deuda en moneda extranjera a 13 meses. El costo financiero se estima por cada moneda, posteriormente se suma la cifra total que se guardará en un vector con los costos de cada escenario.
7. El proceso de pronosticar 13 meses de los tipos de cambio y estimar el costo financiero de la deuda en moneda extranjera se realiza 10,000 veces y con el costo financiero de cada escenario se va llenado la matriz con los costos financieros de cada escenario.
8. Una vez que se han calculado el costos financiero de cada escenario, lo que sigue es calcular el CaR, el CaR relativo y el CaR extremo de acuerdo a las definiciones del capítulo dos, ver ecuaciones 2.1, 2.2 y 2.3, respectivamente. Además, se realiza un histograma para poder observar el comportamiento del costo financiero.

7.2. Ajuste ARIMA a los tipos de cambio

Para poder realizar el ajuste a ARIMA a las series de tiempo generadas por los tipos de cambio es necesario realizar un análisis del comportamiento de la serie con las herramientas presentadas en el capítulo 6. En dicho análisis se tratará de identificar la existencia de tendencia, en caso de existir, su tipo y además si presenta o no estacionalidad la serie. Una vez que se identifique algún modelo a la serie, se verificarán los supuestos de comportamiento de ruido blanco en los residuales de la serie. Este ajuste de modelo fue realizado en el software R version 3.3.3.

El primer paso es graficar la serie para tratar de identificar algún tipo de comportamiento de manera visual. Las gráficas de las distintas series se muestran en la figura 7.3.

Por el momento se centrará la atención en una sola serie, la del tipo de cambio con respecto al dólar americano, ver figura 7.4. En el apéndice se muestra el ajuste para el resto de las series.

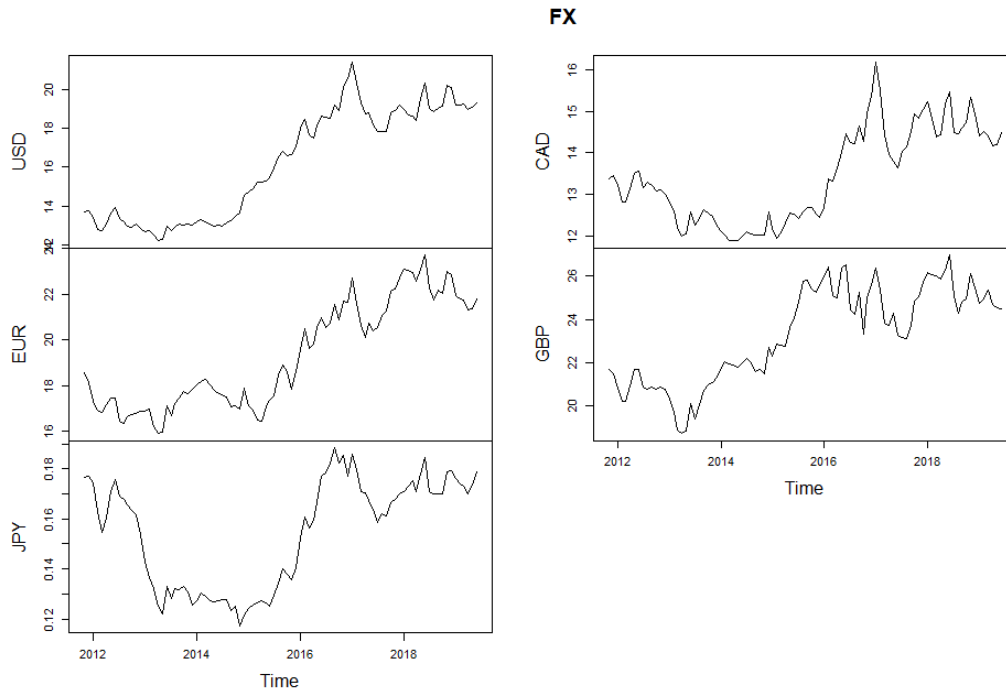


Figura 7.3: Tipos de cambio vistos como series de tiempo

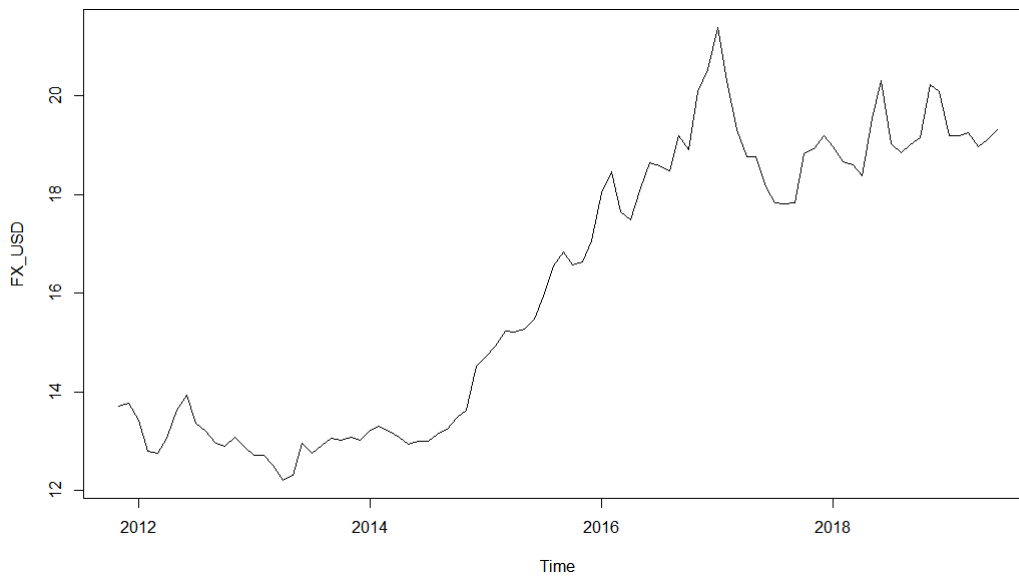


Figura 7.4: Tipo de cambio USD/MXN

Al parecer la serie presenta una una tendencia cuadrática, esto debido al rápido crecimiento en la paridad cambiaria. Se realiza una regresión lineal y una cuadrática para poder observar a la serie con respecto a una tendencia lineal y otra polinómica.

Debido a la similitud de las regresiones, las gráficas, las medidas R^2 , así como percentiles de los residuales (figuras 7.5, 7.7, 7.6 y 7.8), se decide también realizar diferentes diferenciaciones a las series: Se realizará una diferenciación de orden 1, la cual filtra una tendencia lineal, así como una de orden 2, la cual filtra una tendencia cuadrática y finalmente, una serie diferenciada una vez en la parte estacional y luego diferenciada una vez en la parte tendencia (tendencia y estacionalidad lineales) con el fin de visualizar la series para

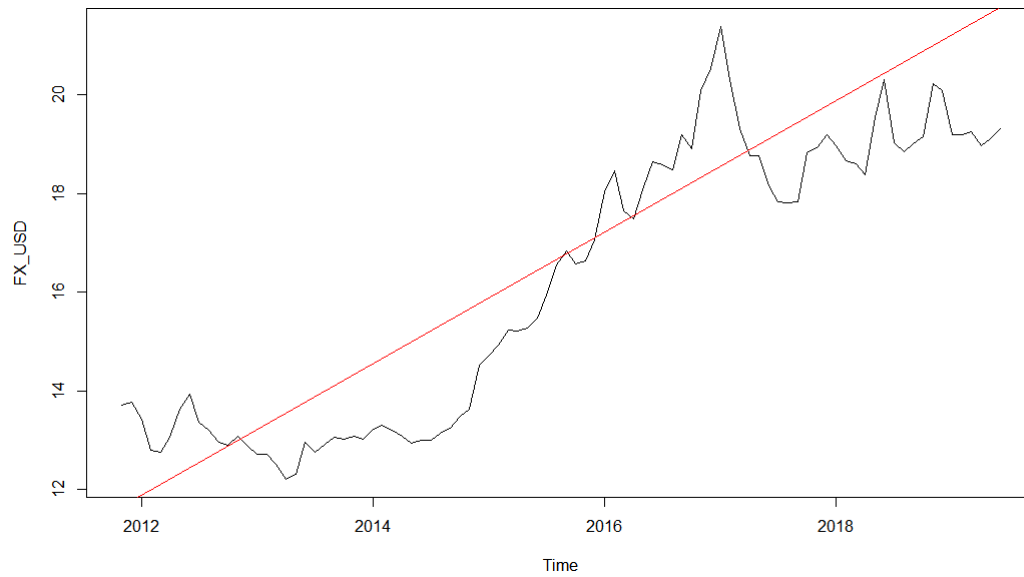


Figura 7.5: Regresión lineal al tipo de cambio USD/MXN

```
> summary(regres)
Call:
lm(formula = FX_USD ~ t)

Residuals:
    Min       1Q   Median       3Q      Max
-1.8576 -0.8651 -0.2430  0.7273  3.6381

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.673e+03  1.293e+02  -20.67  <2e-16 ***
t            1.335e+00  6.416e-02   20.80  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.182 on 90 degrees of freedom
Multiple R-squared:  0.8278,    Adjusted R-squared:  0.8259
F-statistic: 432.6 on 1 and 90 DF,  p-value: < 2.2e-16
```

Figura 7.6: Resumen de regresión lineal al tipo de cambio USD/MXN

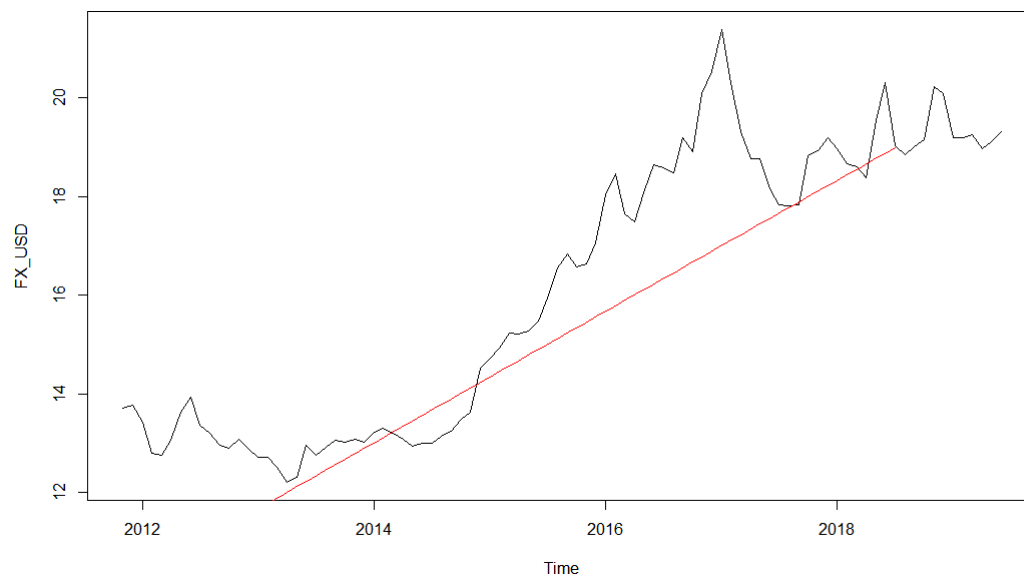


Figura 7.7: Regresión cuadrática al tipo de cambio USD/MXN

```

> summary(regres2)

Call:
lm(formula = FX_USD ~ t + I(t^2))

Residuals:
    Min       1Q   Median       3Q      Max
-1.8617 -0.8628 -0.2410  0.7276  3.6348

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -8.642e+03  1.525e+05  -0.057   0.955
t             7.258e+00  1.514e+02   0.048   0.962
I(t^2)       -1.470e-03  3.756e-02  -0.039   0.969

Residual standard error: 1.189 on 89 degrees of freedom
Multiple R-squared:  0.8278,    Adjusted R-squared:  0.8239
F-statistic: 213.9 on 2 and 89 DF,  p-value: < 2.2e-16

```

Figura 7.8: Resumen de regresión cuadrática al tipo de cambio USD/MXN

poder determinar cual transformación es mejor candidata, en el sentido de que hace la serie estacionaria, para así poder después realizar un ajuste ARIMA.

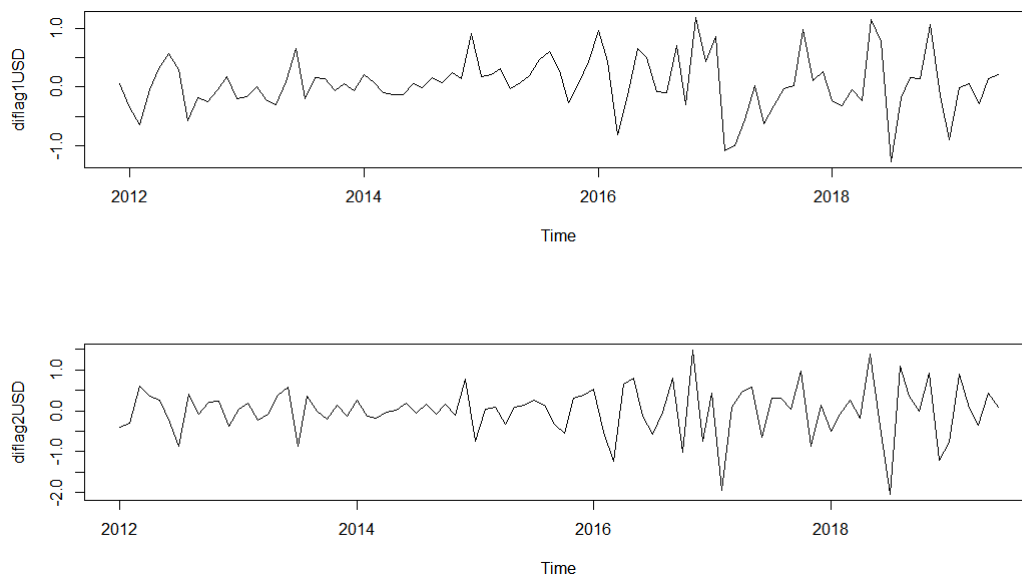


Figura 7.9: De arriba a abajo: Serie diferenciada una vez, serie diferenciada dos veces.

Las series diferenciadas presentadas que de manera visual lucen como mejores candidatas a ser series estacionales son la series diferenciada 2 veces, llamada *diflag2USD* y la serie diferenciada en la parte estacional y luego diferenciada una vez en su tendencia, es decir, la serie llamada *diflag12-1USD*. A estas dos series se les realizó la prueba de Dickey-Fuller, recordando la hipótesis nula plantea la existencia de alguna raíz unitaria (y por lo tanto la aceptación de que la serie no es estacionaria), ambas pruebas, ver 7.6, indican que no se debe rechazar la hipótesis nula por lo que se puede decir que ambas series son estacionarias. Un modelo sobre una serie diferenciada en su parte estacional y luego una diferencia simple para filtrar tendencia, sería un modelo SARIMA, el cual se encuentra fuera del alcance de este trabajo y del tipo de modelos ajustados por la SHCP no los contempla, por lo cual se elige a la serie diferenciada 2 veces para ajustar un modelo.

Para tratar de identificar el modelo se graficaron tanto las sucesiones de autocorrelaciones como las sucesiones de autocorrelaciones parciales muestrales de la serie diferenciada, el resultado sobre las ACF arroja que es recomendable utilizar un orden de medias móviles igual a uno ya que a partir de este lag ninguna ACF se sale de la banda de confianza, ver figura 7.12. Así mismo las PACF que se salen de la banda de confianza son aquellas con lag

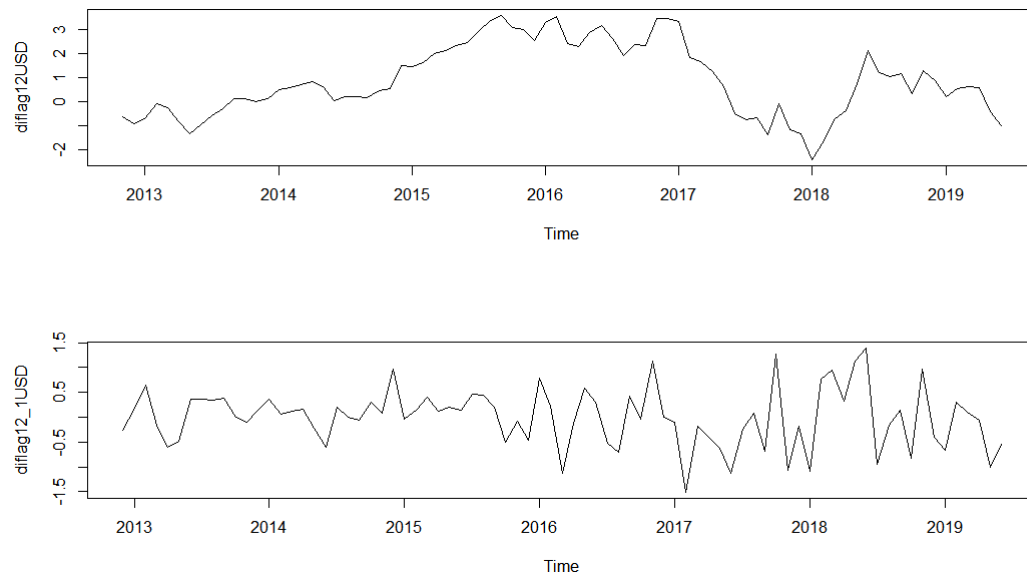


Figura 7.10: De arriba a abajo: Serie diferenciada en la parte estacional, Serie diferenciada en la parte estacional y luego una vez más.

```
> adf.test(diflag2USD, k=0)
Augmented Dickey-Fuller Test
data: diflag2USD
Dickey-Fuller = -12.722, Lag order = 0, p-value = 0.01
alternative hypothesis: stationary
Warning message:
In adf.test(diflag2USD, k = 0) : p-value smaller than printed p-value
> adf.test(diflag12_1USD, k=0)
Augmented Dickey-Fuller Test
data: diflag12_1USD
Dickey-Fuller = -8.4655, Lag order = 0, p-value = 0.01
alternative hypothesis: stationary
Warning message:
In adf.test(diflag12_1USD, k = 0) : p-value smaller than printed p-value
```

Figura 7.11: Prueba de Dickey-Fuller. De arriba a abajo: Serie diferenciada en la parte estacional, Serie diferenciada en la parte estacional y luego una vez más.

2,3,4,5 y 10, ver figura 7.13, por lo que se probó con los distintos valores. Se decidió elegir aquel modelo que minimizó el criterio de Akaike, el cual es el default en la función *arima()*.

ACF del proceso USD DIFF2

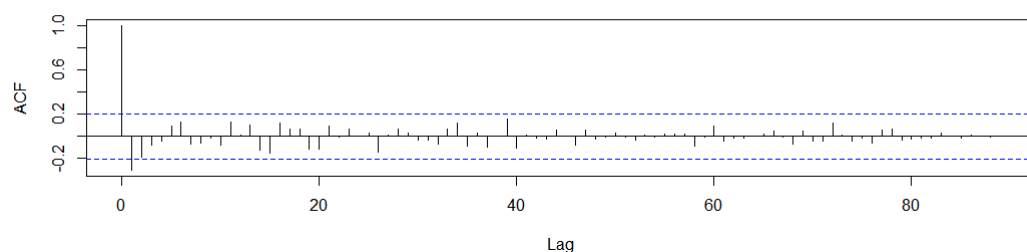


Figura 7.12: Sucesión de autocorrelaciones muestrales de la serie Diff1ag2USD.

Se realizó el análisis para las distintas combinaciones $p = \{2, 3, 4, 5, 10\}$ y $q = \{1\}$, se compararon sus criterios de Akaike y finalmente se determinó que el modelo que lo minimizaba, así como tener la ventaja de no ser excesivamente complejo, fue el modelo ARIMA(2,2,1), con $\phi_1 = 0,2068$, $\phi_2 = -0,1821$ y $\theta_1 = -1$.

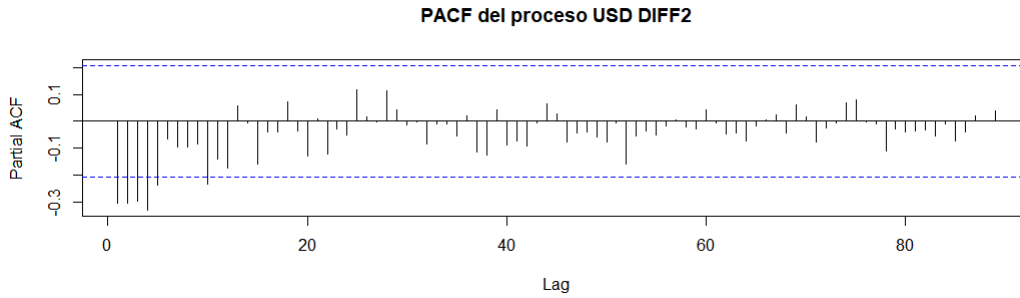


Figura 7.13: Sucesión de autocorrelaciones parciales muestrales de la serie Diffflag2USD.

```

> ARIMA121=arima(FX_USD, order=c(1,2,1), include.mean = TRUE)
> ARIMA121

Call:
arima(x = FX_USD, order = c(1, 2, 1), include.mean = TRUE)

Coefficients:
      ar1      ma1
    0.1784 -1.0000
s.e. 0.1042 0.0384

sigma^2 estimated as 0.2095: log likelihood = -59.45, aic = 124.89
> ARIMA221=arima(FX_USD, order=c(2,2,1), include.mean = TRUE)
> ARIMA221

Call:
arima(x = FX_USD, order = c(2, 2, 1), include.mean = TRUE)

Coefficients:
      ar1      ar2      ma1
    0.2068 -0.1821 -1.0000
s.e. 0.1035 0.1031 0.0504

sigma^2 estimated as 0.2016: log likelihood = -57.92, aic = 123.84

```

Figura 7.14: Ejemplo de distintos ajustes ARIMA la serie Diffflag2USD.

Después de realizado el ajuste, se deben verificar el supuesto de comportamiento de ruido blanco en el modelo, por lo que se realizó un histograma de los residuales del ajuste ARIMA, recordando que los residuales pueden verse como estimadores de estas variables aleatorias. Además del histograma, ver figura 7.15, se muestra la curva generada por una distribución normal con la media y varianza generadas por los residuales.

También se muestran las ACF de los residuales y el estadístico de Ljung-Box para los distintos lags. Se puede observar, ver figura 7.16, que las autocorrelaciones se mantienen dentro de las bandas de confianza y así mismo los estadísticos de Ljung-Box se encuentran por encima del nivel de significancia por lo que no se rechaza la hipótesis nula de que los residuales son conjuntamente cero.

De manera visual el histograma se ajusta a la curva de la distribución normal, además la media generada por los residuales es 0.03694858, que es un número *cercano* a cero, por lo que haciendo este análisis preliminar, se podría decir que los residuales siguen un comportamiento de ruido blanco. Sin embargo, se realizará una prueba de bondad de ajuste para tener un soporte más robusto acerca del comportamiento de los residuales. Se utilizaron primero la prueba de Kolmogorov-Smirnov y luego la de Anderson-Darling, mencionadas en el capítulo seis.

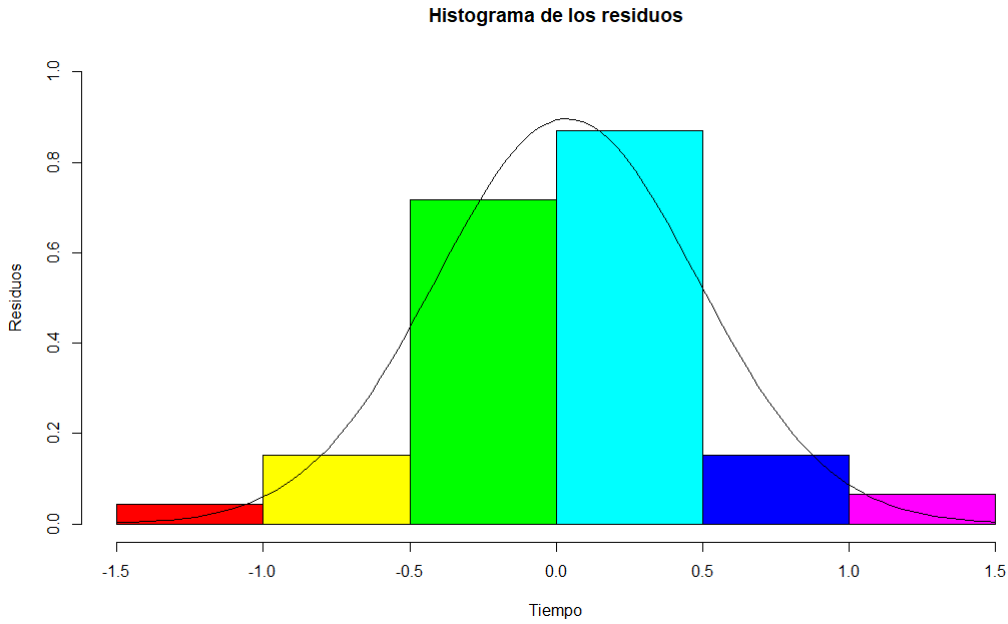


Figura 7.15: Residuales del ajuste ARIMA(2,2,1) a la serie DiffIag2USD .

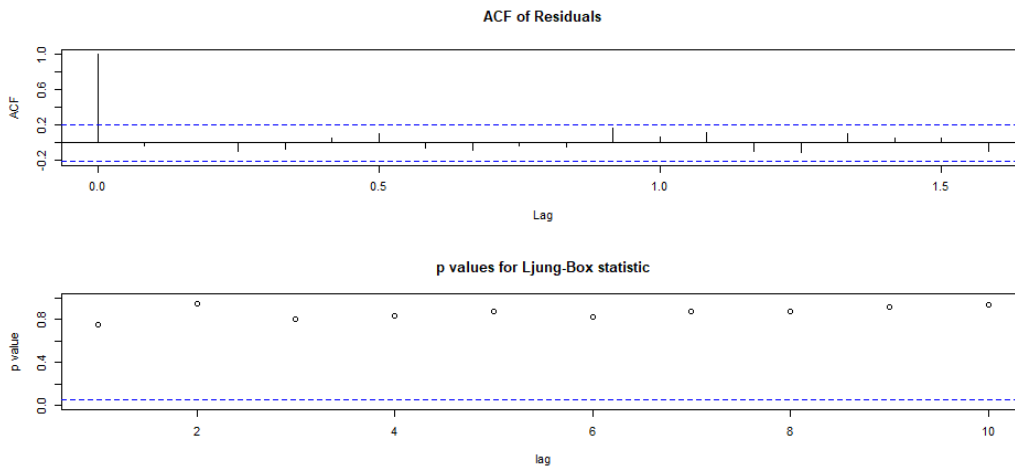


Figura 7.16: ACF y el estadístico de Ljung-Box para los distintos lags de los residuales

Para la prueba de Kolmogorov-Smirnov que el p-value es más grande que el nivel de significancia α de 5 %, sin embargo, para la prueba de Anderson-Darling (cuyo uso se prefiere para probar normalidad) el p-value es más pequeño que el nivel de significancia, por tanto no se puede concluir que los residuales tienen comportamiento normal. Sin embargo, la definición de ruido blanco no contempla la normalidad, solo que sean idénticamente distribuidos, ver figura 7.17.

En el apéndice se encuentra la Tabla Resumen de los ajustes ARIMA de todas las series de los tipos de cambio y adicionalmente del precio de la UDI que se utilizará más adelante. Señalar que para la mayoría de los ajustes, no se rechaza la normalidad en los residuales de los modelos.

Para concluir con el ajuste ARIMA, solo faltaría el pronóstico, cabe señalar que en este momento ya se cuenta con todo lo necesario para correr el código que estima el CaR


```

> ks.test(residuos, pnorm, mean(residuos),sd=sqrt(var(residuos)))

One-sample Kolmogorov-Smirnov test

data:  residuos
D = 0.093562, p-value = 0.3732
alternative hypothesis: two-sided

> lillie.test(residuos)

Lilliefors (Kolmogorov-Smirnov) normality test

data:  residuos
D = 0.093562, p-value = 0.04543

> ad.test(residuos)

Anderson-Darling normality test

data:  residuos
A = 1.292, p-value = 0.002214

```

Figura 7.17: Prueba Kolmogorov-Smirnov para una distribución normal. La primera con la media muestral y la segunda con media cero.

de tipo de cambio y no es necesario realizar el pronóstico, sin embargo se realizará para concluir el proceso común de este tipo de análisis. El forecast realizado por el programa devuelve la predicción que minimiza el error cuadrático, este podría ser el mejor pronóstico que se puede realizar con un ajuste ARIMA, sin embargo, lo que se requiere en el código son los parámetros ya que se van a simular una serie de realizaciones para determinar el comportamiento del CaR, este pronóstico se puede ver en la figura 7.18.

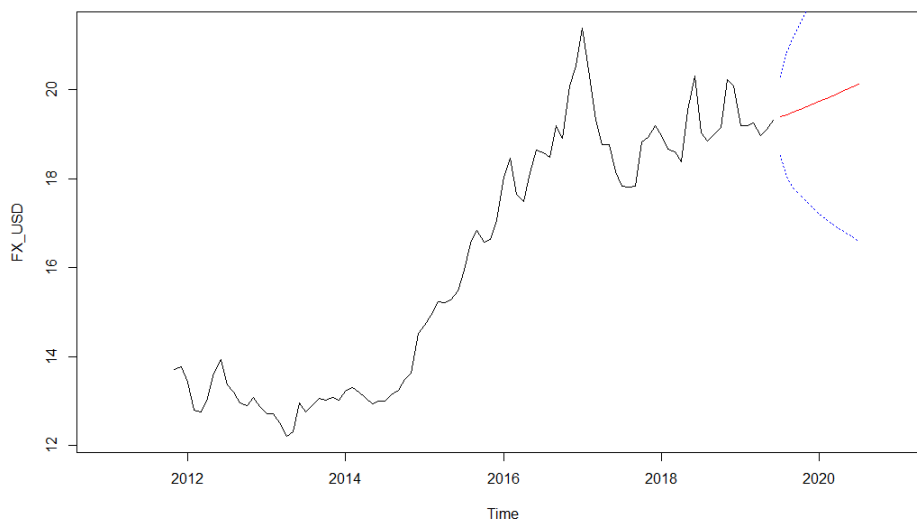


Figura 7.18: Pronostico de un ajuste ARIMA(2,2,1) a la serie de tipos de cambio USD/MXN

Es importante mencionar nuevamente, que los ajustes bajo un modelo SARIMA, el cual considera la parte estacional quizá se hubieran asemejado más al comportamiento de la serie original, sin embargo, el tipo de modelos implementados por la SCHP son solamente ARIMA así que este tipo de modelos no se utilizaron para la estimación del CaR. Sin embargo, se realizó un pronóstico utilizando el modelo SARIMA para comparar el ajuste con el modelo ARIMA, ver figura 7.19. Se puede observar que el pronóstico en el modelo SARIMA se adapta de mejor manera a la naturaleza de la serie original, ya que toma en cuenta la parte estacional de esta.

Finalmente, se ajustan los modelos ARIMA para el resto de series de tipo de cambio, replicando el proceso mismo. Recapitulando primero se realiza un análisis visual para tratar

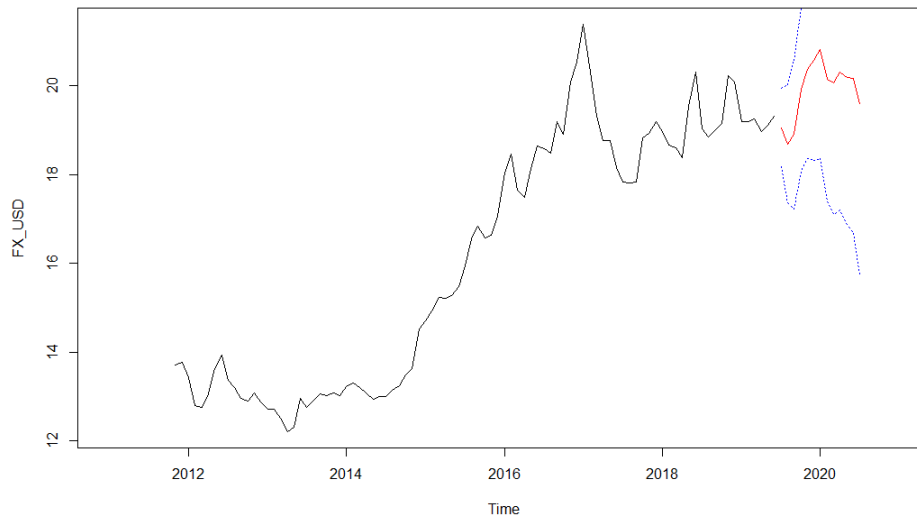


Figura 7.19: Pronostico de un ajuste SARIMA a la serie de tipos de cambio USD/MXN

de identificar la tendencia y estacionalidad de la serie, después se realizan una serie de transformaciones a las series para tratar de obtener una serie estacionaria a la cual por medio de la prueba de Dickey-Fuller se determina si lo es o no. Una vez que se obtiene una serie estacionaria se grafican sus sucesiones de autocorrelaciones y autocorrelaciones parciales para tratar de determinar el orden las partes autoregresivas y de medias móviles de la serie. Una vez determinado el orden se debe realiza un análisis para la validación del modelo, por medio de pruebas de bondad de ajuste (Anderson-Darling y Kolmogorov-Smirnov) se determina el comportamiento de los residuales y se debe concluir si es que provienen de un ruido blanco. Finalmente, se mostró el pronóstico tanto del ajuste ARIMA como el de un modelo SARIMA, esto último solamente con el fin de mostrar cómo luce un pronostico aunque para el fin de este trabajo solo basta con determinar el modelo, así como sus parámetros. A continuación se muestra el código utilizado en R con el que se realizó el ajuste para la serie de tiempo del tipo de cambio USD/MXN.

7.2.1. Código en R para ajuste ARIMA a la serie de tiempo USD/MXN

```
rm(list = ls())

#Establece el directorio trabajo
setwd("C:/Users/Rodrigo García/Documents/Mi Tesis/PARTE PRÁCTICA/
PARAMETRICO/TIPOS DE CAMBIO")

#Lee los archivos .csv y los almacena
datos<-read.table("FX.csv",sep=";",header=TRUE)

#Los tipos de cambio en la matriz son, de izquierda derecha
#USD, EUR, JPY, CAD, GBP. Los tipos de cambio son promedios mensuales del
MID market
#La serie inicia en noviembre de 2015 y termina en noviembre de 2018 (37
meses)

FX_USD<-ts(datos[,1], frequency = 12, start = c(2011, 11))
FX_USD
library(stats)
library(tseries)
library(nortest)
```

```

#Gráfica de la serie de tiempo
plot(FX_USD)

# Primero se genera una secuencia de tiempo
t<-seq(2011.917,2018.5, length=length(FX_USD)) #2015.917 representa
  noviembre de 2017
t

#Regresión FX como respuesta del tiempo
regres<-lm(FX_USD~t)
fitted.values(regres)
regres

#Esta gráfica proporciona la regresión ajustada (en rojo) junto con la
  serie de datos original (USD).
plot(FX_USD)
abline(regres,col=2)

#Residuales de la regresión
residuals(regres)
# Grafica de los residuales
plot(residuals(regres))

#Regresion cuadratica

regres2<-lm(FX_USD~t+I(t^2))
fitted.values(regres2)
regres2
plot(residuals(regres2))
plot(FX_USD)
lines(t,-8.642e+03+(7.258*t)-(1.470e-03*t^2), add = TRUE,col="red")

#Ya que pareciera que los tipos de cambio presentan tendencia (positiva)
  se realizará una diferencia de orden 1

diflag12USD<-diff(FX_USD, lag = 12, differences = 1)
diflag12_1USD<-diff(diflag12USD, lag = 1, differences = 1)

diflag1USD<-diff(FX_USD, lag = 1, differences = 1)
diflag2USD<-diff(FX_USD, lag = 1, differences = 2)

par(mfrow=c(2,1))
plot(diflag1USD)
plot(diflag2USD)
plot(diflag12USD)
plot(diflag12_1USD)

#Las candidatas a parecer estacionariorias son las serie difenciada 2
  veces (con lag 1) es decir
#la serie diflag2USD y la serie diferneciada en la parte estacional y
  luego diferenciada un vez
#es deci la serie diflag12_1USD

#Pruebas de Dickey-Fuller para ambas series
adf.test(diflag2USD, k=0)
adf.test(diflag12_1USD, k=0)
p<-length(diflag2USD)^{1/4}+1
p #p=4.088 por tanto se utilizara la prueba de DF aumentada para 4 raices
adf.test(diflag2USD, k=4)
adf.test(diflag12_1USD, k=4)

#A pesar de que las cuatro pruebas indican que se deben rechazar la

```

```

hipotesis nula
#de que existen raices unitarias , el p.value de para la prueba sobre la
serie
# diflag12_1USD es un tanto mayor por lo que probaremos con la serie
diflag2USD que ademas tiene
#el estadistico mas pequeño

seriemodelar=scale(diflag2USD , scale = FALSE)
mean(diflag2USD) # Media de la serie diferenciada se extrae para trabajar
,
# es el primer renglo de la matriz ARMA

# Sucesion de autocorrelaciones
acf(ts(seriemodelar , freq=1), lag.max=200, main="ACF del proceso USD
DIFF2")
#Se recomienda utilizar la p=1

# Sucesion de autocorrelaciones parciales
pacf(ts(seriemodelar , freq=1), lag.max=200, main="PACF del proceso USD
DIFF2")
#Se salen de las bandas las autocorrelaciones parciales con lag 2,3,4,5 y
10 por lo que se probaran con
# los distintos valores

ARIMA121=arima(FX_USD, order=c(1,2,1), include.mean = TRUE)
ARIMA121
ARIMA221=arima(FX_USD, order=c(2,2,1), include.mean = TRUE)
ARIMA221
ARIMA321=arima(FX_USD, order=c(3,2,1), include.mean = TRUE)
ARIMA321
ARIMA421=arima(FX_USD, order=c(4,2,1), include.mean = TRUE)
ARIMA421
ARIMA521=arima(FX_USD, order=c(5,2,1), include.mean = TRUE)
ARIMA521
ARIMA1021=arima(FX_USD, order=c(10,2,1), include.mean = TRUE)
ARIMA1021
#El modelo que minimiza el Criterio de Akaike es el modelo ARIMA(2,2,1)

#Verificacion de residuos

tsdiag(ARIMA221)
residuos=ARIMA221$residuals
plot(residuos , xlab = "Tiempo", ylab = "Residuos", main = "Residuos" )
hist(residuos , col=rainbow(6), xlab = "Tiempo", ylab = "Residuos", main =
"Histograma de los residuos", ylim = c(0,1), probability = TRUE)
curve(dnorm(x, mean(residuos), sd(residuos)), add = TRUE, col="black")
#Pruebas de normalidad - Kolmogorov-Smirnoff -

source("Kolmogorov.R")
Kolmogorov_Smirnov(residuos)
ks.test(residuos , pnorm, mean(residuos),sd=sqrt(var(residuos)))

#Como el estadistico Dn es mas chico que el cuantil 1-alpha (95 %) se
decide no rechazar H0
#Equivalente a decir que p-value es mas grande que el nivel de
significancia alpha (5 %)
#Por tanto se puede asumir que los residuos se comportan como una normal
(ruido blanco) ya que
#tienen media 0.03 (cerca a cero) y varianza constante

#Pruebas de normalidad - Anderson-Darling -

ad.test(residuos)

```

```

#Predicciones
ARIMA_USD.pred<-predict (ARIMA221,n. ahead=13)
plot (FX_USD, xlim=c(2011,2021))
lines (ARIMA_USD.pred$pred , col="red ")
lines (ARIMA_USD.pred$pred+1.96*ARIMA_USD.pred$se , col="blue ", lty=3)
lines (ARIMA_USD.pred$pred-1.96*ARIMA_USD.pred$se , col="blue ", lty=3)

ARIMA_USD.pred

#Exportar a csv las predicciones
write.csv(x = ARIMA_USD.pred$pred , file = "ARIMA_USD.csv", row.names =
FALSE)

#AJUSTE ALTERNATIVO MODELO SARIMA
# seriemodelar2=scale(diflag12_1USD , scale = FALSE)
# acf(ts(seriemodelar2 , freq=1), lag.max=200, main="ACF del proceso USD
DIFF2")
# pacf(ts(seriemodelar2 , freq=1), lag.max=200, main="PACF del proceso USD
DIFF2")
# SARIMA111=arima(FX_USD, order=c(10,1,10),seasonal = list(order = c(0,
1, 0), period = NA), include.mean = TRUE)
# tsvdiag(SARIMA111)
# residuos=SARIMA111$residuals
# plot(residuos , xlab = "Tiempo", ylab = "Residuos", main = "Residuos" )
# hist(residuos , col=rainbow(6), xlab = "Tiempo", ylab = "Residuos", main
= "Histograma de los residuos", probability = TRUE, ylim = c(0,1))
# curve(dnorm(x, mean(residuos), sd(residuos)), add = TRUE, col="black")
# source("Kolmogorov.R")
# Kolmogorov_Smirnov(residuos)
# ks.test(residuos , pnorm, mean(residuos),sd=sqrt(var(residuos)))
# SARIMA_USD.pred<-predict (SARIMA111,n. ahead=13)
# plot (FX_USD, xlim=c(2011,2021))
# lines (SARIMA_USD.pred$pred , col="red ")
# lines (SARIMA_USD.pred$pred+1.96*SARIMA_USD.pred$se , col="blue ", lty=3)
# lines (SARIMA_USD.pred$pred-1.96*SARIMA_USD.pred$se , col="blue ", lty=3)
# SARIMA111
# SARIMA_USD.pred

```

Como resultado de los diversos ajustes que se realizaron a las series de tipos de cambio se obtuvieron los siguientes modelos:

- a) Para el dólar americano se determinó utilizar un modelo ARIMA(2,2,1).
- b) Para el euro se determinó utilizar un modelo ARIMA(3,2,1).
- c) Para el yen se determinó utilizar un modelo ARIMA(1,2,1).
- d) Para el dólar canadiense se determinó utilizar un modelo ARIMA(2,2,1).
- e) Para la libra se determinó utilizar un modelo ARIMA(1,2,1).

A continuación se muestra la matriz ARIMA que se utilizará para la estimación del costo en riesgo, esta se construye a partir de ajustar un modelo a cada serie de tipo de cambio y los parámetros se colocan en la matriz de acuerdo a la ecuación 7.1. El primer renglón es la media de la serie ajustada, los renglones 2 a 4 son los parámetros autorregresivos (ya que el orden máximo fue de 3), el quinto renglón es el parámetro de medias móviles (todos los ajustes fueron de orden 1) y el último renglón es la desviación estándar de los errores asociados a los residuales.

ARIMA					
	1	2	3	4	5
1	0.0017435	0.0097191	0.00053425	0.0022729	0.0013958
2	0.2068	-0.0025	0.1211	0.217	0.006
3	-0.1821	-0.1289	0	-0.2243	0
4	0	-0.1538	0	0	0
5	-1	-0.9873	-0.9526	-1	-1
6	0.2016	0.282	2.766e-05	0.1049	0.4943
7					
8					

Figura 7.20: Matriz ARIMA

7.2.2. Código en Matlab para CaR de tipo de cambio

```

%-----
%                               Estimación CaR TC
%-----

clear ; close all; clc

pkg load statistics

load BASE_FX.mat % CARGAR INPUTS

%1) Se generan las predicciones para el tipo de cambio. Se usa una
    serie de modelos ARIMA que se calcularon por separado y se pasaron a
    la matriz TC

TC; % Matriz de la serie de los tipos de cambio utilizados en el modelo
    ARIMA
TC_d; % Matriz de la serie diferenciada de los tipos de cambio
    utilizados en el modelo ARIMA

t0=TC_d(end-2:end,:); % Se requieren ultimos 3 Tipos de Cambio ya que el
    modelo AR de mayor orden tiene orden 3

calen=calen(2:end,:);

pred=13;

e(1,:)= [0,0,0,0,0]; % Inicialización del vector de errores en el modelo.
tcf_d=zeros;
for j=1:10000 % Indica el número de simulaciones del costo.
for i=1:pred
tcf_d(1:3,1:5)=t0; % Valores semilla (ultimos 3 Tipos de Cambio)
e(i+1,:)=ARIMA(6,:) .* (normrnd(0,1)*ones(1,5)); % Desv Est de los
    errores X N(0,1)
tcf_d(i+3,:)=ARIMA(1,:)+(ARIMA(2,:).*tcf_d(i+2,:))+(ARIMA(3,:).*tcf_d
    (i+1,:))+(ARIMA(4,:).*tcf_d(i,:))+(ARIMA(5,:).*e(i,:))+e(i+1,:);
tcf=diffinv2(tcf_d,TC(end-2:end,:)); % La funcion diffinv2 regresa
    los valores de una serie diferenciada 2 veces a sus valores sin
    diferenciar

end

tcf=tcf(4:end,:); % Los primeros 3 valores semilla se omiten

```

```

% 2) Estimación del costo financiero
% 2.1) Tasa fija
costo=calen.*tcf(1:end-2,:); % Devuelve el producto del modelo ARMA (Del
    tipo de Cambio) por el costo en moneda extranjera, para obtener el
    costo en moneda local. Además se eliminan 2 valores debido a el numero
    de pronosticos de la serie original es n+2
costo2018(j,1)=sum(sum(costo(1,:),2);
costo2019(j,1)=sum(sum(costo(2:13,:),2);

end

% 3) CaR al 95%
costoesperado2019=mean(costo2019);
CaR952019=prctile(costo2019,95);

% 4) CaR relativo al 95%
CaR_relativo_2019=CaR952019/costoesperado2019;

% 5) CaR Relativo Extremo
AAA=sort(costo2019);
BBB=AAA(.95*length(AAA)+1:end,:);
CaRExtremo2019=mean(BBB)/costoesperado2019;

hist(costo2019,1000)

#-----Regreso Series diferenciada lag 2-----
#Esta funcion sirve para que a partir de una serie (que presenta algun
    pronostico)
#diferenciada 2 veces, vuelva a la serie original con los pronosticos
    ajustados

function [S]=diffinv2(diff2, Xt)
    diff1_1=Xt(2,:)-Xt(1,:);
    diff1_2=Xt(1,:)+cumsum(diff2);
    diff1=[diff1_1;diff1_2];
    S=diffinv1(diff1, Xt);
end

#-----Regreso Series diferenciada lag 1-----
#Esta funcion sirve para que a partir de una serie (que presenta algun
    pronostico)
#diferenciada una vez, vuelva a la serie original con los pronosticos
    ajustados

function [P]=diffinv1(diff1, Xt)
    Xt1=Xt(1,:);
    Xt2=Xt(1,:)+cumsum(diff1);
    P=[Xt1;Xt2];
end

```

Luego de correr el código en Matlab con todos los inputs de manera correcta, los resultados obtenidos son los siguientes:

1. Costo en riesgo esperado que es la esperanza de la variable aleatoria que mide el costo financiero, es decir, $\mathbb{E}[\Lambda]$.
2. Costo en riesgo al 95 % dado por la ecuación 2.1.
3. Costo en riesgo relativo dado por la ecuación 2.2.

4. Costo en riesgo extremo dado por la ecuación 2.3.
5. Histograma de la variable aleatoria que mide el costo financiero.

El costo esperado para los próximos 13 meses es de 82,997 millones de pesos, así mismo el CaR al 95 % es de 85,724 millones de pesos por lo que se puede decir que el costo financiero generado por la deuda en moneda extranjera no deberá superar los 86 billones de pesos. El Costo en riesgo tanto relativo al 95 % como extremo fueron, respectivamente, 1.0329 y 1.0409 veces el valor con respecto al CaR al 95 %. Ver figura 7.22. Con estas medidas se puede tener idea de la cantidad de dinero necesaria, en los siguientes 13 meses⁹, para que el gobierno haga frente a sus obligaciones por bonos en moneda extranjera para los meses futuros y poder así planificar la administración. El histograma de la variable aleatoria que mide el costo financiero se muestra a continuación. Ver figura 7.21.

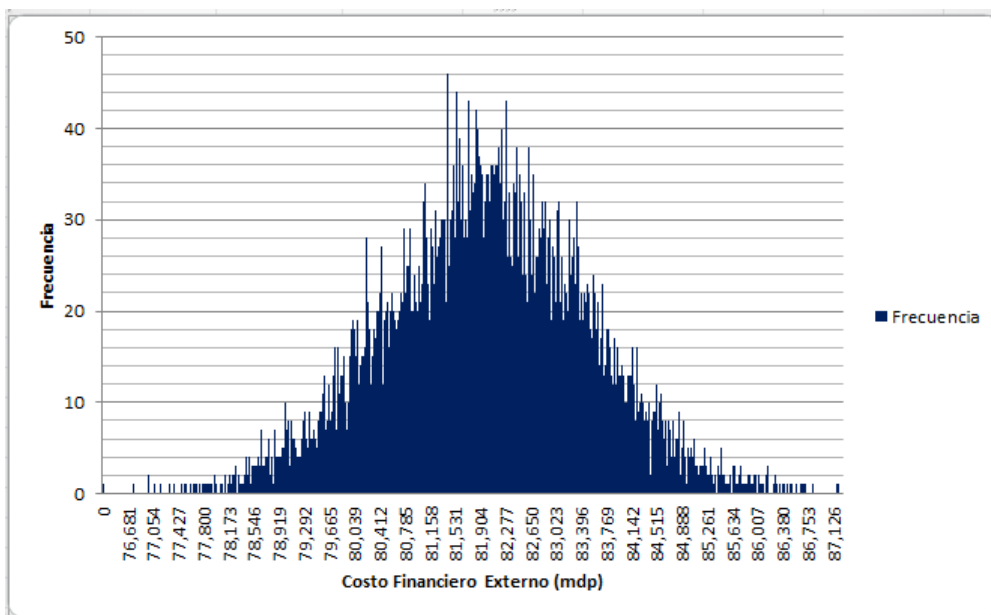


Figura 7.21: Histograma de la variable aleatoria que mide el costo financiero generado por la deuda pública en moneda extranjera.

```
>> costoesperado2019=mean(costo2019)
costoesperado2019 = 82997705032.81319
>> CaR952019=prtile(costo2019,95)
CaR952019 = 85724260531.42645
>> CaR_relativo_2019=CaR952019/costoesperado2019
CaR_relativo_2019 = 1.0329
>> CaRExtremo2019=mean(BBB)/costoesperado2019
CaRExtremo2019 = 1.0409
>>
```

Figura 7.22: Resultados del código para el CaR en moneda extranjera.

7.3. CaR de Tasas de interés

En esta sección se describirá paso a paso el método de estimación del costo en riesgo generado por los movimientos en las tasas de interés, es decir aquel que mide el costo

⁹Suponiendo que el ejercicio se realizó en Diciembre de 2018

financiero de la deuda interna ya sea por los instrumentos que pagan a tasa flotante o las colocaciones futuras de los bonos a tasa fija.

Para la estimación del CaR en tasas de interés, al igual que en el costo en riesgo de tipo de cambio, se deberá realizar simulación Monte Carlo de 10,000 escenarios de comportamientos de tasas de interés, para luego estimar el costo financiero por cada instrumento de deuda, de acuerdo a las notas técnicas del Banco de México[2],[3],[4],[5]. Y finalmente se proporciona el valor del CaR al un nivel $(1 - \alpha)$ % de confianza con los pronósticos de las tasas, así como el CaR relativo y el CaR extremo, de acuerdo a la ecuaciones 2.1, 2.2 y 2.3 respectivamente.

Para la predicción de las tasas se utilizará el modelo de Cox-Ingersol-Ross, mientras que para la estimación de los parámetros que utiliza el modelo se utiliza una función adicional al código que calcula el CaR, la cual estima los parámetros de acuerdo a la propuesta de Overbeck y Rydén, mencionado en la ecuación 4.13

Se abre un paréntesis en este momento para dedicar un momento a la función mencionada en el párrafo anterior, la función programada en Matlab llamada *cirparameters*. Esta función recibe dos parámetros, el primero es un vector de tasas de interés ¹⁰ mientras que el segundo parámetro es el δt , es decir indica la frecuencia de las observaciones en un año. Lo primero que realiza este código es una regresión lineal de las primeras $n - 1$ observaciones sobre las observaciones que van de la segunda a la n -ésima observación, es decir, se considera a la tasa al tiempo t como respuesta de la tasa retrasada en una unidad de tiempo, esto es la tasa al tiempo $t - 1$, para posteriormente estimar \hat{a} y \hat{b} dados por la ecuaciones 4.19 y 4.17 respectivamente, se calcula además el error cuadrático medio, dado por $\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$. Así mismo se calcula el error absoluto medio de la regresión, de acuerdo a la siguiente expresión $\frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$. Después se estiman los parámetros η_0 y η_1 que servirán para finalmente estimar $\tilde{\sigma}_n^2$ y $\hat{\sigma}_n^2$ dados por las ecuaciones 4.25 y 4.26. La función devuelve un vector los parámetros $(\hat{a}/\hat{b}, \hat{b}, \tilde{\sigma}_n^2, \hat{\sigma}_n^2, e_{cm}, e_{am})$ con los cuales es posible pronostica en el modelo CIR. Más adelante se muestra el código, en Matlab, de esta función a fin de mostrar la manera en como lo visto en el capítulo cuatro se aplica para estimar los parámetros del modelo.

Una vez que se cuentan con los parámetros estimados \hat{a} y \hat{b} con la función anterior, se pronostican las tasas discretizando el modelo dado por la ecuación 4.13. El modelo discretizado queda de la siguiente manera:

$$r(\Delta t) \approx r(t) + [a + br(t)]\Delta t + \sigma\sqrt{r(t)}\sqrt{\Delta t}w \quad (7.2)$$

Donde w es una variable aleatoria normal $(0, 1)$.

Los primeros inputs que se utilizarán para la estimación son las matrices con los históricos mensuales de las tasas de los instrumentos que conforman la deuda (CETEs a 28, 91, 182 y 364 días, Bonos a tasa fija a 3, 5, 10, 20 y 30 años, Udibonos a 3, 10 y 30 años y Bonde D a 5 años). Señalar que el Bonde D está referenciado a TIIIE más una sobre tasa, lo que se utilizará es la tasa TIIIE ya que es el componente aleatorio de la tasa.

Adicional a las matrices con las tasas de interés, otro de los insumos utilizados son las matrices de emisiones futuras, éstas contienen el monto (en millones de pesos) de las

¹⁰Los valores se encuentran en forma porcentual, así si una tasa del cinco punto cinco por ciento aparecerá como 5.5 y no como 0.055.

colocaciones de los bonos del año 2019, es decir, se necesitan la matriz con las colocaciones de los Bondes D, de CETEs de Udibonos y los bonos a tasa fija, para que una vez realizadas las proyecciones de las tasas se pueda hacer una estimación del costo financiero, es decir estas matrices fungirán como un nominal para las proyecciones¹¹.

Además de las matrices de emisiones futuras, se necesita la matriz de emisión (ya realizada) de Bondes D, ya que estos bonos se encuentra vivos y al desconocer la tasa con la que pagará (pagan a tasa flotante) es que forman parte también del costo financiero.

La Figura 7.23 muestra el monto en millones de pesos de la colocación de bonos para el año 2019. De arriba a abajo: En la primer matriz se encuentra la colocación de Bondes D, seguido de los bonos a tasa fija a 3, 5 10, 20 y 30 años, seguido de CETEs a 28, 91, 182 y 364 días y finalmente Udibonos a 3, 10 y 30 años.

bonde					
	1	2	3	4	5
1	7500				
2	16500				
3	25500				
4	39000				
5	49000				
6	59000				
<					
btf					
	1	2	3	4	5
1	8000	8000	8500	0	3500
<					
cete					
	1	2	3	4	5
1	28000	44000	46000	12500	
<					
udibono					
	1	2	3	4	5
1	750	600	400		
2	950	700	400		
3	950	700	400		
<					

Figura 7.23: Matrices de colocación de Bonos. De arriba a abajo: Bonde D, Bonos a Tasa Fija, CETEs, Udibonos

Además de las matrices de tasas de interés y emisiones, otro insumo utilizado son los calendarios de pagos de los bonos, estas matrices se pueden calcular con base en el número de cupones que paga al año cada instrumento financiero. Estas matrices indican el número de cupones que la emisión de cada mes paga en el año 2019 (en su defecto el año de la valuación). Como una medida más conservadora, las matrices contienen la parte proporcional de cupones que pagan anualmente. Por ejemplo, los bonos que serán emitidos en agosto de 2019 en la matriz se colocará el número 0.8333 que es la parte proporcional que corresponde al número de cupones anuales, de lo contrario, estrictamente ese bono no pagará ese año. Esta es una medida que tienen como propósito darle un peso mayor al costo financiero y ser más conservadores con el costo financiero estimado.

¹¹Cabe señalar que debido a la imposibilidad de contar con las subasta programadas para el año 2019 (Subastas futuras), se consideraron los resultados de las subastas del año 2018 como las subastas programadas. El resultado de las subastas realizadas se pueden encontrar en <http://www.banxico.org.mx/SieInternet/>

En la imagen 7.24 se puede apreciar la estructura de una de las matrices de calendario de pago de cupón para 13 meses (diciembre y año calendario siguiente).

1	2	3	4	5
2	2	2	2	2
2	2	2	2	2
1.83333	1.83333	1.83333	1.83333	1.83333
1.66667	1.66667	1.66667	1.66667	1.66667
1.5	1.5	1.5	1.5	1.5
1.33333	1.33333	1.33333	1.33333	1.33333
1.16667	1.16667	1.16667	1.16667	1.16667
1	1	1	1	1
0.833333	0.833333	0.833333	0.833333	0.833333
0.666667	0.666667	0.666667	0.666667	0.666667
0.5	0.5	0.5	0.5	0.5
0.333333	0.333333	0.333333	0.333333	0.333333
0.166667	0.166667	0.166667	0.166667	0.166667

Figura 7.24: Calendario de cupones de Bonos a tasa fija para un periodo de 13 meses. De izquierda a derecha: 3 años, 5 años, 10 años, 20 años y 30 años.

El último input es la matriz Udi que contiene pronósticos de los siguientes 13 meses del valor de la udi. Esta matriz se requiere para la estimación del costo financiero generado por lo udibonos, ya que estos tienen valor nominal de 100 udis por lo que es necesario estimar su valor además. Para su estimación, se realizó el mismo proceso que en la serie de tipos de cambio¹² y se vio que el modelo adecuado que ajusta los datos fue un modelo ARIMA(2,2,2) a la serie generada y se realizó el pronóstico de este modelo. A diferencia de los ajustes realizados previamente donde se requerían solamente los parámetros para realizar las simulaciones, en este caso sí se utiliza el mejor pronóstico mencionado en el capítulo seis. El pronóstico de ajuste ARIMA se puede ver a continuación, ver figura 7.25.

Una vez que se cuenta con los inputs iniciales, el algoritmo a seguir para poder realizar la estimación del CaR es el siguiente:

1. El primer paso que se realiza en el código es utilizar la función que calcula los parámetros \hat{a} , \hat{b} y $\hat{\sigma}$ del modelo dado por la ecuación 4.13 para cada serie (se considera una serie al vector con las tasas mencionadas como primeros inputs). Con la información generada por la función se construye una matriz, con la cual se calculará posteriormente el costo en riesgo.
2. Construir una matriz que contendrá en su primera columna el parámetro \hat{a} , en la segunda al parámetro \hat{b} , en la tercer y cuarta columna a los parámetros $\tilde{\sigma}$ y $\hat{\sigma}$ dados por la ecuaciones 4.25 y 4.26. Adicionalmente se añadirá una quinta columna la cual contendrá el último rendimiento conocido (es decir, antes de realizar los pronósticos) de las matrices de rendimientos de los bonos. La matriz se llena renglón a renglón y tiene la siguiente forma:

¹²Ver el ajuste en el apéndice.

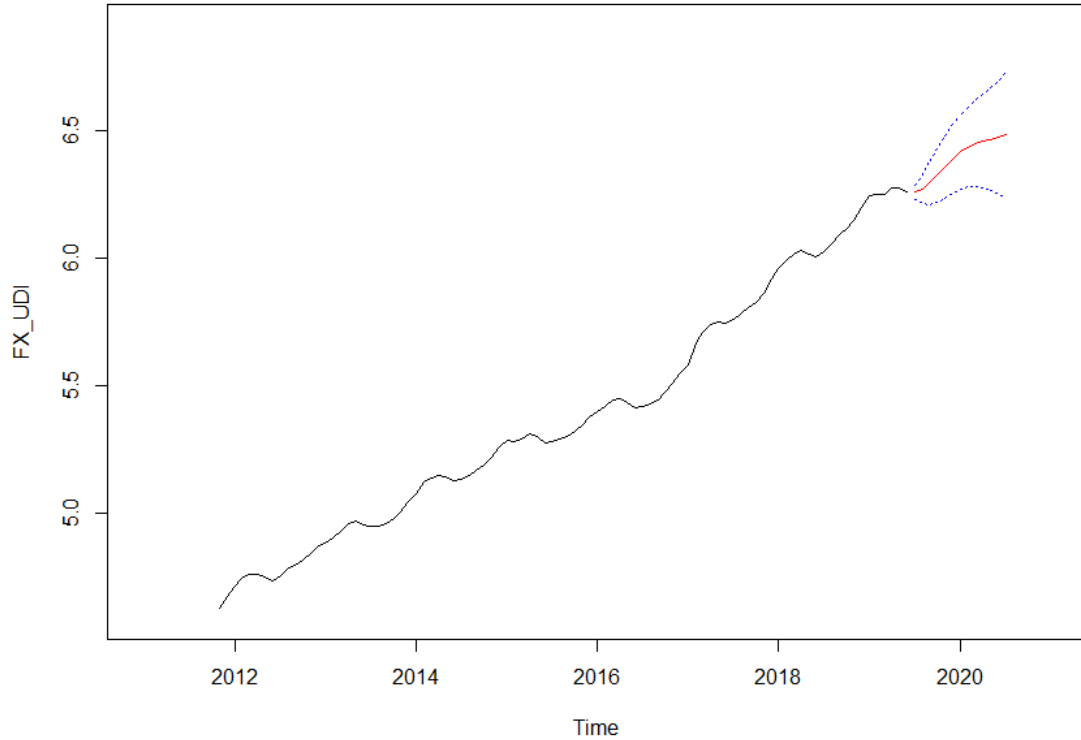


Figura 7.25: Pronóstico de modelo ARIMA(2,2,2) a la serie de tiempo generada por los valores de la udi.

$$\begin{bmatrix} \hat{a}_{cete28} & \hat{b}_{cete28} & \tilde{\sigma}_{cete28} & \hat{\sigma}_{cete28} & r_{cete28} \\ \hat{a}_{cete91} & \hat{b}_{cete91} & \tilde{\sigma}_{cete91} & \hat{\sigma}_{cete91} & r_{cete91} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \hat{a}_{Bonde5} & \hat{b}_{Bonde5} & \tilde{\sigma}_{Bonde5} & \hat{\sigma}_{Bonde5} & r_{Bonde5} \end{bmatrix} \quad (7.3)$$

3. Una vez construida esta matriz se realiza el pronóstico a 13 meses de las tasas de interés de acuerdo al modelo CIR ajustado, es decir al dado por la ecuación 4.13, mediante la discretización dada por la ecuación 7.2 y se guarda en una matriz con los pronósticos de las tasas.
4. Con la matriz de tasas se estima, mediante el producto de las matrices de emisiones por las tasa correspondientes¹³, el costo financiero de la deuda en moneda local a 13 meses de cada tipo de instrumento financiero que se colocará en una serie de vectores auxiliares, para finalmente calcular y almacenar el costo financiero como la suma de los costos por tipo de instrumento en un vector único.
5. El proceso de pronosticar 13 meses de tasas de interés y estimar el costo financiero de la deuda en moneda local se realiza 10,000 veces y con el costo financiero de cada escenario se va llenado la matriz con los costos financieros de cada escenario.
6. Una vez que se han calculado el costos financiero de cada escenario, lo que sigue es calcular el CaR, el CaR relativo y el CaR extremo de acuerdo a las definiciones del capítulo dos, 2.1, 2.2 y 2.3, respectivamente. Además, se realiza un histograma para poder observar el comportamiento del costo financiero.

¹³Dependiendo del tipo de tasa (compuesta o simple) y temporalidad del cupón se realiza la expresión en el código para la estimación del costo financiero. Ver capítulo 3.

7.3.1. Código para CaR de tasas de interés

```

%
-----

%                               Estimación del CaR con CIR
%
-----

clear ; close all; clc

pkg load statistics

load BASE_IR.mat % CARGAR INPUTS

pred=13; % Indica el número de periodos a pronosticar
T=13/12; % Indica el periodo de tiempo (años)
dt=1/12; % Indica la frecuencia de las observaciones en un año

Tasas=cat(2,cete28 , cete91 , cete182 , cete364 , btf3 , btf5 , btf10 , btf20 ,
          btf30 , udi3 , udi10 , udi30 ,fondeo)

% 1) Simulación de las tasas de interés
% 1.1) Estimación de parámetros

for c=1:13
    para=cirparameters(Tasas(:,c),dt);
    para2(c,:)=[para(1,1:4),Tasas(end,c)];
end

% 1.2) Pronósticos de las tasas de interés
alpha=(-para2(:,2))'; % Reversión
bita=para2(:,1)'; % Media
sigma=para2(:,4)'; % Desviación
R(1,:,1)=(para2(:,5)./100)'; % Punto de partida de la simulación
[row,column]=size(R);
N=T/dt;

% Simulación de las tasas de interés
for j=1:10000
    for i=1:N
        R(1,:)=R(1,:);
        R(i+1,:)=R(i,:)+(alpha.*(bita-R(i,:)).*(1/(1/dt)))+sigma.*abs(sqrt(R(
            i,:))).*sqrt(1/(1/dt)).*(normrnd(0,1).*ones(1,column)));
    end
    rendfut=R*100;

% 2) Estimación del costo financiero de la Deuda EMITIDA de Gobierno
% 2.1) Costo de la deuda que tiene tasa fija (se calculó en Excel)
costotf2018=83738131630.4279; %No actualizado. No es necesario ya que es
una cantidad determinista
costotf2019=211676408616.04; %No actualizado. No es necesario ya que es
una cantidad determinista
% 2.2) Costo de la deuda a tasa variable
costotv=((1.+(rendfut(2:end,13)./36000)).^28)-1).*bondeemitido;
costotv2018=sum(costotv(1:2,:));
costotv2019=sum(costotv(3:end,:));
% 2.3) Costo Financiero de la deuda Emitida del Gob Federal
costoemitido2018(j,1)=costotf2018+costotv2018;

```

```

costoemitido2019(j,1)=costotf2019+costotv2019;

% 3) Estimación del Costo Financiero de la Deuda POR EMITIRSE del Gob
% 3.1) Costo de la Deuda a tasa fija
% El precio futuro de la UDI se calculó con un ARIMA
% 3.1.1) Cetes
vencimiento=[(28/36000)*ones(pred,1),(91/36000)*ones(pred,1),(182/36000)*
ones(pred,1),(364/36000)*ones(pred,1)];
costocete=(cete./(1.+(rendfut(2:end,1:4).*vencimiento)))*1000000;
costocete2018=sum(sum(costocete.*calencete2018),2);
costocete2019=sum(sum(costocete.*calencete2019),2);
% 3.1.2) BTF
costobtf=btf.*(182/36000).*rendfut(2:end,5:9).*1000000;
costobtf2018=sum(sum(costobtf.*calenbtf2018),2);
costobtf2019=sum(sum(costobtf.*calenbtf2019),2);
% 3.1.3) Udibonos
costoudi=udibono.*(182/36000).*rendfut(2:end,10:12).*udi.*1000000;
costoudi2018=sum(sum(costoudi.*calenudi2018),2);
costoudi2019=sum(sum(costoudi.*calenudi2019),2);
% 3.2) Costo de la Deuda a tasa variable (bonde)
costobonde=((1.+(rendfut(2:end,13)./36000)).^28)-1).*bonde.*1000000;
costobonde2018=sum(costobonde(1:2,:)); %No requiere calendario de
cupones ya que paga cada 28 días
costobonde2019=sum(costobonde(3:end,:));
% 3.3) Costo Financiero de la Deuda por emitirse
costoemitirse2018(j,1)=costocete2018+costobtf2018+costoudi2018+
costobonde2018;
costoemitirse2019(j,1)=costocete2019+costobtf2019+costoudi2019+
costobonde2019;

% 4) Estimación de los costos totales
costototal2018(j,1)=costoemitido2018(j,1)+costoemitirse2018(j,1);
costototal2019(j,1)=costoemitido2019(j,1)+costoemitirse2019(j,1);

end

% 5) CaR al 95%
costoesperado2019=mean(costototal2019);
CaR952019=prctile(costototal2019,95);

% 6) CaR relativo al 95%
CaR_relativo_2019=CaR952019/costoesperado2019;

% 7) CaR Relativo Extremo
AAA=sort(costototal2019);
BBB=AAA(.95*length(AAA)+1:end,:);
CaRExtremo2019=mean(BBB)/costoesperado2019;

hist(costototal2019,1000)

7.3.2. Código para hallar parámetros en modelo CIR

function [parametros]=cirparameters(r100, dt)
r=r100/100; % Convierte las tasas a forma decimal
n=length(r);
Y=r(2:n,1);
X=r(1:n-1,1);
X=[ones(n-1,1),X];

Beta=(inv(X'*X))*X'*Y; % Estimadores por mínimos cuadrados de los
parámetros de la regresión lineal (beta0;beta1)
Yest=X*Beta; % Y estimada

```

```

e=Y-Yest; % Estimacion de los errores
errorcuadmedio=mean((e*100).^2); %ECM=(1/n)*sum[(Yest-Y)^2]
errorabsmedio=mean(abs(e)); %ECM=(1/n)*sum[|Yest-Y|]

B=(1/dt)*(log(Beta(2,1)));
A=(B/(exp(B*dt)-1))*Beta(1,1);

n0=(A/(2*B^2))*((exp(B*dt)-1)^2);
n1=(1/B)*exp(B*dt)*(exp(B*dt)-1);

z=X*[n0;n1]; % z matriz de (n-1)x1
ez=e./z;

sigmaa=sqrt((ez'*e)/(length(e))); % (1/n*sum(eolsz)^2)^(1/2)
sigmab=sqrt((inv(z'*z))*z'*(e.^2)); % [(1/sum(z)^2)*(sum(z*eols^2))]^(1/2)

parametros(1,:)=[-A/B,B,sigmaa,sigmab,errorcuadmedio,errorabsmedio];

end

```

Luego de correr el código en Matlab con todos los inputs correctamente, los resultados obtenidos son los siguientes:

1. Costo en riesgo esperado que es la esperanza de la variable aleatoria que mide el costo financiero, es decir, $\mathbb{E}[\Lambda]$.
2. Costo en riesgo al 95 % dado por la ecuación 2.1.
3. Costo en riesgo relativo dado por la ecuación 2.2.
4. Costo en riesgo extremo dado por la ecuación 2.3.
5. Histograma de la variable aleatoria que mide el costo financiero.

El costo esperado para los próximos 13 meses es de 299,980 millones de pesos, así mismo el CaR al 95 % es de 305,740 millones de pesos por lo que se puede decir que el costo financiero generado por la deuda en moneda local no deberá superar los 306 billones de pesos. El Costo en riesgo tanto relativo al 95 % como extremo fueron respectivamente 1.0192 y 1.0245 con respecto CaR al 95 %. Ver figura 7.27. Con estas medidas, se puede tener idea de la cantidad de dinero necesaria para que el gobierno haga frente a sus obligaciones por bonos en moneda local para los meses futuros y poder así planificar la administración. El histograma de la variable aleatoria que mide el costo financiero se muestra a continuación, figura (7.26).

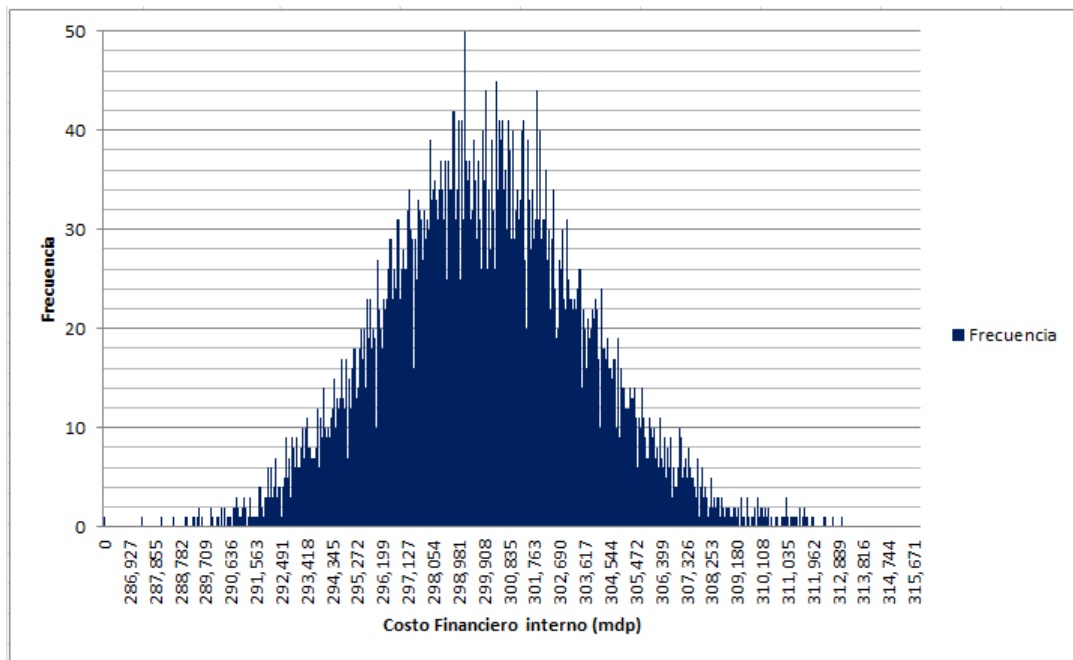


Figura 7.26: Histograma de la variable aleatoria que mide el costo financiero generado por la deuda pública en moneda local.

```

>> % 5) CaR al 95%
>> costoesperado2019=mean(costototal2019)
costoesperado2019 = 2.9998e+11
>> CaR952019=prctile(costototal2019,95)
CaR952019 = 3.0574e+11
>> % 6) CaR relativo al 95%
>> CaR_relativo_2019=CaR952019/costoesperado2019
CaR_relativo_2019 = 1.0192
>> % 7) CaR Relativo Extremo
>> AAA=sort(costototal2019);
>> BBB=AAA(.95*length(AAA)+1:end,:);
>> CaRExtremo2019=mean(BBB)/costoesperado2019
CaRExtremo2019 = 1.0245
>>

```

Figura 7.27: Resultados del código para el CaR en moneda local.

Capítulo 8

Conclusiones

Con el objetivo de realizar la estimación del costo en riesgo, en este trabajo se presentaron diversos modelos para realizar pronósticos acerca de tasas de interés y de tipo de cambio. Además de los modelos presentados en este trabajo, se pueden considerar modelos multivariados[34] e incluso se pueden realizar modelos no paramétricos conjuntos o individuales[14]. La estimación de un costo en riesgo no está limitada a deuda gubernamental, puede ser también para la deuda de una empresa en moneda local o extranjera, así como para medir modelos que involucren tasas de interés o tipos de cambio.

De manera general, el presente trabajo ha servido para presentar los fundamentos teóricos sobre el cual descansa la estimación del costo en riesgo (ARIMA para los tipos de cambio y Cox-Ingersoll-Ross para las tasas de interés).

En el primer capítulo se presentaron de manera general el concepto de deuda pública, mencionando como se divide. Así mismo, se discutió la composición de la Deuda Federal, la deuda del Sector Público y el SHRFSP. Finalmente, se presentaron algunas cifras, como el saldo de la deuda neta y la comparación con respecto a otros miembros del la OCDE.

En el segundo capítulo, se presentó el concepto de costo en riesgo (al $(1 - \alpha)\%$, relativo y extremo), así como sus respectivas definiciones matemáticas, se hizo mención de su utilidad para poder medir los riesgos asociados a las variables de mercado y poder así tener una herramienta con la cual se puedan tomar decisiones. Se mencionó la estructura y características de la deuda pública del gobierno federal y se realizó una descripción de los principales instrumentos que componen el portafolio de la deuda pública y con los cuales se estimó el CaR de la deuda pública.

El capítulo tres se presentaron los modelos básicos para modelar tasas de interés, tanto de manera básica (a tasa determinista) como de manera estocástica. Adicionalmente, se habló de la ausencia de arbitraje y las consecuencia de la existencia de una medida equivalente con lo cual se pudo transformar un proceso a otro que sea más sencillo de valorar, lo que proporcionó una herramienta con la cual posteriormente se pudieron encontrar los estimadores de los parámetros del modelo de tasa de interés.

En el cuarto capítulo se presentó la dinámica del modelo de Cox–Ingersoll–Ross, así como sus principales propiedades. Posteriormente, mediante un cambio de medida el modelo es transformado en un nuevo modelo el cual es utilizado para la estimación de los parámetros

mediante la propuesta de Rydén y Overbeck, la cual fue descrita a detalle.

En el quinto capítulo se presentaron los conceptos básicos del estudio de las series de tiempo, después se mencionaron los conceptos de proceso causales e invertibles. Finalmente se presentaron los modelos MA(q), AR(p) y ARIMA(p,q) así como sus principales propiedades.

En el capítulo seis, primeramente se presentan herramientas, como las ACF y PACF, con el fin de identificar el modelo adecuado a la serie de tiempo. Posteriormente se presentan algunas técnicas para lograr realizar pronósticos, las cuales se utilizan con frecuencia en los paquetes estadísticos. Luego se presentó el proceso de estimación de los parámetros de los modelos de series de tiempo mediante el método de máxima verosimilitud. Finalmente, se mencionó la manera en que se debe realizar la validación del modelo con el fin de contar con la seguridad de que el ajuste es adecuado. Así mismo, se hizo una mención de manera muy general de los modelos ARIMA y SARIMA, cuya brevedad tiene que ver con no querer perderse del objetivo del presente trabajo.

El penúltimo capítulo se aterriza todo lo visto durante el trabajo para finalmente proceder a realizar la estimación del CaR, primero se mencionan los pasos a seguir de manera general. Posteriormente, se describieron los insumo utilizados para estimación y finalmente se presentan los códigos utilizados tanto en R, para el ajuste ARIMA a las series de tiempo generadas por los tipos de cambio, como en Matlab con los códigos que realizan la estimación, así como los códigos auxiliares que estiman los parámetros del modelo CIR y aquellos que regresan los datos diferenciados a la serie original.

Todo lo visto en este trabajo tiene el objeto de darle un soporte teórico a la estimación del costo en riesgo, ya que en el documento *Estimación del Costo en Riesgo para la Deuda Pública de México*[14] que es el documento que presenta la descripción de la estimación del CaR no presenta el trasfondo teórico que conlleva el proceso. Debido a la gran extensión que abarcan los temas se presentaron de manera muy concreta solamente los elementos necesarios para poder llevar a cabo la estimación del CaR y es de señalar que la teoría de los modelos presentada en este trabajo es solo una pequeña parte de la gran variedad de modelos que existen tanto en el estudio de series de tiempo como en el de tasas de interés. Ya que en lo que al estudio de las series de tiempo concierne, existen los modelos de heterocedasticidad condicional en los errores como los modelos ARCH y GARCH, así mismo modelos de series de tiempo en varias variables, como son, los modelos VAR, SVAR, SVARMA y modelos no lineales. Por otra parte, en lo que respecta al análisis de tasas de interés existen modelos de un factor como el Vasicek, Dothal son otros que se pueden ajustar así como modelos más complejos o generalizaciones como el modelo de Hull-White. Además, se puede agregar que ya que la los movimientos en las tasas de interés se encuentran correlacionados con los tipos de cambio es natural pensar que se pueden modelar tanto los tipos de cambio como las dinámicas de tasas conjuntamente. Todas estos modelos teóricos hacen que la estimación del CaR o de cualquier otra medida de sobre costo financiero o valuación se pueda realizar en formas muy diversas. En este trabajó se presentó la manera en que lo realiza la SHCP y a pesar de la existencia de modelos con mayor complejidad no quiere decir que estos sean mejores modelos, de hecho ya que se validaron los supuestos, los modelos se consideran adecuados.

La base teórica presentada es lo suficientemente robusta para respaldar la estimación del costo en riesgo de la deuda pública generando números coherentes que proporcionan una buena referencia de la cantidad de dinero que necesitaría el gobierno para solventar sus

obligaciones, en lo que respecta al pago de deuda.

El CaR al 95 % de tipo de cambio (deuda externa) fue de 86 mil millones de pesos mientras que de tasa de interés (deuda interna) fue de 306 mil millones pesos, sumados son aproximadamente 392 mil millones de pesos que valorizados al tipo de cambio promedio de diciembre de 2018 son aproximadamente 19,508 millones de dólares lo que representa 1,6 % del PIB del año 2018 de acuerdo al Banco mundial¹.

Las medidas como el CaR son necesarias para que las entidades financieras puedan contar con herramientas que ayuden en la toma de decisiones que permitan optimizar los recursos, ya sea replanteando la estrategia de endeudamiento u optando por laguna otra alternativa. Este trabajo no solo ha permitido conocer esta medida, sino además conocer el trasfondo teórico que esta conlleva.

¹Sitio web: <https://datos.bancomundial.org/pais/mexico>

Capítulo 9

Apéndice

Espacio de probabilidad filtrado

Un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ es una terna en donde Ω es el espacio muestral, \mathcal{F} es una σ -álgebra de subconjuntos de Ω y \mathbb{P} es una medida de probabilidad sobre \mathcal{F} .

Una filtración es una colección de σ -álgebras $\{\mathcal{F}_n\}_{n \geq 1}$ tal que $\mathcal{F}_n \subseteq \mathcal{F}_m$ si $n \leq m$. Una filtración puede ser interpretada como la acumulación de información en el tiempo. Un proceso genera una filtración.

Un espacio de probabilidad filtrado es un espacio de probabilidad equipado con una filtración y es escrito como $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathbb{P})$. A $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ puede ser vista como historia del proceso hasta el tiempo t .

Proceso adaptado

Sea $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathbb{P})$ un espacio de probabilidad filtrado, un proceso $\{x(t)\}_{t \geq 0}$ se dice que es adaptado a la filtración \mathcal{F}_t (también llamado \mathcal{F}_t -adaptado) si es \mathcal{F}_t -medible, es decir, si para algún $\omega \in \Omega$ entonces $x^{-1}(\omega) \in \mathcal{F}_t$.

Proceso de Wiener

Un proceso estocástico a tiempo real, $\{W(t)\}_{t \geq 0}$ se llama proceso de Wiener o movimiento browniano si satisface:

1. $W(0) = 0$ c.s.
2. Los incrementos de $W(t)$ son independientes, esto es, para cualesquiera conjunto de tiempos finito, $0 \leq t_1 < t_2 < \dots < t_n$ entonces las variables aleatorias:

$$W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_n} - W_{t_{n-1}}$$

son independientes.

3. Para cualesquiera $0 \leq s \leq t$ el incremento $W_s - W_t$ tiene distribución gaussiana con media 0 y varianza $t - s$.

4. W_t tiene trayectorias continuas c.s.

Un proceso de Wiener es un proceso especial con drift cero y varianza proporcional a la longitud del intervalo de tiempo. Tsay[34] explica que esto significa que la tasa de cambio en la esperanza es cero, mientras que el cambio en la tasa de varianza es igual a uno. En la práctica la media y varianza de un proceso estocástico puede evolucionar de una manera más compleja. De esta manera se considera el **proceso de Wiener generalizado**, el cual tiene tasa de drift μ y tasa de cambio de varianza σ^2 . Por tanto, un proceso de Wiener generalizado x_t satisface:

$$dx_t = \mu dt + \sigma dW_t$$

donde W_t es un proceso de Wiener. Considerando la versión discretizada de la ecuación anterior se obtiene lo siguiente:

$$x_t - x_0 = \mu t + \sigma \epsilon \sqrt{t}$$

para el incremento de 0 a t . Donde ϵ es una v.a. normal(0,1). En consecuencia:

$$\mathbb{E}[x_t - x_0] = \mu t \quad \text{Var}(x_t - x_0) = \sigma^2 t$$

Lo cual muestra que el proceso tiene tasa de drift μ y tasa de cambio de varianza σ^2 .

Proceso de Ito

Los parámetros de drift (deriva) y volatilidad de un proceso de Wiener generalizado son constantes en el tiempo. Pero el modelo puede ser extendido si μ y σ se convierten en funciones del proceso $x(t)$, entonces se tiene un proceso de Ito. Específicamente si $x(t)$ es un proceso de Ito, entonces satisface:

$$dx(t) = \mu(t, x(t)) + \sigma(t, x(t))dW(t)$$

Donde $W(t)$ es un proceso de Wiener. El proceso también puede ser escrito en su forma integral de la siguiente forma:

$$x(t) = x(0) + \int_0^t \mu(s, x(s))ds + \int_0^t \sigma(s, x(s))dW(s)$$

Donde $x(0)$ denota el valor inicial del proceso al tiempo 0. Un proceso de Wiener es un caso particular de un proceso de Ito con $\mu(t, x(t)) = 0$ y $\sigma(t, x(t)) = 1$.

Lema de Ito

Recordatorio de diferenciación:

Sea $G(x) : \mathbb{R} \rightarrow \mathbb{R}$ tal que $G \in C^{\infty 1}$. Usando la serie de Taylor alrededor del cero se tiene:

$$\Delta G \equiv G(x + \Delta x) - G(x) = \frac{dG}{dx} \Delta x + \frac{1}{2} \frac{d^2 G}{dx^2} (\Delta x)^2 + \frac{1}{6} \frac{d^3 G}{dx^3} (\Delta x)^3 + \dots \quad (9.1)$$

Tomando el límite $\Delta x \rightarrow 0$ e ignorando los términos de orden dos y más entonces:

$$\Delta G = \frac{dG}{\delta x} \Delta x$$

Ahora si $G(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$, $G \in C^{\infty}$, entonces:

$$\Delta G = \frac{dG}{dx} \Delta x + \frac{dG}{dy} \Delta y + \frac{1}{2} \frac{d^2 G}{dx^2} (\Delta x)^2 + \frac{1}{2} \frac{d^2 G}{dy^2} (\Delta y)^2 + \frac{1}{2} \frac{d^2 G}{dxdy} (\Delta x \Delta y) + \dots$$

Tomando los límites $\Delta x \rightarrow 0$ y $\Delta y \rightarrow 0$ e ignorando los términos de orden superior entonces:

$$\Delta G = \frac{dG}{dx} \Delta x + \frac{dG}{dy} \Delta y$$

Ahora se verá el caso cuando $G(x_t, t) \in C^{\infty}$ para x_t, t pero x_t es un proceso de Ito. El proceso de Ito con parámetros constantes discretizado tiene la siguientes forma:

$$\Delta x = \mu \Delta t + \sigma \sqrt{\Delta t} \epsilon_t$$

Donde ϵ_t se distribuye como una v.a. normal(0,1). De la ecuación anterior se tiene que:

$$(\Delta x)^2 = \mu^2 (\Delta t)^2 + \sigma^2 \Delta t \epsilon_t^2 + 2\mu\sigma \Delta t^{3/2} = \sigma^2 \epsilon_t^2 \Delta t + H(\Delta t) \quad (9.2)$$

Donde $H(\Delta t)$ denota los términos de orden superior de Δt . Este resultado muestra que $(\Delta x)^2$ contiene términos de $(\Delta t)^2$ que no puede ser ignorados, cuando se toma el límite $\Delta x \rightarrow 0$. Tsay [34] muestra las siguientes dos propiedades del primer término del resultado anterior:

$$\mathbb{E}[\sigma^2 \epsilon_t^2 \Delta t] = \sigma^2 \Delta t$$

$$Var(\sigma^2 \epsilon_t^2 \Delta t) = \mathbb{E}[\sigma^4 \epsilon_t^4 (\Delta t)^2] - \mathbb{E}^2[\sigma^2 \epsilon_t^2 \Delta t] = 2\sigma^4 (\Delta t)^2$$

¹Se dice que una función f es C-infinito, $f \in C^{\infty}$, si es diferenciable una infinidad de veces.

De estas dos propiedades se puede demostrar que $\sigma^2 \epsilon_t^2 \Delta t$ converge a la cantidad no estocástica $\sigma^2 \Delta t$ y por lo tanto de 9.2 se tiene que:

$$(\Delta t)^2 \rightarrow \sigma^2 dt \text{ cuando } \Delta t \rightarrow 0.$$

Por lo tanto, usando la ecuación 9.1 y la definición del lema de Ito, se obtiene lo siguiente:

$$\begin{aligned} dG &= \frac{dG}{dx} \Delta x + \frac{dG}{dt} \Delta t + \frac{1}{2} \frac{d^2 G}{dx^2} \sigma^2 \Delta t + \frac{1}{2} \frac{d^2 G}{dt^2} \Delta t^2 + \frac{1}{2} \frac{d^2 G}{dx dt} \Delta x \Delta t + \dots \\ &= \frac{dG}{dx} \Delta x + \frac{dG}{dt} \Delta t + \frac{1}{2} \frac{d^2 G}{dx^2} \sigma^2 \Delta t \text{ ya que los terminos de orden superior convergen a 0.} \\ &= \left(\frac{dG}{dx} \mu + \frac{dG}{dt} + \frac{1}{2} \frac{d^2 G}{dx^2} \right) \Delta t + \frac{dG}{dx} \sigma \Delta W_t \text{ ya que } dx = \mu dt + \sigma dW_t. \end{aligned} \tag{9.3}$$

Este es el lema de Ito en cálculo estocástico

A continuación se enunciará de el resultado completo de manera breve, cambiando la notación de Δt a dt :

Sea x_t un proceso de Ito. Si $G(x_t, t) \in C^2$ entonces:

$$dG = G_t(x_t, t)dt + G_x(x_t, t)dx + \frac{1}{2} G_{xx}(x_t, t)\sigma^2 dt \tag{9.4}$$

o en su forma integral:

$$G(x_t, t) = G(0, 0) + \int_0^t \frac{dG(x_s, s)}{dx} dx_s + \int_0^t \frac{dG(x_s, s)}{dt} ds + \frac{1}{2} \int_0^t \frac{d^2 G(x_s, s)}{dx^2} ds dx_s \tag{9.5}$$

Teorema de Girsanov y derivada de Radon-Nikodim

Sean \mathbb{P} y \mathbb{Q} dos medidas de probabilidad sobre el mismo espacio filtrado $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T})$. Se dice que \mathbb{Q} es equivalente con respecto a \mathbb{P} si tienen los mismos conjuntos nulos, esto es, $\mathbb{Q}[A] = 0 \iff \mathbb{P}[A] = 0$.

El teorema de Girsanov dice que cuando dos medidas son equivalentes, es posible expresar la primera en términos de la segunda a través de la derivada de Radon-Nikodim, de hecho existe una martingala m_t en $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathbb{P})$ tal que:

$$\mathbb{Q}[A] = \int_A m_t d\mathbb{P}[w] = \mathbb{E}_{\mathbb{P}}[I_A m_t]$$

La cual puede ser escrita de manera más concisa como:

$$\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_t} = m_t$$

Donde \mathcal{F}_t es la filtración generada por m hasta el tiempo t .

Distribución Xi-Cuadrada No Central

En el capítulo 4 se mencionó que la función de distribución de las tasas de interés $r(t)$ que sigue el modelo de Cox-Ingersoll-Ross sigue una distribución Xi-Cuadrada no central $\chi^2_{(2q+2, 2u)}$, con $2q + 2$ grados de libertad y parámetro de no centralidad $2u$. En este momento se enunciará la naturaleza de esta distribución.

Sea $X \sim N_n(\mu, I_n)$, entonces la v.a. $Y = X^T X$ tiene función de densidad:

$$f(X) = \sum_{k=0}^{\infty} \exp\{-\mu^*\} \frac{\mu^{*k}}{k!} \frac{X^{\frac{n+2k}{2}-1} \exp^{-X/2}}{\Gamma(\frac{n+2k}{s}) 2^{\frac{n+2k}{s}}} \text{ si } X > 0 \text{ y } 0 \text{ e.o.c.}$$

Siendo $\mu^* = \frac{1}{2} \mu^T \mu$. Entonces se dirá que Y se distribuye Xi-Cuadrada no central de parámetros n (grados de libertad) y μ^* (parámetro de no centralidad) y se escribe $Y \sim \chi^2(n, \mu^*)$.

Función de densidad condicional

Sea (X, Y) un vector aleatorio con función de densidad $f_{X,Y}(x, y)$, y sea y tal que $f_Y(y) \neq 0$. La función:

$$f_{X|Y} x|y = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

es conocida como la función de densidad condicional de la v.a. X condicionada a la v.a. Y .

Ajuste de modelos ARIMA

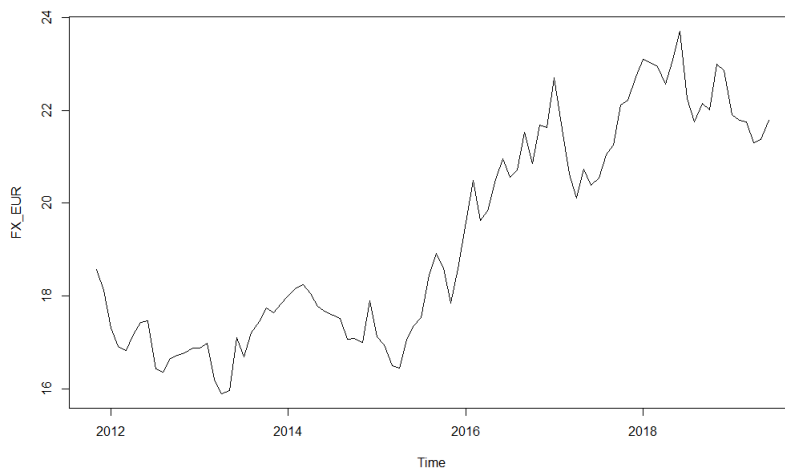


Figura 9.1: Serie de tiempo generada por el tipo de cambio EUR/MXN.

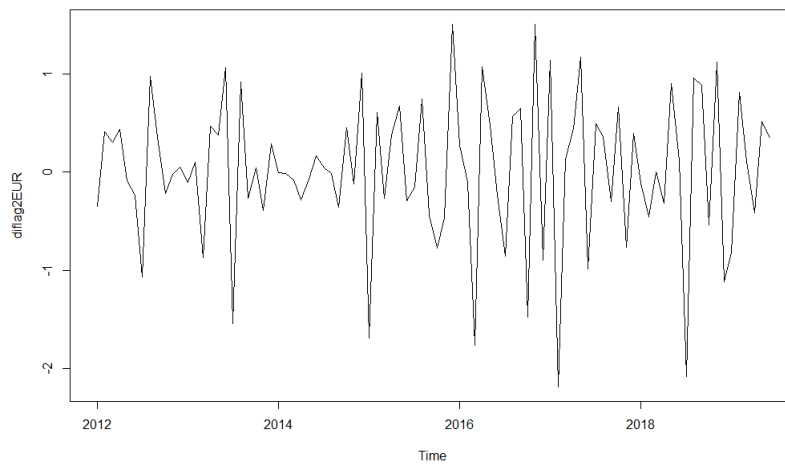


Figura 9.2: Serie diferenciada dos veces del tipo de cambio EUR/MXN.

```
> adf.test(diflag2EUR, k=0)

Augmented Dickey-Fuller Test

data: diflag2EUR
Dickey-Fuller = -14.745, Lag order = 0, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(diflag2EUR, k = 0) : p-value smaller than printed p-value
```

Figura 9.3: Prueba de Dickey-Fuller para determinar la existencia de raíces unitarias a la serie diferenciada del tipo de cambio EUR/MXN.

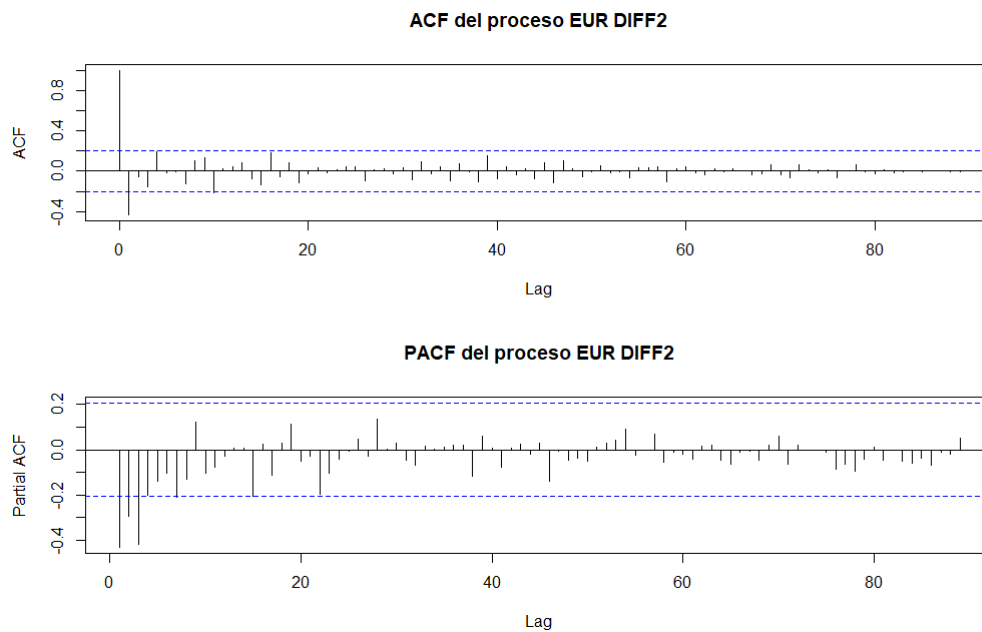


Figura 9.4: ACF y PACF de la serie de tiempo diferenciada del tipo de cambio EUR/MXN para determinar modelo.

```
> ARIMA321=arima(FX_EUR, order=c(3,2,1), include.mean = TRUE)
> ARIMA321

Call:
arima(x = FX_EUR, order = c(3, 2, 1), include.mean = TRUE)

Coefficients:
      ar1      ar2      ar3      ma1
    -0.0025 -0.1289 -0.1538 -0.9873
s.e.    0.1257  0.1159  0.1179  0.2358

sigma^2 estimated as 0.282:  log likelihood = -72.83,  aic = 155.67
```

Figura 9.5: Ajuste de modelo ARIMA a la serie de tiempo del tipo de cambio EUR/MXN.

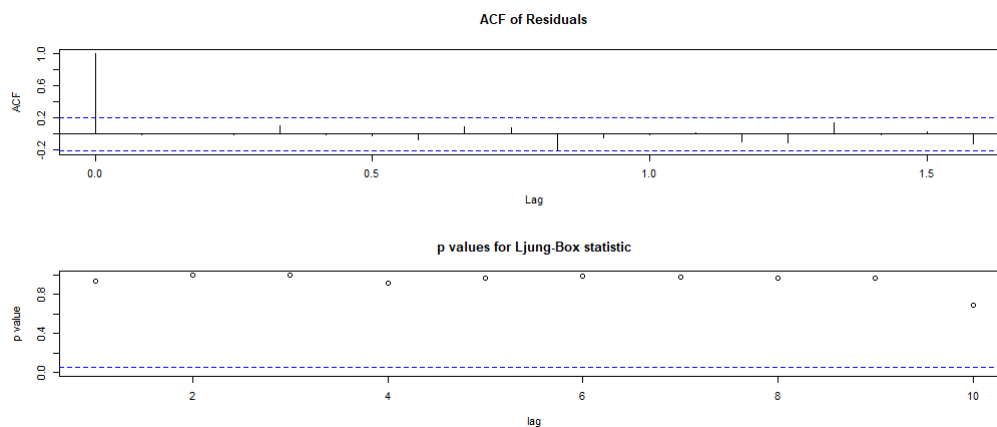


Figura 9.6: ACF y estadísticos de Ljung-Box para los distintos lags de los residuales del modelo ajustado.

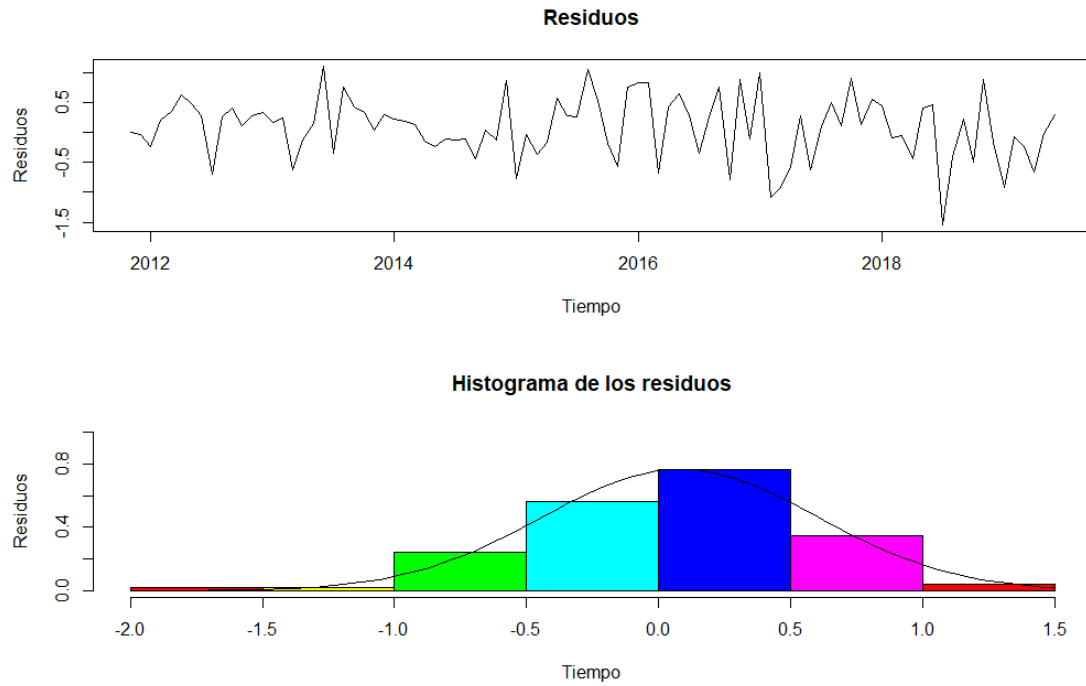


Figura 9.7: Serie generada de los residuales e histograma con densidad de una distribución normal ajustada.

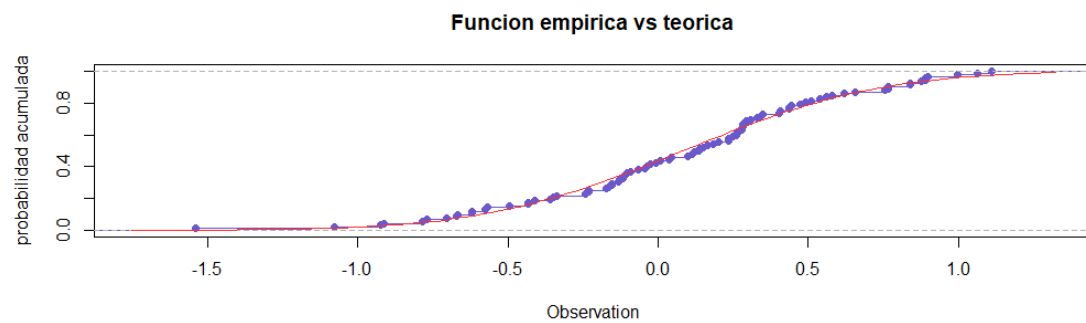


Figura 9.8: Distribución empírica generada por los residuales contra distribución teórica normal ajustada.

```
> ks.test(residuos, pnorm, mean(residuos),sd=sqrt(var(residuos)))

One-sample Kolmogorov-Smirnov test

data:  residuos
D = 0.062789, p-value = 0.8388
alternative hypothesis: two-sided

> lillie.test(residuos)

Lilliefors (Kolmogorov-Smirnov) normality test

data:  residuos
D = 0.062789, p-value = 0.4976

> ad.test(residuos)

Anderson-Darling normality test

data:  residuos
A = 0.35801, p-value = 0.4462
```

Figura 9.9: Pruebas de Kolmogorov-Smirnov, Lilliefors y Anderson-Darling para verificar normalidad en residuales.

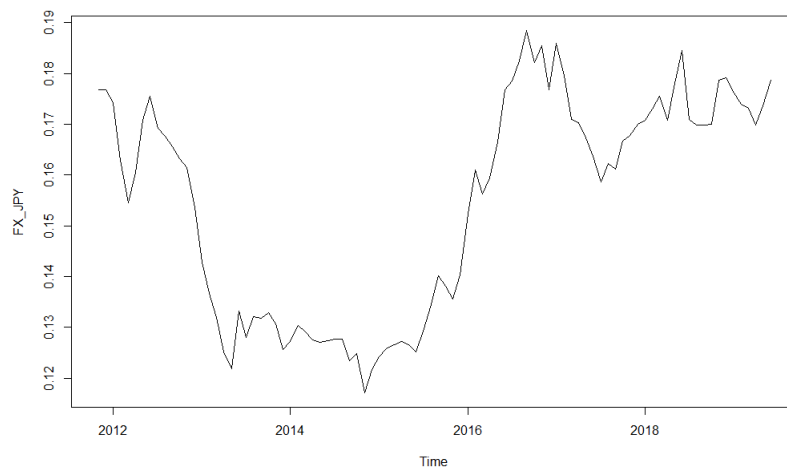


Figura 9.10: Serie de tiempo generada por el tipo de cambio JPY/MXN.

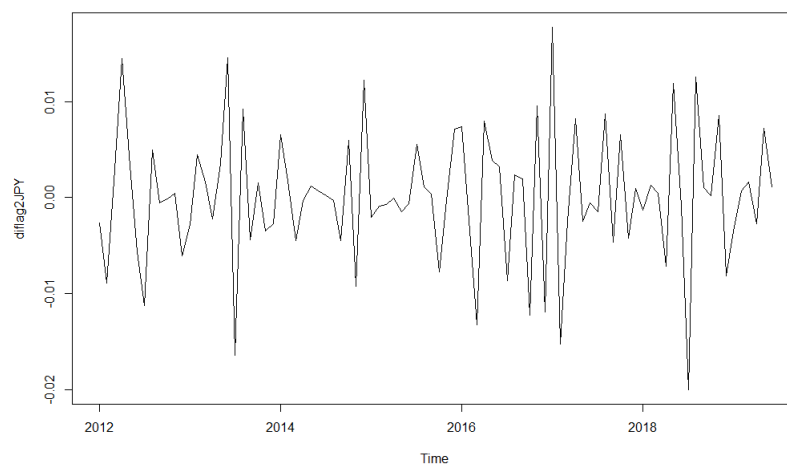


Figura 9.11: Serie diferenciada dos veces del tipo de cambio JPY/MXN.

```
> adf.test(diflag2JPY, k=0)

Augmented Dickey-Fuller Test

data: diflag2JPY
Dickey-Fuller = -14.523, Lag order = 0, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(diflag2JPY, k = 0) : p-value smaller than printed p-value
```

Figura 9.12: Prueba de Dickey-Fuller para determinar la existencia de raíces unitarias a la serie diferenciada del tipo de cambio JPY/MXN.

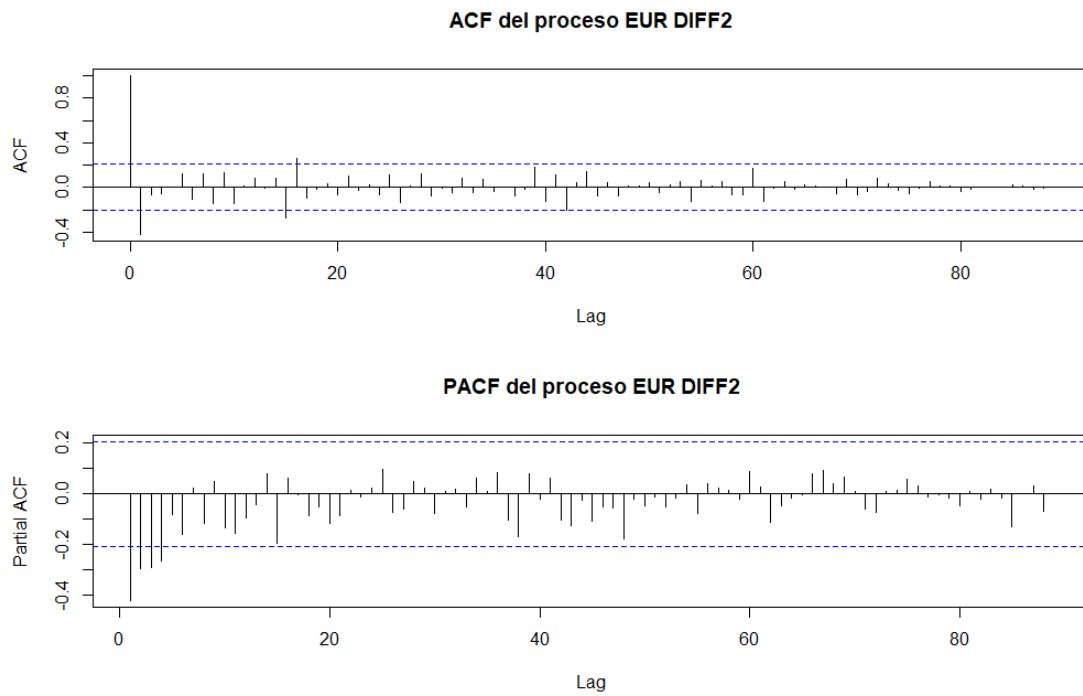


Figura 9.13: ACF y PACF de la serie de tiempo diferenciada del tipo de cambio JPY/MXN para determinar modelo.

```
> ARIMA121=arima(FX_JPY, order=c(1,2,1), include.mean = TRUE)
> ARIMA121

Call:
arima(x = FX_JPY, order = c(1, 2, 1), include.mean = TRUE)

Coefficients:
      ar1      ma1
  0.1211 -0.9526
s.e.  0.1149  0.0508

sigma^2 estimated as 2.766e-05:  log likelihood = 343.51,  aic = -681.03
```

Figura 9.14: Ajuste de modelo ARIMA a la serie de tiempo del tipo de cambio JPY/MXN.

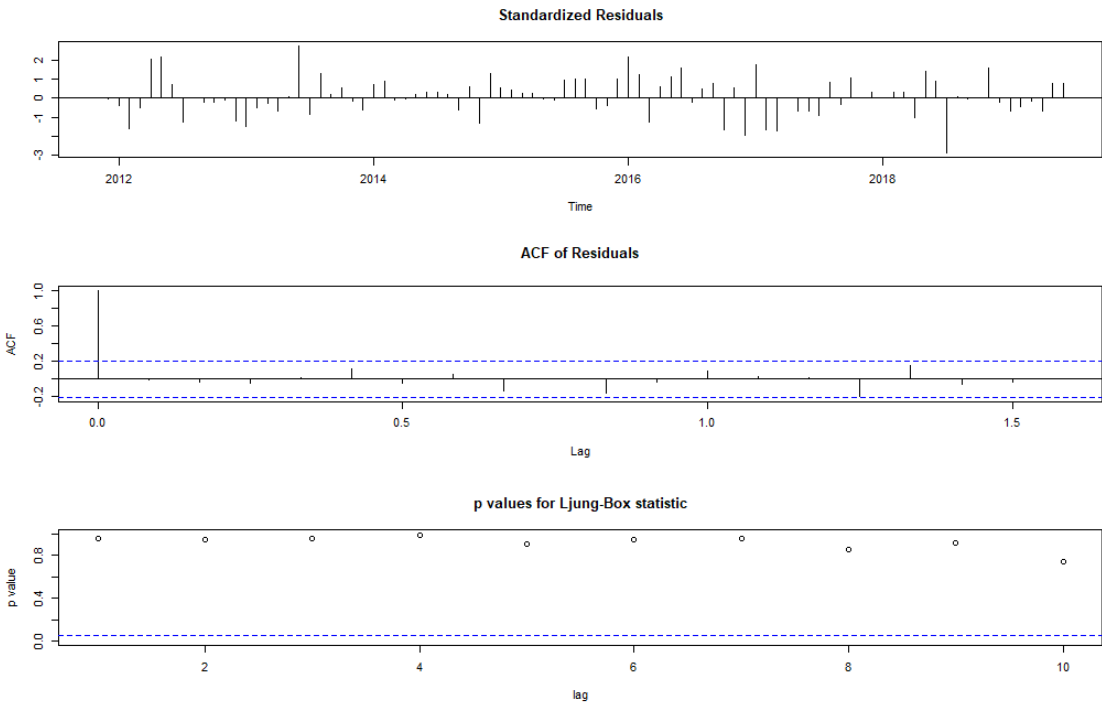


Figura 9.15: ACF y estadísticos de Ljung-Box para los distintos lags de los residuales del modelo ajustado.

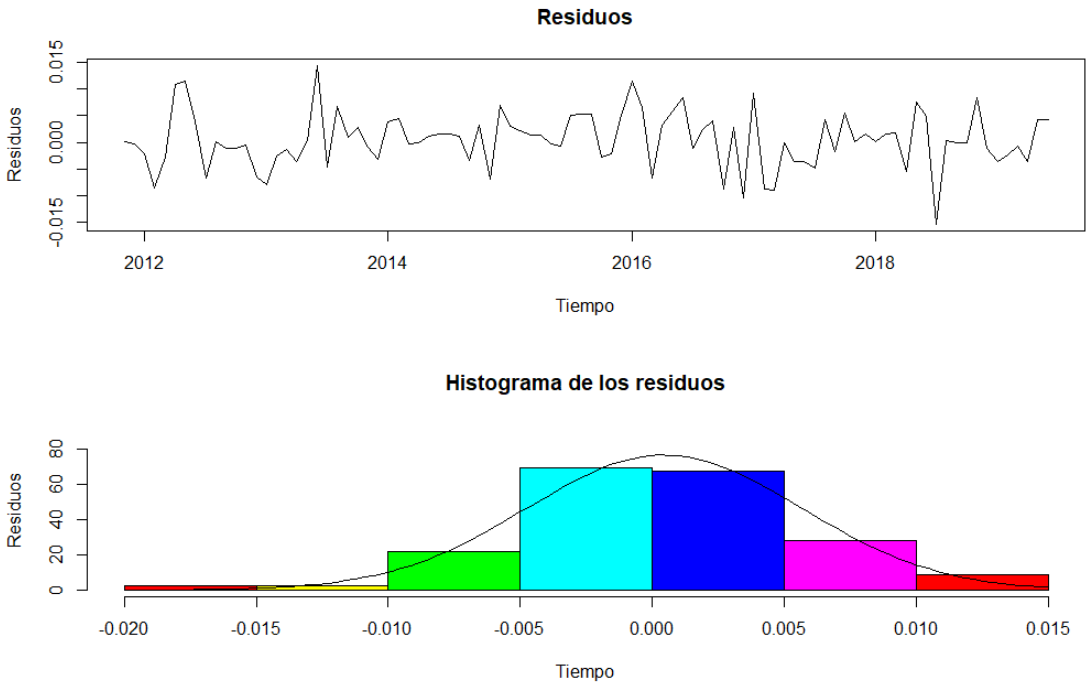


Figura 9.16: Serie generada de los residuales e histograma con densidad de una distribución normal ajustada.

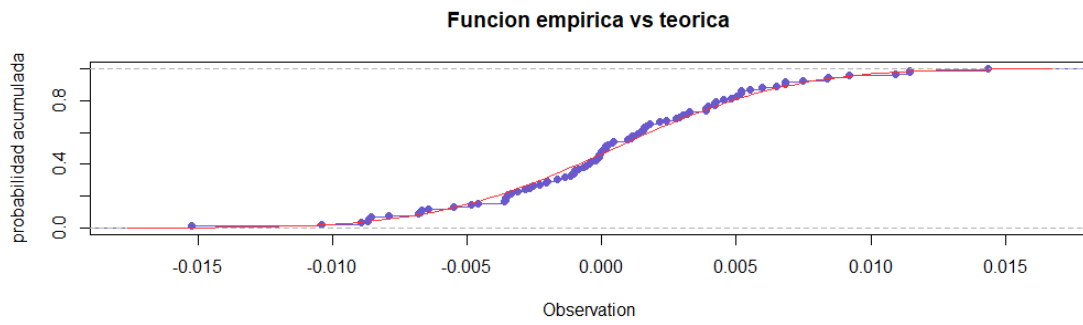


Figura 9.17: Distribución empírica generada por los residuales contra distribución teórica normal ajustada.

```
> ks.test(residuos, pnorm, mean(residuos),sd=sqrt(var(residuos)))

One-sample Kolmogorov-Smirnov test

data:  residuos
D = 0.063722, p-value = 0.8259
alternative hypothesis: two-sided

> lillie.test(residuos)

Lilliefors (Kolmogorov-Smirnov) normality test

data:  residuos
D = 0.063722, p-value = 0.4734

> ad.test(residuos)

Anderson-Darling normality test

data:  residuos
A = 0.35341, p-value = 0.4572
```

Figura 9.18: Pruebas de Kolmogorov-Smirnov, Lilliefors y Anderson-Darling para verificar normalidad en residuales.

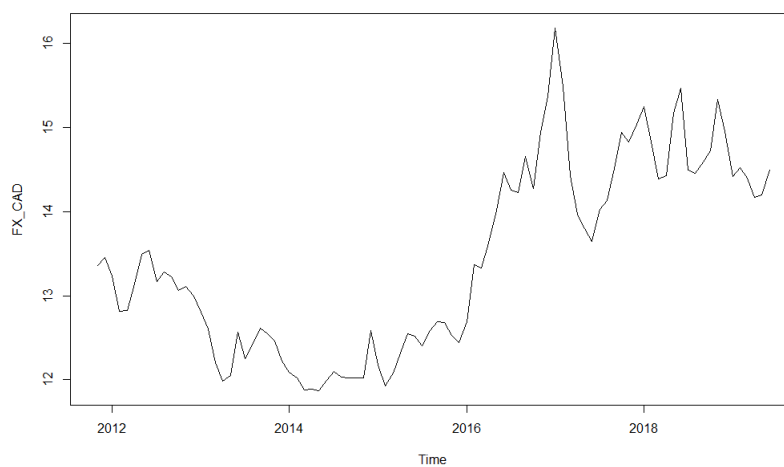


Figura 9.19: Serie de tiempo generada por el tipo de cambio CAD/MXN.

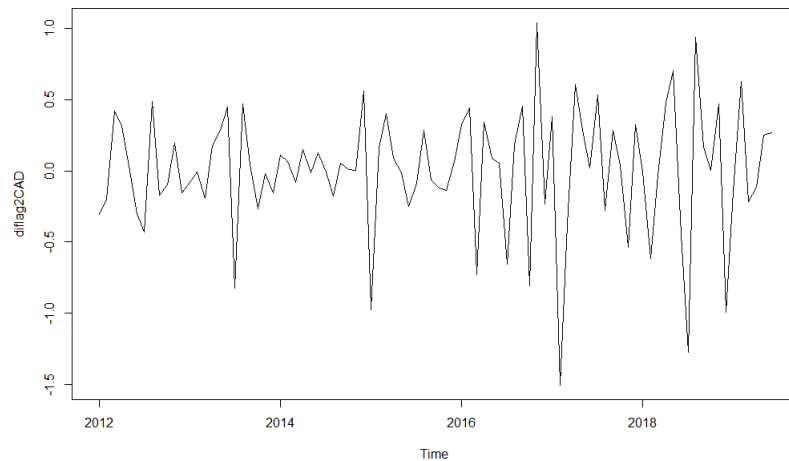


Figura 9.20: Serie diferenciada dos veces del tipo de cambio CAD/MXN.

```
> adf.test(diflag2CAD, k=0)

Augmented Dickey-Fuller Test

data: diflag2CAD
Dickey-Fuller = -12.317, Lag order = 0, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(diflag2CAD, k = 0) : p-value smaller than printed p-value
```

Figura 9.21: Prueba de Dickey-Fuller para determinar la existencia de raíces unitarias a la serie diferenciada del tipo de cambio CAD/MXN.

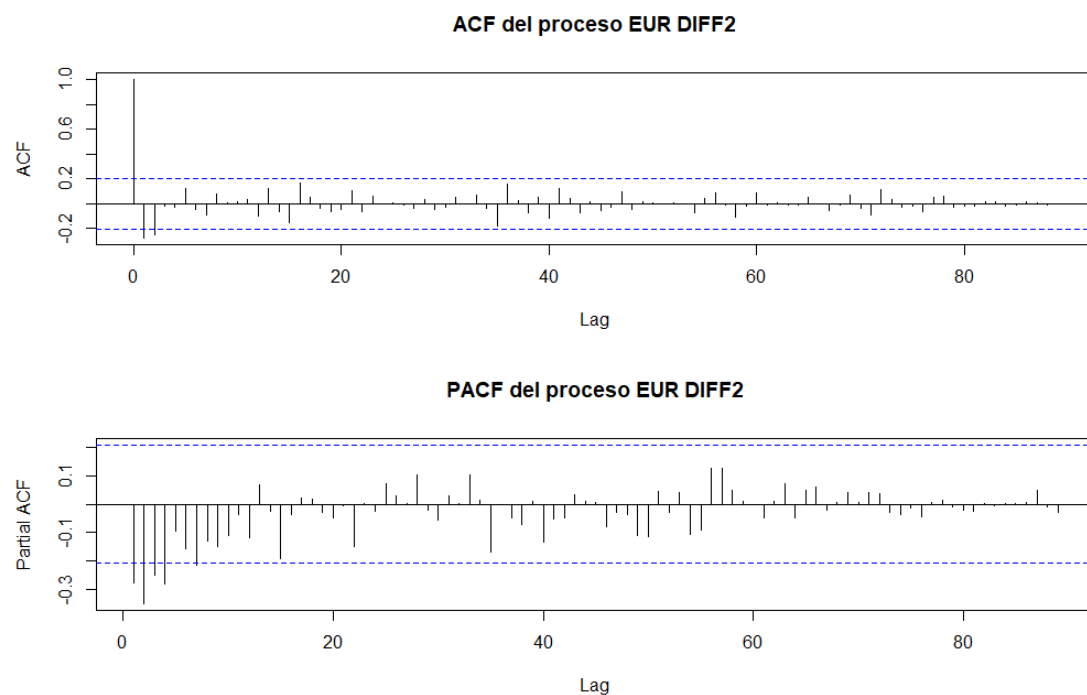


Figura 9.22: ACF y PACF de la serie de tiempo diferenciada del tipo de cambio CAD/MXN para determinar modelo.


```
> ARIMA221=arima(FX_CAD, order=c(2,2,1), include.mean = TRUE)
> ARIMA221

Call:
arima(x = FX_CAD, order = c(2, 2, 1), include.mean = TRUE)

Coefficients:
      ar1      ar2      ma1
  0.2170 -0.2243 -1.0000
s.e.  0.1029  0.1023  0.0471

sigma^2 estimated as 0.1049:  log likelihood = -28.58,  aic = 65.16
```

Figura 9.23: Ajuste de modelo ARIMA a la serie de tiempo del tipo de cambio CAD/MXN.

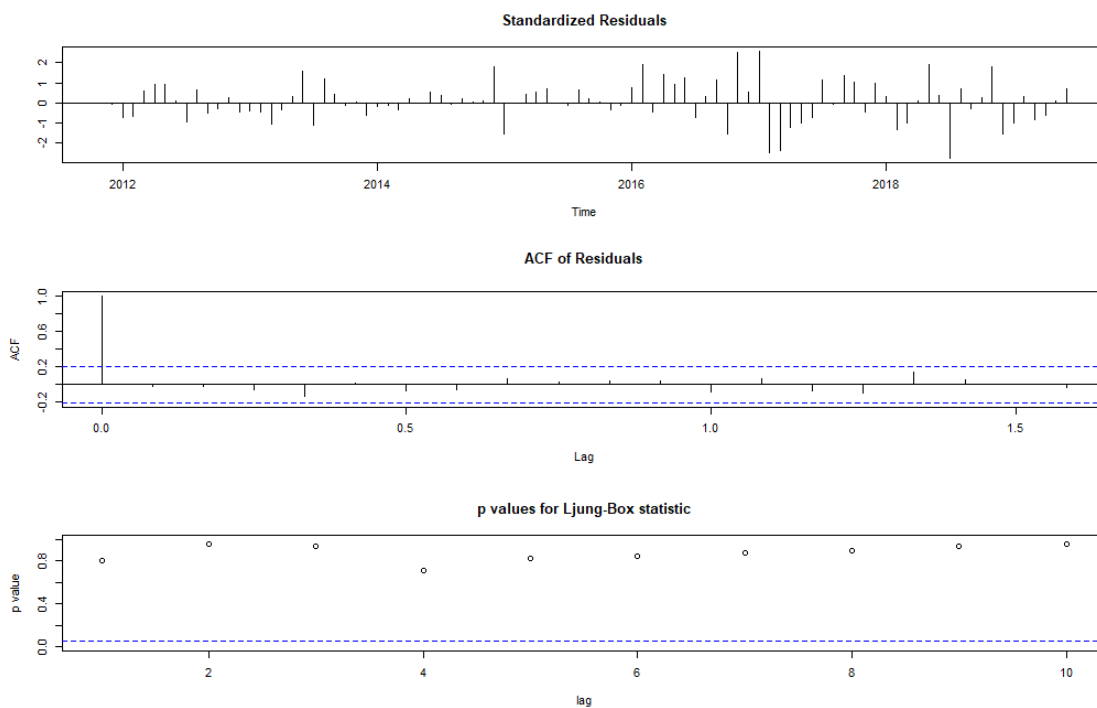


Figura 9.24: ACF y estadísticos de Ljung-Box para los distintos lags de los residuales del modelo ajustado.

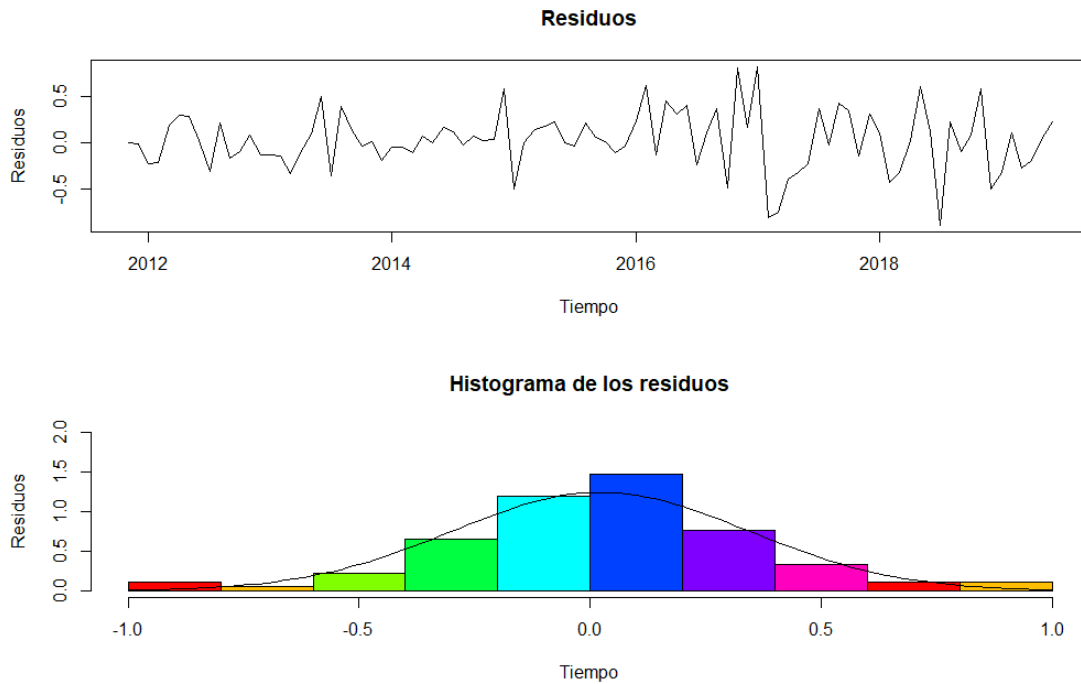


Figura 9.25: Serie generada de los residuales e histograma con densidad de una distribución normal ajustada.

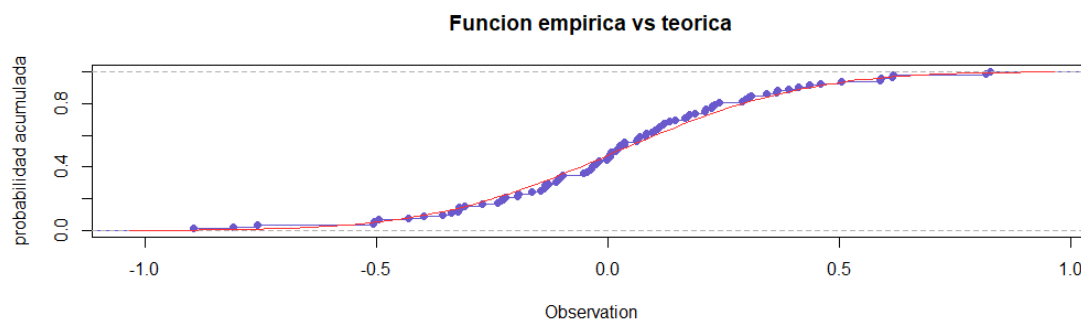


Figura 9.26: Distribución empírica generada por los residuales contra distribución teórica normal ajustada.

```
> ks.test(residuos, pnorm, mean(residuos),sd=sqrt(var(residuos)))

One-sample Kolmogorov-Smirnov test

data:  residuos
D = 0.062343, p-value = 0.8449
alternative hypothesis: two-sided

> lillie.test(residuos)

Lilliefors (Kolmogorov-Smirnov) normality test

data:  residuos
D = 0.062343, p-value = 0.5093

> ad.test(residuos)

Anderson-Darling normality test

data:  residuos
A = 0.45144, p-value = 0.2681
```

Figura 9.27: Pruebas de Kolmogorov-Smirnov, Lilliefors y Anderson-Darling para verificar normalidad en residuales.

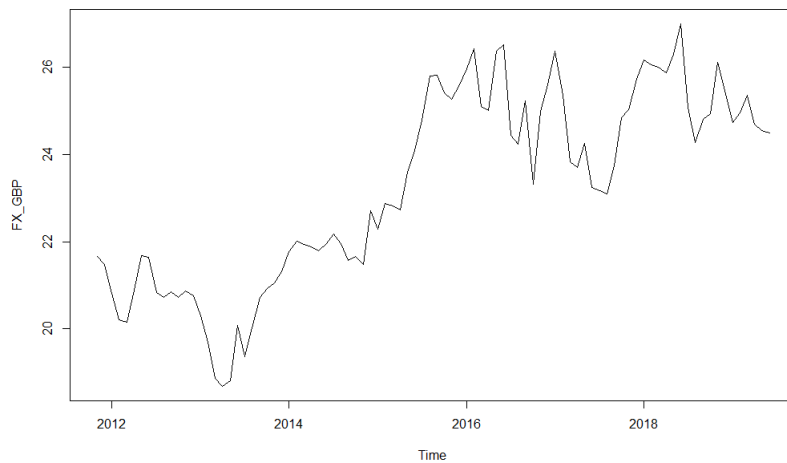


Figura 9.28: Serie de tiempo generada por el tipo de cambio GBP/MXN.

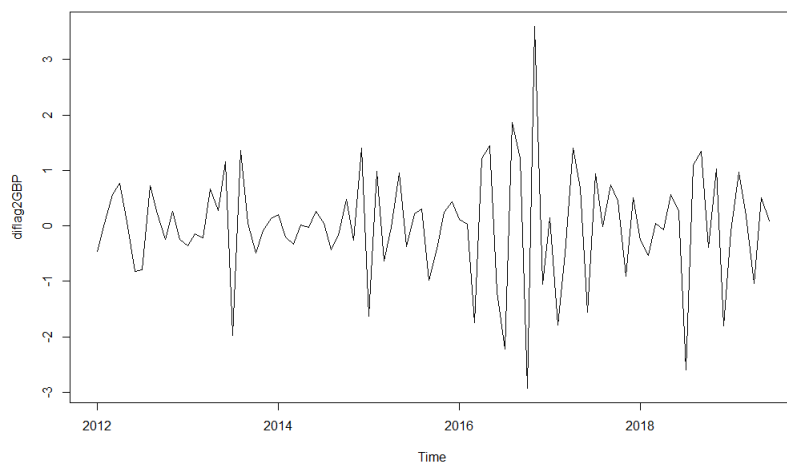


Figura 9.29: Serie diferenciada dos veces del tipo de cambio GBP/MXN.

```
> adf.test(diflag2GBP, k=0)

Augmented Dickey-Fuller Test

data: diflag2GBP
Dickey-Fuller = -14.192, Lag order = 0, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(diflag2GBP, k = 0) : p-value smaller than printed p-value
```

Figura 9.30: Prueba de Dickey-Fuller para determinar la existencia de raíces unitarias a la serie diferenciada del tipo de cambio GBP/MXN.

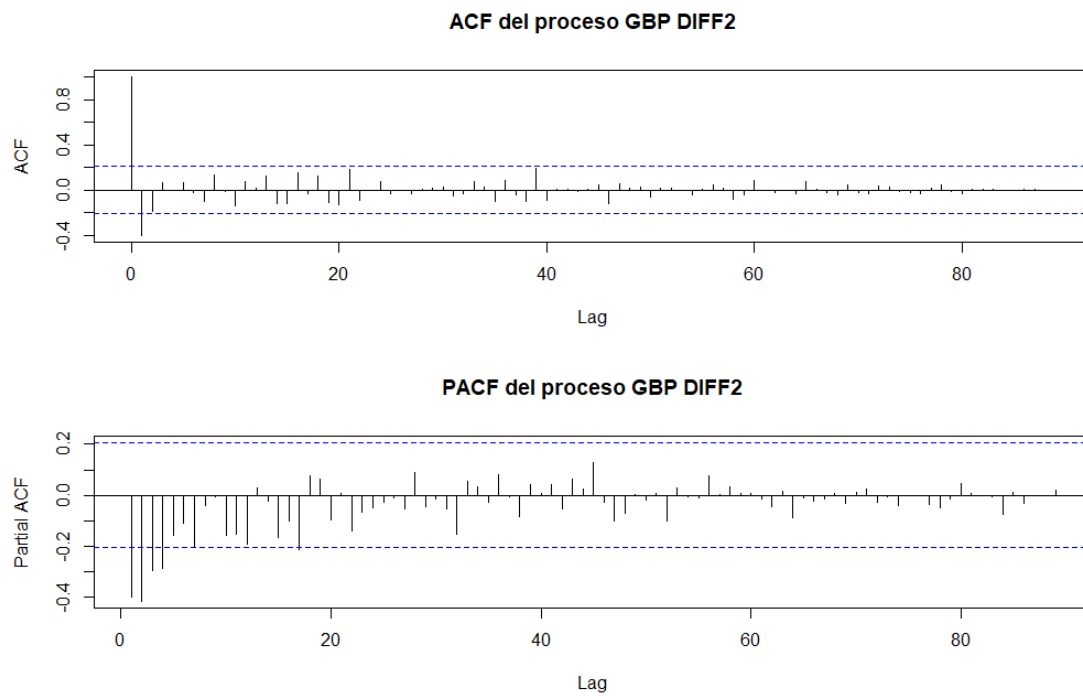


Figura 9.31: ACF y PACF de la serie de tiempo diferenciada del tipo de cambio GBP/MXN para determinar modelo.

```
> ARIMA121=arima(FX_GBP, order=c(1,2,1), include.mean = TRUE)
> ARIMA121

Call:
arima(x = FX_GBP, order = c(1, 2, 1), include.mean = TRUE)

Coefficients:
      ar1      ma1
  0.0060  -1.0000
s.e.  0.1055   0.0333

sigma^2 estimated as 0.4943:  log likelihood = -98.25,  aic = 202.5
```

Figura 9.32: Ajuste de modelo ARIMA a la serie de tiempo del tipo de cambio GBP/MXN.

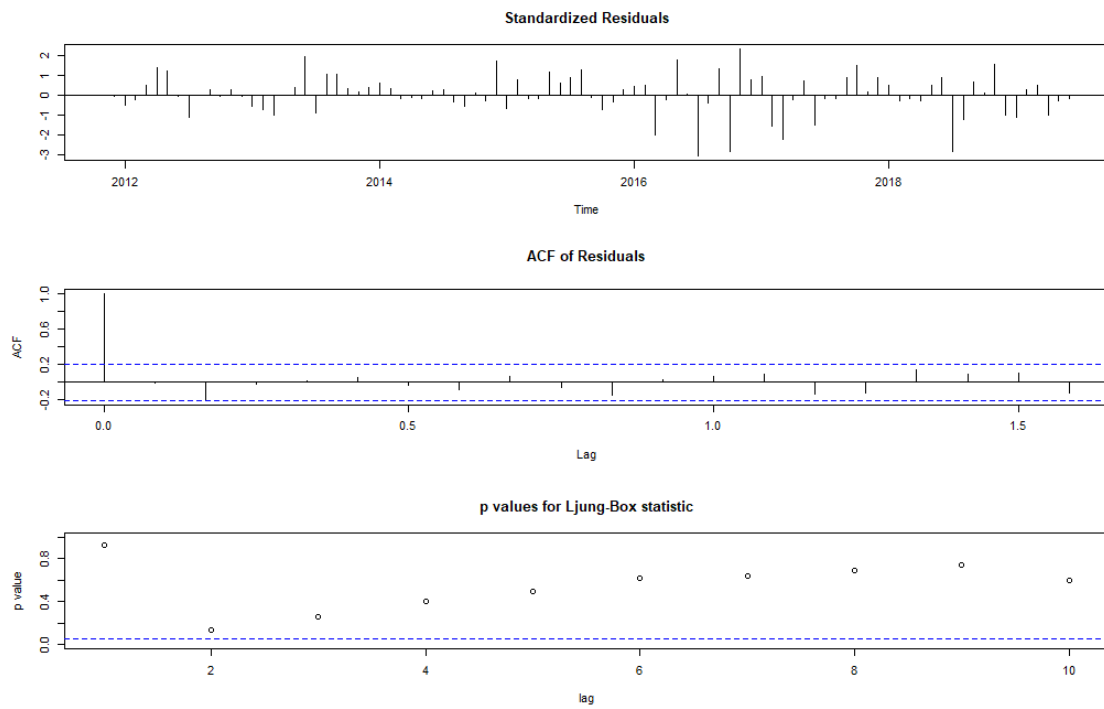


Figura 9.33: ACF y estadísticos de Ljung-Box para los distintos lags de los residuales del modelo ajustado.

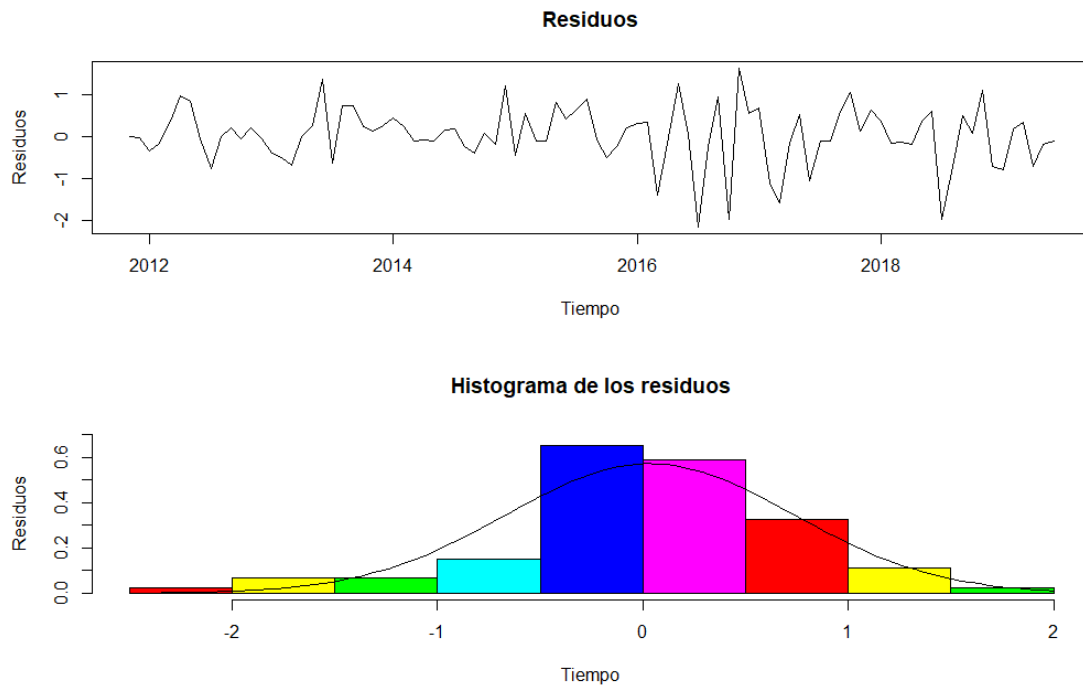


Figura 9.34: Serie generada de los residuales e histograma con densidad de una distribución normal ajustada.

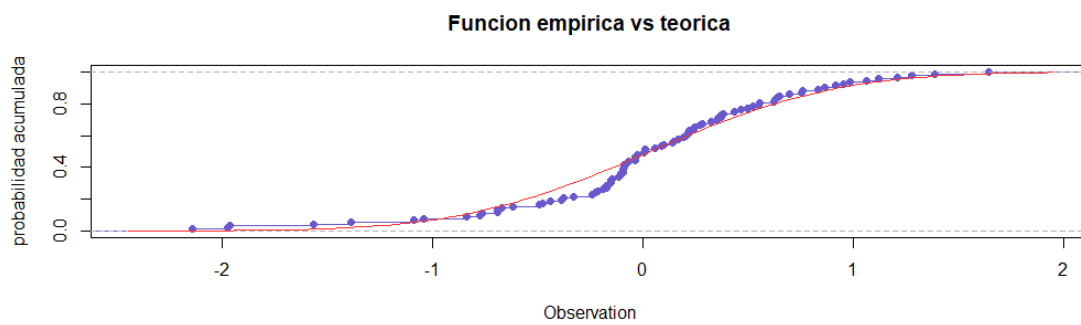


Figura 9.35: Distribución empírica generada por los residuales contra distribución teórica normal ajustada.

```
> ks.test(residuos, pnorm, mean(residuos),sd=sqrt(var(residuos)))

One-sample Kolmogorov-Smirnov test

data:  residuos
D = 0.12918, p-value = 0.08475
alternative hypothesis: two-sided

> lillie.test(residuos)

Lilliefors (Kolmogorov-Smirnov) normality test

data:  residuos
D = 0.12918, p-value = 0.0006579

> ad.test(residuos)

Anderson-Darling normality test

data:  residuos
A = 1.198, p-value = 0.003786
```

Figura 9.36: Pruebas de Kolmogorov-Smirnov, Lilliefors y Anderson-Darling para verificar normalidad en residuales.

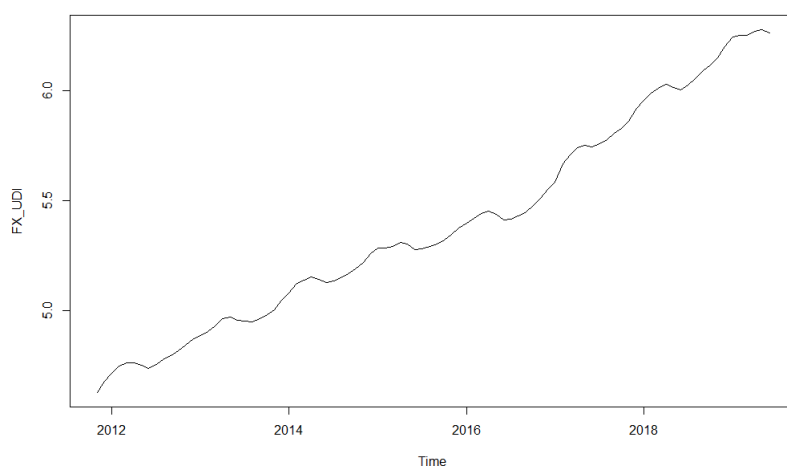


Figura 9.37: Serie de tiempo generada por el precio de la UDI.

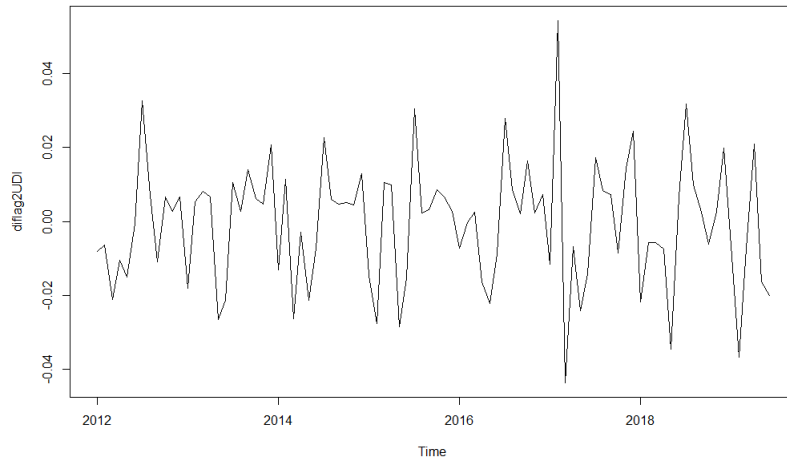


Figura 9.38: Serie diferenciada dos veces del precio de la UDI.

```
> adf.test(diflag2UDI, k=0)

Augmented Dickey-Fuller Test

data: diflag2UDI
Dickey-Fuller = -9.4038, Lag order = 0, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(diflag2UDI, k = 0) : p-value smaller than printed p-value
```

Figura 9.39: Prueba de Dickey-Fuller para determinar la existencia de raíces unitarias a la serie diferenciada del precio de la UDI.

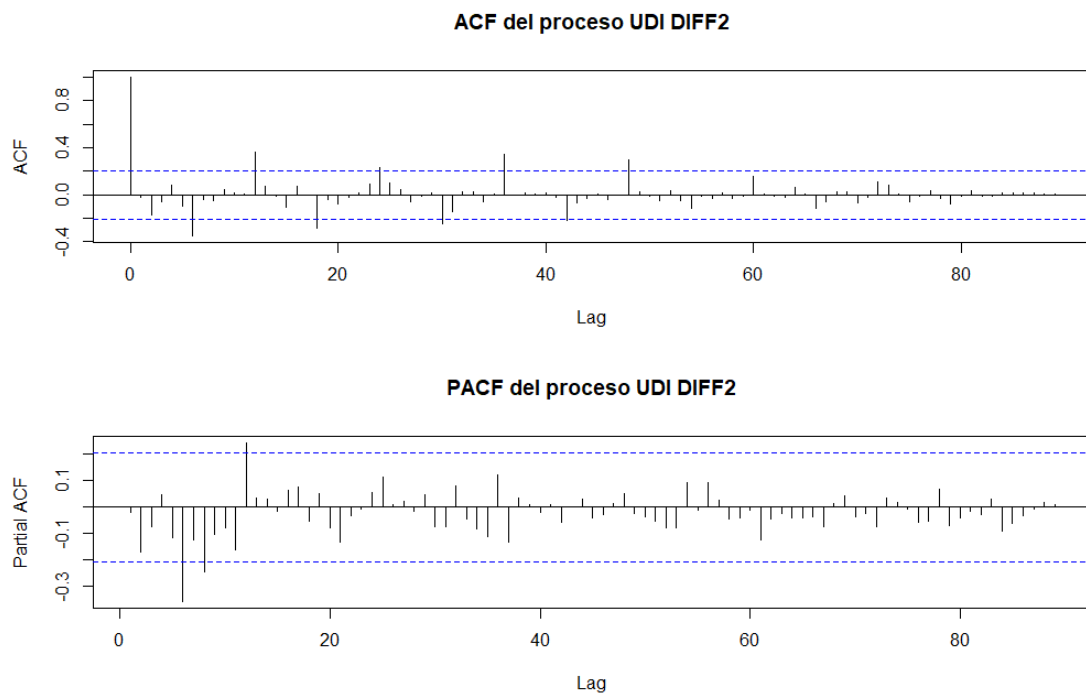


Figura 9.40: ACF y PACF de la serie de tiempo diferenciada del precio de la UDI para determinar modelo.

```

> ARIMA222=arima(FX_UDI, order=c(2,2,2), include.mean = TRUE)
> ARIMA222

Call:
arima(x = FX_UDI, order = c(2, 2, 2), include.mean = TRUE)

Coefficients:
      ar1      ar2      ma1      ma2
  1.5550 -0.7880 -1.8255  0.8892
s.e.  0.0862  0.0784  0.0740  0.0792

sigma^2 estimated as 0.0002081:  log likelihood = 252.46,  aic = -494.91

```

Figura 9.41: Ajuste de modelo ARIMA a la serie de tiempo del precio de la UDI.

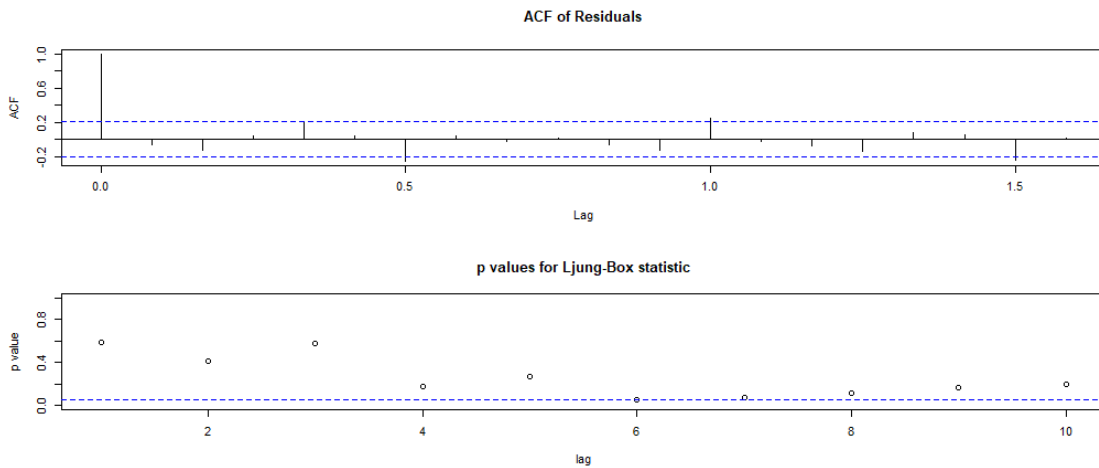


Figura 9.42: ACF y estadísticos de Ljung-Box para los distintos lags de los residuales del modelo ajustado.

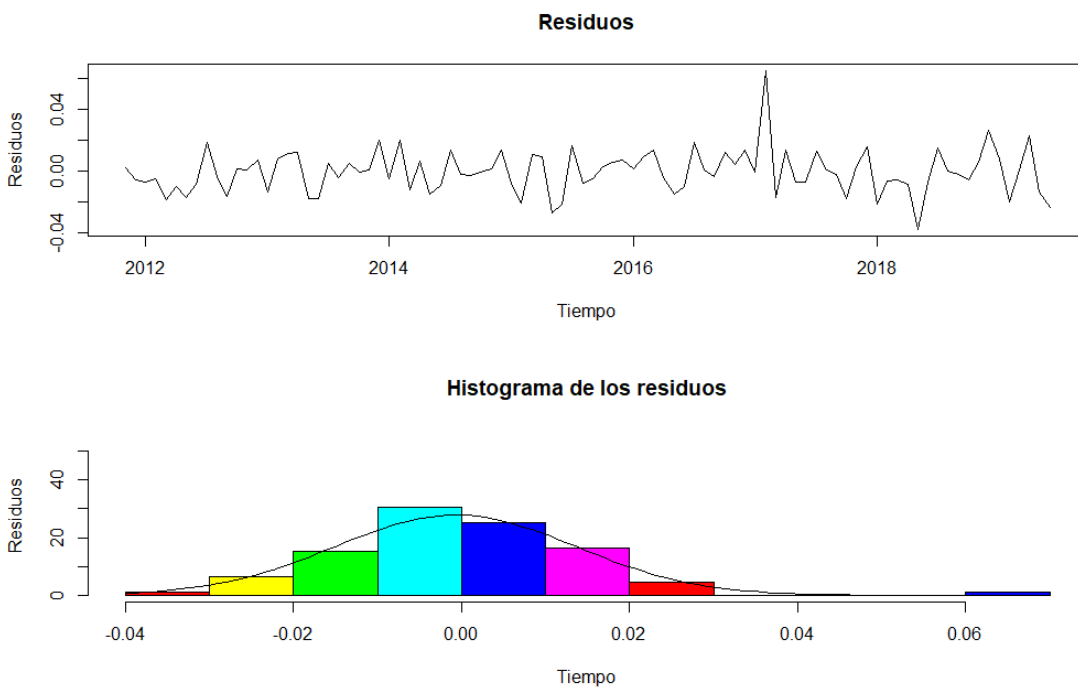


Figura 9.43: Serie generada de los residuales e histograma con densidad de una distribución normal ajustada.

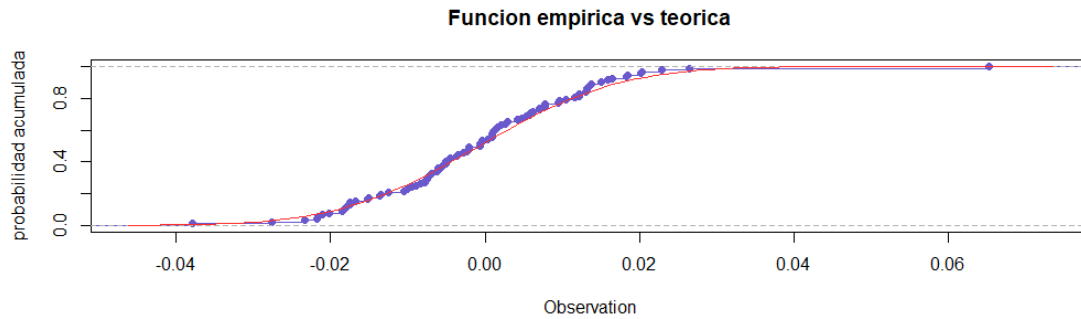


Figura 9.44: Distribución empírica generada por los residuos contra distribución teórica normal ajustada.

```

> ks.test(residuos, pnorm, mean(residuos),sd=sqrt(var(residuos)))

      One-sample Kolmogorov-Smirnov test

data:  residuos
D = 0.054045, p-value = 0.9376
alternative hypothesis: two-sided

> lillie.test(residuos)

      Lilliefors (Kolmogorov-Smirnov) normality test

data:  residuos
D = 0.054045, p-value = 0.7307

> ad.test(residuos)

      Anderson-Darling normality test

data:  residuos
A = 0.46782, p-value = 0.2446

```

Figura 9.45: Pruebas de Kolmogorov-Smirnov, Lilliefors y Anderson-Darling para verificar normalidad en residuos.

Tabla Resumen de modelos ajustados

	Serie	p	d	q	Mu	Phi1	Phi1_SÉ	Phi2	Phi2_SÉ	Phi3	Phi3_SÉ	Theta1	Theta1_SÉ	Theta2	Theta2_SÉ	Sigma.2	Dickey.Fuller	AIC	Ljung.Box	Kolmogorov	Anderson	Lilliefors	Normalidad
1	USD	2	2	1	0.001743465	0.2068	0.1035	-0.1821	0.1031	0.0000	N.A.	-1.0000	0.0504	0.0000	N.A.	2.016e-01	0.01	123.84	Si	0.37320	0.002214	0.0454300	No
2	EUR	3	2	1	0.009719095	-0.0025	0.1257	-0.1289	0.1159	-0.1538	0.1179	-0.9873	0.2358	0.0000	N.A.	2.820e-01	0.01	155.67	Si	0.83880	0.497600	0.4462000	Si
3	JPY	1	2	1	0.000534250	0.1211	0.1149	0.0000	N.A.	0.0000	N.A.	-0.9526	0.0508	0.0000	N.A.	2.766e-05	0.01	-681.03	Si	0.82590	0.457200	0.4734000	Si
4	CAD	2	2	1	0.002272918	0.2170	0.1029	-0.2243	0.1023	0.0000	N.A.	-1.0000	0.0471	0.0000	N.A.	1.049e-01	0.01	65.16	Si	0.84490	0.268100	0.5093000	Si
5	GBP	1	2	1	0.001395792	0.0060	0.1055	0.0000	N.A.	0.0000	N.A.	-1.0000	0.0333	0.0000	N.A.	4.943e-01	0.01	202.50	Si	0.08475	0.003786	0.0006579	No
6	UDI	2	2	2	-0.000700901	1.5550	0.0862	-0.7880	0.0784	0.0000	N.A.	-1.8255	0.0740	0.8892	0.0792	2.081e-04	0.01	-494.91	Si	0.93760	0.244600	0.7307000	Si

Figura 9.46: Resumen de los modelos ajustados a las series de tiempo.

Capítulo 10

Bibliografía

Bibliografía

- [1] Ansley, C.F. (1979), An algorithm for the exact likelihood of a mixed autoregressive-autoregressivemoving- average process, *Biometrika*, 66, 59-65.
- [2] Banco de México, *Descripción Técnica de los Bonos de Desarrollo del Gobierno Federal con Tasa De Interés Fija*. Obtenido de: <https://www.banxico.org.mx/mercados/d/7B77F7E7D0-8EDF-B4C1-406A-32968005EBC97D.pdf>
- [3] Banco de México, *Descripción Técnica de los Certificados de la Tesorería de la Federación*. Obtenido de: <https://www.banxico.org.mx/mercados/d/7B0DE0044F-662D-09D2-C8B3-4F1A8E43655F7D.pdf>
- [4] Banco de México, *Descripción Técnica de los Bonos de Desarrollo del Gobierno Federal con Pago Semestral de Interés y Protección Contra la Inflación*. Obtenido de: <https://www.banxico.org.mx/mercados/d/7B7F0B7876-6747-4AB3-E9C7-F6BF83EBECAC7D.pdf>
- [5] Banco de México, *Descripción Técnica de los Udibonos de Desarrollo del Gobierno Federal Denominados en Unidades de Inversión*. Obtenido de: <https://www.banxico.org.mx/mercados/d/7B52319AD4-4B78-6F95-E313-7AC67498B7287D.pdf>
- [6] Bolder, D. J. 2008. *The Canadian Debt-Strategy Model*. Bank of Canada Review. Summer 2008.
- [7] Black F., Scholes M. (1973). The pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81, 637-654.
- [8] Brigo D., Mercurio F. (2006). *Interest Rate Models- Theory and Practice*. Springer, 2a edición.
- [9] Brockwell P.J., Davis R.A., (2002). *Introduction to Time Series and Forecasting* New York Springer-Verlag, 2a edición.
- [10] Carroll R., Ruppert D. (1988). *Transformation and Weighting in Regression*. New York: Chapman and Hall.
- [11] Casar J. (2017). Deuda Pública y Política Fiscal en México. *Revista de Economía Mexicana*. Núm. 2. pp. 233-264.
- [12] Cox, J.C., Ingersoll, J.E., and Ross, S.A. (1985) A Theory of the Term Structure of Interest Rates, *Econometrica* 53, 385-407.

- [13] Dothan, L.U. (1978) On the Term Structure of Interest Rates, *Journal of Financial Economics* 6, 59-69.
- [14] Díaz de León A., Gutiérrez R., Vela O. (2012). Estimación del Costo en Riesgo para la Deuda Pública de México. *Gaceta de Economía*. Núm. 30. pp. 305-317.
- [15] Galindo M., Ríos V. (2015) Deuda Pública. *Serie de Estudios Económicos*, Vol. 1, Agosto 2015. México DF: México ¿Cómo vamos?
- [16] German H., El Karoui N., Rochet J., (1995). Change of Numeraire, Changes of Probability Measures and Pricing Options, *Journal of Applied Probability* 32, 443-458.
- [17] González J, (2015). *Introducción del Factor Humano al Análisis de Riesgo*. (Tesis Doctoral). Centro de Estudios del Riesgo Tecnológico, Universitat Politècnica de Catalunya, España.
- [18] Harrison J.M., Kreps, D. M. (1979). Martingales and Arbitrage in Multiperiod Securities Markets, *Journal of Economic Theory* 20, 381-408.
- [19] Harrison J.M., Pliska, S. R. (1981). Martingales and Stochastic Integrals in the Theory of Continuous Trading, *Stochastic Processes and their Applications* 11, 215-260.
- [20] Harrison J.M., Pliska, S. R. (1981). A Stochastic Calculus Model of Continuous Trading: Complete Markets, *Stochastic Processes and their Applications* 15, 313-316.
- [21] Jamieson D., Deeley S., (2011). *The Canadian Debt-Strategy Model: An Overview of the Principal Elements*. Bank of Canada. Paper 2011-3.
- [22] Kladívko K. (2007). *Maximum Likelihood Estimation of the Cox-Ingersoll-Ross: The Matlab Implementation*. Department of Statistics and Probability Calculus, University of Economics Prague and Debt Management Department, Ministry of Finance of the Czech Republic.
- [23] Ortega Sánchez, J. Material Didáctico, capítulo 6, liga: <https://www.cimat.mx/jortega/MaterialDidactico/ST2013/Prediccion.pdf>
- [24] Overbeck L., Rydén T. (1997). Estimation in the Cox-Ingersoll-Ross Model. *Economic Theory*. Vol. 13. pp. 430-489.
- [25] Revuz D., Yor M. (1991) *Continuous Martingales and Brownian Motion*. Berlin: Springer-Verlag.
- [26] Salinas L. (2011). *Modelo de tasas de Interés Vasicek y algunas generalizaciones*. (Tesis de maestría). IPN, México.
- [27] SHCP. (29 de Mayo de 2015b). *Deuda Pública Metodología tradicional*. Obtenido de Estadísticas Oportunas de Finanzas Públicas: <http://www.shcp.gob.mx/POLITICAFINANCIERA/FINANZASPUBLICAS>
- [28] SHCP. *Plan Anual de Financiamiento 2018*. Obtenido de Finanzas Públicas y Presupuesto: https://www.finanzaspublicas.hacienda.gob.mx/work/models/FinanzasPublicas/docs/paquete_economico/paf/paf_2018.pdf

- [29] Sisniega G. (2015). *Una Introducción al Análisis de Series de Tiempo y sus Aplicaciones*. (Tesis de pregrado). UNAM, México.
- [30] Steele J.M. (2001). *Stochastic Calculus and financial applications*. New York: Springer-Verlag.
- [31] The World Bank. (2015). *Data, Central government debt*. Obtenido de <http://data.worldbank.org/indicator/GC.DOD.TOTL.GD.ZS/countries>
- [32] Tsay R. (2002). *Analysis of Financial Time Series*. Canadá: John Wiley and Sons. University of Chicago.
- [33] Vasicek, O. (1977) An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics* 5, 177-188.
- [34] Wei W. (2006). *Time Series Analysis: Univariate and Multivariate Methods*. Pearson Addison Wesley, 2a edición.