



**UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MÉXICO**



FACULTAD DE ECONOMÍA

**“UNA INTRODUCCIÓN AL VALOR EN RIESGO:
TEORÍA Y APLICACIONES”**

TESINA

QUE PARA OBTENER EL TÍTULO DE:

LICENCIADO EN ACTUARÍA

PRESENTA:

PAOLA GABRIELA DÍAZ GÓMEZ

ASESOR:

MAT. MAURICIO GARCÍA MARTÍNEZ

REVISORES:

M.EN MAT.APL VERÓNICA ÁNGELES MORALES

DR. EN C. MIGUEL ÁNGEL DÍAZ CARREÑO

TOLUCA, ESTADO DE MÉXICO

NOVIEMBRE 2019

Índice

Introducción.....	6
Casos y Aplicaciones	12
I. Álgebra Lineal.....	18
I.1. Álgebra Lineal	18
1.1.1. Preliminares.....	18
1.1.1.1. Vectores.....	18
1.1.1.2. Producto cruz.....	18
1.1.1.3. Sistemas de ecuaciones	19
1.1.1.4. Polinomios.....	19
1.1.1.5. Matrices.....	19
1.1.1.6. Funciones.....	20
1.1.1.7. Alternativa matricial para Polinomios	21
1.1.2. Aproximaciones Gradiente y gradiente Hessiano	22
1.1.2.1. Aproximaciones Univariadas.....	22
1.1.2.2. Aproximaciones Multivariadas: aproximaciones Gradiente y gradiente Hessiano	22
1.1.3. Interpolación	22
1.1.4. Mínimos cuadrados ordinarios.....	23
1.1.5. Factorización de Cholesky.....	23
1.1.5.1. Formas cuadráticas.....	23
1.1.5.2. Matriz de raíces cuadradas.....	24
1.1.5.3. Matriz de Cholesky.....	24
II. Probabilidad.....	25
II.1. Espacios de Probabilidad.....	25
II.2. Variables Aleatorias	26
2.2.1. Definición de variable aleatoria.....	26
2.2.2. Funciones de probabilidad y funciones de distribución acumulada	26
2.2.3. Momentos empíricos	27
2.2.3.1. Esperanza	27
2.2.3.2. Esperanza de una función de una variable aleatoria	28

2.2.4.	Medidas de dispersión.....	28
2.2.4.1.	Varianza.....	28
2.2.5.	Sesgo.....	28
2.2.6.	Curtosis.....	29
2.2.7.	Cuantiles.....	29
2.2.8.	Momentos.....	29
II.3.	Vectores aleatorios.....	29
2.3.1.	Vector aleatorio.....	29
2.3.2.	Función de distribución conjunta, marginal y condicional.....	30
2.3.3.	Función de probabilidad conjunta, marginal y condicional.....	30
2.3.4.	Independencia de variables aleatorias.....	31
2.3.5.	Parámetros de vectores aleatorios.....	32
2.3.5.1.	Esperanza.....	32
2.3.5.2.	Covarianza.....	33
2.3.6.	Correlación.....	33
2.3.7.	Momentos conjuntos.....	34
2.3.8.	Polinomios lineales de vectores aleatorios.....	34
II.4.	Distribuciones de probabilidad.....	35
2.4.1.	Distribución normal.....	35
2.4.2.	Distribución t - Student.....	36
2.4.3.	Distribución ji-cuadrada.....	37
II.5.	Teoremas adicionales.....	37
2.5.1.	Polinomios cuadráticos de vectores aleatorios normales conjuntos.....	37
2.5.2.	La ley de los grandes números.....	38
2.5.3.	Teorema central del límite.....	38
III.	Estadística y series de tiempo.....	39
III.1.	Conceptos de estadística.....	39
3.1.1.	Estadística descriptiva.....	39
3.1.1.1.	Media.....	40
3.1.1.2.	Varianza.....	40
3.1.2.	Muestras aleatorias.....	40
3.1.2.1.	Estadísticas.....	40
3.1.3.	Estimación puntual.....	40

3.1.3.1.	Estimador puntual	40
3.1.4.	Pruebas de hipótesis	41
3.1.4.1.	Errores de prueba de hipótesis	41
3.1.4.2.	Prueba de hipótesis acerca de la media de una población t-Student usando puntos de rechazo	42
3.1.4.3.	Prueba de hipótesis usando los valores p (p-value).....	43
III.2.	Series de tiempo	44
3.2.1.	Series de tiempo.....	44
3.2.2.	Procesos estocásticos	44
3.2.3.	Procesos estacionarios.....	44
3.2.4.	Autocorrelación.....	45
3.2.4.1.	Función de autocorrelación	46
3.2.4.2.	Función de autocorrelación parcial	47
3.2.5.	Modelos de series de tiempo	47
3.2.5.1.	Proceso ruido blanco	47
3.2.5.1.1.	Prueba de Autocorrelación	47
3.2.5.2.	Caminata Aleatoria.....	48
3.2.6.	Modelos usuales de series de tiempo.....	48
3.2.6.1.	Modelos autorregresivos	48
3.2.6.2.	Modelos de promedios móviles.....	49
3.2.6.3.	Modelos autorregresivos de media móvil.....	49
3.2.7.	Modelos de series financieras	50
3.2.7.1.	Modelos ARCH	51
3.2.7.2.	Modelos GARCH.....	52
3.2.7.2.1.	Estimación de los parámetros del modelo $GARCH(1,1)$	53
3.2.7.2.2.	Ventajas del modelo $GARCH(1,1)$	53
3.2.7.2.3.	Desventajas del modelo $GARCH(1,1)$	53
3.2.7.3.	Modelo EWMA	54
3.2.8.	Modelando correlaciones.....	54
3.2.9.	Análisis estadístico de Series de Tiempo	55
3.2.10.	Análisis estadístico de Series de Tiempo aplicado a Series financieras ..	57
3.2.10.1.	Rendimientos financieros	57
3.2.10.2.	Distribución condicional y no condicional.....	58

3.2.10.3. Propiedades de los rendimientos financieros.....	59
IV. Valor en Riesgo.....	62
IV.1. Introducción al VaR.....	62
4.1.1. Definición de VaR.....	63
4.1.2. Enfoques del VaR.....	65
4.1.2.1. El VaR para distribuciones generales.....	65
4.1.2.2. El VaR para distribuciones paramétricas.....	65
4.1.3. Esquema general del VaR.....	66
4.1.4. Desarrollo del Esquema general del VaR.....	70
4.1.4.1. Procedimiento de Inferencia.....	70
4.1.4.2. Procedimiento de asignación.....	70
4.1.4.2.1. Asignaciones primarias.....	71
4.1.4.2.1.1. Reasignaciones de portafolio.....	72
4.1.4.2.1.2. Reasignaciones lineales con aproximaciones gradiente.....	73
4.1.4.2.1.3. Reasignaciones cuadráticas con aproximaciones gradiente Hessiano	74
4.1.4.2.1.4. Interpolación y el método de mínimos cuadrados.....	74
4.1.5. Procedimiento de Transformación.....	76
4.1.5.1. Procedimiento de Transformación lineal.....	76
4.1.5.2. Procedimiento de Transformación cuadrático.....	77
4.1.5.3. Procedimiento de Transformación Monte Carlo.....	78
IV.2. Aproximaciones VaR.....	79
4.2.1. Simulación Histórica.....	80
4.2.1.1. Ventajas del modelo Simulación Histórica.....	80
4.2.1.2. Desventajas del modelo Simulación Histórica.....	81
4.2.1.3. Extensiones del modelo Simulación Histórica.....	81
4.2.1.3.1. Aproximación Histórica por Bootstrap.....	81
4.2.1.3.2. Ponderación de observaciones.....	81
4.2.1.3.3. Ponderación de volatilidad.....	82
4.2.1.4. Simulación Histórica Filtrada.....	82
4.2.2. Aproximación paramétrica.....	82
4.2.2.1. Delta Normal.....	82
4.2.2.1.1. Ventajas del modelo Delta Normal.....	84

4.2.2.1.2. Desventajas del modelo Delta Normal	84
4.2.3. Simulación Monte Carlo.....	84
4.2.3.1. Simulación con una variable aleatoria.....	85
4.2.3.2. Simulación con múltiples variables	85
4.2.3.3. Ventajas del modelo Monte Carlo	86
4.2.3.4. Desventajas del modelo Monte Carlo	87
4.2.5. VaR mediante modelos GARCH con Innovaciones normales y t-Student..	87
IV.3. Backtesting	88
4.3.1. Test de Kupiec.....	88
IV.4. Críticas al VaR	89
IV.5. Ejemplos aplicados	89
4.5.1. Estimación del VaR: modelo paramétrico Normal, modelo de Simulación Histórica, modelo Monte Carlo.....	89
4.5.1.1 Backtesting: Prueba de Kupiec.....	91
4.5.2. Estimación del VaR dinámico: Modelo paramétrico Normal, Modelo EWMA, Modelo Simulación Histórica, Modelo Simulación Histórica Filtrada.....	94
4.5.3. Análisis estadístico de las Series de Tiempo.....	95
V. Conclusiones	101
Apéndice	102
A.1. Propiedades empíricas de las series financieras	102
A.2. Modelación de datos	104
A.3. Cálculo del VaR dinámico: Modelo paramétrico Normal, Modelo EWMA, Modelo Simulación Histórica, Modelo Simulación Histórica Filtrada.....	104
A.4. Cálculo del VaR utilizando modelos GARCH con innovaciones normales y t- Student.....	106
Bibliografía	108

Introducción

El origen de medidas de riesgo de mercado se remonta a principios del siglo XX a los requerimientos de capital para las sociedades de valores en Estados Unidos. Su desarrollo fue influenciado directa o indirectamente por la teoría del portafolio que presentó Markowitz en 1952 acerca de la optimización de un portafolio. En 1980 se actualizaron algunas disposiciones de requerimientos de capital debido a la volatilidad que presentaron las tasas de interés de Estados Unidos, este proceso se basó en un análisis estadístico utilizando información de mercado y reflejaba un cuantil al 95% de la cantidad de dinero que una firma podría perder en un periodo de liquidación de un mes (Holton, 2014).

En 1986, *Bankers Trust* presentó medidas de riesgo de mercado para mercados de renta fija en las que se empleaban funciones lineales y modelos de componentes principales para simplificar los cálculos. Al mismo tiempo, el banco *Chase Manhattan* desarrolló una medida basada en el modelo Monte Carlo y *Citibank* implementó una medida conocida como “Monto de pérdida potencial” (Holton, 2014).

Durante los años noventa, la globalización y el uso de derivados financieros transformaron e innovaron los mercados financieros propiciando el desarrollo de herramientas estadísticas como el Valor en Riesgo (VaR, por sus siglas en inglés) para la medición del riesgo de mercado. En 1993, se publicó el reporte G-30 de Derivados: Prácticas y principios, el cual incluye recomendaciones para apoyar a *brokers* y a usuarios finales a gestionar sus actividades en derivados (Hull, 2007). Además, una encuesta reportada en 1993 mostraba el uso del VaR como soporte para determinar límites de riesgo de mercado y las intenciones de los *traders* por hacer uso de esta herramienta (Holton, 2014).

De acuerdo con Jorion (2009) el VaR resume la pérdida máxima estimada de una posición de mercado o de un portafolio de inversión en horizonte de tiempo determinado y dentro de un nivel de confianza dado. El VaR se presentó como una medida de riesgo de mercado que estableció un estándar en toda la industria, buscando proporcionar en un número el riesgo total de un portafolio de activos financieros. Hull (2007) describe que,

una institución financiera suele calcular medidas como delta, gamma y vega¹ en un portafolio de derivados para describir diferentes aspectos del riesgo. Estos cálculos son diarios para cada variable de mercado a la que se encuentra expuesta. Algunas veces hay cientos o miles de estas variables de mercado. Un análisis delta-gamma-vega puede llevar a que cada día se produzcan un vasto número de medidas de riesgo diferentes. Sin embargo, no proporcionan una forma de medir el riesgo total al que la institución financiera está expuesta.

Hull y White (1998) detallan que las tres razones por las que la herramienta VaR ganó popularidad fue la decisión de JP Morgan de extender gratuitamente el servicio *Riskmetrics*², permitiendo que su uso se extendiera y no fuera conocido solo por los administradores de fondos de las instituciones financieras; la segunda hace referencia a los desastres financieros que se llevaron a cabo por utilizar derivados en donde se menciona a *Procter and Gamble*, *Orange Country* y *Barings*³ y la tercer razón fue la decisión de utilizar el VaR para determinar el capital requerido para los bancos por parte de los reguladores del banco central.

En 1995, el comité de Basilea implementó los requerimientos de capital de los bancos, basados en una medida VaR estándar⁴ y permitió el uso de medidas VaR propias en ciertas circunstancias (Hull, 2007). El interés por medir el riesgo de mercado se incrementó debido a que los agentes económicos empezaron a presentar oscilaciones en los rendimientos reportados y por lo tanto se buscó implementar sistemas de monitoreo de riesgo de mercado que proporcionaran a la gerencia información oportuna sobre las posiciones y la pérdida potencial estimada de cada posición.

Los enfoques que han evolucionado como un medio para medir el riesgo de mercado son el enfoque estadístico y el enfoque de análisis de escenarios. El primero implica pronosticar la distribución de rendimientos usando probabilidad y utilizar modelos estadísticos. El segundo enfoque reevalúa la cartera con diferentes valores de tasas y precios de mercado y no necesariamente hace uso de un modelo estadístico, además,

¹ Estas medidas de riesgo se conocen como "las griegas" y proporcionan información de una variable de mercado que es de utilidad para el *trader* que está encargado de administrar la parte de los portafolios de la institución financiera

² El servicio *Riskmetrics* permitía modelar una gran cantidad de factores de riesgo, generando una matriz de correlación y actualizándola trimestralmente, además, utilizaba una distribución normal para determinar el VaR

³ Procter & Gamble en Estados Unidos presenta una pérdida por pérdidas de swaps sobre diferenciales, el Condado de Orange County en California anunció. pérdidas en fondos gubernamentales y el banco Barings en Reino Unido pierde 1,330 MM en USD por futuros sobre índices accionarios.

⁴ Bajo Basilea I el riesgo de mercado se basa en un VaR a 10 días a un nivel de confianza de 99%.

las tasas de interés y los precios que son utilizados pueden ser escogidos arbitrariamente (J.P. Morgan, 1996).

Debido a que los administradores de riesgo buscan estimar el riesgo de mercado utilizando medidas estadísticas sofisticadas como el VaR, se han interesado por estudiar las propiedades de las series de rendimientos de los activos financieros. La literatura empírica ha demostrado que las series de rendimientos de los activos presentan propiedades comunes. Cont (2001) las denomina “hechos estilizados” de los rendimientos de los activos. Por ejemplo, describe que las distribuciones de los rendimientos de las series financieras presentan el efecto de leptocurtosis, lo que significa que la distribución se encuentra centrada en la media con colas pesadas y largas como consecuencia de la cantidad de valores atípicos existentes en las series financieras. Otro ejemplo es que presentan asimetría en la volatilidad lo cual se asocia al impacto de los movimientos positivos y negativos del mercado (Tseng y Li, 2011).

En este sentido, la precisión de la estimación del VaR dependerá de la distribución de los factores de riesgo y de elegir un modelo adecuado de volatilidad y correlación. Algunos de los modelos para estimar la volatilidad son los modelos autorregresivos con heterocedasticidad condicional (ARCH, por sus siglas en inglés), el modelo de heterocedasticidad condicional autorregresiva generalizada (GARCH, por sus siglas en inglés) y el modelo de promedio móvil ponderado exponencialmente (EWMA, por sus siglas en inglés). Por otra parte, para estimar las correlaciones se puede mencionar al modelo EWMA y GARCH (Hull, 2007). La eficiencia de los modelos dependerá de los supuestos financieros y estadísticos.

Algunos ejemplos de la literatura existente que aplica las diversas aproximaciones para estimar el VaR, como los modelos EWMA, delta normal, modelos GARCH y variantes de los modelos GARCH, Simulación Monte Carlo y Simulación Histórica, se presenta a Aloui y Mabrouk (2010); Chen y Chen (2013); Hammoudeh, Malik y McAleer (2011); Holton (2014); Hung, Lee y Liu (2007); Nilsson (2017) y Pritsker (2001). En cuanto a las técnicas para evaluar el modelo VaR (*backtesting*) algunas investigaciones se presentan en Campbell (2005); Christoffersen, Hahn y Inoue (2001); Corbetta y Peri (2018); Kupiec (1995); Zhang y Nadarajah (2017).

Sin embargo, los periodos de inestabilidades financieras a nivel mundial han puesto a prueba los modelos VaR, demostrando que los modelos presentan limitaciones, por

ejemplo, son ineficientes para capturar eventos de pérdidas grandes con probabilidad baja de ocurrencia. En este sentido, a lo largo del tiempo se han implementado los modelos y además se comenzaron a utilizar nuevas medidas de riesgo como la pérdida esperada conocida como CVaR.

No obstante, en la actualidad, el VaR sigue siendo una herramienta utilizada a nivel mundial. Muchos de los agentes económicos se enfocan únicamente en los resultados arrojados y poco se habla de una explicación detallada de la construcción de estos modelos.

En este sentido, la presente tesina pretende dar una visión general teórica del VaR. Se describen los elementos básicos para lograr un entendimiento de los modelos VaR y poder estudiar los modelos existentes para entender la viabilidad de su uso y aprobar o desaprobado su uso. Se busca que sea un trabajo que sirva como una guía general para los interesados en el tema y de esta manera busquen profundizar su investigación en algún subtema del VaR, o en su caso, si se busca implementar las herramientas existentes o crear nuevas.

Documentar sobre VaR implica estudiar los conceptos técnicos utilizados para llevar a cabo los modelos de estimación. En consecuencia, se tendría que desarrollar un documento riguroso y largo, sin embargo, en la presente tesina no se profundizará en el estudio de las propiedades de todos los modelos VaR. Por lo cual, este trabajo se enfocará en lo siguiente:

- Proporcionar al lector una visión general y comprensible sobre el VaR y algunos de sus modelos de estimación.
- Llevar a cabo una recopilación de conceptos y definiciones básicas y avanzadas suficientes para facilitar la comprensión de textos que hablan sobre la definición de VaR y los modelos de estimación. Se incluyen temas de Álgebra Lineal, Probabilidad, Estadística y Series de Tiempo.
- Describir algunos de los modelos para estimar el VaR.
- Describir algunos ejemplos de manera general de los modelos aplicándolos a series financieras.

Se espera que este trabajo sirva de motivación para adentrarse a un estudio más profundo sobre el VaR y que al mismo tiempo esta tesina les sea de apoyo cuando consulten otros textos relacionados con esta herramienta estadística.

El primer capítulo de la presente tesina incluirá conceptos y definiciones de Álgebra Lineal. Las variables aleatorias frecuentemente se definen como polinomios lineales de vectores aleatorios, en este sentido, es relevante mencionar conceptos tales como las características de vectores y las operaciones de matrices. Para llegar a una simplificación del portafolio, se utilizan modelos matemáticos por lo cual, se hará mención del concepto de gradientes, los cuales aproximan funciones utilizando polinomios. También en este capítulo se presentará el método de mínimos cuadrados cuyo objetivo es ajustar una función a un conjunto de datos y obtener el valor de parámetros. Para abordar procedimientos tales como la transformación cuadrática y el modelo Monte Carlo para la estimación del VaR, se menciona el concepto de factorización de la matriz de Cholesky.

En el segundo capítulo se describirá la teoría de probabilidad que se aplica al utilizar los modelos de estimación del VaR, además se presentarán definiciones que servirán de soporte para abordar temas de estadística y series de tiempo. En este sentido, se incluyen las definiciones de variable aleatoria, funciones de densidad de probabilidad y funciones de distribución acumulada. También se describen los momentos empíricos y medidas de dispersión. Asimismo, se mencionan las funciones de probabilidad y de distribución de un vector aleatorio, así como la definición de correlación. Además, se describe uno de los conceptos centrales en la teoría de probabilidad, la cual es la independencia entre variables. En este capítulo se detallan algunas de las principales distribuciones de probabilidad utilizadas en el análisis de las series financieras y se describen teoremas importantes de la teoría de probabilidad.

En el tercer capítulo se describen conceptos de estadística y series de tiempo, los cuales soportan los procedimientos de inferencia para la estimación del VaR. En primer lugar, se hace mención de las propiedades que comúnmente presentan las series financieras y de las pruebas que se llevan a cabo para poder trabajar con las series de tiempo. También se presentan los modelos estándar de series de tiempo y se describen los procesos ARCH, GARCH y EWMA que modelan la volatilidad cambiante en el tiempo.

Una vez descritos las definiciones de Álgebra Lineal, Probabilidad, Estadística y Series de Tiempo se tendrán las herramientas para el entendimiento de los modelos para estimar el VaR. En este sentido, en el cuarto capítulo de la presente tesina, se presenta la definición de VaR, los distintos modelos VaR y se hace mención de los métodos de

validación del VaR. Para complementar el entendimiento de los modelos se describirán algunos ejemplos de estos modelos aplicados a series financieras con datos históricos.

Debido a que toda la teoría descrita en los próximos capítulos puede ser encontrada en diversos libros y monografías, no se incluyen todas las pruebas para los resultados expuestos en este trabajo pues el desarrollo de cada uno de estos sale del alcance de la presente tesina. Dentro del presente texto se hará uso de la siguiente notación:

v.a. variable aleatoria

i.i.d. independiente e idénticamente distribuidas

f.d. Función de distribución

f.p.d. Función de densidad de probabilidad

Casos y Aplicaciones

En las últimas décadas, los mercados financieros se han caracterizado por ser más volátiles y riesgosos, generando pérdidas potenciales a los participantes del mercado. En este sentido, los participantes del mercado han demandado mejores herramientas y modelos para medir el riesgo e identificar límites de riesgo. De modo que, con el fin de generar y poseer información más completa para tomar decisiones óptimas en la asignación de capital y recursos, una herramienta estadística que ganó popularidad fue la medida Valor en Riesgo.

La principal ventaja del VaR es que permite a los ejecutivos, directores y tenedores de activos evaluar su nivel de riesgo debido a que resume en un solo número, la exposición total de una institución al riesgo de mercado. En la actualidad, el VaR juega un rol importante en determinar el capital que las instituciones deben mantener como una proporción de un nivel de activos, conforme a lo que establecen las autoridades reguladoras. Por otra parte, las entidades financieras y autoridades encargadas de formular políticas utilizan el VaR para medir el riesgo, determinar límites de riesgo y crear informes de control. Además, los informes financieros requieren información cuantitativa del riesgo de mercado que surge de las tasas de interés, divisas, *commodities* la cual puede ser presentada con un informe de VaR (Holton, 2014). A continuación, se detallan algunas de las principales aplicaciones del VaR:

Requerimientos del capital regulatorio de acuerdo con la regulación

Dentro de las instituciones financieras, se estima el VaR para medir el riesgo de las inversiones, con el fin de soportar este riesgo con niveles de capital adecuados y además inducir a su revelación. Por ejemplo, Cuoco y Liu (2005), estudian el comportamiento de una institución financiera que está sujeta a requerimientos de capital, basadas en la estimación del VaR de acuerdo con el enfoque de modelos internos del Comité de Basilea. Por otra parte, Weber (2018) describe que bajo Solvencia II, el cálculo del requerimiento de capital se basa en el VaR, sin embargo, esta medida es deficiente al capturar el impacto de riesgos extremos, por lo cual propone alternativas para reducir el

requerimiento de capital total dentro de las firmas y concluye que la regulación de capital debería estar basada en medidas de riesgo coherentes⁵ como los expectiles.

Reportes internos (control y reporte del riesgo), adecuación de capital económico y límites de VaR de instituciones financieras y organizaciones de *trading*

Las entidades financieras y autoridades encargadas de formular políticas hacen uso del VaR para medir el riesgo, determinar límites de riesgo y crear informes de control. Por ejemplo, Al Janabi (2007) examina los métodos y procedimientos de gestión de riesgos de mercado que se deberían considerar al formular los objetivos diarios de gestión de riesgos para una perspectiva de mercados emergentes y no líquidos. Además, estudia los procedimientos para el cálculo de los límites máximos autorizados del VAR.

Revelación de información corporativa

Los reguladores y la competencia entre los diferentes participantes del mercado han impulsado a las instituciones a reportar información cuantitativa del riesgo de mercado que surge de las tasas de interés, divisas, commodities en informes financieros. Esta información puede ser presentada como datos sobre posiciones individuales, análisis de sensibilidad o un informe de VaR. Jorion (2002) proporciona resultados empíricos de llevar a cabo revelaciones sobre el VAR. Por ejemplo, en su análisis incluye que los analistas e inversores pueden utilizar las revelaciones del VAR para comparar los perfiles de riesgo de las carteras comerciales.

Modelos cuantitativos de asignación de activos – modelos de optimización de cartera⁶

Los inversionistas institucionales y los gestores de fondos buscan generar rendimientos superiores de inversión y además están interesados en su nivel de exposición al riesgo. Pearson (2002) describe que utilizar técnicas de medición de riesgo como el VAR les muestra a los inversionistas institucionales de qué manera pueden distribuir el riesgo y administrar de manera más eficiente sus carteras de inversión. Por su parte, Mensi, Hussain, Hammoudeh, Zeitun y Rehman, (2017) llevaron a cabo un análisis de diversificación de portafolio (participación de economías emergentes BRIC, mercados accionarios de Asia del Sur y mercados accionarios más desarrollados(Estados Unidos, Londres y Japón) y utilizaron la aproximación VaR basada en *wavelet*

⁵ Las medidas de riesgo cumplen con las siguientes propiedades: Monótona, Invariante por translación, Positiva homogénea y sub-aditiva

⁶ En inglés Holton (2014) hace referencia al término "Risk budgeting"

para determinar el riesgo del portafolio. Entre los resultados se detalla que al añadir a China en el portafolio se reduce el riesgo en el periodo previo y durante la crisis financiera global y la crisis de deuda soberana en Europa. Además, describe que, el análisis tiene implicaciones para los inversores institucionales si se asignan pesos óptimos a los diferentes activos en la formulación del portafolio.

Coberturas óptimas

El VaR se utiliza para determinar estrategias de cobertura óptima. Una cobertura⁷ financiera busca reducir o eliminar las posibles pérdidas generadas por llevar a cabo inversiones financieras. En instituciones financieras, bancos de inversión y empresas surgen situaciones en las que se utiliza algún instrumento líquido para cubrir una posición menos líquida. Sin embargo, puede existir riesgo de mercado en caso de que la correlación entre los dos sea imperfecta. Para saber cuántas unidades del instrumento de cobertura proporcionan una cobertura óptima, se puede aplicar una medida VaR a la posición total y ajustar la cobertura buscando minimizar el VaR (Cong, Seng Tan, Weng, 2013). Además, Hung, Su, Chang y Wang (2019) presentan un estudio en el que involucran el impacto de la liquidez al estimar el VaR de un portafolio, utilizando modelos GARCH multivariados GARCH -t y modelos GJR-GARCH-t. Además, destacan que el hecho de no tomar en cuenta el impacto de la liquidez en la estimación del VaR del portafolio, puede resultar en una cobertura inadecuada y en requerimientos de capital insuficientes.

Por otra parte, diversos autores han llevado a cabo investigaciones en las que se comparan los modelos para estimar el VaR, por ejemplo, Yu, Yang, Wei y Lei (2018) determinan el VaR y el CVaR de un portafolio de 4 tipos de crudos utilizando la teoría de valores extremos y de modelos de cópulas. Los resultados de backtesting presentan que la combinación de modelos GARCH con la teoría de valores extremos y los modelos de cópulas *vine* pueden producir modelos adecuados para el portafolio compuesto de activos de petróleo. En tanto, Del Brio, Mora-Valencia y Perote (2017) evalúan la estimación del VaR y CVaR de una modelación de ETFs de commodities y recomiendan utilizar distribuciones leptocúrticas y para mitigar las inquietudes que marca la regulación sobre la estabilidad financiera del negocio de commodities (Acuerdos de Basilea)

⁷ Se hace referencia al término en inglés "Hedging"

Laporta, Merlo y Petrella (2018) llevaron a cabo estimaciones del VaR para rendimientos diarios de commodities de energía utilizando modelos GARCH, EGARCH, GJR-GARCH, un modelo VaR condicional autorregresivo (CAViaR, por sus siglas en inglés), *Generalized Autorregressive Score* (GAS, por sus siglas en inglés), además se desarrolla un modelo *Dynamic Quantile Regression* (DQR, por sus siglas en inglés).

Por otra parte, Buchner (2017) describe que existe un entendimiento limitado al llevar a cabo la administración de riesgos en los fondos de capital privado por lo cual menciona las fuentes principales de riesgo a las que se enfrentan los inversionistas de capital privado y desarrolla tres medidas dinámicas de riesgo: el VaR, el VaR ajustado por liquidez y el Riesgo de flujo de efectivo, además presenta un análisis de sensibilidad que muestra el impacto de los cambios en los parámetros de las medidas de riesgo. Guo, Ding y Guo (2018) en consecuencia al crecimiento del mercado de paladio, determinaron el VaR para los rendimientos diarios del paladio utilizando 5 tipos diferentes de distribuciones estadísticas, los resultados indicaron que la distribución t sesgada llevó a cabo un ajuste adecuado y genera valores VaR cercanos a valores VaR mediante un modelo histórico no paramétrico.

El VaR también es utilizado en planeación de proyectos, por ejemplo, Levine, Lung, Nimbalkar, Wenning, Kelly y Power (2019) incorporaron metodologías como el VaR a inversiones en eficiencia energética para que los agentes de negocio de energía puedan interpretar de mejor manera los beneficios financieros de su capital invertido. Kong, Huang, Fan, Li, Zeng y Zhu (2017) llevan a cabo un análisis de riesgo para la administración de recursos hidrológicos en el cual se incorpora el análisis factorial y un VaR *fuzzy random* en un entorno de programación estocástico de dos etapas. El VaR se determina para reflejar el riesgo potencial del costo financiero bajo incertidumbre dual. Asimismo, Hosseini y Verma (2017) proponen una medida VaR para evaluar el riesgo del envío de materiales peligrosos por ferrocarril. Rahmanpour, Osanloo (2016) determinan el VaR para la planeación de producción a largo plazo en minas a cielo abierto (*open pit*) debido principalmente a que el precio de los commodities son el principal punto de la planeación de la mina y su incertidumbre pone en riesgo los beneficios de la operación de una mina. Pinheiro, Geraldo, Pacheco y Alves (2018) proponen una metodología de análisis de riesgo de inversión para una planta de energía solar en el mercado brasileño.

Se obtuvo el VaR y el CVaR, los resultados generan información para la toma de decisiones.

Otras aplicaciones del VaR se presentan en Patton, Ziegel y Chen (2019) describen que la implementación del Acuerdo de Basilea III (posterior al año 2019), pone atención al CVAR, debido a que no existe tanta información para la modelación de esta medida de riesgo comparado con el VaR, los autores proponen modelar el CVaR y el VaR conjuntamente y describen modelos dinámicos para esta medida de riesgo conjunta.

Por su parte, Ahmadi-Javid y Fallah-Tafti (2019) aplican una medida de riesgo nueva que es un límite por encima del VaR y del CVar, conocida como *Entropic Value-at-Risk* para la optimización de un portafolio, se determina que conforme aumenta el tamaño de la muestra, las aproximaciones EVaR presenta eficiencias computacionales sobre la medida CVaR.

En tanto, Du, Li y Wu (2019) proponen una transformación de variable para obtener el reaseguro por exceso de pérdidas óptimo utilizando VaR y CVaR. Se define la variable X como la pérdida inicial de un asegurador. Se transforma la función de supervivencia que utiliza de la función de distribución (generando la variable Y). Se obtienen los montos de retención óptimos de X y de Y.

Por otra parte, Woerner y Egger (2019) llevan a cabo un análisis de riesgo mediante un algoritmo cuántico el cual es más eficiente que las simulaciones tradicionales Monte Carlo. En el análisis se emplea una estimación con amplitud cuántica para determinar el precio de los activos y evaluar medidas de riesgo como el VaR y CVar. Se describe la implementación del algoritmo, su convergencia y la profundidad del circuito. Para presentar el algoritmo se utilizan dos modelos, el primer modelo utiliza hardware como *IBM Q Experience* y valua un *T-Bill* expuesto a un incremento en la tasa de interés. El segundo modelo, se simula un algoritmo que muestra como una computadora cuántica puede determinar el riesgo financiero para un portafolio de dos activos de deuda gubernamental con diferentes fechas de maduración

El VaR también se ha aplicado para llevar a cabo análisis del riesgo cibernético, el cual surgió como una amenaza a la estabilidad financiera. Bouveret (2018) analiza los diferentes tipos de incidentes cibernéticos y se identifican patrones utilizando una variedad de datos. Además, utiliza un VaR para evaluar el riesgo cibernético en el sector financiero. Por su parte, Malhorta (2015) entre las contribuciones de su tesis postdoctoral

lleva a cabo un modelo de seguros cibernéticos y desarrolla un estudio de un método empírico del VaR y de Inferencia Bayesiana.

Además, existe literatura sobre la estimación del VaR a las criptomonedas. Gkillas y Katsiampa (2018) aplican la teoría de valores extremos a la cola de la serie de precios de las 5 criptomonedas más relevantes a nivel mundial: *Bitcoin*, *Ethereum*, *Ripple*, *Bitcoin Cash* y *Litecoin*. Se estima el VaR y el CVar como medidas de riesgo y los resultados presentan que la criptomoneda más riesgosa es *Bitcoin Cash*, mientras que *Bitcoin* y *Litecoin* son las criptomonedas menos riesgosas. Por su parte, Stavroyiannis (2018) determina el VaR para *Bitcoin* y lo compara con el índice SPX 500 y la serie de tiempo de precios del oro, resultando en que *Bitcoin* es una moneda volátil y que excede el VaR en más ocasiones comparado con los otros activos. Asimismo, Troster, Kumar Tiwari, Shahbaz y Nicolás Macedo (2019) llevan a cabo un análisis de rendimiento y riesgo de *Bitcoin* utilizando modelos GARCH y GAS con colas pesadas, destacando que estos análisis son reelevantes al utilizar bitcoin en coberturas óptimas y estrategias de inversión.

I. Álgebra Lineal

En este primer capítulo se presentan definiciones de Álgebra Lineal como: *vectores, matrices, funciones y polinomios*. En el presente capítulo se seguirá un tratamiento de Algebra lineal basado en Holton (2014) y Poole (2011).

I.1. Álgebra Lineal

1.1.1. Preliminares

1.1.1.1. Vectores

\mathbb{R}^n se define como el conjunto de todas las n – *adas ordenadas* de números reales escritos como vectores renglón o columna.

Es decir, un vector \mathbf{v} en \mathbb{R}^n es de la forma $[v_1, v_2, \dots, v_n]$ o $\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$. [1.1]

Las entradas individuales de \mathbf{v} son sus componentes; v_i se llama el componente i -ésimo. En \mathbb{R}^n , se tiene que si $\mathbf{u} = [u_1, u_2, \dots, u_n]$ y $\mathbf{v} = [v_1, v_2, \dots, v_n]$, en la definición de suma vectorial el componente i -ésimo de $\mathbf{u} + \mathbf{v}$ es $u_i + v_i$ y en la multiplicación escalar el componente i -ésimo de $c\mathbf{v}$ es cv_i .

Definición: Un vector \mathbf{v} es una **combinación lineal** de vectores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ si existen escalares c_1, c_2, \dots, c_k tales que $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_k\mathbf{v}_k$. Los escalares c_1, c_2, \dots, c_k se llaman coeficientes de la combinación lineal.

1.1.1.2. Producto cruz

Se busca que dados dos vectores no paralelos \mathbf{u} y \mathbf{v} , se produzca un tercer vector \mathbf{n} que sea ortogonal⁸ tanto a \mathbf{u} como a \mathbf{v} . Un método, es usar una construcción conocida como producto cruz de vectores. Para \mathbb{R}^3 , se tiene que:

⁸ Término que hace referencia a la idea de perpendicularidad para los vectores en \mathbb{R}^n

Definición: El producto cruz de $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ es el vector $\mathbf{u} \times \mathbf{v}$ definido por

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix}$$

1.1.1.3. Sistemas de ecuaciones

Definición: Una **ecuación lineal** en las n variables x_1, x_2, \dots, x_n es una ecuación que puede escribirse en la forma:

$$a_1 x_1 + a_2 x_2 + \dots + a_n x_n = b \quad [1.2]$$

donde los coeficientes a_1, a_2, \dots, a_n y el término b son constantes. Un sistema de ecuaciones lineales es un conjunto finito de ecuaciones lineales, cada una con las mismas variables. Una solución de un sistema de ecuaciones lineales es un vector que simultáneamente es una solución de cada ecuación en el sistema. El conjunto solución de un sistema de ecuaciones lineales es el conjunto de todas las soluciones del sistema. Al proceso de encontrar el conjunto solución de un sistema de ecuaciones lineales se le conocerá como “resolver el sistema”.

1.1.1.4. Polinomios

Un polinomio de una sola variable x es una expresión de la siguiente forma:

$$p(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n \quad [1.3]$$

donde a_0, a_1, \dots, a_n son constantes ($a_0 \neq 0$), llamados coeficientes de p . El entero n se llama grado de p , que se denota al escribir grado de $p = n$. Un polinomio de grado cero se llama polinomio constante.

1.1.1.5. Matrices

Definición: Una **matriz** es un arreglo rectangular de números llamados entradas, o elementos de la matriz.

Una matriz de $m \times n$ tiene m renglones y n columnas. Una matriz de $1 \times n$ se llama matriz renglón (o vector renglón), y una matriz de $n \times 1$ se llama matriz columna (o vector columna). La entrada de la matriz A en el renglón i y la columna j se denota mediante a_{ij} . Los componentes de la matriz se asumen como números reales a menos que se establezca lo contrario.

La matriz escalar es una matriz identidad denotada por I si el escalar en la diagonal es 1. Una matriz cero se indica como $\mathbf{0}$.

Para generalizar la suma de vectores, se define la suma de matrices por componentes. Si $A = [a_{ij}]$ y $B = [b_{ij}]$ son matrices de $m \times n$, su suma $A + B$ es la matriz de $m \times n$ que se obtiene al sumar las entradas correspondientes. Por lo tanto, $A + B = [a_{ij} + b_{ij}]$. La suma y multiplicación de matrices requiere que las matrices sean de dimensiones compatibles.

Definición: Si A es una matriz de $m \times n$ y B es una matriz de $n \times r$, entonces el producto $C = AB$ es una matriz de $m \times r$. La entrada (i, j) del producto se calcula del modo siguiente: $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}$

Definición: La transpuesta de una matriz A de $m \times n$ es A^T de $n \times m$, que se obtiene al intercambiar los renglones y columnas de A . Esto es: la i -ésima columna de A^T es el i -ésimo renglón de A para toda i .

Definición: Si A es una matriz de $n \times n$ una inversa de A es una matriz denotada como A^{-1} de $n \times n$ con la siguiente propiedad $AA^{-1} = I$ y $A^{-1}A = I$ donde $I = I_n$ es la matriz identidad. Si tal A^{-1} existe, entonces A es **invertible**.

Teorema: Si A es una matriz invertible de $n \times n$, entonces el sistema de ecuaciones lineales dado por $Ax = \mathbf{b}$ tiene la solución única $\mathbf{x} = A^{-1}\mathbf{b}$ para cualquier \mathbf{b} en \mathbb{R}^n . En particular, si $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, entonces A es invertible cuando $ad - bc \neq 0$, en cuyo caso, $A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Si $ad - bc = 0$ entonces A no es invertible. La expresión $ad - bc$ se llama determinante de A y se denota como $\det A$

1.1.1.6. Funciones

Una función T es una regla que asigna elementos del conjunto A a elementos del conjunto B . A es el dominio de T y B contiene el rango. Algunos tipos de funciones son:

1. Funciones de \mathbb{R} a \mathbb{R} , por ejemplo, la función logaritmo.
2. Funciones de \mathbb{R}^n a \mathbb{R} .
3. Funciones de \mathbb{R}^n a \mathbb{R}^n .

Si una función T tiene inversa, se denota como T^{-1} . La derivada de la función real T se denotará como T' . Se indica el valor de una función T en un punto en particular a como $T(a)$ o $T|_a$. Si se considera $T: \mathbb{R}^p \rightarrow \mathbb{R}^m$ y $S: \mathbb{R}^n \rightarrow \mathbb{R}^p$ como funciones, entonces se forma $T \circ S$, la composición de las dos funciones de \mathbb{R}^n a \mathbb{R}^m , definida como sigue: $T \circ S(x) = T(S(x))$.

El gradiente ∇T y el Hessiano $\nabla^2 T$ de una función $T: \mathbb{R}^n \rightarrow \mathbb{R}$ es el vector de sus primeras derivadas parciales y la matriz de las segundas derivadas parciales:

$$\nabla T = \begin{pmatrix} \frac{\partial T}{\partial x_1} \\ \frac{\partial T}{\partial x_2} \\ \vdots \\ \frac{\partial T}{\partial x_n} \end{pmatrix} \quad \text{y} \quad \nabla^2 T = \begin{pmatrix} \frac{\partial^2 T}{\partial x_1^2} & \frac{\partial^2 T}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 T}{\partial x_1 \partial x_n} \\ \frac{\partial^2 T}{\partial x_2 \partial x_1} & \frac{\partial^2 T}{\partial x_2^2} & & \frac{\partial^2 T}{\partial x_2 \partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2 T}{\partial x_n \partial x_1} & \frac{\partial^2 T}{\partial x_n \partial x_2} & & \frac{\partial^2 T}{\partial x_n^2} \end{pmatrix} \quad [1.4]$$

El Jacobiano JT de la función $T: \mathbb{R}^n$ a \mathbb{R}^m es la matriz de sus primeras derivadas:

$$JT = \begin{pmatrix} \frac{\partial T_1}{\partial x_1} & \frac{\partial T_1}{\partial x_2} & \cdots & \frac{\partial T_1}{\partial x_n} \\ \frac{\partial T_2}{\partial x_1} & \frac{\partial T_2}{\partial x_2} & & \frac{\partial T_2}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial T_m}{\partial x_1} & \frac{\partial T_m}{\partial x_2} & & \frac{\partial T_m}{\partial x_n} \end{pmatrix} \quad [1.5]$$

El Hessiano de una función $T: \mathbb{R}^n \rightarrow \mathbb{R}$ es el Jacobiano de su gradiente.

1.1.1.7. Alternativa matricial para Polinomios

La [ecuación 1.3.] se puede generalizar a dimensiones mayores: Un polinomio en varias variables es lineal si tiene la forma:

$$p(x) = \mathbf{b}x + a \quad [1.6]$$

con \mathbf{b} un vector renglón $n -$ dimensional y a un escalar. Por otra parte, la forma cuadrática tiene la forma:

$$p(x) = x' \mathbf{c}x + \mathbf{b}x + a \quad [1.7]$$

con \mathbf{c} una matriz $n \times n$, \mathbf{b} un vector renglón $n -$ dimensional y a un escalar, comúnmente se asume que \mathbf{c} es simétrica.

1.1.2. Aproximaciones Gradiente y gradiente Hessiano

Los polinomios se utilizan frecuentemente para aproximar funciones. A continuación, se presentan algunas formas de llevar a cabo lo anterior.

1.1.2.1. Aproximaciones Univariadas

Se considera una función $f: \mathbb{R} \rightarrow \mathbb{R}$ que es diferenciable en un intervalo abierto sobre un punto $x^{[0]} \in \mathbb{R}$. El polinomio lineal lleva a cabo una buena aproximación para f , al menos para un intervalo pequeño en $x^{[0]}$.

$$p_1(x) = f(x^{[0]}) + f'(x^{[0]})(x - x^{[0]}) \quad [1.8]$$

Esto se debe a que:

1. p_1 es igual a f en $x^{[0]}$;
2. p_1 tiene la misma primera derivada que f en $x^{[0]}$.

Si f es dos veces diferenciada en un intervalo abierto sobre $x^{[0]}$, se puede mejorar la aproximación con un polinomio cuadrado

$$p_2(x) = f(x^{[0]}) + f'(x^{[0]})(x - x^{[0]}) + \frac{f''(x^{[0]})}{2}(x - x^{[0]})^2 \quad [1.9]$$

1.1.2.2. Aproximaciones Multivariadas: aproximaciones Gradiente y gradiente Hessiano

Las aproximaciones polinomiales de las ecuaciones [1.8.] y [1.9.] se pueden generalizar a múltiples dimensiones. Para $f: \mathbb{R}^n \rightarrow \mathbb{R}$ los gradientes reemplazan a las primeras derivadas y los Hessianos reemplazan a las segundas derivadas, por lo cual el polinomio lineal y cuadrático de las ecuaciones 1.8 y 1.9 se convierten en:

$$p_1(x) = f(x^{[0]}) + \nabla f(x^{[0]})'(x - x^{[0]}) \quad [1.10]$$

$$p_2(x) = f(x^{[0]}) + \nabla f(x^{[0]})'(x - x^{[0]}) + \frac{1}{2}(x - x^{[0]})' \nabla^2 f(x^{[0]})(x - x^{[0]}) \quad [1.11]$$

1.1.3. Interpolación

La interpolación es un procedimiento para ajustar una función a un conjunto de puntos de tal manera que la función intercepte cada uno de los puntos. Considérese m puntos $(x^{[k]}, y^{[k]})$ donde $x^{[k]} \in \mathbb{R}^n$, $y^{[k]} \in \mathbb{R}$, y $x^{[k]}$ son valores distintos.

Se desea construir una función $f: \mathbb{R}^n \rightarrow \mathbb{R}$ tal que $y^{[k]} = f(x^{[k]})$ para todo k , lo cual se lleva a cabo con interpolaciones lineales, interpolaciones cuadráticas.

1.1.4. Mínimos cuadrados ordinarios

Se consideran l puntos $(x^{[k]}, y^{[k]})$ donde $x^{[k]} \in \mathbb{R}^n, y^{[k]} \in \mathbb{R}$. Se desea construir una función $f: \mathbb{R}^n \rightarrow \mathbb{R}$ de la forma $f(x) = \beta_m f_m(x) + \beta_{m-1} f_{m-1}(x) + \dots + \beta_1 f_1(x)$ a los datos de tal manera que se minimice la suma de cuadrados $\sum_{k=1}^l [x^{[k]} - y^{[k]}]^2$.

De manera matricial se expresa la fórmula de la suma de cuadrados como sigue:

$$\sum_{k=1}^l [x^{[k]} - y^{[k]}]^2 = (f\beta - y)'(f\beta - y) = \beta' f' f \beta - 2y' f \beta + y' y$$

siendo un polinomio cuadrado en β . La suma de cuadrados tiene un mínimo único cuando

$$\beta = (f' f)^{-1} f' y \quad [1.12]$$

1.1.5. Factorización de Cholesky

1.1.5.1. Formas cuadráticas

Definición: Una forma cuadrática con n variables es una función $f: \mathbb{R}^n \rightarrow \mathbb{R}$ de la forma

$$f(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} \quad [1.13]$$

Donde A es una matriz simétrica de $n \times n$ y \mathbf{x} está en \mathbb{R}^n . Una forma cuadrática es una suma de términos, cada uno de los cuales tiene grado dos. A se conoce como matriz asociada con f

Se llama forma cuadrática en x y y a una expresión de la forma

$$ax^2 + by^2 + cxy = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & c/2 \\ c/2 & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

De igual modo, se llama forma cuadrática x, y y z

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a & d/2 & e/2 \\ d/2 & b & f/2 \\ e/2 & f/2 & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Definición: Una forma cuadrática $f(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ se clasifica como una de las siguientes:

1. Definida positiva si $f(\mathbf{x}) > 0$ para toda $\mathbf{x} \neq \mathbf{0}$
2. Semidefinida positiva si $f(\mathbf{x}) \geq 0$ para toda \mathbf{x}
3. Definida negativa si $f(\mathbf{x}) < 0$ para toda $\mathbf{x} \neq \mathbf{0}$
4. Semidefinida negativa si $f(\mathbf{x}) \leq 0$ para toda \mathbf{x}
5. Indefinida si $f(\mathbf{x})$ toma valores tanto positivos como negativos

Una matriz simétrica A se llama definida positiva, semidefinida positiva, definida negativa, semidefinida negativa o indefinida si la forma cuadrática asociada $f(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ tiene la propiedad correspondiente.

Teorema: Sea A una matriz simétrica de $n \times n$. La forma cuadrática $f(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ es

- a. Positiva definida si y solo si todos los eigenvalores de A son positivos
- b. Semidefinida positiva si y solo si todos los eigenvalores de A son no negativos
- c. Definida negativa si y solo si todos los eigenvalores de A son negativos
- d. Semidefinida negativa si y solo si todos los eigenvalores de A son no positivos
- e. Indefinida si y solo si A tiene eigenvalores tanto positivos como negativos

Se podría decir que una matriz definida positiva es análoga a números positivos y semidefinida positiva análogo a los números negativos. La diferencia entre matrices semidefinida y definidas es que las primeras son singulares y las otras no. Esto es así debido a que una matriz es singular si el eigenvalor es cero.

1.1.5.2. Matriz de raíces cuadradas

Cualquier matriz semidefinida positiva \mathbf{h} puede ser factorizada de la forma $\mathbf{h} = \mathbf{k}\mathbf{k}'$ para alguna matriz cuadrada real \mathbf{k} , que se puede pensar como una matriz raíz cuadrada de \mathbf{h} . La matriz \mathbf{k} no es única, por lo cual múltiples factorizaciones son posibles para \mathbf{h} . Si \mathbf{h} es no singular (definida positiva), \mathbf{k} será no singular.

1.1.5.3. Matriz de Cholesky

La factorización de Cholesky es una factorización sencilla de la forma $\mathbf{h} = \mathbf{k}\mathbf{k}'$. Cualquier matriz semidefinida positiva tiene una factorización de la forma

$$\mathbf{h} = \mathbf{g}\mathbf{g}' \quad [1.14]$$

donde \mathbf{g} es una matriz triangular inferior. El algoritmo envuelve dos cálculos:

1. Calcular los elementos de la diagonal $g_{i,i}$ implica sacar la raíz cuadrada
2. Calcular los elementos fuera de la diagonal $g_{i,i} \ i > j$ implica dividir algún número sobre el último elemento de la diagonal calculado.

Para una matriz definida positiva \mathbf{h} , resolviendo para cada elemento de la diagonal $g_{i,i}$ implica obtener la raíz cuadrada de un número no negativo. En la práctica se eligen las raíces positivas. El algoritmo de Cholesky es un medio para comprobar si la matriz esta semidefinida positiva.

II. Probabilidad

Para llevar a cabo el estudio formal de los modelos para estimar el VaR, es necesario establecer las bases probabilísticas sobre las cuales están planteados los supuestos de dichos modelos. Para ello, en este capítulo se expondrán las definiciones como: *espacios de probabilidad, variables aleatorias, vectores aleatorios y distribuciones de probabilidad*. En el presente capítulo se seguirá un tratamiento de probabilidad basado en Bowerman, O'Connell y Koehler (2007); Hernández (2016); Holton (2014); Rincón (2006); Rincón (2007,2014)

II.1. Espacios de Probabilidad

Sea Ω un conjunto arbitrario no vacío el cual contienen los resultados “ $\{A\}$ ” de un experimento aleatorio⁹ y donde $P\{A\}$ denota la probabilidad p asociada con el conjunto de resultados $\{A\}$ ¹⁰. Ahora bien el modelo matemático para estudiar a $P\{A\}$, es conocido como **espacio de probabilidad**; el cual consta de la terna ordenada (Ω, \mathcal{A}, P) , donde Ω es llamado espacio muestral¹¹, \mathcal{A} una σ – álgebra¹² de subconjuntos de Ω , y P una medida de probabilidad definida sobre \mathcal{A} .

Una σ – álgebra es una estructura que nos permitirá coleccionar subconjuntos de Ω de interés y que a su vez satisface las siguientes propiedades:

i. Que al realizar el experimento algo ocurra, es decir, $\Omega \in \mathcal{A}$.

ii. Si A ocurre, también pediremos que “no ocurra A ”, es decir,

$$A \in \mathcal{A} \text{ entonces } A^c \in \mathcal{A}$$

iii. Si A_1, A_2, \dots son elementos de \mathcal{A} , entonces que “ocurra alguno de los A_n ” también es un elemento de \mathcal{A} , es decir,

$$\{A_1, A_2, \dots\} \in \mathcal{A} \text{ entonces } \bigcup_{i=1}^{\infty} A_i \in \mathcal{A} \text{ y } \bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$$

⁹Se entiende por experimento aleatorio todo aquel experimento tal que cuando se le repite n veces y bajo las mismas condiciones iniciales, el resultado que se obtiene no siempre es el mismo.

¹⁰ Los $\{ \}$ indican que el argumento $P\{A\} = F\{A\}$ es la probabilidad de un intervalo o conjunto y los $()$ indican que el argumento $P(a) = f(a)$ es la probabilidad de asociada a un punto.

¹¹ El espacio muestral es el conjunto que agrupa a todos los posibles resultados del experimento aleatorio en cuestión.

¹² Se lee como “sigma-álgebra”.

Una **medida de probabilidad** P , es una función $P : \mathcal{A} \rightarrow [0,1]$ que cumple las siguientes condiciones:

I. $P\{\Omega\} = 1 \Rightarrow$ el evento cierto tenga probabilidad igual a 1.

II. $P\{A\} \geq 0$ para toda $A \in \mathcal{A} \Rightarrow$ la probabilidad de cualquier evento sea un \mathbb{R} no negativo.

III. Si $\{A_n\}_{n \geq 1}$ son conjuntos mutuamente disjuntos (o ajenos)¹³ en \mathcal{A} entonces P es llamada σ -aditiva, si

$$P\{\cup_{n=1}^{\infty} A_n\} = \sum_{n=1}^{\infty} P\{A_n\}$$

II.2. Variables Aleatorias

Cuando se habla de una colección muy grande de elementos, se está interesado en ciertas cantidades descriptivas de todo el conjunto de los elementos. Tales cantidades descriptivas serán mejor conocidas como variables aleatorias.

2.2.1. Definición de variable aleatoria

Se define a una variable aleatoria (**v.a.**) como una función $X: \Omega \rightarrow \mathbb{R}$. Las variables aleatorias se clasifican en discretas y continuas. Se tiene una v.a. discreta cuando el rango de X es finito o infinito numerable y se tiene una v.a. continua si el rango es un intervalo contenido en los números reales. Por otra parte, y de aquí en lo sucesivo a los elementos de \mathcal{A} los denotaremos por $\{X = x_j\}$ y serán llamados eventos.

2.2.2. Funciones de probabilidad y funciones de distribución acumulada

Definición: Sea X tiene una v.a discreta con valores x_0, x_1, \dots . La **función de masa de probabilidad** de X , denotada por $f(x): \mathbb{R} \rightarrow \mathbb{R}$, se define como sigue:

$$f(x_j) = \begin{cases} P\{X = x_j\}, & \text{si } x = x_0, x_1, \dots \\ 0, & \text{otro caso} \end{cases} \quad [2.1]$$

Definición: Sea X una v.a continua. Decimos que la función integrable y no negativa $f(x): \mathbb{R} \rightarrow \mathbb{R}$ es la **función de densidad de probabilidad** de X si para cualquier intervalo $[a, b]$ de \mathbb{R} se cumple que:

$$P(X \in [a, b]) = \int_a^b f(x) dx \quad [2.2]$$

¹³ conjuntos mutuamente disjuntos son aquellos cuya intersección es el conjunto vacío, esto es, $A_i \cap A_j = \emptyset$ para valores de i y j distintos.

Las funciones de probabilidad cumplen con la propiedad de no negatividad y la suma extendida de estas funciones sobre el eje x son la unidad; es decir, para el caso discreto tenemos que $\sum_i f(x_i) = 1$ y para el caso continuo $\int_{-\infty}^{+\infty} f(x)dx = 1$.

Definición: Sea X una v.a cualquiera. La función de distribución de X , denotada por $F(x): \mathbb{R} \rightarrow \mathbb{R}$, se define como la probabilidad

$$F(x) = P\{X \leq x\} \quad [2.3]$$

Para el caso discreto, suponiendo que $f(x)$ es la función de probabilidad de una v.a. discreta X , se tiene que la **función de distribución o función de probabilidad acumulada**, se define como la probabilidad:

$$F(x) = \sum_{u \leq x} f(u) \quad [2.4]$$

Para el caso continuo se define como sigue:

$$F(x) = \int_{-\infty}^x f(u) du$$

Proposición: Toda función de distribución satisface las siguientes propiedades:

- i. $\lim_{n \rightarrow \infty} F_n(x) = 1$
- ii. $\lim_{n \rightarrow -\infty} F_n(x) = 0$
- iii. Si $x_1 \leq x_2$ implica $F(x_1) \leq F(x_2)$
- iv. $F(x) = F(x+)$

2.2.3. Momentos empíricos

En los siguientes apartados se definen algunas características de las v.a.

2.2.3.1. Esperanza

Definición: Sea X una v.a. discreta con función de probabilidad $f(x)$. La esperanza de X se define como el número:

$$E(X) = \sum_x xf(x) \quad [2.5]$$

Suponiendo que esta suma es absolutamente convergente, es decir, cuando la suma de los valores absolutos es convergente. Por otro lado, si X es continua con función de densidad $f(x)$ entonces la esperanza es:

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

Suponiendo que esta integral es absolutamente convergente, es decir, cuando la integral de los valores absolutos es convergente.

2.2.3.2. Esperanza de una función de una variable aleatoria

Proposición: Sea X una v.a. discreta y sea $g: \mathbb{R} \rightarrow \mathbb{R}$ una función tal que $g(x)$ es una v.a. con esperanza finita. Entonces se tiene que:

$$E[g(x)] = \sum_x g(x)f_X(x) \quad [2.6]$$

Proposición: Sea X una v.a. continua y sea $g: \mathbb{R} \rightarrow \mathbb{R}$ una función tal que $g(x)$ es una v.a. con esperanza finita. Entonces se tiene que:

$$E[g(x)] = \int_{-\infty}^{\infty} g(x)f_X(x)dx \quad [2.7]$$

2.2.4. Medidas de dispersión

2.2.4.1. Varianza

Definición: Sea X una v.a. discreta con función de probabilidad $f(x)$. La varianza de X se define como el número:

$$Var(x) = \sum_x (x - \mu)^2 f(x) \quad [2.8]$$

Cuando esta suma es convergente y en donde μ es la esperanza de X . Para una v.a. continua X con función de densidad $f(x)$ se define como sigue:

$$Var(x) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx$$

Cuando esta integral es convergente.

La **desviación estándar** es la raíz cuadrada de la varianza y se denota por σ ó $std(X)$

2.2.5. Sesgo

El sesgo es una medida de asimetría de la distribución de probabilidad de las variables aleatorias. El sesgo se define como sigue:

$$Sesgo(x) = \frac{E[(X - \mu)^3]}{\sigma^3} \quad [2.9]$$



Figura 2.1: Se ilustra lo que es el sesgo con dos f.d.p. que tienen la misma media y desviación estándar. La f.d.p. de la izquierda presenta sesgo positivo mientras que la de derecha esta sesgada negativamente. Fuente: Holton (2014)

2.2.6. Curtosis

La curtosis es un parámetro que describe la forma que toma la f.d.p. de las variables aleatorias. Este parámetro se compara con el de una distribución normal, que tiene un valor de 3. Se define como sigue:

$$\text{Curtosis}(x) = \frac{E[(X - \mu)^4]}{\sigma^4} \quad [2.10]$$

2.2.7. Cuantiles

Sea X una v.a. con función de distribución denotada por F_X . Un α – cuantil de X es cualquier número x_α que cumpla que

$$F_X(x_\alpha) = P(X \leq x_\alpha) = \alpha \quad [2.11]$$

Al aplicar el VaR, los α -cuantiles existen y son únicos con $\alpha \in (0,1)$. Para este caso, un α -cuantil es un parámetro y es el valor correspondiente al valor de la función de distribución inversa evaluado en α y se denota por $F^{-1}(\alpha)$.

2.2.8. Momentos

Se define el n -ésimo momento de una v.a. X , cuando existe, como el número $E(X^n)$ para cualquier número natural n . El n -ésimo momento central de X , cuando existe, es el número $E[(X - \mu)^n]$, en donde $\mu = E(X)$.

El primer momento de X es la media y el segundo momento central es la varianza.

II.3. Vectores aleatorios

2.3.1. Vector aleatorio

Definición: Un **vector aleatorio** de dimensión dos es un vector de la forma (X, Y) en donde cada coordenada es una v.a. De manera análoga, se definen vectores aleatorios multidimensionales (X_1, X_2, \dots, X_n) , de aquí en adelante se denotará como \mathbf{X} .

2.3.2. Función de distribución conjunta, marginal y condicional

Definición: La función de distribución de un vector (X, Y) denotada por $F(x, y): \mathbb{R}^2 \rightarrow [0, 1]$, se define de la siguiente manera:

$$F(x, y) = P(X \leq x, Y \leq y) \quad [2.12]$$

Para encontrar la función de distribución en el caso continuo se tiene que

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) dv du. \quad [2.13]$$

Mientras que en el caso discreto se tiene que

$$F(x, y) = \sum_{u \leq x} \sum_{v \leq y} (u, v) \quad [2.14]$$

Definición: Sea (X, Y) un vector aleatorio, continuo o discreto con función de distribución $F(x, y)$. La **función de distribución marginal** de la variable X se define como la función de una variable $F(x) = \lim_{y \rightarrow \infty} F(x, y)$. Análogamente, la función de distribución marginal de la variable Y se define como la función $F(y) = \lim_{x \rightarrow \infty} F(x, y)$.

Definición: Sea (X, Y) un vector aleatorio discreto o continuo con función de masa de probabilidad o de densidad $f_{X,Y}(x, y)$. Sea y un valor de la variable Y tal que $f_Y(y) \neq 0$. La función de distribución condicional de X dado que $Y = y$, es la función:

$$x \rightarrow F_{X|Y}(x|y) = \begin{cases} \sum_{u \leq x} f_{X|Y}(u|y) & \text{en el caso continuo} \\ \int_{-\infty}^x f_{X|Y}(u|y) du & \text{en el caso discreto} \end{cases} \quad [2.15]$$

2.3.3. Función de probabilidad conjunta, marginal y condicional

Definición: Se le llama **función de probabilidad conjunta** de las variables X y Y a la función de probabilidad del vector aleatorio discreto (X, Y) en donde X toma los valores x_1, x_2, \dots y Y toma los valores y_1, y_2, \dots es la función $f(x, y): \mathbb{R}^2 \rightarrow [0, 1]$ dada por:

$$f(x, y) = \begin{cases} P(X = x, Y = y) & \text{si } (x, y) \in \{x_1, x_2, \dots\} \times \{y_1, y_2, \dots\} \\ 0 & \text{otro caso} \end{cases} \quad [2.16]$$

Definición: Sea (X, Y) un vector aleatorio continuo. Se dice que la función integrable y no negativa $f(x, y): \mathbb{R}^2 \rightarrow [0, \infty)$ es la función de densidad del vector (X, Y) , o bien, que es la **función de densidad conjunta** de las variables X y Y si para todo par (x, y) en \mathbb{R}^2 se cumple la igualdad:

$$P(X \leq x, Y \leq y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) dv du. \quad [2.17]$$

Toda función de densidad $f(x, y)$ satisface las siguientes propiedades: 1) $f(x, y) \geq 0$; 2) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$.

Además, cuando $f(x, y)$ es continua, $f(x, y) = \frac{\partial^2}{\partial y \partial x} F(x, y)$

Definición: Sea $f(x, y)$ la función de densidad del vector aleatorio continuo (X, Y) . Se define la función de **densidad marginal** de la variable X como la siguiente integral:

$$f(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad [2.18]$$

y a la función de densidad marginal de Y como:

$$f(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

La definición de función de probabilidad marginal para vectores discretos involucra una suma en lugar de una integral.

Definición: Sea (X, Y) un vector aleatorio discreto o continuo con función de probabilidad o de densidad $f_{X,Y}(x, y)$.

Sea y un valor de la variable Y tal que $f_Y(y) \neq 0$. A la función $x \rightarrow f_{X|Y}(x|y)$ definida a continuación, se le llama la **función de probabilidad o densidad condicional** de X dado que $Y = y$:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)} \quad [2.19]$$

2.3.4. Independencia de variables aleatorias

Una interpretación de independencia de eventos es que la ocurrencia de uno de los eventos no proporciona información que modifique la probabilidad del segundo evento. Dicho de otro modo, la ocurrencia de un evento “ B ”, no afecta la probabilidad del evento “ A ”, i.e.,

$$P(A|B) = P(A). \quad [2.20]$$

Si se traslada la definición de la ecuación 2.20 a variables aleatorias, intuitivamente podríamos decir que, las variables aleatorias $X \wedge Y$ son independientes si cada acontecimiento que implica solamente a X es independiente de cada evento que involucra sólo Y .

La condición de independencia de n variables aleatorias X_1, \dots, X_n con distribución conjunta $F_{X_1, \dots, X_n}(x_1, \dots, x_n)$ y con sus respectivas distribuciones marginales $F_{X_1}(x_1) \cdot F_{X_n}(x_n)$ describe que estas variables son independientes si para cualesquiera valores x_1, \dots, x_n se cumple la igualdad:

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = F_{X_1}(x_1) \cdot F_{X_n}(x_n) \quad [2.21]$$

La independencia de v.a. también se puede definir en términos de función de densidad. Además, la transformación de una variable aleatoria independiente (v.a.i) resulta nuevamente en v.a.i. Hernández (2016) ilustra con un ejemplo el hecho de que dadas las v.a.i. X_1, \dots, X_n , y las v.a. Y_1, \dots, Y_n ; si Y_i es determinada solamente por X_i , entonces Y_i heredará la independencia de X_i ; y se llega a la siguiente definición:

Definición. Sea X_1, X_2, \dots, X_n variables aleatorias independientes. Para cada i , sea $h_i: R \rightarrow R$ una función y definamos la variable aleatoria

$$Y_i = h_i(X_i) \quad [2.22]$$

Luego entonces Y_1, Y_2, \dots, Y_n también serán independientes.

Por último, decimos que una sucesión de v.a. $\{X_n\}$ es *independiente e idénticamente distribuida* (en lo futuro *i.i.d.*) si todas las variables X_1, X_2, \dots son mutuamente disjuntas y si tienen la misma distribución de probabilidad. Una sucesión i.i.d. no siempre implica que las probabilidades para todos los elementos del espacio muestral deben ser iguales. Por ejemplo, repetidos tiros de dados cargados producirán una sucesión de v.a. i.i.d., a pesar de que los resultados estén sesgados.

2.3.5. Parámetros de vectores aleatorios

2.3.5.1. Esperanza

Definición: Sea el vector aleatorio X . De acuerdo con Holton (2014), cuando cada coordenada del vector tiene esperanza finita se define la **esperanza** de X como el vector numérico :

$$E(X) = (E(X_1), \dots, E(X_n)) \quad [2.23]$$

Esperanza de una función de un vector aleatorio: Sea X un vector aleatorio n -dimensional y f una función de \mathbb{R}^n a \mathbb{R} que define a una v.a. $f(X)$. Para calcular la media de $f(X)$, si X es discreta con función de masa de probabilidad ϕ se tiene que

$$E[f(X)] = \sum_x f(x) \phi(x) \quad [2.24]$$

Para el caso continuo con f.d.p ϕ se tiene que

$$\begin{aligned} E[f(X)] &= \int_{\mathbb{R}^n} f(x) \phi(x) dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n) \phi(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \end{aligned} \quad [2.25]$$

2.3.5.2. Covarianza

Para el i -ésimo y j -ésimo componente de X , se denota la covarianza como $Cov(X_i, X_j)$. Por definición, la covarianza es simétrica con $Cov(X_i, X_j) = Cov(X_j, X_i)$. Además, la covarianza de cualquier componente X_i consigo mismo es la varianza de ese componente:

$$\begin{aligned} Cov(X_i, X_i) &= E[(X_i - E(X_i))(X_i - E(X_i))] \\ &= E[(X_i - \mu_i)^2] = \sigma_i^2 \end{aligned} \quad [2.26]$$

Se tiene la matriz de varianzas covarianzas del vector X :

$$Var(X) = \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) & Cov(X_1, X_n) \\ \vdots & \vdots & \vdots \\ Cov(X_n, X_1) & Cov(X_n, X_2) & Var(X_n) \end{pmatrix}$$

Proposición: La matriz de varianzas covarianzas es simétrica y positiva definida, lo cual significa que para cualquier vector $\theta = (\theta_1, \dots, \theta_n)$ de \mathbb{R}^n se cumple la desigualdad $\langle Var(X)\theta, \theta \rangle \geq 0$ en donde $\langle \cdot, \cdot \rangle$ denota el producto interior usual de \mathbb{R}^n .

La covarianza es una medida de la tendencia de como varían juntos dos componentes de un vector aleatorio. La magnitud de la covarianza depende de las desviaciones estándar de los dos componentes. Para obtener más información de como varían juntos dos componentes, se escala la covarianza para obtener la correlación.

2.3.6. Correlación

La correlación $\rho_{i,j}$ del componente i -ésimo y j -ésimo de un vector aleatorio X se define como $\rho_{i,j} = \frac{Cov(X_i, X_j)}{\sigma_i \sigma_j}$

La matriz de correlación del vector X se define como sigue

$$\rho(X) = \begin{pmatrix} \rho(X_1) & \dots & \rho(X_1, X_n) \\ \vdots & & \vdots \\ \rho(X_n, X_1) & \dots & \rho(X_n) \end{pmatrix} \quad [2.27]$$

Cumple con la propiedad de ser simétrica y tener todos los elementos de la diagonal iguales a uno.

2.3.7. Momentos conjuntos

Sea el vector (X, Y) con función generadora de momentos $M_{X,Y}(s, t) = E(e^{sX} e^{tY})$ para valores reales s y t con esperanza convergente. Las variables X y Y son independientes si y solo si $M_{X,Y}(s, t) = M_X(s)M_Y(t)$.

2.3.8. Polinomios lineales de vectores aleatorios

De acuerdo con Holton (2014), las variables aleatorias se pueden definir como polinomios lineales de vectores aleatorios. Sea un vector aleatorio X con media μ y matriz de covarianza Σ . Se define una v.a. Y como un polinomio lineal de X , de la forma:

$$Y = \mathbf{b}X + a \quad [2.28]$$

Donde \mathbf{b} es un vector renglón de n componentes y a pertenece a los números reales.

La media y la varianza se definen como:

$$\begin{aligned} E(Y) &= \mathbf{b}\mu + a \\ Var(Y) &= \mathbf{b}\Sigma\mathbf{b}' \end{aligned} \quad [2.29]$$

Las fórmulas anteriores se pueden generalizar en donde Y es un vector aleatorio con m componentes definido como un polinomio lineal de un vector X de n componentes, es decir,

$$Y = \mathbf{b}X + a \quad [2.30]$$

Donde \mathbf{b} es una matriz de $m \times n$ y \mathbf{a} es un vector de m componentes. Si X tiene media μ_X y matriz de covarianza Σ_X , entonces la media y la matriz de covarianza de Y se obtiene mediante la fórmula

$$\begin{aligned} \mu_Y &= \mathbf{b}\mu_X + \mathbf{a} \\ \Sigma_Y &= \mathbf{b}\Sigma_X\mathbf{b}' \end{aligned} \quad [2.31]$$

II.4. Distribuciones de probabilidad

2.4.1. Distribución normal

Una v.a. X tiene una distribución normal, con función de densidad definida de la siguiente forma:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in \mathbb{R}; \mu \in \mathbb{R} \text{ y } \sigma > 0 \quad [2.32]$$

La distribución de probabilidad normal se denota como $X \sim N(\mu, \sigma^2)$, lo cual quiere decir que la forma de la curva normal dependerá de la media y de la desviación estándar. Algunas de las propiedades de esta curva son:

1. La curva normal tiene su centro en la media poblacional μ
2. La media μ corresponde al punto más alto de la curva normal
3. La curva normal es simétrica con respecto a la media de la población
4. La curva normal es una distribución de probabilidades, por lo cual el área total bajo la curva es igual a 1.

Un polinomio lineal de una v.a. es normal si $X \sim N(\mu, \sigma^2)$, $bX + a \sim N(b\mu + a, (b\sigma)^2)$ para cualesquiera constantes $a, b \in \mathbb{R}$, esto significa que cualquier v.a. $X \sim N(\mu, \sigma^2)$ puede ser expresada como un polinomio lineal de una v.a. $Z \sim N(0,1)$:

$$X = \sigma Z + \mu \quad [2.33]$$

Proposición: Sea X con distribución $X \sim N(\mu, \sigma^2)$, entonces

$$Z = \frac{X - \mu}{\sigma} \sim N(0,1) \quad [2.34]$$

Comúnmente se usa la letra Z para denotar una v.a. con distribución normal estándar. El valor z que corresponde a X , se define como la ecuación [2.33]. Este resultado denota la cantidad de desviaciones estándar a las que está X a partir de una media μ .

Si $X \sim N(\mu, \sigma^2)$ y se requiere calcular la probabilidad de que X tome un valor en el intervalo (a, b) , es decir, $P(a < X < b)$ es decir el área bajo la curva normal con media μ y desviación estándar σ que corresponde al intervalo (a, b) . Se tiene que:

$$P(a < X < b) = \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds = P\left(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right).$$

Se denota $\Phi(x)$ a la función de distribución de una variable normal estándar:

$$\Phi(x) = P(Z \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \quad [2.35]$$

y geoméricamente se ve de la siguiente forma:

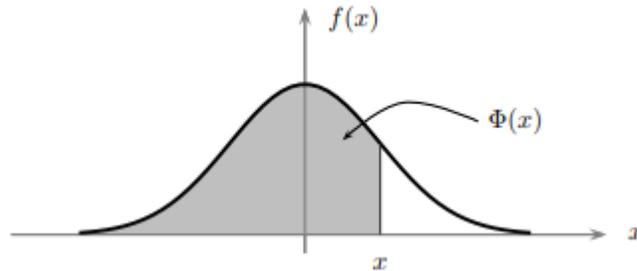


Figura 2.2: Representación geométrica de la función de distribución de una variable normal estándar. FUENTE: Rincón (2007).

Para resolver la integral y aproximar $\Phi(x)$ se usan métodos numéricos. Estos valores se obtienen de las tablas que contienen la probabilidad asociada a cada valor z para distintos valores de x .

Tomando en cuenta la ecuación [2.33] y sea $Z \sim N(0,1)$, se define el número z_α que denota el cuantil al $100 \cdot (1 - \alpha)\%$ de la distribución normal estándar, es decir, $\Phi(z_\alpha) = 1 - \alpha$; con $\alpha \in (0,1)$. Ver ecuación [2.11]

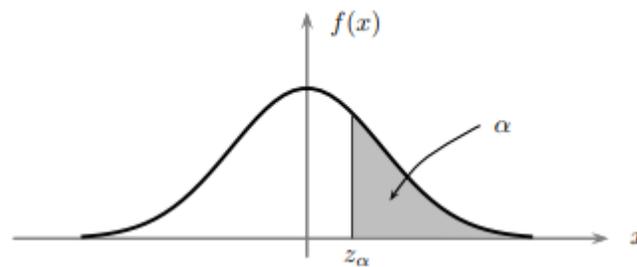


Figura 2.3: Significado gráfico del número z_α . FUENTE: Rincón (2007).

2.4.2. Distribución *t*-Student

Una v.a. continua X tiene una distribución *t*-Student con $n > 0$ grados de libertad con función de densidad:

$$f(x) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)} (1 + x^2/n)^{-(n+1)/2}, x \in \mathbb{R} \quad [2.36]$$

Donde Γ representa la función gamma. Además, tiene las siguientes propiedades:

1. La curva es simétrica y tiene forma de campana

2. La curva es simétrica con respecto a cero, que es la media de la distribución t .
3. La desviación estándar σ de la distribución t es siempre mayor que 1
4. La dispersión exacta, o desviación estándar σ de la distribución t depende del parámetro que se denomina número de grados de libertad
5. A medida que el número de grados de libertad se aproxima a infinito, la desviación estándar σ de la distribución t se aproxima a 1.
6. A medida que el número de grados de libertad se aproxima a infinito, la curva se aproxima a la curva de probabilidad de una distribución normal estándar.

El punto $t_{\gamma}^{(df)}$ se refiere al punto en la escala de la distribución t que tiene df grados de libertad tal que el área bajo la curva a la derecha de este punto es γ . El punto $t_{\gamma}^{(df)}$ se puede determinar mediante la tabla t .

2.4.3. Distribución ji-cuadrada

Una v.a. continua X tiene distribución ji-cuadrada con n grados de libertad y las siguientes propiedades:

1. La curva $f(v)$ de la distribución ji-cuadrada es positiva para todos los valores de $v > 0$.
2. La curva está sesgada en el extremo derecho
3. La forma exacta de la curva depende de un parámetro que se denomina número de grados de libertad y se denota con df . El punto $\chi_{\gamma}^2(df)$ es el punto en la escala de la distribución ji-cuadrada que tiene df grados de libertad tales que el área bajo esta curva a la derecha de este punto es γ .

II.5. Teoremas adicionales

2.5.1. Polinomios cuadráticos de vectores aleatorios normales conjuntos

La ecuación [2.28] presenta expresiones generales de la media y de la varianza de un polinomio lineal. Para obtener la media y la varianza de un polinomio cuadrático no existe una fórmula general, sin embargo, si se toma en cuenta los polinomios cuadráticos de vectores aleatorios conjuntos, existen expresiones para la media, la varianza y los momentos. Esto se lleva a cabo mediante distribuciones ji-cuadrada.

Un polinomio cuadrático de cualquier vector aleatorio normal conjunto puede ser expresado como un polinomio lineal de variables aleatorias independientes normales y ji-cuadradas.

Específicamente, sea $X \sim N_m(\mu, \Sigma)$ con Σ definida positiva. Se define Y como un polinomio cuadrático de X :

$$Y = X^T c X + b X + a \quad [2.37]$$

Sea z una matriz de Cholesky de Σ , además se define u una matriz cuadrada cuyos renglones comprenden eigenvectores¹⁴ ortonormales de $z^T c z$.

Por construcción, u es ortogonal $u^{-1} = u^T$.

Se define el cambio de variables $\dot{X} = uz^{-1}(X - \mu)$, se tiene que

$$Y = \dot{X}^T \dot{c} \dot{X} + \dot{b} \dot{X} + \dot{a} \quad [2.38]$$

En donde $\dot{c} = uz^T c zu^T$

$$\dot{b} = (2\mu^T c + b)zu^T$$

$$\dot{a} = \mu^T c \mu + b\mu + a$$

2.5.2. La ley de los grandes números

Este teorema establece que, bajo ciertas condiciones, el promedio aritmético de variables aleatorias converge a una constante cuando el número de sumandos crece a infinito

Teorema: Sea X_1, X_2, \dots una sucesión infinita de v.a.i.i.d con media finita μ . Entonces cuando $n \rightarrow \infty$, $\frac{1}{n} \sum_{i=1}^n X_i \rightarrow \mu$. En donde la convergencia se verifica en el sentido casi seguro (ley fuerte) y en probabilidad (ley débil).

2.5.3. Teorema central del límite

Una de las aplicaciones de este teorema es simplificar el cálculo de ciertas probabilidades y aproximar algunas distribuciones.

Teorema: Sea X_1, X_2, \dots una sucesión infinita de v.a.i.i.d con media finita μ y varianza finita σ^2 . Entonces la función de distribución de la v.a. Z_n tiende a la función de distribución normal estándar cuando n tiende a infinito, $Z_n = \frac{(X_1 + \dots + X_n) - n\mu}{\sqrt{n\sigma^2}}$

¹⁴ El problema del eigenvalor es saber si para una matriz cuadrada A existen vectores distintos de cero x tales que Ax es un múltiplo escalar de x . Definición: Sea A una matriz de $n \times n$. Un escalar λ se llama eigenvalor de A si existe un vector x distinto de cero tal que $Ax = \lambda x$. Tal vector x se llama eigenvector de A correspondiente a λ .
¿Cómo se encuentran por primera vez los eigenvalores de una matriz dada? La clave es la observación de que λ es un eigenvalor de A si y solo si el espacio nulo de $A - \lambda I$ es no trivial. El teorema fundamental de las matrices invertibles garantiza que una matriz tiene un espacio nulo no trivial si y sólo si no es invertible; es decir, si y sólo si su determinante es cero.
Definición: Los eigenvalores de una matriz cuadrada A son precisamente las soluciones λ de la siguiente ecuación: $\det(A - \lambda I) = 0$

III. Estadística y series de tiempo

Al calcular las medidas de riesgo, es importante entender el comportamiento de las series de tiempo de las cuales se parte, debido a que con lo que se observa se hace inferencia de cómo se cree que se comportara el fenómeno más adelante. Por lo cual, se crean modelos matemáticos, que tratan de ser lo más parecido posible a la serie de tiempo estudiada.

Debido a que las series financieras presentan volatilidad cambiante en el tiempo, los expertos han implementado metodologías que capturen estos cambios y su impacto. En este sentido, intentan estudiar el conjunto de propiedades o “hechos estilizados” que presentan las series financieras y poder ajustar un modelo de series de tiempo para llevar a cabo pronósticos (Cont, 2001).

En el presente capítulo comenzaremos por exponer algunas definiciones asociadas a la modelación de la información de mercado de las variables financieras, tales como: *autocorrelación, proceso estacionario, residuales, proceso ruido blanco, modelos AR, MA, ARMA, ARCH, GARCH, EWMA y estimadores de máxima verosimilitud.*

Se describirán brevemente algunos de los hechos estilizados de las series de tiempo de los rendimientos de las variables financieras y se expondrán los modelos ARCH, GARCH y EWMA. Debido a que este último tema es bastante amplio, solo se presentarán de manera general aquellas definiciones que nos serán de utilidad para comprender y exponer de forma general los modelos para estimar el VaR. En caso de que el lector quiera adentrarse más en este tema se sugiere ver Bollerslev (1986) y Engle (1982).

En el presente capítulo se seguirá un tratamiento de estadística y series de tiempo basado en Bowerman, et al. (2007); Cont (2001); Hernández-Lobato, J. M., Hernández-Lobato, D y Suárez (2007); Holton (2014); Jorion (2009); Montgomery, Jennings y Kulahci (2008); Procesos ARMA CIMAT y Rincón (2006).

III.1. Conceptos de estadística

3.1.1. Estadística descriptiva

Se tiene un conjunto de datos numéricos x_1, x_2, \dots, x_n que representan mediciones de alguna variable de interés. Para conocer las características de esta variable se pueden

calcular las medidas de tendencia central (media, moda y mediana) y las medidas de dispersión (varianza y desviación estándar)

3.1.1.1. Media

La media de los datos x_1, x_2, \dots, x_n es:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad [3.1]$$

3.1.1.2. Varianza

La varianza de la muestra, denotada por s^2 se define como sigue:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad [3.2]$$

La desviación estándar es la raíz cuadrada positiva de s^2 y se denota por s .

3.1.2. Muestras aleatorias

Se considera el conjunto o “población” de n elementos a_1, a_2, \dots, a_n , y se tiene un subconjunto finito y ordenado, será llamado muestra aleatoria. Son el elemento básico para llevar a cabo inferencias estadísticas.

3.1.2.1. Estadísticas

Una estadística es una función cualquiera de una muestra aleatoria (X_1, X_2, \dots, X_n) .

Algunos ejemplos son los siguientes:

La estadística $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ es la media muestral [3.3]

La estadística $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ es la varianza muestral.

3.1.3. Estimación puntual

La estimación puntual consiste en encontrar un valor aproximado a los parámetros con base a los datos de una muestra aleatoria.

3.1.3.1. Estimador puntual

Un estimador puntual es una función de una muestra aleatoria X_1, X_2, \dots, X_n que sirve para aproximar el valor de uno de los parámetros de la distribución.

3.1.4. Pruebas de hipótesis

Un contraste de hipótesis es el procedimiento para determinar si una propiedad que se supone en una población estadística es compatible con lo observado en una muestra de dicha población. Con una muestra representativa de una población se hace inferencia sobre parámetros poblacionales que no son observables. Para establecer si la estimación corresponde al parámetro real (que es desconocido) se hacen supuestos y se lleva a cabo la prueba de hipótesis. Esta prueba evalúa la probabilidad de que la estimación sea igual al parámetro real.

Se puede partir de la idea de que siempre existe una situación predeterminada: es el escenario conservador, el cual sería mejor mantener si no se está lo suficientemente seguro de la suposición y se le conoce como **hipótesis nula** (H_0). Por otro lado, existe el escenario alternativo, el cual, si es aceptado, indicaría que se debería reajustar el proceso. Se define una **prueba de hipótesis** como una regla para decidir si se acepta la H_0 o se rechaza en favor de la hipótesis alternativa H_1 .

3.1.4.1. Errores de prueba de hipótesis

En las pruebas de hipótesis existe la posibilidad de llegar a una conclusión incorrecta. A continuación, se describen los errores a los que se puede llegar al llevar a cabo una prueba de hipótesis (Soporte de Minitab 18).

Se comete un error de tipo I si se rechaza la hipótesis nula cuando es verdadera. Se denota como α a la probabilidad de cometer un error de tipo I, que es el nivel de significancia que se establece para la prueba de hipótesis. Al decir que α vale 0.05 esto indica que se acepta una probabilidad de 5% de estar equivocado al rechazar la hipótesis nula (cuando la hipótesis nula es verdadera).

Un error de tipo II se comete si no se rechaza la hipótesis nula cuando es falsa. La probabilidad de cometer un error de tipo II es β .

En el siguiente cuadro se consideran los errores que se pueden cometer en la prueba de hipótesis:

Verdad acerca de la población		
Decisiones basadas en la muestra	H_0 es cierta: la medida del proceso está en el nivel correcto	H_0 es falsa: la medida del proceso es incorrecta
Rechazar H_0 : Reajustar el proceso	Error tipo I: Reajustar el proceso cuando la medida del proceso está en el nivel correcto. <i>Probabilidad = α</i>	Acción correcta: Reajustar el proceso cuando la medida de este es incorrecta. <i>Probabilidad = $1 - \beta$</i>
No rechazar H_0 : No reajustar el proceso	Acción correcta: No reajustar el proceso cuando la media de este está en el nivel correcto. <i>Probabilidad = $1 - \alpha$</i>	Error tipo II: No reajustar el proceso cuando la media del proceso es incorrecta. <i>Probabilidad = β</i>

3.1.4.2. Prueba de hipótesis acerca de la media de una población t-Student usando puntos de rechazo

Para probar la hipótesis nula $H_0: \mu = c$ contra la hipótesis alternativa $H_a: \mu \neq c$ la condición del punto de rechazo requiere definir $t = \frac{\bar{y}-c}{s/\sqrt{n}}$ que mide la distancia entre \bar{y} (el estimador de μ) y c (el valor que hace que H_0 se cumpla). y encontrar el valor $t_{[a/2]}^{(n-1)}$ en una tabla t .

Para llevar a cabo una prueba de hipótesis se debe conocer la distribución de los datos. Se tiene que \bar{y} y s son la media y la desviación estándar de una muestra de tamaño n que se seleccionó de manera aleatoria de una población cuya media es μ y c es una constante arbitraria.

Se toma el criterio de decisión de rechazar H_0 cuando $|t| > k$ para cierta constante k , haciendo que la probabilidad de un error Tipo I sea igual a α .

El número k se obtiene de una tabla de la distribución t podemos encontrar un valor $t_{[a/2]}^{(n-1)}$ tal que $P(|t| > t_{[a/2]}^{(n-1)}) = a$. Este valor $t_{[a/2]}^{(n-1)}$ es la constante k pues con esto se garantiza que la región de rechazo sea de tamaño a .

Al determinar los puntos de rechazo se tiene que:

Hipótesis alternativas	Condición del punto de rechazo: rechazar H_0 y aceptar H_a si y solo si
$H_a: \mu \neq c$	$ t > t_{[a/2]}^{(n-1)}$
$H_a: \mu > c$	$t > t_{[a]}^{(n-1)}$
$H_a: \mu < c$	$t < -t_{[a]}^{(n-1)}$

Además, cabe destacar que para 30 o más grados de libertad, nos aproximamos a los puntos $|t| > t_{[a/2]}^{(n-1)}$ y $t > t_{[a]}^{(n-1)}$ por medio de los puntos distribuidos normalmente $z_{[a/2]}$ y $z_{[a]}$. Cuando n es por lo menos 30, el teorema del límite central nos indica que la población de todos los valores posibles de $\frac{\bar{y}-c}{s/\sqrt{n}}$ tiene aproximadamente una desviación normal estándar, sin importar que distribución de probabilidad describa a la población que se está muestreando. Por lo tanto, se obtienen las pruebas de muestras grandes de $H_0: \mu = c$ utilizando el valor z .

3.1.4.3. Prueba de hipótesis usando los valores p (p-value)

Otra forma de probar la prueba de hipótesis es calcular la probabilidad de ocurrencia de la media muestral actual o de un valor mayor que este. En otras palabras, si se busca la probabilidad de obtener una media muestral igual o mayor a X , la condición de valor de probabilidad (o bien, el valor p) requiere el cálculo de un área bajo la curva, es decir:

$$P(X \geq \bar{x}) = P\left(\frac{X - \mu}{\frac{\sigma}{\sqrt{n}}} \geq \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(Z \geq \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}\right)$$

Si la población de la cual se extrae la muestra tiene distribución normal con media μ , se puede rechazar H_0 y aceptar H_a haciendo que la probabilidad de un error tipo I sea igual a a , si y solo si *valor* $p < a$.

III.2. Series de tiempo

Al analizar una serie de tiempo, se observa si la serie es estacionaria, es decir, se presente cierta estabilidad estadística en los datos, en otras palabras, que las observaciones tengan media igual a 0¹⁵ y que la varianza y covarianza no dependan del paso del tiempo, es decir, que no estén correlacionadas entre sí. Al ajustar un modelo, para observar lo anteriormente mencionado, se lleva un análisis de los residuales¹⁶ del modelo ajustado. A continuación, se presentan definiciones importantes sobre este tema.

3.2.1. Series de tiempo

Se define una serie de tiempo como una sucesión de N observaciones en función del tiempo ó una sucesión cronológica (Bowerman, et al., 2007).

De manera alterna, se define una serie de tiempo como un proceso estocástico $\{X_t \ t \in \mathbb{R}\}$, es decir una familia de variables aleatorias sobre un espacio de probabilidad (Ω, A, P) (Cervantes, 2014).

3.2.2. Procesos estocásticos

Un proceso estocástico es una sucesión $\{X_t\}$ de variables aleatorias, ordenadas y con características de una unidad observable en diferentes momentos. Se puede suponer que los momentos de las variables aleatorias existen.

Dado $\{X_t\}$ se define la autocovarianza para cualquier retraso k como:

$$\gamma_k = Cov(X_t, X_{t+k}) = E[(X_t - E(X_t)) \cdot (X_{t+k} - E(X_{t+k}))]$$

La secuencia $\gamma_k: k = 0, 1, 2, \dots$ considerada como una función del retardo¹⁷ k se denomina **función de autocovarianza**. La autocovarianza en el retardo $k = 0$ es la varianza de la serie de tiempo.

3.2.3. Procesos estacionarios

De acuerdo con Massó y Pons (2018), las series temporales se pueden clasificar en:

1. Estacionaria: Una serie es estacionaria cuando presenta una estabilidad en los datos a lo largo del tiempo, por ejemplo, cuando la media y varianza son

¹⁵ Además, se debe cumplir que en cualquier combinación dada de valores de x_1, x_2, \dots, x_k , la media de la población del término de error es igual a cero.

¹⁶ Se le llama residual (denotado por ε_t) a la diferencia entre una observación y_t y el valor estimado (\hat{y}_t) donde ε_t se debe comportar como una realización de ruido blanco

¹⁷ Lapso de tiempo k , denota el periodo de tiempo entre los valores de la serie.

constantes en el tiempo. Además, se puede observar en la serie un patrón sistemático que se repite periódicamente.

- 1.1. Estacionaria en media: Cuando la serie temporal no tiene tendencia.
- 1.2. Estacionaria en varianza: Cuando se trata de una serie estable a lo largo del tiempo, sin que se aprecien aumentos o disminuciones de sus valores.
2. No estacionaria: Son series en las cuales la tendencia y/o variabilidad cambian en el tiempo. La serie no oscila alrededor de un valor constante, por ejemplo, los cambios en la media y/o en la varianza presentan una tendencia a crecer o decrecer a largo plazo.
 - 2.1. No estacionarias en media: El cambio en la media implica tendencia (a crecer o decrecer).
 - 2.2. No estacionaria en varianza: El cambio en la varianza implica que la dispersión (variabilidad) no es constante en el tiempo (cuando la varianza se incrementa o disminuye). No se repite periódicamente un patrón.

Para definir estacionariedad (o estacionariedad débil), generalmente se recomienda considerar los primeros dos momentos de la serie de tiempo.

Definición: Proceso estacionario débil

Un proceso $\{X_t\}$, tal que $E(X_t^2) < \infty \forall t \in T$ es un proceso estacionario débil si:

1. $E(X_t) = m$ para todo $t \in T$
2. La función de autocovarianza es solo una función de k y no del tiempo. Esto es:
$$\gamma_k = Cov(X_t, X_{t+k}).$$

Definición: Proceso estacionario estricto

Un proceso $\{X_t\}$, tal que $E(X_t^2) < \infty \forall t \in T$ es un proceso estacionario estricto si $\forall n \in \mathbb{N}, \forall t \in T$ y para todo $\{t_1, t_2, \dots, t_n\} \in T$ las v.a. $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ tienen la misma distribución conjunta que $(X_{t_1+k}, X_{t_2+k}, \dots, X_{t_n+k})$.

3.2.4. Autocorrelación

Los términos de error ordenados en el tiempo pueden estar positivamente correlacionados, negativamente correlacionados o no correlacionados:

- a) Positivamente correlacionados. Si un término de error positivo en el periodo i le sigue otro término de error positivo en algún periodo después (el periodo $i + k$) y si

un término de error negativo en el periodo i le sigue otro término de error negativo en el periodo siguiente. Se presenta si en una gráfica de los residuos en el tiempo se produce un patrón cíclico.

b) Negativamente correlacionados: Si un término de error positivo en el periodo i le sigue un término de error negativo en algún periodo después (el periodo $i + k$) y si un término de error negativo en el periodo i le sigue un término de error positivo en el periodo siguiente. Se presenta si una gráfica de los datos en el tiempo manifiesta un patrón alternante.

c) No correlacionados. Un valor en el período actual no proporciona ninguna información sobre el valor que se observará en el próximo período. Se presenta si los términos de error producen un patrón aleatorio en el tiempo.

Una prueba para evaluar la estacionariedad es observar el comportamiento de la función de autocorrelación, esto se puede llevar a cabo mediante gráficas de la función autocorrelación o correlogramas.

3.2.4.1. Función de autocorrelación

Definición: La **autocorrelación simple de orden k** ($k > 0$) de un proceso X_t estacionario se representa con el símbolo ρ_k y se define como

$$\rho_k = \frac{Cov[Y_t, Y_{t+k}]}{\sqrt{Var[Y_t]}\sqrt{Var[Y_{t+k}]}} = \frac{\gamma_k}{\gamma_0} \quad [3.4]$$

La sucesión $\rho_k: k = 0, 1, 2, \dots$ considerada como una función del retardo k se denomina **función de autocorrelación simple** (ACF, por sus siglas en inglés) del proceso estacionario X_t .

Si se tienen observaciones de una serie de tiempo x_1, \dots, x_n , un posible enfoque para ajustar un modelo es usar la ACF empírica. Para un conjunto de observaciones $\{x_1, \dots, x_n\}$

la función de autocorrelación empírica $\hat{\rho}_k$ se define como $\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}$

La ACF de una serie temporal puede presentar comportamientos diferentes, por ejemplo: la ACF se puede truncar después del retardo k , si no hay espigas en los desfases mayores que k en la ACF; o puede decrecer en forma exponencial, en seno amortiguada o en forma en que domina una de las anteriores o por una combinación.

3.2.4.2. Función de autocorrelación parcial

La ACF da información sobre la estructura de dependencia de un proceso de promedio móvil $MA(q)$, debido a que la función de autocorrelación se anula para retardos mayores que q y como $\theta_q \neq 0$, se tiene que $\rho(q) \neq 0$. Sin embargo, para un proceso autorregresivo $AR(p)$ la ACF se anula y por lo tanto no se tiene la misma información que para un MA . Por esta razón se considera otra función que permite analizar la estructura de dependencia de un proceso AR . Esta función se conoce como la autocorrelación parcial (PACF).

Definición: La función de autocorrelación parcial (PACF) de un proceso estacionario X_t denotado por ϕ_{hh} para $h = 1, 2, \dots$ es

$$\phi_{hh} = Cor(X_{t+h}, X_t) = \rho(h) \quad \text{y} \quad \phi_{hh} = Cor(X_{t+h} - \hat{X}_{t+h}, X_t - \hat{X}_t), h \geq 2$$

Ni $X_{t+h} - \hat{X}_{t+h}$ ni $X_t - \hat{X}_t$ están correlacionados con $\{X_{t+h-1}, X_{t+h-2}, \dots, X_{t+1}\}$

La PACF, ϕ_{hh} es la correlación entre X_{t+h} y X_t cuando se ha eliminado la dependencia de cada una de ellas en $\{X_{t+h-1}, X_{t+h-2}, \dots, X_{t+1}\}$.

3.2.5. Modelos de series de tiempo

En el análisis de series de tiempo se busca ajustar un modelo a las series de tiempo tal que los residuales de las series sea indistinguible de un proceso ruido blanco

3.2.5.1. Proceso ruido blanco

Una serie de tiempo es un proceso de ruido blanco si los errores no presentan correlación serial, tienen varianza constante y el valor esperado de esos errores es cero. Esto quiere decir que los residuales presentan una estructura aleatoria, es decir son independientes e idénticamente distribuidos

Definición: Un proceso de ruido blanco es una sucesión $\{X_n\}, n \geq 1$ de v.a.i.i.d con $E(X_n) = 0, Var(X_n) = \sigma^2$ lo cual suele representarse como $X_n \sim IID(0, \sigma^2)$.

Cuando cada X_n sigue una distribución Normal, $\{X_n\}$ se denomina un proceso de ruido blanco normal o gaussiano.

3.2.5.1.1. Prueba de Autocorrelación

Una prueba que se utiliza para determinar si la serie de tiempo es ruido blanco, es el estadístico Ljung-Box. Si una serie tiene m observaciones, el estadístico Ljung-Box es:

$$Q_{LB} = m \sum_{k=1}^K w_k \eta_k^2 \sim \chi_K^2 \quad [3.5]$$

En donde η_k^2 es la autocorrelación de un retardo de k , K es el número de retardos considerado y $w_k = \frac{m+2}{m-k}$.

Se tiene que: $H_0: \rho_1 = \dots = \rho_K = 0$; $H_1: \rho_i \neq 0$ para algún $i \in \{1, \dots, K\}$. Se rechaza H_0 si $Q_{LB} > \chi_K^2$, donde χ_K^2 indica el $100(1 - \alpha)$ -ésimo percentil de una distribución ji-cuadrada con K grados de libertad¹⁸. Se rechaza H_0 si el p -value de Q_{LB} es menor o igual que el nivel de significancia α .

3.2.5.2. Caminata Aleatoria

La idea de una caminata aleatoria es que el valor futuro de una serie es el valor de hoy más un cambio impredecible. Se considera que un proceso $\{y_t\}$ sigue una caminata aleatoria si:

$$y_t = y_{t_1} + \epsilon_t \quad [3.6]$$

Con $\epsilon_t \sim IID(0, \sigma^2)$. Además, $E[y_t | y_{t_1}, y_{t_2}, \dots] = y_{t_1}$. Si y_t sigue una caminata aleatoria, el mejor pronóstico del valor futuro es el valor de hoy (Massó y Pons, 2018).

3.2.6. Modelos usuales de series de tiempo

A continuación, se describen los modelos usuales de series de tiempo: Modelos de promedios móviles, modelos autorregresivos y modelos autorregresivos de media móvil.

3.2.6.1. Modelos autorregresivos

Un modelo autoregresivo de orden p , denotado por $AR(p)$, es de la forma

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + a_t \quad [3.8]$$

La idea de estos modelos es que los valores actuales de la serie X_t dependen de los p valores previos X_{t-1}, \dots, X_{t-p} . El término "autorregresivo" se refiere al hecho de que este modelo expresa el valor actual X_t de la serie temporal en función de los valores anteriores de la serie temporal $X_{t-1}, X_{t-2}, \dots, X_{t-p}$. Donde X_t es estacionario $\phi_1, \phi_2, \dots, \phi_p$ son

¹⁸ En caso de que se considere un modelo ARMA(p,q), proceso de autocorrelación que se quiere probar, los grados de libertad se determinan como sigue: K-p-q,

parámetros desconocidos que relacionan X_t con $X_{t-1}, X_{t-2}, \dots, X_{t-p}$ y a_t es un ruido blanco con media 0 y varianza σ_a^2 .¹⁹

3.2.6.2. Modelos de promedios móviles

Se denomina modelo de promedios móviles de orden q , $MA(q)$, al modelo

$$X_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad [3.9]$$

A diferencia del modelo $AR(p)$, que es una combinación lineal de observaciones pasadas, el modelo $MA(q)$, es una combinación lineal de términos de error aleatorios previos. Aquí $\theta_1, \theta_2, \dots, \theta_q$ son parámetros desconocidos que relacionan x_t con $a_t, a_{t-1}, a_{t-2}, \dots, a_{t-q}$ es un valor que supuestamente tiene que ser seleccionado en forma aleatoria entre una distribución normal cuya media es cero y su varianza es la misma para cada periodo t . Además, se supone que los errores aleatorios $a_1, a_2, a_3 \dots$ son estadísticamente independientes.

En un momento determinado, en un modelo MA de orden finito, solo un número finito de los errores aleatorios previos contribuye al valor actual de la serie de tiempo y la ventana de las que contribuyen al modelo se mueve en el tiempo, haciendo los errores más antiguos irrelevantes para la siguiente observación.

3.2.6.3. Modelos autorregresivos de media móvil

Hasta ahora, se han descrito dos modelos de series de tiempo, el modelo $AR(p)$, y el modelo $MA(q)$. Los modelos $AR(p)$, intentan explicar el impulso y los efectos de reversión a la media comúnmente observados en los mercados, por ejemplo, efectos de los participantes del mercado. Los modelos $MA(q)$, intentan explicar los efectos de los choques observados en los términos de ruido blanco. Estos choques pueden considerarse como eventos inesperados que afectan el proceso en observación, por ejemplo, ganancias sorpresa, un ataque terrorista, etc.

Al modelar series de tiempo, el modelo autorregresivo de media móvil $ARMA(p, q)$, intenta capturar ambos aspectos, es decir es una fusión de los modelos ya descritos. Se define como sigue:

¹⁹ El proceso X_t tiene media 0. Si queremos analizar un proceso Y_t con media $\mu \neq 0$ podemos considerar el proceso $Y_t - \mu$. Una notación muy útil para considerar este tipo de modelos se basa en el uso del operador de retardo B . El modelo $AR(p)$ es $(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)X_t = w_t$ y se puede escribir de manera resumida como $\phi(B)X_t = w_t$

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad [3.10]$$

Con $\phi_p \neq 0$, $\theta_q \neq 0$ y a_t es un ruido blanco con media 0 y varianza σ_a^2 .

Para una adecuada estimación de los parámetros de un modelo ARMA, es necesario que la serie utilizada para la estimación sea estacionaria en media y varianza^{20 21}

3.2.7. Modelos de series financieras

En un entorno de turbulencia financiera, la volatilidad es considerada el principal componente para el desarrollo de medidas de riesgo, por lo cual los administradores de riesgo están interesados en monitorear la volatilidad diariamente. En los capítulos anteriores se describieron los modelos AR, MA y ARMA para modelar el comportamiento de las series financieras, sin embargo, fallan principalmente al capturar el agrupamiento de volatilidad o efecto de heteroscedasticidad. Para intentar incorporar este efecto a los modelos de series de tiempo y poder estimar la varianza condicional, se desarrollaron los modelos ARCH, EWMA y GARCH, los cuales se definen a continuación.

Se tiene que:

1. El valor de la variable de mercado al final del día i es S_i
2. Se define la variable u_i como el rendimiento compuesto de forma continua durante

el día i , de la forma $u_i = \ln \frac{S_i}{S_{i-1}}$. Donde \bar{u} es la media de u_i y se define como:

$$\bar{u} = \frac{1}{m} \sum_{i=1}^m u_{n-i} \quad [3.11]$$

3. La estimación de σ_n , se lleva a cabo con una aproximación de la volatilidad histórica. Cuando hay m días de observación de u_i , la varianza es σ_n^2 , se tiene que:

$$\sigma_n^2 = \frac{1}{m-1} \sum_{i=1}^m (u_{n-i} - \bar{u})^2 \quad [3.12]$$

²⁰ Se busca que la serie no tenga tendencia y que en cualquier momento de tiempo presente un grado de dispersión similar - homogénea en varianza-

²¹ Para homogeneizar la varianza se aplican logaritmos (no elimina la tendencia) y para que sea estacionaria en media se obtienen diferencias de la serie original objeto de estudio (si elimina la tendencia).

Además, por razones de administración de riesgo²², la fórmula de σ_n^2 se puede ver de la siguiente forma²³:

$$\sigma_n^2 = \frac{1}{m} \sum_{i=1}^m u_{n-i}^2 \quad [3.13]$$

La estimación de volatilidad de la ecuación [3.13] se conoce como un promedio móvil. Este método se basa en incluir información del día anterior y omitir información de $(m + 1)$ días previos. Las ponderaciones de los rendimientos pasados son iguales $\frac{1}{m}$. Por lo cual su mayor desventaja es que no toma en cuenta la ordenación dinámica de las observaciones, es decir, la información reciente y antigua tienen el mismo peso, no dando mayor relevancia a los datos recientes. Además, el desplazamiento de la ventana puede afectar la estimación de la volatilidad debido a que no se podría tomar en cuenta observaciones que podrían ser relevantes.

En consecuencia, se desarrollaron modelos para estimar el nivel de volatilidad que tomen en cuenta el efecto de dar más peso a los datos recientes, los cuales parten de la siguiente ecuación:

$$\sigma_n^2 = \sum_{i=1}^m \alpha_i u_{n-i}^2 \quad [3.14]$$

La variable α_i es el peso que se le proporciona a la observación de i - días anteriores, deben ser positivas. Para dar menor peso a las observaciones antiguas, se escogen $\alpha_i < \alpha_j$ donde $i > j$, entonces la suma de los pesos es 1. A continuación, se presentan extensiones a la fórmula de la ecuación [3.14].

3.2.7.1. Modelos ARCH

En los modelos autorregresivos condicionales heterocedásticos (ARCH por sus siglas en inglés) la estimación de la varianza se basa en m observaciones y en un factor de varianza a largo plazo, al cual se le asigna un peso específico. La varianza evoluciona como una función de las innovaciones de los rendimientos, es decir:

²² Las razones son las siguientes: 1. u_i se define como el cambio porcentual en la variable de mercado entre lo que sucede en el final del día $i - 1$ y al final del día i , $u_i = \frac{S_i - S_{i-1}}{S_{i-1}}$ por lo cual se tiene una diferencia muy pequeña a u_i compuesta 2. Se asume que \bar{u} es cero. La justificación es que el cambio esperado en la variable en un día es muy pequeño cuando se compara con la desviación estándar de los cambios 3. Se reemplaza $m - 1$ por m , se busca un estimador bajo máxima verosimilitud.

²³ En la mayoría de las series financieras, para las estimaciones de la volatilidad, el ignorar los rendimientos esperados constituye una diferencia pequeña

$$\sigma_n^2 = \gamma V_L + \sum_{i=1}^m \alpha_i u_{n-i}^2 \quad [3.15]$$

Donde V_L es la varianza a largo plazo; γ es el peso asignado a V_L . Debido a que los pesos deben sumar 1, se tiene que $\gamma + \sum_{i=1}^m \alpha_i = 1$. Además, con $\omega = \gamma V_L$ la ecuación [3.15] se puede escribir como:

$$\sigma_n^2 = \omega + \sum_{i=1}^m \alpha_i u_{n-i}^2 \quad [3.16]$$

Una de las características de este modelo es que si los valores pasados de los rendimientos hasta el periodo $i - 1$ son grandes, entonces la varianza de la observación en el periodo i condicionada al conjunto de información hasta $i - 1$ también será grande. En otras palabras, entre más fuerte sea la tendencia de los rendimientos cuadrados, mayor será la varianza condicional.

3.2.7.2. Modelos GARCH

De la familia de modelos $GARCH(p, q)$ el más utilizado en la práctica es el modelo $GARCH(1,1)$. En los modelos GARCH, la varianza condicional depende de los cuadrados de las innovaciones y de las varianzas condicionales de periodos anteriores. La ecuación del modelo $GARCH(1,1)$ es la siguiente:

$$\sigma_n^2 = \gamma V_L + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2. \quad \text{Si } \omega = \gamma V_L, \text{ se tiene que:}$$

$$\omega = \gamma V_L + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2 \quad [3.17]$$

Donde, σ_n^2 es la varianza condicional que depende de la última innovación u_{n-1}^2 , σ_{n-1}^2 es la varianza condicional previa; ω es un término determinístico en la ecuación de la varianza que indica que la volatilidad condicional debe converger a un valor positivo.

Para asegurar una volatilidad condicional estrictamente positiva y que sea estacionario en varianza, los parámetros ω , α y β deben ser estrictamente positivos. Al valor obtenido como la suma de $\alpha + \beta$ se denomina persistencia²⁴ de volatilidad y su valor suele estar en torno a uno. Los valores que toma β muestran el efecto que tiene la volatilidad previa sobre la volatilidad del día posterior. Los parámetros del modelo $GARCH(1,1)$ se pueden estimar por el método de máxima verosimilitud.

²⁴ Se denomina persistencia a que tanto tiempo se necesita para que los shocks de volatilidad desaparezcan

3.2.7.2.1. Estimación de los parámetros del modelo *GARCH(1,1)*

Los parámetros de los modelos se estiman con información histórica, una aproximación que se utiliza es el método de máxima verosimilitud. Esto implica escoger valores para los parámetros que maximicen la probabilidad de que los datos ocurran.

En el siguiente ejemplo se utiliza el método de máxima verosimilitud para estimar los parámetros de un modelo *GARCH(1,1)*²⁵. Se define $v_i = \sigma_i^2$ como la volatilidad estimada al día i . Se asume que la distribución de probabilidad de u_i condicionada a la varianza es normal. Los mejores parámetros son los que maximizan:

JHGG

$$\prod_{i=1}^m \frac{1}{\sqrt{2\pi v_i}} \exp\left(\frac{-u_i^2}{2v_i}\right)$$

Si se utilizan las propiedades de los logaritmos, esto es equivalente a maximizar

$$\sum_{i=1}^m \left[-\ln(v_i) - \frac{-u_i^2}{v_i} \right] \quad [3.18]$$

Se llevan a cabo iteraciones para encontrar los parámetros en el modelo que maximicen la expresión en la ecuación [3.18]

3.2.7.2.2. Ventajas del modelo *GARCH(1,1)*

- a) Asume volatilidad cambiante en el tiempo
- b) La varianza tiende a un nivel promedio.
- c) Simplicidad matemática y computacional para estimar los respectivos parámetros.
- d) Captura el fenómeno conocido como volatilidad en aglomeraciones.

3.2.7.2.3. Desventajas del modelo *GARCH(1,1)*

- a) Los parámetros se estiman por máxima verosimilitud, lo cual implica el desarrollo de una optimización numérica. En la práctica se asume que los residuales siguen una distribución normal, por lo cual, el inconveniente de los modelos GARCH es su no-linealidad.

²⁵ Puede ser utilizado otro esquema de volatilidad

b) En los modelos GARCH, los términos en la varianza condicionada se determinan al cuadrado por lo cual, no se puede diferenciar entre los residuales positivos y los negativos. Lo anterior presenta la limitación de explicar los efectos de la asimetría asociados a los movimientos positivos y negativos del mercado.

Las desventajas que presenta el modelo GARCH llevó a que los expertos en la materia desarrollaran extensiones del modelo GARCH para capturar apropiadamente la dinámica del riesgo financiero. Por ejemplo, el modelo EGARCH captura los efectos asimétricos de la volatilidad y el modelo de dos componentes o CGARCH que captura el fenómeno de memoria larga en la volatilidad. Otros modelos son: GJR-GARCH, FIGARCH. Para más detalle de este tema, se recomienda ver Hans y Van Dijk (2000).

3.2.7.3. Modelo EWMA

Este modelo también es un caso particular de la ecuación [3.14]. En el modelo EWMA los pesos decrecen exponencialmente a medida que se retrocede al pasado. Particularmente, $\alpha_{i+1} = \lambda \alpha_i$, donde λ es un número entre 0 y 1. Además, es un caso particular del modelo $GARCH(1,1)$, en donde $\gamma = 0, \alpha = 1 - \lambda$ y $\beta = \lambda$. El sistema Riskmetrics, utiliza el enfoque EWMA para modelar el riesgo, con $\lambda = .94$.

La fórmula del modelo EWMA lleva a cabo una actualización de volatilidades, es decir, si se desea estimar σ_n , la volatilidad para el día n (calculada al final del día $n - 1$), esta se calcula de la suma de σ_{n-1} (la estimación de la volatilidad para el día $n - 1$ que se calculó el día $n - 2$) y u_{n-1} (el cambio en porcentaje diario más reciente), lo cual se puede ver de la siguiente forma:

$$\sigma_n^2 = \lambda \sigma_{n-1}^2 + (1 - \lambda) u_{n-1}^2 \quad [3.19]$$

La principal ventaja de este modelo es que está diseñado para poder observar los cambios en la volatilidad. Además, tiene solo un parámetro para estimar, lo cual es más robusto para el error de estimación que otros modelos. También es un modelo recursivo y la historia se resume en un número, σ_{n-1}^2 .

3.2.8. Modelando correlaciones

La precisión de la medición del VaR depende del pronóstico de volatilidad y de las correlaciones. Por ejemplo, en periodos de turbulencia global, si las correlaciones

incrementan al mismo tiempo que una volatilidad creciente, se anulan las propiedades de diversificación de portafolios. Por lo cual el VaR basado en datos históricos subestima el riesgo y las correlaciones, lo cual podría llevar a obtener rendimientos muy alejados del rango de pronóstico. Algunos de los modelos para capturar la variación de la correlación en el tiempo son: promedios móviles, GARCH y EWMA.

Asimismo, Dr. (2002) describe que, si se reemplaza el concepto de covarianza o correlación que generalmente se usa para describir la dependencia entre variables, se pueden usar cópulas para describir la estructura de dependencia entre dos o más variables. Este tema puede llegar a ser exhaustivo y no es el objetivo de esta tesina por lo cual únicamente se hace énfasis en que la correlación es de fundamental importancia para el VaR.

Además, es relevante destacar que los datos históricos utilizados para el pronóstico de la volatilidad y de las correlaciones pueden no proporcionar los mejores pronósticos disponibles de los riesgos futuros, excluyendo situaciones como cambios en los regímenes cambiarios²⁶. En su caso se pueden utilizar parámetros implícitos de las opciones que incluyen los últimos datos del mercado (por ejemplo, volatilidad implícita). Sin embargo, una desventaja de esta información es la disponibilidad de la información, debido a que no existen opciones de todos los precios financieros

3.2.9. Análisis estadístico de Series de Tiempo

Al llevar a cabo el análisis estadístico de las series de tiempo, se podrían seguir los siguientes pasos:

PASO 1: Descargar los datos y graficarlos: comprobar si los datos presentan tendencia, patrones cíclicos, variaciones estacionales o fluctuaciones irregulares²⁷. Se busca determinar si es razonable asumir un modelo estacionario²⁸.

²⁶ Cuando se trabaja en series temporales es habitual encontrarse con cambios institucionales o de contexto socioeconómico que cambian la estructura interna de funcionamiento de fenómenos.

²⁷ Se puede llevar a cabo la descomposición ETS (Error, trend, Seasonality, por sus siglas en inglés) para verificar la estacionalidad en los datos. Si la estacionalidad es diminuta en comparación con el componente de tendencia, no se considera. Una función que se puede utilizar en Python es `seasonal_decompose()`.

²⁸ Se puede aplicar la prueba de Dickey Fuller (para confirmar estacionariedad en los datos, H_0 : la serie de tiempo tiene raíz unitaria, por lo cual es no estacionaria. Tiene alguna estructura dependiente del tiempo). Una función que se puede utilizar en Python es `adf_fuller()`

PASO 2: Llevar a cabo correlogramas y pruebas formales a los datos brutos, valores absolutos (o cuadrados) y probar si es ruido blanco.

PASO 3: Ajuste de modelo. Identificar el modelo de series de tiempo adecuado para los datos, por ejemplo $ARMA(p, q)$ para datos estacionarios, $ARIMA(p, d, q)$ ²⁹ para datos no estacionarios con componente de tendencia o un modelo $SARIMA(p, d, q, P, D, Q)$ ³⁰ en caso de que los datos presenten estacionalidad significativa. Además, identificar el orden del modelo seleccionado. Estos coeficientes pueden ser determinados con las gráficas ACF (determina el orden del modelo MA) y PACF (determina el orden del modelo AR). Frecuentemente es difícil determinar p y q utilizando esta técnica, por lo cual, puede ser mejor utilizar métricas como el criterio de Akaike (AIC, por sus siglas en inglés) o el criterio de información Bayesiano BIC (BIC, por sus siglas en inglés)³¹.

PASO 4: Llevar a cabo una revisión del modelo seleccionado para poder observar si el modelo sobre ajusta los datos. Esto se lleva a cabo mediante una separación del conjunto de datos: se determina un conjunto de entrenamiento y un conjunto de prueba. Se ajustará el modelo seleccionado sobre el conjunto de entrenamiento

PASO 5: Llevar a cabo pruebas del modelo ajustado sobre el conjunto de pruebas: predecir los valores de prueba, y se lleva a cabo el análisis residual se comparan los valores obtenidos con los valores actuales del conjunto de pruebas. Se pueden observar los correlogramas de ε_t y se puede utilizar una prueba formal para revisar si es un proceso ruido blanco (Ljung box), de otra forma no se ha encontrado un modelo adecuado. Si los cuadrados no son un proceso ruido blanco, tal vez sea necesario aplicar un modelo GARCH.

PASO 6: Ajustar el modelo seleccionado sobre todo el conjunto de datos

PASO 7: Se sabe X_0, \dots, X_t . Se desea predecir X_{t+k}

²⁹ $ARIMA(p, d, q)$ ó Modelo Autorregresivo-Integrado de Medias Móviles de orden p, d, q , que es un modelo $ARMA(p, q)$ aplicado a una serie integrada de orden d : ha sido necesario diferenciar d veces para eliminar la tendencia.

³⁰ $SARIMA$ ó Modelo estacional autorregresivo integrado de media móvil

³¹ Son modelos de selección de modelos que llevan a cabo diferentes combinaciones de valores p, q y comparan su desempeño, generalmente se utiliza el Criterio de Akaike

3.2.10. Análisis estadístico de Series de Tiempo aplicado a Series financieras

3.2.10.1. Rendimientos financieros

A lo largo de este capítulo se desarrollará un ejemplo (*Ejemplo 3.1*) en el lenguaje de programación Python y se irán aplicando las definiciones que se van describiendo. (ver código utilizado en el Apéndice 1). Se descarga la serie de tiempo en forma de un Data Reader con la librería pandas. Para este ejemplo se utilizó la serie histórica de precios diarios obtenida de Yahoo Finance de Alsea, S.A.B. de C.V. con código accionario (ticker: ALSEA.MX). Los datos abarcan el periodo del 3 de enero del 2012 al 31 de diciembre del 2018.

Sea el precio de un activo financiero denotado como P_i en el tiempo i . El rendimiento neto del activo entre el periodo $i - 1$ y i , se denota como u_i y se define como:

$$u_i = \frac{P_i}{P_{i-1}} - 1 \quad [3.20]$$

Los rendimientos compuestos continuamente asumen implícitamente que las ganancias temporales son reinvertidas de forma continua. Se define el rendimiento compuesto continuo o “el logaritmo de los rendimientos” entre el periodo $i - 1$ y i como sigue:

$$r_i = \ln(1 + u_i) = \ln \frac{P_t}{P_{t-1}} = \ln(P_t) - \ln(P_{t-1}) \quad [3.21]$$

En la Figura 3.1, (gráfica superior) se presenta la serie de tiempo de los precios diarios³². Se puede apreciar que hay una tendencia general creciente con épocas de descenso. Se observan caídas pronunciadas, sin embargo, las variaciones relativas son difíciles de apreciar. Para apreciar mejor los cambios relativos, se utiliza la serie de rendimientos.

En la gráfica de rendimientos (Figura 3.1, gráfica inferior) se muestra que la mayor caída relativa ocurre a finales del 2016. El valor medio de la serie parece ser estable, con oscilaciones alrededor de este valor medio que en ocasiones es pronunciada y no es uniforme. Los periodos de esta volatilidad tienen a agruparse, por lo cual se considera que presenta características particulares de las series financieras.

³² En la presente tesina se utilizan los precios ajustados al cierre debido a que ya contemplan los ajustes de desdoblamiento (Se hace referencia al término en inglés “splits” de acciones) y dividendos.

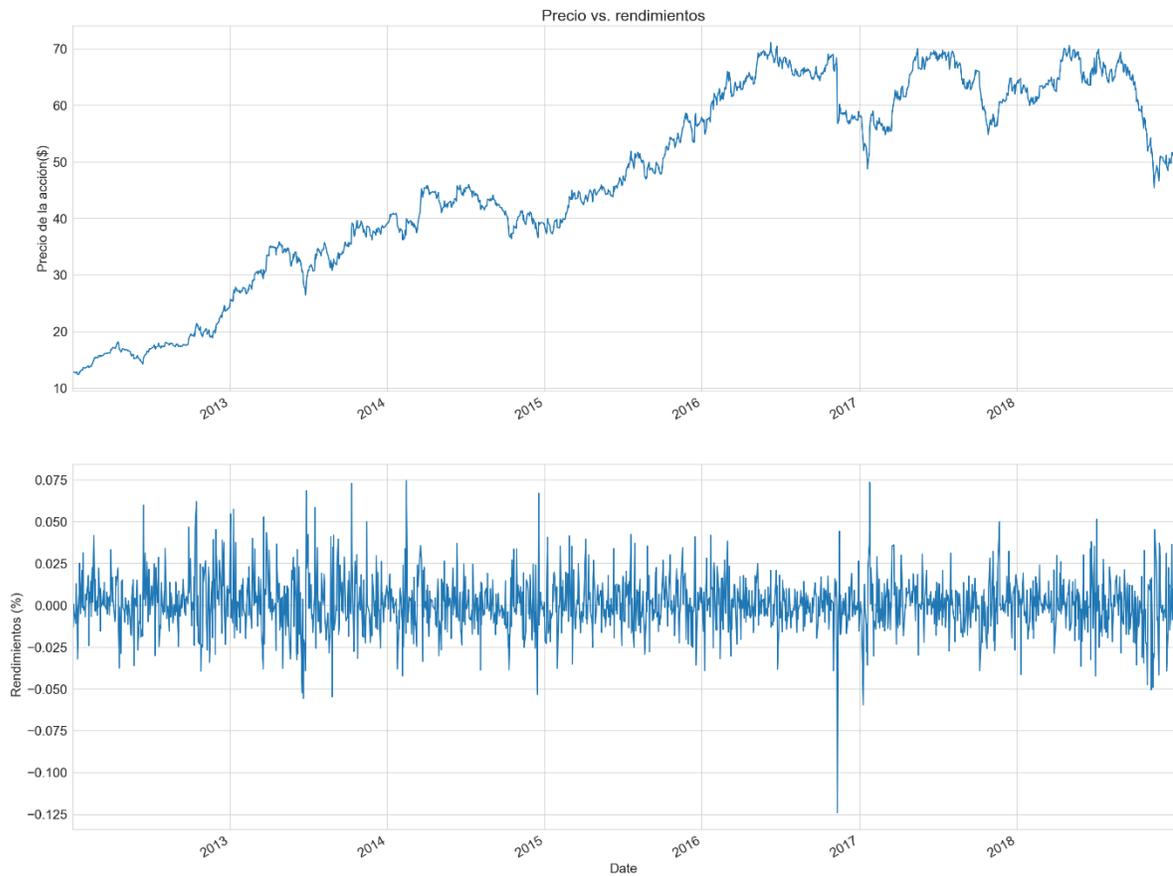


Figura 3.1: Precios diarios (gráfica superior) y el logaritmo de los rendimientos de la acción de ALSEA.MX (gráfica inferior) del periodo que abarca del 3 de enero de 2012 al 31 de diciembre del 2018. FUENTE: Estimación propia.

3.2.10.2. Distribución condicional y no condicional

Para este apartado, cabe aclarar que el análisis de series de tiempo es diferente de la estimación estadística. Los estimadores estadísticos se aplican a realizaciones de muestras. Las series de tiempo se aplican a una parte del proceso estocástico. Debido a que los X términos de un proceso estocástico no necesariamente son independientes, se debe distinguir entre los términos “distribución condicional” y “distribución no condicional”.

Los modelos con distribución de rendimientos no condicional asumen que los rendimientos son independientes y que el proceso generado es lineal, con parámetros que son independientes de las realizaciones pasadas. Un ejemplo de la distribución no condicional es asumir que los rendimientos logarítmicos siguen una distribución normal no condicional. Por otra parte, la distribución condicional (distribuciones dependientes del tiempo) surge de la evidencia de refutar la suposición i.i.d. Los modelos GARCH son un

ejemplo de esta categoría, en donde la volatilidad es cambiante en el tiempo (Penza y Bansak, 2001):

3.2.10.3. Propiedades de los rendimientos financieros

Al construir modelos, es importante tener en cuenta los “hechos estilizados” que comúnmente presentan las series de tiempo de rendimientos de los activos financieros. Los modelos que se construirán representarán la dinámica de los precios de los activos por lo cual deben capturar o replicar algunas o todas de estas propiedades. A continuación, se presentan algunas de las más importantes:

1. Colas pesadas: La distribución de rendimientos presenta colas pesadas y alta curtosis en comparación a una distribución normal.

2. Autocorrelación en los rendimientos: La autocorrelación mide el grado de similitud entre una serie de tiempo y la misma serie desfasada en intervalos sucesivos de tiempo. Contar con una muestra autocorrelacionada positivamente (por mínima que esta sea) puede generar estimaciones no fiables si los datos presentan colas pesadas

3. Agrupamiento de volatilidad: las series de tiempo de los rendimientos de los activos presentan el efecto de heteroscedasticidad, es decir, la volatilidad o la desviación estándar de los rendimientos presentan una estructura dependiente en el tiempo. Además, la volatilidad tiende a agruparse en grupos, es decir, a altas variaciones del precio (ya sean positivos o negativos) es probable que le sigan variaciones del precio que serán altas (Esto se puede observar en un correlograma de los rendimientos cuadrados absolutos)

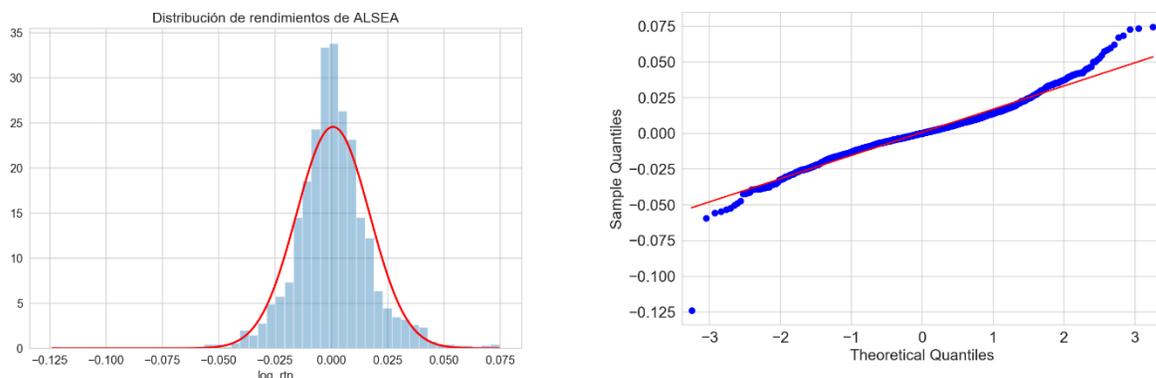


Figura 3.2: Comportamiento de la serie de datos. FUENTE: Estimación propia.

En la Figura 3.2 del lado izquierdo se compara la distribución normal y la distribución empírica. En relación con la distribución normal, la distribución empírica presenta más observaciones en el centro y en las colas. Se puede observar que la distribución empírica de los rendimientos presenta curtosis alta. Del lado derecho se presenta un gráfico normal cuantil-cuantil de los rendimientos diarios de la acción de Alsea. Se puede apreciar que las colas de la densidad empírica son más pesadas que las de la distribución normal³³.

Para detallar el punto 2 y 3, en la Figura 3.3 se presentan las gráficas de autocorrelación de los logaritmos de los rendimientos de la acción de Alsea, también de los rendimientos al cuadrado y de los valores absolutos (Figura 3.4). El área rectangular en azul fuerte indica el intervalo de confianza del 95%, en donde los puntos fuera de esta región son estadísticamente significativos. Se puede ver que, para los rendimientos logarítmicos (Figura 3.3), hay pocos puntos significativos y se observa que existe autocorrelación. Por otra parte, para ejemplificar el punto 3, en la Figura 3.4, se observa que el decrecimiento es mayor en los rendimientos al cuadrado en comparación a los valores absolutos de los rendimientos.

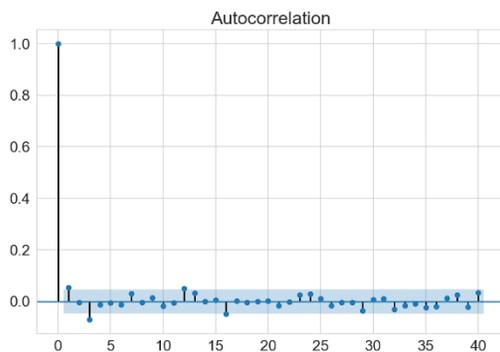


Figura 3.3: Función de autocorrelación de los rendimientos
FUENTE: Estimación propia.

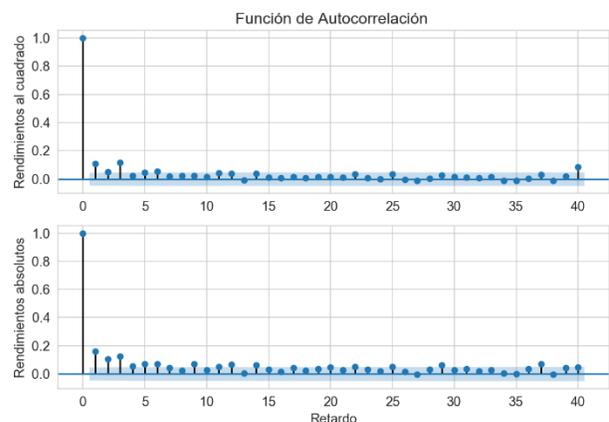


Figura 3.4: Función de autocorrelación: de los rendimientos al cuadrado y rendimientos absolutos. FUENTE: Estimación propia.

Para resolver el punto de la autocorrelación, es importante modelar dos momentos: la media condicional y la varianza condicional. La media condicional se modela utilizando los modelos ARIMA, propuestos por Box y Jenkins (1976). Sin embargo, estos modelos no son adecuados para explicar el hecho de que la varianza condicional cambie en el

³³ Además, se llevó a cabo una prueba formal conocida como prueba de Jarque-Bera (prueba de bondad de ajuste para comprobar si una muestra de datos tiene la asimetría y la curtosis de una distribución normal), cuyo p value fue menor a .05 por lo cual se rechazó la hipótesis nula de que los rendimientos siguen una distribución normal, para funciones utilizadas en Python, ver código en el Apéndice 1

tiempo, además debido al efecto de volatilidad agrupada se hace uso del modelo GARCH. En este sentido, a continuación, se lleva a cabo el ejercicio de utilizar el lenguaje de programación R y Python para obtener los parámetros del modelo GARCH aplicados a la serie de rendimientos de la acción de Asea. En la siguiente tabla se muestran los resultados obtenidos (ver código utilizado en el Apéndice 2):

Estimación de los parámetros del modelo <i>GARCH</i>(1, 1)			
	MS Excel con procedimiento de acuerdo con Hull (2007)	R	Python
ω	0.0000328303	0.00002148 (3.469e-06)	0.000026863 (7.352e-07)
α	0.121921	0.1028 (9.593E-03)	0.1016 (2.728e-02)
β	0.759574	0.8201 (1.804E-03)	0.79986 (2.112e-02)

Tabla 3.1: Se presenta un comparativo de los valores de los parámetros obtenidos con R y Python. En paréntesis se presentan los errores estándar. FUENTE: Estimación propia.

La relevancia de presentar la tabla 3.1 es mostrar que hacer uso de diferentes lenguajes de programación puede llevar a obtener diferentes resultados. Por lo cual, se debe tener cuidado con lo que se asume en las funciones de las paqueterías de cada lenguaje de programación debido a que puede afectar a la efectividad de los resultados. Entre los resultados de R y Python no parece ver una diferencia muy grande, sin embargo, al compararlos con el método que presenta Hull (2007), se pueden observar diferencias y esto se debe a lo que se considera para obtener el primer valor de la varianza, en Hull (2007) es la primera innovación al cuadrado, mientras que en R es la varianza de los rendimientos de todo el periodo.

Una vez que se tienen los resultados del modelo, se analizan los residuales. Por ejemplo, se observan los valores p (p-value) de la prueba Ljung Box para poder saber si el modelo sugerido remueve la autocorrelación. Una vez que se elige el modelo se puede estimar el riesgo dentro y fuera de la muestra.

IV. Valor en Riesgo

En los capítulos anteriores se describieron algunas definiciones para el entendimiento de los modelos de estimación del VaR. En el presente capítulo se presentan las características, ventajas y desventajas de algunos modelos VaR y además se mencionan algunas pruebas de validación de los modelos VaR. Entre los modelos que se exponen se encuentran los modelos paramétricos, el modelo de Simulación Histórica y el modelo Monte Carlo. Debido a que es un tema bastante amplio y complejo, solo se exponen de manera general los supuestos y fundamentos en los que se basan los modelos. En caso de que el lector quiera adentrarse más en este tema se sugiere ver Aloui y Mabrouk (2010); Angelidis y Degiannakis (2005); Campbell (2005); Christoffersen et al. (2001); Chen y Chen (2013); Corbetta y Peri (2018); Hammoudeh, et al. (2011); Holton (2014); Hung, et al. (2007); J.P. Morgan (1996); Kupiec (1995); Nilsson (2017); Pritsker (2001) y Zhang y Nadarajah (2017).

En el presente capítulo se seguirá un tratamiento basado en Holton (2014, 2017); Hull (2007); Jorion (2007, 2009); Malik (2019a); Malik (2019b) y Rozga y Arneric (2009).

IV.1. Introducción al VaR

La exposición al riesgo de mercado puede causar pérdidas, afectando a instituciones financieras, empresas no financieras y administradores de activos. Las pérdidas pueden ocurrir por la combinación de dos factores: la exposición que se tenga sobre la fuente de riesgo (representan las decisiones de los *traders* ó de los administradores de portafolios) y la volatilidad de los factores de riesgo (distribución de los factores de riesgo). Las empresas no tienen control sobre la volatilidad de las variables financieras, sin embargo, pueden ajustar su exposición a estos riesgos, por ejemplo, haciendo uso de los derivados. El VaR captura el efecto combinado de la volatilidad y la exposición o posición abierta.

Los modelos de exposición se clasifican en modelos de valuación local y global. Los primeros, valúan el portafolio a un tiempo determinado y hacen uso de derivadas para inferir posibles movimientos, por ejemplo, modelos lineales como aproximaciones delta-

normal y los modelos no lineales, como la aproximación delta- gamma³⁴. Por su parte, el modelo global, mide el riesgo revaluando el portafolio en un rango de escenarios, por ejemplo, el modelo de Simulación Histórica y Monte Carlo ³⁵. Por otra parte, las distribuciones de los factores de riesgo se pueden modelar utilizando aproximaciones paramétricas (por ejemplo, utilizando una distribución normal) ó aproximaciones basadas en datos históricos o llevando a cabo simulaciones Monte Carlo.

El VaR se puede estimar calculando el cuantil de la distribución de pérdida del portafolio. Si se tiene que p es el valor actual del portafolio y P es el valor futuro del portafolio al final del horizonte de tiempo, la pérdida del portafolio es³⁶:

$$L = p - P \quad [4.1]$$

Sin embargo, no se sabe el valor futuro del portafolio P , por lo tanto, no se sabe el valor de la pérdida L . Las dos son variables aleatorias y se les puede asignar distribuciones de probabilidad. Para estimar el VaR se asigna una distribución a P y/o a L y se calcula el cuantil de la pérdida. Comúnmente, las medidas VaR trabajan directamente con la distribución de P y la utilizan para inferir el cuantil de L . Por lo cual, el reto de cualquier medida VaR es construir la distribución de P .

4.1.1. Definición de VaR

De acuerdo con Jorion (2009), el VaR resume la pérdida máxima esperada en un horizonte de tiempo objetivo y con un nivel de confianza dado. Jorion (2007) puntualiza que el VaR sobre un horizonte de tiempo determinado es el cuantil de la distribución de pérdidas y ganancias.

³⁴ Enfoque de aproximación no lineal que implica sensibilidades del factor de riesgo de orden superior, tal como gamma.

³⁵ La medición de la exposición lineal a los movimientos en las variables subyacentes de riesgo se presenta en diferentes mercados. En el mercado de renta fija, la exposición a los movimientos de las tasas de interés se le conoce como duración. En el mercado accionario, esta exposición se denomina riesgo sistemático o beta. En los mercados de derivados, la exposición a los movimientos en el valor del activo subyacente se conoce como delta. Las segundas derivadas o mediciones de segundo orden, que miden la exposición de segundo orden o cuadrática a la variable financiera se conocen como convexidad en el mercado de renta fija y gamma en el mercado de derivados

En donde, delta se define como la tasa de cambio que presenta el valor del portafolio con respecto a la variable de mercado subyacente; Gamma se define como la tasa de cambio que presenta la delta con respecto a la variable de mercado; La convexidad mide el cambio en delta ante cambios del precio del subyacente

³⁶ Si $p < P$, la pérdida será negativa lo cual significa que se tiene una ganancia en el portafolio. Si $P < p$, la pérdida es positiva.

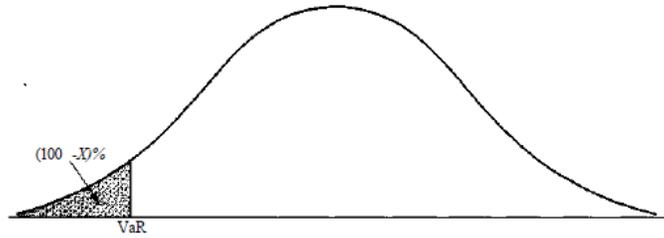


Figura 4.1. Representación gráfica del VaR en la distribución de probabilidad de los cambios en el portafolio. El nivel de confianza es X%. Fuente: Hull (2007).

Un ejemplo es el siguiente: el VaR a un día de un portafolio es de \$30,000,000 pesos mexicanos, a un horizonte de tiempo de 100 días y con un nivel de confianza del 95%. Esto significa que se pueden tener pérdidas superiores al VaR cinco veces cada 100 días ($5\% \cdot 100 = 5$). Asimismo, las pérdidas en el portafolio serán inferiores al VaR 95 veces de los 100 días.

El primer paso para la medición del VaR es la elección de tres factores cuantitativos:

1. El horizonte de tiempo. Para cuantificar el VaR se requiere definir un periodo sobre el cual medirlo. Este periodo puede corresponder, por ejemplo, para un administrador de inversiones al periodo de regular de presentación de informes, mientras que para un administrador bancario puede ser lo suficientemente largo para detectar a los operadores que toman posiciones que exceden sus límites. Puede ocurrir que se tenga que llevar a cabo una transformación para tener horizontes de tiempo comparables (debido a la forma en cómo se presenta la información histórica y el horizonte de inversión)³⁷. Se tiene que:

$$\begin{aligned} \mu &= \mu_{anual} * T & [4.2] \\ \sigma &= \sigma_{anual} * \sqrt{T} \end{aligned}$$

Donde T es el número de periodos por año³⁸.

2. El nivel de confianza. Su elección se basa en objetivos. El nivel de confianza se puede determinar llevando a cabo la validación del modelo (*backtesting*) en el que se estima el número de excepciones³⁹ que se experimentan en un periodo de tiempo, lo cual refleja

³⁷ Esta transformación sugiere el supuesto de que los rendimientos no están correlacionados en intervalos sucesivos de tiempo. Este supuesto es consistente con los mercados eficientes, en donde el precio actual incluye toda la información acerca de un activo en particular. Todos los cambios en el precio son originados por noticias que no pueden ser anticipados y por lo tanto no deben estar correlacionados en el tiempo, es decir los precios siguen una caminata aleatoria.

³⁸ En horizontes largos, la media dominará a la volatilidad debido a que la volatilidad crece con la raíz cuadrada del tiempo y la media con el tiempo. En horizontes corto (por ejemplo, un día), la volatilidad domina por lo cual las mediciones de riesgo se basan solo en la volatilidad sin tomar en cuenta los rendimientos esperados.

³⁹ Una excepción ocurre cuando el VaR que se estimó no es capaz de cubrir la pérdida observada

que tan confiable es el modelo y las suposiciones que se han tomado en cuenta al determinar los parámetros

3. Definir la moneda base. Se refiere a la moneda en la que se denomina la posible pérdida.

4.1.2 Enfoques del VaR

Para este apartado, se define W_0 como la inversión inicial y R como su tasa de rendimiento. El rendimiento esperado y la volatilidad de R son μ y σ . El valor del portafolio al final del horizonte objetivo es W . El número W^* es denominado el cuantil muestral de la distribución.

4.1.2.1. El VaR para distribuciones generales

En su forma más general, el VaR puede derivarse de la distribución de probabilidad del valor futuro del portafolio $f(w)$.

En un nivel de confianza dado, c , se busca encontrar:

- la peor realización posible W^* tal que la probabilidad de exceder dicho valor sea c :

$$c = \int_{W^*}^{\infty} f(w)dw \quad [4.3]$$

- Ó tal que la probabilidad de un valor inferior a W^* , $p = P(w \leq W^*)$, sea $1 - c$:

$$1 - c = \int_{-\infty}^{W^*} f(w)dw = P(w \leq W^*) = p \quad [4.4]$$

En la ecuación [4.4], el área $-\infty$ a W^* debe sumar $p = 1 - c$.

$$VaR = E[W] - W^* \text{ (medido en relación de la media)} \quad [4.5]$$

$$VaR = W_0 - W^* \text{ (en términos de la pérdida absoluta en dólares)}$$

Este enfoque no utiliza la desviación estándar para encontrar el VaR. Por ejemplo, si se supone que los ingresos diarios están i.i.d., se puede estimar el VAR a un nivel de confianza del 95% a partir del 5% del lado izquierdo de “la cola de perdida” del histograma, tomando en cuenta el número total de observaciones.

4.1.2.2. El VaR para distribuciones paramétricas

Si los rendimientos siguen una distribución normal, el VaR puede calcularse directamente de la desviación estándar del portafolio utilizando un factor multiplicativo que depende del nivel de confianza.

Esto se debe a que la distribución general $f(w)$ se traduce en una distribución normal estándar $\Phi(\epsilon)$ donde ϵ tiene media cero y desviación estándar 1. Tomando en cuenta la ecuación [2.33], se tiene que:

$$1 - c = \int_{-\infty}^{W^*} f(w)dw = \int_{-\infty}^{-z_c} \Phi(\epsilon)d\epsilon \quad [4.6]$$

Encontrar el VaR es equivalente a encontrar la desviación z_c tal que el área a su izquierda sea igual a $1 - c$. Esto se lleva a cabo tomando los datos de las tablas de la función de distribución normal estándar acumulada, que muestran el área a la izquierda de una variable normal estándar. Ver ecuación [2.11]

Si el rendimiento más bajo es: $R^* = \mu + z_c\sigma$ y se asume que los parámetros μ y σ están expresados en una base anual, el intervalo de tiempo considerado es T , en años, sustituyendo en la ecuación [4.5], se tiene que:

$$VaR = E[W] - W^* = W_0 * z_c * \sigma * \sqrt{T} \text{ (medido en relación de la media)}$$

$$VaR = -W_0 * R^* = W_0 * (z_c * \sigma * \sqrt{T} - \mu * T) \text{ (en términos de la pérdida absoluta en dólares)}^{40}$$

Una forma de denotar lo anterior, utilizando la ecuación [2.33], tomando en cuenta la ecuación [2.28] y $Z \sim N(0,1)$, para estimar el VaR medido como un α -cuantil de pérdida, se tiene que:

$$\begin{aligned} \Phi_L^{-1}(\alpha) &= p - \Phi_P^{-1}(1 - \alpha) \\ &= p - [E(P) + \Phi_Z^{-1}(1 - \alpha) * \sigma_P] \\ &= p - E(P) + \Phi_Z^{-1}(\alpha) * \sigma_P \end{aligned} \quad [4.7]$$

Si $p = E(P)$, se tiene que:

$$\Phi_L^{-1}(\alpha) = \Phi_Z^{-1}(\alpha) * \sigma_P^{41} \quad [4.8]$$

4.1.3. Esquema general del VaR

Las ideas en las que se basa el esquema general del VaR son:

- El riesgo se compone de dos elementos: incertidumbre y exposición. Además, el VaR combina los dos componentes del riesgo (incertidumbre- con la distribución

⁴⁰ Muchas veces se asume que μ es cero, como se comentó anteriormente

⁴¹ Si el periodo al que se estima el VaR es corto (a un día o a una semana), se podría asumir que $E(P) = p$

condicional de \mathbf{R} construida por medio del procedimiento de inferencia- y exposición- caracterizado con la función de asignación θ del portafolio, construido a través del procedimiento de asignación, con $P = \theta(R)$, la cual puede ser una asignación primaria o una reasignación).

- Las variables financieras que ocupan las medidas de VaR son:

1. Valores de los activos S : Sea S un vector que contiene los m valores de los activos, en donde m es el número total de activos que serán modelados. Estos valores representan el valor acumulado al tiempo 1 de los activos individuales que son parte del portafolio en el tiempo 0. Los activos individuales pueden ser acciones, bonos, futuros, opciones u otros instrumentos. Además, su valor actual se representa como s y es el conjunto ordenado del valor de los activos actuales.

2. Factores de riesgo R : Sea R un vector que se compone por n factores de riesgo los cuales representan valores al tiempo 1 de las variables financieras que pueden ser utilizadas para determinar el valor de los activos. Dependiendo de la composición del portafolio, los factores de riesgo pueden representar tipos de cambio, tasas de interés, precios de *commodities*, *spreads*, volatilidades implícitas, etc. Los valores del vector requieren información pasada del vector de factor de riesgos, que en conjunto con la información actual se conoce como información de mercado histórica.

Se sabe que las cantidades P , S y R son variables aleatorias. El valor del portafolio P es una función de los valores S de los activos que lo componen. Además, los valores S están en función de los factores de riesgo R . Esto quiere decir que P es una función θ de R . Un ejemplo que describe y que hace muy claro el uso de estas v.a. es el siguiente: el valor de un portafolio de bonos del tesoro es P y está en función de los valores S de los bonos individuales que lo conforman. Además, los valores están en función de las tasas de interés aplicables R .

Las medidas VaR aplican análisis de series de tiempo a los datos históricos para construir una distribución de probabilidad conjunta de R . Por otra parte, determinan la relación θ que se lleva a cabo entre P y R para convertir esa distribución conjunta en una distribución para P . Finalmente, de la distribución de P se calcula el VaR

Una vez descritas las ideas principales, a continuación, se presenta un resumen del esquema general para obtener el VaR:

1. Los dos tipos de insumos utilizados son:

1.1. Datos históricos para R .

1.2. El número de unidades de cada instrumento que se tienen en el portafolio

2. Los dos tipos de insumos son tratados por separado por dos diferentes procedimientos:

2.1. El proceso de inferencia aplica modelos de series de tiempo a la información de mercado histórica para construir una distribución conjunta de probabilidad para R .

2.2. El procedimiento de asignación utiliza el vector de número de posiciones de cada instrumento para construir una función tal que se tenga una distribución para P .

3. Finalmente, se lleva a cabo un procedimiento de transformación⁴²

Ejemplo 4.1.(Tipo de cambio) Se tiene un portafolio de acciones australianas. Se desea obtener el VaR a un día en libras esterlinas (GBP) a un nivel de confianza del 95%. Se define lo siguiente:

1.El vector S , que contiene el valor acumulado (que reflejan los rendimientos de los precios) de las acciones en GBP.

2.El número de acciones por emisora que se tiene en el portafolio. Representado como un vector renglón y denotado por ω

3. Para calcular el valor futuro del portafolio P , el cual es un polinomio lineal de S , $P = \omega S$ se tiene que llevar a cabo la matriz de covarianza de S en libras, sin embargo, la información histórica de los precios se presenta en dólares australianos. Por lo cual, S estará en función de R , donde R es el valor acumulado de los rendimientos de los precios de las acciones en dólares australianos y además al vector R se le incluirá el componente de tipo de cambio libra/dólar australiano.

6.Se construye la función $\theta = \omega \circ \varphi$ la cual relaciona P con R y se tiene un polinomio cuadrático.

⁴² Los procedimientos de transformación se pueden clasificar en dos grupos: a) transformaciones lineales o cuadráticas; b) Transformaciones numéricas

7. Se construye la matriz de covarianza condicional de R .

Se tiene la matriz de covarianza condicional de R sin embargo se tiene que P , el valor futuro del portafolio se expresa como un polinomio cuadrático de R , por lo cual, no se puede aplicar la fórmula de la ecuación [2.28] para obtener la desviación estándar condicional de P . Se debe buscar un método apropiado para encontrar la desviación estándar condicional.

La asignación de los portafolios no lineales son un reto para estimar el VaR. Algunas de las soluciones son las siguientes:

1. Utilizar el modelo Monte Carlo para aproximar el cuantil deseado

2. Llevar a cabo una Reasignación: Aproximar el polinomio cuadrático θ a un polinomio lineal basados en el gradiente de θ (y poder obtener la desviación estándar de P con la fórmula de un polinomio lineal). Ver ecuación [2.28].

2.1. Se escoge un punto para obtener el gradiente, se escoge el valor esperado de R .

3. Asumir que R está condicionada a una distribución normal conjunta y aplicar técnicas probabilísticas apropiadas para polinomios cuadrados de vectores aleatorios que siguen una distribución normal conjunta.

3.1. Cualquier polinomio cuadrático de un vector aleatorio normal conjunto con valores reales se puede expresar como un polinomio lineal de variables aleatorias independientes normales y ji-cuadrada.

3.2. Para determinar el cuantil de la pérdida del portafolio se pueden utilizar dos aproximaciones:

3.2.1. Utilizando la expansión Cornish-Fisher⁴³

3.2.2. Invertir la función característica de P usando integración numérica.

Mas adelante se desarrollarán la mayoría de los puntos mencionados.

⁴³ La expansión Cornish-Fisher puede ser utilizada para estimar los cuantiles de una distribución de probabilidad, utilizando los momentos

4.1.4. Desarrollo del Esquema general del VaR

Siguiendo con el esquema anterior, a continuación, se detallan los procedimientos para modelar los componentes de riesgo, determinar un proceso de transformación y estimar el VaR.

4.1.4.1. Procedimiento de Inferencia

El componente de incertidumbre ocupa datos históricos para diseñar procedimientos de inferencia y determinar una distribución para R condicionados a la información disponible al tiempo cero. Holton (2014) describe que esto se puede llevar a cabo de tres formas:

1. Especificar una distribución condicional
2. Asumir un vector de medias condicionado que dependa de lo que cada factor de riesgo representa individualmente, de la misma forma para la matriz de covarianzas condicionada. Por ejemplo, si un factor de riesgo son los rendimientos, se asume que $\mu = 0$.

3. Una manera de estimar la matriz de covarianzas condicionada es tratar los datos como una realización de una muestra y aplicar modelos como EWMA. El detalle de algunos de estos procedimientos se presenta en el capítulo 3 de la presente tesina.

Además, para llevar a cabo la construcción de algunas de las tres caracterizaciones se debe tener en cuenta que:

- Se debe cuidar la calidad de la información histórica del factor de riesgo.
- Si se van a aplicar técnicas de análisis de series de tiempo, es conveniente que se apliquen a un proceso ruido blanco.

4.1.4.2. Procedimiento de asignación

Por otra parte, el componente de exposición se representa con un proceso de asignación al portafolio, buscando simplificar el portafolio. Se construye un proceso de asignación especificando, por ejemplo, una asignación primaria $P = \theta(R)$. Se busca una función de asignación θ y un vector de factores de riesgo R que pueda reflejar como el valor de mercado del portafolio dependerá de las variables del mercado.

La función θ puede llegar a ser compleja. Para ciertas medidas de VaR, solo se utilizan asignaciones primarias como una herramienta práctica y aplican directamente un procedimiento de transformación. Otras medidas VaR aproximan las asignaciones

primarias bajo un proceso de reasignación. En este apartado se describen todos estos conceptos de simplificación.

4.1.4.2.1. Asignaciones primarias

La construcción de asignaciones primarias se lleva a cabo de la siguiente forma:

1. Se define $P = \omega S$ en donde ω es un vector fila y S es un vector de activos (Si S es el vector de factores de riesgo, entonces $P = \omega S$ es la asignación primaria)

2. Si se selecciona otro vector de factores de riesgo R , el procedimiento de asignación utiliza un conjunto de funciones que valúan cada activo S en términos de R . Utilizando fórmulas de valuación φ para cada activo S , se especifica una asignación $S = \varphi(R)$. La función de asignación φ es un vector de funciones de componente φ_i , cada una valúa algún activo: $S_i = \varphi_i(R)$. La composición $\theta = \omega \circ \varphi$ define una asignación primaria $P = \theta(R)$

Ejemplo 4.2: Desarrollando el ejemplo de las acciones australianas (Ejemplo 4.1) en donde se busca el VaR a un día en GBP con un 95% de confianza. Un *trader* tiene posiciones en acciones de tres diferentes empresas (empresa A, empresa B y empresa C)

1. Se tiene el valor actual del portafolio en libras esterlinas
2. Se construye una asignación primaria

$$P = \theta(R) \tag{4.9}$$

3. Se especifica un vector S en donde el valor acumulado puede reflejar cambios en los precios, dividendos y cambios en el tipo de cambio desde el tiempo 0. Llevando a cabo un vector de componentes se tiene que:

$$S = \begin{pmatrix} S_1^{GBP} \\ S_1^{GBP} \\ S_1^{GBP} \end{pmatrix} \sim \begin{pmatrix} \text{valor acumulado en GBP de una acción de la empresa A} \\ \text{valor acumulado en GBP de una acción de la empresa B} \\ \text{valor acumulado en GBP de una acción de la empresa C} \end{pmatrix}$$

4. Se define a P como el vector producto $P = \omega S$, en donde ω representa el número de acciones del portafolio donde X es número de acciones de la empresa A, Y es número de acciones de la empresa B, Z es número de acciones de la empresa C (en posición corta). Por lo que se tiene que:

$$\omega = (X \quad Y \quad -Z)$$

5. Debido a que los componentes de S están denominados en GBP y la información histórica en dólares australianos (AUD), se lleva a cabo un cambio de variable. Para

expresar S en términos de R , es decir, $S = \phi(R) = R_4 \begin{pmatrix} R_1 \\ R_2 \\ R_3 \end{pmatrix}$

Donde

$$R = \begin{pmatrix} R_1^{AUD} \\ R_2^{AUD} \\ R_3^{AUD} \\ R_4^{FX} \end{pmatrix} \sim \begin{pmatrix} \text{valor acumulado en AUD de una acción de la empresa A} \\ \text{valor acumulado en AUD de una acción de la empresa B} \\ \text{valor acumulado en AUD de una acción de la empresa V} \\ \text{tipo de cambio spot GBP/AUD} \end{pmatrix}$$

6. Llevando a cabo la composición de ω con ϕ se tiene la función $\theta = \omega \circ \phi$ que relaciona P con R :

$$P = \theta(R) = R_4^{FX} (XR_1^{AUD} + YR_2^{AUD} - ZR_3^{AUD})$$

7. La función compuesta ω, ϕ y θ tiene una asignación del portafolio que es cuadrática.

4.1.4.2.1.1. Reasignaciones de portafolio

Si el portafolio está conformado por instrumentos complejos como derivados o títulos inmobiliarios, una asignación primaria podría resultar en una valuación computacionalmente cara. Muchos procedimientos de asignación reemplazan una asignación primaria θ con una aproximación simple, conocida como reasignación.

Las razones principales por las cuales se lleva a cabo la reasignación son las siguientes:

1. Facilitar un procedimiento de transformación:

Si una función de asignación primaria θ puede ser aproximada por un polinomio $\tilde{\theta}$ lineal o cuadrático, esto facilitará una transformación lineal o cuadrática. Si se tiene que utilizar una transformación numérica, se puede reemplazar una función de asignación primaria θ con otra función $\tilde{\theta}$ que sea más fácil de valorar, reducirá el tiempo requerido para llevar a cabo todas las valuaciones que las transformaciones numéricas requieren. Se tienen reasignaciones de interpolación, mínimos cuadrados, gradiente y gradiente Hessiano.

2. Facilitar un proceso de inferencia: se introduce un vector de factores de riesgo alternativo en caso de que no haya información histórica disponible del seleccionado o se requiera reemplazar un vector con gran cantidad de componentes por uno con menor número de componentes.

3. Facilitar otra reasignación: se reasignan el número de activos que se consideran en el portafolio, se buscan activos con características similares y se agrupan para reducir el número de activos que deben ser modelados. Un ejemplo es utilizar el análisis de componentes principales.

4.1.4.2.1.2. Reasignaciones lineales con aproximaciones gradiente

Las reasignaciones lineales se construyen con aproximaciones gradiente. Una aproximación gradiente puede ser construida sobre cualquier punto. Sin embargo, la esperanza condicional $\mu = E(R)$ es una opción razonable.

La reasignación tiene la forma

$$\tilde{\theta}(R) = \theta(\mu) + [\nabla\theta(\mu)]'(R - \mu) \quad [4.10]$$

Debido a que las fórmulas de fijación de precios para instrumentos financieros son diferenciables, es viable valorar analíticamente el gradiente. De otra forma, tiene que ser valuado con diferencias finitas. Esto requerirá $n + 1$ valuaciones de θ .

Ejemplo 4.3. Función de reasignación: Reasignación lineal- Aproximaciones con gradiente.

Siguiendo con el ejemplo de las acciones australianas **Ejemplo 4.1** en donde se tiene una asignación primaria cuadrática, se busca aproximar θ con un polinomio lineal $\tilde{\theta}$ basado en el gradiente de θ . Esto lleva a cabo de la siguiente forma:

1. El punto escogido para obtener el gradiente es $E(R)$, la esperanza de R , condicionada a la información disponible en el tiempo 0.

2. Al aplicar la ecuación [4.10] se obtiene una reasignación lineal $\tilde{P} = \tilde{\theta}(R)$.

3. Como $\tilde{P} = \tilde{\theta}(R)$ es un polinomio lineal, se puede aplicar la ecuación [2.28] para obtener la desviación estándar condicional de P .

4. Si se asume que \tilde{P} es distribución normal, se obtienen los insumos para obtener el VaR⁴⁴. Se aplica la fórmula de la ecuación [4.8].

4.1.4.2.1.3. Reasignaciones cuadráticas con aproximaciones gradiente

Hessiano

Una reasignación cuadrática puede ser construida con una aproximación gradiente Hessiano. Construida sobre el punto $\mu = E(R)$, la reasignación tiene la forma

$$\tilde{\theta}(R) = \theta(\mu) + [\nabla\theta(\mu)]'(R - \mu) + \frac{1}{2}(R - \mu)'[\nabla\theta(\mu)][\nabla\theta(\mu)] \quad [4.11]$$

Las reasignaciones cuadráticas generalmente no se construyen de esta manera. Esto se debe a que las aproximaciones son locales debido a que se basan en la primera y segunda derivada. Para construir una aproximación cuadrática que sea viable para una región amplia de valores de R , se puede aplicar el método ordinario de interpolación o el método de mínimos cuadrados.

4.1.4.2.1.4. Interpolación y el método de mínimos cuadrados

A continuación, se describe brevemente la metodología para ajustar una reasignación cuadrática aplicando el método de mínimos cuadrados.

1. Se selecciona un conjunto de realizaciones $r^{[1]}, r^{[2]}, \dots, r^{[l]}$ de R
2. Se llevan a cabo valuaciones de la función θ a cada realización, es decir, $p^{[k]} = \theta(r^{[k]})$
3. Se interpola o se aplica mínimos cuadrados a los puntos $(r^{[k]}, p^{[k]})$.

Además, se deben tener en cuenta las siguientes consideraciones:

- Cuando R tenga un gran número de componentes, si es posible, buscar una reasignación que simplifique θ . Ver apartado 4.1.4.2.1.1. punto 3.
- Las realizaciones r para la interpolación o mínimos cuadrados deben ser seleccionadas con cuidado debido a que de esto dependerán los resultados.

⁴⁴ También se puede evaluar que tanto se aproxima \tilde{P} a P , sin embargo, el desarrollo de este tema no va en el alcance de este proyecto de tesina

- Una aproximación para especificar realizaciones es escoger a μ como un punto para seleccionar otras realizaciones que estén dispersas de alguna manera sobre un elipsoide.

Ejemplo 4.4. Función de reasignación: Reasignación Cuadrática: Interpolación y método de mínimos cuadrados

Se busca el VaR de un portafolio de un comerciante de cocoa que tiene 39 opciones call pactadas a 1,050 en el futuro. Los contratos se pactan a 1,077 USD/ton. Cada futuro es por 10 toneladas, se tiene que:

1. La asignación primaria es

$$P = \theta(R) = 390B^{call}(R_1, 1050, R_2, R_3, \frac{28}{365})$$

Donde $B^{call}()$ es la fórmula de fijación de precios de una opción y

$$R = \begin{pmatrix} R_1 \\ R_2 \\ R_3 \end{pmatrix} \sim \begin{pmatrix} \text{precio futuro} \\ \text{volatilidad implícita} \\ \text{Libor a 1 - mes} \end{pmatrix}$$

2. Utilizando información histórica del mercado, se asume que $R \sim N_n(\mu, \Sigma)$ y se tiene los datos de μ, Σ
3. Se utiliza el método de mínimos cuadrados ordinarios con 10 realizaciones $r^{[k]}$ para construir una reasignación cuadrática de la forma

$$\theta(R) = c_{1,1}R_1^2 + c_{2,2}R_2^2 + b_1R_1 + b_2R_2 + b_3R_3 + a \quad [4.12]$$

4. La realización $r^{[10]}$ se iguala a μ . Para obtener las realizaciones restantes, se generan puntos pseudoaleatorios en la esfera unidad y se cambian iterativamente hasta cumplir con el principio de energía mínima. Los puntos deben estar proyectados en el elipsoide. Con esto se obtienen las 9 realizaciones $r^{[k]}$ y los valores $p^{[k]}, p^{[k]} = \theta(r^{[k]})$
5. Se ajustan los puntos $(r^{[k]}, p^{[k]})$ al polinomio de la ecuación, utilizando el método de mínimos cuadrados ordinarios para obtener los coeficientes $c_{1,1}, b_1$ y a , se obtiene \tilde{P} . Ver ecuación [1.12].

4.1.5. Procedimiento de Transformación

De acuerdo con Holton (2014), el objetivo principal de un proceso de transformación es construir una distribución para P , en donde $P = \theta(R)$ y con esto calcular el VaR del portafolio. Esta caracterización se lleva a cabo combinando los dos componentes de riesgo: incertidumbre, mediante la distribución condicional de R y exposición, mediante la función de asignación del portafolio, en algunas ocasiones denotado por θ .

Los tres tipos básicos de procedimientos de transformación son los Procedimientos de transformación lineales, Procedimientos de transformación cuadráticos y procedimiento de Transformación Monte Carlo.

4.1.5.1. Procedimiento de Transformación lineal

Estos procedimientos aplican si la función θ del portafolio es un polinomio lineal. Emplean una fórmula estándar para calcular la varianza del polinomio lineal del vector aleatorio. Algunos ejemplos cuyas asignaciones primarias pueden ser polinomios lineales son las acciones y futuros.

Ejemplo 4.5. Transformación lineal En este ejemplo se desea calcular el VaR a una semana bajo un 90% de confianza de un portafolio con posiciones en dólares americanos en diversas industrias de metales. Se define:

1. El vector S que refleja el cambio de los precios de los n -metales.
2. El vector s , que representa los precios actuales en USD/ton de los n –metales
3. El vector representa el total de toneladas que se tiene por cada metal
4. Se determina el valor futuro del portafolio, $P = \omega S$.
5. Partiendo de lo anterior, P es un polinomio lineal de S .
6. Con lo anterior, se puede aplicar la ecuación [2.28] para obtener la matriz de covarianza de S y determinar la desviación estándar de P . Esto se lleva a cabo utilizando información histórica.
7. Si se asume que P sigue una distribución normal y debido a que se tiene la media (valor del portafolio actual) y la desviación estándar, se puede especificar una función de distribución acumulada para P y obtener el cuantil (.1).

8. Se determina el cuantil (.9) de la pérdida del portafolio como la resta entre el precio actual del portafolio y el cuantil (.1) de P . Aplicar la ecuación [4.8].

4.1.5.2. Procedimiento de Transformación cuadrático

Se considera un portafolio (p, P) con una asignación cuadrática

$$P = \theta(R) = R'cR + bR + a \quad [4.13]$$

En donde $R \sim N_n(\mu, \Sigma)$, c es una matriz $n \times n$, b es un vector renglón n -dimensional y a es un escalar. Se asume que Σ es definida positiva

De acuerdo con la ecuación [2.37], P se puede expresar como un polinomio lineal de variables aleatorias independientes normales y ji cuadrada. Tomando en cuenta esta representación se puede calcular la desviación estándar

Sin embargo, para obtener el VaR es más complicado, por lo cual se utiliza la expansión de Cornish-Fisher.

El procedimiento de transformación cuadrático aplica si la asignación θ del portafolio es un polinomio cuadrático y la distribución conjunta de R es normal conjunta. Esta transformación tiene el potencial de calcular el VaR de portafolios no lineales en un tiempo real inmediato. Sin embargo, la aplicación más importante es aplicarlos para facilitar la reducción de la varianza en las transformaciones Monte Carlo.

Ejemplo 4.6. Derivados de platino. Se busca la desviación estándar del portafolio en yen japonés (JPY) y el VaR en JPY a 1 día a un 90% de nivel de confianza.

1. El tiempo se mide en días de negociación. Una firma japonesa de metales presenta una exposición de forwards y opciones de platino. Algunas de las posiciones son denominadas en USD. Las tasas de interés son constantes.
2. Se modelan 3 factores de riesgo:

$$R = \begin{pmatrix} R_1 \\ R_2 \\ R_3 \end{pmatrix} \sim \begin{pmatrix} \text{Precio del platino spot en JPY} \\ \text{Volatilidad implícita representativa del platino} \\ \text{tipo de cambio JPY/USD} \end{pmatrix}$$

3. Se asume que $R \sim N_3(\mu, \Sigma)$. Se tienen los datos de c, a, b, μ y Σ

4. Se construye una asignación primaria del portafolio $P = \theta(R)$ basadas en fórmulas de fijación de precios de forwards y opciones. Se hace una reasignación cuadrática siguiendo la ecuación [4.13].
5. Se construye la matriz de Cholesky z de Σ . Se construye la matriz u con renglones iguales a eigenvectores ortonormales de $z'cz$. Se define un cambio de variables. Se tiene que $\tilde{P} = \hat{R}'c\hat{R} + b\hat{R} + a$ y se obtienen los valores de \hat{c}, \hat{b} y \hat{a} . Ver ecuación [2.38].
7. Para obtener la desviación estándar condicional de \tilde{P} se calculan los momentos centrales y condicionales de \tilde{P} y los momentos centrales y acumulativos⁴⁵ de la normalización \tilde{P}^* de \tilde{P} . Se utiliza el primer y el segundo momento para calcular la desviación estándar condicional de \tilde{P} .
8. Para obtener el valor en riesgo se utiliza la expansión Cornish Fisher para encontrar el cuantil (.1) de \tilde{P}^* y poder aplicar la ecuación [4.8].

4.1.5.3. Procedimiento de Transformación Monte Carlo

El procedimiento de Transformación Monte Carlo es el siguiente:

Se considera un portafolio (p, P) con una asignación $P = \theta(R)$. Si se tiene una muestra aleatoria para R , se define una muestra aleatoria para P con $P^{[k]} = \theta(R^{[k]})$. Se aplica un estimador adecuado a P para estimar el VaR. Para aplicar el estimador, se genera una secuencia de números pseudo⁴⁶ aleatorios los cuales pueden ser contruidos a partir de diferentes algoritmos.

El error estándar del análisis Monte Carlo dependerá de: el tamaño de la muestra (m), la métrica de riesgo que se está estimando, la distribución condicional de P que depende de la distribución condicional de R y de la composición del portafolio. Algunas de las técnicas de reducción de varianza⁴⁷ viables para el VaR es el método de variables de control, el muestreo estratificado para estimar la desviación estándar de pérdida y para calcular el VaR y la valuación selectiva de las realizaciones.

⁴⁵ Los acumulativos de una distribución de probabilidad son un conjunto de cantidades y funcionan como una alternativa a los momentos de una distribución

⁴⁶ Una secuencia de números "pseudo" aleatorios hace referencia a que la generación de estos datos pasa ciertas pruebas empíricas consideradas importantes por el usuario.

⁴⁷ En la presente tesina no se profundiza en el tema de las técnicas de reducción de varianza.

Una subcategoría de esta transformación son los procedimientos de transformación histórica los cuales no generan aleatoriamente los escenarios, en cambio se utilizan los datos históricos del vector R .

Ejemplo 4.7. Siguiendo con el **Ejemplo 4.1**, a continuación, se presenta el desarrollo para obtener el VaR utilizando la transformación de Monte Carlo:

1. Se asume que R sigue una distribución normal conjunta con media y matriz de covarianza dada
2. Se generan 10,000 observaciones aleatorias $r^{[1]}, r^{[2]}, \dots, r^{[10,000]}$ de R .
3. Se construyen 10,000 realizaciones $p^{[1]}, p^{[2]}, \dots, p^{[10,000]}$ de P , en donde $p^{[k]} = \theta(r^{[k]})$
4. Se puede aproximar cualquier parámetro de P con la muestra de las realizaciones generada en el punto 3.
5. Se obtiene el cuantil muestral (.05) de las realizaciones $p^{[k]}$ de P y se utiliza como una aproximación para el cuantil de P .
6. Para calcular el VaR a un día, se calcula el cuantil (.95) de la pérdida del portafolio como la resta del valor del portafolio actual y el cuantil (.5) de P , se aplica la ecuación [4.8].

IV.2. Aproximaciones VaR

Holton (2014) menciona que los modelos que usualmente se utilizan para calcular el VaR, hacen referencia a los procedimientos de transformación. Por ejemplo, los términos "modelo paramétrico" o "modelo de varianza-covarianza" se refieren a medidas VaR que emplean un procedimiento de transformación lineal. La "aproximación delta-gamma" usa un procedimiento de transformación cuadrática. El "modelo Monte Carlo" y el "modelo de simulación histórica" se refieren a las medidas VaR que utilizan procedimientos de transformación Monte Carlo o histórica. A continuación, se detallan las características de estos modelos.

4.2.1. Simulación Histórica

El modelo de simulación histórica estima la distribución del cambio del valor del portafolio basado en un número finito de datos históricos, es decir, utiliza la información histórica como una guía para proyectar el futuro. La metodología para estimar el VaR mediante el modelo de simulación histórica es la siguiente:

1. Contar con datos históricos de las variables del mercado que afecten al portafolio
2. Generar un vector de los cambios de un día a otro de las variables del mercado que afecten al portafolio
3. Se forma un vector de precios futuros multiplicando lo que se generó en el punto 2 y el precio actual de la variable de mercado (última observación).
4. Se calcula el vector de pérdidas y ganancias, con la diferencia entre el valor del portafolio futuro y el valor actual del portafolio
5. Se calcula el percentil de la distribución de probabilidad del cambio del valor futuro del portafolio, ordenando el vector de pérdidas y ganancias en forma ascendente. Se escoge la observación que corresponda a $[(1 - c) * K]$, en donde c =nivel de confianza y K es el número total de observaciones. Por ejemplo, para una distribución de pérdidas y ganancias de 10,000 observaciones, el VaR a un nivel de confianza del 99% una vez ordenado en forma ascendente, está determinado por la entrada $(10,000 * .01) = 100$ del vector de pérdidas y ganancias del portafolio.⁴⁸

En resumen, el modelo de SH busca lograr un pronóstico correcto de los cuantiles de la distribución empírica de los rendimientos del portafolio lo cual dependerá de la construcción de la base histórica y del tamaño de la ventana de observación.

4.2.1.1. Ventajas del modelo Simulación Histórica

- a) El uso de la distribución empírica captura el efecto de colas pesadas en el valor futuro del portafolio, lo cual se dará siempre y cuando hayan existido eventos previos similares en la historia de los factores de riesgo.
- b) Su implementación es relativamente simple si se tiene la información histórica de los factores de riesgo adecuada.

⁴⁸ Para el ejemplo, cada día se puede actualizar la estimación del VaR utilizando las 10,000 observaciones más recientes

4.2.1.2. Desventajas del modelo Simulación Histórica

- a) Supone que la historia se va a repetir, lo cual no suele suceder en los mercados financieros. Algunos eventos que sucedieron en el pasado y no necesariamente ocurrirán en el futuro aportarán información y probablemente sesgarán el cálculo del VaR.
- b) Los resultados dependerán en gran medida al conjunto de datos y al tamaño de la muestra.

4.2.1.3. Extensiones del modelo Simulación Histórica

El modelo básico de simulación histórica asume que todas las observaciones pasadas tienen el mismo peso sobre los resultados. Sin embargo, se busca que las observaciones recientes tengan mayor importancia y que reflejen efectos de un periodo actual.

4.2.1.3.1. Aproximación Histórica por Bootstrap

Es una variación del modelo básico de simulación histórica. La idea principal de este modelo es el muestreo con reemplazo que se aplica: Se hace un muestreo de los rendimientos y se calculan los precios futuros y el valor del portafolio. Esto se lleva a cabo por el número del horizonte de tiempo que se busca obtener el VaR (por ejemplo, si es un VaR a 10 días, se hacen diez subconjuntos de muestreos y se obtienen 10 VaR). Para determinar el VaR, se ordenan los resultados (valores del portafolio obtenidos) de menor a mayor. Se obtiene el VaR tomando en cuenta el menor valor del portafolio obtenido y el valor del portafolio actual.

4.2.1.3.2. Ponderación de observaciones

Este modelo aplica un peso a los rendimientos, tomando en cuenta el retraso en días en los que ocurrió determinado rendimiento respecto a n que es el número total de observaciones (la última observación). En primer lugar, se acomoda el vector de rendimientos de peor a mejor. Al comenzar a ponderar, la suma de todos los pesos debe sumar 1, por lo cual el peso que se le asignara a un rendimiento que está entre el día $n - i$ y $n - i + 1$, se define como sigue: $\frac{\lambda^{i-1}(1-\lambda)}{1-\lambda^n}$. El VaR se obtiene cuando se alcance el cuantil requerido de la columna de pesos acumulados. El valor del parámetro λ se puede obtener de llevar a cabo un backtesting.

4.2.1.3.3. Ponderación de volatilidad

En este modelo se incorpora un esquema de actualización de la volatilidad en el modelo de simulación histórica. Se utilizan los modelos GARCH o EWMA para estimar la volatilidad. Por lo cual, el rendimiento del activo se ajusta en función de su volatilidad actual y pronosticada.

4.2.1.4. Simulación Histórica Filtrada

El modelo de Simulación Histórica Filtrada (SHF) utiliza la técnica de *bootstrapping*⁴⁹ y las observaciones son ajustadas por información reciente de volatilidades, lo cual permite estimar un VaR que es sensible a las condiciones cambiantes del mercado. Una limitante para este modelo puede ser la disponibilidad de datos históricos y que los resultados dependerán en gran medida al conjunto de datos seleccionado: Además, debido al ajuste por la ponderación de volatilidades, el VaR puede ser alto o demasiado bajo.

4.2.2. Aproximación paramétrica

En los modelos paramétricos los factores de riesgo asumen una distribución de probabilidad específica. En este sentido, la precisión de la estimación del VaR dependerá de que tanto la distribución elegida describa el comportamiento real del portafolio.

4.2.2.1. Delta Normal

Un portafolio se compone de posiciones de mercado sobre cierto número de factores de riesgo. Si se determina la descomposición del portafolio, el rendimiento del portafolio es una combinación lineal de los rendimientos de los activos subyacentes. Por lo tanto, el VaR de un portafolio puede reconstruirse a partir de una combinación de los riesgos de los valores subyacentes.

Esta aproximación se define de la siguiente forma:

Se tiene un portafolio con un valor de P dólares, que consiste de n activos en el que se invierte un monto α_i en el activo i ($1 \leq i \leq n$) al inicio del periodo. Se define Δx_i como el rendimiento diario del activo i . Por lo cual, $\alpha_i \Delta x_i$ es el cambio en el valor de la inversión en dólares del activo i en un día y ΔP es el cambio en el valor total del portafolio en dólares en un día, denotado como sigue:

⁴⁹ Técnica de aleatorización no paramétrica que se infiere a partir de la distribución de los datos observados, para modelar la distribución de un estadístico de interés.

$$\Delta P: \sum_{i=1}^n \alpha_i \Delta x_i \quad [4.14]$$

Para llevar a cabo la medición del VaR, se calcula la media y la desviación estándar de ΔP . Para calcular la media, se asume que el valor esperado para cada Δx_i es cero, lo cual implica que el cambio esperado en el portafolio es cero⁵⁰. Por otra parte, para calcular la desviación estándar de ΔP , se define la volatilidad diaria del i -ésimo activo, denotada por σ_i y ρ_{ij} la correlación del rendimiento del i -ésimo activo con el rendimiento del j -ésimo activo. Es decir, σ_i es la desviación estándar de Δx_i y ρ_{ij} la correlación entre Δx_i y Δx_j .

Por lo tanto, la varianza del portafolio es igual a:

$$\sigma_p^2 = \sum_{i=1}^n \sum_{j=1}^n \rho_{ij} \alpha_i \alpha_j \sigma_i \sigma_j \quad [4.15]$$

La ecuación [4.15] se puede ver de la siguiente forma:

$$\sigma_p^2 = \sum_{i=1}^n \alpha_i^2 \sigma_i^2 + 2 \sum_{i=1}^n \sum_{j<i}^n \rho_{ij} \alpha_i \alpha_j \sigma_i \sigma_j \quad [4.16]$$

Esta suma contiene los valores individuales σ_i^2 y todos los distintos productos cruzados (covarianzas). Definiendo Σ como la matriz de covarianza, la varianza del portafolio puede escribirse como $\sigma_p^2 = w' \Sigma w$

En el caso de un portafolio conformado por una posición en un solo activo y en el que asume normalidad $N(\mu, \sigma)$ en el cambio del valor del portafolio a lo largo del horizonte de tiempo determinado, el cálculo del VaR se obtiene mediante la siguiente fórmula:

$$\text{VaR} = \sigma N^{-1}(X) \quad [4.17]$$

Donde σ es la desviación estándar del cambio en el valor del portafolio; X es el intervalo de confianza; $N^{-1}(\cdot)$ es la función de distribución normal acumulada inversa; μ es cero. Si todos los Δx_i se distribuyen normal, el rendimiento del portafolio como una combinación lineal de variables normales seguirá una distribución normal. En este

⁵⁰ Es posible asumir que el cambio esperado en una variable de mercado sobre el periodo de tiempo considerado, es igual a cero. El cambio esperado en el precio de una variable de mercado sobre un periodo de tiempo corto, es generalmente menor cuando se compara a la desviación estándar del cambio.

sentido, bajo el supuesto de normalidad, el VaR del portafolio se calcula tomando en cuenta la fórmula de las ecuaciones [4.15] y [4.17].

El modelo lineal se aplica principalmente a portafolios que consisten de posiciones en acciones, bonos, tipo de cambio y *commodities*. Este modelo también puede aplicarse para algunos derivados como futuros de moneda y swaps de tasa de interés. A pesar de que el modelo presenta limitaciones para obtener resultados precisos, los administradores de riesgos de las instituciones financieras e inversionistas utilizan el modelo lineal para calcular el valor en riesgo de un portafolio.

4.2.2.1.1. Ventajas del modelo Delta Normal

- a) El supuesto de normalidad facilita la estimación del percentil de la distribución de probabilidad de los cambios en el valor del portafolio, ya que los niveles de confianza se pueden traducir en un número específico de desviaciones estándar que representa el rango donde el valor del portafolio va a estar fluctuando en un día, semana o mes.

4.2.2.1.2. Desventajas del modelo Delta Normal

- a) La simplicidad de los supuestos que se asumen, son limitantes del modelo respecto a la precisión y complejidad para medir y controlar el riesgo de mercado.
- b) El VaR está condicionado a estimaciones de la correlación de las fuentes de riesgo obtenidos por datos históricos, cuya elección es extremadamente volátil
- c) La mayoría de los rendimientos financieros presentan exceso de curtosis, es decir, colas más pesadas. Asumir normalidad en los rendimientos probablemente subestime la proporción de datos atípicos y por ende, el VaR.
- d) Cuando el portafolio incluye instrumentos no lineales tales como las opciones financieras, el modelo lineal no toma en cuenta el efecto gamma del portafolio, impactando en el sesgo de la distribución y por lo tanto puede subestimar-sobreestimar el riesgo.

4.2.3. Simulación Monte Carlo

Una alternativa a las aproximaciones descritas hasta ahora para generar la distribución de probabilidad de los rendimientos es la simulación Monte Carlo. Estimar el VaR mediante el modelo de simulación Monte Carlo tiene como fundamento la creación de

escenarios de posibles valores del portafolio y su cálculo se basa en un percentil de la distribución del valor del portafolio simulado.

En este apartado se presenta el modelo Monte Carlo para obtener el VaR. Primero se describirá para una sola v.a. Después se detallará con múltiples fuentes de riesgo.

4.2.3.1. Simulación con una variable aleatoria

La simulación se lleva a cabo tomando en cuenta los siguientes pasos:

1. Seleccionar un proceso estocástico para el comportamiento de los precios y determinar parámetros (Un modelo comúnmente utilizado es el movimiento geométrico browniano)
2. Generar una secuencia de números pseudo aleatorios $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ con las cuales se obtienen los precios $S_{t+1}, S_{t+2}, \dots, S_{t+n}$
3. Calcular el valor del activo con los datos generados en el Paso 2.
4. Obtener la diferencia entre el valor del Paso 3 y el valor actual del activo para tener el cambio del valor del activo.
5. Repetir los pasos 1 al 4 un número finito de veces para construir una distribución de probabilidad del cambio en el valor del activo.
6. El VaR se calcula como el percentil de la distribución de probabilidad del cambio en el valor de activo obtenido en el Paso 5 dado un intervalo de confianza. Por ejemplo, si se tienen 6000 observaciones generadas en el Paso 5, el VaR a un día con un 99% de confianza será la (peor) observación número 60.

4.2.3.2. Simulación con múltiples variables

La Simulación Monte Carlo puede extenderse al caso más general de variables múltiples. A continuación, se definirá el procedimiento para transformar N fuentes independientes de riesgo en N variables correlacionadas de ϵ ⁵¹.

1. Supongamos que se tiene un vector N de valores ϵ , para los cuales se desea desplegar alguna estructura de correlación $V(\epsilon) = E(\epsilon\epsilon') = R$

⁵¹ ϵ es una variable aleatoria normal estándar, con media cero y varianza igual a uno

1.1. Como la matriz R es una matriz simétrica real, puede descomponerse utilizando la factorización de Cholesky, obteniendo una matriz triangular menor con ceros en las esquinas superiores derechas, tal que $R = TT'$

1.2. Se inicia a partir de un N vector η , compuesto por variables independientes, todas con varianzas unitarias, es decir, $V(\eta) = I$, donde I es la matriz identidad con unos sobre la diagonal principal y ceros alrededor.

1.3. Después se construye la variable $\epsilon = T\eta$. Su matriz de covarianzas es $V(\epsilon) = E(\epsilon\epsilon') = E(T\eta\eta'T') = TE(\eta\eta')T' = TIT' = R$. Por lo tanto, se confirma que los valores de ϵ tienen las correlaciones deseadas.

Una vez que se ha llevado a cabo el procedimiento anterior, los pasos para obtener el VaR son los siguientes:

2. Estimar la serie alternativa: Se utilizan los valores obtenidos en el paso 1, el precio del último día y la ecuación para determinar el valor de cada variable de mercado al final del día.

3. Reevaluar el portafolio al final del día (Precio*No. De activos i en el portafolio).

4. Obtener la diferencia entre el valor del Paso 4 y el valor actual del portafolio para tener una muestra del vector del cambio del valor del portafolio

5. Repetir la simulación un número finito de veces para construir una distribución de probabilidad del cambio en el valor del portafolio.

6. El VaR se calcula como el percentil de la distribución de probabilidad del cambio en el valor del portafolio dado un intervalo de confianza.

4.2.3.3. Ventajas del modelo Monte Carlo

a) Genera diversos escenarios considerando variaciones de la volatilidad y escenarios extremos que no han ocurrido en el pasado y por lo tanto proporciona una estimación del riesgo más confiable y precisa.

b) Permite modelar portafolios que contienen tanto instrumentos lineales como no lineales y basa sus cálculos no solo en datos históricos sino en la proyección de posibles acontecimientos futuros.

c) Se puede escoger una distribución diferente a la distribución normal y poder capturar efectos como el efecto de colas pesadas.

4.2.3.4. Desventajas del modelo Monte Carlo

- a) Aplicar generadores de números aleatorios que no produzcan series que cumplan las pruebas convencionales de independencia, puede llevar a que el rango de valores posibles del portafolio pueda estar incompleto y en consecuencia generar una medida VaR incorrecta.
- b) Es propenso al riesgo de modelo, esto quiere decir que, si no se elige un proceso estocástico realista para proyectar el precio, tampoco lo será la estimación del VaR.
- c) Su implementación requiere de un alto nivel de sofisticación y por lo tanto, altos costos computacionales. Esto se debe a que requiere llevar a cabo una gran cantidad de simulaciones de precios, que se incrementan considerablemente conforme el tamaño del portafolio. En este sentido, el acceso a un software tan especializado tiene un costo considerablemente elevado.

4.2.5. VaR mediante modelos GARCH con Innovaciones normales y t-Student

Se puede llevar a cabo la estimación del VaR utilizando modelos GARCH, aplicando el modelo descrito anteriormente para obtener la volatilidad. Sin embargo, la elección de una adecuada distribución es primordial para lograr una estimación más precisa del VaR con modelos GARCH. En este sentido, a continuación, se describe la estimación del VaR aplicando modelos GARCH con Innovaciones normales y t-Student.

El VaR con innovaciones normales, se obtiene de la siguiente forma:

$$\text{VaR}^{\text{GARCH}} = Z_p * \sigma_t \quad [4.18]$$

Donde Z_p denota el percentil de la normal estándar y σ_t la desviación estándar pronosticada por el modelo GARCH. Por otra parte, el VaR con innovaciones t-Student se estima de la siguiente forma:

$$\text{VaR}^{\text{GARCH-T}} = T(v_t) * \sqrt{\frac{v_t - 2}{v_t}} * \sigma_t \quad [4.19]$$

Donde $T(v_t)$ denota el percentil de la v.a. t-student con v_t grados de libertad y σ_t representa la desviación estándar pronosticada por el modelo GARCH.

IV.3. Backtesting

La prueba de *backtesting* se desarrolla para validar el modelo que se está utilizando para estimar el VaR. Esta técnica valida el modelo de riesgo comparando el VaR estimado con las pérdidas y/o ganancias reales (resultados reales) por un periodo determinado.

El Comité de Basilea en 1986 recomendó llevar a cabo pruebas de *backtesting* a los modelos VaR debido a que permite analizar si se está midiendo el VaR de manera apropiada. Los principios fundamentales para hacer el *backtesting* son:

1. Un grado de tolerancia para la magnitud del error entre la estimación y la realidad
2. Una tolerancia para el número de veces que puede fallar el modelo.

Las técnicas de *backtesting* se pueden dividir en tres categorías:

1. Las pruebas de cobertura: evalúan si la frecuencia de excepciones es coherente con lo que el cuantil de pérdida del VaR pretende reflejar.
2. Pruebas de distribución, son pruebas de bondad de ajuste que se aplican a la distribución de pérdida.
3. Pruebas de independencia, evalúan si los resultados parecen ser independientes de un periodo al siguiente.

4.3.1. Test de Kupiec

Una manera para comprobar la bondad del modelo VaR propuesto es comprobar cuál es la proporción de períodos de la muestra en que se observa una pérdida superior al VaR (es decir, superior al valor estimado). Se busca evaluar si la probabilidad de excepciones está en sintonía con la probabilidad p , en donde $p = 1 - \text{nivel de confianza del VaR}$.

Kupiec desarrollo el siguiente estadístico de prueba, que sigue una distribución t-Student con 1 grado de libertad:

$$LR = -2 \log \left(\frac{(p^*)^x (1 - p^*)^{n-x}}{(\tilde{p})^x (1 - \tilde{p})^{n-x}} \right) \sim \chi_1^2 \quad [4.20]$$

En donde n es el total de observaciones, x es el número de excepciones. Se contrasta la hipótesis nula de que la frecuencia observada de excepciones $\frac{x}{n} \approx \tilde{p}$ es igual a la probabilidad de error considerada para el VaR, α . Si se rechaza H_0 , el modelo VaR se rechaza.

La prueba consiste en realizar el conteo de las pérdidas y/o ganancias que exceden al VaR (excepciones) durante un periodo determinado, de forma que pueda concluirse si estadísticamente las proporciones son iguales o no, lo que determina la calidad del modelo evaluado.

IV.4. Críticas al VaR

Algunas de las críticas al VaR se presentan en los diferentes supuestos y fundamentos que se especifican en cada modelo para estimar el VaR, siendo ventajosos o limitantes para capturar la evolución dinámica del mercado. Por otra parte, Hull (2007) describe que el VaR no se considera una medida de riesgo coherente. Esto se debe a que se contradice la idea de que la diversificación reduce el riesgo, por ejemplo, si dos portafolios se combinan en un solo el VaR resultante será mayor en comparación a la suma de los dos VaR de los portafolios originales. El no cumplir con la condición de subaditividad puede tener serias consecuencias en los modelos de riesgo como asumir demasiado riesgo o a no cubrirse cuando sea necesario (Danielsson, Jorgensen, Samorodnitsky, Sarma y G.de Vries, 2005).

Otra crítica al VaR, es que dos portafolios pueden tener el mismo VaR y tener diferentes CVaR, siendo uno más riesgoso que el otro debido a que la distribución de probabilidad presenta colas pesadas. Además, se debe revisar la eficiencia del modelo dentro y fuera de la muestra debido a que un modelo puede ser válido dentro de la muestra y no serlo fuera de la muestra, lo cual puede corresponder a cambios profundos del sistema que el modelo trata de representar. Más críticas al VaR se presentan en Taleb (1997).

IV.5. Ejemplos aplicados

4.5.1. Estimación del VaR: modelo paramétrico Normal, modelo de Simulación Histórica, modelo Monte Carlo

Microsoft Excel es una herramienta básica que puede utilizarse para llevar a cabo cálculos financieros. En el siguiente ejemplo se lleva a cabo la estimación del VaR a un 95% y 99% de confianza de un portafolio eficiente utilizando el modelo paramétrico Normal, modelo de Simulación Histórica, modelo Monte Carlo.

Para llevar a cabo la estimación del VaR del portafolio se tomaron precios de cierre históricos en pesos mexicanos (MXN) de Yahoo Finance del 4 de enero de 2016 al 24 de

septiembre de 2018. Las empresas que conforman el portafolio son Alsea, S.A.B. de C.V. con código accionario ALSEA.MX, Kimberly-Clark de Mexico, S. A. B. de C. V. con código accionario KIMBERA.MX, Grupo Bimbo S.A.B de C.V. con código accionario BIMBOA.MX y Wal-Mart de México S.A.B. de C.V. con código accionario WALMEX.MX.

Se llevaron a cabo las siguientes consideraciones:

- Se consideró una inversión inicial de 1,000,000 MXN
- Se obtiene el VaR a un nivel de confianza al 95% y 99%

A continuación, se presentan los resultados obtenidos al 24 de septiembre de 2018:

Estimación del VaR			
Nivel de confianza	Modelo de Simulación Histórica	Modelo Delta Normal	Modelo Monte Carlo
95%	-\$ 15,074	-\$ 16,121	-\$15,672
99%	-\$ 24,130	-\$ 22,801	-\$20,579

Tabla 4.1: Resultados de la estimación del VaR en MXN del portafolio calculado en Excel.

Los resultados de la tabla 4.1 presentan que la mayor pérdida esperada dentro de la muestra se registró con el modelo Delta Normal al 95% de confianza y en un nivel de confianza de 99% con el modelo de Simulación Histórica. Se puede observar que el modelo de simulación histórica al 95% y el Modelo Monte Carlo al 99% resulta ser menos conservador. A un 99% de confianza el modelo de Simulación Histórica sobrestima el riesgo obtenido por el Modelo Delta Normal y Monte Carlo. Además, se puede observar que la pérdida aumenta conforme el nivel de confianza es mayor en todos los modelos.

A continuación, se presenta el VaR del portafolio en el periodo del 28 de agosto del 2018 al 24 de septiembre del 2019 a un nivel de confianza del 95%.

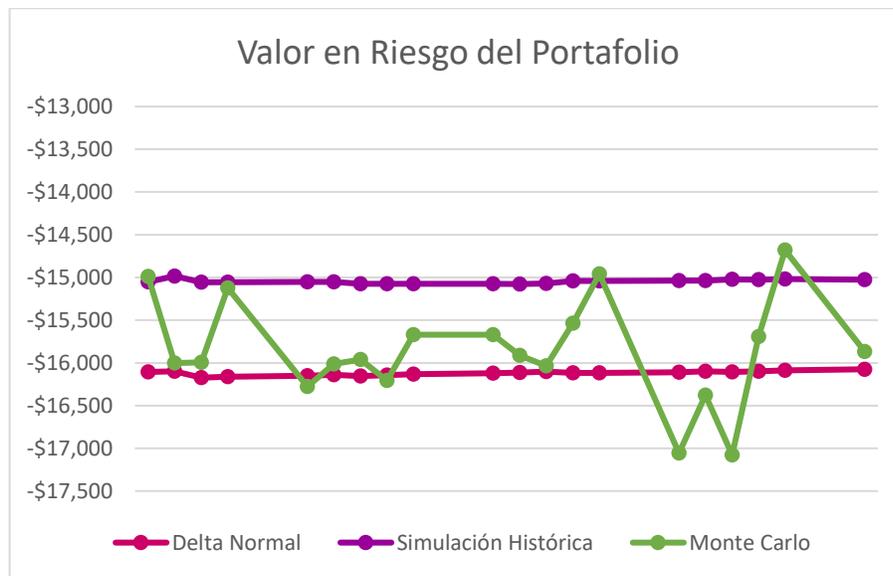


Figura 4.2: Var del portafolio. FUENTE: Estimación Propia

En la figura 4.2 se observa que el modelo Monte Carlo no marca una trayectoria constante como lo hacen los otros dos modelos.

4.5.1.1 Backtesting: Prueba de Kupiec

Adicionalmente se analizó la efectividad de los modelos con la prueba de Kupiec, con las siguientes consideraciones:

- La inversión inicial es de 1,000,000 MXN
- $\alpha = 5\%$
- La prueba se llevó a cabo sobre 20 observaciones de las 688.

Los resultados son los siguientes:

Periodo	Precio histórico de la acción				Pérdida/ Ganancia (P/G)
	ALSEA	KIMBERA	BIMBOA	WALMEX	
28/08/2018	\$69	\$33	\$39	\$54	\$7,464
29/08/2018	\$69	\$33	\$40	\$53	-\$26,164
30/08/2018	\$68	\$32	\$39	\$52	\$182
31/08/2018	\$68	\$32	\$39	\$52	-\$2,482
03/09/2018	\$67	\$32	\$39	\$52	-\$1,866
04/09/2018	\$67	\$32	\$39	\$52	-\$14,909
05/09/2018	\$66	\$31	\$38	\$52	-\$1,573
06/09/2018	\$66	\$31	\$38	\$52	\$3,332
07/09/2018	\$67	\$31	\$38	\$51	\$2,845
10/09/2018	\$67	\$31	\$38	\$52	-\$5,480

11/09/2018	\$66	\$31	\$38	\$52	-\$2,121
12/09/2018	\$65	\$31	\$38	\$52	\$14,349
13/09/2018	\$66	\$31	\$39	\$54	\$9,852
14/09/2018	\$67	\$31	\$39	\$54	-\$6,304
17/09/2018	\$66	\$31	\$40	\$54	\$2,670
18/09/2018	\$65	\$31	\$40	\$54	\$12,262
19/09/2018	\$66	\$31	\$40	\$55	-\$5,498
20/09/2018	\$65	\$31	\$40	\$55	\$982
21/09/2018	\$65	\$32	\$40	\$55	\$1,831
24/09/2018	\$64	\$32	\$40	\$55	

Tabla 4.2. Precios y vector de pérdidas y ganancias utilizados para la prueba de Backtesting.
FUENTE: Estimación Propia

PARA MODELO DELTA NORMAL

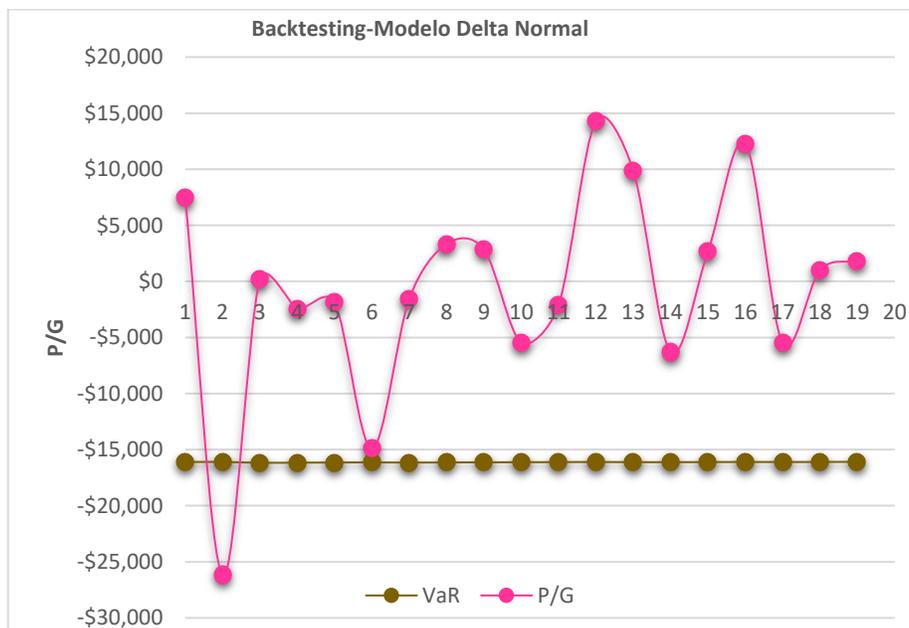


Figura 4.3: Backtesting con Modelo Delta Normal. FUENTE: Estimación Propia

PARA MODELO DE SIMULACIÓN HISTÓRICA:

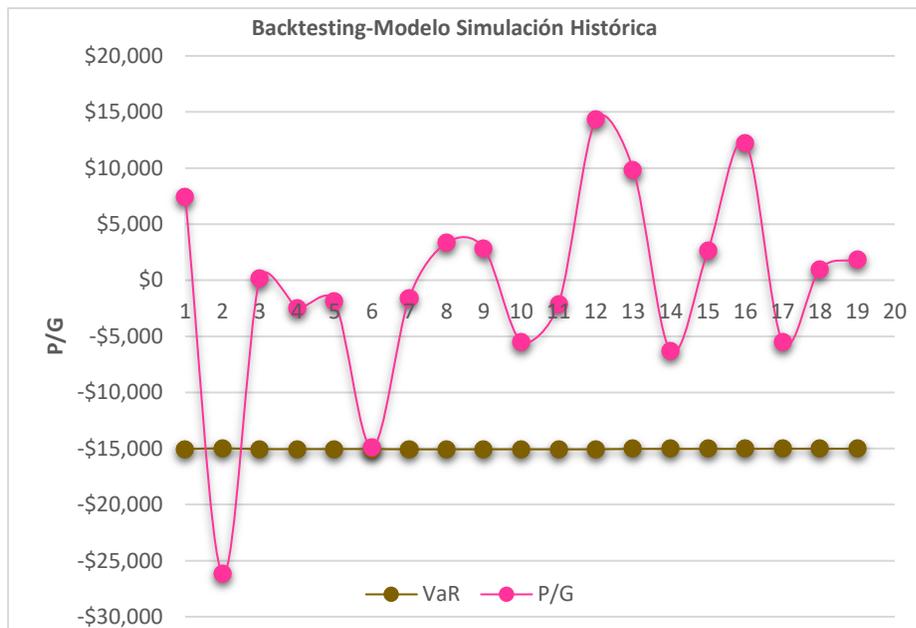


Figura 4.4: Backtesting - Modelo Simulación Histórica. FUENTE: Estimación Propia

- **PARA MODELO MONTE CARLO**

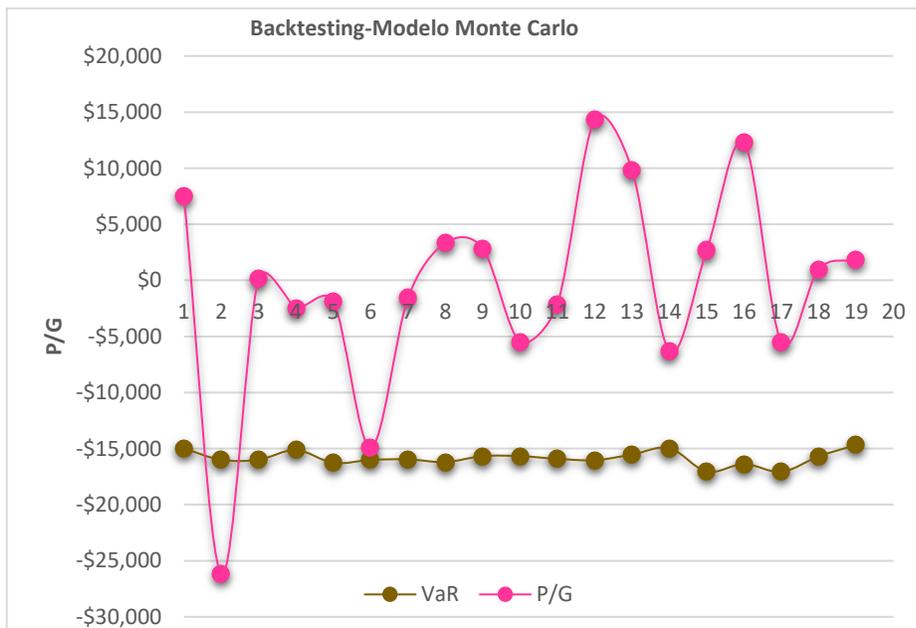


Figura 4.5: Backtesting con Modelo Simulación Monte Carlo. FUENTE: Estimación Propia

De la figura 4.3 a la figura 4.5 se presentan las pérdidas y ganancias que se obtuvieron con cada modelo sobre el periodo seleccionado. Además, se graficó el VaR para poder

observar el número de excepciones o fallos que se presentaron. En los resultados de la prueba se obtuvo solo una falla entre las 20 observaciones seleccionadas, esto ocurrió para todas las metodologías. Con base en la prueba de Kupiec, se observa que los modelos no excedieron el VaR ajustado en más de 5% del total de las observaciones (p-value mayor a 0.05). Esto se debería cumplir para todo el periodo seleccionado, sin embargo, en la presente tesina no se llevó a cabo el ejercicio para las 688 observaciones. Además, cabe destacar que los resultados de la prueba pueden depender de la cantidad de inversión inicial que se tenga.

4.5.2. Estimación del VaR dinámico: Modelo paramétrico Normal, Modelo EWMA, Modelo Simulación Histórica, Modelo Simulación Histórica Filtrada

Python es un lenguaje de programación gratuito que ha ganado popularidad en la industria financiera debido a las paqueterías avanzadas y completas con las que cuenta, su velocidad y la facilidad de codificar. A continuación, se presenta un comparativo del comportamiento dinámico del riesgo con diversos modelos de la acción ALSEA en Python. El código utilizado se presenta en el Apéndice 3 de la presente tesina y presenta los siguientes modelos para estimar el VaR en MXN: Modelo paramétrico, EWMA, Simulación Histórica y Simulación Histórica Filtrada. Los datos empleados para llevar a cabo los ejemplos corresponden a la serie de los precios de cierre diarios en MXN de Alsea, S.A.B. de C.V. con código accionario ALSEA.MX para el periodo que comprende del 3 de enero de 2012 al 31 de diciembre de 2018. Dichos datos fueron obtenidos a partir de la serie de datos de YAHOO FINANCE.

Además, se toman las siguientes especificaciones:

1. En la función principal "VaR", se define un nivel de confianza de 95% para todos los modelos.
2. Se obtiene el VaR paramétrico, llamando una función normal y la desviación estándar.
3. Para el modelo EWMA se establece el parámetro $\lambda=0.94$, que es el valor que se utiliza en el sistema Riskmetrics.
4. Se obtiene el modelo de Simulación Histórica
5. Debido a que en el modelo de simulación histórica filtrada se ajusta un modelo de volatilidad condicional para los datos históricos, para este ejemplo es una combinación del modelo de simulación histórica y el modelo EWMA

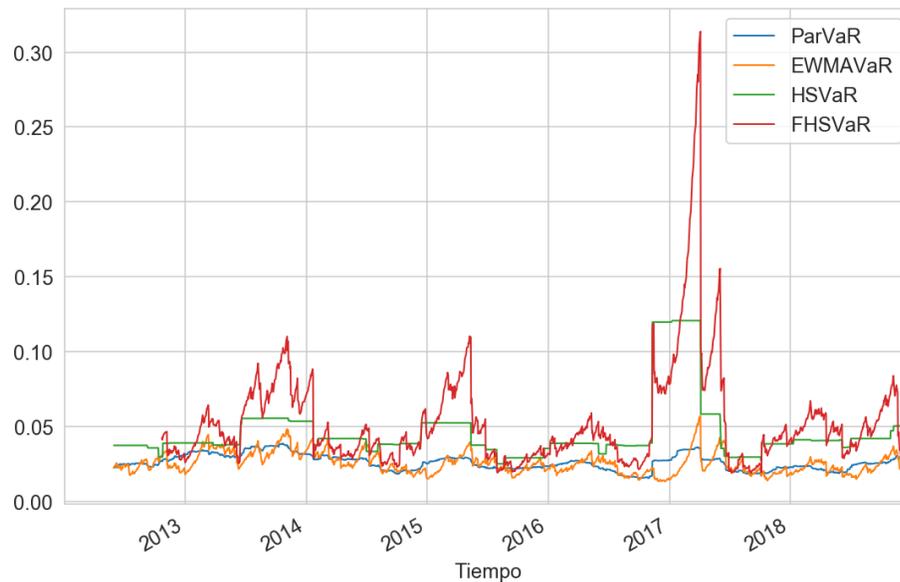


Figura 4.6: Resultados de la estimación del VaR con las diferentes metodologías en Python.
 FUENTE: Estimación Propia.

En la gráfica de la Figura 4.6 se presentan las estimaciones dinámicas del VaR dentro de la muestra del modelo paramétrico, el modelo EWMA, la simulación histórica y la simulación histórica filtrada. Se observa que el VaR de Alsea obtenido por SHF resulta ser conservador con respecto a las demás metodologías. En términos de la evaluación del comportamiento dinámico del VaR, los resultados empíricos señalan que el modelo de SHF captura el fenómeno de heteroscedasticidad condicional presente en los rendimientos del índice, es decir, modela las condiciones cambiantes del mercado sin permanecer en ningún momento constante por lo cual proporciona información más completa y real del riesgo al que se encuentran expuestos los agentes económicos.

Llevar a cabo este tipo de comparativos, contribuye a la toma de decisiones que llevan a cabo los agentes económicos y para la administración de riesgos se puede determinar que metodología es una herramienta eficiente para diseñar estrategias de cobertura optima ante cambios inesperados en los mercados financieros que afecten el comportamiento del precio de los instrumentos financieros.

4.5.3 Análisis estadístico de las Series de Tiempo

The R Project for Statistical Computing (R) es un lenguaje de programación gratuito con una gran variedad de paqueterías estadísticas y además genera gráficos de alta calidad. A continuación, se lleva a cabo un ejemplo de un análisis estadístico de la serie de

rendimientos del índice accionario NIKKEI en R. El código utilizado se presenta en el Apéndice 4 de la presente tesina. Los datos empleados para llevar a cabo los ejemplos corresponden a la serie de los precios de cierre diarios del índice accionario NIKKEI en yen japonés (JPY) para el periodo que comprende del 1 de enero de 2006 al 31 de diciembre de 2011. Dichos datos fueron obtenidos utilizando la librería `qrmdata` de R. Se llevan a cabo los siguientes pasos:

1. Se sabe que las series financieras presentan “hechos estilizados” (como ya se comentó anteriormente). Para modelar la correlación serial de la media condicional se aplicó un proceso $AR(1)$. Además, para capturar la heterocedasticidad de la serie financiera se llevó a cabo un ajuste $AR(1) - GARCH(1,1)$ con innovaciones normales y t -Student.

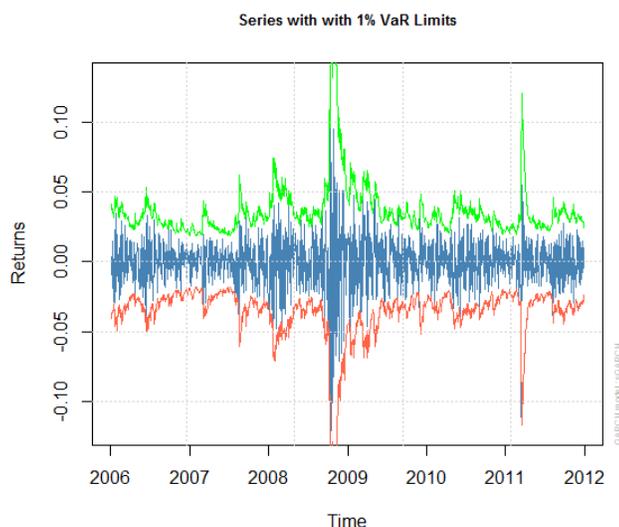


Figura 4.7: Rendimientos del índice NIKKEI y los límites del VaR en el nivel de confianza del 99%. FUENTE: Estimación Propia.

En la gráfica 4.7 se compara el comportamiento entre los rendimientos del índice NIKKEI contra los límites del VaR en el nivel de confianza del 99%. Se puede observar que para el nivel de confianza del 99% el VaR estimado por $AR(1) - GARCH(1,1)$ con innovaciones normales no presenta muchos fallos, entendiendo que un fallo ocurre cuando el rendimiento actual excede el valor del VaR estimado. Por lo anterior, se puede asumir que el modelo captura las condiciones del mercado en su mayoría.

2. Se lleva a cabo un ajuste $AR(1) - GARCH(1,1)$ con innovaciones normales y con innovaciones t-Student. Se observa que la distribución de los rendimientos presenta colas pesadas y que con las innovaciones t-Student esto mejora. En las siguientes gráficas se presentan los resultados del ajuste (Ver QQ PLOT en figuras 4.7 y 4.8):

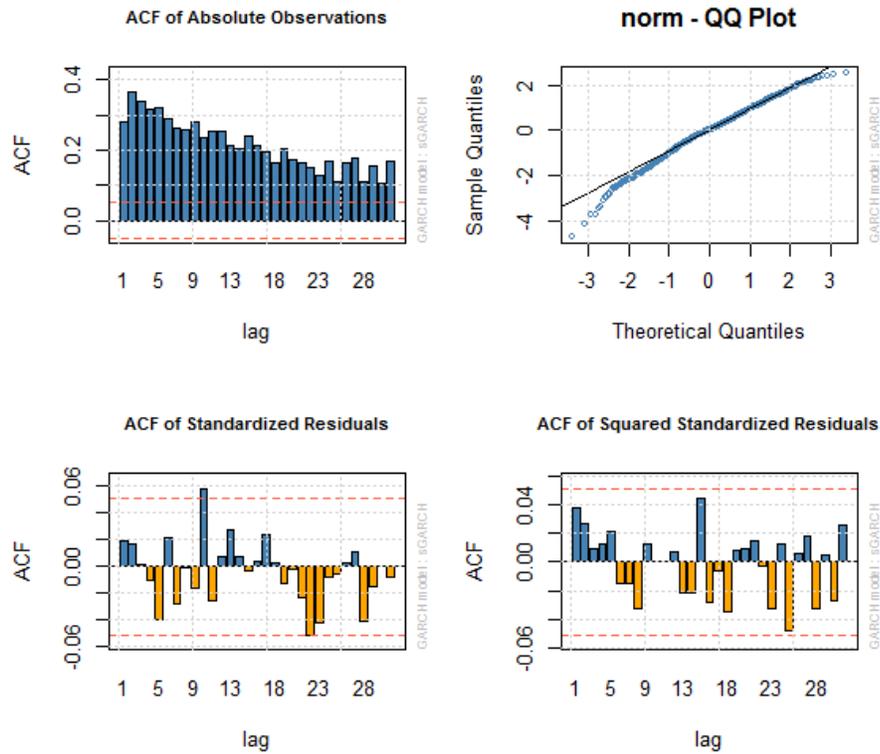


Figura 4.8: Análisis con innovaciones normales FUENTE: Estimación Propia.

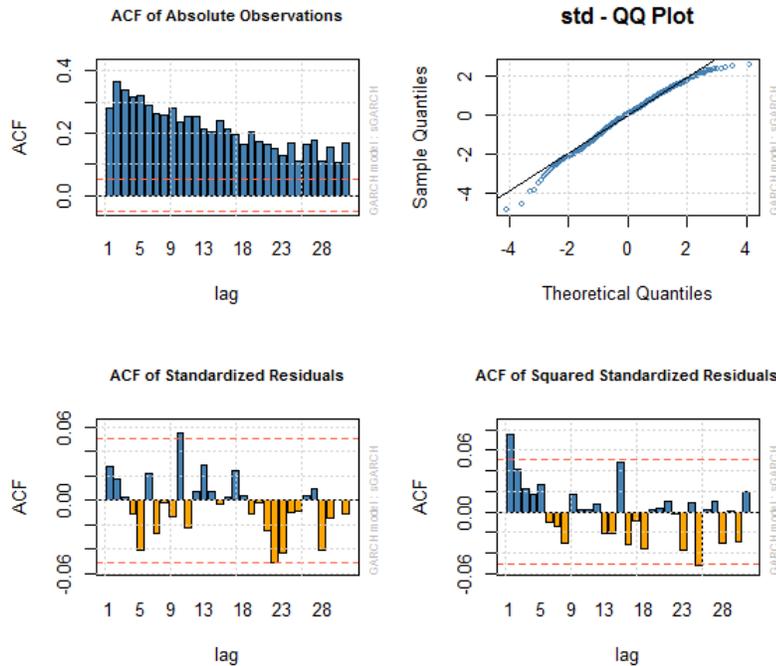


Figura 4.9: Análisis con innovaciones T-Student. FUENTE: Estimación Propia.

En la Figura 4.8 y 4.9 se presenta (partiendo de la esquina superior izquierda) la ACF de los valores absolutos de los datos, las cuales presentan correlación serial, se presenta la gráfica de los cuantiles de la distribución normal contra los cuantiles de la distribución empírica (QQ-PLOT) se observa la distribución de los residuales presenta leptocurtosis por lo cual se rechaza el supuesto de normalidad. En la parte inferior derecha se observa la ACF de los residuales estandarizados, que muestran si la dinámica AR explica la media condicional. También se presenta la ACF de los residuales estandarizados al cuadrado, los cuales muestran si la dinámica GARCH explica la desviación estándar condicional.

Una vez obtenidos los residuales estandarizados independientes e idénticamente distribuidos y la volatilidad condicional, se determina si los residuales son adecuados para estimar el VaR por simulación histórica de una forma dinámica.

3. Para seleccionar un modelo, se escogió el modelo que minimizó el criterio de información de Akaike En este caso se prefirió al modelo con innovaciones t-Student, y se trabajó con este modelo. Se obtiene la desviación estándar condicional y el VaR a un nivel de confianza del 95%, los cuales se presentan a continuación:

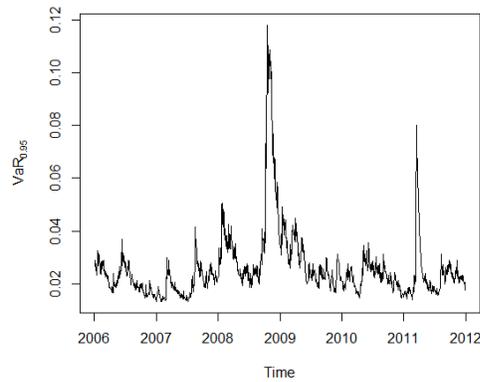


Figura 4.10: VaR, 95% de confianza en R. FUENTE:
Estimación Propia.

. En la Figura 4.10 el VaR al 95% de confianza ajustado a un modelo AR-GARCH utilizando R. Se pueden observar picos a finales del 2008 y en el 2011 y valles a mediados del 2008 y en el 2011.

4. Se revisan los residuales del modelo, se busca cumplir:

4.1. Media, ≈ 0

4.2. Desviación estándar ≈ 1

4.3. Sesgo, tiene que ser 0 (si es $< 0 \Rightarrow$ sesgada a la izquierda)

4.4. Curtosis tiene que ser $6/(df-4)$

4.5. Se lleva a cabo el estadístico Anderson Darling, con el cual se busca saber si la distribución t-Student se ajusta a los datos.

```

> ## Residuales
> (mu.Z <- mean(Z))
[1] -0.05866842
> (sd.Z <- as.vector(sd(Z)))
[1] 0.9991891
> skewness(Z)
      X
-0.4412051
> hist(Z)
> kurtosis(Z)
      X
 3.69418
> nu <- param[["shape"]]
> 6/(nu-4)
[1] 0.6915911
> pt.hat <- function(q) pt((q-mu.Z)/sd.Z, df = nu)
> ad.test(as.numeric(Z), distr.fun = pt.hat)

Anderson-Darling GoF Test

data: as.numeric(Z) and pt.hat
AD = 4.9344, p-value = 0.003082
alternative hypothesis: NA

```

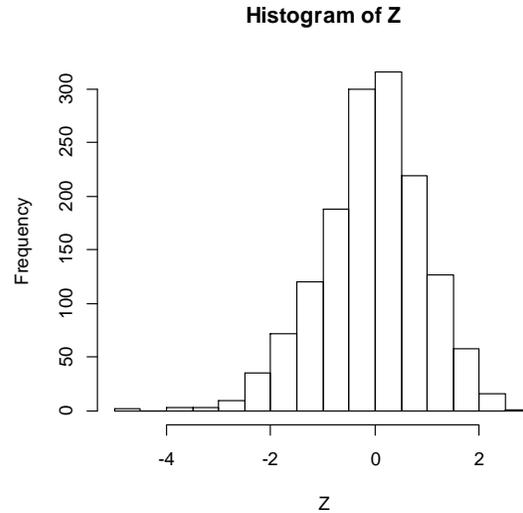


Figura 4.11: Análisis residual. FUENTE: Estimación Propia.

En la figura 4.11 se presenta que la distribución de los residuales está sesgada a la izquierda (ver histograma del lado derecho). Además, el estadístico Anderson Darling, arrojo que la distribución t-Student no se ajusta a los datos. Por lo cual, para mayor precisión se podría probar llevar a cabo un Modelo GJR-GARCH, un modelo GARCH con innovaciones con una distribución asimétrica o llevar a cabo especificaciones diferentes para la media condicional.

5. Predicción a partir del modelo ajustado, Fecha inicial=2011-12-30

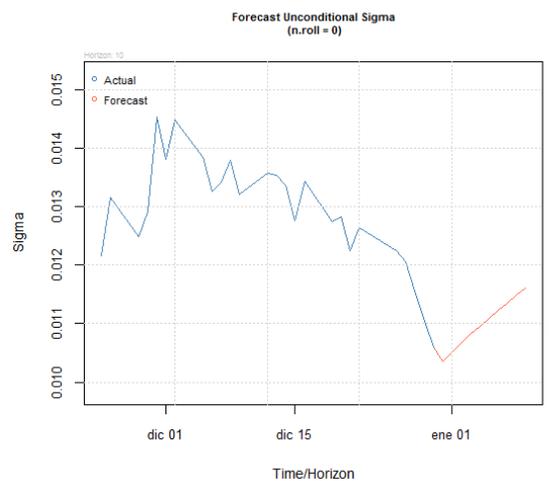


Figura 4.12: Pronóstico de Volatilidad. FUENTE: Estimación Propia.

V. Conclusiones

La medición del riesgo se ha vuelto esencial en las instituciones financieras, además, las regulaciones financieras buscan una mejor comprensión del riesgo. El VaR sigue siendo una herramienta estadística muy popular ya que sus ventajas prácticas superan sus deficiencias teóricas. El VaR provee información valiosa para identificar pérdidas potenciales durante periodos estables del mercado, sin embargo, es recomendable completar cualquier análisis de riesgo con pruebas de estrés, backtesting, mediciones de riesgo para escenarios extremos y otras medidas como CVAR.

El presente trabajo contribuye a tener una idea general del proceso para estimar el VaR. Además se describen algunas de las distintas metodologías para la estimación de la herramienta estadística VaR. De esta forma se cumple con el objetivo primordial, que es presentar y desarrollar los conceptos primordiales para llegar al entendimiento de la estimación del VaR. Además, deja la ventana abierta para el desarrollo pro activo de las metodologías y entender la viabilidad de su uso al aplicarlo a series financieras, en su caso refutarlo y/o buscar implementar las herramientas existentes o crear nuevas. Cabe aclarar que las metodologías aplicadas en la presente tesina simplemente se aplicaron como ejemplos por lo cual no necesariamente se debe aplicar estas metodologías para otras series de datos.

Finalmente, la presente tesina contiene una gran variedad de temas sobre el VaR en los que se podría profundizar y con esto poder generar diversas investigaciones en el campo del Álgebra Lineal, Estadística, Series de Tiempo y Programación.

Apéndice

A.1. Propiedades empíricas de las series financieras

Cabe destacar que algunas funciones para graficar en Python, traen los títulos en inglés ya predeterminados, por lo cual no se presentan en español.

```
#Librerías ----
#Librerías para gráficos
import pandas as pd
import numpy as np
import seaborn as sns

import matplotlib.pyplot as plt
from matplotlib.pyplot import figure
import datetime as dt

import statsmodels.formula.api as smf
import statsmodels.tsa.api as smt
import statsmodels.api as sm
import scipy.stats as scs

#Librerías para datos

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from matplotlib import mlab
import datetime as dt
from matplotlib import style
pd.core.common.is_list_like = pd.api.types.is_list_like
import pandas_datareader as web
from scipy.stats import norm
import math

#Descarga de series financieras de yahoo
df = web.DataReader('ALSEA.MX','yahoo', start='2012-01-01', end='2018-12-31')
df = df[['Adj Close']]
df.columns = ['adj_close']

# Se obtienen los rendimientos
df['simple_rtn'] = df.adj_close.pct_change()
df['log_rtn'] = (np.log(df.adj_close) - np.log(df.adj_close.shift(1))).dropna()
# dropping NA's in the first row
df.dropna(how = 'any', inplace = True)

#Opcional, solo para observar los datos
df.head()
print(df.log_rtn)

# Configuraciones ----
%config InlineBackend.figure_format = 'retina'
%matplotlib inline
sns.set_style('darkgrid')
sns.mpl.rcParams['figure.figsize'] = (10.0, 6.0)
sns.mpl.rcParams['savefig.dpi'] = 90
```

```

sns.mpl.rcParams['font.size'] = 14

# Grafica de series de tiempo ----
fig,ax =plt.subplots(2, 1, figsize=(24, 20))
# Precio ----
df.adj_close.plot(ax=ax[0])
ax[0].set_ylabel('Precio de la acción($)')
ax[0].set_xlabel("")
ax[0].set_title('Precio vs. rendimientos')
# Rendimientos ----
df.log_rtn.plot(ax=ax[1])
ax[1].set_ylabel(' Rendimientos (%)')
fig.show()

# Gráfica de la distribución de rendimientos ----
ax = sns.distplot(df.log_rtn, kde = False, norm_hist=True)
xx = np.linspace(min(df.log_rtn), max(df.log_rtn), num=1000)
yy = scs.norm.pdf(xx, loc=df.log_rtn.mean(), scale=df.log_rtn.std())
ax.plot(xx, yy, 'r', lw=2)
ax.set_title('Distribución de rendimientos de IBM')
# QQ plot ----
qq = sm.qqplot(df.log_rtn.values, line='s')
qq.show()

# Estadística descriptiva ----
print('Range of dates:', min(df.index.date), '-', max(df.index.date))
print('Number of observations:', df.shape[0])
print('Mean: {0:.4f}'.format(df.log_rtn.mean()))
print('Median: {0:.4f}'.format(df.log_rtn.median()))
print('Min: {0:.4f}'.format(df.log_rtn.min()))
print('Max: {0:.4f}'.format(df.log_rtn.max()))
print('Standard Deviation: {0:.4f}'.format(df.log_rtn.std()))
print('Skewness: {0:.4f}'.format(df.log_rtn.skew()))
print('Kurtosis: {0:.4f}'.format(df.log_rtn.kurtosis())) #Kurtosis of std. Normal dist = 0
print('Jarque-Bera statistic: {stat:.2f} with p-value: {p_val:.2f}'.format(stat =
scs.jarque_bera(df.log_rtn.values)[0],
p_val = scs.jarque_bera(df.log_rtn.values)[1]))

# Gráfica de autocorrelación de los rendimientos, cuadrados y absolutos ----
# Especificar el número máximo de retardos
# Gráfica de Autocorrelation de los rendimientos ----
acf_r = smt.graphics.plot_acf(df.log_rtn, lags=40 , alpha=0.05)
acf_r.show()

# Gráfica de Autocorrelation parcial de los rendimientos ----

acf_r = smt.graphics.plot_pacf(df.log_rtn, lags=40 , alpha=0.05)
acf_r.show()

# Especificar número de retardos
lags = 40

fig, ax =plt.subplots(2, 1, figsize=(9, 6))
# squared returns ----
smt.graphics.plot_acf(df.log_rtn**2, lags=lags, alpha=0.05, ax = ax[0])
ax[0].set_ylabel('Rendimientos al cuadrado')
ax[0].set_xlabel("")

```

```

ax[0].set_title('Función de Autocorrelación ')
# absolute returns ----
smt.graphics.plot_acf(np.abs(df.log_rtn), lags=lags, alpha=0.05, ax = ax[1])
ax[1].set_ylabel('Rendimientos absolutos')
ax[1].set_title("")
ax[1].set_xlabel('Retardo')

```

A.2. Modelación de datos

Modelo GARCH

-----Python:

```

from arch import arch_model

am = arch_model(df.log_rtn)
res = am.fit(update_freq=5)
print(res.summary())

fig = res.plot(annualize='D')
fig=res.plot()

```

-----R programming:

```

install.packages("tseries")
install.packages("foreign")
library(tseries)
library(foreign)

mydata<-read.csv(file.choose())
attach(mydata)
X<-Return
summary(X)

plot(X)
arch.X<-garch(X,order=c(0,1))
summary(arch.X)

GARCH.X<-garch(X,order=c(1,1))
summary(GARCH.X)

```

A.3. Cálculo del VaR dinámico: Modelo paramétrico Normal, Modelo EWMA, Modelo Simulación Histórica, Modelo Simulación Histórica Filtrada

#Para descargar el código completo se puede visitar el repositorio de GitHub EN donde se encuentra el código: <http://github.com/BSIC/VaR>

#En la primera sección se llaman a las librerías necesarias. Se define una función que maneja los parámetros que se llegarán a utilizar en las subfunciones. También se lleva a cabo una condición para no considerar los datos vacíos. En cada subfunción se hará una distinción entre obtener el VaR con datos del intervalo del último periodo u obtener el VaR con toda la serie de datos.

```

from scipy.stats import norm
import numpy as np
import pandas as pd

```

```

import math
def VaR>Returns, Formula = 'Parametric Normal', Confidence_Interval = 0.95 , Period_Interval = None ,
EWMA_Discount_Factor = 0.94, Series = False ,removeNa = True):

    # Se quitan los NA
    if removeNa==True: Returns = Returns[pd.notnull>Returns)]

    if (Series == True and Period_Interval == None):
        Period_Interval = 100
    elif Period_Interval == None:
        Period_Interval = len>Returns)

#La serie de datos se descargó de la siguiente forma:

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from matplotlib import mlab
import datetime as dt
from matplotlib import style
pd.core.common.is_list_like = pd.api.types.is_list_like
import pandas_datareader as web
from scipy.stats import norm
import math

# Se descargan los datos de yahoo
df = web.DataReader('ALSEA.MX','yahoo', start='2012-01-01', end='2018-12-31')
df = df[['Adj Close']]*100
df.columns = ['adj_close']

# Se obtienen los rendimientos
df['simple_rtn'] = df.adj_close.pct_change()
df['log_rtn'] = (np.log(df.adj_close) - np.log(df.adj_close.shift(1))).dropna()
# dropping NA's in the first row
df.dropna(how = 'any', inplace = True)

df_VaR_compare = VaR_Compare(df.log_rtn)
plt.show()

```

A.4. Cálculo del VaR utilizando modelos GARCH con innovaciones normales y t-Student

#Código extraído de QRM_R_CODES: "04_Financial_Time_Series"- "04_GARCH_Estimation". Para ver el código completo ir a la referencia.

By Alexander McNeil and Marius Hofert

Librerías

```
library(rugarch)
library(zoo)
library(ADGofTest) # for ad.test()
library(moments) # for skewness(), kurtosis()
library(qrmdata)
library(qrmtools)
```

1. Datos

```
# NIKKEI
data("NIKKEI")
plot.zoo(NIKKEI, xlab = "Time")
```

```
# Descargar datos y obtener los rendimientos
SPdata <- NIKKEI['2006-01-01/2011-12-31']
plot.zoo(SPdata, xlab = "Time")
```

```
X <- returns(SPdata)
plot.zoo(X, xlab = "Time")
```

2 Ajustar un modelo AR(1)--GARCH(1,1) con innovaciones normales

```
uspec.N <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1,1)),
  mean.model = list(armaOrder = c(1,0), # parte AR(1)
  include.mean = TRUE), # con media
  distribution.model = "norm") # innovaciones normales
(fit.N <- ugarchfit(spec = uspec.N, data = X))## Se obtiene la estimación de los parámetros
```

```
# Análisis de la serie
plot(fit.N, which = 2)
layout(matrix(1:4, ncol = 2, byrow = TRUE))
plot(fit.N, which = 6)
plot(fit.N, which = 9)
plot(fit.N, which = 10)
plot(fit.N, which = 11)
layout(1)
```

3 Ajustar un modelo AR(1)--GARCH(1,1) con innovaciones t Student

```
uspec.t <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1,1)),
  mean.model = list(armaOrder = c(1,0), include.mean = TRUE),
  distribution.model = "std")
fit.t <- ugarchfit(spec = uspec.t, data = X)
```

```
layout(matrix(1:4, ncol = 2, byrow = TRUE))
plot(fit.t, which = 6)
```

```

plot(fit.t, which = 9)
plot(fit.t, which = 10)
plot(fit.t, which = 11)
layout(1)

# 4 Se lleva a cabo una prueba para saber que modelo se ajusta mejor a los datos
LL.N <- fit.N@fit$LLH
LL.t <- fit.t@fit$LLH
LRT <- 2*(fit.t@fit$LLH-fit.N@fit$LLH)
LRT > qchisq(0.95, 1)
1-pchisq(LRT, df = 1)

# 5 Se utiliza el Criterio de Akaike para comparar los modelos

(AIC.N <- 2*length(fit.N@fit$coef) - 2*LL.N)
(AIC.t <- 2*length(fit.t@fit$coef) - 2*LL.t)

#Se obtiene el VaR
class(fit.t)
getClass("uGARCHfit")
getSlots("uGARCHfit")
fit.t.fit <- fit.t@fit
names(fit.t.fit)
fit.t.model <- fit.t@model
names(fit.t.model)
fit.t.model$modeldesc
fit.t.model$pars

(param <- coef(fit.t))
sig <- sigma(fit.t)
VaR.95 <- quantile(fit.t, probs = 0.95)
Z <- residuals(fit.t, standardize = TRUE)

# Plots
plot.zoo(sig, xlab = "Time", ylab = expression(hat(sigma)[t]))
plot.zoo(VaR.95, xlab = "Time", ylab = expression(widehat(VaR)[0.95]))
plot.zoo(Z, xlab = "Time", ylab = expression(hat(Z)[t]))

# 6 Se lleva a cabo el análisis residual del modelo ajustado
(mu.Z <- mean(Z))
(sd.Z <- as.vector(sd(Z)))
skewness(Z)
hist(Z)
kurtosis(Z)
nu <- param[["shape"]]
6/(nu-4)
pt.hat <- function(q) pt((q-mu.Z)/sd.Z, df = nu)
ad.test(as.numeric(Z), distr.fun = pt.hat) ## => Anderson--Darling

# 7 Pronóstico a partir del modelo ajustado

(forc <- ugarchforecast(fit.t, n.ahead = 10))
plot(forc, which = 1) # prediction of X_t
plot(forc, which = 3) # prediction of sigma_t

```

Bibliografía

Ahmadi-Javid, A. y Fallah-Tafti, M., (2019). Portfolio optimization with entropic value-at-risk. *European Journal of Operational Research*, volume 279, pp. 225-241.

Aloui, C y Mabrouk S., (2010). Value-at-risk estimations of energy commodities via long-memory, asymmetry and fat-tailed GARCH models. *Energy Policy*, volume 38, Issue 5, pp. 2326-2339.

Al Janabi, M., (2007). Risk analysis, reporting and control of equity trading exposure: Viable applications to the Mexican financial markets. *Journal of Derivatives & Hedge Funds*, Volume 13, Issue 1, pp 33–58

Angelidis, T y Degiannakis, S., (2005). Modeling risk for long and short trading positions. *The Journal Risk of Finance*. Vol 6. No 3. 2005.

Bollerslev, T., (1986). Generalized Autoregressive Conditional heteroskedasticity. *Journal of Econometrics*, *Journal of Econometrics*, vol. 31, Issue 3, pp. 307-327.

Bouveret, A., (2018). Cyber Risk for the Financial Sector: A Framework for Quantitative Assessment. IMF Working Paper, Strategy, Policy & Review Department.

Bowerman, B., O'Connell , T., Koehler, A., (2007).Pronósticos, series de tiempo y regresión. Cengage Learning

Box, G., Jenkins, G., (1976), Time series analysis, forecasting and control. Holden-Day series in time series analysis and digital processing

Buchner, A., (2017). Risk Management for Private Equity Funds. *Journal of Risk*, volume 19, No. 6.

Campbell, S., (2005). A review of backtesting and backtesting procedures. Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.

Cervantes, D., (2014). Series de tiempo de memoria larga con volatilidad estocástica y su aplicación a la inflación Mexicana. Tesis de Licenciatura. México, Facultad de Ciencias, Universidad Autónoma de México.

- Chen, Q. y Chen, R., (2013). Method of Value-at-Risk and empirical research for Shanghai stock market. *Procedia Computer Science*, volume 17, pp 671,677
- Christoffersen, P., Hahn J. y Inoue, A., (2001). Testing and Comparing Value-at-Risk Measures. *Journal of Empirical Finance*, volume 8, pp.325-342.
- Cong, J. Seng Tan, K., Weng, C., (2013). VaR-Based Optimal Partial Hedging. Published online by Cambridge University Press. Volume 43, Issue 3, pp. 271-299. Recuperado de <https://www.cambridge.org/core/journals/astin-bulletin-journal-of-the-iaa/article/varbased-optimal-partial-hedging/713D560A7BCC7D8A43AD4BB4E786F36A>
- Cont, R., (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, volume 1, pp. 223–236
- Corbetta, J y Peri, I., (2018). A New Approach to Backtesting and Risk Model Selection
- Cuoco,D y Liu, H., (2005). An análisis of VaR-based capital requirements. *Journal of Financial Intermediation*. Volume 15, Issue 3, pp. 362-394.
- Danielsson, J., Jorgensen, B., Samorodnitsky, G., Sarma, M y G.de Vries, C., (2005). Subadditivity Re-Examined: the Case for Value-at-Risk. Recuperado de <https://faculty.washington.edu/ezivot/econ589/VaRsubadd.pdf>
- Del Brio, E.B., Mora-Valencia, A. y Perote, J., (2017). Risk quantification for commodity ETFs: Backtesting value-at-risk and expected shortfall. *International Review of Financial Analysis*. Recuperado de <https://www.sciencedirect.com/science/article/abs/pii/S1057521917301801>
- Dr. J. R., (2002). Improving Value at Risk Calculations by Using Copulas and Non-Gaussian Margin. MSc in Mathematical Finance. University of Oxford
- Du, J., Li, Z. y Wu, L., (2019). Optimal Stop-Loss Reinsurance Under the VaR and CTE Risk Measures: Variable Transformation Method. *Computational Economics*, volume 53, Issue 3, pp. 1133-1151.
- Engle, R., (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom Inflation. *Econometrica*, volume 50, Issue 4, PP.987-1008
- Gkillas, K. y Katsiampa, P., (2018). An application of extreme value theory to cryptocurrencies. *Economics Letters*, volume 164, pp. 109-111.

Guo, T., Ding, J., Guo, B., (2018). Fat Tails Value at Risk and the Palladium Returns. *Global Journal of Management And Business Research*. Recuperado de <https://journalofbusiness.org/index.php/GJMBR/article/view/2490>.

Hammoudeh, S., Malik, F. y McAleer, M., (2011). Risk management of precious metals. *The Quarterly Review of Economics and Finance*, Volume 51, Issue 4, pp. 435-441.

Hans, P. y Van Dijk, D., (2000). Non-linear time series models in empirical finance. Cambridge University Press.

Hernández, A., (2016). Una introducción a la teoría de valores extremos. Tesina de Licenciatura. México, Facultad de Economía, Universidad Autónoma del Estado de México.

Hernández-Lobato, J. M., Hernández-Lobato, D y Suárez, A., (2007). Time Series Models for Measuring Market Risk. Technical Report. Departamento de Ingeniería Informática, Universidad Autónoma de Madrid.

Holton, G. (2014) Value-at-Risk: Theory and practice. Published by the autor

Holton, G., (2017). How to Calculate Value-at-Risk – Step by Step. Recuperado de https://www.glynholton.com/blog/risk-measurement/var_measure/

Hosseini, S.D. y Verma, M., (2017). A Value-at-Risk (VAR) approach to routing rail hazmat shipments. *Transportation Research Part D: Transport and Environment*, volume 54, pp. 191-211.

Hull, J. y White, A., (1998). Value at Risk when daily changes are not normally distributed. *Journal of Derivatives*, v5.

Hull, J. (2007). Risk Management and Financial Institutions. Wiley

Hung, J.C., Lee, M.C, Liu, H., (2007). Estimation of value-at-risk for energy commodities via fat-tailed GARCH models. *Energy Economics*, volume 30, Issue 3, pp. 1173-1191.

Hung, J. C., Su, J. B., Chang, M.C. y Wang, Y.H., (2019). The impact of liquidity on portfolio value-at-risk forecasts. *Applied economics*. Recuperado de <https://www.tandfonline.com/doi/abs/10.1080/00036846.2019.1644442>

Jorion, P., (2002). How Informative Are Value-at-Risk Disclosures?. The accounting review, Vol. 77, No.4, pp. 911-931.

Jorion, P., (2007). Value at Risk. The New Benchmark for Managing Financial Risk. McGraw Hill. Tercera edición.

Jorion, P., (2009) El nuevo paradigma para el control de riesgos con derivados. Limusa.

J.P. Morgan (1996). RiskMetrics Technical Document. Fourth edition, J.P. Morgan/Reuters.

Kong, X. M., Huang, G. H., Fan, Y.R., Li, Y.P., Zeng, X. T. y Zhu, Y., (2017). Risk analysis for water resources management under dual uncertainties through factorial analysis and fuzzy random value-at-risk. *Stochastic Environmental Research and Risk Assessment*, volume 31, Issue 9, pp. 2265-2280.

Kupiec, P., (1995). Techniques for Verifying the Accuracy of Risk Measurement Models. *Journal of Derivatives*, volume 3, issue 2, pp. 73-84.

Laporta, A., Merlo, L. y Petrella, L., (2018). Selection of Value at Risk Models for Energy Commodities: *Energy Economics*, volume 74, pp. 628-643.

Levine, E., Lung, B., Nimbalkar, S.U., Wenning, T., Kelly, N. y Power, Z., (2019). Applying Value at Risk Analysis towards Energy Efficiency Investments. Recuperado de <https://www.osti.gov/biblio/1559641>

Malhotra, Y., (2015). Stress Testing for Cyber Risks: Cyber Risk Insurance Modeling beyond Value-at-Risk (VaR): Risk, Uncertainty, and, Profit for the Cyber Era. *Insurance: Global Risk Management Network*, LLC.

Malik, F., (2019). Market Risk: Explanation Of VaR Calculations. Recuperado de <https://medium.com/fintechexplained/market-risk-explanation-of-var-calculations-e31734f26d93>

Malik, F., (2019). What Is Value At Risk?. Recuperado de <https://medium.com/fintechexplained/what-is-value-at-risk-fc1af4fdec48>

Massó, V y Pons, E., (2018). Caminata aleatoria o "Random Walk". Recuperado de https://rstudio-pubs-static.s3.amazonaws.com/372908_51086b0fd9af49e3966d372046e7969a.html

Mensi, W., Hussain, S. J., Hammoudeh, S., Zeitun, R. y Rehman, M. U., (2017). Diversification potential of Asian frontier, BRIC emerging and major developed stock markets: A wavelet-based value at risk approach. *Emerging Markets*, volume 32, pp. 130-147.

Montgomery, D., Jennings, C., Kulahci, M., (2008). Introduction to Time Series Analysis and Forecasting. John Wiley & Sons, INC.

Nilsson, C. (2017). Forecasting Swedish Stock Market Volatility and Value-at-Risk: A Comparison of EWMA and GARCH Models. Department of Economics. Recuperado de <https://lup.lub.lu.se/student-papers/search/publication/8923808>

Patton, A., Ziegel, J. y Chen, R., (2019). Dynamic semiparametric models for expected shortfall (and Value-at-Risk). *Journal of Econometrics*, volume 211, Issue 2, pp. 388-413.

Pearson, N., (2002). Risk Budgeting: Portfolio Problem Solving with Value-at-Risk. Wiley Finance

Penza, P y Bansak, V.K., (2001). Measuring Market Risk with Value at Risk. John Wiley & Sons, Inc.

Pinheiro, D., Geraldo, E., Pacheco, W. y Alves, A., (2018). Methodology of Investment Risk Analysis for Wind Power Plants in the Brazilian Free Market. *Electric Power Components and Systems*, volume 46, Issue 3, pp. 316-330.

Poole, D., (2011). Álgebra Lineal. Una introducción moderna. Tercera Edición. CENGAGE Learning

Pritsker, M., (2001). The hidden dangers of historical simulation. *Journal of Banking & Finance*, volume 30, Issue 2, pp. 561-582

Procesos ARMA CIMAT. Recuperado de <https://www.google.com/search?q=jortega+ARMA&oq=jortega+ARMA&aqs=chrome..69i57j33.4925j0j7&sourceid=chrome&ie=UTF-8>

QRM_R_CODES: "04_Financial_Time_Series"- "04_GARCH_Estimation". Recuperado de https://gitlab.com/alejandro.gomez.henao/qrm_r_codes/tree/rmarkdown/code

Rahmanpour, M. y Osanloo, M., (2016). Determination of value at risk for long-term production planning in open pit mines in the presence of price uncertainty. *Journal of the*

Southern African Institute of Mining and Metallurgy. Department of Mining and Metallurgical Engineering, Amirkabir University of Technology, Tehran, Iran. Recuperado de http://www.scielo.org.za/scielo.php?script=sci_arttext&pid=S2225-62532016000300007

Rincón, L. (2006) Una introducción a la probabilidad y estadística. Departamento de Matemáticas. Facultad de ciencias UNAM. Recuperado de <http://lya.fciencias.unam.mx/lars/libros/pe-agosto-2006.pdf>

Rincón, L. (2007). Curso intermedio de probabilidad. Departamento de Matemáticas. Facultad de ciencias UNAM. Recuperado de <http://lya.fciencias.unam.mx/lars/Publicaciones/cip2010.pdf>

Rincón, L. (2014) Introducción a la probabilidad. Departamento de Matemáticas. Facultad de ciencias UNAM. Recuperado de <http://lya.fciencias.unam.mx/lars/Publicaciones/Prob1-2014.pdf>

Rozga, A. y Arneric, J., (2009) Dependence between volatility persistence, kurtosis and degrees of freedom. *Investigación Operacional*. Vol., 30, No.1, 32-39

Soporte de Minitab 18. Di Recuperado de sponible en <https://support.minitab.com/es-mx/minitab/18/help-and-how-to/statistics/basic-statistics/supporting-topics/basics/type-i-and-type-ii-error/>

Stavroyiannis, S., (2018). Value-at-Risk and related measures for the Bitcoin. *The Journal of Risk Finance*.

Taleb, N., (1997). Against Value-at-Risk: Nassim Taleb Replies to Philippe Jorion. Recuperado de <https://www.fooledbyrandomness.com/jorion.html>

Troster, V., Kumar Tiwari, A., Shahbaz, M. y Nicolás Macedo, D., (2019). Bitcoin returns and risk: A general GARCH and GAS analysis. *Finance Research Letters*, volume 30, pp. 187-193.

Tseng, J.J. y Li S.P., (2011) Asset returns and volatility clustering in financial time series. *Physica A: Statistical Mechanics and its Applications*

Weber, S., (2018). Solvency II, or how to sweep the downside risk under the carpet. *Insurance: Mathematics and Economics*, volume 82, pp. 191-200.

Woerner, S. y Egger, D., (2019). Quantum risk analysis. *npj Quantum Information*, volume 5. Recuperado de <https://www.nature.com/articles/s41534-019-0130-6>

Yu, W., Yang, K., Wei, Y. y Lei, L., (2018). Measuring Value-at-Risk and Expected Shortfall of crude oil portfolio using extreme value theory and vine copula. *Physica A: Statistical Mechanics and its Applications*, volume 490, pp. 1423-1433.

Zhang Y. y Nadarajah S.(2017). A review of backtesting for value at risk. *Communications in Statistics - Theory and Methods*