



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

MODELOS OCULTOS DE MARKOV PARA
SERIES DE TIEMPO FINANCIERAS: UNA
APLICACIÓN EN ÍNDICES BURSÁTILES

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
ACTUARIA

PRESENTA:
YADIRA RIVAS GODOY

DIRECTORA DE TESIS:
DRA. LIZBETH NARANJO ALBARRÁN

CIUDAD UNIVERSITARIA, CD. MX., 2019.





Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

1. Datos del alumno

Rivas

Godoy

Yadira

55 68 05 48 99

Universidad Nacional Autónoma de México

Facultad de Ciencias

Actuaría

312097895

2. Datos del tutor

Dra.

Lizbeth

Naranjo

Albarrán

3. Datos del sinodal 1

Dr.

Eduardo Arturo

Gutiérrez

Peña

4. Datos del sinodal 2

Dra.

Ruth Selene

Fuentes

García

5. Datos del sinodal 3

Dr.

Carlos Erwin

Rodríguez

Hernández-Vela

6. Datos del sinodal 4

M. en C.

Christian Gabriel

Miranda

Ruíz

7. Datos del trabajo escrito

Modelos ocultos de Markov para series de tiempo financieras: una aplicación en índices bursátiles

104 p

2019

*“Eines tages wirst du diese welt hinter dir lassen,
lebe also ein leben, an das du dich erinnern wirst.”*

Agradecimientos

Heme aquí, a punto de concluir una gran etapa, un enorme reto que logré gracias al apoyo de mi gente, familia y amigos que en todo momento me alentaron a seguir, por más complicado que fuera.

El mayor de mis agradecimientos y respeto es a mis padres, gracias por siempre estar y creer en mí, cuando ni siquiera yo lo hacía. Padre, todos los que te conocen saben que eres un hombre duro, pero sobre todo muy responsable y agradezco que siempre me has apoyado. Madre, ¿qué sería de mí sin ti? Simplemente, gracias por todo.

A mis hermanos Uriel y Marco, de los que he aprendido y a lo largo de la vida siempre me han cuidado. A mi familia, esa unión que tenemos me hace sentir fuerte, se que estarán siempre en las buenas y malas.

Tengo la fortuna de tener muy buenos amigos, hemos estado juntos desde la primaria y sin duda forman parte importante en mi vida: Narda, Pepe, Anyhe, Karen, gracias por escucharme, aconsejarme y en general por todo.

En el momento en que me propuse hacer tesis, vino a mi mente un solo nombre

para pedirle ser mi asesora y esa fuiste tú Lizbeth, muchas gracias por enseñarme, por confiar en mí e impulsarme, eres una parte fundamental en mi vida escolar y siempre estaré en deuda contigo. Gracias Liz, eres la mejor.

Ruth, sin duda alguna trabajar contigo es increíble, he aprendido mucho de ti, estoy muy agradecida contigo.

A mis sinodales, la Dra. Ruth Selene, al Dr. Eduardo Arturo, al Dr. Carlos Erwin y al M. en C. Christian Gabriel, trabajar con ustedes ha sido todo un placer, muchas gracias por sus correcciones y comentarios.

De igual forma, mi agradecimiento es a todos aquellos maestros, que a lo largo de mí vida me han transmitido sus conocimientos. A mis amigos de CCH, de la universidad, a mis alumnitos y en general a todos aquellos que han dejado algo en mí.

Por último, quiero agradecer a la Universidad Nacional Autónoma de México y en especial a la Facultad de Ciencias, por darme una profesión y ayudarme a desarrollarme como persona.

Yadira

1. Introducción	1
2. Series de tiempo financieras	4
2.1. Series de tiempo	4
2.1.1. Componentes de una series de tiempo	6
2.1.2. Series de tiempo financieras	6
2.2. Índices bursátiles	8
2.2.1. Índice de Precios y Cotizaciones	9
2.2.2. The Standard & Poor's 500	10
2.2.3. Deutscher Aktienindex	12
2.2.4. Nikkei Heikin Kabuka	14
3. Medidas de riesgo financiero	16
3.1. Valor en Riesgo	17
3.1.1. VaR por simulación histórica sin alisado	19
3.2. Medidas coherentes de riesgo	20
3.3. Valor en Riesgo Condicional	21
3.4. <i>Backtesting</i>	23
3.4.1. Prueba de Kupiec	24

4. Modelos ocultos de Markov	26
4.1. Introducción	27
4.1.1. Cadenas de Markov	27
4.2. Modelos ocultos de Markov	29
4.2.1. Elementos de un <i>HMM</i>	30
4.3. Problemas para el HMM	32
4.3.1. El problema de evaluación	32
Algoritmo Forward-Backward	33
4.3.2. El problema de decodificación	36
Algoritmo de Viterbi	37
4.3.3. El problema de aprendizaje	38
Algoritmo de Baum-Welch	39
4.4. Selección de modelos	42
4.4.1. Criterio de Información de Akaike	43
4.4.2. Criterio de Información Bayesiano	44
4.5. Predicción	44
5. HMM aplicado a series de tiempo financieras	46
5.1. Algoritmo	46
5.2. Índice de Precios y Cotizaciones	47
5.3. Standard & Poor's 500	54
5.4. Deutscher Aktienindex	61
5.5. Nikkei Heikin Kabuka	69
6. Conclusiones y trabajo futuro	78
6.1. Conclusiones	78
6.2. Trabajo futuro	81
A. Código de R	82
Bibliografía	93

Índice de figuras

4.1. Estructura de un <i>HMM</i>	30
5.1. IPC: serie original y transformada.	47
5.2. IPC: diagnóstico por estado.	49
5.3. IPC: probabilidad de estar en cada uno de los estados.	50
5.4. IPC: clasificación por estado.	50
5.5. IPC: clasificación en los precios al cierre.	51
5.6. IPC: predicción.	51
5.7. IPC: predicción y valor real.	53
5.8. IPC: métricas de riesgo.	53
5.9. S&P500: serie original y transformada.	54
5.10. S&P500: diagnóstico por estado.	56
5.11. S&P500: probabilidad de estar en cada uno de los estados.	57
5.12. S&P500: clasificación por estado.	57
5.13. S&P500: clasificación en los precios al cierre.	58
5.14. S&P500: predicción.	59
5.15. SP&500: predicción y valor real.	60
5.16. S&P500: métricas de riesgo.	61
5.17. DAX: serie original y transformada.	62

5.18. DAX: diagnóstico por estado.	64
5.19. DAX: probabilidad de estar en cada uno de los estados.	64
5.20. DAX: clasificación por estado.	65
5.21. DAX: clasificación en los precios al cierre.	66
5.22. DAX: predicción.	67
5.23. DAX: predicción y valor real.	68
5.24. DAX: métricas de riesgo.	68
5.25. Nikkei225: serie original y transformada.	69
5.26. Nikkei225: diagnóstico por estado.	72
5.27. Nikkei225: probabilidad de estar en cada uno de los estados.	72
5.28. Nikkei225: clasificación por estado.	73
5.29. Nikkei225: clasificación en los precios al cierre.	74
5.30. Nikkei225: predicción.	75
5.31. Nikkei225: predicción y valor real.	76
5.32. Nikkei225: métricas de riesgo.	76

Índice de tablas

2.1. Diferencia horaria entre índices bursátiles.	9
2.2. IPC: algunas empresas emisoras.	10
2.3. S&P500: algunas empresas emisoras.	11
2.4. DAX: algunas empresas emisoras.	13
2.5. Nikkei225: algunas empresas emisoras.	15
5.1. IPC: selección del mejor modelo.	48
5.2. IPC: descripción por estado del mejor modelo.	49
5.3. IPC: predicción.	52
5.4. IPC: prueba de Kupiec.	54
5.5. S&P500: selección del mejor modelo.	55
5.6. S&P500: descripción por estado.	56
5.7. S&P500: predicción.	60
5.8. S&P500: prueba de Kupiec.	61
5.9. DAX: selección del mejor modelo.	63
5.10. DAX: descripción por estado.	63
5.11. DAX: predicción.	67
5.12. DAX: prueba de Kupiec.	69
5.13. Nikkei225: selección del mejor modelo.	70

5.14. Nikkei225: descripción por estado.	71
5.15. Nikkei225: predicción.	75
5.16. Nikkei225: prueba de Kupiec.	77
6.1. Error cuadrático medio en la predicción.	79
6.2. Medidas de riesgo financiero.	80

CAPÍTULO 1

Introducción

Desde siempre, el trabajar con índices bursátiles ha representado un reto para expertos de la materia que buscan ajustar un modelo que haga buenas predicciones; regularmente ajustan un modelo de series de tiempo, por ejemplo los modelos ARCH o GARCH, entre otros. Sin embargo, existen modelos totalmente diferentes que se pueden adaptar a un conjunto de datos secuenciales y aprender de los mismos; estamos hablando de los modelos ocultos de Markov (*Hidden Markov Models, HMM*).

Un modelo oculto de Markov es un modelo gráfico probabilístico que representa un proceso dinámico. Se compone de dos procesos estocásticos: el primero corresponde a una cadena de Markov, en el cual se encuentran los estados ocultos que actúan como una sucesión de variables latentes que se adaptan a la naturaleza estocástica de las observaciones; y el segundo proceso justamente son las observaciones.

Estos modelos han tenido aplicaciones en diversas áreas ya que tienen una buena coincidencia con datos secuenciales; por ejemplo con los precios al cierre de los

índices bursátiles. Además los HMM representan un modelo novedoso en el área interdisciplinaria de las finanzas computacionales.

Es por ello que analizaremos cuatro de los principales índices bursátiles del mundo que son el *IPC* de México, el *S&P500* de Estados Unidos de América, el *DAX* de Alemania y el *Nikkei225* de Japón, con un periodo de entrenamiento de enero de 1992 a diciembre del 2017, y un conjunto de evaluación de enero del 2018 a abril del 2019.

El objetivo será ajustar un HMM, analizarlo, predecir el valor del índice y calcular el valor en riesgo, mejor conocido como VaR_α , que es un método para cuantificar la exposición al riesgo de mercado y esto para analizar si el aplicar un HMM sería una buena alternativa para el mundo laboral.

El presente trabajo está compuesto de 4 capítulos, más la introducción y las conclusiones.

En el capítulo 2 veremos a grandes rasgos qué son las series de tiempo financieras, y en particular describiremos los aspectos más importantes de los índices con los que estamos trabajando, al igual que la transformación *log return* que se le aplicó a todos los datos.

El capítulo 3 contiene los aspectos más importantes de las medidas de riesgo financiero en las que entraremos a detalle en la teoría y cálculo del VaR_α y también del $CVaR_\alpha$, que es una medida complementaria del valor en riesgo ya que ésta no es una medida coherente de riesgo, al no cumplir la propiedad de subaditividad; de lo que hablaremos más adelante. De igual forma describiremos la prueba de hipótesis no paramétrica de Kupiec, la que permite determinar si la proporción de observaciones que exceden el VaR_α de un modelo es consistente con la proporción esperada.

En el capítulo 4 encontraremos toda la información relacionada con el modelo oculto de Markov, empezando por una introducción de conceptos de procesos estocásticos que necesitaremos para entender el modelo. Seguiremos con la descripción de los elementos de un HMM, hasta llegar a los tres grandes problemas de este modelo, que son: la evaluación, la decodificación y el aprendizaje. Después de cada uno describiremos la forma de solucionar el problema, que será con tres algoritmos: el algoritmo Forward-Backward, el algoritmo de Viterbi y el algoritmo de Baum-Welch que es una forma particular del algoritmo EM. Como siguiente punto veremos criterios de bondad de ajuste para elegir el mejor modelo, y por último observaremos la predicción con ese modelo.

En el capítulo 5 presentaremos los resultados de aplicar el HMM a los cuatro índices bursátiles. Comenzaremos describiendo los datos al cierre y con la transformación *log return*, después daremos las estimaciones de los parámetros que ajustaron ese modelo, seguido de la descripción por estado junto con su diagnóstico, y analizaremos la decodificación local y global por estado. Seguiremos con la predicción que compararemos con el conjunto de prueba, y como último punto obtendremos el VaR_α , el $CVaR_\alpha$ y el estadístico de Kupiec.

Por último, en el capítulo 6 presentaremos las conclusiones a las que llegamos y una propuesta de trabajo futuro.

CAPÍTULO 2

Series de tiempo financieras

En este capítulo, analizaremos los aspectos más generales de una serie de tiempo financiera, daremos además una breve descripción de los índices bursátiles que se analizarán en capítulos posteriores.

El estudio de las series de tiempo será muy importante ya que puede ayudar a obtener pronósticos de valores futuros de la serie y también ayuda a entender el proceso de generación de datos.

2.1. Series de tiempo

Una Serie de Tiempo es una sucesión de observaciones correspondientes a una variable tomada en periodos de tiempo regulares y de duración constante; estos periodos suelen ser semanales, mensuales, trimestrales o anuales.

Entre los modelos de series de tiempo, tenemos los tradicionales en los que encontramos los modelos de medias móviles (*MA*), los modelos autorregresivos (*AR*) y los modelos autorregresivos de medias móviles (*ARMA*); sin embargo, última-

mente los modelos autorregresivos integrados de medias móviles (ARIMA) se han popularizado, debido a su utilidad para hacer predicciones, y puede generalizarse para tomar en cuenta el efecto de estacionalidad, y lo conocemos como (*SARIMA*). Puede encontrar información a detalle de cada modelo en [Chatfield \(2016\)](#), [Shumway & Stoffer \(2017\)](#), [Brockwell et al. \(2002\)](#), [Kirchgässner & Wolters \(2007\)](#).

Las series de tiempo tienen una infinidad de aplicaciones en diversas áreas, por ejemplo:

- Economía

Ventas mensuales de un producto en una empresa en los últimos 4 años.

- Demografía

La población década por década de un estado en el siglo anterior.

- Meteorología

Temperatura máxima diaria, en la alcaldía Coyoacán, los últimos 6 meses.

- Finanzas

Los índices bursátiles diarios de los últimos 10 años.

Esta última será de nuestro interés, dado que a lo largo de este trabajo estudiaremos índices bursátiles que se describirán más adelante.

El objetivo del estudio de una serie de tiempo puede ser a nivel de descripción, en el que se utilizan medidas descriptivas para dar un panorama general de la serie; a nivel de explicación, que es cuando las observaciones se toman de varias variables y busca explicar la variación en otra serie; a nivel de predicción, en el que se usan los valores observados para predecir los valores futuros de una serie de tiempo; y a nivel de control que es donde las series de tiempo se analizan para mejorar el control sobre algún sistema específico.

2.1.1. Componentes de una series de tiempo

El análisis de series de tiempo es interesante, ya que podemos encontrar diversos problemas, por ejemplo: en los datos se nos pueden presentar ciclos, que los datos sigan cierta tendencia, movimientos estacionales, entre otros, y debemos ser capaces de considerar cualquier imprevisto para que al momento de hacer una predicción, ésta sea buena.

Las series de tiempo pueden estar caracterizadas por distintos componentes:

- **Tendencia:** Se define como una función T_t del tiempo t que describe la evolución lenta y a largo plazo del nivel medio de la serie. Éste es un componente de dirección, que puede presentarse de forma creciente, decreciente o ambas.
- **Ciclos:** Es un conjunto de fluctuaciones en torno a la dirección de la tendencia.
- **Estacional:** Es un movimiento periódico s_t en el tiempo t a corto plazo y recoge las oscilaciones que se producen en esos periodos de repetición.
- **Aleatoriedad:** Describe el comportamiento irregular que está compuesto por fluctuaciones causadas por sucesos impredecibles.

2.1.2. Series de tiempo financieras

Este tipo de series de tiempo, son una sucesión de observaciones X_t en un periodo de tiempo determinado, los cuales pueden ser, por ejemplo, precios, rendimientos de activos financieros o tasas de interés; la volatilidad es una característica inherente de las mismas. Cabe destacar que la volatilidad es la desviación estándar de los rendimientos del proceso, ver [Kim et al. \(1998\)](#).

El problema con este tipo de series de tiempo es que por lo general los modelos tradicionales no se ajustan de forma correcta, ya que se tiene una amplia variedad, poca disponibilidad, diferente frecuencia y sobre todo una alta volatilidad de los

datos. Entonces una opción es ajustar un modelo de heteroscedasticidad condicional tal como un *GARCH*, *EGARCH*; puede encontrar la descripción de cada uno en [Francq & Zakoian \(2011\)](#).

Cuando se trabaja con datos financieros, regularmente se recurre a utilizar el *return* y hay más de una razón para utilizarlo; por ejemplo, para el inversionista promedio, los mercados financieros se consideran cerrados o competitivos, entonces lo utilizan para que el tamaño de la inversión no afecte las variaciones de precios. Por otro lado los retornos presentan características estadísticas más atractivas que los precios como la estacionariedad. Existen varias transformaciones; sin embargo, todas parten de obtener los rendimientos.

Sea P_t el precio de un activo al tiempo t ; el rendimiento se define como:

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}}.$$

Pero esta fórmula es para un solo periodo; para una serie de tiempo financiera de k períodos, el rendimiento al final del período es un rendimiento acumulado que se compone de todos los períodos, como se muestra a continuación:

$$\begin{aligned} 1 + P_t[k] &= (1 + R_1)(1 + R_2) \dots (1 + R_k) \\ &= \prod_{t=1}^k (1 + R_t). \end{aligned}$$

Ahora vamos a introducir otro concepto que es el logaritmo natural del rendimiento simple y se denomina *Continuously Compounded Return* o simplemente *log return* y se define como:

$$r_t = \log(1 + R_t) = \log \left[\frac{P_t}{P_{t-1}} \right].$$

Más adelante veremos cómo se ven las gráficas de las series de tiempo de los índices bursátiles que utilizaremos junto con la gráfica del *log return*.

Para más información de esta transformación dentro de las series de tiempo financieras, ver [Campbell et al. \(1997\)](#).

2.2. Índices bursátiles

Un índice bursátil es el valor promedio de un conjunto de acciones determinado. La finalidad de estos índices es evidenciar los cambios en el tiempo de los precios de las acciones que lo componen. Usualmente, estas acciones comparten ciertas características como pertenecer a una misma bolsa de valores o pertenecer a la misma industria. La importancia de estos índices radica en que es un referente para que un inversionista decida invertir o no en alguna bolsa, en particular poder decidir en qué país comprar acciones y en cuál no. Hay toda una teoría de estos índices; para más información se recomienda leer [Stoll & Whaley \(1990\)](#).

Vamos a estudiar 4 diferentes índices bursátiles, el *IPC*, el *S&P500*, el *DAX* y el *Nikkei225*. La información de los índices será obtenida del primero de enero de 1992 al primero de enero del 2018, por lo que abarcamos un periodo de 26 años o 9497 días; sin embargo, recordemos que estos índices no son calculados en fines de semana, ni en días feriados, por lo que contamos con diferente número de datos dependiendo el índice que se exhiben a continuación:

- *IPC*: Se cuentan con 6507 datos.
- *S&P500*: Se cuentan 6550 datos.
- *DAX*: Se cuentan con 6581 datos.
- *Nikkei225*: Se cuentan con 6392 datos.

Es por ello que tomaremos sólo los precios de final de mes para tener en todos los casos 312 valores y con ello, al sacar los *log return*, tendremos 311 valores que son con los que trabajaremos. Los datos son de dominio público y se pueden descargar desde el software [R Core Team \(2017\)](#) con el paquete *stockSymbols* (*TTR*).

Otro detalle que debemos tomar en cuenta es el desfase en horario de los índices. Tomando como referencia a México la diferencia horaria es:

Mercado	Hora local	Hora México	Diferencia horaria
México	8:30 am - 15:00 pm	8:30 am - 15:00 pm	–
Nueva York	9:30 am - 16:00 pm	8:30 am - 15:00 pm	–
Frankfurt	9:00 am - 17:30 pm	2:00 am - 10:30 am	6.5 horas
Tokio	9:00 am - 15:00 pm	19:00 pm - 1:00 am	10.5 horas

Tabla 2.1: Diferencia horaria entre índices bursátiles.

La Tabla 2.1 es importante porque podemos ver la hora de apertura y cierre de las bolsas; en este horario es donde se efectúan las renegociaciones de compra y venta de acciones. La mayor volatilidad se produce al cierre ya que se producen cambios bruscos porque una vez que se cierra el mercado ya no hay operaciones hasta el día siguiente. Otro detalle es que en Japón al igual que en Hong Kong se toman una hora de almuerzo a la mitad de la operación, es decir abren la bolsa de 9:00 am a 11:30 am y de 12:30 pm a 15:00 pm.

2.2.1. Índice de Precios y Cotizaciones

El Índice de Precios y Cotizaciones conocido como IPC es el principal índice de la Bolsa Mexicana de Valores (BMV), que expresa el rendimiento del mercado accionario mexicano; este tipo de índice es importante ya que refleja cuánto gana o pierde la BMV y eso es un referente mundial. El IPC comenzó a ser registrado el 30 de octubre de 1978 y en él cotizan las 35 emisoras más bursátiles que contribuyen a la [Bolsa Mexicana de Valores \(2016\)](#). El IPC se calcula de la siguiente manera:

$$I_t = I_{t-1} \left(\frac{\sum_{i=1}^n P_{it} \times (Q_{it} \times FAF_i)}{\sum_{i=1}^n P_{i,t-1} \times (Q_{i,t-1} \times FAF_i) \times f_{i,t-1}} \right),$$

donde:

I_t corresponde al IPC al día t .

P_{it} corresponde al precio de las acciones de la empresa i al día t .

Q_{it} corresponde al número de acciones inscritas de la empresa i al día t .

FAF_i corresponde al factor de ajuste por acciones flotantes de la empresa i .

f_i corresponde al factor de ajuste por ex-derechos de la empresa i al día t .

++

En la Tabla 2.2 veremos algunas empresas emisoras en el IPC; el resto puede encontrarlos en [Bolsa Mexicana de Valores \(2018\)](#).

Clave	Nombre
ALFA	Alfa, S.A.B. de C.V.
ALSEA	Alsea, S.A.B. de C.V.
AMX	America Movil, S.A.B. de C.V.
BIMBO	Grupo Bimbo, S.A.B. de C.V.
BOLSA	Bolsa Mexicana De Valores, S.A.B. de C.V.
CEMEX	Cemex, S.A.B. de C.V.
CHDRAUI	Grupo Comercial Chedraui, S.A.B. de C.V.
ELEKTRA	Grupo Elektra, S.A.B. de C.V.
FEMSA	Fomento Económico Mexicano, S.A.B. de C.V.
GENTERA	Compartamos, S.A.B. de C.V.

Tabla 2.2: IPC: algunas empresas emisoras.

2.2.2. The Standard & Poor's 500

El siguiente índice bursátil que vamos a estudiar es el *Standard & Poor's 500* (S&P500) que es uno de los índices más importantes de Estados Unidos. Se basa en la capitalización de mercado de 500 empresas que cotizan en *New York Stock Exchange (NYSE)*. El S&P500 comenzó a registrarse el 4 de junio de 1968; este índice es de suma importancia a nivel mundial ya que juega en la plaza financiera más importante del mundo, es decir Wall Street; ver [Indices \(2019\)](#).

El cálculo es sencillo y se realiza de la siguiente manera:

$$Index = \frac{\sum_{i=1}^n (P_i * Q_i)}{DIVISOR},$$

donde:

P_i corresponde al precio de cada acción.

Q_i corresponde al número de acciones disponibles públicamente para cada acción.

DIVISOR corresponde a una cifra patentada por *Standard & Poor's*, que se ajusta dependiendo de si se trata de divisiones de acciones, dividendos especiales o una escisión, esto con el fin de que no se altere el índice por factores no económicos. La mayor parte de las fuentes toman este valor como 8.9 mil millones.

Una de las condiciones para formar parte de este índice es tener una capitalización de 4 mil millones de dólares o más, un volumen promedio de 250 mil o más acciones negociadas por un período de seis meses antes de su fecha de evaluación y estar basada en los Estados Unidos, por lo que estar dentro del índice representa un reto para las empresas y por lo mismo la duración de las acciones suele ser por períodos cortos. Un detalle interesante es que de las 500 empresas que iniciaron a cotizar, más de 414 han salido.

Clave	Nombre
AAPL	Apple Inc.
T	AT&T Inc.
KO	The Coca-Cola Company
K	Kellogg Co.
WHR	Whirlpool Corp.
DIS	The Walt Disney Company
BAC	Bank of America Corp
FBHS	Fortune Brands Home & Security
AAL	American Airlines Group
FB	Facebook, Inc.

Tabla 2.3: S&P500: algunas empresas emisoras.

En la Tabla 2.3 vemos una lista de algunas empresas integrantes del índice S&P500, el resto se puede encontrar en [Insider \(2018\)](#).

A pesar de tener 500 integrantes, hay 5 emisoras que explican más del 75%

del rendimiento del índice en 2018 y son: Amazon, Apple, Microsoft, Netflix e Intel. Notemos que todos forman parte del sector tecnológico, lo que nos muestra su poderío en el mercado.

2.2.3. Deutscher Aktienindex

Otro índice que vamos a analizar será el *DAX*, que es el índice bursátil más importante de Alemania y que cotiza en la bolsa de Fráncfort (*Frankfurter Wertpapierbörse*); se compone de las 30 empresas más bursátiles de Alemania. Es un índice extremadamente líquido, con pautas y reglas transparentes que están disponibles para todos, escritas de manera clara y entendible, que puede ver en [Börse \(2018\)](#).

Comenzó a ser registrado el 1 de julio de 1988, y desde el 21 de junio de 1999 las operaciones comenzaron a realizarse en una plataforma electrónica llamada *XETRA* que es un sistema caracterizado por su elevada transparencia, flexibilidad y liquidez. Este sistema se actualiza cada segundo mientras el mercado está abierto. Para la elaboración del índice se utilizan los precios registrados de la plataforma *XETRA* con la condición de que la ponderación de cada emisora no sea mayor del 10%.

Para el cálculo del *DAX* se utiliza la fórmula del índice de Laspeyres que toma en cuenta precios y cantidades; sin embargo, tienen una modificación para no sobreestimar la inflación. Tomemos en cuenta que sólo se consideran las acciones en *free float*, que son las que están en libre circulación disponibles al público, es decir, del total de acciones en circulación se le restan las acciones restringidas. Con ello el cálculo del índice se ve de la siguiente manera:

$$Index_t = K_t \left(\frac{\sum_{i=1}^n p_{it} \times f_{iT} \times q_{iT} \times c_{it}}{\sum_{i=1}^n p_{i0} \times q_{i0}} \times Base \right),$$

donde:

c_{it} corresponde a el factor de ajuste de la empresa i al tiempo t .

ff_{iT} corresponde al factor *free float* de la clase de acciones i al tiempo T .

n corresponde al número de acciones en el índice.

p_{i0} corresponde al precio de cierre de la acción i del día de negociación de su inclusión en el índice.

p_{it} corresponde al precio de la acción i al tiempo t .

q_{i0} corresponde al número de acciones de la empresa i del día de negociación de su inclusión en el índice.

q_{iT} corresponde al número de acciones de la empresa i al tiempo T .

K_T corresponde al factor de encadenamiento a partir del tiempo T .

Base corresponde a una constante, que aparece en las especificaciones del índice, que tiene que ver con los puntos base.

En la Tabla 2.4 veremos las primeras empresas emisoras del DAX, el resto de la lista puede encontrarlo en [Frankfurt \(2018\)](#), junto con su ponderación.

Clave	Nombre
SAP	SAP SE
SIE	Siemens AG
BAYN	Bayer AG
BAS	BASF SE
ALV	Allianz SE
DAI	Daimler AG
DTE	Deutsche Telekom AG
ADS	Adidas AG
LIN	Linde plc
VOW3	Volkswagen AG

Tabla 2.4: DAX: algunas empresas emisoras.

En este índice se encuentran empresas representantes del sector automovilístico, farmacéutico, industria química, aseguradoras e incluso energéticas, entre otros que impulsan la economía alemana.

2.2.4. Nikkei Heikin Kabuka

Por último, vamos a estudiar el principal índice bursátil del mercado japonés, llamado *Nikkei Heikin Kabuka*, con su traducción oficial al inglés como *Nikkei Stock Average* o simplemente *Nikkei225*. Este índice es uno de los principales de la Bolsa de Tokio, que es la tercera bolsa más importante del mundo y la más importante de Asia. Se calculó por primera vez en 1971 y se compone por 225 emisoras altamente líquidas, con el fin de que tenga continuidad a largo plazo y por ello son muy estrictos con las emisoras; puede ver todas las reglas de ingreso y mucha más información en [Nikkei \(2017\)](#).

El Nikkei es ponderado por precios, y no por la capitalización, lo calcula el periódico japonés económico y de negocios *Nihon Keizai Shinbun* y se hace de la siguiente manera:

$$Nikkei225 = \frac{\sum_{i=1}^n PA_i}{Divisor},$$

donde:

Divisor corresponde a un denominador que ha sido ajustado a lo largo del tiempo para asegurar la continuidad del índice.

PA_i corresponde al precio de acciones ajustado de la empresa i y se calcula como

$$PA_i = \frac{P_i * 50(\text{yen})}{VN},$$

donde:

P_i corresponde al precio de la acción i .

VN corresponde al valor nominal del yen.

Notamos que el cálculo del índice tiene similitudes con el S&P500 en la parte del *DIVISOR* y eso es porque para el Nikkei225 se utilizó la misma metodología que los editores del *Wall Street Journal*; sin embargo, en ambos casos no se sabe su valor exacto, la finalidad de esta cantidad es la de estabilizar el índice, lo cual puede ayudar pero también hace que el cálculo de los índices no sea posible

analíticamente para el público.

En la Tabla 2.5 veremos un listado de emisoras del Nikkei225. Notemos que a diferencia de los demás índices las claves se componen de números; para ver los demás integrantes junto con su ponderación y separado por sector vea [Site \(2018\)](#).

Clave	Nombre
4503	Astellas Pharma Inc.
6503	Mitsubishi Electric Corp.
6752	Panasonic Corp.
7201	Nissan Motor Co. Ltd.
7267	Honda Motor Co. Ltd.
6758	Sony Corp.
7731	Nikon Corp.
9432	Nippon Telegraph & Telephone Corp.
8411	Mizuho Financial Group, Inc.
8628	Matsui Securities Co.Ltd.

Tabla 2.5: Nikkei225: algunas empresas emisoras.

CAPÍTULO 3

Medidas de riesgo financiero

Una parte importante dentro del análisis financiero es medir el riesgo en el que podría incurrir el instrumento. Vamos a entender al riesgo como la exposición a la posible pérdida, daño, o alguna otra circunstancia desafortunada.

Existen varios tipos de riesgo, por ejemplo el de mercado, el de crédito, de liquidez, operativo, entre otros; sin embargo, nos centraremos exclusivamente en el riesgo de mercado, que de acuerdo con los acuerdos en *Basilea II*, se define como la probabilidad de sufrir pérdidas en posiciones dentro y fuera del balance a causa de oscilaciones en los precios de mercado. Los riesgos sujetos a este requerimiento de capital son los siguientes:

- Riesgos inherentes a las acciones y a instrumentos relacionados con los tipos de interés en la cartera de negociación.
- Riesgo de divisas.
- Riesgo de productos básicos en todo el banco.

Puede encontrar más información en [Banco de Pagos Internacionales \(2004\)](#).

El objetivo será maximizar los rendimientos y minimizar el riesgo de pérdida. Puede encontrar la definición del resto de los riesgos así como sus características en [Crouhy \(2005\)](#).

Una medida de riesgo es una función $\rho : V \rightarrow \mathbb{R}$, donde V es un espacio de probabilidad de factores de riesgo de una entidad.

En 1952, Harry Markovitz propuso utilizar la varianza de los rendimientos de los activos como medida de riesgo y fue aceptada hasta finales de la década de los ochentas y principio de los noventas, cuando finalmente se hizo evidente que ésta es una medida de incertidumbre y no de riesgo.

La razón más importante por lo que ya no se utilizó la varianza de los rendimientos de los activos fue por la gran crisis financiera de 1987 en Estados Unidos. Esto provocó pérdidas muy grandes, y los llevó a suponer que la medida de riesgo debía tomar en cuenta las pérdidas potenciales y la probabilidad de ocurrencia. Es por ello que JP Morgan propuso el VaR_α . Sin embargo, como veremos más adelante, ni la varianza de los rendimientos de los activos ni el VaR_α , son medidas coherentes de riesgo y por lo tanto es necesario considerar otras como el *expected shortfall*, mejor conocido como $CVaR_\alpha$.

3.1. Valor en Riesgo

El valor en riesgo VaR_α , es una medida de riesgo que cuantifica y resume todo el riesgo de una cartera en una sola cifra; es por ello que es la más utilizada ya que se expresa en términos de dinero perdido, lo cual resulta ser simple y universal para los involucrados.

Formalmente, como lo define [McNeil et al. \(2005\)](#), si se conoce la distribución de pérdidas de la cartera L y dado un nivel de confianza $\alpha \in (0, 1)$, el VaR_α se define como:

$$VaR_{\alpha}(L) = \inf\{l \in \mathbb{R} : \mathbb{P}(L > l) \leq 1 - \alpha\},$$

que es equivalente a:

$$VaR_{\alpha}(L) = \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\},$$

donde $F_L(\cdot)$ es la función de distribución de probabilidad acumulada de las pérdidas de la cartera. En términos probabilísticos, el VaR_{α} es un cuantil de la distribución de pérdidas y contesta a la pregunta, ¿cuál es la mínima pérdida en que incurrimos en los $100(1-\alpha)\%$ peores casos de la cartera?

Por ejemplo, si tuviéramos un $VaR_{95\%}=4400$, calculado en un horizonte temporal de 10 días, significa que la pérdida de la posición no debería ser superior a 4400 pesos en 95 de 100 casos a lo largo de los diez días. Ahora bien, si bajo las mismas condiciones tuviéramos un $VaR_{95\%}=-4400$, significa que el beneficio de la posición no debería ser superior a 4400 pesos en 95 de 100 casos a lo largo de los diez días. Lo que significa que si el VaR es positivo implica una pérdida, pero si es negativo implica una ganancia.

En términos generales, para estimar la función de probabilidad y la medida de riesgo deseada, en este caso el VaR_{α} , es necesario seguir cuatro etapas, según [Jorion \(2003\)](#):

1. Identificar los factores de riesgo que pueden influir en el valor de mercado del portafolio de inversión.
 2. Estimar la distribución de probabilidad de los cambios de los factores de riesgo que podrían ocurrir durante el horizonte de inversión.
 3. Construir la distribución de probabilidad de los cambios en el valor de mercado del portafolio, a partir de la combinación de las distribuciones de probabilidad estimadas en la fase anterior.
-

4. Calcular la medida de riesgo deseada de las posiciones individuales y de todo el portafolio de inversión.

En función de los supuestos y alcances que se consideran para realizar las fases anteriores, podemos calcular el valor en riesgo de diferentes maneras. Iniciaremos a partir de de la siguiente clasificación:

- Modelos de portafolio o de varianza-covarianza.
 - Delta-Normal.
 - Delta-Gamma.
- Simulación.
 - Histórica.
 - Monte Carlo.
- Otros modelos.

La clasificación podría ser mucho más extensa, incluyendo modelos particulares; sin embargo, no es el objetivo de este trabajo pero puede encontrar información a detalle en [Jorion \(2003\)](#).

Por otra parte, si en el cálculo del VaR se le da el mismo peso a las observaciones durante todo el periodo se le conoce como VaR sin alisado.

Para efectos de esta tesis calcularemos el valor en riesgo por medio de simulación histórica sin alisado, ya que es el más utilizado en el mercado.

3.1.1. VaR por simulación histórica sin alisado

Este método para calcular el VaR_α consiste en generar escenarios de los factores de riesgo a partir de la información observada en un determinado número de días. Se calcula como el α -ésimo cuantil de la distribución empírica de pérdidas y

ganancias ($P&L$).

Veamos los pasos a seguir para calcular el VaR_α bajo este método:

1. Debemos tener información diaria del activo y tomaremos el precio de cierre como el precio representativo de ese día.
2. Calculamos los rendimientos diarios.
3. Si tenemos una muestra de n observaciones de precios, tendremos $n - 1$ rendimientos. Entonces calculamos la rentabilidad, que será el rendimiento al día $t - 1$ por la posición del día t y esto lo hacemos para toda la muestra.
4. Basta calcular el cuantil deseado para obtener el VaR_α .

El atractivo de este método es que el cálculo es muy simple; además no tenemos que hacer ningún supuesto de normalidad en los datos, tampoco es necesario calcular varianzas ni covarianzas. En cambio, esta alternativa puede no ser la mejor cuando se tiene información histórica limitada, pero no es nuestro caso.

3.2. Medidas coherentes de riesgo

Durante años, el VaR_α fue la solución de los expertos de Wall Street; sin embargo, no se percataron de que esa medida sólo mide probabilidades, entonces al tomar el 99 % de confianza dejan el 1 % a la deriva y es ahí donde tuvieron pérdidas millonarias, ya que pasó lo que creían o esperaban que no pasaría. Entonces fue evidente que necesitaban tener otra medida que fuera mucho más robusta; entonces Artzner et al. (1999), describieron formalmente algunas características que debían de satisfacer las medidas de riesgo y definieron lo que es una medida coherente de riesgo.

Sea $\rho : V \rightarrow \mathbb{R}$ una medida de riesgo. Decimos que ρ es una medida de riesgo coherente si cumple lo siguiente:

- *Monotonía*: Para toda $X, Y \in V$ tal que $X \leq Y$, donde “ \leq ” representa el orden estocástico; se cumple casi seguramente que $\rho(X) \leq \rho(Y)$.
- *Homogeneidad positiva*: Para todo $\lambda \geq 0$, y $X \in V$, $\rho(\lambda X) = \lambda \rho(X)$. Es decir, aumentar el tamaño de una cartera por λ debería simplemente escalar su riesgo por el mismo factor.
- *Invarianza ante traslaciones*: Para todo $X \in V$, y todo $\zeta \in \mathbb{R}$, $\rho(X + \zeta) = \rho(X) + \zeta$. Es decir, una medida de riesgo invariante ante traslaciones implica que si tenemos un activo con riesgo y formamos una cartera agregando un activo libre de riesgo, el riesgo de la cartera disminuye en la cantidad del mismo.
- *Subaditividad*: Para todo $X, Y \in V$, $\rho(X + Y) \leq \rho(X) + \rho(Y)$. Es decir la fusión de carteras no puede aumentar el riesgo.

En particular, la última hace mención al principio de diversificación; además garantiza que se puede encontrar una combinación óptima dentro del portafolio tal que minimice el riesgo por su relación con la convexidad de la medida de riesgo.

Sin embargo, a pesar de que el VaR es la medida que más se ocupa en el mercado, no cumple con ser una medida coherente de riesgo tal y como lo exhiben Artzner et al. (1999). Es por ello que se reformuló el VaR_α , buscando una medida sustituta para la cuantificación del riesgo de mercado y es como nació el $CVaR_\alpha$, del que hablaremos a continuación.

3.3. Valor en Riesgo Condicional

El valor en riesgo condicional ($CVaR_\alpha$), también conocido como *Expected Shortfall* (ES_α), es una medida complementaria del VaR_α , que satisface las propiedades necesarias para ser una medida de riesgo coherente tal y como expone Pflug (2000). En la práctica, el $CVaR_\alpha$ puede ser entendido como el valor esperado de los $100(1 - \alpha)\%$ peores casos y puede ser calculado a través del promedio

de los valores que exceden el VaR con un nivel de confianza α .

Para una distribución de pérdidas L , con $E(|L|) < \infty$ y función de distribución F_L , el $CVaR_\alpha$ a un nivel de confianza α se define como:

$$CVaR_\alpha(L) = \frac{1}{1-\alpha} \int_\alpha^1 q_u(F_L) du,$$

donde:

$$q_u(F_L) = \inf\{x \in \mathbb{R} : F_L(x) \geq \alpha\}.$$

Por lo que el $CVaR_\alpha$ se reduce a

$$CVaR_\alpha(L) = \frac{1}{1-\alpha} \int_\alpha^1 VaR_u(F_L) du,$$

y de esa forma llegamos a la expresión más usada

$$CVaR_\alpha(L) = E(L|L \geq VaR_\alpha). \quad (3.1)$$

Esta última expresión fue discutida detalladamente por [Acerbi & Tasche \(2002\)](#). El $CVaR_\alpha$ debería ser la que se utilice por excelencia; sin embargo, no se ha aceptado de la mejor forma dado que se necesita un requerimiento mayor de capital respecto al VaR_α , esto se debe a que por construcción el $CVaR_\alpha(Y) \geq VaR_\alpha(Y)$. Para mayor información de la optimización del $CVaR_\alpha$, consulte [Rockafellar & Uryasev \(2000\)](#).

Ahora, para el cálculo del $CVaR_\alpha$ basta con calcular el VaR_α para poder utilizar la ecuación (3.1) y así obtener el valor. Recordemos que para efectos de esta tesis lo calcularemos por simulación histórica sin alisado, por lo que una vez que aplicamos los pasos descritos en la sección 3.1.1, calcularemos el $CVaR_\alpha$.

3.4. *Backtesting*

Como lo describimos anteriormente, hay muchas técnicas diferentes de obtener el VaR_α ; entonces la pregunta es ¿cómo saber cuál es mejor?, y en nuestro caso ¿cómo poder asegurar que la técnica que elegimos es buena? Es por ello que debemos poner a prueba nuestro resultado a través de una prueba de *backtesting*.

El *backtesting* es un procedimiento estadístico utilizado para validar la calidad y precisión del VaR_α , a través de la comparación de los resultados reales de las posiciones y las medidas de riesgo. Es muy importante que se realice esta prueba ya que por medio de ella se verifica periódicamente el modelo que se está utilizando y si no la pasa se debe recalibrar o cambiar.

Los pasos para obtener esta prueba son:

1. Calcular pérdidas y ganancias.
2. Comparar el VaR_α observado ajustado a un día con las pérdidas y ganancias diarias.
3. Calcular cuántas veces las pérdidas y ganancias exceden el VaR_α observado; lo llamaremos errores.
4. Calcular la eficiencia, que es el número de errores entre el número de observaciones.

Lo ideal sería que la eficiencia estuviera cerca al nivel de significancia propuesto, y para asegurarnos de ello lo probaremos con una de las pruebas más utilizadas del *backtesting*, que es la prueba de *Kupiec*.

Puede revisar más información del *backtesting* en [De Lara Haro \(2005\)](#) y por supuesto el marco de supervisión del *backtesting* de Basilea en [Basel Committee on Banking Supervision \(1996\)](#).

3.4.1. Prueba de Kupiec

Paul H. Kupiec en 1995 publicó la prueba de proporción de fallas; este método mide si el nivel de significancia propuesto para el VaR es consistente con la proporción de fallas que presenta el modelo. Es decir, lo que hace la prueba es modelar diferencias entre los resultados reales y el resultado del VaR_α .

La hipótesis nula de la prueba de Kupiec es:

$$H_0 : \rho = \frac{e}{N},$$

donde:

ρ corresponde a la probabilidad de falla.

e corresponde al número de errores.

N corresponde al número de observaciones.

El estadístico de prueba es:

$$L_{Kupiec} = -2 \log \left(\frac{(1 - \rho)^{N-e} \times \rho^e}{\left[1 - \left(\frac{e}{N}\right)\right]^{N-e} \times \left(\frac{e}{N}\right)^e} \right),$$

que se construye como un cociente de verosimilitudes y bajo H_0 , L_{Kupiec} se distribuye asintóticamente como una χ_1^2 . La regla de decisión nos dice que si el valor de la estadística de Kupiec es mayor que el valor crítico de la Ji-cuadrada, se rechazará la hipótesis nula y con ello se concluye que el modelo es incorrecto.

De igual forma nos basaremos en el p-valor que es la probabilidad, cuando suponemos la hipótesis nula cierta, de obtener un resultado muestral tan ó más extremo que el resultado muestral observado. Este valor también llamado valor p , tiene la siguiente regla de decisión, si el p-valor de una prueba es menor que el nivel de significancia κ se rechazaría la hipótesis nula. Entonces esperaríamos un p-valor mayor que κ ya que nos interesa no rechazar H_0 . Utilizaremos este valor

para contrastar la hipótesis y no tener dudas del resultado del estadístico de Kupiec.

Cabe destacar que ésta no es la única prueba de *backtesting*; sin embargo, al tener bastante información, lo más recomendable es utilizarla. Para mayor detalle puede consultar el artículo original en [Kupiec \(1995\)](#).

CAPÍTULO 4

Modelos ocultos de Markov

En este capítulo empezaremos con el estudio del modelo oculto de Markov (*Hidden Markov Model, HMM*). Este tipo de modelos fue estudiado por primera vez en una serie de artículos publicados por *Leonard E. Baum* y sus compañeros del Instituto de Análisis de Defensa (IDA) en Princeton, Nueva Jersey, a fines de los años sesenta. Sin embargo, fue varios años después que fue publicada la primera aplicación del modelo y fue en reconocimiento de voz, publicado por *Rabiner (1989)*, a partir del cual se han utilizado ampliamente los modelos ocultos de Markov en varias aplicaciones, como por ejemplo: reconocimiento de actividad de video, (*Niu & Abdel-Mottaleb 2005*); búsqueda de genes, (*Lukashin & Borodovsky 1998*); rastreo de gestos manuales, (*Chen et al. 2003*); e incluso para predecir el comportamiento de acciones del mercado de valores, (*Gupta & Dhingra 2012, Hassan & Nath 2005*).

Comenzaremos definiendo conceptos importantes que serán de utilidad para entender el modelo y posteriormente veremos las componentes de los *HMM* junto con los problemas que conlleva su implementación.

4.1. Introducción

Los modelos ocultos de Markov son modelos gráficos probabilísticos que representan procesos dinámicos; en particular, modelan cómo el estado de un proceso cambia con el tiempo. Estos modelos se utilizan para describir datos secuenciales o series de tiempo, tales como datos financieros; en específico, los podemos aplicar a índices bursátiles ver [Ramaprasad Bhar \(2004\)](#).

La idea principal es que un *HMM* es un modelo estadístico en el cual el sistema a estimar es un proceso de Markov con parámetros desconocidos, en donde la ocurrencia de los estados está asociada con una distribución de probabilidad y las transiciones entre los estados están gobernadas por un conjunto de probabilidades de transición. Un *HMM* es un caso particular de las redes Bayesianas dinámicas de primer orden, [Sucar \(2015\)](#).

Más adelante veremos mucho más a detalle este modelo, cabe destacar que cada autor maneja una notación diferente, entonces para efectos de esta tesis utilizaremos la notación de [Zucchini et al. \(2016\)](#).

4.1.1. Cadenas de Markov

Las cadenas de Markov fueron introducidas por el matemático ruso *Andrey Markov* en 1905, quien las define como un proceso estocástico discreto que tiene un número de estados discretos $\{C_t : t = 0, 1, \dots\}$. Además hay una probabilidad de transitar de un estado C_i a un estado C_j ; aunado a esto debe satisfacer la **propiedad de Markov**, esto es, que para cualquier entero, $t \geq 0$ y para cualesquiera estados se cumple que:

$$P(C_{t+1}|C_1, C_2, \dots, C_t) = P(C_{t+1}|C_t).$$

Esto significa que la probabilidad del siguiente estado solo depende del estado actual, lo cual resulta muy útil ya que no debemos ir cargando la información pasada.

Probabilidad de transición

Otro concepto importante es el de probabilidad de transición. Para ello tenemos dos estados, el j e i de una cadena de Markov, y se obtiene como:

$$P(C_{s+t} = j | C_s = i).$$

Esta probabilidad se denota por $\gamma_{ij}(s, s+t)$, que representa la probabilidad de estar en el estado i en el tiempo s , y pasar al estado j en el tiempo $s+t$. Además cuando $\gamma_{ij}(s, s+t)$ no depende de s se dice que la cadena es homogénea.

Entonces si aplicamos la definición anterior sobre todo el conjunto de estados obtendríamos la matriz de transición de un paso tal como se muestra a continuación:

$$\Gamma(s, s+t) = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} \end{pmatrix}.$$

Esta matriz satisface tres propiedades principalmente:

- $\gamma_{ij} \geq 0$.
- $\sum_j \gamma_{ij} = 1$.
- $\Gamma(t+s) = \Gamma(t)\Gamma(s)$,

donde $\Gamma(t) = (\gamma_{ij}(0, t))$ si la cadena es homogénea. La última propiedad se conoce como las ecuaciones de Chapman–Kolmogorov.

Cabe destacar que $\Gamma(1) = \Gamma$ y que $\Gamma(t)$ representa la matriz de transición a t pasos.

Ahora bien será de nuestro interés definir las probabilidades no condicionales $P(c_t = j)$ de una cadena de Markov:

$$u(t) = (P(C_t = 1), \dots, P(C_t = N)); \quad t \in \mathbb{N},$$

que cumple que, para el tiempo $t + 1$,

$$u(t + 1) = u(t)\Gamma.$$

Distribución de probabilidad inicial

Por último, necesitamos saber la probabilidad de comenzar desde un estado determinado. Se le conoce como la distribución de probabilidad inicial y se define de la siguiente manera:

$$\pi_i = P(C_1 = i),$$

además cumple que:

- $\sum_{i=1}^n \pi_i = 1.$

- Coincide con $u(1).$

Estos conceptos son de suma importancia para nuestro estudio. Puede encontrar más información del tema en [Ross et al. \(1996\)](#) y [Rincón \(2012\)](#).

4.2. Modelos ocultos de Markov

Un modelo oculto de Markov (*Hidden Markov Model*, HMM) consiste en un proceso doblemente estocástico. El primer proceso estocástico es una cadena de Markov que se caracteriza por estados y probabilidades de transición. Los estados de la cadena no son observables, por lo tanto, están “*ocultos*” y los denominaremos $\{C_t : t = 1, 2, \dots\}$. El segundo proceso estocástico se define con las observaciones que tenemos, a partir de la distribución de probabilidad que depende del estado, que denominaremos $\{Y_t : t = 1, 2, \dots\}$.

Podemos ver lo anterior expresado en la Figura 4.1.

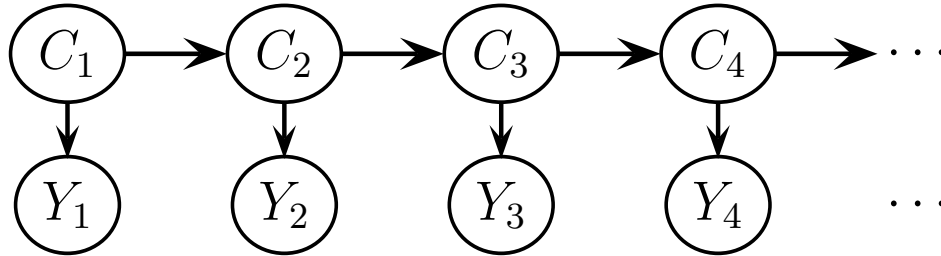


Figura 4.1: Estructura de un *HMM*.

Es importante mencionar que la denominación de “oculto” dentro de los *HMM* se refiere a los estados de la cadena de Markov, no a los parámetros del modelo.

Todo comienza de un estado inicial, a partir del cual se genera una sucesión de estados que se mueve de un estado a otro según las probabilidades de transición, hasta que elegimos un estado final óptimo, con base en los criterios de selección de modelos, de lo que hablaremos más adelante. Además cada estado emite símbolos de acuerdo con la distribución de probabilidad de emisión de ese estado, creando una sucesión de símbolos.

4.2.1. Elementos de un *HMM*

Los elementos que constituyen un Modelo Oculto de Markov con espacio de observaciones discreto son cinco, y se muestran a continuación:

1. Espacio de estados: $\{C = 1, 2, \dots, N\}$, el estado en el tiempo t se denota c_t .
2. Espacio de observaciones: $\{Y_t : t = 1, 2, \dots, T\}$.
3. Vector de probabilidades iniciales: $\Delta = \{\delta_i\}$ donde $i \in C$ y δ_i se define como:

$$\delta_i = P(C_1 = i).$$

4. Matriz de transición a un paso: $\Gamma = \{\gamma_{ij}\}$, con $i, j \in C$ de forma que:

$$\gamma_{ij} = P(C_{t+1} = j | C_t = i).$$

5. Distribución de probabilidades de observación: $\mathbf{P} = \{p_j(k)\}$ y $p_j(k)$ es la probabilidad de que la observación v_k sea emitida por el estado j es decir:

$$p_j(k) = P(Y_t = v_k | C_t = j).$$

Por lo que un *HMM* se representa por $\lambda = \{\Gamma, \mathbf{P}, \delta\}$.

Ahora bien, para las distribuciones univariadas, con observaciones discretas y_t , tenemos que:

$$\begin{aligned} P(Y_t = v_k) &= \sum_{i=1}^n P(C_t = i) P(Y_t = v_k | C_t = i) \\ &= \sum_{i=1}^n u_i(t) p_i(k). \end{aligned}$$

Que podemos reescribir de forma matricial de la siguiente forma:

$$\begin{aligned} P(Y_t = k) &= u(t) \begin{pmatrix} p_1(k) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & p_n(k) \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \\ &= u(t) \mathbf{P}(k) \mathbf{1}, \end{aligned}$$

donde $u(t)$ representa el vector de probabilidades no condicionales, $\mathbf{P}(k)$ es la matriz diagonal con i -ésimo elemento en la diagonal $p_i(k)$ e $\mathbf{1}$ es el vector de unos. Pero también podemos ver a $u(t) = u(1) \Gamma^{t-1}$, por lo que:

$$P(Y_t = k) = u(1) \Gamma^{t-1} \mathbf{P}(k) \mathbf{1}. \quad (4.1)$$

Más aún, como la distribución inicial es estacionaria y $u(1)$ coincide con la distribución inicial δ_i , podemos ver la ecuación (4.1) como:

$$P(y_t = k) = \delta \mathbf{P}(k) \mathbf{1}.$$

De esta forma podemos escribir la verosimilitud del *HMM* como:

$$L_T = \delta P(y_1) \Gamma P(y_2) \Gamma P(y_3) \dots \Gamma P(y_T) \mathbf{1}.$$

Puede encontrar mucha más información del modelo en [Zucchini et al. \(2016\)](#), [Dymarski \(2011\)](#) y [MacDonald & Zucchini \(1997\)](#). Cabe destacar que cada autor maneja una notación diferente y en algunos casos utilizan el mismo símbolo pero se trata de cosas completamente diferentes. Nosotros nos basaremos en lo más posible en la notación de *Walter Zucchini*, autor del libro titulado: “*Hidden Markov Models for Time Series, An Introduction Using R*”.

4.3. Problemas para el HMM

La dificultad de este modelo recae en estimar los parámetros $\{\Gamma, \mathbf{P}, \delta\}$; además debemos asegurar que estamos tomando la mejor sucesión de estados y sobre todo tener los parámetros que maximicen la verosimilitud del modelo. Entonces debemos hacer uso de al menos tres algoritmos para solucionar estos problemas, pero antes de hablar de las soluciones veamos los problemas básicos para aplicar un modelo oculto de Markov, que son:

- Evaluación
- Decodificación
- Aprendizaje

4.3.1. El problema de evaluación

Supongamos que tenemos un modelo oculto de Markov de la forma $\lambda = \{\Gamma, \mathbf{P}, \delta\}$ y un conjunto de observaciones $\{Y = y_1, \dots, y_T\}$; deseamos encontrar $P(Y|\lambda)$. Pa-

ra ello notamos que Y se puede generar por un conjunto diferente de estados C_i de la siguiente manera:

$$P(Y|\lambda) = \sum_{i=1}^N P(Y, C_i|\lambda), \quad (4.2)$$

mientras que para $P(Y, C_i|\lambda)$, basta:

$$P(Y, C_i|\lambda) = \delta_1 p_1(y_1) \gamma_{12} p_2(y_2) \dots \gamma_{(T-1)T} p_T(y_T).$$

Por lo que la ecuación (4.2) se puede ver de la siguiente manera:

$$P(Y|\lambda) = \sum_C \delta_1 p_1(y_1) \gamma_{12} p_2(y_2) \dots \gamma_{(T-1)T} p_T(y_T).$$

Entonces para un modelo con N estados y $t = 1, \dots, T$ observaciones tendríamos N^T sucesiones de estados posibles. Como resultado final, la evaluación requiere una serie de operaciones del orden de $2T \times N^T$, lo cual lo hace muy ineficiente de calcular directamente por lo que debemos recurrir a un algoritmo que optimice el proceso, el cual será descrito a continuación.

Algoritmo Forward-Backward

El algoritmo que se utiliza para resolver el problema de evaluación es el llamado algoritmo *forward-backward*, que se basa en utilizar variables auxiliares que pueden verse de manera recursiva y de esa forma agilizar el procedimiento para obtener $P(Y|\lambda)$.

Vamos a empezar por introducir las variables *forward* que serán denotadas por α_t y se obtienen de la siguiente manera:

$$\begin{aligned} \alpha_t &= \delta P(y_1) \Gamma P(y_2) \Gamma P(y_3) \dots \Gamma P(y_T) \\ &= \delta P(y_1) \prod_{s=2}^T \Gamma P(y_s). \end{aligned}$$

Entonces podemos hacer un proceso que al tiempo t dependa del $t - 1$ es decir:

$$\alpha_t = \alpha_{t-1} \Gamma P(y_t); \quad \text{para } t \geq 2,$$

o en forma escalar podemos verlo como:

$$\alpha_t(j) = \left(\sum_{i=1}^n \alpha_{t-1}(i) \gamma_{ij} \right) p_j(y_t); \quad \text{para } t \geq 2.$$

Para el caso de α_1 se reduce a $\delta P(y_1)$ por lo que:

$$L_T = \alpha_T \mathbf{1}.$$

Ahora vamos a definir la variable conocida como *backward* que la denotaremos con β_t y se obtiene de la siguiente manera:

$$\begin{aligned} \beta_t &= \Gamma P(y_{t+1}) \Gamma P(y_{t+2}) \dots \Gamma P(y_T) \mathbf{1} \\ &= \left(\prod_{s=t+1}^T \Gamma P(y_s) \right) \mathbf{1}, \end{aligned}$$

y de forma recursiva podemos verlo como:

$$\beta_t(i) = \sum_{j=1}^N \gamma_{ij} p_j(y_{t+1}) \beta_{t+1}(j); \quad \text{para } t = 1, 2, \dots, T.$$

Ambas variables las podemos ver como probabilidades y de hecho comparten una relación muy especial con $P(Y|\lambda)$. Notemos que para $t = 1, 2, \dots, T$ y para $i = 1, 2, \dots, n$,

$$\alpha_t(i) = P(Y^{(t)} = y^{(t)}, C_t = i),$$

por lo que con las variables *forward* obtenemos una primera solución que es:

$$P(Y|\lambda) = \sum_{i=1}^N \alpha_T(i).$$

Ahora para $t = 1, 2, \dots, T - 1$ y $i = 1, 2, \dots, n$,

$$\beta_t(i) = P(Y_{t+1}^T = y_{t+1}^T | C_t = i),$$

donde:

$$Y_{t+1}^T = (Y_{t+1}, Y_{t+2}, \dots, Y_T).$$

Una segunda solución con las variables *backward* es:

$$P(Y|\lambda) = \sum_{i=1}^N \delta_i \beta_1(i).$$

Sin embargo, podemos obtener una solución óptima al problema operando ambos resultados, ya que al hacer el producto de $\alpha_t(i) \times \beta_t(i)$ tendríamos lo siguiente:

Para $t = 1, 2, \dots, T$ y para $i = 1, 2, \dots, N$,

$$\alpha_t(i) \times \beta_t(i) = P(Y^{(T)} = y^{(T)}, C_t = i), \quad (4.3)$$

y si operamos $\alpha_t \times \beta_t$ tendríamos:

$$\alpha_t \times \beta_t = P(Y^{(T)} = y^{(T)}) = L_T, \quad (4.4)$$

que coincide con la verosimilitud del modelo para cada una de las t .

Entonces para obtener $P(Y|\lambda)$ basta:

$$P(Y|\lambda) = \sum_i \alpha_t(i) \beta_t(i).$$

En conclusión, el algoritmo *forward-backward* consiste en:

- Para cada $\{i : 1, 2, \dots, N\}$ comienza el proceso *forward* es decir $\alpha_i(1) = \delta P_i(y_1)$.
 - Recurrentemente calcula $\alpha_{t+1}(j) = (\sum_{i=1}^m \alpha_t(i) \gamma_{ij}) p_j(y_{t+1})$.
-

- Ahora debemos iniciar la variable *backward*, esto es para cada $\{i : 1, 2, \dots, N\}$ hacemos $\beta_T(i) = 1$.
- Recurrentemente calcula $\beta_t(i) = \sum_{j=1}^N \gamma_{ij} p_j(y_{t+1}) \beta_{t+1}(j)$.
- Por último calcula $P(Y|\lambda) = \sum_i \alpha_t(i) \beta_t(i)$.

Para más información del algoritmo vea [Yu & Kobayashi \(2003\)](#) y [Cappé et al. \(2005\)](#).

4.3.2. El problema de decodificación

Ahora que ya hemos solucionado el problema de la evaluación y obtuvimos $P(Y|\lambda)$, supongamos que tenemos un HMM, $\lambda = \{\Gamma, \mathbf{P}, \delta\}$ y una sucesión de observaciones $\{Y_t : t = 1, 2, \dots, T\}$, la pregunta es: ¿cuál es la sucesión de estados ocultos $\{C = 1, 2, \dots, N\}$ que se ajusta mejor a las observaciones dadas?

Eso es lo que conocemos como el problema de decodificación, que se refiere a determinar la sucesión de estados más probables; sin embargo, antes de obtener la sucesión debemos saber a cada instante del tiempo t cuál es el estado más probable, para ello vamos a definir una variable auxiliar que denominaremos $\Upsilon_i(t)$ y está definida por:

$$\begin{aligned} \Upsilon_i(t) &= P(C_t = i | Y^{(T)} = y^{(T)}) \\ &= \frac{P(C_t = i, Y^{(T)} = y^{(T)})}{P(Y^{(T)} = y^{(T)})}, \end{aligned}$$

y recordando la ecuación (4.3) y (4.4) podemos ver a $\Upsilon_i(t)$ como:

$$\Upsilon_i(t) = \frac{\alpha_i(t) \times \beta_i(t)}{L_T},$$

donde $\alpha_i(t)$ y $\beta_i(t)$ corresponden a las variables *forward* y *backward*, descritas en la sección 4.3.1.

Entonces la forma de encontrar el estado más probable i_t^* para cada observación en el tiempo t se resume a:

$$i_t^* = \operatorname{argmax}_{i=1,\dots,n} \Upsilon_i(t),$$

Una vez obtenida la decodificación por cada observación, podríamos establecer la sucesión de estados “óptima”; sin embargo, tiene dos inconvenientes, el primero es que tendríamos que hacer N^T operaciones y el segundo es que no tomaríamos en cuenta las probabilidades de transición, lo cual puede ocasionar que realmente no sea la mejor sucesión de estados ocultos para el modelo. La solución para este problema es utilizar el algoritmo de Viterbi, que describiremos a continuación.

Algoritmo Viterbi

En 1967, *Andrew Viterbi* publicó un artículo, donde describió un algoritmo de programación dinámica que soluciona el problema de decodificación, (*Viterbi 1967*), el cual busca hacer una recursión entre un tiempo t y el tiempo $t - 1$ para optimizar el proceso de encontrar la ruta más probable de estados ocultos, entonces vamos a definir ξ_{ti} de la siguiente manera:

$$\xi_{1i} = P(C_1 = i, Y_1 = y_1) = \delta_i p_i(y_1);$$

para $t = 2, \dots, T$,

$$\xi_{ti} = \max_{c_1, c_2, \dots, c_{t-1}} P(C^{(t-1)} = c^{(t-1)}, C_t = i, Y^{(T)} = y^{(T)}).$$

Por lo que para $t = 2, 3, \dots, T$ y para $i = 1, 2, \dots, n$, la recursión queda de la siguiente manera:

$$\xi_{tj} = (\max_i (\xi_{t-1,i} \gamma_{ij})) p_j(y_t).$$

Ahora que tenemos la variable encargada de hacer la recursión, obtendremos la mejor sucesión de estados con las siguientes ecuaciones:

$$i_T = \operatorname{argmax}_{i=1,\dots,n} \xi_{Ti},$$

y si tomamos a $t = T - 1, T - 2, \dots, 1$, obtendríamos la sucesión de la siguiente manera:

$$\phi_i(t) = \operatorname{argmax}_{i=1, \dots, n} (\xi_{ti} \gamma_{i, i+1}).$$

Notemos que tiene parecido con el algoritmo *forward-backward* ya que de igual forma iteraremos hacia atrás y hacia adelante para obtener el resultado óptimo.

En resumen, el algoritmo de Viterbi consiste en:

- Para cada $\{i : 1, 2, \dots, N\}$, definamos las condiciones iniciales, es decir: $\xi_{1i} = \delta_i p_i(y_1)$ y $\phi_i(1) = 0$.
- Recurrentemente calcula $\xi_{tj} = (\max_i (\xi_{t-1, i} \gamma_{ij})) p_j(y_t)$.
- De igual forma calcula $\phi_i(t) = \operatorname{argmax}_{i=1, \dots, n} (\xi_{ti} \gamma_{i, i+1})$.
- Ahora obtén $i_T = \operatorname{argmax}_{i=1, \dots, n} \xi_{Ti}$.
- La sucesión de estados óptimos será $i_t = \phi_{t+1}(i_{t+1})$.

Puede encontrar más información de este algoritmo en [Forney \(1973\)](#).

4.3.3. El problema de aprendizaje

Ya pudimos solucionar la evaluación y la decodificación; sin embargo, falta lo que podría ser lo más importante, saber cómo vamos a estimar los parámetros del HMM $\lambda = \{\Gamma, \mathbf{P}, \delta\}$, de forma que maximicemos la $P(Y|\lambda)$ a partir de un conjunto de observaciones $\{Y_t : t = 1, 2, \dots, T\}$ dadas.

Podemos formular este problema de la siguiente manera:

$$\operatorname{argmax}_{\lambda} P(Y|\lambda).$$

Ese es el problema del aprendizaje, pero hay que tomar en cuenta que solo vamos a ir calibrando el modelo, con base en las observaciones que tengamos, pero también supone que el número de estados ocultos del modelo ya está dado, por lo que iremos probando el modelo con diferente número de estados y decidiremos cuál es el mejor tomando en cuenta el *AIC* y el *BIC*, de los cuales hablaremos más adelante.

Entonces, ya que no existe una fórmula analítica de resolver ese problema [Baum et al. \(1970\)](#), publicaron un artículo donde desarrollaron un algoritmo basado en la estructura de un algoritmo EM para solucionar el problema del aprendizaje de un HMM.

El algoritmo EM (*Expectation–maximization*) es un método iterativo que calcula estimadores máximo verosímiles cuando hacen faltan algunos datos. Este algoritmo se separa en dos pasos:

1. **Paso E:** en el cual se calcula la esperanza de la log-verosimilitud con respecto a la distribución de los datos faltantes, dada una aproximación de la solución.
2. **Paso M:** este paso da una nueva aproximación maximizando la log-verosimilitud que resultó del paso E.

Estos pasos se realizan hasta que el algoritmo converge.

Para más información del tema vea [Dempster et al. \(1977\)](#).

Algoritmo de Baum-Welch

El algoritmo de *Baum-Welch* está especialmente diseñado para estimar los parámetros $\Gamma, \mathbf{P}, \delta$ que componen a un HMM; para ello requerimos dos variables auxiliares que llamaremos $u_j(t)$ y $v_{jk}(t)$ y se definirán de la siguiente manera:

$$u_j(t) = P(C_t = j | y^{(T)}) = \frac{\alpha_t(j) \times \beta_t(j)}{P(Y | \lambda)},$$

$$\begin{aligned}
v_{jk}(t) &= P(C_t = j, C_{t+1} = k | y^{(T)}) \\
&= \frac{\alpha_t(j) \times \gamma_{jk} \times p_k(y_t) \times \beta_{t+1}(k)}{P(Y|\lambda)}.
\end{aligned} \tag{4.5}$$

Notamos que de nuevo aparecen las variables *forward*, *backward* pero además debemos ver la relación que existe entre las variables $u_j(t)$ y $v_{jk}(t)$:

$$\begin{aligned}
u_j(t) &= P(C_t = j | y^{(T)}) \\
&= \sum_{k=1}^N P(C_t = j, C_{t+1} = k | y^{(T)}) \\
&= \sum_{k=1}^N v_{jk}(t).
\end{aligned}$$

Ahora, para comenzar con el algoritmo, tomaremos valores aleatorios para los parámetros (paso E) y empezaremos a iterar hasta que logre encontrar los valores óptimos que maximizan la verosimilitud; para ello los parámetros re-estimados serán (paso M):

$$\hat{\delta}_j = \frac{u_j(1)}{\sum_{j=1}^n u_j(1)} = u_j(1), \tag{4.6}$$

$$\hat{\gamma}_{jk} = \frac{\sum_{t=1}^T v_{jk}(t)}{\sum_{t=1}^T u_j(t)}, \tag{4.7}$$

$$\hat{p}_j(k) = \frac{\sum_{t=1}^T v_{jk}(t) \mathbb{1}_{(y_t=k)}}{\sum_{t=1}^T v_{jk}(t)}, \quad 1 \leq t \leq T. \tag{4.8}$$

Entonces tendremos un HMM, $\hat{\lambda} = \{\hat{\Gamma}, \hat{\mathbf{P}}, \hat{\delta}\}$, cuyos parámetros maximizan la log-verosimilitud a partir de las observaciones dadas y el número de estados que

hace que tengamos el mejor modelo.

Sin embargo, antes de eso veamos que la ecuación (4.8) depende directamente de la distribución dependiente del estado que se asuma, la cual puede ser discreta, continua o incluso una mezcla, por lo que para efectos de esta tesis utilizaremos la distribución normal. Entonces el problema se reduce a encontrar los estimadores máximo verosímiles de la distribución normal para poder llevar a cabo el paso M.

Lo anterior lo podemos ver en la ecuación (4.5), entonces para una distribución normal tenemos que:

$$p_j(y) = \left(2\pi\sigma_j^2\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma_j^2}(y - \mu_j)^2\right).$$

Por último, los estimadores máximo verosímiles para σ_j^2 y μ_j son:

$$\hat{\mu}_j = \frac{\sum_{t=1}^T \hat{u}_j(t)y_t}{\sum_{t=1}^T \hat{u}_j(t)}, \quad (4.9)$$

$$\hat{\sigma}_j^2 = \frac{\sum_{t=1}^T \hat{u}_j(t)(y_t - \hat{\mu}_j)^2}{\sum_{t=1}^T \hat{u}_j(t)}. \quad (4.10)$$

En resumen para el algoritmo de *Baum-Welch* debemos hacer lo siguiente:

- Iniciar aleatoriamente los parámetros del HMM $\{\Gamma, \mathbf{P}, \delta\}$.
 - Calcular los estimadores máximo verosímiles así como se presenta en las ecuaciones (4.9) y (4.10).
 - Comenzar a volver a estimar los parámetros, tal como se muestra en las ecuaciones (4.6), (4.7) y (4.8).
-

- Hacer el paso anterior hasta que el algoritmo converja y así obtengamos los parámetros $\{\hat{\Gamma}, \hat{\mathbf{P}}, \hat{\delta}\}$ que maximicen la log-verosimilitud bajo las condiciones que tenemos.

Puede encontrar más información en [Baggenstoss \(2001\)](#).

4.4. Selección de modelos

En el análisis de datos, siempre que se proponen varios modelos, la pregunta inmediata es ¿cuál es el mejor de todos?, y en los modelos ocultos de Markov, no será la excepción.

En el momento que implementamos el algoritmo, iremos variando el número de estados, así generaremos muchos modelos y debemos de decidir cuál es el mejor; es por ello que debemos tener criterios para poder decidir con qué modelo nos quedaremos.

Existen muchos criterios que se ocupan para la selección de modelos; entre los más populares encontramos el criterio de información de Akaike (*AIC*) y el criterio de información Bayesiano (*BIC*), los cuales son extensiones del método de máxima verosimilitud. El objetivo será elegir el modelo cuyo *AIC* y *BIC* sean mínimos y ese modelo lo propondremos como el mejor modelo. Es por ello que es importante hablar de estos criterios, ya que serán la base del siguiente capítulo donde presentaremos los resultados de la implementación del modelo en índices bursátiles.

Puede encontrar información de otros criterios como el criterio de información de la devianza (*DIC*), criterio de información de Hannan-Quinn (*HQ*), entre otros, en [Burnham & Anderson \(2002\)](#), [Burnham & Anderson \(2004\)](#) y [Read & Cressie \(2012\)](#).

4.4.1. Criterio de Información de Akaike

En 1973 *Hirotsugu Akaike* propuso el criterio de información de Akaike, mejor conocido como *AIC*. Su introducción fomentó el reconocimiento de la importancia de un buen modelo estadístico, por lo que su uso se ha popularizado en los últimos años.

El AIC es una medida que penaliza el número excesivo de parámetros estimados de un modelo estadístico para un conjunto de datos conocidos, se deriva de minimizar el valor esperado insesgado del criterio de Kullback-Leibler, [Akaike et al. \(1998\)](#) y se obtiene de la siguiente manera:

$$AIC = -2\log(L(\hat{\theta}|x)) + 2k,$$

donde: $\log(L(\hat{\theta}|x))$ corresponde a la log-verosimilitud evaluada en el estimador $\hat{\theta}$ y k el número de parámetros estimados.

Además, al término $-2\log(L(\hat{\theta}|x))$ se le conoce como la devianza. Dentro de ella se encuentra la medida de bondad de ajuste, la cual es $2\log(L(\hat{\theta}|x))$ que es una función creciente y por ello se le aplica el signo menos para hacerla decreciente, de esta forma cuando la verosimilitud es muy grande la expresión de la bondad de ajuste tiende a un valor muy chico. De esta forma este término mide el desajuste entre una distribución hipotética y una distribución teórica.

Ahora el término de $2k$, es el que penaliza el número de parámetros estimados. Esta expresión es la que mide la complejidad del modelo.

Notemos que mientras tengamos más parámetros, el modelo gana complejidad, pero a la vez el desajuste disminuye, por lo que esta medida es un equilibrio entre ambas cosas.

En resumen, cuando tengamos el AIC de cada modelo debemos de tomar el mínimo y elegirlo como el mejor modelo.

4.4.2. Criterio de Información Bayesiano

El criterio de información Bayesiano, conocido comúnmente como *BIC* y también como criterio de información de *Schwarz (SIC)*, es una medida de bondad de ajuste para un modelo estadístico. Fue introducida por Schwarz en 1978 pero como su nombre lo indica le dio una perspectiva bayesiana; sin embargo, comparte similitudes con el AIC, solo difiere en cuanto a la penalización de los parámetros.

El cálculo se realiza de la siguiente manera:

$$BIC = -2\log(L(\hat{\theta}|x)) + k\log(n),$$

donde el primer término y k corresponden a lo equivalente que en el AIC, solo que este criterio introduce a n que es el tamaño de muestra. Entonces mientras más complejo sea el modelo el BIC tendrá una mayor penalización que el AIC.

La interpretación de ambos términos del BIC sigue la idea del AIC, por lo que debemos elegir el modelo cuyo BIC sea el mínimo.

Al elegir el mejor modelo mediante estos criterios, será el que tenga menor número de parámetros pero lo suficientemente complejo tal que hagan las mejores estimaciones.

Para efectos de esta tesis realizaremos ambas medidas. En el caso en que las dos medidas coincidan vamos a seleccionarlo como el mejor modelo para ese índice bursátil, pero en caso contrario haremos dos ajustes, uno que minimice al AIC y otro que minimice al BIC.

4.5. Predicción

Una vez que tenemos un modelo adecuado, implica que sabemos exactamente cuántos estados necesitamos, cuáles son los parámetros del modelo y además sabemos cuál es la decodificación por estado, por lo cual ya podemos empezar a

preguntarnos si nuestro mejor modelo es un buen modelo como predictor para el tiempo $T + h$, para ello necesitamos conocer la distribución de predicción que es:

$$\begin{aligned}
 P(Y_{T+h} = y | Y^{(T)} = y^{(T)}) &= \frac{P(Y^{(T)} = y^{(T)}, Y_{T+h} = y)}{P(Y^{(T)} = y^{(T)})} \\
 &= \frac{\delta P(y_1) \times B_2 \times \cdots \times B_T \Gamma^h P(y) \mathbb{1}}{\delta P(y_1) \times B_2 \times \cdots \times B_T \mathbb{1}} \\
 &= \frac{\alpha_T \Gamma^h P(y) \mathbb{1}}{\alpha_T \mathbb{1}},
 \end{aligned}$$

donde $B_i = \Gamma P(y_i)$.

De esa forma podremos hacer que el modelo prediga la observación del tiempo siguiente.

A lo largo de este capítulo se presentó bibliografía general del modelo y también de cada algoritmo que usaremos para obtener un HMM, así que puede encontrar más información y más específica en ella, al igual que las demostraciones de las fórmulas.

HMM aplicado a series de tiempo financieras

A continuación mostaremos los resultados de aplicar HMM a los índices bursátiles. Cabe destacar que todos los programas fueron realizados en el lenguaje de programación *R*.

5.1. Algoritmo

La forma en la que se realizó el análisis estadístico de los datos y también la forma de presentar los resultados será la siguiente:

1. Como primer paso revisaremos cómo se ve la serie junto con la transformación descrita en la sección 2.2.
2. Una vez que tenemos los datos transformados, ajustaremos el modelo para $C = \{2, 3, 4, 5, 6, 7, 8\}$, con una distribución normal, recordando que la C corresponde al número de estados ocultos del modelo.
3. Obtendremos los valores AIC y BIC de los modelos para detectar al candidato con mejores estimaciones como mejor modelo.

4. Ya que tenemos el mejor modelo revisaremos las estimaciones de los parámetros que lo ajustan al igual que la sucesión de estados que mejor se adaptan a las observaciones.
5. Una vez que tenemos la sucesión de estados ocultos presentaremos cómo se ve en los datos transformados, al igual que en los precios al cierre.
6. Después mostraremos la predicción.
7. Por último mostraremos el VaR_α , el $CVaR_\alpha$ y aplicaremos la prueba de Kupiec para ver si los resultados son consistentes.

Estos siete pasos se realizarán para el análisis de los cuatro índices bursátiles: Índice de Precios y Cotizaciones (IPC), Standard & Poor's (S&P500), Deutscher Aktienindex (DAX) y Nikkei Heikin Kabuka (Nikkei225).

5.2. Índice de Precios y Cotizaciones

Comenzaremos con el Índice de Precios y Cotizaciones (IPC). En la Figura 5.1 se muestra cómo se ven los datos y la serie al aplicarle el *log return*.

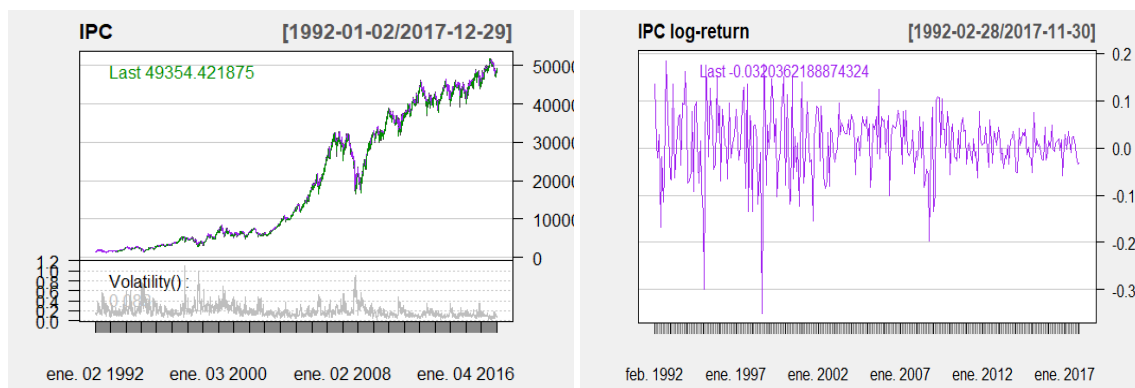


Figura 5.1: IPC: serie original y transformada.

El valor al cierre más bajo reportado en el periodo que elegimos es el 30 de septiembre de 1992. Esto se debe a que ese año se reportó una temporada de huracanes

en el océano Pacífico y en esa fecha en particular, estaba el huracán Roslyn el cual provocaba pérdidas muy grandes. Por otro lado, el IPC llegó a su nivel máximo el 30 de septiembre del 2017, lo que tiene sentido ya que después de la crisis del 2008 el IPC tuvo una tendencia creciente.

Ahora, recordando el algoritmo presentado en la sección 5.1, correremos el modelo para todos los estados deseados. Los resultados se muestran en la tabla 5.1. Por lo que basándonos en el AIC y BIC, el mejor modelo para el IPC será con 2 estados, que además converge con solo 10 iteraciones, lo que hace que el proceso sea rápido.

IPC México				
Estado	Iteraciones	LogLik	AIC	BIC
2	10	437.06	-860.11	-833.96
3	85	441.38	-854.76	-802.44
4	95	448.80	-851.59	-765.65
5	95	452.74	-837.47	-710.43
6	83	466.85	-839.70	-664.08
7	122	469.33	-814.67	-583.00
8	163	479.38	-800.76	-505.57

Tabla 5.1: IPC: selección del mejor modelo.

En este punto ya hemos resuelto los tres problemas del modelo, implementando los algoritmos antes descritos, así que daremos los resultados más importantes. Comenzaremos dando las estimaciones de los parámetros que ajustan el mejor modelo que son, la probabilidad inicial

$$\hat{\Delta} = \{1.921e - 20, 1\}, \quad (5.1)$$

y matriz de transición

$$\hat{\Gamma} = \begin{pmatrix} 0.99104571 & 0.008954293 \\ 0.01764542 & 0.982354579 \end{pmatrix}. \quad (5.2)$$

Ahora describiremos los parámetros que ajustan cada una de las distribuciones por estado, esto gracias al algoritmo de Viterbi que irá clasificando cada observación en un estado que puede ser el estado 1 o el estado 2. De igual forma veremos el número de observaciones que las componen.

	Número de observaciones	Media	Varianza
Estado 1	168	0.013278069	0.001436689
Estado 2	142	0.007967425	0.008634246

Tabla 5.2: IPC: descripción por estado del mejor modelo.

Como se puede observar en la tabla 5.2, el estado 1 contiene a las observaciones que presentan mayor media y varianza, además tiene un poco de más datos que el estado 2. Con las estimaciones de los parámetros que se presentan podemos obtener el diagnóstico.

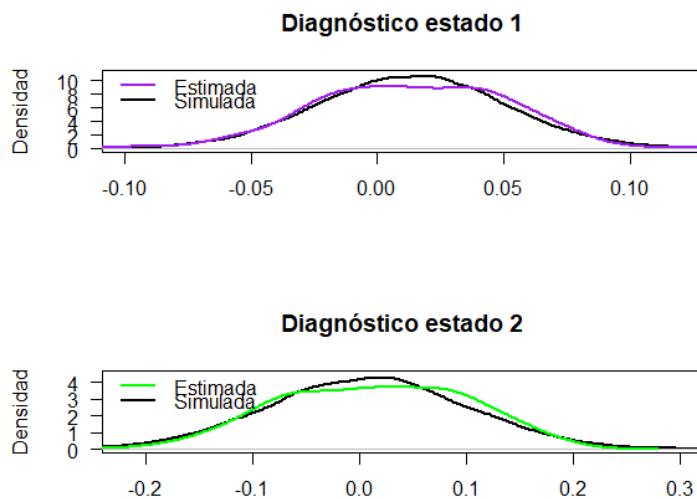


Figura 5.2: IPC: diagnóstico por estado.

Lo primero que podemos notar en la Figura 5.2 es que el rango donde se concentran las observaciones del estado 2 es mayor que el del estado 1 a pesar que el segundo estado contiene un número menor de observaciones. La distribución en morado y verde respectivamente representan la densidad estimada de cada estado, mientras que la de negro se construyó usando una distribución normal con media y varianza iguales a las estimaciones que aparecen en la tabla 5.2.

Para determinar a qué estado pertenece cada una de las observaciones debemos seleccionar el más probable, en la Figura 5.3 podemos observar la probabilidad de permanecer a un estado u otro.

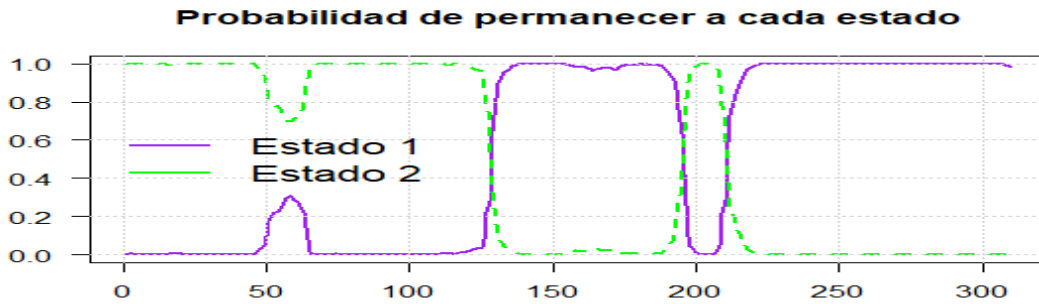


Figura 5.3: IPC: probabilidad de estar en cada uno de los estados.

A partir de este punto podemos obtener la decodificación de los datos del IPC que se muestra en la Figura 5.4. Podemos notar que en la primera parte se presenta la decodificación por observación, presentando en morado el estado 1 y en verde el estado 2, que además el estado 2 predomina desde el inicio de la serie y cambia al estado 1 en octubre del 2002 pero es en junio del 2008 donde aparece de nuevo el estado 2, lo cual tiene mucho sentido por que justamente entre 2008 y 2009 se vivió una de las peores crisis financieras a las que se ha enfrentado nuestro país y justamente el cambio al estado 1 se da en agosto del 2009. En la segunda parte se muestra la serie por estado. Para obtener dicha serie basta con unir la codificación por estado.

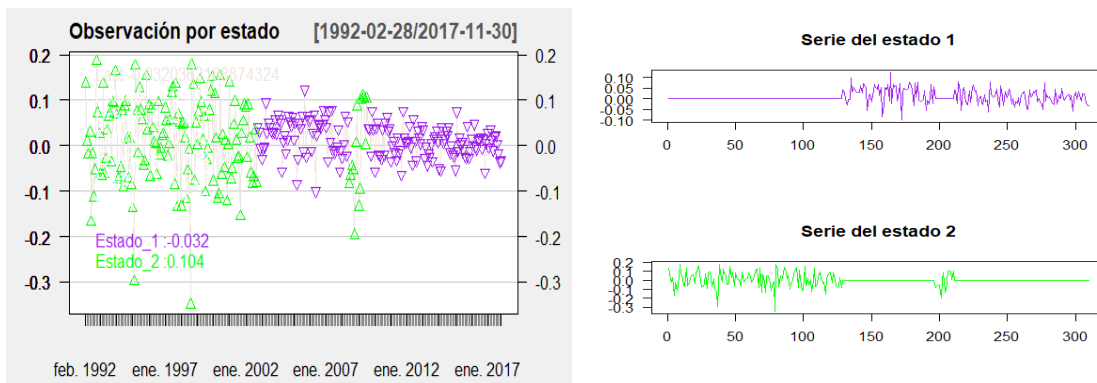


Figura 5.4: IPC: clasificación por estado.

Más aún, como ya tenemos la decodificación de la serie podemos ver cómo se ve en los precios al cierre. En la Figura 5.5 confirmamos lo que se observa en la Figura 5.4 y de hecho podemos observar que el primer cambio de estado se presenta cuando el índice comienza a tener una tendencia creciente y vuelve a cambiar de estado cuando hay una caída en el mismo.

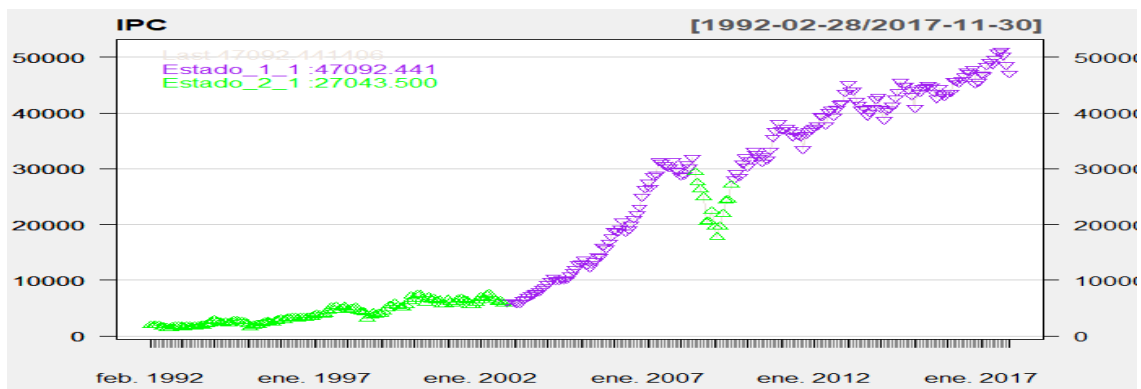


Figura 5.5: IPC: clasificación en los precios al cierre.

Ahora a partir de la media y varianza de cada estado (tabla 5.2), la probabilidad inicial (5.1) y la matriz de transición (5.2) de este modelo podemos hacer una predicción, la cual se hará sobre 60 periodos empezando por enero del 2018. Los resultados se presentan en la Figura 5.6, donde el color verde corresponde a la predicción que abarca de enero del 2018 a diciembre del 2022.

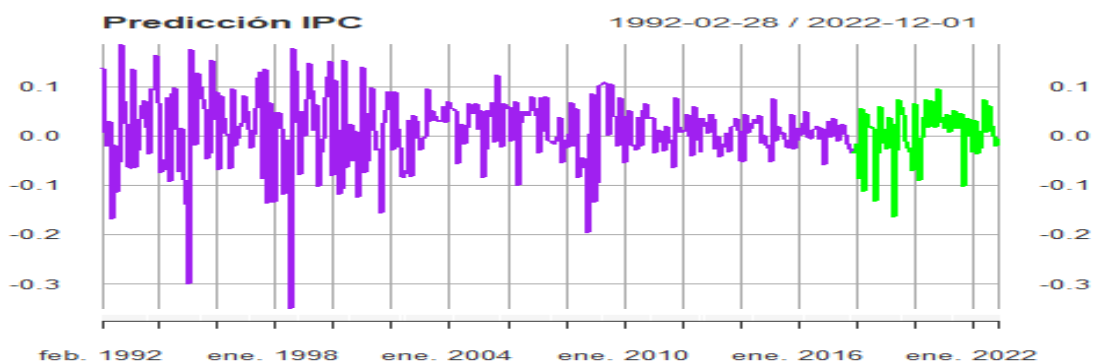


Figura 5.6: IPC: predicción.

De esas 60 nuevas observaciones, 41 se clasificaron en el estado 1 y 19 en el estado 2. Recordando que el estado 1 es el que presenta mayor varianza podríamos decir que en los próximos años el índice bursátil mexicano sufrirá cambios importantes.

Para ver qué tan buena fue la predicción, vamos a obtener los datos del 2018 y del primer cuatrimestre del 2019 para poder evaluar nuestro modelo; con ello haremos el mismo tratamiento que le hicimos a los datos para obtener el *log return* con lo que entrenamos el modelo para poder comparar el valor de la predicción con el valor real.

IPC México		
Fecha	Logret	Predicción
28/02/2018	-0.06168	-0.01567
28/03/2018	-0.02807	-0.08637
30/04/2018	0.04728	0.05449
31/05/2018	-0.07950	-0.11263
29/06/2018	0.06502	0.04506
31/07/2018	0.04181	0.01815
31/08/2018	-0.00303	0.01950
28/09/2018	-0.00088	0.01621
31/10/2018	-0.11917	-0.13271
30/11/2018	-0.05160	-0.03581
31/12/2018	-0.00222	0.06013
31/01/2019	0.05485	-0.01446
28/02/2019	-0.02682	0.01432
29/03/2019	0.01063	-0.02646
30/04/2019	0.02995	0.03714

Tabla 5.3: IPC: predicción.

En la tabla 5.3 y en la Figura 5.7 notamos que los resultados de la predicción son parecidos a los del *log return* y al calcular el error cuadrático medio obtenemos un resultado del 0.396 %, el cual es un excelente resultado, por lo que concluimos que el modelo es bueno, aprendió bien y da buenas predicciones.

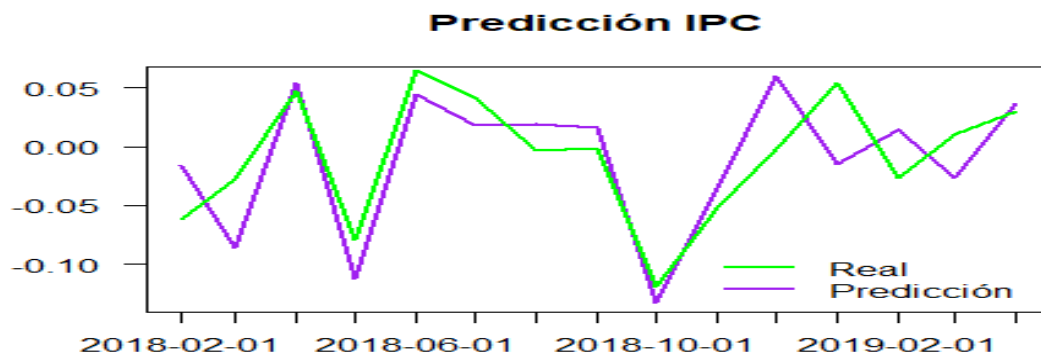


Figura 5.7: IPC: predicción y valor real.

Ahora presentaremos el resultado de las métricas de riesgo. Utilizaremos un nivel de confianza del 95 %.

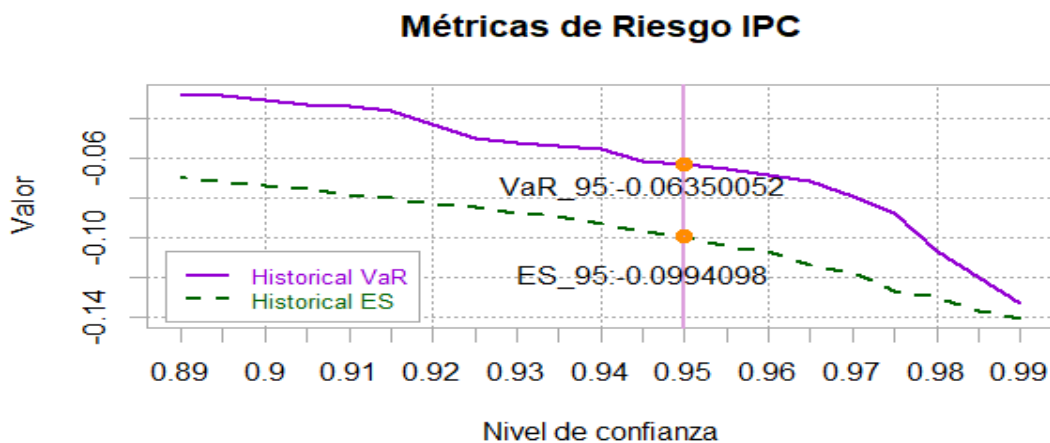


Figura 5.8: IPC: métricas de riesgo.

Notamos que el resultado en ambos casos es negativo por lo que estamos esperando una ganancia, financieramente esto es muy importante ya que ambas cantidades son referentes para inversionistas.

Por último presentaremos el resultado de aplicar la prueba de Kupiec con el fin de dar a conocer si el resultado del $VaR_{0.05}$ es consistente.

Índice	$\chi_{0.05}^2$	Test	P-valor
IPC	0.04	0.4456401	0.5044127

Tabla 5.4: IPC: prueba de Kupiec.

Como podemos observar en la tabla 5.4, el valor de la prueba es mayor que el cuantil 0.05 de la distribución χ_1^2 y además el p-valor es mayor que el nivel de significancia 0.05, por lo que no podemos rechazar H_0 . Entonces concluimos que el modelo es consistente.

5.3. Standard & Poor's 500

Ahora mostraremos los resultados del índice bursátil *S&P500*. En la Figura 5.9 se presentan las series de datos originales y al aplicarle la transformación *log return*.

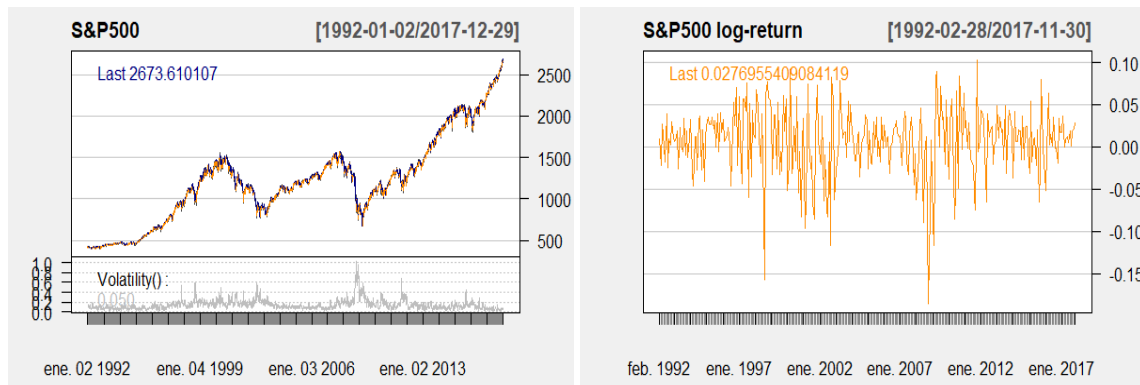


Figura 5.9: S&P500: serie original y transformada.

En la gráfica del *log return* notamos que el valor mínimo dentro del periodo que elegimos es en septiembre del 2008, que es la fecha en la que estalló la crisis de liquidez en Estados Unidos, en donde este índice cayó casi 60 puntos bases, una de las peores caídas en la historia por lo que el gobierno tuvo que proveer de reservas

para poder estabilizar el índice y de hecho comenzó a tener una tendencia creciente.

Por otro lado, el valor mínimo en el valor al cierre se da al principio del periodo y dado que tiene una tendencia creciente a partir del 2016, el máximo es en el último periodo que analizamos.

Al igual que en el índice anterior vamos a proceder a ajustar el modelo variando el número de estados para poder determinar cuál es el mejor modelo. En la tabla 5.5 podemos observar que el estado que minimiza el AIC y el BIC es con tres estados que converge con 79 iteraciones.

S&P500 EUA				
Estado	Iteraciones	LogLik	AIC	BIC
2	17	592.59	-1171.20	-1118.43
3	79	601.64	-1175.29	-1122.98
4	165	610.05	-1174.11	-1088.17
5	86	618.41	-1168.83	-1041.78
6	202	616.47	-1138.94	-963.32
7	131	628.30	-1132.61	-900.94
8	181	635.39	-1112.79	-817.60

Tabla 5.5: S&P500: selección del mejor modelo.

Ahora veamos las estimaciones de los parámetros $\hat{\Delta}$ y $\hat{\Gamma}$ que calibran este modelo. Ya que tiene tres estados, tendremos una distribución inicial con tres entradas y una matriz de transición de 3×3 :

$$\hat{\Delta} = \{2.415958e - 260, 9.898887e - 141, 1\}, \quad (5.3)$$

$$\hat{\Gamma} = \begin{pmatrix} 9.773861e - 01 & 1.552975e - 47 & 0.02261391 \\ 2.728605e - 05 & 1.004127e - 01 & 0.89956006 \\ 2.736613e - 02 & 2.283089e - 01 & 0.74432497 \end{pmatrix}. \quad (5.4)$$

Las estimaciones de los parámetros que corresponden a la distribución de cada

uno de los estados al igual que el número de observaciones que contiene cada estado lo veremos en la tabla 5.6.

	Número de observaciones	Media	Varianza
Estado 1	146	0.001537236	0.0028665526
Estado 2	34	-0.022907363	0.0001055407
Estado 3	130	0.018678866	0.0002141778

Tabla 5.6: S&P500: descripción por estado.

Ahora analicemos el diagnóstico de cada una de las distribuciones por estado.

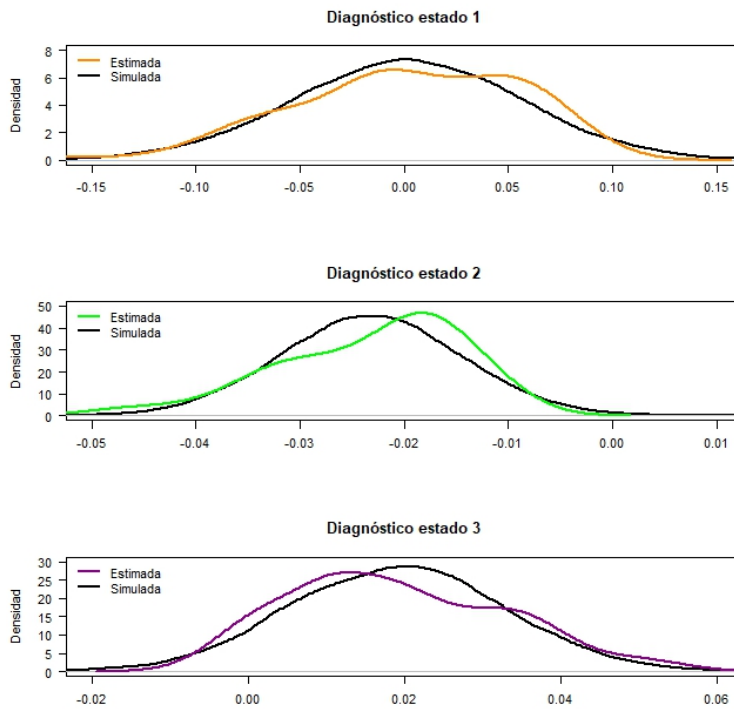


Figura 5.10: S&P500: diagnóstico por estado.

En la Figura 5.10 notamos que en el estado 1 las observaciones tienen una mayor dispersión que en el resto. En el estado 2 los datos se concentran entre -0.04 y -0.01 , mientras en el estado 3 la media es 0.018 y los datos se concentran entre 0 y 0.04 por lo que el estado 3 contiene los datos con *log return* mayor. El estado 1 es el que contiene casi el 50% de las observaciones, mientras el 2 solo tiene el 10.2% de los datos.

Ahora veremos la probabilidad de que cada observación esté en alguno de los estados; esto se presenta en la Figura 5.11.

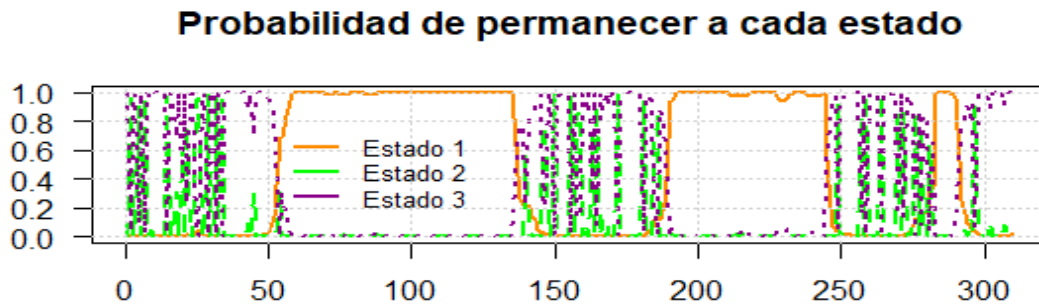


Figura 5.11: S&P500: probabilidad de estar en cada uno de los estados.

En esta gráfica podemos observar que el estado 1 se separa completamente de los otros dos estados, pero en el estado 2 y 3 hay intervalos en las que se difícilmente se ve cuál es el estado más probable. Además, tenemos 34 observaciones en el estado dos y 130 en el tercero por lo que sí hay diferencia, así que veamos cómo se clasificaron las observaciones por estado y cuáles son las características de cada uno.

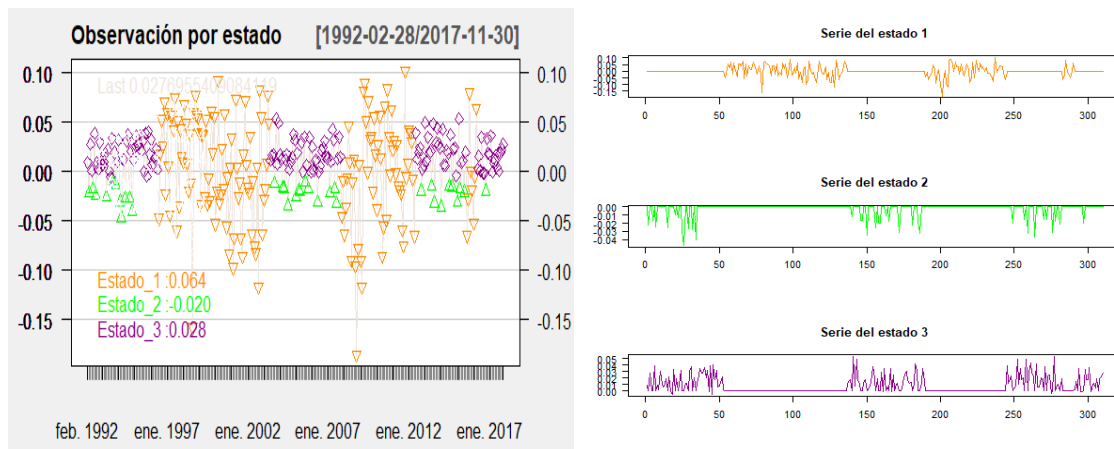


Figura 5.12: S&P500: clasificación por estado.

La Figura 5.12 es muy interesante porque podemos ver que el estado uno corresponde a las observaciones con mayor volatilidad, justo como se veía en el

diagnóstico. Sin embargo, dentro de los intervalos en los que no hay mucha volatilidad está clasificando dentro del estado tres a los que tienen un *log return* mayor a cero, mientras que el estado dos contiene los que tienen un *log return* menor a cero.

Por otra parte, veamos cómo se clasifican las observaciones en los precios al cierre, esto se muestra en la Figura 5.13, y pareciera que el estado 2 y 3 están combinados, sin embargo ya hemos descrito la diferencia entre uno y otro.

Cabe destacar que el estado 1 se compone de tres intervalos, el primero va de julio de 1996 a mayo del 2003, en el cual se presentó la crisis de Asia conocida como la primera crisis global que golpeó fuertemente al mercado de EUA; este periodo también abarca el atentado a las torres gemelas que fue el 11 de septiembre del 2001. El segundo va de julio del 2007 a mayo del 2012, que abarca el periodo conocido como la gran recesión que fue una de las peores crisis financieras en la historia de EUA. Por último, tenemos el periodo de agosto del 2015 a marzo del 2016 que fue un año muy violento para el país al igual que el empresario Donald Trump comenzó su campaña presidencial con discursos de rechazo hacia los migrantes.

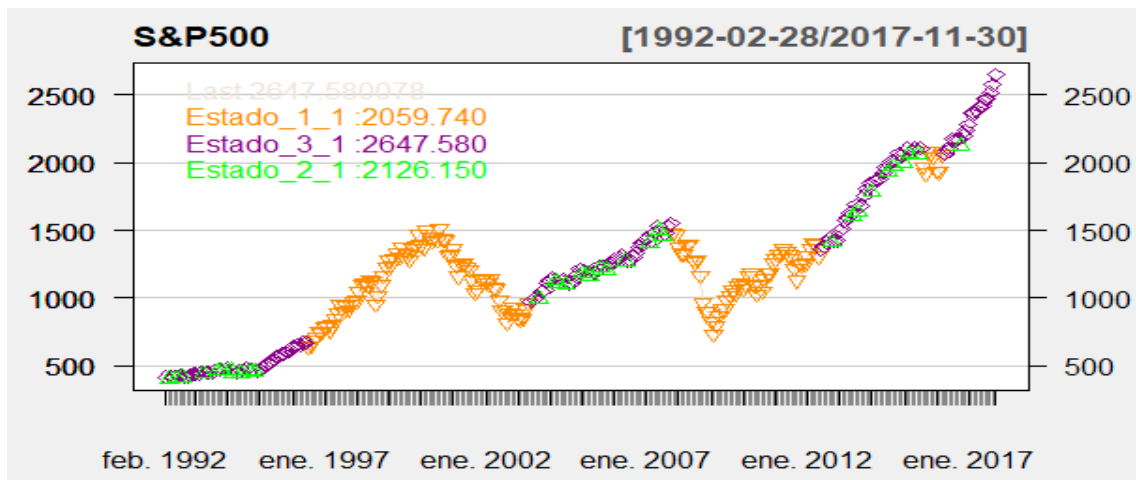


Figura 5.13: S&P500: clasificación en los precios al cierre.

De igual forma que con el IPC, se realizó una predicción a 60 periodos empezando por enero del 2018 y terminando en diciembre del 2022 a partir de la media y varianza de cada estado (tabla 5.6), la probabilidad inicial (5.3) y la matriz de transición (5.4). Los resultados se presentan en la Figura 5.14, en la cual se presenta de color naranja la serie del *log return* y de magenta la predicción.

De las 60 observaciones que predcimos, el algoritmo de Viterbi clasifica 22 en el estado 1, 5 en el estado 2 y 33 del estado 3, por lo que se ve que a partir del 2020 habría mayor variabilidad en el índice; sin embargo, más de la mitad de las nuevas observaciones las clasifica en el estado donde el *log return* es mayor a cero, lo que significaría un buen panorama para EUA.

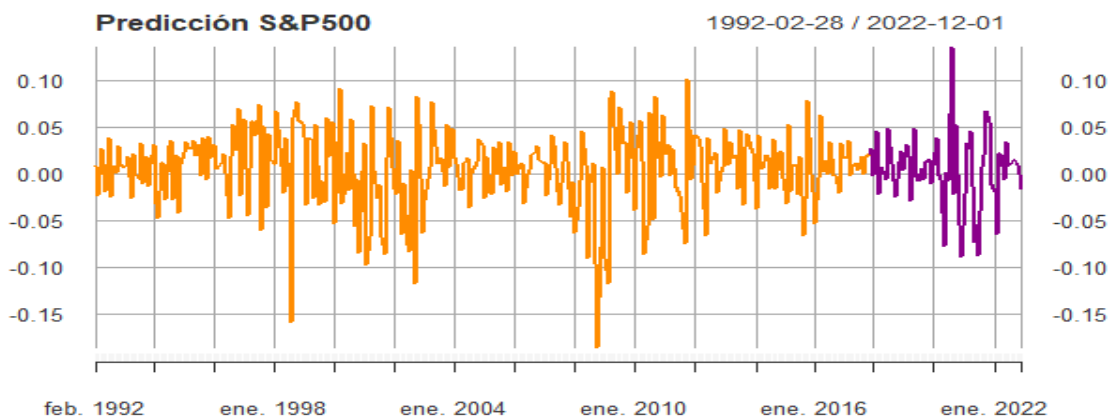


Figura 5.14: S&P500: predicción.

Para evaluar qué tan buena es la predicción obtendremos los datos del 2018 y del primer cuatrimestre del 2019, los transformaremos y compararemos con el valor de la predicción que arroja nuestro modelo.

En la Tabla 5.7 vemos que los resultados son parecidos y al calcular el error cuadrático medio obtenemos un valor de 0.2093% el cual es un muy buen valor.

S&P500 EUA		
Fecha	Logret	Predicción
28/02/2018	-0.03973	-0.00098
28/03/2018	-0.02725	0.04605
30/04/2018	0.00272	-0.02123
31/05/2018	0.02138	0.00783
29/06/2018	0.00483	0.00044
31/07/2018	0.03539	-0.00455
31/08/2018	0.02981	0.04961
28/09/2018	0.00429	0.01238
31/10/2018	-0.07193	-0.00651
30/11/2018	0.01770	-0.02362
31/12/2018	-0.09627	0.01299
31/01/2019	0.07574	0.02429
28/02/2019	0.02930	0.00529
29/03/2019	0.01777	0.03259
30/04/2019	0.03856	0.02038

Tabla 5.7: S&P500: predicción.

Ahora en la Figura 5.15, se presentan los resultados anteriores y notamos que el valor real del índice entre octubre del 2018 y enero del 2019, sube y baja. Ese cambio no lo predijo el modelo, sin embargo, fuera de ese periodo, la predicción es buena.

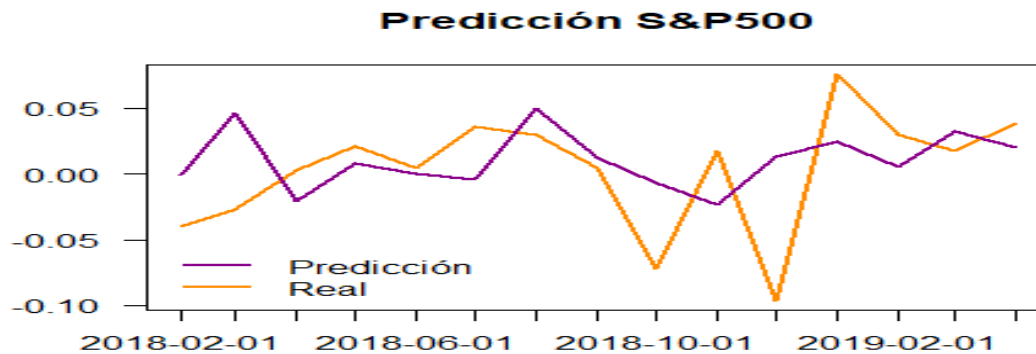


Figura 5.15: SP&500: predicción y valor real.

Ahora presentaremos el resultado de las métricas de riesgo con un nivel de confianza del 95 %. Ver Figura 5.16.

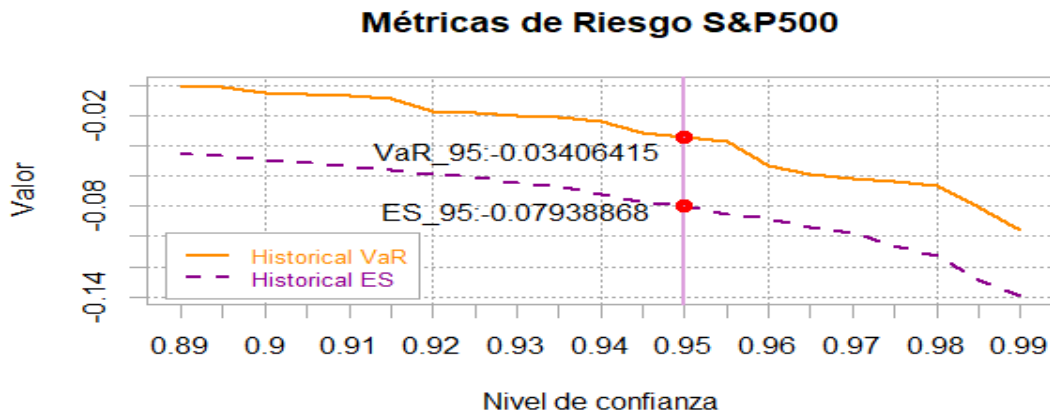


Figura 5.16: S&P500: métricas de riesgo.

El resultado en ambos casos es negativo, lo que es una excelente señal para inversionistas. Además, si fuéramos más estrictos y tomáramos un nivel de confianza mayor el resultado seguiría siendo favorable.

Por último presentaremos los resultados al aplicar la prueba de Kupiec con el $VaR_{95\%}$. La Tabla 5.8 confirma que el resultado es consistente por lo que la estimación del $VaR_{95\%}$ es confiable.

Índice	$\chi^2_{0.05}$	Test	P-valor
SP500	0.04	0.9768591	0.3229755

Tabla 5.8: S&P500: prueba de Kupiec.

5.4. Deutscher Aktienindex

En tercer lugar presentaremos los resultados de aplicar un HMM al Deutscher Aktienindex (DAX). En la Figura 5.17 se muestra la serie original y la serie al

aplicar la transformación *log return*.

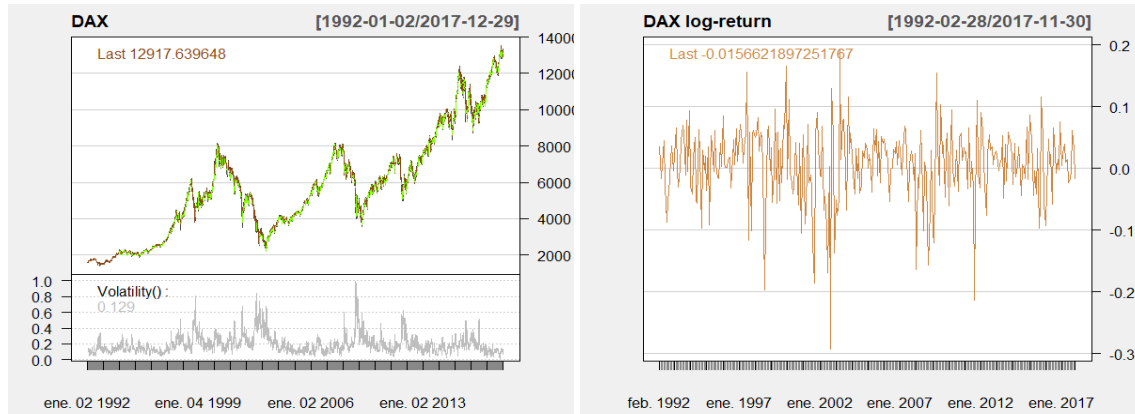


Figura 5.17: DAX: serie original y transformada.

En la primera parte de la gráfica podemos notar que el DAX comienza con una tendencia creciente, esto ya que en Alemania los bancos Deutsche Bank y Dresdner estaban a punto de fusionarse para convertirse en la segunda mayor entidad bancaria del mundo. Esto se dio en febrero del 2000, que coincide con el primer punto de inflexión importante del DAX; después tuvo una tendencia decreciente hasta marzo del 2003, año en el que pasó la llamada agenda 2010, que se refiere a un conjunto de reformas implementadas por Gerhard Schröder, planeadas justamente para deducir costos en el modelo social alemán y funcionó porque el índice tuvo tendencia creciente de nuevo.

Tuvo otro retroceso en el 2008 y eso por la fuerte crisis a la que se enfrentó Estados Unidos, después de esa fecha ha tenido pequeñas caídas, pero hasta ahora se mantiene a la alta.

Ahora vamos a implementar el HMM a este índice con diferente número de estados. En la Tabla 5.9 podemos notar que el mejor modelo será con dos estados, que converge con solo 24 iteraciones.

DAX Alemania				
Estado	Iteraciones	LogLik	AIC	BIC
2	24	465.18	-916.36	-890.21
3	269	469.34	-910.68	-858.37
4	136	477.16	-908.31	-822.37
5	204	478.91	-889.83	-762.78
6	424	488.68	-883.37	-707.75
7	222	490.27	-856.54	-624.87
8	398	492.16	-826.32	-531.13

Tabla 5.9: DAX: selección del mejor modelo.

Las estimaciones de los parámetros $\hat{\Delta}$ y $\hat{\Gamma}$ que ajustan este modelo son, probabilidad inicial

$$\hat{\Delta} = \{1.955293e - 17, 1\}, \quad (5.5)$$

y la matriz de transición

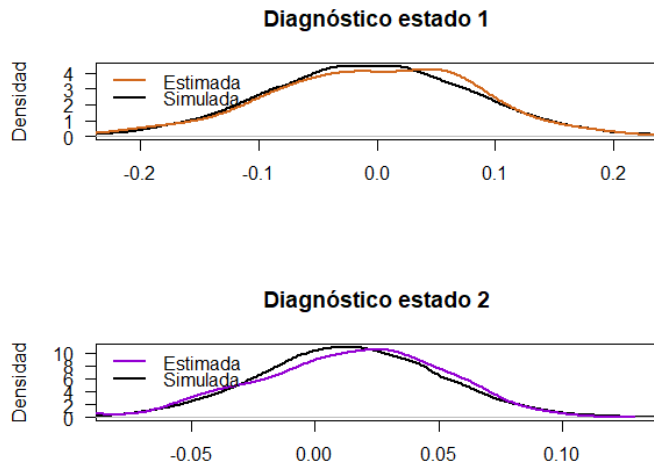
$$\hat{\Gamma} = \begin{pmatrix} 0.92463718 & 0.07536282 \\ 0.04882395 & 0.95117605 \end{pmatrix}, \quad (5.6)$$

y las estimaciones de los parámetros que corresponden a la distribución de cada uno de los estados, junto con el número de observaciones de cada estado se presenta en la Tabla 5.10.

	Número de observaciones	Media	Varianza
Estado 1	111	-0.009161226	0.006859127
Estado 2	199	0.016689481	0.001289495

Tabla 5.10: DAX: descripción por estado.

Podemos ver que a diferencia del IPC en que la media de cada una de las distribuciones era positiva, en el DAX sí hay un cambio respecto al signo. Veamos el diagnóstico de cada una de las distribuciones al igual que el análisis de cada uno de los estados.



En la Figura 5.18 podemos ver que las observaciones del estado 1 tienen mayor dispersión que los del estado 2, de hecho en el estado dos se concentran entre -0.05 y 0.1. Lo interesante es que el estado dos tiene 88 datos más que el otro estado.

Figura 5.18: DAX: diagnóstico por estado.

Ahora veremos la probabilidad de que cada observación esté en alguno de los estados; esto se presenta en la Figura 5.19.

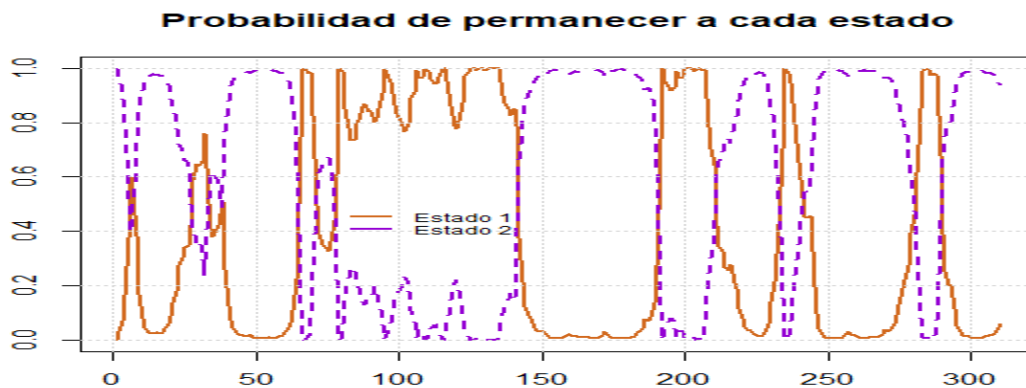


Figura 5.19: DAX: probabilidad de estar en cada uno de los estados.

Notamos que a diferencia del SP500, en el DAX sí podemos observar cuál es el estado más probable por cada uno de los datos. Una vez que tenemos estas probabilidades podemos aplicar el algoritmo de Viterbi, ver la decodificación y

dar un análisis por estado.

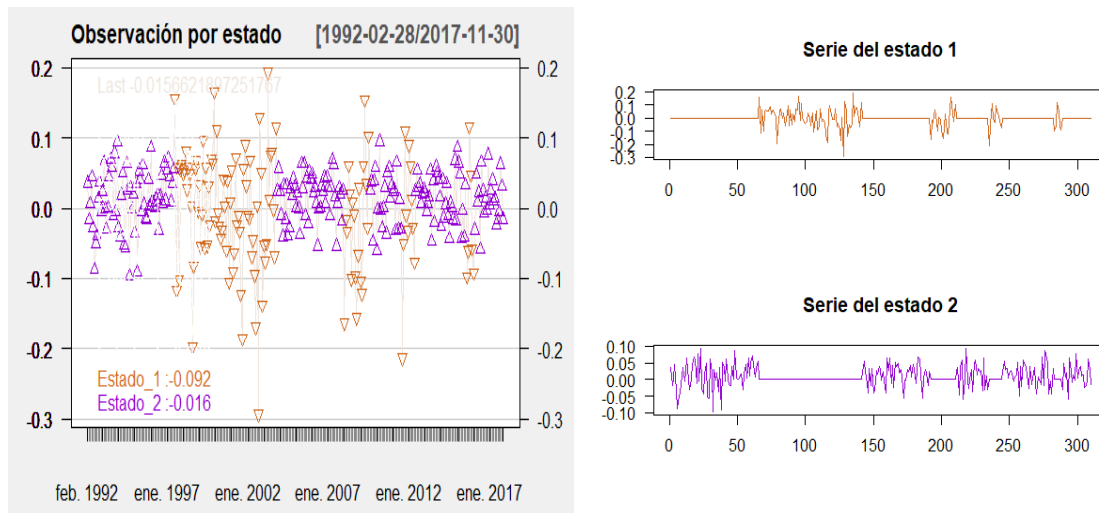


Figura 5.20: DAX: clasificación por estado.

En la Figura 5.20 podemos observar que el estado 1 se caracteriza por ser más volátil ya que el estado 2 se encuentra entre la franja -0.1 y 0.1 para todos los casos.

El primer periodo que va de enero de 1992 hasta junio de 1997 se encuentra en el estado 2, lo que implica que el índice tenía una estabilidad pero para 1998 la coalición entre el Partido Socialdemócrata de Alemania (SDP) y Alianza 90 ganó las elecciones para gobernar el país. Con ello vinieron grandes cambios y también empezaron las alianzas del 2000 que mencionamos anteriormente, por lo que el índice se desestabilizó y es hasta noviembre del 2003 donde vuelve a cambiar al estado 2.

Cinco años después y gracias a la crisis económica de Estados Unidos, se desestabiliza el índice y fue hasta que el gobierno inyectó cientos de miles de millones de euros en el sistema financiero para respaldar a los bancos alemanes en septiembre del 2009, que se estabilizó de nuevo pero a finales del 2011 y el primer semestre del 2012 se desestabiliza debido a que el presidente electo Christian Wulff tuvo escándalos de corrupción y tráfico de influencias, sin embargo, su mandato estaba

por terminar y en marzo Joachim Gauck fue electo presidente de Alemania.

Después de eso podemos ver que entra el estado 2, salvo en un pequeño intervalo que va de agosto del 2015 a enero del 2016 y fue debido a que las exportaciones cayeron un 5.22% comparadas con julio del 2015.

Una vez descrito cada uno de los estados, podemos observar los estados en los datos al cierre y lo presentaremos en la Figura 5.21.

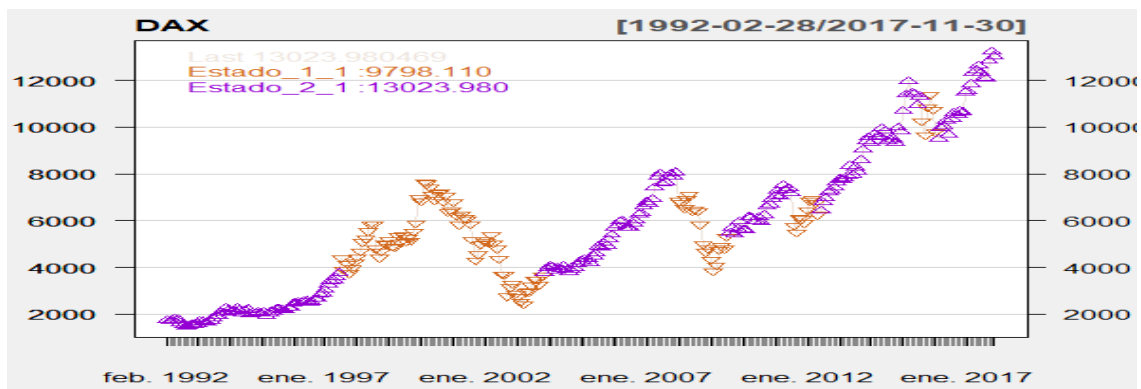


Figura 5.21: DAX: clasificación en los precios al cierre.

Por lo que podemos decir que el estado 1 corresponde a las caídas que ha tenido el índice, lo que involucra mayor volatilidad.

Por otro lado, haremos una predicción a 60 periodos a partir de enero del 2018 con las estimaciones de los parámetros de media y varianza de cada una de las distribuciones presentadas en la Tabla 5.10, la probabilidad inicial (5.5) y la matriz de transición (5.6).

Los resultados se muestran en la Figura 5.22, el color café representa a la serie de *log return* del DAX y de color morado la predicción que abarca de enero del 2018 a diciembre del 2022. De esa gráfica, de los 60 periodos se clasificó a 12 en el estado 1 y a 48 en el estado 2, lo que implicaría un panorama positivo.

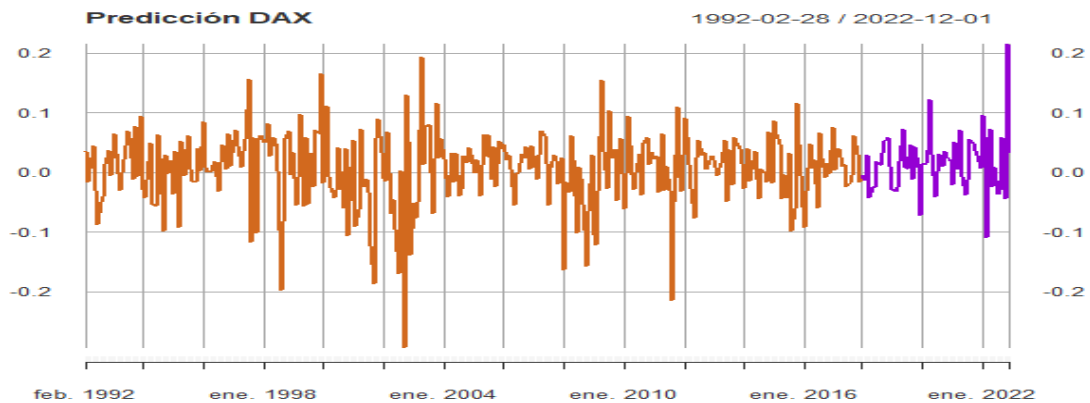


Figura 5.22: DAX: predicción.

Del mismo modo que los índices anteriores, obtendremos los datos reales del DAX de todo el 2018 y el primer cuatrimestre del 2019 para poder evaluar nuestro modelo y los resultados se muestran a continuación.

DAX Alemania		
Fecha	Logret	Predicción
28/02/2018	-0.05884	-0.00439
28/03/2018	-0.02765	-0.01285
30/04/2018	0.04172	0.02963
31/05/2018	-0.00057	-0.04193
29/06/2018	-0.02400	-0.02575
31/07/2018	0.03979	-0.02559
31/08/2018	-0.03508	0.01805
28/09/2018	-0.00953	0.01180
31/10/2018	-0.06749	0.03236
30/11/2018	-0.01676	0.05201
31/12/2018	-0.06404	0.05911
31/01/2019	0.05653	0.01103
28/02/2019	0.03020	-0.02829
29/03/2019	0.00090	-0.03122
30/04/2019	0.06857	-0.02523

Tabla 5.11: DAX: predicción.

En la Tabla 5.11 y en la Figura 5.23 se presentan los resultados y al calcular un error cuadrático medio obtenemos un resultado de 0.385 %, lo que es un resultado muy bueno, por lo que concluimos que el modelo es bueno.

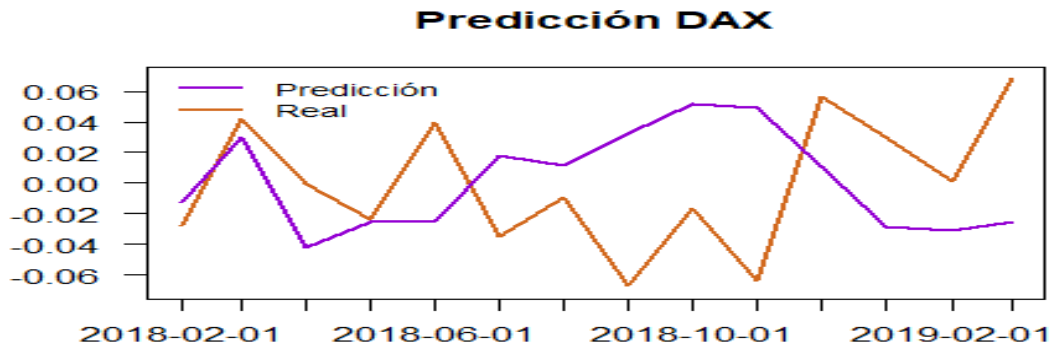


Figura 5.23: DAX: predicción y valor real.

Como siguiente punto presentaremos el resultado de las métricas de riesgo financiero con un nivel de confianza del 95 %.

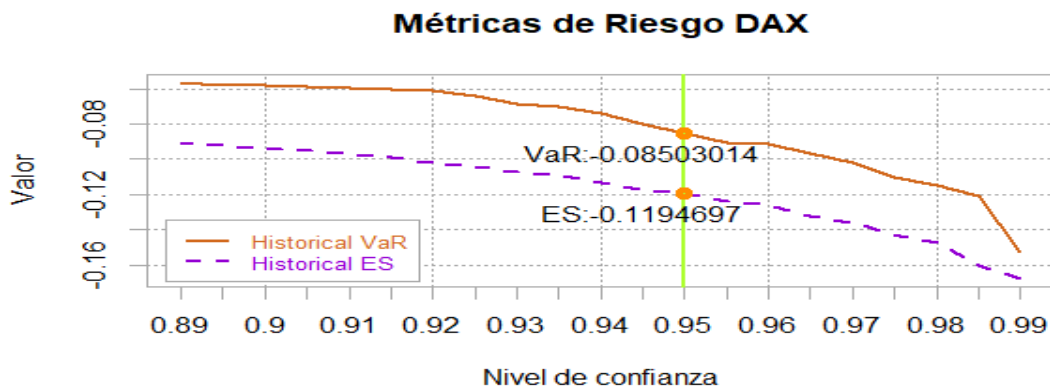


Figura 5.24: DAX: métricas de riesgo.

En la Figura 5.24 notamos que tanto el $VaR_{95\%}$ como el $CVaR_{95\%}$ son ne-

gativos, lo cual implica un buen panorama para inversionistas ya que estaríamos esperando una ganancia.

Por último, presentamos el resultado de aplicar la prueba de Kupiec para poder conocer si el resultado que nos arroja el $VaR_{95\%}$ es consistente. Con los resultados de la Tabla 5.12, concluimos que con un p-valor de 0.171 y un valor de la prueba de 1.87 no podemos rechazar H_0 a un nivel de significancia del 5%.

Índice	$\chi^2_{0.05}$	Test	P-valor
DAX	0.04	1.8721442	0.1712297

Tabla 5.12: DAX: prueba de Kupiec.

5.5. Nikkei Heikin Kabuka

Por último daremos los resultados del índice Nikkei Heikin Kabuka (Nikkei225). Primero observamos los precios al cierre y al aplicar la transformación $\log return$, esto en la Figura 5.25.

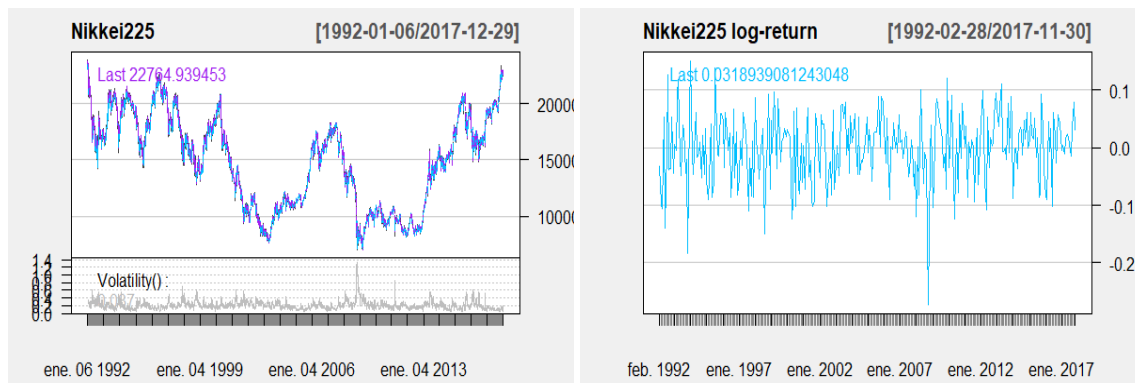


Figura 5.25: Nikkei225: serie original y transformada.

A simple vista, notamos que a diferencia del resto de los índices éste no tiene una tendencia creciente al principio del periodo; al contrario en febrero de 1992 tenemos un valor al cierre de 21338.81, el cual es altísimo pues el máximo de todo

el periodo es noviembre del 2017 con un valor de 22724.96.

La razón por la que el índice comienza tan fuerte es que entre 1986 y 1992 ocurrió el acontecimiento conocido como “la burbuja financiera e inmobiliaria”, con la cual el índice estaba muy elevado, pero justamente a principios de 1992 esta burbuja estalló y Japón entró en la llamada década perdida, en la que el índice iba en bajada con un valor crítico en abril del 2003.

Después se fortaleció el índice, en parte gracias a la copa mundial de futbol que se festejó en Japón y Corea del Sur. El índice se volvió a fortalecer hasta marzo del 2006, año en el que el gobernador del banco de Japón, Toshihiko Fukui, aceptó haber invertido en un fondo bursátil ilegal de Yoshiaki Murakami, quien era el gestor más famoso de la Bolsa de Tokio, lo que provocó una caída importante en el índice.

Por otro lado, la gran recesión golpeó al índice e hizo que se alcanzara el valor mínimo de todo el periodo que elegimos en febrero del 2009. A partir de ese punto el índice se ha fortalecido alcanzando por poco el valor que se tenía antes de la década perdida.

Ahora aplicaremos el HMM a los datos *log return* del Nikkei225 y los resultados los presentaremos a continuación.

Nikkei225 Japón				
Estado	Iteraciones	LogLik	AIC	BIC
2	45	448.92	-883.85	-857.70
3	213	451.69	-875.39	-823.07
4	387	457.11	-868.23	-782.29
5	144	463.64	-859.28	-732.24
6	123	469.43	-844.86	-669.24
7	191	478.33	-832.67	-601.01
8	213	482.12	-806.24	-511.05

Tabla 5.13: Nikkei225: selección del mejor modelo.

Como podemos notar en la Tabla 5.13, el mejor modelo para estos datos se da con 2 estados, que converge con 45 iteraciones, lo que hace que el proceso sea rápido respecto a más estados. Las estimaciones de los parámetros $\hat{\Delta}$ y $\hat{\Gamma}$ que ajustan este modelo son, probabilidad inicial

$$\hat{\Delta} = \{1.507832e - 47, 1\}, \quad (5.7)$$

y matriz de transición

$$\hat{\Gamma} = \begin{pmatrix} 0.90439625 & 0.09560375 \\ 0.07126585 & 0.92873415 \end{pmatrix}. \quad (5.8)$$

Los parámetros de media y varianza que describen el comportamiento de cada uno de los estados al igual que el número que los conforman se muestran en la Tabla 5.14.

	Número de observaciones	Media	Varianza
Estado 1	121	0.02077252	0.001477693
Estado 2	189	-0.01435506	0.004342669

Tabla 5.14: Nikkei225: descripción por estado.

Podemos observar que en el estado 1 la media es positiva aunque la varianza es menor que en el estado 2, lo cual puede ser por el número de integrantes de cada estado, así que comencemos analizando el diagnóstico de cada estado y veamos qué pasó.

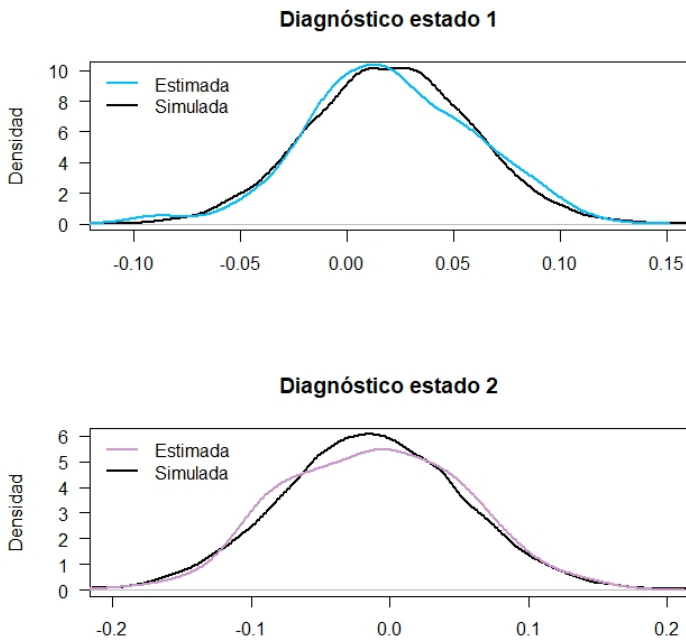


Figura 5.26: Nikkei225: diagnóstico por estado.

Ahora veremos la probabilidad de que cada uno de los datos que se encuentren en cada uno de los estados, esto en la Figura 5.27, que podemos observar que va por bloques la sucesión de estados más probables.

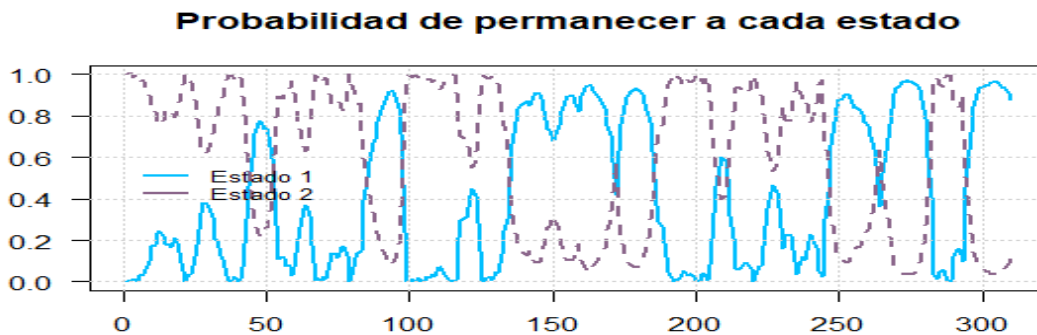


Figura 5.27: Nikkei225: probabilidad de estar en cada uno de los estados.

En este índice podemos notar cuál es el estado más probable por observación, salvo en algunos puntos en los que se interceptan y la diferencia es muy pequeña; sin embargo, sí se nota que el estado 2 predomina.

Una vez que tenemos estas probabilidades podemos aplicar el algoritmo de Viterbi, ver la decodificación y dar un análisis por estado.

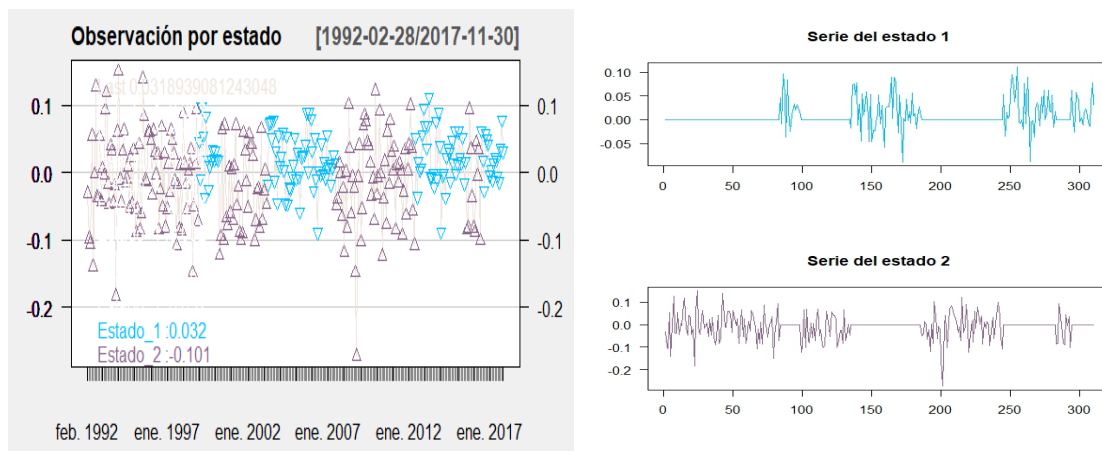


Figura 5.28: Nikkei225: clasificación por estado.

En la Figura 5.28 podemos observar que el estado dos representa el estado más volátil y bajo contexto tiene sentido, pues como ya lo habíamos notado la situación que enfrentaba Japón a principios del periodo era muy complicada, estaba en declive y justo lo decodifica en el estado 2.

El cambio de estado se da en enero de 1999 que se estabilizó hasta marzo del 2000 pero vuelve a cambiar debido a los escándalos políticos; después se volvió a estabilizar hasta que llegó la crisis en Estados Unidos que golpeó la economía mundial, en particular el Nikkei225 se estabilizó de nuevo hasta mayo del 2012, pero en agosto del 2015 se registró su mayor subida en 7 años, lo que desestabilizó el índice porque después volvió a caer hasta que en julio del 2016 se estabilizó hasta que terminó el periodo.

Una vez que vimos la decodificación por cada uno de los datos, observamos las trayectorias en los precios de cierre.

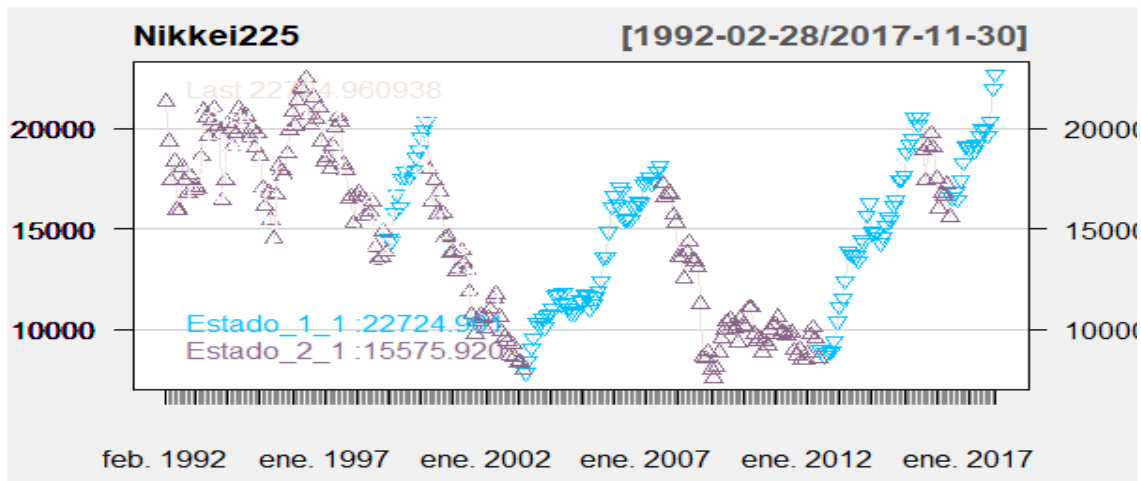


Figura 5.29: Nikkei225: clasificación en los precios al cierre.

En la Figura 5.29 confirmamos lo que veníamos planteando pero lo bonito de esta gráfica es que notamos que a partir de enero de 1999 el estado 1 involucra a las subidas del índice mientras que el estado 2 representa a las caídas del mismo, por lo que la mayor volatilidad se da cuando el índice va a la baja.

Como siguiente paso haremos la predicción a 60 periodos de este índice a partir de la media y varianza de cada una de las distribuciones presentadas en la Tabla 5.14, la probabilidad inicial (5.7) y la matriz de transición (5.8). La haremos desde enero del 2018 hasta diciembre del 2022.

En la Figura 5.30 se presenta la predicción, de los 60 periodos que predijo el modelo clasificó a 36 en el estado 1 y a 24 en el estado 2, entonces a pesar de que hay más observaciones en el estado 1 no es una diferencia significativa. Se presenta en color azul la serie de *log return* del Nikkei225 y de morado la predicción.

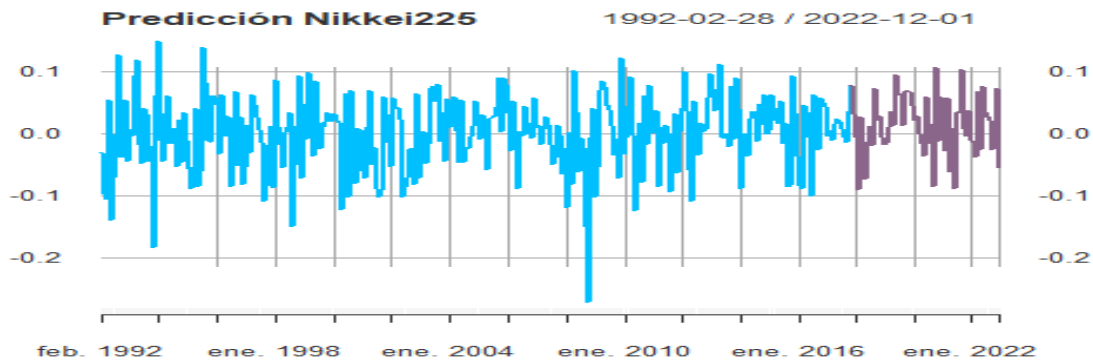


Figura 5.30: Nikkei225: predicción.

Al igual que los demás índices obtendremos los valores al cierre del 2018 y del primer cuatrimestre, los transformaremos y los utilizaremos como el conjunto de valores de prueba y con ello evaluaremos nuestro modelo.

NIKKEI225 Japón		
Fecha	Logret	Predicción
28/02/2018	-0.04562	-0.00479
28/03/2018	-0.02821	-0.09102
30/04/2018	0.04616	0.02625
31/05/2018	-0.01191	-0.04749
29/06/2018	0.00461	-0.07296
31/07/2018	0.01111	-0.01407
31/08/2018	0.01371	0.01061
28/09/2018	0.05343	-0.01967
31/10/2018	-0.09562	0.07230
30/11/2018	0.01945	0.02196
31/12/2018	-0.11040	0.02845
31/01/2019	0.03721	-0.00514
28/02/2019	0.02902	-0.01830
29/03/2019	-0.00842	-0.01503
30/04/2019	0.04846	0.00894

Tabla 5.15: Nikkei225: predicción.

En la Tabla 5.15, obtenemos el valor de la predicción junto con el valor real y al

calcular el error cuadrático medio obtenemos un valor de 0.689% el cual es muy bueno, por lo que podríamos decir que a pesar de la dificultad de la dinámica de este índice, el modelo predice bien.

En la Figura 5.31 podemos ver que en octubre y diciembre del 2018 se da la mayor diferencia entre el valor real y la predicción.

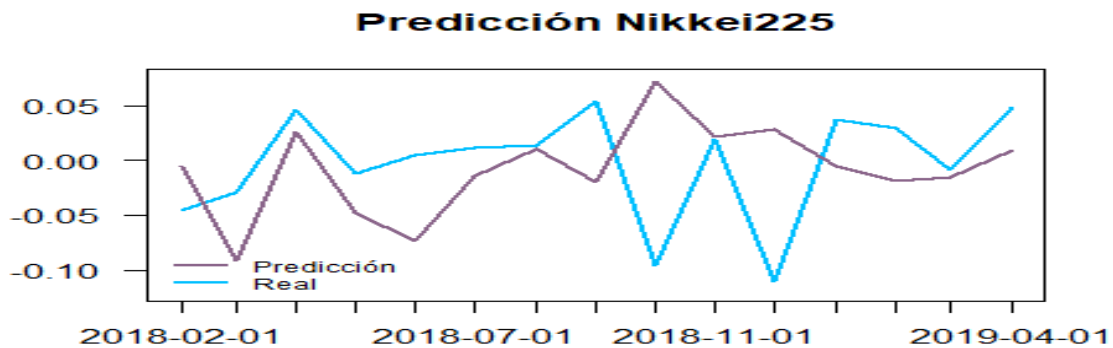


Figura 5.31: Nikkei225: predicción y valor real.

Ahora presentamos el resultado de las métricas de riesgo financiero.

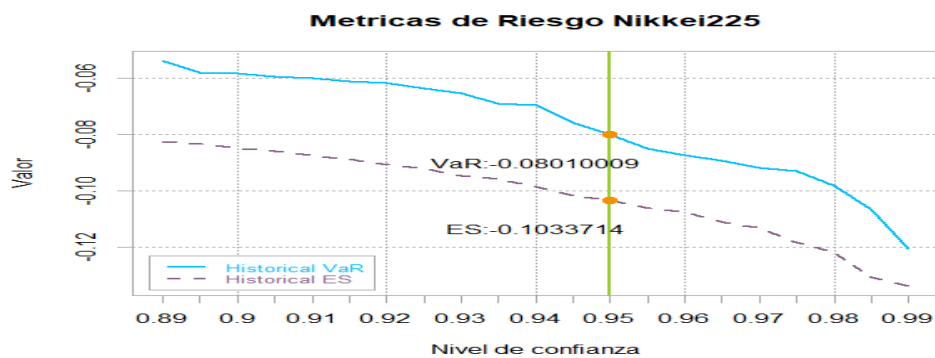


Figura 5.32: Nikkei225: métricas de riesgo.

En la Figura 5.32 notamos que ambas medidas son negativas lo cual financiera-

mente es muy bueno para el índice. En la gráfica se muestra el resultado de un nivel de confianza del 95 %; sin embargo, mostramos la trayectoria de ambas métricas a partir de un nivel de confianza del 89 % al 99 %, dependiendo de qué tan estricto sea la persona interesada puede conocer el VaR_α y el $CVaR_\alpha$ que deseé.

Por último presentamos el resultado de aplicar la prueba de Kupiec para poder conocer si el resultado que nos arroja el $VaR_{95\%}$ es consistente.

Índice	$\chi_{0.05}^2$	Test	P-valor
Nikkei225	0.04	0.4737990	0.4912446

Tabla 5.16: Nikkei225: prueba de Kupiec.

Por lo que podemos asegurar que con valor de prueba de 0.4737 y un p-valor de 0.4912 aceptamos H_0 a un nivel de confianza del 95 %, lo que significa que el cálculo del $VaR_{95\%}$ es consistente.

6.1. Conclusiones

A lo largo de este trabajo hemos analizado las principales características de los cuatro índices bursátiles elegidos y las medidas más ocupadas de riesgo financiero, pero sobre todo la estructura, componentes y forma de aplicar los modelos ocultos de Markov.

Aprendimos que la dificultad de este modelo es poder obtener la mejor decodificación, maximizar la verosimilitud y al mismo tiempo ajustar los parámetros. Sin embargo, solucionamos esto al implementar el algoritmo de Viterbi, *Forward-Backward* y el de Baum-Welch, respectivamente. El costo computacional que requiere el aplicar estos algoritmos se incrementa considerablemente al aumentar el número de observaciones y/o estados, por lo que al trabajar con 310 datos y 2 ó 3 estados no representa una dificultad.

Presentamos la aplicación del modelo en el capítulo 5 y descubrimos ciertos detalles interesantes:

- Todos los modelos están programados hasta 8 estados, pero en 3 casos el mejor modelo fue con dos estados, salvo el S&P500 que es con tres estados. El mejor modelo de cada índice converge rápido.
- La dinámica de cada índice es completamente diferente; sin embargo, en todos los casos se adapta de una manera adecuada. En el S&P500 vemos que primero hace una diferencia entre baja o alta volatilidad y después a los que tienen poca volatilidad los separa entre un *log return* positivo o negativo, lo cual hace que tenga una interpretación sencilla. Para los tres casos restantes se divide entre volatilidad alta y baja.
- Una vez que hacemos la predicción y calculamos el error cuadrático medio teníamos los resultados en la Tabla 6.1.

	IPC	S&P500	DAX	Nikkei225
ECM	0.396%	0.209%	0.385%	0.689%

Tabla 6.1: Error cuadrático medio en la predicción.

Por lo que en todos los casos el error es pequeño, pero en donde más se equivoca es en el Nikkei225, que tiene sentido porque es el índice en el que la tendencia cambia todo el tiempo y por contexto histórico e influencia de su gobierno es muy complicado comprenderlo. A pesar de esto el HMM aprendió bien, capta los movimientos bruscos del *log return* y la predicción no se ve mal tal como se muestra en la Figura 5.30. Para mayor seguridad de las predicciones habría que ampliar el conjunto de datos de prueba y obtener el ECM.

- Con base en la predicción y utilizando el algoritmo de Viterbi para decodificar de nuevo las nuevas observaciones:
 - IPC: 41 nuevos datos los clasifica en el estado uno que representa mucha volatilidad.

- S&P500: 33 observaciones se clasifican en el estado donde el *log return* es positivo, lo que podría significar un buen panorama para EUA pero tenemos 22 en el estado de mucha volatilidad y 5 en poca volatilidad pero con *log return* negativo.
 - DAX: se clasificó a 48 nuevos datos en el estado 2 por lo que Alemania tiene un panorama positivo.
 - Nikkei225: tenemos 36 observaciones en el estado 1 que significa poca volatilidad pero no es una gran diferencia respecto al estado 2.
- Después de tener el mejor modelo, analizarlo y describirlo, obtuvimos las medidas de riesgo financiero y obtuvimos los siguientes resultados presentados en la Tabla 6.2.

	IPC	S&P500	DAX	Nikkei225
$VaR_{95\%}$	-0.06350	-0.03406	-0.08503	-0.08010
$CVaR_{95\%}$	-0.09940	-0.07938	-0.11946	-0.10337

Tabla 6.2: Medidas de riesgo financiero.

En todos los casos tenemos resultados negativos, lo que significa que en 5 de los peores escenarios esperaríamos una ganancia. Debemos notar que el S&P500, a pesar de tener el error de predicción más chico, también tiene medidas de riesgo financiero más pequeño. Por otro lado el DAX tiene el segundo error más pequeño pero tiene las medidas de riesgo financiero más negativas.

- Por último al aplicar la prueba de Kupiec para ver si los resultados del $VaR_{95\%}$ son consistentes, notamos que los cuatro índices bursátiles pasan la prueba de Kupiec, por lo que los resultados del $VaR_{95\%}$ son consistentes.

Con base en los resultados del modelo, su clasificación, la predicción y las medidas de riesgo financiero concluimos que el mejor mercado para invertir durante los próximos 3 años es **Alemania**, además que cuenta con el sistema *XETRA* que hace transparente sus resultados.

6.2. Trabajo futuro

Como continuación de esta tesis se podría extender el período de prueba y comparar la predicción con el valor real del índice, esto con el fin de seguir evaluando nuestro modelo.

Por otra parte, se podría ajustar un modelo de tipo GARCH con el mismo conjunto de datos para obtener las predicciones y ver la eficiencia de los modelos ocultos de Markov respecto al modelo de series de tiempo; con ello probaríamos si los HMM son una buena alternativa para estudiar los índices bursátiles.

APÉNDICE A

Código de R

A continuación presentamos el código de R de manera general, así se obtienen los resultados presentados en esta tesis para cada índice.

```
#Cargamos las librerías.
library(RHmm)
library(quantmod)
library(PerformanceAnalytics)
library(TTR)
library(ldhmm)
library(xts)
library(stats)
library(plotrix)
library(ggplot2)
library(VaR)
library(GAS)
par(mfrow=c(1,1))
#Obtenemos los datos
Sys.setenv(tz = "UTC")
INDICE<-getYahooData("clave_indice", start = 19920101, end = 20180101)
INDICE <- na.exclude(INDICE)
length(INDICE$Close)
#Graficamos los precios al cierre
chartSeries(INDICE,name="INDICE", lwd=3, up.col="COLOR1", dn.col="COLOR2",
```

```
TA=c(addVolatility()), theme = "white")
#Obtenemos el log return con el que trabajaremos
f_meses <- endpoints(INDICE, on = "months", k = 1)
INDICE <- INDICE[f_meses[2:(length(f_meses)-1)]]
INDICE$logret <- log(INDICE$Close) - lag(log(INDICE$Close))
#Graficamos el log return
chartSeries(INDICE$logret, name="INDICE_log-return", theme = "white")
#Comprobamos que no haya datos faltantes
table(is.na(INDICE))
#Ajustamos el modelo para:
#dos estados
HMM_DOS<-HMMFit(obs=INDICE$logret, nStates = 2)
vit2<-viterbi(HMM_DOS, INDICE$logret)
summary(vit2)
table(vit2$states)
#tres estados
HMM_TRES<-HMMFit(obs=INDICE$logret, nStates = 3)
vit3<-viterbi(HMM_TRES, INDICE$logret)
summary(vit3)
table(vit3$states)
#cuatro estados
HMM_CUATRO<-HMMFit(obs=INDICE$logret, nStates = 4)
vit4<-viterbi(HMM_CUATRO, INDICE$logret)
summary(vit4)
table(vit4$states)
#cinco estados
HMM_CINCO<-HMMFit(obs=INDICE$logret, nStates = 5)
vit5<-viterbi(HMM_CINCO, INDICE$logret)
summary(vit5)
table(vit5$states)
#seis estados
HMM_SEIS<-HMMFit(obs=INDICE$logret, nStates = 6)
vit6<-viterbi(HMM_SEIS, INDICE$logret)
summary(vit6)
table(vit6$states)
#siete estados
HMM_SIETE<-HMMFit(obs=INDICE$logret, nStates = 7)
vit7<-viterbi(HMM_SIETE, INDICE$logret)
summary(vit7)
table(vit7$states)
#ocho estados
```

```
HMM_OCHO<-HMMFit(obs=INDICE$logret, nStates = 8)
vit8<-viterbi(HMM_OCHO, INDICE$logret)
summary(vit8)
table(vit8$states)
#Vemos el AIC y BIC de cada modelo
(c(HMM_DOS$AIC,HMM_TRES$AIC,HMM_CUATRO$AIC,HMM_CINCO$AIC,HMM_SEIS$AIC,
HMM_SIETE$AIC,HMM_OCHO$AIC))
(c(HMM_DOS$BIC,HMM_TRES$BIC,HMM_CUATRO$BIC,HMM_CINCO$BIC,HMM_SEIS$BIC,
HMM_SIETE$BIC,HMM_OCHO$BIC))
#Elegimos el mejor modelo e imprimimos su contenido
HMM_BM<-HMMFit(obs=INDICE$logret, nStates = 3)
vit_BM<-viterbi(HMM_BM, INDICE$logret)
summary(vit_BM)
table(vit_BM$states)
HMM_BM
```

Dentro de los 4 modelos que tenemos el de mayor número de estados es 3 por lo que se presentará el código de esa forma. Sigamos con el diagnóstico por estado.

```
INDICE_aux<-cbind(INDICE$logret,vit_BM$states)
Estado_1=INDICE_aux[INDICE_pred[,2]==1,1]
Estado_2=INDICE_aux[INDICE_pred[,2]==2,1]
Estado_3=INDICE_aux[INDICE_pred[,2]==3,1]
norma1<-rnorm(40000,mean(Estado_1),sd(Estado_1))
norma2<-rnorm(40000,mean(Estado_2),sd(Estado_2))
norma3<-rnorm(40000,mean(Estado_3),sd(Estado_3))
par(mfrow=c(3,1))
plot(density(norma1), xlab="", main="Estado1", ylab="Densidad")
lines(density(Estado_1), lwd=2, col="COLOR1")
legend("topleft", legend=c("Estimada", "Simulada"),
col=c("COLOR1", "black"),
box.lty=2, box.lwd=2, bty='n', lwd=2)
plot(density(norma2), xlab="", main="Estado2", ylab="Densidad")
lines(density(Estado_2), lwd=2, col="COLOR2")
legend("topleft", legend=c("Estimada", "Simulada"),
col=c("COLOR2", "black"),
box.lty=2, box.lwd=2, bty='n', lwd=2)
plot(density(norma3), xlab="", main="Estado3",)
lines(density(Estado_3), lwd=2, col="COLOR3")
legend("topleft", legend=c("Estimada", "Simulada"),
col=c("COLOR3", "black"),
```

```

box.lty=2, box.lwd=2, bty='n', lwd=2)
#Ahora graficaremos la serie por estado
aux=NULL
for (i in 1:310) { if(vit_BM$states[i]==1) {
aux[i]=INDICE$logret[i]} else{aux[i]=0}}
aux2=NULL
for (i in 1:310) { if(vit_BM$states[i]==2) {
aux2[i]=INDICE$logret[i]} else{aux2[i]=0}}
aux3=NULL
for (i in 1:310) { if(vit_BM$states[i]==3) {
aux3[i]=INDICE$logret[i]} else{aux3[i]=0}}
par(mfrow=c(3,1))
plot(aux, type="l", main="Serie del estado 1", col="COLOR1", las=1)
plot(aux2, type="l", main="Serie del estado 2", col="COLOR2", las=1)
plot(aux3, type="l", main="Serie del estado 3", col="COLOR3", las=1)
par(mfrow=c(1,1))
#Como siguiente paso vemos la probabilidad de estar en cada estado
fb = forwardBackward(HMM_BM, INDICE$logret)
matplot(fb$Gamma, type='l', main='Probabilidad', col=c("COLOR1",
"COLOR2", "COLOR3"), lwd=2, las=1)
grid()
legend(x=51,y=0.8, legend=c("Estado 1", "Estado 2","Estado 3"),
col=c("COLOR1","COLOR2", "COLOR3"), cex=0.8, box.lty=2,
box.lwd=2, bty='n', lwd=2)
#Ahora pondremos cada una de las observaciones en el estado que
#corresponda
chartSeries(INDICE_aux[,1], theme = "white",up.col="seashell2")
addTA(Estado_1, on=1, type="p",col="COLOR1", pch=25)
addTA(Estado_2, on=1, type="p",col="COLOR2", pch=24)
addTA(Estado_3, on=1, type="p",col="COLOR3", pch=23)
#De igual forma lo hacemos sobre los precios al cierre
INDICE_aux_1<-cbind(INDICE$Close,vit_BM$states)
Estado_1_1=INDICE_aux_1[INDICE_pred_1[,2]==1,1]
Estado_2_1=INDICE_aux_1[INDICE_pred_1[,2]==2,1]
Estado_3_1=INDICE_aux_1[INDICE_pred_1[,2]==3,1]
chartSeries(INDICE_aux_1[,1], theme = "white", name="INDICE")
addTA(Estado_1_1, on=1, type="p",col="COLOR1", pch=25)
addTA(Estado_2_1, on=1, type="p",col="COLOR2", pch=24)
addTA(Estado_3_1, on=1, type="p",col="COLOR3", pch=23)
#VaR
hmmRets <- HMMCont::logreturns(HMM_BM$obs)

```

```
chart.VarSensitivity(hmmRets,
methods=c("HistoricalVaR", "HistoricalES"),
colorset=c("COLOR1", "COLOR2"), lwd=2)
abline(v=0.95, col="greenyellow", lwd=2)
points(0.95, VAR, col="COLOR1", lwd=4)
points(0.95, CVAR, col="COLOR2", lwd=4)
VaR(hmmRets, method="historical")
ES(hmmRets, method="historical")
#Prueba de Kupiec
Kupiec <- function(datos_ret, p) {
N = length(datos_ret)
x = sum(datos_ret)
p_gorro = x/N
test = -2 * log(((1 - p)^(N - x) * p^x)/
((1 - p_gorro)^(N - x) * p_gorro^x))
pvalue = 1 - pchisq(test, 1)
resultados = c(test, pvalue)
names(resultados) = c("Test", "P-valor")
return(resultados)
}
#Prediccion
SetHMM=distributionSet("NORMAL", mean=HMM_BM$HMM$distribution$mean,
var=HMM_BM$HMM$distribution$var)
initProb=HMM_BM$HMM$initProb
transMat=HMM_BM$HMM$transMat
HMM_comp=HMMSet(initProb, transMat, SetHMM)
pred_HMM=HMMSim(60, HMM_comp)
table(pred_HMM$states)
aux<-c(HMM_BM$obs, pred_HMM)
plot(aux, type="n")
lines(aux[1:310], col="COLOR1", lwd=2)
lines(aux[311:370], col="COLOR2", lwd=2)
#Veamos el valor real
INDICE_comprobar <- getYahooData("INDICE", start = 20180101,
end = 20221201)
INDICE_comprobar <- na.exclude(INDICE_comprobar)
length(INDICE_comprobar$Close)
#\textit{log return}
f_meses_1 <- endpoints(INDICE_comprobar, on = "months", k = 1)
INDICE_comprobar <- INDICE_comprobar[f_meses_1[2:(length(
f_meses_1)-1)]]
```

```
INDICE_comprobar$logret <- log(INDICE_comprobar$Close)
- lag(log(INDICE_comprobar$Close))
#View(INDICE_comprobar)
#Valor real vs lo que nos dio el modelo
predi_INDICE<- ts(aux[c(311:326)], frequency = 12, start = c(2018,2))
real_INDICE <- ts(INDICE_comprobar$logret[c(1:15)], frequency = 12,
start = c(2018,2))
predi_HMM_1=coredata(predi_INDICE)
real_obs=coredata(real_INDICE)
sec_aux=seq(as.Date("2018-02-01"),as.Date("2019-04-01"),by="months",
format="%m-%Y")
plot(sec_aux,real_obs, xlab = '',
ylab = '', type="l", las=1, lwd=2, col="COLOR1", xaxt="n")
invisible(lines(sec_aux,predi_HMM_1, type="l", lwd=2, col="COLOR2"))
legend("bottomleft", legend=c("Real","Prediccion"),
col=c("COLOR1","COLOR2"), cex=1, box.lty=2, box.lwd=2, bty='n', lwd=2)
lab=sec_aux
axis.Date(1,at=sec_aux, labels =sec_aux)
```

Bibliografía

- Acerbi, C. & Tasche, D. (2002), ‘On the coherence of expected shortfall’, *Journal of Banking & Finance* **26**(7), 1487–1503.
- Akaike, H., Parzen, E., Tanabe, K. & Kitagawa, G. (1998), *Selected Papers of Hirotugu Akaike*, Springer Science & Business Media.
- Artzner, P., Delbaen, F., Eber, J.-M. & Heath, D. (1999), ‘Coherent measures of risk’, *Mathematical finance* **9**(3), 203–228.
- Baggenstoss, P. M. (2001), ‘A modified Baum-Welch algorithm for hidden Markov models with multiple observation spaces’, *IEEE Transactions on speech and audio processing* **9**(4), 411–416.
- Banco de Pagos Internacionales (2004), ‘Convergencia internacional de medidas y normas de capital’, *Comité de Supervisión Bancaria de Basilea, Marco Revisado junio* .
- Basel Committee on Banking Supervision (1996), *Supervisory framework for the use of “backtesting” in conjunction with the internal models approach to market risk capital requirements*, Bank for International Settlements Basel.

-
- Baum, L. E., Petrie, T., Soules, G. & Weiss, N. (1970), 'A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains', *The Annals of Mathematical Statistics* **41**(1), 164–171.
- Bolsa Mexicana de Valores (2016), *Índice de Precios y Cotizaciones (IPC), Nota Metodológica*, Grupo BMV.
- Bolsa Mexicana de Valores (2018), *IPC Bolsa Mexicana – Cotización de Acciones*, Grupo BMV.
- Börse, D. (2018), 'Guide to the equity indices of Deutsche Börse', *Deutsche Börse AG* **60485**.
- Brockwell, P. J., Davis, R. A. & Calder, M. V. (2002), *Introduction to time series and forecasting*, Vol. 2, Springer.
- Burnham, K. P. & Anderson, D. R. (2002), *Model Selection and Multimodel Inference: A Practical Information-theoretic Approach*, 2nd edn, Springer.
- Burnham, K. P. & Anderson, D. R. (2004), 'Multimodel inference: Understanding AIC and BIC in model selection', *Sociological Methods & Research* **33**(2), 261–304.
- Campbell, J. Y., Lo, A. W., MacKinlay, A. C. et al. (1997), *The econometrics of financial markets*, Vol. 2, Princeton University press Princeton, NJ.
- Cappé, O., Moulines, E. & Rydén, T. (2005), *Inference in Hidden Markov Models*, Springer series in statistics, 1st edn, Springer.
- Chatfield, C. (2016), *The analysis of time series: an introduction*, Chapman and Hall/CRC.
- Chen, F.-S., Fu, C.-M. & Huang, C.-L. (2003), 'Hand gesture recognition using a real-time tracking method and hidden Markov models', *Image and vision computing* **21**(8), 745–758.
-

-
- Crouhy, M. (2005), *The Essentials of Risk Management*, 1st edn, McGraw-Hill.
- De Lara Haro, A. (2005), *Medición y control de riesgos financieros*, Editorial Limusa.
- Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977), ‘Maximum likelihood from incomplete data via the EM algorithm’, *Journal of the Royal Statistical Society: Series B (Methodological)* **39**(1), 1–22.
- Dymarski, P. (2011), *Hidden Markov models, theory and applications*, InTechOpen.
- Forney, G. D. (1973), ‘The Viterbi algorithm’, *Proceedings of the IEEE* **61**(3), 268–278.
- Francq, C. & Zakoian, J.-M. (2011), *GARCH models: structure, statistical inference and financial applications*, John Wiley & Sons.
- Frankfurt, B. (2018), *Index Composition Report*, Börse Frankfurt Group.
- Gupta, A. & Dhingra, B. (2012), Stock market prediction using hidden Markov models, in ‘2012 Students Conference on Engineering and Systems’, IEEE, pp. 1–4.
- Hassan, M. R. & Nath, B. (2005), Stock market forecasting using hidden Markov model: a new approach, in ‘5th International Conference on Intelligent Systems Design and Applications (ISDA’05)’, IEEE, pp. 192–196.
- Indices, S. D. J. (2019), *S&P U.S. Indices Methodology*, S&P Global.
- Insider, M. (2018), *S&P 500 STOCKS*, Insider Inc. and finanzen.net.
- Jorion, P. (2003), *Financial risk manager handbook*, 2nd edn, Wiley.
- Kim, S., Shephard, N. & Chib, S. (1998), ‘Stochastic volatility: likelihood inference and comparison with ARCH models’, *The Review of Economic Studies* **65**(3), 361–393.
-

- Kirchgässner, G. & Wolters, J. (2007), *Introduction to modern time series analysis*, Springer Science & Business Media.
- Kupiec, P. (1995), 'Techniques for verifying the accuracy of risk measurement models', *The J. of Derivatives* **3**(2), 24–95.
- Lukashin, A. V. & Borodovsky, M. (1998), 'Genemark. HMM: new solutions for gene finding', *Nucleic acids research* **26**(4), 1107–1115.
- MacDonald, I. L. & Zucchini, W. (1997), *Hidden Markov and other models for discrete-valued time series*, Monographs on statistics and applied probability 70, 1st edn, Chapman & Hall.
- McNeil, A. J., Frey, R., Embrechts, P. et al. (2005), *Quantitative Risk management: Concepts, techniques and tools*, Vol. 3, Princeton University Press.
- Nikkei (2017), *Nikkei Stock Average Index Guidebook*, Nikkei Inc.
- Niu, F. & Abdel-Mottaleb, M. (2005), HMM-based segmentation and recognition of human activities from video sequences, in '2005 IEEE International Conference on Multimedia and Expo', IEEE, pp. 804–807.
- Pflug, G. C. (2000), Some remarks on the value-at-risk and the conditional value-at-risk, in 'Probabilistic constrained optimization', Springer, pp. 272–281.
- Posada, D. & Buckley, T. R. (2004), 'Model Selection and Model Averaging in Phylogenetics: Advantages of Akaike Information Criterion and Bayesian Approaches Over Likelihood Ratio Tests', *Systematic Biology* **53**(5), 793–808.
- R Core Team (2017), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: <https://www.R-project.org/>
- Rabiner, L. R. (1989), 'A tutorial on hidden Markov models and selected applications in speech recognition', *Proceedings of the IEEE* **77**(2), 257–286.
-

-
- Ramaprasad Bhar, S. H. (2004), *Hidden Markov Models: Applications to Financial Economics*, Advanced Studies in Theoretical and Applied Econometrics 40, 1st edn, Springer US.
- Read, T. R. & Cressie, N. A. (2012), *Goodness-of-fit statistics for discrete multivariate data*, Springer Science & Business Media.
- Rincón, L. (2012), 'Introducción a los procesos estocásticos', *Ciudad de México: Universidad Nacional Autónoma de México* .
- Rockafellar, R. T. & Uryasev, S. (2000), 'Optimization of conditional value-at-risk', *Journal of risk* **2**, 21–42.
- Ross, S. M., Kelly, J. J., Sullivan, R. J., Perry, W. J., Mercer, D., Davis, R. M., Washburn, T. D., Sager, E. V., Boyce, J. B. & Bristow, V. L. (1996), *Stochastic processes*, Vol. 2, Wiley New York.
- Shumway, R. H. & Stoffer, D. S. (2017), *Time series analysis and its applications: with R examples*, Springer.
- Site, N. . O. (2018), 'Components: Nikkei stock average', url <https://indexes.nikkei.co.jp/en/nkave/index/component>.
- Stoll, H. R. & Whaley, R. E. (1990), 'The dynamics of stock index and stock index futures returns', *Journal of Financial and Quantitative Analysis* **25**(4), 441–468.
- Sucar, L. E. (2015), 'Probabilistic graphical models', *Advances in Computer Vision and Pattern Recognition. London: Springer London*. **10**, 63–82.
- Viterbi, A. (1967), 'Error bounds for convolutional codes and an asymptotically optimum decoding algorithm', *IEEE transactions on Information Theory* **13**(2), 260–269.
- Yu, S.-Z. & Kobayashi, H. (2003), 'An efficient forward-backward algorithm for an explicit-duration hidden Markov model', *IEEE signal processing letters* **10**(1), 11–14.
-

Zucchini, W., MacDonald, I. L. & Langrock, R. (2016), *Hidden Markov Models for Time Series: An Introduction Using R*, Monographs on statistics and applied probability 150, 2nd edn, CRC Press.
