



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



FACULTAD DE ADMINISTRACIÓN Y
DIRECCIÓN DE EMPRESAS. UPV

MODELIZACIÓN ESTOCÁSTICA PARA REALIZAR PREDICCIONES DEL SUBYACENTE BANKIA (BKIA.MC) DEL IBEX-35.

*Realización de una página web para mostrar las
predicciones diariamente*

CURSO 2015-2016

Autor:

D. Iván Gimeno Polo

Directores:

Dr. Juan Carlos Cortés López

Dr. Rafael Jacinto Villanueva Micó

ÍNDICE

Agradecimientos.....	4
Índice de Figuras.....	5
Índice de Gráficos.....	7
Índice de Tablas.....	8
Prólogo.....	10
Resumen del trabajo de fin de grado.....	12
Justificación de las asignaturas relacionadas.....	15
Capítulo 1: Antecedentes y situación actual.....	15
Capítulo 2: Bankia S.A.....	16
Capítulo 3 y 4: Cálculo Estocástico de Itô y Modelo Log-Normal.....	16
Capítulo 5: Aplicación del Modelo Log-Normal al Activo Subyacente BANKIA (BKIA.MC): Metodología Aplicada.....	17
Capítulo 6: Diseño y Aplicación de la Página Web.....	17
Objetivo del trabajo de fin de grado.....	20
1. Antecedentes y situación actual.....	23
2. Bankia S.A.....	27
2.1 Bankia S.A.....	27
Contexto.....	27
Origen.....	28
La Marca.....	29
Áreas de Negocio.....	29
Red Comercial.....	31
Modelo de Negocio.....	32
Organización.....	35
Bankia en la Bolsa.....	36
Principales Ratios Financieros e Información Relevante.....	39
3. Cálculo estocástico de Itô.....	42
3.1 Movimiento Browniano.....	42
3.2 Propiedades estadísticas del Movimiento Browniano.....	44
3.3. Simulación del Movimiento Browniano.....	46

3.4. El Cálculo de Itô.....	47
4. El Modelo Browniano Geométrico	55
4.1 Motivación del modelo para subyacentes en condiciones de certidumbre.....	56
4.2 Motivación del modelo para un subyacente cotizado: El Modelo Log-Normal.....	57
4.3 Solución del Modelo Log-Normal.....	58
4.4 La solución del Modelo Log-Normal y sus propiedades estadísticas.....	64
4.5 Calibración de los parámetros del Modelo Log-Normal.	65
4.6 Validación del Modelo Log-Normal.....	74
4.7 Estimaciones.....	78
5. Aplicación del modelo log-normal al activo subyacente BANKIA (Bkia.mc): metodología aplicada	80
5.1 <i>PYTHON</i> : Lenguaje de programación.....	81
5.2 Funcionamiento del código	82
6 Diseño y aplicación de la página web	102
6.1 Requisitos previos.	102
6.3 Diseño de la página web.	103
Conclusiones.....	114
Bibliografía.....	118
Libros:	118
Artículos:	119
Págnas web:	119

Agradecimientos

Un camino se acaba y otro empieza.

A mis padres, por todo, a Ana por su sonrisa, a mis amigos por aguantarme este verano tan... estocástico.

A Tino que me hizo civilizado, a Merche por su 'enfado' a tiempo, a Llorenç por lo que he llegado (al final me sirvió de algo la optativa a la que nunca quería ir), a Juan Carlos y Rafa por dirigir este Trabajo Fin de Grado.

- Hola! ¿Com estás?

- Fotut però bé...

- Va, que estás millor que jo.

- Quan era fadrí, un día....

A tots Gràcies.

ÍNDICE DE FIGURAS

Figura 1.1 : Dax.....	24
Figura 2.1 Bankia.....	30
Figura 2.2 Oficinas.....	22
Figura 2.2 Modelo de Negocio.....	33
Figura 1.4 Diseño.....	34
Figura 2.4 Distribución.....	34
Figura 2.6 Páginas Vistas	35
Figura 3.1: Simulación MB.....	51
Figura 4.1 : Capitalización a Interés Compuesto Continuo	61
Figura 5.1: Paquetes.....	88
Figura 5.2: Cotización BKIA.MC.....	89
Figura 5.3: Web scrapping.....	90
Figura 5.4: Resultado Web <i>Scrapping</i>	90
Figura 5.5: Media Est. , Varianza y Media2.....	91
Figura 5.6: Funciones de Estimación según Método 1 y 2.....	92
Figura 5.7: Errores.....	93
Figura 5.8: Fechas.....	94
Figura 5.9: Fechas definidas.....	94
Figura 5.10: IC Datos.....	95
Figura 5.11: Estimación Medias.....	96
Figura 5.12: Gráfico M1 y M2.....	97
Figura 5.13: Gráfico Comparativo.....	98

Figura 5.14: Tabla Datos actual.....	100
Figura 5.15: Tabla Comparación.....	102
Figura 5.16: Captura Datos.....	104
Figura 5.17: Método 1 I.....	105
Figura 5.18: Método 2 I.....	106
Figura 6.1: Inicio a.....	109
Figura 6.2: Inicio b.....	110
Figura 6.3: Método 1a.....	111
Figura 6.4: Método 1b.....	112
Figura 6.5: Método 2a.....	113
Figura 6.6: Método 2b.....	114
Figura 6.7: Comparativa 1.....	115
Figura 6.8: Comparativa 2.....	116
Figura 6.9: Conclusiones.....	117
Figura 7.1: Diagrama.....	120

ÍNDICE DE GRÁFICOS

Gráfico 2.1: Cotización BKIA 1.....	40
Gráfico 2.1: Cotización BKIA 2.....	40
Figura 3.1: Simulación MB.....	51
Gráfico 1.1: Cotización histórica Inditex.....	32
Gráfico 2.1: Simulación del Movimiento Browniano.....	42
Gráfico 5.1: Previsión M1.....	98
Gráfico 5.2: Comparativa M1.....	99

ÍNDICE DE TABLAS

Tabla 2.1 Ratios Financieros.....	41
Tabla 2.2 Ratings Emisor.....	41
Tabla 1.1: Principales Ratios Inditex.....	33
Tabla 2.1: Justificación de la identidad (3.3.1) para la simulación del Movimiento Browniano.	50
Tabla 4.1: Tabla comparación métodos de Estimación.....	83
Tabla 4.2: Resultados medidas de bondad de ajuste	89
Tabla 4.3: Resultados Predicciones.....	91
Tabla 5.1: Estimación mu y sigma.....	93
Tabla 5.2: Estimación MAPE y ECM.....	93
Tabla 5.3: Antes, después, hoy, ia.....	95
Tabla 5.4: Predicción, IC 95% Mape y Ecm Método de Máxima Verosimilitud.....	101
Tabla 5.5 Predicción e IC 95% Mape y Ecm Método no Paramétrico.....	101
Tabla 5.6: Comparación predicción y cotización real.....	103
Tabla 5.7 Comparación predicciones métodos y valor real.	103

PRÓLOGO

Conocer y entender el mundo que le rodea es uno de los grandes retos del ser humano. Es posible considerar la matemática, como ciencia que surgió con el fin de hacer los cálculos en el comercio, para medir la Tierra y para predecir los acontecimientos astronómicos. Estas tres necesidades pueden ser relacionadas en cierta forma a la subdivisión amplia de la matemática en el estudio de la estructura, el espacio y el cambio. Sin embargo, hay un concepto que prevalece sobre el resto: el poder que supone poder cuantificar los acontecimientos.

A lo largo de los siglos, las matemáticas han mostrado que son capaces de modelizar múltiples aspectos de la naturaleza que posteriormente se han extrapolado a otras materias. Tal es el caso de las Finanzas. Actualmente, existen numerosos modelos para predecir, entre otras cosas, el valor de activos subyacentes. Todos estos modelos matemáticos han evolucionado a lo largo de los años mejorando la robustez y precisión de los mismos.

La actual coyuntura económica mundial favorece el crecimiento de las principales bolsas mundiales, con bajos tipos de interés y una política por parte del Banco Central Europeo y la Reserva Federal que aplica un gran plan de compra de deuda para devolver la estabilidad económica. La estabilidad que se pretende lograr incide directamente en los inversores que, ante la caída de los tipos de interés ven insatisfechas sus expectativas respecto al rendimiento de la renta fija. Es por ello que la renta variable es una vía que, a pesar del mayor riesgo asociado, ofrece unas plusvalías más satisfactorias para el inversor.

Tomar decisiones que conllevan riesgo de manera certera es una tarea utópica, sin embargo, es posible desarrollar ciertos mecanismos que reduzcan el riesgo. Por este motivo, con el aumento de la inversión en renta variable aumentan también las herramientas para lograr una plusvalía satisfactoria. Este incremento de la inversión en renta variable hace que se desarrollen y estudien modelos para predecir el valor de activos subyacentes y así reducir el riesgo al que se somete el inversor en la elección de los valores más adecuados.

RESUMEN DEL TRABAJO DE FIN DE GRADO

En el Trabajo de Fin de Grado que se presenta a continuación, se estudia, aplica y valida el modelo estocástico Log-Normal al activo subyacente Bankia (BKIA.MC). El resultado final que se espera de la modelización validada de este activo subyacente, será la obtención de sus predicciones a corto plazo.

Como se anuncia en el Prólogo son diversos los modelos que son capaces de predecir el valor de un activo subyacente. Uno de ellos es el Modelo Log-Normal. Este modelo atiende a modelos estocásticos de un factor, quedando estos representados por una ecuación diferencial estocástica de tipo Itô que contiene en su formulación la tendencia y la volatilidad del subyacente. La aleatoriedad se introduce en dicha ecuación mediante la derivada del Movimiento Browniano (o proceso estocástico de Wiener), llamado proceso de ruido blanco (*white noise process*).

La aplicación del modelo, requiere la estimación de sus parámetros. En el presente trabajo se aplicarán dos métodos, por un lado, el método de Máxima Verosimilitud y, por otra parte, un método de momentos no paramétrico.

Estimados los parámetros del modelo, se llevará a cabo la validación del mismo. Posteriormente se calcularán, las estimaciones puntuales (a través de la función media) y las estimaciones probabilísticas (a través de los intervalos de confianza del 95%) y se determinarán algunas medidas de bondad de ajuste. En este proyecto, como medidas de bondad de ajuste se han utilizado el error cuadrático medio (ECM) y el error porcentual absoluto medio (MAPE).

Las estimaciones puntuales (función media) y por intervalos (intervalos de confianza del 95%), quedarán representadas gráficamente junto con los valores observados de la muestra. Para validar correctamente la aplicación del modelo se presentarán los resultados tanto numéricamente (a través de los resultados obtenidos en las medidas de bondad de ajuste), como gráficamente. Una vez, validado el modelo, se realizarán las correspondientes predicciones (puntuales y por intervalos) del activo subyacente en cuestión para los 5 días posteriores al último de los datos de la muestra seleccionada.

Además, se han empleado dos programas informáticos (uno por cada método de estimación de parámetros) utilizando el lenguaje de programación Python en los que aplicaremos el Modelo Browniano Geométrico sobre el activo subyacente escogido, en este caso Bankia (BKIA.MC), aunque es importante señalar que los programas son aplicables a cualquier otro activo subyacente. Dichos programas conectan con la página web <https://es.finance.yahoo.com>, de la cual se puede extraer el histórico de cotizaciones del subyacente, y aplican el modelo y generan ficheros con los resultados obtenidos en tablas y gráficos.

Para lograr una mayor comprensión de los resultados presentados, se representarán diariamente los resultados obtenidos en una web realizada para el presente Trabajo Fin de Grado ("http://cotizacion.imm.upv.es/"). Con esta web se consigue ofrecer los resultados obtenidos con el Modelo Log-Normal actualizados diariamente a todo el público interesado. Todo este proceso se ha realizado con el lenguaje de programación *Python*.

JUSTIFICACIÓN DE LAS ASIGNATURAS RELACIONADAS

Capítulo 1: Antecedentes y situación actual.

En el primer capítulo, debido a que es un capítulo de toma de contacto con el presente trabajo, es de reducida dimensión. Sin embargo, al ser un capítulo introductorio es necesario sintetizar grandes cantidades de información. Se han utilizado diversos conocimientos de distintas asignaturas cursadas durante el Grado en Administración y Dirección de Empresas. Se destacan:

- **“Inglés I y II”** (Tercer y Cuarto curso): Gran parte de la bibliografía necesaria se encuentra en artículos en inglés por lo que ha sido necesario la aplicación de los conocimientos obtenidos en estas asignaturas.
- **“Macroeconomía I y II”** (Primero y Segundo curso): Proporciona la visión de cuestiones esenciales como el papel de las políticas fiscal y monetaria, y sus efectos sobre la fluctuación de la actividad económica, dado que es un punto esencial de controversia entre las distintas visiones de la realidad económica. Igualmente, la comprensión de fenómenos tales como el desempleo y la inflación que tanto ha evolucionado en estas últimas décadas, de cómo dichos fenómenos interaccionan entre sí a lo largo del tiempo, y del papel que juegan las expectativas y los ajustes de los mercados en este contexto. Cómo estas cuestiones de la realidad y su interpretación van configurando los desarrollos teóricos y, a su vez, cómo éstos configuran la enseñanza de esta disciplina es esencial para su comprensión.

- **“Economía Mundial”** (Segundo curso): Proporciona una visión general del contexto económico internacional en el que se desenvuelven las empresas. La economía internacional supone algo más que presiones competitivas e inestabilidad financiera y cambiaria.

Capítulo 2: Bankia S.A.

- **“Análisis y Consolidación Contable”** (Tercer curso): A partir de los conocimientos adquiridos en esta asignatura somos capaces de realizar un análisis de la empresa a partir de los principales ratios financieros.
- **“Introducción a la Administración de Empresas”** (Primer curso): Proporciona una visión global de la gestión en la empresa y de los principales departamentos, con el fin de que el alumno adquiera los conocimientos básicos sobre cada una de las funciones de la administración y del funcionamiento empresarial.
- **“Introducción a las Finanzas”** (Primer curso): Esta asignatura proporciona conocimientos sobre la renta variable y el funcionamiento del sistema financiero.
- **“Contabilidad Financiera y de Sociedades”** (Segundo curso): Ofrece una visión general sobre la normativa contable : Constitución de sociedades; ampliaciones y reducciones de capital; inmovilizado; instrumentos financieros (activos y pasivos financieros); ingresos imputados a patrimonio neto (activos financieros clasificados como disponibles para la venta y subvenciones de capital); fusiones de sociedades de capital; escisión, disolución y liquidación de sociedades de capital; cuentas anuales (balance, cuenta de pérdidas y ganancias, memoria, estado de cambios en el patrimonio neto y estado de flujos de efectivo). De esta manera es posible interpretar los estados contables de la empresa analizada.

Capítulo 3 y 4: Cálculo Estocástico de Itô y Modelo Log-Normal.

- **“Métodos Estadísticos en Economía”** (Segundo curso): La estadística es una ciencia que analiza series de datos y trata de extraer conclusiones sobre el comportamiento de las variables. Por ello, esta asignatura proporciona múltiples

herramientas que se utilizaran para la implementación y validación de modelo Log-Normal.

- **“Econometría”** (Segundo curso): La aplicación exhaustiva de modelos matemáticos y estadísticos así como de la programación lineal y la teoría de juegos para analizar, interpretar y hacer predicciones sobre sistemas económicos. La asignatura se centra en los modelos de regresión y en los modelos de series temporales para describir la naturaleza y predecir los valores futuros de una variable. Aunque en este Trabajo Fin de Grado no se han utilizado series temporales para modelizar la dinámica del activo subyacente objeto de análisis, si es importante señalar que la asignatura ofrece una formación muy interesante que podría constituir una alternativa al enfoque utilizado en este Trabajo Fin de Grado a tener en cuenta en un futuro, así como conceptos básicos de utilidad en el contexto de la modelización.
- **“Modelos Matemáticos para ADE I y II”** (Primer curso): En esta asignatura se establecen las bases apropiadas del proceso de modelización en Economía y en particular de las ecuaciones diferenciales. El presente Trabajo de Fin de Grado se inspira en las bases proporcionadas por esta asignatura para el desarrollo del mismo.

Capítulo 5: Aplicación del Modelo Log-Normal al Activo Subyacente BANKIA (BKIA.MC): Metodología Aplicada.

La formación práctica cursada durante el grado se expone en este capítulo. De especial interés, resultan muy útiles los conocimientos adquiridos en las prácticas de laboratorio de diferentes asignaturas (MADE, Métodos Estadísticos, Econometría, etc.) en las que se han utilizado programas como Wolfram Mathematica o el software libre R; este conocimiento ha resultado muy importante para poder hacer uso del lenguaje de programación Python en este trabajo.

Capítulo 6: Diseño y Aplicación de la Página Web.

Para el desarrollo de este capítulo se ha tenido que adquirir conocimientos que no se desarrollan propiamente en el Plan de Estudios de la titulación, como es el diseño de

páginas webs. Cabe mencionar que en la asignatura **“Inglés II”** se presentó como proyecto semestral una página web denominada *BeLove*.

- **“Ingles I y II”** (Segundo y Tercer curso): La mayoría de los conocimientos aplicados que no se han desarrollado durante la carrera se han adquirido por medio de manuales en inglés por lo que estas asignaturas han sido de gran ayuda.

OBJETIVO DEL TRABAJO DE FIN DE GRADO

El objeto principal del presente trabajo es el estudio de técnicas cuantitativas para la predicción de valores de subyacentes a corto plazo, incluyendo su posterior aplicación a un caso práctico.

Por ello, una vez detallada la parte teórica se ha aplicado la modelización al activo subyacente Bankia (BKIA.MC), valor cotizado en el IBEX-35.

El esquema general que sigue el presente trabajo parte del modelo estocástico de predicción de subyacentes Log-Normal, cuya resolución requiere del conocimiento y aplicación del Cálculo de Itô. Los resultados obtenidos a partir del modelo de predicción de subyacentes se presentarán en una web creada para este Trabajo de Fin de Grado. Es necesario para la correcta realización del trabajo que se detallen los objetivos a conseguir.

- Estudiar y aplicar las herramientas estocásticas pertenecientes al Cálculo de Itô necesarias para obtener la solución del Modelo Log-Normal, así como las propiedades estadísticas de las mismas.
- Modelizar a partir de una muestra de datos reales y en un horizonte temporal adecuado, el subyacente Bankia (BKIA.MC)
- Validar la modelización realizada sobre el subyacente Bankia (BKIA.MC)
- Estudiar técnicas estadísticas para construir predicciones puntuales y probabilísticas (vía de intervalos de confianza) y medidas de bondad de ajuste, en el contexto de las ecuaciones diferenciales de tipo Itô, en las cuales se basa el Modelo Log-Normal.
- Aplicar los resultados teóricos para obtener las predicciones probabilísticas del subyacente Bankia (BKIA.MC).

- Evaluar los resultados del modelo aplicado comparando cada una de las predicciones resultantes con el correspondiente dato real del subyacente.
- Utilizar diferentes técnicas para construir las estimaciones de los parámetros del modelo garantizando la robustez de las estimaciones.
- Adaptar un programa con el lenguaje de programación *Python* para realizar predicciones con el Modelo Log-Normal para cualquier subyacente.
- Crear una página web dinámica donde mostrar las predicciones de Bankia (BKIA.MC) actualizadas diariamente.

1. ANTECEDENTES Y SITUACIÓN ACTUAL

La situación económica en 2016 es mejor que la de los últimos años. La tasa de crecimiento ha aumentado considerablemente, desde el 1,4% en 2014 hasta el 2,8 % en la actualidad. Algunos desequilibrios estructurales (de déficit exterior han mejorado), el sistema bancario se ha saneado y a pesar de no haber finalizado, el proceso de desapalancamiento de familias y empresas ha progresado notablemente.

La mejora de la economía española se debe, en parte, a tres factores exógenos. El primero, la devaluación del euro frente al dólar, que mejora el comportamiento frente a las exportaciones de la economía, reduciendo el déficit exterior. El segundo, los bajos precios del petróleo, permiten una inyección de renta importante debido a la gran dependencia que España tiene de los hidrocarburos. En tercer lugar, la política monetaria QE (Quantitative Easy) instrumentada desde mediados del 2014 por el Banco Central Europeo (BCE) que propició la caída de la prima de riesgo.

Para el año 2017, la Organización para el Desarrollo y la Cooperación Económica (OCDE) prevé una ralentización de la economía española, con una tasa de crecimiento de 2,3 % aproximadamente.

La previsible frenada de la evolución de la economía española tiene diversas causas. Por un parte, el precio del petróleo repunta levemente, las importaciones se ven influidas por la creciente debilidad de los países latinoamericanos, los efectos de la política monetaria del BCE son decrecientes. Por otra parte, la OCDE prevé que no se mantenga la política expansiva aprobada por el gobierno interino en vista a las elecciones. Estímulos que aumentaron el crecimiento económico y que ahora ya no están presentes. Como punto negativo, la deuda pública alcanza ya el 100,5% del PIB.

En cuanto a los valores más representativos de los mercados desde un punto de vista técnico, destacan el S&P500¹ que se presenta con 2.071,22 puntos a 20-06-2016 (variación de +62,02% respecto a 20-06-2011).

El Dax² abre a 20-06-2016 con 9.850,59 puntos tal y como muestra el *Gráfico 1- Valor de Dax*



Figura 1.1 Dax.

Fuente: https://www.google.es/webhp?hl=es&sa=X&ved=0ahUKEwj_zb7xtdvOAhXEfxoKHUo1A2sQPAGD#hl=es&q=dax.

Del año 2014 a la fecha actual de elaboración de este Trabajo Fin de Grado (segundo trimestre de 2016), la situación ha cambiado notablemente. Como ya advirtió Barclays³

¹ El índice Standard & Poor's 500, también conocido como S&P 500, es uno de los índices bursátiles más importantes de Estados Unidos y se le considera el índice más representativo de la situación real del mercado.

² El índice DAX (Deutscher Aktienindex) es el índice bursátil de referencia de la Frankfurter Wertpapierbörse (Bolsa de Frankfurt).

³ Barclays PLC es una compañía de servicios financieros con sede en Londres, que opera a nivel mundial. Es un holding que cotiza en la Bolsa de Londres, la Bolsa de Nueva York y la Bolsa de Tokyo. Actúa a través de su subsidiaria Barclays Bank PLC

las ganancias fáciles en los mercados se han acabado. Ya no se garantiza que la bolsa y la deuda vayan a subir como ha ocurrido en los últimos años.

En un escenario de subida de tipos en Estados Unidos, de la desaceleración económica de China y de grandes incertidumbres respecto a los países emergentes y las materias primas es necesario analizar con más cuidado si cabe, dónde y cómo invertir.

Una nueva regla ha entrado en escena y todos los activos se ven afectados en mayor o menor medida, la rentabilidad es más baja con volatilidad más elevada.

El uso de herramientas que proporcionen apoyo en la toma de decisiones resulta vital. Por este motivo los modelos matemáticos de predicción del valor de activos subyacentes pueden resultar muy efectivos.

Realizar una modelación de la evolución del valor de un activo subyacente es una tarea compleja debida fundamentalmente a la gran volatilidad que acompaña al activo. La evolución que ha experimentado la economía y el interés en conocer herramientas útiles y aplicables supone un desarrollo continuo que ha perfeccionado el modelo a lo largo de los años.

Las finanzas cuantitativas comenzaron en Estados Unidos en los años setenta, cuando algunos inversores comenzaron a utilizar fórmulas matemáticas para la asignación de precios de acciones y bonos. Harry Markowitz⁴, en su tesis de doctorado "Portfolio Selection" publicada en 1952, fue uno de los primeros en adoptar formalmente conceptos matemáticos a las finanzas.

El principal problema de la aplicación de las fórmulas matemáticas para la predicción de los precios de acciones y bonos era la no contemplación de la aleatoriedad. Si se tiene en cuenta la cantidad de factores que determinan el precio de un activo subyacente, estas previsiones resultaban insuficientes.

Previo a introducir la aleatoriedad en el proceso fue el matemático japonés Kiyoshi Itô⁵ fue el encargado de desarrollar el cálculo estocástico requerido en los nuevos modelos de activos subyacentes.

⁴ Harry Max Markowitz, (Chicago, 1927 -) recibió el premio Nobel de Economía en 1990. Markowitz es uno de los numerosos economistas laureados con el Nobel de Economía durante el siglo XX producido por la prestigiosa Escuela de Economía de Chicago de la Universidad de Chicago.

⁵ ITÔ, K. (1951). "On stochastic differential equations" en *Memoirs American Mathematical Society*, vol. 4, p. 1-51. ITÔ, K. (1944). "Stochastic Integral" en *Proc. Imperial Acad. Tokyo* 20, p. 519-524. ITÔ, K. (1961). *Lectures on stochastic processes*. Tata Institute, Bombay.

La aleatoriedad fue introducida formalmente en 1969, por el ganador del premio Nobel de Economía Robert C. Merton⁶ dónde introdujo y aplicó el cálculo estocástico en el estudio de las finanzas⁷ a partir de un proceso estocástico denominado proceso estocástico Wiener o Movimiento Browniano. Desde entonces, las ecuaciones diferenciales estocásticas pasan a representar los modelos para la predicción de los valores de los activos subyacentes. Existen modelos alternativos basados en series temporales, pero la exitosa incidencia del trabajo de K. Itô en la valoración de activos subyacentes cotizados y productos derivados (opciones tipo Call y Put europeas, americanas, exóticas, etc.) ha influido de modo que la mayor parte de los modelos dinámicos sobre subyacentes financieros se realizan hoy usando ecuaciones diferenciales estocásticas y, en particular, el Modelo Log-Normal.

⁶ Robert C. Merton (31 de julio de 1944) es un economista estadounidense que recibió el Premio Nobel de Economía en 1997, compartido con Myron Scholes, por sus trabajos para calcular el precio de las opciones financieras. Merton ayudó a introducir el cálculo estocástico en la economía financiera, lo que permitió que el comportamiento de los precios fuese descrito con el lenguaje preciso de la probabilidad.

⁷ MERTON, ROBERT C. (1973). "Theory of Rational Option Pricing". Bell Journal of Economics and Management Science, nº 4, pp. 141–183.

2. BANKIA S.A.

En el presente capítulo se realiza una visión general de la situación de la economía española y de Bankia S.A, de la cual se va a realizar la predicción de su valor bursátil a lo largo del Trabajo Fin de Grado.

2.1 Bankia S.A.

Bankia es la marca comercial del denominado Banco Financiero y de Ahorros, que surgió de la unión de Caja Madrid, Bancaja y las cajas Insular de Canarias, Laietana, Ávila, Segovia y Rioja.

En el momento de su constitución (2010) se situó como el primer banco español por volumen de activos, con 328.000 millones y el primero, por volumen de negocio doméstico, con 485.900 millones.

Bankia entra en escena por consecuencia de la reestructuración del sistema de cajas de ahorro establecido por el Banco de España con la instauración del Sistema Institucional de Protección (SIP).

Contexto.

La creación de la entidad viene ligada directamente al Sistema Institucional de Protección, por lo que antes de profundizar en su historia se va a proceder a explicar el contexto que causó su creación.

La crisis económica española sufrida en el periodo de 2008-2010 causó una inestabilidad del sistema financiero español que iba en aumento. Particularmente afectó a las cajas de ahorro, que quedan en una posición de debilidad.

Las entidades bancarias, ante situaciones de peligro pueden recurrir a formar un *holding* con el objetivo de agrupar a las entidades y autoprotgerse. El problema de las cajas de ahorro reside en la dependencia de la comunidad autónoma de origen, que dificulta las posibles fusiones con otras entidades por motivos políticos y de control regional.

Los SIPs aparecen en el marco financiero español en 2010 y sus principales objetivos son mejorar la solvencia y estabilidad financiera de las entidades así como su eficiencia a

través de un proceso de concentración empresarial. También se les denomina como modelos de *fusión fría*, es decir, una unión en la que sus entidades integrantes mantienen su marca, su forma jurídica, su capacidad de decisión individual, así como su independencia económica. Pero, por otra parte, comparten ventajas comunes como los servicios de información, facilitan las tareas operativas y aumenta la posibilidad de economías de escala.

Los SIPs surgieron con la Directiva 2006/48/CE, de 14 de junio y provocaron una reestructuración radical del sistema financiero español. Este medio de concentración empresarial es aplicable a diversos tipos de entidades bancarias pero se adapta mejor a las cajas de ahorro y a las cooperativas de crédito por poseer una cultura más solidaria.

Existen tres enfoques principales vinculados a los objetivos de un SIP:

- Solvencia: Se establece un sistema de garantías cruzadas dónde se ponen en común entre el 40-50% de los recursos propios y hasta el 100% de los beneficios.
- Estrategia: Se establece una duración mínima del contrato de integración de 10 años con un preaviso de 2 años si una entidad planea su salida.
- Gestión del riesgo: La constitución de un órgano central que se vincule a un plan de eficiencia de obligado cumplimiento por las entidades.

Origen.

El 30 de julio de 2010, Caja de Ahorros y Monte de Piedad de Madrid, Caja de Ahorros de Valencia, Castellón y Alicante (Bancaja), Caja Insular de Ahorros de Canarias, Caja de Ahorros y Monte de Piedad de Ávila, Caixa d'Estalvis Laietana, Caja de Ahorros y Monte de Piedad de Segovia y Caja de Ahorros de la Rioja suscribieron un Contrato de Integración del que surge un Sistema Institucional de Protección.

El Contrato de Integración establecía principalmente el reconocimiento de la organización como un grupo consolidable desde el punto de vista contable, y como una concentración desde el punto de vista de la competencia. Además establecía la integración de la gestión y de la titularidad de las inversiones empresariales y la centralización de las decisiones de inversión o desinversión de las carteras presentes y futuras.

El 3 de diciembre de 2010 se creó la Sociedad Central del SIP bajo el nombre de Banco Financiero y de Ahorros, S.A. (BFA). Esta sociedad se constituye como la matriz del Grupo Banco Financiero y como cabeza del Sistema Institucional de Protección.

El 1 de enero de 2011, Bankia S.A. se establece como una entidad integrada en el Grupo Banco Financiero y encabeza el Grupo Bankia que a 31 de diciembre de 2011 cuenta con 385 sociedades en su perímetro de consolidación.

Bankia, S.A. tiene su domicilio social en Valencia mientras que su sede operativa reside en Madrid. Cotiza en la Bolsa española desde el 20 de julio de 2011 y desde el 23 de diciembre de 2013 vuelve al índice Ibex-35 ya que debido a su proceso de reestructuración se excluyó del índice durante casi un año.

La Marca.

Bankia es el nombre comercial con el que se da a conocer la entidad financiera nacida el 1 enero de 2011.

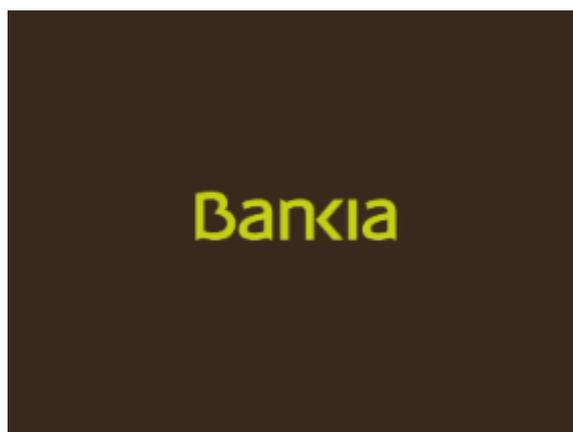


Figura 2.1: Bankia.

Fuente: <https://www.bankia.es>.

Áreas de Negocio.

Las principales áreas de negocio en las que Bankia y sus participadas desarrollan su actividad son las siguientes:

Banca de Particulares.

Es un negocio estratégico de Bankia. A través de una amplia red multicanal, es un modelo de negocio orientado hacia el cliente. Está formado por actividades de banca

minorista que factura anualmente menos de 6 millones de euros. Bankia es una de las principales instituciones financieras en esta área de negocio en España.

Banca de Particulares se centra en la actividad minorista, bajo un modelo de banca universal. Su objetivo es lograr la vinculación y retención de clientes, aportando a los clientes valor en los productos y servicios ofrecidos, en el asesoramiento y calidad en la atención

Banca de Empresas.

Se configura como una actividad estratégica en el desarrollo de las actividades de Bankia. Comprende la actividad bancaria dirigida a personas jurídicas con una facturación anual superior a los 6 millones de euros. Actualmente, más de 20.000 empresas de estas características mantienen relaciones financieras con Bankia, lo que sitúa a la Entidad como una de las más relevantes del mercado nacional en este segmento de negocio.

Banca Privada.

Dirigida a clientes con elevados patrimonios y altas rentas (superiores a los 600.000) presenta una oferta de productos y servicios personalizada y de alta confianza para lograr un servicio de máxima calidad y rapidez.

Gestión de Activos y Bancaseguros.

Engloba los negocios de fondos de inversión, planes y fondos de pensiones y seguros. Para la puesta en distribución de estos productos y servicios se utilizan las plataformas multicanal de grupo.

El modelo de distribución se establece de la siguiente forma:

- Seguros de no vida: acuerdo de distribución con Mapfre.
- Seguros de vida: acuerdo de distribución a través de la joint venture Bankia Mapfre Vida (51% Mapfre Vida – 49% Bankia), resultado de la integración de las compañías de seguros en las que Bankia tenía participación accionarial (Mapfre Caja Madrid Vida, Aseval y Laietana Vida).

Mercado de Capitales.

Engloba varias actividades referentes a los productos financieros tales como originar, diseñar y estructurar, además de su venta directa. Comprende el conjunto de grandes cuentas del Banco, con varios denominadores comunes: el tamaño de las compañías (más de 300 millones de euros de facturación), grupos de empresas con elevado número

de sociedades, y la mayor complejidad y sofisticación en su demanda de servicios financieros.

Participadas.

Bankia es la matriz del Grupo Bankia, por lo que la gestión del conjunto de filiales dependientes es una actividad necesaria en el negocio. Además cuenta con participaciones en diversas sociedades de diversos sectores que deben ser gestionadas correctamente.

Red Comercial.

La red comercial de Bankia se extiende por toda España tal como se muestra en el siguiente mapa. Además cuenta con una oficina de representación en Shanghai.



Figura 2.2: Oficinas.

Fuente: <https://www.bankia.es>.

Modelo de Negocio.

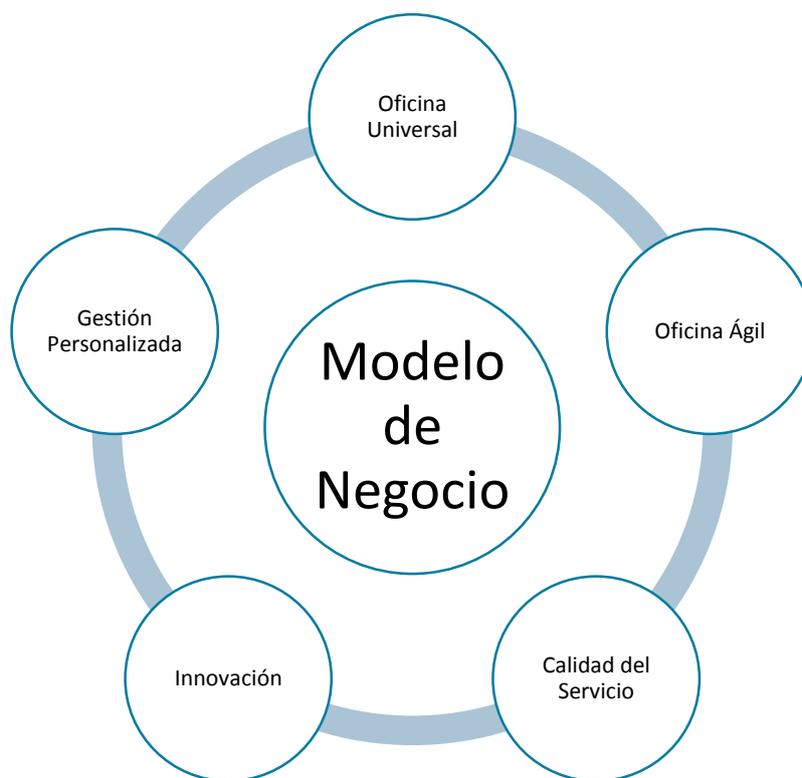


Figura 2.3: Modelo de Negocio.

Fuente: Elaboración propia.

➤ Oficina Universal.

Bankia pretende optimizar la banca minorista implantando una red de oficinas universales donde se optimiza el espacio para proporcionar un mejor servicio al cliente y facilitar el desempeño del trabajo a sus empleados.

Así se pretende mejorar tanto la calidad del servicio como la relación con el cliente.

➤ Oficina Ágil.

Mejorar la velocidad y calidad de atención al cliente es el objetivo de la reestructuración de las oficinas de Bankia que pretende facilitar al cliente las operaciones sencillas.



Figura 2.4: Diseño.

Fuente: <https://www.bankia.es>.

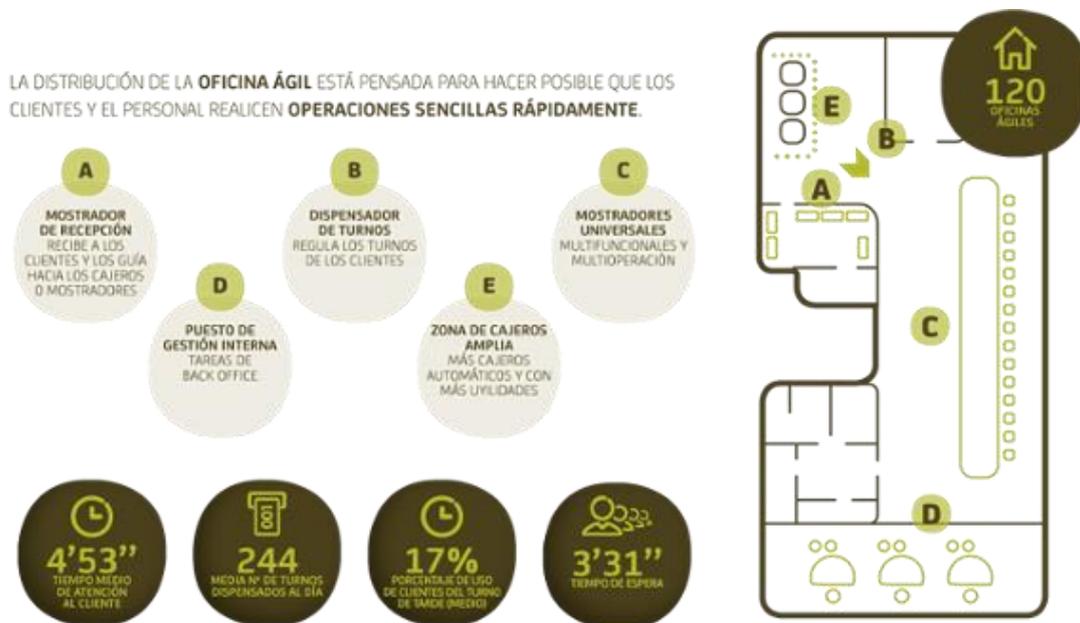


Figura 2.5: Distribución.

Fuente: <https://www.bankia.es>.

➤ Innovación.

Para facilitar el desempeño del trabajo se han instaurado dos sistemas informáticos que agilizan y mejoran la relación con el cliente.

NEO o Nuevo Escritorio de Oficinas permite visualizar en todo momento la información más relevante del cliente en todos los canales para ofrecer un servicio adecuado y personalizado.

SOS o Servicio Operativo de Soporte facilita a la red comercial la operativa de los procesos para hacer más sencillo el desempeño de los procesos.

Páginas vistas en SOS por tipo de operativa

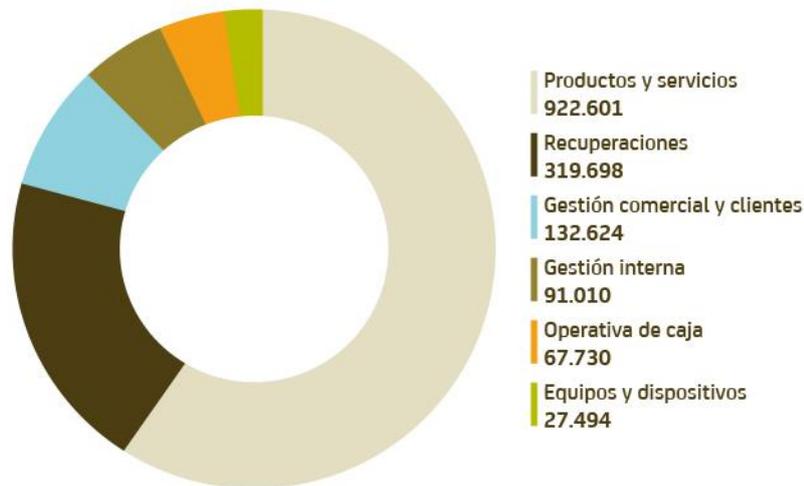


Figura 2.6: Páginas Vistas.

Fuente: <https://www.bankia.es>.

➤ Gestión Personalizada.

Para cada cliente de la Banca Personal o Privada se establecen gestores especializados con cursos de formación que les permite ofrecer productos y servicios adecuados al perfil del cliente.

Para el usuario de la Banca Privada se ofrece un asesoramiento individualizado según el perfil de riesgo que tenga el cliente.

➤ Calidad del Servicio.

Mejorar la calidad del servicio es una prioridad para Bankia que de acuerdo con su Plan de Gestión de Calidad utiliza el *Mystery Shopping* para evaluar la calidad del servicio proporcionado. Los últimos datos reflejan un aumento del 12% en la percepción de la calidad del servicio. Según los últimos índices, Bankia se situaría por encima del sector en cuanto a calidad ofrecida.

Organización.

A fecha de 30 de Junio de 2016 el perímetro de consolidación estaba formado por 60 sociedades. Del conjunto total de sociedades, 33 sociedades son empresas del grupo, 2 sociedades son entidades multigrupo y 25 son sociedades asociadas.

La sociedad se rige por dos órganos de gobierno: la Junta General de Accionistas y el Consejo de Administración.

La Junta General de Accionistas es el máximo órgano de decisión en las materias de su competencia atribuidas legal o estatutariamente. Nombrar consejeros, aprobar las cuentas anuales y especificar la política de distribución de dividendos son algunas de sus tareas.

Por otra parte, Bankia S.A. está administrada por un Consejo de Administración que dispone de las más altas facultades para administrar la sociedad con la excepción de las competencias de la Junta General.

El Consejo de Administración asume las competencias legalmente reservadas a su actuación directa así como cualquier actuación necesaria en relación al ejercicio de supervisión. Es el órgano de representación de la Sociedad y dispone de las más amplias atribuciones para su administración.

Los Consejeros tienen como objetivo fundamental que la Sociedad cumpla con la legalidad vigente, la buena praxis y los principios de responsabilidad social que se hayan acordado previamente.

Ente sus actividades destacan la aprobación del Plan estratégico o de negocio y los objetivos de gestión y presupuesto anuales, así como las políticas de inversiones y financiación.

Actualmente el Consejo de Administración está formado por:

(8 consejeros independientes y 3 consejeros ejecutivos).

- D. José Ignacio Goirigolzarri Tellaeché. Presidente Ejecutivo.
- D. José Sevilla Álvarez. Consejero Delegado.
- D. Antonio Ortega Parra. Vocal Ejecutivo.
- D. Joaquín Ayuso García. Consejero Independiente Coordinador.
- D. Francisco Javier Campo García. Vocal Independiente.
- D^a Eva Castillo Sanz. Vocal Independiente.

- D. Jorge Cosmen Menéndez-Castañedo. Vocal Independiente.
- D. José Luis Feito Higuera. Vocal Independiente.
- D. Fernando Fernández Méndez de Andés. Vocal Independiente.
- D. Antonio Greño Hidalgo. Vocal Independiente.
- D. Álvaro Rengifo Abbad. Vocal Independiente.

Bankia en la Bolsa.

La evolución de Bankia en la Bolsa española ha tenido varios altibajos por lo que se va a hacer referencia al contexto en el que la entidad decidió salir a bolsa.

Antes de la salida a bolsa, en el proceso de integración de las sociedades para establecer el SIP, se solicitó un préstamo de 4.465 millones de euros para cubrir los costes de integración.

La salida a bolsa ha estado envuelta en polémica por las declaraciones de los peritos del Banco de España que actuaron para la Audiencia Nacional. Afirman que las cuentas de Bankia estaban modificadas para hacer más atractiva la salida a bolsa. Sin embargo, el FROB (Fondo de Reestructuración y Ordenación Bancaria) cuestionó a estos peritos postulando que las cuentas de Bankia podrían ser correctas a la hora de la salida a Bolsa.

El Banco de España ha mandado un informe para responder a las preguntas del FROB pero consciente de que contradecir a los dos inspectores elegidos supone contradecir a la CNMV (Comisión Nacional del Mercado de Valores) como regulador de la salida a Bolsa de la entidad.

El problema principal de Bankia a la hora de su salida a Bolsa es una dotación a provisiones inferior a la que realmente necesitaba. Además, el gobierno español, como mecanismo para mejorar la estabilidad del sistema financiero, incrementó las dotaciones a provisiones.

Tras un primer decreto de Guindos, el ratio de capital principal se elevaba al 10% por lo que el grupo necesitaba 5.775 millones de euros para poder soportar dicho ratio según los informes de 31 de diciembre de 2010.

Debido a las turbulencias en los mercados financieros Bankia retrasa su salida a Bolsa hasta el 20 de julio de 2011 con un precio de 3,75€ la acción, un 15% menos que el precio anunciado en el folleto entregado a la CNMV.

Los primeros meses de cotización el valor se mantiene estable gracias al esfuerzo de las entidades 'cuidadoras'.

En febrero de 2012 empieza una etapa de declive en la cotización de las acciones de la entidad. El día 2 de febrero el Gobierno anuncia una reforma financiera que impone a las entidades una dotación adicional de provisiones a los activos inmobiliarios. Rodrigo Rato anuncia que Bankia puede hacer frente a la nueva obligación por si sola. Sin embargo, las dudas empiezan a envolver al mercado financiero en cuanto a la capacidad de Bankia para solventar esta medida. El 10 de febrero la cotización de la acción se desploma desde los 3,535 (cierre de 10-02-2012) hasta los 0,5 euros por acción.

El FMI publica un informe en el que demanda a Bankia que mejore su balance y la gestión de sus activos. Deloitte se niega a firmar las cuentas del ejercicio y Rodrigo Rato dimite.

El 7 de mayo de 2012, José Ignacio Goirigolzarri se haría con el control de la presidencia. Dos días después Bankia es nacionalizada a través de su matriz Banco Financiero y de Ahorro poseedora del 45% de la entidad.

Ese mismo mes se producen varios hechos de controversia. La entidad es rescatada con 19.000 millones de euros de las arcas públicas (sumándose a los 4.465 millones del préstamo). Son presentadas las cuentas de 2011 (que son consideradas de dudosa regularidad). Bankia declaró unas ganancias de 300 millones de euros cuando soportaba pérdidas de 3000 millones de euros aproximadamente.

En junio de 2012 Guindos anuncia un plan de rescate a la banca española de 50.000 millones de euros. Junto con el plan se exigen unas condiciones de saneamiento que obligan a la entidad a una recapitalización. Bankia reduce el valor de las acciones de 2 euros a 0.01 euros por acción (no es posible un valor más bajo).

Actualmente la cotización es de 1,0410 euros por acción.

A continuación se muestra el histórico de cotizaciones desde la salida a bolsa hasta la actualidad (5 de enero de 2016) siendo "S" la división de las acciones en proporción 1:100 y "D" como el reparto de dividendos de 0,0176 euros por acción



Gráfico 2.1: Cotización BKIA 1.

Fuente: <https://es.finance.yahoo.com/q/bc?s=BKIA.MC>.

Desde la división:



Gráfico 2.2: Cotización BKIA 2.

Fuente: <https://es.finance.yahoo.com/q/bc?s=BKIA.MC>.

Principales Ratios Financieros e Información Relevante.

Año	2012	2013	2014	2015	2016	2017
BPA (en €)	-116,78	0,04	0,07	0,1	0,1	0,11
Rentabilidad por dividendo	0,00%	0,00%	1,63%	2,79%	3,72%	4,39%
Dividendo neto (en €)	0	0	0,02	0,03	0,04	0,05
PER (veces)	0	27,76	17,69	10,86	10,9	9,76
Beneficio neto (millones de €)	-19.057,00	512	747	1.135,83	1.103,71	1.240,50
Deuda neta (millones de €)	0	0	0	0	0	0
Valor de los activos (por nº de acciones)	-31,8	1,01	1,09	1,1	1,14	1,18
Ebitda (millones de €)	0	0	0	0	0	0
Rentabilidad sobre recursos propios	0,00%	4,40%	6,43%	8,98%	8,64%	9,32%
EV / Ebitda (veces)(*)	0	0	0	0	0	0

Tabla 2.1: Ratios Financieros.

Fuente: Elaboración propia.

(*) Capitalización más deuda entre beneficio antes de impuestos y amortizaciones.

Fecha	02/12/2015	19/05/2015
Largo Plazo	BB	BB+
Corto Plazo	B	B
Perspectiva	Positiva	Positiva
Emisor	Standar & Poor's	FitchRatings

Tabla 232: Ratings Emisor.

Fuente: <http://www.eleconomista.es>.

En el caso de las entidades bancarias, los estados contables poseen unas características diferentes del resto de empresas e incluso están sometidos a una normativa contable distinta.

Por ello se presenta a continuación un extracto del *Informe de Gestión del Grupo Bankia* a fecha de junio del 2016 en el que se presenta información de interés respecto a la situación actual del Grupo:

“Los aspectos más relevantes de la evolución del Grupo Bankia en el primer semestre del ejercicio 2016 se resumen en los siguientes puntos:

Resultados

En la primera mitad del ejercicio 2016 el Grupo Bankia ha obtenido un resultado atribuible de 481 millones de euros, lo que representa un descenso del 13,4% con respecto al mismo periodo de 2015. Estos resultados han venido marcados, en líneas generales, por la continuidad de factores penalizadores para el negocio bancario en

general como son el entorno de tipos de interés en mínimos históricos y el consiguiente descenso de la rentabilidad de la cartera hipotecaria y las carteras de renta fija. A estos factores hay que añadir, en el caso concreto de Bankia, la salida del perímetro de consolidación del Grupo del City National Bank of Florida (en adelante CNBF), cuya venta se formalizó en octubre de 2015, y el efecto de la puesta en marcha de políticas comerciales destinadas a fidelizar a los clientes de la Entidad, que han impactado en el volumen de comisiones obtenido por el Grupo en el primer semestre de 2016.

En este contexto, los puntos fuertes de la gestión de Bankia han continuado siendo una política de moderación de costes, que ha mantenido el ratio de eficiencia del Grupo en el 46,6%, uno de los mejores entre las grandes entidades financieras españolas, y los avances en la gestión del riesgo, que se han traducido en menores provisiones y en una mejora significativa del coste del riesgo del Grupo.

Fortaleza en capital

La capacidad de generación de capital sitúa al Grupo Bankia entre las entidades más solventes del sistema financiero español. Al cierre de junio de 2016 el Grupo Bankia ha alcanzado una ratio CET 1 phased-in del 14,57%, que mejora en 61 puntos básicos el nivel alcanzado en diciembre de 2015. Esta nueva mejora ha sido impulsada, principalmente, por la generación de capital vía resultados, el paulatino proceso de desapalancamiento del balance y la mejora de la calidad de la cartera crediticia, permitiendo al Grupo mantener un elevado exceso de capital sobre el mínimo regulatorio establecido por el BCE.

Crecimiento de la nueva financiación y de los recursos minoristas más rentables

Por lo que hace referencia a la inversión, ha continuado creciendo la nueva producción de crédito a segmentos estratégicos como son empresas, pymes y consumo. Este avance de la inversión, unido a la desaceleración del desapalancamiento del sector privado en España, está contribuyendo a estabilizar el volumen de la cartera crediticia del Grupo, que en el primer semestre de 2016 ya no ha registrado una caída relevante respecto al cierre de diciembre de 2015.

En cuanto a los recursos de clientes, cabe destacar el favorable comportamiento que han tenido los depósitos estrictos de clientes y los recursos fuera de balance (fundamentalmente los fondos de inversión), que de forma conjunta han crecido en un 1,6% (1.951 millones de euros) respecto a diciembre de 2015. Por otra parte, en la primera mitad de 2016 se ha consolidado la transferencia del ahorro hacia depósitos

vista o fondos de inversión, que están captando el trasvase del ahorro de los clientes desde depósitos a plazo.

Nuevas mejoras en la calidad de los activos y la gestión del riesgo

Los principales indicadores de riesgo del Grupo han evolucionado de forma muy positiva en el primer semestre de 2016, en línea con la senda mostrada en el ejercicio anterior. Los saldos dudosos han disminuido un 9,6% desde el cierre de 2015, tanto por el refuerzo de las actividades de seguimiento y recuperaciones como por la venta de carteras de crédito realizada en el primer semestre del año. Todo ello se ha traducido en una disminución de la ratio de morosidad de un punto porcentual con respecto a diciembre de 2015 hasta el 9,6% en junio de 2016, mejorando también el nivel de cobertura, que se ha situado en el 60,8%, 0,8 puntos porcentuales más que en 2015.

La activa gestión del riesgo, tanto en las carteras de crédito como en los activos inmobiliarios adjudicados ha permitido al Grupo reducir de forma significativa el volumen de dotación de provisiones en comparación con el primer semestre de 2015.

1.5.- Sólida estructura de financiación y elevada liquidez

La solidez financiera del Grupo Bankia se apoya en la financiación de la actividad crediticia fundamentalmente a través de recursos de clientes. Prueba de ello es que, al cierre de junio de 2016, el Grupo Bankia está operando con un LTD ratio del 100,7%, lo que pone de manifiesto el equilibrio alcanzado entre su volumen de crédito y de depósitos.

Para cubrir las necesidades de liquidez adicionales, el Grupo acude a los mercados de capitales. El entorno de mercado, con tipos más bajos, unido al respaldo obtenido por la gestión y solvencia del Grupo, han posibilitado que en el primer semestre de 2016 Bankia haya colocado con éxito dos nuevas emisiones de cédulas hipotecarias por un importe global de 2.000 millones de euros.

Asimismo, en el primer semestre de 2016, Bankia ha accedido a 11.316 millones de euros en las nuevas subastas TLTRO II del Banco Central Europeo, lo que ha permitido alargar los vencimientos desde el ejercicio 2016 hasta el año 2020 y sustituir emisiones mayoristas que estaban previstas y se habrían realizado a un coste superior.”

3. CÁLCULO ESTOCÁSTICO DE ITÔ

En el presente capítulo se presentan las herramientas estocásticas necesarias para el cálculo de la solución del Modelo Browniano Geométrico. Además, se presentan las principales propiedades del mismo. En primer lugar, se introduce el Movimiento Browniano o proceso de Wiener, un proceso estocástico con una gran importancia en el posterior desarrollo teórico del Modelo Browniano Geométrico. En segundo lugar, se presentan las propiedades estadísticas del proceso, que se necesitarán posteriormente. En tercer lugar, se explicará cómo obtener simulaciones del proceso de Wiener. Por último, es necesario enunciar y comprender las integrales estocásticas de Itô y sus propiedades estadísticas, ya que el Modelo Browniano Geométrico se fundamenta en una ecuación diferencial estocástica de Itô cuya solución se representa a través de integrales de este tipo.

3.1 Movimiento Browniano.

Definimos como proceso estocástico a aquel que permite describir la evolución temporal de una variable aleatoria, formalmente, un proceso estocástico es una familia de variables aleatorias indexadas por uno o más índices (habitualmente, denotado por el tiempo). En el presente trabajo se va a utilizar uno de los procesos más influyentes en la Estadística, el movimiento Browniano. A continuación se dará una breve descripción de dicho proceso y por qué es adecuado utilizarlo para resolver el problema planteado.

Citando al físico francés Jean Perrin (1870-1942) que describió este fenómeno como: *"En un fluido en equilibrio, como el agua dentro de un vaso, todas sus partes aparecen completamente sin movimiento. Si ponemos en el agua un objeto de mayor densidad, cae. La caída, es cierto, será más lenta si el objeto es menor; pero un objeto visible siempre termina en el fondo del vaso y no tiende a subir. Sin embargo, sería difícil examinar durante mucho tiempo una preparación de partículas muy finas en un líquido sin observar un movimiento perfectamente irregular. Se mueven, se detienen, empiezan de nuevo, suben, bajan, suben otra vez, sin que se vea que tiendan a la inmovilidad."*

En el año 1828 un botánico inglés llamado Robert Brown (1773-1858) observó como en una solución de agua con polen de *Clarkia pulchella*, el polen realizaba un movimiento continuo con constantes cambios de dirección parecido a un zigzag. El propio Brown lo describió con las siguientes palabras: "Al examinar la forma de estas partículas inmersas en agua, vi muchas de ellas evidentemente en movimiento; éste consistía no solamente en un cambio de lugar en el fluido, manifestado por alteraciones en sus posiciones relativas, sino que también, con no poca frecuencia, por un cambio en la forma de la misma partícula [...] .En algunos casos se veía a la partícula girar sobre su eje mayor. Estos movimientos eran tales que me convencieron, después de observaciones repetidas, de que no surgían de corrientes en el fluido, ni de su gradual evaporación, sino que pertenecían a la misma partícula."

Durante las tres décadas siguientes el interés por el movimiento browniano decayó (denominado así por su descubridor). La primera explicación científica de este fenómeno la realizó Albert Einstein en 1905.

El movimiento browniano es un proceso estocástico gaussiano. Fue en el siglo XX cuando se profundizó en el estudio de este proceso estocástico y se aplicó con éxito en diversos campos. En el caso particular del contenido de este trabajo, se debe hacer referencia a la tesis doctoral del matemático Louis Bachelier⁸ al cuál se le atribuye el ser el primero en modelar el movimiento browniano en su tesis: "La teoría de la especulación" publicada en 1900.

Sin embargo el trabajo de L. Bachelier no fue un referente en su época y tuvieron que pasar unos años hasta que a partir de 1926, Norbert Wiener, con su entrada en el MIT, consiguió formalizar el movimiento browniano. Por ello, el movimiento browniano también se denomina como proceso de Wiener. Dicho proceso se denota como $\{B(t; \omega): t \geq 0, \omega \in \Omega\}$ ó $\{W(t; \omega): t \geq 0, \omega \in \Omega\}$ también es posible observar la notación: $\{B(t): t \geq 0\}$ en su forma más abreviada.

El proceso de Wiener es necesario para la posterior aplicación del proceso estocástico para modelizar activos subyacentes denominado como Modelo Log-Normal, ya que como se ha anunciado previamente la solución del modelo es un proceso estocástico denominado Modelo Browniano Geométrico (MBG).

⁸ BACHELIER, L. (1900). "Théorie de la spéculation". en Annales Scientifiques de l'École Normale Supérieure, vol. 3, p. 21-86.

Así pues, conocer el proceso de Wiener resulta fundamental ya que el Modelo Browniano Geométrico describe un proceso estocástico en el que el logaritmo natural de una variable aleatoria sigue un proceso Wiener generalizado.

El movimiento browniano, cuyas trayectorias muestrales son muy irregulares, introduce en el MBG la componente aleatoria. Para definir el proceso correctamente se va a enunciar una serie de propiedades que lo caracterizan:

MB.1. Comienza en el origen con probabilidad 1: $P[B(0) = 0] = 1$.

MB.2. Los incrementos del Browniano dados por $B(t) - B(s)$, son variables aleatorias independientes:

$$B(t_1) - B(t_0); B(t_2) - B(t_1); \dots; B(t_{n+1}) - B(t_n),$$

con $0 \leq t_1 \leq t_2 \leq \dots \leq t_{n+1} \leq +\infty$

MB.3. Tienen incrementos estacionarios:

$$B(t + \Delta t) - B(t) = B(s + \Delta t) - B(s), \quad \forall s, t : 0 \leq s \leq t \in [0, +\infty],$$

donde la igualdad anterior es en distribución.

MB.4. Los incrementos del proceso son gaussianos de media 0 y varianza $t - s$:

$$B(t) - B(s) \sim N(0; \sqrt{t - s}), \quad \forall s, t : 0 \leq s \leq t.$$

Considerando la propiedad MB.4., en el caso particular que $s = 0$, se deduce que:

$$B(t) \sim N(0; \sqrt{t}),$$

es decir, que fijado t , la variable aleatoria $B(t)$ sigue una distribución normal o gaussiana de media 0 y de desviación típica \sqrt{t} .

3.2 Propiedades estadísticas del Movimiento Browniano.

En este apartado se enuncian las propiedades estadísticas principales del Movimiento Browniano $\{B(t), t \geq 0\}$. Para las propiedades las cuales no se presenta demostración es posible obtenerla en libro de Øksendal⁹.

⁹ ØKSENDAL, B.K. (2003). Stochastic Differential Equations: An Introduction with Applications. Berlin: Springer.

P.1. Función Media: A partir de la condición MB.4, se deduce que la función media del Movimiento Browniano es idénticamente nula:

$$\mu_{B(t)} = 0, \forall t \geq 0,$$

P.2. Función Covarianza: Mide el grado de relación lineal entre las variables aleatorias $B(s)$ y $B(t)$, que se obtienen al fijar dos instantes s y t , respectivamente.

$$\text{Cov} [B(t), B(s)] = \min(s, t), \quad \forall s, t \geq 0.$$

Si tomamos $0 \leq s \leq t$, entonces utilizando las propiedades MB.4, P.1 y las propiedades básicas del operador esperanza se tiene:

$$\begin{aligned} \text{Cov} [B(t), B(s)] &= E [B(t)B(s)] - E[B(t)]E[B(s)] \\ &= E [B(t)B(s) - (B(s))^2 + (B(s))^2] \\ &= E [(B(t) - B(s))B(s) + (B(s))^2] \\ &= E [(B(t) - B(s))B(s)] + E [(B(s))^2] \\ &= E [(B(t) - B(s))(B(s) - B(0))] + E [(B(s))^2] \\ &= E [B(t) - B(s)] E [B(s) - B(0)] + E [(B(s))^2] \\ &= (E [B(t)] - E [B(s)])(E [B(s)] - E [B(0)]) + E [(B(s))^2] \\ &= E [(B(s))^2] \\ &= \text{Var}[B(s)] \\ &= s. \end{aligned}$$

Obsérvese que si en la relación anterior tomamos, se obtiene la propiedad MB.4, es decir, que la varianza del Movimiento Browniano es t .

P.3. $B(t)$ es $\frac{1}{2}$ autosemejante: Esta es, únicamente, una propiedad geométrica que formalmente se denota de la siguiente manera:

$$B(T \cdot t) = \sqrt{T} \cdot B(t), \forall T \geq 0.$$

P.4. Las trayectorias muestrales de $B(t)$ son continuas, pero no son diferenciables en ningún punto. Se puede demostrar que $B(t)$ tiene trayectorias que no son de variación acotada, esto es, que no son derivables (las trayectorias del Movimiento Browniano tienen puntos angulosos, es decir, con pico, para todo instante). Este comportamiento muestral se observa en el Gráfico 3.1.

3.3. Simulación del Movimiento Browniano.

En este apartado se va a realizar una simulación del Movimiento Browniano. Posteriormente será necesario, ya que cuando se analice y se resuelva el Modelo Log-Normal, la solución quedará expresada en términos Brownianos y, para poder realizar predicciones del activo subyacente debemos poder realizar simulaciones. A continuación se propone un método para simular el proceso.

Este método es viable, debido a su sencillez, para simular el Movimiento Browniano es a través de variables aleatorias normales tipificadas, usando para ello la identidad:

$$B(t) = \sqrt{t}Z, \quad Z \sim N(0; 1). \quad (3.3.1)$$

En la siguiente Tabla 3.1 se justifica esta identidad en distribución. Para ello, es suficiente con probar que ambos miembros de la identidad tienen la misma distribución. Más específicamente, es sencillo ver que dicha distribución es gaussiana y justificar que la media y varianza de ambos miembros coinciden.

	$B(t)$	$\sqrt{t}Z$
<i>Distribución</i>	Gaussiana [MB.4]	Gaussiana Transformación lineal de una variable Gaussiana
<i>Media</i>	0[MB.4]	$E[\sqrt{t}Z] = \sqrt{t} \cdot E[Z] = 0$
<i>Varianza</i>	t[MB.4]	$Var[\sqrt{t}Z] = (\sqrt{t})^2 \cdot Var[Z] = t$

Tabla 3.1: Justificación de la identidad (3.3.1) para la simulación del Movimiento Browniano.

Fuente: Elaboración Propia.

A continuación se muestra el Gráfico 3.1, dónde se visualiza una simulación de Movimiento Browniano sobre la ventana temporal [0,1].

```

T = 1;
NN = 100;
HW[ t_ ] :=  $\sqrt{t}$  * RandomVariate[NormalDistribution[0, 1]]
           |variable aleatoria |distribución normal
datat = Range[0, T,  $\frac{T}{NN}$ ];
           |rango
y = Map[HW, datat];
           |aplica a todos
listaW = Thread[{datat, y}];
           |atraviesa
ListPlot[listaW]
           |representación de lista

```

Figura 3.1: Simulación MB.

Fuente: Elaboración propia.

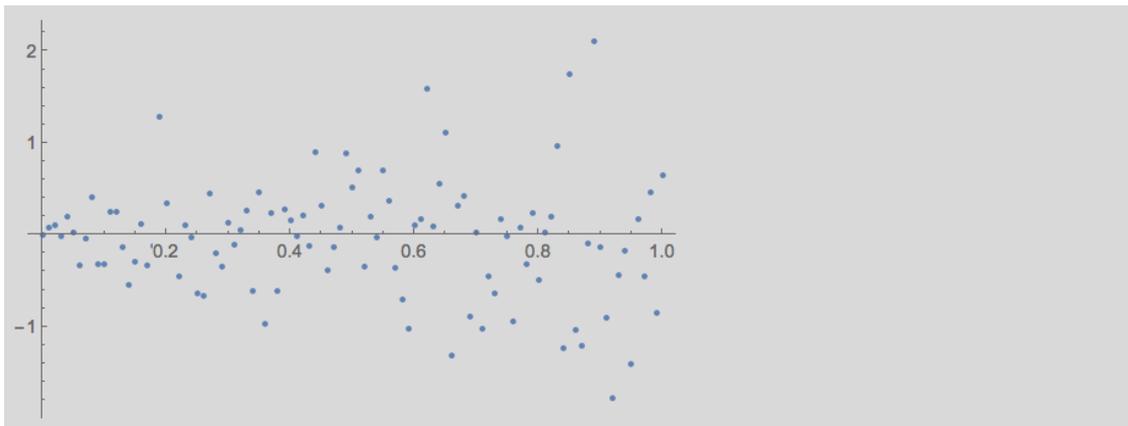


Gráfico 3.1: Simulación MB.

Fuente: Elaboración propia.

3.4. El Cálculo de Itô.

Kiyoshi Itô fue un matemático japonés que desarrolló una teoría para la diferenciación e integración de procesos estocásticos. En 1942, formuló una teoría de "ecuaciones diferenciales estocásticas" creando una herramienta esencial para el estudio de los fenómenos aleatorios en general, y de los fenómenos *brownianos* en particular. Es uno de los pioneros de la teoría de la probabilidad y el creador del Cálculo de Itô cuyo núcleo es la Integral de Itô y el resultado más importante, el Lema de Itô. Su trabajo facilita la comprensión matemática de fenómenos aleatorios y difiere de la teoría matemática clásica de integración y diferenciación.

El Cálculo de Itô tiene diversas aplicaciones de interés como por ejemplo la biología, la física, ingeniería y finanzas. En el presente trabajo de fin de grado se utilizará la versión adaptada al objetivo propuesto, las finanzas. Se centra en el cálculo exacto de la solución de una ecuación diferencial estocástica definida de la siguiente forma:

$$dX(t) = f(t, X(t))dt + g(t, X(t))dB(t). \quad (3.4.1)$$

Dicha solución se obtiene, en algunos casos que dependen de la forma específica de los coeficientes $f(t, X(t))$ y $g(t, X(t))$ utilizando el Lema de Itô. El Lema de Itô es una versión estocástica de la regla de la cadena para procesos estocásticos $X(t)$. En el caso del modelo utilizado en el presente trabajo (Modelo Browniano Geométrico) en el siguiente capítulo se determinará la solución. Con este objetivo, a continuación se enuncia a continuación el Lema de Itô en su versión integral:

Lema de Itô (versión integral)

Hipótesis: Sea $X(t)$ un proceso estocástico que satisface la siguiente ecuación diferencial tipo Itô con condición inicial determinista x_0 :

$$dX(t) = f(t, X(t))dt + g(t, X(t))dB(t), t \geq 0,$$

$$X(0) = x_0$$

Y sea $F(t, x)$ una función $F: [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ tal que las siguientes derivadas parciales existen y son continuas:

$$\frac{\partial F(t, x)}{\partial t} = F_1(t, x), \quad \frac{\partial F(t, x)}{\partial x} = F_2(t, x), \quad \frac{\partial^2 F(t, x)}{\partial x^2} = F_{22}(t, x),$$

Tesis: Entonces para $t > 0$ se cumple:

$$\begin{aligned} F(t, x(t)) - F(s, x_0) &= \int_0^t \{F_1(r, x(r)) + f(r, x(r))F_2(r, x(r))\}dr \\ &+ \int_0^t \frac{1}{2} (g(r, x(r)))^2 F_{22}(r, x(r))dr \\ &+ \int_0^t g(r, x(r))F_2(r, x(r))dB(r). \end{aligned}$$

➤ **3.4.1 La integral de Itô.**

En esta sección se introduce el concepto de Integral de Itô de un proceso estocástico. Una de las particularidades del Cálculo de Itô es que difiere con la teoría matemática clásica de integración y diferenciación. A continuación se proporcionará una interpretación de la expresión:

$$\int_0^t X(s)dB(s), \quad (3.4.2)$$

dónde $X(s)$ cumple unas determinadas condiciones. La expresión (3.4.2) es la Integral de Itô con respecto al Movimiento Browniano. Las ideas proporcionadas en el libro de E. Allen¹⁰ servirán de referencia a la explicación.

En primer lugar, supongamos que $X(s)$ y $B(s)$ no fueran procesos estocásticos, sino dos funciones deterministas.

Supongamos que $f(s)$ y $g(s)$ son dos funciones deterministas suaves¹¹ respecto la variable tiempo s y consideremos la siguiente integral:

$$\int_0^t g(s)df(s). \quad (3.4.3)$$

Cuando $f(s)$ tiende a una función diferenciable se escribe $\frac{df(s)}{ds} = f'(s)$, o de manera equivalente, $df(s) = f'(s)ds$. Si se sustituye la equivalencia en la integral (3.4.3) se obtiene:

$$\int_0^t g(s)df(s) = \int_0^t g(s)f'(s)ds.$$

Esta equivalencia queda definida como una integral de Riemann¹². En el caso de que $f(s)$ no fuese diferenciable, sería posible utilizar la teoría determinista para definir la integral. Cuando $f(s)$ es una función de variación acotada, es decir, que no es irregular en exceso, es posible demostrar que la integral está definida como el límite:

$$\int_0^t g(s)df(s) = \lim_{n \rightarrow \infty} \sum_{i=1}^{n-1} g(s_i)(f(s_{i+1}) - f(s_i)).$$

Por ello, al ser $f(s)$ de variación acotada, $f(s_{i+1})$ está próximo a $f(s_i)$. Es por ello que el límite anterior siempre existe si $g(s)$ no varíe en exceso. En caso contrario, cuando $g(s)$ tenga grandes variaciones en diferentes puntos en el tiempo, es posible que el límite sea divergente.

La integral (3.4.2) se define de forma análoga con el límite:

¹⁰ ALLEN, E. (2007). Modelling with Itô Stochastic Differential Equations. Springer (Series Mathematical Modelling: Theory and Applications).

¹¹ O que existe una fuerte correlación entre el estado presente y los estados futuros, aun admitiendo la influencia de sucesos esencialmente aleatorios e impredecibles.

¹² Georg Friedrich Bernhard Riemann (Breselenz, Alemania, 17 de septiembre de 1826 - Verbania, Italia, 20 de julio de 1866) fue un matemático alemán que realizó contribuciones muy importantes al análisis y la geometría diferencial, algunas de las cuales allanaron el camino para el desarrollo más avanzado de la relatividad general

$$\int_0^t X(s, \omega) dB(s, \omega) = \lim_{n \rightarrow \infty} X(s_i, \omega)(B(s_{i+1}, \omega) - B(s_i, \omega)), \quad \omega \in \Omega \quad (3.4.4)$$

Cabe mencionar que se toma el límite para cada ω fijo en el espacio muestral Ω de la variable aleatoria $B(s)$. Hay que tener en cuenta que el límite para cada ω en general no existe ya que empieza en $\pm\infty$ para una gran cantidad de procesos estocásticos $X(s)$. Para cada ω , la función $s \rightarrow B(s, \omega)$ es extremadamente volátil e irregular (recuérdese la propiedad estadística P.4 del Movimiento Browniano, es un proceso estocástico cuyas trayectorias muestrales de $B(t)$ son continuas, pero no son diferenciables en ningún punto). El Movimiento Browniano como función del tiempo no es de variación acotada para cada ω , tal y como se requiere para $f(s)$. Por ello, y con el objetivo de compensar la irregularidad de las trayectorias del Movimiento Browniano, se deben exigir dos condiciones. La primera condición consiste en asumir que $X(s)$ es independiente de los incrementos del Movimiento Browniano. La segunda condición guarda relación con la variación del integrando (similar a la condición de que $g(s)$ en (3.4.3) no debe variar demasiado). Si se observa el límite bajo estas condiciones a pesar de la irregularidad ya dicha integral existirá.

A partir de la propiedad MB.2 del Movimiento Browniano se sabe que la variación de un incremento del mismo, está dada por:

$$E[B(s_{i+1}) - B(s_i)]^2 = s_{i+1} - s_i.$$

Si $X(s_i)$ es independiente del incremento $B(s_{i+1}) - B(s_i)$, se obtiene:

$$\begin{aligned} E \left[\left(X(s_i)(B(s_{i+1}) - B(s_i)) \right)^2 \right] &= E[(X(s_i)^2)] E \left[(B(s_{i+1}) - B(s_i))^2 \right] \\ &= E[(X(s_i)^2)](s_{i+1} - s_i). \end{aligned}$$

Considerando el segundo momento de la variable aleatoria,

$$\sum_{i=1}^{n-1} X(s_i)(B(s_{i+1}) - B(s_i)), \quad (3.4.5)$$

y asumiendo que $X(s_i)$ es independiente de $B(s_{i+1}) - B(s_i)$ para cada $i = 1, \dots, n$, puede verse por la independendencia de los incrementos del Movimiento Browniano, que:

$$E \left[\left(\sum_{i=1}^{n-1} X(s_i)(B(s_{i+1}) - B(s_i)) \right)^2 \right] = \sum_{i=1}^{n-1} E[(X(s_i)^2)] (s_{i+1} - s_i).$$

La suma del miembro derecho es una aproximación de la integral $\int_0^t E [(X(s))^2] ds$. Por tanto, si esta integral existe, se deduce:

$$\lim_{n \rightarrow \infty} E \left[\left(\sum_{i=1}^{n-1} X(s_i)(B(s_{i+1}) - B(s_i)) \right)^2 \right] = \int_0^t E [(X(s))^2] ds.$$

Lo cual conduce a la conclusión de que la varianza de la suma en (3.4.5) converge a $\int_0^t E [(X(s))^2] ds$. Asumiendo que esta integral existe, se demuestra:

$$E \left[\left(\lim_{n \rightarrow \infty} \sum_{i=1}^{n-1} X(s_i)(B(s_{i+1}) - B(s_i)) \right)^2 \right] = \int_0^t E [(X(s))^2] ds. \quad (3.4.6)$$

Hay que señalar que $\int_0^t E [(X(s))^2] ds$ puede no existir en algunos casos. Siempre que sea finita, la integral existirá, es decir, siempre que el proceso estocástico $X(s)$ sea tal que su segundo momento puede ser integrado de 0 a t .

Por ejemplo, para el proceso $X(s) = s^{-1}B(s)$, por la condición MB.2 del Movimiento Browniano, se tiene:

$$\int_0^t E[(X(s))^2] ds = \int_0^t s^{-1} ds = \ln(t) - \ln(0) = +\infty.$$

Si se supone que $X(s) = B(s)$, se puede reconocer que $X(s)$ satisface la condición de integralidad. Una gran cantidad de procesos estocásticos cumple esta condición.

Fíjese en la relación (3.4.6) de nuevo, considerando que el término $X(s_i)$, que aparece en el miembro izquierdo, debe ser independiente de incrementos $B(s_{i+1}) - B(s_i)$ para todos los valores de $i = 1, \dots, n - 1$, propiciará la motivación de la adaptabilidad del proceso integrador.

Definición 1. Una variable aleatoria X es llamada \mathcal{F}_s -adaptada si X puede ser escrita como (límite de una sucesión de) funciones de $B(\tau)$ para uno o más $\tau \leq s$, pero no como función de cualquier $B(u)$ con $u > s$. Un proceso estocástico $X(s)$ se dice que es adaptado si para cada tiempo $s \in [0, t]$, la variable aleatoria $X(s)$ es \mathcal{F}_s -adaptada.

Para aclarar en mayor medida la definición anterior, en primer lugar, cabe señalar algunos procesos que se derivan de composiciones simples del Movimiento Browniano tales como $X(s) = f(s, B(s))$, los cuales son adaptados; mientras que procesos tales como $X(s) = B(s + 1)$, no lo son.

La integral $X(s) = \int_0^s B(\tau) d\tau$, también define un proceso estocástico adaptado, puesto que la integral es el límite de sumas del Movimiento Browniano en diferentes tiempos menores que s . En definitiva, por la definición de la integral de Itô, se tiene

$$X(s) = \int_0^s B(\tau) d\tau = \lim_{n \rightarrow \infty} \sum_{i=1}^{n-1} B(\tau_i) (\tau_{i+1} - \tau_i).$$

Por este motivo si se incide sobre la Integral de Itô, obsérvese que siempre que el proceso integrando $X(s)$ sea adaptado, la integral $\int_0^t X(s) dB(s)$ tiene sentido como el límite puntual en (3.4.4). Además, este límite puede probarse que converge en media cuadrática y por lo tanto, también para cada $\omega \in \Omega$. Se concluye la exposición con la definición rigurosa de la Integral de Itô.

Definición 2. Un proceso estocástico $X(s)$, es integrable en el intervalo $[0, t]$ en el sentido de Itô si

1. $X(s)$ es adaptado para $s \in [0, t]$, y
2. $\int_0^t E \left[(X(s))^2 \right] ds < \infty$.

La Integral de Itô se define como la variable aleatoria

$$\int_0^t X(s, \omega) dB(s, \omega) = \lim_{n \rightarrow \infty} \sum_{i=1}^{n-1} X(s_i, \omega) (B(s_{i+1}, \omega) - B(s_i, \omega)), \quad (2.4.7)$$

donde el límite es considerado por cada $\omega \in \Omega$.

3.4.2 Propiedades de la Integral de Itô.

Ya se ha introducido el concepto de Integral de Itô. A continuación se enunciarán sus propiedades estadísticas. Estas propiedades se distinguen si el integrando es una función determinista o un proceso estocástico $X(t)$.

Supongamos que $h(t)$ es una función determinista tal que $\int_0^t (h(s))^2 ds$, entonces se cumplen las propiedades:

P.1. Media:

$$E \left[\int_0^t h(s) dB(s) \right] = 0.$$

P.2. Varianza:

$$\text{Var} \left[\int_0^t h(s) dB(s) \right] = \int_0^t (h(s))^2 ds.$$

P.3. Covarianza:

$$\text{Cov} \left[\int_0^t h_1(\tau) dB(\tau), \int_0^s h_2(\tau) dB(\tau) \right] = \int_0^{t \wedge s} h_1(\tau) h_2(\tau) d\tau,$$

siendo $t \wedge s = \min(t, s)$. En particular:

$$E \left[\left(\int_0^t h_1(\tau) dB(\tau) \right) \left(\int_0^s h_2(\tau) dB(\tau) \right) \right] = \int_0^{t \wedge s} h_1(\tau) h_2(\tau) d\tau.$$

P.4. Normalidad:

$$\int_0^t h(s) dB(s) \sim N \left(0; \int_0^t (h(s))^2 ds \right).$$

Si el integrando es un proceso estocástico $X(t)$ que satisface las condiciones de la Definición 2 anterior, entonces se verifican las siguientes propiedades:

P.5. Media: $E \left[\int_0^t X(s) dB(S) \right] = 0.$

P.6. Varianza: $V \left[\int_0^t X(s) dB(s) \right] = \int_0^t E[(X(s))^2] ds.$

Obsérvese que a partir de las propiedades I.5 e I.6 se deduce

$$E \left[\left(\int_0^t X(s) dB(s) \right)^2 \right] = \int_0^t E[(X(s))^2] ds,$$

llamada isometría de Itô.

4. EL MODELO BROWNIANO GEOMÉTRICO

En el presente capítulo se presentará la base y el fundamento matemático principal sobre el que se realizarán las predicciones del activo subyacente, el Modelo Browniano Geométrico o distribución Log-Normal. La siguiente ecuación diferencial estocástica tipo Itô es la base del modelo:

$$dS(t) = \mu S(t)dt + \sigma S(t)dB(t), \quad (4.0.1)$$

$$S(0) = s_0,$$

dónde:

- $S(t)$ es el valor del subyacente en el instante t .
- s_0 representa el valor del subyacente en el instante inicial $t = 0$ y es conocido.
- $\mu \in \mathbb{R}$ es un parámetro del modelo denominado tendencia.
- $\sigma > 0$ es un parámetro del modelo denominado volatilidad local.
- $B(t)$ es un proceso estocástico de Wiener o Movimiento Browniano.

Para desarrollar correctamente el modelo este capítulo se va a dividir en varias secciones. En primer lugar (sección 4.1), a partir de modelos deterministas sobre activos financieros se motivará la consideración del modelo (4.0.1). En segundo lugar (sección 4.2), se hallará la solución del modelo utilizando el cálculo de Itô, además de obtener sus principales propiedades estadísticas. En tercer lugar (sección 4.3), se estimarán los parámetros μ y σ usando para ello técnicas estadísticas. Por último (sección 4.4), se explicará cómo disponer de los resultados del modelo para realizar la predicción del activo subyacente.

4.1 Motivación del modelo para subyacentes en condiciones de certidumbre.

En esta sección se presenta un modelo determinístico clásico para representar la trayectoria temporal de una inversión en un entorno de certidumbre (sin riesgo), y servirá de inicio para la motivación, en la siguiente sección, del Modelo Log-Normal.

Sea S_0 un principal que se invierte a un interés μ compuesto durante un intervalo temporal $[0, T]$ dividido en K periodos de longitud $\Delta t > 0$ cada uno (ver Figura 4.1).



Figura 4.1 Capitalización a interés compuesto continuo.

Fuente: Elaboración propia.

Se denota a $\hat{S}^{(j)}, j = 0, 1, \dots, K$, el capital al cabo de j periodos teniendo en cuenta que, por términos de notación, $\hat{S}^{(0)} = S_0$ entonces:

$$\hat{S}^{(1)} = S_0 e^{\mu \Delta t}, \quad \hat{S}^{(2)} = S^{(1)} e^{\mu \Delta t}, \quad (4.1.1)$$

y en general:

$$\hat{S}^{(j)} = \hat{S}^{(j-1)} e^{\mu \Delta t}, \quad j = 0, 1 \dots K, \quad (4.1.2)$$

y por tanto, razonando recursivamente se tiene:

$$\hat{S}^{(j)} = \hat{S}^{(j-1)} e^{\mu \Delta t} = \hat{S}^{(j-2)} e^{2\mu \Delta t} = \dots = S_0 e^{\mu j \Delta t}, \quad j = 0, 1 \dots K. \quad (4.1.3)$$

Véase que cuando $j = K$, se cumple que: $\hat{S}(T) = \hat{S}^{(K)} = S_0 e^{\mu K \Delta t} = S_0 e^{\mu T}$, donde se supone que $K \Delta t > T$ y de ha introducido la notación: $\hat{S}(T)$ que representa el capital al cabo de K periodos. Es decir, el efecto de K pasos de longitud Δt equivale a un paso de tamaño $T = K \Delta t$. En realidad. Este modelo de capitalización resulta de resolver el problema de valor inicial (ecuación diferencial ordinaria junto a una condición inicial) siguiente:

$$\frac{\Delta \hat{S}(t)}{\Delta t} = \mu \hat{S}(t), \hat{S}(0) = S_0, \quad (4.1.4)$$

cuya solución es:

$$\hat{S}(t) = S_0 e^{\mu t}, \quad (4.1.5)$$

la cual, para $t = T$ captura la solución del problema anterior: $\hat{S}(T) = S_0 e^{\mu T}$. Es importante observar que el modelo (4.0.1), el modelo estocástico de partida, contiene al modelo clásico determinista de capitalización continua, pues con hacer $\sigma = 0$ en (4.0.1) para obtener a partir de la ecuación diferencial estocástica la ecuación diferencial ordinaria (4.1.4).

4.2 Motivación del modelo para un subyacente cotizado: El Modelo Log-Normal.

La característica principal del modelo presentado anteriormente es la no contemplación de la aleatoriedad debido a las condiciones de certidumbre que se contempla. Realmente, el valor de un activo subyacente depende de un gran número de factores externos que influyen en mayor o menor medida en los mercados donde cotiza la acción. Determinar el nivel de influencia que ejerce sobre el activo es una tarea compleja que introduce una gran incertidumbre en el modelo. Por este motivo, si se pretende modelizar el valor del activo lo más acertado es introducir la aleatoriedad en el modelo.

El parámetro μ es ahora una magnitud aleatoria en vez de una magnitud determinada. Si se tienen en cuenta las herramientas estocásticas presentadas en el capítulo anterior el Movimiento Browniano $B(t)$, concretamente con su diferencial o derivada, el proceso de Ruido Blanco $dB(t)$. Por ese motivo, se motiva la siguiente expresión estocástica del parámetro μ :

$$\mu \rightarrow \mu + \sigma dB'(t), \quad \sigma > 0. \quad (4.2.1)$$

En la expresión (4.2.1) se han introducido varios cambios:

- Se sigue denotando al parámetro μ como el rendimiento medio del subyacente.
- Se han añadido variaciones aleatorias siendo $B(t)$ el proceso de Ruido Blanco y $\sigma > 0$ su intensidad.

A continuación, se introduce en el modelo la notación diferencial de la derivada:

$$S'(t) = \frac{dS(t)}{dt}.$$

Si se considera la aleatoriedad aplicada al parámetro μ , el modelo se reescribe de la siguiente forma:

$$\frac{dS(t)}{dt} = (\mu + \sigma B'(t))S(t) \Rightarrow dS(t) = (\mu + \sigma B'(t))S(t)dt$$

Si se agrupan los diferentes parámetros se obtiene:

$$dS(t) = \mu S(t)dt + \sigma S(t)B'(t)dt. \quad (4.2.2)$$

Si se observa la expresión anterior:

- $\mu S(t)dt$ representa la parte determinista.
- $\sigma S(t)B'(t)dt$ representa la parte estocástica, en la cual se introduce el Ruido Blanco.
- Dado que $dB(t) = B'(t)dt$, el modelo se puede reescribir como (4.0.1), es decir, el Modelo Browniano Geométrico.

4.3 Solución del Modelo Log-Normal.

El objetivo de este apartado es resolver la ecuación diferencial estocástica (4.0.1). En el presente trabajo se va a resolver mediante dos métodos. El primero de ellos, es de un enfoque más técnico que hace uso del Cálculo de Itô presentado en el Capítulo 3, particularmente el Lema de Itô. El segundo enfoque introduce la aleatoriedad de forma más artificial que el primero, pero evita el uso del Cálculo de Itô. De esta manera no se requieren herramientas de diferenciación e integración de procesos estocásticos, facilitando así su resolución.

➤ 4.3.1 Solución del Log-Normal mediante la aplicación del Cálculo de Itô.

En esta sección se resolverá la ecuación diferencial estocástica de tipo Itô (4.0.1) con condición inicial:

$$\left. \begin{aligned} dS(t) &= \mu S(t)dt + \sigma S(t)dB(t), \\ S(0) &= s_0, \end{aligned} \right\}$$

aplicando el Lema de Itô introducido en el Apartado 3.4. Para ello, escribimos (4.0.1) en forma integral:

$$\int_0^t dS(r) = \int_0^t \mu S(r) dr + \int_0^t \sigma S(r) dB(r),$$

o equivalentemente:

$$S(t) - S(0) = \int_0^t \mu S(r) dr + \int_0^t \sigma S(r) dB(r). \quad (4.3.1)$$

A continuación, aplicamos el Lema de Itô con la siguiente identificación:

$$X(t) = S(t), f(t, x(t)) = f(t, S(t)) = \mu S(t), g(t, x(t)) = g(t, S(t)) = \sigma S(t),$$

tomando

$$F(t, x) = \ln(x).$$

Para la aplicación del Lema de Itô se necesitan calcular las siguientes derivadas parciales:

$$F_1(t, x) = \frac{\partial F(t, x)}{\partial t} = 0, \quad F_2(t, x) = \frac{\partial F(t, x)}{\partial x} = \frac{1}{x},$$

$$F_{22}(t, x) = \frac{\partial^2 F(t, x)}{\partial x^2} = -\frac{1}{x^2}.$$

Por tanto:

$$\ln(S(t)) - \ln(s_0) = \int_0^t \left(\mu S(r) \frac{1}{S(r)} + \frac{1}{2} (\sigma S(r))^2 \left(\frac{-1}{(S(r))^2} \right) \right) dr$$

$$+ \int_0^t \sigma S(r) \frac{1}{S(r)} dB(r).$$

Simplificando tenemos:

$$\ln\left(\frac{S(t)}{s_0}\right) = \int_0^t \left(\mu - \frac{1}{2} \sigma^2 \right) dr + \int_0^t \sigma dB(r).$$

$$\ln\left(\frac{S(t)}{s_0}\right) = \left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma(B(t) - B(0)).$$

Observemos que por la definición del Movimiento Browniano, $B(0) = 0$ con probabilidad 1, por tanto obtenemos:

$$\ln\left(\frac{S(t)}{s_0}\right) = \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma B(t),$$

$$\frac{S(t)}{s_0} = e^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma B(t)},$$

$$S(t) = s_0 e^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma B(t)}, \quad (4.3.2)$$

que representa el proceso estocástico solución, llamado Movimiento Browniano Geométrico del Modelo Log-Normal.

➤ 4.3.2 Solución del Modelo Log-Normal sin la aplicación del Cálculo de Itô.

Determinar el valor de una acción en el mercado es una tarea compleja debido a la multitud de factores que afectan a la misma. Tal y como se mostró en el Apartado 4.2 en el modelo (4.0.1) se recoge, en su formulación, una parte determinista que cuantifica la evolución del precio del activo subyacente que se comporta como cualquier otra inversión libre de riesgo, y la parte aleatoria que cuantifica la variación del valor del activo a causa de los factores externos inciertos.

Partamos de la expresión (4.1.2) con $j = 1$, introduzcamos la aleatoriedad como sigue:

$$\tilde{S}^{(1)} = \tilde{S}^{(0)} e^{\mu\Delta t} e^{c Z_1}, \quad (4.3.3)$$

siendo c una constante libre y no nula (que más tarde fijaremos) y Z_1 una variable aleatoria normal o gaussiana tipificada, i.e., $Z_1 \sim N(0; 1)$, y en general, introduciendo la aleatoriedad en (4.1.2) del mismo modo se tiene:

$$\tilde{S}^{(j)} = \tilde{S}^{(j-1)} e^{\mu\Delta t} e^{c Z_j}, \quad (4.3.4)$$

siendo $Z_j \sim N(0; 1)$. Si hacemos esto para cada, $j = 1, 2 \dots K$ de modo que, $Z_1, \dots, Z_j, \dots, Z_k$ sean variables aleatorias $N(0; 1)$ e independientes, desde (4.3.4) con $j = K$, obtenemos por recursividad que:

$$\tilde{S}^{(K)} = S_0 e^{\mu K \Delta t} e^{c(Z_1 + \dots + Z_K)}, \quad (4.3.5)$$

o equivalentemente:

$$\tilde{S}(T) = S_0 e^{\mu T} e^{c(Z_1 + \dots + Z_K)}.$$

La aleatoriedad introducida así en (4.3.1) (o más generalmente en (4.3.4)), parece coherente porque garantiza que $\tilde{S}^{(1)} > 0$ ($\tilde{S}^{(j)} > 0$); pero no es totalmente satisfactoria, ya que, es de esperar que cuando no haya incertidumbre, $\tilde{S}^{(1)}$ en (4.3.3) se comporte como en (3.1.3) con $j = 1$ al menos en media. Sin embargo, se sabe¹³:

$$E[e^{cZ}] = e^{\frac{c^2}{2}}, \quad Z \sim N(0; 1), \quad (4.3.6)$$

y por tanto:

$$E[\tilde{S}^{(1)}] = E[S_0 e^{\mu \Delta t} e^{cZ_1}] = S_0 e^{\mu \Delta t} E[e^{cZ_1}] = S_0 e^{\mu \Delta t} e^{\frac{c^2}{2}} \neq S_0 e^{\mu \Delta t} = \tilde{S}^{(1)}.$$

Para lograr introducir la aleatoriedad considérese:

$$S^{(1)} = S_0 e^{\mu \Delta t} e^{cZ_1 - \frac{c^2}{2}}. \quad (4.3.7)$$

Obsérvese que en media este proceso estocástico se comporta como el modelo determinista:

$$E[S^{(1)}] = S_0 e^{\mu \Delta t},$$

y en general en (4.1.2) se introduce la aleatoriedad del mismo modo:

$$S^{(j)} = S^{(j-1)} e^{\mu \Delta t} e^{cZ_j - \frac{c^2}{2}}, \quad j = 1, 2, \dots, K. \quad (4.3.8)$$

Si se razona recursivamente como en (3.1.3) se obtiene:

$$S(T) = S^{(j-1)} e^{\mu \Delta t} e^{c(Z_1 + \dots + Z_K)} e^{-K \frac{c^2}{2}}. \quad (4.3.9)$$

¹³ QUESADA, V., GARCÍA A. (1988) "Lecciones de Cálculo de Probabilidades". Ed. Díaz de Santos.

Se denota por:

$$B_K = Z_1 + \dots + Z_K, \quad (4.3.10)$$

entonces dado que $Z_j \sim N(0; 1)$, $j = 1, 2 \dots K$, son variables aleatorias independientes se tiene $B_K \sim N(0; 1)$ y como $K\Delta t = T$ y $S^{(K)} = S(T)$, (4.3.9) equivale a:

$$S(T) = S_0 e^{\mu T} e^{cB_K - K \frac{c^2}{2}}, \quad (4.3.11)$$

donde S_0 es el precio o subyacente de la acción en el instante inicial; $e^{\mu T}$ es la componente determinista del valor de la acción (ligado al valor μ del tipo de interés); B_K representa la introducción de la aleatoriedad en el modelo y $e^{-K \frac{c^2}{2}}$ es un término (determinista) de corrección.

La modelización dada (4.3.11) tiene una gran ventaja frente a otro tipo de modelos como los basados en árboles binomiales¹⁴, y es que permite que el valor de la acción no sólo tome dos valores posibles, sino cualquier valor positivo. Sin embargo, sigue conteniendo un defecto importante en relación con el objetivo de este trabajo; si se fija T y se van considerando subintervalos más pequeños (particiones más finas del intervalo $[0, T]$), es decir, se hace $K \rightarrow +\infty$ ó equivalentemente $\Delta t \rightarrow +0$, con T fijo, para la varianza del término que introduce la aleatoriedad en el modelo (4.3.11) se obtiene por (4.3.10) que:

$$Var[cB_K] = c^2 Var[B_K] = c^2 K \rightarrow +\infty, \quad (4.3.12)$$

Por ello, la varianza de este término que forma parte del modelo de la acción aumentará cuantas más discretizaciones se tomen del intervalo objeto de análisis, con independencia del valor T que define su extremo superior, el cual está fijo. Esto no tiene sentido desde el punto de vista financiero, ya que fijado un intervalo durante el cual se observa una acción, la volatilidad de la misma es un valor fijo que no debería cambiar aunque en lugar de observar la acción intra-semana se hiciera intra-día. Para resolver

¹⁴ HULL, J.C. (2002). "Introducción a los Mercados de Futuros y Opciones". Pearson Education. 4ª edición.

este inconveniente, se va a utilizar la constante c introducida en (4.3.8) ya que es libre y por tanto es posible adaptarla a conveniencia. Se toma entonces c de modo que:

$$c^2 K = \sigma^2 T, \quad (4.3.13)$$

siendo σ^2 un parámetro fijo identificativo del modelo particular se utiliza (es menester indicar que en la práctica se calculará a partir de los datos reales observados). De esta forma:

$$\text{Var}[cW_K] = \sigma^2 T < +\infty. \quad (4.3.14)$$

Obsérvese que esto nos indica algo coherente que se aprecia en la práctica: a mayor longitud T del intervalo $[0, T]$, se tiene mayor volatilidad, es decir, variabilidad de la acción.

Ahora se va a relacionar B_K dada en (3.3.9) con el Movimiento Browniano. Para ello, obsérvese que por la propiedad P.3 del Capítulo 2 sabemos que:

$$B(T) = \sqrt{\Delta t} B_K. \quad (4.3.15)$$

Entonces despejando de (4.3.13) se tiene:

$$c = \sigma \sqrt{\frac{T}{K}} \xrightarrow{T=K\Delta t} c = \sigma \sqrt{\Delta t}, \quad (4.3.16)$$

y entonces por (4.3.15) se tiene:

$$cB_K = \sigma \sqrt{\Delta t} B_K = \sigma B(T). \quad (4.3.17)$$

Finalmente, sustituyendo (4.3.13) y (4.3.17) en (4.3.11) llegamos a:

$$S(T) = s_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)T} e^{\sigma B(T)}. \quad (4.2.18)$$

Con ello, a través de una motivación basada en la introducción de la aleatoriedad en el modelo determinista, se ha obtenido la solución del Modelo Log-Normal.

4.4 La solución del Modelo Log-Normal y sus propiedades estadísticas.

La solución aportada en la Sección 4.3.1 mediante la aplicación del Cálculo de Itô (4.3.2) se utiliza para realizar predicciones puntuales basadas en la probabilidad del activo subyacente a partir de la función media y el ajuste mediante intervalos de confianza en un tiempo $t = T$.

Por este motivo, se deben obtener de (4.3.2) las funciones media y varianza en $t = T$.

➤ Función Media.

Si se aplica la propiedad de $\frac{1}{2}$ -autosemejanza del Movimiento Browniano estudiada en el Capítulo 3:

$$B(T \cdot t) = \sqrt{T} \cdot B(t), \quad t \geq 0, \quad T \geq 0,$$

aplicada al caso particular $t = 1$, es decir:

$$B(T) \stackrel{d}{=} \sqrt{T} \cdot B(1). \quad (4.4.1)$$

Se recuerda que por definición el Movimiento Browniano:

$$B(1) \sim N(0; 1).$$

Por otra parte, además se aplica la propiedad:

$$E[e^{\lambda Z}] = e^{\frac{\lambda^2}{2}}, \quad Z \sim N(0; 1), \quad (4.4.2)$$

que se deduce del cálculo de la siguiente integral:

$$E(e^{\lambda Z}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{\lambda z} e^{-\frac{z^2}{2}} dz.$$

Si se utiliza el operador esperanza en (3.3.18) se tiene:

$$\begin{aligned} E[S(T)] &= S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)T} E[e^{\sigma B(T)}] \\ &= S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)T} E[e^{\sigma \sqrt{T} B(1)}] \quad (4.4.3) \\ &= S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)T} e^{\frac{(\sigma \sqrt{T})^2}{2}}, \end{aligned}$$

por tanto:

$$E[S(T)] = S_0 e^{\mu T}. \quad (4.4.4)$$

Cabe decir que en el segundo paso de (4.4.3) se ha aplicado la propiedad (4.4.1) y en el segundo paso se ha aplicado la relación (4.4.2) con $\lambda = \sigma\sqrt{T}$, $B(1) = Z \sim N(0; 1)$.

La fórmula (4.4.4) indica que el modelo Log-Normal posee la característica de que el comportamiento medio del precio de la acción es el mismo que el valor en el caso determinista (véase (4.1.5) con $t = T$).

➤ Función Varianza.

Predecir la volatilidad del precio del activo subyacente es necesario. Si se razona como en la Función Media, la varianza es:

$$\begin{aligned} Var [S(T)] &= E \left[(S(T))^2 \right] - (E[S(T)])^2 \\ &= E \left[(S_0)^2 e^{2\sigma B(T) + 2\left(\mu - \frac{\sigma^2}{2}\right)T} \right] - (S_0)^2 e^{2\mu T} \\ &= (S_0)^2 e^{2\mu T} (E[e^{2\sigma B(T)}] e^{-\sigma^2 T}) - 1, \end{aligned}$$

es decir:

$$\begin{aligned} Var [S(T)] &= (S_0)^2 e^{2\mu T} (E[e^{2\sigma\sqrt{T}B(1)}] e^{-\sigma^2 T}) - 1 \\ &= (S_0)^2 e^{2\mu T} \left(e^{\frac{4\sigma^2 T}{2}} e^{-\sigma^2 T} \right) - 1 \quad (4.4.5) \\ &= (S_0)^2 e^{2\mu T} (e^{\sigma^2 T} - 1), \end{aligned}$$

tal y como es de esperar, a medida que avanza el horizonte temporal T la varianza aumenta.

4.5 Calibración de los parámetros del Modelo Log-Normal.

A la hora de estimar los parámetros del modelo Log-Normal (μ y σ) se van a utilizar dos técnicas diferentes. El motivo, asegurar que se trata de estimaciones lo suficientemente robustas para aceptarlas como válidas. Si ambas técnicas proporcionan resultados similares significará que las estimaciones son aceptables.

Las dos técnicas son:

- Método de Máxima Verosimilitud.
- Método no paramétrico.

➤ 4.5.1 Método de Máxima Verosimilitud.

En esta sección, se va a utilizar una técnica estadística denominada como Método de Máxima Verosimilitud para estimar los parámetros μ y σ . Pevio a la aplicación del método es necesario conocer una serie de conceptos y definiciones que facilitan la correcta aplicación del mismo.

Fue recomendado, analizado y popularizado por R. A. Fisher entre 1912 y 1922, aunque había sido utilizado antes por Carl Friedrich Gauss, Pierre-Simon Laplace, Thorvald N. Thiele y Francis Edgeworth.

Supóngase que se ha obtenido una muestra $\{x_K: 0 \leq K \leq N\}$ de una población descrita por una variable aleatoria continua con una función de densidad $p(x; \vec{\theta})$ dependiente del vector de parámetros $\vec{\theta} \in \mathbb{R}^n$. Los resultados son análogos si la población es discreta. Si se consideran muestras aleatorias simples, la probabilidad de que ocurra la muestra (x_1, x_2, \dots, x_N) para un vector dado del vector de parámetros $\vec{\theta}$ viene descrita por la función de densidad de probabilidad conjunta (f.d.p) de la muestra y que se denota por L .

$$L(x_1, x_2, \dots, x_n; \vec{\theta}) = \prod_{i=1}^N p(x_i; \vec{\theta}). \quad (4.5.1)$$

Se pretende que dada una muestra fija, se estime el valor del vector de parámetros $\vec{\theta}$ que es desconocido. Hay que tener en cuenta que si $(x_1, x_2 \dots x_N)$ es una muestra dada, se llama *función de verosimilitud* a la función de $\vec{\theta}$ dada por:

$$L(x_1, x_2, \dots, x_n; \vec{\theta}) = \prod_{i=1}^N p(x_i; \theta).$$

Además, esta función es considerada como función solo de $\vec{\theta}$ y no de la muestra, que ya es fija, al construirse esta función después de haber obtenido la muestra. Por este motivo se podría utilizar simplemente la notación $L(\vec{\theta})$.

El Método de Máxima Verosimilitud considera como mejor estimación del vector de parámetros $\vec{\theta}$ el valor que haga máxima L para la muestra dada. Es decir, dada (x_1, x_2, \dots, x_N) consideremos mejor estimación $\vec{\theta}_1$ que $\vec{\theta}_2$ si:

$$L(x_1, x_2, \dots, x_n; \vec{\theta}_1) > L(x_1, x_2, \dots, x_n; \vec{\theta}_2),$$

es decir, el valor de $\vec{\theta}_1$ es más probable que el valor de $\vec{\theta}_2$ dada la muestra $((x_1, x_2, \dots, x_N))$. La premisa clave en la utilización de la técnica es elegir la moda de la distribución a posteriori de $\vec{\theta}$. Si se tiene en cuenta que la función de verosimilitud es mayor o igual que cero para todo valor de $\vec{\theta}$ y que la función logaritmo es monótona creciente resulta que:

$$\max[L(x_1, x_2, \dots, x_N; \vec{\theta})] = \max[\ln(L(x_1, x_2, \dots, x_N; \vec{\theta}))].$$

Como se puede observar en la igualdad, es equivalente utilizar el $\ln L$ en lugar de L . El cambio se debe a que es más sencilla la derivación de una suma que de un producto si se pretende calcular el máximo de la función.

Si la muestra depende únicamente de un parámetro, θ entonces $\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$ es un estimador de máxima verosimilitud de θ para la distribución poblacional definida por $p(x; \theta)$ si $\hat{\theta}$ hace máximo el valor de la función de verosimilitud, o lo que es lo mismo, hace máximo:

$$\ln L(x_1, x_2, \dots, x_n; \theta) = \sum_{i=1}^N \ln(p(x_i, \theta)).$$

Por ello, el cálculo efectivo de $\hat{\theta}$ se reduce al resolver la ecuación:

$$\frac{d \ln L}{d \theta} = 0,$$

de forma que si $\hat{\theta}$ es solución de esta ecuación $\hat{\theta}$ será máximo si la segunda derivada es negativa.

Si la distribución poblacional depende de N parámetros $\theta_1, \theta_2, \dots, \theta_N$ habrá que resolver el sistema:

$$\begin{aligned} \frac{\partial \ln L(\theta_1, \theta_2, \dots, \theta_N)}{\partial \theta_1} &= 0, \\ &\vdots \\ \frac{\partial \ln L(\theta_1, \theta_2, \dots, \theta_N)}{\partial \theta_N} &= 0. \end{aligned}$$

Posteriormente se debe comprobar que si $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_n)$ verifica este sistema, ya que hace el hessiano correspondiente definido negativo.

Ya que se ha avanzado en los conceptos y definiciones necesarios para la estimación, es aconsejable recapitular la información presentada hasta el momento. Dada una muestra $\{x_k : 0 \leq k \leq N\}$ de un proceso estocástico $x(t)$ en los momentos $\{t_k : 0 \leq k \leq N\}$ y suponiendo que el proceso $x(t)$ depende de un vector de parámetros $\vec{\theta} \in \mathbb{R}^n$, deseamos determinar el vector $\vec{\theta}$ tal que:

$$\max L(\theta) = f(t_0, x_0, t_1, x_1, \dots, t_n, x_n; \vec{\theta}).$$

Siendo $f(t_0, x_0, t_1, x_1, \dots, t_n, x_n; \vec{\theta})$ la función de densidad conjunta de la muestra. En el contexto del presente trabajo, no se puede asumir que la muestra sea simple (que existe independencia) por lo que no es posible legitimar la factorización dada en (4.5.1). Sin embargo, si se desea aplicar el Método de Máxima Verosimilitud es necesario tener una expresión de la función de densidad conjunta que sea preferiblemente algebraica.

Se presenta a continuación el modelo de subyacente (véase 4.0.1) de la forma:

$$\begin{aligned} dx(t) &= f(t, x(t); \vec{\theta}) dt + g(t, x(t); \vec{\theta}) dB(t), \\ x(0) &= x_0, \end{aligned} \quad (4.5.2)$$

con:

Para expresar el valor de $p(t_K, x_K | t_{K-1}, x_{K-1}, \vec{\theta})$ es necesario partir de la ecuación diferencial estocástica de Itô dada en (4.5.2) y aplicar el esquema de Euler-Maruyama para aproximar dicha ecuación en los instantes de la muestra que se suponen de igual distancia en el tiempo siendo Δt el paso temporal.

Si se considera un subintervalo de tiempo genérico, $[t_{K-1}, t_K]$ y vía la discretización de Euler-Maruyama obtenemos:

$$x(t_K) - x(t_{K-1}) = f(t_{K-1}, x(t_{K-1}); \vec{\theta}) \Delta t + g(t_{K-1}, x(t_{K-1}); \vec{\theta}) \{B(t_K) - B(t_{K-1})\}. \quad (4.5.4)$$

Denótese por x_K la aproximación de la solución obtenida por el esquema anterior en el instante t_K :

$$x_K \sim x(t_K),$$

es posible reescribir el esquema (4.5.4) como:

$$x(t_K) \approx x_{K-1} + f(t_{K-1}, x_{K-1}; \vec{\theta}) \Delta t + g(t_{K-1}, x_{K-1}; \vec{\theta}) \{B(t_K) - B(t_{K-1})\}, k = 1, \dots, N. \quad (4.5.5)$$

Si se hace referencia al Capítulo 3 una de las propiedades del Movimiento Browniano consiste en:

$$B(t_K) - B(t_{K-1}) \sim N(0; \sqrt{\Delta t}), \quad k = 1, 2, \dots, N.$$

Por lo tanto de (4.5.5) se sigue que la aproximación, vía Euler-Maruyama de $x(t_K)$ dado $x(t_{K-1})$, sigue la siguiente distribución normal:

$$x_K | x_{K-1} \sim N(\mu_K, \sigma_K), \quad \begin{cases} \mu_K = x_{K-1} + f(t_{K-1}, x_{K-1}; \vec{\theta}) \Delta t, \\ \sigma_K = g(t_{K-1}, x_{K-1}; \vec{\theta}) \Delta t. \end{cases} \quad (4.5.6)$$

Por lo tanto:

$$p(t_K, x_K | t_{K-1}, x_{K-1}, \vec{\theta}) = \frac{1}{\sqrt{2\pi}\sigma_K} e^{-\frac{1}{2}\left(\frac{x_K - \mu_K}{\sigma_K}\right)^2}, \quad k = 1, 2, \dots, N, \quad (4.5.7)$$

siendo μ_K y σ_K las expresiones dadas en (3.5.6). Esto permite expresar explícitamente la función de log- verosimilitud dada en (3.5.3) de la siguiente forma:

$$\begin{aligned}
 D(\vec{\theta}) &= -\ln(p(t_0, x_0; \vec{\theta})) \\
 &- \sum_{K=1}^N \ln\left(\frac{1}{\sqrt{2\pi}\sigma_K} e^{-\frac{1}{2}\left(\frac{x_K-\mu_K}{\sigma_K}\right)^2}\right) \\
 &= -\ln(p(t_0, x_0; \vec{\theta})) \quad (4.5.8) \\
 &- N\ln\left(\frac{1}{\sqrt{2\pi}\sigma_K}\right) + \sum_{K=1}^N \frac{1}{2}\left(\frac{x_K-\mu_K}{\sigma_K}\right)^2 \\
 &= -\ln(p(t_0, x_0; \vec{\theta})) + \frac{N}{2}\ln(2\pi) \\
 &+ \sum_{K=1}^N \ln(\sigma_K) + \frac{1}{2} \sum_{K=1}^N \left(\frac{x_K-\mu_K}{\sigma_K}\right)^2.
 \end{aligned}$$

Hay que tener en cuenta que el modelo de subyacente $x_0 = s_0$ utilizado es determinista, por tanto:

$$p(t_0, x_0; \vec{\theta}) = p(0, s_0; \mu, \sigma) = 1, \quad (4.5.9)$$

además las funciones μ_K y σ_K dadas en (4.5.6) toman la forma:

$$\mu_K = S_{K-1} + \mu S_{K-1} \Delta t; \sigma_K = \sigma S_{K-1} \sqrt{\Delta t}. \quad (4.5.10)$$

Entonces si se sustituye (4.5.9) y (4.5.10) en (4.5.8) obtenemos:

$$\begin{aligned}
 D(\mu, \sigma) &= \frac{N}{2}\ln(2\pi) + \sum_{K=1}^N \ln(\sigma S_{K-1} \sqrt{\Delta t}) + \frac{1}{2} \sum_{K=1}^N \left(\frac{S_K - S_{K-1} - \mu S_{K-1} \Delta t}{\sigma S_{K-1} \sqrt{\Delta t}}\right)^2 \\
 &= \frac{N}{2}\ln(2\pi) + \frac{N}{2}\ln(\Delta t) + N \ln(\sigma) \\
 &+ \sum_{K=1}^N \ln(S_{K-1}) + \frac{1}{2\sigma^2 \Delta t} \sum_{K=1}^N \left(\frac{S_K}{S_{K-1}} - 1 - \mu \Delta t\right)^2.
 \end{aligned}$$

Así pues, el programa de minimización de la función de Log-Verosimilitud es:

$$\begin{aligned}
 \text{Min } D(\mu, \sigma) \\
 (\mu, \sigma) \in \mathbb{R} \times]0, +\infty[& \quad \frac{N}{2}\ln(2\pi) + \frac{N}{2}\ln(\Delta t) + N \ln(\sigma) \\
 & + \sum_{K=1}^N \ln(S_{K-1}) + \frac{1}{2\sigma^2 \Delta t} \sum_{K=1}^N \left(\frac{S_K}{S_{K-1}} - 1 - \mu \Delta t\right)^2.
 \end{aligned}$$

Con el objetivo de calcular el mínimo de $D(\mu, \sigma)$ se deben obtener, en primer lugar, los puntos críticos, que son aquellos en los que el gradiente es nulo:

$$\begin{aligned}\frac{\partial D(\mu, \sigma)}{\partial \mu} &= -\frac{1}{\sigma^2} \sum_{K=1}^N \left(\frac{S_K}{S_{K-1}} - 1 - \mu \Delta t \right) = 0, \\ \frac{\partial D(\mu, \sigma)}{\partial \sigma} &= \frac{N}{\sigma} - \frac{1}{\sigma^3 \Delta t} \sum_{K=1}^N \left(\frac{S_K}{S_{K-1}} - 1 - \mu \Delta t \right)^2 = 0.\end{aligned}\quad (4.5.11)$$

Resolviendo el sistema de ecuaciones (4.5.11) se obtienen como puntos críticos:

$$\hat{\mu} = \frac{1}{N \Delta t} \sum_{K=1}^N \left(\frac{S_K}{S_{K-1}} - 1 \right), \quad \hat{\sigma} = \frac{1}{N \Delta t} \sum_{K=1}^N \left(\frac{S_K}{S_{K-1}} - 1 - \mu \Delta t \right)^2. \quad (4.5.12)$$

Se puede comprobar que la matriz hessiana formada por las cuatro derivadas parciales segundas de la función $D(\mu, \sigma)$, es decir,

$$H(D(\mu, \sigma)) = \begin{bmatrix} \frac{\partial^2 D(\mu, \sigma)}{\partial \mu^2} & \frac{\partial^2 D(\mu, \sigma)}{\partial \mu \partial \sigma} \\ \frac{\partial^2 D(\mu, \sigma)}{\partial \sigma \partial \mu} & \frac{\partial^2 D(\mu, \sigma)}{\partial \sigma^2} \end{bmatrix},$$

evaluada en el punto crítico $(\hat{\mu}, \hat{\sigma})$ dado en (4.5.12) tiene valores propios positivos, lo que muestra que $(\hat{\mu}, \hat{\sigma})$ es un mínimo de la función opuesta de Log-Verosimilitud y por tanto son los estimadores máximos verosímiles del Modelo Log-Normal.

➤ 4.5.2 Método no paramétrico.

En la presente sección se presenta la segunda técnica para estimar los parámetros μ y σ , un método no paramétrico¹⁶.

A partir de de una ecuación diferencial estocástica tipo Itô general dada en (4.5.2), donde $\vec{\theta} \in \mathbb{R}^n$ el vector de parámetros que se desea estimar. En el caso del Modelo Log-Normal, $\vec{\theta} \in \mathbb{R}^2$. A continuación se discretiza la ecuación (4.5.2) por algún método numérico, por ejemplo, mediante un esquema de Euler-Maruyama:

¹⁶ STANTON, R (1997). "A non parametric model of term structure dynamic and the market price of interest rate risk". Journal of Finance.

$$x(t + \Delta t) - x(t) = f(t, x(t); \vec{\theta})\Delta t + g(t, x(t); \vec{\theta})\Delta B(t),$$

$$x(t + \Delta t) = x(t) + f(t, x(t); \vec{\theta})\Delta t + g(t, x(t); \vec{\theta})\Delta B(t). \quad (4.5.13)$$

Fijada una partición del intervalo de tiempo $[0, T]$ donde se quiere aproximar la ecuación (3.5.13):

$$0 = t_0 < t_1 < t_2 < \dots < t_{N-1} = T,$$

con paso fijo $\Delta t = \frac{T}{N}$, obsérvese que se está dividiendo el intervalo $[0, T]$ en N subintervalos de la forma:

$$[t_i, t_{i+1}], \quad 0 \leq i \leq N - 2,$$

siendo $\Delta t = t_{i+1} - t_i$, y denotando por x_i la aproximación de $x(t)$ en el punto t_i , por ejemplo, $x_i \cong x(t_i)$, $0 \leq i \leq N - 1$, según (3.5.13) se obtiene:

$$x_{i+1} = x_i + f(t_i, x_i; \vec{\theta})\Delta t + g(t_i, x_i; \vec{\theta})\sqrt{\Delta t}Z_i, \quad Z_i \sim N(0; 1), \quad 0 \leq i \leq N - 1,$$

donde se ha utilizado que (véase propiedad MB.4 del Capítulo 3):

$$B(t_i + \Delta t) - B(t_i) \sim N(0; \sqrt{\Delta t}).$$

Ahora utilizando que la solución de la ecuación diferencial estocástica (4.5.13) es un proceso de difusión¹⁷, se cumplen las siguientes relaciones:

$$E \left[\frac{x_{i+1} - x_i}{\Delta t} - f(t_i, x_i; \vec{\theta}) \right] = \vartheta(\Delta t), \quad (4.5.14)$$

$$E \left[\frac{(x_{i+1} - x_i)^2}{\Delta t} \right] - \left(g(t_i, x_i; \vec{\theta}) \right)^2 = \vartheta(\Delta t), \quad (4.5.15)$$

que, en su notación de la derecha, indican que el tamaño de los órdenes de los momentos de los términos que aparezcan en los miembros de la izquierda.

¹⁷ ØKSENDAL, B.(1980) "Stochastic Diffencial Equations". Nueva York: Springer-Verlag, 2nd Edition.

Para estimar el vector de parámetros a partir de una muestra $\{x_0, x_1, \dots, x_{N-1}\}$ se utilizan las relaciones (4.5.14)-(4.5.15) en su versión muestral:

$$\sum_{i=0}^{N-1} f(t_i, x_i; \vec{\theta}) = \frac{1}{\Delta t} \sum_{i=0}^{N-1} (x_{i+1} - x_i), \quad (4.5.16)$$

$$\sum_{i=0}^{N-1} (g(t_i, x_i; \vec{\theta}))^2 = \frac{1}{\Delta t} \sum_{i=0}^{N-1} (x_{i+1} - x_i)^2. \quad (4.5.17)$$

En el caso del Modelo Log-Normal se escribe a partir de la muestra $\{x_0, x_1, \dots, x_{N-1}\}$ que en el Modelo Log-Normal son las cotizaciones $\{s_0, s_1, \dots, s_{N-1}\}$, resolviendo el sistema de ecuaciones (3.5.16) y (3.5.17), se estimaría el vector de parámetros $\vec{\theta} \in \mathbb{R}^2$. Obsérvese que el contexto del Modelo Log-Normal:

$$\begin{aligned} f(t_i, x_i; \vec{\theta}) &= \mu s_i, \\ g(t_i, x_i; \vec{\theta}) &= \sigma s_i, \end{aligned}$$

con la identificación $x_i = s_i$.

4.6 Validación del Modelo Log-Normal.

Realizar el ajuste del modelo es una tarea que requiere de diversas herramientas para conseguir un buen resultado. En el presente apartado se presentan una selección de medidas de bondad del ajuste que se utilizarán para la validación del Modelo Log-Normal aplicado al activo subyacente de Bankia (BKIA.MC). Una vez aplicadas, y en base a los resultados obtenidos, se evaluará la adecuación y ajuste del modelo para poder realizar las predicciones del activo. Las medias empleadas son:

- Error cuadrático medio (Root Mean Squared Error, ECM).
- Error porcentual absoluto medio (Mean Absolute Percentage Error, MAPE)
- Construcción de Intervalos de Confianza del 95%.
- Gráfico de observaciones y predicciones (puntuales y por intervalos).

➤ 4.6.1 Herramientas para la medición de la bondad del ajuste.

Para medir la bondad del ajuste del modelo se utilizarán dos medidas numéricas, el error cuadrático medio (ECM) y el error porcentual absoluto medio (MAPE).

➤ Error cuadrático medio (ECM).

La medida de bondad de ajuste ECM, mide la distancia euclídea (distancia ordinaria entre dos puntos de un espacio euclídeo) por término medio entre los valores observados y los estimados. Cuanto menor sea el resultado del ECM, entonces se podrá decir que el error en el modelo presentado será menor. Para calcular el ECM se emplearán las predicciones puntuales (función media) que se obtengan de aplicar el modelo teórico Log-Normal con los dos métodos de estimación de los parámetros planteados en el Apartado 4.5. Una vez calculados y contrastados los resultados del ECM que muestre cada método, será posible observar cuál de los presenta un mayor ajuste. Su valor está definido de la siguiente manera:

$$RMSE = \sqrt{\frac{\sum_{i=1}^K (S_i - \hat{S}_i)^2}{K}}.$$

➤ Error porcentual absoluto medio (MAPE).

El MAPE es la segunda medida de precisión de pronóstico a emplear. El error porcentual absoluto medio se calcula como el promedio de las diferencias absolutas entre los valores pronosticados y los observados divididos por las observaciones correspondientes, quedando expresado como un porcentaje de los valores reales.

El MAPE será calculado considerando las predicciones puntuales dadas a partir de la aplicación del modelo teórico Log-Normal con cada uno de los métodos de estimación de los parámetros explicados. Además de realizar el correspondiente contraste de los resultados que esta medida presente con cada uno de los métodos, individualmente, se considerará un error global de diagnóstico tolerable, cuando el MAPE obtenido sea inferior al 5-10%.

Su expresión matemática si se pronostican k periodos y los valores reales corresponden a k periodos, el MAPE se expresa como:

$$MAPE = \frac{100}{k} \sum_{i=1}^k \frac{|S_i - \hat{S}_i|}{S_i}$$

➤ 4.6.2. Validación por intervalos de confianza.

La tercera herramienta empleada para validar la aplicación del Modelo Log-Normal es la construcción y estudio de intervalos de confianza (IC) del 95%. A la hora de construir los IC hay dos elementos clave a tener en cuenta. En primer lugar, la amplitud del intervalo da la precisión de la estimación y por tanto, deberá ser la menor posible. En segundo lugar, la probabilidad de que el intervalo contenga el verdadero valor del parámetro a estimar (BKIA.MC), es decir, el nivel de confianza (interesa que sea el mayor posible).

Esta sección se centra en la estimación de la variable en cuestión (Bankia (BKIA.MC)) a partir del intervalo que tiene por extremos los valores de dos funciones para una muestra, siendo esta una estimación de la variable por intervalos de confianza.

A continuación, se presentan una serie de conceptos para facilitar la aplicación de los IC.

Supóngase que se selecciona una muestra aleatoria simple de una población descrita por la función $f(x; \theta)$ dependiente del parámetro θ que se pretende estimar. La cuestión se plantea tal que: se fija un nivel de confianza, que se denota por $1 - \alpha$, en donde $0 < \alpha < 1$, y se trata de determinar dos funciones $\theta_1(x_1, x_2, \dots, x_n)$ y $\theta_2(x_1, x_2, \dots, x_n)$ de forma que:

$$P[\theta_1(x_1, x_2, \dots, x_n) \leq \theta \leq \theta_2(x_1, x_2, \dots, x_n)] = 1 - \alpha .$$

Al intervalo $[\theta_1(x_1, x_2, \dots, x_n), \theta_2(x_1, x_2, \dots, x_n)]$ se le denomina *intervalo de confianza* del parámetro θ al nivel de confianza del $(1 - \alpha)100\%$. Es importante observar que sería un error afirmar que la probabilidad indicada anteriormente, es la probabilidad de que θ esté entre los números reales $\theta_1(x_1, x_2, \dots, x_n)$ y $\theta_2(x_1, x_2, \dots, x_n)$, puesto que θ no es una variable aleatoria sino un parámetro que tendrá un valor concreto. Las

variables aleatorias son $\theta_1(x_1, x_2, \dots, x_n)$ y $\theta_2(x_1, x_2, \dots, x_n)$ al variar la muestra, entonces la probabilidad anterior debe considerarse como la probabilidad de que el intervalo aleatorio $[\theta_1(x_1, x_2, \dots, x_n), \theta_2(x_1, x_2, \dots, x_n)]$ contenga el verdadero valor de θ .

Dicho en términos de frecuencias esto significa que de cada 100 muestras aleatorias que se tomen, cabe esperar que el $(1 - \alpha)100\%$ de ellas contenga el verdadero valor de θ entre $\theta_1(x_1, x_2, \dots, x_n)$ y $\theta_2(x_1, x_2, \dots, x_n)$.

Supóngase que (x_1, x_2, \dots, x_n) es una muestra aleatoria simple y que \bar{x} es la media muestral. La variable aleatoria \bar{x} media muestral se distribuye según una normal $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$, suponiendo que la población es $N[\mu, \sigma]$ con σ conocida. Esta propiedad permite asegurar que la variable tipificada $\frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$ sigue una $N[0, 1]$, y por lo tanto, fijando un nivel de confianza $1 - \alpha$, obtener el valor $\lambda_{\alpha/2}$ tal que:

$$P\left[\left|\frac{\bar{x} - \mu}{\sigma/\sqrt{n}}\right| \leq \lambda_{\alpha/2}\right] = 1 - \alpha.$$

Operando, resulta:

$$\begin{aligned} 1 - \alpha &= P\left[-\lambda_{\alpha/2} \leq \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \leq \lambda_{\alpha/2}\right], \\ &= P\left[-\lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \bar{x} - \mu \leq \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right], \\ &= P\left[-\bar{x} - \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq -\mu \leq -\bar{x} + \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right], \\ &= P\left[\bar{x} - \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right]. \end{aligned}$$

Con lo que las funciones deseadas son:

$$\begin{aligned} \theta_1 &= \bar{x} - \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \\ \theta_2 &= \bar{x} + \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}}. \end{aligned}$$

El intervalo de confianza para la media poblacional al nivel de confianza del $(1 - \alpha) 100\%$ es:

$$\left[\bar{x} - \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right].$$

Cabe decir que es posible ganar precisión si:

- Se pierde confianza (no interesa).
- Se aumenta el tamaño muestral n .

➤ 4.6.3. Gráficos comparativos.

Por último, calculada la función media (predicción puntual) y los intervalos de confianza (predicción por intervalos) para cada uno de los instantes de tiempo que comprenden el horizonte temporal observado y bajo cada uno de los métodos de estimación dados para la validación del modelo, se realizarán sus representaciones gráficas.

Los gráficos comparativos se emplearán para realizar un estudio visual del grado de ajuste entre las observaciones de la muestra (valor real de cotización del activo) y las predicciones puntuales obtenidas en cada uno de los métodos presentados. Para complementar el análisis cualitativo se presentarán también las medidas de bondad de ajuste anteriormente señaladas y una comparativa de los intervalos de confianza obtenidos en ambos casos.

4.7 Estimaciones.

Los dos modelos presentados a lo largo del trabajo tienen como objetivo realizar predicciones del activo subyacente. Con el modelo y su correspondiente ajuste debe ser posible predecir el valor medio (estimación puntual), y definir un intervalo que contenga el valor real con una probabilidad tolerable (estimación por intervalos).

Es posible obtener una estimación puntual mediante el valor esperado de la variable estudiada Y , asociado a determinados valores de las variables explicativas del modelo

ajustado. Por ello, es necesario calcular la estimación de Y correspondiente a su valor medio condicionado $E(Y|x_1, \dots, x_k)$, como se expresa:

$$\hat{Y} = s_0 e^{(\hat{\mu} - \frac{1}{2}\hat{\sigma}^2)t + \sigma\sqrt{t}Z}, \quad Z \sim N(0; 1).$$

Por otra parte, la predicción puntual debe completarse con la varianza de la estimación, conociéndose así el error de la predicción puntual y obteniendo con ambos valores, el intervalo de confianza. El intervalo de confianza a nivel α para la estimación de Y se muestra a continuación:

$$E[S(t)] \pm 1,96 \cdot S$$

5. APLICACIÓN DEL MODELO LOG-NORMAL AL ACTIVO SUBYACENTE BANKIA (BKIA.MC): METODOLOGÍA APLICADA

El activo subyacente escogido para la aplicación y validación del modelo Log-Normal es Bankia (BKIA.MC). En el presente trabajo la metodología aplicada se basa en el lenguaje de programación *Python* para la aplicación del Modelo Log-Normal.

A continuación, y a través de los diferentes apartados del capítulo, se realizará la estimación de los parámetros mediante los dos modelos planteados en el apartado 4.5. En segundo lugar, se validará el Modelo Log-Normal utilizando los resultados de las estimaciones obtenidas por el método de Máxima Verosimilitud. De igual manera, se seguirá la misma metodología para las estimaciones del método no paramétrico. Para validar el modelo, se utilizarán las estimaciones puntuales, media y varianza, de los Intervalos de Confianza (I.C.) del 95% y por medio de las medidas de bondad de ajuste del modelo explicadas en el Apartado 4.6.1, Error Cuadrático Medio y MAPE. Por último, se realizarán las predicciones para el activo subyacente Bankia (BKIA.MC)

5.1 PYTHON: Lenguaje de programación.

En este apartado se introducirá el lenguaje de programación *Python* utilizado para la aplicación y validación del Modelo Log-Normal.

Python fue creado a finales de los años ochenta por Guido van Rossum en el Centro para las Matemáticas y la Informática (CWI, Centrum Wiskunde & Informatica), en los Países Bajos, como un sucesor del lenguaje de programación ABC, capaz de manejar excepciones e interactuar con el sistema operativo Amoeba¹⁸

Python fue concebido para establecer una sintaxis que proporcione un código legible. La *Filosofía Python* es bastante análoga a la filosofía *Unix* y se basa en los principios de legibilidad y transparencia.

Se trata de un lenguaje de programación multiparadigma, es decir, soporta orientación a objetos, programación imperativa y, en menor medida, programación funcional. Es un lenguaje interpretado, usa tipado dinámico y es multiplataforma. Además, es posible añadir extensiones que interactúen con otros lenguajes.

Los principios del lenguaje fueron descritos por el desarrollador de *Python* Tim Peters en “*El Zen de Python*”:

- Bello es mejor que feo.
- Explícito es mejor que implícito.
- Simple es mejor que complejo.
- Complejo es mejor que complicado.
- Plano es mejor que anidado.
- Disperso es mejor que denso.
- La legibilidad cuenta.
- Los casos especiales no son tan especiales como para quebrantar las reglas.
- Lo práctico gana a lo puro.
- Los errores nunca deberían dejarse pasar silenciosamente.
- A menos que hayan sido silenciados explícitamente.

¹⁸ [The Making Of Python](#)

- Frente a la ambigüedad, rechaza la tentación de adivinar.
- Debería haber una -y preferiblemente sólo una- manera obvia de hacerlo.
- Aunque esa manera puede no ser obvia al principio a menos que usted sea holandés.
- Ahora es mejor que nunca.
- Aunque *nunca* es a menudo mejor que *ya mismo*.
- Si la implementación es difícil de explicar, es una mala idea.
- Si la implementación es fácil de explicar, puede que sea una buena idea.
- Los espacios de nombres (*namespaces*) son una gran idea ¡Hagamos más de esas cosas!

La elección de *Python* para la aplicación y posterior validación del modelo se debe a varios factores.

En primer lugar, ofrece una sintaxis simple, clara y sencilla y posee una gran cantidad de bibliotecas disponibles. En segundo lugar, su lenguaje está adaptado a los usuarios y es Multiplataforma por lo que puede ser utilizado para Linux, Windows o Mac. Por último, la facilidad para incorporar extensiones y los módulos de procesamiento HTML / XML y envío de correo ofrecen un conjunto de características que favorecen el desarrollo del presente trabajo.

Actualmente, diversas empresas y proyectos emplean el lenguaje con éxito; Google, Yahoo, la NASA, Industrias Ligh & Magic, y todas las distribuciones Linux, en las que *Python* cada vez representa un tanto por ciento mayor de los programas disponibles son buenos ejemplos. Tal y como se ha mencionado anteriormente, este Trabajo Fin de Grado culmina con el diseño de una página web, por ello, contar con una herramienta que posee bibliotecas potentes de cálculo, representación gráfica y facilidad para generar archivos que se pueden mostrar en páginas web resulta imprescindible.

5.2 Funcionamiento del código

El código base sobre el que se cimienta el trabajo fue proporcionado por el departamento de matemática multidisciplinar.

El código se divide en dos archivos .py; "*Funciones*" e "*Ivan_EXE*".

En el primero de ellos se trata de un archivo de funciones auxiliares que proporcionan todos los datos y parámetros necesarios para la generación de los resultados. El segundo de ellos proporciona los gráficos necesarios para comprender los resultados y la generación de tablas para la página web. El código se renueva cada día para evitar errores de acumulación a lo largo del paso de las semanas.

El flujo de aplicación del código en formato *Python (.py)* es el siguiente:

1. Extracción de la información necesaria: el histórico de cotizaciones.
2. Funciones auxiliares.
3. Generación de resultados.

Para poder realizar los cálculos necesarios se deben importar diversos paquetes.

A continuación, se muestran los paquetes importados en el archivo *“Funciones.py”*.

```
Imports que vamos a necesitar en este fichero de funciones auxiliares
"""
# urllib is a package that collects several modules for working with URLs
import urllib
# Con re podemos manipular con python expresiones con comodines *, ?, etc.
import re
# Funciones numericas de calculo
import numpy as np
# The datetime module supplies classes for manipulating dates and times in both simple and complex ways
import datetime

# Algoritmo de minimizacion. No hace falta. Lo dejamos comentado por si
# from scipy.optimize import minimize

# Funciones para dibujar (graficos) con python
# matplotlib is a python 2D plotting library
import matplotlib
# Force matplotlib to not use any Xwindows backend.
# Esta instrucción debe aparecer antes de cargar pylab
matplotlib.use( 'Agg' )
# Cargamos pylab
import pylab as pl
```

Figura 5.1: Paquetes

Fuente: Elaboración propia

➤ 5.2.1 Extracción de la información necesaria: el histórico de cotizaciones.

Se denomina como *Web scraping* a la técnica utilizada mediante programas de software para extraer información de sitios web. Estos programas simulan la navegación de un humano en la *World Wide Web* ya sea utilizando el protocolo HTTP manualmente, o mediante una extensión en el propio navegador predeterminado. Esta técnica es la utilizada para obtener un histórico de 30 cotizaciones del activo.

Concretamente, se enfoca más en la transformación de datos sin estructura en datos estructurados que pueden ser almacenados y analizados en un formato apropiado para su uso.

Así pues, gracias al código diseñado con *Python*, es posible obtener los datos necesarios de la página de referencia deseada: <https://es.finance.yahoo.com>.

Como ejemplo de aplicación del código se presenta a continuación el histórico de cotizaciones directamente extraído de *Yahoo Finance* en la Figura 5.2

PRECIOS						
Fecha	Apertura	Máximo	Mínimo	Cerrar	Volumen	Ajustes de Cierre*
8 de jul de 2016	0,60	0,63	0,60	0,62	42.286.200	0,62
7 de jul de 2016	0,61	0,62	0,59	0,60	28.909.400	0,60
6 de jul de 2016	0,63	0,63	0,60	0,61	42.128.000	0,61
5 de jul de 2016	0,64	0,65	0,63	0,63	27.736.700	0,63
4 de jul de 2016	0,65	0,66	0,64	0,65	21.919.600	0,65
1 de jul de 2016	0,65	0,66	0,64	0,65	53.865.500	0,65
30 de jun de 2016	0,63	0,68	0,62	0,65	101.347.300	0,65
29 de jun de 2016	0,67	0,68	0,64	0,64	56.942.600	0,64
28 de jun de 2016	0,63	0,68	0,62	0,66	79.967.900	0,66
27 de jun de 2016	0,63	0,68	0,58	0,61	131.377.000	0,61
24 de jun de 2016	0,55	0,64	0,53	0,57	94.631.500	0,57
23 de jun de 2016	0,71	0,72	0,70	0,72	33.286.500	0,72
22 de jun de 2016	0,71	0,72	0,70	0,70	27.264.000	0,70
21 de jun de 2016	0,72	0,72	0,70	0,70	28.424.200	0,70
20 de jun de 2016	0,72	0,72	0,71	0,72	20.154.300	0,72
17 de jun de 2016	0,69	0,71	0,68	0,70	27.576.600	0,70
16 de jun de 2016	0,68	0,68	0,66	0,68	35.604.000	0,68
15 de jun de 2016	0,70	0,71	0,68	0,69	34.811.700	0,69
14 de jun de 2016	0,71	0,72	0,68	0,68	37.035.300	0,68
13 de jun de 2016	0,72	0,73	0,71	0,72	22.003.400	0,72
10 de jun de 2016	0,75	0,75	0,72	0,73	26.698.600	0,73
9 de jun de 2016	0,75	0,76	0,75	0,75	25.200.200	0,75
8 de jun de 2016	0,76	0,77	0,75	0,75	16.315.600	0,75
7 de jun de 2016	0,77	0,78	0,76	0,77	15.056.100	0,77
6 de jun de 2016	0,77	0,78	0,76	0,77	13.116.100	0,77
3 de jun de 2016	0,80	0,80	0,76	0,76	23.110.700	0,76
2 de jun de 2016	0,78	0,81	0,78	0,79	26.534.500	0,79
1 de jun de 2016	0,78	0,79	0,77	0,78	11.874.300	0,78
31 de may de 2016	0,80	0,81	0,78	0,78	28.226.100	0,78
30 de may de 2016	0,81	0,82	0,80	0,81	10.959.300	0,81

* Precio de cierre ajustado por dividendos y splits.

Figura 5.2: Cotización BKIA.MC.

Fuente: <https://es.finance.yahoo.com/q/hp?s=BKIA.MC>.

Para evitar redundancias se ha decidido denominar al Método de Máxima Verosimilitud como Método 1 y al uso del Método No Paramétrico como Método 2.

A continuación se muestra el código que, a su ejecución, permite extraer la información requerida de la página web.

El código queda reflejado en la Figura 5.3.

```

"""
Captura

Esta función lee los datos correspondientes a los valores de la cotización del activo solicitado desde la web
de yahoo finance y devuelve los valores numéricos.
El activo debe darse como cadena y con el formato técnico.
"""
def captura( activo ):

    # Direccion de la pagina web de donde extraeremos Los datos
    url = "https://es.finance.yahoo.com/q//hp?s=" + 'BKIA.MC'
    # Abrimos la pagina web
    htmlfile = urllib.urlopen( url )
    # Leemos los datos que hay en la pagina web abierta
    htmltext = htmlfile.read( )
    # Definimos el patron en el que tenemos que buscar para obtener los datos de las cotizaciones
    regex = '<td class="yfnc_tabledata1" align="right">(.*?)</td>'
    # Compile a regular expression pattern into a regular expression object, which can be used for matching
    pattern = re.compile( regex )
    # Ahora es donde buscamos y extraemos de la pagina web los valores de cotizacion
    price = re.findall( pattern, htmltext )

    # Cribamos los datos. Desde el dato 5 al 240 (180 para sólo 30 datos) con saltos de 6 en 6, que corresponden
    # a los 'Ajustes de cierre'
    share = price[ 5 : 240 : 6 ]
    # Ponemos la lista del revés, para poner los valores más antiguos los primeros
    share = share[ : : -1 ]
    # Cambiamos las comas separadoras de los números por puntos
    share = [ x.replace( ',', '.' ) for x in share ]
    # Y ahora pasamos las cadenas por floats
    share = [ float( x ) for x in share ]
    # Y finalmente a array de numpy
    share = np.array( share )

    # Devolvemos los datos de cotización
    return share

```

Figura 5.3: Web scrapping.

Fuente: Elaboración propia.

Como resultado de la ejecución del *script* obtenemos las últimas treinta cotizaciones de Bankia (BKIA.MC):

Fecha 25/08/2016

Variable Explorer Spyder (Python 2.7)

```

[ 0.67 0.67 0.67 0.66 0.68 0.68 0.7 0.69 0.68 0.7 0.68 0.69
 0.68 0.65 0.65 0.66 0.68 0.69 0.7 0.72 0.72 0.71 0.71 0.7
 0.71 0.71 0.7 0.7 0.71 0.73]

```

Figura 5.4: Resultado web scrapping.

Fuente: Elaboración propia.

➤ 5.2.2 Funciones auxiliares.

Una vez extraído el histórico de cotizaciones es posible empezar a aplicar el Modelo Log-Normal para el subyacente.

A continuación se muestran las funciones auxiliares necesarias. En cada una de ellas se explican las variables que se incluyen.

En la Figura 5.6 se muestran las funciones necesarias para el cálculo de la media, la desviación típica, la varianza, la media estimada y la media ± 2 veces la desviación típica.

```

"""
Función de la media estimada de la solución
Los parámetros son
* t -> instante temporal
* cond_ini -> primer valor de la lista share de cotizaciones
* mu_estimada -> valor estimado de la media
"""
def media( t, cond_ini, mu_estimada ):

    # Devolvemos el valor inicial por la exponencial del valor estimado por t
    return cond_ini * np.exp( mu_estimada * t )

#####
"""
Función que calcula la varianza
Los parámetros son
* t -> instante temporal
* cond_ini -> primer valor de la lista share de cotizaciones
* mu_estimada -> valor estimado de la media
* sigma_estimada -> valor estimado de la varianza
"""
def varianza( t, cond_ini, mu_estimada, sigma_estimada ):

    # Devolvemos  $S_0^2 * \exp( 2 * \mu * t ) * ( \exp( \sigma^2 * t ) - 1 )$ 
    return np.power( cond_ini, 2 ) * np.exp( 2 * mu_estimada * t ) * ( np.exp( t * np.power( sigma_estimada, 2 ) ) - 1 )

#####
"""
Función que calcula la media menos/mas 2 veces la desviación típica
Los parámetros son
* t -> instante temporal
* cond_ini -> primer valor de la lista share de cotizaciones
* mu_estimada -> valor estimado de la media
* sigma_estimada -> valor estimado de la varianza
"""
def mediaVdt( t, cond_ini, mu_estimada, sigma_estimada ):

    dt = 2.0 * np.sqrt( varianza( t, cond_ini, mu_estimada, sigma_estimada ) )
    me = media( t, cond_ini, mu_estimada )
    return [ me - dt, me + dt ]

"""
Media de cada paso del esquema de Euler
Los parámetros son
* t -> instante temporal
* share -> la lista con los valores de la cotización
* drift -> el drift
* D -> Incremento del paso temporal del esquema numérico de Euler - Maruyama
"""
def mu( t, share, drift, D ):

    # Valor de cotización en el instante actual
    valor_actual = share[ t - 1 ]
    # Devolvemos
    return valor_actual + valor_actual * drift * D

#####
"""
Desviacion tipica de cada paso del esquema de Euler
Los parámetros son
* t -> instante temporal
* share -> la lista con los valores de la cotización
* volatility -> la volatility
* D -> Incremento del paso temporal del esquema numérico de Euler - Maruyama
"""
def sigma( t, share, volatility, D ):

    # Valor de cotización en el instante actual
    valor_actual = share[ t - 1 ]
    # Devolvemos
    return valor_actual * volatility * np.sqrt( D )

```

Figura 5.5: Media Est. , Varianza y Media2.

Fuente: Elaboración propia.

La Figura 5.7 corresponde a las funciones de estimación de los parámetros μ y σ .

```

"""
Definición de la función de estimación exacta de mu y sigma a partir de la función de verosimilitud (METODO 1)
Los parámetros son
* share -> la lista con los valores de la cotización
* D -> Incremento del paso temporal del esquema numérico de Euler - Maruyama
"""
def estimacion_modelo1( share, D ):

    # Número de shares
    n = len( share )
    # Rango de valores para Los sumatorios
    rg = xrange( 1, n )

    # Constante que va a multiplicar a ambas estimaciones
    cte = 1.0 / ( n * D )
    # Cociente de shares sucesivos - 1
    quo = [ ( share[ t ] / share[ t - 1 ] ) - 1 for t in rg ]

    # Estimación de mu
    mu = cte * np.sum( quo )

    # Estimación de sigma
    sigma = np.sqrt( cte * np.sum( np.power( quo - mu * D, 2 ) ) )

    # Devolvemos
    return [ mu, sigma ]

#####
"""
Definición de la función de estimación exacta de mu y sigma a partir de la función de verosimilitud (METODO 2)
Los parámetros son
* share -> la lista con los valores de la cotización
* D -> Incremento del paso temporal del esquema numérico de Euler - Maruyama
"""
def estimacion_modelo2( share, D ):

    # Rango para Los for
    rg = xrange( 1, len( share ) - 1 )
    # Definimos la constante
    const = 1 / D

    ##### Para obtener el drift
    # Obtenemos el denominador
    denom = np.sum( share )
    # Ahora el numerador
    h = np.array( [ share[ i + 1 ] - share[ i ] for i in rg ] )
    num = np.sum( h )
    # Y ahora el drift
    drift = num / denom

    ##### Para obtener la volatilidad
    # Denominador
    b = np.array( share ** 2 )
    denom = np.sum( b )
    # Numerador
    j = np.array( [ ( share[ i + 1 ] - share[ i ] ) ** 2 for i in rg ] )
    num = np.sum( j )
    # Y ahora la volatilidad
    volatility = np.sqrt( num / denom )

    # Devolvemos
    return [ drift, volatility ]

```

Figura 5.6: Funciones de estimación según Método 1 y 2.

Fuente: Elaboración propia

Como resultado de la aplicación del código obtenemos la Tabla 5.1 donde se estiman los valores de μ (*est_mu* para el Método 1 y *est_mu2* para el Método 2) y σ (*est_sigma* para el Método 1 y *est_sigma2* para el Método 2).

est_mu	float64	1	0.0020972669201433664
est_mu2	float64	1	0.0028985507246376782
est_sigma	float64	1	0.018012318186727945
est_sigma2	float64	1	0.018701686403357384

Tabla 5.1: Estimación mu y sigma.

Fuente: Elaboración propia.

En la Figura 6.8 se definen las funciones necesarias para el cálculo del ECM (ó ECM) y el MAPE.

```

"""
Errores. Esta función calcula y devuelve el error cuadrático medio (ECM) y el MAPE en formato adecuado
Los parámetros son
* share -> lista de cotizaciones
* mu_estimada -> valor estimado de la media
"""
def errores( share, mu_estimada ):

    # Primer y último share
    primershare = share[ 0 ]
    # Longitud del share
    n = len( share )

    # Calculamos los valores de la media
    m = np.array( [ media( t, primershare, mu_estimada ) for t in xrange( n ) ] )

    # ECM
    ecm = np.linalg.norm( m - share ) / np.sqrt( n )

    # MAPE
    mape = 100 * np.sum( np.abs( ( m - share ) / share ) ) / n

    # Devolvemos
    return [ ecm, mape ]

```

Figura 5.7: Errores.

Fuente: Elaboración propia.

Como resultado se obtiene la Tabla 5.2.

ecm	float64	1	0.019058655111757656
ecm2	float64	1	0.018129000875533791
mape	float64	1	2.0637909278992734
mape2	float64	1	1.9329259515685098

Tabla 5.2: Estimación MAPE y ECM.

Fuente: Elaboración propia.

En la Figura 6.9 se define la función *fechas_dia* que permite establecer las cabeceras de las tablas de manera cronologica teniendo en cuenta los días que abre la Bolsa.

```

"""
fechas_dia. Esta función devuelve las fechas de los 5 días anteriores y los 5 días posteriores, incluyendo hoy,
quitando los sábados y domingos que no tienen cotización. De esta forma, las cabeceras de las tablas estarán
correctas.
"""
def fechas_dia( dummy ):

    # El día de hoy: 0 es Lunes, 1 martes, 2 miércoles, 3 jueves y 4 viernes
    dia = datetime.datetime.today().weekday()

    if dia == 0:
        des = [ 0, 1, 2, 3, 4 ]

    elif dia == 1:
        des = [ 0, 1, 2, 3, 6 ]

    elif dia == 2:
        des = [ 0, 1, 2, 5, 6 ]

    elif dia == 3:
        des = [ 0, 1, 4, 5, 6 ]

    elif dia == 4:
        des = [ 0, 3, 4, 5, 6 ]

    # Ahora calculamos ant
    ant = [ 7 - des[ i ] for i in xrange( 5 ) ]

    # Devolvemos Las Listas
    return [ ant, des ]

```

Figura 5.8: Fechas.

Fuente: Elaboración propia.

En la Figura 6.10 se definen los momentos temporales necesarios; la hora actual, el día en el que nos encontramos, y los cinco días anteriores y posteriores.

```

fechas. Devuelve la hora actual, el día de hoy, los 5 días anteriores y los 5 días posteriores
"""
def fechas( dummy ):

    # Calculamos Las Listas para tener Los 5 días anteriores y posteriores sin sábados y domingos
    [ ant, des ] = fechas_dia( 0 )

    # Ahora
    ahora = datetime.datetime.now()
    # Instante actual
    ia = ahora.strftime("%H:%M, %d/%m/%y")

    # 5 días anteriores, en formato adecuado, y la ponemos del revés para que el orden sea el adecuado
    antes5 = [ ( ahora - datetime.timedelta( days = i ) ).strftime('%d/%m/%y') for i in ant ]
    # 5 días después, en formato adecuado
    despues5 = [ ( ahora + datetime.timedelta( days = i ) ).strftime('%d/%m/%y') for i in des ]

    # Hoy
    hoy = ahora.strftime('%d/%m/%y')

    # Devolvemos [ instante actual, hoy, antes5, despues5 ]
    return [ ia, hoy, antes5, despues5 ]

```

Figura 5.9: Fechas definidas.

Fuente: Elaboración propia.

Como ejemplo y a fecha de 05/08/16 se obtiene la Tabla 5.3.

antes5	list	5	['18/08/16', '19/08/16', '22/08/16', '23/08/16', '24/08/16']
despues5	list	5	['25/08/16', '26/08/16', '29/08/16', '30/08/16', '31/08/16']
hoy	str	1	25/08/16
ia	str	1	01:20, 25/08/16

Tabla 5.3: Antes, después, hoy, ia.

Fuente: Elaboración propia.

Una vez definidas las funciones necesarias, es menester representar los resultados obtenidos en gráficos y tablas que ayuden al usuario de la página web a identificar de manera rápida y eficaz la situación del activo en el mercado y el valor de las predicciones estimadas.

➤ 6.2.3 Generación de resultados.

Una vez ejecutado el código, se generan las gráficas y tablas correspondientes y que son directamente aplicables a la página web. Sin embargo, a continuación se presentan los datos en bruto obtenidos del *Variable Explorer* de *Spyder*.

```
Método 1 25/08/2016 IC 97.5%
[ 0.67          0.64703713  0.63861468  0.63255774  0.62771563  0.62364411
  0.62011604  0.61699694  0.61419942  0.61166297  0.60934365  0.6072084
  0.6052316   0.60339293  0.60167597  0.60006722  0.59855542  0.59713109
  0.59578613  0.59451358  0.59330739  0.59216228  0.59107362  0.59003727
  0.58904957  0.58810726  0.58720738  0.58634729  0.58552455  0.584737
  0.58398263  0.58325961  0.58256627  0.58190105  0.58126254]

[ 0.67          0.6970376   0.70954718  0.71970367  0.72865781  0.73685386
  0.74451899  0.75178774  0.75874753  0.76545891  0.77196584  0.77830144
  0.78449137  0.79055597  0.79651172  0.80237215  0.80814855  0.81385045
  0.81948599  0.82506217  0.83058508  0.83606002  0.8414917   0.84688426
  0.85224141  0.85756647  0.86286242  0.86813197  0.87337755  0.8786014
  0.88380556  0.8889919   0.89416214  0.89931786  0.90446053]

Método 2 25/08/2016 IC 97.5%
[ 0.67          0.64680964  0.63824249  0.63205527  0.62709057  0.6229015
  0.61925955  0.61602938  0.61312304  0.61047964  0.60805496  0.60581569
  0.60373605  0.60179558  0.59997774  0.59826892  0.5966578   0.59513482
  0.59369181  0.59232177  0.5910186   0.58977698  0.58859223  0.58746019
  0.58637717  0.58533988  0.58434534  0.58339086  0.58247403  0.58159263
  0.58074464  0.57992823  0.5791417   0.57838348  0.57765215]

[ 0.67          0.69708005  0.70954818  0.71964771  0.72853607  0.7366602
  0.74424862  0.75143673  0.75831248  0.76493682  0.77135401  0.77759736
  0.78369271  0.78966055  0.79551745  0.80127705  0.8069507   0.81254802
  0.81807719  0.82354525  0.82895834  0.8343218   0.83964037  0.84491821
  0.85015907  0.85536628  0.86054284  0.86569147  0.87081464  0.87591458
  0.88099335  0.88605283  0.89109475  0.8961207   0.90113216]
```

Figura 5.10: IC Datos.

Fuente: Elaboración propia.

Como se puede observar en la Figura 5.12 los datos en bruto están divididos según el método.

```

Estimación Medias 25/08/2016

Método 1
[ 0.67          0.67203737  0.67408093  0.67613071  0.67818672  0.68024898
 0.68231752  0.68439234  0.68647347  0.68856094  0.69065474  0.69275492
 0.69486148  0.69697445  0.69909384  0.70121968  0.70335199  0.70549077
 0.70763606  0.70978788  0.71194623  0.71411115  0.71628266  0.71846076
 0.72064549  0.72283687  0.7250349  0.72723963  0.72945105  0.7316692
 0.7338941  0.73612576  0.7383642  0.74060946  0.74286154]

Método 2
[ 0.67          0.67194485  0.67389534  0.67585149  0.67781332  0.67978085
 0.68175409  0.68373305  0.68571776  0.68770823  0.68970448  0.69170653
 0.69371438  0.69572806  0.69774759  0.69977298  0.70180425  0.70384142
 0.7058845  0.70793351  0.70998847  0.71204939  0.7141163  0.7161892
 0.71826812  0.72035308  0.72244409  0.72454117  0.72664433  0.7287536
 0.730869  0.73299053  0.73511822  0.73725209  0.73939216]

```

Figura 5.11: Estimación Medias.

Fuente: Elaboración propia.

A continuación, en la Figura 5.11 se presenta el código correspondiente a la elaboración del gráfico de ajuste y predicción de cada método.

Se genera un gráfico (véase Gráfico 5.1) para facilitar la progresión de los resultados a lo largo del tiempo con la media, los Intervalos de Confianza del 95% y las predicciones para los próximos 5 días.

El gráfico se genera para los dos métodos. Se ha decidido presentar uno de los dos gráficos ya que ambos están a disposición del usuario en la página web para cualquier consulta.

```

Gráfico
Con esta función generamos las gráficas de ajuste y predicción de cada método
Parámetros
* share -> datos de cotización
* est_medias -> estimación de la media
* est_ic_2_5 -> IC inferior
* est_ic_97_5 -> IC superior
* titulo_grafico -> titulo
* nombre_imagen -> nombre del fichero imagen que se va a generar
"""
def grafico( share, est_medias, est_ic_2_5, est_ic_97_5, titulo_grafico, nombre_imagen ):

    # Para el eje X
    ejeX = np.array( xrange( 35 ) )

    # Plot de Los datos
    pl.plot( ejeX[ : 30 ], share, '-bo', label = 'Cotizacion real' )
    # Plot de la media de Los datos y previsión
    pl.plot( ejeX, est_medias, 'g', label = 'Media' )
    # Intervalo de confianza inferior
    pl.plot( ejeX, est_ic_2_5, 'r', label = 'IC 95%' )
    # Intervalo de confianza superior
    pl.plot( ejeX, est_ic_97_5, 'r' )
    # Previsión para Los próximos 5 días
    pl.plot( ejeX[ 30 : ], est_medias[ 30 : ], '-yo', label = 'Prevision' )
    # Línea vertical que separa Los datos de Las previsiones
    pl.axvline( 29, color = 'k' )

    # Título del gráfico
    pl.title( titulo_grafico )
    # Título de Los ejes
    pl.xlabel( 'Tiempo (días)' )
    pl.ylabel( 'Cotizacion' )
    # Creamos La Leyenda del gráfico, abajo a La izquierda
    pl.legend( loc = 'upper left' )

    # Vamos a determinar Los topes de Los ejes.
    # El horizontal está claro
    pl.xlim( 0.0, 36.0 )
    # El vertical va a depender de Los valores máximos y mínimos de Los intervalos de confianza
    mx = np.max( est_ic_97_5 )
    mn = np.min( est_ic_2_5 )
    pl.ylim( 0.7 * mn, 1.3 * mx )

    # Dibujamos y guardamos en un fichero y chapamos el gráfico
    pl.savefig( nombre_imagen )
    pl.close()

    # Devolvemos nada
    return True

```

Figura 5.12: Gráfico M1 y M2.

Fuente: Elaboración propia.

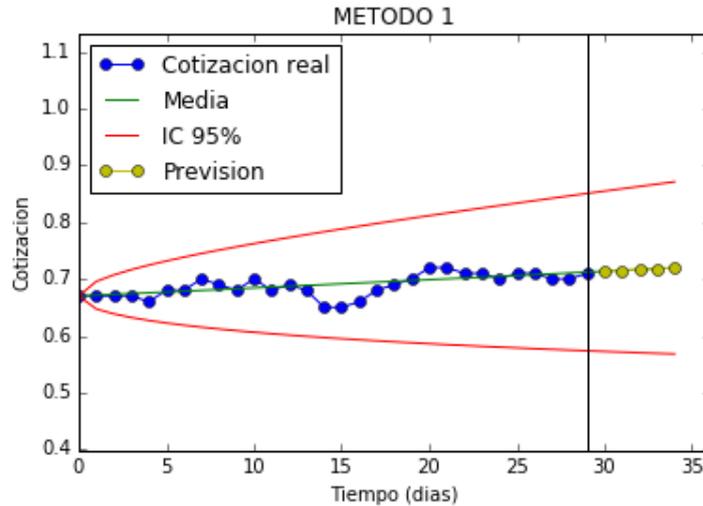


Gráfico 5.1: Previsión M1.

Fuente: Elaboración propia.

Por otra parte se presenta también el código correspondiente al gráfico generado para observar la comparativa entre las predicciones y el valor real a lo largo de los 10 últimos días. Son representados mediante la Figura 5.12 y el Gráfico 5.2.

```
grafico_comparativo
Gráfica de la comparativa entre las predicciones y los valores reales de los 10 últimos días
"""
def grafico_comparativo( reales, historico_media, historico_ic275, historico_ic975, titulo_grafico, nombre_imagen ):

    # Para el eje X
    ejeX = np.array( xrange( 10 ) )

    # Dibujamos los valores reales
    pl.plot( ejeX, reales, '-bo', label = 'Cotizacion real' )
    # Dibujamos las predicciones hechas
    pl.plot( ejeX, historico_media, 'g', label = 'Predicciones' )
    # Dibujamos el intervalo de confianza inferior
    pl.plot( ejeX, historico_ic275, 'r' )
    # Dibujamos el intervalo de confianza superior
    pl.plot( ejeX, historico_ic975, 'r', label = 'IC 95%' )

    # Título del gráfico
    pl.title( titulo_grafico )
    # Título de los ejes
    pl.xlabel( 'Tiempo (días)' )
    pl.ylabel( 'Cotizacion' )
    # Creamos la leyenda del gráfico, abajo a la izquierda
    pl.legend( loc = 'upper left' )

    # Vamos a determinar los topes de los ejes.
    # El horizontal está claro
    pl.xlim( 0.0, 9.0 )
    # El vertical va a depender de los valores máximos y mínimos de los intervalos de confianza
    mx = np.max( historico_ic975 )
    mn = np.min( historico_ic275 )
    pl.ylim( 0.7 * mn, 1.3 * mx )

    # Dibujamos y guardamos en un fichero y chapamos el gráfico
    pl.savefig( nombre_imagen )
    pl.close()

    # Devolvemos nada
    return True
```

Figura 5.13 Gráfico comparativo.

Fuente: Elaboración propia.

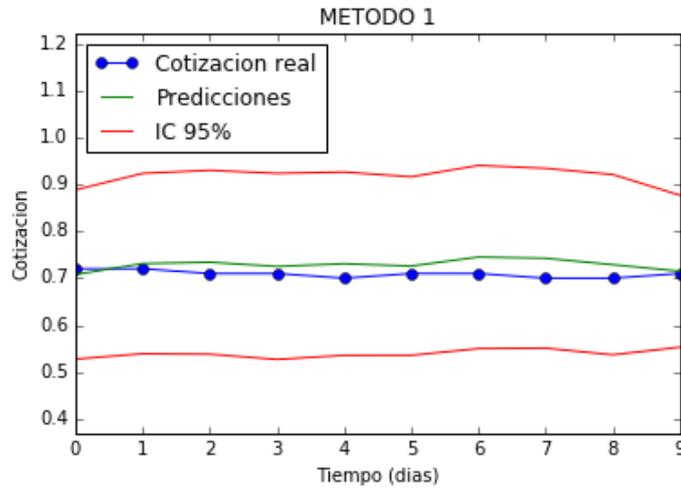


Gráfico 5.2: Comparativa M1.

Fuente: Elaboración propia.

➤ Predicciones e Intervalos de Confianza 95%

El primer archivo que genera el programa proporciona dos tablas. En la primera tabla (véase Tabla 5.4 y Tabla 5.5) aparecen los resultados generados por el procedimiento descrito en el Apartado 5.4 es decir, las predicciones y los Intervalos de Confianza del 95% para los próximos 5 días del subyacente Bankia (BKIA.MC). Esta tabla se genera con los resultados para los dos métodos propuestos). Además se incluyen las medidas de bondad de ajuste del modelo, MAPE y ECM.

Cabe señalar que en las tablas proporcionadas se incluye una fecha de actualización de los datos. Los activos que cotizan en el IBEX-35 varían su valor diariamente. Por esta razón, el valor de los datos presentados en la versión física de la presente memoria pueden variar considerablemente a lo largo del tiempo.

Por este motivo, es necesario el diseño y aplicación de una página web, ya que permanecerá actualizada a lo largo de las semanas y permitirá ser una referencia robusta el día que sea necesario consultar la robustez del presente trabajo.

El código que genera las tablas se describe a continuación en la Figura 5.13

```

"""
Tablas de datosactual?.js
"""
def tabla1( nom_empr, nombre_fichero, despues5, ic_sup, prediccion, ic_inf, ia, ecm, mape ):

# Primera tabla
cadena = '<center> '
cadena = cadena + '<table cellpadding = "4" border = "1" cellspacing = "0"> '
cadena = cadena + '<tbody> '

# CABECERA
cadena = cadena + '<tr align = "center"> '
cadena = cadena + '<th style = "height: 30px;"> ' + nom_empr + ' </th> '
for i in xrange( 5 ):
    cadena = cadena + '<th style = "width: 80px;"> ' + despues5[ i ] + ' </th> '
cadena = cadena + '</tr> '

# IC 95% SUPERIOR
cadena = cadena + '<tr align = "center"> '
cadena = cadena + '<th style = "height: 30px;"> IC 95% superior </th> '
for i in xrange( 5 ):
    cadena = cadena + '<td> ' + '{0:.3f}'.format( ic_sup[ i ] ) + ' </td> '
cadena = cadena + '</tr> '

# PREDICCIÓN
cadena = cadena + '<tr align = "center"> '
cadena = cadena + '<th style = "height: 30px;"> Predicción </th> '
for i in xrange( 5 ):
    cadena = cadena + '<td bgcolor="#bdbdbd"> ' + '{0:.3f}'.format( prediccion[ i ] ) + ' </td> '
cadena = cadena + '</tr> '

# IC 95% INFERIOR
cadena = cadena + '<tr align = "center"> '
cadena = cadena + '<th style = "height: 30px;"> IC 95% inferior </th> '
for i in xrange( 5 ):
    cadena = cadena + '<td> ' + '{0:.3f}'.format( ic_inf[ i ] ) + ' </td> '
cadena = cadena + '</tr> '

# Cierre de la 1a primera tabla
cadena = cadena + '</tbody> '
cadena = cadena + '</table> '

# Corte
cadena = cadena + '<br> '

# Segunda tabla de ecm y mape
cadena = cadena + '<table cellpadding = "4" border = "1" cellspacing = "0"> '
cadena = cadena + '<caption align = "bottom" > Last update: ' + ia + ' </caption> '
cadena = cadena + '<tbody> '

# Títulos
cadena = cadena + '<tr align = "center"> '
cadena = cadena + '<td style = "width: 100px;" > MAPE </td> '
cadena = cadena + '<td style = "width: 100px;" > ECM </td> '
cadena = cadena + '</tr> '

# Errores
cadena = cadena + '<tr align = "center"> '
cadena = cadena + '<td style = "width: 120px;"> ' + '{0:.3f}'.format( round( mape, 3 ) ) + ' % </td> '
cadena = cadena + '<td style = "width: 120px;"> ' + '{0:.3f}'.format( round( ecm, 3 ) ) + ' </td> '
cadena = cadena + '</tr> '

# Cierre de la 1a tabla
cadena = cadena + '</tbody> '
cadena = cadena + '</table> '
cadena = cadena + '</center> '
cadena = cadena + '<br> '

# Añadimos el comando de js
cadena = 'document.write(' + chr(39) + cadena + chr(39) + ') '

# Abrimos, escribimos, cerramos y adiós
f = open( nombre_fichero, 'w' )
f.write( cadena )
f.close()

# Devolvemos nada
return True

```

Figura 5.14: Tabla Datosactual.

Fuente: Elaboración propia.

BANKIA	25/08/16	26/08/16	29/08/16	30/08/16	31/08/16
IC 95% superior	0.884	0.889	0.894	0.899	0.904
Prediccion	0.734	0.736	0.738	0.741	0.743
IC 95% inferior	0.584	0.583	0.583	0.582	0.581

MAPE	ECM
2.064 %	0.019

Last update: 01:20, 25/08/16

Tabla 5.4: Predicción, IC 95% Mape y Ecm Método de Máxima Verosimilitud.

Fuente: <http://cotizaccion.imm.upv.es/>.

BANKIA	25/08/16	26/08/16	29/08/16	30/08/16	31/08/16
IC 95% superior	0.881	0.886	0.891	0.896	0.901
Prediccion	0.731	0.733	0.735	0.737	0.739
IC 95% inferior	0.581	0.580	0.579	0.578	0.578

MAPE	ECM
1.933 %	0.018

Last update: 01:20, 25/08/16

Tabla 5.5 Predicción e IC 95% Mape y Ecm Método no Paramétrico.

Fuente: <http://cotizaccion.imm.upv.es/>.

➤ Comparación de predicciones con el valor real.

La tabla que se muestra a continuación (véase Tabla 5.6) se genera a través del código mostrado en la Figura 5.14 y compara las 5 últimas predicciones proporcionadas por cada método con la cotización real del activo subyacente. De esta manera, es posible comparar rápidamente la efectividad del modelo.

```

"""
Tablas de compara?.js
"""
def tabla_compara( nom_empr, nombre_fichero, antes5, ic_sup, prediccion, real, ic_inf, ia ):

    # Empezamos
    cadena = '<center> '
    cadena = cadena + '<table cellpadding = "4" border = "1" cellspacing = "0"> '
    cadena = cadena + '<caption align = "bottom" > Last update: ' + ia + ' </caption> '

    # CABECERA
    cadena = cadena + '<tr align = "center"> '
    cadena = cadena + '<th style = "height: 30px;"> ' + nom_empr + ' </th> '
    for i in xrange(5):
        cadena = cadena + '<th style = "width: 80px;"> ' + antes5[i] + ' </th> '
    cadena = cadena + '</tr> '

    # IC 95% SUPERIOR
    cadena = cadena + '<tr align = "center"> '
    cadena = cadena + '<th style = "height: 30px;"> IC 95% superior </th> '
    for i in xrange(5):
        cadena = cadena + '<td> ' + '{0:.3f}'.format(ic_sup[i]) + ' </td> '
    cadena = cadena + '</tr> '

    # PREDICCION
    cadena = cadena + '<tr align = "center"> '
    cadena = cadena + '<th style = "height: 30px;"> Prediccion </th> '
    for i in xrange(5):
        cadena = cadena + '<td bgcolor="#E6E6E6"> ' + '{0:.3f}'.format(prediccion[i]) + ' </td> '
    cadena = cadena + '</tr> '

    # VALORES REALES
    cadena = cadena + '<tr align = "center"> '
    cadena = cadena + '<th style = "height: 30px;"> Real </th> '
    for i in xrange(5):
        cadena = cadena + '<td bgcolor="#A4A4A4"> ' + '{0:.3f}'.format(real[i]) + ' </td> '
    cadena = cadena + '</tr> '

    # IC 95% INFERIOR
    cadena = cadena + '<tr align = "center"> '
    cadena = cadena + '<th style = "height: 30px;" > IC 95% inferior </th> '
    for i in xrange(5):
        cadena = cadena + '<td> ' + '{0:.3f}'.format(ic_inf[i]) + ' </td> '
    cadena = cadena + '</tr> '

    # Cierre de la 1a primera tabla
    cadena = cadena + '</tbody> '
    cadena = cadena + '</table> '
    cadena = cadena + '</center>'

    # Añadimos el comando de js
    cadena = 'document.write(' + chr(39) + cadena + chr(39) + ')'

    # Abrimos, escribimos, cerramos y adiós
    f = open(nombre_fichero, 'w')
    f.write(cadena)
    f.close()

    # Devolvemos nada
    return True

```

Figura 5.15: Tabla Compara.

Fuente: Elaboración propia.

MÉTODO 1	Jul 5, 2016	Jul 6, 2016	Jul 7, 2016	Jul 8, 2016	Jul 11, 2016
I.C. 95%	0.29	0.29	0.29	0.29	0.29
PREDICCIÓN	0.62	0.62	0.64	0.64	0.64
REAL	0.63	0.61	0.60	0.62	0.63
I.C. 95%	0.95	0.95	0.99	0.99	0.99

Última Actualización: 15:31 12-07-2016

Tabla 5.6: Comparación predicción y cotización real.

Fuente: <http://cotizaccion.imm.upv.es/>.

Tanto el código como las tablas también se generan para el Método no Paramétrico.

➤ Comparación entre métodos

La comparación de los dos métodos sigue el esquema de las anteriores predicciones.

En la Tabla 5.7 se muestran las 5 últimas predicciones realizadas por ambos métodos de estimación de los parámetros comparados con las cotizaciones reales.

Última Actualización: Tue Jul 12 15:31:13 2016

FECHA	Jul 5, 2016	Jul 6, 2016	Jul 7, 2016	Jul 8, 2016	Jul 11, 2016
MÉTODO 1	0.62	0.62	0.64	0.64	0.64
MÉTODO 2	0.60	0.60	0.64	0.64	0.64
REAL	0.63	0.61	0.60	0.62	0.63

Tabla 5.7 Comparación predicciones métodos y valor real.

Fuente: <http://cotizaccion.imm.upv.es/>.

➤ Archivo Ivan_EXE.py.

La generación de las gráficas y las tablas se apoya en la ejecución del archivo Ivan_EXE.py. Para facilitar la explicación, se ha decidido incluir las gráficas y las tablas junto con la explicación del archivo Funciones.py. Por este motivo, a continuación se presenta el código del archivo Ivan_EXE.py ya que este es el archivo ejecutable que permite el estudio y desarrollo del modelo.

En primer lugar se importan los paquetes necesarios (incluido el archivo *Funciones.py*) y acto seguido se procede a la captura de los datos de cotización.

```
Importamos los paquetes que necesitamos
"""
# Funciones numericas de calculo
import numpy as np

# Funciones definidas por nosotros
from Funciones import *

# Funciones para el depurado del codigo. En condiciones normales estaran comentadas
# import pdb

# Paso temporal del esquema numérico de Euler - Maruyama
D = 1.0

#####
"""
CAPTURA DE DATOS DE COTIZACION
Hay que recordar que, en las tablas, hay que cambiar el nombre de
"""
# Nombre de la empresa sobre la que vamos a trabajar y su código IBEX-35 para recuperar Los datos
NOM = 'BANKIA'
CODE = 'BKIA.MC'

# Empezamos capturando Los datos de cotización. En particular, cogemos 40, Los 30 últimos para ajuste
# y predicción, y Los 10 anteriores para históricos.
shareTodo = captura( CODE )
# Seleccionamos Los 30 últimos datos
share = shareTodo[ 10 : ]

# Número de datos de cotizaciones
n = len( share )
# Rango para Los for
rg30 = xrange( n )
rg35 = xrange( n + 5 )

# Fechas para que salgan en Las tablas
[ ia, hoy, antes5, despues5 ] = fechas( 0 )
```

Figura 5.16: Captura datos.

Fuente: Elaboración propia.

A continuación se presenta la Figura 5.16 donde se estiman los parámetros necesarios para la realización de las gráficas y tablas referentes al Método 1.

```

"""
MODELO 1
"""
# Estimamos la mu y la sigma
[ est_mu, est_sigma ] = estimacion_modelo1( share, D )

# Calculamos los errores cometidos
[ ecm, mape ] = errores( share, est_mu )

# Con las estimaciones, calculamos las medias y los IC estimados
est_medias = np.array( [ media( t, share[ 0 ], est_mu ) for t in rg35 ] )
# media + - dt
aux = np.array( [ mediaYdt( t, share[ 0 ], est_mu, est_sigma ) for t in rg35 ] )
# Intervalo de confianza del 2.5
est_ic_2_5 = aux[ :, 0 ]
# Intervalo de confianza del 97.5
est_ic_97_5 = aux[ :, 1 ]

# Gráfico de ajuste y estimación
grafico( share, est_medias, est_ic_2_5, est_ic_97_5, 'METODO 1', 'graficoM1.png' )

"""
Tabla 1
"""
# Escribimos el fichero con las predicciones
tabla1( NOM, 'datosactual.js', despues5, est_ic_97_5[ 30 : ], est_medias[ 30 : ], est_ic_2_5[ 30 : ], ia, ecm, mape )

"""
Datos históricos
"""
# Inicializamos la media y los IC históricos
media_hist = np.zeros( 10 )
ic_2_5_hist = np.zeros( 10 )
ic_97_5_hist = np.zeros( 10 )

for i in xrange( 10 ):

    # Seleccionamos los datos
    datos = shareTodo[ i : 30 + i ]
    # Estimamos el mu y el sigma para estos datos
    [ est_mu, est_sigma ] = estimacion_modelo1( datos, D )
    # Calculamos la media estimada para t = 30 + i + 1
    media_hist[ i ] = media( 30 + i + 1, datos[ 0 ], est_mu )
    # Calculamos el IC para t = 30 + i + 1
    aux = mediaYdt( 30 + i + 1, datos[ 0 ], est_mu, est_sigma )
    # Separamos y guardamos los valores del IC
    ic_2_5_hist[ i ] = aux[ 0 ]
    ic_97_5_hist[ i ] = aux[ 1 ]

"""
Tabla comparativa entre las predicciones de los últimos 5 días y los valores reales
"""
tabla_compara( NOM, 'compara.js', antes5, ic_2_5_hist[ 5 : ], media_hist[ 5 : ], share[ 25 : 30 ], ic_97_5_hist[ 5 : ], ia )

"""
Gráfico de estimaciones de los últimos 10 días
"""
grafico_comparativo( share[ 20 : 30 ], media_hist, ic_2_5_hist, ic_97_5_hist, 'METODO 1', 'graficoC1.png' )

```

Figura 5.17: Método 1 I.

Fuente: Elaboración propia.

Por otra parte se presenta la Figura 5.17 donde se estiman los parámetros necesarios para la realización de las gráficas y tablas referentes al Método 2.

```

MODELO 2
"""
# Estimamos la mu y la sigma
[ est_mu2, est_sigma2 ] = estimacion_modelo2( share, D )

# Calculamos los errores cometidos
[ ecm2, mape2 ] = errores( share, est_mu2 )

# Con las estimaciones, calculamos las medias y los IC estimados
est_medias2 = np.array( [ media( t, share[ 0 ], est_mu2 ) for t in rg35 ] )
# media + - dt
aux = np.array( [ mediaYdt( t, share[ 0 ], est_mu2, est_sigma2 ) for t in rg35 ] )
# Intervalo de confianza del 2.5
est_ic_2_5_bis = aux[ :, 0 ]
# Intervalo de confianza del 97.5
est_ic_97_5_bis = aux[ :, 1 ]

# Gráfico de ajuste y estimación
grafico( share, est_medias2, est_ic_2_5_bis, est_ic_97_5_bis, 'METODO 2', 'graficoM2.png' )

"""
Tabla 1
"""
# Escribimos el fichero con las predicciones
tabla1( NOM, 'datosactual1.js', despues5, est_ic_97_5_bis[ 30 : ], est_medias2[ 30 : ], est_ic_2_5_bis[ 30 : ], ia, ecm2, mape2 )

"""
Datos históricos
"""
# Inicializamos la media y los IC históricos
media_hist2 = np.zeros( 10 )
ic_2_5_hist = np.zeros( 10 )
ic_97_5_hist = np.zeros( 10 )

for i in xrange( 10 ):

    # Seleccionamos los datos
    datos = shareTodo[ i : 30 + i ]
    # Estimamos el mu y el sigma para estos datos
    [ est_mu, est_sigma ] = estimacion_modelo1( datos, D )
    # Calculamos la media estimada para t = 30 + i + 1
    media_hist2[ i ] = media( 30 + i + 1, datos[ 0 ], est_mu2 )
    # Calculamos el IC para t = 30 + i + 1
    aux = mediaYdt( 30 + i + 1, datos[ 0 ], est_mu2, est_sigma2 )
    # Separamos y guardamos los valores del IC
    ic_2_5_hist[ i ] = aux[ 0 ]
    ic_97_5_hist[ i ] = aux[ 1 ]

"""
Tabla comparativa entre las predicciones de los últimos 5 días y los valores reales
"""
tabla_compara( NOM, 'compara1.js', antes5, ic_2_5_hist[ 5 : ], media_hist2[ 5 : ], share[ 25 : 30 ], ic_97_5_hist[ 5 : ], ia )

"""
Gráfico de estimaciones de los últimos 10 días
"""
grafico_comparativo( share[ 20 : 30 ], media_hist2, ic_2_5_hist, ic_97_5_hist, 'METODO 2', 'graficoC2.png' )

#####
"""
Tabla de la comparativa
"""
tabla_combinada( NOM, 'compara2.js', antes5, media_hist[ 5 : ], media_hist2[ 5 : ], share[ 25 : 30 ], ia, ecm, ecm2, mape, mape2 )

# San se acabó
exit()

```

Figura 5.18: Método 2 I.

Fuente: Elaboración propia.

6 DISEÑO Y APLICACIÓN DE LA PÁGINA WEB

Para facilitar la presentación de los resultados de la aplicación del Modelo Log-Normal se ha diseñado una página web donde se exponen las predicciones obtenidas de activo subyacente Bankia (BKIA.MC). De esta manera, se presentarán las estimaciones de cada modelo y una comparativa de ambos respecto al valor real del activo. Se recuerda que en la página web, la notación “Método 1” corresponde a la estimación de los parámetros por el Método de Máxima Verosimilitud y “Método 2” corresponde a la estimación de los parámetros por el Método No Paramétrico

6.1 Requisitos previos.

Para la realización del presente trabajo se han dos ordenadores. El primero de ellos, con sistema operativo Windows 10 sirve de campo de pruebas de la aplicación del modelo. Para poder editar y adaptar el código a las necesidades del activo subyacente se han utilizado diversos programas.

- El lenguaje de programación *Python* en un entorno en el que se agregan librerías científicas necesarias como *Scipy* (<http://www.scipy.org>) y *Numpy* (<http://www.numpy.org>) . Por ello, y para evitar incompatibilidades, se ha descargado *Anaconda* (<https://www.continuum.io/why-anaconda>) que se caracteriza porque es un software libre, instala varias versiones e instala los paquetes más importantes para tratamiento de grandes volúmenes de datos, cálculos científicos y análisis de predicciones.
- Para la revisión y búsqueda de errores en los ficheros *html* se ha utilizado *Notepad++* (<https://notepad-plus-plus.org/>).

El segundo ordenador, con sistema operativo Ubuntu server (<http://www.ubuntu.com>) 14.04 LTS (Long Time Support) y con la instalación del servidor web Apache (<http://httpd.apache.org>) se utiliza para ser el servidor que ejecute diariamente el script que permita tener actualizada la página web de dirección <http://cotizacion.imm.upv.es/>.

Una vez configurados los diferentes programas en *Python* a utilizar, se alojan en el segundo ordenador(a partir de ahora denominado como servidor) y se lanza la web.

6.3 Diseño de la página web.

La página web está diseñada sobre una plantilla de *HTML5 UP* modificada mediante *Notepad++*. La web se compone de 5 páginas; *Inicio*, *Método 1*, *Método 2*, *Comparativa* y *Conclusiones*. A continuación se muestran las páginas.

El lenguaje de programación escogido para la realización de la página es *HTML5* junto con *CSS*.

Escribir páginas usando HTML 5 nos brinda varios beneficios, por ejemplo:

- Al ser un código más sencillo y simplificado, las páginas cargan más rápido en el navegador.
- Las páginas y los elementos que contienen, se ven perfectamente en todos los navegadores. La gran mayoría de los navegadores de los teléfonos Smartphone y las tabletas, son compatibles con HTML5
- Los plugins, widgets y botones que ofrecen los desarrolladores de las redes como Facebook, Twitter y otras, escritos en HTML 5 funcionan excelentemente, con más opciones que los clásicos en XHTML o que los iframes.
- Es posible insertar directamente videos en las páginas sin tener que acudir a los iframes o usar la etiqueta *object*.
- HTML 5 incluye etiquetas orientadas principalmente a los buscadores, para facilitarles comprender el contenido de las páginas.

A continuación se presenta una vista en fase *alpha* de la web en su totalidad a *01/09/2016*.. La página *Inicio* (ver Figura 6.15 y 6.16) sirve de punta de partida de la web y sirve de introducción.

En términos generales se presentan los Agradecimientos, el motivo del trabajo y de la web. Así como una descripción del Instituto de Matemática Multidisciplinar de la UPV.

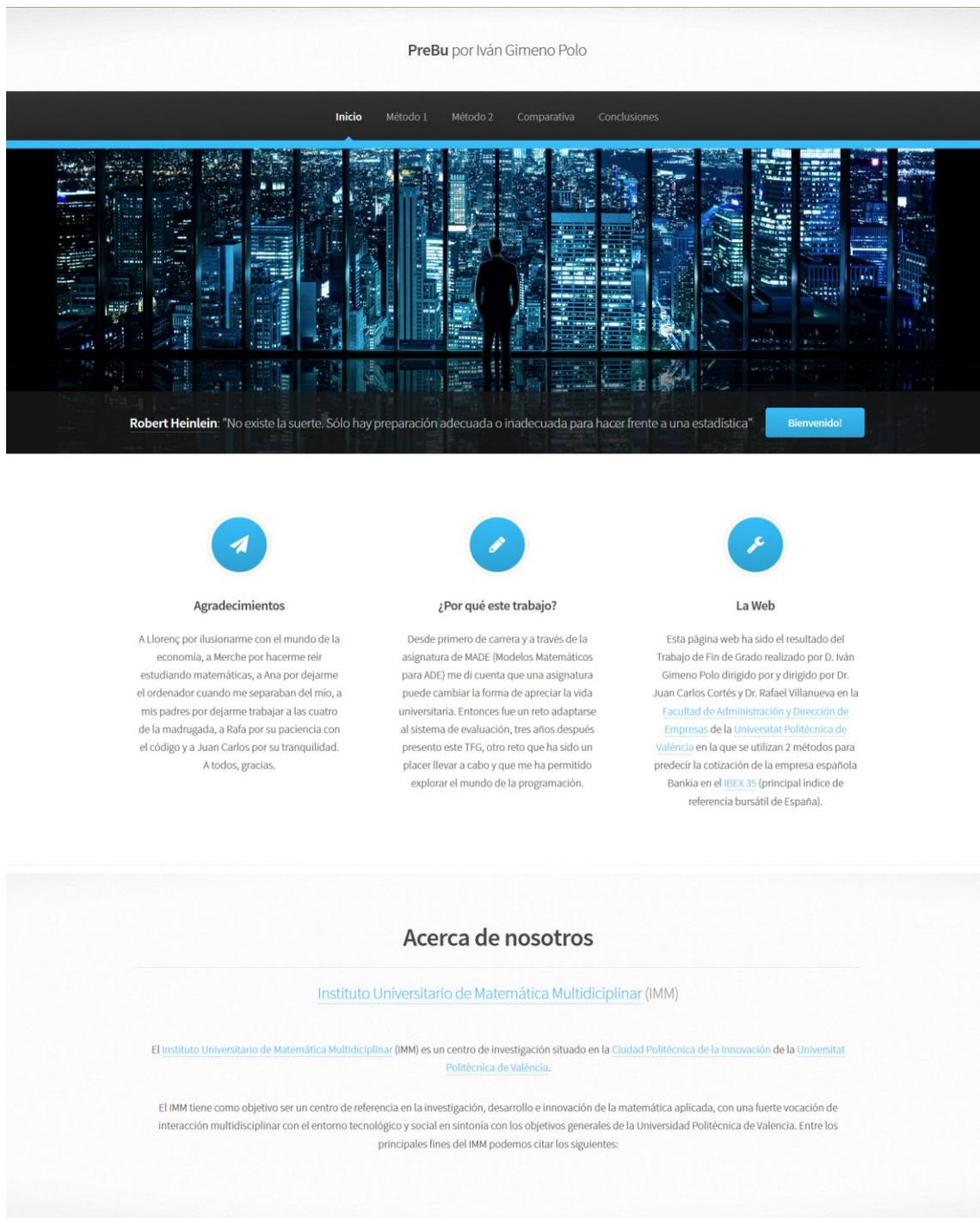


Figura 6.1: Inicio a.

Fuente: Elaboración propia.



En primer lugar

Promover la participación coordinada de los grupos en proyectos de I+D+i a nivel nacional e internacional.



En segundo lugar

Apoyar la participación de los grupos en proyectos de alta complejidad mediante un soporte de gestión profesionalizada.



En tercer lugar

Ofrecer oportunidades de formación de postgrado a través de cursos, trabajos de final de grado, trabajos finales de máster y tesis doctorales, que tengan un núcleo de desarrollo dentro del Instituto.



En cuarto lugar

Establecer vínculos de colaboración con otros centros de investigación y desarrollo, con las empresas e instituciones y promover el debate interno y externo sobre temas de innovación docente e investigadora de carácter matemático.

¿Dispuesto a seguir? Pasemos a las predicciones

PreBu

© Untitled. All rights reserved | Design: HTML5 UP adaptado por Iván Gimeno

Figura 6.2: Inicio b.

Fuente: Elaboración propia.

Método 1

Este método se basa en que la dinámica del subyacente $S(t)$ sigue el **modelo log-normal** basado en la siguiente ecuación diferencial estocástica de tipo Itô:

$$dS(t) = \mu S(t) dt + \sigma S(t) dB(t) \quad (1)$$

donde $B(t)$ denota el Movimiento Browniano o también llamado Proceso de Wiener.

Para ajustar el modelo a datos reales, se requiere estimar los parámetros μ (denominada tendencia o drift) y σ (volatilidad) del subyacente.

Para ello, en primer lugar se utiliza el esquema de Euler-Maruyama para discretizar la ecuación (1) y a continuación se utiliza el método de Máxima Verosimilitud para estimar ambos parámetros.

Resultados!

Resultados

Estimación a través del Método de Máxima Verosimilitud.



En la primera tabla que se muestra a continuación, se proporcionan los valores numéricos tanto de las **predicciones puntuales** (PREDICCIÓN) como por **intervalos de confianza al 95%** (IC 95%) para los cuales se especifican el extremo superior e inferior.

Asimismo, en la segunda tabla se muestran el **MAPE** (Mean Absolute Percentage Error) y el **ECM** (Error Cuadrático Medio), que son medidas de bondad de ajuste del método.

BANKIA	06/09/1607/09/1608/09/1609/09/1612/09/16				
IC 95% superior	0.947	0.954	0.961	0.969	0.976
Prediccion	0.794	0.798	0.802	0.806	0.811
IC 95% inferior	0.641	0.642	0.643	0.644	0.645

MAPE	ECM
3.339 %	0.028

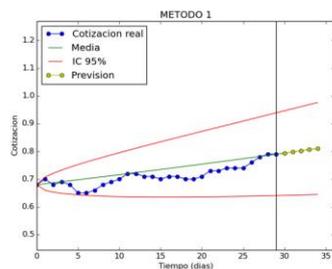
Last update: 01:01, 06/09/16

Figura 6.3: Método 1a.

Fuente: Elaboración propia.

En la primera parte de la página *Método 1* se presenta una breve explicación del Método 1 así como la presentación de resultados del Método de Máxima Verosimilitud. Se aprecia la tabla que contiene la predicción, el MAPE y el ECM.

En el siguiente gráfico, los puntos azules son la cotización real, la línea verde es la media y las líneas rojas determinan el intervalo de confianza del 95% que proporciona el método. Los puntos amarillos a la derecha de la línea vertical corresponden a la media de la previsión para los próximos 5 días.



En la siguiente tabla se comparan las predicciones realizadas para los últimos cinco días con las cotizaciones reales que tuvo el valor.

BANKIA	30/08	1631/08	1601/09	1602/09	1605/09	16
IC 95% superior	0.593	0.595	0.610	0.627	0.644	
Predicción	0.762	0.759	0.788	0.811	0.832	
Real	0.740	0.760	0.780	0.790	0.790	
IC 95% inferior	0.931	0.924	0.966	0.995	1.020	

Last update: 01:01, 06/09/16

Método 2

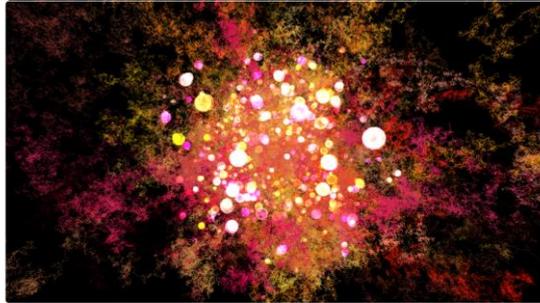
Figura 6.4: Método 1b.

Fuente: Elaboración propia.

En la segunda parte de la página se presenta un gráfico con la estimación y la previsión así como una tabla con la comparación entre realidad y estimación.

Resultados

Estimación a través de un Método No Paramétrico.



En la primera tabla que se muestra a continuación, se proporcionan los valores numéricos tanto de las **predicciones** puntuales (PREDICCIÓN) como por **intervalos de confianza al 95%** (IC: 95%) para los cuales se especifican el extremo superior e inferior.

Asimismo, en la segunda tabla se muestran el **MAPE** (Mean Absolute Percentage Error) y el **ECM** (Error Cuadrático Medio), que son medidas de bondad de ajuste del método.

BANKIA	06/09/16	07/09/16	08/09/16	09/09/16	10/09/16
IC 95% superior	0.917	0.924	0.930	0.936	0.942
Predicción	0.772	0.775	0.778	0.781	0.785
IC 95% inferior	0.626	0.626	0.627	0.627	0.627

MAPE	ECM
2.583 %	0.022

Last update: 01:01, 06/09/16

Método 2

En base al modelo **log-normal**, la estimación de los parámetros μ y σ de la ecuación

$$dS(t) = \mu S(t) dt + \sigma S(t) dB(t) \quad (1)$$

Se obtienen mediante una estimación no paramétrica basada en la magnitud de crecimiento de los momentos estadísticos de primer y segundo orden de los incrementos instantáneos de las cotizaciones respecto de las funciones $\mu S(t)$ y $\sigma S(t)$ que definen los coeficientes de la ecuación diferencial estocástica (1).

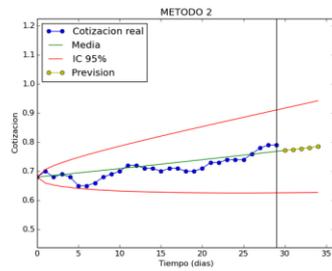
Resultados

Figura 6.5: Método 2a .

Fuente: Elaboración propia.

En la primera parte de la página *Método 2* se presenta una breve explicación del Método 1 así como la presentación de resultados del Método de Máxima Verosimilitud.

En este gráfico, los puntos azules son la cotización real, la línea verde es la media y las líneas rojas determinan el intervalo de confianza del 95% que proporciona el método. Los puntos amarillos a la derecha de la línea vertical corresponden a la media de la previsión para los próximos 5 días.



En la siguiente tabla se comparan las predicciones realizadas para los últimos cinco días con las cotizaciones reales que tuvo el valor.

BANKIA	30/08/16	31/08/16	01/09/16	02/09/16	03/09/16
IC 95% superior	0.609	0.628	0.629	0.647	0.639
Prediccion	0.768	0.795	0.798	0.825	0.817
Real	0.740	0.760	0.780	0.790	0.790
IC 95% inferior	0.927	0.961	0.968	1.003	0.995

Last update: 01:01, 06/09/16

Comparativa

Figura 6.6: Método 2b.

Fuente: Elaboración propia.

En la segunda parte de la página se presenta un gráfico con la estimación y la previsión así como una tabla con la comparación entre realidad y estimación.

Comparativa

Es hora de comparar la robustez de los dos métodos frente a las predicciones reales



Vamos a comparar los resultados obtenidos en ambos métodos.

Tablas comparativas

Como medidas de bondad de ajuste de las predicciones, utilizamos el MAPE (Mean Absolute Percentage Error) y el ECM (Error Cuadrático Medio). El MAPE y el ECM de cada método se muestran en la segunda tabla.

BANKIA	30/08/16	31/08/16	01/09/16	02/09/16	03/09/16
METODO 1	0.762	0.759	0.788	0.811	0.832
METODO 2	0.768	0.795	0.798	0.825	0.817
Real	0.740	0.760	0.780	0.790	0.790

	MAPE	ECM
METODO 1	3.339 %	0.028
METODO 2	2.583 %	0.022

Last update: 01:01, 06/09/16

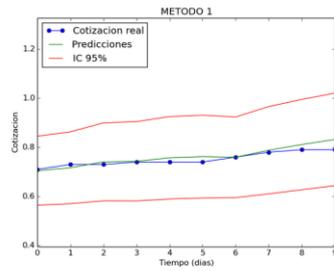
Figura 6.7: Comparativa 1.

Fuente: Elaboración propia.

En la página *Comparativa* se muestran dos tablas comparativas de los respectivos métodos con la realidad, así como el MAPE y el ECM de cada método.

En la segunda parte de la página se observan las comparativas respecto a la realidad de las predicciones de los dos métodos.

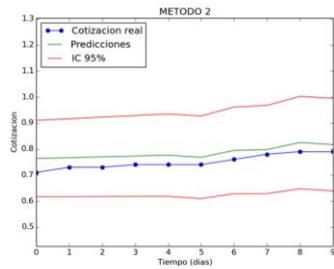
La línea azul con puntos muestra las cotización real, la línea verde y las líneas rojas indican la media e intervalo de confianza del 95%, respectivamente, que se calculó para estimar cada cotización los días previos.



A continuación mostramos en una gráfica cómo se han ido estimando las predicciones los últimos 10 días con el MÉTODO 2.

La línea azul con puntos muestra las cotización real, la línea verde y las líneas rojas indican la media e intervalo de confianza del 95%, respectivamente, que se calculó para estimar cada cotización los días previos.

Habremos tenido una buena estimación si las líneas azul y verde permanecen cercanas y siempre dentro de las líneas rojas.



Conclusiones

Figura 6.8: Comparativa 2.

Fuente: Elaboración propia.

Conclusiones

Qué hemos aprendido a lo largo del proyecto?



El **objetivo** del presente Trabajo de Fin de Grado ha sido el estudio, desarrollo y aplicación del modelo estocástico Log-Normal para realizar predicciones de activos subyacentes. El desarrollo de la página web y la aplicación del código *Python* ha sido una ardua tarea que me ha permitido explorar nuevas fronteras e iniciarme firmemente en el mundo de la programación.

La principal aplicación del presente trabajo radica en reducir la situación de incertidumbre que causa la evolución diaria del IBEX-35 y facilitar la toma de decisiones de un inversor.

La página web se proporciona como herramienta para facilitar la **comprensión** de los resultados generados por el programa. La versión *alpha* de esta página presenta un sólo activo Bankia (BKIA.MC), con la posibilidad, en un futuro, de añadir más predicciones de diversos activos.

Como contenido adicional, se ha utilizado **Python** para realizar las estimaciones necesarias y plasmarlas gráficamente para facilitar su comprensión visual. Por otra parte, se ha diseñado una página web para poder presentar las predicciones de manera clara y visual al usuario que desee consultar el trabajo llevado a cabo.

Para la realización del trabajo ha sido necesario el conocimiento adquirido en las diferentes asignaturas del Grado de Administración y Dirección de Empresas además de conocimientos adicionales de programación y diseño web.

Para concluir el presente trabajo...

Kiyoshi Itô: "Es construyendo estructuras matemáticas como los matemáticos encuentran el mismo tipo de belleza que otros encuentran en la música o en la arquitectura. Pero hay una gran diferencia: la música de Mozart puede ser disfrutada incluso sin conocer la teoría musical. Sin embargo, la belleza de las estructuras matemáticas no se puede apreciar sin entender las fórmulas: sólo los matemáticos pueden leer las partituras matemáticas y tocar esa música en sus corazones. Yo creía que sin las fórmulas nunca podría transmitir la dulce melodía que tocaba mi corazón. Las ecuaciones diferenciales estocásticas, llamadas la fórmula de Itô, se usan para describir las fluctuaciones aleatorias. Al principio, mi artículo no atrajo mucha atención, y fue 10 años después cuando otros matemáticos comenzaron a leer mis partituras y a tocar mi música".

Gracias, Iván.

Figura 6.9: Conclusiones.

Fuente: Elaboración propia.

Para la actualización de la página web se utiliza el segundo equipo, alojado en el *imm*². Para ello se utiliza una aplicación alojada en el sistema operativo Linux, Se denomina *Cron*, y es un administrador regular de procesos en segundo plano que ejecuta procesos o guiones a intervalos regulares (por ejemplo, cada minuto, día, semana, etc.). Los procesos que deben ejecutarse y la hora en la que deben hacerlo se especifican en el fichero *crontab*. *Cron*, en un entorno Windows se podría definir como el "equivalente" a Tareas Programadas.

Para el programa se configura el *crontab* para que ejecute de Lunes a Viernes a la 1 a.m. el archivo IVAN_EXE de *Python* que leen los datos de Bankia, los procesan y generan los resultados. Se ha que el programa *cron* se active a la 1 a.m. debido a que el histórico en la web <http://yahoo.finance.com> en algunas ocasiones no se actualiza hasta después de las 12 a.m. De este modo se soluciona ese problema y se evita trabajar con un histórico incorrecto.

7. CONCLUSIONES

En el Trabajo de Fin de Grado presentado, se ha estudiado, aplicado y validado el modelo estocástico Log-Normal al activo subyacente Bankia (BKIA.MC). Como resultado final, ha sido posible predecir a corto plazo la evolución del activo a partir de un histórico de cotizaciones.

Como contenido adicional, se ha utilizado *Python* para realizar las estimaciones necesarias y plasmarlas gráficamente para facilitar su comprensión visual.

Por otra parte, se ha diseñado una página web para poder presentar las predicciones de manera clara y visual al usuario que desee consultar el trabajo llevado a cabo.

Para la realización del trabajo ha sido necesario el conocimiento adquirido en las diferentes asignaturas del Grado de Administración y Dirección de Empresas además de conocimientos adicionales de programación y diseño web.

A lo largo del Capítulo 1 se ha realizado una visión breve del pasado, presente y futuro del panorama financiero actual con tal de contextualizar la realización del presente trabajo.

En el Capítulo 2 de esta memoria se ha descrito la historia y la evolución de Bankia S.A. La elección viene motivada por el gran impacto que causó su fundación sobre el sistema financiero español.

En el Capítulo 3 y en el Capítulo 4 se ha realizado el estudio, desarrollo y proceso de validación del modelo estocástico Log-Normal desde un punto de vista teórico. Se incluye una introducción al Cálculo Estocástico de Itô que es imprescindible para la

resolución teórica del modelo y el cálculo de sus principales funciones estadísticas, las cuales son imprescindibles para la puesta en práctica del modelo sobre el activo subyacente Bankia (BKIA.MC).

En el Capítulo 5 se pone en práctica el Modelo Log-Normal con objeto de modelizar el activo subyacente y realizar predicciones. Para la aplicación del Modelo Log-Normal ha sido necesario la aplicación de un código en el lenguaje de programación *Python* con el que se han realizado las estimaciones necesarias. Se ha seguido el siguiente diagrama de actuación.



Figura 7.1: Diagrama.

Fuente: Elaboración propia.

Por último, en el Capítulo 6 se presenta en forma de captura de pantalla el diseño de la página web dónde se presentan los resultados.

La principal aplicación del presente trabajo radica en reducir la situación de incertidumbre que causa la evolución diaria del IBEX-35 y facilitar la toma de decisiones de un inversor.

En este sentido, en futuros estudios que se quisieran realizar y/o incluso para ampliar y perfeccionar el trabajo de la presente memoria, se podrían aplicar el Modelo Log-Normal por medio de simulaciones de Monte Carlo. También podría utilizarse un histórico de cotizaciones mayor e incluso realizar predicciones para intervalos temporales más largos. Sería interesante estudiar la viabilidad de otros modelos estocásticos más complejos disponibles en la literatura especializada con el objetivo de contrastar los resultados obtenidos con el Modelo Log-Normal. Ejemplos interesantes en este sentido son los basados en ecuaciones diferenciales estocásticas cuyo proceso

conductor de la aleatoriedad es una mezcla de un proceso continuo como el movimiento Browniano, y un proceso discontinuo o de salto tipo *Poisson*¹⁹.

Por otra parte, es posible actualizar la página web y añadir la predicción de más activos.

¹⁹BENTH, F.E.; KHOLODNYI, V.; LAURENCE, P. (2014). *“Quantitative Energy Finance: Modeling, pricing and hedging in energy and commodity markets”*. Springer Science+Business.

BIBLIOGRAFÍA

Libros:

ALLEN, E. (2007) *Modelling With Itô Stochastic Differential Equations*. Springer, Series Mathematical Modelling: Theory and Applications.

BAXTER, M.; REMIE, A. (2012) *Financial Calculus: An Introduction to Derivate Pricing*. Cambridge Univ. Press, 20th Edition.

BENTH, F.E.; KHOLODNYI, V.; LAURENCE, P. (Ed. 2014) *Quantitative Energy Finance: Modeling, Pricing and Hedging in Energy and Commodity Markets*. Springer, Science+Business. Media B.V.

GONZÁLEZ DUQUE, R. (2014) *Python para todos*.

HULL, J.C. (2002) *Introducción a los Mercados de Futuros y Opciones*. Pearson Education, 4^º Edición.

ITÔ, K. (1944) *Stochastic Integral*. Tokio. Proc. Imperial Acad.

ITÔ, K. (1961) *Lectures on Stochastic Processes*. Bombay. Tata Institute.

KLEBANER, F. C. (2011) *Introduction to Stochastic Calculus With Applications to Finance*. Imperial College Press.

LAMBERTON, D.; LAPEYRE, B. (1996) *Introduction to Stochastic Calculus to Finance*. Chapman & Hall/CRC.

YANAJARA, H. (2012) *Manual Sage para principiantes*.

ØKSENDAHL, B. (1980) *Stochastic Differential Equations*. Nueva York. Springer.

ØKSENDAHL, B. (1998) *Stochastic Differential Equations: An Introduction with Applications*. Springer.

Artículos:

BACHELIER, L. (1900) Théorie de la spéculation. *Annales du Science de L'Ecole Normale Supérieure*, 17, pp. 21-86.

ITÔ, K. (1951) On stochastic differential equations. *Memoirs, American Mathematical Society*, fascículo 4, pp. 1–51.

MERTON, R.C. (1973) An inter temporal capital asset pricing model. *Econometrica*, fascículo 41 pp. 867-887.

MERTON, R.C. (1973) Theory of rational option pricing. *Bell Journal of Economics and Management Science (The RAND Corporation)*, volumen 4, fascículo 1 pp. 141–183.

NELDER, J.A and MEAD R. (1965)“A simplex method for function minimization”, *Computer Journal* 7, pp 308-313.

Págnas web:

El Economista (Última Fecha de Acceso: 17/08/2016).

http://economia.elpais.com/economia/2016/06/01/actualidad/1464778881_012072.html

El Confidencial (Última Fecha de Acceso: 15/08/2016).

http://www.elconfidencial.com/mercados/2015-11-21/en-que-invertir-en-2016-con-este-panorama_1102782/

Las Provincias (Última Fecha de Acceso: 17/08/2016).

<http://www.lasprovincias.es/20120605/economia/bankia-banco-201206051032.html>

Expansión (Última Fecha de Acceso: 17/08/2016).

<http://www.expansion.com/diccionario-economico/sistema-institucional-de-proteccion-sip.html>

Bankia (Última Fecha de Acceso: 17/08/2016)

<http://www.bankia.com/es/quienes-somos/sobre-nosotros/origen/>

Bankia S.A. (Última Fecha de Acceso: 01/09/2016).

<http://www.bankia.com/recursos/img/memoriaanual/20150408/clientes/distribucion-y-medios-especificos-mini-580x337.png>

Bankia S.A. (Última Fecha de Acceso: 01/09/2016).

<http://www.bankia.com/recursos/img/memoriaanual/20150408/clientes/planta-y-distribucion-de-una-oficina-mini-500x334.png>

Vozpopuli (Última Fecha de Acceso: 01/09/2016).

<http://vozpopuli.com/economia-y-finanzas/53875-asi-maniobro-rato-para-que-bankia-pudiera-salir-a-bolsa>

ABC (Última Fecha de Acceso: 01/09/2016).

<http://www.abc.es/asesoria-legal-arriaga/20150213/abci-reclamarconexito-acciones-bankia-201502131552.html>

El País (Última Fecha de Acceso: 01/09/2016).

http://economia.elpais.com/economia/2015/11/12/actualidad/1447329193_327238.html

Quimicaweb (Última Fecha de Acceso: 01/09/2016).

http://www.quimicaweb.net/albert_einstein/einstein/movimiento_browniano/movimiento_browniano.htm

Instituto Latinoamericano de la Comunicación Educativa (Última Fecha de Acceso: 01/09/2016).

http://bibliotecadigital.ilce.edu.mx/sites/ciencia/volumen1/ciencia2/13/htm/sec_5.html

Bankia S.A. (Última Fecha de Acceso: 01/09/2016).

<http://www.bankia.com/recursos/doc/corporativo/20120927/anual/cuentas-anuales-consolidadas-informe-de-gestion-e-informe-del-auditor-del-ejerci.pdf>

Bankia S.A. (Última Fecha de Acceso: 01/09/2016).

<http://www.bankia.com/recursos/doc/corporativo/20120927/anual/informe-anual-cuentas-anuales-consolidadas-2013.pdf>

Python Software Foundation (Última Fecha de Acceso: 01/09/2016).

<https://www.python.org>

SciPy developers (Última Fecha de Acceso: 01/09/2016).

<http://www.scipy.org>

Canonical Ltd (Última Fecha de Acceso: 01/09/2016).

<http://www.ubuntu.com>

Yahoo! (Última Fecha de Acceso: 02/09/2016).

<http://www.yahoo.finance.com>

