



Universidad de Valladolid

Facultad de Ciencias Económicas y Empresariales

Trabajo de Fin de Grado

Grado en Finanzas, Banca y Seguros

**Características de la distribución de los
rendimientos bursátiles: Un análisis con
datos diarios y datos intradiarios**

Presentado por:

Ignacio Fernández Velázquez

Tutelado por:

Ana Pérez Espartero

Valladolid, 14 de Julio de 2017

RESUMEN

En este trabajo examinamos la distribución marginal de los rendimientos bursátiles para 2 índices europeos, 2 índices americanos y 2 índices asiáticos. Realizamos un análisis con datos diarios e intradiarios y consideramos diversas distribuciones como son la distribución Normal, la Hiperbólica Generalizada, la Logística y la mixtura Gaussiana. Estimamos los parámetros de estas distribuciones por el método de máxima verosimilitud. Por último, utilizamos distintos contrastes de bondad de ajuste para analizar si estas distribuciones son una aproximación apropiada para la distribución empírica de los rendimientos bursátiles.

ABSTRACT

In this paper we examine the marginal distribution of stock index returns for 2 European indexes, 2 American indexes and 2 Asian indexes. We carry out an analysis with daily data and intraday data and we consider many probability distributions, such as the Normal distribution, the Hyperbolic Generalized, the Logistic and the Gaussian mixture. We estimate the parameters of these distributions for the maximum likelihood method. In the end, we use different goodness of fit test to analyze if these distributions are an appropriate approximation for the empirical distribution of stock returns.

Palabras clave: Mercado bursátil, distribución de probabilidad, rendimientos bursátiles, R.

Clasificación JEL: C87, C88, G11.

ÍNDICE

1. Introducción.....	1
2. Descripción de los datos.....	2
2.1. Índices bursátiles utilizados: definición y representación gráfica.....	2
2.2. Análisis descriptivo de los rendimientos diarios.....	5
3. Contrastes de Normalidad sobre la distribución de los rendimientos diarios.....	9
3.1. Análisis gráfico: histogramas y Q-Q plots.....	9
3.2. Contrastes de Normalidad: definición y propiedades.....	12
3.3. Análisis de resultados.....	15
4. Distribuciones alternativas para los rendimientos diarios: Hiperbólica Generalizada, Logística y Mixtura Gaussiana.....	16
4.1. Descripción de las distribuciones utilizadas: definición y propiedades.....	16
4.1.1. Distribución hiperbólica generalizada: casos particulares.....	16
4.1.2. Distribución logística.....	20
4.1.3. Mixtura Gaussiana.....	21
4.2. Resultados de la estimación.....	23
4.3. Contrastes de bondad de ajuste.....	27
5. Sensibilidad de los resultados a la frecuencia de observación: análisis de datos intradiarios.....	31
5.1. Descripción de los datos intradiarios: definición y propiedades.....	31
5.2. Estimación de modelos y contrastes de bondad de ajuste.....	34
6. Conclusiones.....	40
7. Bibliografía.....	41
8. Anexo	44

1. INTRODUCCIÓN

En este trabajo analizaremos la distribución de los rendimientos bursátiles de los principales índices bursátiles europeos, americanos y asiáticos.

Importantes modelos de valoración de activos financieros utilizados en economía financiera se basan en el supuesto de normalidad de los rendimientos bursátiles, como son el modelo de selección de carteras de Markowitz, el CAPM y el modelo de Black-Scholes. Por tanto, el supuesto de no normalidad de los rendimientos supondría que las distribuciones no estarían definidas completamente mediante la media y la desviación típica, y tendríamos que estudiar otras características como son la asimetría y la curtosis. Por lo que los anteriores modelos de economía financiera no estarían completos.

A lo largo de los años se han realizado numerosos estudios sobre la distribución de los rendimientos bursátiles. Así, Bachelier (1900) propuso para éstos un proceso gaussiano. Hasta los años 60 este resultado era aceptado.

Mandelbrot (1963) en un estudio sobre el mercado de algodón, afirmaba que la distribución de rendimientos era leptocútica, por lo que no era gaussiana. Para modelizar esta distribución propuso la familia de distribuciones Pareto-estables, con las que encontró un mejor ajuste a la distribución marginal de los rendimientos que la distribución gaussiana. El problema de esta familia de distribuciones es que posee varianza infinita, por lo que muchas herramientas estadísticas no se pueden utilizar; además, no poseen una expresión para su función de densidad, provocando la imposibilidad del uso de métodos de máxima verosimilitud.

Fama (1965) en su estudio realizado sobre los rendimientos de los 30 valores que componen el índice Dow-Jones, llega a unas conclusiones similares a las de Mandelbrot: la distribución marginal de los rendimientos es picuda y la familia de distribuciones Pareto-estables se ajusta mejor que la normal.

Posteriormente, tanto Praetz (1972) y Blattberg y Gonedes (1974) utilizando la distribución t de Student, como Clark (1973) usando la distribución lognormal-normal, obtuvieron mejores resultados que las distribuciones Pareto-

estables. Además estas dos distribuciones presentan la ventaja, frente a la Pareto-estable, de que está definida su función de densidad.

Además de éstas, se han utilizado otras distribuciones para modelizar el rendimiento de las acciones de distintos mercados bursátiles. Así, Kon (1984) utilizó la mixtura gaussiana, Smith (1981), Gray y French (1990) y Peiró (1992) han utilizado la distribución logística y otros autores, como Eberlein y Keller (1995), Prause (1997, 2000) o Necula (2009) han abordado este problema utilizando la distribución hiperbólica generalizada con buenos resultados.

Algunos artículos, como Aparicio y Estrada (2001), han señalado que a medida que las frecuencias observadas son mayores, la distribución se aproxima a una normal, mientras que con frecuencias temporales pequeñas dista mucho de seguir una distribución gaussiana, ya que se obtienen distribuciones picudas y con un cierto grado de asimetría.

Por todas estas razones, estudiaremos la posible no normalidad de los rendimientos bursátiles, con frecuencias diarias e intradiarias, así como analizaremos distribuciones probabilísticas que se ajusten mejor a dichos rendimientos, como son la mixtura gaussiana, la logística y la hiperbólica generalizada. Para éste análisis utilizaremos tanto el software Matlab como R.

2. DESCRIPCIÓN DE LOS DATOS

2.1. Índices bursátiles utilizados: definición y representación gráfica

En este trabajo vamos a analizar los rendimientos, tanto diarios como intradiarios, de diversos índices bursátiles, como son:

- CAC 40: índice francés formado por las 40 empresas más significativas de las 100 empresas más grandes de la bolsa de París.
- FTSE 100: índice formado por las 100 empresas de mayor capitalización de la bolsa de Londres, siendo el índice más importante de Reino Unido.
- NASDAQ 100: índice bursátil estadounidense en el que se recogen las 100 principales empresas del sector industrial no financieras que están inscritas en la bolsa de Nueva York (NYSE) y ofrecidas en el Nasdaq Stock Market.
- NIKKEI 225: índice japonés formado por las 225 empresas que gozan de mayor liquidez en la Bolsa de Tokio.

- S&P 500: índice estadounidense considerado como el índice más representativo del mercado bursátil de E.E.U.U., formado por las 500 mayores empresas de los mercados NYSE y el NASDAQ (NASDAQ Global Select Market, NASDAQ Select Market, o el NASDAQ Capital Market).
- SSE Composite: índice chino que recoge todos los valores que se negocian en la Bolsa de Shanghái.

Los cotizaciones diarias consideradas (obtenidas de www.investing.com) corresponden al valor de cierre de las correspondientes bolsas durante el periodo de 31/12/2006 hasta el 31/12/2016, por lo que se dispone de una muestra de 2.561 valores para el índice CAC 40, 2.526 para el FTSE 100, 2.518 para el NASDAQ 100, 2.446 para el NIKKEI 225, 2.516 para el S&P 500 y 2.433 para el SSE Composite.

A partir de las cotizaciones diarias a cierre de cada índice se ha obtenido la serie de rendimientos diarios mediante diferencia logarítmica, es decir:

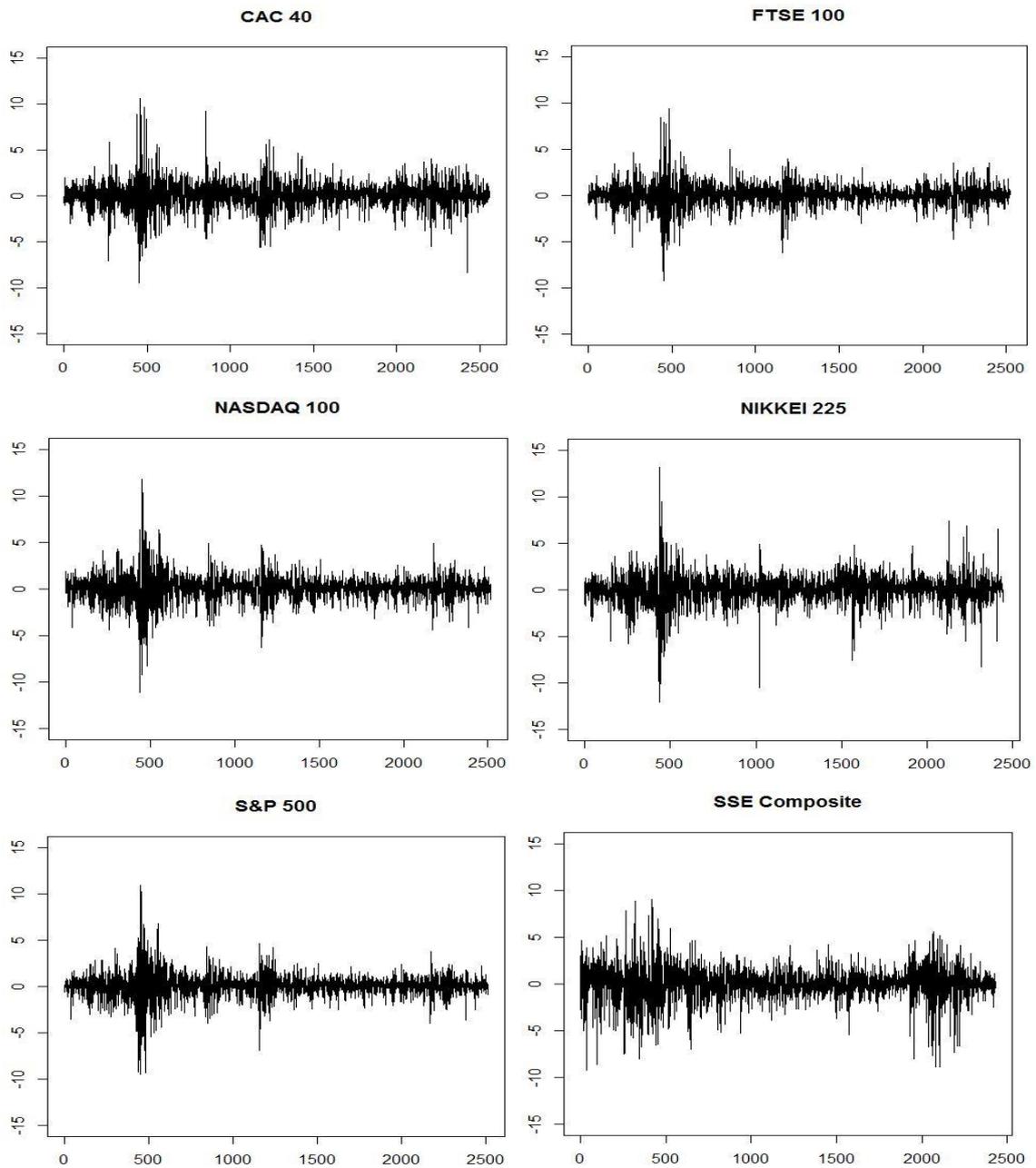
$$X_t = 100 \log(P_t/P_{t-1}) = 100 \cdot (\log P_t - \log P_{t-1}),$$

donde X_t denota el rendimiento para cada día t , P_t es el valor de cotización de cierre del índice para el día t y P_{t-1} la cotización de cierre en el día $t-1$.

Para el cálculo de rendimientos utilizamos la diferencia logarítmica y no la rentabilidad simple porque es la más utilizada en la literatura y presenta varias ventajas; por ejemplo, el rendimiento de un mes es la suma de los rendimientos de los correspondientes días y el rendimiento diario es la suma de los correspondientes rendimientos intradiarios; véase Törnqvist *et al.* (1985) y Peiró (1997). Además se puede aproximar la fórmula de la rentabilidad simple a la diferencia logarítmica basándose en una expansión de Taylor de segundo orden de la función $\log(P_t/P_{t-1})$ alrededor de 1.

A continuación, analizaremos gráficamente la serie de rendimientos históricos de estos índices. Para ello, el siguiente gráfico incluye una representación gráfica de los valores históricos de nuestras 6 series para el periodo 31/12/2006 – 31/12/2016.

Gráfico 2.1: Rendimientos diarios de los índices bursátiles.



Como podemos ver en el Gráfico 2.1, todos los índices presentan representaciones similares, ello es debido a la fuerte interrelación entre los mercados bursátiles, provocado por el efecto contagio. Hay periodos en los que todos los índices presentan grandes oscilaciones, mientras que en otros periodos los índices están en “calma”.

A continuación detallaremos aquellos momentos en los que las rendimientos bursátiles presentaron mayor volatilidad.

En torno a $t \approx 500$ se observa un periodo de gran volatilidad, donde los rendimientos son más extremos, que corresponde al mes de octubre de 2008, momento en el que se inició la crisis mundial.

Otro hito importante se da en torno a $t \approx 850$, que corresponde al 6/5/2010, conocido como el día del “Flash Crash”, especialmente notable en los mercados europeos y estadounidenses.

El siguiente periodo de grandes oscilaciones, que son más acusadas en los mercados europeos y americanos, se produce en torno a $t \approx 1.200$. En concreto, el 4/8/2011 tuvo lugar un importante descenso del valor de los índices bursátiles debido a la deuda europea y al incremento de la deuda pública estadounidense.

En torno a $t \approx 1.600$ (23/5/2013) tanto en el Nikkei 225 como en el SSE, descendieron las cotizaciones, tras una mala noticia de la actividad manufacturera china.

En torno a $t \approx 2.100$ para el caso del SSE Composite y el Nikkei 225, y en torno a $t \approx 2.200$ para el resto de los índices, se produjeron grandes fluctuaciones, correspondientes a los días 24 y 25 de agosto de 2015, debido a la desaceleración de la economía china.

Para finalizar este análisis destacaremos el día 24/6/2016, lo que podemos ver en torno a $t \approx 2.300$ en el caso de SSE Composite y NIKKEI 225, y en torno a $t \approx 2.400$ para el resto de índices. Dicho día se aprobaba el referéndum de salida de la Unión Europea por parte de Reino Unido.

Esta descripción ilustra la fuerte interrelación entre unos mercados financieros y otros, especialmente entre los mercados europeos y americanos, y por otra parte, entre los mercados asiáticos.

A continuación analizaremos las principales características de la distribución marginal de los rendimientos diarios que acabamos de comentar.

2.2. Análisis descriptivo de los rendimientos diarios.

En primer lugar definiremos las principales medidas (posición central, dispersión y forma) que tendremos en cuenta (véase Peña, 2014). De aquí en adelante T denotará el número de rendimientos observados disponibles:

- La media es la medida de posición central más utilizada y se define como:

$$\bar{x} = \frac{x_1 + \dots + x_T}{T} = \frac{\sum x_t}{T},$$

- La mediana (Me) es otra medida de posición central y representa el valor central de la muestra, ordenando los datos de menor a mayor.
- La desviación típica es una medida de dispersión y se define como:

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{T}}$$

- El coeficiente de asimetría es una medida de forma y se define como:

$$g_1 = \frac{\sum (x_t - \bar{x})^3}{T S^3}. \quad (1)$$

- El coeficiente de curtosis o de apuntamiento es una medida de forma y mide el grado de concentración de probabilidad en torno a la media y en las colas, en comparación con la distribución normal. Se define como:

$$g_2 = \frac{\sum (x_i - \bar{x})^4}{T S^4}. \quad (2)$$

Estos dos últimos coeficientes son los más relevantes para el análisis de la distribución de los rendimientos bursátiles, ya que son los primeros indicadores de la posible no Normalidad de dichos rendimientos. Su interpretación es la siguiente:

El signo de g_1 indica la forma de la distribución:

- a) Si $g_1=0$, la distribución será simétrica.
- b) Si $g_1<0$, la distribución se alargará en los valores a la izquierda de la media (asimétrica a la izquierda).
- c) Si $g_1>0$, la distribución se alargará para los valores de la derecha de la media (asimétrica a la derecha).

Bajo el supuesto de normalidad, el coeficiente estandarizado de asimetría, que se define como:

$$g_1^* = \frac{g_1 - 0}{\sqrt{6/T}},$$

se distribuye asintóticamente como una $N(0,1)$; véase Peiró (1992, 1997) y Aparicio y Estada (2001). Por tanto, si g_1^* toma un valor fuera del intervalo $[-1.96, 1.96]$ se rechaza que la población correspondiente siga una distribución normal para un nivel de significación del 5%.

Otro indicio de asimetría nos lo da la posición de \bar{x} y Me. Estas dos medidas coinciden si la distribución es simétrica, mientras que \bar{x} será menor que Me en una distribución asimétrica a la izquierda (asimetría negativa) y \bar{x} será mayor que Me en una distribución asimétrica a la derecha (asimetría positiva).

Por lo que respecta al coeficiente de curtosis, g_2 , en función de los valores que tome, diremos que la distribución es leptocúrtica, platicúrtica o mesocúrtica:

- a) Si $g_2 > 3$, la distribución será más puntiaguda que la normal, y con colas más anchas (leptocúrtica).
- b) Si $g_2 < 3$, la distribución será más achatada que la normal, con colas más estrechas (platicúrtica).
- c) Si $g_2 = 3$, la distribución tiene un grado de apuntamiento igual que la normal (mesocúrtica).

Bajo normalidad, el coeficiente estandarizado de curtosis, definido como:

$$g_2^* = \frac{g_2 - 3}{\sqrt{24/T}},$$

se distribuye asintóticamente como una $N(0,1)$; véase Peiró (1992). Por tanto, si g_2^* toma un valor fuera del intervalo $[-1.96, 1.96]$ se rechaza que la población correspondiente siga una distribución normal para un nivel de significación del 5%.

La siguiente tabla recoge los valores de estos estadísticos para las 6 series analizadas, así como el valor mínimo y máximo que toman estas series en el periodo muestral.

Tabla 2.1: Estadísticos de los rendimientos diarios.

	\bar{x}	Me	s	Min	Max	g_1	g_1^*	g_2	g_2^*
CAC 40	-0,006	0,032	1,527	-9,472	10,595	-0,009	-0,179	8,599	57,830*
FTSE 100	0,005	0,042	1,273	-9,265	9,384	-0,152	-3,114*	10,152	73,359*
NASDAQ 100	0,040	0,096	1,405	-11,115	11,849	-0,183	-3,756*	10,893	80,831*
NIKKEI 225	0,004	0,065	1,650	-12,111	13,235	-0,481	-9,716*	10,248	73,156*
S&P 500	0,018	0,062	1,320	-9,470	10,957	-0,327	-6,689*	12,910	101,446*
SSE	0,005	0,090	1,802	-9,256	9,034	-0,604	-12,161*	6,714	37,387*

Las principales conclusiones que se extraen de la Tabla 2.1 son las siguientes: los seis índices estudiados presentan unas medidas de posición central (media y mediana) similares, prácticamente iguales a cero, indicando que el rendimiento medio en todos los índices es prácticamente nulo, por lo que podemos considerar a los mercados bursátiles como un juego de suma 0, ya que las pérdidas y las ganancias se ven compensadas con el paso del tiempo. Pero no ocurre lo mismo con la dispersión, que es mayor en el SSE Composite, por eso su gráfica de rendimientos históricos presenta mayor volatilidad.

La Tabla 2.1 también muestra que todas las distribuciones presentan una cierta asimetría a la izquierda, que se ve reflejada, por una parte, en que la media es menor que la mediana siempre y, por otra parte, en que el coeficiente de asimetría tenga signo negativo siempre. Además, los coeficientes estandarizados de asimetría indican que la asimetría observada es significativa al 5% en todos los índices excepto en el CAC 40.

En esta misma tabla podemos ver que todas las distribuciones son leptocúrticas, ya que el coeficiente de curtosis toma un valor mayor a 3 en todos los casos. Adicionalmente, los coeficientes estandarizados de curtosis muestran que la curtosis es significativa al 5% para todos los índices.

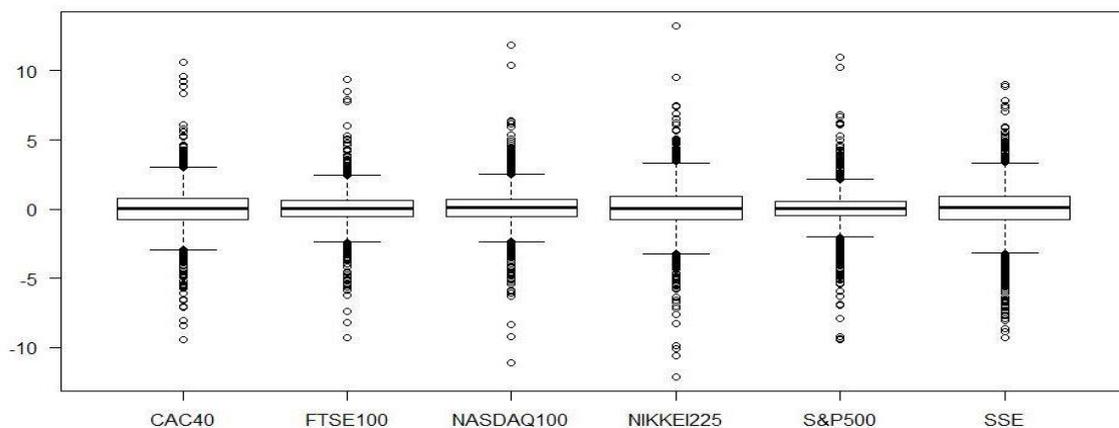
Por tanto, podemos concluir que los coeficientes recogidos en las últimas 4 columnas de la Tabla 2.1 manifiestan los primeros indicios de ausencia de normalidad en la distribución marginal de los rendimientos diarios.

Para analizar estas características con más detalle, en el Gráfico 2.2 representamos los box-plots de las distribuciones marginales de las 6 series analizadas (estos gráficos se han hecho con el software R). Un box-plot o diagrama de caja es un recurso gráfico en el que se representan las características principales de la distribución basadas en cuantiles, así como los posibles datos atípicos. En primer lugar se representan el primer (Q_1), segundo (Q_2) y tercer cuartil (Q_3), mediante una caja; después se dibujarán unos bigotes que se extienden 1,5 veces el rango intercuartílico (diferencia entre el tercer cuartil y el primero) a partir de cada extremo de la caja. Llamaremos valores atípicos a aquellas observaciones mayores que $Q_3+1.5(Q_3-Q_1)$, o menores que $Q_1-1.5(Q_3-Q_1)$, y las representaremos mediante círculos.

En el Gráfico 2.2 podemos extraer unas conclusiones similares a la Tabla 2.1. En primer lugar vemos que la mediana (Q_2) es próxima a 0. En

segundo lugar, este gráfico pone en evidencia la leptocurtosis de las distribuciones por 2 hechos, el primero es que todas las cajas sean muy estrechas (el rango intercuartílico es pequeño), lo que quiere decir que el 50% de las observaciones están contenidas en un intervalo pequeño y, por tanto, la concentración de valores en torno a la media es muy alto; el segundo hecho es la gran cantidad de valores atípicos que encontramos en nuestras distribuciones, lo que provocará colas más anchas que la distribución Normal.

Gráfico 2.2: Boxplots de las rendimientos diarios.



3. CONTRASTES DE NORMALIDAD SOBRE LA DISTRIBUCIÓN DE LOS RENDIMIENTOS DIARIOS.

En esta sección analizaremos si la distribución gaussiana se ajusta a las distribuciones marginales de los rendimientos estudiados anteriormente. Para ello, en primer lugar haremos una representación gráfica de los histogramas de dichos rendimientos junto con la correspondiente función de densidad gaussiana, después analizaremos los Q-Q plots y, por último, realizaremos los correspondientes contrastes de normalidad formalmente, todo ello realizado mediante el software Matlab.

3.1. Análisis gráfico: histogramas y Q-Q plots.

Para adentrarnos en este análisis, primero estimaremos los parámetros de la distribución normal para cada uno de nuestros índices mediante el método de máxima verosimilitud. Los parámetros estimados coinciden con la

media y la desviación típica muestral correspondientes a cada índice, es decir, $\hat{\mu}=\bar{X}$ y $\hat{\sigma}=S$. En segundo lugar analizaremos gráficamente si la distribución de frecuencias (relativas) de cada índice es similar a la función de densidad de una normal con parámetros estimados $\hat{\mu}$ y $\hat{\sigma}$. Para ello, el Gráfico 3.2 recoge la comparativa entre la distribución empírica del rendimiento de cada índice (representada por un histograma), y las correspondientes distribuciones normales teóricas con la misma media y desviación típica que los datos.

Como podemos observar en el Gráfico 3.2, la distribución normal no presenta un buen ajuste ya que la leptocurtosis de los datos provoca que se infraestimen los valores centrales si utilizamos una distribución normal, ello indica que es mucho más probable observar rendimientos bursátiles diarios en torno a 0 que en la distribución normal. Al igual ocurre con los extremos de las colas, donde se aprecia que los rendimientos, negativos y positivos, son mucho más frecuentes de lo que lo serían en un modelo gaussiano.

A esta misma conclusión se llega analizando los Q-Q plots representados en el Gráfico 3.3. Este gráfico nos permite analizar si las distribuciones empíricas son similares o no a la distribución teórica mediante la comparación de sus cuantiles. Para ello, se ordenan los datos de menor a mayor y se comparan los cuantiles teóricos (en este caso bajo el supuesto de normalidad) con los correspondientes cuantiles empíricos de los datos. En el caso de que la representación gráfica diera como resultado la diagonal de un cuadrado, ambas distribuciones tendrían cuantiles muy próximos y diríamos que la distribución empírica se ajusta bien a la teórica.

En el Gráfico 3.3 destaca la diferencia existente entre las distribuciones empíricas y las correspondientes distribuciones normales en ambos extremos, tanto el izquierdo como el derecho. En el extremo izquierdo la gráfica muestra que para encontrar el cuantil z_q en la distribución empírica (eje de ordenadas) tendríamos que irnos a valores más alejados de la media con respecto al valor en el que nos encontraríamos este mismo cuantil z_q en la distribución teórica normal (eje de abscisas). Esto muestra que la cola izquierda de la distribución empírica es más pesada que la cola izquierda de la normal. En el extremo derecho ocurre lo mismo que en el extremo izquierdo, por lo que la cola derecha de la distribución empírica es más pesada que la cola derecha de la distribución gaussiana. En los valores centrales nos encontramos con que la

gráfica se posiciona sobre la línea discontinua, lo que quiere decir que los cuantiles centrales de los rendimientos están muy próximos a los correspondientes cuantiles de la normal, por lo que el centro de las distribuciones se alcanza en un punto muy similar en ambas.

Gráfico 3.2: Comparativa de las distribuciones empíricas de las rentabilidades diarias, y las correspondientes distribuciones normales.

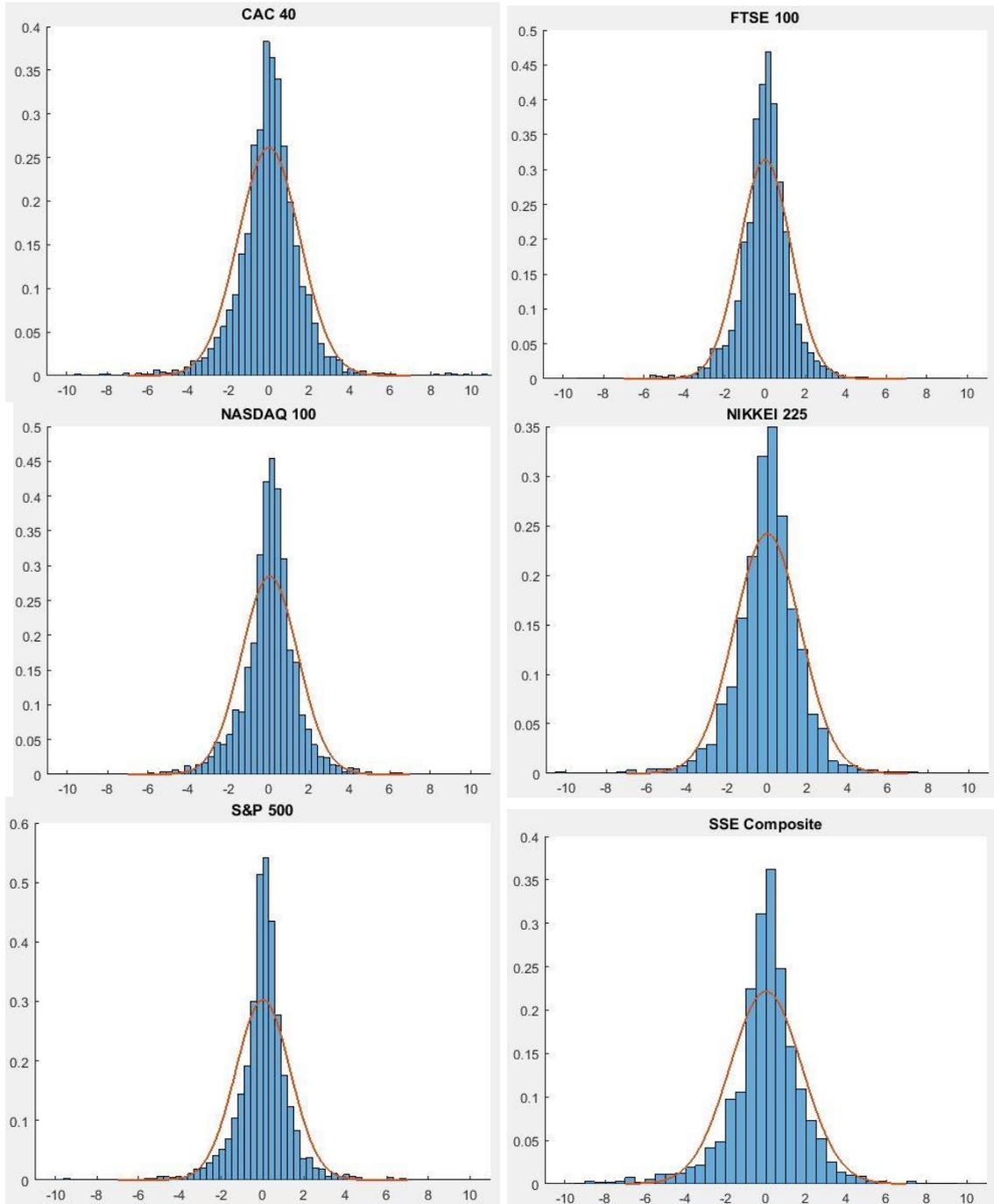
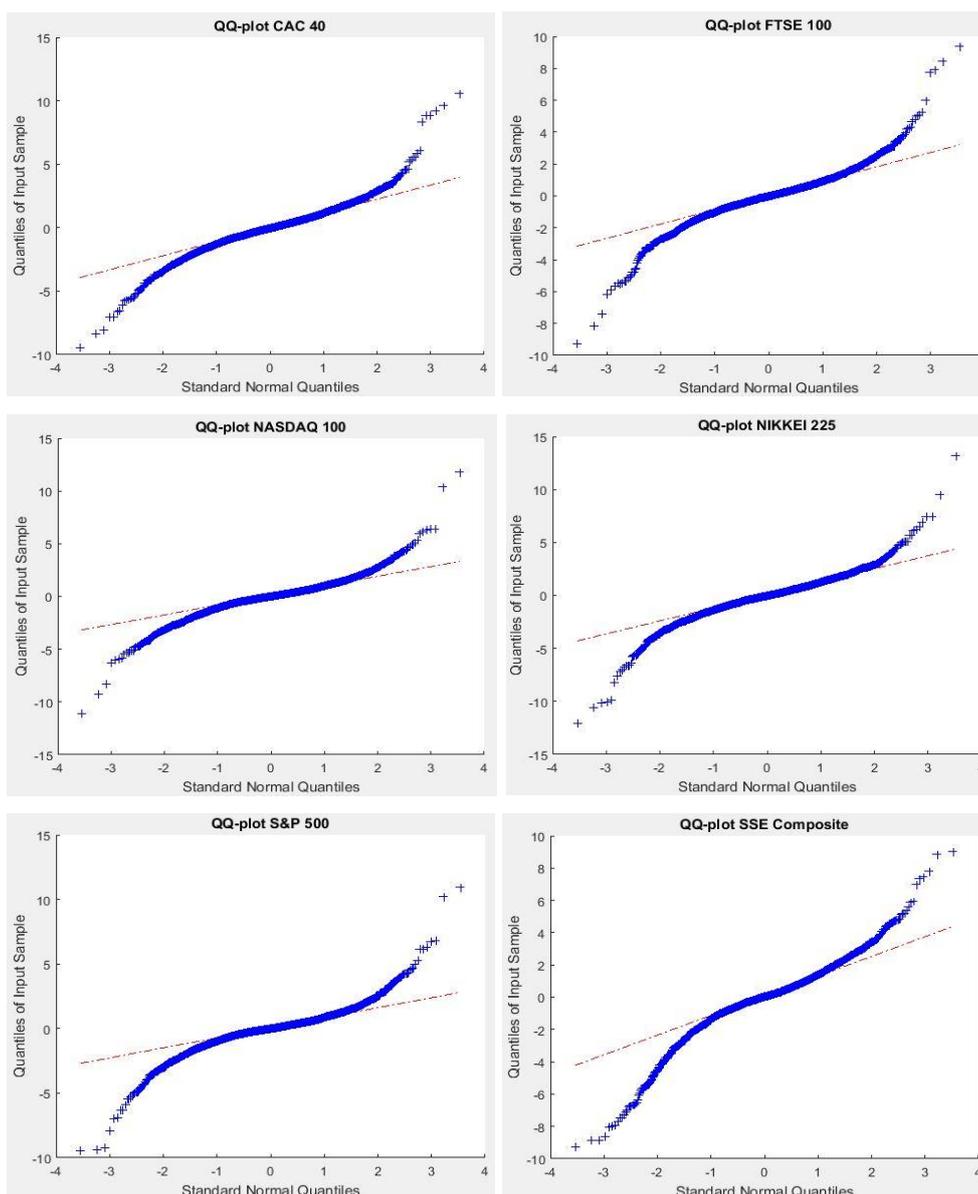


Gráfico 3.3: Q-Q plots de los rendimientos diarios



Por tanto, llegamos a la misma conclusión que en los apartados anteriores: la distribución marginal de los rendimientos parece no seguir una distribución normal debido al exceso de curtosis y una cierta asimetría izquierda. A continuación contrastaremos formalmente esta hipótesis.

3.2. Contrastes de Normalidad: definición y propiedades.

En esta sección aplicaremos distintos contrastes de bondad de ajuste para ver si nuestros datos empíricos pueden seguir una distribución normal.

Sea X la variable que representa los rendimientos diarios de un determinado índice bursátil. Las hipótesis a contrastar son:

$$H_0: X \rightarrow N(\mu, \sigma)$$

$$H_1: X \not\rightarrow N(\mu, \sigma)$$

Realizaremos cuatro contrastes: el de Kolmogorov-Smirnov-Lilliefors, el contraste de χ^2 de Pearson, el de Jarque Bera y el de Anderson Darling.

Kolmogorov-Smirnov-Lilliefors

El estadístico del contraste de Kolmogorov-Smirnov se define como:

$$D_T = \max |F_T(x) - F(x)|,$$

donde $F_T(x)$ es la función de distribución empírica y $F(x)$ es la función de distribución teórica bajo H_0 . Por tanto, D_T es la máxima discrepancia entre ambas funciones de distribución, y si ésta es suficientemente grande (mayor que el valor crítico para un nivel de significación dado y con ese número de observaciones), se rechazará la hipótesis nula.

Lilliefors tabuló la distribución del estadístico D_T para el caso de normalidad, estimando μ y σ mediante los datos muestrales, \bar{x} y S (Peña, 2014). En este caso, para el cálculo de $F(x)$ bajo H_0 , se toma:

$$F(x) = \Phi\left(\frac{x - \bar{x}}{s}\right).$$

Chi Cuadrado

Se basa en comparar las frecuencias relativas observadas de ciertos intervalos (histograma de los datos observados) con las correspondientes probabilidades de dichos intervalos bajo la distribución teórica establecida en H_0 . Para ello, se agrupan los T datos en k clases, se calcula para cada clase i , con $i=1, \dots, k$, la frecuencia observada en la muestra de dicha clase (O_i), que coincide con el número de datos muestrales en esa clase i , y la frecuencia esperada teórica de dicha clase (E_i). La discrepancia entre las frecuencias observadas y las esperadas las calcularemos mediante el siguiente estadístico:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

donde $O_i=T_i$, y $E_i=Tp_i$, siendo p_i la probabilidad que el modelo asigna a la clase i , y T_i el número de observaciones en dicha clase.

Bajo H_0 , este estadístico se distribuye aproximadamente como una χ^2 con $k-1$ grados de libertad si el modelo bajo H_0 está totalmente especificado, o con $k-r-1$ grados de libertad si para calcular las probabilidades p_i deben estimarse r parámetros desconocidos del modelo bajo H_0 . Para el caso en el que se contrasta la normalidad, la distribución χ^2 tendrá $k-3$ grados de libertad, ya que se estiman 2 parámetros (μ y σ^2).

Rechazaremos H_0 , para un nivel de significación α cuando:

$$\chi^2 \geq \chi_{\alpha}^2(k - r - 1),$$

siendo $\chi_{\alpha}^2(k - r - 1)$ el percentil $(1-\alpha)*100\%$ de la distribución χ^2 .

Una cuestión importante es elegir el número de clases en las que se van a agrupar los datos, ya que debe haber un número elevado de clases (se recomienda $k \geq 5$) para una correcta utilización del contraste (Peña, 2014). En nuestro caso, al realizar los contrastes con Matlab, ya crea él por defecto el número de clases.

Jarque Bera

Este estadístico sirve para comparar si la distribución empírica tiene los coeficientes de asimetría y curtosis de una normal (asimetría=0 y curtosis=3), y se define mediante la siguiente expresión:

$$JB = T \left(\frac{g_1^2}{6} + \frac{(g_2 - 3)^2}{24} \right)$$

donde g_1 es el coeficiente de asimetría definido en (1) y g_2 es el coeficiente de curtosis definido en (2). Bajo el supuesto de normalidad este estadístico se distribuirá aproximadamente como una χ^2 con 2 grados de libertad (Peiró, 1992). Por tanto, se rechazará la normalidad para un nivel de significación α si JB es mayor que el valor crítico para ese nivel α dado.

Anderson Darling

Es uno de los contrastes más potentes para detectar desviaciones del supuesto de normalidad. El estadístico del contraste es:

$$A^2 = -T - S$$

donde:

$$S = \sum_{i=1}^T \frac{(2i-1)}{T} [\ln \Phi(Y_i) + \ln(1 - \Phi(Y_{T+1-i}))],$$

donde $\Phi(Y_i)$ es la función de distribución normal evaluada en el punto Y_i , siendo Y_i las observaciones estandarizadas, $Y_i = (X_i - \hat{\mu})/\hat{\sigma}$, y T es el número de observaciones de la muestra.

Si A^2 excede de un valor crítico para un nivel de significación α dado, la hipótesis de normalidad será rechazada para ese nivel de significación α .

Este contraste es una modificación del test de Kolmogorov-Smirnov, e incrementa su eficacia en las colas de la distribución, por lo que es muy útil para los valores extremos (Marques, 2001).

A continuación se mostrarán los resultados obtenidos por los contrastes para cada una de nuestras muestras.

3.3. Análisis de resultados

La Tabla 3.1 recoge los resultados de los contrastes mencionados anteriormente para cada índice. Los resultados obtenidos no son una sorpresa: todos ellos rechazan la hipótesis nula de Normalidad con gran contundencia (los p-valores son inferiores incluso al 5%).

Tabla 3.1: Contrastes de Normalidad para las series de rendimientos diarios.

		CAC 40	FTSE 100	NASDAQ 100	NIKKEI 225	S&P 500	SSE
KSL	estadístico	0,071	0,079	0,096	0,072	0,110	0,096
	(p-valor)	(0,001)	(0,001)	(0,001)	(0,001)	(0,001)	(0,001)
CHI 2	estadístico	126,504	30,964	21,588	123,187	303,497	167,153
	(p-valor)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)
JB	estadístico	3344,300	5391,600	6547,800	5445,900	10337,000	1545,700
	(p-valor)	(0,001)	(0,001)	(0,001)	(0,001)	(0,001)	(0,001)
AD	estadístico	26,616	34,510	Inf	27,261	62,729	37,520
	(p-valor)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)

Estos resultados concuerdan con los numerosos estudios realizados por otros autores sobre la distribución de los rendimientos diarios, ya que en toda la literatura al respecto se rechazaba el supuesto de normalidad para dichas distribuciones y además todas ellas eran leptocúrticas, al igual que las

nuestras; véase, por ejemplo, Martínez y Arnillas (1992); Hurst y Platen (1997); Aparicio y Estrada (2001); Behr y Pötter (2009).

Por lo tanto, a continuación propondremos distribuciones que se ajusten mejor a dichos rendimientos, utilizando como referencia las distribuciones propuestas en otros trabajos que perseguían este mismo objetivo; véase, por ejemplo, Peiró (1992, 1997); Hurst y Platen (1997); Aparicio y Estrada (2001); Harris y Küçüközmen (2001); Behr y Pötter (2009); Necula (2009).

4. DISTRIBUCIONES ALTERNATIVAS PARA LOS RENDIMIENTOS DIARIOS: HIPERBÓLICA GENERALIZA, LOGÍSTICA Y MIXTURA GAUSSIANA

En esta sección presentaremos varias distribuciones de probabilidad que permiten reflejar la leptocurtosis (con grandes picos y colas alargadas) y la asimetría observadas en los rendimientos.

4.1. Descripción de las distribuciones utilizadas: definición y propiedades

4.1.1. Distribución hiperbólica generalizada: casos particulares

La distribución hiperbólica generalizada (GH) fue introducida por Barndorff-Nielsen (1977) para modelar la forma que tomaban las dunas debido a la arena arrastrada por el viento. A causa de su flexibilidad, la distribución hiperbólica generalizada es considerada como un buen modelo para modelizar la distribución de rendimientos; véase Behr y Pötter (2009); Necula (2009).

Si la variable aleatoria X sigue una distribución hiperbólica generalizada, se denota como $X \sim GH(\lambda, \alpha, \beta, \delta, \mu)$, donde μ es el parámetro de localización, δ el de escala, α el de forma, β determina la asimetría (si $\beta=0$ la distribución será simétrica, si $\beta < 0$ será asimétrica a la izquierda, y si $\beta > 0$ será asimétrica a la derecha) y λ determina la curtosis (Necula, 2009; Hammerstein, 2010). Estos parámetros deben cumplir las siguientes condiciones (Aas y Haff, 2005):

$$\delta \geq 0, |\beta| < \alpha \text{ si } \lambda > 0$$

$$\delta > 0, |\beta| < \alpha \text{ si } \lambda = 0$$

$$\delta > 0, |\beta| \leq \alpha \text{ si } \lambda < 0$$

La función de densidad viene definida por:

$$f(x) = \kappa \{\delta^2 + (x - \mu)^2\}^{\frac{1}{2}(\lambda - \frac{1}{2})} K_{\lambda - \frac{1}{2}}(\alpha \sqrt{\delta^2 + (x - \mu)^2}) e^{\beta(x - \mu)}, \quad (3)$$

donde $\kappa = \frac{(\alpha^2 - \beta^2)^{\frac{\lambda}{2}}}{\sqrt{2\pi} \alpha^{\lambda - \frac{1}{2}} \delta^{\lambda} K_{\lambda}(\delta \sqrt{\alpha^2 - \beta^2})}$, y la función $K_{\lambda}(t)$ es la función Bessel modificada del tercer tipo con índice λ , también conocida como la función MacDonald, que se define como

$$K_{\lambda}(t) = \frac{1}{2} \int_0^{\infty} u^{\lambda - 1} e^{-\frac{1}{2}t(u + u^{-1})} du, \quad t > 0.$$

La función de distribución no tiene expresión conocida y normalmente se haya integrando numéricamente la función de densidad. La función de densidad es unimodal y se pueden hallar todos los momentos. La forma de la función de densidad se puede acomodar para que sea leptocúrtica, platicúrtica, asimétrica a la derecha o a la izquierda (Behr y Pötter, 2009).

La media y la varianza de una variable aleatoria X con distribución hiperbólica generalizada son (Aas y Haff, 2005):

$$E(X) = \mu + \frac{\beta \delta K_{\lambda+1}(\delta \gamma)}{\gamma K_{\lambda}(\delta \gamma)},$$

$$Var(X) = \delta^2 \left(\frac{K_{\lambda+1}(\delta \gamma)}{\delta \gamma K_{\lambda}(\delta \gamma)} + \frac{\beta^2}{\gamma^2} \left(\frac{K_{\lambda+2}(\delta \gamma)}{K_{\lambda}(\delta \gamma)} - \left(\frac{K_{\lambda+1}(\delta \gamma)}{K_{\lambda}(\delta \gamma)} \right)^2 \right) \right),$$

donde $\gamma = \sqrt{\alpha^2 - \beta^2}$. Los coeficientes de asimetría y curtosis vienen definidos por (Hammerstein, 2010):

$$g_1(X) = Var(X)^{-3/2} \left[\frac{\beta^3 \delta^6}{(\delta \gamma)^3} \left(\frac{K_{\lambda+3}(\delta \gamma)}{K_{\lambda}(\delta \gamma)} - \frac{3K_{\lambda+2}(\delta \gamma)K_{\lambda+1}(\delta \gamma)}{K_{\lambda}^2(\delta \gamma)} + \frac{2K_{\lambda+1}^3(\delta \gamma)}{K_{\lambda}^3(\delta \gamma)} \right) + \frac{3\beta \delta^4}{(\delta \gamma)^2} \left(\frac{K_{\lambda+2}(\delta \gamma)}{K_{\lambda}(\delta \gamma)} - \frac{K_{\lambda+1}^2(\delta \gamma)}{K_{\lambda}^2(\delta \gamma)} \right) \right],$$

$$g_2(X) = -3 + Var(X)^{-2} \left[\frac{\delta^8 \beta^4}{(\delta \gamma)^4} \left(\frac{K_{\lambda+4}(\delta \gamma)}{K_{\lambda}(\delta \gamma)} - \frac{4K_{\lambda+3}(\delta \gamma)}{K_{\lambda}^2(\delta \gamma)} + \frac{6K_{\lambda+2}(\delta \gamma)K_{\lambda+1}^2(\delta \gamma)}{K_{\lambda}^3(\delta \gamma)} - \frac{3K_{\lambda+1}^4(\delta \gamma)}{K_{\lambda}^4(\delta \gamma)} \right) + \frac{\delta^6 \beta^2}{(\delta \gamma)^3} \left(\frac{6K_{\lambda+3}(\delta \gamma)}{K_{\lambda}(\delta \gamma)} - \frac{12K_{\lambda+2}(\delta \gamma)K_{\lambda+1}(\delta \gamma)}{K_{\lambda}^2(\delta \gamma)} + \frac{6K_{\lambda+1}^3(\delta \gamma)}{K_{\lambda}^3(\delta \gamma)} \right) + \frac{3\delta^4}{(\delta \gamma)^3} \frac{K_{\lambda+2}(\delta \gamma)}{K_{\lambda}(\delta \gamma)} \right].$$

En función del valor que tomen sus parámetros, la distribución hiperbólica generalizada incluye, como casos particulares, una serie de distribuciones interesantes, entre las que destacan las siguientes, como las más utilizadas en la literatura:

Distribución hiperbólica (simple)

Se obtiene si $\lambda=1$ en (3), como aparece reflejado en Necula (2009) y Behr y Pötter (2009) y se denota por $\text{Hyp}(\alpha, \beta, \delta, \mu)$. Su función de densidad es:

$$f(x) = \frac{\sqrt{\alpha^2 - \beta^2}}{2\alpha\delta K_1(\delta\sqrt{\alpha^2 - \beta^2})} \exp\left(-\alpha\sqrt{\delta^2 + (x - \mu)^2} + \beta(x - \mu)\right).$$

La media y la varianza de esta distribución son:

$$E(X) = \mu + \frac{\beta\delta K_2(\delta\gamma)}{\gamma K_1(\delta\gamma)},$$

$$\text{Var}(X) = \delta^2 \left(\frac{K_2(\delta\gamma)}{\delta\gamma K_1(\delta\gamma)} + \frac{\beta^2}{\gamma^2} \left(\frac{K_3(\delta\gamma)}{K_1(\delta\gamma)} - \left(\frac{K_2(\delta\gamma)}{K_1(\delta\gamma)} \right)^2 \right) \right),$$

mientras que los coeficientes de asimetría y curtosis vienen dados por:

$$g_1(X) = \text{Var}(X)^{-\frac{3}{2}} \left[\frac{\beta^3 \delta^6}{(\delta\gamma)^3} \left(\frac{K_4(\delta\gamma)}{K_1(\delta\gamma)} - \frac{3K_3(\delta\gamma)K_2(\delta\gamma)}{K_1^2(\delta\gamma)} + \frac{2K_2^3(\delta\gamma)}{K_1^3(\delta\gamma)} \right) + \frac{3\beta\delta^4}{(\delta\gamma)^2} \left(\frac{K_3(\delta\gamma)}{K_1(\delta\gamma)} - \frac{K_2^2(\delta\gamma)}{K_1^2(\delta\gamma)} \right) \right],$$

$$g_2(X) = -3 + \text{Var}(X)^{-2} \left[\frac{\delta^8 \beta^4}{(\delta\gamma)^4} \left(\frac{K_5(\delta\gamma)}{K_1(\delta\gamma)} - \frac{4K_4(\delta\gamma)}{K_1^2(\delta\gamma)} + \frac{6K_3(\delta\gamma)K_2^2(\delta\gamma)}{K_1^3(\delta\gamma)} - \frac{3K_2^4(\delta\gamma)}{K_1^4(\delta\gamma)} \right) + \frac{\delta^6 \beta^2}{(\delta\gamma)^3} \left(\frac{6K_4(\delta\gamma)}{K_1(\delta\gamma)} - \frac{12K_3(\delta\gamma)K_2(\delta\gamma)}{K_1^2(\delta\gamma)} + \frac{6K_2^3(\delta\gamma)}{K_1^3(\delta\gamma)} \right) + \frac{3\delta^4}{(\delta\gamma)^3} \frac{K_3(\delta\gamma)}{K_1(\delta\gamma)} \right].$$

Normal Inversa Gaussiana

Se obtiene con $\lambda=-1/2$ en (3) y se denota por $\text{NIG}(\alpha, \beta, \delta, \mu)$. Su función de densidad es:

$$f(x) = \frac{\alpha\delta}{\pi} \exp\left(\delta\sqrt{\alpha^2 - \beta^2} + \beta(x - \mu)\right) \frac{K_1(\alpha\sqrt{\delta^2 + (x - \mu)^2})}{\sqrt{\delta^2 + (x - \mu)^2}}.$$

La Distribución NIG es la única de las subclases de la distribución GH que permite convoluciones, propiedad importante para modelar rendimientos financieros. Concretamente, si $X_1 \sim \text{NIG}(\alpha, \beta, \delta_1, \mu_1)$ y $X_2 \sim \text{NIG}(\alpha, \beta, \delta_2, \mu_2)$ entonces la distribución de la convolución de las 2 variables aleatorias será una distribución NIG con $X_1 + X_2 \sim \text{NIG}(\alpha, \beta, \delta_1 + \delta_2, \mu_1 + \mu_2)$; véase Necula (2009).

La media y la varianza de la distribución son:

$$E(X) = \mu + \frac{\delta\beta}{\sqrt{\alpha^2 - \beta^2}}, \quad \text{Var}(X) = \frac{\delta\alpha^2}{(\alpha^2 - \beta^2)^{3/2}},$$

mientras que los coeficientes de asimetría y curtosis vienen dados por:

$$g_1(X) = 3 \frac{\beta}{\alpha} \frac{1}{\delta^{1/2}(\alpha^2 - \beta^2)^{1/4}}, \quad g_2(X) = 3 \left(1 + 4 \left(\frac{\beta}{\alpha}\right)^2\right) \frac{1}{\delta(\alpha^2 - \beta^2)^{1/2}},$$

donde g_1 y g_2 deben satisfacer la condición $|g_1| \leq \sqrt{\frac{3}{5}g_2}$.

Skew Student's t

Se obtiene cuando $\lambda = -v/2$ y $\alpha \rightarrow |\beta|$ en (3) y se denota por $t\text{Skew}(v, \beta, \delta, \mu)$, donde $v = -2\lambda$. Su densidad es (Aas y Haff, 2005):

$$f(x) = \frac{2^{\frac{1-v}{2}} \delta^v |\beta|^{\frac{v+1}{2}} K_{\frac{v+1}{2}}(\sqrt{\beta^2(\delta^2 + (x - \mu)^2)}) \exp(\beta(x - \mu))}{\Gamma\left(\frac{v}{2}\right) \sqrt{\pi} \left(\sqrt{\delta^2 + (x - \mu)^2}\right)^{\frac{v+1}{2}}}.$$

Los principales momentos de esta distribución son (Aas y Haff, 2005):

$$E(X) = \mu + \frac{\beta\delta^2}{v-2}, \quad \text{Var}(X) = \frac{2\beta^2\delta^4}{(v-2)^2(v-4)} + \frac{\delta^2}{v-2},$$

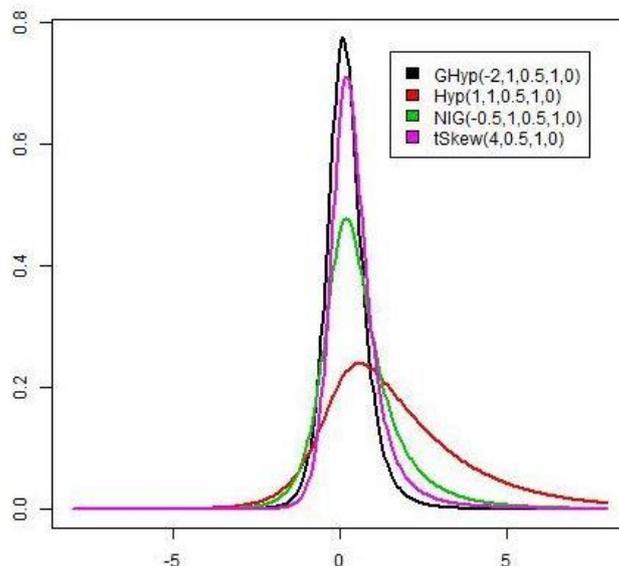
$$g_1(X) = \frac{2(v-4)^{1/2}\beta\delta}{[2\beta^2\delta^2 + (v-2)(v-4)]^{3/2}} \left[3(v-2) + \frac{8\beta^2\delta^2}{v-6}\right],$$

$$g_2(X) = \frac{6}{[2\beta^2\delta^2 + (v-2)(v-4)]^2} \left[(v-2)^2(v-4) + \frac{16\beta^2\delta^2(v-2)(v-4)}{v-6} + \frac{8\beta^4\delta^4(5v-22)}{(v-6)(v-8)}\right],$$

donde la varianza será sólo finita cuando $v > 4$ y g_1 y g_2 no existen cuando $v \leq 6$ y $v \leq 8$, respectivamente.

A continuación se muestran las gráficas de las diferentes clases de la Distribución Hiperbólica Generalizada para distintos parámetros:

Gráfica 4.1: Hiperbólica generalizada.



En la Gráfica 4.1 podemos ver que en función de los valores de los parámetros, la función de densidad de la distribución hiperbólica generalizada puede tomar distintas formas, acomodándose tanto a la curtosis como a la asimetría, lo que es muy útil para modelizar los rendimientos bursátiles.

4.1.2. Distribución logística

Se trata de una distribución simétrica con colas más largas y un coeficiente de curtosis más alto que la normal; véase Peiró (1992) y Aparicio y Estrada (2001). Su función de densidad se define como:

$$f(x) = \frac{\exp\left(\frac{x-\mu}{\alpha}\right)}{\alpha \left[1 + \exp\left(\frac{x-\mu}{\alpha}\right)\right]^2},$$

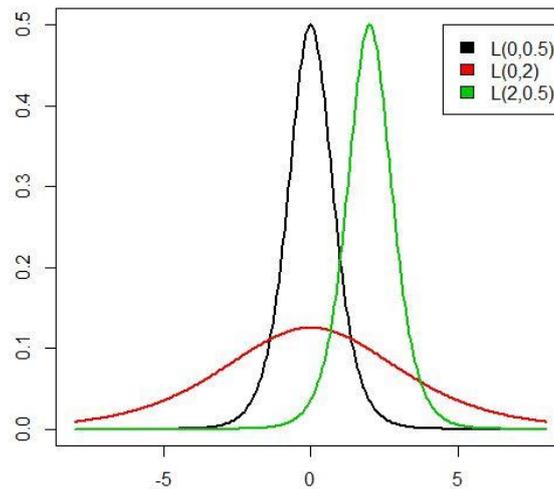
donde μ ($-\infty < \mu < \infty$) es el parámetro de localización o media y α ($\alpha > 0$) es el parámetro de escala o dispersión.

Los principales momentos de la distribución Logística, $L(\mu, \alpha)$, son:

$$E(X) = \mu, \text{Var}(X) = \left(\frac{\pi^2}{3}\right) \alpha^2, g_1(X) = 0, g_2(X) = \frac{21}{5}.$$

La Gráfica 4.2 representa la función de densidad de esta distribución variando los parámetros y pone de manifiesto que la distribución Logística se puede adaptar a la curtosis, pero es simétrica, por lo que no se ajustaría a la asimetría de los rendimientos financieros.

Gráfica 4.2: Distribución Logística



4.1.3. *Mixtura Gaussiana*

Se trata de una distribución que es la mezcla de varias distribuciones normales. En nuestro análisis sólo tendremos en cuenta la mixtura de 2 distribuciones normales, con ello consideramos que los rendimientos se obtienen con una probabilidad ρ de una distribución normal con unos determinados parámetros, y con una probabilidad $1-\rho$ de otra distribución normal con otros parámetros, véase Peiró (1992, 1997); Aparicio y Estrada (2001); Behr y Pötter (2009). Tendríamos:

$$X_1 \sim N_1(\mu_1, \sigma_1) \text{ con probabilidad } \rho.$$

$$X_2 \sim N_2(\mu_2, \sigma_2) \text{ con probabilidad } 1-\rho.$$

La función de densidad de esta distribución está definida como:

$$f(x) = \frac{\rho}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(x - \mu_1)^2}{2\sigma_1^2}\right) + \frac{1 - \rho}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(x - \mu_2)^2}{2\sigma_2^2}\right),$$

donde μ_i ($-\infty < \mu_i < \infty$) es el parámetro de localización y σ_i^2 ($\sigma_i^2 > 0$), $i = 1,2$, es el parámetro de dispersión de cada una de las normales de la mixtura.

Los principales momentos de esta mixtura de dos normales son:

$$E(X) = \mu = \rho\mu_1 + (1 - \rho)\mu_2 ,$$

$$Var(X) = \sigma^2 = \rho[(\mu_1 - E(X))^2 + \sigma_1^2] + (1 - \rho)[(\mu_2 - E(X))^2 + \sigma_2^2] ,$$

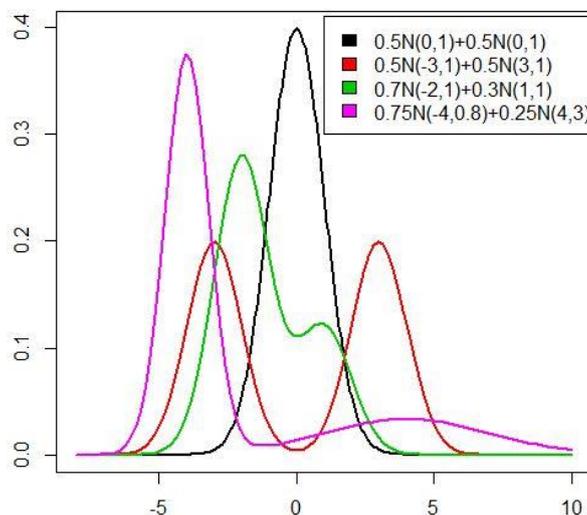
$$g_1(X) = \frac{1}{\sigma^3} [(\rho(\mu_1 - \mu)(3\sigma_1^3 + (\mu_1 - \mu)^2)) + ((1 - \rho)(\mu_2 - \mu)(3\sigma_2^3 + (\mu_2 - \mu)^2))],$$

$$g_2(X) = \frac{1}{\sigma^4} [\rho(3\sigma_1^4 + 6(\mu_1 - \mu)^2\sigma_1^2 + (\mu_1 - \mu)^4) + (1 - \rho)(3\sigma_2^4 + 6(\mu_2 - \mu)^2\sigma_2^2 + (\mu_2 - \mu)^4)].$$

Este modelo puede dar lugar a diversos tipos de distribuciones en función de los valores de los parámetros y de la probabilidad ρ . Así, nos podemos encontrar con distribuciones simétricas (cuando $\mu_1 = \mu_2$ ó $\rho=0.5$ y $\sigma_1=\sigma_2$), asimétricas, unimodales, bimodales, o con la propia distribución normal (cuando $\mu_1 = \mu_2$ y $\sigma_1 = \sigma_2$ ó ρ toma un valor extremo; con $\rho \approx 0$ ó $\rho \approx 1$); véase Peiró (1992). Por otra parte, utilizando 2 componentes en la mixtura gaussiana se consigue captar los rasgos más importantes de la distribución empírica, sin caer en la sobreparametrización que nos llevaría al utilizar más componentes; véase Peiró (1997).

En la Gráfica 4.3 representaremos la distribución mixtura gaussiana de 2 componentes, con distintos valores de los parámetros. Podemos ver que la distribución mixtura gaussiana se puede adaptar tanto a la curtosis cómo a la asimetría, además su función de densidad permite que haya más de 1 moda.

Gráfica 4.3: Distribución Mixtura Gaussiana



4.2. Resultados de la estimación

En esta sección se estimarán los modelos anteriormente explicados para cada una de las series de rendimientos descritas en la sección 2.1.

Para estimar los parámetros de las distribuciones se utilizará el método de máxima verosimilitud. Para ello, hemos utilizado el software R y Matlab para las distribuciones mixtura gaussiana y logística, obteniendo en ambos casos resultados muy similares, y para la distribución hiperbólica generalizada hemos utilizado sólo R, debido a la imposibilidad de obtenerlo por Matlab. Para éste cálculo, en R hemos utilizado las librerías “ghyp”, “mixtools”, y “fitdistrplus”. Los parámetros estimados son los que aparecen en la Tabla 4.1.

Tabla 4.1: Parámetros estimados para los rendimientos diarios.

		CAC 40	FTSE 100	NASDAQ 100	NIKKEI 225	S&P 500	SSE
ghyp	λ	-1,000	-1,052	-0,071	-1,525	-0,210	0,612
	α	0,402	0,371	0,616	0,241	0,537	0,663
	δ	1,526	1,189	0,661	1,951	0,557	0,332
	β	-0,046	-0,047	-0,063	-0,061	-0,062	-0,046
	μ	0,104	0,081	0,164	0,165	0,122	0,154
hyp	λ	1,000	1,000	1,000	1,000	1,000	1,000
	α	0,996	1,188	1,072	0,932	1,184	0,812
	δ	0,344	0,180	0,068	0,366	0,009	0,000
	β	-0,050	-0,061	-0,069	-0,066	-0,063	-0,046
	μ	0,104	0,095	0,161	0,169	0,108	0,146
NIG	λ	-0,500	-0,500	-0,500	-0,500	-0,500	-0,500
	α	0,549	0,596	0,464	0,526	0,414	0,378
	δ	1,246	0,934	0,899	1,351	0,706	1,230
	β	-0,047	-0,050	-0,063	-0,067	-0,061	-0,057
	μ	0,102	0,084	0,163	0,178	0,123	0,193
Skew t	ν	3,513	3,128	2,702	3,626	2,298	2,865
	δ	1,937	1,432	1,368	2,118	1,047	1,899
	β	-0,041	-0,040	-0,051	-0,060	-0,046	-0,055
	μ	0,092	0,073	0,151	0,166	0,113	0,201
mixtura	p	0,821	0,792	0,697	0,888	0,757	0,674
	$1-p$	0,179	0,208	0,303	0,112	0,243	0,326
	$\mu 1$	0,061	0,053	0,145	0,080	0,102	0,169
	$\mu 2$	-0,311	-0,178	-0,202	-0,595	-0,244	-0,332
	$\sigma 1$	1,048	0,796	0,740	1,199	0,685	0,979
	$\sigma 2$	2,806	2,308	2,273	3,534	2,371	2,794
logística	μ	0,018	0,024	0,073	0,044	0,052	0,071
	σ	0,780	0,636	0,695	0,835	0,628	0,914

* ghyp denota la distribución hiperbólica generalizada, hyp es la distribución hiperbólica (simple), NIG es la Normal Inversa Gaussiana y Skew t es la distribución Skew Student's t.

Como podemos ver, el parámetro β estimado para la distribución hiperbólica generalizada (ghyp) y sus tres casos particulares (hyp, NIG, Skew t), es siempre negativo y pequeño, lo que indica una cierta asimetría a la izquierda. Además, en la mixtura gaussiana siempre se cumple que: para la mayor proporción le corresponde una media positiva y cercana a 0, mientras que para la menor proporción le corresponde una media negativa; la desviación típica de la mayor proporción siempre es menor que la desviación típica de la menor proporción. Los valores de los parámetros de la mixtura nos indican una asimetría a la izquierda, al igual que ocurría con la distribución hiperbólica y con las distribuciones empíricas marginales de rendimientos estudiadas.

Para la comparación de la función de densidad empírica de las observaciones con la función de densidad teórica de las distribuciones estimadas, se ha utilizado la densidad kernel para representar la función de densidad empírica, en lugar del histograma. El estimador kernel de la densidad $f(x)$ es un estimador no paramétrico que se define como:

$$\hat{f}_T(t) = \frac{1}{Th_T} \sum_{i=1}^T K\left(\frac{t - x_i}{h_T}\right),$$

donde t es el punto donde se estima la función de densidad, x_i es el valor i -ésimo de la variable x , h_T es el parámetro de suavizado (también conocido como ventana o ancho de ventana) que debe tender a 0 cuando $T \rightarrow \infty$, y K es una función de densidad conocida como núcleo o kernel, y es la que da forma a las “campanas” que estimamos. En la práctica se suele utilizar la función de densidad de la normal estándar como K , aunque se pueden utilizar otras funciones de densidad; véase Silverman (1990).

La mayor dificultad en la estimación kernel es elegir el parámetro h_T . Este es similar a la longitud de los intervalos de los histogramas, por lo que si elegimos un h_T demasiado grande obtenemos una función de densidad “muy plana” y si elegimos un h_T demasiado pequeño, obtenemos una función de densidad con muchos picos. Silverman (1990) propone 2 anchos de banda óptimos, en función de la presencia o no de valores atípicos en la muestra y del kernel utilizado:

$$h_{opt} = 1.06sT^{-1/5}, \quad (4)$$

$$h_{opt} = 0.79RT^{-1/5}, \quad (5)$$

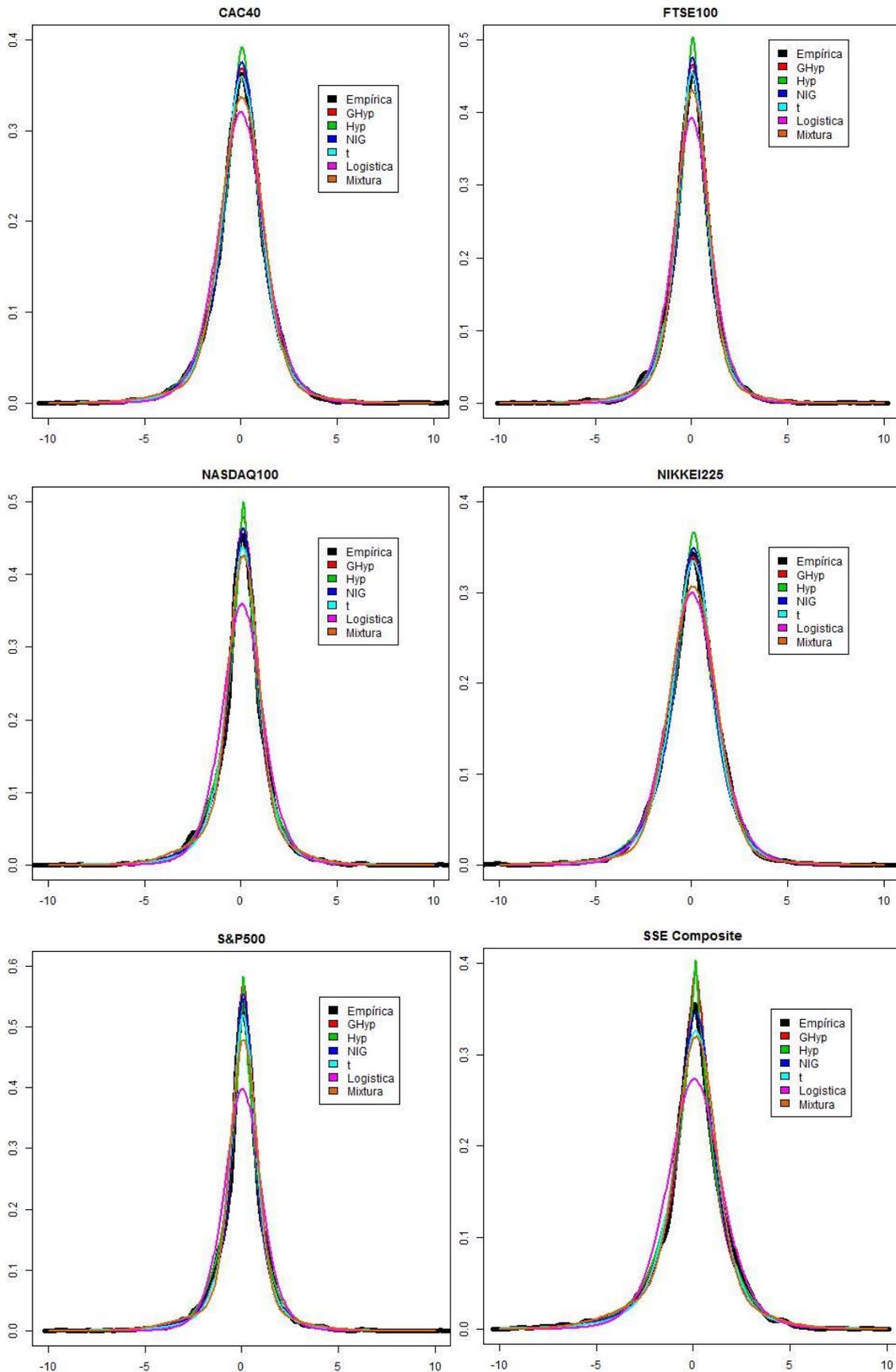
donde s es la desviación típica muestral y R es el rango intercuartílico. El primero de ellos se recomienda cuando se usa el kernel K Gaussiano, mientras que el segundo tiene la ventaja de ser menos sensible a la presencia de datos atípicos.

El estimador kernel de la densidad lo calcularemos mediante R , que utiliza un kernel K Gaussiano por defecto. Además, en el Gráfico 2.2 hemos visto la enorme presencia de datos atípicos en las distribuciones de rendimientos diarios. Por estos dos motivos calcularemos el ancho de ventana óptimo por los dos procedimientos (4) y (5). El ancho de banda óptimo calculado por la expresión (5) refleja unos mejores resultados que el calculado por (4), por lo que para los posteriores gráficos de la función de densidad kernel utilizaremos los anchos de banda calculados por (5).

En el Gráfico 4.4 se representa la densidad kernel correspondiente a cada índice y las distribuciones teóricas correspondientes con los parámetros estimados de la tabla 4.1. Para la elaboración de dicho Gráfico se ha utilizado el software R junto con las librerías “ghyp”, “KernSmooth”, y “RColorBrewer”.

En el Gráfico 4.4 podemos observar que la distribución logística y la distribución mixtura gaussiana se ajustan peor a la distribución empírica que la distribución hiperbólica generalizada y sus casos particulares. Esto se debe a que la logística y la mixtura gaussiana no consiguen ajustarse a la elevada curtosis existente en la distribución de rendimientos. Similares conclusiones se obtienen en Peiró (1992) y Aparicio y Estrada (2001). Destacamos también que la distribución *Skew t* se ajusta peor para los índices NASDAQ 100 y SSE Composite, ya que no consigue captar la elevada curtosis de la distribución empírica. También podemos ver que la alta curtosis que posee la distribución hiperbólica (simple) provoca que, en el pico de la campana de todas las distribuciones empíricas, se ajuste peor que el resto de distribuciones de la familia GH. Mientras que la distribución hiperbólica generalizada se ajusta bien a todos los índices excepto a S&P 500 y SSE Composite, la Normal Inversa Gaussiana consigue adaptarse correctamente a la distribución empírica en todos los casos. Otra conclusión a la que podríamos llegar es que en la mayoría de los casos se modeliza mejor la cola derecha que la izquierda.

Gráfico 4.4: Comparativa de las densidades kernel de los índices y las densidades teóricas estimadas.



A través del Gráfico 4.4 podemos observar la bondad del ajuste en el pico central de la campana de la distribución, pero no podemos sacar una conclusión clara del ajuste en las colas. Para ello, en el Gráfico 4.5 haremos una representación similar en escala logarítmica, comparando el logaritmo de las funciones de densidad teóricas estudiadas, $\log \hat{f}(x)$, y el logaritmo de la función de densidad empírica de los rendimientos diarios, $\log f_T(x)$. Mediante este método conseguimos “poner una lupa en las colas” ya que gracias al cálculo de los logaritmos se hace visible las diferencias en aquellas partes de la función de densidad donde ésta toma valores pequeños, que son precisamente las colas; véase Behr y Pötter (2009) y Necula (2009). El Gráfico 4.5 se ha obtenido utilizando el software R con las mismas librerías del Gráfico 4.4.

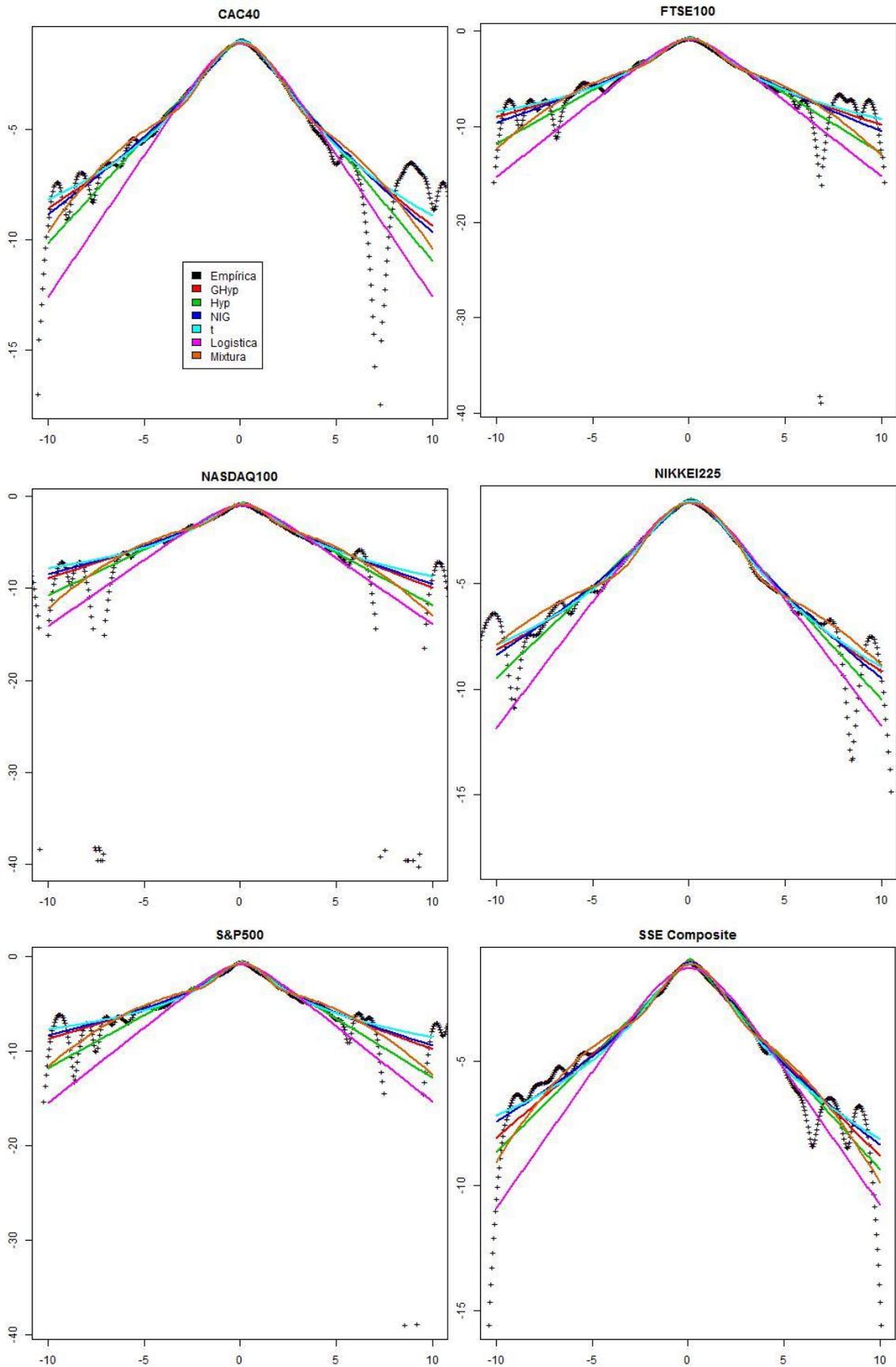
Visualizando el comportamiento de las colas para los rendimientos diarios (Gráfico 4.5), podemos observar que la distribución logística nos proporciona la peor aproximación. Además, tanto la mixtura gaussiana como la distribución hiperbólica (simple) no parecen ser apropiadas para representar las colas de la distribución empírica. El resto de distribuciones nos proporcionan una mejor aproximación. Estos resultados concuerdan con los resultados de otros estudios sobre el comportamiento de las colas de la distribución de los rendimientos diarios; véase, por ejemplo, Behr y Pötter (2009) y Necula (2009).

Para confirmar estas conclusiones, procedemos a realizar formalmente los contrastes de bondad de ajuste para las distribuciones teóricas expuestas.

4.3. Contrastes de bondad de ajuste

En esta sección aplicaremos los contrastes de Kolmogorov-Smirnov (KS) y Anderson-Darling (AD), explicados en la sección 3.2, para analizar la bondad del ajuste que presentan las distribuciones teóricas expuestas anteriormente sobre las distribuciones empíricas de rendimientos diarios. Para ello, utilizaremos el software R junto con los paquetes “ghyp” y “ADGofTest”. En esta sección, el modelo bajo H_0 no es la $N(\mu, \sigma)$, como en la sección 3.2, sino que, en cada caso, será el modelo teórico correspondiente. En este caso, los valores críticos del test KS, ya no son los de Lilliefors, y ahora suelen ser más conservadores.

Gráfico 4.5: Comparativa del comportamiento de colas entre la distribución empírica y las distribuciones teóricas para los rendimientos diarios.



Los resultados del test KS se recogen en la Tabla 4.2. Las conclusiones más relevantes que se extraen de esta tabla son las siguientes: primero, observamos que todas las distribuciones expuestas se adaptan mejor a la distribución empírica que la distribución normal (compárense los p-valores con los correspondientes de la Tabla 3.1). En segundo lugar, podemos comprobar que, en todos los índices se rechaza la hipótesis nula de que los rendimientos siguen una distribución logística, para un nivel de significación del 5%, mientras que para ese mismo nivel de significación, el resto de los modelos no se rechazan, en general, excepto para el índice SSE Composite que se rechaza la mixtura gaussiana, y para el S&P 500, que se rechaza la distribución hiperbólica (simple) y la mixtura gaussiana. Finalmente, observamos que la distribución que mejor se adapta a los índices CAC 40, FTSE 100 y NIKKEI 225 parece ser la distribución Normal Inversa Gaussiana, mientras que para los otros 3 índices, es la Hiperbólica Generalizada.

Tabla 4.2: Contraste de Kolmogorov-Smirnov para los rendimientos diarios.

		CAC 40	FTSE 100	NASDAQ 100	NIKKEI 225	S&P 500	SSE
ghyp	estadístico	0,009	0,010	0,011	0,013	0,012	0,016
	(p-valor)	(0,983)	(0,968)	(0,937)	(0,819)	(0,883)	(0,535)
hyp	estadístico	0,009	0,013	0,023	0,012	0,031	0,024
	(p-valor)	(0,987)	(0,805)	(0,155)	(0,902)	(0,015)	(0,111)
NIG	estadístico	0,008	0,010	0,013	0,009	0,014	0,019
	(p-valor)	(0,999)	(0,971)	(0,816)	(0,982)	(0,730)	(0,329)
Skew t	estadístico	0,011	0,012	0,019	0,014	0,020	0,026
	(p-valor)	(0,933)	(0,875)	(0,320)	(0,696)	(0,254)	(0,078)
mixtura	estadístico	0,018	0,015	0,019	0,027	0,029	0,028
	(p-valor)	(0,397)	(0,634)	(0,322)	(0,057)	(0,032)	(0,047)
logística	estadístico	0,030	0,033	0,054	0,033	0,061	0,048
	(p-valor)	(0,022)	(0,009)	(0,000)	(0,011)	(0,000)	(0,000)

Finalmente, en la Tabla 4.3 se presentan los resultados del contraste de Anderson Darling para las mismas distribuciones teóricas.

Como podemos comprobar, ambos contrastes (KS y AD) presentan unos resultados similares. La distribución Hiperbólica Generalizada y sus casos particulares modelizan de manera aceptable los rendimientos diarios, especialmente, las distribuciones que mejor se ajustan son la Normal Inversa

Gaussiana y la Hiperbólica Generalizada; en cambio la distribución Mixtura Gaussiana y la Logística no gozan de ese grado de ajuste, esta última se rechaza con ambos contrastes para todos los índices analizados.

Tabla 4.3: Contraste de Anderson Darling para los rendimientos diarios.

		CAC 40	FTSE 100	NASDAQ 100	NIKKEI 225	S&P 500	SSE
ghyp	estadístico	0,216	0,160	0,275	0,397	0,370	0,992
	(p-valor)	(0,986)	(0,998)	(0,955)	(0,852)	(0,878)	(0,362)
hyp	estadístico	0,346	0,616	1,035	0,610	2,925	1,578
	(p-valor)	(0,900)	(0,633)	(0,339)	(0,638)	(0,030)	(0,159)
NIG	estadístico	0,169	0,170	0,404	0,393	0,506	1,048
	(p-valor)	(0,997)	(0,997)	(0,845)	(0,856)	(0,740)	(0,333)
Skew t	estadístico	0,354	0,283	1,173	0,400	1,376	1,720
	(p-valor)	(0,893)	(0,951)	(0,278)	(0,848)	(0,209)	(0,132)
mixtura	estadístico	1,240	0,810	1,222	1,583	2,631	1,928
	(p-valor)	(0,253)	(0,474)	(0,259)	(0,158)	(0,042)	(0,101)
logística	estadístico	4,421	6,138	11,701	4,074	17,465	11,045
	(p-valor)	(0,005)	(0,001)	(0,000)	(0,008)	(0,000)	(0,000)

Estos resultados concuerdan además con las conclusiones de la sección 4.2. Cabe destacar que, aunque en el Gráfico 4.4 concluíamos que la hiperbólica generalizada no conseguía un óptimo grado de ajuste en los índices S&P 500 y SSE Composite, a la vista de los resultados obtenidos en la Tabla 4.2 y la Tabla 4.3, podemos decir que quizá no se conseguía ese ajuste en esos índices debido a la amplitud de ventana utilizada para dichos índices. Basándonos en la Tabla 4.2 y en la Tabla 4.3, podemos concluir que las distribuciones Normal Inversa Gaussiana y la hiperbólica generalizada son las que mejor se ajustan a la distribución empírica.

Los resultados que acabamos de presentar corroboran los obtenidos por otros autores. Así, Prause (1997) también llegó a la conclusión de que la distribución NIG y la hiperbólica generalizada conseguían un buen grado de ajuste para los rendimientos diarios obtenidos desde Enero de 1988 hasta Mayo de 1994 para el NYSE y el mercado alemán. Eberlein y Prause (1998), examinando los rendimientos diarios del índice DAX correspondientes al periodo 15/12/1993-26/11/1997, concluyeron que la distribución que mejor se ajustaba a estos rendimientos era la hiperbólica generalizada, y tras ésta, la

distribución NIG. Behr y Pötter (2009) determinaron que la distribución hiperbólica generalizada era la distribución que proporcionaba un mejor ajuste en los rendimientos diarios del S&P 500 para el periodo 2001-2005. Necula (2009) concluyó que la distribución hiperbólica generalizada modelizaba correctamente los rendimientos diarios para los índices S&P 500, NIKKEI 225, Hang Seng, DAX, FTSE 100, CAC 40, IBEX 35, PX 50, BUX y BETC, entre Enero de 1998 y Septiembre de 2008.

En resumen, parece que la distribución hiperbólica generalizada y la Normal Inversa Gaussiana se ajustan correctamente a los rendimientos diarios de diversos índices bursátiles sea cuál sea el periodo temporal estudiado, ya que ambas distribuciones consiguen un buen grado de ajuste incluso cuando se produce un gran “crash” en los mercados bursátiles, como el “crash” de Septiembre de 2008.

5. SENSIBILIDAD DE LOS RESULTADOS A LA FRECUENCIA DE OBSERVACIÓN: ANÁLISIS DE DATOS INTRADIARIOS

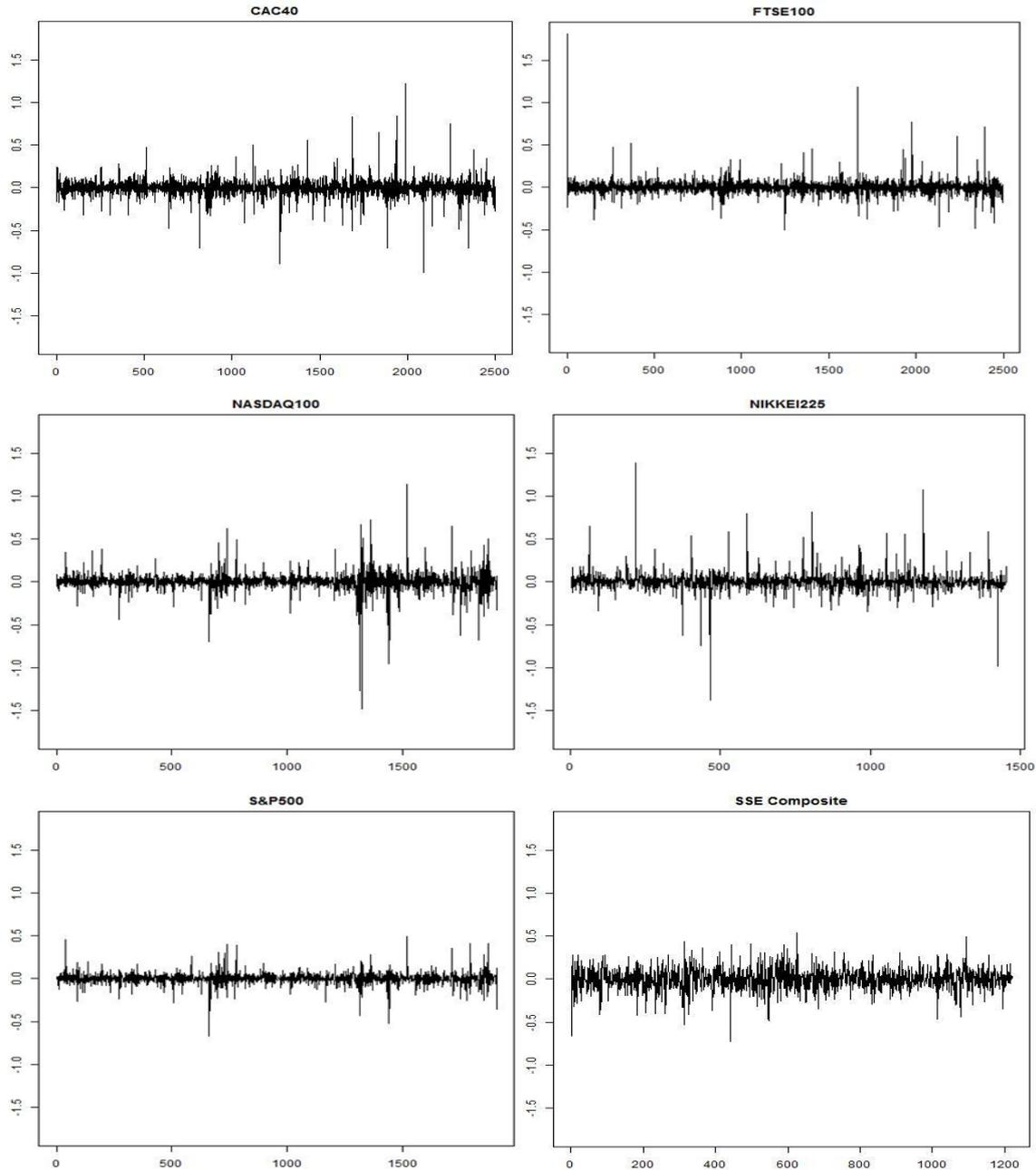
5.1. Descripción de los datos intradiarios: definición y propiedades.

En esta sección, utilizando el software R, analizaremos los rendimientos intradiarios de los mismos índices bursátiles estudiados con anterioridad, es decir, CAC 40, FTSE 100, NASDAQ 100, NIKKEI 225, S&P 500 y SSE Composite. Para ello, analizaremos las cotizaciones consideradas con una frecuencia de 10 minutos (obtenidas de Google Finance) durante el periodo de 23/04/2017 hasta el 02/07/2017, por lo que se dispone de una muestra de 2.499 valores para el índice CAC 40, 2.496 para el FTSE 100, 1.911 para el NASDAQ 100, 1.457 para el NIKKEI 225, 1.911 para el S&P 500 y 1.222 para el SEE Composite.

A partir de estas cotizaciones calcularemos los rendimientos de cada índice mediante diferencia logarítmica, por las razones expuestas anteriormente. El Gráfico 5.1 incluye una representación gráfica de los rendimientos históricos intradiarios (con frecuencia de 10 minutos) de nuestras 6 series para el periodo 23/04/2017 - 02/07/2017. Como podemos ver, a diferencia de los rendimientos diarios, los rendimientos intradiarios no

presentan representaciones similares, es decir, en cortos periodos de tiempo, no se aprecia el efecto contagio entre los mercados bursátiles.

Gráfico 5.1: Rendimientos intradiarios de los índices bursátiles.



Para analizar con más detalle la distribución marginal de los rendimientos intradiarios, utilizaremos las mismas medidas de posición central, dispersión y forma que en la sección 2.2. La Tabla 5.1 recoge los valores de estos estadísticos para los rendimientos intradiarios de las 6 series analizadas, así como el valor mínimo y máximo que toman estas series en el periodo

muestral. Para su elaboración hemos utilizado el software R junto con la librería “moments”.

Tabla 5.1: Estadísticos de los rendimientos intradiarios.

	\bar{x}	Me	s	Min	Max	g1	g1*	g2	g2*
CAC 40	-0,001	0,000	0,102	-0,990	1,223	0,176	3,582*	26,760	242,401*
FTSE 100	0,001	0,001	0,086	-0,499	1,807	5,245	106,953*	102,089	1010,306*
NASDAQ 100	0,001	0,004	0,115	-1,480	1,138	-1,577	-28,128*	37,814	310,569*
NIKKEI 225	0,004	0,003	0,116	-1,375	1,384	0,872	13,580*	45,018	327,272*
S&P 500	0,001	0,003	0,067	-0,665	0,490	-0,286	-5,111*	18,069	134,431*
SSE	0,001	0,001	0,136	-0,723	0,540	-0,338	-4,823*	5,204	15,722*

Las principales conclusiones que se extraen de la Tabla 5.1 son las siguientes: los 6 índices estudiados presentan unas medidas de posición central prácticamente iguales a 0. En cuanto a la desviación típica, podemos ver que para el SEE Composite es mayor que para el resto de índices, al igual ocurría en los rendimientos diarios. Los valores mínimo y máximo para los rendimientos intradiarios distan mucho de los valores que se alcanzaban en los rendimientos diarios para los correspondientes índices.

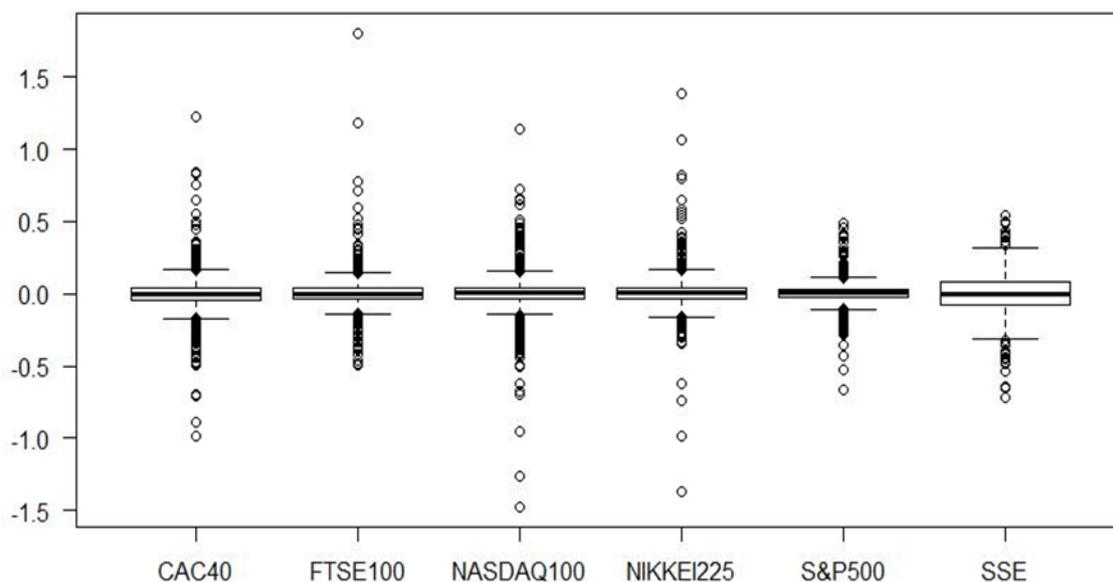
En cuanto a la asimetría, 3 índices presentan asimetría negativa (NASDAQ 100, S&P 500 y SSE Composite), mientras que otros 3 presentan asimetría positiva (CAC 40, FTSE 100, y NIKKEI 225), esto supone una diferencia con respecto a los rendimientos diarios, ya que todos presentaban asimetría negativa. Los coeficientes estandarizados de asimetría para los rendimientos intradiarios indican que la asimetría observada es significativa para todos los índices.

Finalmente, en la Tabla 5.1 podemos ver que todas las distribuciones son leptocúrticas. Además los coeficientes estandarizados de curtosis muestran que la curtosis es significativa para todos los índices. La curtosis observada en los rendimientos intradiarios es mayor que la curtosis observada en los rendimientos diarios para los correspondientes índices, excepto para el SSE Composite, que presenta una curtosis menor.

Para completar este análisis, en el Gráfico 5.2 representamos los boxplots de las distribuciones marginales para los rendimientos intradiarios. Estos gráficos corroboran los resultados que acabamos de comentar. Además,

ponen de manifiesto que en todos los índices encontramos un gran número de valores atípicos excepto en el SSE Composite.

Gráfico 5.2: Boxplots de las rentabilidades intradiarias



5.2. Estimación de modelos y contrastes de bondad de ajuste

En esta sección se analizará si la distribución normal, las distribuciones de la familia GH, la distribución logística y la mixtura gaussiana, utilizadas en la sección 4, se ajustan a los rendimientos intradiarios con frecuencia de 10 minutos. Para ello, en primer lugar estimaremos estas distribuciones teóricas mediante el método de máxima verosimilitud. Los parámetros estimados para las distribuciones estudiadas se muestran en la Tabla 5.2.

Las primeras conclusiones que podemos obtener en la Tabla 5.2 son las siguientes. En primer lugar, en el parámetro β encontramos la primera diferencia con respecto a los rendimientos diarios, ya que entonces éste era negativo siempre. Sin embargo, ahora el parámetro β de las distribuciones de la familia GH tiene valor positivo en los índices FTSE 100 y NIKKEI 225, índices que presentaban asimetría positiva. Mientras que es negativo en los índices que presentaban asimetría negativa (NASDAQ 100, S&P 500 y SSE Composite) y en el CAC 40, a pesar de que este último presentaba asimetría positiva en la sección 5.1.

Tabla 5.2: Parámetros estimados de los rendimientos intradiarios.

		CAC 40	FTSE 100	NASDAQ 100	NIKKEI 225	S&P 500	SSE
Normal	μ	-0,001	0,001	0,001	0,004	0,001	0,001
	σ	0,102	0,086	0,115	0,116	0,067	0,136
ghyp	λ	-1,390	-1,515	-0,896	-1,134	-1,224	-2,586
	α	0,540	0,511	1,512	0,308	2,292	3,096
	δ	0,093	0,078	0,062	0,074	0,056	0,255
	β	-0,348	0,177	-0,344	0,207	-0,693	-0,241
	μ	0,002	0,000	0,006	0,001	0,004	0,004
hyp	λ	1,000	1,000	1,000	1,000	1,000	1,000
	α	16,062	20,700	15,547	15,823	23,990	13,360
	δ	0,009	0,011	0,001	0,000	0,005	0,119
	β	-0,440	0,296	-0,764	0,257	-1,401	-0,584
	μ	0,002	0,000	0,008	0,002	0,006	0,011
NIG	λ	-0,500	-0,500	-0,500	-0,500	-0,500	-0,500
	α	6,710	8,875	4,104	4,586	9,532	9,793
	δ	0,060	0,049	0,047	0,048	0,039	0,179
	β	-0,379	0,173	-0,403	0,234	-0,870	-0,589
	μ	0,002	0,000	0,006	0,002	0,005	0,012
Skew t	ν	2,798	3,018	2,018	2,277	2,614	5,996
	δ	0,093	0,077	0,067	0,074	0,059	0,272
	β	-0,335	0,144	-0,267	0,201	-0,626	-0,661
	μ	0,002	0,000	0,006	0,001	0,004	0,013
mixtura	p	0,919	0,953	0,859	0,934	0,875	0,867
	$1-p$	0,081	0,047	0,141	0,066	0,125	0,133
	$\mu 1$	0,002	0,000	0,005	0,000	0,003	0,006
	$\mu 2$	-0,030	0,019	-0,021	0,056	-0,012	-0,035
	$\sigma 1$	0,065	0,053	0,054	0,063	0,040	0,111
	$\sigma 2$	0,282	0,315	0,273	0,381	0,156	0,240
logística	μ	0,000	0,000	0,004	0,002	0,002	0,002
	σ	0,046	0,036	0,048	0,047	0,031	0,074

También podemos observar que los parámetros δ , β , y μ de la distribución *Skew t* son muy similares a estos mismos parámetros en la distribución hiperbólica generalizada (ghyp) para los índices correspondientes, a excepción de los índices NASDAQ 100 y SSE Composite. Además, si tenemos en cuenta que en la distribución *Skew t* $\nu = -2\lambda$, también llegamos a la conclusión que el parámetro λ es muy similar para ambas distribuciones, salvo en los índices NASDAQ 100 y SSE Composite. Por lo que podemos concluir que la forma que tomarán las distribuciones *Skew t* e hiperbólica generalizada será muy parecida para los índices CAC 40, FTSE 100, NIKKEI 225 y S&P 500.

En cuanto a la distribución mixtura gaussiana, la Tabla 5.2 muestra para los índices FTSE 100 y NIKKEI 225 una asimetría positiva, mientras que para el resto de los índices muestra una asimetría negativa.

En el Gráfico 5.3 se representa la función de densidad kernel de los rendimientos intradiarios para cada índice, utilizando el ancho de ventana óptimo dado por la expresión (5), junto con las correspondientes funciones de densidad de las distribuciones teóricas estimadas.

Las conclusiones que obtenemos de este gráfico son las siguientes: para todos los índices, excepto para el SSE Composite, las distribuciones teóricas que mejor se ajustan a la empírica son la *Skew t* y la hiperbólica generalizada, que consiguen un buen grado de ajuste tanto en el centro como en las colas. La distribución hiperbólica simple y la Normal Inversa Gaussiana no consiguen modelizar correctamente el pico de la campana ya que presentan una curtosis mayor que ésta. Mientras que la distribución logística y la mixtura gaussiana no consiguen captar la elevada curtosis presente en la distribución empírica, además la mixtura gaussiana presenta un mal ajuste en la cola izquierda en los índices CAC 40, FTSE 100, NASDAQ 100 y NIKKEI 225. Como era de esperar, la distribución normal nos da el peor ajuste de las distribuciones teóricas. Además, podemos observar que la distribución del índice SSE Composite es claramente diferente al resto: es menos picuda y más ancha. Esto concuerda con las conclusiones a las que llegamos en el apartado 5.1, en el que veíamos que el SSE Composite tenía menos curtosis y menos valores atípicos que el resto de índices. Esto hace que todas las distribuciones teóricas a excepción de la normal, consigan un grado de ajuste similar y aceptable en este índice.

Finalmente, en el Gráfico 5.4 compararemos gráficamente los logaritmos de las densidades teóricas y de la distribución empírica de los rendimientos intradiarios para determinar el ajuste en las colas. De este gráfico podemos concluir que tanto la distribución *Skew t*, como la hiperbólica generalizada, proporcionan un buen ajuste en las colas (tanto derecha como izquierda) de la distribución empírica. La distribución Normal Inversa Gaussiana también proporciona un óptimo ajuste, pero en menor medida que las anteriores. El resto de distribuciones no parecen ser apropiadas para representar las colas de la distribución empírica, especialmente la distribución normal, que nos proporciona el peor ajuste de todas las distribuciones teóricas.

Gráfico 5.3: Comparativa de las densidades kernel de los rendimientos intradiarios y las densidades teóricas estimadas

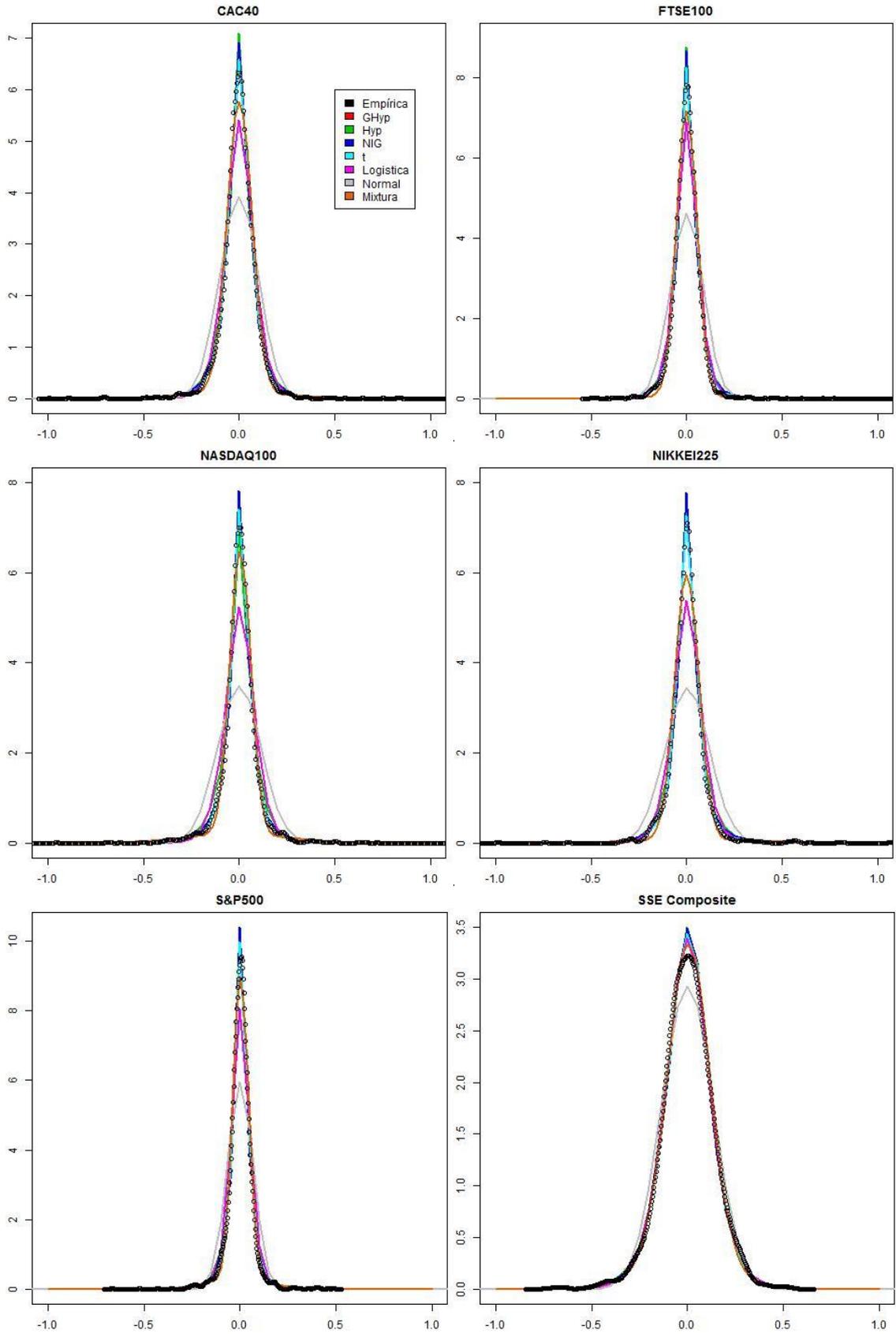
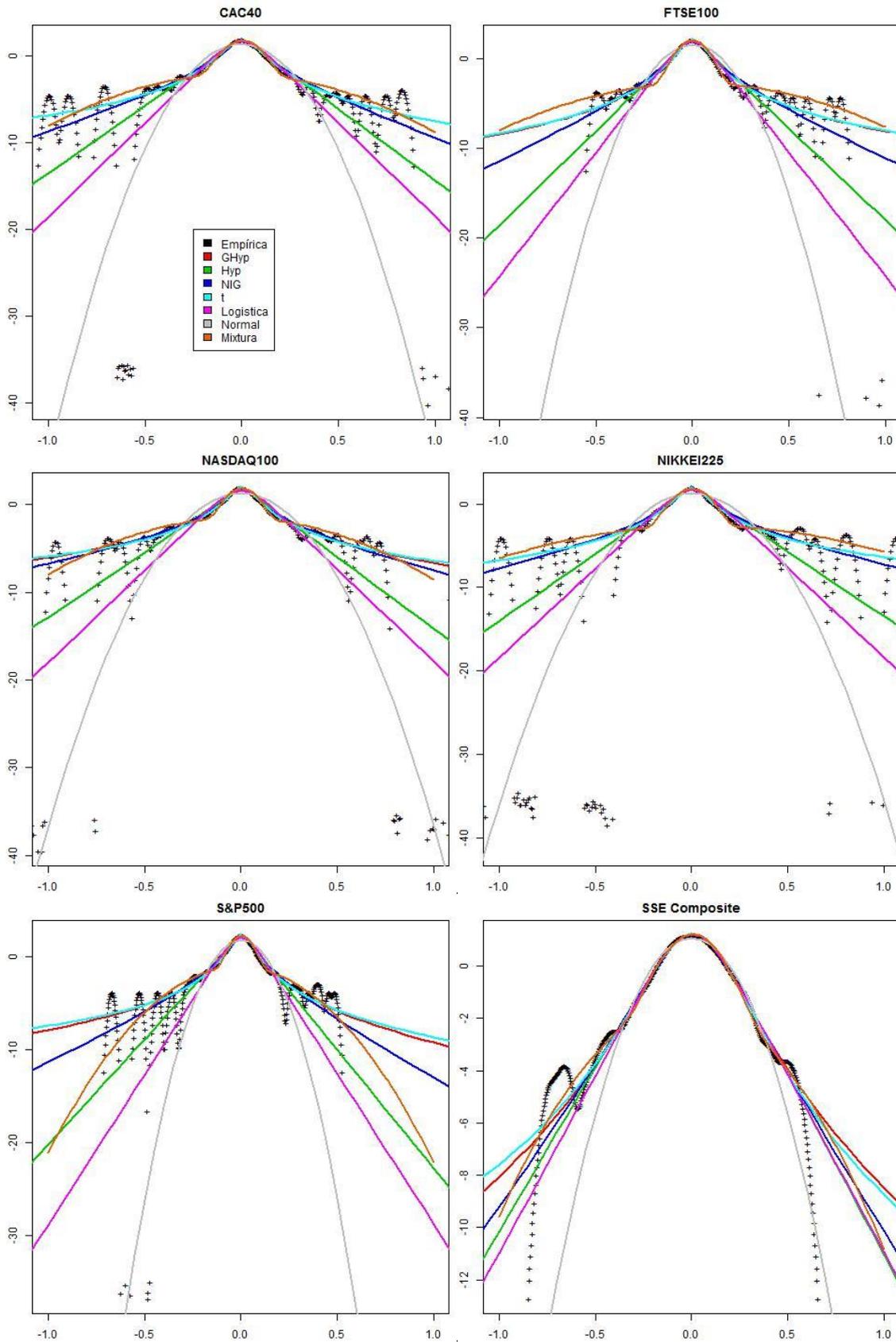


Gráfico 5.4: Comparativa del comportamiento de las colas entre la distribución empírica y las distribuciones teóricas para los rendimientos diarios.



Tras esto, realizaremos los contrastes de bondad de ajuste para las distribuciones teóricas expuestas. Para ello, al igual que en la sección 4.3, utilizaremos los contrastes de Kolmogorov-Smirnov y de Anderson Darling explicados en la sección 3.2, utilizando la versión de Lilliefors en el contraste KS cuando el modelo bajo H_0 sea la distribución normal. Para este análisis, el modelo bajo H_0 será el modelo teórico correspondiente. Para la elaboración del contraste KS hemos utilizado el software R junto con la librería “nortest”.

Tabla 5.3: Contraste de Kolmogorov-Smirnov para los rendimientos intradiarios

		CAC 40	FTSE 100	NASDAQ 100	NIKKEI 225	S&P 500	SSE
Normal	estadístico	0,11	0,12	0,15	0,15	0,10	0,04
	(p-valor)	(0,00)	(0,00)	(0,00)	(0,00)	(0,00)	(0,00)
ghyp	estadístico	0,012	0,014	0,012	0,017	0,011	0,016
	(p-valor)	(0,886)	(0,728)	(0,939)	(0,793)	(0,973)	(0,931)
hyp	estadístico	0,025	0,026	0,042	0,045	0,024	0,017
	(p-valor)	(0,080)	(0,067)	(0,002)	(0,005)	(0,293)	(0,871)
NIG	estadístico	0,014	0,021	0,015	0,024	0,015	0,016
	(p-valor)	(0,704)	(0,237)	(0,786)	(0,389)	(0,763)	(0,903)
Skew t	estadístico	0,012	0,013	0,011	0,017	0,011	0,016
	(p-valor)	(0,877)	(0,768)	(0,971)	(0,794)	(0,976)	(0,930)
mixtura	estadístico	0,026	0,021	0,027	0,042	0,018	0,015
	(p-valor)	(0,065)	(0,211)	(0,135)	(0,012)	(0,539)	(0,932)
logística	estadístico	0,043	0,032	0,067	0,060	0,043	0,015
	(p-valor)	(0,000)	(0,010)	(0,000)	(0,000)	(0,002)	(0,942)

La Tabla 5.3 proporciona conclusiones similares al Gráfico 5.3 y al Gráfico 5.4. Para todos los índices, excepto para el SSE Composite, las distribuciones que mejor se ajustan a la distribución empírica son la hiperbólica generalizada y la *Skew t* (para estos modelos no se rechaza la hipótesis nula para un nivel de significación del 5% en ningún índice). Para todos los índices (excepto el SSE Composite) se rechaza al 5% la hipótesis nula de que la distribución empírica siga una distribución logística o una distribución normal. Cabe destacar que el contraste de KS no permite rechazar el modelo NIG para este nivel de significación (5%) en ningún índice. Para el índice SSE Composite, todas las distribuciones teóricas, a excepción de la distribución normal, se ajustan correctamente a la distribución empírica (especialmente la hiperbólica generalizada, *Skew t*, mixtura gaussiana y logística).

A estas mismas conclusiones se llega analizando los resultados del contraste de Anderson-Darling recogidos el Anexo (Tabla A1), que también concuerdan con las conclusiones obtenidas del Gráfico 5.3 y del Gráfico 5.4.

Si comparamos estos resultados con los resultados obtenidos en la sección 4.3 para los rendimientos diarios, nos encontramos con la similitud de que los rendimientos para frecuencia diaria y con una frecuencia de 10 minutos, se pueden modelar correctamente con la distribución hiperbólica generalizada (para todos los índices), y que la distribución normal no consigue ajustarse a ambos rendimientos para ningún índice.

6. CONCLUSIONES

En el presente trabajo hemos estudiado el supuesto de normalidad de los rendimientos bursátiles. Para ello, hemos considerado 6 índices, 2 de ellos son europeos (CAC 40 y FTSE 100), otros 2 son americanos (NASDAQ 100 y S&P 500), y los 2 últimos son asiáticos (NIKKEI 225 y SSE Composite).

En primer lugar, hemos realizado varios contrastes de normalidad para los rendimientos diarios de los 6 índices, que nos han conducido a rechazar la hipótesis de distribución gaussiana para todos los índices debido a que la distribución marginal de los rendimientos diarios posee un gran grado de apuntamiento y colas pesadas. En consecuencia, hemos ajustado las distribuciones de la familia GH, la distribución mixtura gaussiana y la logística, a las mismas series. Así, hemos encontrado que todas estas distribuciones teóricas se ajustaban mejor a los rendimientos diarios que la distribución normal, especialmente, las distribuciones que han logrado un mejor ajuste han sido la hiperbólica generalizada y la Normal Inversa Gaussiana.

Hemos repetido este análisis para los rendimientos intradiarios con una frecuencia de 10 minutos, obteniendo unas conclusiones similares: se rechaza la hipótesis de normalidad; todas las distribuciones teóricas analizadas se ajustan mejor que la distribución normal; y la distribución hiperbólica generalizada y *Skew t*, logran un buen grado de ajuste en todos los índices.

Por consiguiente, podemos concluir que la distribución hiperbólica generalizada es apropiada para modelizar los rendimientos bursátiles con un frecuencia pequeña (en este caso con frecuencia diaria y de 10 minutos).

7. BIBLIOGRAFÍA

- Aas, K. y Haff, I. H. (2005): «The Generalized Hyperbolic Skew Student's t-Distribution». *Journal of Financial Econometrics*, 4(2), pp. 275-309.
- Aparicio, F. M. y Estrada, J. (2001): «Empirical distributions of stock returns: European securities markets, 1990-1995». *The European journal of Finance*, 7, pp. 1-21.
- Bachelier, L. (1900): «Theorie de la Speculation». *Annales de l'Ecole Normale Supérieure*, 3, pp. 21-86.
- Barndorff-Nielsen, O. (1977): «Exponentially decreasing distributions for the logarithm of particle size». *Proceedings of the Royal Society London*, 353, pp. 401-419.
- Behr, A. y Pötter, U. (2009): «Alternatives to the normal of stock returns: Gaussian mixture, generalised logF and generalised hyperbolic models». *Annals of Finance*, 5, pp. 49-68.
- Blattberg, R. C. y Gonedes, N. J. (1974): «A Comparison of the Stable and Student Distribution as Statistical Models for Stock Prices». *Journal of Business*, 47, pp. 247-280.
- Clark, P. K. (1973): «A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices». *Econometrica*, 41, pp. 135-155.
- Eberlein, E., y U. Keller, (1995): «Hyperbolic distributions in finance». *Bernoulli*, 1, pp. 281-299.
- Eberlein, E., y Prause, K. (1998): «The Generalized hyperbolic Model: Financial Derivates and Risk Measures». FDM-Preprint 56, University of Freiburg.
- Fama, E. F. (1965): «The Behaviour of Stock Market Prices». *Journal of Business*, 38, pp. 34-105.
- Gray, J. B., y French, D. W. (1990): «Empirical Comparisons of Distributonal Models for Stock Index Returns». *Journal of Business Finance & Accounting*, 17, pp. 451-459.
- Hammerstein, E. A. (2010): «Generalized hyperbolic distributions: Theory and applications to CDO pricing». PhD thesis, Universidad de Freiburg.
- Harris, R. y Küçüközmen, C. (2001): «The Empirical Distribution of UK and US Stock Returns». *Journal of Business Finance & Accounting*, 28, pp. 715-740.
- Hurst, S. R. y Platen, E. (1997): «The Marginal Distributions of Returns and Volatility». *Lecture Notes-Monograph Series*, 31, pp. 301-314.
- Kon, S. J. (1984): «Models of Stock Returns-A comparison». *Journal of Finance*, 39, pp. 147-165.
- Mandelbrot, B. (1963): «The Variation of Certain Speculative Prices». *Journal of Business*, 36, pp. 394-419.
- Marques, M. J. (2001): *Estadística Básica: un enfoque no paramétrico*. Editorial FES-Zaragoza. UNAM, México.

- Martínez, E. y Arnillas, M. L. (1992): «Normalidad de los rendimientos bursátiles». IESE. DI-237.
- Necula, C. (2009): «Modeling Heavy-Tailed Stock Index Returns Using the Generalized Hyperbolic Distribution». Romanian Journal of Economic Forecasting, 2, pp. 118-131.
- Peiró, A (1992): «Distribución de los rendimientos de acciones». Estadística Española, 34, pp. 431-453.
- Peiró, A. (1997): «La propiedad de simetría en los rendimientos financieros españoles». Estadística Española, 39, pp. 41-65.
- Peña, D. (2014): Fundamentos de Estadística. Editorial Alianza, Madrid.
- Praetz, P. D. (1972): «The Distribution of Share Price Changes». Journal of Business, 45, pp. 49-55.
- Prause, K. (1997): «Modelling financial data using generalized hyperbolic distributions». FDM Preprint 48, University of Freiburg.
- Prause, K., (2000): «The Generalized Hyperbolic Model: Estimation, Financial Derivates, and Risk Measures». PhD. Dissertation, University of Freiburg.
- R Package 'ADGofTest'*. [package]. Versión 0.3. CRAN. Fecha de publicación: 19 de Febrero de 2015. Disponible en: <https://cran.r-project.org/web/packages/ADGofTest/ADGofTest.pdf>
- R Package 'fitdistrplus'* [package]. Versión 1.0-9. CRAN. Fecha de publicación: 24 de Marzo de 2017. Disponible en: <https://cran.r-project.org/web/packages/fitdistrplus/fitdistrplus.pdf>
- R Package 'ghyp'*. [package]. Versión 1.5.7. CRAN. Fecha de publicación: 17 de Agosto de 2016. Disponible en: <https://cran.r-project.org/web/packages/ghyp/ghyp.pdf>
- R Package 'KernSmooth'*. [package]. Versión 2.23-15. CRAN. Fecha de publicación: 29 de Junio d 2015. Disponible en: <https://cran.r-project.org/web/packages/KernSmooth/KernSmooth.pdf>
- R Package 'mixtools'*. [package]. Versión 1.1.0. CRAN. Fecha de publicación: 10 de Marzo de 2017. Disponible en: <https://cran.r-project.org/web/packages/mixtools/mixtools.pdf>
- R Package 'moments'*. [package]. Versión 0.14. CRAN. Fecha de publicación: 20 de Febrero de 2015. Disponible en: <https://cran.r-project.org/web/packages/moments/moments.pdf>
- R. Package 'nortest'*. [package]. Versión 1.0-4. CRAN. Fecha de publicación: 30 de Julio de 2015. Disponible en: <https://cran.r-project.org/web/packages/nortest/nortest.pdf>
- R Package 'RColorBrewer'*. [package]. Versión 1.1-2. CRAN. Fecha de publicación: 19 de Febrero de 2015. Disponible en: <https://cran.r-project.org/web/packages/RColorBrewer/RColorBrewer.pdf>
- Silverman, B. W. (1990): Density Estimation for Statistics and Data Analysis. 2ª ed. Chapman and Hall, Londres.

Smith, J. B. (1981): «The Probability Distribution of Market Returns: A Logistic Hypothesis». PhD dissertation. University of Utah.

Törnqvist, L., Vartia, P. y Vartia Y. O. (1985): «How Should Relative Changes be Measured?». *The American Statistician*, 39, pp. 43-496.

ANEXO

Tabla A.1: Contraste de Anderson Darling para los rendimientos intradiarios.

		CAC 40	FTSE 100	NASDAQ 100	NIKKEI 225	S&P 500	SSE
Normal	estadístico	Inf	Inf	Inf	Inf	49,485	3,735
	(p-valor)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,012)
ghyp	estadístico	0,393	0,676	0,162	0,517	0,259	0,316
	(p-valor)	(0,855)	(0,579)	(0,998)	(0,729)	(0,965)	(0,926)
hyp	estadístico	2,449	-	5,986	3,835	2,005	0,358
	(p-valor)	(0,053)	(-)	(0,001)	(0,011)	(0,091)	(0,890)
NIG	estadístico	1,026	1,869	0,495	1,174	0,576	0,325
	(p-valor)	(0,344)	(0,109)	(0,752)	(0,278)	(0,671)	(0,918)
Skew t	estadístico	0,394	0,642	0,190	0,513	0,264	0,266
	(p-valor)	(0,855)	(0,609)	(0,993)	(0,734)	(0,962)	(0,961)
mixtura	estadístico	2,300	2,140	2,013	3,364	1,101	0,282
	(p-valor)	(0,063)	(0,077)	(0,090)	(0,018)	(0,308)	(0,951)
logística	estadístico	8,920	Inf	19,930	11,421	8,636	0,282
	(p-valor)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,951)