Midiendo el riesgo en mercados financieros

Alfonso Novales
Departamento de Economía Cuantitativa
Universidad Complutense

Octubre 2017
Versión preliminar
No citar sin permiso del autor

Contents

1 La medición del riesgo inherente a un activo

1.1 Distintos tipos de riesgo financiero

La medición del riesgo incorporado en un determinado activo es, sin duda, uno de los problemas más importantes de la economía financiera. El nivel de riesgo es una de las características de un activo que, junto con su rentabilidad esperada, su liquidez, etc..determinan las decisiones óptimas de inversión de los agentes. Es habitual identificar la medición del riesgo con la varianza que ofrece la serie temporal de rentabilidad del activo. En el caso de un mercado financiero, el riesgo suele medirse mediante la varianza de las variaciones en el índice correspondiente (rentabilidades) observadas con una determinada frecuencia (hora, día, semana, mes). A veces, utilizamos la varianza con datos de alta frecuencia, como son los datos intradía (dentro del día de negociación); por ejemplo, utilizando las ctoizaciones cada hora, cada 5 minutos, o incluso los precios de todas las operaciones cruzadas.

Sin embargo, pocas veces reflexionamos suficientemente acerca de lo que estamos midiendo. Conviene pensar acerca de qué queremos medir, y si la varianza es una medida adecuada de riesgo.

Además del riesgo-precio o riesgo de reinversión, tenemos el riesgo de mercado, debido a la evolución del mercado en que cotiza nuestro activo durante el plazo de la inversión, el riesgo de liquidez, debido a las potenciales dificultades de deshacernos del activo si así lo deseamos, el riesgo de crédito o de contrapartida, etc.; a nivel institucional tenemos también el riesgo operacional o riesgo operativo, que se deriva de sucesos exogenos: un fallo informático, un incendio, un fraude financiero, etc.. En mercados de renta variable, el *riesgo-beta* que definimos más abajo, es útil para muchos fines. En otras ocasiones, todo lo que queremos es un umbral máximo de pérdidas en la forma de un Valor en Riesgo, es decir, un determinado percentil de la distribución de probabilidad de la rentabilidad esperada de una cartera en un horizonte estipulado previamente. Por tanto, es importante saber qué tipo de riesgo queremos medir en cada caso. Frecuentemente incurriremos en riesgo de modelo, al tener que escoger entre

distintas opciones, una representación paramétrica para la evolución temporal de la rentabilidad de un activo o del vector de rentabilidades de varios activos.

El riesgo de precio en renta fija, se debe al desconocimiento de los tipos de interés futuros a que podremos invertir los cupones recibidos sobre un bono. Hablamos entonces de riesgo precio, o riesgo de reinversión. A igualdad de condiciones, un bono cupón cero tiene un menor componente de riesgo, debido a la ausencia de reinversiones, si bien está sujeto en cualquier caso a riesgo-precio, por cuanto que desconocemos las posibles fluctuaciones que pueda experimentar su precio. En general, el riesgo de precio o de reinversión se produce cuando compramos un activo cuyo vencimiento no coincide con el horizonte de nuestra inversión (tanto si es inferior somo si es un plazo más largo). El riesgo de precio en renta variable puede deberse a la incertidumbre acerca del valor de la empresa percepción debido a acciones recientes, las inversiones que ha asumido, la gestión de sus directivos, etc.. En el caso de una divisa, puede deberse a que un fuerte deterioro de su balanza por cuenta corriente, o de sus cuentas públicas, su situación política, etc., pueden sugerir una posible devaluación, lo que reduciría significativamente la rentabilidad de un inversor extranjero. Una liquidez reducida es otro componente del riesgo específico de un activo, si bien en ocasiones es todo un mercado el que está sujeto a una reducida liquidez. Por ej., la mayor parte de una emisión de deuda privada puede estar en manos de un gran fondo, que no la saca al mercado, por lo que los inversores privados que poseen el resto de la emisión se enfrentan a un riesgo de liquidez.

Por supuesto que un activo de renta variable está sujeto a estas consideraciones, además de las propias de su emisor, por lo que tiene riesgo de mercado o riesgo-precio, riesgo de emisor, etc..

La primera cuestión es que existen distintos tipos de riesgo, que requieren medidas diferentes: *riesgo sistemático o no diversificable* dentro del mercado, *riesgo especifico del activo o riesgo diversificable en el mercado*. Cuando analizamos un mercado concreto, el *Riesgo total* de un activo que cotiza en dicho mercado puede descomponerse en un componente de *Riesgo sistemático o de mercado*, y un componente de *Riesgo específico*. Por ej., las acciones del mercado continuo de Madrid, tienen un componente de riesgo explicado por el propio mercado, representado por el índice. Tienen también un segundo componente de riesgo que no puede explicarse por el riesgo del mercado. Algo similar ocurre con cada una de las referencias que cotiza en el mercado secundario de deuda pública español. También en este caso podríamos hablar de un componente de riesgo global o "de mercado", así como de un componente de riesgo específico de cada índice.

Distinguir entre estos tipos de riesgo y disponer de procedimientos para la estimación de cada uno de ellos es un aspecto importante de la gestión de carteras.El componente de riesgo de mercado es un *riesgo sistemático*, que no puede eliminarse mediante la inversión en activos distintos del mismo mercado. Por eso decimos que dicho riesgo no es diversificable. Este componente no diversificable del riesgo del activo está determinado por la covariación de su rentabilidad con la rentabilidad del índice del mercado al que pertenece.Viene generalmente caracterizado por la beta del activo, que suele estimarse mediante

procedimientos de regresión entre las rentabilidades del activo y del mercado, ambas descontadas de la rentabilidad ofrecida por el activo libre de riesgo. En el caso de la renta variable, este es el modelo CAPM y hablamos del riesgo beta, cuya cuantía viene medida por: a) la beta del activo y, b) la volatilidad del mercado. Por el contrario, el componente de *riesgo específico* mide un riesgo no vinculado al mercado al que pertenece el activo. Este es un riesgo que puede eliminarse por diversificación, si existe una variedad de activos suficientemente rica en el mercado. Este componente del riesgo puede deberse, en unos casos, a las características del emisor, y en otras, a las características técnicas del activo.

1.2 El riesgo en las decisiones de inversión financiera

Disponer de medidas numéricas del nivel de riesgo asociado a la inversión en un determinado activo financiero durante un determinado período de tiempo es una herramienta clave en muchos aspectos de la gestión de carteras. Algunos ejemplos notables son,

Gestión de carteras mediante el análisis rentabilidad/riesgo: Markowitz.

Este enfoque, supone que los inversores tienen preferencias dependientes de dos argumentos: riesgo y rentabilidad esperada. La teoría de carteras muestra que, cuando un inversor se enfrenta a la posibilidad de invertir en dos activos alternativos, no se trata de seleccionar aquel que ofrezca mejores posibilidades. Dada la estructura de preferencias mencionada, podria pensarse que si un activo ofrece mayor rentabilidad (esperada) que otro, y menor varianza, debería ser un activo preferido, con lo que un inversor compraría solo este activo, y nada del otro. Por el contrario, es importante la complementariedad que pueda existir entre ellos. Precisamente, la teoría de carteras óptimas nos enseña que ésta puede no ser la estrategia óptima, dependiendo del signo y magnitud de la correlación que exista entre ambos activos, supuesto que en el momento de llevar a cabo la inversión estemos en un contexto de incertidumbre acerca de sus rentabilidades.

Pero ¿qué combinación de los dos activos sería óptima? Un inversor puede estar dispuestos a asumir un mayor nivel de riesgo, si recibe también una mayor rentabilidad, aunque no cualquier combinación es preferible: el inversor tendrá un mapa de curvas de utilidad constante en el plano (riesgo, rentabilidad esperada). Cada una de estas curvas, convexa hacia el origen y posiblemente crecientes, es el lugar geométrico de los pares de valores (riesgo, rentabilidad esperada) que ofrecen un mismo nivel de utilidad. Curvas más elevadas en dicho plano corresponden a niveles de utilidad superiores, y curvas por debajo de una dada corresponden a niveles de utilidad inferiores.

Por otra parte, la frontera eficiente es el lugar geométrico de pares de valores (riesgo, rentabilidad esperada) con la siguiente propiedad: Consideremos todas las carteras con un determinado nivel de riesgo/varianza; todas ellas ofrecen una rentabilidad (esperada) inferior a la dada por la frontera eficiente para ese nivel de riesgo/varianza. La frontera eficiente tiene la forma de media elipse creciente, desde un vértice, que corresponde a la cartera de mínima varianza.

La cartera óptima para un inversor está determinada gráficamente en el punto de tangencia entre las curvas de indiferencia de utilidad, y la frontera eficiente. Otro inversor tendrá unas preferencias diferentes, y la tangencia se producirá en un punto distinto de la frontera eficiente, por lo que la cartera óptima para este inversor será distinta de la del primero. Esta es la base del análisis de carteras propuesto por Markowitz. Para el cálculo de la frontera eficiente, es imprescindible contar con una estimación de la matriz de varianzas y covarianzas, o de las varianzas y correlaciones entre los activos disponibles al inversor. Cuando un inversor puede distribuir su riqueza entre varios mercados, puede simlpificarse el problema (perdiendo algo de precisión), resolviendo le problema de aset allocation primero entre los distintos mercados, conotruyendo la cartera óptima entre el conjunto de índices de estos mercados, para luego, en una segunda etapa, construir para cada mercado la cartera óptima en la que invertir la cuantía previamente seleccionada para dicho mercado.

Pero antes de poder escoger una inversión (activo o cartera), hemos de hacer frente a dos dificultades: 1) por un lado, lo que interesa al inversor es la rentabilidad *esperada*, para cada activo, a lo largo del período en que se va a llevar a cabo la inversión, 2) por otro, *el riesgo no es observable*, por lo que hemos de utilizar alguna medida del mismo, para lo que generalmente se identifica riesgo con volatilidad, y ésta con varianza. Es muy importante observar que, desde el punto de vista de la teoría financiera, ambas deberían ser medidas hacia el futuro y, sin embargo, suelen ser inadecuadamente sustituidas por medidas históricas.

Valoración de opciones:

El precio de una opción depende de: a) el precio de ejercicio de la opción, b) el tiempo que resta hasta su vencimiento, c) el tipo de interés del activo sin riesgo, d) los dividendos ofrecidos por el activo subyacente, si los hay, e) el precio del activo subyacente, f) su *volatilidad, que no es observable.*

Para evaluar si el *precio de mercado* de una opción es correcto ha de disponerse de una estimación de la volatilidad del activo subyacente. Para ello, se necesita la volatilidad estimada del precio del subyacente durante el período residual hasta el vencimiento de la opción. Por tanto, necesitamos una *predicción* de la volatilidad. Con dicha medida, podríamos utilizar alguno de los modelos disponibles que, condicionado en la validez de las hipótesis en él incorporadas, nos proporcionaría el *precio teórico* de la opción. La comparación con su precio de mercado nos permitiría evaluar el interés que pueda tener tomar posiciones cortas o largas en la misma.

Cobertura de riesgos en inversiones a largo plazo:

El diseño de estrategias de cobertura de carteras depende crucialmente de la estimación del riesgo de los activos que configuran la cartera. Además, en este caso, tan importante como las medidas de volatilidad de los mercados del activo subyacente y del activo que se utiliza en la cobertura, es la medida de *covariación* entre ambos. De hecho, es ya habitual hablar de un *riesgo de correlación* entre activos.

La utilización de medidas de volatilidad y de covariación alternativas puede conducir a estrategias de cobertura bastante diferentes, lo que implicará a)

costes bastante distintos para las mismas y b) resultados asimismo diferentes, que pueden depender del tipo de evolución temporal seguido por la cotización del activo subyacente.

Varianza y covarianzas/correlaciones

En una población estadística, la varianza es el promedio ponderado de la desviación cuadrática entre un punto extraído al azar del soporte de la distribución y la esperanza matemática. La ponderación que recibe cada punto del soporte de la distribución es igual a la masa de probabilidad en cada punto. En una muestra, la varianza es el promedio equiponderado de las desviaciones cuadráticas de los datos muestrales respecto a la media muestral. Las ponderaciones son las frecuencias relativas de observación de los datos. Alternativamente, puede pensarse que no prestamos atención a los datos que puedan aparecer repetidos, y que cada uno de ellos recibe en el cálculo de la varianza una ponderación igual al inverso del tamaño muestral $1/N$, el mismo que recibiría en el cálculo de la media muestral.

La *desviación típica*, definida como raíz cuadrada de la varianza, puede interpretarse aproximadamente como el tamaño medio de las desviaciones de una variable alrededor de un valor de referencia, ya sea su esperanza matemática (en el caso de la población), o su media muestral (en el caso de la muestra). *En el caso de una variable aleatoria para la que se disponen de observaciones a través del tiempo, la desviación típica puede interpretarse como el tamaño medio de sus fluctuaciones* alrededor de un nivel de referencia, definido por la media muestral. Por consiguiente, cuando se trabaja con variables aleatorias de esperanza (o media muestral) igual a cero, la desviación típica es un buen indicador del *tamaño* de dicha variable.

Las matrices de covarianzas entre activos son necesarias en la mayoria de las aplicaciones financieras, como:

- estimación y predicción de la volatilidad de una cartera,
- estimación del Value at Risk (VAR) y Expected Shortfall (ES) en carteras con flujos de caja lineales,
- determinación de la asignación de recursos entre un conjunto de activos para configurar una cartera óptima,
- simular rentabilidades de un conjunto de activos con determinada estructura de correlaciones,
- estimación del VaR de carteras con pagos no lineales,
- estimación de precios en carteras de opciones sobre multiples activos,
- cobertura del riesgo de una cartera.

1.3 La varianza como indicador de riesgo: Limitaciones

La varianza y la desviación típica (poblacional o muestral) sólo tienen sentido frente a una medida de posición central de la distribución de probabilidad, que sirve de referencia. Sin embargo, no siempre las medidas de posición son estables en el tiempo. Cuando no lo son, el uso de la varianza como indicador de volatilidad queda en entredicho, como iremos viendo sucesivamente.

- la rentabilidad que interesa al inversor es la rentabilidad que espera obtener durante el horizonte de su inversión, por lo que, en realidad, debería utilizar una *predicción de la rentabilidad* durante dicho período. Generalmente, los modelos teóricos (selección de cartera de Markovitz, valoración de opciones de Black-Scholes) se basan en una medida de riesgo esperado durante el horizonte de la inversión, que es substituida generalmente por una medida histórica de riesgo, y ésta es calculada como la varianza muestral, sin llevar a cabo el tipo de predicción requerido por el modelo teórico. Para ello, el análisis de series temporales es imprescindible: especificando y estimando un modelo estadístico para la serie temporal de rentabilidades, podríamos obtener tal previsión. El modelo en cuestión debería incorporar todas aquellas variables que se considera que pueden influir sobre la rentabilidad del activo, si bien entonces necesitaremos prever asimismo el comportamiento de tales factores durante el horizonte de inversión. Una posibilidad consiste en utilizar un modelo univariante de series temporales (por ej., según el enfoque Box-Jenkins), confiando en que dicho modelo capture suficientemente bien la dinámica de la evolución temporal de la rentabilidad a lo largo del horizonte de inversión; otra posibilidad consistiría en utilizar modelos vectoriales autoregresivos (VAR), que incorporasen variables adicionales que puedan influir sobre el precio del activo en cuestión.

- *tampoco el nivel de riesgo del activo es observable,* pero se *identifica riesgo con volatilidad.* Ha sido asimismo tradicional asociar la volatilidad a un momento de segundo orden de la distribución de probabilidad o de frecuencias de una determinada rentabilidad. Así, la identificación entre volatilidad y varianza o, más precisamente, entre volatilidad y desviación típica, es habitual. Por tanto, *la volatilidad se define con respecto a un nivel de referencia,* generalmente la esperanza matemática de la rentabilidad analizada, que es una medida de posición central. Pero hay otras medidas de posición que pueden ser útiles bajo condiciones de asimetría: mediana, moda, percentiles, etc. De hecho, la identificación entre volatilidad y desviación típica no conduce a una medida adecuada del riesgo asumido en la inversión en un determinado activo,

- en la *práctica habitual* se entiende que *el riesgo es una característica relativamente estable* de un activo (en el caso del riesgo específico, no diversificable) o de un mercado (en el caso del riesgo sistemático, diversificable) que puede, por tanto, estimarse a partir de datos históricos, utilizando la

desviación típica de la rentabilidad de un activo. Sin embargo, debemos hacernos varias preguntas: ¿es el nivel de riesgo o de volatilidad estable en el tiempo? ¿deberíamos medir volatilidad sobre períodos relativamente breves de tiempo, obteniendo así una medición numérica que evoluciona de manera más o menos suave?

- el uso que habitualmente se hace de la desviación típica como indicador de volatilidad/riesgo, se fundamenta en el supuesto de Normalidad de la variable cuya volatilidad hemos calculado. Por ejemplo, la varianza estimada en un instante determinado puede utilizarse para construir un intervalo de confianza para los valores que puede tomar la rentabilidad que está siendo objeto de análisis. Sin embargo, la gran mayoría de las rentabilidades de activos financieros no siguen una distribución Normal, con clara evidencia de asimetría y exceso de curtosis.

- si llevamos a cabo una inversión con un determinado horizonte, es habitual considerar que el riesgo asumido viene medido por la varianza de la suma de las variaciones diarias en precio (rentabilidades continuas diarias). Asimismo, es habitual aproximar dicha varianza multiplicando la varianza diaria, supuesta constante, por el número de días contenidos en el horizonte de inversión. Sin embargo, este procedimiento no es correcto si el proceso con el que trabajamos (precios o rentabilidades) presenta autocorrelación, en cuyo caso la varianza no es aditiva temporalmente. Esta práctica conduce a sobre-estimacion (bajo autocorrelación negativa) o sub-estimación (bajo autocorrelación positiva) de la volatilidad de la rentabilidad en estudio. En tal situación, conviene utilizar una estructura temporal de volatilidades (volatilidad como función del horizonte), más que trabajar con una volatilidad constante para todos los plazos de inversión. Esta cuestión será analizada en más detalle más adelante.

Hay varias situaciones en las cuales utilizar la varianza como indicador de riesgo parece especialmente problemático:

- en presencia de una tendencia determinista, el nivel seleccionado como referencia para el comportamiento de la variable, que habitualmente es la media o la mediana muestrales, no será representativo de la evolución de la variable: si la tendencia es creciente, la primera parte de la muestra estará sistemáticamente por debajo de la media, mientras que la segunda parte estará sistemáticamente por encima. El estadístico de posición central no representa ni la primera ni la segunda parte de la muestra. Si calculamos la varianza muestral como indicador de volatilidad en este caso, imputaremos como tal lo que no es sino tendencia, y podríamos llegar a afirmar, erróneamente, que una variable es muy volátil, cuando lo que presenta es una fuerte tendencia determinista. De hecho, la varianza de una variable tendencial puede ser elevada incluso si ésta apenas experimenta fluctuaciones. En presencia de una tendencia lineal, la varianza está midiendo la

tasa de crecimiento; lo sorprendente es que este aspecto, que es positivo si estamos hablando del precio o cotización de un activo, será considerado negativo, al ser imputado como volatilidad y, por tanto, como riesgo asociado a la inversión en el mismo.

- algo similar ocurre en presencia de tendencias estocásticas (es decir, de raíces unitarias), un caso típico de no estacionariedad de variables financieras. En tales procesos la varianza crece con el número de observaciones utilizado en su cálculo, por lo que no tiene sentido hablar de la varianza, pues este momento no está bien definido. Su análogo muestral tiende a infinito si utilizamos en su cálculo un número de observaciones arbitrariamente grande, por lo que no proporciona información acerca del riesgo inherente a la toma de posición en un activo cuyo precio presenta tal comportamiento tendencial. En tal caso, deberíamos trabajar con la primera diferencia de la variable. Si se trata de un precio, su logaritmo presentará las mismas carcterísticas tendenciales, aunque quizá algo más amortiguadas. Su primera diferencia es la rentabilidad del activo, por lo que es muy intuitivo trabajar con esta transformación de la variable original. Este es el caso del comportamiento de los precios en muchos mercados.

- cuando, aun no existiendo tendencia, se ha producido un cambio de nivel en la media. En este caso, la media calculada con toda la muestra no representará ni la primera parte de ella, ni la segunda. Lo que ocurre es que la media ha sido distinta en la primera y segunda submuestras, y deberíamos recoger este hecho. De lo contrario, estaremos imputando como volatilidad lo que no es sino una ruptura en la media de la variable en estudio. Ni la media muestral es representativa de la rentabilidad ofrecida por el activo, ni la varianza muestral es una medida de incertidumbre.

- el último comentario está basado en el hecho de que la varianza mide *toda la fluctuación* que experimenta una variable (sea precio o rentabilidad), y seguramente querremos pensar que el riesgo es sólo una parte (quizá la parte no predecible) de dicha fluctuación (esto será analizado en detalle en la Sección 4.i). Como caso extremo, una función trigonométrica como $y_t = A.sen\left(2\pi\frac{t}{T}\right), t = 1, 2, ..., T$, para una constante A dada, experimenta fluctuaciones de un tamaño arbitrario, determinado por el valor de A, pero son de naturaleza puramente determinista. Ello significa que el valor de y_{t+s} en cualquier período futuro es perfectamente predecible en el instante t. Perfectamente predecible significa que el error de predicción es cero; además, la información muestral disponible en el instante t sería irrelevante, pues no necesitaríamos utilizarla para obtener dicha predicción. Las fluctuaciones en este proceso podrían ser arbitrariamente grandes, pues bastaría para ello con alterar el valor de las constantes. A pesar de que un activo cuyo precios siguiese tal comportamiento, no implicaría riesgo alguno para el inversor, la varianza de dicho proceso podría resultar arbitrariamente grande.

2 Momentos de una distribución de probabilidad

Toda variable aleatoria está caracterizada por su distribución de probabilidad, que no es sino el conjunto de valores posibles de la variable aleatoria, acompañados de sus respectivas probabilidades. El modo en que se representa la distribución de probabilidad depende de que la variable aleatoria en cuestión sea de naturaleza discreta o continua.

Si denotamos por $P(x_i)$ la masa de probabilidad en cada punto x_i del soporte Ω de la distribución de probabilidad de una variable aleatoria X, (conjunto de valores posibles de la variable aleatoria X), y por $f(x_i)$ la función de densidad que la representa, cuando ésta existe (distribuciones de tipo continuo), la esperanza matemática de la variable X se define:

$$E(X) = \mu_x = \int_{-\infty}^{\infty} x f(x) dx;$$

si la medida de probabilidad es continua, o:

$$E(X) = \mu_x = \sum_{x_i} x_i dP(x_i)$$

si la medida de probabilidad es discreta. En este último caso, x_i denota cada uno de los valores posibles de la variable aleatoria X, en número finito o no.

La *mediana* m está definida por el punto del soporte valor numérico para el cual se cumple:

$$\int_{-\infty}^{m} f(x) dx = \frac{1}{2}$$

en el caso de una variable aleatoria o distribución de probabilidad continuas, y:

$$Med(X) = \inf \left\{ m \mid \sum_{x_i}^{m} dP(x_i) = \frac{1}{2} \right\}$$

en el caso de una variable discreta. Esta formulación de la definición se debe a que en distribuciones discretas puede aparecer alguna ambigüedad en su cálculo.

La *moda* es el valor más probable de una distribución, es decir, el punto x_M del soporte Ω de la distribución, tal que:

$$P(X = x_M) \geq P(X = x) \; \forall x \in \Omega,$$

La moda puede no ser única. No existen condiciones bajo las cuales la mediana o la moda deban preferirse a la esperanza matemática como medida representativa de la distribución, pero hay que considerar tal posibilidad, dependiendo de las características de la distribución de probabilidad.

La esperanza matemática [suma de los valores numéricos ponderada por probabilidades] de las desviaciones entre los valores del soporte de la distribución y su esperanza matemática es igual a cero:

$$E(X-\mu_x) = E(X) - E(\mu_x) = \mu_x - \mu_x = 0$$

El valor numérico que minimiza la expresión: $E\left[(X-a)^2\right]$ es: $a = \mu_x$. El valor minimizado es la varianza de X.

El valor numérico que minimiza la expresión: E($\mid X - a \mid$) es: $a = m$.

La *varianza* de una variable aleatoria (cuando existe), es la esperanza matemática del cuadrado de las desviaciones entre los valores de la variable y su esperanza matemática:

$$\begin{aligned} \sigma_x^2 &= E\left(X-\mu_x\right)^2 = \int_{-\infty}^{\infty} \left(x-\mu_x\right)^2 f(x)dx \\ \sigma_x^2 &= \sum_{x_i} \left(x_i-\mu_x\right)^2 dP(x_i) \end{aligned}$$

en distribuciones continuas y discretas, respectivamente.

La varianza puede escribirse también:

$$\begin{aligned} \sigma_x^2 &= E\left[(X-\mu)^2\right] = E\left(X^2 - 2\mu X + \mu^2\right) = E\left(X^2\right) - \mu^2 \\ \sigma_x^2 &= \sum_{x_i} \left(x_i-\mu_x\right)^2 dP(x_i) = \sum_{x_i} x_i^2 dP(x_i) - 2\sum_{x_i} x_i\mu_x dP(x_i) + \sum_{x_i} \mu_x^2 dP(x_i) = \\ &= \sum_{x_i} x_i^2 dP(x_i) - 2\mu_x \sum_{x_i} x_i dP(x_i) + \mu_x^2 \sum_{x_i} dP(x_i) = E(x_i^2) - 2\mu_x^2 + \mu_x^2 = E(x_i^2) - \mu_x^2 \end{aligned}$$

Algunas propiedades de la varianza de una variable aleatoria X:

$$\begin{aligned} Var(aX+b) &= a^2 Var(X) \\ Var(a) &= 0 \end{aligned}$$

Como en muchas ocasiones se quiere poner dicho indicador en relación con el valor medio de la variable, se prefiere un indicador que tenga unidades comparables a las de la rentabilidad por lo que, cuando hablamos de volatilidad solemos referirnos a la *desviación típica*: raíz cuadrada de la varianza, tomada con signo positivo:

$$DT(X) = \sigma_x = \sqrt{\sigma_x^2}$$

Otros momentos poblacionales son el *coeficiente de variación*:

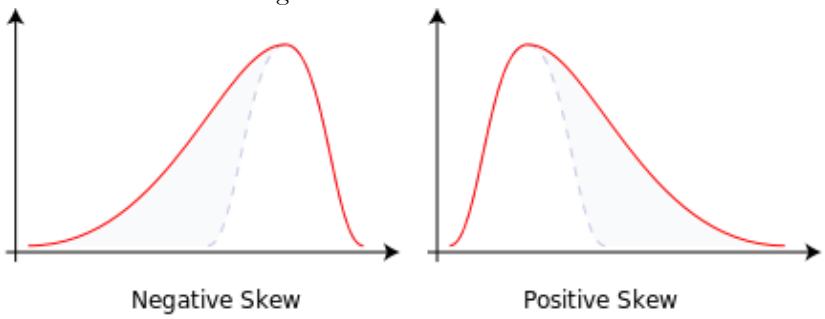
$$Coeficiente\ de\ variación = 100\frac{\sigma_x}{\mu_x}$$

que considera la desviación típica (volatilidad) como porcentaje del nivel alrededor del cual fluctúa la variable, lo cual es útil al comparar la volatilidad de variables que tienen una esperanza matemática diferente; por ej., al comparar la volatilidad de dos índices bursátiles distintos.

El *coeficiente de asimetría*:

$$Coeficiente\ de\ asimetría = \frac{E\left[(x-\mu_x)^3\right]}{\sigma_x^3}$$

que es positivo cuando la distribución es asimétrica hacia la derecha, en cuyo caso la moda es inferior a la mediana, y ésta es, a su vez, inferior a la media aritmética. El coeficiente de asimetría es negativo cuando la distribución es asimétrica hacia la izquierda, en cuyo caso la moda es mayor que la mediana, y ésta es, a su vez, superior a la media aritmética. Toda distribución simétrica tiene coeficiente de asimetría igual a cero.



El *coeficiente de curtosis*:

$$Coeficiente\ de\ curtosis = \frac{E\left[(x-\mu_x)^4\right]}{\sigma_x^4}$$

también llamado *coeficiente de apuntamiento*, es un indicador del peso que en la distribución tienen los valores más alejados del centro. Toda distribución Normal tiene coeficiente de curtosis igual a 3. Un coeficiente de curtosis superior a 3 indica que la distribución es más apuntada que la de una Normal teniendo, en consecuencia, menos dispersión que dicha distribución. Se dice entonces que es *leptocúrtica*, o *apuntada*. Lo contrario ocurre cuando el coeficiente de curtosis es inferior a 3, en cuyo caso la distribución es *platicúrtica* o *aplastada*. A veces se utiliza el *Exceso de curtosis*, que se obtiene restando 3 del coeficiente de curtosis.

La *covarianza* entre dos variables mide el signo de la asociación entre las fluctuaciones que experimentan ambas. Esencialmente, nos dice si, cuando una de ellas está por encima de su valor de referencia, p.ej., su media, la otra variable tiende a estar por encima o por debajo de su respectiva media:

$$Cov(X,Y) = E\left[(X-EX)(Y-EY)\right] = E(XY) - E(X)E(Y)$$

Siempre se cumple que:

$$Cov(X,Y) = E\left[X(Y - EY)\right] = E\left[(X - EX)Y\right]$$

Cuando alguna de las dos variables tiene esperanza cero, entonces:

$$Cov(X,Y) = E\left(XY\right)$$

Otras propiedades de la covarianza que no requieren ningun supuesto:

$$Cov(aX + b, mY + n) = amCov(X,Y)$$

El *coeficiente de correlación lineal* entre dos variables es el cociente entre su covarianza, y el producto de sus desviaciones típicas:

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

Mientras que la covarianza puede tomar cualquier valor, positivo o negativo, el coeficiente de correlación solo toma valores numéricos entre -1 y +1. Esto ocurre porque, por la desigualdad de Schwarz, la covarianza está acotada en valor absoluto por el producto de las desviaciones típicas de las dos variables.

Un caso importante es el de la covariación entre los valores de una variable con sus propios valores pasados. Así, tenemos, para cada valor entero de k:

$$\gamma_k = Cov(X_t, X_{t-k}),\ k = 0, 1, 2, 3, ...$$

sucesión de valores numéricos que configura la función de autocovarianza de la variable X_t, así como su función de autocorrelación:

$$\rho_k = \frac{Cov(X_t, X_{t-k})}{Var(X_t)} = \frac{\gamma_k}{\gamma_0}$$

El primer valor de la función de autocovarianza, $\gamma_{0,}$ es igual a la varianza de la variable. El primer valor de su función de autocorrelación, ρ_0, es siempre igual a 1.

Dos variables aleatorias son independientes si su función de densidad conjunta es igual al producto de sus funciones de densidad marginales:

$$f(x,y) = f_1(x).f_2(y)$$

dentro del rango de variación de ambas variables.

En el caso de distribuciones discretas (aquéllas en las que la variable en estudio toma valores en un conjunto discreto de puntos, que puede ser infinito), dos distribuciones son independientes si:

$$P(X = x, Y = y) = P(X = x).P(Y = y)$$

En general, en el caso continuo, la función de densidad de una variable Y, condicionada en otra variable X viene dada por:

$$f(y/x) = \frac{f(x,y)}{f_2(x)}$$

pudiendo definirse de modo similar la función de densidad de la variable X, condicionada por la variable Y.

En el caso discreto, se tiene (ver más adelante (2.7)):

$$P(Y = y/X = x) = \frac{P_{XY}(X = x, Y = y)}{P_X(X = x)}$$

Es fácil probar que si dos variables aleatorias son independientes, entonces su covarianza es cero.

La varianza de una suma o de una diferencia de dos variables aleatorias es:

$$\begin{aligned} Var(X+Y) &= Var(X) + Var(Y) + 2Cov(X,Y) \\ Var(X-Y) &= Var(X) + Var(Y) - 2Cov(X,Y) \end{aligned}$$

de modo que solo si ambas variables son *independientes* se tiene que la varianza de su suma es igual a la varianza de su diferencia:

$$Var(X+Y) = Var(X) + Var(Y)$$

En tal caso, el riesgo (medido por la desviación típica) de una cartera sería función de las ponderaciones con que entran en ella cada uno de los activos que la configuran y del riesgo de cada uno de dichos activos, pero no dependería de si la posición adoptada en cada activo es *corta* o *larga*, es decir, de si estamos comprados o vendidos en cada uno de ellos.

Estas expresiones pueden extenderse análogamente a cualquier combinación lineal de n variables. Un ejemplo sería la suma de dichas n variables.

Desigualdad de Chebychev:

$$E\left[g(X)\right] = \int_{-\infty}^{\infty} g(x)f(x)dx \geq \varepsilon^2 \int_S f(x)dx = \varepsilon^2 P\left[g(X) \geq \varepsilon^2\right]$$

siendo S el conjunto de puntos del soporte de X donde la función g es superior o igual a ε^2. Por tanto,

$$P\left[g(X) \geq \varepsilon^2\right] \leq \frac{E\left[g(X)\right]}{\varepsilon^2}$$

Cambio de variable

Si y es una función monótona y diferenciable de la variable aleatoria x, la función de densidad de y viene dada por: $g(y) = |dx/dy|\, f(x)$.

Ejemplo: Consideremos la relación entre variables: $y = x^2$, entonces $x = \sqrt{y} \Rightarrow |dx/dy| = (2\sqrt{y})^{-1}$, por lo que $g(y) = h(x)/\left(2\sqrt{y}\right)^{-1}$.Si, por ejemplo, $x \sim N(0,1)$, entonces: $f(x) = (2\pi)^{-1/2}\exp(-x^2/2)$, por lo que $g(y) = \frac{1}{2\sqrt{2\pi y}}\exp(-y/2),\ y \geq 0$

2.1 Momentos muestrales

En general, contamos con observaciones históricas acerca de una o varias variables (precios, rentabilidades, etc.) y queremos calcular medidas de posición central, de dispersión y de correlación con el objeto de resumir las propiedades básicas de dichos datos. El conjunto de datos observados define un histograma de frecuencias, o distribución muestral de frecuencias, que contiene *toda* la información disponible acerca de la variable considerada. Un histograma de frecuencias es similar a una distribución de frecuencias, pero es diferente de ella. Para entender la diferencia entre ambos, hemos de comprender el concepto de proceso estocástico, y el modo de utilizarlo en el análisis de datos de series temporales.

Un *proceso estocástico* $X_t, t = 1, 2, 3, ...$es una sucesión de variables aleatorias, indexadas por la variable tiempo. Las variables aleatorias pueden ser independientes entre sí o no, y pueden tener la misma distribución de probabilidad, o una distribución de probabilidad diferente. Cada dato de una serie temporal debe interpretarse como una muestra de tamaño 1 de la distribución de probabilidad correspondiente a la variable aleatoria de ese instante. Por ej., el dato de cierre del IBEX35 (suponiendo que disponemos de datos de cierre diarios) de hoy es una *realización*, es decir, una muestra de tamaño 1 de la variable aleatoria "precio de la cesta IBEX35" (como índice) el día de hoy. La distribución de probabilidad de esta variable puede ser diferente de la variable aleatoria IBEX35 hace un año por tener, por ejemplo, una esperanza matemática menor, una volatilidad mayor, o no ser Normal, mientras que hace un año sí lo era.

Vamos a suponer inicialmente que las variables X_t tienen todas la misma distribución de probabilidad, y son independientes entre sí. Este es el caso más sencillo, y constituye un proceso de *ruido blanco*. Sólo en este caso está totalmente justificado la utilización de momentos muestrales como características de "la variable X". Esta observación debe servir como llamada de atención al lector, dada la excesiva frecuencia con que se calculan estadísticos muestrales, calculados con datos históricos, para representar características de una variable; por ej., la desviación típica de la rentabilidad bursátil de un determinado mercado. Bajo este supuesto, podemos considerar los análogos muestrales de los momentos poblacionales o teóricos que vimos en la sección anterior. Los momentos muestrales consitiutyen estimaciones numéricas de los momentos poblacionales o teóricos.

Las medidas de posición central y dispersión análogas a la esperanza, varianza y desviación típica son:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n};\ S_x^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1};\ DT_x = \sqrt{S_x^2}$$

mientras que la covarianza y coeficiente de correlación muestrales son:

$$Cov(X,Y) = \frac{1}{T}\sum_{t=1}^{T} (x_t - \bar{x})(y_t - \bar{y}) = \frac{1}{T}\sum_{t=1}^{T} x_t y_t - \bar{x}\bar{y}$$

La media, varianza, mediana, covarianza y coeficiente de correlación mues-

trales satisfacen propiedades similares a las ya mencionadas para sus análogos poblacionales. Entre ellas:

- La suma de las desviaciones de la variable respecto de su media, es igual a cero:
$$\sum_{i=1}^{n}(x_i-\bar{x})=\sum_{i=1}^{n}x_i-\sum_{i=1}^{n}\bar{x}=n\bar{x}-n\bar{x}=0$$

- Como consecuencia de lo anterior, la media muestral de las diferencias $x_i-\bar{x}, i=1,2,...,n$ es igual a cero.

- Si una de las dos variables, X o Y tiene esperanza cero, tenemos:

$$Cov(X,Y)=\frac{1}{T}\sum_{t=1}^{T}x_t y_t=E(XY)$$

- La varianza de X puede escribirse:

$$\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2=\frac{1}{n}\sum_{i=1}^{n}x_i^2-2\frac{1}{n}\sum_{i=1}^{n}x_i\bar{x}+\frac{1}{n}\sum_{i=1}^{n}\bar{x}^2=\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2$$

Al igual que en el caso de una distribución de probabilidad, otras medidas utilizadas en la representación de una muestra son:

$$Coeficiente\ de\ variación=100\frac{DT_x}{\bar{x}}$$

$$Coeficiente\ de\ asimetría=\frac{\frac{1}{T}\sum_{t=1}^{T}(x_t-\bar{x})^3}{DT_x^3}$$

$$Coeficiente\ de\ curtosis=\frac{\frac{1}{T}\sum_{t=1}^{T}(x_t-\bar{x})^4}{DT_x^4}$$

siendo T el tamaño muestral.

El *recorrido* o *rango* es la diferencia entre el mayor y el menor valor observados de una variable. Los cuartiles son los datos que dividen a la muestra, una vez ordenada crecientemente, en cuatro submuestras de igual tamaño (aproximadamente). El segundo cuartil es la *mediana*. El *rango intercuartílico* es la distancia entre los cuartiles primero y tercero. Estos estadísticos tienen la virtud de no verse afectados por la presencia de valores atípicos. De modo análogo se definen los *deciles* y *percentiles*.

En una variable temporal, las funciones de autocovarianza y autocorrelación muestrales se definen:

$$\gamma_k=Cov(X_t,X_{t-k})=\frac{1}{T}\sum_{t=k+1}^{T}(x_t-\bar{x})(x_{t-k}-\bar{x})$$

$$\rho_k = Corr(X_t, X_{t-k}) = \frac{Cov(X_t, X_{t-k})}{\sqrt{S_x^2}\sqrt{S_x^2}} = \frac{\frac{1}{T}\sum_{t=k+1}^{T} x_t x_{t-k} - \bar{x}^2}{S_x^2}$$

siendo siempre: $\gamma_0 = Var(X_t)$ y $\rho_0 = 1$.

[Case Study II:3] Volatilidades a lo largo de una estructura temporal.

2.2 Semi-desviación típica y momentos parciales inferiores de segundo orden

En la medición del riesgo estamos interesados en los posibles resultados negativos de una cartera. Por tanto, nos preocupa especialmente medir las características de la distribución de rentabilidades en su cola inferior. Los momentos que examinamos en esta sección miden las propiedades de las rentabilidades cuando éstas son inferiores a una determinada referencia, como pudiera ser la rentabilidad de un indice de renta variable, o un determinado porcentaje previamente estipulado. Es decir, caracterizan las propiedades en la cola de la distribución, frente a momentos tradicionales, que caracterizan las propiedades de la distribución de rentablidades a lo largo de todo el rango de valores observados de dicha rentabilidad. Por ejemplo, el bien conocido ratio de Sharpe relaciona la rentabilidad esperada, en exceso de la ofrecida por el activo libre de riesgo, con la volatilidad de la rentabilidad. Por tanto, considera la informacion relativa a la distribución de probabilidad a lo largo de toodo el rango de valores de la rentabilidad. Para algunas aplicaciones este tipo de ratios puede ser importante, mientras que para otro tipo de usos, como pueda ser la medida del riesgo en la cola de la distribución, son claramente inapropiados.

Veamos alguno de los momentos en la cola de una distribucion:

La semivarianza es una medida de dispersión de las observaciones inferiores a la esperanza matemática de una variable:

$$SVar(X) = E\left[\min(X - E(X), 0)^2\right]$$

Puesto que $E\left[\min(X - E(X), 0)\right] \neq 0$, excepto que se tuviese $X = E(X)$ con probabilidad 1, tenemos que: $E\left[\min(X - E(X), 0)^2\right] \neq Var\left(\min(X - E(X), 0)\right)$ .Es decir, la semivarianza no es la varianza de las observaciones inferiores a la referencia escogida en su definición, que en este caso es $E(X)$, pero que podría ser un umbral α.

La semi-desviación típica estimada a partir de una muestra es:

$$\hat{\sigma}_{semi} = \sqrt{\frac{1}{T}\sum_{t=1}^{T} \min\left(R_t - \bar{R}, 0\right)^2}$$

en cuyo cálculo suelen incluirse los "ceros", es decir las observaciones con $R_t = \bar{R}$. Al igual que el resto de los momentos que vemos en esta sección, la semi-desviación típica se especifica en términos anualizados, para lo que debe hacerse

el habitual ajuste: si se trata de rentabilidades diarias, se multiplica por raiz de 252, mientras que si se tratase de rentabilidades mensuales, se multiplicaría por la raiz de 12.

El momento parcial inferior de orden 2 es análogo a la semi-desviación típica, pero calculado con respecto a una referencia τ :

$$LPM_{2,\tau}(X) = \sqrt{E\left[\left|\min(X-\tau,0)\right|^2\right]} = \sqrt{E\left[\max(\tau - X, 0)^2\right]}$$

En general, el momento parcial inferior de orden k , siendo k un número no necesariamente entero, (*Lower partial moment*) es:

$$LPM_{k,\tau}(X) = \left(E\left[\left|\min(X-\tau,0)\right|^k\right]\right)^{1/k} = \left(E\left[\max(\tau - X, 0)^k\right]\right)^{1/k}$$

El LPM de orden 1 respecto del origen, $\text{LPM}_{1,0}$ es:

$$LPM_{k,\tau}(X) = E\left[\max(\tau - X, 0)\right]$$

que es el pago esperado en una opción put con precio de ejercicio igual a τ, y puede interpretarse como el coste de asegurarse respecto del riesgo de pérdidas de una cartera.

Según aumenta k, el LPM va asignando mayor peso a las rentabilidades más negativas. El momento $LPM_{2,0}$ es la semi-desviación típica respecto de τ, $LPM_{3,0}$ se denomina la semi-asimetría (semi-skewness), mientras que $LPM_{4,0}$ es la semi-curtosis. Ejercicios: [EIV.1.1], [EIV.1.2].

Los *lower order partial moments* intervienen en el calculo de muchas medidas de performance de una cartera de activos financieros.

El Example IV.1.1 calcula la semi desviación típica y el momento inferior de segundo orden de una muestra de 36 rentabilidades mensuales, respecto de un rendimiento activo del 2% anual. El ejemplo IV.1.2 calcula los momentos de orden inferior de orden 1, 2, 3, 4, 5, 10 y 20 de los mismos rendimientos del ejemplo previo, tanto respecto del origen como respecto de un rendimiento activo del 2%.

Una vez introducido este tipo de estadisticos para la cola de una distribucion de rentabilidades, nos centramos en lo sucesivo en el uso de cuantiles de dicha distribución y otros estadísticos derivados de los mismos como medidas de riesgo. Sin duda,el mas conocido es el cuantil Valor en Riesgo, asi como un estadístico derivado del mismo, el VaR condicional o Pérdida Esperada en la Cola de la distribución, que introduciremos más adelante.

2.3 Estacionariedad

Hay varias razones estadísticas que justifican el uso de rentabilidades, en vez de precios o cotizaciones, al analizar los mercados financieros. La más importante, es la general ausencia de estacionariedad en los precios de los activos financieros, así como en los índices de los principales mercados, que puede reflejarse de

diversas formas: presencia de tendencias estocásticas, presencia de *tendencias deterministas* en los precios de mercado, volatilidad cambiante en el tiempo, etc.. Una *tendencia determinista* es una función exacta del tiempo, generalmente lineal o cuadrática. Una *tendencia estocástica* es un componente estocástico cuya varianza tiende a infinito con el paso del tiempo.

Si una variable presenta una *tendencia determinista*, su valor esperado tenderá a aumentar o disminuir continuamente, con lo que será imposible mantener el supuesto de que la esperanza matemática de la sucesión de variables aleatorias que configura el proceso estocástico correspondiente a dicha variable, es constante. En consecuencia, tampoco podrá mantenerse que la distribución de probabilidad de dichas variables es la misma a través del tiempo. Sin embargo, si efectuamos una correcta especificación de la estructura de dicha tendencia, podrá estimarse y extraerse del precio, para obtener una variable estacionaria, que no presentaría las dificultades antes mencionadas. Un ejemplo claro es la aparente tendencia cuadrática en el índice S&P500, que puede estimarse mediante un polinomio de grado 2 del tiempo, con coeficiente positivo en la segunda potencia,

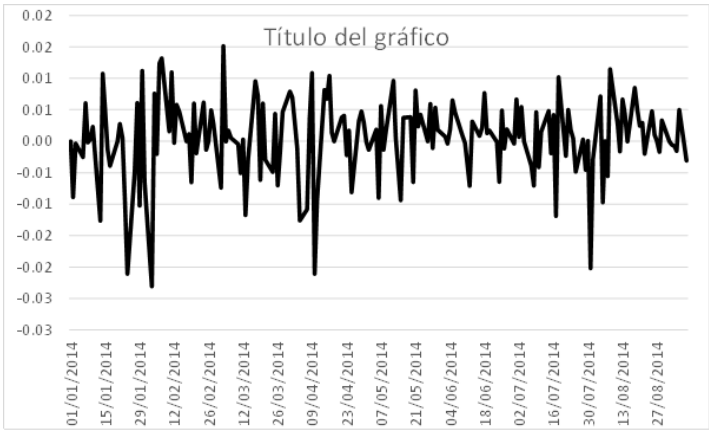
$$SP500_t = a + bt + ct^2 + u_t$$

Las diferencias entre los valores del índice y los que toma dicha función determinista del tiempo podrían servirnos como la versión sin tendencia del índice S&P500 y, como se ve en los gráficos de la pestaña *SP500 trend* en el archivo *Indices.xls*, ambas versiones de la variable son de naturaleza muy diferente. En este caso, el grafico ilustra que la eliminación de la tendencia cuadrática determinista deja un comportamiento un tanto extraño, que podemos admitir de carácter estocástico, que habría que modelizar. La volatilidad de la serie S&P500 hacia el final de la muestra, que es enorme en términos históricos, queda claramente reflejada al eliminar la tendencia determinista.

Antes comentamos la conveniencia de considerar toda serie temporal de precios o de rentabilidades de activos financieros como una realización de un proceso estocástico. Pues bien, dicho proceso se dice que es estacionario cuando su distribución de probabilidad es invariante en el tiempo. Recordemos que un proceso es una sucesión de variables aleatorias, y que el precio o la rentabilidad que observamos en un periodo determinado es una realización (una muestra de tamaño 1) de la variable aleatoria correspondiente a dicho periodo. Pues bien, cuando se examina una serie de precios como puede ser el indice S$&P500, es dificil pensar que cada una de las cotizaciones de fin de mes, que aparecen en el gráfico, provienen de una misma distribución de probabilidad. No es probable que los precios de final de 2002, en torno a 800, provengan de la misma distribución de probabilidad que generá las cotizaciones de mitad de 2007, en torno a 1.600, el doble del anterior. Los valores medios de las distribuciones de probabilidad correspondientes a dichos dias parecen claramente diferentes, como pudo suceder asimismo con otros momentos, como la varianza o asimetría.

S&P500: enero 2000 - julio 2013

S&P500: rentabilidades diarias



En definitiva, se trata de un proceso no estacionario. Mayor dificultad entraña el caso en que una variable precio incluye una tendencia estocástica pues, en tal caso, su esperanza y varianza no están definidas. La presencia de una tendencia estocástica requiere transformar la variable, generalmente en primeras diferencias temporales, o tomando las diferencias entre las observaciones correspondientes a una misma estación cronológica, en el caso de una variable estacional. La transformación mediante diferencias resulta bastante natural en el análisis de datos financieros, por cuanto que la primera diferencia del logaritmo de un precio, es la *rentabilidad* del activo. Por esto es que también la transformación logarítmica es utilizada habitualmente cuando se trabaja con precios o índices de mercado. Solemos pensar que las rentabilidades son procesos estacionarios, es decir, que sus observaciones provienen de distribuciones de probabilidad con igual distribución. Esto es relativamente razonable por

cuanto que, a diferencia de lo que sucede con el propio índice, las rentabilidades fluctúan en torno a su media. Sin embargo, en el caso del S&P500, el gráfico de la rentabilidad, obtenida como primera diferencia logarítmica muestra períodos de mayor y de menor volatilidad, como sucede con todo activo financiero. Por tanto, tambien es razonable pensar que si esta transformacion ha podido generar una variable que tiene una misma media a lo largo del tiempo, sin embargo su varianza no lo es, lo que tendremos en cuenta a lo largo de las próximas secciones. A pesar de este contratiempo, es habitual transformar los precios en retabilidades y trabajar con éstas, modelizando la variación temporal que pueda existir en otros momentos, como la varianza, e incluso la asimetria y la curtosis.

La transformacion a rentabilidades es de gran importancia, pues hemos de tener presente que si la media de la serie temporal no es constante en le tiempo, entonces ni el calculo de la media muestral, ni el de ningun momento que se calcula tomando a la media como referencia, como sucede con la varianza, desviación tipica, asimetria o curtosis, tienen sentido. Por ejemplo, como prácticamente ningún precio o índice financiero es estacionario, el uso indiscriminado de un estadístico como la varianza o la desviación típica como indicador de riesgo conduce a medidas sesgadas al alza.

2.4 Ejercicio práctico: Estadísticos descriptivos en la estimación de la volatilidad

El archivo *indices.xls* presenta datos históricos de un conjunto de indices de renta variable. Queremos comparar en terminos de volatilidad estos mercados. El lector debe estar atento al hecho de que tal análisis, que en ocasiones se efectúa, está sujeto a un problema fundamental, cual es el carácter no estacionario de los indices de bolsa. Ello hace que sus momentos muestrales no estén bien definidos, negando validez al análisis que se presenta, que debería realizarse en términos de las rentabilidades diarias, por ejemplo, de los mercados. El lector debe pregunatrse cuáles de los cálculos que presentamos sobre los indices está justificado.

En la pestaña se presentan las observaciones muestrales para los índices de Bolsa, así como para sus rentabilidades logarítmicas (a la derecha). Debajo tenemos la asimetria y exceso de curtosis de cada índice, así como el cálculo del estadístico Bera_Jarque para contrastar la Normalidad de la distribución de probabilidad seguida por cada índice. Por último, se calcula la desviación típica de cada indice sobre submuestras de amplitud creciente. De hecho, se presenta la volatilidad anualizada. Como se aprecia, la estimación de la desviación típica aumenta con la amplitud de la muestra. Esto no debe suceder. No estamos estimando nada que sea proporciona a la amplitud del intervalo sobre el que se calcula. Si estimamos la desviación típica con 1 o con 2 años tendremos distintos resultados numéricos, pero comparables. En este ejercicio, la desviación tipica es creciente porque en procesos no estacionarios, como estos, la varianza es creciente con la longitud de de la muestra utilizada, tendiendo a infinito con ésta. De hecho, este es un procedimiento para contrastar la ausencia de estacionariedad de una serie temporal.

Como se aprecia, esto no sucede al trabajar con rentabilidades (a la derecha). Más a la derecha se estima la volatilidad anual como el promedio de las rentabilidades al cuadrado. Como se ve, las diferencias respecto del cálculo con desviaciones típicas son minimas.

Al trabajar con indices, las cotizaciones máxima o mínima, por sí solas, no son nada informativas respecto de establecer comparaciones entre índices. Ni siquiera el rango muestral lo es, a pesar de que ya establece un intervalo de valores cubierto por la variable. Parece evidente que el posible interés de estos estadísticos descansa en expresarlos como porcentaje de una medida de posición central. En esta comparación, ya aparecen Nasdaq, IPC Mexico y Athens como los índices de mayor aparente variabilidad.

Hay que observar, sin embargo, que un rango amplio no implica volatilidad si los valores separados de la media no aparecen apenas en la muestra; por tanto, una limitación del rango es que sólo utiliza como información los valores máximo y mínimo. No estamos considerando todavía la frecuencia con que aparece en la muestra cada valor numérico del rango de variación de cada índice.

Una medida similar es la relación entre rango centrado del 80% y media. Ahora hemos descartado los valores muy separados de la media, tanto por encima como por debajo, y estamos analizando la amplitud del rango en el que recaen el 80% de los valores muestrales. Establecemos así una diferencia entre *valores normales* y *valores extremos*. Si los valores extremos del rango de variación aparecen con relativa frecuencia, entonces un rango como el del 80% tenderá a ser más amplio que si los valores separados de la media aparecen infrecuentemente. Por tanto, si un índice que tiene un rango total amplio pasa a tener un rango del 80% relativamente más estrecho ello se debe a que los valores extremos ocurren con poca frecuencia. Si la amplitud del rango del 80% es relativamente mayor que la del rango total, en relación con otros índices, se deberá a que si bien los valores separados de la media *no son demasiado extremos*, ocurren con una relativa frecuencia.

Nuevamente, la desviación típica ni la varianza por sí solas nos proporcionan información relevante. Si examinamos las desviaciones tipicas muestrales, ya apreciamos que no pueden utilizarse para obtener una estimacion de la volatilidad anual, pues habriamos de multiplicar por el numero de dias de mercado en un año, unos 250, obteniendo valores de volatilidad disparatados.

Dejando aparte este problema, podríamos comparar las desviaciones típicas, como proporción del nivel medio respecto al cual se han calculado. El uso de desviaciones típicas como porcentaje de la media permite la comparación entre mercados o activos, o también comparar la volatilidad en un mismo mercado en distintos instantes de tiempo. Este es el *coeficiente de variación:*

$$\upsilon = 100\frac{s_x}{\bar{x}}$$

Avanzando en esta línea de poner los estadísticos en términos relativos, cuando se pretende comparar variables medidas en diferentes unidades, es útil *tipificar* o *estandarizar* las variables, restando de cada observación la media muestral, y dividiendo por la desviación típica. Mediante esta transformación, eliminamos

las unidades de cada variable, por lo que pueden ser comparables entre sí, en términos de volatilidad. De hecho, bajo el supuesto de que la serie temporal relativa a cada una de las variables está compuesta de observaciones independientes, extraídas de una determinada distribución, con esperanza μ y varianza σ^2 constantes, las variables tipificadas tienen esperanza cero y varianza igual a uno. El carácter de la distribución no juega ningún papel en este resultado.

Cuando se pretende inspeccionar en un gráfico la posible correlación entre variables, es asimismo útil utilizar esta transformación. Esto corrige, además el efecto que produciría el que los distintos índices toman magnitudes diferentes, lo que haría que, en un gráfico de sus niveles, se observasen las fluctuaciones de tan sólo uno o dos de ellos, apareciendo los demás como líneas suaves.

Estos últimos cálculos que hemos hecho, al presentar estadísticos como porcentajes de la media, resuelven en parte un problema, que es el de que las distintas series temporales que consideramos tienen muy diferente promedio. Pero el problema es más grave: como hemos advertido, los cálculos anteriores adolecen de serias dificultades cuando se aplican a índices de Bolsa, debido a su naturaleza no estacionaria. Cuando esto sucede, los momentos de la distribución de probabilidad no están bien definidos, pues cambian en el tiempo. Es necesario entonces trabajar con una transformación que sea estacionaria. La rentabilidad logarítmica generalmente lo es.

El análisis de estadisticos debe repetirse ahora para las rentabilidades, como se hace en la pestaña "*indices*". En ese analisis, ahora con plena justificación estadistica, se aprecia que Bovespa es el índice más volatil, seguido de Nasdaq y Athens. Podemos ver que algunos estadisticos ahora carecen de sentido para las rentablidades, por tener media proxima a cero. Por ejemplo, el coeficiente de variación, o los rangos, calculados en términos relativos de la media. En cambio, los rangos por si mismos, tienen pleno sentido, a diferencia de lo que sucedia trabajando con los indices.

2.5 Varianza de la cartera

Si tenemos una cartera formada por n activos definida mediante pesos $\omega_i, i = 1, ..., n$, el VaR de la cartera se calcula de modo análogo a como lo hemos definido para un solo activo foinanciero. La rentabilidad de la cartera puede escribirse:

$$R_{c,t+1} = \sum_{i=1}^{n} w_i R_{i,t+1}$$

mientras que la varianza de dicha rentabilidad de la cartera será:

$$Var(R_{c,t+1}) = Var\left(\sum_{i=1}^{n} w_i R_{i,t+1}\right) = \sum_{i=1}^{n} w_i^2 Var(R_{i,t+1}) + \sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n} w_i w_j Cov(R_{i,t+1}, R_{j,t+1}) =$$

$$= \sum_{i=1}^{n} w_i^2 Var(R_{i,t+1}) + 2\sum_{i=1}^{n}\sum_{\substack{j=1\\ i<j}}^{n} w_i w_j Cov(R_{i,t+1}, R_{j,t+1})$$

que puede escribirse en notación matricial,

$$\sigma^2_{c,t+1} = Var(R_{c,t+1}) = \omega' \Sigma_{t+1} \omega$$

siendo ω el vector de pesos relativos de la cartera: $\omega = (\omega_1, \omega_2, ..., \omega_n)$, y Σ_{t+1} la matriz simétrica nxn de varianzas y covarianzas de las rentabilidades de los n activos que componen la cartera.

Si denotamos por $x = (x_1, x_2, ..., x_n)$ las cantidades invertidas en cada activo y por W la cuantía total, $W = \sum_{i=1}^{n} x_i$, de modo que $\omega_i = x_i/W$, tenemos, en términos monetarios:[1]

$$\sigma^2_{c,t+1} W^2 = x' \Sigma_{t+1} x$$

En el caso de una cartera compuesta por dos activos, (con $n = 2$), tendríamos:

$$\sigma^2_{c,t+1} \equiv Var(R_{c,t+1}) = w_1^2 Var(R_{1,t+1}) + w_2^2 Var(R_{2,t+1}) + 2w_1 w_2 Cov(R_{1,t+1}, R_{2,t+1}) =$$

$$= (\omega_1, \omega_2) \begin{pmatrix} \sigma^2_{1,t+1} & \sigma^2_{12,t+1} \\ \sigma^2_{12,t+1} & \sigma^2_{2,t+1} \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}$$

El riesgo de una cartera (su volatilidad) puede reducirse introduciendo en ella N activos con correlación reducida, o bien ampliando el número de activos de la cartera. Para hacernos una idea, si formamos una cartera equiponderada con activos, todos con igual volatilidad σ, y cada dos de ellos con la misma correlación, ρ, tenemos una volatilidad:

$$\sigma_c = \sigma \sqrt{\frac{1}{N} + \left(1 - \frac{1}{N}\right)\rho}$$

que disminuye rapidamente al aumentar el número de activos.

2.6 El coeficiente de correlación lineal como medida de asociación entre variables

El coeficiente de correlación lineal de Pearson, habitualmente utilizado, está basado en supuestos que pueden limitar su utilización con carácter general.

[1] Puesto que: $\sigma^2_{c,t+1} = \omega' \Sigma_{t+1} \omega = \left(\frac{1}{W}x\right)' \Sigma_{t+1} \left(\frac{1}{W}x\right) = \frac{1}{W^2} x' \Sigma_{t+1} x$

- El coeficiente de correlación entre variables no estacionarias puede ser muy próximo a 1 (en general lo será), sin que necesariamente implique que el componente puramente aleatorio de ambas variables (sus innovaciones) están relacionadas, o se mueven en sintonía. Es lo que conocemos como *correlación espúrea*.

- El coeficiente de correlación no es invariante bajo transformaciones no lineales (por ejemplo, logarítmicas),

- Los valores factibles del coeficiente de correlación dependen de las distribuciones marginales de las dos variables. Por ejemplo, si $\ln(X)$ es $N(0,1)$ y $ln(Y)$ es $N(0,4)$, entonces el coeficiente de correlación no puede valer más de 2/3 ni menos de -0,09

- La dependencia perfecta y positiva entre dos variables aleatorias no implica necesariamente un coeficiente de correlación lineal igual a 1, ni la correlación perfecta negativa implica un coeficiente de correlación lineal igual a -1, como muestra el ejemplo anterior.

- Un coeficiente de correlación lineal igual a cero no implica independencia. Solo en el caso en que las variables sigan una distribución conjunta Normal bivariante, ausencia de correlación (coeficiente de correlación lineal igual a cero) implicaría independencia. Por ejemplo, una dependencia perfecta en que las dos variables toman valores alineados a lo largo de un círculo implicará un coeficiente de correlación lineal muy reducido. Un coeficiente de correlación lineal nulo entre dos variables con distribución t de Student no implica independencia, pues ambas variables pueden estar relacionadas a través de sus momentos de orden superior.

2.7 Distribuciones marginales y condicionadas: Un ejemplo

Consideremos la distribución de probabilidad bivariante,

		X_1				
		-2	-1	0	1	2
	-1	2/24	0	2/24	4/24	
X_2	0	0	1/24	2/24	0	2/24
	2	0	3/24	2/24	0	6/24

donde X_1 puede tomar valores -2,-1,0,1,2, mientras que X_2 puede tomar valores -1, 0,2. El cuadro recoge probabilidades; por ejemplo, $P[X_1 = -1, X_2 = 0] = 1/24$. Las 15 probabilidades del cuadro suman 1.

La distribución marginal de X_1 es,

$Valores\ de\ X_1$	-2	-1	0	1	2
Pr$obabilidades$	2/24	4/24	6/24	4/24	8/24

con $E(X_1) = 1/2, Var(X_1) = 7/4$,siendo la distribución de X_2,

$Valores\ de\ X_2$	-1	0	2
Pr$obabilidades$	8/24	5/24	11/24

con $E(X_2) = 7/12, Var(X_2) = 263/144$.

La distribución de probabilidad de X_1 condicional en un valor numérico de X_2 es,

$Valores\ de\ X_1$	-2	-1	0	1	2
Si $X_2 = -1$	$1/4$	0	$1/4$	$1/2$	0
Si $X_2 = 0$	0	$1/5$	$2/5$	0	$2/5$
Si $X_2 = 2$	0	$3/11$	$2/11$	0	$6/11$

con $E(X_1/X_2 = -1) = 0, E(X_1/X_2 = 0) = 3/5, E(X_1/X_2 = 2) = 9/11$.

Luego $E(X_1/X_2)$ es una variable aleatoria que toma valores 0, 3/5, 9/11, con probabilidades respectivas: 8/24, 5/24, 11/24. Por tanto, su esperanza matemática es 1/2, que coincide con $E(X)$. Este es un resultado general, pues siempre se tiene,

$$E\left[E\left(X_1/X_2\right)\right] = E(X_1)$$

Las dos variables que hemos analizado no son independientes, pues ninguna de ellas satisface la condición de que su distribución marginal coincida con su distribución condicionada en cualquier valor de la otra. Dicho de otro modo, el valor que toma una variable X_2 es *informativo* acerca de los posibles valores de la otra variable X_1.

3 Rentabilidades

3.1 Rentabilidad continua equivalente

Distinguimos entre rentabilidades porcentuales y rentabilidades logarítmicas. Estas últimas se conocen asimismo como *rentabilidades continuas.*

Rentabilidad porcentual:

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

Rentabilidad logarítmica:

$$r_t = \ln(P_t/P_{t-1}) = \ln P_t - \ln P_{t-1}$$

donde vemos la diferencia en la transformación logarítmica a que antes nos referíamos.

Ambas rentabilidades son aproximadamente iguales si R_t es pequeña, puesto que:

$$r_t = \ln P_t - \ln P_{t-1} = \ln \frac{P_t}{P_{t-1}} = \ln\left(1 + R_t\right) \sim R_t$$

mientras que la relación exacta entre ambas, siempre válida, está dada por:

$$\ln\left(1 + R_t\right) = r_t$$

y r_t se dice que es la *rentabilidad continua equivalente* a R_t.

La transformación logarítmica hace que podamos obtener rentabilidades continuas compuestas mediante sumas. Supongamos que queremos calcular la rentabilidad sobre dos períodos. Observando que:

$$\begin{aligned} r_t^1 + r_{t-1}^1 &= \ln \frac{P_t}{P_{t-1}} + \ln \frac{P_{t-1}}{P_{t-2}} = \ln P_t - \ln P_{t-1} + \ln P_{t-1} - \ln P_{t-2} = \\ &= \ln P_t - \ln P_{t-2} = \ln \frac{P_t}{P_{t-2}} = r_t^2 \end{aligned}$$

vemos que la rentabilidad continua a 2 períodos es, simplemente, la suma de las rentabilidades continuas a 1 período obtenidas durante los dos últimos períodos. Algo similar ocurre para inversiones llevadas a cabo durante n y m períodos de tiempo, respectivamente, siendo n un múltiplo de m ($n = km$), pero siempre que las rentabilidades sean continuas. En ese caso, la suma de las rentabilidades continuas obtenidas durante los últimos k intervalos de tiempo, cada uno de ellos de duración n períodos, es igual a la rentabilidad continua obtenida durante los últimos m períodos.

Por el contrario, la suma de rentabilidades porcentuales sobre k períodos de tiempo de longitud m no proporciona exactamente la rentabilidad porcentual sobre un intervalo de longitud n, y el error de aproximación va aumentando con k.

Es importante observar que, para realizar la agregación temporal de las rentabilidades de tipo continuo no es preciso suponer independencia temporal de las mismas.

3.2 Rentabilidad en mercados cotizados en tipos de interés

En mercados que cotizan en TIRes o en tipos de interés, como sucede con un mercado interbancario, no existe un activo negociable. Sin embargo, en ocasiones queremos hablar de rentabildades. La rentabilidad en dicho mercado no es el tipo de interés cotizado, excepto si se mantiene el activo a vencimiento. Para ello, calculamos la rentabilidad considerando la variación en el precio de una cartera invertida en el mismo. Generamos un índice de precios sobre un nominal de 100 unidades monetarias como si se tratase de un bono cupon cero, y calculamos la variación porcentual o logarítmica en dichos precios. n el caso de tipos a 1 o más años, utilizamos la expresión: $P = 100/(1 + r_t)^a$, siendo a el vencimiento en años, mientras que en plazos menores a un año utilizamos: $P = 100/(1 + br_t)$, donde b denota la fracción de año de la inversión. Por ejemplo, si una rentabilidad cotizada a vencimiento 1 año se ha reducido de 5,32% a 4,25%, la cartera habrá incrementado su valoración en el mercado. El descenso de tipos se puede evaluar por medio de:

$$\frac{P_t - P_{t-1}}{P_{t-1}} = \frac{\frac{100}{1+r_t} - \frac{100}{1+r_{t-1}}}{\frac{100}{1+r_{t-1}}} = \frac{\frac{100}{1,00425} - \frac{100}{1,0532}}{\frac{100}{1,0532}} = \frac{95,9233 - 94,9487}{94,9487} = \frac{0,9476}{94,9486} = 0,010264$$

y la revalorización habrá sido del 1,02%. A vencimientos más cortos, la misma variación en tipos de interés generaría una rentabilidad más reducida, mientras que a plazos superiores a un año, la rentabilidad sería superior a la obtenida si el vencimiento de los tipos cnosiderados es de 1 año.

Nótese que:

$$\frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1 = \left(\frac{100/P_t}{100/P_{t-1}}\right)^{-1} - 1 = \left(\frac{1+r_t}{1+r_{t-1}}\right)^{-1} - 1$$

por lo que un procedimiento más simple, aunque quizá más dificil de recordar, consiste en sumar 1 a los tipos de interés cotizados r_t, r_{t-1}, y calcular la inversa de la tasa de variación de $1 + r_t$:

$$\left(\frac{1+r_t}{1+r_{t-1}}\right)^{-1} - 1 = \left(\frac{1,0425}{1,0532}\right)^{-1} - 1 = 1,010264 - 1 = 0,010264 = \frac{R_t}{100}$$

obteniéndose en ambos casos la misma rentabilidad, R =1,0264%.

La expresión $\left(\frac{1+r_t}{1+r_{t-1}}\right)^{-1}$ puede aproximarse en serie de Taylor alrededor de: $r_t = 0, r_{t-1} = 0$ mediante:

$$\left(\frac{1+r_t}{1+r_{t-1}}\right)^{-1} \simeq 1 + r_{t-1} - r_t$$

por lo que la variación diaria en el tipo de interés puede tomarse como una aproximación a la rentabilidad de la inversión en el activo que cotiza en tipos de interés.

Las expresiones $\left(\frac{(1+r_t)^a}{(1+r_{t-1})^a}\right)^{-1}$ y $\left(\frac{1+br_t}{1+br_{t-1}}\right)^{-1}$ pueden aproximarse ambas mediante:

$$\left(\frac{(1+r_t)^a}{(1+r_{t-1})^a}\right)^{-1} \simeq 1 + a(r_{t-1} - r_t);\ \left(\frac{1+br_t}{1+br_{t-1}}\right)^{-1} \simeq 1 + b(r_{t-1} - r_t)$$

Puesto que $a > 1$ y $b < 1$, es evidente que una misma serie de tipos de interés genera rentabilidades con una mayor varianza cuanto mayor es el vencimiento.

Más adelante veremos que las sensibilidades de los tipos de interés a los factores de riesgo que puedan establecerse vienen medidas en puntos básicos. Por eso, es habitual trabajar con variaciones diarias de los tipos de interés (que serán generalmente estacionarias) e interpretar las volatilidades en puntos basicos, no como porcentajes.

3.3 El supuesto de rendimientos lognormales

Se dice que una variable aleatoria X, definida sobre el subespacio de números reales positivos, sigue una distribución de probabilidad Lognormal cuando la

variable aleatoria que se obtiene como su logaritmo neperiano, $Y = ln(X)$, sigue una distribución Normal(μ,σ^2). En tal caso, la función de densidad de Y es:

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}},\ -\infty < y < \infty$$

y la función de densidad de X,

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}},\ x > 0$$

La esperanza y varianza de X son:

$$E(X) = e^{\mu+\frac{1}{2}\sigma^2};\ Var(X) = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right)$$

Es habitual suponer que el proceso seguido por el precio o cotización de un activo es tal que el rendimiento porcentual bruto correspondiente a un período sigue una distribución lognormal, es decir, que su logaritmo, el tipo continuo, tiene una distribución Normal:

$$r_t = \ln(1 + R_t) \sim N\left(\mu, \sigma^2\right)$$

Una ventaja de suponer una distribución lognormal para el rendimiento porcentual es que asegura que $1 + r_t$ sea no negativo, lo que no ocurriría si supusiéramos Normalidad de R_t.

Pero conviene recordar que la distribución lognormal no es simétrica de modo que bajo este supuesto, el tamaño medio de las rentabilidades por encima de la media es superior al promedio de las rentabilidades por debajo de la media.

Bajo este supuesto, $1 + R_t$ sigue una distribución lognormal, por lo que la esperanza y varianza de la rentabilidad porcentual R_t , en función de los momentos de la rentabilidad logarítmica, son:

$$\begin{aligned} E(1 + R_t) &= e^{\mu+\frac{1}{2}\sigma^2};\ E(R_t/100) = e^{\mu+\frac{1}{2}\sigma^2} - 1 \\ Var(1 + R_t) &= Var(R_t/100) = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right) \end{aligned}$$

estas fórmulas son muy útiles para obtener predicciones a partir de modelos estimados para los logaritmos de los rendimientos, pues si μ es la predicción para el logaritmo del rendimiento y σ^2 es la varianza condicional estimada para dicho logaritmo del rendimiento (es decir, la varianza de la innovación del proceso para el logaritmo del rendimiento), entonces la predicción para el propio rendimiento y la varianza asociada, que nos servirá para construir intervalos de confianza para dicha predicción, se obtienen a partir de las expresiones anteriores.

En la otra dirección, los momentos de la rentabilidad logarítmica en función de m_1 y m_2, la esperanza y varianza del proceso de rentabilidades R_t, son,

$$E(r_t) = \ln\left(\frac{m_1+1}{\sqrt{1+\frac{m_2}{[1+m_1]^2}}}\right);\ Var(r_t) = \ln\left(1+\frac{m_2}{[1+m_1]^2}\right)$$

Nótese que aparece $1+m_1$ porque r_t es el logaritmo de $1+R_t$.

En el modelo de Black y Scholes se supone que el precio del activo subyacente sigue un proceso de difusión lognormal:

$$dP_t = \mu P_t dt + \sigma P_t dW_t$$

y aplicando Ito:

$$d\ln P_t = \left(\mu - \frac{1}{2}\sigma^2\right)dt + \sigma dW_t$$

por lo que la rentabilidad logaritmica de una inversión en el activo subyacente tiene una media de $\mu - \frac{1}{2}\sigma^2$ y una varianza de σ^2. Por tanto, podemos asociar la volatilidad de la formula de Black y Scholes con la volatilidad de la rentabilidad logarítmica. No podemos hacer lo mismo con la volatilidad de la rentabilidad porcentual. De hecho, utilizando las relaciones anteriores, tenemos: $\sigma^2 = \ln\left(1+\frac{m_2}{[1+m_1]^2}\right)$, por lo que m_2 es únicamente una aproximación a σ^2. Con datos frecuentes, m_1 será pequeño y puede aproximarse por cero. Por otra parte la varianza diaria tambien será numericamente reducida, por lo que la aproximación en serie de Taylor a $\ln(1+m_2)$ sería m_2. Como puede verse, m_2 es solo una aproximación a σ^2, por lo que no debe utilizars ene la expresión de valoracion de opciones de Black Scholes, pues el pequeño error de aproximación entre ambas varianzas puede dar lugar a diferencias notables en los precios de las opciones a que dan lugar.

3.4 Contrastes de Normalidad

Bera y Jarque propusieron el contraste de Normalidad que lleva su nombre, que utiliza los coeficientes de asimetría AS y de curtosis K:

$$BJ = T\left(\frac{AS^2}{6} + \frac{(K-3)^2}{24}\right)$$

que se distribuye como una chi-cuadrado con 2 grados de libertad.

Este es un contraste paramétrico de la hipótesis de Normalidad, existiendo asimismo varios contrastes no paramétricos, quizá más aconsejables:

- el contraste de Kolmogorov-Smirnov, que se basa en el supremo de los valores absolutos de las diferencias entre la función de distribución empírica y la función de distribución teórica de una variable Normal de esperanza y varianza iguales a las muestrales. Para ello se divide el rango observado en intervalos pequeños, y se comparan los valores de ambas funciones en uno de los extremos de cada intervalo.

- el contraste chi-cuadrado o de Pearson, basado en la comparación de las frecuencias teórica y empírica en cada uno de los subintervalos en que se ha dividido previamente el rango de valores observados.

Al igual que muchos contrastes cuyo estadístico hace intervenir al tamaño muestral de modo multiplicativo, el contraste de Bera-Jarque tiene una peculiaridad, y es que para tamaños muestrales elevados, el estadístico del contraste toma un valor alto, que puede conducir al rechazo de la hipótesis nula en "demasiadas ocasiones". Dicho de otro modo, para muestras grandes, el contraste tiene un tamaño muy superior al teórico.

3.5 Extrapolación temporal de la varianza

En el análisis financiero, la estimación de la volatilidad obtenida a lo largo de un período de tiempo breve suele extrapolarse a un determinado período de referencia, generalmente anual. La anualización de la volatilidad permite comparar el riesgo de varios activos, independientemente del intervalo de tiempo considerado en su análisis.

La anualización puede conseguirse a partir de la volatilidad calculada para cada período de una determinada frecuencia, sin más que multiplicar por la raíz cuadrada del número de datos de dicha frecuencia que hay en un año.

La volatilidad de una rentabilidad continua no satisface la misma propiedad de extrapolación temporal que antes vimos para la rentabilidad. En efecto, la varianza de una rentabilidad sobre dos períodos es:

$$Var\left(\frac{r_t^2}{100}\right) = Var\left(\frac{r_t^1}{100} + \frac{r_{t-1}^1}{100}\right) = Var\left(\frac{r_t^1}{100}\right) + Var\left(\frac{r_{t-1}^1}{100}\right) + 2Cov\left(\frac{r_t^1}{100}, \frac{r_{t-1}^1}{100}\right)$$

por lo que la varianza de la rentabilidad durante un período amplio no es igual a la suma de las varianzas de las rentabilidades durante los períodos más cortos comprendidos en el intervalo amplio. La diferencia entre ambos cálculos estriba en que el segundo ignora las covarianzas entre cada par de rentabilidades sobre períodos cortos. Por tanto, si dichas rentabilidades fuesen independientes, sus covarianzas serían nulas, y tendríamos que la varianza sobre el horizonte largo sería igual a la varianza de las rentabilidades sobre los períodos cortos.

La práctica habitual es la siguiente: si se utilizan datos diarios, y σ^2 denota una estimación de la variabilidad diaria (varianza u otra medida), entonces se toma $252\sigma^2$ como estimación de la variabilidad (varianza) anual (252 es el número aproximado de días de mercado dentro de un año), y $\sqrt{252}\sigma$ como estimación de la volatilidad anual. Con datos semanales, la volatilidad anual se obtiene a partir de la desviación típica de los datos semanales mediante: $\sqrt{52}\sigma$, mientras que si se dispone de datos mensuales, la volatilidad típica anual se obtiene a partir de la desviación típica de los datos mensuales mediante: $\sqrt{12}\sigma$. Se procede de igual modo si se trabaja con indicadores de volatilidad alternativos a la desviación típica. Una vez obtenido un indicador de volatilidad, se extrapolaría a una medida anual del modo que acabamos de describir.

En general, dada una desviación típica calculada con datos de una determinada frecuencia, si queremos obtener la estimación de la desviación típica sobre un intervalo de tiempo que comprende N observaciones de las utilizadas en el cálculo de dicha desviación típica, multiplicamos por $\sqrt{N}$. Esto es lo que hicimos en el párrafo anterior.

Así, si hemos estimado σ^2 con datos diarios, entonces: la Volatilidad semanal se estima por: $\sqrt{5}\sigma$, la Volatilidad mensual se estima por: $\sqrt{21}\sigma$, y la Volatilidad anual se estima por: $\sqrt{252}\sigma$.

De modo más general, si hemos calculado rentabilidades utilizando datos observados cada n períodos, y hemos estimado su varianza σ_1^2, la volatilidad sobre T períodos se estima mediante $\sigma_T^2 = \frac{T}{n}\sigma_n^2$. Las covarianzas se convierten en términos anuales utilizando los mismos factores de escala que las varianzas.

Example 1 *Si la varianza de las rentabilidades diarias de un activo es 0,001, y considerando 250 dias de riesgo (mercado) a lo largo del año, la volatilidad del activo sería 50%. Si la volatilidad anual del activo es 36%, la desviación típica de sus rentabilidades semanales es 0,05.*

Como ya discutimos en la Sección 3.d, esta práctica habitual de extrapolar una estimación de la volatilidad a un intervalo amplio de tiempo es aplicable en rigor sólo al cálculo de la *volatilidad de rentabilidades continuas*, y se basa en la hipótesis de que: 1) los datos básicos utilizados, ya sean rentabilidades mensuales, diarias, horarias, etc. son *independientes; 2)* también supone que la varianza es constante en el tiempo. Es importante observar que no hay que hacer ningun supuesto sobre el carácter de la distribución de rentabilidades. En particular, para esta extrapolación temporal de la varianza no es preciso que la rentabilidades sean Normales.

Si se está calculando la varianza de las rentabilidades, deben ser independientes éstas, no necesariamente los precios o cotizaciones que las generaron. Esto se corresponde con la extendida idea de que el *logaritmo del precio* de un activo financiero tiene una estructura estocástica de camino aleatorio. En tal caso, la *rentabilidad* de dicho activo, definida como la primera diferencia del logaritmo del precio, es un *ruido blanco*. Es decir, la serie temporal de rentabilidades obedece a un proceso formado por variables aleatorias independientes e idénticamente distribuidas, y el método de extrapolación de la varianza es correcto.

Varianza de rentabilidades porcentuales

Recordemos, además, que la suma de variables aleatorias Normales, independientes o no, sigue asimismo una distribución de probabilidad Normal. Por tanto, si suponemos que las rentabilidades continuas durante un período son independientes y obedecen a la misma distribución

Normal, tendremos, a lo largo de T períodos:

$$\frac{r_t + r_{t-1} + r_{t-2} + ... + r_{t-T+1}}{100} \sim N(T\mu, T\sigma^2)$$

de modo que la rentabilidad porcentual (o simple) a lo largo del intervalo de tiempo $(t-T, t)$ tiene por esperanza y varianza:

$$E\left(\frac{R_t}{100}\right) = e^{T\mu+\frac{1}{2}T\sigma^2} - 1;\ Var\left(\frac{R_t}{100}\right) = e^{2T\mu+T\sigma^2}\left(e^{T\sigma^2} - 1\right)$$

Agregación temporal de varianza con rentabilidades autocorrelacionadas

Si las rentabilidades no fuesen independientes a lo largo del tiempo, su varianza no sería tan sencilla como $T\sigma^2$. Como antes, un análisis similar aplica a intervalos de tiempo n y m, con $n = km$. Cuando las rentabilidades están autocorrelacionadas, la acumulación temporal de varianzas del modo que hemos descrito es un estimador sesgado del riesgo. En el caso de tipos de interés, existe generalmente elevada autocorrelación positiva, mientras que en rentabilidades bursátiles diarias de valores individuales se detecta, en ocasiones, autocorrelación negativa. Como la varianza de una suma de variables es igual a la suma de varianzas más el doble de su covarianza, tenderemos a subestimar la varianza de la rentabilidad sobre el horizonte temporal amplio en el caso de autocorrelación positiva (creeremos que, sobre el período amplio, la rentabilidad es menos volátil de lo que realmente es), y a sobre-estimarla en el caso de autocorrelación negativa (creeremos que es más volátil de lo que realmente es).

Si la rentabilidad de un activo tiene una estructura de autocorrelación representada por un proceso autoregresivo de primer orden (AR(1), $r_t = \alpha + \rho r_{t-1} + \varepsilon_t$, con $\varepsilon_t \sim N(0, \sigma_\varepsilon^2), i., i.d.$, y $|\rho| < 1$, la varianza de la rentabilidad continua sobre h periodos: $r_{ht} = \sum_{i=0}^{h-1} r_{t+i}$ es:

$$Var(r_{ht}) = \sum_{i=0}^{h-1} Var(r_{t+i}) + 2\sum_{i\neq j}^{h-1} Cov(r_{t+i}, r_{t+j}) = \sigma_\varepsilon^2\,(h+2)\sum_{i=0}^{h-1}(h-i)\rho^i$$

y utilizando la identidad:

$$\sum_{i=1}^{n}(n-i+1)x^i = \frac{x}{(1-x)^2}[n(1-x) - x(1-x^n)],\ |x| < 1.$$

Si hacemos $x = \rho$ y $n = h-1$, tenemos:

$$Var(r_{ht}) = \sigma_\varepsilon^2\left(h + 2\sum_{i=1}^{h-1}(h-1)\rho^i\right) = \sigma_\varepsilon^2\left(h + 2\frac{\rho}{(1-\rho)^2}[(h-1)(1-\rho) - \rho(1-\rho^{h-1})]\right)$$

Esta expresión nos puede dar pautas acerca de la posible infraestimación de la volatidad en presencia de rentabilidades positivamente autocorrelacionadas. Pero está basada en el supeusto de Normalidad y en una determinada estructura de autocorrelación, y no seria estrictamente aplicable en otra situación.

Example 2 *Suponga que las rentabilidades mensuales de un hedge fund sobre los tres últimos años tienen una desviación típica de 5% ¿cual sería su estimacion de la volatilidad del fondo bajo el supuesto de rentabilidades independientes? ¿y bajo el supuesto de que tengan una autocorrelación de 0,25? R=17,32%; 21,86%.*

Esta consideración sugiere, por tanto, que un modo de contrastar la independencia de rentabilidades consiste en analizar si la varianza muestral aumenta linealmente con la amplitud de la ventana muestral utilizada en su cálculo. En variables con covariación positiva, al agregar temporalmente tendremos un crecimiento más que lineal de la varianza, y lo contrario ocurrirá bajo covariación negativa. En definitiva, la agregación de volatilidades descansa en la independencia temporal de las rentabilidades del activo en cuestión, lo que equivale a que el precio de dicho activo obedezca a una estructura de camino aleatorio. Existen distintos enfoques estadísticos para el contraste de dicha hipótesis, que pueden verse en la notas de clase correspondientes al tema de Series Temporales.

La existencia de autocorrelación en rentabilidades no es necesariamente un resultado negativo, pues ofrece la posibilidad de predecir rentabilidades. Por esta razón, sería contraria a la noción de mercados eficientes, según la cual, el precio actual resume toda la información relevante acerca del precio futuro, lo que equivale a decir que el precio sigue un comportamiento de camino aleatorio.

3.6 Desviación típica de los estimadores de la varianza, la desviación típica y el coeficiente de correlación lineal

3.6.1 Intervalo de confianza para la varianza de una población Normal

Si la población de la que se extrae una muestra aleatoria simple es Normal, con esperanza μ y varianza σ^2, ambas constantes, y $\hat{\sigma}^2$ denota la cuasi-varianza muestral, el cociente $\frac{(T-1)\hat{\sigma}^2}{\sigma^2}$ sigue una distribución χ^2_{T-1}. Por tanto, si observamos una muestra de 25 observaciones sucesivas de una rentabilidad que estamos dispuestos a suponer que evoluciona independientemente en el tiempo, y calculamos una cuasi-varianza muestral de 12,5, tendremos que $\frac{(24)(12,5)}{\sigma^2}$se distribuye como una χ^2_{24}.

Por tanto, tendremos:

$$0,95 = P\left[12,4 \leq \frac{(24)(12,5)}{\sigma^2} \leq 39,4\right] = P\left(7,61 \leq \sigma^2 \leq 24,19\right)$$

un intervalo no muy preciso, que tendríamos que tener en cuenta al establecer nuestras conclusiones acerca de la volatilidad de un mercado. Por supuesto, que el número de datos utilizados es muy importante para la precisión de la estimación y, como consecuencia, para la amplitud del intervalo de confianza. Si la cuasi-varianza de 12,5 hubiese sido obtenida a partir de 10 datos, entonces $\frac{125}{\sigma^2}$se distribuiría como una χ^2_{10}, y tendríamos:

$$0,95 = P\left[3,25 \leq \frac{125}{\sigma^2} \leq 20,5\right] = P\left(6,10 \leq \sigma^2 \leq 38,46\right)$$

Example 3 *EII.3.8: Suponga que las rentabilidades diarias del FTSE100 siguen una distribución Normal. Utilice la muestra de 11 datos diarios: {6038.3,*

6219.0, 6143.5, 6109.3, 5858.9, 6064.2, 6078.7, 6086.1, 6196.0, 6196.9, 6220.1} para construir un intervalo de confianza del 95% para la varianza de las rentabilidades. R: (0,000211; 0,001333) .

3.6.2 Desviación tipica de la varianza

Los cuantiles de una distribución de probabilidad son invariantes por transformaciones estrictamente monótonas. De modo que si f es una transformación estrictamente monótona de una variable aleatoria X, tenemos:

$$P(c_1 < X < c_2) = P(f(c_1) < f(X) < f(c_2))$$

Example 4 *Por tanto, para obtener el intervalo de confianza del 95% para la volatilidad a partir de la estimación de la varianza del ejemplo anterior, tendriamos:*

$$0,95 = P\left(\sqrt{250(0,000211)} < volatilidad < \sqrt{250(0,001333)}\right) = (23,0\%; 57,7\%)$$

Este ejemplo, asi como el analisis que antes hicimos para la distribución de probabilidad del estimador de la varianza, son válidos en muestras finitas. Vamos a presentar ahora otro análisis que es válido bajo Normalidad y que puede verse como el caso limite al cual converge el resultado que antes obtuvimos.

De la definición de varianza, suponiendo que el periodo de tiempo es suficientemente pequeño como para considerar que la rentabilidad media es cero, y que las rentabilidades son independientes en el tiempo e igualmente distribuidas tenemos, tomando varianzas:

$$\hat{\sigma}^2 = T^{-1}\sum_{t=1}^{T} r_t^2 \Rightarrow Var(\hat{\sigma}^2) = T^{-2}\sum_{t=1}^{T} Var(r_t^2)$$

Como $Var(X) = E(X^2) - (E(X))^2$,tenemos: $Var(r_t^2) = E(r_t^4) - \left[E(r_t^2)\right]^2$ .Pero: $E(r_t^2) = \sigma^2$. Por otra parte, si las rentabilidades siguen una distribución Normal, entonces: $E(r_t^4) = 3\sigma^4$. Por tanto,

$$Var(\hat{\sigma}^2) = T^{-2}\sum_{t=1}^{T}\left[3\sigma^4 - (\sigma^2)^2\right] = 2\frac{\sigma^4}{T}$$

por lo que en una muestra de rentabilidades con esperanza matemática cero, independientes y con distribución Normal, la desviación típica de la varianza muestral es:

$$DT(\hat{\sigma}^2) = \sqrt{\frac{2}{T}}\sigma^2$$

o, en términos porcentuales:

$$\frac{DT(\hat{\sigma}^2)}{\sigma^2} = \sqrt{\frac{2}{T}}$$

Por ejemplo, si $T = 32$,entonces la desviación típica es un 25% de la varianza poblacional que queremos estimar. Si $T = 200$, la desviación típica será un 10% de la varianza.

Este mismo resultado puede obtenerse volviendo al resultado: $\frac{(T-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{T-1}$. Si la muestra es suficientemente larga ($T > 20$), la distribución chi-cuadrado con k grados de libertad converge a una Normal$(k, 2k)$, por lo que: $\frac{(T-1)\hat{\sigma}^2}{\sigma^2} \to N(T, 2T)$, y tendremos: $\hat{\sigma}^2 \sim N\left(\sigma^2, \sigma^4 \frac{2}{T}\right)$, y: $DT(\hat{\sigma}^2) = \sqrt{\frac{2}{T}}\sigma^2$.

3.6.3 Desviación típica del estimador de la desviación típica

Veamos ahora como calcular *la desviación típica del estimador de la desviación típica*. Esto no es un juego de palabras: necesitamos estimar la desviación típica para estimar la volatilidad de una determinada rentabilidad. Y la desviación típica del estimador es la medida de la precisión con que hemos estimado la volatilidad, cuestión sin duda de toda importancia para el analista de riesgos como para el gestor de carteras. Podriamos pensar que basta con tomar la raiz cuadrada de $DT(\hat{\sigma}^2)$ que acabamos de calcular, para obtener la desviación típica de $\hat{\sigma}$. Pero no es tan sencillo, debido a la desigualdad de Jensen.[2]

Aproximando $f(x)$ alrededor de $E(x) = \mu_x$, tenemos:

$$f(x) \simeq f(\mu) + f'(\mu)(x-\mu) + \frac{1}{2}f''(\mu)(x-\mu)^2 + ...$$

por lo que, tomando esperanzas:

$$\begin{aligned} E(f(x)) &\simeq f(\mu) + \frac{1}{2}f''(\mu)Var(x) \\ E(f(x)^2) &\simeq f(\mu)^2 + \left[f'(\mu)^2 + f(\mu)f''(\mu)\right]Var(x) + ... \end{aligned}$$

y, despreciando términos en $Var(x)^2$ que, en datos diarios de rentabilidades, serán muy pequeños, tenemos:[3]

$$Var(f(x)) \simeq E(f(x)^2) - [E(f(x))]^2 = (f'(\mu_x))^2 Var(x) = [f'(E(x))]^2 Var(x)$$

[2] Si x es una variable aleatoria y φ es una función convexa, entonces: $\varphi(E(x) \leq E(\varphi(x))$. Consideremos: $x^2 = \varphi(x)$,una función convexa. La desigualdad de Jensen implica: $(Ex)^2 \leq E(x^2)$. Es decir, la esperanza del cuadrado de una variable aleatoria es mayor o igual que el cuadrado de la esperanza de dicha variable.

Consideremos ahora: $\sqrt{x}$, una función cóncava. Por tanto, $-\sqrt{x} = \varphi(x)$ es una función convexa. En consecuencia: $-\sqrt{Ex} \leq E(-\sqrt{x}) = -E(\sqrt{x})$, es decir: $\sqrt{Ex} \geq E(\sqrt{x})$

[3] Estamos despreciando términos en $E(x-\mu)^3$ y $E(x-\mu)^4$ que serán igual a los coeficientes de asimetría y curtosis multiplicados por $Var(x)^3, Var(x)^4$, que en rentabilidades frecuentes deberian ser muy pequeños.

Cuando $y = f(x) = \sqrt{x}$, tenemos:

$$Var(y) = Var(\sqrt{x}) = \left(\frac{1}{2\sqrt{\mu_x}}\right)^2 Var(x)$$

Por tanto, en particular, si $y = \hat{\sigma}, x = \hat{\sigma}^2$ denotan la cuasi-desviación típica y cuasivarianza muestrales, tendremos $E(x) = E(\hat{\sigma}^2) = \sigma^2$:

$$Var(\hat{\sigma}) \approx \frac{1}{4\sigma^2} Var(\hat{\sigma}^2) = \frac{1}{4\sigma^2}\frac{2\sigma^4}{T} = \frac{\sigma^2}{2T}$$

por lo que,

$$DT(\hat{\sigma}) \approx \sqrt{\frac{1}{2T}}\sigma$$

y, como porcentaje de la volatilidad:

$$\frac{DT(\hat{\sigma})}{\hat{\sigma}} = \sqrt{\frac{1}{2T}}$$

Otra aplicación del resultado anterior sería: $Var(\ln(x)) \simeq \frac{1}{\mu_x^2} Var(x)$. Por ejemplo, sabemos que si $x \sim \log Normal(\mu, \sigma^2)$, entonces $\ln(x) \sim N(\mu, \sigma^2)$. Si utilizamos la aproximación anterior, tendremos: $Var(\ln(x)) \simeq \frac{1}{\left(e^{\mu+\frac{1}{2}\sigma^2}\right)^2} e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right) = \left(e^{\sigma^2}-1\right) \simeq \left[1+\sigma^2+\frac{1}{2!}\left(\sigma^2\right)^2+\frac{1}{3!}\left(\sigma^2\right)^3+...-1\right] \simeq \sigma^2$ si la varianza es pequeña. Asimismo, $Var(x^2) \simeq (2\mu_x)^2 Var(x)$.

Example 5 *Utilizando rentabilidades mensuales del índice S&P500 para el período 9/1994 a 9/2014,* $T = 241$, *observamos:* $\hat{\mu} = 0,60\%$ *y* $\hat{\sigma} = 5,17\%$. *Sus desviaciones típicas son:* $DT(\hat{\mu}) = \sigma/\sqrt{T} = (5,17\%)/\sqrt{241} = 0,33\%$ *y* $DT(\hat{\sigma}) = \sigma/\sqrt{2T} = (5,17\%)/\sqrt{482} = 0,235\%$, *por lo que la rentabilidad media es apenas significativamente diferente de cero, debido a que se mide con poca precisión, mientras que la desviación típica se estima con bastante precisión.*

Obsérvese que ni se cumple: $E(\hat{\sigma}) = \sqrt{E(\hat{\sigma}^2)}$, debido a la desigualdad de Jensen, ni tampoco se cumple: $Var(\hat{\sigma}) = \sqrt{Var(\hat{\sigma}^2)}$. En efecto, si para estimar la desviación típica calculamos la raíz cuadrada de la cuasivarianza muestral, no podemos afirmar que dicha estimación sea insesgada, puesto que de acuerdo con la desigualdad de Jensen: $E\left(\sqrt{(\hat{\sigma}^2)}\right) \leq \sqrt{E(\hat{\sigma}^2)} = \sqrt{\sigma^2} = \sigma$.

3.6.4 Desviación típica del coeficiente de correlación

Para el coeficiente de correlación tenemos que, si las rentabilidades de dos activos carecen de autocorrelación y son idénticamente distribuidas, con distribución Normal y media cero, entonces el coeficiente de correlación, dividido por su

desviación típica tiene una distribución t-Student con T grados de libertad. Por otro lado, su varianza es:[4]

$$Var(\hat{\rho}) = (T-2)^{-1}(1-\rho^2)$$

de modo que:

$$\frac{\hat{\rho}\sqrt{T-2}}{\sqrt{1-\hat{\rho}^2}} \sim t_T$$

Example 6 *Una correlación lineal de 0,20, calculada con 36 rentabilidades de dos activos ¿es significativamente distinta de cero? R: No.*

3.7 Funcion generatriz de momentos. Funcion caracteristica

Funcion generatriz de momentos

La función generatriz de momentos $M_X(t)$ de una variable aleatoria X se define:

$$M_X(t) = E(e^{tX})$$

siempre que dicha esperanza matemática exista. $M_X(0)$ existe siempre y es igual a 1.

Utilizando el desarrollo en serie de la función exponencial:

$$e^{tX} = 1 + tX + \frac{t^2X^2}{2!} + \frac{t^3X^3}{3!} + ... + \frac{t^nX^n}{n!} + ...$$

tenemos:

$$\begin{aligned} M_X(t) &= E(e^{tX}) = 1 + tE(X) + \frac{t^2E(X^2)}{2!} + \frac{t^3E(X^3)}{3!} + ... + \frac{t^nE(X^n)}{n!} + ... = \\ &= 1 + tm_1 + \frac{t^2m_2}{2!} + \frac{t^3m_3}{3!} + ... + \frac{t^nm_n}{n!} + ... \end{aligned}$$

por lo que es facil probar para los momentos respecto al origen que:

$$E(X^a) = \frac{d^a(M_X(t))}{dt^a} \mid_{t=0}$$

Cuando no sea necesario, ignoraremos el subindice X en $M_X(t)$.

Para el cálculo de momentos respecto de la media a partir de la Función Generatriz, puede tenerse en cuenta que,

[4] Si suponemos que la esperanza matemática de ambas variables es igual a cero, entonces el número de grados de libertad es T, ya que en ese caso no hay que restar 2 por el número de parametros que es preciso estimar antes de calcular el coeficiente de correlación.

$$
\begin{aligned}
E\left[(X-\mu)^2\right] &= E\left(X^2\right) - 2E\left(X\right)\mu + \mu^2 \\
E\left[(X-\mu)^3\right] &= E\left(X^3\right) - 3E\left(X^2\right)\mu + 3E\left(X\right)\mu^2 - \mu^3 = E\left(X^3\right) - 3E\left(X^2\right)\mu + 2\mu^3 = \\
&= M'''(0) - 3M''(0)M'(0) + 2\left[M'(0)\right]^3 \\
E\left[(X-\mu)^4\right] &= E\left(X^4\right) - 4\mu E\left(X^3\right) + 6E\left(X^2\right)\mu^2 - 4E\left(X\right)\mu^3 + \mu^4 = \\
&= M''''(0) - 4M'''(0)M'(0) + 6M''(0)\left[M'(0)\right]^2 - 4\left[M'(0)\right]^3 M'(0) + \left[M'(0)\right]^4
\end{aligned}
$$

En el caso de una Normal $N(\mu, \sigma^2)$, tenemos $M(t) = \exp(\mu t + 1/2\sigma^2 t^2)$, con

$$
\begin{aligned}
M'(t) &= (\mu + \sigma^2 t)M(t) \\
M''(t) &= \left[\sigma^2 + (\mu + \sigma^2 t)^2\right] M(t) \\
M'''(t) &= \left[3\sigma^2 + (\mu + \sigma^2 t)^2\right] (\mu + \sigma^2 t)M(t) \\
M''''(t) &= \left[6\sigma^2(\mu + \sigma^2 t)^2 + 3\sigma^4 + (\mu + \sigma^2 t)^4\right] M(t)
\end{aligned}
$$

por lo que

$$
\begin{aligned}
M(0) &= 1;\ M'(0) = \mu;\ M''(0) = \mu^2 + \sigma^2; \\
M'''(0) &= 3\mu\sigma^2 + \mu^3;\ M''''(0) = 6\mu^2\sigma^2 + 3\sigma^4 + \mu^4
\end{aligned}
$$

de donde se prueba con facilidad que el coeficiente de asimetría de toda $N(\mu, \sigma^2)$ es igual a cero, mientras que su coeficiente de curtosis es igual a 3.

La función característica de una variable aleatoria X se define:

$$\varphi_X(t) = E(e^{itX}) = M_{iX}(t) = M_X(it)$$

Una función $\varphi : R^n \longrightarrow C$ es función característica de alguna variable aleatoria si y solo si es definida positiva, continua en el origen, y $\varphi(0) = 1$ [Teorema de Bochner].

La función característica es la transformada de Fourier de la función de densidad de la variable aleatoria X. La función característica *caracteriza* una determinada distribución de probabilidad. Si calculamos la función característica de una determinada variable aleatoria y reconocemos dicha función característica como la de una determinada distribución de probabilidad, habremos probado que la variable aleatoria de la que partimos tiene dicha distribución. La función caracteristica siempre existe, incluso en casos cuando la función generatriz de momentos no existe.

Para un vector de variables, la función característica se definede igual modo, pero con t siendo un vector de parámetros: $\varphi(t) = E(e^{it^T X})$. La función característica de una distribución Normal multivariante $N(\mu, \Sigma)$ es:

$$\varphi(t) = e^{it\mu - \frac{1}{2}t^T \Sigma t}$$

Si la función caracteristica es diferenciable de orden k en el origen, entonces la variable aleatoria X tiene momentos hasta de orden k (si k es par), o hasta orden $k-1$ (si k es impar), y se tiene:

$$E(X^k) = (-i)^k \varphi_X^{(k)}(0)$$

Si $X_1, X_2, ..., X_n$ son variables aleatorias independientes, no necesariamente igualmente distribuidas, entonces:

$$\varphi_{a_1X_1+a_2X_2+...+a_kX_k} = \varphi_{X_1}(a_1t)...\varphi_{X_k}(a_kt)$$

Por ejemplo, la función característica de una distribución Gamma $\Gamma\ (\theta, k)$ es: $\varphi(t) = (1-\theta it)^{-k}$. Consideremos dos variables aleatorias X_1, X_2, con distribucion de probabilidad Gamma con igual parámetro θ, independientes entre sí. Si queremos conocer la distribucion de probabilidad de la suma, calculamos la función característica:

$$\varphi_{X_1+X_2} = \varphi_{X_1}.\varphi_{X_2} = (1-\theta it)^{-k_1}(1-\theta it)^{-k_2} = (1-\theta it)^{-k_1-k_2}$$

lo que demuestra que la suma de ambas variable sigue una distribución de probabilidad $\Gamma(\theta, k_1+k_2)$.

La función característica de X es el complejo conjugado de la transformada continua de Fourier de la función de densidad $f_X(x)$:

$$\varphi_X(x) = \int_R e^{itx} f(x)dx = \overline{\left(\int_R e^{-itx} f(x)dx\right)} = \overline{P(t)}$$

donde $P(t)$ denota la transformada de Fourier continua de la función de densidad $f_X(x)$.

De modo equivalente, cuando X es escalar, si la función caracteristica $\varphi_X(t)$ es integrable, entonces la función de distribución F_X es absolutamente continua y por tanto la función de densidad de X puede recuperarse a partir de $\varphi_X(x)$ mediante la inversa de la transformada de Fourier:

$$f_X(x) = F'_X(x) = \frac{1}{2\pi}\int_R e^{itx} P(t)dt = \frac{1}{2\pi}\int_R e^{itx}\overline{\varphi_X(t)}dt$$

4 Otras medidas de volatilidad

4.1 Volatilidad implícita versus volatilidad histórica

La *volatilidad implícita* es la estimación de volatilidad que se obtiene a partir de precios observados en el mercado para derivados que cotizan en volatilidad, como las opciones out y call, entre otros. Para ello, necesitaremos un modelo de valoración de dicho derivado. Fijando el precio teórico igual al precio observado en el mercado, podremos invertir dicha expresión y obtener una estimación de la volatilidad implicita del activo subyacente.

Puesto que las fórmulas teóricas de valoración de un producto derivado son funciones altamente no lineales de sus argumentos y, en particular, de la volatilidad, la resolución de la ecuación que iguala el precio teórico (es decir, el que se obtiene de la fórmula) con el precio observado en el mercado para obtener la volatilidad no puede llevarse a cabo analíticamente, siendo preciso recurrir a algoritmos numéricos, del tipo de los que analizaremos al estudiar la estimación de modelos no lineales.

Al efectuar este ejercicio, se está suponiendo que el modelo teórico de valoración del activo es correcto, y que el mercado forma expectativas de volatilidad utilizando eficientemente la información de que dispone. Ello hace que el precio de mercado resuma de manera adecuada toda la información disponible acerca del activo.

Estamos interesados en obtener volatilidades implícitas por dos razones:

- Una vez determinada la *volatilidad cotizada* en el mercado para un determinado subyacente, podremos poder evaluar si una determinada opción está subvalorada, correctamente valorada o sobrevalorada por el mercado, lo que podría sugerir diversas estrategias de inversión, y

- Generalmente, nos interesa utilizar la volatilidad implícita en un sentido temporal, pues si podemos obtener buenas previsiones de la volatilidad implícita futura de un determinado activo, dispondremos de previsiones de precios futuros de las opciones sobre dicho subyacente. En esta línea, pueden establecerse diversos ejercicios:

a) obtener un indicador de la *percepción del mercado* acerca de la volatilidad de un activo y analizar el modo en que dicha percepción cambia en el tiempo y cuáles son sus determinantes,

b) utilizar la *serie temporal* de la volatilidad implícita para especificar un modelo univariante predictivo de la *volatilidad implícita futura;* para ello necesitaremos haber calculado la volatilidad implícita durante todos los días a través de un largo período de tiempo,

c) establecer una relación entre volatilidad implicita y alguna de las medidas de *volatilidad histórica*: estas son las medidas que se basan exclusivamente en precios de mercado históricos del subyacente, sin utilizar modelo de valoración alguno, como ocurre con una desviación típica estimada a través de ventanas muestrales, o las medidas de Parkinson y Garman-Klass, que se basan en datos intradía, que luego veremos.

La volatilidad implícita no hace sino reflejar la visión del mercado acerca del grado de incertidumbre que entraña la evolución temporal de la rentabilidad que ofrece un activo. Cambios en la información disponible (resultados de una empresa, intervenciones de política económica, publicación de algún dato clave sorprendente) pueden incidir sobre tal percepción.

Existe una importante distinción entre ambos tipos de volatilidad: por un lado, tenemos l*a volatilidad histórica, que mira hacia el pasado*, y se basa exclusivamente en información histórica del precio o de la rentabilidad cuya volatilidad se pretende calcular. Por otra parte, *la volatilidad implícita* afecta a la

valoración de un producto derivado y, en consecuencia, *mira hacia el futuro*, tratando de estimar una característica no observable, por cuanto que aún no se ha realizado, como es la volatilidad futura del subyacente. La mayor diferencia estriba en la forma de calcular las volatilidades. La volatilidad histórica, que tiene la forma de una desviación típica; por otro lado, la volatilidad implicita se obtiene resolviendo una formula como la de Black-Scholes, de modo que el precio teórico resultante coincida con el observado en el mercado. Conviene recordar asimismo que la volatilidad implicita es una medida riesgo-neutro, a diferencia de la volatilidad histórica.

Sólo si pensáramos que la volatilidad futura es igual a la pasada estaríamos estimando el mismo concepto, si bien por método distintos, que nos proporcionarán valores numéricos diferentes. Por otra parte, la volatilidad histórica, calculada en forma de serie temporal a través de ventanas móviles, como describimos anteriormente, también podría utilizarse para *predecir* la volatilidad futura. Por tanto, ambos conceptos pueden ponerse en relación, aunque sus niveles medios pueden ser diferentes.

Una hipótesis interesante estriba en si la volatilidad implícita responde a variaciones en la volatilidad histórica. Si se produce una variación en la rentabilidad de un activo que modifica su volatilidad histórica, el mercado puede percibir un mayor riesgo futuro, lo que debería elevar el precio de las opciones sobre el mismo, conduciendo a una mayor volatilidad implícita. Este es un ejercicio empirico interesante, cuya respuesta no es evidente.

La existencia de una relación estadística estable entre volatilidad histórica e implícita no precisa que los niveles de volatilidad estimados por cada uno de los dos procedimientos coincidan. De hecho, esperamos más bien lo contrario; en todo caso, no importa que ambos niveles de volatilidad sean los mismos, sino que variaciones en el nivel de volatilidad histórica anticipen cambios en el nivel de volatilidad implícita, que puedan utilizarse para la gestión de carteras.

4.2 Utilización de información intradía en la medición de la volatilidad de un activo financiero

Si observamos los precios negociados de un activo a intervalos regulares de tiempo, podemos definir,

$$R_{t+j/m} = \ln(S_{t+j/m}) - \ln(S_{t+(j-1)/m})$$

donde suponemos m observaciones diarias, para estimar la varianza diaria,

$$\sigma^2_{m,t+1} = \sum_{j=1}^{m} R^2_{t+j/m}$$

que podría utilizarse, nuevamente, en la validación de modelos de previsión de volatilidad, en sustitución del cuadrado de la rentabilidad diaria, o utilizarse directamente en la predicción de volatilidad. Nótese que estamos acumulando los cuadrados de las rentabilidades observadas dentro de un mismo día, por lo que

no hay que promediar. Según aumenta el número de observaciones intradia m, la medida de varianza realizada anterior converge a la verdadera varianza diaria.

El uso de rentabilidades intradia se ve condicionado en el caso de activos poco líquidos por la imposibilidad de observar el precio de operaciones cruzadas con mucha frecuencia. Lo que obtenemos entonces no es el precio fundamental del activo, que no es observable, sino una secuencia de precios bid y ask [ver simulación en Figure 2.7 en Christoffersen]. Los precios diarios intradía pueden contener mucha volatilidad espúrea, que no existe en el precio fundamental del activo, por los rebotes observados en las transacciones entre precios bid y ask. Como consecuencia, las medidas de varianza realizada basadas en rentabilidades intradía pueden tener también este problema, especialmente en mercados poco líquidos. En un contexto de limitada liquidez, el máximo puede calcularse como el máximo realmente observado menos la mitad del spread bid-ask, mientras el mínimo es calculado como el mínimo realmente observado más la mitad del spread bid-ask. En ausencia de fricciones se estima que las medidas de varianza basadas en el rango diario de precios cruzados en el mercado contienen información equivalente únicamente a la contenida en 4 rendimientos horarios intradía. Lamentablemente, es difícil extender la idea a la estimación de covarianzas y correlaciones, a diferencia de lo que sucede con las medidas de varianza realizada como veremos más adelante.

Varianzas log-Normales

La hipótesis de Normalidad del logaritmo de $\sigma^2_{m,t+1}$ suele no rechazarse en datos intradía, por lo que podemos utilizar un modelo de predicción basado en la volatilidad realizada,

$$\ln \sigma^2_{m,t+1} = \alpha + \rho \ln \sigma^2_{m,t} + u_{t+1},\ con\ u_{t+1}/\Omega_t \sim N(0, \sigma^2_u)$$

Pero estamos interesados en la previsión del nivel de volatilidad, no de su logaritmo. Por tanto, conviene recordar que,

$$u_{t+1} \sim N(0, \sigma^2_u) \Rightarrow E\left(e^{u_{t+1}}\right) = e^{\sigma^2_u/2}$$

por lo que en un modelo autoregresivo como el anterior,

$$E_t \sigma^2_{m,t+1} = E_t e^{\alpha + \rho \ln \sigma^2_{m,t} + u_{t+1}} = e^{\alpha + \rho \ln \sigma^2_{m,t}} . E_t e^{u_{t+1}} = \left(\sigma^2_{m,t}\right)^{\rho} e^{\alpha + \sigma^2_u/2}$$

4.2.1 Estacionalidad intra-día en volatilidad

Tratar de caracterizar pautas de estacionalidad, tanto en rentabilidad como en volatilidad, puede producir información de enorme interés para un inversor. Ha sido muy popular durante mucho tiempo buscar efectos estacionales en las rentabilidades ofrecidas por los mercados de valores. Así, existe el denominado *efecto Enero*, mes en el que las Bolsas tienden a ofrecer una rentabilidad superior a la de otros meses, debido a la recomposición de carteras de muchos inversores, que liquidaron parte de las mismas antes de final de año por razones fiscales.

Asimismo, se ha debatido durante mucho tiempo la existencia de efectos estacionales entre semana o *efectos días de la semana*, afirmando algunos autores que existe *efecto lunes* en algunos mercados.

Menos estudiada ha sido la posible existencia de pautas estacionales en volatilidad. Evidentemente, la posible existencia de tales pautas sería asimismo un fenómeno muy a tener en cuenta por todos los que gestionan riesgo de uno u otro modo.

Parece, sin embargo, bastante probada la existencia de pautas "estacionales" de volatilidad intradía, que se reflejan en una mayor volatilidad en el período siguiente a la apertura del mercado, un descenso en las horas centrales del día, y un incremento posterior, según se acerca la hora de cierre.

A este perfil en forma de U de la volatilidad a lo largo del día de negociación suele venir unido un perfil similar de los volúmenes negociados. Por tanto, las pautas de negociación tienen mucho que ver con esta posible regularidad horaria en la volatilidad de algunos mercados. Una de las figuras adjuntas, acompañada de una tabla, tomadas de Daigler (19xx), muestra el perfil medio de la volatilidad intra-día, cuando se agrupan los precios en intervalos de 15 minutos. Se utiliza como medidas de volatilidad: la desviación típica de las rentabilidades, la medida de Garman-Klass (que veremos más adelante), y el número de ticks observados en cada intervalo de tiempo. En todos los casos se tiene un perfil en forma de U, si bien el máximo local de volatilidad no se produce en el instante de cierre del propio mercado de futuros, sino algo antes, coincidiendo con el cierre del mercado de contado. La tabla que se acompaña es de este mismo trabajo. Dos gráficos tomados de Lafuente (1999) presentan la volatilidad del IBEX 35, así como del futuro sobre este índice, en dos tramos horarios: 11 a 12 de la mañana, y 12 a 13 horas, apreciándose claramente la mayor volatilidad al comienzo del día. En las tablas que se acompañan, se presenta nuevamente evidencia a favor de un perfil de volatilidad en forma de U a lo largo del día. Chan, Chan y Karolyi (19xx), presentan una evidencia de estacionalidad intra-día similar a la mencionada.

4.2.2 Medidas de Parkinson y Garman-Klass

Generalmente, entendemos por volatilidad de un activo financiero el valor anualizado de un indicador de variabilidad de su tasa de rendimiento. Tradicionalmente, se ha tomado como como indicador de variabilidad la desviación típica aunque, posteriormente, se han ido introduciendo otras medidas alternativas de volatilidad que se consideran superiores en términos de eficiencia informativa, algunas de las cuales discutimos en esta sección, dejando las restantes para capítulos sucesivos. Se entiende que la volatilidad es una medida del riesgo del activo, aunque ya hemos adelantado algunas razones para tomar con precaución dicha interpretación.

Enlazando con los estadísticos hasta ahora considerados, extendamos el cálculo de la volatilidad histórica de una variable, que puede hacerse, disponiendo de la información relativa a un día de negociación, a través de:

1) Con precios de cierre (u otro dato representativo del día)

2) Con precios de apertura y cierre
3) Con los precios máximo y mínimo
4) Con el máximo, mínimo, apertura y cierre
5) Con precios bid y ask (en otro sentido)

Si disponemos de precios cotizados continuamente, como ocurre cuando hemos almacenado todas las transacciones realizadas a lo largo de un día de mercado:

$$Volatilidad: \ V = \sqrt{252}\sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(r_t - \bar{r}\right)^2}$$

donde $\bar{r}$ denota la rentabilidad media del día. La segunda raíz calcula la desviación típica de la rentabilidad a lo largo de dicho día de mercado, mientras que el producto por la raíz de 252 anualiza dicha volatilidad.

La rentabilidad media sobre un período reducido de tiempo, como el transcurrido entre dos transacciones, será muy pequeña, en cuyo caso, podemos calcular la volatilidad, muy aproximadamente, como:

$$Volatilidad: \ V = \sqrt{252}\sqrt{\frac{1}{T-1}\sum_{t=1}^{T} r_t^2}$$

En el caso de que dispongamos de *precios de cierre* (o cualquier otro *dato único* por día) observados con regularidad inferior a la diaria:

$$Volatilidad: \ V = \sqrt{\frac{N}{T}}\sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(r_t - \bar{r}\right)^2}$$

donde $\bar{r}$ denota la rentabilidad media durante los T días considerados en el cálculo. Si, por ejemplo, son datos de cierre observados el último dia de negociación de cada mes, tendremos $N = 252, T = 21$.

Medida de volatilidad de Parkinson

Supongamos que el precio de un activo sigue un camino aleatorio continuo con difusión constante σ^2. Al cabo de t periodos, el precio cumplirá:

$$P\left(P_0 \leq P_t \leq P_0 + dP\right) = \frac{dP}{\sqrt{2\pi\sigma^2 t}} \exp\left(-\frac{(P - P_0)^2}{2t\sigma^2}\right)$$

Como σ^2 es la varianza del desplazamiento tras una unidad de tiempo (con datos diarios, un dia), esto sugiere estimar:

$$\begin{aligned} d_t^2 &= \left(P_t - P_{t-1}\right)^2; \ \hat{\sigma}_d^2 \equiv \overline{d^2} = T^{-1}\sum_{t=1}^{T} d_t^2 \\ Var\left(\hat{\sigma}_d^2\right) &= \frac{1}{T} Var(\overline{d^2}) = \frac{1}{T}\left(\frac{1}{T-1}\sum_{t=1}^{T}(d_t^2 - \overline{d^2})^2\right) \end{aligned}$$

siendo $d_t = P_t - P_{t-1}$ la variación diaria en el precio. Si P_t es el logaritmo del precio, x_t será rentabilidad.

Supongamos que observamos unicamente los desplazamientos máximo y mínimo durante cada intervalo de tiempo, $H_t = \max(P_t), L_t = \min(P_t)$. Con el *máximo* y *mínimo* de la sesión [Parkinson (1980)], el *rango* del logaritmo del precio se define como la diferencia entre los logaritmos de los precios máximo H_t y mínimo L_t diarios,

$$l_t = \ln(H_t) - \ln(L_t)$$

y mide, aproximadamente, el porcentaje en el que el precio máximo excede del mínimo.

Puede probarse [Feller (1951)] que al cabo de un intervalo de tiempo de t periodos:[5]

$$\begin{aligned} E(l_t) &= \sqrt{\frac{8}{\pi}\sigma^2 t}; \\ E\left(l_t^2\right) &= 4\ln(2)\sigma^2 t = 4\ln(2)\sqrt{\frac{\pi}{8}}\left[E(l_t)\right]^2 = 1,088\left[E(l_t)\right]^2 \end{aligned}$$

que es una relación que podría utilizarse en un contraste de estructura de camino aleatorio. Por tanto, para un intervalo de tiempo de un periodo, tenemos:[6]

$$\sigma^2 = \frac{\pi}{8}\left[E(l)\right]^2 = 0,393\left[E(l)\right]^2 = 0,361E\left(l^2\right)$$

que sugiere un estimador del Método Generalizado de Momentos para la volatilidad de un periodo. El estimador, basado en el rango observado, es:

$$\begin{aligned} \hat{\sigma}_l^2 &= \frac{1}{4\ln(2)}\left(\frac{1}{T}\sum_1^T l^2\right) = 0,361\left(\frac{1}{T}\sum_1^T l^2\right) \\ Var\left(\hat{\sigma}_l^2\right) &= \frac{1}{T}Var(\overline{l^2}) = \frac{1}{T}\left(\frac{1}{T-1}\sum_{t=1}^T (l_t^2 - \overline{l^2})^2\right) \end{aligned}$$

lo cual nos lleva a la medida de volatilidad de Parkinson:[7]

$$Volatilidad: \; V = \sqrt{252}\sqrt{\frac{\sum_{t=1}^T \left[\ln\left(H_t/L_t\right)\right]^2}{T4\ln 2}}$$

[5] $E(l^p) = \frac{4}{\sqrt{\pi}}\Gamma\left(\frac{p+1}{2}\right)\left(1 - \frac{4}{2^p}\right)\varsigma(p-1)\left(2Dt\right)^{p/2}$, donde $\varsigma(.)$ denota la función z de Riemann.

[6] Para dejar evidente que se trata de un solo periodo de observación, hemos suprimido el subindice t .

[7] Al utilizar 252 estamos suponiendo que hay dias que no son dias hábiles de mercado. En otro caso, el factor de escala deberia ser 360.

Para un día cualquiera, podríamos utilizar como proxy de la volatilidad:

$$\hat{\sigma}_l^2 = \frac{1}{4\ln(2)} l^2 = .361 l^2$$

Comparación de varianzas de los dos estimadores $\hat{\sigma}_d^2$ y $\hat{\sigma}_l^2$:

$$\begin{aligned}
Var(\hat{\sigma}_d^2) &= E\left[\left(\hat{\sigma}_d^2 - \sigma^2\right)^2\right] = Var\left(\overline{d^2}\right) = \frac{1}{T_d} Var(d_t^2) = \frac{1}{T_d} E\left(d_t^2 - E(d_t^2)\right)^2 = \\
&= \frac{1}{T_d} E\left(\frac{\left(d_t^2\right)^2}{\left[E(d_t^2)\right]^2} + 1 - 2\frac{d_t^2}{E(d_t^2)}\right)\left[E(d_t^2)\right]^2 = \left(\frac{E(d_t^4)}{\left[E(d_t^2)\right]^2} - 1\right)\frac{\sigma^2}{T_d} = 2\frac{\sigma^2}{T_d} \\
Var(\hat{\sigma}_l^2) &= E\left[\left(\hat{\sigma}_l^2 - \sigma^2\right)^2\right] = (0,361)^2 Var\left(\overline{l^2}\right) = (0,361)^2 \frac{1}{T_l} E\left(l_t^2 - E(l_t^2)\right)^2 = \\
&= \frac{1}{T_l} E\left(\frac{\left(l_t^2\right)^2}{\left[E(l_t^2)\right]^2} + 1 - 2\frac{l_t^2}{E(l_t^2)}\right)\left[E(l_t^2)\right]^2 = (0,361)^2\left(\frac{E(l_t^4)}{\left[E(l_t^2)\right]^2} - 1\right)\frac{0,361\sigma^2}{T_l} = 0,41\frac{\sigma^2}{T_l}
\end{aligned}$$

donde hemos utilizado que bajo Normalidad (asintóticamente) y media cero: $E(x^4) = 3\left[E(x^2)\right]^2$, y tambien que: $E(d^2) = \sigma^2, E(l^2) = \sigma^2$.

Hacen falta 5 veces más datos para tener con σ_d^2 la misma precisión en la estimación de σ^2 que con σ_l^2.

Ejercicio de simulación: Nótese que el ejercicio consiste en obtener ambas estimaciones de la varianza de las rentabilidades en T dias de mercado. Luego, calcularemos la desviacion tipica de cada estimador utilizando las T estimaciones disponibles. Vamos a comprobar que el estimador $\hat{\sigma}_l^2$ es más preciso que $\hat{\sigma}_d^2$. Obtenemos los precios negociados cada dia generando M pasos de una distribución Uniforme en (-1/2,1/2), con una varianza teórica: $\sigma^2 = \frac{10.000}{12} = 833$. Los valores uniformes representan las variaciones a partir de un precio inicial ($P_0 = 500$). En el articulo de Garman-Klass, se toma $T = 100, M = 10000$.

Con los datos simulados para cada trayectoria diaria, se calculan las rentabilidades diarias al cierre y los dos estimadores de la varianza. El estimador $\hat{\sigma}_d^2$ es el promedio para los T dias de las variaciones diarias en precio al cuadrado, d_t^2 (precio de cierre menos precio de apertura al cuadrado): $\hat{\sigma}_x^2 = T^{-1}\sum_{t=1}^{T} d_t^2 = 812,63$, con una varianza, a lo largo de los 100 dias, de $T^{-1}\sum_{t=1}^{t}\left(d_t^2 - 812,63\right)^2 = 1164,60$; por tanto, con una desviación típica de $116,46$. El intervalo de confianza para la estimación de la varianza, de una desviación típica de amplitud, sería: $\sigma^2 \in (812,63 \pm 116,46)$.

Por otra parte, la media de las distancias l_t entre el máximo y el mínimo precio cada dia (sin logaritmos) es $\bar{l} = 45,0$. En este caso, en lugar de utilizar la desviación típica de $\bar{l}$ utilizamos el argumento:[8] $\hat{\sigma}_l^2 = .361E\left(l^2\right) = 0,393(E(l))^2 = 0,393(45)^2 = 795$, de modo que: $Var(\hat{\sigma}_l^2) = 0,41\frac{\hat{D}^2}{N} = 0,41\frac{\hat{\sigma}_l^2}{N} = 0,41\frac{795^2}{N} = 2591,30$, por lo que la desviación tipica del estimador es $50,90$.El

[8] La desviación del rango l resultó ser $14,3$, por lo que la estimación de la desviación típica del rango medio $\bar{l}$ es $1,43$.

intervalo de confianza para la estimación de la varianza, de una desviación típica de amplitud, es ahora: $\sigma^2 \in (795 \pm 50,9)$, mucho más estrecho. Recordemos que el valor teórico es $\sigma^2 = 833$.

Medida de volatilidad de Garman-Klass

Si disponemos de datos de apertura, cierre, máximo y mínimo la medida de volatilidad Garman-Klass sería:

$$Volatilidad: \ V = \sqrt{252}\sqrt{\frac{\sum_{t=1}^{T} \frac{1}{2}\left[\ln\left(H_t/L_t\right)\right]^2 - 0,386\sum_{t=1}^{T}\left[\ln\left(C_t/A_t\right)\right]^2}{T}}$$

Nota: 2 ln(2)-1 = 0,386.

El trading de activos es un proceso que, en muchos casos, tiene lugar de modo continuo a lo largo del día pero, sin embargo, se observa en momentos discretos de tiempo. Los valores de trading overnight no se registran, por lo que los valores realmente observados como máximo y mínimo no coinciden necesariamente con el máximo y mínimo realmente producidos a lo largo de las 24 horas. Esto produce un sesgo a la baja en el estimador del máximo, y un sesgo al alza en la estimación del mínimo. El rango de precios queda subestimado, siendo un subintervalo del verdadero rango de precios. Este sesgo que se produce por generar un proceso discreto a partir de un proceso que es realmente continuo, es significativo en la medida de Parkinson, y queda bastante atenuado en la medida de Garman-Klass.

Para tratar de tener en cuenta que las medidas de apertura/cierre y máximo/mínimo se obtienen en un intervalo inferior a 24 horas, suele ajustarse la volatilidad resultante por $\sqrt{24/8,5}$.El sesgo puede ser importante si se utiliza la volatilidad estimada para valoración de opciones. Ejercicios de simulación sugieren que una mayor liquidez del mercado tiende a reducir el sesgo, lo que presta en tal situación mayor justificación al uso de las medidas de Parkinson y Garman-Klass [ver Wiggins, J. (1992)]. La dirección del sesgo no es evidente cuando se utilizan exclusivamente datos de cierre.

Por el contrario, las medidas basadas en el rango observado son relativamente inmunes a este problema. En todo caso, dado que la existencia de ticks impide que los precios fluctúen de modo continuo, haciendo que se tienda a sobreestimar la volatilidad [ver Ball, C.A., (1988)], este sesgo se suele corregir en la medida de Parkinson por medio del ratio $c = d/vP$, siendo d el tamaño del tick, v la volatilidad diaria estimada, y P el precio del activo. Si $c \geq$1,77, se utiliza k=0,50$\sqrt{\frac{\pi}{c}}$, mientras que si $c<$1,77, se utiliza $k=\sqrt{1-c^2/6}$.

Example 7 *El archivo ACCIONES.xls presenta la comparación entre estas medidas de volatilidad y la volatilidad más estándar, calculadas para las cotizaciones de BBV, TELEFONICA, ENDESA y REPSOL, desde 9/10/97 a 10/11/99.*

Con la corrección $\sqrt{24/8,5}$, las medidas Parkinson y Garman-Klass de volatilidad *diaria* de BBV aumentan desde 2,26% a 3,80% y 3,75%, respectivamente.

Las medidas de volatilidad de Parkinson y Garman-Klass producen importantes ganancias de eficiencia: con un número de datos 5 ó 7 veces menor, generan estimaciones de la varianza poblacional con una precisión estadística similar a la estimación que se obtiene utilizando únicamente datos diarios de cierre.

Estos estimadores son, generalmente, menos erráticos que la estimación de la varianza basada en la rentabilidad diaria al cuadrado, y tienen más persistencia que las rentabilidades diarias, lo cual parece deseable. Ello sugiere la posibilidad de utilizar el estimador basado en el rango para validar un modelo de predicción de varianza que haya generado estimaciones $\hat{\sigma}_t^2$, [Ejercicio 2.5 en Christoffersen]

$$\sigma_{r,t+1}^2 = \alpha + \beta\hat{\sigma}_{t+1}^2 + u_{t+1}$$

Si estamos interesados en la predicción de la volatilidad un período hacia adelante, podríamos utilizar el rango en el modelo de predicción de volatilidades:

$$\sigma_{t+1}^2 = \omega + \alpha R_t^2 + \beta\sigma_t^2 + \gamma D_t^2$$

4.3 Conos de volatilidad

Una vez que se dispone de la serie temporal de rentabilidades de un activo, puede calcularse su volatilidad muestral sobre intervalos de distinta amplitud temporal. Queremos representar el modo en que la volatilidad varía con la amplitud de dichos intervalos temporales. Ya sabemos que, bajo supuestos de independencia temporal de las rentabilidades, la varianza sería una función lineal de la amplitud del intervalo. Sin embargo, esta es una hipótesis que no siempre se cumple.

Para ello, calculamos la volatilidad de la rentabilidad ofrecida por el activo utilizando ventanas moviles de una determinada amplitud, por ejemplo, una semana. Vamos desplazando dicha ventana a lo largo de la muestra y en cada periodo (cada dia, si disponemos de datos diarios), calculamos la volatilidad de las rentabilidades diarias observadas en la ventana semanal que finaliza ese dia. Luego, repetimos el ejrcicio con distintas amplitudes para las ventanas: quincena, mes, trimestre, semestre, año, etc.. De este modo, construimos una serie temporal de volatilidades para cada amplitud de ventana.

Asi, fijada una determinada amplitud temporal, por ej., un mes, vemos cómo ha ido cambiando la volatilidad a través del tiempo: si estamos a 15 de noviembre de 2001, y disponemos de datos desde comienzos de enero de 1996, empezaríamos calculando la volatilidad muestral para todo el mes de enero de 1996, por ejemplo (el comienzo es relativamente arbitrario), e iríamos añadiendo un día al final de la muestra, y quitando un día al comienzo de la misma, para volver a calcular la volatilidad registrada a lo largo de un mes de mercado. El procedimiento puede seguir hasta el último dato disponible. De este modo habremos generado una serie temporal de volatilidad, a lo largo de un mes, desde el 1 de enero de 1996, hasta el 15 de octubre de 2001.

En esta ocasión, sin embargo, no nos detenemos en analizar la variación temporal de la volatilidad, sino en estudiar algunos de sus estadísticos descriptivos, pues queremos analizar cómo cambian las propiedades de la volatilidad al cambiar la amplitud del intervalo de tiempo considerado. De hecho, vamos a considerar los valores que configuran la serie temporal de volatilidades de una determinada ventana como valores extraídos al azar de la distribución de probabilidad de la varianza correspondiente a dicha ventana. Inicialmente, tomamos los valores máximo y mínimo de las volatilidades así calculadas, y los representamos en la vertical sobre el eje de abscisas, en el punto correspondiente a 1 mes. El mismo procedimiento puede llevarse a cabo para cada una de las ventanas escogidas: para intervalos de 1 semana, comenzaríamos nuevamente al inicio de enero de 1996, obteniendo un máximo y un mínimo de las volatilidades calculadas sobre un rango temporal de una semana.

De este modo tendríamos una serie temporal para cada una de las volatilidades calculadas sobre intervalos de: una semana, dos semanas, un mes, trimestre, semestre, o año, y podríamos calcular su máximo y su mínimo. Cuando se representan dichos máximos y mínimos, se observa generalmente, que la volatilidad máxima es mayor en los intervalos menores (una semana) que en los intervalos amplios de tiempo. Por otra parte, la volatilidad mínima es menor asimismo cuando se calcula sobre intervalos breves de tiempo que cuando se calcula sobre intervalos amplios. Algo similar ocurre cuando tomamos percentiles simétricos para cada una de las series temporales de volatilidad, por ejemplo, percentiles 5% y 95%: el primero será menor para ventanas de una semana que para las de un mes, mientras que el percentil 95% será generalmente superior en las ventanas más cortas.

Esto se debe a que, como hemos visto en las secciones anteriores, la precisión en la estimacion de la volatilidad aumenta con el tamaño de la muestra utilizada en su calculo. Esto hará asimismo que las distribuciones de frecuencias de las distintas estimaciones de la volatilidad tengan más curtosis cuanto menor sea la amplitud de la ventana correspondiente. Dicho de otro modo, el rango de valores de volatilidad calculados sobre una semana, tiende a incluir al rango de volatilidades calculado sobre un mes, éste al rango calculado sobre tres meses, y así sucesivamente. De este modo, habremos obtenido un *cono de volatilidad.*

Los conos de volatilidad desempeñan un papel importante cuando se quiere apreciar si una opción está relativamente cara o barata en el mercado. Para ello, se trata de comparar la volatilidad implícita en el precio de mercado de la opción, con el rango de volatilidades que históricamente se ha estimado sobre un período de tiempo igual al que queda hasta la expiración de la opción.

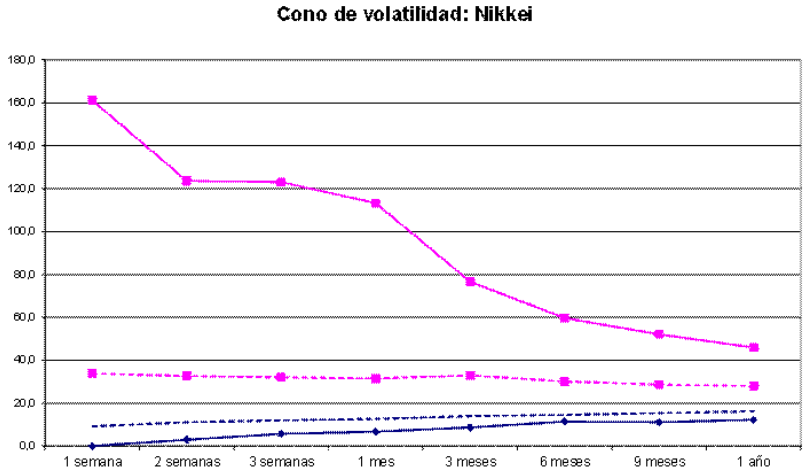
Si, por ejemplo, la volatilidad implícita cuando queda un mes para la expiración de la opción está por debajo del percentil 10 de la distribución de frecuencias de las volatilidades que hemos calculado sobre intervalos de un mes, diremos que el mercado está infravalorando dicha opción, puesto que está dando un precio que se corresponde con una volatilidad que es poco creíble que se produzca, por lo reducido de su cuantía.

Esto significa que, en base a la experiencia histórica, la volatilidad que cabría esperar es superior a la que el mercado espera. Si no hay razones para que así

resulte, habría que pensar que la opción está barata. Lo contrario ocurriría si la volatilidad implícita fuese superior al percentil 90, por ej.. Salvo que hubiese razones para esperar una volatilidad excesivamente alta para los registros históricos, habría que pensar que el mercado está sobrevalorando dicha opción. Al construir un cono de volatilidad cabría introducir un ajuste si realmente existe una discrepancia permanente entre los niveles de volatilidad histórica e implícita.

Los percentiles escogidos determinan el número de señales que puedan obtenerse acerca de posibles situaciones de *mispricing* (error en precio). Percentiles menos extremos producirán más señales de precio incorrecto, pero también mayor nivel de riesgo, porque las señales tenderán a ser incorrectas más frecuentemente. Seleccionar unos determinados percentiles es similar a seleccionar un determinado nivel de riesgo para el cálculo del VaR.

Las pestañas *Cono* 1 y *Cono* 2 presentan los cono de volatilidad para el SP500 y el DAX, calculados con datos de enero 1997 a agosto 1997, con percentiles 10 y 90.



5 Varianzas cambiantes en el tiempo

Case Study II.3: Varianzas y covarianzas a lo largo de la estructura temporal de tipos de interés.

En las siguientes secciones vamos a analizr la estimación cambiante en el tiempo de varianzas, covarianzas y correlaciones. Varias preguntas que habremos de tener presente al estimar una serie temporal de volatilidades o de correlaciones:

- ¿Qué frecuencia de observación de datos debe utilizarse en la estimación de varianzas y correlaciones?
- ¿Qué longitud de periodo muestral debe utilizarse? ¿Largo, corto?

- ¿Deben utilizarse variaciones absolutas o variaciones relativas? Si estamos estimando momentos que suponemos constantes en el tiempo, deberiamos utilizar aquellas variaciones que generen estimaciones más estables de las volatilidades y correlaciones.

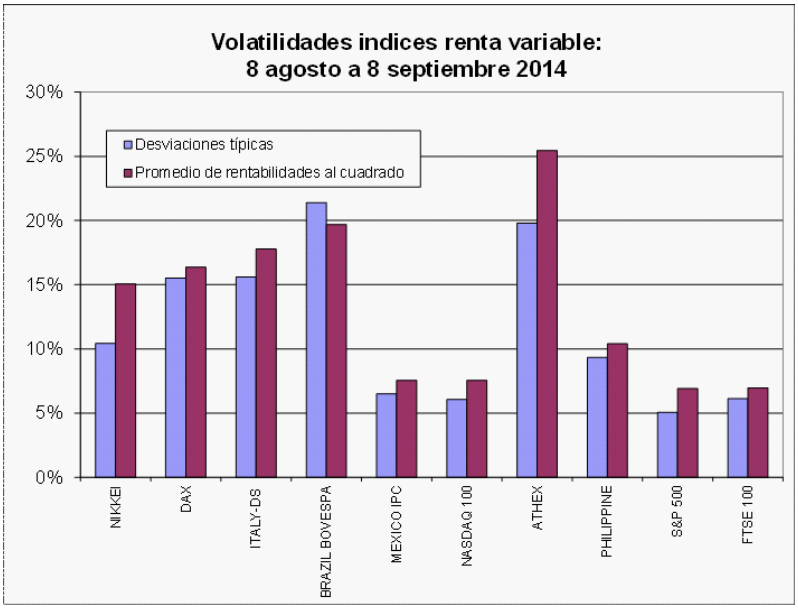
- ¿Debe utilizarse equiponderacion cdelas observaciones muestrales, ponderacion exponencial, o estructuras GARCH?

5.1 El uso de ventanas móviles en la estimación de la varianza (equiponderadas)

Una generalización importante en el análisis de datos financieros, consiste en considerar los estadísticos muestrales no como constantes, sino siendo a su vez funciones del tiempo, en cuyo caso estaremos interesados en disponer de *series temporales* de los mismos. Si identificamos volatilidad con desviación típica, sólo generando series temporales de la varianza de su rentabilidad podremos hablar de variaciones en la volatilidad de dicho activo.

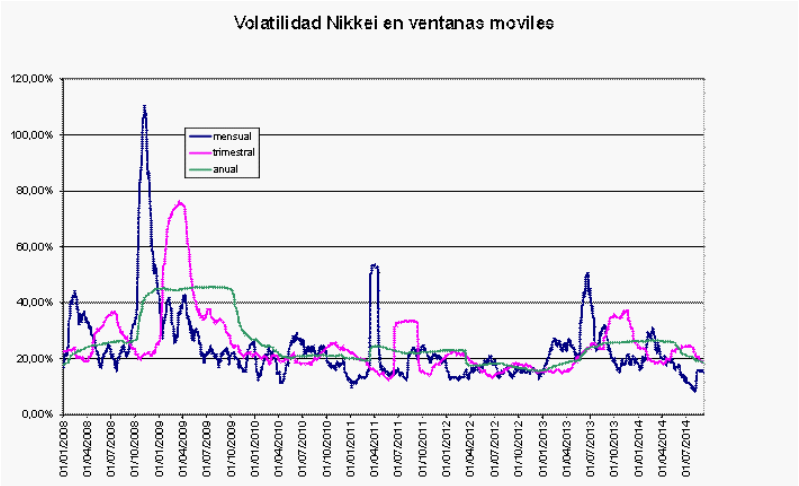
Sin embargo, la varianza es un momento poblacional o muestral y, como tal es constante. ¿Cómo podemos generar una serie temporal para la varianza? Utilizando las denominadas *ventanas muestrales*, que son submuestras cortas, cada una de las cuales se obtiene a partir de la previa, añadiendo un último dato, y prescindiendo del primero. La amplitud de la ventana ha de ser suficiente como para creer que, con cada una de ellas podemos estimar el parámetro en cuestión (por ej., la varianza) con suficiente aproximación. De este modo, estaremos generando un valor numérico de la varianza en cada instante para el cual tenemos un dato. Sólo perderemos un número de observaciones iniciales, igual al número de ellas incluidas en cada ventana. Si, por ej., cada ventana consta de 20 datos, entonces podremos generar datos de varianza a partir de la observación 21.

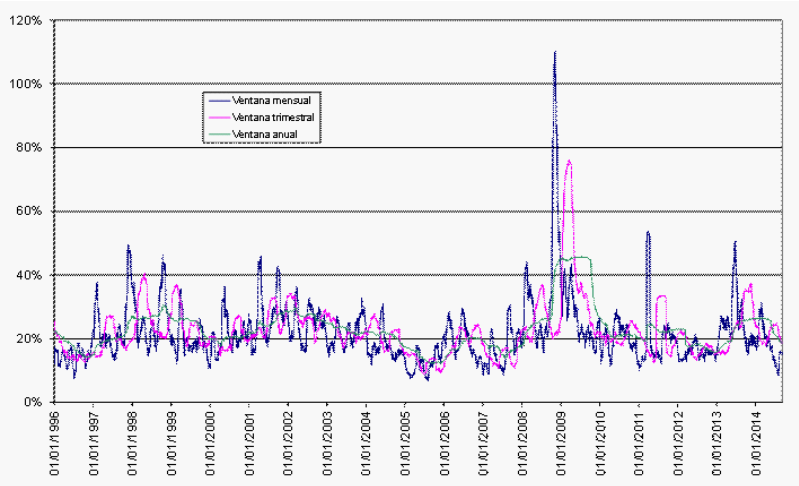
Hay que mantener un equilibrio, no siempre fácil, al decidir la amplitud de la ventana que se utiliza en el cálculo de la varianza: por un lado, una ventana más corta tendrá más posibilidad de utilizar una media estable, y representará mejor la situación actual, pero la varianza resultante será bastante volátil, entre otras cosas, porque no la estimaremos con suficiente precisión. Por otro, una ventana amplia proporcionará una medida de volatilidad suave, pero calculada respecto a una medida de referencia posiblemente no constante. En la valoración de opciones, se recomienda generalmente utilizar una ventana de longitud igual al período que resta hasta el vencimiento de la opción.



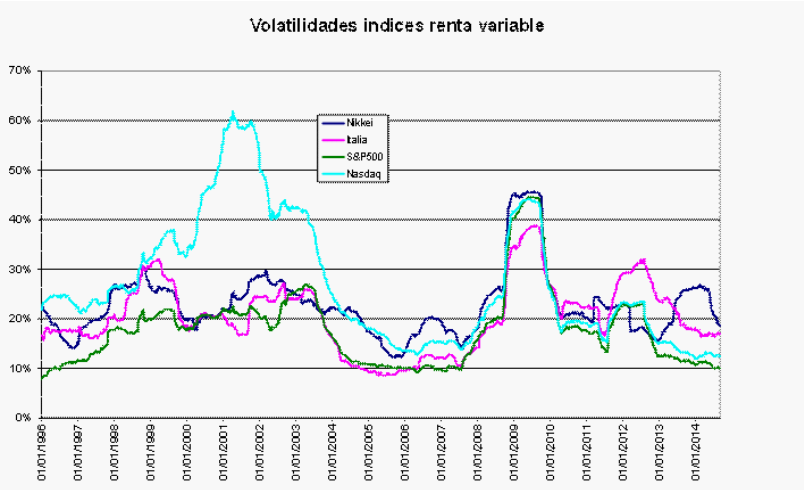
Example 8 *En la pestaña "indices" a la derecha de las rentabilidades logaritmicas, se presentan las rentabilidades al cuadrado. Debajo de las mismas se presentan las volatilidades de los distintos indices de Bolsa, medidas como promedio del cuadrado de las rentabilidades diarias. Se hace el cálculo para distintos intervalos de tiempo. Las desviaciones típicas son anualizadas. Se comparan con las volatilidades calculadas como desviaciones típicas de las rentabilidades diarias.*

Example 9 *Este es un mercado con un nivel de volatilidad relativamente alto. Pero lo más significativo son las fluctuaciones que experimenta su nivel de volatilidad (por ej., mensual). Es decir, la volatilidad puede ser bastante errática. Las pestañas "Volatilidad Nikkei" y "Vol. Nikkei 2008-14" muestran la serie temporal de volatilidades del Nikkei, calculadas sobre ventanas móviles de 1 y 3 meses (22 y 66 sesiones). Puede apreciarse que la serie temporal de volatilidad calculada con una ventana muestral más amplia es más suave que la calculada con una ventana muestral más corta. Esto siempre ocurre así, por construcción. Este gráfico continúa mostrando notables variaciones en el nivel de volatilidad. Sin embargo, si nos interesa el grado de asociación existente entre los niveles de volatilidad en dos mercados, es difícil apreciarlo en un gráfico temporal. Es mucho más útil considerar nubes de puntos de volatilidades para los mismos pares de índices.*





Sería interesante estimar modelos estadísticos de relación (regresión) entre estos pares de volatilidades: por un lado, un modelo que relacionase sus niveles en cada día de la muestra podría conducir a una pendiente próxima a la unidad en el caso de las comparaciones entre mercados europeos. Ello sería, además, consistente con la idea de que existe quizá un factor de volatilidad que explica una buena parte de los niveles de volatilidad en estos mercados. En la comparación Nikkei vs. S&P 500, la relación parece estar bastante condicionada por los episodios de alta volatilidad que han sido comunes a ambos mercados. Por último, puede verse que la asociación entre los niveles de volatilidad trimestral es bastante mayor que la existente entre los niveles de volatilidad mensual, como quizá cabría esperar.



El mayor interés que presenta el cálculo de la varianza en la forma de una serie temporal es que, con ella, podemos plantearnos la *predicción de la volatilidad futura*, que discutiremos en detalle más adelante. Un segundo aspecto de importancia reside en la capacidad que nos prestan las series temporales de cuantificar el grado de asociación de la volatilidad en distintos mercados, así como las características dinámicas de su relación. Si detectamos que una mayor volatilidad en un índice de mercado, como Dow Jones, *anticipa* un aumento de volatilidad en otro índice, como el DAX, quizá podamos utilizar dicha información para mejorar nuestras predicciones de la volatilidad en este último mercado.

Ahora bien, ¿sobre qué intervalo de tiempo debe estimarse la volatilidad? Ya hemos dicho que la elección de una longitud para la ventana muestral dista de ser trivial. En algunos casos, como cuando se quiere extrapolar hacia el futuro (predecir) volatilidad, es habitual utilizar una misma longitud en su cálculo que la del período sobre el que se quiere predecir. Esto es más evidente en algunos casos, como los conos de volatilidad que veremos más adelante, que en otros. Que no haya unanimidad sobre cuestiones de este tipo ayuda a generar mercado, pues distintos agentes valorarán la volatilidad de distinta manera, entre otras cosas, porque estén interesados en distintos horizonte de inversión.

5.2 Volatilidad ponderando más el pasado reciente (ponderación exponencial)

Pero, incluso fijado un intervalo temporal ¿debemos de dar a todas las observaciones pasadas la misma relevancia en el cálculo de la volatilidad? Puede parecer razonable ponderar más las observaciones más recientes. Utilizando las potencias de un factor λ, $0<\lambda<1$, conseguimos que las observaciones vayan perdiendo importancia cuanto más se alejan en el tiempo.

La medida de volatilidad es entonces:

$$\sigma = \sqrt{\frac{1}{\Lambda}\sum_{i=1}^{n} \lambda^i x_{t-i}^2}$$

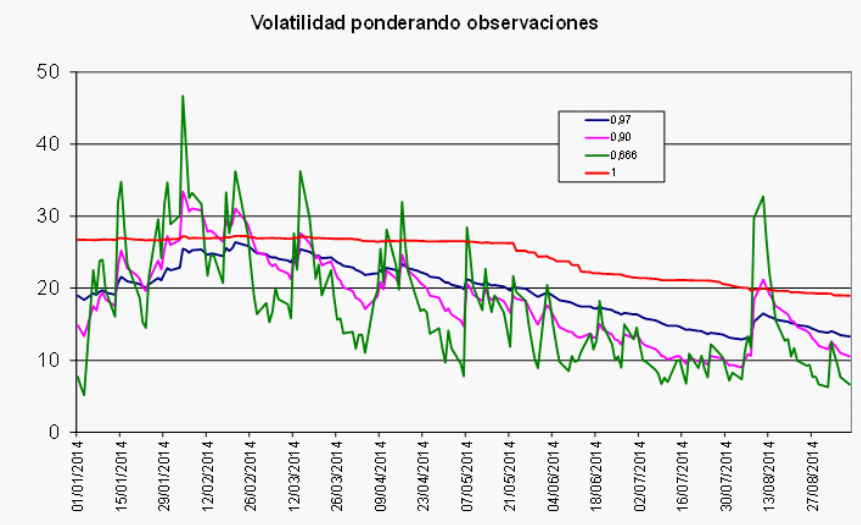
donde n es el número de datos utilizado en el cálculo de la volatilidad, y: $\Lambda = \sum_{i=1}^{n} \lambda^i = \lambda\frac{1-\lambda^n}{1-\lambda}$, que se reduce a: $\Lambda = \frac{1}{1-\lambda}$ cuando no ponemos un límite al número de datos utilizado. En tal caso, el uso del factor λ substituye a la necesidad de fijar de antemano un número de observaciones para el cálculo de la volatilidad. Reducir el valor de λ equivale a acortar el intervalo temporal utilizado en la estimación.

Un análisis similar podría aplicarse al cálculo de la correlación entre dos rentabilidades:

$$\rho = \frac{1}{\Lambda}\sum_{i=1}^{n} \lambda^i x_{t-i} y_{t-i}$$

con el objeto de aminorar el efecto de acontecimientos relativamente alejados.

Este esquema de ponderaciones en forma de sucesión geométrica constituye el modelo de suavizado exponencial de varianzas y covarianzas que analizaremos en detalle más adelante.



Example 10 *La pestaña "Ponderaciones" muestra este cálculo, aplicado a la volatilidad del Nikkei, utilizando como pesos las potencias de 0,97, 0,90 y 0,66, alternativamente. Puede apreciarse que la volatilidad es más variable para ponderaciones más reducidas, en las que el presente pesa maás que el pasado. Asimismo, los niveles de volatilidad disminuyen, en este caso, al aplicar las ponderaciones relativas y tanto más cuanto menor es la ponderación, es decir, cuanto más se descuentan los valores más alejados en el tiempo. Esto se debe a*

que durante el período considerado, los niveles de volatilidad fueron superiores al comienzo que al final de la muestra.

5.3 Bandas de confianza para precios y rentabilidades bajo el supuesto de Normalidad

Si la rentabilidad de un activo obedece a una distribución Normal, la probabilidad de que dicha rentabilidad se sitúe entre su esperanza matemática y un rango alrededor de ella de más o menos una desviación típica, es del 68,26%. Pasa a ser del 95,46% cuando el intervalo tiene dos desviaciones típicas de amplitud, y es del 99,87% para tres desviaciones típicas. El intervalo de confianza del 95% está delimitado por la esperanza matemática más y menos 1,96 veces la desviación típica, mientras que el intervalo de confianza del 99% está delimitado por la esperanza matemática más y menos 2,33 veces la desviación típica.

Si creemos que el proceso de cotizaciones no es estacionario, entonces este ejercicio es bastante cuestionable, puesto que se basa en la hipótesis de que la distribución de probabilidad del proceso de cotizaciones que se analiza es relativamente estable. En general, la ausencia de estacionariedad va a aparecer en la forma de esperanza y varianza cambiantes en el tiempo, por lo que intervalos centrados alrededor de una cotización media histórica pueden ser muy poco representativos de la evolución futura del mercado.

Por tanto, si la variable que nos interesa no es estacionaria, hemos de transformarla en una variable estacionaria (por ejemplo, tomando primeras diferencias, o primeras diferencis de su logaritmo) y construir los intervalos de confianza para dicha transformación estacionarias para despues transformarlos en intervalos de confianza para la variable que nos interesa. Esta transformación de intervalos de confianza puede hacerse porque los percentiles se transforman adecuadamente al aplicar transformaciones monótonas.

Existe un modo razonable de construir *intervalos de valores esperados* bajo los supuestos que hemos hecho acerca de la distribución de probabilidad de las *rentabilidades continuas*. Retomemos la hipótesis de que $\ln(1 + r_t)$ o, lo que es lo mismo, $\ln(P_t/P_{t-1})$, se distribuye como una Normal(μ,σ^2) y, por estabilidad temporal, lo mismo ocurre con $\ln(P_{t+1}/P_t)$. Ello significa que, una vez que $\ln(P_t)$ es conocido, entonces podemos considerar que $\ln(P_{t+1})$ se distribuye como una Normal($\mu + \ln(P_t)$, σ^2).

La cotización media del IBEX35 durante diciembre de 1997 fue de 7.152,52. A lo largo del mismo mes, la volatilidad diaria *de las cotizaciones*, medida por su desviación típica, fue de 91,93. Bajo el supuesto de que la cotización del índice sigue una distribución Normal con *esperanza y varianza constantes*, los fundamentos estadísticos que acabamos de recordar nos permitirían construir intervalos de confianza para las cotizaciones de días futuros, llevando a izquierda y derecha de la cotización mensual media, tomada como predicción de la cotización en días sucesivos, un determinado número de veces su desviación típica. Esto nos produciría intervalos de confianza que cambiarían a través del tiempo según fuesen variando la predicción puntual de la cotización futura, y la desviación típica muestral.

Para ello, es importante observar que la desviación típica de las rentabilidades fue, a lo largo de diciembre de 1997, de 0,0135363. Como primera aproximación, vamos a ignorar la rentabilidad diaria media durante diciembre de 1997, que fue de 0,198%, y en cuya repetición quizá el inversor no quiera confiar. Este es el parámetro μ de la expresión anterior, que supondremos igual a cero, por lo que centraremos nuestro intervalo exclusivamente alrededor de $\ln(P_t)$.

En tales condiciones, el rango de cotizaciones del 68,26% para el día siguiente de mercado (primer día de mercado de enero) es de:

$$\begin{aligned} \ln(7152,52) - \sqrt{1}(0,0135363) &< \ln P_{t+1} < \ln(7152,52) + \sqrt{1}(0,0135363) \Rightarrow \\ 8,8616387 &< \ln P_{t+1} < 8,888756 \Rightarrow 7.056,4 < P_{t+1} < 7.250,0 \end{aligned}$$

siendo el último un cálculo aproximado.

El rango de cotizaciones del 95,46% para el día siguiente de mercado es de:

$$\begin{aligned} \ln(7152,52) - 1,96\sqrt{1}(0,0135363) &< \ln P_{t+1} < \ln(7152,52) + 1,96\sqrt{1}(0,0135363) \Rightarrow \\ 8,848147 &< \ln P_{t+1} < 8,902293 \Rightarrow 6.961,5 < P_{t+1} < 7.348,8 \end{aligned}$$

mientras que el rango del 99% está determinado por:

$$\begin{aligned} \ln(7152,52) - 2,33\sqrt{1}(0,0135363) &< \ln P_{t+1} < \ln(7152,52) + 2,33\sqrt{1}(0,0135363) \Rightarrow \\ 8,843680 &< \ln P_{t+1} < 8,906760 \Rightarrow 6.930,5 < P_{t+1} < 7.381,7 \end{aligned}$$

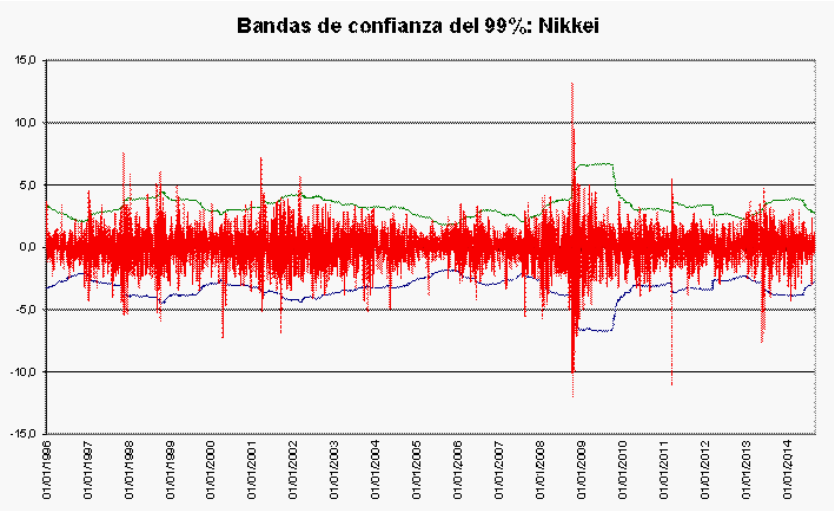
que es, lógicamente, más amplio que el anterior.

Por último, el rango del 99% para cinco días de negociación (una semana) después, es:

$$\begin{aligned} \ln(7152,52) - 2,33\sqrt{5}(0,0135363) &< \ln P_{t+1} < \ln(7152,52) + 2,33\sqrt{5}(0,0135363) \Rightarrow \\ 8,804695 &< \ln P_{t+1} < 8,945745 \Rightarrow 6.666,5 < P_{t+1} < 7.675,2 \end{aligned}$$

Puede observarse que:

- los intervalos construidos no son centrados en torno a la cotización del día, 7.152,52, como consecuencia del supuesto de lognormalidad, que hace más probables aumentos importantes que descensos importantes (es decir, el incremento medio esperado es mayor que el descenso medio esperado),
- la amplitud de los intervalos aumenta con el grado de confianza que queremos tener en que el intervalo construido contenga a la cotización que se pretende anticipar,
- la amplitud de los intervalos aumenta con el horizonte temporal para el cual establecemos la predicción.



Con este procedimiento podemos aprovecharnos de la aditividad de las rentabilidades continuas. Recordemos que esta propiedad garantiza que la rentabilidad continua sobre un determinado período de tiempo puede obtenerse agregando las rentabilidades continuas sobre subperíodos del mismo. Además, si las rentabilidades continuas son independientes, y cada una de ellas sigue una distribución Normal, todas ellas con igual esperanza y varianza, entonces su suma obedece asimismo una distribución Normal, con esperanza y varianza igual a la esperanza y varianza de cada una de las rentabilidades sobre un subperíodo, multiplicadas por el número de rentabilidades incluido en el período amplio.

Por tanto, si quisiésemos tomar en consideración el incremento diario medio en rentabilidad, estimado en un 0,198%, cuando calculamos un rango admisible para dentro de una semana, lo que haríamos sería añadir 5 veces 0,00198 al logaritmo de la cotización actual, 7.152,52, antes de tomar 2,33 veces a su izquierda y a su derecha, la desviación típica, de 0,0135363.

Así,

$$\begin{aligned} [ln(7152,52) + 5\,(0,00198)] - 2,33\sqrt{5}(0,0135363) &< lnS < \\ [\ln(7152,52) + 50\,(,00198)] + 2,33\sqrt{5}(0,0135363) &\Rightarrow \\ 6.731,8 &< S < 7.751,5 \end{aligned}$$

Exercise 11 *Este tipo de análisis proporciona una primera evaluación acerca de si un dato de mercado de un día concreto, puede considerarse como anómalo. La pestaña "Gráfico Bandas" en indices.xls muestra la rentabilidad del índice S&P 500 (variación en cotización), junto con las bandas de confianza del 99%.*

En segundo lugar, este tipo de evaluaciones es claramente importante al valorar la conveniencia de establecer estrategias de cobertura, pues estimamos el rango esperado de fluctuación de la rentabilidad de un determinado activo.

En este sentido, debe notarse que, aunque nos hemos limitado a calcular intervalos de confianza, en realidad disponemos de una distribución de probabilidad centrada alrededor de la última cotización o precio observados. Por consiguiente, no sólo podemos construir el rango de fluctuación esperado a un determinado nivel de confianza, sino que también podemos asociar probabilidades a cada uno de los posibles eventos, dentro o fuera de dicho rango.

Este análisis se ha basado en dos supuestos:

- Independencia de las rentabilidades sobre subperíodos no solapados. Este supuesto facilita enormemente el cálculo. Cuando no se cumple, debe establecerse una modelización de la correlación temporal, construir el intervalo de confianza para la innovación y deducir de dicho intervalo el correspondiente para la rentabilidad. Por ejemplo, si creemos que la estructura de autocorrelación es: $r_t = a + \beta r_{t-1} + \varepsilon_t$,con $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$, estimaremos dicho modelo por mínimos cuadrados, y tendremos: $0,95 = P(-1,96\hat{\sigma}_\varepsilon \leq \varepsilon_t \leq 1,96\hat{\sigma}_\varepsilon)$, con $\hat{\sigma}_\varepsilon = \frac{1}{T-1}\sum_{t=2}^{T}(r_t - (\hat{a} + \hat{\beta} r_{t-1}))^2$, por lo que:

$$0,95 = P\left(\hat{a} + \hat{\beta} r_{t-1} - 1,96\hat{\sigma}_\varepsilon \leq r_t \leq \hat{a} + \hat{\beta} r_{t-1} + 1,96\hat{\sigma}_\varepsilon\right)$$

  obteniendo así el intervalo de confianza para r_t.

- Normalidad de las rentabilidades continuas. En muchas ocasiones, tal supuesto no resulta admisible para variables de rentabilidad financiera, como hemos visto ya en algunos ejemplos. En ocasiones, las distribuciones de frecuencias presentan cierto grado de asimetría. Más frecuentemente, en el caso de variables financieras, la distribución muestral o de frecuencias de las rentabilidades observadas presenta desviaciones respecto de su valor central que son mayores de lo que la Normalidad podría explicar. Dicho de otro modo, las colas de la distribución son muy gruesas o los valores extremos demasiado frecuentes, en relación con la distribución Normal.

- De hecho, este ejercicio puede utilizarse como back-testing para contrastar el supuesto mantenido de Normalidad, con varianza constante, a lo largo de toda la muestra histórica. Con un tamaño muestral grande, T, el número de excepciones, es decir, el número días en los que la rentabilidad cae fuera del intervalo de confianza de $\alpha\%$ debe ser aproximadamente igual al producto $T\alpha\%$. Un número de excepciones mayor podría indicar la existencia de colas gruesas (leptocurtosis) en la distribución de rentabilidades. Además, las excepciones deberían repartirse equilibradamente entre rentabilidades excepcionalmente positivas y negativas, es decir, debería haber tantas excepciones por la izquierda de los intervalos de confianza diarios, como por la derecha de los mismos.

6 Modelización y predicción de la volatilidad

Generalmente, consideramos que trabajamos con series temporales de rentabilidades de activos financieros observadas frecuentemente, pues es entonces cuando resulta habitual observar volatilidades cambiantes en el tiempo. En el caso más sencillo, consideramos que la rentabilidad obedece al proceso estocástico,

$$R_{t+1} = \sigma_{t+1} z_{t+1}, \text{ con } z_{t+1} \sim i., i.d., N(0,1)$$

Una característica de los mercados financieros es que suelen observarse intervalos concretos de tiempo en los que sistemáticamente se produce cada día una alta volatilidad, seguidos de períodos de reducida volatilidad. Esto se manifiesta en que el cuadrado de las rentabilidades diarias tenga generalmente una alta autocorrelación.

En muchas ocasiones, la distribución de probabilidad de las rentabilidades es dependiente en el tiempo a través de los segundos momentos, aunque las propias rentabilidades no presenten autocorrelación. Cuando esto sucede, podria aparecer autocorrelación en las rentabilidades al cuadrado a través de la dependencia en la evolución temporal de σ_t^2. Por tanto, al corregir las rentabilidades del posible efecto persistente de la volatilidad, la autocorrelación en las rentabilidades al cuadrado debería desaparecer. El objetivo fundamental de la modelización de la evolución temporal de la volatilidad consiste precisamente en lograr que las rentabilidades estandarizadas al cuadrado R_t^2/σ_t^2 estén libres de correlación temporal.

Si las propias rentabilidades tienen autocorrelación (lo que es poco habitual en datos frecuentes), aplicaremos estas consideraciones a las innovaciones del modelo que explica la evolución temporal de R_t. A modo de ejemplo, supongamos que dicho modelo es:

$$R_{t+1} = \alpha + \beta R_t + \varepsilon_t$$

donde $Var(\varepsilon_t) = \sigma_t^2$ es la varianza de dichas innovaciones. Siendo innovaciones, los valores normalizados de ε_t deben carecer de autocorrelación. En este caso, es ε_t quien tiene la representación que antes propusimos:

$$\varepsilon_{t+1} = \sigma_{t+1} z_{t+1}, \text{ con } z_{t+1} \sim i., i.d., N(0,1)$$

En definitiva, la función de autocorrelación de las rentabilidades estandarizadas al cuadrado o de sus innovaciones al cuadrado es un estadístico fundamental en este análisis de modelización de la varianza. Si la modelización de la varianza ha sido adecuada no debe existir autocorrelacion en el cuadrado de las innovaciones.

Cuando se trabaja con varianzas cambiantes en el tiempo, una primera intuición para su estimación consiste en utilizar la expresión habitual de la varianza, pero sobre ventanas móviles, es decir, sobre submuestras que vamos moviendo a lo largo de toda la muestra. Por ejemplo, calculamos la varianza entre los datos 1 y 250 y la tomamos como varianza en el dia 250 de la muestra; luego, tomamos la varianza entre los dias 2 y 251 como varainza el dia 251, y asi sucesivamente.

Si queremos estimar el nivel de volatilidad el próximo día de mercado, una sencilla posibilidad es utilizar la volatilidad media observada en los últimos m días de mercado,

$$\sigma_{t+1}^2 = \frac{1}{m} \sum_{i=0}^{m-1} R_{t-i}^2$$

En esta expresión se ha incorporado ya el supuesto habitual de que, trabajando con rentabilidades en frecuencias altas, la rentabilidad media es prácticamente cero. Además, su estimación numérica podría introducir mayor distorsión que la incorporación directa de una valor medio nulo. Una ventaja de esta expresión es que nos permite generar una estimación del nivel de volatilidad al cierre del mercado en el período t, sin necesidad de efectuar ningún cálculo adicional. Tiene varias desventajas:

- no es claro cómo debe elegirse el número de días m . Este número suele denominarse *amplitud de la ventana de datos*. Un número reducido tenderá a generar una serie temporal de volatilidad muy errática, mientras que un número elevado de días generará una serie de volatilidad que puede considerarse excesivamente suave. La elección de la amplitud de ventana debe depender de la utilización que quiera hacerse de la previsión de volatilidad resultante.

- la serie de volatilidades reacciona al alza sólo después de que se haya observado en el mercado una rentabilidad diaria elevada. En este sentido, su naturaleza no es tanto la de anticipar el comportamiento futuro de la volatilidad, como el de reflejar el comportamiento reciente de la misma.

- precisamente por esta razón, es un indicador que va reaccionando a incrementos de volatilidad con cierto retraso, pues se trata de un promedio de los niveles de volatilidad en los últimos m días de mercado.

- pondera por igual cada uno de los m días utilizados en su calculo. Ello hace que la presencia de un día de alta volatilidad elevará la previsión de volatilidad la primera ocasión en que dicha rentabilidad se utilice en el cálculo, y tenderá a mantener la volatilidad elevada durante m días, reduciéndose nuevamente de manera drástica. La función de autocorrelación de muchas rentabilidades financieras al cuadrado sugiere bastante persistencia, siendo por tanto contraria a estas variaciones bruscas al inicio y al final del período de m dias.

6.1 El modelo de suavizado exponencial (EWMA)

Remark 12 *En la presentación de los modelos EWMA y GARCH(1,1) suponemos que estamos interesados en modelizar la varianza condicional de una rentabilidad, es decir, la evolucion de su varianza a lo largo del tiempo. Cuando la*

rentabilidad presenta autocorrelación, estos mismos modelos se aplican a la varianza condicional de la innovación del proceso de rentabilidades, y asi los utilizaremos más adelante a lo largo del curso. En ese caso en todas las expresiones de las varianzas de ambos modelos hay que sustituir las rentabilidades por los valores actuales y pasados de las innovaciones del proceso de rentabilidad. La presentación que hacemos inicialmente equivale a suponer que el proceso de rentabilidades no tiene structura dinamica, siendo del tipo: $r_t = \varepsilon_t$.

La varianza muestral otorga a todas las observaciones disponibles la misma ponderación. Por tanto, las desviaciones respecto del nivel de referencia tienen la misma importancia tanto si se produjeron recientemente como si se produjeron hace ya algún tiempo. Esto puede no ser totalmente deseable en el análisis de mercados financieros. En ocasiones, es conveniente abandonar este supuesto, dando pie a esquemas con ponderaciones del tipo,

$$\sigma^2_{t+1} = \frac{1}{\sum_{i=0}^{m-1} \alpha_i} \sum_{i=0}^{m-1} \alpha_i r^2_{t-i}$$

donde $R_s = \ln(P_s/P_{s-1})$, es la rentabilidad de un determinado activo financiero, $\alpha_i > 0$, $\sum_{i=0}^{m-1} \alpha_i = 1$, y $1 < i < j < m-1 \Rightarrow \alpha_i > \alpha_j$. Esta expresión calcula la varianza como media ponderada de las rentabilidades al cuadrado. No utiliza las desviaciones respecto de la rentabilidad media porque se supone que en datos de alta frecuencia dicha media es despreciable. Si los pesos no suman uno, hay que dividir en la expresión anterior por su suma.

En ocasiones se utilizan como pesos las potencias de una constante λ comprendida entre 0 y 1, $\alpha_i = \lambda^i$, lo que conduce a,

$$\sigma^2_{t+1} = (1-\lambda) \sum_{i=0}^{\infty} \lambda^i r^2_{t-i} \tag{1}$$

La constante que aparece fuera del sumatorio, $1-\lambda$, hace que las ponderaciones sumen 1, ya que $\sum_{i=0}^{\infty} \lambda^i = \frac{1}{1-\lambda}$, que es equivalente a la expresión recursiva:

$$\sigma^2_{t+1} = \lambda \sigma^2_t + (1-\lambda) r^2_t$$

Un esquema similar puede utilizarse para la covarianza entre las rentabilidades de dos activos:

$$\sigma_{12,t+1} = (1-\lambda) \sum_{i=0}^{\infty} \lambda^i r_{1,t-i} r_{2,t-i}$$

Pero al trabajar con datos, no podemos utilizar sumas infinitas. Es preciso truncar estas expresiones mediante:

$$\sigma^2_{t+1} = (1-\lambda) \sum_{i=0}^{t-1} \lambda^i r^2_{t-i} + \lambda^t \sigma^2_0$$

donde el último término, que es función de la volatilidad en un período inicial, σ_0^2, pierde relevancia con el paso del tiempo. Nótese que la suma de los pesos en el primer término a la derecha de la igualdad es,

$$(1-\lambda)\sum_{i=0}^{t-1}\lambda^i = (1-\lambda)\frac{1-\lambda^t}{1-\lambda} = 1-\lambda^t$$

por lo que la suma total de los pesos en el miembro derecho de la expresión anterior es igual a 1, como debería suceder. Tras el truncamiento, el modelo continua satisfaciendo la misma expresion recursiva:

$$\begin{aligned}\sigma_t^2 &= (1-\lambda)\left[r_{t-1}^2 + \lambda r_{t-2}^2 + ...\right] + \lambda^t\sigma_0^2 \\ \sigma_{t+1}^2 &= (1-\lambda)\left[r_t^2 + \lambda r_{t-1}^2 + ...\right] + \lambda^{t+1}\sigma_0^2\end{aligned}$$

por lo que:

$$\sigma_{t+1}^2 = \lambda\sigma_t^2 + (1-\lambda)r_t^2$$

Este es un modelo simple, que sólo tiene un parámetro para estimar, λ. Además, una vez calculada la volatilidad para un determinado día, la fórmula de actualización no precisa utilizar nuevamente los datos históricos, sino unicamente la ultima rentabilidad. Nótese que estamos estimando la varianza cada dia en función de la varianza y del cuadrado de la rentabilidad del dia anterior. En realidad, el modelo considera esencialmente un horizonte infinito, sin estar sujeto a los problemas de elección del número de días m que se utilizan en la estimación de volatilidad mediante ventanas móviles.

Cuando se utiliza este modelo para generar una serie temporal de volatilidad histórica, es necesario comenzar a partir de una volatilidad inicial σ_0^2 , lo que puede hacerse de dos modos: 1) mediante la varianza de las rentabilidades previas a dicha fecha, que pasaría a ser tomada como origen de tiempo; es decir, utilizamos una primera submuestra (por ejemplo, 200 observaciones) para calcular dicha varianza y comenzamos a extrapolar la varianza en el tiempo partir de la observación 201; 2) alternativamente, suele partirse de un valor inicial igual a la varianza muestral de la serie temporal, es decir, se substituye σ_0^2 en la expresión anterior por la varianza muestral, $\frac{1}{T}\sum_{s=1}^{T} r_s^2$.

Elección del valor de λ

El modelo EWMA es el utilizado por *RiskMetrics* que calcula la volatilidad del próximo dia, mediante un promedio ponderado del nivel de volatilidad que calculamos previamente para hoy, y el cuadrado de la rentabilidad del mercado hoy. RiskMetrics utiliza sistemáticamente un valor numérico $\lambda = 0.94$, por considerar que las estimaciones no difieren mucho entre diferentes activos.

El parámetro λ es la *persistencia* en volatilidad o en covarianza. Mide la inercia en estos estadísticos: cuanto más presistentes sean, mayor será el tipo de rentabilidad extrema que deba observarse para que la estimación de la varianza

o de la covarianza se altere significativamente. Por el contrario, $1-\lambda$ mide la *capacidad de reacción* de estos momentos, varianza y covarianza, a rentabilidades extremas, positivas o negativas. Con $\lambda = 1$, la serie temporal de varianzas seria constante, Con $\lambda = 0$, la persistencia en varianza seria nula y la varianza estaría respondiendo cada periodo a los shock de mercado.

Ejemplo: De acuerdo con este modelo, si el nivel de volatilidad estimado un determinado día es del 1%, y la variación porcentual en precio dicho día es del 2%, utilizando un parámetro $\lambda = 0.90$, estimaríamos una volatilidad para el día siguiente de 1,14%. El parámetro puede estimarse por máxima verosimilitud, bajo un determinado supuesto acerca de la distribución de probabilidad de rentabilidades. Por ejemplo, si estamos dispuestos a suponer que siguen una distribución Normal y que son independientes, tendremos la función de verosimilitud:

$$\ln L(\lambda; r) = -\frac{1}{2}\sum_{t=1}^{T}\left(\ln \sigma_t^2(\lambda) + \frac{r_t^2}{\sigma_t^2(\lambda)}\right)$$

donde hemos hecho explícito que la serie temporal de varianzas dependerá del valor escogido para λ. Si la rentabilidad de la cartera tuviera autocorrelación, entonces tendriamo que utilizar en la función de verosimilitud las innovaciones ε_t en lugar las rentabilidades observadas r_t . Distintos valores de λ en el intervalo $(0, 1)$ generarán valores diferentes de la verosimilitud, y podremos encontrar el valor numérico de λ que maximiza $\ln L(\lambda; r)$.

Precisión en la estimación de la desviación típica del método EWMA

Supongamos que las rentabilidades siguen una distribución Normal con media cero. Tomando varianzas en (1) tenemos:

$$Var\left(\hat{\sigma}_\lambda^2\right) = \frac{(1-\lambda)^2}{1-\lambda^2}Var(r_t^2) = 2\frac{1-\lambda}{1+\lambda}\sigma^4$$

ya que, bajo Normalidad: $Var(r_t^2) = E(r_t^4) - E(r_t^2)^2 = 3\sigma^4 - \sigma^4 = 2\sigma^4$.
En términos porcentuales,

$$\frac{DT(\hat{\sigma}_\lambda^2)}{\sigma^2} = \sqrt{2\frac{1-\lambda}{1+\lambda}}$$

que disminuye al aumentar λ :

$$\lambda = 0.90 \Rightarrow \frac{DT(\hat{\sigma}_\lambda^2)}{\sigma^2} = 0.32;\ \lambda = 0.99 \Rightarrow \frac{DT(\hat{\sigma}_\lambda^2)}{\sigma^2} = 0.10$$

Como vimos anteriormente, también tenemos la aproximación que relaciona las varianzas de los estimadores de la varianza y de la desviación típica:

$$Var(\hat{\sigma}) = \frac{1}{4\hat{\sigma}^2}Var(\hat{\sigma}^2)$$

por lo que tenemos una desviación típica:

$$DT(\hat{\sigma}_\lambda) = \frac{1}{2\hat{\sigma}_\lambda} DT(\hat{\sigma}_\lambda^2) = \frac{1}{2\hat{\sigma}_\lambda}\sqrt{2\frac{1-\lambda}{1+\lambda}}\hat{\sigma}_\lambda^2 = \frac{1}{\sqrt{2}}\sqrt{\frac{1-\lambda}{1+\lambda}}\hat{\sigma}_\lambda \Rightarrow \frac{DT(\hat{\sigma}_\lambda)}{\hat{\sigma}_\lambda} = \sqrt{\frac{1-\lambda}{2(1+\lambda)}}$$

$$DT(\hat{\sigma}_\lambda) \simeq \hat{\sigma}_\lambda\sqrt{\frac{1-\lambda}{2(1+\lambda)}}$$

que da lugar a mayor precisión que la estimación de la varainza que antes obtuvimos:

$$\lambda = 0.90 \Rightarrow \frac{DT(\hat{\sigma}_\lambda^2)}{\sigma^2} = 0.23;\ \lambda = 0.99 \Rightarrow \frac{DT(\hat{\sigma}_\lambda^2)}{\sigma^2} = 0.07$$

Covarianzas cambiantes en el tiempo

Para covarianzas podemos utilizar una expresión recursiva similar:

$$\sigma_{12,t+1} = \lambda\sigma_{12,t} + (1-\lambda)\, r_{1t}r_{2t}$$

que suele denominarse como modelo de *suavizado exponencial,* en inglés: EWMA (exponentially weighted moving average).

El esquema EWMA nos puede permitir estimar determinados estadísticos importantes en el análisis de riesgos utilizando momentos muestrales calculados mediante EWMA. Así, tenemos betas:

$$\beta_{t\lambda} = \frac{Cov_\lambda(X_t, Y_t)}{Var_\lambda(X_t)}$$

y coeficientes de correlación:

$$\rho_{t\lambda} = \frac{Cov_\lambda(r_{1t}, r_{2t})}{\sqrt{Var_\lambda(r_{1t})Var_\lambda(r_{2t})}}$$

utilizando el mismo valor numérico de λ en el cálculo de las covarianzas y varianzas. Para convertir una varianza EWMA en volatilidad anual, se procede como con la varianza habitual.

6.2 Exponentially weighted moving average (EWMA) version of the one-factor model

Risk management requires monitoring on a frequent basis (daily and even intra-daily) and parameter estimates must be left to vary to reflect current risk conditions. So we consider:

$$r_t = a_t + \beta_t I_t + u_t$$

The simplest possible way to estimate time varying parameters is through an Exponentially Weighted Moving Average mechanism (EWMA), using a smoothing constant λ :

$$
\begin{aligned}
\beta_{\lambda,t} &= \frac{Cov_\lambda(r_t, I_t)}{Var_\lambda(I_t)} \\
Cov_\lambda(r_t, I_t) &\equiv \sigma_{12_t} = (1-\lambda) r_{t-1} I_{t-1} + \lambda \sigma_{12_{t-1}} = (1-\lambda) \sum_{i=1}^{\infty} \lambda^{i-1} r_{t-i} I_{t-i} \\
Var_\lambda(I_t) &\equiv \sigma_t^2 = (1-\lambda) I_{t-1}^2 + \lambda \sigma_{t-1}^2 = (1-\lambda) \sum_{i=1}^{\infty} \lambda^{i-1} I_{t-i}^2
\end{aligned}
$$

where we are assuming that the asset's return and the factor have zero expectation. A time varying correlation coefficient could similarly be defined by division of the covariance of both returns by the square root of the product of variances, both statistics defined as above. The value of λ,between 0 and 1, determines the persistence of the process of covariance or variance. A zero value would produce immediate reactions to events, while a value close to one would make the variance or covariance almost constant. The higher the value of λ, the longer it will take for the effects on moments of events to die away. The EWMA mechanism is justified only if returns are $i., i.d.$.

The value of λ can also be chosen to optimize a measure of fit, like the value of the log-likelihood function under Normality. It is sometimes chosen subjectively a it is the case with the 0.94 value used in Riskmetrics with daily data or the 0.97 value used with monthly data. A value of $\lambda = 0.97$ amounts to a half-life of 23 days, close to one month. That is the length of time needed for the process to close half the initial distance to its long-run level.

Exercise: For assets of different nature, compute covariances and variances for alternative values of λ. Compare with moments computed with rolling windows of different length. Estimate the value of λ.

Under the EWMA specification, systematic risk is estimated by:

$$
Systematic\ Risk = \sqrt{h} \hat{\beta}_{\lambda,t} \sqrt{Var_\lambda(I_t)}
$$

where h denotes the number of returns per year, which will be around 250 when working with daily data. This analysis produces time varying betas and correlations. It is obviously interesting to observe the time changes in beta, one of the two components of systematic risk of the asset. Systematic risk will change over time as a function of changes in beta and changes in factor variance. Systematic risk may be low even for assets with beta above one if factor risk is high, and the opposite can also happen.

There is an interesting relationship between the equity beta and the relative volatility of the asset and the market:

$$
\beta_{\lambda,t} = \rho_{\lambda,t} \sqrt{\frac{Var_\lambda(r_t)}{Var_\lambda(I_t)}} = \rho_{\lambda,t} \nu_{\lambda,t}
$$

where $\rho_{\lambda,t}$ is the correlation coefficient between the asset and the market and $\nu_{\lambda,t}$ is relative volatility, both calculated with the weighted moments.

A graphical comparison of time variation in correlation and relative volatility across assets will usually provide interesting information on the time evolution of beta-risk.

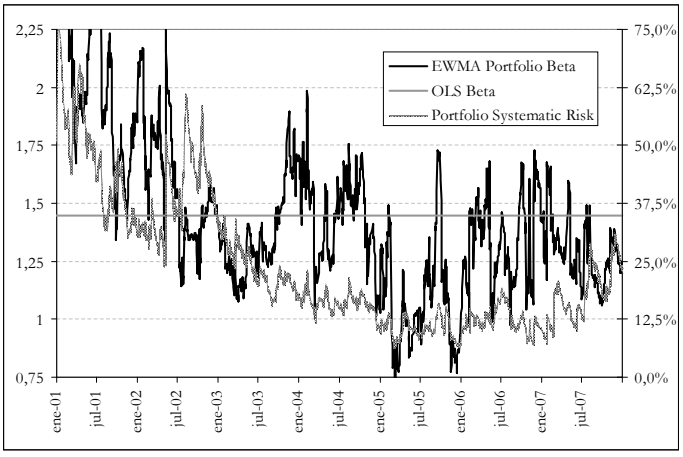
So,

$$Systematic\ Risk_t \equiv SR_t = \sqrt{h}\rho_{\lambda,t}\nu_{\lambda,t}\sqrt{Var_\lambda(I_t)}$$

Si el riesgo sistematico aumenta por aumentos en la volatilidad del indice, podriamos cubrirnos con futuros sobre dicho indices,mientras que si se debe a aumentos en la volatilidad relativa, podriamos utilizar futuros sobre la volatilidad del activo, si los hubiera. Aumentos en la correlacion podrian sugerir utilizar futuros sobre el indice, asimismo.

Example 13 Example (Figures II.1.1 to II.1.3 in file Figures II.1). *The example considers a portfolio made up by Amex and Cisco. It is obvious that Cisco has a greater systematic risk than Amex (larger β). The average market correlation is similar for Cisco and Amex, but Cisco is much more volatile than Amex, relative to the market and hence, EWMA correlation is much more unstable and Cisco beta is often considerably higher than Amex beta.*

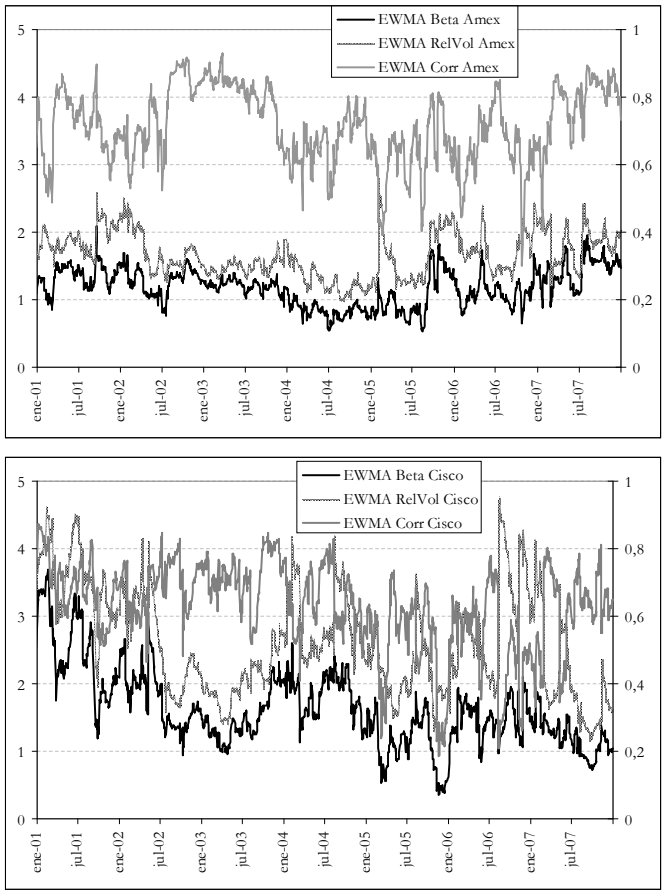
During 2001, the portfolio had a much higher β (left axis) than the OLS β estimate, and the opposite happens in the second part of the sample. Systematic risk (right axis) is the product of the portfolio β times the volatility of the market factor. This was low in the second part of the sample in spite of the fact that the β was larger than 1, because of low market volatility. High systematic risk in the summer of 2002 was not due to a high β but to a high market volatility. The OLS estimate of systematic risk is unable to reflect such tiume variation. Average market volatilty over the sample was 18.3%, so the OLS estimate of systematic risk is (18.3%)(1.448)=26.6%, an average of the level of systematic risk over the sample period.



Beta and relative volatility (left axis) and correlation (right axis) are shown in the next graphs. Clearly, Cisco has higher systematic risk than Amex. The

average market correlation is slightly higher for Amex (0.713 versus 0.658) but Cisco is much more volatile than Amex, relative to the market. As a consequence, EWMA correlation is more volatile for Cisco and its EWMA beta is considerable higher than Amex beta.

Beta, relative volatility and correlations for each asset



6.3 El modelo GARCH(1,1)

Un inconveniente del modelo previo es que no incluye una constante, por lo que el modelo no proporciona un nivel de referencia para la volatilidad a largo plazo. El modelo mejora si se incorpora un nivel de volatilidad de largo plazo, σ^2, que recibe una cierta ponderación γ en la expresión de la varianza,

$$\sigma_{t+1}^2 = \gamma\sigma^2 + \sum_{i=0}^{m-1} \alpha_i r_{t-i}^2$$

donde ahora, la suma de los pesos α_i debería ser igual a $1-\gamma$. Este es un modelo $ARCH(m)$. Si denotamos $\omega = \gamma\sigma^2$, tenemos,

$$\sigma_{t+1}^2 = \omega + \sum_{i=0}^{m-1} \alpha_i r_{t-i}^2$$

El modelo GARCH(1,1) combina las dos ideas anteriores proponiendo la siguiente expresión para la varianza condicional:

$$\sigma_t^2 = \gamma\sigma^2 + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2 = \omega + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2$$

que requiere que $\alpha + \beta < 1$ para que la varianza sea estable. En caso contrario, el peso aplicado a la varianza de largo plazo sería negativo. El alisado exponencial de la sección anterior, utilizado en RiskMetrics, es un caso particular del modelo $GARCH(1,1)$, cuando $\alpha + \beta = 1$ y $\gamma = 0$.

De acuerdo con este modelo,

$$\sigma^2 \equiv E\left(\sigma_{t+1}^2\right) = \omega + \alpha E\left(r_t^2\right) + \beta E\left(\sigma_t^2\right) = \omega + \alpha\sigma^2 + \beta\sigma^2$$

por lo que, la *volatilidad incondicional*, o *volatilidad de largo plazo* es,

$$\sigma^2 = \frac{\omega}{1 - \alpha - \beta} \tag{2}$$

de modo que la ponderación γ con que la volatilidad de largo plazo entra en el modelo GARCH(1,1) es:

$$\gamma = 1 - \alpha - \beta$$

Como en el modelo RiskMetrics, el modelo GARCH(1,1) nos permite prever la volatilidad del próximo día a partir de la volatilidad prevista para hoy y de la rentabilidad observada al cierre del mercado:

$$\sigma_{t+1}^2 = \omega + \alpha r_t^2 + \beta\sigma_t^2$$

De hecho, mediante sustituciones reiteradas, el modelo puede escribirse en la forma,

$$\sigma_t^2 = \omega + \omega\beta + \omega\beta^2 + \alpha r_{t-1}^2 + \alpha\beta r_{t-2}^2 + \alpha\beta^2 r_{t-3}^2 + \beta^3\sigma_{t-3}^2$$

Tomando límites, tenemos,

$$\sigma_t^2 = \frac{\omega}{1 - \beta} + \alpha \sum_{s=1}^{\infty} \beta^{s-1} r_{t-s}^2$$

que hace depender la volatilidad σ_t^2 de una constante y de las rentabilidades históricas al cuadrado, con ponderaciones decrecientes, según nos alejamos hacia el pasado. Es similar al modelo de suavizado exponencial EWMA, excepto en que asigna una ponderación también a la varianza de largo plazo. El parámetro β es la tasa a la cual el tamaño de las rentabilidades pasadas (o su volatilidad, si se quiere, pues están al cuadrado) inciden sobre la volatilidad actual del activo.

El modelo GARCH también puede escribirse,

$$\sigma_{t+1}^2 = (1-\alpha-\beta)\,\sigma^2 + \alpha r_t^2 + \beta\sigma_t^2 = \sigma^2 + \alpha(r_t^2 - \sigma^2) + \beta(\sigma_t^2 - \sigma^2)$$

que expresa que la previsión de la varianza el próximo día se obtiene corrigiendo el nivel de volatilidad de largo plazo (VLP), σ^2, en función de que la rentabilidad al cuadrado y el nivel de volatilidad en t hayan estado por encima o por debajo del nivel de largo plazo. Si ambas han estado por encima de la VLP, también σ_{t+1}^2 estará por encima de VLP. Si ambas han estado por deabjo de la VLP, también σ_{t+1}^2 estará por debajo de VLP. Si una de elllas ha estado por encima y otra por debajo de VLP, entonces σ_{t+1}^2 puede estar por encima o por debajo de la volatilidad de largo plazo.

En todo caso, la volatilidad incondicional (un número) es la media de la volatilidad condicional (una variable).

Example 14 *Si hemos estimado el modelo:* $\sigma_t^2 = 0,000002 + 0,13r_{t-1}^2 + 0,86\sigma_{t-1}^2$ *, tendríamos:* $\gamma = 1-\alpha-\beta = 0,01$; $\sigma^2 = \frac{\omega}{1-\alpha-\beta} = 0,0002$. *Al considerar las fluctuaciones en volatilidad, si la volatilidad estimada para un determinado día es de* $\sigma_t = 1,6\%$*, y ese día el precio del activo financiero varía un 1% al alza o a la baja, estimaríamos para el día siguiente una volatilidad:* $\sigma_t^2 = 0,000002 + 0,13\,(0,0001) + 0,86\,(0,000256) = 0,00023516$ *, que equivale a una volatilidad diaria del 1,53%. Este modelo implica un nivel de volatilidad de largo plazo* σ^2*, de* $\sqrt{0,0002} = 1,41\%$*. La volatilidad anual seria* $\sqrt{252}(1.41\%) = 22.38\%$.

La expresión (2) muestra que el nivel de volatilidad a largo plazo no está bien definida en el modelo de RiskMetrics, que impone $\alpha + \beta = 1$. Ello afectará más a las previsiones de volatilidad a largo plazo que a corto plazo, que vamos a calcular en la sección siguiente. Que esto sea o no importante depende de que creamos que existe un nivel de volatilidad media relativamente estable, a la cual revierte el mercado cada vez que se separa del mismo al alza o a la baja.

6.4 La rentabilidad diaria como estimación de volatilidad

En situaciones en que la rentabilidad no tiene estructura de autocorrelación, suele tomarse la rentabilidad diaria al cuadrado como una buena aproximación a la varianza condicional dado que $E_t r_{t+1}^2 = \sigma_{t+1}^2$, de modo que es un estimador insesgado de dicha varianza, pero puede ser poco preciso[9], pues su varianza puede ser grande:

$$\begin{aligned} Var_t r_{t+1}^2 &= E_t\left[(r_{t+1}^2 - \sigma_{t+1}^2)^2\right] = E_t\left[\left(\sigma_{t+1}^2\left(z_{t+1}^2 - 1\right)\right)^2\right] = \sigma_{t+1}^4 E_t\left[\left(z_{t+1}^2 - 1\right)^2\right] \\ &= \sigma_{t+1}^4\,(\kappa - 1) \Rightarrow DT_t r_{t+1} = \sigma_{t+1}^2\sqrt{(\kappa-1)} \end{aligned}$$

[9] Cuando la rentabilidad tiene estructura de autocorrelación, su varianza condicional es la varianza de la innovación del proceso autoregresivo que recoja dicha autocorrelación. En ese caso, es claro que la rentabilidad al cuadrado no seria una buena aproximacion a la misma.

siendo κ la curtosis de la innovación del proceso GARCH, z_t,que sería igual a 3 si suponemos Normalidad condicional: $z_t \sim i.,i.d.,N(0,1)$, pero que puede ser elevada en caso contrario. La segunda igualdad utiliza la representación $r_t = \sigma_t z_t$, que suponemos válida. La tercera igualdad utiliza el hecho de que en los modelos de volatilidad que estamos considerando, RiskMetrics y GARCH, σ^2_{t+1} es función de σ^2_t y r^2_t y por tanto, σ^2_{t+1} es conocida en t. La última igualdad utiliza: $E_t\left[\left(z^2_{t+1}-1\right)^2\right] = E_t\left(z^4_{t+1} - 2z^2_{t+1} + 1\right) = \kappa - 2E_t z^2_{t+1} + 1 = \kappa - 2 + 1 = \kappa - 1$.

El valor numérico de la expresión anterior puede ser elevado, por lo que el cuadrado de la rentabilidad de un período r^2_{t+1} es, generalmente, una proxy muy contaminada de la varianza condicional σ^2_{t+1}, que es la esperanza $E_t r^2_{t+1}$. Al poder ser la varianza muy grande, la realización numérica de la variable aleatoria es un mal indicador de su valor esperado. Por tanto, si disponemos de datos para ello, puede ser preferible utilizar medidas intradía en la estimación de la volatilidad.

6.5 Predicción de volatilidad: EWMA y GARCH(1,1)

En el modelo de *suavizado exponencial* utilizado por RiskMetrics $\alpha + \beta = 1$, y la predicción a cualquier horizonte coincide con la varianza actual. En efecto: $\sigma^2_{t+1} = \lambda\sigma^2_t + (1-\lambda)r^2_t$ es conocido en t . Para el periodo siguiente, utilizando $E_t\sigma^2_{t+1} = E_t\left(r_{t+1} - E_t r_{t+1}\right)^2 = E_t r^2_{t+1}$, tenemos:

$$\begin{aligned} \sigma^2_{t+2} &= \lambda\sigma^2_{t+1} + (1-\lambda)r^2_{t+1} \Rightarrow \\ E_t\sigma^2_{t+2} &= \lambda E_t\sigma^2_{t+1} + (1-\lambda)E_t(r^2_{t+1}) = \lambda E_t\sigma^2_{t+1} + (1-\lambda)E_t\sigma^2_{t+1} = E_t\sigma^2_{t+1} = \sigma^2_{t+1} \end{aligned}$$

De modo similar, puesto que $E_t r^2_{t+k} = E_t\sigma^2_{t+k}, \forall k > 1$, tenemos:

$$\sigma^2_{t+3} = \lambda\sigma^2_{t+2} + (1-\lambda)r^2_{t+2} \Rightarrow E_t\sigma^2_{t+3} = \lambda E_t\sigma^2_{t+2} + (1-\lambda)E_t(r^2_{t+2}) = E_t\sigma^2_{t+2}$$

El mismo razonamiento se aplicaría para todos los períodos sucesivos, por lo que:

$$E_t\left(\sigma^2_{t+k}\right) = \sigma^2_{t+1}\ \forall k$$

de modo que la volatilidad de este modelo tiene persistencia igual a 1. Se espera que todo shock en volatilidad persista para siempre, y cualquier incremento observado en volatilidad elevará la previsión de volatilidad de todos los períodos futuros en la cuantía del shock.

El modelo de RiskMetrics extrapola la situación de volatilidad actual a todos los períodos en el futuro, mientras que el modelo GARCH

$$\sigma^2_t = \omega + \alpha r^2_{t-1} + \beta\sigma^2_{t-1}$$

genera una reversión al nivel medio de volatilidad a largo plazo, como vamos a ver a continuación.

Recordemos que la suma $\alpha + \beta$ se denomina *persistencia* en volatilidad. Vamos a ver que:

- si la volatilidad actual es más alta que el nivel de largo plazo σ^2, la previsión será a la baja, y lo contrario ocurrirá si el nivel actual es de reducida volatilidad.
- cuando $\alpha + \beta < 1$, el último término va perdiendo importancia, y la predicción converge a la varianza de largo plazo, al aumentar el horizonte de la predicción.
- la velocidad de convergencia está inversamente relacionada con la proximidad de $\alpha + \beta$ a 1. Se dice que este modelo tiene *reversión al nivel medio de volatilidad,* σ^2, a una tasa $1 - \alpha - \beta$.

En efecto, de acuerdo con el modelo GARCH(1,1), tenemos para el periodo $t + 1$:

$$\begin{aligned} \sigma_{t+1}^2 &= \omega + \alpha r_t^2 + \beta\sigma_t^2 \Rightarrow \sigma_{t+1}^2 - \sigma^2 = \alpha(r_t^2 - \sigma^2) + \beta(\sigma_t^2 - \sigma^2) \\ &\Rightarrow E_t\sigma_{t+1}^2 - \sigma^2 = \alpha(r_t^2 - \sigma^2) + \beta(\sigma_t^2 - \sigma^2) \end{aligned}$$

En $t + 2$:

$$\sigma_{t+2}^2 = \omega + \alpha r_{t+1}^2 + \beta\sigma_{t+1}^2 \Rightarrow \sigma_{t+2}^2 - \sigma^2 = \alpha(r_{t+1}^2 - \sigma^2) + \beta(\sigma_{t+1}^2 - \sigma^2)$$

Tomando esperanzas condicionales y teniendo en cuenta que: $E_t r_{t+1}^2 = E_t\sigma_{t+1}^2$, tenemos:

$$E_t\left(\sigma_{t+2}^2 - \sigma^2\right) = (\alpha + \beta)\left(E_t\sigma_{t+1}^2 - \sigma^2\right)$$

e iterando:

$$E\left(\sigma_{t+k}^2 - \sigma^2\right) = (\alpha + \beta)^{k-1}\left(E_t\sigma_{t+1}^2 - \sigma^2\right) = (\alpha + \beta)^{k-1}\left[\alpha(r_t^2 - \sigma^2) + \beta(\sigma_t^2 - \sigma^2\right])$$

Por tanto, segun alejamos el horizonte de predicción, la predicción de volatilidad converge a:

$$E_t\sigma_{t+k}^2 = \sigma^2 = \frac{\omega}{1 - (\alpha + \beta)}$$

que es la varianza incondicional o varianza de largo plazo (como se vio en la sección anterior al hablar del modelo GARCH(1,1)). Esto es lo que cabe esperar que suceda con las predicciones de toda variable estacionaria. La convergencia de la predicción de varianza a su nivel de largo plazo es ademas monótona, a una tasa $\alpha + \beta$, aunque puede llevar mucho tiempo alcanzar dicho nivel si la persistencia es muy elevada.

Si hacemos la predicción en términos de esperanza matemática incondicional, tendríamos, utilizando $E(r_t^2) = \sigma_t^2$:

$$\sigma_{t+1}^2 - \sigma^2 = \alpha(r_t^2 - \sigma^2) + \beta(\sigma_t^2 - \sigma^2) \Rightarrow E\left(\sigma_{t+1}^2 - \sigma^2\right) = (\alpha + \beta)\left(\sigma_t^2 - \sigma^2\right)$$

e iterando:

$$E\left(\sigma_{t+k}^2 - \sigma^2\right) = (\alpha + \beta)^k \left(\sigma_t^2 - \sigma^2\right)$$

Por tanto, podemos escribir también la expresión para la predicción de la varianza condicional del modelo GARCH(1,1) como:

$$E_t\sigma_{t+k}^2 = \sigma^2 + (\alpha + \beta)^{k-1}\left(\sigma_t^2 - \sigma^2\right)$$

donde podemos ver que si la varianza condicional actual es superior a la varianza de largo plazo, entonces la predicción de la varianza condicional para los proximos periodos seguirá una senda descendente. Lo contrario sucederá si la varianza actual es inferior a la varianza de largo plazo.

Algo más delicada es la *predicción de la volatilidad de la rentabilidad acumulada* a lo largo de k días de mercado. Nótese la diferencia con el ejercicio anterior, en el que se preveía la rentabilidad diaria k-períodos hacia adelante. En términos de rentabilidades continuas, sabemos que dicha rentabilidad acumulada será, por construcción, la suma de las rentabilidades continuas obtenidas para cada uno de los días del período. Bajo el supuesto de que las rentabilidades son temporalmente independientes, tendremos,

$$E_t\left(\sum_{i=1}^{k} R_{t+i}\right)^2 = \sum_{i=1}^{k} E_t\sigma_{t+i}^2$$

de manera que con RiskMetrics tenemos,

$$E_t\left(\sum_{i=1}^{k} R_{t+i}\right)^2 = k\sigma_{t+1}^2$$

mientras que con el modelo GARCH tenemos,

$$E_t\left(\sum_{i=1}^{k} R_{t+i}\right)^2 = k\sigma^2 + \sum_{i=1}^{k}(\alpha + \beta)^{i-1}\left(\sigma_t^2 - \sigma^2\right) = k\sigma^2 + \left(\sigma_t^2 - \sigma^2\right)\frac{1 - (\alpha + \beta)^k}{1 - (\alpha + \beta)}$$

Si partimos de un mismo nivel de σ_{t+1}^2, la predicción del modelo GARCH será superior a la de RiskMetrics si y sólo si el nivel de σ_{t+1}^2 es inferior al nivel de largo plazo, σ^2, y lo contrario sucederá si $\sigma_{t+1}^2 > \sigma^2$.

6.6 Estimación del modelo de volatilidad por máxima verosimilitud

Como ya mencionamos al describir el modelo de suavizado exponencial, bajo el supuesto de Normalidad para las rentabilidades logarítmicas, $r_t = \sigma_t z_t$, con $z_t \sim i., i.d.N(0,1)$, tenemos la función de verosimilitud,

$$L = \prod_{t=1}^{T} \left[\frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(\frac{-r_t^2}{2\sigma_t^2} \right) \right]$$

y su logaritmo neperiano,

$$\ln L = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^{T} \left(\ln \sigma_t^2 + \frac{r_t^2}{\sigma_t^2} \right)$$

que se puede maximizar bien mediante algoritmos numéricos, o bien mediante procedimientos de búsqueda. Evidentemente, el vector de valores paramétricos que maximiza $\ln L$ es el mismo que el que maximiza la función de verosimilitud, L.

El supuesto de Normalidad no es facilmente sostenible cuando se trabaja con rentabilidades de activos financieros. El método de quasi-máxima verosimilitud consiste en maximizar el logaritmo de la función de verosimilitud bajo el supuesto de Normalidad, pues el estimador resultante es consistente incluso cuando la verdadera distribución condicional de r_t no es Normal, siempre que las ecuaciones de la media y la varianza condicionales de r_t estén bien especificadas. Unicamente hay que prestar atención al cálculo de la matriz de covarianzas de los estimadores resultantes.

En todo caso, lo primero que hemos de hacer es substitutir en la expresión anterior la volatilidad σ_t^2 por un determinado modelo dependiente de un vector de parámetros θ. En las próximas secciones veremos cómo se lleva a cabo este proceso. Como en cualquier otro problema de estimación, hemos de tener en cuenta que estamos maximizando la verosimilitud bajo el supuesto de estabilidad paramétrica, lo que puede condicionar el número de observaciones utilizado en dicho proceso de estimación.

Primer caso: rentabilidades incorrelacionadas con media cero Supongamos que las rentabilidades obtenidas en la unidad temporal de observación carecen de autocorrelación, lo que puede contrastarse a partir de un examen de sus funciones de autocorrelación simple y parcial, así como llevando a cabo contrastes formales del tipo Ljung-Box o Box-Pierce.

Para estimar los parámetros del modelo en una hoja de cálculo, se estima inicialmente $\sigma_{t_0}^2$ por alguno de los dos procedimientos que mencionamos antes, y comienza la recursión a partir de dicho instante temporal, después de haber fijado valores iniciales para los parámetros α, β, ω. Una vez evaluada la función de verosimilitud para los valores parámetricos inicialmente escogidos (*condiciones*

iniciales), se trata de buscar en el espacio paramétrico con el objeto de obtener los valores que maximizan la función de verosimilitud,

$$\ln L(\omega, \alpha, \beta) = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\ln \sigma_t^2(\omega, \alpha, \beta) + \frac{r_t^2}{\sigma_t^2(\omega, \alpha, \beta)}\right)$$

Finalmente, la varianza de largo plazo, σ^2, se estima a partir de las expresiones anteriores y las estimaciones obtenidas para α, β, ω: $\sigma^2 = \omega/(1-\alpha-\beta)$.

La alternativa denominada *variance targetting* consiste en fijar un nivel de volatilidad de largo plazo σ^2, por ejemplo igual a la varianza muestral, y utilizando la expresión analítica de la varianza a largo plazo para fijar el valor de la constante del modelo de varianza en función de los otros dos parámetros: $\omega = \sigma^2(1-\alpha-\beta)$, estimando así sólo 2 parámetros, α y β.

Si queremos estimar un modelo de alisado exponencial como el utilizado en RiskMetrics, se fija $\omega = 0, \alpha = 1-\lambda, \beta = \lambda$, y se efectúa una búsqueda sobre el valor numérico de $\lambda, \lambda \in (0,1)$,en la función

$$\ln L(\lambda) = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\ln \sigma_t^2(\lambda) + \frac{r_t^2}{\sigma_t^2(\lambda)}\right)$$

Segundo caso: rentabilidades posiblemente correlacionadas, con media no nula Como alternativa, consideremos la posibilidad de que las rentabilidades obedezcan al modelo

$$r_t = \rho_0 + \rho_1 r_{t-1} + \varepsilon_t$$

que recoge la presencia de autocorrelación, es decir, de dependencia temporal en las rentabilidades. Tendría sentido entonces hacer el supuesto de estructura GARCH de volatilidad, pero ahora sobre la innovación ε_t del proceso estocástico de rentabilidades, por lo que σ_t^2 sería ahora: $\sigma_t^2 = Var(\varepsilon_t)$, con función de verosimilitud,

$$L = \prod_{t=1}^{T}\left[\frac{1}{\sqrt{2\pi\sigma_t^2}}\exp\left(\frac{-\varepsilon_t^2}{2\sigma_t^2}\right)\right]$$

con,

$$\ln L = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\ln\sigma_t^2 + \frac{\varepsilon_t^2}{\sigma_t^2}\right) = cons\tan te - \frac{1}{2}\sum_{t=1}^{T}\left(\ln\sigma_t^2 + \frac{(r_t - \rho_0 - \rho_1 r_{t-1})^2}{\sigma_t^2}\right)$$

y la estimación del modelo se lleva a cabo con respecto a los parámetros $\alpha, \beta, \omega, \rho_0, \rho_1$.

En este caso habría que tener en cuenta que el procedimiento nos daría la evolución temporal de la volatilidad de la innovación ε_t, el componente no

predecible de la rentabilidad, que es la volatilidad de r_t condicional en su pasado. Bajo el supuesto de estructura AR(1) para r_t, la relación entre las volatilidades incondicionales de la Rentabilidad y su innovación es:

$$Var(r_t) = \frac{Var(\varepsilon_t)}{1-\rho_1^2}$$

6.7 Validación del modelo de volatilidad

Un *contraste* del modelo consiste en un test de ausencia de autocorrelación en las rentabilidades al cuadrado, r_t^2. Puesto que hemos pretendido recoger en nuestro modelo la evolución temporal de la volatilidad, no deberia existir tal autocorrelación. Para ello hemos de utilizar las *rentabilidades normalizadas o estandarizadas* al cuadrado, $\frac{r_t^2}{\sigma_t^2}$. Una vez estimada la función de autcorrelación, la desviación tipica de cada uno de sus valores γ_i por separado puede aproximarse por $\frac{1}{\sqrt{T}}$, de modo que, por ejemplo, $\pm\frac{2}{\sqrt{T}}$ podria utilizarse como banda de confianza del 95% para el contraste de la hipótesis nula: $H_0 : \gamma_i = 0$.

Para un contraste riguroso, que considera varios valores de la función de autocorrelación, puede utilizarse el estadistico de Ljung-Box,

$$T\sum_{i=1}^{k}\frac{T-2}{T-i}\gamma_i^2 \sim \chi_k^2$$

donde γ_i denota los sucesivos valores numéricos de la funcion de autocorrelación de R_t^2/σ_t^2. En realidad, se trata de una familia de estadisticos, pues puede y debe utilizarse para diferentes valores de k.

Contrastes relevantes son asimismo los pertenecientes a la familia de *tests de razón de verosimilitudes*, que permiten contrastar un modelo restringido frente a una alternativa más general, del cual el primero se obtiene imponiendo determinadas restricciones. El estadístico del contraste es,

$$LRT = -2\left[\ln L_R - \ln L_{NR}\right]$$

y tiene una distribución asintótica igual a una chi-cuadrado con grados de libertad igual al número de restricciones que transforman el modelo más general en el modelo restringido.

Otro posible contraste consiste en analizar si la serie temporal de varianza, σ_{t+1}^2, cuyo valor numérico es determinado con información disponible en t, es un *predictor insesgado* de la rentabilidad al cuadrado futura [ver Ejercicio 2.4 en Christoffersen],

$$r_{t+1}^2 = \alpha + \beta\sigma_{t+1}^2 + u_{t+1}$$

y se dice que σ_{t+1}^2 , que es determinado con información en t,es un predictor insesgado de r_{t+1}^2 si en esta regresión se tiene: $\alpha = 0$ y $\beta = 1$. Como vimos antes, la rentabilidad diaria al cuadrado es un estimador poco preciso de la varianza diaria. Eso hará que en esta regresión la varianza residual pueda ser alta y el

$R2$ posiblemente menor de lo que cabría esperar. Pero los parámetros α y β debería tomar los valores indicados.

6.8 Estructura temporal de volatilidad

Consideremos una opción que vence en $t + N$. Podemos utilizar la expresión anterior para predecir el nivel medio de volatilidad del activo subyacente durante dicho período, mediante,

$$Volatilidad\ a\ horizonte\ N\ períodos = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1} E_t\sigma_{t+i}^2}$$

Cuando este ejercicio se lleva a cabo para opciones sobre el mismo activo subyacente, con distinta fecha de vencimiento, se tiene una Estructura Temporal de Volatilidades. Esta es la relación entre las volatilidades implícitas de las opciones y su vencimiento residual.

Cuando se utiliza el modelo *GARCH*, se obtiene un perfil creciente o decreciente para las previsiones de varianza, precisamente por su propiedad de *reversión al nivel medio* de volatilidad. Por tanto, este modelo predice una curva de volatilidades que será o bien creciente o bien decreciente respecto del vencimiento de las opciones.

Si denotamos: $V_k = E_t\sigma_{t+k}^2$ y $a = \ln\frac{1}{\alpha+\beta}$, la ecuación de predicción $E\sigma_{t+k}^2 = \sigma^2 + (\alpha+\beta)^k(\sigma_t^2 - \sigma^2)$ se convierte en:

$$V_k = \sigma^2 + e^{-ak}\left(V_0 - \sigma^2\right)$$

siendo $V(k)$ una estimación actual de la volatilidad instantánea dentro de k dias. La varianza media sobre un período de T dias a partir de hoy, será:

$$\frac{1}{T}\int_0^T V_k dk = \sigma^2 + \frac{1-e^{-aT}}{aT}\left[V_0 - \sigma^2\right]$$

que se aproximará más a la varianza a largo plazo σ^2 cuanto más distante sea el vencimiento de la opción. Si denotamos por $\sigma(T)$ la volatilidad anualizada que debemos utilizar para valorar una opción que vence en T dias según un modelo GARCH(1,1), tendremos:

$$\sigma_T^2 = 252\left(\sigma^2 + \frac{1-e^{-aT}}{aT}\left[V_0 - \sigma^2\right]\right)$$

de manera que si la volatilidad actual está por encima de la volatilidad a largo plazo, el modelo GARCH genera una estructura temporal de volatilidades decreciente, mientras que si la volatilidad actual es inferior al nivel de volatilidad a largo plazo, la estructura temporal es creciente.

En el cociente $\frac{1-e^{-aT}}{aT}$, el numerador permanece acotado, mientras que el denominador crece linealmente con T. Por tanto, el cociente tiende a cero y σ_T^2 converge a $252\sigma^2$ cuando el vencimiento T tiende a infinito.

Puede calcularse asimismo cual sería el efecto sobre cada una de dichas previsiones, de una variación por ejemplo, de un 1% en la volatilidad actual del activo subyacente. El efecto no será el mismo para todos los vencimientos, y se obtiene un perfil análogo al de la Estructura Temporal de Volatilidades, lo que debería tenerse en cuenta al estimar la exposición de una cartera de opciones a variaciones en volatilidad del activo subyacente. Al hacer una simulación de este tipo y calcular una *vega*, no debería suponerse una variación análoga en volatilidad a lo largo de todos los vencimientos. Si reescribimos dicha ecuación como:

$$\sigma_T^2 = 252\left(\sigma^2 + \frac{1-e^{-aT}}{aT}\left[\frac{\sigma_0^2}{252} - \sigma^2\right]\right)$$

entonces, cuando la volatilidad actual, σ_0 , cambia en $\Delta\sigma_0$, σ_T cambia en una cantidad:[10]

$$\Delta\sigma_T = \frac{1-e^{-aT}}{aT}\frac{\sigma_0}{\sigma_T}\Delta\sigma_0$$

[Hull: Risk Management and Financial Institutions, Ch.5]

Example 15 *Consideremos un activo con* $\alpha+\beta = 0,99351$, *y* $\sigma^2 = 0.0002075$. *Supongamos que la varianza actual es:* $V_0 = 0,0003$ *por dia. Muestre que la estructura temporal de volatilidades para* $T = 10, 30, 50, 100$ *y* 500 *dias es:*$\sigma_{10} = 27,36\%; \sigma_{30} = 27,10\%; \sigma_{50} = 26,87\%; \sigma_{100} = 26,35\%; \sigma_{500} = 24,32\%$

Example 16 *En el ejemplo anterior, suponga que la volatilidad actual sube 100 puntos básicos:* $\Delta\sigma_0 = 0,01$. *Muestre que el efecto sobre la volatilidad de opciones que vencen en los periodos citados es:*$\Delta\sigma_{10} = 0,97; \Delta\sigma_{30} = 0,92; \Delta\sigma_{50} = 0,87; \Delta\sigma_{100} = 0,77; \Delta\sigma_{500} = 0,33\%$

6.9 Otros modelos GARCH (completar con notas de clase sobre modelos de volatilidad condicional)

- Modelo GARCH(p,q):

$$\sigma_{t+1}^2 = \gamma\sigma^2 + \sum_{i=1}^{p}\alpha_i R_{t+1-i}^2 + \sum_{j=1}^{q}\beta_j\sigma_{t+1-j}^2$$

Modelo GARCH(1,1):

$$\sigma_{t+1}^2 = \gamma\sigma^2 + \alpha R_t^2 + \beta\sigma_t^2$$

- Modelo GARCH de componentes, que permite variación temporal en el nivel de varianza de largo plazo, v_{t+1} :

[10] Diferenciando en la expresión que relaciona σ_T con σ_0, tenemos:
$2\sigma_T d\sigma_T = 252\left(\frac{1-e^{-aT}}{aT}\right)\frac{1}{252}2\sigma_0 d\sigma_0$

$$\begin{aligned}\sigma_{t+1}^2 &= \upsilon_{t+1} + \alpha(R_t^2 - \upsilon_t) + \beta(\sigma_t^2 - \upsilon_t) \\ \upsilon_{t+1} &= \omega + \alpha_\upsilon(R_t^2 - \sigma_t^2) + \beta_\upsilon \upsilon_t\end{aligned}$$

Este modelo es un modelo GARCH(2,2) restringido.

- Efecto apalancamiento (*leverage*): El argumento básico es que una rentabilidad negativa de una acción implica una caida en el valor de mercado de la empresa, lo que aumenta su apalancamiento financiero, aumentando su nivel de riesgo (a igual nivel de deuda). Podemos modificar el modelo GARCH(1,1) para recoger este efecto de varias maneras:

$$\sigma_{t+1}^2 = \omega + \alpha R_t^2 + \alpha\theta\left(I_t^- R_t^2\right) + \beta\sigma_t^2$$

donde I_t^- es una variable ficticia que toma el valor 1 cuando la rentabilidad es negativa, siendo igual a cero en caso contrario. Definida de este modo, una rentabilidad R tiene una contribución al nivel de volatilidad el período siguiente de αR_t^2, si dicha rentabilidad fue positiva, y de $\alpha\left(1+\theta\right)R_t^2$ si fue negativa Este modelo se conoce como GJR-GARCH.

Bajo el supuesto mantenido de que la rentabilidad sigue un proceso: $R_t = \sigma_t z_t$, con $z_t \sim i.,i.d.N(0,1)$, otra posibilidad es el modelo NGARCH,

$$\sigma_{t+1}^2 = \omega + \alpha(R_t - \theta\sigma_t)^2 + \beta\sigma_t^2 = \omega + \alpha\sigma_t^2(z_t - \theta)^2 + \beta\sigma_t^2$$

de modo que, si $\theta > 0$, noticias positivas tienen menos impacto sobre la varianza que noticias negativas. La persistencia de la varianza en este modelo es $\alpha\left(1+\theta^2\right) + \beta$, mientras que el nivel de varianza de largo plazo es: $\sigma^2 = \frac{\omega}{1-\alpha(1+\theta^2)-\beta}$.

Una última posibilidad es el modelo GARCH exponencial, o EGARCH:

$$\ln\sigma_{t+1}^2 = \omega + \alpha(\phi R_t + \gamma\left[\mid R_t \mid -\sqrt{2/\pi}\right] + \beta\ln\sigma_t^2$$

que presenta efecto apalancamiento si $\alpha\phi < 0$. Por otra parte, la especificación logarítmica garantiza que la varianza resultante será positiva en todos los períodos. En la expresión anterior, $\sqrt{2/\pi}$ aparece por ser la esperanza matemática del valor absoluto de la rentabilidad: $\sqrt{2/\pi} = E(\mid R_t \mid)$.

- Inclusión de variables explicativas, como el efecto fin de semana, mediante una variable ficticia que tome el valor 1 los lunes, así como tras días festivos, anuncios macroeconómicos, reuniones de la Fed, etc. También podría considerarse la inclusión de un índice de volatilidad tipo VIX cuando queremos prever la volatilidad del subyacente de las opciones con las que se ha calculado dicho índice.

- Algunos modelos univariantes

La ecuación de la varianza en el modelo $EGARCH(1,1)$ es:

$$\begin{aligned} y_t &= \varepsilon_t h_t \\ \ln h_t^2 &= \omega + \beta \ln h_{t-1}^2 + \delta \varepsilon_{t-1} + \theta \left(\mid \varepsilon_{t-1} \mid - \sqrt{2/\pi} \right) \end{aligned}$$

Modelo *GJR:*

$$h_t^2 = \omega + \sum_{i=1}^{q} \left[\alpha_i \varepsilon_{t-i}^2 + \alpha_i^- I \left(\varepsilon_{t-i} \leq 0 \right) \varepsilon_{t-i}^2 \right] + \sum_{j=1}^{p} \beta_j h_{t-j}^2$$

que en un caso sencillo, sería:

$$h_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \alpha^- I \left(\varepsilon_{t-1} \leq 0 \right) \varepsilon_{t-1}^2 + \beta_1 h_{t-1}^2$$

7 Modelos de correlación condicional

7.1 Estimación de covarianzas y correlaciones condicionales

Recordemos lo que hemos visto en secciones previas acerca de la estiamción de varianzas y covarianzas cambiantes en el tiempo: Supongamos que tenemos N activos, con los cuales podemos configurar una cartera, para lo que necesitaremos estimar su matriz de covarianzas describiendo su evolución a lo largo del tiempo. O, alternativamente, supongamos que queremos caracterizar las matrices de covarianzas y de correlaciones de N factores de riesgo. La representación más sencilla de una covarianza cambiante en el tiempo podría obtenerse a partir de una ventana móvil de m datos, utilizando una expresión similar a la que propusimos para la estimación de la varianza:

$$\sigma_{ij,t+1} = \frac{1}{m} \sum_{s=0}^{m-1} r_{i,t-s} r_{j,t-s}$$

con las limitaciones que ya conocemos por la presencia de la amplitud de ventana m. En esta expresión estamos incorporando el supuesto de que las rentabilidades tienen media cero, lo que será muy aceptable en datos de alta frecuencia.

Este tipo de modelización puede generar excesiva variabilidad en la serie de covarianzas. Por tanto, puede resultar conveniente introducir persistencia mediante un suavizado exponencial en covarianzas (RiskMetrics),

$$\sigma_{ij,t+1} = (1 - \lambda) r_{i,t} r_{j,t} + \lambda \sigma_{ij,t} \tag{3}$$

que tiene la limitación que ya vimos en el caso de estimación de la varianza, en el sentido de que no existe un nivel de referencia que pudiera interpretarse como la covarianza a largo plazo. Por tanto, al igual que en aquél caso, este modelo implica que no existe reversión a la media en las covarianzas. Si la covarianza

es alta y positiva (por ejemplo) un dia, el modelo predice que continuará siendo elevada y positiva. En todo caso, este es el modelo utilizado por RiskMetrics, con $\lambda = 0,94$.

Una vez que disponemos de expresiones correspondientes a una determinada metodología, como en este caso las utilizadas por RiskMetrics, la estimación de una serie temporal de correlaciones puede efectuarse mediante:

$$\rho_{uv,t} = \frac{\sigma_{uv,t}}{\sigma_{u,t}\sigma_{v,t}}$$

Example 17 *Supongamos que las volatilidades diarias estimadas para dos activos son $\sigma_{u,t}$ =1% y $\sigma_{v,t}$ =2%. Con un parámetro de suavizado exponencial: $\lambda = 0,95$, si se producen variaciones diarias en precios de $r_{u,t}$ =0,5% y $r_{v,t}$ =2,5%, respectivamente, las nuevas varianzas serían,*

$$\begin{aligned}\sigma^2_{u,t+1} &= (0.95)(0.01)^2 + (0.05)(0.005)^2 = 0,00009625 \Rightarrow \sigma_{u,t+1} = 0,00981 = 0,981\% \\ \sigma^2_{v,t+1} &= (0.95)(0.02)^2 + (0.05)(0.025)^2 = 0,00041125 \Rightarrow \sigma_{v,t+1} = 0,02028 = 2,028\%\end{aligned}$$

Si la correlación actual es $\rho_{uv,t}$ =0,60, la covarianza estimada entre ambas rentabilidades sería:

$$\begin{aligned}\sigma_{uv,t} &= \rho_{uv,t}\sigma_{u,t}\sigma_{v,t} = (0,60)(0,01)(0,02) = 0,00012 \\ \sigma_{uv,t+1} &= (0.95)\sigma_{uv,t} + (0,05)u_t v_t = (0.95)(0.00012) + (0.05)(0.005)(0.025) = 0,00012025\end{aligned}$$

El nuevo coeficiente de correlación sería:

$$\rho_{uv,t+1} = \frac{\sigma_{uv,t+1}}{\sigma_{u,t+1}\sigma_{v,t+1}} = \frac{0.00012025}{(0.00981)(0.02028)} = 0,60443$$

Una segunda alternativa consistiría en utilizar el esquema GARCH(1,1) de covarianza, que presenta reversión a la media,

$$\sigma_{ij,t+1} = \omega_{ij} + \alpha r_{i,t} r_{j,t} + \beta \sigma_{ij,t}$$

según el cual la covarianza entre los activos i, j revertirá a su nivel de largo plazo,

$$\sigma_{ij} = \frac{\omega_{ij}}{1 - \alpha - \beta}$$

Las expresiones anteriores pueden utilizarse asimismo para estimar covarianzas y correlaciones entre los residuos u_t, v_t de dos modelos estimados, mediante:

$$\begin{aligned}\sigma_{uv,t} &= \lambda\sigma_{uv,t-1} + (1-\lambda)\, u_{t-1}v_{t-1} \text{ (suavizado exponencial)}, \\ \sigma_{uv,t} &= \omega + \beta\sigma_{uv,t-1} + \alpha u_{t-1}v_{t-1} \quad \text{(GARCH(1,1))}\end{aligned}$$

para estimar posteriormenta la correlación mediante: $\rho_{uv,t} = \frac{\sigma_{uv,t}}{\sigma_{u,t}\sigma_{v,t}}$.

Imponer los mismos parámetros de persistencia, α y β en la estimación de las varianzas y covarianzas de los distintos activos considerados garantiza que obtengamos una matriz de covarianzas definida positiva, lo que entendemos como matriz de covarianzas *internamente consistente*.[11] Lo mismo sucede si en el esquema de suavizado exponencial que vimos inicialmente, utilizamos el mismo valor del parámetro λ para todos los activos. Sin embargo, la homogeneidad de valores numéricos de los parámetros de persistencia puede ser una restricción poco razonable, por lo que consideramos a continuación modelos que no la imponen.

Calcular la correlación condicional como cociente entre la covarianza condicional y la raiz cuadrada del producto de varianzas condicionales, utilizando en el numerador y denominador de dicho cociente las expresiones de RiskMetrics o un esquema GARCH(1,1) puede generar algunos problemas. De hecho, es generalmente preferible trabajar directamente con las correlaciones, que son mas facilmente interpretables, y obtener las covarianzas como consecuencia del comportamiento de correlaciones, por un lado, y de varianzas, por otro. Por ejemplo, las covarianzas puden variar en el tiempo incluso si las correlaciones son constantes, simplemente porque las varianzas sean cambiantes en el tiempo (como veremos en el modelo correlación condicional constante, CCC):

La relación entre covarianzas y correlaciones puede escribirse:

$$\sigma_{ij,t+1} = \sigma_{i,t+1}\sigma_{j,t+1}\rho_{ij,t+1}$$

o, en notación matricial,

$$\Sigma_{t+1} = D_{t+1}\Gamma_{t+1}D_{t+1}$$

donde D_{t+1} es una matriz con desviaciones típicas condicionales en la diagonal y ceros fuera de la diagonal, y Γ_{t+1}es una matriz con unos en la diagonal, y con las correlaciones condicionales fuera de dicha diagonal principal. En el caso de dos activos:

$$\Sigma_{t+1} = \begin{pmatrix} \sigma^2_{1,t+1} & \sigma_{12,t+1} \\ \sigma_{12,t+1} & \sigma^2_{2,t+1} \end{pmatrix} = \begin{pmatrix} \sigma_{1,t+1} & 0 \\ 0 & \sigma_{2,t+1} \end{pmatrix} \begin{pmatrix} 1 & \rho_{12,t+1} \\ \rho_{12,t+1} & 1 \end{pmatrix} \begin{pmatrix} \sigma_{1,t+1} & 0 \\ 0 & \sigma_{2,t+1} \end{pmatrix},$$

que es el modo en que generamos la matriz de covarianzas una vez que hayamos estimado una matriz de correlaciones.

Supongamos que las volatilidades de cada activo ya han sido estimadas previamente mediante modelos GARCH univariantes. Si estandarizamos las rentabilidades,

$$z_{i,t+1} = \frac{r_{i,t+1}}{\sigma_{i,t+1}},\ \forall i, t$$

[11] Hay *consistencia interna* cuando $\omega'\Sigma_{t+1}\omega \geq 0$ para toda cartera definida por el vector de ponderaciones ω.

las rentabilidades estandarizadas $z_{i,t+1}$ tienen desviación típica condicional igual a 1. Por tanto, la covarianza condicional entre rentabilidades estandarizadas coincide con la correlación condicional de las rentabilidades originales:

$$E_t(z_{i,t+1}z_{j,t+1}) = E_t\left(\frac{r_{i,t+1}}{\sigma_{i,t+1}}\frac{r_{j,t+1}}{\sigma_{j,t+1}}\right) = \frac{E_t(r_{i,t+1}r_{j,t+1})}{\sigma_{i,t+1}\sigma_{j,t+1}} = \frac{\sigma_{ij,t+1}}{\sigma_{i,t+1}\sigma_{j,t+1}} = \rho_{ij,t+1}$$

donde hemos utilizado el hecho de que en modelos GARCH, las desviaciones típicas $\sigma_{i,t+1}$ y las covarianzas en $t+1$ son conocidas en t.

Por tanto, modelizar la correlacion condicional de las rentabilidades originales equivale a modelizar la covarianza condicional de las rentabilidades estandarizadas. Vamos a considerar 3 modelos: el primero que estima una correlación condicional constante, y dos modelos adicionales, EWMA y GARCH, para aproximar el modo en que la correlación entre dos activos evoluciona en el tiempo:

7.2 Modelo de correlación condicional constante (CCC)

Este modelo consiste en calcular el coeficiente de correlacion lineal estándar (coeficiente de correlación de Pearson) para cada par de rentabilidades estandarizadas. Para ello comenzamos estimando un modelo de volatilidad condicional de cada rentabilidad (de ahi la denominación "condicional" del modelo de correlación), y estandarizamos cada una de ellas dividiendo el dato de cada día por la desviación típica estimada (no la varianza) de ese día.

La correlacion condicional constante es el coeficiente de correlacion lineal entre cada par de rentabilidades estandarizadas. De este modo obtendremos un unico valor numerico para la correlación entre cada par de activos, válido para toda la muestra, y construimos una matriz de correlaciones Γ constante en el tiempo.[12]

Una vez estimada la matriz de correlaciones Γ, las covarianzas condicionales se obtienen:

$$\Sigma_{t+1} = D_{t+1}\Gamma D_{t+1}$$

siendo los elementos de la diagonal en este producto las mismas varianzas condicionales estimadas a partir de modelos univariantes y que han sido utilizadas en la estandarización de las rentabilidades. Nótese que, de acuerdo con este modelo, las covarianzas condicionales variarán en el tiempo, a pesar de que las correlaciones condicionales no lo hacen, debido a que las varianzas condicionales son cambiantes en el tiempo, ya sea mediante un esquema de suavizado exponencial o mediante un modelo GARCH.

[Chapter3Results.xls in Christoffersen, pestañas Question 6, Question 7 y Question 8].

[12] Cada uno de sus elementos puede interpretarse como el promedio de la serie temporal de correlacion condicional (cambiante en el tiempo) que pudiesemos estimar para dicho par de activos en la muestra utilizada mediante un procedimiento que permitiese tal variabilidad temporal.

7.3 Modelo de suavizado exponencial (*Exponential smoother, EWMA, Integrated DCC*)

Suponemos ahora que la evolución dinámica de la correlación está guiada por las variables auxiliares $q_{ij,t+1}$, que juegan el papel de covarianzas condicionales, y que se actualizan a partir de valores iniciales $q_{ij,1}$ mediante:

$$q_{ij,t+1} = (1-\lambda)\, z_{i,t} z_{j,t} + \lambda q_{ij,t} \forall i,j$$

El algoritmo recursivo anterior puede inicializarse tomando como valor inicial $q_{ij,1}$ el promedio de los productos $z_{i,t}z_{j,t}$ a lo largo de toda la muestra. La condición inicial para las varianzas condicionales $q_{ii,1}$debe ser su esperanza matemática, que es 1, pues se trata ahora de la varianza condicional de una rentabilidad estandarizada. Si no queremos imponer como condición inicial la media de toda la muestra, también podemos utilizar el promedio de un número inicial de observaciones, 50 por ejemplo, y actualizar $q_{ij,t}$ a partir de la observación siguiente, desechando los primeros 50 datos.

Si tenemos un conjunto de k activos, en notación matricial tendremos,

$$Q_{t+1} = (1-\lambda)\, z_t z_t' + \lambda Q_t$$

donde Q_t es una matriz simétrica kxk. Este modelo requiere la estimación de un único parámetro, λ, con independencia del número de activos considerados.

Por ejemplo, con dos activos:

$$Q_t = \begin{pmatrix} q_{11,t+1} & q_{12,t+1} \\ q_{12,t+1} & q_{22,t+1} \end{pmatrix} = (1-\lambda) \begin{pmatrix} z_{1,t}^2 & z_{1,t}z_{2,t} \\ z_{1,t}z_{2,t} & z_{2,t}^2 \end{pmatrix} + \lambda \begin{pmatrix} q_{11,t} & q_{12,t} \\ q_{12,t} & q_{22,t} \end{pmatrix}$$

o, en notacion $vech$:

$$\begin{pmatrix} q_{11,t+1} \\ q_{12,t+1} \\ q_{22,t+1} \end{pmatrix} = (1-\lambda) \begin{pmatrix} z_{1,t}^2 \\ z_{1,t}z_{2,t} \\ z_{2,t}^2 \end{pmatrix} + \lambda \begin{pmatrix} q_{11,t} \\ q_{12,t} \\ q_{22,t} \end{pmatrix}$$

Al ser una covarianza entre rentabilidades estandarizadas, la serie temporal $q_{ij,t}$ ya nos proporciona una estimación de la correlación condicional entre dos rentabilidades. Pero para garantizar que dicha correlación esté siempre en el intervalo $(-1,1)$, utilizamos la normalización:

$$\rho_{ij,t+1} = \frac{q_{ij,t+1}}{\sqrt{q_{ii,t+1}}\sqrt{q_{jj,t+1}}}.$$

Una vez estimadas las correlaciones condicionales $\rho_{ij,t}$, la matriz de varianzas y covarianzas condicionales se obtiene de: $\Sigma_{t+1} = D_{t+1}\Gamma_{t+1}D_{t+1}$, donde Γ_{t+1} es la mtriz de correlaciones y D_{t+1} es la matriz diagonal de las desviaciones típicas estimadas para $t+1$. Esto equivale a calcular: $\sigma_{ijt} = \rho_{ijt}\sigma_{it}\sigma_{jt}$, mientras que las varianzas condicionales coincidirán con las de los modelos GARCH univariantes inicialmente estimados, al tener la matriz Γ_t unos en su diagonal principal.

7.4 Correlaciones dinámicas GARCH (*DCC GARCH*)

Para permitir reversión a la media en las correlaciones condicionales, podemos utilizar una especificación del tipo GARCH(1,1):

$$q_{ij,t+1} = \rho_{ij} + \alpha\left(z_{i,t}z_{j,t} - \rho_{ij}\right) + \beta\left(q_{ij,t} - \rho_{ij}\right)$$

y nuevamente utilizamos la normalización,

$$\rho_{ij,t+1} = \frac{q_{ij,t+1}}{\sqrt{q_{ii,t+1}}\sqrt{q_{jj,t+1}}}$$

para obtener la estimación final de los coeficientes de correlación condicional. Las condiciones iniciales para las variables $q_{ij,t+1}$ pueden escogerse como en el modelo anterior.

En este modelo estamos imponiendo que los parámetros de persistencia de las correlaciones, α y β sean los mismos para cualquier par de activos. Lógicamente, eso no significa que sean iguales los coeficientes de correlación. Tampoco condiciona las volatilidades condicionales estimadas, que procederán de modelos previamente estimados, posiblemente utilizando un modelo univariante para cada rentabilidad.

Aunque el parámetro ρ_{ij}, que es específico a cada par de activos, puede tratarse como un parámetro más a estimar, junto con α y β, puede tener sentido imponer en el modelo la reversión a un nivel de correlación de largo plazo, $E(z_{i,t}z_{j,t})$, que podemos denotar por $\bar{\rho}_{ij}$,

$$\bar{\rho}_{ij} = E(z_{it}.z_{jt})$$

que estimaríamos:

$$\bar{\rho}_{ij} = \frac{1}{T}\sum_{t=1}^{T} z_{it}z_{jt}$$

teniendo entonces el modelo:

$$q_{ij,t+1} = \bar{\rho}_{ij} + \alpha\left(z_{i,t}z_{j,t} - \bar{\rho}_{ij}\right) + \beta\left(q_{ij,t} - \bar{\rho}_{ij}\right)$$

Para el caso particular de dos activos, teniendo en cuenta que: $E(z_{it}.z_{jt}) = E\left[\begin{pmatrix} z_{1,t}^2 & z_{1,t}z_{2,t} \\ z_{1,t}z_{2,t} & z_{2,t}^2 \end{pmatrix}\right] = \begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix}$ resulta,

$$\begin{pmatrix} q_{11,t+1} & q_{12,t+1} \\ q_{12,t+1} & q_{22,t+1} \end{pmatrix} = \begin{pmatrix} 1 & \bar{\rho}_{12} \\ \bar{\rho}_{12} & 1 \end{pmatrix}(1-\alpha-\beta)+\alpha\begin{pmatrix} z_{1,t}^2 & z_{1,t}z_{2,t} \\ z_{1,t}z_{2,t} & z_{2,t}^2 \end{pmatrix}+\beta\begin{pmatrix} q_{11,t} & q_{12,t} \\ q_{12,t} & q_{22,t} \end{pmatrix}$$

o, en notacion *vech* :

$$\begin{pmatrix} q_{11,t+1} \\ q_{12,t+1} \\ q_{22,t+1} \end{pmatrix} = \begin{pmatrix} 1 \\ \bar{\rho}_{12} \\ 1 \end{pmatrix}(1-\alpha-\beta) + \alpha\begin{pmatrix} z_{1,t}^2 \\ z_{1,t}z_{2,t} \\ z_{2,t}^2 \end{pmatrix} + \beta\begin{pmatrix} q_{11,t} \\ q_{12,t} \\ q_{22,t} \end{pmatrix}$$

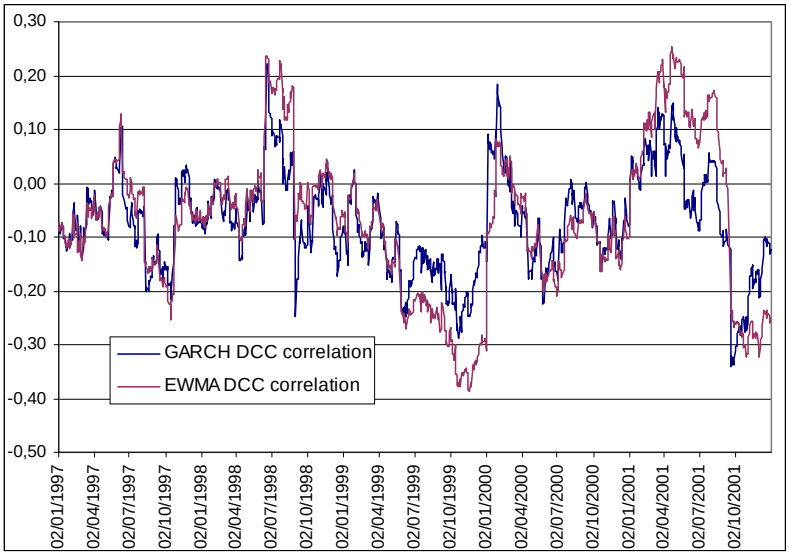
En notación matricial, si estamos calculando las correlaciones condicionales entre varios activos, este modelo es:

$$Q_{t+1} = E(z_t z_t')(1 - \alpha - \beta) + \alpha\,(z_t z_t') + \beta Q_t$$

En ambos casos, la matriz Q_{t+1} es definida positiva por construcción, por lo que también lo serán las matrices de covarianzas Σ_{t+1} y de correlaciones Γ_{t+1}. La matriz de covarianzas se construye como en los modelos anteriores.

Ejercicio: Chapter3 Results.xls, en las pestañas Question 7, Question 8, estima correlaciones EWMA DCC y GARCH DCC entre S&P500 y el tipo de cambio $US/yen. En toras dos pestañas se calculan dichas correlciones para S&P500 y el indice de la bolsa de Toronto:

Correlación condicional: S&P500 y $US/yen



Una ventaja de este modelo es que sus parámetros pueden estimarse en varias etapas: Primero estimamos los parámetros de los modelos de volatilidad condicional univariantes por los procedimientos vistos en secciones previas. A continuación, estandarizamos las rentabilidades y estimamos la matriz de correlaciones incondicionales que, en el caso sencillo de dos activos consta de un sólo parámetro $\bar{\rho}_{ij} = \frac{1}{T}\sum z_{1,t} z_{2,t}$. Finalmente, estimamos los parámetros α y β, que determinan la persistencia en los coeficientes de correlación. Lo importante es que en cada etapa, se estima simultaneamente un número reducido de parámetros.

Nótese que si estimamos las correlaciones de largo plazo, ρ_{ij}, entonces el número de parámetros aumenta linealmente con el número de pares de activos. Si fijamos estos parámetros en los valores que toman las correlaciones muestrales, entonces el número de parámetros es 2, α y β, con independencia del número de activos considerados.

7.5 Estimación por cuasi-máxima verosimilitud

Apelando al procedimiento de cuasi-máxima verosimilitud, tiene sentido trabajar bajo el supuesto de Normalidad teniendo en cuenta que al estar trabajando con rentabilidades estandarizadas, la matriz de covarianzas es una matriz de correlaciones con unos en la diagonal. Recordemos que la función de densidad de una Normal m-variante es: $L = (2\pi)^{-m/2} |\Sigma|^{-1/2} \exp\left[-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right]$, por lo que: $\ln L = -\frac{m}{2}\ln(2\pi) - \frac{1}{2}\ln|\Gamma| - \frac{1}{2}z'\Gamma^{-1}z$.

El logaritmo de la función de verosimilitud es entonces, salvo constantes,

$$\ln L = -\frac{1}{2}\sum_{t=1}^{T}\left[\ln(1-\rho_{12,t}^2) + \frac{z_{1,t}^2 + z_{2,t}^2 - 2\rho_{12,t}z_{1,t}z_{2,t}}{1-\rho_{12,t}^2}\right]$$

en la que la correlación condicional $\rho_{12,t}$ se obtiene a partir del modelo particular de correlación que se utilice, que será distinta en el modelo de alisado exponencial (EWMA-DCC) y en el modelo GARCH-DCC. Como ya dijimos antes, el algoritmo numérico puede inicializarse con $q_{11,0} = q_{22,0} = 1, q_{12,0} = T^{-1}\sum_{t=1}^{T} z_{1,t}z_{2,t}$. Nótese que estamos utilizando en todo momento las rentabilidades estandarizadas, para lo que necesitaremos utilizar varianzas condicionales procedentes de modelos que hayamos estimado previamente. Se trata, por tanto, de una estimación secuencial, que resulta bastante sencilla, aunque a riesgo de perder eficiencia estadística. Pero la estimación simultánea de los parámetros de las varianzas y las correlaciones se puede hacer imposible si disponemos de un número incluso reducido de activos.

En el caso del modelo EWMA-DCC de correlacion estimaremos en la segunda etapa tan solo el parametro λ, mientras que en el modelo GARCH-DCC estimaremos los parámetros α y β. Seria estadísticamente más eficiente estimar simultáneamente el parámetro del modelo EWMA y los parámetros de los modelos de volatilidad condicional, pero el problema es numéricamente más complejo.

En el caso de un vector de n activos, la función a maximizar sería,

$$\ln L = -\frac{1}{2}\sum_{t}\left(\ln \mid \Gamma_t \mid + z_t'\Gamma_t^{-1}z_t\right)$$

siendo Γ_t la matriz de correlaciones condicionales. La función $\ln L$ que presentamos antes para dos variables es claramente un caso particular de esta última cuando tratamos unicamente con dos activos.

8 Relaciones entre contado y futuro

Frecuentemente, se utiliza un modelo DCC para estimar la correlación condicional entre dos o más variables. Para modelizar relaciones entre un activo de contado (por ejemplo, el IBEX35) y un futuro (por ej., el futuro sobre IBEX35), podriamos especificar:

$$\begin{aligned} r_{s,t} &= a_0 + a_{11} r_{s,t-1} + a_{12} r_{f,t-1} + \varepsilon_{s,t} \\ r_{f,t} &= a_1 + a_{21} r_{s,t-1} + a_{22} r_{f,t-1} + \varepsilon_{f,t} \\ (\varepsilon_{s,t}, \varepsilon_{f,t})/\Omega_{t-1} &\sim N(0_2, \Sigma_t) \end{aligned}$$

que es un modelo VAR_1, y estimar las varianzas y las covarianzas de las innovaciones mediante un modelo GARCH. En el caso de especificar un modelo DCC-GARCH seguiriamos el procedimiento que antes expusimos, aplicado a los residuos de la estimacion del modelo VAR de rentabilidades, en lugar de aplicarlo a las propias rentabilidades. De este modo, el procedimiento sería:

1. Estimacion del modelo VAR de rentabilidades, conservando los residuos $(\varepsilon_{s,t}, \varepsilon_{f,t})$. Prreviamente habria que identificar el orden correcto del VAR, ya sea mediante los estadisticos habituales: AIC, BIC, o escogiendo el menor orden que elimine la evidencia de autocorrelación residual,

2. contrastación, si se desea, de causalidad entre contado y futuro,

3. estimación de modelos GARCH univariates, simetricos o asimétricos para cada una de las series de residuos,

4. estandarización de cada serie de residuos dividiendo por la desviación típica condicional estimada en su modelo GARCH univariante,

5. aplicación de la metodología DCC-GARCH a las series tipificadas de residuos.

Como se aprecia, continúa siendo un procedimiento en dos etapas, estimando los parametros del modelo de relación entre contado y futuro por separado de los parámetros del modelo de varianzas-covarianzas DCC-GARCH.

8.0.1 Modelo de corrección de error con estructura DCC GARCH

Volvamos a considerar un problema de cobertura entre un activo financiero y un futuro, con precios S_t, F_t y rentabilidades $r_{s,t}, r_{f,,t}$. Suponiendo la cointegracion entre los precios de ambos activos, el modelo de corrección de error tendria una estructura:

$$\begin{pmatrix} r_{s,t} \\ r_{f,t} \end{pmatrix} = \sum_{i=1}^{n} \begin{pmatrix} \alpha(i)_{11} & \alpha(i)_{12} \\ \alpha(i)_{21} & \alpha(i)_{22} \end{pmatrix} \begin{pmatrix} r_{s,t-i} \\ r_{f,t-i} \end{pmatrix} + \begin{pmatrix} \gamma_s \\ \gamma_f \end{pmatrix} (\ln S_{t-1} - \beta \ln F_{t-1}) + \begin{pmatrix} \varepsilon_{s,t} \\ \varepsilon_{f,t} \end{pmatrix}$$

donde, condicional en Ω_{t-1}, el conjunto de informacion disponible en $t-1$,los términos de innovacion seguirán una determinada distribución, que tendriamos que especificar para la estimación por Máxima Verosimilitud del modelo. Por ejemplo, podriamos suponer, despues de hacer algunos contrastes, que $(\varepsilon_{s,t}, \varepsilon_{f,t})/\Omega_{t-1} \sim D(0_2, \Sigma_t)$, siendo Σ_t la matriz de covarianzas condicional del

proceso bivariante de innovaciones, y D una distribución de probabilidad por especificar.

El modelo se formula en dos etapas. Primero, suponemos que la evolución temporal de las varianzas condicionales viene recogida por un modelo GARCH(p,q), posiblemente asimétrico:

$$\begin{pmatrix} \sigma^2_{s,t} \\ \sigma^2_{f,t} \end{pmatrix} = \begin{pmatrix} \omega_s \\ \omega_f \end{pmatrix} + \sum_{i=1}^{p} \begin{pmatrix} A(i)_{11} & A(i)_{12} \\ A(i)_{21} & A(i)_{22} \end{pmatrix} \begin{pmatrix} \varepsilon^2_{s,t-i} \\ \varepsilon^2_{f,t-i} \end{pmatrix} + \sum_{j=1}^{q} \begin{pmatrix} B(j)_{11} & B(j)_{12} \\ B(j)_{21} & B(j)_{22} \end{pmatrix} \begin{pmatrix} \sigma^2_{s,t-j} \\ \sigma^2_{f,t-j} \end{pmatrix} + \begin{pmatrix} \varepsilon_{s,t} \\ \varepsilon_{f,t} \end{pmatrix} + \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix} \begin{pmatrix} \varepsilon^2_{s,t-1} I_{s,t-1} \\ \varepsilon^2_{f,t-1} I_{f,t-1} \end{pmatrix}, \; con \quad \begin{matrix} I_{k,t-1} = 1 \; si \; \varepsilon_{k,t-1} < 0 \\ I_{k,t-1} = 0 \; si \; \varepsilon_{k,t-1} > 0, \; k = s, f \end{matrix}$$

Por ejemplo, si suponemos, $n = 2, p = 1, q = 1$, tendríamos:

$$\begin{pmatrix} r_{s,t} \\ r_{f,t} \end{pmatrix} = \begin{pmatrix} \alpha(1)_{11} & \alpha(1)_{12} \\ \alpha(1)_{21} & \alpha(1)_{22} \end{pmatrix} \begin{pmatrix} r_{s,t-1} \\ r_{f,t-1} \end{pmatrix} + \begin{pmatrix} \alpha(2)_{11} & \alpha(2)_{12} \\ \alpha(2)_{21} & \alpha(2)_{22} \end{pmatrix} \begin{pmatrix} r_{s,t-2} \\ r_{f,t-2} \end{pmatrix} + + \begin{pmatrix} \gamma_s \\ \gamma_f \end{pmatrix} (\ln S_{t-1} - \beta \ln F_{t-1}) + \begin{pmatrix} \varepsilon_{s,t} \\ \varepsilon_{f,t} \end{pmatrix}$$

$$\begin{pmatrix} \sigma^2_{s,t} \\ \sigma^2_{f,t} \end{pmatrix} = \begin{pmatrix} \omega_s \\ \omega_f \end{pmatrix} + \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} \varepsilon^2_{s,t-1} \\ \varepsilon^2_{f,t-1} \end{pmatrix} + \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} \sigma^2_{s,t-1} \\ \sigma^2_{f,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{s,t} \\ \varepsilon_{f,t} \end{pmatrix} + \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix} \begin{pmatrix} \varepsilon^2_{s,t-1} I_{s,t-1} \\ \varepsilon^2_{f,t-1} I_{f,t-1} \end{pmatrix}, \; con \quad \begin{matrix} I_{k,t-1} = 1 \; si \; \varepsilon_{k,t-1} < 0 \\ I_{k,t-1} = 0 \; si \; \varepsilon_{k,t-1} > 0, \; k = s, f \end{matrix}$$

Una vez que se ha estimado este modelo, se generan las rentabilidades estandarizadas:

$$\begin{aligned} z_{s,t} &= \varepsilon_{s,t}/\sigma_{s,t} \\ z_{f,t} &= \varepsilon_{f,t}/\sigma_{f,t} \end{aligned}$$

y las variables auxiliares $q_{ss,t}, q_{ff,t}, q_{sf,t}$:

$$\begin{aligned} q_{ss,t} &= (1 - \kappa_1 - \kappa_2)\bar{q}_{ss} + \kappa_1 q_{ss,t-1} + \kappa_2 z^2_{s,t-1}, \; \bar{q}_{ss} = T^{-1} \sum_{t=1}^{T} z^2_{s,t} \\ q_{ff,t} &= (1 - \kappa_1 - \kappa_2)\bar{q}_{ff} + \kappa_1 q_{ff,t-1} + \kappa_2 z^2_{f,t-1}, \; \bar{q}_{ff} = T^{-1} \sum_{t=1}^{T} z^2_{f,t} \\ q_{ss,t} &= (1 - \kappa_1 - \kappa_2)\bar{q}_{sf} + \kappa_1 q_{sf,t-1} + \kappa_2 z_{s,t-1} z_{f,t-1}, \; \bar{q}_{sf} = T^{-1} \sum_{t=1}^{T} z_{s,t} z_{f,t} \end{aligned}$$

comenzando a partir de condiciones iniciales:

$$
\begin{aligned}
q_{ss,1} &= 1, \\
q_{ff,1} &= 1, \\
q_{sf,1} &= T^{-1} \sum_{t=1}^{T} z_{s,t} z_{f,t}
\end{aligned}
$$

y estimando el modelo por Máxima Verosimilitud.
La correlación condicional se estima:

$$\rho_{sf,t} = \frac{q_{sf,t}}{\sqrt{q_{ss,t} q_{ff,t}}}$$

y la covarianza condicional:

$$\sigma_{sf,t} = \rho_{sf,t} \sigma_{s,t} \sigma_{f,t}$$

y si suponemos Normalidad de la distribución condicional de rentabilidades,

$$\begin{pmatrix} z_{s,t} \\ z_{f,t} \end{pmatrix} / \Omega_{t-1} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, R \equiv \begin{pmatrix} 1 & \rho_{sf,t} \\ \rho_{sf,t} & 1 \end{pmatrix} \right)$$

por lo que la función logaritmo de la verosimilitud resulta:

$$
\begin{aligned}
\ln L &= -\frac{T}{2} \ln 2\pi - \frac{1}{2} \sum_{t=1}^{T} \left[\ln(|R|) + (z_{s,t} \; z_{f,t}) \, R^{-1} \begin{pmatrix} z_{s,t} \\ z_{f,t} \end{pmatrix} \right] = \\
&= -\frac{T}{2} \ln 2\pi - \frac{1}{2} \sum_{t=1}^{T} \left[\ln(1 - \rho_{sf,t}^2) + \frac{z_{1,t}^2 + z_{2,t}^2 - 2\rho_{sf,t} z_{1,t} z_{2,t}}{1 - \rho_{sf,t}^2} \right]
\end{aligned}
$$

Esta funcion de verosimilitud deberia maximizarse respecto de los valores numéricos de los parámetros:

$$
\begin{aligned}
&\alpha(1)_{11}, \alpha(1)_{12}, \alpha(1)_{21}, \alpha(1)_{22}, \alpha(2)_{11}, \alpha(2)_{12}, \alpha(2)_{21}, \\
&\alpha(2)_{22}, \gamma_s, \gamma_f, A_{11}, A_{12}, A_{21}, A_{22}, B_{12}, B_{12}, B_{21}, B_{22}, D_{11}, \\
&D_{12}, D_{21}, D_{22}, \kappa_1, \kappa_2,
\end{aligned}
$$

lo que ilustra claramente la dificultad del problema numérico de estimación. El procedimiento en dos etapas simplifica la dificultad, puesto que los valores numericos estimados para los parámetros del modelo de corrección del error y los del modelo GARCH bivariate se toman como dados cuando se estiman los parámetros κ_1, κ_2 de la especificación DCC. De hecho, aun a riesgo de perder alguna eficiencia estadística, el procedimiento de estimación suele instrumentarse en tres etapas: 1) estimación del modelo de corrección del error y generación de sus residuos, 2) estimación del modelo GARCH para dichos residuos y generación de las rentabilidades estandarizadas, 3) estimacion de los parámetros $\kappa_1 \kappa_2$ maximizando la función de verosimilitud que aparece arriba respecto de estos dos parámetros.

8.0.2 Cobertura de carteras: una aplicación del modelo DCC-GARCH

Un ejemplo habitual de aplicación de este tipo de modelos es el cálculo de ratios de cobertura. Denotemos por $\sigma_{sf,t}$ la covarianza entre la cartera de contado (s) y el futuro (f) que se va a utilizar en la cobertura, y por $\sigma^2_{f,t}$ la varianza de la rentabilidad de dicho contrato de futuro. Por razones de estacionariedad, en ambos casos estaríamos utilizando la rentabilidad de dichos activos, no sus precios.

Consideremos el problema de minimizar la varianza de la cartera cubierta en dos periodos. Si tenemos una posición de b contratos de contado y cubrimos poniéndonos cortos en h contratos de futuros, la rentabilidad de la posición cubierta será: $br_{ct} - hr_{ft}$. (En general, puede considerarse $b = 1$). El problema será:

$$\min_h Var(br_{ct} - hr_{ft})$$

Como:

$$Var(br_{ct} - hr_{ft}) = b^2 Var(r_{ct}) + h^2 Var(r_{ft}) - 2bhCov(r_{ct}, r_{ft})$$

derivando dos veces respecto a h :

$$2hVar(r_{ft}) - 2bCov(r_{ct}, r_{ft})$$
$$2Var(r_{ft})$$

donde vemos que la segunda derivada es positiva. Igualando la primera derivada a cero y resolviendo, tenemos la expresion para el ratio de cobertura de mínima varianza:

$$\frac{h_t}{b_t} = \frac{Cov(r_{ct}, r_{ft})}{Var(r_{ft})}$$

que coincide con el estimador de mínimos cuadrados de la pendiente de una regresión de r_{ct} sobre r_{ft}.

Estos momentos pueden estimarse con carácter incondicional, mediante técnicas de minimos cuadrados ordinarios, o puede tratarse de momentos condicionales, en cuya estimación utilizaríamos un modelo GARCH. de hecho, si consideramos el problema:

$$\min_h Var_t(b_t r_{c,t+1} - h_t r_{f,t+1})$$

llegaremos a:

$$\frac{h_t}{b_t} = \frac{Cov_t(r_{c,t+1}, r_{f,t+1})}{Var(r_{f,t+1})}$$

En un modelo tipo GARCH, recordemos que:

$$\sigma^2_{f,t+1} = \omega + \beta\sigma^2_{f,t} + \alpha\varepsilon^2_{f,t}$$

por lo que, una vez estimado el modelo con datos hasta T, $\sigma^2_{f,t+1}$ es conocido en T (en realidad no es conocido porque no hemos utilizado los verdaderos valores de los parametros, sino sus estimaciones).

Si los momentos se estiman con toda la muestra decimos que se trata de estimaciones *ex-post*, ya que para estimar la varianza de un dia $t, t < T$, utilizaríamos información que no estaba disponible en dicho dia, sino que fue conocida en un momento futuro. Esto afecta tanto a la estimación MCO, con la que obtenemos una única estimación numerica de los momentos $\sigma_{cf,t}, \sigma^2_{f,t}$, y del ratio de cobertura b_t, como a la estimación mediante un modelo GARCH asimismo con toda la muestra, si bien en este ultimo caso obtendremos varianzas y covarianzas condicionales diferentes para cada dia de la muestra. En ambas estimaciones estariamos utilizando informacion no disponible en cada momento, por lo que este tipo de estimaciones no habrian podido obtenerse en tiempo real, segun se observa la evolución de los mercados.

Las estimaciones *ex-ante* se obtienen utilizando para cada dia t informacion estadistica correspondiente a dias no posteriores a t. Para ello, podemos ir aumentando la muestra cada dia, o podemos utilizar ventanas moviles, manteniendo su amplitud constante. Esto puede hacerse tanto con la estimacion de minimos cuadrados como con la estimacion GARCH.

La diferencia entre uno y otro tipo de momentos estriba en que los momentos incondicionales se consideran constantes en el tiempo, y su estimacion va mejorando segun se dispone de mas informacion muestral. Por tanto, bajo esta interpretacion, tiene sentido estimar los momentos correspondientes al dia t con todos los datos disponibles hasta dicho dia, no con ventanas moviles de amplitud constante. En ese sentido, cabria esperar que la estimacion numerica variase con el tamaño muestral, auque convergiendo al verdadero valor del parametro que se estima (si el estimador es consistente). Si esto no sucede, habrá que pensar que ha habido algun cambio estructural en el valor del parametro durante la muestra utilizada en su estimación, o que la naturaleza incondicional del momento poblacional no es correcta. Los momentos condicionales, por definición, cambian con el tiempo, por lo que cada dia se estiman momentos diferentes, y tiene sentido mantener constante la amplitud de la ventana muestral utilizada en la estimación.

Lafuente y Novales (2003) consideran el problema:

$$\min_h Var(b_t r_{ct} - h_t r_{ft})$$

sujeto a:

$$\begin{aligned} dr_{c,t} &= \mu_{c,t}dt + \sigma_{c,t}dW_{c,t} \\ dr_{f,t} &= \mu_{f,t}dt + \sigma_{c,t}dW_{c,t} + \sigma_{s,t}dW_{s,t} \end{aligned}$$

donde las rentabilidades de la cartera de contado y del contrato de futuro que se va a utilizr en la cobertura, comparten un browniano, $dW_{c,t}$. Pero en la determinación de la rentabilidad del futuro aparece un segundo browniano,

$dW_{s,t}$. Una manera de interpretar este segundo browniano es como *riesgo de base*. Su importancia se medirá por el cociente de volatilidades σ_s/σ_f. En este modelo, si denotamos por $\rho_{cs,t}$ la correlación entre ambos brownianos, entonces la correlación entre las rentabilidades del contado y del futuro es:

$$\rho_{sf,t} = \frac{\sigma_{c,t}^2 + \rho_{cs,t}\sigma_{c,t}\sigma_{s,t}}{\sqrt{\sigma_{c,t}^2\left(\sigma_{c,t}^2 + \sigma_{s,t}^2 + 2\rho_{cs,t}\sigma_{c,t}\sigma_{s,t}\right)}}$$

y el ratio de cobertura de mínima varianza puede expresarse:

$$\frac{h_t}{b_t} = \frac{\sigma_{c,t}^2 + \rho_{cs,t}\sigma_{c,t}\sigma_{s,t}}{\sigma_{c,t}^2 + \sigma_{s,t}^2 + 2\rho_{cs,t}\sigma_{c,t}\sigma_{s,t}} = \frac{1 + \rho_{cs,t}\delta_t}{1 + \delta_t^2 + 2\rho_{cs,t}\delta_t}$$

donde $\delta_t = \sigma_{s,t}/\sigma_{c,t}$.

El ratio de minima varianza solo queda sin determinar cuando $\rho_{cs,t} = -1$ y $\delta_t = 1$. En este caso particular, la función objetivo es igual a la varianza de la posición descubierta, $b_t^2\sigma_{c,t}^2$, con independencia del ratio de cobertura que se utilice. En este caso limite, la rentabilidad del contrato futuro es no estocástica, y no proporciona cobertura alguna.

En el caso general, la varianza condicional mínima de la posición cubierta es:

$$b_t^2\sigma_{c,t}^2\delta_t^2\frac{1-\rho_{cs,t}^2}{1+\delta_t^2+2\rho_{cs,t}\delta_t}$$

siendo $\delta_t^2\frac{1-\rho_{cs,t}^2}{1+\delta_t^2+2\rho_{cs,t}\delta_t}$ el factor de reducción de varianza conseguido con la cobertura.

Proposition 18 *Si $\delta_t \to 0$, entonces el ratio de cobertura de minima varianza converge a 1 y la varianza minima de la posición cubierta converge a cero. Ambos limites se cumplen con independencia de la correlación entre las innovaciones común y específica.*

Proposition 19 *Si $\delta_t \to \infty$, entonces el ratio de cobertura de minima varianza converge a 0 con independencia de la correlación entre las innovaciones común y específica.*

Proposition 20 *Bajo correlación positiva o nula entre ambas innovaciones, el ratio de cobertura de minima varianza es positivo, menor que 1, y decreciente con δ_t.*

Esta última proposición muestra que el ratio de cobertura es menor cuanto mayor sea el riesgo de base.

Proposition 21 *Si $\delta_t < 1$ ($\delta_t > 1$), el ratio de cobertura de minima varianza es monotonicamente creciente (decreciente) en la correlación entre ambas innovaciones.*

Proposition 22 *Si las innovaciones $W_{c,t}, W_{st}$ están correlacionadas positiva y perfectamente, el ratio de cobertura óptimo es: $h_t/b_t = 1/(1+\delta_t)$ y conseguimos una cobertura perfecta. Si la corrrelación es perfecta, pero negativa, entonces el ratio óptimo es $h_t/b_t = 1/(1-\delta_t)$ y alcanzamos de nuevo una cobertura completa.*

Proposition 23 *El ratio de cobertura de minima varianza es negativo si y solo si $\delta_t > 1$ y $-(1+\delta_t^2)/2\delta_t < \rho_{cs,t} < -1/\delta_t$.*

8.0.3 Modelo asimétrico

El modelo DCC GARCH admite una versión asimétrica que puede formularse de distintas maneras. Capiello, Engle y Sheppard (Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns, Journal of Financial Econometrics, 2006, vol.4, no.4, 537-572) proponen el modelo AG-DCC (Asymmetric Generalized DCC):

$$Q_{t+1} = \left(\bar{P} - A'\bar{P}A - B'\bar{P}B - G'\bar{N}G\right) + A'z_t z_t'A + G'\eta_t\eta_t'G + B'Q_tB$$

where $\eta_t = 1_{\varepsilon_t<0}.\varepsilon_t, \bar{P} = E(z_t z_t'), \bar{N} = E(\eta_t\eta_t')$, matrices que se estiman mediante sus análogos muestrales, teniendo en cuenta que en la estimación de $\bar{N}$ hay que promediar sobre toda la muestra, no unicamente sobre los períodos en que $\varepsilon_t < 0$.

En la expresión anterior A, B, G son matrices kxk, siendo k el número de activos. Si sustituimos estas matrices por escalares, tenemos el modelo A-DCC (Asymmetric DCC):

$$Q_{t+1} = \left[(1-\alpha-\beta)\bar{P} - \gamma\bar{N}\right] + \alpha z_t z_t' + \gamma\eta_t\eta_t' + \beta Q_t$$

Así, por ejemplo:

$$\begin{aligned} q_{ii,t+1} &= [(1-\alpha-\beta)Var(\varepsilon_{it}) - \gamma Var\left(1_{\varepsilon_{it}<0}.\varepsilon_{it}\right)] + \alpha\varepsilon_{it}^2 + \gamma\left(1_{\varepsilon_{it}<0}.\varepsilon_{it}^2\right) + \beta q_{ii,t} \\ q_{ij,t+1} &= \left[(1-\alpha-\beta)Cov(\varepsilon_{it},\varepsilon_{jt}) - \gamma Cov\left[\left(1_{\varepsilon_{it}<0}.\varepsilon_{it}\right), \left(1_{\varepsilon_{jt}<0}.\varepsilon_{jt}\right)\right]\right] + \alpha\varepsilon_{it}\varepsilon_{jt} + \\ &\quad + \gamma\left[\left(1_{\varepsilon_{it}<0}.\varepsilon_{it}\right)\left(1_{\varepsilon_{jt}<0}.\varepsilon_{jt}\right)\right] + \beta q_{ij,t} \end{aligned}$$

En el articulo se presenta asimismo una versión diagonal del modelo AG-DCC y se dan las condiciones para que las matrices de covarianzas sean definidas positivas en cada uno de estos modelos.

Es muy sencillo generalizar cualquiera de estos modelos para permitir algun cambio estructural, ya sea en las rentabilidades medias o en las relaciones dinamicas entre rentabilidades, pudiendo contrastar tras la estimación la validez de tal hipótesis mediante un test de Wald habitual.

9 Selección de carteras y evaluación de su gestión

Un inversor quiere obtener una buena rentabilidad sin asumir un elevado nivel de riesgo. El fundamento básico de la teoría financiera es la relación inversa entre ambas características, rentabilidad y riesgo, de tal modo que las carteras que incorporan un mayor nivel de riesgo, o más agresivas, tienden a generar, en media, una mayor rentabilidad, mientras que las carteras más conservadoras tienden a proporcionar una rentabilidad menor. Por eso es que tanto unas carteras como otras pueden resultar atractivas para distintos inversores. Junto con esta relación inversa entre rentabilidad y riesgo, que siempre puede contrastarse empíricamente, la teoría financiera más básica considera que el inversor toma sus decisiones de cartera maximizando su función de utilidad, que depende de dos momentos de la distribución de rentabilidades: media y varianza.

Esta idea se justifica suponiendo que la utilidad del inversor depende positivamente de la rentabilidad de su inversión y depende inversamente de su volatilidad. El inversor supuestamente maximiza tal función de utilidad sujeta a las restricciones de recursos de que disponga obteniendo como solución su cartera óptima. Dicha cartera vendrá definida en términos de las proporciones de la inversión que debe asignar a cada uno de los activos que considera en dicho problema.

Es claro que distintos inversores, que difieran en su función de utilidad, tanto por la importancia que asignan a la rentabilidad de una cartera como a su volatilidad, tendrán una cartera óptima diferente. Por supuesto que si el rango de activos en que consideran invertir es diferente, ello también hará que sus carteras óptimas difieran. Es importante señalar asimismo que la función de utilidad descrita debe depender de la rentabilidad esperada y del nivel de riesgo, que se supone adecuadamente medido por la volatilidad. Pero como sabemos, la volatilidad no es observable y existen diversos modelos para evaluarla. Cada uno de ellos conducirá a una solución diferente al problema de maximización de utilidad y, por tanto, a una cartera óptima distinta. Por eso es tan importante conocer cual puede ser el modelo más apropiado de volatilidad en cada contexto. Por último, como se ha de considerar la rentabilidad esperada y el riesgo (o volatilidad, en su caso) previstos sobre el horizonte de inversión, dos inversores con idénticas preferencias pero con diferentes horizontes de inversión tendrán generalmente carteras diferentes.

Dados n activos, y denotando por w al vector de ponderaciones, por r al vector de rentabilidades, y por V a la matriz simétrica de varianzas y covarianzas de dichas rentabilidades, tendremos una rentabilidad esperada sobre el horizonte de inversión igual a $w'E(r)$, y una varianza igual a: $w\prime Vw$. De acuerdo con estas ideas que acabamos de exponer, la cartera óptima para un inversor con aversión al riesgo γ se obtiene resolviendo el problema:

$$\max_{w} \left(w'E(r) - \frac{1}{2}\gamma w'Vw \right) \tag{4}$$

Como podemos ver, la aversión al riesgo mide el descenso que se produce

en el nivel de utilidad del inversor si la varianza de su cartera aumenta en una unidad. Este método de selección de carteras se conoce como *criterio Media-Varianza* y es muy habitual en la literatura financiera. Vamos a ver en esta sección una justificación de dicho criterio, es decir, un conjunto de condiciones bajo las cuales, seleccionar carteras de este método estaría justificado. Por el contrario, cuando las condiciones no se cumplan, este criterio de selección de carteras no estaría plenamente justificado.

9.1 Aversion al riesgo

La *tolerancia absoluta al riesgo* es la máxima cantidad que estaría dispuesto a invertir en un juego que dobla o reduce a la mitad la cantidad invertida. Si el inversor declara estar dispuesto a invertir hasta 400.000 euros para tener una probabilidad de 1/2 de recibir 200.000 euros o, con igual probabilidad, recibir 800.000 euros, entonces 400.000 euros sería su tolerancia absoluta al riesgo. El *coeficiente de aversión absoluta al riesgo* es el cociente entre su riqueza total y dicha cantidad. Si su riqueza es de 1.000.000 euros, el coeficiente de aversion absoluta al riesgo sería: $1.000.000/400.000 = 2,5$

La *tolerancia relativa al riesgo* es la respuesta a la pregunta ¿cuál es la máxima proporción de su riqueza que estaría dispuesto a invertir en una loteria que dobla o reduce a la mitad dicha proporción? Por ejemplo, si dicha proporción es el 20% de su riqueza, entonces la tolerancia relativa al riesgo es 0,20 y el *coeficiente de aversión relativa al riesgo* es su inverso: $0,20^{-1} = 5$.

El grado de aversión al riesgo está relacionado con la concavidad de la función de utilidad. Dado un nivel de riqueza W, el *coeficiente de aversión absoluta al riesgo* se define:

$$A(W) = \frac{-U''(W)}{U'(W)}$$

y el *coeficiente de aversión relativa al riesgo*:

$$R(W) = \frac{-WU''(W)}{U'(W)}$$

En general, un inversor con $A'(W) > 0$ reducirá la cantidad invertida en activos con riesgo al aumentar su riqueza. Si además tiene $R'(W) > 0$, entonces, cuando aumente su riqueza, reducirá no sólo la cuantía invertida en activos con riesgo sino también la proporción debn su riqueza invertida en tales activos.

Por ejemplo, para una función logaritmica: $U(W) = \ln(W)$, se tiene: $A(W) = W^{-1}$, que decrece con el nivel de riqueza. De modo que el valor absoluto de la inversión en activos con riesgo aumentará con el nivel de riqueza, puesto que la aversión absoluta al riesgo disminuye. La aversión relativa al riesgo para preferencias logaritmicas es; $R(W) = 1$, constante. Por tanto, para tal inversor, la proporción de la riqueza invertida en activos con riesgo será independiente del nivel de riqueza.

Si un inversor tiene una función de utilidad exponencial sobre su riqueza W,

$$U(W) = -e^{-\gamma W}, \gamma > 0$$

tenemos:

$$U'(W) = \gamma e^{-\gamma W},\ U''(W) = -\gamma^2 e^{-\gamma W}$$

por lo que sus *coeficientes de aversión absoluta y relativa al riesgo* son:

$$A(W) = \gamma;\ R(W) = \gamma W$$

de modo que con estas preferencias, el coeficiente de aversión *absoluta* al riesgo es γ , constante. Esta propiedad caracteriza este tipo de preferencias, es decir, la funcion de utilidad exponencial es apropiada para un inversor que quiere mantener en activos con riesgo siempre la misma cantidad, a pesar de que su renta aumente o disminuya.

Bajo preferencias logaritmicas, el coeficiente de aversión al riesgo de un inversor se expresa comparando su tolerancia absoluta al riesgo con la cuantía total de la inversión.

Ejemplo: Consideremos un inversor con función de utilidad exponencial $U(W) = -e^{-\gamma W}, \gamma > 0$, y una tolerancia absoluta al riesgo de 250.000 euros. Si este tiene un patrimonio de 1 millón de euros, su coeficiente de aversión al riesgo sería: $\gamma = \frac{1.000.000}{250.000} = 4$.

9.2 Selección de carteras

9.2.1 Maximización del Equivalente Cierto

El Equivalente Cierto (EC) de una lotería o juego (podríamos decir también de una inversión), es la cantidad de dinero que, otorgada con seguridad, dejaría indiferente al jugador entre percibirla o tomar parte en la lotería. Ante un resultado incierto representado por una variable aleatoria P, el EC es la cantidad x que satisface: $U(x) = E(U(P))$; entonces denotamos: $x = EC(P)$, por lo que: $U\left[EC(P)\right] = E\left(U(P)\right)$.

Si efectuamos un desarrollo en serie de Taylor de $U(P)$ alrededor de la rentabilidad media $\mu_P = E(P)$, tenemos:

$$U(P) \approx U(\mu_P) + U'(\mu_P)(P - \mu_P) + \frac{1}{2}U''(\mu_P)(P - \mu_P)^2$$

y tomando esperanza matemática:

$$E\left[U(P)\right] \approx U(\mu_P) + \frac{1}{2}U''(\mu_P)\sigma_P^2$$

donde $\sigma_P^2 = E\left[(P - \mu_P)^2\right]$ .Esta expresión indica que para un inversor averso al riesgo, $U'' < 0$, se tendrá $E\left[U(P)\right] < U(\mu_P)$, por lo que el Equivalente cierto de una cartera será inferior a su rentabilidad esperada: $EC(P) < \mu_P$.

Consideremos ahora una expansión de Taylor de orden superior para la función de utilidad del inversor, que suponemos dependiente de la *rentabilidad* de la cartera, y tomamos esperanzas:

$$E\left[U(P)\right] \approx U(\mu)+U'(\mu).E(P-\mu)+\frac{1}{2}U''(\mu).E\left[(P-\mu)^2\right]+\frac{1}{6}U'''(\mu).E\left[(P-\mu)^3\right]+$$
$$+\frac{1}{24}U''''(\mu).E\left[(P-\mu)^4\right]+...$$

donde $\mu = E(P)$.Recordando la definición de Equivalente Cierto: $U(EC) = E\left[U(P)\right]$ tendríamos, bajo una utilidad exponencial:

$$\begin{aligned} U(EC) &= -\exp(-\gamma EC) = E\left[U(P)\right] \approx \\ &\approx -\exp(-\gamma\mu)\left(1+\frac{1}{2}\gamma^2 E\left[(P-\mu)^2\right]-\frac{1}{6}\gamma^3 E\left[(P-\mu)^3\right]+\frac{1}{24}\gamma^4 E\left[(P-\mu)^4\right]\right) = \\ &= -\exp(-\gamma\mu)\left(1+\frac{1}{2}(\gamma\sigma)^2-\frac{\tau}{6}(\gamma\sigma)^3+\frac{\kappa}{24}(\gamma\sigma)^4\right) \end{aligned}$$

donde τ, κ denotan los coeficientes de asimetria y curtosis: $\tau = \sigma^{-3}E\left[(R-\mu)^3\right]$, $\kappa = \sigma^{-4}E\left[(R-\mu)^4\right]$. Si tomamos logaritmos y utilizamos la aproximación: $\ln(1+x) = x-\frac{1}{2}x^2$, tenemos:

$$\begin{aligned} -\gamma EC &= -\gamma\mu+\ln\left(1+\frac{1}{2}(\gamma\sigma)^2-\frac{\tau}{6}(\gamma\sigma)^3+\frac{\kappa}{24}(\gamma\sigma)^4\right) \approx \\ &\approx -\gamma\mu+\left(\frac{1}{2}(\gamma\sigma)^2-\frac{\tau}{6}(\gamma\sigma)^3+\frac{\kappa}{24}(\gamma\sigma)^4\right)-\frac{1}{2}\left(\frac{1}{2}(\gamma\sigma)^2-\frac{\tau}{6}(\gamma\sigma)^3+\frac{\kappa}{24}(\gamma\sigma)^4\right)^2 \end{aligned}$$

y, despreciando términos de grado superior a σ^4 :

$$-\gamma EC \approx -\gamma\mu+\left(\frac{1}{2}(\gamma\sigma)^2-\frac{\tau}{6}(\gamma\sigma)^3+\frac{\kappa}{24}(\gamma\sigma)^4\right)-\frac{1}{8}(\gamma\sigma)^4 = -\gamma\mu+\left(\frac{1}{2}(\gamma\sigma)^2-\frac{\tau}{6}(\gamma\sigma)^3+\frac{\kappa-3}{24}(\gamma\sigma)^4\right)$$

llegamos finalmente a:

$$EC \approx \mu-\frac{1}{2}\gamma\sigma^2+\frac{\tau}{6}\gamma^2\sigma^3-\frac{\kappa-3}{24}\gamma^3\sigma^4$$

Como vemos en esta expresión, cuando un inversor tiene una utilidad exponencial, el Equivalente Cierto que aceptaría a cambio de liquidar su cartera invertida bajo incertidumbre[13] sería menor cuanto mayor sea la varianza de la distribución de probabilidad de rentaibilidades, así como también con una asimetría negativa mayor y con un mayor exceso de curtosis.

La media, asimetria y curtosis, μ, τ, κ de una distribución de rentabilidades se refieren al caso en que se invierte una unidad monetaria en la cartera correspondiente. Cuando se inviertan q unidades, el Equivalente Cierto es:

$$EC(q) \approx q\mu-\frac{1}{2}\gamma\sigma^2q^2+\frac{\tau}{6}\gamma^2\sigma^3q^3-\frac{\kappa-3}{24}\gamma^3\sigma^4q^4$$

[13] En general, a cambio de no participar en una determinada lotería.

9.2.2 Criterio de Media-Varianza

Consideremos un inversor con función de utilidad exponencial sobre su riqueza W,

$$U(W) = -e^{-\gamma W}, \gamma > 0$$

Supongamos una riqueza inicial igual a cero, por lo que riqueza y rentabilidad al final del periodo de inversión coinciden. Entonces, de acuerdo con la expresión anterior, si la rentabilidad de la inversión sigue una distribución Normal con esperanza matemática μ y varianza σ^2, tendremos: $\tau = 0, \kappa = 3$, y el Equivalente Cierto de la cartera será aproximadamente igual a:

$$EC \simeq \mu - \frac{1}{2}\gamma\sigma^2$$

Por tanto, para un inversor con una función de utilidad exponencial, que invierte en activos con rentabilidades con distribución Normal, el problema de selección de cartera que maximiza el Equivalente Cierto es:

$$\underset{w}{Max} \left\{ w'E(r) - \frac{1}{2}\gamma w'Vw \right\} \tag{5}$$

donde r es el vector de rentabilidades de los activos disponibles, y V su matriz de covarianzas, que no es sino la maximizacion del criterio Media-Varianza.

Proposition 24 *Suponiendo que todas las alternativas de inversión disponibles para un inversor con función de utilidad exponencial ofrecen una distribución de rentabilidad Normal, el criterio Media-Varianza consiste en escoger la cartera de inversión que proporciona una distribución de rentabilidades con un mayor Equivalente Cierto.*

Bajo otros supuesto, el resultado anterior no se cumplirá. Por ejemplo, si un inversor tiene aversión *relativa* al riesgo constante, entonces una función de utilidad exponencial no será apropiada, y no existirá una relación tan sencilla entre la utilidad esperada de una cartera y sus características de Media-Varianza.

Exercise 25 *(Ex.I.6.3) Supongamos que un inversor con función de utilidad exponencial* $U(W) = -e^{-\gamma W}, \gamma > 0$, *y una tolerancia absoluta al riesgo de 250.000 euros tiene a posibilidad de invertir en un activo que ofrece una rentabilidad con distribucion Normal, esperanza matemática de 10% y volatilidad de 20%, y en un segundo activo con rentabilidad asimismo Normal, con esperanza de 15% y volatilidad de 30%. La correlacion entre ambos activos es de -0,50. La maximización* (4) *conduce a invertir 565.789 euros en el primer activo y 434.211 euros en el segundo activo.*

Exercise 26 *(Ex.I.6.4): Considere un inversor con una función de utilidad exponencial, que puede invertir en dos fondos. El fondo A tiene rentabilidad esperada 10%, con desviación típica 12%, asimetría -0,50 y curtosis 2,5. El*

fondo B tiene rentabilidad esperada 15%, con desviación típica 20%, asimetría -0,75 y curtosis 1,5. Considere un inversor con función de utilidad exponencial, que dispone de 1.000.000 de euros para invertir. ¿En qué fondo invertiría a) 1.000.000 de euros si su tolerancia absoluta al riesgo es 200.000 euros? b) 1.000.000 de euros si su tolerancia absoluta al riesgo es 400.000 euros? c) 1.000.000 de euros si su tolerancia absoluta al riesgo es 100.000 euros? d) 500.000 de euros si su tolerancia absoluta al riesgo es 100.000 euros?. Interprete los resultados que obtiene. R: a) Fondo A, b) Fondo B, c) No deberia invertir en ninguno de los dos fondos, d) Fondo A.

9.2.3 Cartera de mínima varianza

Otro criterio que podría seguirse para seleccionar una cartera con n activos, sería el de formar aquella cartera que tuviese una menor varianza. Es, sin duda, un criterio conservador, que puede ser muy útil para muchos inversores y que puede ser aproximadamente el comportamiento seguido por muchos fondos. Resolvemos entonces el problema:[14]

$$\begin{aligned} &\min_{w} w'Vw \\ \text{sujeto a} \quad &: \quad \sum_{i=1}^{n} w_i = 1, w_i \geq 0 \end{aligned}$$

cuya solución es: $\tilde{w_i} = \frac{\psi_i}{\sum_{j=1}^{n} \psi_j}$, siendo ψ_i la suma de todos los elementos de la columna j de la matriz V^{-1}. La cartera resultante es la cartera de mínima varianza, y dicha varianza es: $\left(\sum_{j=1}^{n} \psi_j \right)^{-1}$.

Si se resuelve directamente el problema de minimizar la varianza de la cartera, en el caso de 2 activos, se obtiene:

$$w^* = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \tag{6}$$

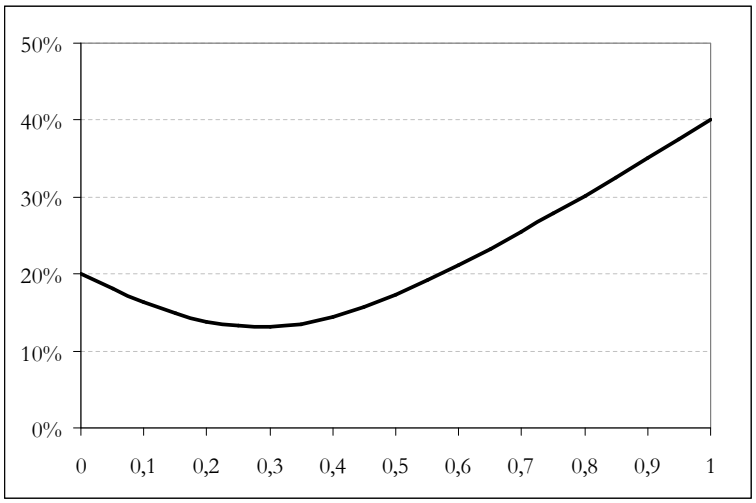
con el significado habitual de los parámetros, siendo w^* el porcentaje invertido en el primer activo. Esta expresión es un caso particular de la anterior. Cuando ambos activos tienen igual varianza, la minima varianza se logra invirtiendo en ambos activos a partes iguales, no importa cuál sea la corrrelación entre ambos. El denominador es siempre positivo, ya que: $\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2 = (\sigma_1 - \rho\sigma_2)^2 + (1-\rho^2)\sigma_2^2$, pero el numerador es positivo solo si $\rho < \sigma_2/\sigma_1$. Solo

[14] Una solución trivial al problema si no imponemos la restricción $\sum_{i=1}^{n} w_i = 1$, sería fijar todas las ponderaciones iguales a cero, lo que generaría una varianza nula, alcanzando su menor valor posible. Desde el punto de vista de la teoría de inversión, esta solución es claramente poco interesante.

entonces es $w^* > 0$. Como además, requerimos que $w^* < 1$, para esto debe cumplirse que $\rho < \sigma_1/\sigma_2$. Por tanto, $0 < w^* < 1$ si y sólo si: $\rho < \min\left(\frac{\sigma_1}{\sigma_2}, \frac{\sigma_2}{\sigma_1}\right)$. Si permitimos posiciones cortas, entonces no es preciso que $\rho < \sigma_2/\sigma_1$.

Ejercicio: (Ex.I.6.5) Tenemos 10.000 euros para invertir en dos activos. La rentabilidad del primero tiene volatilidad 20%, y la del segundo tiene volatilidad 30%. La correlación entre ambas rentabilidades es -0,25. ¿Cual es la cartera de minima varianza y cual seria su volatilidad? Resuelva el ejercicio minimizando directamente la varianza de la cartera, asi como aplicando la regla (6), y compruebe que obtiene los mismos resultados. ¿cual habria sido la cartera de minima varianza con una correlación $\rho = 2/3$? ¿y con $\rho = 0,75$? R: dicha cartera invertiría 6.562,5 euros en el activo 1 y 3.437,5 euros en el activo 2, con volatilidad 14,52%, inferior a la de los dos activos. Con $\rho = 2/3$, se tiene $w^* = 1,0$. Con $\rho = 0,75$, la cartera de minima varianza consiste en invertir 11.250 euros en el activo 1 y ponerse corto en 1.250 euros en el activo 2.

Volatilidad de la cartera como función del % invertido en Activo 1



9.2.4 Problema de Markowitz

El conocido como Problema de Markowitz es similar, pero se plantea de modo distinto: minimización de la varianza de la cartera, para una determinada rentabilidad. Este problema tiene una solución analítica, y según vamos variando los niveles de rentabilidad, las soluciones que se van obteniendo para el problema de minimización de la varianza generan la *frontera eficiente* de carteras. Cada una de dichas carteras es la cartera de menor varianza para una determinada rentabilidad esperada.

El Problema de Markowitz es:

$$\min_w w'Vw$$

$$\text{sujeto a} \quad : \quad w\prime E(r) = \bar{R},\ \sum_{i=1}^{n} w_i = 1, w_i \geq 0$$

En la formulación de ambos problemas hemos ignorado la posible prohibición de adoptar posiciones cortas, es decir, con $w_i < 0$, para algún activo i. Cuando tal restricción existe, hay que añadir las correspondientes restricciones a los problemas de optimización anteriores, lo que dificulta su solución e impide generalmente que exista una solución analítica.

La solución al problema de Markowitz es:

$$\begin{pmatrix} w^* \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 2V & 1 & E(r) \\ 1' & 0 & 0 \\ E(r)' & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ \bar{R} \end{pmatrix}$$

donde λ_1, λ_2 son los multiplicadores de Lagrange de las dos restricciones del problema.

Exercise 27 *(Ex.I.6.7) Consideremos tres activos X, Y, Z, cuyas rentabilidades tienen volatilidad 15%, 20% y 40%, respectivamente. Las correlaciones entre rentabilidades son 0,5 para X e Y, -0,7 para X y Z y -0,4 para Y y Z. Encuentre la cartera de minima varianza. R: w1 = 66,07%, w2=10,41%, w3=23,52%. La volatilidad de la cartera es reducida: 8,09%.*

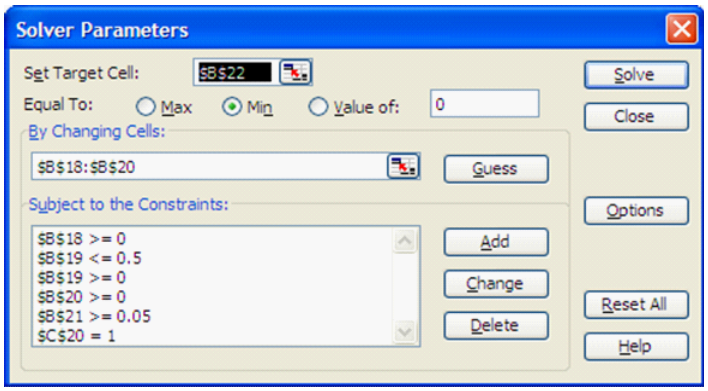
Nótese que a pesar de ser Z el activo más volatil, está incluido en la cartera de minima varianza. Ello se debe a los beneficios que produce su inclusión por diversificación, dado que tiene correlación negativa tanto con X como con Y.

Exercise 28 *(Ex.I.6.8) Si sus rentabilidades esperadas son 5%, 6% y 0%, respectivamente, encuentre una cartera que proporciona una rentabilidad esperada del 6% con la menor varianza posible. ¿Cual es la volatilidad de dicha cartera? R: w1=37,32%, w2=68,90%, w3=-6,22%, 18.75%..*

Una expresión alternativa para la solución al problema de Markowitz es:

$$\begin{aligned} w_i &= \frac{(a\psi_i - b\xi_i) - \bar{R}(b\psi_i - \tilde{V}\xi_i)}{b\tilde{V} - a^2} \\ a &= \sum_{i=1}^{m} \psi_i E(r_i);\ b = E(r)'V^{-1}E(r) \end{aligned}$$

siendo ψ_i la suma de los elementos de la columna i-ésima de V^{-1}, ξ_i la rentabilidad esperada utilizando como ponderaciones la columna i-ésima de V^{-1}, y $\tilde{V}$ es la varianza de la cartera de minima varianza sin restricciones: $\tilde{V} = \left(\sum_{j=1}^{n} \psi_j\right)^{-1}$.



Frecuentemente, es preciso incorporar restricciones adicionales al problema de Markowitz, como cuando no se permiten posiciones en corto en la cartera, o cuando se quiere limitar la posición que se toma en un determinado activo.

Exercise 29 *(I.6.9) Resuelva el problema de selección de cartera con los tres activos de los ejercicios anteriores, si a) queremos obtener una rentabilidad esperada de al menos un 5% , b) limitamos la posición en Y a un máximo de 50% del importe de la cartera y c) no admitimos posiciones cortas. R: 51,2%, 40,7%, 8,1% con rentabilidad esperada de 5% y volatilidad de 11,92%.*

Para este ejercicio el Solver de Excel se formularia:

Exercise 30 *Ex.II.4.10: Se consideran tres activos con volatilidades 25%, 20% y 30% y correlaciones: 0,6 entre X e Y, -0,5 entre X y Z y -0,6 entre Y y Z. Las rentabildades esperadas para el próximo mes son 5%, 4% y -1%, respectivamente, y se pregunta cuál es la cartera que con la minima varianza posible, genera una rentabilidad esperada para el proximo mes de 2,5%. ¿Cual es la volatilidad de la cartera resultante? R: composición de la cartera: 19,51%, 46,59%, 33,90%, volatilidad 10,26%. A continuación se estima dicha cartera teniendo en cuenta que se ha estimado un modelo A-GARCH trivariante en formato vech, y que las volatilidades actuales son: 32%, 26% y 35% y lass correlaciones: 0.75 para X e Y, -0,6 para X y Z y -0,7 para Y y Z. Las volatilidades y correlaciones de largo plazo son las mencionadas antes, y se proporcionan las ecuaciones estimadas para cada varianza y cada covarianza. ¿Cual es la volatilidad de la cartera resultante? R: composición de la cartera: 26,49%, 38,21%, 35,30%, volatilidad: 12,34%*

Es habitual que las decisiones de rebalanceo de una cartera se tomen en horizontes de un mes, y es crucial utilizar una buena estimacion de la matriz de covarianzas. Por eso, la discusión entre utilizar momentos condicionales o incondicionales es de la mayor importancia. Momentos condicionales como los que se obtienen de una especificación GARCH tienen varias ventajas: 1) pueden

recoger respuestas asimétricas de las volatilidades y las correlaciones a shocks en las rentabilidades, 2) recogen agrupamientos en volatilidad como los que se observan en los mercados, 3) convergen a los momentos análogos de largo plazo y no es preciso aplicar la regla de la raiz cuadrada para la extrapolación temporal de las volatilidades, 4) puede utilizarse la matriz de covarianzas de largo plazo que se estime conjuntamente con la especificación GARCH, o puede imponerse en la estimación del modelo GARCH una determinada matriz de covarianzas.

9.2.5 Modelo CAPM

El modelo CAPM es un modelo de equilibrio que postula que en media, todos los activos satisfacen la relación:

$$E(r_{it}) - r_{Ft} = \beta_i \left(E(r_{Mt}) - r_{Ft}\right) + \varepsilon_{it}$$

siendo r_{Mt} la rentabilidad del mercado en el período t, y $\beta_i = Cov(r_{it}, r_{Mt})/Var(r_{Mt})$ la beta del activo i. El mercado al que se refiere r_{Mt} debe ser un mercado común a todos los activos, lo que es bastante dificil de caracterizar.

De acuerdo con este modelo, todos los activos deben estar sobre una linea recta en el plano que representa la rentabilidad esperada en exceso $E(r_{it}) - r_{Ft}$ en función de su beta β_i, ya que el término $E(r_{Mt}) - r_{Ft}$ es común para todos ellos. La pendiente de dicha recta es: $E(r_{Mt}) - r_{Ft}$, y la recta se conoce como *security market line*.

Un contraste del modelo se basa en estimar el sistema:

$$r_{it} - r_{Ft} = \alpha_i + \beta_i \left(r_{Mt} - r_{Ft}\right) + \varepsilon_{it}$$

y contrastar: $H_0 : \alpha_1 = \alpha_2 = ... = \alpha_k = 0$, siendo k el número de activos disponibles para inversión. El estadistico de contraste es:

$$W = \left(1 + \left(\frac{\bar{X}}{s_X}\right)^2\right)^{-1} \hat{\alpha}' \Sigma^{-1} \alpha$$

siendo $\Sigma = Var(\varepsilon_t)$.

Extensiones del modelo CAPM

El contraste anterior supone que el riesgo sistemático de cada activo es constante en el tiempo y que el término de error tiene varianza constante. Jagannathan y Wang (1996) y Bollerslev et al. (1988) extendieron el modelo para permitir que la varianza condicional de los errores tuviera una estructura GARCH, y el riesgo sistemático variase en el tiempo.

El modelo CAPM supone que las rentabilidades de los activos con riesgo quedan representadas por su media y varianza, y el modelo ha sido extendido para permitir que el riesgo sistemático de cada activo pueda depender de momentos de orden superior. Este modelo extendido se deduce maximizando la aproximacion de Taylor de la funcion de utilidad, llegando al modelo de equilibrio:

$$E(r_{it}) - r_{Ft} = \theta_1 \frac{E\left(\tilde{r}_i \tilde{r}_M\right)}{E(\tilde{r}_M^2)} + \theta_2 \frac{E\left(\tilde{r}_i \tilde{r}_M^2\right)}{E(\tilde{r}_M^3)} + \theta_3 \frac{E\left(\tilde{r}_i \tilde{r}_M^3\right)}{E(\tilde{r}_M^4)}$$

donde $\tilde{r}_i = r_i - E(r_i), \tilde{r}_M = r_M - E(r_M)$. Los tres términos que aparecen en los numeradores son la covarianza, coskewness y cokurtosis del activo i con el índice de mercado. Harvey y Siddique (2000), *Conditional skewness in asset pricing tests*, Jornal of Finance, 54, 1263-1296 contrastan la validez de esta versión del modelo CAPM con momentos de orden superior. Un buen trabajo de revisión de la literatura es Jurczenko y Maillet (2006), *Multi-moment asset allocation and pricing models*, John Wiley and Sons.

9.3 Criterios de evaluación de carteras

Es claro de todo lo anterior que todo inversor querrá comparar la rentabilidad que obtiene con respecto al riesgo que asume, que supondremos inicialmente, por simplificar, que está adecuadamente medida por la volatilidad de dicha cartera. Si el inversor tiene varias carteras en distintos mercados, esperará tener rentabilidad más elevada en aquél mercado que le reporta un mayor riesgo. Pero querrá saber si la rentabilidad en exceso que obtiene en el mercado de mayor riesgo, respecto de la que podría obtener en el mercado de menor riesgo, le compensa suficientemente. En esta sección analizamos distintos criterios para responder a esta pregunta. El más habitual es el *ratio de Sharpe,*

$$Sharpe \equiv S = \frac{E(R) - R_f}{\sigma}$$

que relaciona la rentabilidad media, en exceso de la que proporciona el activo sin riesgo, con la volatilidad de la cartera. Entre dos carteras con distinta rentabilidades esperadas y varianzas, preferiríamos aquellas con mayor ratio de Sharpe. Una vez más, tanto la rentabilidad como la varianza deberian aparecer como valores previstos a lo largo del horizonte de futuro de inversión, aunque esta medida, al igual que las siguientes, suelen utilizarse con valores históricos.

Al calcular el ratio de Sharpe, así como los que veremos a continuación, debe tenerse en cuenta la corrección por autocorrelación al anualizar la volatilidad, en el caso de que las rentabilidades no estén incorrelacionadas:

$$Var(r_{ht}) = \sigma^2 \left(h + 2\frac{\rho}{(1-\rho)^2}[(h-1)(1-\rho) - \rho(1-\rho^{h-1})] \right)$$

siendo h el número de rentabilidades observadas por año. La raiz cuadrada del valor numérico de esta expresión es lo que debe aparecer en el denominador del ratio de Sharpe. Por el contrario, no es preciso ajustar la rentabilidad media debido a la presencia de autocorrelación.

Exercise 31 *La rentabilidad diaria de cierto activo tiene una media de 0,05% y una desviación típica de 0,75%. Calcule el ratio de Sharpe suponiendo 250 dias por año, bajo el supuesto de que las rentabilidades diarias sean independientes.*

Vuelva a calcular el ratio de Sharpe suponiendo que la autocorrelación entre rentabilidades diarias es 0,20. R: 1,0541; 0,8612.

El ratio de Sharpe puede incumplir el principio de *dominancia estocástica*. Un activo A domina estocásticamente a otro activo B en sentido fuerte si la probabilidad de que A ofrezca una rentabilidad mayor que τ es mayor que la probabilidad de que B ofrezca una rentabilidad sperior a τ para cualquier τ. Entonces el activo A debería preferirse al B. Un activo A domina estocásticamente a otro activo B en sentido débil si la probabilidad de que A ofrezca una rentabilidad mayor que τ es mayor o igual que la probabilidad de que B ofrezca una rentabilidad sperior a τ para cualquier τ.

Exercise 32 *La distribución de rentabilidades de dos activos A y B es la siguiente: A ofrece rentabilidades 20%, 10%, -20% con probabilidades 0,1; 0,8; 0,1, mientras que B ofrece rentabilidades 40%, 10%, -20% con las mismas probabilidades. Mostrar que el ratio de Sharpe incumple el principio de dominancia estocástica tanto en sentido fuerte como en sentido débil.*

El *ratio de Treynor* tiene en cuenta el interés que presenta incorporar en la cartera activos con alfa positivo, es decir con coeficientes $\alpha > 0$ en el modelo CAPM:

$$E(R_{it}) - R_{ft} = \alpha_i + \beta_i(E(R_{Mt}) - R_{ft}) + u_{it}$$

pues son activos que ofrecen una rentabilidad mayor que la que cabría esperar teniendo en cuenta su sensibilidad a las fluctuaciones del mercado y la volatilidad de éste. Sin embargo, sobreponderar tales activos introduce riesgo especifico, y conviene controlar los resultados obtenidos. El ratio de Treynor se define:

$$Treynor \equiv TR = \frac{\alpha}{\beta}$$

La cantidad óptima de un activo con α positivo que debe incorporarse a la cartera viene definida por el *ratio de Información* (*appraisal ratio*):

$$IR = \frac{\alpha}{\sigma}$$

aunque Jensen (1969) argumenta que, para establecer un ranking de carteras, es preferible utilizar el coeficiente estimado α (α de Jensen) por sí solo antes que los ratios TR o IR .

No conviene olvidar que el uso de estos ratios está estrictamente justificado sólo bajo el supuesto de preferencias representadas por una función de utilidad exponencial y rentabilidades con distribución Normal.

Cuando el inversor es averso no solo a la volatilidad sino también a la asimetría negativa (fuertes pérdidas) y a un exceso de curtosis positivo, el ratio de Sharpe se modifica del siguiente modo:

$$RSM = S + \frac{\tau}{6}S^2 - \frac{\varkappa}{24}S^3$$

donde S denota el ratio de Sharpe tradicional que antes vimos, τ es el coeficiente de asimetría y $\varkappa$ es el exceso de curtosis, $\varkappa = \kappa - 3$.

Rentabilidades no Normales

En contextos no Gaussianos, y bajo utilidad logaritmica, es recomendable utilizar el *Ratio de Sharpe Generalizado*, que se define en función del máximo nivel alcanzable de utilidad esperada EU^*, como:

$$RSG = \sqrt{-2\ln(-EU^*)}$$

teniendo en cuenta que bajo preferencias logaritmicas, esta raiz cuadrada estará bien definida.

Cuando se invierten q unidades, el Equivalente Cierto es:

$$EC(q) \approx q\mu - \frac{1}{2}\gamma\sigma^2 q^2 + \frac{\tau}{6}\gamma^2\sigma^3 q^3 - \frac{\kappa-3}{24}\gamma^3\sigma^4 q^4$$

Despreciando términos de orden superior, el valor de q que maximiza $EC(q)$ es: $q^* = \mu/(\gamma\sigma^2)$, por lo que:

$$\begin{aligned} EC(q^*) &\approx \frac{\mu^2}{\gamma\sigma^2} - \frac{1}{2}\gamma\sigma^2\left(\frac{\mu}{\gamma\sigma^2}\right)^2 + \frac{\tau}{6}\gamma^2\sigma^3\left(\frac{\mu}{\gamma\sigma^2}\right)^3 - \frac{\kappa-3}{24}\gamma^3\sigma^4\left(\frac{\mu}{\gamma\sigma^2}\right)^4 = \\ &= \frac{\mu^2}{\gamma\sigma^2} - \frac{1}{2}\frac{\mu^2}{\gamma\sigma^2} + \frac{\tau}{6}\frac{\mu^3}{\gamma\sigma^3} - \frac{\kappa-3}{24}\frac{\mu^4}{\gamma\sigma^4} \end{aligned}$$

y teniendo en cuenta la identidad: $EU^* = U(EC^*)$, tenemos, bajo una función de utilidad exponencial que el máximo nivel alcanzable de utilidad esperada es, aproximadamente:

$$EU^* \approx -\exp\left(-\frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2 - \frac{\tau}{6}\left(\frac{\mu}{\sigma}\right)^3 + \frac{\kappa-3}{24}\left(\frac{\mu}{\sigma}\right)^4\right)$$

Pero $\mu/\sigma = S$, el ratio de Sharpe estándar, por lo que:

$$RSG = \sqrt{-2\ln(-EU^*)} = \sqrt{S^2 + \frac{1}{3}\tau S^3 - \frac{1}{12}\varkappa S^4}$$

Al igual que sucede con el ratio de Sharpe modificado (RSM), también RSG se reduce al ratio de Sharpe tradicional cuando la rentabilidad tiene una distribución Normal y la función de utilidad es exponencial.

Indices Kappa

Los indices Kappa fueron introducidos por Kaplan y Knowles (2004), *Kappa: A generalized downside risk performance measure*, Journal of Performance Measurement, 8, 42-54. El *índice Kappa de orden* δ respecto al umbral h se define como:

$$K_\delta(h) = \frac{E(R) - h}{LPM_{\delta,h}(R)}$$

donde h denota cierto umbral de rentabilidad, $\delta > 0$ pero no necesariamente entero, y $LPM_{\delta,h}(R)$ es el lower partial moment (momento parcial inferior) de la rentabilidad R, definido como:

$$LPM_{\delta,h}(R) = E\left(|\min(0, R-h)|^{\delta}\right)^{1/\delta} = E\left[|\max(0, h-R)|^{\delta}\right]^{1/\delta}$$

El indice Kappa es el cociente entre la rentabilidad esperada en exceso del umbral y una función (el momento parcial inferior de orden δ) de la rentabilidad esperada cuando ésta es inferior al umbral. El denominador del indice no depende de las rentabilidades por encima del umbral, es como si el inversor no se preocupase de éstas. El umbral puede ser la rentabilidad del activo sin riesgo, aunque en ocasiones también se propone utilizar una rentabilidad muy elevada, ya que ignoramos lo que sucede por encima de ella.

Cuando el umbral es la rentabilidad del activo sin riesgo, se trabaja con *rentabilidades en exceso* de dicha rentabilidad. Cuando el umbral es la rentabilidad del fondo de referencia, se trabaja con *rentabilidades activas*. En todos los casos, se trabaja con numerador y denominador en términos anualizados, para presentar los resultados asimismo en términos anualizados. Un inversor preocupado por la asimetría y la curtosis, con alta aversión al riesgo, deberia utilizar indices Kappa con $\delta = 3$ ó 4, mientras que un inversor menos averso al riesgo puede utilizar órdenes inferiores de este índice. A diferencia del ratio de Sharpe, en esta familia de índices intervienen todos los momentos de la distribución de rentabilidades.

Cuando se calculan estos momentos a partir de series temporales de rentabilidades, intervienen únicamente las rentabilidades por debajo del umbral escogido h, pero no debe olvidarse "contar los ceros", es decir dividir la media de las rentabilidades inferiores a h por la proporción que éstas representan en el total muestral. Se trata de una esperanza incondicional, no de una esperanza condicionada por el suceso $R < h$.

El indice Kappa de primer orden es igual a la rentabilidad en exceso del umbral h, dividida por el momento parcial de primer orden:

$$K_1(h) = \frac{E(R) - h}{E\left[\max(0, h-R)\right]}$$

El indice Kappa de primer orden está relacionado con el *estadístico Omega* introducido por Keating y Shadwick (2002). Denotando por $F(R)$ la función de distribución de la rentabilidad R, el estadístico Omega es:

$$\Omega(h) = \frac{E\left[\max(0, R-h)\right]}{E\left[\max(0, h-R)\right]} = \frac{\int_h^{\infty}(1 - F(R))dR}{\int_{-\infty}^{h} F(R)dR}$$

que es el cociente entre la rentabilidad esperada por encima del umbral h y la rentabilidad esperada por debajo de dicho umbral. Es asimismo el cociente del área por debajo de la curva $F(R)$ a la derecha y a la izquierda del umbral h.

Un valor elevado del índice Omega puede interpretarse como que el activo proporciona más ganancias que pérdidas en relación con un umbral de rentabilidad h. A diferencia del ratio de Sharpe, que considera unicamente los dos primeros momentos de la distribucion de rentabilidades, el ratio Omega considera toda la distribución de probabilidad.

Puede probarse que:[15]

$$K_1(h) = \Omega(h) - 1$$

El indice Kappa de segundo orden, utilizando como umbral h la rentabilidad del activo sin riesgo R_f, se conoce como *ratio de Sortino.* Tambien se conoce como el *Upside-Potential ratio,* si se calcula respecto a lo que se considere la mínima rentabilidad aceptable (supongamos que se toma como tal la rentabilidad del activo sin riesgo R_f):

$$RS \equiv K_2(R_f) = \frac{E(R) - R_f}{E\left[\max(0, R_f - R)^2\right]^{1/2}} = \frac{E(R) - R_f}{LPM_{2,R_f}(R)}$$

Todos los indices Kappa aumentan al disminuir el umbral de rentabilidad escogido, h. Son negativos cuando $E(R) < h$, y positivos cuando sucede lo contrario. Los índices Kappa de orden superior son más sensibles a la elección de umbral h. También son más sensibles a la asimetría y al exceso de curtosis porque los momentos parciales inferiores de orden superior son más sensibles a rentabilidades extremas (ver Example I.6.15).

[15] Nótese que el numerador de $\Omega(h) - 1$ es: $E\left[\max(0, R - h)\right] - E\left[\max(0, h - R)\right] = \max(0, ER - h) - \max(0, h - ER)$

que será igual a $ER - h$, el numerador de $K_1(h)$, tanto si $ER > h$ como si $ER < h$.