Valor en Riesgo

Alfonso Novales
Departamento de Economía Cuantitativa
Universidad Complutense

Enero de 2016
Versión preliminar
No citar sin permiso del autor

Contents

1 Valor en Riesgo: Definiciones

El Valor en Riesgo (VaR) es una de las medidas utilizadas para evaluar el riesgo de una determinada posición o cartera de activos financieros. La definición del VaR puede hacerse en términos de rentabilidades o en términos de Pérdidas y Ganancias P&L (términos nominales); la definición también depende de que se aplique a una posición larga (comprada), como es habitual, o a una posición corta (vendida) en un activo financiero. Consideremos la situación de un inversor que ha comprado un determinado activo o cartera. El VaR responde entonces a la pregunta: dado un determinado horizonte de gestión (se denomina *horizonte de riesgo*) y con una cierta probabilidad reducida, por ejemplo, p =0,1% ó 1%, ¿cuál es la caída en el valor del activo que será sobrepasada sólo con una probabilidad del p%, o un porcentaje p% de los días? Una interpretación equivalente es que con probabilidad $1-p$ el propietario de dicha posición experimentará una pérdida no superior al VaR, o posiblemente un beneficio. Para una posición corta, el VaR sería el aumento en el precio de la cartera que será sobrepasado solo con una probabilidad del p%.

Por consiguiente, el VaR no es sino un determinado percentil de la distribución de probabilidad prevista para las variaciones en el valor de mercado de la cartera en el horizonte de tiempo escogido. Puede estimarse asimismo en términos del percentil correspondiente de la distribución de probabilidad seguida por la rentabilidad de la cartera, y es habitual hacerlo así, pero se pierde información acerca de la cuantía de la posible pérdida. Conviene, sin embargo, hablar en términos nominales, lo que se consigue facilmente cuando se ha estimado un VaR en términos de rentabilidades, sin más que multiplicar su valor numérico por el valor de mercado de la cartera en el momento de estimación del VaR.

El VaR puede calcularse para períodos de inversión de un día o también superiores, como una semana o un mes. De hecho, para el cálculo del VaR hay que especificar el *nivel de significación* p o equivalentemente, *el nivel de confianza* $1-p$, y el *horizonte de riesgo* al cual se está calculando la rentabilidad en cuestión. El VaR dependerá del horizonte de riesgo y evolucionará en el tiempo, según cambie nuestra percepción de la distribución de rentabilidades al ir recibiendo más informacion.

El horizonte debe estar asociado al tiempo durante el cual pensamos que vamos a estar expuestos al riesgo con la posición asumida por el activo o cartera. Ese período de tiempo es menor en los activos muy liquidos, y mayor en los activos poco liquidos. Por eso, al calcular el VaR en un banco, no es extraño hablar del VaR a 1 año. Esto es lo natural si estamos evaluando la posibilidad de insolvencia de una compañía, es decir, al evaluar riesgo de crédito, y tambien lo es al estimar el nivel de riesgo operacional. En cambio, al hablar de activos que cotizan en mercados liquidos como son las acciones, puede ser adecuado calcular el VaR a un día. Esto es más habitual en mesas de Tesorería. Cuando se fijan limites de negociación a los *traders*, se suele trabajar con el VaR a 1 día al nivel de confianza del 95%, umbral que un trader no debe sobrepasar con excesiva frecuencia (por ejemplo, más del 5% de los días). Un nivel de confianza superior daría excesiva libertad al *trader*. Otro ejemplo de utilización: supongamos que una agencia de crédito tiene establecida la concesión a una empresa de una calificación AA solo si la empresa puede probar que la probabilidad de quiebra en el horizonte de un año es de 0,03% o inferior, o de 3,24% o inferior en un horizonte de 10 años. En ese caso, una empresa que aspire a dicha calificación, calculará su VaR a 1 año a 99,97% de confianza, asi como su VaR a 10 años al 96,76% de confianza.

Para fijar el capital regulatorio de los bancos, Basilea II adoptó como criterio estándar el cálculo del VaR 1% para la rentabilidad a 10 días (2 semanas de mercado). Pero, en media, se debería obtener una pérdida superior al VaR 1% en uno de cada 100 dias, es decir, una vez cada cuatro años. Esto puede ser inasumible para el supervisor bancario, de ahí que se multiplique el VaR por un factor de 3 ó 4 para obtener el capital regulatorio exigible a la entidad.

Sea $\Delta V(h)$ la variación en el valor de los activos de una posición financiera entre t y $t+h$, medida en unidades monetarias. En t, el valor que la posición tenga en $t+h$ es aleatorio, y denotamos por $F_h(x)$ la función de distribución de $\Delta V(h)$. Definimos el VaR *nominal* de una posición larga en el horizonte de h días, con probabilidad p, como la cantidad VaR que satisface:

$$p = P\left[\Delta V(h) \leq VaR\right] = F_h(VaR) \tag{1}$$

Puesto que la probabilidad de los valores posibles de $\Delta V(h)$ estará repartida de manera relativamente equilibrada entre valores positivos y negativos se tendrá que, para valores reducidos de la probabilidad p, el VaR será habitualmente negativo, por lo que se proporciona cambiado de signo: $-VaR$. Puesto que V_t está dado, la distribución de probabilidad de $\Delta V(h)$ tiene asociada de modo biunívoco una distribución de probabilidad para V_{t+h}, de modo que una vez estimado el VaR $\Delta V(h)^*$, tendremos asimismo un umbral V_{t+h}^* para el propio valor de la cartera: $V_{t+h}^* = V_t + \Delta V(h)^*$.

Para *posiciones cortas*, la pérdida se producirá ante una elevación del precio de la cartera de magnitud poco habitual, por lo que tendríamos:

$$p = P\left[\Delta V(h) \geq VaR\right] = 1 - P\left[\Delta V(h) \leq VaR\right] = 1 - F_h(VaR)$$

y para una p pequeña, tal cantidad será positiva. Por tanto, la cola izquierda de la distribución de $F_h(x)$ es la relevante para posiciones largas, mientras que la cola derecha es la relevante para las posiciones cortas.

Asimismo, la definición (1) es válida para posiciones cortas si utilizamos la distribución de $-\Delta V(h)$, porque el valor numérico ΔV^* que deja a su derecha un 1% de variaciones mayores, es asimismo el valor numérico $-\Delta V^*$ que deja por debajo un 1% de valores inferiores de $-\Delta V(h)$. Y algo similar puede decirse si aplicamos la definición relativa a la cola derecha para posiciones largas a la distribución de $-\Delta V(h)$. Por tanto, es suficiente analizar los métodos de cálculo del VaR para posiciones largas o para posiciones cortas. Estas afirmaciones no tienen nada que ver con la posible simetría de la distribución de $\Delta V(h)$, que no es preciso suponer para el cálculo del VaR.

Salvo que se diga lo contrario, en lo sucesivo consideramos posiciones largas en un activo o cartera.

Para una distribución univariante, $F_h(x)$ y una probabilidad p, $0 < p < 1$, el cuantil p-ésimo de $F_h(x)$ es:

$$x_p = \inf \{x \mid F_h(x) \geq p\}$$

donde inf denota la menor de las cantidades que satisface la desigualdad indicada. Si se conociese la distribución $F_h(x)$, entonces el VaR de la cartera sería simplemente el cuantil p-ésimo de $F_h(x)$. Sin embargo, esta distribución se desconoce en la práctica, y el cálculo del VaR requiere estimar $F_h(x)$ o su cuantil p-ésimo.

El VaR suele contabilizarse en términos de rentabilidades, aunque hemos de tener en cuenta que si la cartera consta de posiciones largas y posiciones cortas, el concepto de rentabilidad puede perder su sentido (por ej., el valor de la posición puede hacerse cero o ser muy reducido, dificultando la interpretación de los valores numéricos de las rentabilidades), y es preferible trabajar con las Pérdidas y Ganancias ($P\&L$).

Tenemos que la variación en el valor de la cartera entre dos períodos es :

$$\Delta V(h) = V_{t+h} - V_t = V_t(1 + R_{t,t+h}) - V_t = V_t R_{t+h}$$

siendo $R_{t,t+h}$ la rentabilidad alcanzada por la cartera entre t y $t+h$. Puesto que V_t está dado, la distribución de probabilidad de $\Delta V(h)$ tiene asociada de modo biunívoco distribuciones de probabilidad para el valor de la cartera V_{t+h} y para las rentabilidades, $R_{t,t+h}$, y el *VaR* en rentabilidades es el p-cuantil $R^*_{t,t+h}$ de dicha distribución de probabilidad, que para valores reducidos de p será un número negativo, $R^*_{t,t+h} < 0$.

El *VaR nominal respecto al origen o VaR absoluto* puede escribirse, por tanto:

$$VaR(origen) = V_t - V^*_{t+h} = V_t - (1 + R^*_{t,t+h})V_t = -V_t R^*_{t,t+h}$$

que será generalmente positivo.

En ocasiones, es más conveniente estimar el VaR *relativo*, que considera las pérdidas que puedan producirse en la cartera, no con respecto a una rentabilidad nula entre t y $t+h$, sino con respecto a la rentabilidad esperada de la misma. Si denotamos por μ la rentabilidad esperada, $\mu_R = E(R_{t,t+h})$, tenemos:

$$VaR(relativo) = E(V_{t+h}) - V^*_{t+h} = (1+E(R_{t,t+h}))V_t - (1+R^*_{t,t+h})V_t = -V_t \left(R^*_{t,t+h} - \mu_R\right)$$

siendo generalmente positivo.

Cuando el horizonte de inversión es reducido (un dia, una semana), la rentabilidad esperada será prácticamente nula, con lo que el *VaR respecto al origen* y el *VaR relativo* practicamente coincidirán. En general, el *VaR* relativo es más apropiado, puesto que considera las desviaciones en el valor de la cartera con respecto al escenario central. El *VaR* relativo es un enfoque más conservador si, como sucederá en la mayoría de los casos, la cartera tiene una rentabilidad esperada positiva, pues ello hará que el *VaR* relativo sea más elevado en valor absoluto que el *VaR* calculado respecto al origen.

1.1 Metodologías para el cálculo del Valor en Riesgo

Existen distintos enfoques para el cálculo del VaR: *i*) el modelo lineal, *ii*) el VaR histórico, y *iii*) el método de simulación de Monte Carlo, y sólo para el primero de ellos es necesario establecer un supuesto acerca de la distribución de probabilidad de las rentabilidades, lo cual es bastante conveniente.

1. método paramétrico de VaR, en el que suponemos que la distribución de las rentabilidades de los factores sigue una distribución Normal multivariante y la cartera es función lineal de los factores. Su ventaja es que es tratable analíticamente, pero solo se puede generalizar a una pocas formas paramétricas, como la Normal, la t-Student, o mixturas de Normales o de t-Student. Cuando se incluyen tipos de interés entre los factores las relaciones son no lineales, pero la no linealidad ya está incorporada en los términos PV01. El modelo no puede aplicarse, sin embargo, a carteras con opciones. En este modelo no podemos predecir la matriz de covarianzas utilizando un modelo GARCH, porque ello significa que las rentabilidades no son i., i.d.. Como consecuencia, la regla de la raiz cuadrada en la extrapolación de la varianza no es válida. Pero el mayor problema es que en ese caso, desconocemos cual es la distribución de las rentabilidades h dias a partir de hoy.

2. método de simulación histórica, que utiliza un gran cantidad de datos históricos para estimar el VaR pero hace el mínimo de supuestos acerca de la distribución de probabilidad seguida por las rentabilidades de los factores. Supone que todas las variaciones futuras posibles en los precios de los activos ya se han observado en el pasado. Esto impone restricciones no muy realistas en los datos.

3. método VaR Monte Carlo, que hace supuestos similares a los del modelo lineal Normal. Se puede aplicar a carteras no lineales, activos con pagos *path-dependent*, etc.. Pero es computacionalmente intensivo y los errores de simulación pueden ser considerables, por lo que conviene utilizar métodos numéricos de cierto nivel de sofisticación.

[*Case Study: Cálculo del VaR para el índice* S&P500, IV.1.9.4]. La distribución del S&P500 es leptocúrtica, lo que hace que el VaR histórico sea superior al VaR Normal y, generalmente, también mayor que el VaR Monte Carlo.

1.2 Enfoque paramétrico: Estimación del VaR bajo Normalidad

Si denotamos por $f(R)$ la función de densidad de la rentabilidad de una cartera a lo largo de un período, el VaR R^* al nivel de significación $p\%$ es el p-cuantil de la distribución de rentabilidades, es decir, el número real R^* que satisface la igualdad:

$$p = \int_{-\infty}^{R^*} f(R)dR = P(R < R^*)$$

o, equivalentemente, el número real que deja a su derecha un $(1-p)\%$ de probabilidad: $1 - p = \int_{R^*}^{\infty} f(R)dR = P(R > R^*)$. Como p es un valor reducido: $5\%, 1\%$ o incluso 0,1%, entonces R^* será una rentabilidad negativa, y el VaR se proporciona cambiando de signo a R^* para que resulte un número positivo,

$$VaR = -R^*$$

que se interpretará como el peor resultado que puede producirse, dejando aparte los $p\%$ casos peores, en el horizonte de inversión considerado en el cálculo del VaR.

Supongamos que estamos dispuestos a aceptar que la rentabilidad anual de la cartera sigue una distribución Normal: $R \sim N(\mu_R, \sigma_R)$, donde la rentabilidad media y la varianza están en términos anualizados. Entonces: $\frac{R-\mu_R}{\sigma_R} \sim N(0,1)$, y tendremos :

$$p = P(R < R^*) = P\left(\frac{R-\mu_R}{\sigma_R} < \frac{R^*-\mu_R}{\sigma_R}\right) = \Phi\left(\frac{R^*-\mu_R}{\sigma_R}\right)$$

donde Φ denota la función de distribución de una $N(0,1)$.Por la simetría de la distribución Normal, tenemos también:

$$\Phi\left(\frac{R^*-\mu_R}{\sigma_R}\right) = 1 - \Phi\left(-\frac{R^*-\mu_R}{\sigma_R}\right)$$

de modo que:

$$\Phi\left(-\frac{R^*-\mu_R}{\sigma_R}\right) = 1-p \Rightarrow -\frac{R^*-\mu_R}{\sigma_R} = \Phi^{-1}(1-p) \Rightarrow R^* = -\Phi^{-1}(1-p)\sigma_R + \mu_R$$

por lo que si la rentabilidad esperada sobre el periodo de cálculo del VaR es positiva, el VaR del activo o cartera se reducirá, sucediendo lo contrario si la rentabilidad esperada fuese negativa.

y el Valor en Riesgo *respecto al origen* a horizonte 1 periodo y nivel de significación p, será:

$$VaR = -R^* = \Phi^{-1}(1-p)\sigma_R - \mu_R = \alpha\sigma_R - \mu_R$$

Este sería el denominado modelo lineal Normal del VaR de una cartera, que surge cuando hacemos el supuesto de Normalidad sobre la distribución de probabilidad de rentabilidades de la cartera. Este es el enfoque paramétrico para el cálculo del VaR, frente a los enfoques Histórico o de Simulación Monte Carlo, que veremos más adelante.

Ejemplo: Supongamos que estamos interesados en calcular el VaR 1% de una rentabilidad que sigue una distribución $N(\mu_R, \sigma_R)$, con momentos anualizados. La tabla de la distribución $N(0,1)$ nos proporciona: $p = 0,01 \Rightarrow \Phi^{-1}(0,99) = 2,326$, de modo que:

$$R^* = 2,326\sigma_R - \mu_R$$

Si la cartera tiene una rentabilidad positiva, el VaR 1% a 1 año será inferior a $2,326\sigma_R$, siendo mayor que $2,326\sigma_R$ si la rentabilidad esperada de la cartera durante el horizonte de inversión es negativa.

Ejemplo IV.1.4: ¿Cuanto es el VaR al 10% sobre horizonte de un año, de 2 millones de euros invertidos en un fondo cuya rentabilidad anual, en exceso de la ofrecida por el activo sin riesgo se distribuye Normal, con esperanza 5% y volatilidad 12%?

$$0.10 = P(X \leq x) = P\left(Z \leq \frac{x - 0.05}{0.12}\right) \Rightarrow \frac{x - 0.05}{0.12} = -1.2816 \Rightarrow x = -0.1038$$

Luego con probabilidad 90%, el resultado de la inversión será mejor de lo que se obtiene descontando 10.38% de la rentabilidad del activo sin riesgo, lo que generalmente dará un resultado negativo, denotando una pérdida. Como 10.38% de 2 millones es 207.572 euros, podemos decir que la cartera tendra un resultado mejor que el obtenido al restar esta cantidad de la inversión en el activo sin riesgo. Como veremos más adelante, obtener la rentabilidad del activo sin riesgo equivale a obtener una rentabilidad descontada nula, que es la rentabilidad que nos debe interesar si tomamos en cuenta el valor temporal del dinero. Por tanto, alternativamente diríamos que esperamos no perder, en términos descontados al dia de hoy, más de 207.572 euros.

1.3 Extrapolación temporal del VaR

Corrección por autocorrelación

Recordemos que la regla de extrapolación temporal de la varianza debe aplicarse con cuidado si las rentabilidades presentan autocorrelación. Muy frecuentemente, se calcula el VaR bajo el supuesto de Normalidad sobre un horizonte tan breve como 1 día, y se extrapola el valor numérico obtenido al horizonte deseado. Tal extrapolación se basa en el supuesto de que trabajamos con rentabilidades logarítmicas, y que éstas son independientes en el tiempo (carecen de autocorrelación).

Bajo tales condiciones, si hemos estimado una rentabilidad y volatilidad diarias μ_1, σ_1, tenemos:

Volatility	0.1% 10-day	0.1% 1-day	1% 10-day	1% 1-day	5% 10-day	5% 1-day
5%	3,1%	1,0%	2,3%	0,7%	1,6%	0,5%
10%	6,2%	2,0%	4,7%	1,5%	3,3%	1,0%
15%	9,3%	2,9%	7,0%	2,2%	4,9%	1,6%
20%	12,4%	3,9%	9,3%	2,9%	6,6%	2,1%
25%	15,5%	4,9%	11,6%	3,7%	8,2%	2,6%
30%	18,5%	5,9%	14,0%	4,4%	9,9%	3,1%
40%	24,7%	7,8%	18,6%	5,9%	13,2%	4,2%
50%	30,9%	9,8%	23,3%	7,4%	16,4%	5,2%
75%	46,4%	14,7%	34,9%	11,0%	24,7%	7,8%
100%	61,8%	19,5%	46,5%	14,7%	32,9%	10,4%

$$VaR = \sigma_1 \sqrt{h} \Phi^{-1}(1-p) - h\mu_1$$

Si hemos trabajado con rentabilidades porcentuales, como suele ser habitual, la expresión anterior será una aproximación, que dejará de ser exacta para horizontes h largos.

La tabla muestra los valores que puede tomar el VaR según este modelo, en funcion del horizonte de riesgo, el nivel de significación y la volatilidad de la cartera. En ella podemos familiarizarnos con el valor numérico que puede tomar el VaR en rentabilidad.

Por otra parte, si las rentabilidades no son independientes en el tiempo, entonces

$$\sigma_h = \sqrt{\tilde{h}}\sigma_1$$

para una cierta constante $\tilde{h}$. Por ejemplo, si las rentabilidades diarias siguen un proceso autoregresivo de primer orden:

$$\tilde{h} = h + 2\rho(1-\rho)^{-2}[(h-1)(1-\rho) - \rho(1-\rho^{h-1})]$$

[Ver Ejercicios [EIV.1.5] [EIV.2.1] sobre la corrección por autocorrelación]

VaR respecto al origen y VaR relativo

Si los parámetros μ_R y σ_R están expresados en términos anualizados, enla estimación del VaR debemos utilizar los valores numéricos de dichos momentos para el periodo de cálculo del VaR.

Teniendo esto encuenta, el VaR nominal respecto al origen sobre un horizonte Δt será el producto del VaR en rentabilidad y el valor nominal de la cartera:

$$VaR(origen) \equiv V^*_{t+h} = -V_t R^*_{t,t+h} = V_t \left(\Phi^{-1}(1-p)\sigma_R \sqrt{\Delta t} - \mu_R \Delta t \right) = V_t \left(\alpha \sigma_R \sqrt{\Delta t} - \mu_R \Delta t \right)$$

mientras que el VaR nominal *respecto de la rentabilidad media* es:

$$
\begin{aligned}
VaR(relativo) &= E(V_{t+h}) - V^*_{t+h} = -V_t\left(R^*_{t,t+h} - \mu_R \Delta t\right) = \\
&= -V_t\left[-\left(\Phi^{-1}(1-p)\sigma_R\sqrt{\Delta t} - \mu_R\Delta t\right) - \mu_R\Delta t\right] = \\
&= V_t\Phi^{-1}(1-p)\sigma_R\sqrt{\Delta t} = V_t\alpha\sigma_R\sqrt{\Delta t}
\end{aligned}
$$

que se obtiene multiplicando la desviación típica de la rentabilidad por el valor de la cartera, el cuantil de rentabilidad al nivel de confianza p escogido y el horizonte de riesgo, Δt. Notemos que p será reducido, por lo que $\alpha = \Phi^{-1}(1-p) > 0$.

En estas expresiones, una rentabilidad media positiva reducirá el VaR respecto del origen. Como ya vimos en el caso general (sin suponer Normalidad), el VaR relativo (respecto de la rentabilidad esperada de la cartera) será más elevado que el VaR respecto del origen.

El producto $E = \sigma_d V$, donde la desviación típica se refiere al horizonte de riesgo, se conoce como *exposición neta* de la cartera.

Este método es válido para otras distribuciones de probabilidad, siempre que la incertidumbre esté resumida en el parámetro σ_R. La distribución Normal puede resultar válida para carteras amplias, bien diversificadas, pero no cuando existe una concentración de riesgos.

En las versiones previas de las normas de Basilea se exigía como capital regulatorio un múltiplo de 3 ó 4 veces el VaR, lo cual tiene una justificación estadística: la desigualdad de Chebychev asegura que cualquier variable aleatoria, en particular la rentabilidad r de la cartera, satisface: $P\left(\mid r-\mu_r \mid > \lambda\sigma_r\right) \leq \frac{1}{\lambda^2}$. Por tanto, si la distribución de rentabilidades es simétrica, se tendrá para rentabilidades por debajo de la media :

$$
P\left(r - \mu_r < -\lambda\sigma_r\right) \leq \frac{1}{2\lambda^2}
$$

Si para el cálculo del VaR fijamos el miembro derecho de esta desigualdad en 1%, tendremos: $\lambda = 7,071$, y el estimador del VaR según este argumento será: $(7,071)\,\sigma_r$. Supongamos que la entidad presenta al supervisor su VaR al 1% bajo el supuesto de Normalidad. Estará dando la cifra de $(2,326)\,\sigma_r$. Si esta distribución no es correcta, el factor de corrección sobre el VaR calculado debería ser, aproximadamente: $\frac{7{,}071\sigma_r}{2{,}326\sigma_r} = 3.03$, muy próximo al factor definido en Basilea II.

1.4 Descontando el valor futuro de la cartera

Un segundo aspecto que hemos de tener en cuenta cuando estimamos el *VaR* a horizontes largos, es el de utilizar el valor presente de la cartera, pues solo así estaremos teniendo realmente en cuenta el valor temporal del dinero. Para ello descontamos el valor final de la posición V_{t+h}, utilizando el LIBOR, por ejemplo, como tipo libre de riesgo. El resultado de Pérdidas y Ganancias ($P\&L$) es:

$$P\&L(h) = B_{ht}V_{t+h} - V_t$$

siendo B_{ht} el precio de un bono cupón cero con vencimiento dentro de h dias de negociación, el factor descuento habitualmente utilizado para actualizar flujos de caja, que será un factor inferior a 100, que utilizamos en porcentaje. Si R_F es la tasa anual libre de riesgo y el plazo es inferior a un año, $B_{ht} = \frac{1}{1+hR_F} < 1$. Cuando hablemos de Pérdidas y Ganancias, lo haremos en términos descontados, salvo que el horizonte de riesgo sea tan corto que no sea necesario considerar el descuento.

Como ejemplo, observemos los factores descuento:

R_F	$1año$	$6meses$	$4meses$	$3meses$	$1mes$	$10dias$	$5dias$
5%	$0,951$	$0,975$	$0,983$	$0,987$	$0,996$	$0,998$	$0,999$
3%	$0,971$	$0,985$	$0,990$	$0,993$	$0,998$	$0,999$	$0,999$

por lo que el trabajo con rentabilidades descontadas será necesario únicamente para tipos de interés relativamente elevados y horizontes largos, pudiendo evitarse, como aproximación, en los demás casos.

En lo sucesivo, denotamos este factor descuento por $d : d \equiv B_{ht}$.

La *rentabilidad descontada* de un activo o cartera es:

$$R_d = \frac{dV_{t+h} - V_t}{V_t}$$

Respecto de la rentabilidad habitual, $R = \frac{V_{t+h}-V_t}{V_t}$ tenemos la relación:

$$1 + R_d = (1+R)d = \frac{1+R}{1+R_F} \tag{2}$$

de modo que si denotamos por $\mu_d, \sigma_d, \mu, \sigma$ la esperanza y desviación típica de las rentabilidades descontada y sin descontar, respectivamente, tendremos:

$$\begin{aligned} \mu_d &= E(R_d) = dE(R) - (1-d), \\ Var(R_d) &= Var(R)d^2 \end{aligned}$$

por lo que: *La rentabilidad descontada es menor que la rentabilidad calculada sin descontar el valor futuro de la cartera y tiene una varianza inferior.*

La diferencia entre ambas rentabilidades es:

$$R - R_d = \frac{V_{t+h} - V_t}{V_t} - \frac{dV_{t+h} - V_t}{V_t} = \frac{V_{t+h} - dV_{t+h}}{V_t a <} = \frac{V_{t+h}}{V_t}(1-d) = (1+R)\frac{R_F}{1+R_F}$$

Aunque no solemos tenerlo en cuenta, realmente sufrimos una pérdida cuando obtenemos una rentabilidad descontada negativa en el valor de nuestra cartera. Equivalentemente, teniendo en cuenta el valor temporal del dinero, una pérdida significa que nuestra cartera ha obtenido una rentabilidad inferior a la del activo sin riesgo. En efecto: $dV_{t+h} - V_t < 0$, significa: $\frac{1}{1+hR_F}V_{t+h} - V_t < 0$, es decir:

$V_{t+h} < V_t(1+hR_F)$, y $\frac{V_{t+h}}{V_t} < 1+hR_F$, por lo que si tenemos una rentabilidad descontada negativa, se debe a que nuestra cartera ha generado una rentabilidad inferior a la del activo sin riesgo, es decir, un rendimiento en exceso de signo negativo. Una ganancia, es decir, una rentabilidad descontada positiva, equivale a una rentabilidad superior a la ofrecida por el activo sin riesgo. Una cartera ofrece una rentabilidad igual a la del activo sin riesgo, $R = R_F$, si y sólo si su rentabilidad descontada es igual a cero, $R_d = 0$, como puede verse en (2).

Cálculo correcto del VaR con rentabilidades descontadas

Si nuestra cartera ofrece una rentabilidad esperada igual a la del activo sin riesgo, entonces la distribución de $P\&L$ descontadas estará centrada en torno al origen. Si la distribución es Normal, será simétrica en torno a dicho punto, que será además el valor más probable. El $VaR_p(h)$ *nominal* es el p-cuantil de la distribución de Pérdidas&Ganancias ($P\&L$) descontadas. Su estimación es la cantidad $X_{ht,p}$ que resuelve la ecuación:

$$p = P\left[dV_{t+h} - V_t \leq X_{ht,p}\right] \tag{3}$$

Si calculamos el $VaR_p(h)$ en términos de rentabilidades, el VaR es entonces el valor numérico $x_{ht,p}$ que satisface:

$$p = P\left[\frac{dV_{t+h} - V_t}{V_t} \leq x_{ht,p}\right] = P(R_d \leq x_{ht,p})$$

Si, pr el contrario, utilizamos la rentabilidad no descontada, como suele ser habitual en el cálculo del VaR, $y_{ht,p}$, tenemos:

$$p = P\left[\frac{V_{t+h} - V_t}{V_t} = R \leq y_{ht,p}\right]$$

¿Qué relación existe entre ambas estimaciones del VaR, según que descontemos o no descontemos las valoraciones futuras? Tenemos:

$$p = P\left[R_d \leq x_{ht,p}\right] = P\left[(1+R)d - 1 \leq x_{ht,p}\right] = P\left[R \leq \frac{1+x_{ht,p}}{d} - 1\right]$$

Por tanto:

$$y_{ht,p} = \frac{1+x_{ht,p}}{d} - 1, \text{ es decir, } x_{ht,p} = (1+y_{ht,p})d - 1$$

y es lo que determina la diferencia entre ambas estimaciones del VaR porcentual (en rentabilidades). Recordemos que, *cambiados de signo*, $x_{ht,p}$ es el VaR de las rentabilidades descontadas, e $y_{ht,p}$ es el VaR de las rentabilidades sin descontar, por lo que:

VaR con descuento = *d.VaR sin descuento* + (1 - *d*)

En el caso paramétrico (Normal) que antes analizamos se debe cumplir también esta relación. En efecto, en tal contexto, el VaR correctamente calculado sobre rentabilidades descontadas sería:

$$VaR_d = \Phi^{-1}(1-p)\sigma_d - \mu_d$$

donde μ_d, σ_d son la esperanza matemática y varianza de la rentabilidad descontada. Como $\mu_d = (1+\mu)d - 1$, y $\sigma_d = d\sigma$, tenemos:

$$VaR_d = \Phi^{-1}(1-p)d\sigma - [(1+\mu)d - 1] = d[\Phi^{-1}(1-p).\sigma - \mu] + (1-d) = d.VaR + (1-d)$$

como acabamos de probar en el caso general.

Remark 1 *En el caso particular de que la rentabilidad descontada esperada* μ_d *fuese igual a cero, es decir, que* $\mu = R_F$*, calcularíamos el VaR mediante* $VaR_d = \Phi^{-1}(1-p)\sigma_d$*, y entonces tendríamos:*

$$VaR_d^{\mu_d=0} = \Phi^{-1}(1-p)\sigma_d = d.\Phi^{-1}(1-p)\sigma = d.VaR^{\mu=0}$$

que es correcto como igualdad algebraica, pero carece de interpretación, puesto que compara dos situaciones diferentes: una con rentabilidad descontada igual a cero, y otra con rentabilidad sin descontar igual a cero. Sería erroneo, por tanto, concluir de esta expresión que cuando la rentabilidad descontada es igual a cero, el VaR se puede obtener descontando el VaR estimado con las rentabilidades sin descontar. En dicha situación de mercado, con una rentabilidad descontada igual a cero, el VaR estimado con rentabilidades sin descontar sería el VaR correspondiente a una rentabilidad igual a la del activo sin riesgo:

$$VaR^{\mu=R_F} = \Phi^{-1}(1-p)\sigma - \mu = \Phi^{-1}(1-p)\sigma - R_F$$

y la relación correcta sería:

$$VaR_d^{\mu_d=0} = \Phi^{-1}(1-p)\sigma_d = d\left[VaR^{\mu=R_F} + R_F\right]$$

como se comprueba al final de la hoja de cálculo del ejercicio EIV.1.6.

En los ejercicios [EIV.1.6] y [EIV.2.11] analizamos el impacto numérico del descuento sobre el cálculo del VaR en el caso de una distribución Normal.

Ejemplos:

- Si la rentabilidad descontada tiene una esperanza matemática nula, $\mu_d = 0$ y una volatilidad: $\sigma_d = 9\%$, tendríamos para $p = 1\%$, y un horizonte anual: $VaR_d = (0,09)(2,33) = 0,21$ ó $21,00\%$. Supongamos que el tipo de interés sin riesgo es 4%, lo que implica un factor de descuento: $d = 1/1,04$. La volatilidad de la rentabilidad si no tenemos en cuenta el descuento es: $d^{-1}\sigma_d = (1.04)(0.09)$, y la rentabilidad sin descontar, 4%. Por tanto, para $p = 1\%$, y un horizonte anual, utilizando las rentabilidades sin descontar habríamos calculado (incorrectamente) : $VaR = (0.09)(1.04)(2.33) - 0.04 = 0.178\,1$ ó 17,81%, frente al 21,00% correcto.

- Si la rentabilidad descontada esperada fuese del 3,5%, entonces: $VaR_d = (0.09)(2.33) - 0.035 = 0.174\,7$ ó 17,47%. La rentabilidad anual sin descontar habria sido: $\mu = \frac{1+\mu_d}{d} - 1 = (1.035)(1.04) - 1 = 0.076\,4$, o 7,64%, por lo que (incorrectamente) habríamos calculado el VaR: $VaR = (1.04)(0.09)(2.33) - 0.0764 = 0.141\,69$ ó 14,17%, frente al 17,47% correcto.

- En una cartera de 1 millón de euros cuya rentabilidad descontada tuviese un valor esperado de 3.5% anual y una volatilidad de 9%, tendríamos a 3 meses un VaR nominal al 1% respecto a la rentabilidad media de: $VaR_d = 10^6\left((0.09)(2.33)\sqrt{0.25} - (0.035)(0.25)\right) = 96100$ euros. Con rentabilidades sin descontar, habríamos observado una volatilidad $d^{-1}\sigma$ y una rentabilidad anual: $\frac{1+\mu_d}{d} - 1 = 0.076\,4$, por lo que habriamos estimado (incorrectamente) un VaR a 3 meses: $VaR = 10^6\left((1.04)(0.09)(2.33)\sqrt{0.25} - (0.076\,4)(0.25)\right) =$ 8161.0 euros.

- En el ejemplo anterior, el VaR (correcto) respecto al origen sería: $VaR_d = 10^6(0.09)(2.33)\sqrt{0.25} = 104850$ euros, siempre más conservador que el VaR respecto a la media, mientras que el (incorrecto) VaR respecto al origen, sin descontar, sería: $VaR = 10^6(1.04)(0.09)(2.33)\sqrt{0.25} = 109040$ euros.

Debe observarse que en este enfoque paramétrico estamos suponiendo que la distribución de rentabilidades y los parámetros que en ella aparecen permanecen invariantes en el tiempo. En muchos b casos, querremos evitar tal supuesto, y deberemos utilizar momentos cambiantes en el tiempo.

1.5 Limitaciones del VaR

- El cálculo del VaR no entra en consideraciones sobre cuál pueda ser la pérdida esperada en caso de que la rentabilidad del activo o cartera caigan por encima del nivel indicado por el VaR. Este es el objetivo de otras medidas de riesgo, como el *Expected Shortfall*, o la *Pérdida Esperada* en la cola de la distribución, que analizamos más adelante.También proporcionan informacion útil sobre el riesgo de pérdidas los *Lower Partial Moments* que vimos en la primera sección.

- Es necesario siempre pensar acerca del tipo de pregunta a la que se quiere responder con la estimación de una medida de riesgo. Por ejemplo, en ocasiones interesa estimar la magnitud de la pérdida esperada en relación a un benchmark, lo que da lugar al uso del denominado *Benchmark VaR*. En tales casos, también es interesante estimar el *Expected Shortfall* respecto del benchmark, o pérdida esperada en caso de que la rentabilidad de la cartera esté por debajo de la rentabilidad del benchmark en una cuantía superior al *Benchmark VaR*.

- Cuando el VaR se calcula a un horizonte determinado, por ejemplo, en un mes, se está suponiendo que la composición de la cartera va a quedar inalterada durante dicho período, lo cual no es muy razonable.

- Asimismo, se supone que la estructura de la matriz de covarianzas de los activos que componen la cartera es invariante a lo largo del horizonte temporal de cálculo del VaR. Cuando no es asi, es preciso reconstruir históricamente el precio de la cartera cada vez que se cambia su composición, para modelizar su varianza. Alternativamente, hemos de modelizar la volatilidad de los activos individuales que pueden entrar a formar parte de nuestra cartera. Sin embargo, precisamente en periodos de crisis pueden incrementar drásticamente las correlaciones entre activos, incluso cambiando de signo si eran negativas en periodos normales de mercado.

- Disponer de un buen modelo de predicción de la varianza es crucial para el cálculo del VaR. Si, por no disponer de tal modelo, se utiliza en el cálculo del VaR una varianza histórica, puede ser conveniente no utilizar la expresión habitual que pondera todas las observaciones por igual y está contaminada por valores extremos, extendiendo hacia el futuro dicho efecto. La utilización de un esquema EWMA para el cálculo de la varianza puede ser preferible.

- No es muy evidente en algunos casos como seleccionar el horizonte de cálculo o el umbral de probabilidad para el cálculo del VaR.

- No tiene en cuenta el riesgo de liquidez, que puede ser importante en alguns mercados

1.6 Benchmark VaR

A un gestor que sigue una estrategia activa se le exige habitualmente un benchmark no solo para sus rentabilidades, sino también para sus riesgos. El benchmark podria ser el bono soberano, o el indice de bolsa, dependiendo de la naturaleza de la cartera cuyo benchmark VaR estamos interesados en medir. La gestión activa requiere medir las rentabilidades obtenidas por la cartera en relación con las obtenidas por un activo benchmark. Se define como *Rentabilidad Activa* la diferencia entre la rentabilidad de la cartera y la rentabilidad del benchmark. El *Tracking Error* es la volatilidad de la rentabilidad activa, es decir, la volatilidad de las diferencias de rentabilidad entre la cartera y su objetivo.

En términos porcentuales, el Benchmark VaR es el p-cuantil de la distribución de la rentabilidad activa a h dias, descontada hasta hoy, t. En términos nominales el Benchmark VaR, BVaR, se estima:

$$BVaR = \left(\Phi^{-1}(1-p)\sigma_{RA}\sqrt{\frac{h}{250}} - \mu_{RA}\frac{h}{250} \right) W$$

donde σ_{RA} denota el *tracking error* anual (desviación típica de la discrepancia entre la rentabilidad del fondo y su objetivo de gestión) y μ_{RA} es la Rentabilidad Activa anual (rentabilidad esperada o rentabilidad objetivo), siendo W la cantidad invertida, y h el horizonte de gestión, en días. Esta expresión es la habitual para el cálculo del VaR, sólo que utilizando ahora el tracking error como medida de volatilidad.

Nótese que las fluctuaciones de rentabilidad de la cartera alrededor del objetivo entran en el cálculo del BVaR a través del tracking error, pero el Benchmark VaR no depende del nivel de rentabilidad que se toma como objetivo o benchmark, sino solo de las diferencias alrededor del mismo.

En comparacion con el tracking error, el Benchmark VaR tiene dos ventajas: 1) el tracking error mide las desviaciones respecto del benchmark, no importa si es con una rentabilidad superior o inferior al mismo; por el contrario, el benchmark VaR mide el riesgo de tener un resultado peor que el benchmark, 2) la rentabilidad activa esperada afecta al benchmark VaR, mientras que el tracking error no guarda relación con la rentabilidad activa esperada del fondo (por ejemplo, si la cartera generase una rentabilidad sistemáticamente inferior a la del activo benchmark en 3 puntos, su tracking error sería cero).

Ejemplo [EIV.1.9] *: Supongamos un fondo de 10 millones de euros que tiene una rentabilidad activa esperada igual al tipo de interés sin riesgo, y un tracking error del 3%. Calcule su Benchmark VaR al 1% sobre un horizonte de 1 año, suponiendo que las rentabilidades de dicho fondo siguen una distribución Normal.*

R: Como la Rentabilidad Activa esperada es igual al tipo de interés sin riesgo, la rentabilidad activa descontada tiene en este caso esperanza matemática igual a cero: $\mu = 0$*. El VaR solicitado es:* $BVaR = [\Phi^{-1}(0,99)(0,03)\sqrt{1}] = 6.98\%$*, y multiplicada por el valor de la cartera, proporciona un Benchmark VaR nominal de* 697.904 *euros. Es decir, en el horizonte de gestión de 1 año, con probabilidad del 99%, la rentabilidad del fondo no será inferior a la rentabilidad del benchmark en más de 697.904 euros.*

La hoja de cálculo del ejercicio [EIV.1.9] muestra, además, que la rentabilidad activa esperada tiene un efecto lineal sobre el benchmark VaR. A mayor rentabilidad objetivo o rentabilidad esperada μ, menor Benchmark VaR. Si se espera que la cartera ofrezca mejor resultado que el activo benchmark, el VaR de la cartera será relativamente pequeño, ocurriendo lo contrario si se espera que la cartera proporcione una rentabilidad inferior a la del activo benchmark.

1.7 VaR condicional: Expected Tail Loss y Expected Shortfall

El VaR proporciona información acerca del número de pérdidas que puede exceder de dicho nivel, pero no acerca de su cuantía. Sin embargo, dicha magnitud es muy importante en la gestión de riesgos, pues es la que, en definitiva, puede determinar el resultado de la gestión de cartera. De hecho, un mismo VaR al 1%, por ejemplo, puede ser compatible con perfiles muy diferentes en la cola de la densidad. En realidad, querríamos tener información acerca de toda la cola de la distribución, pero eso tampoco sería útil.

Como solución intermedia, existen dos medidas de VaR condicional, dependiendo de que midamos el VaR en relación con un benchmark o no. La *Pérdida Esperada en la Cola de la Distribución* (*Expected Tail Loss, ETL, o también TailVaR*) al 100 p% es el VaR condicional definido por:

$$ETL_h^p = -E\left(R_h \mid R_h < -VaR_h^p\right) W$$

donde R_h denota la rentabilidad *descontada* de la cartera a h dias y W es su valor actual.

Tomando como referencia el horizonte de un día, tenemos:

$$ES_{t+1}^p = -E_t\left[R_{t+1} \mid R_{t+1} < -VaR_{t+1}^p\right]$$

medido en términos de rentabilidades logarítmicas, no en términos nominales.

El *Expected Shortfall (ES)* es el benchmark VaR condicional definido por:

$$ES_h^p = -E\left(RA_h \mid RA_h < -BVaR_h^p\right) W$$

donde RA denota ahora la rentabilidad activa de la cartera, y $BVaR_{h,p}$ es el benchmark VaR.

Ejercicio [EIV.1.10] Supongamos que tenemos una cartera de 1000 valores diarios de $P\&L$ obtenidos por una determinada cartera que pretende replicar el índice Dow Jones. El VaR 1% será la décima mayor pérdida absoluta, el ETL 1% es la media de las 10 mayores pérdidas absolutas, el 1% benchmark VaR es la décima perdida relativa más grande, es decir, el décimo mayor valor de las diferencias de rentabilidad entre el fondo y el índice Dow Jones, y el 1% ES es el promedio de las 10 mayores pérdidas relativas.

Veamos algunas propiedades de la Pérdida Esperada.

1.7.1 VaR condicional en el modelo Normal

La ETL es una esperanza matemática condicional, es decir, es la esperanza matemática respecto de una distribución condicional. Dicha distribución es la que sigue la rentabilidad R_{t+1} condicional en que este a la izquierda del VaR, es decir, condicional en $R_{t+1} < -VaR_{t+1}^p$. En este caso, la distribución condicional es una distribución truncada en el punto $-VaR_{t+1}^p$. La función de densidad de R_{t+1} truncada en $-VaR_{t+1}^p$, $f^*(R_{t+1})$, es igual a la función de densidad ordinaria de $R_{t+1}, f(R_{t+1})$, dividida por la probabilidad de que: $R_{t+1} < -VaR_{t+1}^p$, es decir: $f^*(R_{t+1}) = f(R_{t+1})/P\left(R_{t+1} < -VaR_{t+1}^p\right)$

Por tanto:

$$ETL_{t+1}^p = -E_t\left[R_{t+1} \mid R_{t+1} < -VaR_{t+1}^p\right] = -\frac{\int_{-\infty}^{-VaR_{t+1}^p} R_{t+1} f(R_{t+1}) dR_{t+1}}{P\left(R_{t+1} < -VaR_{t+1}^p\right)}$$

En el caso de una distribución Normal,[1] tenemos:

$$ETL_{t+1}^p = -E_t\left[R_{t+1} \mid R_{t+1} < -VaR_{t+1}^p\right] = \sigma_{t+1}\frac{\phi(-VaR_{t+1}^p/\sigma_{t+1})}{\Phi(-VaR_{t+1}^p/\sigma_{t+1})}$$

donde ϕ denota la función de densidad y Φ la función de distribución de una $N(0,1)$. Pero en el caso de la Normal sabemos que: $VaR_{t+1}^p = -\sigma_{t+1}\Phi^{-1}(p)$, por lo que:

$$ETL_{t+1}^p = \sigma_{t+1}\frac{\phi(-VaR_{t+1}^p/\sigma_{t+1})}{\Phi(-VaR_{t+1}^p/\sigma_{t+1})} = \sigma_{t+1}\frac{\phi\left(\Phi^{-1}(p)\right)}{p}$$

La ratio entre *Pérdida Esperada* y *VaR* bajo Normalidad es:

$$\frac{ETL_{t+1}^p}{VaR_{t+1}^p} = -\frac{\phi\left(\Phi^{-1}(p)\right)}{p\Phi^{-1}(p)}$$

Si, por ejemplo, $p = 0.01$, tenemos: $\Phi_p^{-1} \approx -2,33$, por lo que:

$$\frac{ETL_{t+1}^p}{VaR_{t+1}^p} = -\frac{\phi\left(\Phi_p^{-1}\right)}{p\Phi_p^{-1}} = -\frac{(2\pi)^{-1/2}\exp[-(-2.33)^2/2]}{(0.01)(-2.33)} \approx 1,145$$

por lo que la Pérdida Esperada sería superior al VaR en un 14,5%. En la distribución Normal, esta ratio converge a 1 según converge p a cero.

En general, para distribuciones con cola gruesa, la ratio $\frac{ETL_{t+1}^p}{VaR_{t+1}^p}$ será superior al valor de la Normal. En la distribución de EVT (Extreme Value Theory), que veremos más adelante, cuando p tiende a 0, dicha ratio converge a:

$$\frac{ETL_{t+1}^0}{VaR_{t+1}^0} = \frac{1}{1-\xi}$$

donde ξ es un parámetro de dicha distribución que mide el grosor de las colas de la distribución de rentabilidades de modo que, cuanto más gruesa sea la cola en esta distribución EVT, mayor será ξ y, por tanto, mayor será la ratio de Pérdida Esperada a VaR.

1.8 Medidas coherentes de riesgo

Artzner et al. (1999), "*Coherent Measures of Risk*", Mathematical Finance, 9, 203-228, expusieron cuatro propiedades que es conveniente que cumpla toda medida de riesgo $\rho(V)$:

[1] $N(0,1) : \int_{-\infty}^{a} xf(x)dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} xe^{-x^2/2}dx = \frac{1}{\sqrt{2\pi}}\left[-e^{-x^2/2}\right]_{-\infty}^{a} = -\frac{1}{\sqrt{2\pi}}e^{-a^2/2} = -\phi(a)$

$N(0,\sigma^2) : \int_{-\infty}^{a} xf(x)dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{a} xe^{-x^2/2\sigma^2}dx = \frac{1}{\sigma\sqrt{2\pi}}\sigma^2\left[-e^{-x^2/2\sigma^2}\right]_{-\infty}^{a} = -\frac{1}{\sqrt{2\pi}}\sigma e^{-a^2/2\sigma^2} = -\sigma\phi(a/\sigma)$

- Monotonía: si $V_1 \leq V_2$, entonces $\rho(V_1) \leq \rho(V_2)$
- Invarianza por traslación. Si k es determinista: $\rho(V+k) = \rho(V) - k$. Si añadimos liquidez (cash) a una cartera, su riesgo debe reducirse
- Homogeneidad: $\rho(kV) = k.\rho(V)$. Si incrementamos el tamaño de la cartera por un factor k, su riesgo aumenta proporcionalmente
- Subaditividad: $\rho(V_1+V_2) \leq \rho(V_1)+\rho(V_2)$. Fusionar dos carteras no puede aumentar el riesgo conjunto.

El *VaR* incumple la última condición, mientras que el *Expected Shortfall* satisface las cuatro condiciones mencionadas.

1.9 Varianzas cambiantes en el tiempo

Las expresiones que hemos visto hasta ahora suponen que la varianza de las rentabilidades permanece constante en el tiempo, pero el análisis se extiende sin ninguna dificultad a la consideración de varianzas cambiantes en el tiempo.

Para ello, deberemos utilizar un procedimiento que permita estimar una varianza cambiante en el tiempo, ya sea mediante un esquema de ventanas móviles muestrales, o mediante modelos como el suavizado exponencial (EWMA) de RiskMetrics, o un enfoque GARCH. El VaR es el percentil de la distribución de probabilidad al término del horizonte de gestión, en $T+h$, distribución que hoy (en T) desconocemos. La estimación del VaR requiere, en definitiva, prever la evolución de la distribución de probabilidad de la rentabilidad del activo o cartera a lo largo del horizonte de inversión, entre T y $T+h$. Se necesita una predicción porque como tantos otros conceptos financieros, el VaR se refiere a la distribución prevista para las rentabilidades de la cartera durante el horizonte de inversión fijado, no al dia en que se calcula el VaR.

En el caso de rentabilidades Normales, para el cálculo del Valor en Riesgo necesitamos, además de la rentabilidad esperada en exceso, una predicción de la volatilidad de la rentabilidad del activo en el horizonte para el cual se quiere calcular el VaR. El cálculo del VaR es, precisamente, una de las razones por las que conviene disponer de un buen modelo de predicción de volatilidad. Para ello, hay varias posibilidades: una consistiría en tomar la serie temporal de varianzas estimada mediante EWMA o GARCH, y predecir entre T (hoy) y $T+h$ (el final del horizonte de gestión), y tomar el promedio de dichas predicciones; otra consistiría en tomar la última varianza observada, σ_T^2, como predicción. Una vez estimada una serie temporal de varianzas a partir de cualquiera de los procedimientos ya vistos (ventanas móviles, RiskMetrics, suavizado exponencial, GARCH) el VaR se calcularía utilizando la predicción de la varianza que surge de dicho modelo sobre el horizonte de cálculo del VaR.

Por ejemplo, recordemos que el modelo RiskMetrics, que supone que la rentabilidad diaria continua de la cartera sigue una distribución Normal: $r_t \mid F_{t-1} \sim N(\mu_t, \sigma_t^2)$, con:

$$\begin{aligned}\mu_t &= 0; \\ \sigma_t^2 &= \alpha\sigma_{t-1}^2 + (1-\alpha)r_{t-1}^2;\ 0 < \alpha < 1\end{aligned}$$

o, equivalentemente, que el logaritmo del precio: $p_t = \ln(P_t)$, obedece un proceso IGARCH(1,1) sin constante: $p_t - p_{t-1} = a_t$, con $a_t = \sigma_t\varepsilon_t$, y $\varepsilon_t \sim N(0,1)$. El valor de α suele tomarse en el intervalo $(0.90; 1)$, siendo 0,94 un valor bastante habitual, con el objeto de proyectar la varianza hacia el futuro y aplicar la expresión del VaR para el caso lineal bajo distribución Normal. En este caso, calculamos el VaR utilizando para el horizonte de riesgo $(t, t+h)$ la varianza estimada para el último dia de la muestra. Si la rentabilidad no es ruido blanco, como por ejemplo, en:

$$r_t = \beta_0 + \beta_1 r_{t-1} + u_t$$

entonces el modelo debería ajustarse: $\sigma_t^2 = \alpha\sigma_{t-1}^2 + (1-\alpha)u_{t-1}^2$, si bien RiskMetrics no considera esta posibilidad.

En otros modelos, obtener la predicción de la varianza sobre el horizonte deseado para el VaR es algo más complejo, pero se dispone de expresiones analíticas, como es el caso de los modelos de la familia GARCH. Pero este modelo implica que las rentabilidades presentan dependencia temporal, a través de sus segundos momentos, lo que no permite aplicar la regla de la raiz cuadrada para la extrapolación temporal.

Una gran ventaja del modelo EWMA es que no se precisan series temporales largas para su cálculo. Otra ventaja es el mayor peso que otorga a las observaciones más recientes. La correlación lineal puede calcularse dividiendo covarianzas por la raiz cuadrada del producto de varianzas, si los tres momentos se han estimado mediante una especificación EWMA.

[Figure IV.2.15]. [EIV.2.25]

Riskmetrics utiliza esta metodología. Riskmetrics calcula tres tipos de matrices de covarianzas:

- *Regulatoria*: una matriz de covarianzas con iguales ponderaciones, basada en los últimos 250 días,
- *Matriz diaria*: una matriz de covarianzas EWMA calculada utilizando datos diarios con $\lambda = 0.94$
- *Matriz mensual*: una matriz de covarianzas EWMA calculada utilizando datos diarios con $\lambda = 0.97$ para todos sus elementos, y multiplicando el resultado por 25.

En el ejercicio [EIV.2.26] se comparan los VaR resultantes de las dos primeras opciones.

2 El VaR de una cartera: diversificación de riesgos

Sea ω el vector de pesos relativos de una cartera: $\omega = (\omega_1, \omega_2, ..., \omega_n)$, y Σ 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 $\omega_i = x_i/W$, la varianza de una unidad monetaria invertida en la cartera es: $\sigma_c^2 = \omega' \Sigma \omega$.

Hacer un supuesto para la distribución de probabilidad de las rentabilidades, como la Normal, nos permite traducir el nivel de significación del VaR p de modo sencillo en un umbral α tal que la probabilidad de observar una pérdida superior a $-\alpha$ sea igual a p. En el resto de esta sección suponemos que la rentabilidad de cada activo de la cartera sigue una distribución Normal, lo que puede ser un supuesto válido en el caso de carteras bien diversificadas. Esto hace que la rentabilidad de la cartera, al ser una combinación lineal de las rentabilidades de los activos que la componen, siga asimismo una distribución Normal.

Suponemos asimismo un horizonte de riesgo corto, con una rentabilidad esperada igual a cero, por lo que el VaR de la cartera en términos de rentabilidades es:

$$VaR_c = \Phi^{-1}(1-p)\sigma_c = \alpha\sqrt{\omega'\Sigma\omega} \tag{4}$$

donde $\alpha = \Phi^{-1}(1-p)$, es tal que la probabilidad de tener una pérdida más elevada que $-\alpha$ es igual a p. Como se ve, tratamos la volatilidad de la rentabilidad de la cartera de igual modo a como trataríamos la volatilidad de un activo individual.

En términos monetarios, la varianza de la inversión es el producto $\sigma_c^2 W^2$ que puede expresarse:[2]

$$\sigma_c^2 W^2 = x'\Sigma x$$

por lo que el VaR de la cartera en términos nominales,

$$VaR_c = \sigma_c W \; \Phi^{-1}(1-p) = \alpha\sqrt{x'\Sigma x}$$

Estas expresiones proporcionan el *VaR diversificado* de la cartera, que tiene en cuenta los efectos de la diversificación a través de la presencia de la matriz de covarianzas en su cálculo.

En el caso de una cartera, tiene sentido calcular el VaR de la misma únicamente si la composición de la cartera va a mantenerse invariante durante el horizonte temporal para el cual se calcula el VaR. No tendría mucho sentido calcular en T (hoy) el VaR a horizonte h si en $T+h$ la actual cartera no va a existir porque ya hayamos cambiado la composición de nuestras inversiones.

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

El *VaR individual* del activo i-ésimo de la cartera, se define por analogía con la expresión del VaR para una cartera:

$$VaR_i = \alpha\sigma_i \mid W_i \mid = \alpha\sigma_i \mid \omega_i \mid W \tag{5}$$

donde admitimos posiciones cortas y largas en dicho activo. El factor $\alpha\sigma_i \mid \omega_i \mid$ sería el VaR individual en rentabilidad, mientras que la expresion anterior nos da el VaR nominal.

El *VaR no diversificado* es la suma de los VaR individuales, que será siempre superior, por ignorar los efectos de la diversificación sobre el riesgo. De hecho, los beneficios de la diversificación se observan comparando el *VaR diversificado* y el *VaR no diversificado* de una misma cartera.

Ejemplo (Jorion). Consideremos una cartera invertida en dos divisas frente al euro: dolar y yen, que suponemos incorrelacionadas. Las cantidades invertidas son 2 millones y 1 millón, respectivamente, por lo que el vector de ponderaciones es $\omega = (2/3, 1/3)$. *Las volatilidades de ambas divisas son 5% y 12%. La varianza de la cartera es, por tanto:* $\sigma_c^2 = \omega'\Sigma\omega = \omega' \begin{pmatrix} 0,05^2 & 0 \\ 0 & 0,12^2 \end{pmatrix} \omega = (2/3, 1/3) \begin{pmatrix} 0.05^2 & 0 \\ 0 & 0.12^2 \end{pmatrix} \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix} = (27.111)10^{-4}$, *y su volatilidad:* $\sigma_c = \sqrt{(27.111)10^{-4}}$ *=5,21%.*

En términos monetarios: $x'\Sigma x = (\omega'\Sigma\omega)\, W^2 = (27,11)10^{-4}9(10^{12}) = 244(10^8)$, *con una volatilidad de 156.205 euros. El VaR 95% a 1 año es (1,65)(156.205)=257.738 euros (suponemos que las rentabilidades diarias tienen media cero). Los VaR individuales (VaR no diversificado) en cada divisa son* $VaR_1 = (1.65)(0.05)2.10^6$ *=165.000 y* $VaR_2 = (1.65)(0.12)10^6$ *=198.000 euros, respectivamente, que suman 363.000 euros, un 26% más que el VaR diversificado.*

Para el cálculo del VaR 95% a 1 dia tendríamos: $x'\Sigma x = (\omega'\Sigma\omega)\, W^2 = 27.11 \times 10^{-4} \times 9 \times 10^{12}/250 = 9.759\,6 \times 10^7$, *con una volatilidad de* $\sqrt{9.759\,6 \times 10^7} = 9879.1$ *euros. El VaR a 1 dia sería:* $1.65 \times 9879.1 = 16301$ euros. Los VaR individuales a 1 dia en cada divisa son: $VaR_1 = 1.65 \times 0.05 \times 2 \times 10^6/\sqrt{250} = 10435.5$ *y* $VaR_2 = 1.65 \times 0.12 \times 2 \times 10^6/\sqrt{250} = 12522.6$ *euros, respectivamente, que suman* 22958.1 *euros, un* 40,8% *más que el VaR diversificado.*

El cálculo del VaR de una cartera requiere disponer de estimaciones de las covarianzas o de las correlaciones entre las rentabilidades de los activos que la integran, lo que ha suscitado la necesidad de generar métodos que simplifiquen la alta dimensionalidad de este problema, dado que en carteras como las de los fondos de inversión, el número de activos que pueden incluirse es excesivamente grande como para estimar todas sus varianzas y covarianzas. Los métodos factoriales que hemos examinado en un capítulo previo son muy útiles a tales efectos.

En el caso de dos activos, con $\omega_2 = 1 - \omega_1$:

$$\begin{aligned} VaR_c &= \Phi^{-1}(1-p)\sqrt{\omega_1^2 Var(R_1)+\omega_2^2 Var(R_2)+2\omega_1\omega_2 Cov(R_1,R_2)} = \\ &= \Phi^{-1}(1-p)\sqrt{\omega_1^2\sigma_1^2+\omega_2^2\sigma_2^2+2\omega_1\omega_2\sigma_1\sigma_2\rho_{12}} \end{aligned}$$

donde es fácil relacionar cada uno de los sumandos con la expresión (5) del VaR individual.

En el caso particular en que los dos activos estén incorrelacionados:

$$VaR_c = \Phi^{-1}(1-p)\sqrt{\omega_1^2\sigma_1^2+\omega_2^2\sigma_2^2} \Longrightarrow (VaR_c)^2 = (VaR_1)^2 + (VaR_2)^2$$

por lo que, cuando contamos con activos que se mueven independientemente, podemos constituir carteras de menor riesgo que cualquiera de los activos que la componen. En efecto, dados dos activos con rentabilidades independientes, supongamos, sin pérdida de generalidad, que $\sigma_1^2 < \sigma_2^2$. Sobre períodos cortos, el VaR de la cartera será inferior al VaR del activo 1 si la volatilidad de la cartera también lo es: $\sqrt{\omega_1^2\sigma_1^2+\omega_2^2\sigma_2^2} = \sqrt{\omega^2\sigma_1^2+(1-\omega)^2\sigma_2^2} < \sigma_1$, es decir, si: $\omega^2\sigma_1^2+(1-\omega)^2\sigma_2^2 < \sigma_1^2$, para lo que necesitamos: $(1-\omega)^2\sigma_2^2 < (1-\omega^2)\sigma_1^2$, o $(1-\omega)\sigma_2^2 < (1+\omega)\sigma_1^2$, es decir: $\omega > \frac{\sigma_2^2-\sigma_1^2}{\sigma_2^2+\sigma_1^2}$. Si los dos activos tuvieran una volatilidad similar, esto se conseguiría casi para cualquier valor positivo del peso ω . Si σ_2^2 fuese mucho mayor que σ_1^2 se lograría con valores numéricos de ω próximos a 1.

Si los dos activos estuviesen perfectamente y positivamente correlacionados, y los pesos son todos positivos,[3] tendríamos:

$$\begin{aligned} VaR_c &= \Phi^{-1}(1-p)\sqrt{\omega_1^2\sigma_1^2+\omega_2^2\sigma_2^2+2\omega_1\omega_2\sigma_1\sigma_2} = \Phi^{-1}(1-p)\sqrt{(\omega_1\sigma_1+\omega_2\sigma_2)^2} = \\ &= \Phi^{-1}(1-p)\,(\omega_1\sigma_1+\omega_2\sigma_2) = VaR_1 + VaR_2, \end{aligned}$$

siendo este el único caso en que el VaR de la cartera es igual a la suma de los VaR individuales de los dos activos, lo que muestra la imposibilidad de diversificar el riesgo con activos perfectamente correlacionados.

En general, dado que:

$$\sqrt{\omega_1^2\sigma_1^2+\omega_2^2\sigma_2^2+2\omega_1\omega_2\rho_{12}\sigma_1\sigma_2} < \omega_1\sigma_1+\omega_2\sigma_2 \equiv \sqrt{\omega_1^2\sigma_1^2+\omega_2^2\sigma_2^2+2\omega_1\omega_2\sigma_1\sigma_2}$$

tendremos que:

$$VaR_c < VaR_1 + VaR_2$$

[3] De modo que $|\omega_i| = \omega_i, i = 1, 2, ..., n$. [Jorion]

por lo que el VaR de una cartera será menor que la suma de los VaR individuales, lo que ilustra las ventajas de la diversificación.

Ejemplo (Peña): Consideremos una cartera que mantiene una posición corta de -767$ en dolar canadiense, y posiciones largas de 117$ en dólares USA y 108$ en yen japonés, con matriz de correlaciones:

$$M = \begin{pmatrix} 1 & -0,21 & -0,21 \\ -0,21 & 1 & 0,79 \\ -0,21 & 0,79 & 1 \end{pmatrix}$$

Las volatilidades de los 3 activos son: 5,54%, 12,82% y 16,63%. Calcule los VaR mensuales de las posiciones individuales (R: 20,24$, 7,14$, 8,55$) y el VaR 95% de la cartera (R: 27,6$), que es inferior a la suma de los VaR individuales, debido a las ganancias que se producen por la diversificación de la inversión.

VaR individuales:

$$\begin{aligned} VaR(Canada) &: \quad 767 \times 0.0554 \times 1.65/\sqrt{12} = 20.239 \\ VaR(US\$) &: \quad 117 \times 0.1282 \times 1.65/\sqrt{12} = 7.1444 \\ VaR(yenes) &: \quad 108 \times 0.1663 \times 1.65/\sqrt{12} = 8.5548 \end{aligned}$$

Matriz de covarianzas:

$$\begin{pmatrix} 0.0554 & 0 & 0 \\ 0 & 0.1282 & 0 \\ 0 & 0 & 0.1663 \end{pmatrix} \begin{pmatrix} 1 & -0.21 & -0.21 \\ -0.21 & 1 & 0.79 \\ -0.21 & 0.79 & 1 \end{pmatrix} \begin{pmatrix} 0.0554 & 0 & 0 \\ 0 & 0.1282 & 0 \\ 0 & 0 & 0.1663 \end{pmatrix}$$
$$= \begin{pmatrix} 3.0692 \times 10^{-3} & -1.4915 \times 10^{-3} & -1.9347 \times 10^{-3} \\ -1.4915 \times 10^{-3} & 1.6435 \times 10^{-2} & 1.6843 \times 10^{-2} \\ -1.9347 \times 10^{-3} & 1.6843 \times 10^{-2} & 2.7656 \times 10^{-2} \end{pmatrix}$$

Varianza de la cartera:

$$\begin{pmatrix} -767 & 117 & 108 \end{pmatrix} \begin{pmatrix} 3.0692 \times 10^{-3} & -1.4915 \times 10^{-3} & -1.9347 \times 10^{-3} \\ -1.4915 \times 10^{-3} & 1.6435 \times 10^{-2} & 1.6843 \times 10^{-2} \\ -1.9347 \times 10^{-3} & 1.6843 \times 10^{-2} & 2.7656 \times 10^{-2} \end{pmatrix} \begin{pmatrix} -767 \\ 117 \\ 108 \end{pmatrix} = 3367.0$$

VaR de la cartera a 1 mes:

$$VaR = 1.65\sqrt{3367.0}/\sqrt{12} = 27.639$$

2.1 Descomposición del VaR

Inicialmente, el VaR fue concebido como un indicador del nivel de riesgo asumido en una determinada posición. Pero gradualmente, resultó evidente que era asimismo un instrumento que podía utilizarse para la gestión del riesgo en carteras de activos financieros. Para ello, resulta fundamental la capacidad de agregar y desagregar el VaR en distintos componentes. La agregación de VaR es la herramienta fundamental en el llamado capital budgeting, que consiste en asignar el capital económico entre actividades, la asignación de límites de VaR para los *traders*, o la estimación del requisito de capital regulatorio. La desagregación de VaR ayuda al analista de riesgos a entender cuáles son las pricipales fuentes de riesgo en su cartera, determinar que elementos deben cubrirse, qué limites imponer a los *traders*, o qué riesgos pueden derivarse de una nueva inversión.

En las secciones que siguen, utilizaremos en ocasiones distintos procedimientos estadísticos para tratar de caracterizar las principales fuentes de riesgo que pueden afectar al valor de una cartera. Se trata de una estrategia de reducción de dimensionalidad porque, en general, toda cartera se ve sometida a muchas más influencias de las que un gestor puede controlar, por lo que es importante poder seleccionar un número reducido de ellas que expliquen un alto porcentaje de las fluctuaciones posibles en el valor de la cartera. Según el método de selección de factores utilizado, llegaremos a un modelo factorial de uno u otro tipo (método de regresión, componentes principales, etc,...).

Tengamos en cuenta que el *VaR total* de una cartera podría calcularse a partir de una serie temporal de rentabilidades de la misma. Tendríamos que tener en cuenta, sin embargo, si al generar dichos datos se ha mantenido constante la composición de la cartera o no.

- *VaR sistematico y VaR especifico*: Una vez que se dispone de un modelo factorial, que explica el riesgo de una cartera en funcion de un conjunto de factores previamente seleccionados, el VaR total puede descomponerse entre un componente de *VaR sistemático*, el componente del VaR total que está explicado por los factores de riesgo previamente identificados, y un *VaR específico o idiosincrático*, que es el no explicado por ellos, es decir, el residual en la estimación del modelo factorial. En el modelo lineal Normal, el VaR total es igual a la raiz cuadrada de la suma del VaR sistematico, al cuadrado, y el VaR especifico, también al cuadrado. Es decir, incluso si el modelo de factores es lineal, el VaR total no será igual a la suma de los VaR sistemático y específico; ello sólo sucede cuando ambos componentes, *VaR sistemático y VaR específico*, están perfectamente correlacionados. El componente de VaR sistemático a un determinado horizonte puede calcularse a partir de la extrapolación temporal de la matriz de varianzas y covarianzas de los factores si las rentabilidades de todos los factores son i., i.d. y son condicionalmente homocedasticos. Si la rentabilidad esperada en exceso es nula, el propio VaR puede extrapolarse utilizando el factor habitualmente usado para las varianzas.

- *VaR individual* o *Stand-alone VaR*: Una desagregación alternativa del VaR total es la que vimos en la sección anterior, a través del VaR asociado a cada uno de los activos que integran la cartera o de cada una de las clases de factores considerados (componentes *stand-alone*). Así, tendriamos el

VaR equity de una cartera, el VaR de tipos de interés, el VaR de tipo de cambio o *forex*, el VaR commodity, etc. (El oro se trata habitualmente como riesgo *Forex*, no como riesgo commodity). Para calcular el componente stand-alone de un determinado factor de riesgo en el VaR total, se fijan las sensibilidades de todos los demas factores a cero. Esto es importante cuando se negocia en distintas mesas, pues ninguna de ellas deberia verse beneficiada ni perjudicada por las posibles correlaciones entre los factores que afectan a la cartera global. Es como si nos preguntamos qué impacto tiene sobre la rentabilidad del activo la variación en un factor, dejando los otros factores inalterados. La suma de los VaR stand-alone no es igual al VaR sistemático total, excepto si las carteras son lineales, las rentabilidades siguen una distribución Normal y los distintos componentes están perfectamente correlacionados lo cual, por otra parte, indicaría una mala elección de los factores de riesgo. Este es el resultado que vimos al final de la sección anterior. En general, el VaR total es inferior a la suma de los VaR individuales (*stand-alone*), lo que se conoce como propiedad de sub-aditividad del VaR. Matemáticamente, la razón es que la varianza no es un operador lineal. La subaditividad genera un problema para el *risk budgeting*, pues las carteras individuales podrian estar dentro de los limites de riesgo asumibles, y sin embargo la suma de los componentes VaR podría exceder el umbral trazado para el VaR en términos de riesgo admisible, haciendo que la cartera global no fuese admisible.

- *VaR marginal y VaR incremental*: El VaR marginal mide la sensibilidad del VaR total a los parámetros del modelo de riesgo factorial. El *VaR marginal* de cada activo mide el cambio que se produce en el VaR de la cartera cuando varían las sensibilidades a los factores de riesgo. En general, el VaR es función de un vector de parámetros θ :

$$VaR = f(\theta)$$

cuyo vector gradiente es:

$$g(\theta) = \left(\frac{\partial f}{\partial \theta_1}, \frac{\partial f}{\partial \theta_2}, ..., \frac{\partial f}{\partial \theta_n} \right)'$$

por lo que la aproximación de primer orden del VaR en serie de Taylor alrededor del vector θ_0 conduce a:

$$f(\theta_1) = f(\theta_0) + (\theta_1 - \theta_0)' g(\theta) = f(\theta_0) + \sum_{i=1}^{n} (\theta_{i1} - \theta_{i0})' \frac{\partial f}{\partial \theta_i} \tag{6}$$

En el caso particular en que $\theta_0 = 0$, tendremos:

$$f(\theta) \approx \theta' g(\theta) = \sum_{i=1}^{n} \theta_i \frac{\partial f}{\partial \theta_i} \tag{7}$$

una expresión que descompone el VaR total en suma de las contribuciones de los *VaR marginales* de cada factor.

Si consideramos la composición de una cartera invertida en un conjunto de activos, una posibilidad es considerar el vector de ponderaciones ω como vector θ. Esto es lo que hacemos en la sección siguiente. En este caso, el vector gradiente del VaR, interpretado como función dependiente de un vector de parámetros, puede utilizarse para aproximar el impacto sobre el VaR de un pequeño *trade,* así como el efecto que tendría el cambio en el limite máximo de VaR permisible para un *trader.* Estaríamos entonces pensando en pasar de una cartera, caracterizada por un vector θ_0 a otra, caracterizada por el vector θ_1, como se hace en (6). El cambio en VaR, $f(\theta_1) - f(\theta_0)$, se conoce como *VaR incremental.*

Cuando la expresión (7) es aplicable, el producto $\theta_i f_i(\theta)$ nos proporciona los componentes del VaR, que constituyen una descomposición aditiva del VaR, que es exacta en modelos lineales del VaR. No es una descomposición exacta en carteras con pagos no lineales (por ejemplo, cuando se incluyen opciones), o cuando se calcula el VaR por otros procedimientos. Por ejemplo, cuando se estima el VaR por simulación, la suma de los VaR marginales es sólo aproximadamente igual al VaR total.

Más generalmente, cuando consideremos carteras con un amplio número de activos, el vector θ estará formado por las sensibilidades del valor de la cartera a los diferentes factores de riesgo, como veremos exhaustivamente en los ejemplos que siguen.

2.2 VaR marginal y VaR por componentes en una cartera

Cada una de las descomposiciones del VaR que hemos descrito trata de responder a diferentes cuestiones, y es importante entender bien estas diferencias. Aunque el VaR fue introducido como indicador del nivel de riesgo de una posición, pronto resultó evidente que podía asimismo utilizarse para gestionar eficientemente el riesgo de una cartera. Así, una pregunta que cabe hacerse es: ¿qué posición de nuestra cartera deberíamos modificar para reducir el VaR de la cartera de modo más eficaz?

En la estimación del VaR marginal hemos de tener en cuenta que la variación en una posición determinada tiene un efecto directo sobre el VaR de la cartera, pero también efectos indirectos, a través de las correlaciones entre la rentabilidad de dicho activo y las rentabilidades de los demás activos de la cartera.

En lo sucesivo, en ocasiones consideraremos el VaR marginal en términos porcentuales, y en otras ocasiones, en términos nominales, es decir, en unidades monetarias. La diferencia entre ambos estriba en que el VaR nominal es igual al VaR porcentual multiplicado por el valor absoluto de la cuantía de la posición tomada en dicho activo, si el intervalo de tiempo es breve y podemos suponer una rentabilidad igual a cero.[4]

[4] Algnos autores (como Jorion, "Value at Risk") reservan la denominación de VaR marginal para su expresión porcentual, denominando Componente VaR al VaR marginal expresado en términos nominales.

Una posible interpretación del VaR marginal surge al considerar como vector de parámetros θ el vector de ponderaciones ω, que es el caso que consideramos en esta sección. Diferenciando en la ecuación de la varianza de la cartera, tenemos:

$$\frac{\partial \sigma_c^2}{\partial \omega_i} = 2\omega_i \sigma_i^2 + 2 \sum_{j=1, j \neq i}^{n} \omega_j \sigma_{ij} = 2Cov\left(R_i, \omega_i R_i + \sum_{j \neq i}^{n} \omega_j R_j \right) = 2Cov(R_i, R_c)$$

donde R_i, R_c denotan las rentabilidades del activo $i-ésimo$ y de la cartera, respectivamente.

La derivada de la volatilidad se obtiene utilizando: $\frac{\partial \sigma_c^2}{\partial \omega_i} = 2\sigma_c \frac{\partial \sigma_c}{\partial \omega_i}$, por lo que tenemos:

$$\frac{\partial \sigma_c}{\partial \omega_i} = \frac{Cov(R_i, R_c)}{\sigma_c}$$

Para transformar esta medida en un VaR, definimos el VaR marginal del activo i como el cambio en el VaR de la cartera cuando cambia la ponderacion del activo. Si el VaR de la cartera está en términos nominales habrá que dividir dicha derivada por el nominal invertido en la cartera:

$$\Delta VaR_i = \frac{1}{W} \frac{\partial VaR_c}{\partial \omega_i} = \alpha \frac{\partial \sigma_c}{\partial \omega_i} = \alpha \frac{Cov(R_i, R_c)}{\sigma_c} \tag{8}$$

que carece de unidades, al ser un cociente de cantidades nominales.

Proposition 2 *Tenemos la relación entre VaR marginal y betas:*

$$\Delta VaR_i = \alpha \beta_i \sigma_c = \frac{VaR_c}{W} \beta_i$$

por lo que *el valor relativo de los VaR marginales de dos activos de la cartera es igual al cociente de sus betas.*

Proof. Recordemos que la beta del activo i respecto de la cartera se define: ■

$$\beta_i = \frac{Cov(R_i, R_c)}{\sigma_c^2} = \rho_{ic} \frac{\sigma_i}{\sigma_c},$$

y mide el riesgo sistemático del activo i en relación con la cartera, siendo el resto del riesgo del activo debido a otros componentes no incluidos en la composición de la cartera. Por otra parte, el vaR de la cartera es: $VaR_c = \alpha \sigma_c W$. Sustituyendo en (8) se obtiene el resultado.

Proposition 3 *El VaR marginal de cada activo es igual al VaR de dicho activo multiplicado por el coeficiente de correlación entre las rentabilidades de dicho activo y la cartera.*

Proof. La beta del activo i respecto de la cartera tambien puede expresarse: $\beta_i = \rho_{ic}\frac{\sigma_i}{\sigma_c}$, de modo que, en términos porcentuales: ■

$$\Delta VaR_i = \alpha\beta_i\sigma_c = \alpha\rho_{ic}\sigma_i = VaR_i\rho_{ic}$$

Si un inversor quiere reducir el VaR de su cartera, reduciendo el importe invertido en una cierta cuantía, debe calcular el VaR marginal de cada activo y reducir la posición en la cuantía deseada en aquel activo de la cartera que tenga el mayor VaR marginal.

Para la gestión del riesgo, es útil tener una descomposición del VaR que nos permita conocer la contribución al riesgo de cada activo de la cartera. Para esto no sirve el VaR individual, pues su suma no coincide con el VaR de la cartera, por ignorar la diversificación. En cambio, consideremos el *VaR marginal en términos nominales*, que algunos autores (Jorion, *Value at Risk,* McGraw-Hill) denominan *VaR por componentes*:

$$CVaR_i = (\Delta VaR_i)\, w_i W = VaR_c\beta_i w_i$$

cuya relación con el VaR individual es:

$$\begin{aligned} CVaR_i &= (VaR_c)\,\beta_i w_i = (\alpha\sigma_c W)\,\beta_i w_i = (\alpha\sigma_c W)\,\frac{\rho_i\sigma_i}{\sigma_c} w_i = \\ &= \alpha W\rho_i\sigma_i w_i = VaR_i\rho_i \end{aligned}$$

Si agregamos los VaR por componentes, tenemos:

$$\sum_{i=1}^{N} CVaR_i = VaR_c\left(\sum_{i=1}^{N}\omega_i\beta_i\right) = VaR_c$$

ya que la suma que aparece dentro del paréntesis no es sino la beta de la cartera con respecto a sí misma, que es lógicamente, igual a uno.[5]

2.3 VaR incremental

También podemos estar interesados en el impacto total que tendrá una determinada operación financiera sobre el riesgo de nuestra cartera. Representamos dicha operación por un vector b que recoge los cambios que se producirían con dicha operación en las exposiciones a cada activo. El vector b puede representar un cambio en la posición en un sólo activo o en varios, quizá todos, de los activos de la cartera. En este caso estamos considerando cambios en las posiciones que pueden ser de cuantía importante, lo que hace que el VaR marginal no pueda utilizarse para responder a esta pregunta, por ser una aproximación válida solo para variaciones muy pequeñas.

[5] $\sigma_c^2 = w_1 cov(R_1, R_c) + ... + w_N cov(R_N, R_c) = w_1(\beta_1\sigma_c^2) + ... + w_N(\beta_N\sigma_c^2) = \sigma_c^2(\Sigma_{i=1}^N w_i\beta_i)$, por lo que, necesariamente: $\Sigma_{i=1}^N w_i\beta_i$.

La idea sería calcular el VaR_{c+b} que se tendría con la nueva posición y compararlo con el VaR de la posición antigua para tener el VaR incremental:

$$VaR\, incremental = VaR_{c+b} - VaR_c$$

Si el VaR se reduce, decimos que la operación representa una cobertura. El problema con este enfoque es que requiere una nueva evaluación del VaR con cada nueva operación, lo que puede resultar impracticable por el elevado número de activos o por el alto número de operaciones realizadas a lo largo del dia. Se puede utilizar una aproximación mediante un desarrollo en serie alrededor de la cartera original:

$$VaR_{c+b} = VaR_c + (\Delta VaR)'\, b + ...$$

de modo que el VaR incremental, siempre que la operacion propuesta sea relativamente pequeña (b pequeño) se aproximaría, utilizando (8) por:

$$VaR\, incremental = (\Delta VaR)'\, b = \alpha \frac{\{Cov(R_i, R_c)\}_{i=1}^n}{\sigma_c} b$$

donde debemos observar que tanto $\{Cov(R_i, R_c)\}_{i=1}^n$ como b son vectores de dimensión igual al número de activos de la cartera

Best hedge

Supongamos ahora que la nueva operación implica variar la posición únicamente en uno de los activos i que configuran la cartera. El valor de la cartera es ahora $W_N = W + b$, siendo b la cantidad invertida en dicho activo. La varianza nominal de la nueva cartera puede escribirse:

$$\sigma_N^2 W_N^2 = \sigma_c^2 W^2 + b^2 \sigma_i^2 + 2bW\sigma_{ic}$$

y un gestor de cartera puede plantearse cuál será el cambio en la posición de su cartera que genere una mayor reducción del nivel de riesgo. Para ello, derivamos $\sigma_N^2 W_N^2$ respecto de b , igualamos a cero y resolvemos, obteniendo:

$$b^* = -W \frac{\sigma_{ic}}{\sigma_i^2} = -W \beta_i \frac{\sigma_c^2}{\sigma_i^2}$$

una operación que se conoce como el *best hedge* y que nos proporciona la cuantía en que debemos reducir la exposición en el activo i para lograr la mayor reducción posible en el VaR. Comparando el resultado obtenido con distintos activos podemos saber cuál de las posiciones deberíamos variar.

En el siguiente ejercicio vamos a utilizar el hecho de que en notación matricial, el vector de betas de los n activos que configuran la cartera puede obtenerse:

$$\beta = \frac{\Sigma \omega}{\omega' \Sigma \omega}$$

Example 4 *(Jorion): Con los datos de las dos divisas del ejemplo anterior, consideramos aumentar la posición en dolares en 10.000 euros. Las betas son:*[6]

$$\begin{aligned}
\Sigma\omega &= \begin{pmatrix} 0.05^2 & 0 \\ 0 & 0.12^2 \end{pmatrix} \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix} = \begin{pmatrix} 1.6667 \times 10^{-3} \\ 0.0048 \end{pmatrix} \\
\sigma_c^2 &= \omega' \Sigma\omega = \begin{pmatrix} 2/3 & 1/3 \end{pmatrix} \begin{pmatrix} 1.6667 \times 10^{-3} \\ 0.0048 \end{pmatrix} = 2.7111 \times 10^{-3} \\
\beta &= \begin{pmatrix} 1.6667 \times 10^{-3} \\ 0.0048 \end{pmatrix} \frac{1}{2.7111 \times 10^{-3}} = \begin{pmatrix} 0.61477 \\ 1.7705 \end{pmatrix}
\end{aligned}$$

por lo que el VaR marginal de ambas divisas, en términos porcentuales, es:

$$\Delta VaR = -\alpha\beta\sigma_c = -\alpha \frac{\Sigma\omega}{\omega'\Sigma\omega}\sigma_c = (1.65) \begin{pmatrix} 0.61477 \\ 1.7705 \end{pmatrix} \left(\sqrt{2.7111 \times 10^{-3}}\right) = \begin{pmatrix} 5.2816 \times 10^{-2} \\ 0.15211 \end{pmatrix}$$

y el VaR marginal en términos nominales, o VaR por componentes:

$$\begin{pmatrix} CVaR_1 \\ CVaR_2 \end{pmatrix} = \begin{pmatrix} (0,0528)2.10^6 \\ (0,1521)10^6 \end{pmatrix} = \begin{pmatrix} 105.630 \text{ } euros \\ 152.108 \text{ } euros \end{pmatrix}$$

cuya suma es igual al VaR total de la cartera. El 41% del riesgo, medido por el VaR, se debe a la primera divisa, y el 59% se deba a la posición en la segunda divisa.

Por tanto, s*i aumentamos la posición en $US en 10.000 euros, tenemos un VaR incremental:* $(\Delta VaR)' b = (0.0528\ 0.1521) \begin{pmatrix} 10000 \\ 0 \end{pmatrix} = 528\ euros$

Calculemos ahora exactamente el aumento que se produciría en el VaR de la cartera, a partir de la completa evaluación del riesgo de la cartera, tenemos:

$$\begin{aligned}
\sigma_{c+a}^2 &= \begin{pmatrix} 2.01 & 1 \end{pmatrix} \begin{pmatrix} 0.05^2 & 0 \\ 0 & 0.12^2 \end{pmatrix} \begin{pmatrix} 2.01 \\ 1 \end{pmatrix} = 0.0245 \\
&\Rightarrow \quad VaR_{c+a} = (1.65)\sqrt{0.0245}10^6 = 258.267 \text{ } euros
\end{aligned}$$

[6] La misma estimación del VaR marginal puede hacerse trabajando con el vector x de posiciones nominales en los distintos activos, en vez del vector w (ponderaciones de la cartera). Notese, sin embargo, que las betas son diferentes:

$$\beta = \frac{Cov(R_i, R_c)}{\sigma_c^2} = \frac{1}{0.156^2} \begin{pmatrix} 0.05^2 & 0 \\ 0 & 0.12^2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0,205 \\ 0,592 \end{pmatrix}$$

por lo que el VaR marginal:

$$\Delta VaR_i = -\alpha\beta\sigma_c = (1.65) \begin{pmatrix} 0.205 \\ 0.592 \end{pmatrix} (0.156) = \begin{pmatrix} 5.2767 \times 10^{-2} \\ 0.15238 \end{pmatrix}$$

Si comparamos con el VaR inicial, que era de 257.738 euros, tenemos un incremento de 529 euros, similar al calculado por el método aproximado.

Un incremento de la posicion en yenes en 10.000 euros el VaR aumentaría en:

$$(\Delta VaR)' \, b = (0.0528 \; 0.1521) \begin{pmatrix} 0 \\ 10000 \end{pmatrix} = 1.521 \; euros$$

El VaR sería ahora:

$$\sigma^2_{c+a} = \begin{pmatrix} 2 & 1.01 \end{pmatrix} \begin{pmatrix} 0.05^2 & 0 \\ 0 & 0.12^2 \end{pmatrix} \begin{pmatrix} 2 \\ 1.01 \end{pmatrix} = 2.4689 \times 10^{-2}$$
$$\Rightarrow \quad VaR_{c+a} = (1.65)\sqrt{2.4689 \times 10^{-2}}10^6 = 259.260 \; euros$$

con un aumento en VaR de la cartera de: $259.260 \; euros - 257.738 \; euros = 1.422$ *euros, muy similar a la estimación obtenida con el VaR marginal.*

Si eliminásemos la posición en yenes, ya que es el activo que más riesgo genera, tendriamos unicamente el VaR de la posición en dólares, $VaR_1 = 165.000 \; euros$. Por tanto, el VaR incremental de tomar la posición en yen es igual a: 257.738 euros-165.000 euros = 92.738 euros. El VaR marginal (152.108 euros) es, sin embargo, significativamente menor. El error de aproximación al utilizar el componente VaR se debe a que con solo dos activos, eliminar uno de ellos implica una variación muy notable en la composición de la cartera.

Si eliminasemos la posición en $US tendriamos el VaR de la posición en yenes, $VaR_2 = 198.000$ euros, por lo que el VaR incremental de la posición en $US es: 257.738-198.000=59.738 euros, de nuevo bastante inferior a su VaR marginal, que es 105.630 euros.

El *best hedge* respecto de cada uno de los dos activos sería:

$$b^* = -W\beta_i \frac{\sigma_c^2}{\sigma_i^2} = 3(10^6) \begin{pmatrix} 0.61477/0.05^2 \\ 1.7705/0.12^2 \end{pmatrix} 2.7111 \times 10^{-3} = \begin{pmatrix} 2.0 \times 10^6 \\ 1.0 \times 10^6 \end{pmatrix}$$

es decir, consistiría en deshacer toda la posición ¿Por qué se obtiene este resultado tan drástico?

Descomposición del VaR de la cartera del ejemplo					
	Posición euros	VaR individual euros	VaR marginal	Componente VaR euros	Contribución porcentual
Divisa	x_i ó $\omega_i W$	$-\Phi^{-1}_{1-p}\sigma_i\omega_i W$	$\Delta VaR_i =$	$CVaR_i =$	$CVaR_i/VaR$
			$= VaR\frac{\beta_i}{W}$	$= \Delta VaR_i x_i$	
US $	2.10^6	165.000	0,0528	105.630	41,0%
Yen	10^6	198.000	0,1521	152.108	59,0%
Total	3.10^6				
VaR no diversificado		363.000			
VaR diversificado				257.738	

Example 5 *ySupongamos ahora que los dos activos tienen una correlacion* $\rho = 0.65$. *Su covarianza será:* $\sigma_{12} = (0.05)(0.12)(0.65) = +0.003\,9$.

La varianza de la cartera es ahora:
$\sigma_c^2 = \omega' \Sigma \omega = (2/3, 1/3) \begin{pmatrix} 0.05^2 & 0.0039 \\ 0.0039 & 0.12^2 \end{pmatrix} \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix} = 4.\,444\,4 \times 10^{-3}$, *y su volatilidad:* $\sigma_c = \sqrt{4.\,444\,4 \times 10^{-3}} = 6.\,66\%$

En términos monetarios: $x' \Sigma x = (\omega' \Sigma \omega) W^2 = (4.4444) 10^{-3} 9 (10^{12}) = 4.0 \times 10^{10}$, *con una volatilidad de 200.000 euros. El VaR 95% a 1 año es* $(1.65)(200000) = 330.000$ *euros (suponemos que las rentabilidades diarias tienen media cero). Los VaR individuales (VaR no diversificado) en cada divisa son* $VaR_1 = (1.65)(0.05)2.10^6 =$*165.000 y* $VaR_2 = (1.65)(0.12)10^6 =$*198.000 euros, respectivamente, que suman 363.000 euros, un 10% más que el VaR diversificado.*

$$\begin{aligned} \Sigma\omega &= \begin{pmatrix} 0.05^2 & 0.0039 \\ 0.0039 & 0.12^2 \end{pmatrix} \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix} = \begin{pmatrix} 2.\,966\,7 \times 10^{-3} \\ 0.007\,4 \end{pmatrix} \\ \beta &= \begin{pmatrix} 2.\,966\,7 \times 10^{-3} \\ 0.007\,4 \end{pmatrix} \frac{1}{4.\,444\,5 \times 10^{-3}} = \begin{pmatrix} 0.667\,50 \\ 1.\,665\,0 \end{pmatrix} \end{aligned}$$

por lo que el VaR marginal de ambas divisas es:

$$\Delta VaR = -\alpha\beta\sigma_c = (1.65) \frac{1}{\sqrt{4.\,444\,5 \times 10^{-3}}} \begin{pmatrix} 2.\,966\,7 \times 10^{-3} \\ 0.007\,4 \end{pmatrix} = \begin{pmatrix} 7.\,342\,5 \times 10^{-2} \\ 0.183\,15 \end{pmatrix}$$

de modo que un incremento de 10.000 euros en la posición en dólares generaría un incremento aproximado de 734,25 euros, mientras que un incremento de 10.000 euros en la posición en yenes produciría un aumento de 1.831,50 euros.

Calculemos ahora exactamente el aumento que se produciría en el VaR de la cartera, a partir de la completa evaluación del riesgo de la cartera, tenemos:

$$\begin{aligned} \sigma_{c+a}^2 &= \begin{pmatrix} 2.01 & 1 \end{pmatrix} \begin{pmatrix} 0.05^2 & 0.0039 \\ 0.0039 & 0.12^2 \end{pmatrix} \begin{pmatrix} 2.01 \\ 1 \end{pmatrix} = 4.\,017\,8 \times 10^{-2} \\ &\Rightarrow \quad VaR_{c+a} = (1.65)\sqrt{4.\,017\,8 \times 10^{-2}} 10^6 = 330.730 \; euros \end{aligned}$$

Si comparamos con el VaR inicial, que era de 330.000 euros, tenemos un incremento de 730 euros, similar al calculado por el método aproximado.

Con un incremento de la posicion en yenes en 10.000 euros el VaR aumentaría en:

$$(\Delta VaR)' b = (0.073425 \; 0.18315) \begin{pmatrix} 0 \\ 10000 \end{pmatrix} = 1.831,5 \; euros$$

El VaR sería ahora:

$$\begin{aligned} \sigma_{c+a}^2 &= \begin{pmatrix} 2 & 1.01 \end{pmatrix} \begin{pmatrix} 0.05^2 & 0.0039 \\ 0.0039 & 0.12^2 \end{pmatrix} \begin{pmatrix} 2 \\ 1.01 \end{pmatrix} = 4.\,044\,5 \times 10^{-2} \\ &\Rightarrow \quad VaR_{c+a} = (1.65)\sqrt{4.\,044\,5 \times 10^{-2}} 10^6 = 331.830 \; euros \end{aligned}$$

con un aumento en VaR de la cartera de: 331.830 *euros*−330.000 *euros* = 1.830 *euros, muy similar a la estimación obtenida con el VaR marginal.*

El best hedge respecto de cada uno de los dos activos sería:

$$b^* = -W\beta_i \frac{\sigma_c^2}{\sigma_i^2} = 3(10^6) \begin{pmatrix} 0.6675/0.05^2 \\ 1.6650/0.12^2 \end{pmatrix} 4.444\,5 \times 10^{-3} = \begin{pmatrix} 3.56 \times 10^6 \\ 1.541\,7 \times 10^6 \end{pmatrix}$$

Example 6 *Supongamos, por último que los dos activos tienen una correlacion* $\rho = -0.25$. *Su covarianza será:* $\sigma_{12} = -(0.05)(0.12)(0.25) = -0.001\,5$.

La varianza de la cartera es ahora:
$\sigma_c^2 = \omega'\Sigma\omega = (2/3, 1/3) \begin{pmatrix} 0.05^2 & -0.0015 \\ -0.0015 & 0.12^2 \end{pmatrix} \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix} = 2.044\,4 \times 10^{-3}$,
y su volatilidad: $\sigma_c = \sqrt{2.044\,4 \times 10^{-3}} = 4.52\%$.

En términos monetarios: $x'\Sigma x = (\omega'\Sigma\omega)\, W^2 = (2.044\,4 \times 10^{-3}) 9(10^{12}) = 1.840\,0 \times 10^{10}$, *con una volatilidad de* $\sqrt{1.840\,0 \times 10^{10}} = 1356\,50$ *euros. El VaR 95% a 1 año es* $(1.65)(1356\,50) = 2\,23.8\,20$ *euros (suponemos que las rentabilidades diarias tienen media cero). Los VaR individuales (VaR no diversificado) en cada divisa son* $VaR_1 = (1.65)(0.05)2.10^6 =$*165.000 y* $VaR_2 = (1.65)(0.12)2.10^6 =$*198.000 euros, respectivamente, que suman 363.000 euros, un 62% más que el VaR diversificado, ya que ignora los beneficios de la diversificación que son ahora importantes, dada la correlación negativa.*

$$\begin{aligned} \Sigma\omega &= \begin{pmatrix} 0.05^2 & -0.0015 \\ -0.0015 & 0.12^2 \end{pmatrix} \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix} = \begin{pmatrix} 1.166\,7 \times 10^{-3} \\ 0.003\,8 \end{pmatrix} \\ \beta &= \begin{pmatrix} 1.166\,7 \times 10^{-3} \\ 0.003\,8 \end{pmatrix} \frac{1}{2.044\,4 \times 10^{-3}} = \begin{pmatrix} 0.570\,68 \\ 1.858\,7 \end{pmatrix} \end{aligned}$$

por lo que el VaR marginal de ambas divisas es:

$$\Delta VaR = -\alpha\beta\sigma_c = (1.65) \frac{1}{\sqrt{2.044\,4 \times 10^{-3}}} \begin{pmatrix} 1.166\,7 \times 10^{-3} \\ 0.003\,8 \end{pmatrix} = \begin{pmatrix} 0.042578 \\ 0.138\,67 \end{pmatrix}$$

de modo que un incremento de 10.000 euros en la posición en dólares generaría un incremento aproximado de 425,78 euros, mientras que un incremento de 10.000 euros en la posición en yenes produciría un aumento de 1.386,70 euros.

Calculemos ahora exactamente el aumento que se produciría en el VaR de la cartera, a partir de la completa evaluación del riesgo de la cartera, tenemos:

$$\begin{aligned} \sigma_{c+a}^2 &= \begin{pmatrix} 2.01 & 1 \end{pmatrix} \begin{pmatrix} 0.05^2 & -0.0015 \\ -0.0015 & 0.12^2 \end{pmatrix} \begin{pmatrix} 2.01 \\ 1 \end{pmatrix} = 0.018\,47 \\ &\Rightarrow \quad VaR_{c+a} = (1.65)\sqrt{0.018\,47}10^6 = 224.240 \text{ } euros \end{aligned}$$

Si comparamos con el VaR inicial, que era de 223.820 euros, tenemos un incremento de 420 euros, similar al calculado por el método aproximado.

Un incremento de la posicion en yenes en 10.000 euros el VaR aumentaría en 1.386,7 euros. El VaR sería ahora:

$$\sigma_{c+a}^2 = \begin{pmatrix} 2 & 1.01 \end{pmatrix} \begin{pmatrix} 0.05^2 & -0.0015 \\ -0.0015 & 0.12^2 \end{pmatrix} \begin{pmatrix} 2 \\ 1.01 \end{pmatrix} = 1.8629 \times 10^{-2}$$
$$\Rightarrow \quad VaR_{c+a} = (1.65)\sqrt{1.8629 \times 10^{-2}}10^6 = 225.210 \; euros$$

con un aumento en VaR de la cartera de: $225.210 \; euros - 223.820 \; euros = 1.390$ *euros, muy similar de nuevo a la estimación obtenida con el VaR marginal.*

El best hedge respecto de cada uno de los dos activos sería:

$$b^* = -W\beta_i \frac{\sigma_c^2}{\sigma_i^2} = 3(10^6) \begin{pmatrix} 0.57068/0.05^2 \\ 1.8587/0.12^2 \end{pmatrix} 2.0444 \times 10^{-3} = \begin{pmatrix} 1.4 \times 10^6 \\ 7.9165 \times 10^5 \end{pmatrix}$$

que en ningun caso agota las posiciones en los mismos.

3 Enfoques alternativos para el cálculo del VaR

Los enfoques para el cálculo del VaR pueden clasificarse en dos grupos: métodos de valoración local y métodos de valoración global. Los métodos locales requieren valorar la cartera una sola vez, en los precios observados, y utiliza derivadas en ese punto para inferir los posibles cambios en el nivel de riesgo. Dentro de estos, el método delta-Normal utiliza primeras derivadas ("deltas") y supone Normalidad de las rentabilidades. Es sencillo de aplicar, y una variante suya "el método de las Griegas" utiliza aproximaciones a las derivadas de primer y segundo orden.

Los métodos de valoracion global valoran la cartera en el punto inicial y tambien bajo distintos escenarios, para medir los posibles cambios en riesgo. Entre estos se encuentran el método de Monte Carlo de cálculo del VaR. Los métodos de simulación histórica generan directamente una valoración global. Consisten en generar hacia el pasado una serie temporal de rentabilidades de la actual cartera, aplicando a los precios de los activos que la componen las ponderaciones con que entran actualmente en la cartera. Esta serie de datos no representa a ninguna cartera real. A continuación, se construyen precios futuros hipotéticos bajo cada escenario, aplicando variaciones en precio observadas históricamente. Por eso es que a este procedimiento se le conoce también como *bootstrapping*. Una ventaja del metodo es que, puesto que solo necesita la serie temporal de rentabilidades de la cartera, no precisa estimar la matriz de varianzas y covarianzas de los activos que componen la cartera. No necesita hacer ningun supuesto sobre el tipo de distribución de probabilidad seguida por los precios de mercado. Incorpora de modo natural además la existencia de colas pesadas en las rentabilidades. Su mayor limitación es el supuesto de que el pasado representa fielmente lo que cabe esperar del futuro. Pondera por igual las observaciones más recientes que las más alejadas en el tiempo. Puede calcularse a distintos horizontes de inversión.

El método de Monte Carlo incorpora supuestos acerca de la distribución de probabilidad conjunta de los factores, y utiliza dichos supuestos para simular trayectorias futuras. La diferencia con el método histórico es que utiliza variaciones en precios no necesariamente observadas en la muestra. Las correlaciones entre factores están totalmente incorporadas a través del modelo supuesto para los factores. Es un método muy potente, pero tiene un alto riesgo de modelo. Puede incorporar facilmente distintos tipos de no linealidad, y cualquier horizonte de inversión. Si el modelo se establece correctamente, este es el método de cálculo del VaR más potente.

3.1 Valoración local: el método Delta-Normal

El primer paso consiste en evaluar la cartera en el punto inicial:

$$V_0 = V(S_0)$$

donde suponemos la existencia de un único factor de riesgo, representado por el precio S del activo subyacente a la cartera que se está valorando. Denotamos por Δ_0 la primera derivada parcial, la sensibilidad de la cartera al precio del subyacente, evaluado en la posición inicial, V_0. En el caso de un activo de renta fija, se trataría de su *duración modificada*, mientras que en el caso de una opción sería su *Delta*.

La pérdida potencial en el valor de la cartera es:

$$dV = \left(\frac{\partial V}{\partial S} \mid_{S_0} \right) dS = \Delta_0 dS$$

Si la distribución es Normal, el VaR de la cartera será:

$$VaR_c = |\Delta_0| \, VaR_S = -\Phi_{1-p}^{-1} \, |\Delta_0| \, \sigma_{dS/S} S_0$$

siendo $\sigma_{dS/S}$ la desviación típica de las variaciones porcentuales en el precio del subyacente. Vemos que solo precisamos evaluar la cartera en la posición inicial, para obtener V_0.

En una cartera de renta fija, el factor de riesgo sería su TIR, y, siendo la relación precio-TIR :

$$dV = -D^* V \; dy$$

donde D^* denota la duración modificada. En este caso, el VaR de la cartera será:

$$VaR = -\Phi_{1-p}^{-1} \, (D^* V) \, \sigma_{dy}$$

donde σ_{dy} denota ahora la volatilidad de cambios en el nivel de la TIR.

En el caso de una opción call comprada, la mayor pérdida que puede producirse a un nivel de confianza dado surge cuando $S^* = S_0 - \Phi_{1-p}^{-1} \sigma S_0$:

$$VaR_c = V(S_0) - V(S_0 - \Phi_{1-p}^{-1}\sigma S_0)$$

En su aplicación práctica, el método Delta-Normal sigue los pasos:

1. especificación de un conjunto de factores de riesgo
2. mapear la exposición lineal de los activos incluidos en la cartera con respecto a tales factores
3. agregar las exposiciones con respecto a un mismo factor
4. estimar la matriz de covarianzas de los factores de riesgo
5. calculo del riesgo total de la cartera

3.2 Aproximaciones Delta-Gamma ("las Griegas")

Podemos mejorar la aproximación lineal del método Delta-Normal añadiendo términos en el desarrollo en serie de Taylor:

$$dV = \frac{\partial V}{\partial S}dS + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}dS^2 + \frac{\partial V}{\partial t}dt + ... = \Delta dS + \frac{1}{2}\Gamma dS^2 + \Theta dt + ...$$

donde Γ es la segunda derivada del valor de la cartera respecto del factor de riesgo, y Θ es la tendencia temporal (*time drift*).

En un contexto multivariante (varias fuentes de riesgo), tendríamos:

$$dV_S = \Delta' dS + \frac{1}{2}(dS)'\,\Gamma(dS) + ...$$

siendo ahora dS un vector de variaciones en los factores, Δ un vector de Deltas, y Γ una matriz simétrica de Gammas con respecto a los distintos factores de riesgo. Los elementos fuera de la diagonal serían cross-gammas: $\Gamma_{ij} = \partial^2 V/\partial S_i \partial S_j$. Esto tendría en cuenta, por ejemplo, que en una opción, la delta depende también de la volatilidad implícita. Cuando hay muchas fuentes de riesgo, este método se hace impracticable.

En una cartera de renta fija, tendríamos:

$$dV = -D^* V\ dy + \frac{1}{2}CV dy^2$$

donde C denota la convexidad de la cartera, y es similar a Γ.

Para calcular el VaR de una opción call comprada podemos utilizar:

$$\begin{aligned} VaR_c &= V(S_0) - V(S_0 - \Phi_{1-p}^{-1}\sigma S_0) = \\ &= V(S_0) - \left[V(S_0) + \Delta\left(-\Phi_{1-p}^{-1}\sigma S\right) + 1/2\Gamma(-\Phi_{1-p}^{-1}\sigma S^2)\right] = \\ &= |\Delta|\ \Phi_{1-p}^{-1}\sigma S - 1/2\Gamma(\Phi_{1-p}^{-1}\sigma S^2) \end{aligned}$$

que puede utilizarse en realidad para posiciones cortas y largas.

Para funciones V más complejas, esto nos sería suficiente y habría que trabajar con términos cuadráticos en la expansión de Taylor, lo que nos fuerza a considerar las variables aleatorias dS y dS^2. Describamos primero el método *Delta-Gamma-Delta*: Tomando varianzas en el desarrollo de Taylor de segundo orden, tendríamos:

$$\sigma_{dV}^2 = \Delta^2 \sigma_{dS}^2 + (1/2\Gamma)^2 \sigma_{dS^2}^2 + 2\Delta(1/2\Gamma) cov(dS, dS^2)$$

Bajo Normalidad de dS, el ultimo termino es cero, y $\sigma_{dS^2}^2 = 2\left(\sigma_{dS}^2\right)^2$, de modo que:

$$\sigma_{dV}^2 = \Delta^2 \sigma_{dS}^2 + 1/2\left(\Gamma \sigma_{dS}^2\right)^2$$

Si las variables dS y $(dS)^2$ se distribuyesen conjuntamente como una Normal bivariante, entonces:

$$VaR = -\Phi_{1-p}^{-1} \sqrt{\Delta^2 \sigma_{dS}^2 + 1/2\left(\Gamma \sigma_{dS}^2\right)^2}$$

aunque esto no puede suceder, pues si dS fuese Normal, su cuadrado seguiría una distribución χ^2.

Tambien puede lograrse una mejor aproximación al VaR si tenemos en cuenta la existencia de asimetría ξ en rentabilidades, lo que podemos hacer mediante la expansión de Cornish-Fisher. El VaR se obtiene entonces reemplazando $-\Phi_{1-p}^{-1}$ en la expresión anterior por:

$$-\Phi_{1-p}^{-1} - \frac{1}{6}\left[\left(-\Phi_{1-p}^{-1}\right)^2 - 1\right]\xi$$

de modo que si existe asimetría negativa, el VaR aumentará con esta corrección.

Como alternativa, el método *Delta-Gamma-Monte Carlo* genera simulaciones aleatorias de los factores de riesgo S, y utiliza la aproximación de Taylor para generar movimientos simulados en el valor de la opción. Este método se conoce como un *enfoque parcial de simulación*.

3.3 Algunas aplicaciones del método Delta-Normal

La aplicación práctica del método Delta-Normal se basa en el proceso de "mapeo" o proyección de las exposiciones de la cartera sobre los factores de riesgo seleccionados, generalmente un conjunto reducido dentro del amplio conjunto de factores de riesgo que inciden sobre toda cartera de activos financieros. El otro elemento importante es la matriz de covarianzas Σ de los factores de riesgo sobre el horizonte de cálculo del VaR. El VaR es:

$$VaR = -\Phi_{1-p}^{-1} \sqrt{x' \Sigma x}$$

Dicha matriz puede expresarse: $\Sigma = S'RS$ donde R es la matriz de correlaciones de los factores y S es la matriz diagonal de sus desviaciones típicas (por lo que $S' = S$). El VaR es entonces:

$$VaR = -\sqrt{x'(\Phi_{1-p}^{-1} S'RS\Phi_{1-p}^{-1})x} = \sqrt{(x'V)R(x'V)'}$$

donde V es el factor de riesgo: $V = \Phi_{1-p}^{-1}S$.

4 Cálculo del VaR a partir de un modelo factorial lineal

Una estrategia habitual que facilita la gestión de riesgos, dada la enorme cantidad de posibles fuentes de riesgo, consiste en establecer un número reducido de factores que explique las correlaciones entre todos los activos que configuran una cartera o un fondo. Habitualmente, se dispone de un mapping de la rentabilidad de la cartera sobre un conjunto de factores, que habremos identificado previamente. Habremos estimado asimismo las sensibilidades de la cartera con respecto a cada factor, y las propiedades estadísticas de cada uno de los factores. Esencialmente, querriamos conocer su distribución de probabilidad multivariante: esperanzas matemáticas, varianzas y covarianzas y correlaciones. La proyección de la rentabilidad de la cartera sobre las rentabilidades de los factores:

$$r_{ct} = \alpha + \beta_1 f_{1t} + \beta_2 f_{2t} + \beta_3 f_{3t} + \beta_4 f_{4t} + u_t, t = 1, 2, ..., T$$

nos proporcionará una descomposición del VaR total en VaR sistemático, debido a los factores, y VaR específico, idiosincrático o residual.

Los modelos factoriales nos ayudan a reducir mucho la dimension del espacio de activos que tendríamos que considerar para la gestión de nuestra cartera. Su elección genera riesgo de modelo a partir de: a) la elección de los factores, b) la estimación de las sensibilidades, c) que ignoramos el riesgo específico en tal caracterización.

Un supuesto habitual es que las rentabilidades de cada factor son independientes en el tiempo y tienen una distribución, generalmente de tipo Normal. Las rentabilidades de los factores pueden estar correlacionadas entre sí. Algunos procedimientos de selección de factores generan factores incorrelacionados, y ello simplifica todavía más los cálculos, como veremos. Una generalización permitiría que las varianzas fueseen cambiantes en el tiempo. Otra generalización de más impacto eliminaría el supuesto de Normalidad. El método paramétrico genera expresiones analíticas para el cálculo del VaR cuando las rentabilidades tienen distribución Normal, t-Student, o mixturas de Normales o de t-Student. Ello no impide que utilicemos otras distribuciones, aunque sin disponer de expresiones analíticas para el VaR.

Los factores pueden ser tipos de interés, tipos de cambio (en carteras internacionales), índices de renta variable (en carteras con un componente de bolsa), indicadores macroeconómicos, etc.. Las sensibilidades son habitualmente coeficientes estimados en regresiones lineales, en cuyo casos suelen denominarse betas de la cartera respecto de cada factor. En carteras de activos sensibles a variaciones en los tipos de interés, los factores son generalmente tipos de interés a distintos vencimientos, como los que se utilizan para descontar los flujos de caja de la cartera. Si se descuentan flujos de caja entre bancos, se utilizará una estructura temporal del LIBOR como factores de riesgo; si tratamos con una contraparte que tiene un rating inferior a AA, añadiremos en ese caso el spread de crédito BBB a algunos vencimientos relevantes como factores de riesgo. En el caso de tener un tipo de interés como factor de riesgo, la sensibilidad respecto del mismo es el PV01 (valor presente de un punto básico). Como veremos, el componente de riesgo producido por los factores de tipos de interés suele ser de importancia menor.

El precio de un bono o de un swap es una función no lineal de los tipos de interés. Esta no linealidad queda recogida en el concepto de Valor Presente de 1 punto básico, como sensibilidad a los distintos factores de riesgo, por lo que su uso nos permitirá utilizar el método paramétrico lineal del VaR en estas carteras. Para carteras de acciones se establecen modelos lineales de factores de riesgo, de naturaleza macroeconómica o financiera.[7] Las únicas carteras para las cuales no puede aplicarse el método paramétrico del VaR es para carteras de opciones, o en cualquier caso en que las $P\&L$ de la cartera son función no lineal de los factores de riesgo.

Pueden darse dos situaciones: que estemos interesados en estimar el VaR de una única cartera, o en estimar el VaR de un amplio número de activos. En el primer caso, el modelo de factores nos puede servir para elaborar escenarios de evolución temporal de la cartera. Un escenario consiste en un determinado supuesto acerca de la evolucion de cada factor durante el horizonte de gestión. Estos escenarios pueden elaborarse mediante técnicas de predicción para cada factor (mediante modelos ARIMA, por ejemplo) o simultáneamente (mediante un modelo VAR, por ejemplo) para todos los factores. Alternativamente, los escenarios pueden responder a la creencia del analista. Por ejemplo, en el momento actual llevamos un tiempo con tipos de interés muy reducidos de modo estable, por lo cual, cualquier predicción econométrica del tipo a corto plazo, si este es un factor, tenderá a reproducir tal comportamiento. Sin embargo, podríamos estar en una situacion en que se han recibido indicaciones de que la economía de la zona euro está repuntando, y que una subida de tipos, incluso quiza notable, está próxima. Preferíriamos entonces escribir directamente el escenario previsto del tipo a corto que confiarlo a un modelo econométrico.

Es importante, finalmente, asignar a cada escenario una verosimilitud, en la forma de una probabilidad. La suma de las probabilidades asignadas a los distintos escenarios debe ser igual a uno. El modelo factorial puede servirnos para simular trayectorias posibles para la rentabilidad de la cartera durante el horizonte de gestión, de las cuales podríamos deducir, si queremos, trayec-

[7] Las carteras de commodities se mapean como cash-flows sobre estructuras temporales de forwards o tipos de interés a vencimientos constantes.

torias para el valor de mercado de la cartera. Para ello, bastaría especificar una trayectoria para cada factor durante el horizonte de gestión $(T, T+h)$, y extraer aleatoriamente h observaciones para $u_t, t = T+1, ..., T+h$. Con ello tendríamos una trayectoria para $r_{ct}, t = T+1, ...T+h$. Repitiendo este ejercicio un número amplio de veces (5.000, por ejemplo) tendríamos una distribución de probabilidad para $r_{c,T+h}$.

Pero, además, podriamos considerar que la propia trayectoria de cada factor es aleatoria, extrayendo aleatoriamente sus valores numéricos en el intervalo $T+1, .., T+h$ de una distribución de probabilidad alrededor de una senda central. En un caso sencillo, podríamos fijar el valor final de un factor, por ejemplo, en f_{1T+h} desde su valor inicial f_{1T}, y suponer una senda central consistente en seguir una trayectoria uniformemente creciente (o decreciente, segun que el valor final sea superior o inferior a su valor incial) desde f_{1T} hasta f_{1T+h}. Si suponemos adicionalmente que la trayectoria observada se desviará de esta senda central con una varianza determinada y de acuerdo con una distribución Normal, podriamos incorporar esta información en la generación de las 5.000 trayectorias que antes mencionamos. En este caso, en cada trayectoria, no solo la senda de u_t, sino tambien la senda seguida por los factores, serían diferentes. En ambos casos, terminamos con 5.000 realizaciones de la rentabilidad o del valor de mercado a vencimiento de nuestra cartera, con las que podriamos construir la distribución empirica de frecuencias, y tomar de ella el percentil adecuado como Valor en Riesgo. Este es un posible enfoque del método Monte Carlo de cálculo del Valor en Riesgo

La segunda situación en que podemos considerar un modelo factorial es cuando tratamos de modelizar el comportamiento de un amplio número de activos. El modelo de factores puede simplificar entonces el problema de modo importante, pues nos bastaría con tener previsiones o escenarios de evolución temporal de unos pocos factores, para deducir escenarios para los activos individuales, que quizá sean muchos, para hacer un seguimiento individualizado.

Para el cálculo del Valor en Riesgo precisamos estimar la varianza de la cartera.

Para un solo activo, tenemos:

$$Var(r_{it}) = \beta_i' Var(f_t)\beta_i$$

En el caso de tratar con un amplio número de activos, el modelo factorial nos permitiría obtener la matriz de covarianzas de las rentabilidades de los activos mediante:

$$Var(r_c) = B' Var(f) B$$

donde ahora, la matriz de betas B contiene en cada una de sus filas las betas de un activo. Será por tanto una matriz nxk, siento n el número de activos y k el número de factores de riesgo. El resultado sería una estimación de la matriz nxn de varianzas y covarianzas de los activos. Nótese que incluso si los factores estuviesen incorrelacionados, como sucede cuando se utiliza la técnica de Componentes Principales para su estimación, los activos individuales tendrían correlación no nula. Si con estos activos configuramos una cartera con ponderaciones recogidas en el vector ω, entonces la varianza de dicha cartera se estimaría:

$$Var(r_c) = (\omega B)' Var(f)(\omega B)$$

Dicho de otro modo, las sensibilidades de la cartera respecto de cada factor vendrían estimadas por el n-vector ωB. Denotando por:

$$\beta_c = \omega B$$

el vector de betas de la cartera, tendríamos:

$$Var(r_{ct}) = \beta_c' Var(f_t)\beta_c$$

Este cálculo permite desglosar la relevancia de cada factor por sí solo, a través de su varianza, así como la importancia de las correlaciones entre factores, para determinar la varianza de la rentabilidad de la cartera.

En un caso extremo, si se cuenta con una cartera de activos de renta variable bien diversificada dentro de un mismo mercado, podemos utilizar la varianza del índice de mercado como factor,

$$\sigma^2_{c,t+1} = \beta_c^2 \sigma^2_{M,t+1}$$

donde $\sigma^2_{M,t+1}$denota la predicción de la varianza de la rentabilidad del índice de mercado, y β_c es la beta de la cartera.

Para calcular el VaR en unidades monetarias, es decir, el VaR de la distribución $P\&L$, multiplicamos el VaR de la distribución de rentabilidades por el valor de la posición:

$$VaR = \left[\sigma_{c,t}\Phi^{-1}(1-p) - \mu\right] W_t.$$

Cuando contamos con las exposiciones $E \quad (E = \sigma_c W)$ a los distintos activos de la cartera, podemos estimar su VaR nominal (por tanto, en unidades monetarias) mediante:

$$VaR = \Phi^{-1}(1-p)\sqrt{E'ME}$$

siendo M la matriz de correlaciones de las rentabilidades de los distintos activos que entran en la cartera (en esta última expresión hemos supuesto rentabilidades nulas).[8]

[8] En ocasiones se utiliza la aproximación:

$$\begin{aligned} VaR &= (Valor\ cartera) \, . \left[VaR\left(rentabilidades\ \log arítmicas\right)\right] \cong \\ &\cong (Valor\ cartera) \, . \left(\exp\left[VaR\left(rentabilidades\ \log arítmicas\right)\right] - 1\right) \end{aligned}$$

4.1 Etapas en la construcción de un modelo factorial de VaR

1. Definir la cartera e identificar sus factores de riesgo
2. Fijar los parámetros básicos del modelo: nivel de confianza y horizonte de riesgo
3. Proyectar (mapping) la cartera sobre sus factores de riesgo
4. Modelizar la evolución de los factores de riesgo a lo largo del horizonte de riesgo. Cada uno de los tres enfoques sigue aqui una estrategia diferente
5. Revaluar la cartera para cada realización de los factores de riesgo. Habitualmente se supone que las sensibilidades permanecen constantes a lo largo del horizonte de riesgo, lo que equivale al supuesto de rebalanceo continuo de la cartera, que logicamente, no siempre se sigue.
6. Construir una distribución para la rentabilidad o para la $P\&L$ de la cartera. Con tipos de interés o con carteras que contemplan posiciones cortas y largas, es mejor lo segundo. La rentabilidad o el $P\&L$ sobre los h dias debe expresarse en términos de valor presente. Si la rentabilidad esperada de la cartera es muy distinta de la tasa de descuento, entonces la distribución de rentabilidades debe modificarse para ajustar esta diferencia. Cuando el VaR se mide sobre un año, este ajuste puede ser notable, por lo que es importante para hacer un seguimiento de dichos fondos, pero no es tan importante en el caso de fondos de gestión activa, en los que puede no estar tan justificado que la rentabilidad activa vaya a tener una media positiva.
7. Calcular el VaR y el ETL

4.2 Descomposición del VaR en el modelo factorial lineal Normal

Vamos a analizar en esta sección las desagregaciones del VaR que antes vimos, en un contexto particular, muy utilizado, como es el del modelo factorial bajo supuestos de distribución Normal para los factores de riesgo que se hayan identificado para el modelo.

4.2.1 VaR sistemático en el modelo lineal Normal

Para calcular el VaR de una cartera a horizonte de h dias, necesitamos conocer las predicciones de su valor esperado y su varianza, $E(R_h), Var(R_h)$. Si disponemos de un modelo de m factores, tendremos:

$$\begin{aligned} E(R_h) &= \theta' \mu_h \\ Var(R_h) &= \theta' \Omega_h \theta \end{aligned}$$

siendo μ_h el vector $mx1$ de rentabilidades esperadas para los factores, Ω_h la predicción de la matriz de covarianzas mxm de los factores sobre el horizonte de inversión, y θ el vector de sensibilidades de la rentabilidad de la cartera con respecto a los distintos factores. Tendremos:

$$VaR_{h,\alpha}\ sistem\acute{a}tico = \Phi^{-1}(1-\alpha)\sqrt{\theta'\Omega_h\theta} - \theta'\mu_h$$

por lo que el VaR sistemático puede calcularse a partir de la representación factorial de la cartera. Sólo necesitamos las sensibilidades θ y la predicción de la matriz de covarianzas, Ω_h. Bajo los supuestos de independencia temporal de las rentabilidades y de ausencia de heterocedasticidad condicional, tendremos:

$$\Omega_h = h\Omega$$

por lo que, bajo tales supuestos:

$$VaR_{h,\alpha}\ sistem\acute{a}tico = \sqrt{h}.VaR_{1,\alpha}\ sistem\acute{a}tico$$

4.2.2 VaR individuales (*Stand-alone VaR*)

Supongamos que tenemos una cartera por importe W con sensibilidades: $\theta = (\theta'_E, \theta'_R, \theta'_x)$, cada uno de los cuales es un vector de sensibilidades a equity, tipos de interés y forex. Por simplicidad, de momento consideramos que la sensibilidad de tipos de interés es con respecto a una única curva cupón cero. Los tres vectores de sensibilidades pueden estar en términos porcentuales o en términos nominales.

- En términos porcentuales, las sensibilidades serian: las betas (respecto al equity), los PV01 respecto de los tipos de interés divididos por W^{-1}, y un vector de unos (respecto de los tipos de cambio).
- En términos nominales, las sensibilidades serían W multiplicado por las betas (equity), los PV01 (tipos), y las cuantías de las posiciones en divisas (forex).

Particionamos la matriz de varianzas y covarianzas de los factores:

$$\Omega = \begin{pmatrix} \Omega_{Eh} & \Omega_{ERh} & \Omega_{EXh} \\ \Omega'_{ERh} & \Omega_{Rh} & \Omega_{RXh} \\ \Omega'_{EXh} & \Omega'_{RXh} & \Omega_{Xh} \end{pmatrix}$$

cada bloque tiene una dimensión (filas y columnas) en función del número de factores de cada tipo. Ignorando los posibles ajustes en las rentabilidades medias (por ejemplo, porque el horizonte del VaR sea corto), tenemos tres medidas de VaR *stand-alone*:

$$\begin{aligned}
\theta_R &= 0, \theta_X = 0 \Rightarrow Equity\ VaR_{h,p} = \Phi^{-1}(1-p)\sqrt{\theta'_E \Omega_E \theta_E} \\
\theta_E &= 0, \theta_X = 0 \Rightarrow Interest\ rate\ VaR_{h,p} = \Phi^{-1}(1-p)\sqrt{\theta'_R \Omega_R \theta_R} \\
\theta_R &= 0, \theta_E = 0 \Rightarrow Forex\ VaR_{h,p} = \Phi^{-1}(1-p)\sqrt{\theta'_X \Omega_X \theta_X}
\end{aligned}$$

incluso si todas las matrices de covarianzas fuesen igual a cero, la suma de los VaR *stand-alone* no coincidiría con el VaR total (por ser raiz cuadrada de sumandos que sí serían aditivos).

4.2.3 VaR marginal y VaR incremental

Para ambas medidas, VaR marginal y VaR incremental, necesitamos el vector gradiente del VaR. Recordemos que: $VaR = f(\theta) = \Phi^{-1}(1-\alpha)\sqrt{\theta'\Omega_h\theta}$, donde θ es el vector de sensibilidades de la rentabilidad de la cartera con respecto de cada factor de riesgo, y Ω_h es la matriz kxk de covarianzas de los factores en el horizonte de h periodos. Suponiendo un horizonte corto, de modo que los valores esperados de los factores de riesgo puedan suponerse igual a cero, y derivando en la expresión del VaR, tenemos:

$$g(\theta) = \frac{\partial VaR}{\partial \theta} = \Phi^{-1}(1-p)\left(\Omega_h\theta\right)\left(\theta'\Omega_h\theta\right)^{-1/2}$$

Nótese que $\Phi^{-1}(1-p)$ y $\left(\theta'\Omega_h\theta\right)^{-1/2}$ son escalares 1x1, mientras que $(\Omega_h\theta)$ es un vector columna kx1, siendo k el número de factores de riesgo. Si denotamos por m al producto de ambos escalares, tenemos:

$$g(\theta) = \frac{\partial VaR}{\partial \theta} = m\left(\Omega_h\theta\right) = \left(m\Omega_h\right)\theta = \left(\frac{\partial VaR}{\partial \theta_1}, \frac{\partial VaR}{\partial \theta_2}, ..., \frac{\partial VaR}{\partial \theta_k}\right)$$

y el producto: $\theta' g(\theta)$ es en este caso:

$$\theta' g(\theta) = \theta'\left(m\Omega_h\right)\theta = \Phi^{-1}(1-p)\ \left(\theta'\Omega_h\theta\right)^{-1/2}\left(\theta'\Omega_h\theta\right) = \Phi^{-1}(1-p)\sqrt{\theta'\Omega_h\theta}$$

de modo que la aproximación habitual a una función:

$$f(\theta) \approx \theta' g(\theta) = \sum_{i=1}^{n} \theta_i \frac{\partial f}{\partial \theta_i}$$

es en esta caso una igualdad exacta que descompone el VaR en función de los *VaR marginales* de cada factor, $\frac{\partial f}{\partial \theta_i} = \frac{\partial VaR(\theta)}{\partial \theta_i}$.

Otra descomposición alternativa:

$$\Delta f(\theta) \approx (\Delta\theta)' g(\theta) = \sum_{i=1}^{n} \left(\Delta\theta_i\right)\frac{\partial f}{\partial \theta_i}$$

permite calcular el $+VaR$ *incremental*.

Con esperanzas matemáticas no nulas para los factores de riesgo, tendriamos:

$$\begin{aligned} VaR &= \Phi^{-1}(1-\alpha)\sqrt{\theta'\Omega_h\theta} - \theta'\mu_h \Rightarrow \\ &\Rightarrow \quad g(\theta) = \Phi^{-1}(1-p)\left(\Omega_h\theta\right)\left(\theta'\Omega_h\theta\right)^{-1/2} - \mu_h \end{aligned}$$

En las secciones 6 y 7 examinamos distintas aplicaciones de los conceptos anteriores a carteras de renta fija, en primer lugar, y a carteras de renta variable, posteriormente.

En los ejemplos EIV.2.4, EIV.2.5, EIV.2.6 se descompone el VaR para carteras de renta fija, IV.2.14 a IV.2.17 para carteras de acciones, y el Caso IV.2.7 para commodities. En el caso del VaR histórico, los Casos IV.3.5.3 descomponen el VaR para acciones y divisas, IV.3.5.4 para tipos de interés y forex, y IV.3.5.5 para commodities. En estos ejemplos calculamos asimismo los VaR marginales.

5 Medición del riesgo en carteras de renta fija

5.1 Algunos conceptos básicos

Un bono es una secuencia de flujos de caja de igual cuantía, un porcentaje$100c$ del nominal N, seguida de un último cash flow de cuantía $100(1+c)N$. La constante c es el cupón del bono, cuyo nominal se toma en unidades de 100. Si, por ejemplo, el nominal de un bono es de 100.000 euros, el precio del bono en cada instante, en base 100, por ejemplo 98,76, se multiplicaría por 1.000 para obtener su valor de mercado.

Con descuento de tipo discreto, el *valor presente* de un bono de nominal N es:

$$PV = \frac{N}{100}\sum_{i=1}^{n} C_{t_i}\frac{1}{\left(1+R_{t_i}\right)^{t_i}}$$

donde n es el numero de pagos pendientes de realizarse, C_{t_i} la cuantia de los mismos, que será $100cN$ excepto a vencimiento, que será $100(1+c)N$. El valor presente de un bono se conoce como *Fair Value* del bono, que es un concepto teórico, mientras que el precio de mercado de dicho bono es el resultado de las interacciones de oferta y demanda. Los tipos de interés R_{t_i} a que se descuentan los flujos de caja con serán constantes, y su representación, en función de su vencimiento, se conoce como *estructura temnporal* de los tipos de interés.Los factores de descuento son los términos $\frac{1}{1+R_{t_i}}$. Podemos pensar en un tipo de interés constante que generase un valor presente igual al precio de mercado del bono:

$$P^M = \frac{N}{100}\sum_{i=1}^{n} C_{t_i}\frac{1}{\left(1+y\right)^{t_i}}$$

este tipo constante y se denomina *tasa interna de rentabilidad* (TIR) del bono. La TIR es la rentabilidad que obtendría el tenedor del bono si lo mantuviese hasta su vencimiento. El cupón apenas tiene efecto sobre la TIR del bono. Claramente, la TIR del bono está inversamente relacionada con su precio. Cuando el precio del bono sube, la TIR disminuye. Por eso es que una elevación de los tipos de interés de mercado conduce a una caida en los precios de la renta fija.

Con descuento continuo, el valor presente y la TIR del bono vendrian definidos por:

$$
\begin{aligned}
PV &= \frac{N}{100}\sum_{i=1}^{n} C_{t_i}\exp\left(-r_{t_i}t_i\right) \\
P^M &= \frac{N}{100}\sum_{i=1}^{n} C_{t_i}\exp\left(-yt_i\right)
\end{aligned}
$$

Dadas las caracteristicas de un bono, es instructivo variar su TIR y representar la relación entre el precio del bono y la TIR. *Ejercicio: Dibuje las curvas Precio-TIR para a) un bono cupón 5%, vencimiento a 3 años, b) un bono cupón 10%, vencimiento a 3 años, c) un bono cupón 10%, vencimiento a 5 años, y observe:*

- un aumento en el cupón eleva el precio del bono, para cada TIR
- cuando la TIR es igual al cupón del bono, su precio es igual a 100
- un alargamiento del vencimiento aumenta la pendiente y la convexidad de la relación precio-TIR del bono.

Ejercicio: Considere dos bonos. El primero paga cupón anual del 5% y tiene un vencimiento de 3 años. El segundo bono paga cupón del 10% y tiene vencimiento 5 años. Los tipos de interés de mercado a 1, 2, 3, 4 y 5 años son: 4,0%, 4,25%, 4,50%, 4,25% 4,20%. Calcule el valor presente de cada bono y su TIR ¿Que efecto tiene el cupón sobre la TIR del bono?: *R: 101,42 y 125,59. TIRes: 4,48% y 4,22%.*

Un bono cupón cero tiene un único pago, a su vencimiento. Tales bonos se venden generalmente a descuento. Es decir, el comprador paga $N(1+R_T)^{-T}$, siendo T el tiempo a vencimiento y R_T el tipo cupón cero anual correspondiente a dicho vencimiento. En realidad, no existen tales tipos cupón cero, sino que se deducen de los precios de mercado observados para productos cupón cero, resolviendo:

$$
\begin{aligned}
P^M &= N(1+R_T)^{-T} \\
P^M &= N(1+TR_T)
\end{aligned}
$$

según que el vencimiento sea superior a un año (arriba) o inferior a un año (abajo).

Como verá en el siguiente ejercicio:

- los bonos con cupon por encima de los tipos de mercado cotizan por encima de la par (100), y lo contrario sucede con los bonos con cupón por debajo de los tipos de mercado,
- el precio de un bono aumenta con el vencimiento si el cupón está por encima de los tipos de mercado, y lo contrario sucede si el cupón e sinferior a los tipos de mercado,

Ejercicio: Considere tres conjuntos de bonos, cada uno de ellos con bonos de vencimientos desde 1 año a 20 años. Los cupones del primer conjunto de bonos son 0%, los del segundo son 5% y los del tercero son 10%. Los tipos de interés de mercado a vencimientos anuales desde 1 a 20 años son: 6,00%, 6,50%, 7,00%, 7,20%, 7,20%, 6,60%, 6,00%, 5,60%, 5,40%, 5,25%, 5,25%, 5,20%, 5,20%, 5,00%, 5,10%, 5,00%, 4,90%, 4,90%, 4,80%, 4,80%. Represente gráficamente los Precios Justos (Fair Value) de cada bono mediante una curva para cada clase, representando precio contra vencimiento. Haga lo mismo con las Yield (TIR).

5.1.1 Duración y convexidad

La *duración de Macaulay* de un bono es el promedio del valor presente de los distintos flujos, ponderados por vencimiento:

$$D_M = \frac{\sum_{i=1}^{n} t_i P_{t_i}}{P}$$

con:

$$P_{t_i} = \frac{C_{t_i}}{(1+y)^{t_i}} \text{ ó } P_{t_i} = C_{t_i} \exp(-y t_i)$$

Un bono cupón cero tiene duración de Macaulay igual a su vencimiento. Todo bono que paga cupón tienen duración inferior a su vencimiento. La duración representa el tiempo medio sobre el que se reciben los flujos de caja. Si se desplaza la curva de rentabilidades (yield curve), la duración de Maculay es el punto "break-even" en que la renta que se pierde por reinversión de los cupones queda compensada exactamente por la ganancia en el valor del bono. [EIII.1.7], [EIII.1.8]

La duración modificada es, cambiada de signo, la aproximación de primer orden al porcentaje de variación en el precio por unidad de cambio en la TIR (yield). En un bono anual, la duración modificada es la duración de Macaulay dividida por $1 + y$.

Bajo descuento continuo tenemos:

$$\frac{dP}{dy} = -\sum_{i=1}^{n} C_{t_i} t_i \exp(-y t_i) = -\sum_{i=1}^{n} C_{t_i} t_i \exp(-y t_i) = -\sum_{i=1}^{n} P_{t_i} t_i$$

de modo que:

$$D_M \equiv duración\ de\ Macaulay = -\frac{1}{P}\frac{dP}{dy}$$

mientras que bajo descuento discreto,

$$D \equiv duración\ modificada = -\frac{1}{P}\frac{dP}{dy} = \frac{D_M}{1+y}$$

Puesto que en mercados se utiliza el descuento discreto es la duración modificada, y no la duración de Macaulay, la que se utiliza como aproximación a la variación porcentual en el precio del bono ante cambio de una unidad en la TIR.

La duración es la aproximación de primer orden al cambio porcentual en el precio del bono por cambio unitario en la TIR, pero ya hemos visto que la relación entre precio y TIR no es lineal, por lo que tiene sentido considerar la curvatura de dicha relación a través del concepto de *Convexidad*:

$$Convexidad = \frac{1}{P}\frac{d^2P}{dy^2}$$

En el caso de un bono con cupón anual y vencimiento en un número entero T de años, tenemos:

$$\begin{aligned} P &= \sum_{t=1}^{T} \frac{C_t}{(1+y)^t} \Rightarrow \frac{dP}{dy} = -\sum_{t=1}^{T} \frac{tC_t}{(1+y)^{t+1}} \Rightarrow \frac{d^2P}{dy^2} = \sum_{t=1}^{T} \frac{t\,(t+1)\,C_t}{(1+y)^{t+2}} \\ \frac{d^2P}{dy^2} &= \frac{2C_1}{(1+y)^3} + \frac{6C_2}{(1+y)^4} + \frac{12C_3}{(1+y)^5} + ... + \frac{T(T+1)C_T}{(1+y)^{T+2}} \end{aligned}$$

Un bono cupón cero tiene convexidad: $T(T+1)(1+Y)^{-2}$.La convexidad es una variación de segundo orden: es, aproximadamente, la variación en la duración modificada ante cambios de una unidad en la TIR.

Cuando se disponde carteras de bonos, hay que tener en cuenta que la duración modificada no es aditiva para distitos bonos. Tiene interés considerar la duración en valor: $D^{\$} = \frac{dP}{dy}$ y la convexidad en valor: $C^{\$} = \frac{d^2P}{dy^2}$, que están medidas en la misma unidad y la misma divisa que el propio bono. La Value Duration es la Duracion modificada, multiplicada por el precio del bono, mientras que la Value Convexity es la Convexidad multiplicada por el precio del bono. Estas medidas sí que son aditivas para posiciones en distintos bonos, sin mas que tener en cuenta sus magnitudes y el signo de la posición. Sin embargo, los distintos bonos tendrán TIRes diferentes, por lo que la interpretación de la duración y convexidad en valor que resulten no es la misma de antes. Deben interpretarse como la variación en el valor de la cartera ante un desplazamiento paralelo de la curva de TIRes, es decir, una variación igula en las TIRes a todos los vencimientos.

5.1.2 Aproximación en duración y convexidad al precio de un bono

La aproximación de segundo orden al precio de un bono es:

$$P(y) \simeq P(y_0) + \frac{dP}{dy} \mid_{y=y_0} (y - y_0) + \frac{1}{2}\frac{d^2P}{dy^2} \mid_{y=y_0} (y - y_0)^2$$

denotando: $\Delta y = y - y_0$, y por: $\Delta P = P(y) - P(y_0)$, es decir, el cambio en el precio cuando la TIR pasa de y_0 a y, la aproximación de segundo orden puede escribirse:

$$\Delta P \simeq \frac{dP}{dy}\Delta y + \frac{1}{2}\frac{d^2P}{dy^2}(\Delta y)^2 = -D^{\$}\Delta y + \frac{1}{2}C^{\$}(\Delta y)^2 \tag{9}$$

Supongamos que las TIRes de todos los bonos en la cartera cambian en la misma magnitud, Δy. Podemos agregar entonces las variaciones en los precios de los distintos bonos para obtener, para toda la cartera:

$$\Delta P \simeq -D_c^{\$}\Delta y + \frac{1}{2}C_c^{\$}(\Delta y)^2$$

Para un único bono, si dividimos (9) por P, tenemos:

$$\frac{\Delta P}{P} = \frac{1}{P}\frac{dP}{dy}\Delta y + \frac{1}{2}\frac{1}{P}\frac{d^2P}{dy^2}(\Delta y)^2 = -Duración\ \mathrm{mod}ificada.(\Delta y) + (0,5)\,Convexidad\,(\Delta y)^2$$

Los signos de duracion y convexidad son opuestos. Por tanto, dados dos cash flows con la misma duración modificada, aquel que tenga mayor convexidad tendrá un valor menos sensible a cambios en los tipos de interés. Es decir, el valor del cash-flow disminuye menos si suben los tipos, y aumenta más si los tipos de interés descienden. Por tanto, conviene tener una posición con convexidad positiva y elevada.

Inmunización Si prevemos un desplazamiento paralelo en la curva de tipos y queremos que nuestra cartera no experimente variaciones en precio, deberemos configurar la cartera de modo que tenga duración cero y convexidad igual a cero.

Ejercicio: Consideremos de nuevo los dos bonos de los ejemplos anteriores, y supongamos que hemos invertido 1,5 millones de euros en el primero y 1 millón de euros en el segundo. La Value Duration del primer bono es 416,41 euros, y la del segundo bono es igual a 514,00 euros. La Value Convexitiy del primero es 15.678.184 euros y la del segundo bono es 27.938.173 euros. Un cartera que está larga en una unidad en el primer bono y corta en le segundo tendría una Value Duration de -97,59 euros y una Value Convexity de -12.250.989. Supongamos que queremos inmunizar esta cartera utilizando dos bonos, B_1 y B_2. El primero tiene un principal de 1.000 euros, Fair Value de 1.200 euros, Value Duration igual a 5 y Convexity igual a 20.000. El bono B_2 tiene un principal de 10.000 euros, Fair Value de 10.780 euros, Value Duration igual a 2 y Convexity igual a 100.000. La inmunización se consigue comprando 32,05 unidades de B_1 y vendiendo -128,92 unidades de B_2. Por tanto, hemos de invertir $(32,05).(1.200) = 38.459$

euros en el primer bono, y $(128,92)(10.0780) = 1.389.755$ euros en el segundo, Como sus precios son 120,0 euros y 107,80 euros, respectivamente, ello equivale a comprar cantidades nominales de 32.049 euros y 1.289.198 euros, respectivamente.

5.2 Métodos de proyección de cash-flows

Una cartera de deuda tiene un alto número de exposiciones de riesgo, tantas como cash-flows pendientes de pago en la misma. Se hace preciso reducir dicho número, para lo que una posibilidad es la proyección de los cash-flows en unos cuantos vencimientos, que se denominan *vértices* de la proyección. Cuanto mayor sea el número de vértices, más aproximada será la proyección de los cash-flows, pero por otra parte, se necesita que los vértices dispongan de suficiente liquidez, lo que generalmente conducirá a reducir su número.

La idea es asociar cada cash-flow a los vértices más próximos, como veremos más abajo. El objetivo es sustituir nuestra cartera, que tiene posiblemente muchos vencimienots a plazos no estándar, y que además van cambiando con el paso del tiempo, en otra cartera a plazos estandar que además, se mantienen fijos. Cada cash-flow se proyecta sobre los dos vértices adyacentes a su vencimiento, para posteriormente consolidar todos los cash-flows que se han proyectado en cada vértice.

Como los vértices se mantienen constantes, no es preciso estimar sus propiedades estadísticas sino al cabo de un tiempo. Lo que se va haciendo es actualizar la proyección cada cierto tiempo. Las proyecciones deben compararse en términos de Valor Presente, pues se refieren a distintos instantes de tiempo.

Veremos primero dos proyecciones muy simples que transforman la cartera en un único flujo, la proyección Principal y la proyeccion por Duración

- La *proyección Principal,* que asocia el riesgo de un bono con el vencimiento del principal, únicamente. Para aplicarlo, se calcula el vencimiento medio de los bonos de la cartera, y se toma el VaR del bono cupon cero que tenga dicho vencimiento. Es un método simple, pero ignora el pago de cupones, sobreestimando el riesgo de la cartera.

Ejemplo: Con los tipos de interés de la hoja PCA_Short_Spot_new, estimamos las correlaciones entre rentabilidades[9]:

[9] Se trata de correlaciones entre rentabilidades diarias, calculadas como variación porcentual del precio diario implícito en los tipos de interés de los que partimos.

Correlaciones de rentabilidades diarias

1m	3m	6m	12m	24m	36m	48m	60m
1,000	0,737	0,451	0,269	0,117	0,074	0,058	0,051
0,737	1,000	0,854	0,562	0,322	0,257	0,232	0,217
0,451	0,854	1,000	0,854	0,609	0,526	0,488	0,462
0,269	0,562	0,854	1,000	0,899	0,827	0,783	0,746
0,117	0,322	0,609	0,899	1,000	0,985	0,959	0,928
0,074	0,257	0,526	0,827	0,985	1,000	0,992	0,974
0,058	0,232	0,488	0,783	0,959	0,992	1,000	0,994
0,051	0,217	0,462	0,746	0,928	0,974	0,994	1,000

Bajo desplazamientos paralelos de la curva de rentabilidades, las volatilidades de todos los vencimientos deberían ser iguales, y las correlaciones entre vencimientos deberían ser igual a uno lo que, como vemos, no es el caso. Bajo el supuesto de Normalidad, el VaR a un horizonte de tiempo reducido será proporcional al vencimiento: $VaR_h = \Phi_{1-p}^{-1}\sigma_h = \Phi_{1-p}^{-1}h\sigma$. La tabla siguiente muestra que tampoco sucede así: El VaR a 2, 3, 4 y 5 años excede del multiplo correspondiente del VaR a 1 año. Esto se debe, por supuesto, a que la volatilidad aumenta más que proporcionalmente, reflejando autocorrelación positiva entre rentabilidades.

	1m	3m	6m	12m	24m	36m	48m	60m
Tipos de interés último dia muestra	5,49	5,41	5,23	4,75	4,34	4,31	4,36	4,41
Volatilidad %	0,003%	0,006%	0,014%	0,034%	0,078%	0,120%	0,159%	0,197%
Volatilidad puntos básicos	0,015	0,035	0,074	0,160	0,337	0,515	0,693	0,871
VaR Normal 95% 1 mes	0,021%	0,049%	0,107%	0,255%	0,587%	0,905%	1,202%	1,492%

Consideremos una cartera que consta de dos bonos. El primer bono tiene nominal 100 millones de euros, con vencimiento a 5 años y cupón 6%, mientras que el segundo bono tiene vencimiento a un año, cupón 4% y nominal 100 millones. Con la actual estructura de tipos de interés (último dia de la muestra) el Valor Presente de la cartera es $206,277$millones de euros.

El vencimiento medio es 3 años, por lo que la proyección Principal asocia 200 millones de euros al vencimiento de 3 años. Bajo esta proyección Principal, el VaR de la cartera a 1 dia, y al 5% se estima en $206,277 millones \times (0,905\%) = 1,867 millones$. Este enfoque sobreestima el riesgo, ya que ignora el pago de cupones.

	Cash flows				*Proyección(PV)*	
Plazo	*Bono 1*	*Bono 2*	*Tipos*	*Principal*	*Duración*	*Cash Flow*
1	6	104	4,75	0	0	105,01
2	6	0	4,34	0	0	5,51
$2,808$					206,28	
3	6	0	4,31	206,28	0	5,29
4	6	0	4,36	0	0	5,06
5	106	0	4,41	0	0	85,41
Total				206,28 m.	206,28 m.	206,28 m.

- La *proyección por Duración* asocia el riesgo al de un bono cupón cero con vencimiento igual a la duración del bono. Se sustituye la cartera de bonos por un único bono cupón cero, con vencimiento igual a la duración de la cartera. Dicha duración no será, generalmente, un número entero de años, por lo que habrá que calcular su VaR por interpolación de los VaR de los dos bonos cupón cero con vencimientos adyacentes a la duración de la cartera de bonos. Este VaR será generalmente inferior al anterior.

La duración es 2,808 años, por lo que la proyección por Duración sustituye la cartera por un bono cupón cero con vencimiento igual a la Duración. Por tanto, esta proyección asocia 206,277 millones de euros al vencimiento de 2,808 años. Interpolando linealmente los VaR a 2 y 3 años, tenemos una estimación del VaR : $0.587 + (0.905 - 0.587)(2.808 - 2) = 0.844\,\%$ y, por tanto, del VaR de la cartera: $206,277 millones \times (0.844) = 1,741$millones de euros, algo inferior a la anterior estimación, precisamente por incorporar el pago de cupones.

- En *la proyección de Cash-flows* se descompone el riesgo de los instrumentos de renta fija en el componente de riesgo asociado a cada uno de los cash-flows. Los cash-flows, en valor presente, descontados con el tipo de interés cupón cero apropiado, se agrupan en los vértices correspondientes de la estructura temporal. En este caso, no se construye otra distinta, tan solo se distribuye por vencimientos la estimación del riesgo.

$Plazo$	$VP\ Cash\ flows$ (x)	VaR $plazo\ (V)$	VaR $in\operatorname{div}idual : xV$
1	$105,01$	$0,255\%$	$26,78$
2	$5,51$	$0,587\%$	$3,23$
3	$5,29$	$0,905\%$	$4,78$
4	$5,06$	$1,202\%$	$6,08$
5	$85,41$	$1,492\%$	$127,44$
$Total$	$206,28$		$168,32$
	$VaR\ No\ \operatorname{div}ersificado$		$1,68$
	$VaR\ Diversificado$		$1,62$

Una vez calculado el valor presente del cash flow x a un determinado vencimiento, se multiplica por el VaR de dicho vértice, que será $V_j = \Phi^{-1}_{1-p}\sigma_j$. Si los tipos cupón cero a los distintos vencimientos tuvieran correlación perfecta (lo que justificaría el uso de un solo factor, así como el uso de la duración Macaulay), el *VaR no diversificado* de la cartera sería:

$$VaR = \sum_{j=1}^{N} |x_j|\, V_j = \Phi^{-1}_{1-p} \sum_{j=1}^{N} |x_j|\, \sigma_j.$$

que es el que aparece como la suma de los VaR individuales, 1,68 millones de euros. Esta sería la máxima cantidad que esperaríamos perder en el horizonte de 1 mes con una confianza del 95%.

Cuando, como es habitual, las correlaciones no son perfectas, se premultiplica y postmultiplica dicha matriz de correlaciones por el vector de riesgos monetarios $x_j V_j$ de cada vértice:

$$VaR = \sqrt{(xV)'\Gamma(xV)}$$

donde $xV = (x_1V_1, x_2V_2, ..., x_nV_n)$. La raiz cuadrada de dicho producto es el VaR, que resulta ser 1,624 millones de euros para esta cartera de bonos.

La diferencia entre esta cantidad y el VaR Duración, 1,741 millones de euros, es igual a 117.000 euros. Esta diferencia se debe a dos factores. Por un lado, que los movimientos en la curva de tipos no son paralelos, lo que hace que la medida de riesgo no sea lineal en el vencimiento y la volatilidad anualizada no sea constante con el vencimiento. Este factor explica una diferencia de 1,741-1,683=58.000 euros. En segundo lugar, que las correlaciones no son perfectas, lo que explica una diferencia de 1,683-1,624=59.000 euros, similar a la anterior.

La siguiente tabla recoge otro modo de estimar el VaR de la cartera utilizando las variaciones en el valor de los bonos cupón cero, como se hace habitualmente en un ejercicio de *stress testing.* Si todos los bonos cupón cero estuviesen perfectamente correlacionados, podriamos reducir su valor de acuerdo con el VaR de cada plazo. Por ejemplo, el bono a 5 años vale inicialmente 0,80577 (ver tabla siguiente) y su VaR es 1,492%, por lo que tras un descenso igual al VaR, el nuevo precio del bono sería: $0,80577(1-0,01492) = 0,79375$. El nuevo valor de la cartera sería 204,595 millones de euros. El descenso respecto a su valor inicial seria: $206,277 - 204,595 = 1,682$ millones de euros, que es aproximadamente igual al VaR no diversificado que antes encontramos. Sin embargo, el cálculo utilizando la matriz de correlaciones, que nos generó una estimación de 1,741 millones de euros, es más exacto.

Plazo	*Cash flow millones*	*Valor inicial bono c.cero*	*PV inicial cash flows*	*VaR plazo*	*Precio final bono*	*nuevo PV cash flows*
1	110	0,95462	105,009	0,255%	0,95219	104,741
2	6	0,91857	5,511	0,587%	0,91318	5,479
3	6	0,88112	5,287	0,905%	0,87315	5,239
4	6	0,84310	5,059	1,202%	0,83297	4,998
5	106	0,80577	85,412	1,492%	0,79375	84,138
Total			206,278			204,595

A continuacion, pasamos a analizar los distintos procedimientos de proyección de un cash flow en varios vértices. Hay distintos criterios que pueden utilizarse para comparar el cash-flow inicial con la proyección. Un criterio básico es que se mantenga el Valor Presente del cash-flow inicial. Otros criterios podrían ser: mantener la volatilidad del cash-flow, mantener su PV01, etc.. Para tener una solución única al problema de proyección deben utilizarse en ella tantos vértices como criterios queremos que se cumplan.

5.2.1 Proyección con valor presente y duración constantes

Supongamos un cash flow en el período T, con un valor presente de 1 euro, donde $T_1 < T < T_2$, donde T_1 y T_2 son dos vértices previamente seleccionados. Denotemos por x_1 y x_2 la parte del capital de 1 euro inicial que asociaremos a los vértices T_1 y T_2.Todos los cash-flows en este punto están en términos de valor presente.[10]

Aunque no es imprescindible que la suma del valor presente de las proyecciones coincida con el valor presente de la cantidad proyectada, es habitual imponer tal condición, que significa:

$$x_1 + x_2 = 1$$

La duración Macaulay del cash-flow original es T, mientras que la duración Macaulay de la secuencia formada por las proyecciones es: $(x_1T_1 + x_2T_2)/(x_1 + x_2)$, por lo que la proyección mantendrá invariante la duración Macaulay si y solo si,

$$x_1T_1 + x_2T_2 = (x_1 + x_2)T$$

Resolviendo el sistema formado por estas dos ecuaciones:

$$x_1 = \frac{T_2 - T}{T_2 - T_1}$$

Las dos condiciones conducen en este caso a una única solución, mientras que por sí solas, cada una de ellas sería compatible con un continuo de proyecciones (si bien no con cualquier proyección) .

En el ejercicio anterior, la cartera tiene duración 2,808, y si proyectamos un valor presente de 1 euros entre los vertices de 2 y 3 años, tendríamos:

$$x_1 = \frac{3 - 2,808}{3 - 2} = 0,192$$

por lo que asignamos al plazo de 2 años una cantidad: $(206,28)\,(0,192) = 39,65$ millones, asignando el resto, 166,63 millones al plazo de 3 años. En este ejemplo, hemos trabajado con bonos cupon cero, de modo que vencimiento y duración coinciden. Posteriormente, podriamos calcular las cantidades nominales equivalentes a estos valores presentes.

El VaR de la proyección sería: $(39,65)\,(,587) + (166,63)(,905) = 1,74$ millones de euros.

Ejercicio [EIII.5.1]*: La proyección de un cash-flow con vencimiento 1 año y 65 dias con valor presente de 1 millon de euros en vértices de 12 y 18 meses, que mantenga invariante la duración y el valor presente es: 638.889 euros en 12 meses y 361.111 en el vértice de 18 meses, ambos en valor presente. Utilice la convención de 360 dias por año.*

[10] Alternativamente, sin descuento, podriamos considerar que un cash flow de e^{rT} euros se está proyectando en $x_1e^{r_1T_1}$ euros y $x_2e^{r_2T_2}$ euros, respectivamente. Con descuento discreto, y considerando pr simplicidad vencimientos en un número entero de años, pensaríamos en proyectar $(1+R)^T$ en $x_1(1+R_1)^{T_1}$ y $x_2(1+R_2)^{T_2}$ euros.

5.2.2 Proyeción con invarianza del PV01

El modo de imponer esta condición depende de que trabajemos con descuento continuo o discreto. Con descuento continuo, un cash-flow con valor presente de 1 euro tiene un valor no descontado de e^{rT} euros, con un PV01 aproximado de: $Te^{rT}e^{-rT}.10^{-4} = T\ .10^{-4}$euros. Si este capital de 1 euro se proyecta un cash-flow con valor presente de x euros en el vértice i, con vencimiento T_i, su valor no descontado será $e^{r_iT_i}$ siendo r_i el tipo de interés cupón cero a dicho vencimiento, y su PV01 aproximado es x_iT_i (Recordemos que el PV01 se calcula utilizando cantidades en valor presente).

Por tanto, bajo descuento continuo, la invarianza del PV01 requiere,

$$\sum_{i=1}^{n} x_iT_i = T$$

Con descuento discreto, un cash-flow en T con valor presente de 1 euro, tiene un valor no descontado de $(1+R_i)^{T_i}$, y un PV01 aproximado: $T_i(1+R_i)^{T_i}(1+R_i)^{-T_i-1} = T_i(1+R_i)^{-1}$ (suponiendo que los vencimientos son un número entero de años). En este caso, la invarianza de PV01 requiere:

$$\sum_{i=1}^{n} \frac{x_iT_i}{(1+R_i)^{T_i}} = \frac{T}{1+R}$$

Podemos ver que, bajo descuento continuo, una proyección que mantenga inalterados el valor presente y la duración, mantendrá asimismo invariante el PV01, ya que la condición que hemos encontrado también debe ser satisfecha por la proyección que mantiene invariantes valor presente y duración. Esto no sucede bajo descuento discreto.

Además de esto, otra ventaja del descuento continuo es que para mantener la invarianza del PV01 en la proyección, no necesitamos conocer los tipos cupón cero en los vértices, al contrario de lo que sucedería con descuento discreto.

5.2.3 Proyección con invarianza en volatilidad

Denotemos por $\sigma, \sigma_1, \sigma_2$ las volatilidades de las variaciones en los tipos de interés en los tres vencimientos T, T_1, T_2 y por ρ la correlación entre las variaciones en tipos de interés en los vértices. En ocasiones, se desconoce la volatilidad en T, y se aproxima mediante interpolación de las volatilidades en los dos vértices adyacentes.

Si queremos que la proyección mantenga invariante la volatilidad, tendremos que imponer:

$$x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + 2x_1x_2\sigma_1\sigma_2\rho = \sigma^2$$

que no tiene una solución única, pero que puede combinarse con alguna de las condiciones que antes vimos para generar una única solución. En general, con más de dos vértices, la condición sería:

$$x'Vx = \sigma^2$$

donde V es la matriz de covarianzas de las variaciones en los tipos de interés en los vértices seleccionados.

En el ejemplo de la sección anterior, podemos interpolar las volatilidades de los plazos de 2 y 3 años para obtener una volatilidad estimada para el plazo de 2,808 años igual a: 0,112%. Para mantener la volatilidad de la cartera, queremos resolver la ecuación:

$$x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + 2x_1x_2\sigma_1\sigma_2\rho = \sigma_c^2 = 0,112^2$$

y si mantenemos la condición $x_1 + x_2 = 1$(manteniendo con ello la duración) tendremos:

$$\left(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2\right)x^2 + 2\left(-\sigma_2^2 + \rho\sigma_1\sigma_2\right)x + \left(\sigma_2^2 - \sigma_c^2\right) = 0$$

donde $\sigma_1 = 0,078\%, \sigma_2 = 0,120\%$. La única solucion entre 0 y 1 es: $x = 0,177$, similar al resultado que mantuvimos para la proyección que preserva la duración. Esto conduce a asociar $(206.28)\,(0.177) = 36.\,51$ millones al vencimiento de 2 años y el resto, 169,77 millones al plazo de 3 años.

Puede probarse que si la correlación entre vértices es igual a 1 y la volatilidad de cada vértice es proporcional a su duración, entonces la proyección con Duración invariante y la proyección con Volatilidad constante, coinciden.

Ejercicio: Considere las mismas condiciones del ejemplo anterior, con volatilidades de 75 pb. para las variaciones en el tipo a 12 meses, y de 90 pb. para las variaciones en el tipo a 18 meses, con correlación de 0,75 entre ambas variaciones. La proyeción que mantiene invariante la volatilidad y el valor presente es de 321.564 euros en el tipo a 12 meses y 678.436 euros en el tipo a 18 meses, ambos en valor presente. Existe otra solución con un cash-flow negativo. La proyeción que mantiene invariante la volatilidad y la duración es de 683.944 euros en el tipo a 12 meses y 386.577 euros en el tipo a 18 meses, ambos en valor presente. Esta proyección es muy diferente de la anterior. El valor presente de esta proyección es de 1.070.521 euros. Finalmente, la proyeción que mantiene invariante la volatilidad y el PV01 es de 893.542 euros en el tipo a 12 meses y 191.342 euros en el tipo a 18 meses, ambos en valor presente. El valor presente de esta proyección es de 1.084.884 euros.

5.2.4 Proyección sobre varios vértices

Como hemos visto, si proyectamos sobre dos vértices, no podemos mantener invariantes simultáneamente muchas características del cash-flow original. Podemos conseguirlo si proyectamos sobre un conjunto suficientemente amplio de vértices.

Ejercicio: [EIII.5.3] *Supongamos un cash-flow a 1 año y 65 dias, con valor presente de 1 millon de euros, sobre vértices de 6, 12 y 18 meses, con las volatilidades y correlaciones que aparecen en la tabla*:

Correlaciones	6 meses	12 meses	18 meses	Volatilidades (bp.)
6 meses	1			60
12 meses	0,8	1		75
18 meses	0,7	0,75	1	980

Considere 360 dias por año, y construya una cartera que tenga el mismo valor presente, PV01 y volatilidad que la cartera original. Solución: $(x_6, x_{12}, x_{18}) = (-227.054$ *euros*; $1.092.997$ *euros; 134.057 euros), estando estas cantidades en valor presente.*

Indicación: Utilizar Solver con función objetivo: $x'Vx - \sigma^2$ (ver hoja de cálculo).

5.2.5 Benchmarking un fondo

Utilizamos en esta sección la misma idea de las proyecciones para construir fondos réplica de un determinado fondo de renta fija. Supongamos un fondo de renta fija con la estructura de cash-flows que aparece en la tabla siguiente, en millones de euros. Se trata de un fondo excesivamente simple. Logicamente un fondo real tendría posiciones en vencimientos más largos que los indicados. La duración de este fondo es de 3,237 años y su Valor Presente es 59.385.000 euros.

Teniendo en cuenta que la duración essuperior a 3 años, para encontrar una cartera que replicara el fondo podriamos adoptar el criterio de reproducir su duración tomando posiciones en los vencimientos de 3 y 4 años. Para ello, tomamos:

$$x_1 = \frac{T_2 - T}{T_2 - T_1}(59.385) = \frac{4 - 3,237}{4 - 3}(59.385) = (0.763)(59.385) = 45,333$$

en el plazo de 3 años y 14,052 en el plazo de 4 años. En términos de valores nominales, se trata de 51,449 millones de euros en el plazo de 3 años y 16,667 en el plazo de 4 años. El VaR no diversificado de esta cartera es: $(45.333)(0.905) + (14.052)(1.202) = 579.170$ euros, en valor presente.

Si construimos una cartera en los vencimientos extremos, 1 y 5 años, tendriamos:

$$x_1 = \frac{T_2 - T}{T_2 - T_1}(59.385) = \frac{5 - 3.237}{5 - 1}(59.385) = 26,179$$

en el plazo de 1 años y $59,385 - 26,179 = 33,205$ en el plazo de 5 años. En términos nominales, se trataria de $\frac{26.179}{0.95462} = 29,712$ a 1 año y $\frac{33.205}{0.80577} = 39.385$ a 5 años. El VaR no diversificado es: $(26.179)(0.255) + (33.205)(1.492) = 562.180$ euros, en valor presente.

Plazo	*Cash flows*(10^6) *(x)*	*Descuento*	*Cash − flow descontado*	*VaR plazo (V)*	*VaR individual (VaR)(PV)*
1	2, 49	0, 95462	2, 38	0,255%	0,0061
2	13, 96	0, 91857	12,82	0,587%	0,0752
3	24, 83	0, 88112	21,88	0,905%	0,1979
4	15,40	0, 84310	12,98	1,202%	0,1560
5	11, 57	0, 80577	9,32	1,492%	0,1391
Total	68,25		59,38		0,574
	VaR No diversificado				0,574
	VaR Diversificado				0,569

Plazo	*VaR plazo (V)*	*VaR individual (VaR)(PV)*	*Cartera* 1 *(PV)*	*Cartera* 2 *(PV)*	*Cartera* 1	*Cartera* 2
1	0,255%	0,0061		26,179		0,067
2	0,587%	0,0752				
3	0,905%	0,1979	45,333		0,410	
4	1,202%	0,1560	14,052		0,169	
5	1,492%	0,1391		33,205		0,495
Total		0,574				
	VaR no diversificado	0,574			0,579	0,562
	VaR diversificado	0,569			0,578	0,547
	VaR Relativo				0,0051	0,0149

El VaR relativo se mide:

$$VaR\ \mathrm{Re}lativo = \Phi^{-1}_{1-p}\sqrt{(x-x_0)'\Sigma_h(x-x_0)} = \Phi^{-1}_{1-p}\sqrt{h(x-x_0)'DCD(x-x_0)}$$

donde, en nuestro ejemplo, $h = 1, \Phi^{-1}_{1-p} = 1,65, D$ es la matriz diagonal de desviaciones tipicas, y C es la matriz de correlaciones de las rentabilidades a los vencimientos de 1 a 5 años.

Como puede apreciarse, la cartera que invierte a 3 y 4 años tiene un VaR diversificado algo mayor que el fondo, y su VaR relativo es de 5.900 euros. Esta es la cantidad en la que, con confianza del 95%, el fondo que hemos constituido puede dar un peor resultado sobre un período de 10 dias que el fondo que queremos replicar. Por su parte, la cartera que invierte a 1 y 5 años tiene un VaR diversificado menor que el VaR del fondo, pero tiene un VaR relativo mayor, estimado en 16.600 euros. El VaR relativo suele presentarse en la forma de un R2, mediante:

$$\begin{aligned} Cartera\ 1 &: \ 1-\left(\frac{0.0051}{0.0578}\right)^2 = 0.992 \\ Cartera\ 2 &: \ 1-\left(\frac{0.0149}{0.0547}\right)^2 = 0.933 \end{aligned}$$

La primera cartera tiene un menor *tracking error*. Dado que las correlaciones se debilitan para vencimientos lejanos, cabe esperar que una cartera que replica Duración tenga el menor riesgo absoluto cuando combina venccimientos distantes, como resulta de combinar Liquidez y el plazo a 5 años, en este ejemplo. Pero conviene recordar que reducir el riesgo en términos absolutos no es lo mismo que reducirlo en térrminos relativos y que podemos encontrar carteras réplica que tengan un riesgo absoluto menor pero un riesgo relativo mayor que el fondo que quiere replicarse.

6 El modelo lineal Normal de VaR en carteras de renta fija

Vamos a considerar el cálculo del VaR de una cartera caracterizada por una secuencia de cash-flows. La cartera puede constar de préstamos, bonos, swaps, etc.. Los factores de riesgo son tipos de interés a determinados vencimientos fijos. Cada uno de tales tipos puede descomponerse entre el tipo LIBOR a ese vencimiento más un spread de crédito. En tal caso, además de tener como factores de riesgo la curva LIBOR, hay que añadir una o más estructuras temporales de crédito para diferentes ratings, como verems en un ejercicio. El exceso de rentabilidad de tal cartera sobre la tasa de descuento será significativamente distinto de cero solo si la cartera tiene exposiciones a contrapartidas de baja calidad crediticia y el horizonte de riesgo sea largo.

Suponemos que la secuencia de cash-flows ha sido previamente proyectada a tipos de interés a determinados vencimientos en términos de valor presente y manteniendo invariante la volatilidad. Puesto que el vector PV01 se expresa en términos de valor presente, y no hay constante en el mapping de la secuencia de cash-flows a los factores de riesgo, la rentabilidad esperada descontada de la cartera es cero, por lo que solo necesitamos la volatilidad para la determinación del VaR.

Recordemos que si $\delta 01_t$ denota la sensibilidad del factor descuento aplicable al vencimiento t ante una variación de un punto básico en el tipo cupón cero a dicho plazo, $\delta 01_t = \frac{1}{(1+R_t-.01)^t} - \frac{1}{(1+R_t)^t}$, El valor presente de un punto básico de un cash flow C_t a dicho plazo es: $PV01(C_t) = C_t.\delta 01_t$, que depende, por supuesto, del volumen del flujo de caja. En consecuencia, $PV01(C_t)$ mide el cambio que se produce en el valor presente del flujo C_t cuando *desciende* en un punto básico el tipo de interés que se utiliza para calcular el factor descuento.

Puesto que el precio teórico de un cash-flow es el valor presente del mismo, la variación en el precio de una secuencia de cash-flows an periodos $t_1, t_2, ..., t_n$ puede escribirse:

$$P_t - P_{t-1} = -\sum_{i=1}^{n} C_{t_i}.\delta 01_{t_i}(R_{i,t} - R_{i,t-1}) = -\sum_{i=1}^{n} PV01(C_{ti})(R_{i,t} - R_{i,t-1}) = -\sum_{i=1}^{n} PV01_i.\Delta R_{i,t}$$

donde en lo sucesivo, $PV01_i$ denota el PV01 de un cash flow C_{t_i}.

El $P\&L$ descontado de una cartera es la variación en el valor presente de la secuencia de cash-flows de la misma:

$$\Delta PV \approx -\sum_{i=1}^{n} (PV01_i)\,(\Delta R_i) = -\theta' \Delta R \tag{10}$$

donde donde $\theta = (PV01_1, PV01_2, ..., PV01_n)$ denota el vector de sensibilidades del cash-flow respecto de los factores, y ΔR es el vector de variaciones, en puntos básicos, en los vértices del mapping (la proyección) que hayamos aplicado a cada cash-flow. Ambos vectores, θ y ΔR, tienen una dimensión igual al número de cash-flows que se consideran, n. Por tanto, hay un tipo de interés asociado a cada cash-flow. Nótese que un bono de dicha cartera puede generar un cierto número de cash-flows antes de su vencimiento. La unidad de análisis del riesgo en la cartera de renta fija es el cash-flow, no el bono.

Puesto que el PV01 es el valor presente de un cambio en un punto básico, el cambio (10) en el valor de la cartera ya viene medido en términos de valor presente.

Si ΔR tiene una distribución multivariante Normal con esperanza μ y matriz de varianzas y covarianzas Ω, entonces el $P\&L$ descontado tendrá una distribución aproximada $N(-\theta'\mu, \theta'\Omega\theta)$. Por el razonamiento habitual en el cálculo del VaR, tenemos:

$$VaR_p = \Phi^{-1}(1-p)\sqrt{\theta'\Omega\theta} + \theta'\mu$$

Los tipos utilizados para descontar los flujos de caja tendrán una representación en función de los factores de riesgo. Aunque las variaciones en estos últimos fuesen cero, las variaciones en los primeros podrían no serlo y aparecerían en la expresión anterior, pues μ y Ω sn momentos de ΔR. Pero frecuentemente tomamos como factores de riesgo los mismos tipos de interés que se usan para descontar los flujos de caja, en cuyo caso es natural tomar el cambio esperado como igual a cero, $\theta'\mu = 0$.[11]

Si Ω_h denota la matriz de covarianzas calculada sobre h dias, tenemos en este caso:

$$VaR_p = \Phi^{-1}(1-p)\sqrt{\theta'\Omega_h\theta} \tag{11}$$

En el ejercicio [EIV.1.8] la matriz de covarianzas viene dada en puntos básicos, siendo sus elementos las varianzas y covarianzas de las variaciones en los tipos de interés. Esto se debe a que el vector PV01 contiene las sensibilidades a variaciones absolutas, nominales, de 1 p.b., en los factores de riesgo, no a variaciones relativas en los mismos. Por tanto, es preferible tratar con variaciones absolutas, en puntos básicos, en los tipos de interés. Además, la variación en un

[11] Si los tipos de interés utilizados como factores no coinciden con los de descuento pero utilizamos un modelo de representación de estos en función de los factores, sin incluir constantes, de nuevo será natural suponer $\theta'\mu = 0$. En definitiva, trabajando con variacioens diarias en tipos de interés, como es habitual, es natural hacer este supuesto.

tipo de interés puede interpretarse como aproximadamente igual a la rentabilidad porcentual del bono cupón cero correspondiente.[12] Las volatilidades anuales de dichas variaciones diarias son a menudo del orden de 100 puntos básicos, pero en mercados emergente spueden ser incluso bastante más elevadas.

Ejercicio [EIV.1.8]: *Encuentre el VaR 1% a 10 dias de un cash flow que se ha proyectado sobre vértices a 1 y 2 años, con PV01 de 50 y 75 euros. Suponga que los cambios en términos absolutos de los tipos a dichos vencimientos en el horizonte de 10 dias sigue una distribución Normal bivariante con esperanza nula* $\mu = (0;0)$, *correlación 0,9 y volatilidades anuales de 100 puntos basicos para la variación a 10 dias en el tipo a 1 año y de 80 puntos básicos para la variación en el tipo a 2 años. R:* 4.989 euros.

En mercados desarrollados la rentabilidad de una cartera de renta fija se mide en términos de variaciones de tipos, en puntos básicos, no en términos relativos. En mercados emergentes los tipos de interés pueden ser extremadamente altos y volátiles, y sus volatilidades son tan elevadas que se citan en términos porcentuales. Trabajando con datos de dichos mercados, o bien se ajustan las estimaciones de las PV01 para que reflejen sensibilidades ante variaciones porcentuales en los tipos de interés, o hay que ajustar la matriz de covarianzas para expresarla en puntos básicos. En el Ejercicio [EIV.2.2] se propone esta transformación.

Ejercicio [EIV.2.2] *Considere dos tipos de interés con correlación de 0,9. Uno de ellos está en el 10% con volatilidad del 30%, mientras el segundo está en 8% con volatilidad de 25% ¿cuál es su matriz de covarianzas en puntos básicos? Respuesta*: *Una volatilidad 30% en un tipo de interés del 10% equivale a una volatilidad de* $(0,1)(0,3) = 300$ *puntos básicos, y la varianza anual sería 90.000. La varianza diaria sería* $90.000/250 = 360$.*Siguiendo un argumento similar llegamos a una matriz de covarianzas anual:* $\sum = \begin{pmatrix} 90.000 & 54.000 \\ 54.000 & 40.000 \end{pmatrix}$ *o, en términos diarios*: $\sum = \begin{pmatrix} 360 & 216 \\ 216 & 160 \end{pmatrix}$.

6.1 Descomposición del VaR en carteras de renta fija

Los ejercicios [EIV.2.3] y [EIV.2.4] proporcionan aspectos de la gestión de riesgos en un contexto general. En EIV.2.3, la cartera consta de dos cash-flows y se utilizan los tipos de interés a ambos vencimientos como factores de riesgo. Conociendo las volatilidades de los tipos de interés en los vértices, hemos de calcular la volatilidad de la cartera, $\theta'\Omega\theta$ y, con ella, calcular el VaR

Ejercicio [EIV.2.4] *Consider a cash flow with sensitivities:* PV01(1-year)=\$1000, PV01(2-year)=\$1500, PV01(3-year)=\$2000. *Suppose the interest rates at maturities 1, 2 and 3 years have daily volatilities of 75, 60 and 50 basis points, and correlations of* 0.95 (1yr,2yr), 0.9 (1yr,3yr) and 0.975 (2yr,3yr). *Find the*

[12] Por ejemplo, un bono cupón cero a un año, nominal 100, con tipo de interés 6,15% tendría precio 94,2063. Un bono similar con tipo de interés 6,17% tendría precio 94,1886. El tipo de interés habría variado en 2 p.b., mientras que el precio del bono habría descendido 0,0177, un 0,0188%, aproximadamente 0,02%.

1% 10-day normal linear VaR. Now assume that interest rates are 4.0%, 4.5% and 5.0% at the 1-year, 2-year and 3-year vertices and suppose that a trader considers entering into a swap with the following cash flow: ($3million at 1-year, -$3million at 2-year, -$0.25 at 3-year). *What is the incremental VaR of the trade? R:* $120.970; -$6.693.

Este ejercicio requiere el cálculo del VaR incremental si se entra en una inversión en un swap. Algunos de los vencimientos de la nueva inversión pueden coincidir con los ya comprendidos en la cartera actualmente existente. El VaR incremental se obtiene multiplicando el vector PV01 de la nueva inversión, por el vector gradiente del VaR de la inversión ya existente, que pretendemos modificar. Cada elemento de dicho producto es un VaR incremental, correspondiente a cada uno de los vencimientos.

En este ejercicio se considera esa única inversión alternativa a la ya existente, es decir, una modificación de nuestra cartera. Si se quiere comparar varias de tales inversiones posibles entre sí, conviene *normalizar* los VaR incrementales, pues su valor numérico depende de la magnitud de los cash flow de cada operación, con lo que hablaríamos del VaR incremental por unidad de cash-flow invertida. Para ello, podemos dividir cada PV01 por la suma de los valores de todos los PV01 que aparecen en el vector de sensibilidades de la inversión, o dividir cada PV01 por la raiz cuadrada de la suma de los cuadrados de todos los PV01.

El ejercicio [EIV.2.5] analiza el VaR Normal lineal en una exposición a dos curvas de tipos, y calcula el VaR marginal de cada curva. Se trata de una cartera de deuda invertida en Estados Unidos y el Reino Unido.

El ejercicio [EIV.2.6] considera una cartera de deuda corporativa de rating A. Se descompone cada tipo de interés utilizado como factor (vértice del mapping) en dos componentes: un tipo LIBOR y el spread de crédito, y descompone el VaR en los componentes debidos al riesgo LIBOR y al riesgo de crédito. Las posiciones, en términos del PV01 en cada vértice, son iguales para el LIBOR y para el spread de crédito. Se supone que sólo se conocen varianzas y correlaciones para los spread de crédito a 1 y 5 años, por ser los más negociados, asi que para los restantes vencimientos (vértices) hay que calcularlos por interpolación:

$$\sigma_i = \sqrt{\frac{(i-1)\sigma_5^2 + (5-i)\sigma_1^2}{5-1}}$$

Como conocemos las correlaciones entre los spread de crédito a 1 y 5 años y cada tipo LIBOR, podemos interpolar las restantes correlaciones entre spreads de crédito y tipos LIBOR mediante:

$$\rho_{ik} = \sqrt{\frac{(i-1)\rho_{5k}^2 + (5-i)\rho_{1k}^2}{5-1}}$$

donde ρ_{ik} denota la correlación entre el spread de crédito a vencimiento i años, y el tipo LIBOR a k años. Utilizamos interpolación lineal de las correlaciones al cuadrado, bajo el supuesto de que las correlaciones que interpolamos tienen el mismo signo. Nótese que *la interpolación lineal generaría una matriz de correlaciones singular.*

Para estimar las correlaciones entre los spread de crédito a distinto vencimiento, primero interpolamos para obtener las correlaciones con el spread de crédito a 1 año, ρ_{i1}, sabiendo que $\rho_{51} = 0,90$.

$$\rho_{i1} = \sqrt{\frac{(i-1)\rho_{51}^2 + (5-i)\rho_{11}^2}{5-1}} = \sqrt{\frac{(i-1)\rho_{51}^2 + (5-i)}{5-1}}$$

luego, interpolamos para calcular las correlaciones entre el spread de crédito a 5 años y los restantes vencimientos de esta misma variable (la correlación con el spread a 1 año es conocida: 0,90):

$$\rho_{5i} = \sqrt{\frac{(i-1)\rho_{55}^2 + (5-i)\rho_{51}^2}{5-1}} = \sqrt{\frac{(i-1) + (5-i)\rho_{5i}^2}{5-1}}$$

y, posteriormente, interpolamos para estimar las correlaciones con los restantes vencimientos:

$$i > k \Rightarrow \rho_{ik} = \sqrt{\frac{(i-k)\rho_{5i}^2 + (5-i)\rho_{ii}^2}{5-k}} = \sqrt{\frac{(i-k)\rho_{5i}^2 + (5-i)}{5-k}}$$

A partir de aquí, podemos calcular el gradiente del VaR y los stand-alone VaR (respecto del LIBOR y respecto del riesgo de crédito), asi como los VaR marginales respecto de ambos factores.

Si descomponemos el tipo de interés en cada vértice:

$$r_q(t,T) = r(t,T) + s_q(t,T)$$

donde r_q denota el tipo correspondiente a un credit rating q, $r(t,T)$ es el tipo LIBOR en t a vencimiento T, y s_q es el spread de crédito a ese nivel de rating. En términos de varianzas:

$$V\left(r_q(t,T)\right) = V\left(r(t,T)\right) + V\left(s_q(t,T)\right) + 2Cov\left(r(t,T), s_q(t,T)\right)$$

El ejercicio [EIV.2.7], en el archivo Case Study PC VaR.xls, explica como utilizar un mapping the cash-flows para diseñar escenarios de tipos de interés. Se quiere estimar el efecto que sobre el valor de una cartera tendria una elevación paralela de la curva de tipos, así como un cambio de pendiente, lo cual se puede calcular utilizando los PV01 de cada vencimiento y aplicando las supuestas variaciones en cada uno de dichos vencimientos.

En el ejercicio [EIV.2.8], en Case Study PC VaR.xls, se calcula el VaR de una cartera de renta fija, mapeada (proyectada) sobre unos determinados vértices. Se resuelve mediante la forma cuadrática que tiene el vector de PV01 a los distintos vértices y la matriz de varianzas y covarianzas de los tipos de interés a dichos vértices. Posteriormente, en EIV.2.10 se examina la bondad de la aproximacion por CPs, que pasamos a describir a continuación.

6.2 Combinando cash-flow mapping con análisis de Componentes Principales

Si adoptamos la estrategia de utilizar componentes principales de la estructura temporal como factores de riesgo, una vez seleccionado el número k de los mismos que vamos a utilizar, la representación de las *variaciones* ΔR_t en los n tipos de interés que afectan a una cartera, puede expresarse:

$$\underset{Txn}{\Delta R_t} \approx \underset{Txk}{CPs_t^*}\underset{nxk}{(W^*)'}$$

donde CPs_t^* es el vector columna formado por los k primeros componentes principales, y W^* es la matriz nxk de pesos (loadings).

La varianza de las variaciones diarias en tipos de interés cupón cero se aproxima entonces:

$$Var(\Delta R_t) = Var(CPs^* W^{*\prime}) = W^* Var(CPs^*) W^{*\prime} = W^* \Lambda^* W^{*\prime}$$

donde Λ^* denota la matriz diagonal formada con los autovalores asociados a los componentes principales seleccionados (recordemos que dichos autovalores son las varianzas de cada componente principal). W^* es una matriz que tiene por columnas los autovectores corespondientes, ordenados con el mismo criterio que los autovalores.

Para cada tipo de interés tendriamos:

$$Var(\Delta R_{it}) = \sum_{j=1}^{3} \beta_{ij}^2 \lambda_j$$

siendo $(\beta_{i1}, \beta_{i2}, \beta_{i3})$ el vector que permite aproximar R_{it} en función de los tres componentes principales escogidos (suponiendo que hayamos decidido tomar los tres primeros componentes principales).

Por tanto, si $PV01$ es el vector $nx1$ formado por el Valor Presente de 1 p.b. de cada uno de los tipos de interés, tendríamos para el valor de mercado de la cartera:

$$Var(\Delta P_t) = (PV01)' Var(\Delta R_t)(PV01) = (W^{*\prime}\theta)' \underset{kxk}{\Lambda}{}^* (\underset{kxn}{W^{*\prime}} \underset{nx1}{\theta}) = S' \Lambda^* S$$

donde θ denota el vector $PV01, nx1$, y $S = W^{*\prime}\theta, kx1$, son las sensibilidades que habitualmente calculamos para una secuencia de flujos de caja respecto de los k factores.

Dada la estructura diagonal de la matriz Λ^* tenemos finalmente, en el caso de utilizar Componentes Principales como factores:

$$Var(\Delta P_t) = \sum_{j=1}^{3} S_j^2 \lambda_j$$

siendo S_j la sensibilidad de la secuencia de flujos de caja respecto de cada componente principal. Nótese que hemos reducido un vector de sensibilidades β de dimensión n a un vector de sensibilidades de dimensión k, donde n puede ser del orden de 50, mientras que k se reduce generalmente tan sólo a 3 ó 4. Esta es una expresión similar a la que obtuvimos para $Var(\Delta R_{it})$ pero ahora la sensibilidad a cada factor, S_j, es el producto de una matriz nxk de pesos (W^*) por el vector $nx1$ formado por los $PV01$ de cada tipo de interés.

El VaR de la cartera de renta fija basado en el análisis de Componentes Principales es:

$$PC\ VaR_{h,p} = \Phi^{-1}(1-p)\sqrt{S'\Lambda S} = \Phi^{-1}(1-p)\sqrt{\theta' W^* \Lambda W^{*\prime} \theta}$$

El ejercicio [EIV.2.9], en Case Study PC VaR.xls, describe el uso de los componentes principales obtenidos para la estructura temporal de tipos de interés, como factores de riesgo. Primero se estiman las sensibilidades de la secuencia de cash-flows respecto de los tres componentes principales escogidos, y luego se calcula la varianza de la secuencia de cash-flows mediante $\sum_{i=1}^{3} \beta_i^2 Var(p_i^*) = \sum_{i=1}^{3} \beta_i^2 \lambda_i$ y, posteriormente, se calcula el VaR de la cartera en el ejercicio [EIV.2.10], en Case Study PC VaR.xls. El resultado se compara con el cálculo exacto de [EIV.2.8].

Otra de las grandes ventajas del uso de los componentes principales como factores de riesgo es su simplicidad en el diseño de escenarios de tipos de interés que permitan llevar a cabo un análisis de *stress testing*, entre otras cosas.

6.3 Gestión de un fondo de renta fija

Una situación en la que el uso de factores es especialmente útil, es cuando gestionamos un fondo de renta fija. Para ello, hemos de formar predicciones del precio de cada bono que negocia en el mercado para el término del horizonte de gestión, $T+h$. De este modo, podríamos tomar posiciones largas en T (hoy) en aquellos bonos cuyo precio esperamos que en $T+h$ sea más elevado que hoy, y posiciones cortas en caso contrario. O podríamos elaborar una regla de trading que consistiese en tomar posiciones largas cuando esperamos que el bono se revalorice más de un $\alpha\%$, y posiciones cortas cuando esperamos que su precio descienda en más de un $\beta\%$. El procedimiento que proponemos es una extensión del método descrito al final del ejercicio empírico de la sección 5.

Para predecir en T el precio de un bono en $T+h$ necesitamos predecir los factores descuento aplicables a los flujos de caja que en $T+h$ estarán pendientes de pago. Estos pueden ser un número muy elevado, por lo que en general será impracticable mantener modelos predictivos para cada uno de dichos tipos de interés y se hará conveniente utilizar modelos factoriales. Si optamos por aplicar un modelo factorial a este cálculo, deberemos comenzar por identificar los factores.

Para ello, en una primera etapa, a partir de una estructura temporal, identificamos los factores a utilizar. La estructura temporal viene dada por un conjunto de tipos de interés cupón cero a vencimento estándar. Podría tratarse de tipos cupón cero, mes a mes, desde un mes a 60 meses (5 años), si nuestros activos tienen todos los flujos de caja por debajo de ese plazo, o de tipos cupón cero año a año, desde 1 año hasta 30 años, o incluos más lejanos. Estas series estan disponibles en las Web del Bank of England y del Federal Reserve System. Para caracterizar los factores podemos seguir un método de Componentes Principales, o un método de regresión. El primero nos vendrá a decir que con 3 componentes principales, interpretables como el nivel general de tipos de interés, la pendiente y la curvatura de la estructura temporal, podemos explicar un porcentaje muy elevado de la fluctuación a lo largo de toda la curva. Además, tendríamos la representación de los tipos cupón cero en función de los factores, que son componentes principales en este caso:

$$r_{it} = \beta_{0i} + \beta_{1i}CP_{1t} + \beta_{2i}CP_{2t} + \beta_{3i}CP_{3t} + u_{it} \tag{12}$$

estimando de esta forma una matriz B , $nx3$, de sensibilidades, siendo n el número de tipos cupón cero de los que partimos.

En una segunda etapa, estimamos un modelo polinómico para explicar el n-vector de sensibilidades respecto de un determinado factor, por ejemplo, los $\beta_{1i}, i = 1, 2, ..., n$, en función del plazo,$m_i$. Hariamos lo mismo con los términos constantes, β_{0i}. Así tendríamos por ejemplo:

$$\beta_{1i} = \delta_0^1 + \delta_1^1(\ln m_i) + \delta_2^1(\ln m_i)^2 + \delta_3^1(\ln m_i)^3 + \delta_4^1(\ln m_i)^4 + u_i$$

Una vez estimados estos coeficientes, tendriamos con ellos una matriz $4x5$:

$$A = \begin{pmatrix} \delta_0^0 & \delta_1^0 & \delta_2^0 & \delta_3^0 & \delta_4^0 \\ \delta_0^1 & \delta_1^1 & \delta_2^1 & \delta_3^1 & \delta_4^1 \\ \delta_0^2 & \delta_1^2 & \delta_2^2 & \delta_3^2 & \delta_4^2 \\ \delta_0^3 & \delta_1^3 & \delta_2^3 & \delta_3^3 & \delta_4^3 \end{pmatrix}$$

En la tercera etapa, dado un factor descuento d_h, a un plazo h, $d_h = (1+r_{0h})$, siendo r_{0h} el tipo cupón cero a plazo h, llevamos a cabo la interpolación:

$$a_h = A \begin{pmatrix} 1 \\ \ln h \\ (\ln h)^2 \\ (\ln h)^3 \\ (\ln h)^4 \end{pmatrix} = \begin{pmatrix} \beta_{0h} \\ \beta_{1h} \\ \beta_{2h} \\ \beta_{3h} \end{pmatrix}$$

que llevados a la ecuación (12) nos permiten generar un escenario para el tipo cupón cero:

$$r_{0h} = \hat{\beta}_{0h} + \hat{\beta}_{1h}CP_{1t} + \hat{\beta}_{2h}CP_{2t} + \hat{\beta}_{3h}CP_{3t} + u_{h,t},\ t = T+1, ..., T+h$$

y, finalmente, para el factor descuento:

$$d_{ht} = \frac{1}{(1 + r_{0h})^h}$$

si el plazo es superior a un año, o:

$$d_{ht} = \frac{1}{1 + hr_{0h}}$$

si es inferior a un año.

Una alternativa a la puesta en práctica de un procedimiento no lineal de interpolación como el descrito consistiría en "mapear" o proyectar la secuencia de flujos de caja cuyo riesgo estamos tratando de valorar en unos vencimientos estándar, denominados "vértices". Estos se pueden hacer por distintos procedimientos. Cada uno de ellos tratará de cambiar la secuencia de flujos de caja, cuyas fechas de ingreso o pago estarán en $T + h$ a distancias temporales generalmente atípicas (382 días, 1423 días, etc.), en una secuencia de flujos de caja en cuantías distintas, pero todos ellos a pagar o cobrar en fechas estándar, coincidentes con los vencimientos para los que disponemos de tipos cupón cero en la estructura temporal de la que hemos partido. Este cambio de flujos de caja se hace de modo que determinadas propiedades de la secuencia de flujos de caja permanezcan inalteradas, como su duración, su valor presente, etc.. [ver seccion sobre *Métodos de proyeccion de Cash-flows*]

7 El modelo lineal Normal de VaR para carteras de renta variable

Si tenemos una cartera construida con n valores de renta variable, con pesos w, la volatilidad de la rentabilidad de la cartera es: $\sigma = \sqrt{w'Vw}$ donde V denota la matriz de varianzas y covarianzas del conjunto de rentabilidades de las n acciones.

Para horizontes de riesgo cortos, como suele ser el caso en este contexto, podemos ignorar la diferencia entre la rentabilidad esperada de la cartera y la tasa de descuento, y calcular el VaR:

$$VaR_{h,p} = \Phi^{-1}(1 - p)\sigma\sqrt{\frac{h}{250}}P$$

expresion a la que, para horizontes largos, puede incorporarse la corrección debida a la diferencia entre la rentabilidad de la cartera y el tipo de interés sin riesgo, restando del cálculo anterior la rentabilidad esperada.

Así, en el caso general, para aplicar la expresión del VaR Normal bajo linealidad a un horizonte de h periodos, suponiendo que el vector de rentabilidades de los factores de riesgo sigue una distribución Normal multivariante y dichas rentabilidades son i.,i.d., tenemos que predecir tanto el vector de rentabilidades x_h en exceso, como su matriz de covarianzas V_h, y calcular:

$$VaR_{h,p} = \Phi^{-1}(1-p)\sqrt{w'V_h w} - w'E(x_h)$$

Ejercicio: El Ejercicio [EIV.2.11] hace este tipo de cálculos incluyendo y omitiendo la rentabilidad esperada de la cartera. Se observa que si se calcula el VaR teniendo en cuenta el ajuste por rentabilidad esperada (como debe hacerse), se obtiene un VaR 1% a 10 dias que es inferior en un 5% al que se obtendría ignorando este ajuste (259.765 euros frente a 273.738 euros). La diferencia no es demasiado importante por el reducido horizonte de inversión. Por el contrario, en el cálculo del VaR 1% a 1 año la diferencia entre ambas estimaciones del VaR es de un 24% (1.051.530,5 euros frente a 1.384.864 euros). Tan elevada diferencia se debe a que la rentabilidad esperada de la cartera es apreciable, de un 5,5% (calcularla con los datos del ejercicio).

Si, incorrectamente, ignorasemos el valor temporal del dinero y calculásemos el VaR sobre rentabilidades no descontadas, tendriamos un VaR 1% a 10 dias de 260.272,4 euros, frente a los 259.765 mencionados, una diferencia poco relevante. Por el contrario, sin el descuento tendríamos un VaR a 1 año de 1.104.107 euros, frente a los 1.051.530,5 euros mencionados, sobreestimando el VaR en un 5%, una cantidad muy relevante.

Vemos por tanto, una vez más, que tanto la rentabilidad esperada como el valor temporal del dinero juegan un papel muy importante en el cálclulo del VaR sobre horizontes largos, mientras que sobre horizontes cortos, pueden ignorarse como aproximación sin incurrir en un error apreciable.

Ejercicio: El ejercicio [EIV.2.12] descompone el VaR de una cartera doméstica de acciones enVaR sistemático y VaR específico. Por un lado, se calcula la volatilidad de la cartera y, por otro, la volatilidad explicada por los factores: $\beta'\Sigma\beta$, siendo Σ la matriz de covarianzas de los factores. Con ellos, se puede calcular el VaR total y el VaR debido a los factores, respectivamente. El VaR residual es la raiz cuadrada de la diferencia de los cuadrados de ambos VaR.

Para este tipo de cálculos, puede ser preferible la estimación de unas sensibilidades (betas) mediante EWMA, de modo que sean sensibles a la situación actual de riesgo, en vez del cálculo habitual de las betas por MCO.

7.1 VaR factorial para carteras de acciones, bajo Normalidad

Supongamos un único factor, cuya rentabilidad en exceso se distribuye en el horizonte de los próximos h dias según una $N(\mu_h, \sigma_h^2)$. Si el modelo factorial de nuestra cartera es:

$$Y_t = \alpha + \beta X_t + \varepsilon_t \tag{13}$$

La rentabilidad en exceso de la cartera tendrá tambien distribución Normal, con esperanza: $\alpha + \beta\mu_h$, y desviación típica $\beta\sigma_h$.Como el alfa de la cartera es un elemento idiosincrático del mismo, no entra en el cálculo del riesgo sistemático; por el contrario, se contabiliza como riesgo específico de la cartera. El riesgo sistemático de la cartera, que es el debido al factor equity, tendrá tambien distribución Normal, con esperanza: $\beta\mu_h$, y desviación típica $\beta\sigma_h$, y tenemos:

$$p = P\left[Y_s < y_{hp}\right] = P\left[\beta X_t < y_{hp}\right] = P\left[\frac{\beta X_t - \beta\mu_h}{\beta\sigma_h} < \frac{y_{hp} - \beta\mu_h}{\beta\sigma_h}\right]$$

de modo que:

$$\frac{y_{hp} - \beta\mu_h}{\beta\sigma_h} = \Phi^{-1}(p) \Rightarrow VaR_{hp} = \beta\left(\sigma_h\Phi^{-1}(1-p) - \mu_h\right)$$

(Nótese que hemos hecho la transformación habitual al pasar del percentil p al VaR).

Ejercicio: [EIV.1.7]). *A portfolio contains cash positions on two stocks: \$1 million is invested in a stock with a beta of 1.2, and \$2 million is invested in a stock with a beta of 0.8 with respect to a broad market index. If the excess returns on the index are i., i.d., and normally distributed with expectation 5% and volatility 20% per annum, what is the 1% 10-day VaR of the portfolio?* La beta nominal de la cartera es: \$1(1,2)+\$2(0,8)=\$2,8 m. La volatilidad de la rentabilidad en exceso sobre el horizonte de 10 dias es: $\sigma = 0,2\sqrt{10/250} = 0,04$ ó 4%, y la rentabilidad en exceso: μ = (0,05)(10/250)=0,02 ó 2%. El VaR es: (\$2,8m.)[(2,326)(0,04)-0,02] = \$254.951.

Ejemplo (Peña): Considere una cartera de 100 millones de euros invertida a partes iguales en BBVA, Santander y Telefonica. Sus volatilidades anualizadas son 29,41%, 28,16% y 32,94%, respectivamente. La matriz de correlaciones entre las rentabilidades es:

$$M = \begin{pmatrix} 1 & 0.64 & 0.33 \\ 0.64 & 1 & 0.57 \\ 0.33 & 0.57 & 1 \end{pmatrix}$$

Calcule los VaR 95% a un mes de cada acción (R:4,66%, 4,47%, 5,23%) y el VaR de la cartera (R:11,77%). Calcule el VaR de la cartera utilizando las betas de las acciones, que son 0,81, 1,18 y 1,86, respectivamente, y la volatilidad del índice, que es 11,85% (R:7,31%). ¿Por qué es inferior al VaR calculado directamente? Indicación: Para resolver la última parte debe calcular la beta de la cartera, como promedio ponderado de las betas de las acciones que la componen, utilizando como pesos las posiciones relativas de la cartera en cada acción.

VaR de acciones individuales:

$$\begin{aligned} 1/3 \times 0.2941 \times 1.65/\sqrt{12} &= 4.6695 \times 10^{-2} \\ 1/3 \times 0.2816 \times 1.65/\sqrt{12} &= 0.04471 \\ 1/3 \times 0.3294 \times 1.65/\sqrt{12} &= 5.2299 \times 10^{-2} \end{aligned}$$

Varianza de la cartera:

$$\begin{pmatrix} 1/3 & 1/3 & 1/3 \end{pmatrix} \begin{pmatrix} 0.2941 & 0 & 0 \\ 0 & 0.2816 & 0 \\ 0 & 0 & 0.3294 \end{pmatrix} \begin{pmatrix} 1 & 0.64 & 0.33 \\ 0.64 & 1 & 0.57 \\ 0.33 & 0.57 & 1 \end{pmatrix} .$$
$$. \begin{pmatrix} 0.2941 & 0 & 0 \\ 0 & 0.2816 & 0 \\ 0 & 0 & 0.3294 \end{pmatrix} \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix} = 6.1110 \times 10^{-2}$$

VaR de la cartera:

$$1.65\sqrt{6.1110 \times 10^{-2}}/\sqrt{12} = 0.11775$$

A partir del modelo unifactorial (13), tenemos:

$$\begin{aligned} V(Y_t) &= \beta^2 V(X_t) + V(\varepsilon_t) + 2Cov(X_t, \varepsilon_t) \Rightarrow \sigma_Y^2 = \beta^2\sigma_X^2 + \sigma_\varepsilon^2 + 2\rho\beta\sigma_X\sigma_\varepsilon \Rightarrow \\ \sigma_Y^2 &= (\beta\sigma_X + \sigma_\varepsilon)^2 - 2(1-\rho)\beta\sigma_X\sigma_\varepsilon \end{aligned}$$

donde ρ denota la correlación entre el factor y la rentabilidad residual. Por tanto:

$$\begin{aligned} [Volatilidad\ total]^2 &= [Volatilidad\ debida\ al\ mercado + volatilidad\ residual]^2 - \\ & 2(1-\rho)[Volatilidad\ debida\ al\ mercado][Volatilidad\ residual] \end{aligned}$$

pero como en el modelo Normal lineal, sin ajuste en media, el VaR se comporta como lo hace la volatilidad, tenemos una descomposición similar para el VaR, que es válida para horizontes de inversión cortos:

$$\begin{aligned} [VaR\ total]^2 &= [VaR\ sistemático + VaR\ específico]^2 - \qquad (14) \\ & -2(1-\rho)[VaR\ sistemático][VaR\ específico] \qquad (15) \end{aligned}$$

Esta expresión muestra que el VaR total es igual a la suma de los VaR específico y sistemático unicamente si ambos componentes del VaR están perfecta y positivamente correlacionados, lo que no debería suceder. Al contrario, si los factores están bien elegidos, se debería tener $\rho = 0$. En este caso, el VaR total, al cuadrado. sería igual a la suma de cuadrados de los VaR sistemático y específico.

Salvo en dichos casos, el VaR total será inferior a la suma de los VaR sistemático y específico,[13] lo que se debe a la sub-aditividad del modelo lineal paramétrico, que es una condición necesaria para que una medida de riesgo sea

[13] Si hacemos $\rho = 0$ en: $VaR\ total^2 = (VaR\ sistemático + VaR\ específico)^2 - 2(1-\rho)(VaR\ sistemático)(VaR\ específico)$, tenemos:
$VaR\ total^2 = (VaR\ sistemático^2 + VaR\ específico^2) + 2(VaR\ sistemático)(VaR\ específico) - 2(VaR\ sistemático)(VaR\ específico) =$
$= VaR\ sistemático^2 + VaR\ específico^2$.

coherente. Esto implica que el riesgo de invertir en una cartera no es superior al riesgo que resulta de una inversión de tamaño equivalente en cada uno de los valores que integran la cartera, lo que está relacionado con el efecto de diversificación de una cartera. La volatilidad de una cartera no es nunca superior a la volatilidad de una inversión análoga en uno de los activos de la cartera, y la volatilidad de la cartera disminuye según aumentan las correlaciones entre los activos constituyentes. Por tanto, para reducir el riesgo (medido por la volatilidad) los inversores tienen incentivo a diversificar su cartera. La subaditividad es un aspecto similar a la diversificación, desagregando la volatilidad de la cartera en sus componentes sistemático y específico.

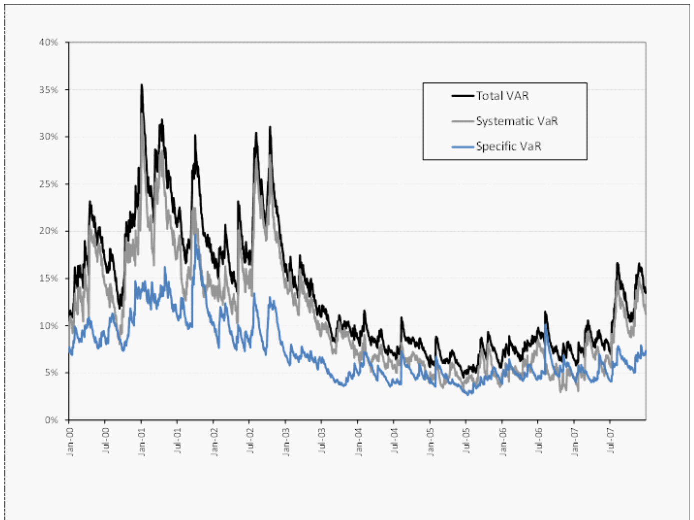
El riesgo específico Normal lineal de una cartera de accciones puede calcularse de tres modos distintos:

- utilizando los residuos de la estimación del modelo factorial y calcular el VaR específico directamente con la varianza residual
- calculando el VaR Normal lineal utilizando la varianza de la rentabilidad de la cartera. Calcular el VaR sistemático mediante $VaR_{sistemático} = \Phi^{-1}(1-p)\sqrt{\beta'\Omega\beta}$ y, posteriormente, calcular el VaR específico utilizando 14 [EIV.2.13]
- utilizando una regla estandarizada, como por ejemplo, fijando el riesgo específico en el 8% del valor de la cartera.

Si se utiliza MCO para estimar la matriz de covarianzas y las betas, los dos primeros procedimientos proporcionan el mismo resultado. Pero MCO no es necesariamente el mejor método, al representar estimaciones promedio sobre todo el horizonte temporal considerado. Si se utiliza EWMA para estimar las betas, entonces deberia utilizarse el segundo procedimiento. No es conveniente mezclar metodologías, por ejemplo utilizando MCO para estimar la matriz de covarianzas, y EWMA para estimar las betas. El enfoque EWMA genera estimaciones más sensibles al riesgo actual y a variaciones en el mismo, lo que será generalmente preferible.

El ejercicio [EIV.2.13] desagrega el VaR de una cartera de acciones en sus componentes sistemático y especifico. La Figura IV.2.5 representa la estimación del VaR específico utilizando betas calculadas mediante un esquema EWMA.

En los demás casos,
$VaR\ total^2 = (VaR\ sistemático + VaR\ específico)^2 - 2(1-\rho)(VaR\ sistemático)(VaR\ específico) <$
$<(VaR\ sistemático + VaR\ específico)^2$, por lo que $VaR\ total < VaR\ sistemático + VaR\ específico$



7.2 Componentes sistemático e idiosincrático del VaR de una cartera de acciones

En esta seción describimos, en una serie de ejemplos, cómo descomponer el riesgo sistemático en los VaR Componentes (stand-alone VaR) y los VaR marginales. A lo largo de los ejemplos, consideraremos diversas fuentes de riesgo.

Supongamos que tenemos una cartera de acciones en un mercado extranjero, en el que consideramos un único factor de riesgo, que pudiera ser el índice de dicho mercado. Tenemos además el tipo de cambio como un factor de riesgo adicional, con una sensibilidad exactamente igual a la posición que tengamos en la divisa extranjera. Queremos descomponer el VaR sistemático en sus componentes de Equity(acciones) y Forex (tipos de cambio). Suponemos por simplicidad que: 1) tanto el tipo de interés doméstico como el extranjero son cero, 2) la rentabilidad esperada descontada de la cartera sobre el horizonte de riesgo es despreciable, y unicamente necesitamos considerar en el cálculo del VaR sistemático la matriz de covarianzas de los factores de riesgo. Denotamos por R_h la rentabilidad en divisa doméstica, $\tilde{R}_h$ la rentabilidad de la cartera en la divisa, y X_h la rentabilidad del tipo de cambio en el período de h dias. La rentabilidad en divisa doméstica es:

$$R_h = \tilde{R}_h + X_h = \beta Y_h + X_h$$

donde Y_h denota la rentabilidad logarítmica del factor extranjero (índice de mercado) durante h dias. Por tanto,

$$\sigma_h = \sqrt{\beta^2\sigma^2_{Y_h} + \sigma^2_{X_h} + 2\rho\beta\sigma_{Y_h}\sigma_{X_h}} = \sqrt{(\beta\ 1)\,\Omega_h \begin{pmatrix} \beta \\ 1 \end{pmatrix}} =$$

$$= \sqrt{(\beta\ 1) \begin{pmatrix} \sigma^2_{Yh} & \rho\sigma_{Y_h}\sigma_{X_h} \\ \rho\sigma_{Y_h}\sigma_{X_h} & \sigma^2_{X_h} \end{pmatrix} \begin{pmatrix} \beta \\ 1 \end{pmatrix}}$$

de donde podemos obtener el VaR sistemático:

$$VaR\ sistemático_{h,p} = \Phi^{-1}(1-p)\sigma_h$$

y sus componentes:

$$\begin{aligned} Equity\ VaR_{h,p} &= \Phi^{-1}(1-p)\sqrt{\beta^2\sigma^2_{Y_h}} = \Phi^{-1}(1-p)\beta\sigma_{Y_h} \\ Forex\ VaR &= \Phi^{-1}(1-p)\sqrt{\sigma^2_{X_h}} = \Phi^{-1}(1-p)\sigma_{X_h} \end{aligned}$$

por lo que la expresión enterior nos permite escribir:

$$\sigma^2_h = (\beta\sigma_{Y_h} + \sigma_{X_h})^2 - 2\beta(1-\rho)\sigma_{Y_h}\sigma_{X_h}$$

es decir, multiplicando por el cuadrado de $\Phi^{-1}(1-p)$:

$$[VaR\ sistemático\ total]^2 = [VaR\ equity + VaR\ forex]^2 - 2(1-\rho)[VaR\ equity][VaR\ forex]$$

una relación similar a la que obtuvimos para la descomposición de l VaR de una cartera de renta variable en riesgo sistemático e idiosincrático.

Siguiendo un razonamiento similar al que hicimos en aquél caso:

$$VaR\ sistemático\ total \leq VaR\ equity + VaR\ forex$$

con igualdad si y solo si $\rho = 1$. Si la correlación quanto (correlación entre acciones y tipo de cambio) es negativa y elevada, el VaR sistemático podría ser menor que ambos componentes, el VaR equity y el VaR forex, como sucede en [EIV.2.14] para correlaciones negativas elevadas. En dicho ejercicio se desarrolla un ejemplo llevando a cabo los cálculos aqui explicados.

7.3 Descomposición en componentes marginales

Ya vimos que en el modelo paramétrico lineal, el vector gradiente del VaR puede escribirse:

$$g(\theta) = \frac{\Phi^{-1}(1-p)}{\sqrt{\theta'\Omega_h\theta}}\Omega_h\theta$$

por lo que a partir de la descomposición en VaR marginales se obtiene exactamente el VaR agregado:

$$\theta' g(\theta) = \Phi^{-1}(1-p)\sqrt{\theta'\Omega_h\theta}$$

En [EIV.2.14] se realiza este ejercicio. Hay dos factores de riesgo: Equity y Forex. La varianza de la cartera se calcula utilizando las betas respecto de los dos factores y las volatilidades de estos, lo que nos permite calcular el VaR total de los factores, o VaR sistemático total. Se dice "total" porque es el VaR de ambos factores, utilizados simultáneamente, no sumando sus efectos individuales. Si utilizamos la desviación tipica de cada factor y la beta de la cartera, calculamos el VaR sistemático de cada factor. Para el cálculo de los VaR marginales, el gradiente se calcula por la expresión que hemos dado arriba, donde el vector θ son las betas sobre los dos factores. Si se utiliza cada elemento del gradiente con la beta correspondiente, se obtiene el VaR marginal de cada factor. Como siempre, la suma de los VaR marginales es el VaR total de la cartera.

En [EIV.2.15] se analiza el VaR cuando una cartera está expuesta a varias divisas y se utiliza un indice general de mercado como factor de riesgo en cada país, teniendo en cuenta que:

$$\theta'\Omega_h\theta = \theta'_E\Omega_{Eh}\theta_E + \theta'_X\Omega_{Xh}\theta_X + 2\theta'_E\Omega_{EXh}\theta_X$$

donde $\theta = (\theta'_E, \theta'_X)$, siendo estas los vectores de sensibilidades de la cartera a los factores de riesgo en Equity y Forex, mientras que Ω_{EXh} es la matriz de covarianzas quanto, que sería generalmente negativa en este ejemplo. Se trata de la cartera de un fondo estadounidense, con posiciones en los mercados de renta variable de Reino Unido, Francia y Alemania. En cada uno de ellos mantiene una cartera, para las que considera un único factor de riesgo, representado por los indices del mercado respectivo: FTSE100, CAC40 y DAX30. Hay 6 factores de riesgo, siendo 4 de mercado y 2 de divisa (libra esterlina y euro). Primero calculamos las *betas netas* de la cartera multiplicando la beta respecto de cada factor por el nominal invertido en el mismo. Dadas las betas netas de la cartera, y tomando distintas submatrices de la matriz de covarianzas de los 6 factores, estimamos la varianza de la *P&L* debida a los 4 factores Equity, la debida a los 2 factores Forex, y la covarianza Quanto, que contiene las covarianzas entre un indice de renta variable y una divisa. La suma de ambas varianzas, más dos veces la covarianza nos da la varianza total de la $P\&L$. Al mismo resultado se llega utilizando los 6 betas y la matriz completa de covarianzas de los factores. Para el cálculo de los VaR marginales, primero obtenemos el producto de la matriz de covarianzas por el vector de betas. Luego, este vector se multiplica por $\Phi^{-1}(1-p)$ y se divide por la varianza total de la $P\&L$. Los VaR marginales se estiman multiplicando los componentes respectivos del vector gradiente por las betas correspondientes.

7.4 VaR cuando hay exposición a tipos de interés extranjeros

En el ejercicio [EIV.2.16] se calcula el VaR de la posición de un inversor estadounidense que compra por 2 millones de US$ una posición forward a 10 días en libras esterlinas, cuando el tipo de interés del Tesoro a 10 dias es 5% y el tipo de contado a 10 dias en UK es 4,5%. Estos tipos de interés tienen volatilidades de 100 puntos básicos (el tipo del Tesoro) y 80 puntos básicos (el tipo UK) con correlación 0,90.

El inversor soporta un riesgo de tipos de interés debido a los cash flows de +2 millones de US$ en el tipo UK y de -2 millones de US$ en el tipo de interés de EEUU. Debe comenzarse calculando el valor presente de 1 pb. de cada cash flow, para lo que se multiplica el valor nominal de cada cash-flow por el cambio que se produce en el factor descuento ante un descenso de 1 pb. en el tipo de interés correspondiente. Así se obtiene el vector: PV01=(5,47; -5,46) en US$, que se multiplica por la matriz de covarianzas a 10 dias, obteniendo un VaR de tipos de interés de $114, que es sin duda muy reducido, como es habitualmente el riesgo de tipos de interés. Sólo se tendría un riesgo de cuantía apreciable si el diferencial de tipos de interés fuese elevado y el plazo de la inversión fuese largo.

7.5 Cobertura de una cartera en acciones extranjeras

En [EIV.2.17] se considera un inversor europeo que ha invertido 5 millones de US$ en una cartera en S&P500, con una beta de 1,5 respecto del índice de mercado. Su horizonte de gestión es de 3 meses. Para cubrir esa posición comprada en dólares vende 5 millones de dólares invirtiendo el importe y asumiendo el riesgo del tipo de interés del euro a 3 meses. Para cubrir el riesgo Equity, vende futuros sobre S&P500 con vencimiento a 3 meses, por 7,5 millones de dólares, asumiendo el riesgo de dividendo correspondiente. Tiene, por tanto, una posición vendida global por 12,5 millones de US$, sobre la que asume riesgo de tipo de interés en dólares, a 3 meses.

8 VaR paramétrico bajo distribuciones de rentabilidad no Gaussianas

8.1 Contrastes de Normalidad: Jarque-Bera, Kolmogorov, QQ-plots

Analizamos en esta sección distintos procedimientos para tratar situaciones en que no podemos suponer que la distribución de probabilidad de la variable aleatoria que estamos analizando sea Normal. Un caso importante es aquél en que las innovaciones de un proceso estimado se consideran no Normales. Una situación en que esto tendría importancia decisiva sería cuando queremos simular dicho proceso para valorar un derivado que tenga como subyacente el activo cuya rentabilidad se ha modelizado.

En tal situación, la primera fase del problema es analizar, ya sea mediante contrastes estadísticos formales o por procedimientos informales, si la variable en cuestión tiene una distribución Normal. Para ello, junto a los contrastes de Normalidad habituales, del tipo Jarque-Bera, o contrastes no-paramétricos, del tipo Kolmogorov-Smirnov o de Fisher, existen los graficos QQ (quantile-quantile), en el que se representa los cuantiles de la muestra de una variable, contra los cuantiles que se obtendrían de una distribución Normal. Para ello, se ordena en orden creciente la muestra y se establece la red de valores i , $0 < i \leq T$. A continuación, el gráfico QQ se obtiene representando el cuantil $\frac{i-.5}{T}$ de la distribución de rentabilidades[14], en ordenadas, contra $\Phi^{-1}(\frac{i-.5}{T})$, en abscisas, siendo Φ la función de distribución de una Normal estándar, o contra $T_\nu^{-1}(\frac{i-.5}{T})$, en abscisas, siendo T_ν la función de distribución de una t-Student con ν grados de libertad.

Si se trabaja con rentabilidades, el contraste se aplicará generalmente a las rentabilidades estandarizadas mediante un modelo de volatilidad previamente estimado, siendo éstas las que se comparan con una Normal(0,1). Dado que la heterocedasticidad es tan habitual en series temporales de rentabilidades financieras, se supone inicialmente la existencia de heterocedasticidad, estimando un modelo para la misma y corrigendo de dicho efecto, pues el QQ-plot contrasta el ajuste con una distribución Normal de varianza constante (de hecho, de varianza unitaria).

8.2 VaR lineal bajo supuestos de t-Student.

La distribución t-Student se utiliza para intentar recoger la leptocurtosis de los datos de rentabilidades diarios. En rentabilidades observadas menos frecuentemente ésta tiende a desaparecer. Esto se debe a que si trabajamos con rentabilidades logarítmicas, la rentabilidad mensual es la suma de las rentabilidades diarias de ese mes, y el Teorema Central del Limite, nos dice que si las rentabilidades diarias son i.i.d., su agregado tenderá a comportarse como una Normal.

Debemos distinguir entre tres versiones de la distribución t-Student. La densidad t-Student estándar o habitual, que es la que generalmente se estudia en cursos de Estadística, tiene esperanza matemática igual a cero y varianza $\nu/(\nu-2)$, siendo ν el número de grados de libertad. La distribución *t-Student habitual* tiene función de densidad:

$$t_\nu(x) = \frac{1}{\sqrt{\nu\pi}}\Gamma\left(\frac{\nu}{2}\right)^{-1}\Gamma\left(\frac{\nu+1}{2}\right)\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

[14] El cuantil $\alpha\%$ de una distribución de probabilidad es el valor numérico del soporte de dicha distribución que deja a su izquierda una probabilidad menor o igual a $\alpha\%$. En variables aleatorias que toman valores numéricos discretos, tal definición puede estar sujerta a ambiguedades.

con esperanza igual a 0, varianza igual a $\nu/(\nu-2)$, y asimetría igual a cero. Su varianza no está definida si $\nu < 2$. Su exceso de curtosis es finito para $\nu > 4$: $\kappa = \frac{6}{\nu-4}$. Denotamos: $X \sim t_\nu$.

La *distribución t-Student estandarizada*, cuyas funciones de densidad y distribución denotaremos $\tilde{t}$, $\tilde{T}$, tiene esperanza 0 y varianza 1. Se obtiene a partir de una distribución *t-Student habitual* mediante la estandarización: $Y = X/\sigma = \sqrt{\frac{\nu-2}{\nu}}X$. Recordando que ante un cambio de variable, la nueva función de densidad es: $g(y) = f(x(y)).(dy/dx)^{-1}$, y como en este caso: $dy/dx = \sqrt{\frac{\nu-2}{\nu}}$, tenemos la función de densidad de la *t-Student estandarizada*:

$$\tilde{t}_\nu(y) = \frac{1}{\sqrt{(\nu-2)\pi}}\Gamma\left(\frac{\nu}{2}\right)^{-1}\Gamma\left(\frac{\nu+1}{2}\right)\left(1+\frac{y^2}{\nu-2}\right)^{-\frac{\nu+1}{2}}$$

que debe utilizarse con rentabilidades estandarizadas, al tener varianza igual a 1. Denotamos en este caso: $Y \sim \tilde{t}_\nu$. Nótese que existe toda una familia de distribuciones t-Student estandarizadas, que difieren en su número de grados de libertad ν, y que se aproximan a la distribución $N(0,1)$ al aumentar ν. En esta distribución no existe conexión entre la varianza y el número de grados de libertad.

Si se utiliza este modelo para calcular el VaR de una distribución de rentabilildades estandarizadas, una vez estimado el VaR, se deshace la estandarización multiplicando el VaR por la desviación tipica de las rentabilidades (o la media de las volatilidades condicionales) y sumando su valor medio en la muestra.

Por último, la distribución *t-Student generalizada*, tiene esperanza μ y varianza σ^2. Podemos pensar acerca de esta distribución como la resultante de hacer un cambio de variable sobre una distribución t-Student estandarizada Y del tipo: $W = \sigma Y + \mu$, o sobre una distribución t-Student habitual X mediante: $W = \sigma\sqrt{\frac{\nu-2}{\nu}}X + \mu$, por lo que su función de densidad resulta:

$$\tilde{t}_\nu(w) = \frac{1}{\sigma\sqrt{(\nu-2)\pi}}\Gamma\left(\frac{\nu}{2}\right)^{-1}\Gamma\left(\frac{\nu+1}{2}\right)\left(1+\frac{1}{\nu-2}\left(\frac{w-\mu}{\sigma}\right)^2\right)^{-\frac{\nu+1}{2}}$$

con esperanza μ, varianza σ^2, asimetria 0 y exceso de curtosis: $6/(\nu-4)$ (*ver pestaña Ex_II.4.5 en Examples II.4, de Alexander, donde se estima un modelo GARCH con innovaciones t-Student*). Al igual que sucede con la distribución t de Student habitual, al aumentar el número de grados de libertad, d , la distribución converge a una Normal(0,1). Denotaremos: $t_\nu(\mu,\sigma)$ y $T_\nu(\mu,\sigma)$ sus funciones de densidad y distribución.

Los cuantiles de una distribución se trasladan correctamente al aplicar transformaciones estrictamente monótonas: Si x_p es el p-cuantil de X, y aplicamos la transformación $Y = F(X)$, con $F' > 0$, entonces el p-cuantil de Y es: $y_p = F(x_p)$. Por tanto, si $t_\nu^{-1}(p)$ denota el p-cuantil de la distribución t-Student habitual, el p-cuantil de la distribución t-Student estandarizada es: $\tilde{t}_\nu^{-1}(p) = \sqrt{\nu^{-1}(\nu-2)}t_\nu^{-1}(p)$.

Remark 7 *Al calcular el VaR de una rentabilidad t-Student, debemos olvidar la conexión entre grados de libertad y varianza que existe en la distribución t-Student habitual. Por ello, si para estimar el VaR comenzamos invirtiendo la distribución t-Student habitual en el nivel de probabilidad p, entonces debemos transformar el cuantil resultante en el cuantil correspondiente de la distribución t-Student estandarizada, para luego multiplicar por la desviacion tipica muestral de la rentabilidad.*

Por otra parte, por simetría, los cuantiles de la distribución estándar satisfacen: $-t_\nu^{-1}(p) = t_\nu^{-1}(1-p)$ por lo que, aplicando un razonamiento similar al del caso Normal, tenemos:

$$VaR_{h,p}\ tStudent = \sigma\sqrt{\nu^{-1}(\nu-2)}t_\nu^{-1}(1-p) - \mu$$

Agregación temporal

La distribución t-Student no es una distribución estable, de modo que la suma de variables i., i.d., con esta distribucion no se distribuye t-Student. De hecho, por el teorema central del limite, dicha suma tenderia a comportarse como una Normal. Si el horizonte de riesgo, h, es reducido, podemos utilizar en el cálculo del VaR una expresión aproximada:

$$VaR_{h,p}\ tStudent = \sigma\sqrt{\nu^{-1}(\nu-2)h}t_\nu^{-1}(1-p) - h\mu$$

pero a partir de 10 dias, o incluso menos, si ν es relativamente alto, la aproximación Normal puede ser suficiente.

Si tenemos una cartera proyectada sobre un conjunto reducido de factores de riesgo, con vector de sensibilidades θ, tenemos el VaR sistemático:

$$VaR_{h,p}\ t-Student\ sistemático = \sqrt{\nu^{-1}(\nu-2)}t_\nu^{-1}(1-p)\sqrt{\theta'\Omega_h\theta} - \theta'\mu$$

donde Ω_h denota la matriz de covarianzas de las rentabilidades de los factores de riesgo y μ es su rentabilidad esperada.

8.2.1 La distribución t-Student en EXCEL

EXCEL no tiene programado los valores numéricos de las funciones de densidad y de distribución t-Student en un punto $t_v(x)$. Dispone de una función DISTR.T que no es dicha función de distribución $T_v(x)$, a pesar de su nombre. La instrucción tiene la estructura: $DISTR.T(x, v,' \mathrm{col}\, as')$, donde v es el número de grados de libertad. Si 'colas'=1, dicha instrucción devuelve: $P(X > x)$, siendo X una variable con distribución T_v, mientras que si 'colas'=2, la instrucción proporciona: $P(|X| > x)$. Por tanto, para calcular el valor numérico de la función de distribución t-Student en un punto, operamos del siguiente modo:

$$\begin{aligned} si\, x &> 0,\ T_v(x) = 1 - DISTR.T(x, v, 1) \\ si\, x &< 0,\ T_v(x) = DISTR.T(-x, v, 1) \end{aligned}$$

de modo que si queremos calcular el valor numérico de la función de distribución $T_v(x)$ en el punto $x = 2,356$, calcularemos $1 - DISTR.T(2,356; v, 1)$, lo que nos proporciona la probabilidad $1 - P(X > 2,356)$, que es,efectivamente, por simetria, igual a: $P(X < 2,356) = T_v(2,356)$.

Por otra parte, si queremos calcular el valor numérico de la función de distribución $T_v(x)$ en el punto $x = -2,356$, calcularemos $DISTR.T(2,356; v, 1)$, que proporciona $P(X > 2,356)$ que, nuevamente por simetria, es igual a $P(X < -2,356) = T_v(-2,356)$.

Para calcular la inversa de la función de distribución t-Student, lo que utilizaremos frecuentemente tanto en el cálculo del VaR como en el trabajo con Cópulas, utilizamos la función EXCEL: $DISTR.T.INV(u, v)$ con $0 < u < 1$ y donde v es un número entero positivo. Si no es entero, EXCEL lo trunca al número entero más próximo. Esta función proporciona el valor numérico t_u: $DISTR.T.INV(u, v) = t_u$ tal que: $P(|X| > t_u) = u$. Por tanto, esta no es realmente la inversa de la función de distribución t-Student que, sin embargo, se puede obtener del siguiente modo:

$$\begin{array}{rcl} (habitual)\ si\ u & \leq & 1/2,\ T_v^{-1}(u) = -DISTR.T.INV(2u, v) \\ si\ u & > & 1/2,\ T_v^{-1}(u) = DISTR.T.INV(2(1-u), v) \end{array}$$

Por ejemplo, si queremos calcular el percentil 0,95 de la distribución, que será positivo, buscaremos $DISTR.T.INV(0,10; v)$, mientras que si queremos calcular el percentil 0,05, que será negativo, buscaremos tambien $DISTR.T.INV(0,10; v)$ y lo cambiaremos de signo.

8.2.2 Estimación de la densidad t de Student

Estimación por Máxima Verosimilitud Si modelizamos las rentabilidades como,

$$r_t = \sigma_t z_t$$

con $z_t \sim \tilde{t}(v)$, e ignoramos el hecho de que la serie temporal de varianzas es una estimación sujeta a error estadístico, podemos tratar el rendimiento estandarizado como una única variable aleatoria. Comenzaríamos con un modelo de volatilidad para generar estimaciones de σ_t, y generar a partir de dicha serie temporal la rentabilidad estandarizada: $z_t = r_t/\sigma_t$. Al tener z_t una varianza unitaria, podemos utilizar la distribución t de Student estandarizada, y tenemos la verosimilitud,

$$\begin{aligned} \ln L_1 &= \sum_{t=1}^{T} \ln\left(\tilde{t}_\nu(z)\right) = T\left[\ln\Gamma\left(\frac{\nu+1}{2}\right) - \ln\Gamma\left(\frac{\nu}{2}\right) - \frac{\ln\pi}{2} - \frac{1}{2}\ln(\nu-2)\right] - \\ &\quad -\frac{1+\nu}{2}\sum_{t=1}^{T}\ln\left(1+\frac{z_t^2}{\nu-2}\right) \end{aligned}$$

Este sería un procedimiento de Quasi-máxima Verosimilitud, al estimar por separado los parámetros del modelo de varianza, que se utilizan para estandarizar las rentabilidades, estimando luego el parámetro de grados de libertad de la función de densidad estandarizada utilizando la verosimilitud anterior.

Si, por el contrario, queremos estimar el parámetro ν simultáneamente con los parámetros de los modelos de varianza, debemos ajustar la distribución para tener en cuenta la varianza. Para ello, suponiendo que las rentabilidades tienen esperanza nula, utilizariamos la distribución t de Student generalizada, ya que la varianza no es unitaria, teniendo, para un valor $\nu > 2$:

$$t_\nu(r_t,\sigma_t^2) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi(\nu-2)\sigma_t^2}}\left(1+\frac{1}{\nu-2}\frac{{r_t}^2}{\sigma_t^2}\right)^{-\frac{1+\nu}{2}}, \ \nu > 2,$$

y, por tanto, la función de verosimilitud,

$$\ln L_2 = \sum_{t=1}^{T} \ln\left(t_\nu(r_t,\sigma_t^2)\right) = \sum_{t=1}^{T} \ln\left(\tilde{t}_\nu(z_t)\right) = \ln L_1 - \sum_{t=1}^{T} \frac{\ln\sigma_t^2}{2}$$

Por ejemplo, supongamos que tratamos con un único activo (quizá una cartera de activos, cuyas ponderaciones se han mantenido constantes durante el período muestral), cuya rentabilidad sigue un proceso GARCH(1,1) con *asimetría*, del tipo:

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

El logaritmo de la función de verosimilitud, que se trataría de maximizar, sería entonces:

$$\begin{aligned}
\ln L_1 \quad = \quad & \sum_{t=2}^{T} \ln\left(t_\nu(r_t,\sigma_t^2)\right) = T\left[\ln\Gamma\left(\frac{\nu+1}{2}\right) - \ln\Gamma\left(\frac{\nu}{2}\right) - \frac{\ln\pi}{2} - \frac{1}{2}\ln(\nu-2)\right] - \\
& -.5\sum_{t=2}^{T} \ln\left(\omega + \alpha\sigma_{t-1}^2(r_{t-1}-\theta)^2 + \beta\sigma_{t-1}^2\right) - .5(1+\nu). \\
& .\sum_{t=2}^{T} \ln\left(1 + \frac{1}{\nu-2}\frac{r_t^2}{\omega + \alpha\sigma_{t-1}^2(r_{t-1}-\theta)^2 + \beta\sigma_{t-1}^2}\right)
\end{aligned}$$

ignorando, en todo caso, la primera observación. El algoritmo numérico de cálculo de la función de verosimilitud debe inicializarse con una elección para σ_1, para lo que puede utilizarse la varianza incondicional a lo largo del período muestral, aunque esta puede ser una opción discutible en algunos casos. Alternativamente, puede estimarse σ_1 utilizando un número inicial de observaciones, que luego no se utilizan en la estimación del modelo. [ver Christoffersen Chapter4Results.Question 5]

Estimación por el Método de Momentos Teniendo en cuenta la expresión de la curtosis que antes vimos para la distribución t de Student de una serie de rentabilidades estandarizadas, podemos utilizar la lógica del método de momentos para estimar el número de grados de libertad de dicha distribución mediante,

$$\nu = \frac{6}{EC} + 4$$

siendo EC el exceso de curtosis muestral.

Si trabajamos con rentabilidades sin estandarizar, podemos utilizar la expresión anterior del exceso de curtosis, pero también la expresión de la varianza, que conduce a:

$$\nu = \frac{2\sigma^2}{\sigma^2 - 1}$$

y podría plantearse alguna estrategia para encontrar el valor de v que satisface con mayor aproximación ambas ecuaciones.

[EIV.2.18], [EIV.2.19].

8.2.3 QQ plots para distribuciones t de Student

Es frecuente disponer de cuantiles para la distribución t-Student habitual, pero no para la distribución t-Student estandarizada, la cual es frecuente utilizar cuando se trabaja con rentabilidades estandarizadas.

En esa situación, conviene recordar que la función de densidad $\tilde{t}_v$ de la distribución t-Student estandarizada puede obtenerse a partir de la función de densidad t_v de la distribución t-Student habitual mediante el cambio de variable:

$$\tilde{y} = y\sqrt{\frac{\nu - 2}{\nu}},$$

donde y sigue una distribución t_v habitual, e $\tilde{y}$ sigue una distribución $\tilde{t}_v$ estandarizada, con $\tilde{t}_\nu(\tilde{y}) = t_\nu(y)\sqrt{\frac{\nu}{\nu-2}}$. La relación entre sus cuantiles es sencilla. El p-cuantil de la distribución t de Student estandarizada es el valor numérico $\tilde{q}$ definido mediante: $p = P\left(\tilde{y} < \tilde{q}\right)$ o, lo que es lo mismo, $\tilde{q} = \tilde{T}_v^{-1}(p)$. Por tanto:

$$p = P\left(\tilde{y} < \tilde{q}\right) = P\left(y\sqrt{\frac{\nu - 2}{\nu}} < \tilde{q}\right) = P\left(y < \tilde{q}\sqrt{\frac{\nu}{\nu - 2}} \equiv q\right),$$

de modo que:

$$q = T_\nu^{-1}(p) = \tilde{q}\sqrt{\frac{\nu}{\nu - 2}} = \tilde{T}_\nu^{-1}(p)\sqrt{\frac{\nu}{\nu - 2}}$$

Por tanto, los cuantiles de la distribución estandarizada $\tilde{t}_\nu$ pueden calcularse, en función de los cuantiles análogos de la distribución t-Student habitual, t_ν

utilizando la relación: $\tilde{T}_\nu^{-1}(p) = \sqrt{\frac{\nu-2}{\nu}} T_\nu^{-1}(p)$, y el QQ-plot para juzgar la adecuación del ajuste proporcionado por una densidad $\tilde{t}_v$ puede construirse tomando en abscisas los valores numéricos $\sqrt{\frac{\nu-2}{\nu}} T^{-1}_{\nu,\frac{i-.5}{n}}(p)$, siendo n el tamaño muestral y tomando en ordenadas las rentabilidades estandarizadas, z_i.

Tambien vimos que la distribución t-Student estandarizada puede obtenerse a partir de la distribución t-Student generalizada mediante el cambio de variable:

$$\tilde{y} = \frac{z-\mu}{\sigma}\sqrt{\frac{\nu-2}{\nu}},$$

donde z sigue una distribución generalizada $T_\nu(\mu,\sigma)$, e $\tilde{y}$ sigue una distribución $\tilde{T}_\nu$ estandarizada, con $\tilde{t}_\nu(\tilde{y}) = t_\nu(z;\mu,\sigma)\sigma\sqrt{\frac{\nu}{\nu-2}}$. El p-cuantil de la distribución t de Student estandarizada es el valor numérico $\tilde{q}$ definido mediante: $\tilde{q} = \tilde{T}_\nu^{-1}(p)$.

$$p = P\left(\tilde{y} < \tilde{q}\right) = P\left(y < \tilde{q}\sigma\sqrt{\frac{\nu}{\nu-2}} + \mu\right)$$

Por tanto, tenemos:

$$q = T_\nu^{-1}(\mu,\sigma;p) = \tilde{T}_\nu^{-1}(p)\sigma\sqrt{\frac{\nu}{\nu-2}} + \mu$$

Por tanto, los cuantiles de la distribución estandarizada $\tilde{t}_\nu(d)$ pueden calcularse, en función de los cuantiles análogos de la distribución t-Student habitual, $t_\nu(d)$ utilizando la relación: $\tilde{T}_\nu^{-1}(p) = \frac{1}{\sigma}\sqrt{\frac{\nu-2}{\nu}}\left[T_\nu^{-1}(p) - \mu\right]$, y el QQ-plot para juzgar la adecuación del ajuste proporcionado por una densidad $\tilde{t}(d)$ puede construirse tomando en abscisas los valores numéricos $\sigma^{-1}\sqrt{\frac{\nu-2}{\nu}} T^{-1}_{\nu,\frac{i-.5}{n}}$ y, en ordenadas, las rentabilidades estandarizadas, z_i.

8.3 VaR lineal bajo mixturas de distribuciones

Las mixturas tratan de recoger diferentes regímenes de mercado, y son distribuciones que pueden generar asimetrias y curtosis importantes, lo que permite ajustar rentabilidades cuyo nivel de curtosis no podría explicarse utilizando una distribución t-Student. Una mixtura de dos distribuciones Normales Φ_1, Φ_2, viene definida por:

$$G(x) = \pi\Phi_1(x;\mu_1,\sigma_1^2) + (1-\pi)\Phi_2(x;\mu_2,\sigma_2^2), 0 < \pi < 1$$

y diferenciando, tenemos una relación similar entre sus funciones de densidad:

$$g(x) = \pi\varphi_1(x;\mu_1,\sigma_1^2) + (1-\pi)\varphi_2(x;\mu_2,\sigma_2^2), 0 < \pi < 1$$

Para distribuciones F_1, F_2 no Normales basta que sustituyamos enlas expresiones anteriores $\left(\mu_1, \sigma_1^2\right)$ por el vector de parámetros relevantes en cada caso y Φ, φ por las funciones de distribución y de densidad de cada caso.

En el caso de una mixtura de distribuciones Normales, sus momentos son:

$$\begin{aligned}
\mu &= \pi\mu_1 + (1-\pi)\mu_2 \\
\sigma^2 &= \pi\sigma_1^2 + (1-\pi)\sigma_2^2 \\
sk &= 0 \\
\kappa &= 3\frac{\pi\sigma_1^4 + (1-\pi)\sigma_2^4}{\left[\pi\sigma_1^2 + (1-\pi)\sigma_2^2\right]^2}
\end{aligned}$$

Si X denota la variable cuya distribución es la mixtura, tenemos:

$$P(X < x_p) = G(x_p) = \pi F_1(x_p; \mu_1, \sigma_1^2) + (1-\pi)F_2(x_p; \mu_2, \sigma_2^2), 0 < \pi < 1$$

por lo que fijado un nivel de significación p, el VaR es el nivel de rentabilidad x_p que satisface esta ecuación.

En el caso de una mixtura de dos distribuciones Normales X_1, X_2, el VaR de una mixtura es el valor numérico x_p que resuelve la ecuación:

$$p = \pi P\left(Y_1 < \frac{x_p - \mu_1}{\sigma_1}\right) + (1-\pi)P\left(Y_2 < \frac{x_p - \mu_2}{\sigma_2}\right)$$

cambiado de signo, donde $Y_i = \frac{X_i - \mu_i}{\sigma_i}, i = 1, 2$, son variables Normales estándar. Si F_1, F_2 fuesen distribuciones t-Student, las variables $Y_i, i = 1, 2$ tendrían una distribución $\tilde{t}$ estandarizada.

El razonamiento se extiende sin dificultad a mixturas de más de dos distribuciones, aunque es raramente necesario utilizar más de dos distribuciones. Como es habitual, el VaR se expresa en porcentaje del valor de la cartera si μ y σ son la esperanza y desviación típica de la distribución de rentabilidades, y se expresa en términos nominales si dichos momentos se refieren a la esperanza y desviación típica de la variable de Pérdidas y Ganancias.

La estimación del modelo de mixturas puede hacerse por el algoritmo EM, o por el Método Generalizado de Momentos (GMM). Para aplicar este último en el caso de una mixtura de distribuciones Normales, hay que utilizar las expresiones analiticas de los cuatro primeros momentos no centrales de la mixtura:

$$
\begin{aligned}
M_1 &= E(X) = \sum_{i=1}^{m} \pi_i \mu_i \\
M_2 &= E(X^2) = \sum_{i=1}^{m} \pi_i \left(\sigma_i^2 + \mu_i^2\right) \\
M_3 &= E(X^3) = \sum_{i=1}^{m} \pi_i \left(3\mu_i \sigma_i^2 + \mu_i^3\right) \\
M_4 &= E(X^4) = \sum_{i=1}^{m} \pi_i \left(3\sigma_i^4 + 6\mu_i^2 \sigma_i^2 + \mu_i^4\right)
\end{aligned}
$$

y la esperanza, varianza, asimetria y curtosis, en función de los mismos resultan:

$$
\begin{aligned}
\mu &= E(X) = M_1 \\
\sigma^2 &= E\left[(X-\mu)^2\right] = M_2 - M_1^2 \\
\tau &= \sigma^{-3} E\left[(X-\mu)^3\right] = \sigma^{-3}\left(M_3 - 3M_1 M_2 + 2M_1^3\right) \\
\kappa &= \sigma^{-4} E\left[(X-\mu)^4\right] = \sigma^{-4}\left(M_4 - 4M_1 M_3 + 6M_1^2 M_2 - 3M_1^4\right)
\end{aligned}
$$

El procedimiento GMM consiste en este caso en igualar los valores numéricos de los cuatro momentos muestrales $(\mu, \sigma^2, \tau, \kappa)$ a las expresiones que arriba se muestran, buscando los valores paramétricos que minimicen la distancia entre ambos vectores. Primero se calculan los momentos respecto del origen M_1, M_2, M_3, M_4 en funcion de pre-estimaciones de los parámetros de la mixtura: $(\pi, \mu_1, \sigma_1, \mu_2, \sigma_2)$. Luego, se calculan los momentos poblacionales en función de los momentos respecto del origen, y se minimiza su distancia respecto de los momentos análogos muestrales. Las esperanzas y desviaciones típicas estimadas se elevan al horizonte de riesgo que se desea, y se calcula el VaR resolviendo la ecuación implícita que antes vimos. [EIV.2.20] parte de una mixtura de Normales estimada para calcular el VaR de cada factor (stand-alone) bajo dicha mixtura. [Case Study I.5.4.4] [EIV.2.21] estima una mixtura por el Método de Momentos y calcula el VaR resultante, comparandolo con el obtenido bajo supuestos de t-Student y de Normal. [EIV.2.22] compara el VaR obtenido bajos supuestos de mixtura de Normales y mixtura de t-Student.[EIV.2.23] analiza el efecto sobre el VaR del ajuste por autocorrelación de ls rentabilidades.

Supongamos ahora una situación en la que tenemos una secuencia de cash-flows proyectada sobre dos factores, cada uno de los cuales sigue una mixtura de Normales, posiblemente con un régimen frecuente, de cierta estabilidad, en el que se encuentra con probabilidad $1 - \pi_i$ y un régimen, menos frecuente, de inestabilidad, en el que se encuentra con probabilidad π_i :

$$
f_1(x_1) = \pi_1 f(x_1; \mu_{11}, \sigma_{11}^2) + (1-\pi_1) f(x_1; \mu_{12}, \sigma_{12}^2), 0 < \pi_1 < 1
$$

$$f_2(x_2) = \pi_2 f(x_2; \mu_{12}, \sigma_{21}^2) + (1-\pi_2) f(x_2; \mu_{22}, \sigma_{22}^2), 0 < \pi_1 < 1$$

La densidad conjunta es una mixtura Normal bivariante de la forma:

$$f(x_1, x_2) = \pi_1 \pi_2 F(x_1, x_2; \mu_1, \Omega_1) + (1-\pi_1)\pi_2 F(x_1, x_2; \mu_2, \Omega_2) +$$

$$+\pi_1(1-\pi_2) F(x_1, x_2; \mu_3, \Omega_3) + (1-\pi_1)(1-\pi_2) F(x_1, x_2; \mu_4, \Omega_4), 0 < \pi_1, \pi_2 < 1$$

con:

$$\mu_1 = \begin{pmatrix} \mu_{11} \\ \mu_{21} \end{pmatrix};\ \mu_2 = \begin{pmatrix} \mu_{12} \\ \mu_{21} \end{pmatrix};\ \mu_3 = \begin{pmatrix} \mu_{11} \\ \mu_{22} \end{pmatrix};\ \mu_4 = \begin{pmatrix} \mu_{12} \\ \mu_{22} \end{pmatrix}$$

$$\begin{aligned} \Omega_1 &= \begin{pmatrix} \sigma_{11}^2 & \rho_1 \sigma_{11} \sigma_{21} \\ \rho_1 \sigma_{11} \sigma_{21} & \sigma_{21}^2 \end{pmatrix};\ \Omega_2 = \begin{pmatrix} \sigma_{12}^2 & \rho_2 \sigma_{12} \sigma_{21} \\ \rho_2 \sigma_{12} \sigma_{21} & \sigma_{21}^2 \end{pmatrix}; \\ \Omega_3 &= \begin{pmatrix} \sigma_{11}^2 & \rho_3 \sigma_{11} \sigma_{22} \\ \rho_3 \sigma_{11} \sigma_{22} & \sigma_{22}^2 \end{pmatrix};\ \Omega_2 = \begin{pmatrix} \sigma_{12}^2 & \rho_4 \sigma_{12} \sigma_{22} \\ \rho_4 \sigma_{12} \sigma_{22} & \sigma_{22}^2 \end{pmatrix} \end{aligned}$$

donde Ω_1 representa las volatilidades y correlaciones en las colas de las dos distribuciones, y Ω_4 las volatilidades y correlaciones en la parte central (estable) de ambas distribuciones. Las otras dos matrices representan volatilidades y correlaciones cuando un factor de riesgo está en la cola y el otro está en la parte central de su distribución. Los parámetros $\rho_i, i = 1, ..., 4$ denotan las correlaciones en funcion de las regiones en que se encuentren ambas rentabilidades.

La cartera tiene como distribución una mixtura de Normales con cuatro componentes, y probabilidades de mezcla: $\pi = (\pi_1\pi_2, (1-\pi_1)\pi_2, \pi_1(1-\pi_2), (1-\pi_1)(1-\pi_2))$.Los componentes tienen medias: $\left\{\theta'\mu_1, \theta'\mu_2, \theta'\mu_3, \theta'\mu_4\right\}$ y matrices de covarianzas: $\left\{\theta'\Omega_1\theta, \theta'\Omega_2\theta, \theta'\Omega_3\theta, \theta'\Omega_4\theta\right\}$. Esta mixtura se puede estimar tratándola como una mixtura de cuatro distribuciones Normales.

[El ejercicio [EIV.2.24] parte de una mixtura de este tipo, ya estimada, para calcular el VaR. *Análisis VaR Normal lineal para carteras de futuros en commodities* (*Case Study*)].

8.4 Expected Tail Loss (conditional VaR) bajo diferentes distribuciones de probabilidad

La Expected Tail Loss (ETL) se define:

$$ETL_p(X) = -E\left(X \mid X < x_p\right) = -p^{-1} \int_{-\infty}^{x_p} x f(x) dx$$

Como se trata de una probabilidad condicional, hemos dividido el promedio ponderado de los valores inferiores a x_p, por la probabilidad de que un valor de X sea inferior a x_p.

En el modelo normal lineal, sea X la rentabilidad aleatoria de una cartera sobre un horizonte de h días. Si $X \sim N(\mu_h, \sigma_h^2)$, :

$$ETL_p(X) = p^{-1}\varphi\left(\Phi^{-1}(p)\right)\sigma_h - \mu_h,$$

donde φ y Φ denotan las funciones de densidad y distribución de una variable $N(0,1)$.[ver ejemplo de cálculo en EIV.2.27]

Remark 8 *Nótese que la expresión anterior está escrita para generar una ETL positiva, y el cuantil p sería 0.001, 0.01, 0.05. Lo mismo sucede en las expresiones que se proporcionan para otras distribuciones. Si la ETS se calculara como una pérdida, con signo negativo, μ_h deberia entrar con signo +.*

Este resultado ya se probó en la seccion 1.9.1 de estas notas. Otra manera de probarlo es, para el caso $h = 1$: supongamos una variable $Z \sim N(0,1)$. Tenemos que el $VaR(p)$ es la rentabilidad $z_p = \Phi^{-1}(p)$ que satisface:

$$\begin{aligned} ETL(Z) &= -p^{-1}\int_{-\infty}^{\Phi^{-1}(p)} zf(z)dz = -\frac{1}{p\sqrt{2\pi}}\int_{-\infty}^{\Phi^{-1}(p)} ze^{-z^2/2}dz = -\frac{1}{p\sqrt{2\pi}}\left[e^{-z^2/2}\right]_{-\infty}^{\Phi^{-1}(p)} = \\ &= p^{-1}\varphi\left(\Phi^{-1}(p)\right) \end{aligned}$$

Si consideramos una variable $X \sim N(\mu, \sigma^2)$, tenemos: $X = \sigma Z + \mu$, de modo que:

$$\begin{aligned} ETL(X) &= -p^{-1}\int_{-\infty}^{\Phi^{-1}(p)} (\sigma z + \mu)f(z)dz = -\sigma p^{-1}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\Phi^{-1}(p)} zf(z)dz - p^{-1}\mu\int_{-\infty}^{\Phi^{-1}(p)} f(z)dz = \\ &= -\sigma p^{-1}\varphi\left(\Phi^{-1}(p)\right) - \mu, \text{ ya que: } \int_{-\infty}^{\Phi^{-1}(p)} f(z)dz = p \end{aligned}$$

En el modelo t-Student, *puede probarse* que:

$$ETL_p(X) = \frac{1}{p(\nu-1)}[\nu - 2 + x_p(\nu)^2]f_\nu(x_p(\nu))\sigma_h - \mu_h,$$

donde $f_\nu(x_p(\nu))$ denota la densidad de una t-Student estandarizada (esperanza=0, varianza=1) con ν grados de libertad y $x_p(\nu)$ es el $p-$cuantil de la distribución t-Student estandarizada. [EIV.2.28].

Supongamos que la rentabilidad descontada de una cartera a lo largo de h dias sigue una mixtura Normal G_0 de distribuciones de probabilidad todas ellas con esperanza matemática igual a cero. La función de densidad de la mixtura es: $\sum_{i=1}^{n} \pi_i f_i(x)$ donde cada $f_i(x)$ es una densidad Normal con esperanza matemática igual a cero y desviación típica σ_{ih}.Tenemos: $x_p = G_0^{-1}(p)$, de modo que $-x_p$ es el VaR $100p\%$ a h dias, y:

$$ETL_p(X) = -p^{-1} \sum_{i=1}^{n} \pi_i \int_{-\infty}^{x_p} x f_i(x) dx = p^{-1} \sum_{i=1}^{n} \left[\pi_i \sigma_{ih} \varphi \left(\sigma_{ih}^{-1} x_p \right) \right]$$

Si la rentabilidad descontada siguiera una mixtura $X \sim NM(\pi, \mu_h, \sigma_h^2)$, donde π, μ_h, σ_h^2 denotan el vector de probabilidades de mezcla y los vectores de esperanzas matemáticas (no nulas) y varianzas de cada componente de la mixtura. En este caso, el valor esperado de la mixtura es $\sum_{i=1}^{n} \pi_i \mu_{ih}$ y tenemos:

$$ETL_p(X) = -p^{-1} \sum_{i=1}^{n} \pi_i \int_{-\infty}^{x_p} x f_i(x) dx = p^{-1} \sum_{i=1}^{n} \left[\pi_i \sigma_{ih} \varphi \left(\sigma_{ih}^{-1} x_p \right) \right] - \sum_{i=1}^{n} \pi_i \mu_{ih}$$

donde x_p es el VaR $100p\%$ a h dias obtenido bajo la mixtura Normal análoga con esperanza cero. [EIV.2.29] [EIV.2.30].

Cuando la distribución de rentabilidades es una mixtura de distribuciones t-Student, *puede probarse* que

$$ETL_p(X) = p^{-1} \sum_{i=1}^{n} \left(\pi_i \frac{1}{\nu_i - 1} [\nu_i - 2 + t_{ip}(\nu)^2] f_{\nu_i}(t_{ip}(\nu)) \sigma_{ih} \right) - \sum_{i=1}^{n} \pi_i \mu_{ih}$$

donde ν denota el vector compuesto por los grados de libertad de cada componente de la mixtura, $t_{ip}(\nu) = x_p(\nu) \nu_i^{-1} (\nu_i - 2) \sigma_{ih}^{-1}$, y $x_p(\nu)$ es el VaR de la mixtura Student, cambiado de signo.

8.5 VaR paramétrico lineal y ETL para spreads de crédito [Case Study IV.2.12]

Este caso muestra el enorme *riesgo de modelo* que puede existir en el cálculo del VaR, incluso si utilizamos una misma metodologia (VaR Normal lineal paramétrico, la misma muestra, y el mismo modelo de riesgo (un único factor, el spread de crédito).

El factor de riesgo que utilizamos es el índice iTraxx a 5 años, un índice equiponderado de spreads de CDS, medido en puntos básicos y construido sobre 125 empresas the rating *investment grade*. Cada 6 meses se cambia el índice, sustituyendo las empresas que han hecho default, se han fusionado con otras, han cambiado de sector o han sido bajadas de rating, por las más liquidas que no formaban parte del índice, cumpliendo sus condiciones. En el periodo 21 de junio de 2004 a 10 de abril de 2008, la desviación típica de las variaciones diarias del índice fue 2.4037, lo que significa una volatilidad anual de 38 puntos básicos por año. Sin embargo, si se tiene en cuenta que el índice tiene una autocorrelación positiva y significativa, de 0.1079, la volatilidad anual es 41.5 puntos básicos. Como muestra la Tabla IV.2.39, el VaR estimado para una exposición de 1000 euros puede oscilar entre 17.683 y 43.784 euros, mientras el ETL oscila entre 20.250 y 47.522 euros.

9 Cálculo del VaR histórico

Este es el segundo de los procedimientos para el cálculo del VaR. El tercer método sería el de simulación Monte Carlo que se ha estado viendo a lo largo del curso, y que requiere la estimación de un modelo de series temporales, ya sea univariante (como ARIMA) o multivariante (como VAR).

Las ventajas del método histórico para la estimación del VaR:

- No precisa hacer supuestos acerca de la forma paramétrica de la distribución de rentabilidades de los factores de riesgo o de la cartera,

- A diferencia del modelo paramétrico, que es de una sola etapa (calcula el VaR al horizonte deseado únicamente), el método de simulación puede utilizarse para el cálculo del VaR en activos cuya rentabilidad es *path-dependent*. Puede asimismo acomodar *volatility clustering*, a diferencia del método paramétrico, que precisa independencia de las rentabilidades diarias. El método Monte Carlo también puede incorporar volatility clustering, pero necesita hacer algún supuesto acerca de la evolución temporal de los factores de riesgo (un modelo VAR-GARCH, por ejemplo).

- No está limitado a carteras en las que los pagos tienen una estructura lineal, por lo que puede utilizarse en carteras que contengan opciones u otro activos con estructuras de pagos no lineales. Tambien el método Monte Carlo puede hacerlo, pero sus resultados están condicionados por el modelo que se establezca para las rentabilidades de los factores.

- El cálculo del VaR debe utilizar períodos de h dias *no solapados*, por lo que si h es largo, las necesidades de datos pueden ser enormes. En este caso es conveniente poder disponer de datos intradía. Por eso es que el método histórico debe utilizarse solo para el cálculo del VaR a un horizonte de muy pocos dias. Para intervalos más amplios se aplica un factor de escala para transformar el VaR a 1 dia en un VaR a h dias, por lo que deben cumplirse los supuestos que justifican tal extensión. Además, la aplicación del factor de escala supone que la cartera se rebalancea durante el horizonte de riesgo de modo que la sensibilidad a los factores de riesgo sea la misma que cuando se estima el VaR.

- Supone que la cartera que hoy se ha escogido es la misma cartera que se habría escogido en cualquier momento en el pasado. Esto se debe a que para su cálculo hemos de utilizar series temporales históricas de rentabilidades para el cálculo de momentos de la cartera, utilizando su composición actual. Tambien el método paramétrico y el Monte Carlo necesitan matrices de correlaciones y otros momentos, y pueden utilizar datos historicos para su estimacion, pero tambien pueden hacer hipotesis (escenarios) al respecto.

- En general, parece que debería preferirse el método de Monte Carlo, que permite simular el comportamiento de la cartera sobre distintos escenarios.

Por el contrario, el enfoque histórico supone que el escenario futuro será el mismo escenario que en el pasado.

El VaR $100p\%$ sobre h dias es el p-cuantil de la distribución empírica de $P\&L$ descontadas, si se expresa en términos nominales, o el p-cuantil de la distribución empírica de rentabilidades, si se expresa como un porcentaje del valor de la cartera. Con carteras que pueden tener posiciones cortas, hemos de trabajar en términos nominales con las $P\&L$, puesto que el concepto de rentabilidad no está entonces muy bien definido.

9.1 Extrapolación temporal de la varianza: Escalado exponencial

Dadas las dificultades que presenta el trabajo con muestras no solapadas, conviene considerar la alternativa de extrapolar el comportamiento de las rentabilidades hacia el futuro.

La regla de la raiz cuadrada en la extrapolación temporal de la varianza no es válida si las rentabilidades no son independientes e identicamente distribuidas. Cuando lo son, si además siguen una distribución Normal, entonces el VaR sigue la misma regla de extrapolación temporal que la desviación tipica (suponiendo rentabilidades esperadas nulas). En el método histórico, no trabajamos con distribuciones de probabilidad paramétricas, por lo que hay que cuestionarse la validez de la extrapolación temporal de la varianza. Supongamos que la rentabilidad X sigue una distribución estable.

9.1.1 Distribuciones estables

Definition 9 *Sean X_1 y X_2 copias independientes de una variable aleatoria X. Se dice que X es estable si para cualquier par de constantes $a > 0$ y $b > 0$ la variable aleatoria $aX_1 + bX_2$ tiene la misma distribución que $cX + d$ para determinadas constantes $c > 0$ y d. La distribución de X es estrictamente estable si lo anterior sucede con $d = 0$.*

Las distribuciones Normal, Cauchy y la distribución de Lévy satisfacen esta condición, por lo que son estables. La distribución de Levy tiene función de densidad y función de distribución:

$$f(x; \mu, c) = \sqrt{\frac{c}{2\pi}} \frac{e^{-\frac{c}{2(x-\mu)}}}{(x-\mu)^{3/2}}; \; F(x; \mu, c) = \operatorname{erf} c \left(\sqrt{\frac{c}{2(x-\mu)}} \right)$$

donde $\operatorname{erf} c$ denota la función de error complementaria:[15] $\operatorname{erf} c = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt$.La distribución de Levy estandarizada (esperanza=0,varianza=1), es: $f(y; 0, 1)$, con $y = \frac{x-\mu}{c}$.

[15] Se denomina funcion de error complementaria porque se relacion con la función de error erf mediante: $\operatorname{erf} c(zx) = 1 - \operatorname{erf}(x)$, siendo $\operatorname{erf} = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$.

La función caracteristica de la distribución de Levy es: $\varphi(t;\mu,c) = e^{i\mu t-\sqrt{-2ict}} = e^{i\mu t-|ct|^{1/2}(1-i.sign(t))}$. Si $X \sim N(0,\sigma^2)$, entonces: $(X-\mu)^{-2} \sim Levy(0, 1/\sigma^2)$.

Una *distribución estable* es invariante por un factor de escala $1/\xi$. Por ejemplo, para el p-cuantil a h dias, tenemos:

$$x_{hp} = h^{1/\xi} x_{1p}$$

Toda la distribución, no solo los cuantiles quedan afectados de una escala como la indicada. La distribución Normal es una distribución estable con $\xi = 2$. Conocer el valor adecuado del escalado exponencial es importante en dos casos en que la estimación del VaR se enfrenta a dificultades: una es cuando la muestra es corta: con 1000 datos, el VaR 0,1% es el valor minimo muestral, por lo que dicho VaR se estimará con poca precisión; asimismo, si queremos calcular un VaR a 10 dias (y seria peor para el VaR a 3 meses) necesitamos calcular el percentil de rentabilidades sobre periodos de 10 dias. Puesto que no conviene utilizar periodos solapados, con 1000 datos, dispondriamos tan solo de 100 rentabilidades sobre periodos de 10 dias. Ahora, el VaR 1% seria el minimo de dichas rentabilidades sobre 10 dias mientras que el VaR 0,1% no podria calcularse. En estos casos, suele estimarse el VaR de rentabilidades diarias, y utilizar su extrapolación temporal. En esta seccion indagamos cual es el parametro de escalado adecuado.

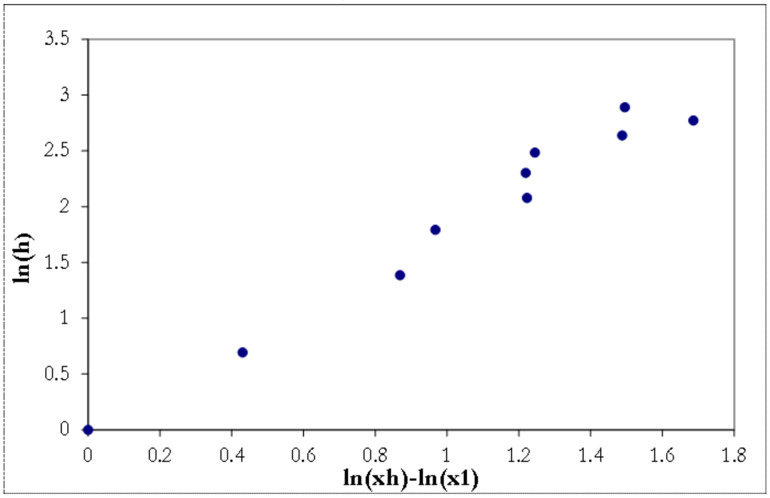
Para tratar de estimar el parámetro de estabilidad, tomamos logaritmos en la expresión anterior,

$$\xi = \frac{\ln(h)}{\ln(x_{hp}) - \ln(x_{1p})}$$

por lo que el escalado exponencial ξ puede estimarse en una regresión del numerador sobre el denominador. En carpeta Case Study IV.3.2 [dentro de Case Study IV.3] se lleva a cabo este análisis para activos de distinta naturaleza. Para ello, se toma la serie temporal de rentabilidades logaritmicas (variaciones diarias en el caso de tipos de interés) y se agregan a lo largo de distintos valores de h , por ejemplo: $h = 2, 5, 10, 21, 63, 126, 250$, utilizando muestras no solapadas. A continuación, se toma para cada una de dicha serie de rentabilidades agregadas en el tiempo, el percentil al nivel deseado, y se resta de él (en logaritmos) el percentil análogo para las rentabilidades a 1 dia $[\ln(x_{hp}) - \ln(x_{1p})]$. La estimación de ξ dependerá de los plazos de h escogidos, pero, sobre todo, del nivel de significación.

El gráfico siguiente representa $\ln(h)$ en función de $\ln(x_{hp}) - \ln(x_{1p})$ para el tipo de interés sobre Treasury Bills a 3 meses, al 5%, siendo el parámetro ξ la pendiente de dicha recta. Tales ejercicios sugieren un valor algo inferior a 0,5 para la extrapolación temporal de la varianza en el caso del S&P500, e inferiores aún, en torno a 0,43 para los índices de volatilidad VIX y VDAX. Para los tipos de interés, el factor de escala es mayor que 0,5 lo que sugiere que si hay reversión a la media, se produce sobre periodos de tiempo muy largos.

Scale exponent for Tbill 3meses al 5%



En el caso de una cartera, no hay forma de agregar las posibles constantes de escala de los activos individuales para lograr una constante de escala aplicable a la cartera. Además, las rentabilidades de la cartera seguirán una distribución estable solo si la siguen todos los activos de la cartera, pero con el mismo exponente. En ese caso, a veces se utiliza el promedio de las constantes de escala, o se estima directamente la constante de escala de la rentabilidad de la cartera, ignorando las constantes de los activos individuales. El primero, que es una aproximación peor, tiene la ventaja de que no hay que estimar una constante para cada una de las carteras que podamos formar a partir de los mismos activos.

Los errores cometidos al ignorar este aspecto pueden ser importantes. En negrilla aparece el factor que extrapolaría el VaR a 1 dia a VaR a períodos más largos, siguiendo el criterio de la raiz cuadrada. Todos los demas factores están expresados como porcentaje de éste.

[$Tabla\ IV.3.2$]

Scale exponent	0,3	0,35	0,4	0,45	0,5	0,55	0,6	0,65	0,7
Holding period									
2	87,1%	90,1%	93,3%	96,6%	**1,41**	103,5%	107,2%	111,0%	114,9%
5	72,5%	78,6%	85,1%	92,3%	**2,24**	108,4%	117,5%	127,3%	138,0%
10	63,1%	70,8%	79,4%	89,1%	**3,16**	112,2%	125,9%	141,3%	158,5%
30	50,6%	60,0%	71,2%	84,4%	**5,48**	118,5%	140,5%	166,6%	197,4%
100	39,8%	50,1%	63,1%	79,4%	**10,00**	125,9%	158,5%	199,5%	251,2%
250	33,1%	43,7%	57,6%	75,9%	**15,81**	131,8%	173,7%	228,9%	301,7%

Imaginemos una muestra de 1000 datos diarios entre las que hay un resultado negativo enorme, de 1 millon, y los siguientes son del orden de 10.000. El VaR 0,1% a 1 dia seria 1 millon, y el VaR 1% a un dia seria 100 veces inferior. ¿Qué podriamos decir sobre los VaR a 10 dias, si son los que nos interesan?

Si tomamos las 100 muestras de 10 dias no solapadas, tendriamos un peor resultado próximo a una pérdida de 1 millon, de modo que el VaR 1% sería ahora 1 millon, y el VaR 0,1% no podria calcularse por falta de datos. Si, para resolver el problema, estimásemos con muestras solapadas, tendriamos de nuevo 1000 datos, y 10 de ellos serían una pérdida aproximada de 1 millon, de modo que ahora, tanto el VaR 0,1% como el VaR 1% serían igual a 1 millon, distorsionando claramente la realidad. Por el contrario, siguiendo la regla de extrapolación, diriamos que $VaR_{10;1\%} = \sqrt{10}VaR_{1;1\%}$ y que $VaR_{10;0.1\%} = \sqrt{10}VaR_{1;0.1\%}$. Es decir, que el VaR 0,1% a 10 dias sería de nuevo 100 veces superior al VaR 1% a 10 dias: *el peor resultado que pueda darse 1 vez cada 4 años*[16] *seria 100 veces peor que el peor resultado que pueda darse una vez cada 40 años.* Esta solución es razonable, suponiendo una distribución estable con $\xi = 2$. La conclusión es que puede no convenir trabajar con muestras no solapadas y pude ser preferible extrapolar el comportamiento de las rentabilidades.

Para una cartera dada, las estimaciones del VaR por los métodos histórico (percentil) y Normal lineal (parametrico estimando la varianza mediante ventanas móviles) suelen estar más próximas que las que se obtengan por el método histórico en dos muestras muy distintas. [Case Study IV.3.3.1, S&P500] Las varianzas que ponderan por igual todas las observaciones, tienen el problema de que extrapolan cualquier dato extremo por un periodo de tiempo aproximadamente igual a la amplitud de la ventana. Esto es lo que explica que los métodos histórico y Normal sean mas similares entre sí que cuando aplicamos cualquiera de ellos a dos ventanas de amplitudes muestrales diferentes.

De este ejercicio deducimos tres conclusiones: *i*) bajo igual ponderación, la amplitud de la ventana determina el error de estimación del VaR, más que el propio método de estimación del VaR, *ii*) hay efectos duraderos sobre el VaR de un shock transitorio, *iii*) el VaR es poco sensible al riesgo (es decir, a condiciones adversas de mercado).

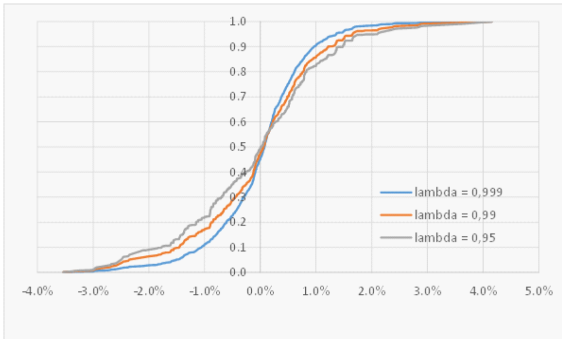
9.1.2 Distribución de probabilidad de rentabilidades ajustadas con pesos exponenciales

[Case Study IV.3.3., S&P500]

Por tanto, en el cálculo del VaR histórico puede ser conveniente aplicar ponderaciones EWMA a las rentabilidades, con objeto de ponderar más el pasado más reciente. [Case Study IV.3.3.1, *S&P500, en carpeta Case Study IV.3.3*]

[16] Si en 1 año hay 250 observaciones, hay 25 periodos de 10 dias. El VaR 1% a 10 dias mide la perdida que cabe esperar cada 100 periodos de 10 dias, es decir, cada 4 años. El VaR 0,1% a 10 dias mide la perdida que cabe esperar cada 1000 periodos de 10 dias, es decir, cada 40 años.

Funciones de distribución con ponderaciones EWMA



El problema es que la estimación del VaR depende del valor numérico escogido para eel parámetro λ, que es arbitrario.

9.2 Ajustes de volatilidad

Este procedimiento se utiliza para reflejar mejor en la estimación del VaR las condiciones de mercado actuales. Al utilizar una muestra larga en el análisis histórico, podemos encontrarnos con que la evolución media a lo largo de la muestra no sea representativa de la situación de volatilidad actual, bien porque esta sea excesivamente alta o baja para los estándares históricos. Pero para nuestros cálculos de riesgo, queremos trabajar bajo el supuesto de rentabilidades que se distribuyen igual e independientemente. Para aproximarnos a una distribución constante a lo largo de la muestra, en ocasiones se ajusta la series histórica de rentabilidades mediante:

$$\tilde{r}_{t,T} = \frac{\hat{\sigma}_T}{\hat{\sigma}_t} r_t$$

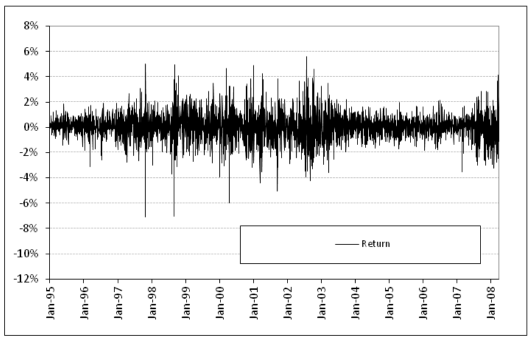
donde $\{\hat{\sigma}_t\}_{t=1}^{T}$ denota una estimación temporal de la volatilidad para la serie histórica utilizando un esquema EWMA o GARCH,[17] y $\hat{\sigma}_T$ es la estimación para el último dato de la muestra.

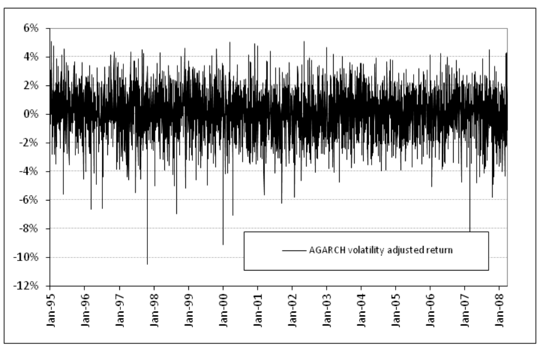
Lo que se hace es generar una serie temporal de volatilidades EWMA, σ_t, para la rentabilidad del SP500 y luego se estima el factor de escala sobre las rentabilidades $r_t\sigma/\sigma_t$, siendo σ la volatilidad deseada. Dicha σ puede ser la ultima observada en la muestra, o una volatilidad de largo plazo que creemos que puede prevalecer durante el periodo de evaluación del riesgo. Este ajuste permite estimar con bastante precisión el VaR para cuantiles muy elevados.

Ejercicio: [EIV.3.1]. Se ajustan modelos GARCH y AGARCH a rentabilidades diarias del S&P500 bajo el supuesto de Normalidad. El modelo GARCH ignora la asimetría en volatilidad, por lo que subestima la volatilidad a largo plazo, que sitúa próxima al 20%, frente a más del 36% del modelo AGARCH. El Ratio de Verosimilitudes favorece al modelo AGARCH. *Ver ambos gráficos*

[17] El uso de un modelo GARCH para estimar la volatilidad evita la elección arbitraria de la constante λ en el esquema EWMA.

en hoja de cálculo. Al estimar el VaR tanto con rentabilidades no ajustadas, como con rentabilidades ajustadas con ambos modelos GARCH, se observa una clara subestimación del riesgo cuando no se ajustan las volatilidades. Esto se debe a que el periodo final de la muestra fue volatil, por lo que el ajuste $\frac{\hat{\sigma}_T}{\hat{\sigma}_t} r_t$ amplifica la distribución de rentabilidades. Si el final de la muestra se refiriese a un periodo tranquilo en el mercado, el resultado sería el opuesto.





El ajuste de volatilidad puede utilizarse asimismo para estimar el factor de escala al que antes nos referimos para convertir un VaR a 1-dia en un VaR a h dias. Para ello debe estimarse una serie temporal de volatilidades (por ejemplo, EWMA o GARCH), y luego aplicar a cada dato el factor σ/σ_t . Así se genera la tabla siguiente (Tabla IV.3.7 de Alexander), cuyo contenido se calcula cambiando los valores de λ y α en Case Study IV.3.3. Scale_Index_S&P500_Volatility Adjusted. Cuando se efectúa el ejercicio puede apreciarse que la estimación del exponente de escala ξ no varía con el nivel de volatilidad al cual se ajusta la serie de rentabilidades.

Valores estimados del exponente de escala para S&P500				
Constante suavizado	Cuantiles			
λ	0.1%	1%	5%	10%
0.98	0.4848	0.5402	0.5368	0.5192
0.95	0.4527	0.5236	0.5366	0.5250
0.90	0.3972	0.5057	0.5334	0.5335
1	0.4662	0.5186	0.5001	0.4853

9.3 Simulación histórica filtrada: horizonte > 1 dia

Una alternativa para el cálculo del VaR a h dias consiste en la simulación histórica filtrada, que utiliza elementos del VaR histórico y del VaR Monte Carlo. En la simulación histórica filtrada, se utiliza un modelo de volatilidad como el GARCH para las rentabilidades logaritmicas, y se estandarizan utilizando las desviaciones tipicas $\{\hat{\sigma}_t\}_{t=2}^T$ estimadas. Entonces, se simula el modelo GARCH extrayendo en la simulación las innovaciones de la distribución de innovaciones *estandarizadas:* $u_t = \varepsilon_t/\hat{\sigma}_t$.

Supongamos que las rentabilidades no tienen estructura estocástica:

$$\begin{aligned} r_t &= \mu + \varepsilon_t,\ E(\varepsilon_t) = 0, Var(\varepsilon_t) = \sigma_t^2, \\ \sigma_t^2 &= \delta_0 + \delta_1 \varepsilon_{t-1}^2 + \delta_2 \sigma_{t-1}^2 \end{aligned}$$

Una vez estimado el modelo, guardamos las innovaciones *estandarizadas:*

$$u_t = \frac{\hat{\varepsilon}_t}{\hat{\sigma}_t} = \frac{r_t - \hat{\mu}}{\sigma_t}$$

Utilizando la rentabilidad del último dia en la muestra, r_T, asi como su varianza estimada $\hat{\sigma}_T^2$, la simulación comienza con:

$$\begin{aligned} \hat{\sigma}_{T+1}^2 &= \hat{\delta}_0 + \hat{\delta}_1 \left(r_T - \hat{\mu}\right)^2 + \hat{\delta}_2 \hat{\sigma}_T^2 \\ \hat{r}_{T+1} &= \hat{\mu} + \hat{u}_{T+1}\hat{\sigma}_{T+1} \\ \hat{\sigma}_{T+2}^2 &= \hat{\delta}_0 + \hat{\delta}_1 \left(\hat{r}_{T+1} - \hat{\mu}\right)^2 + \hat{\delta}_2 \hat{\sigma}_{T+1}^2 \\ \hat{r}_{T+2} &= \hat{\mu} + \hat{u}_{T+2}\hat{\sigma}_{T+2} \\ &\ldots \end{aligned}$$

siendo $\hat{u}_{T+1}, \hat{u}_{T+2}, ...$ extracciones aleatorias de la muestra de innovaciones estandarizadas $\{u_t\}_{t=2}^T = \{\varepsilon_t/\hat{\sigma}_t\}_{t=2}^T$.

Finalmente, se agregan las rentabilidades logaritmicas sobre el horizonte de riesgo de h dias para obtener las rentabilidades logaritmicas simuladas sobre dicho periodo de tiempo, y se calcula el VaR como el p-cuantil de la distribución de rentabilidades a h dias:

$$\hat{r}_{T,T+h} = \hat{r}_{T+1} + \hat{r}_{T+2} + ... + \hat{r}_{T+10}$$

Si efectuamos un amplio número (por ejemplo, 5.000) de estas realizaciones, podremos calcular su distribución empírica y el cuantil deseado, como estimación del VaR.

Si las rentabilidades logaritmicas tiene estructura estocástica, habria que simular a partir de la distribución de innovaciones de r_t obtenidas en la estimación del modelo, y una vez estandarizadas:

$$\begin{aligned} r_t &= \mu + \beta r_{t-1} + \varepsilon_t,\ E(\varepsilon_t) = 0, Var(\varepsilon_t) = \sigma_t^2, \\ \sigma_t^2 &= \delta_0 + \delta_1 \varepsilon_{t-1}^2 + \delta_2 \sigma_{t-1}^2 \end{aligned}$$

trabajaríamos con las rentabilidades *estandarizadas:*

$$u_t = \frac{\hat{\varepsilon}_t}{\hat{\sigma}_t} = \frac{r_t - \hat{\mu} - \hat{\beta} r_{t-1}}{\sigma_t}$$

simulando posteriormente:

$$\hat{r}_{T+1} = \hat{\mu} + \hat{\beta} r_T + \hat{u}_{T+1} \hat{\sigma}_{T+1}$$

y asi sucesivamente.

Aunque hemos utilizado en esta descripción un modelo GARCH simétrico, pueden utilizarse otros esquemas. En [EIV.3.2] se utiliza un modelo AGARCH. [ver tambien Christoffersen Chapter5Results.Question 3 and Question 4]

9.4 Precisión del VaR histórico para cuantiles extremos

Cuando se trabaja con cuantiles no más extremos del 1%, basta con utilizar el ajuste de volatilidad y la simulación histórica filtrada. Cuando utilizamos cuantiles más extremos, como 0,01%, en general no tendremos muchas observaciones para estimar, por lo que la precisión en la estimación del VaR y de la pérdida esperada se reduce mucho. Para resolver el problema necesitamos ajustar distribuciones continuas a las escasas observaciones discretas con que contamos.

9.4.1 Kernel fitting

En algunos casos no tenemos suficiente información como para especificar de antemano la relación no lineal que pueda existir entre dos variables, Y y X. En tal situación pueden ser útiles los métodos no paramétricos. Su principal limmitación es que dependen totalmente de los datos observados en la muestra y que es fácil que puedan resultar en overfitting. Esta es la situacion que se produce cuando una función altamente no lineal permite un buen ajuste a los datos muestrales, pero es tan especifica de dichas observaciones, que predice mal los datos que puedan observarse en el futuro.

La esencia de estos métodos es el *smoothing.* Supongamos dos variables relacionadas por la función no lineal:

$$Y_t = m(X_t) + \varepsilon_t$$

siendo $m(.)$ una función no lineal suave, desconocida, y ε_t un ruido blanco. Si para $X = x$ tenemos observaciones repetidas $y_1, y_2, ..., y_T$ formando un conjunto Ω, tendríamos:

$$y_t = m(x) + \varepsilon_t,\ t =, 2, ..., T$$

y tomando promedios: $\bar{y} = m(x) + \bar{\varepsilon}$.Por la ley de los grandes números, el promedio de las innovaciones converge en probabilidad a cero, por lo que la media de las observaciones de Y para $X = x$ sería un estimador consistente de $m(x)$:

$$\hat{m}(x) = \frac{1}{size(\Omega)} \sum_{y_j \in \Omega} y_j$$

Pero lo habitual en Finanzas es que no tengamos observaciones repetidas. En ese caso podemos utilizar los valores observados de Y_t para valores de X_t proximos a x. Pero parece natural dar mayor peso a las observaciones de Y_t cuanto más cercano está el correspondiente X_t a x. Para ello utilizamos un promedio ponderado:

$$\hat{m}(x) = \sum_{(y_j, x_j) \in \Omega} w_j(x_j, x).y_j \tag{16}$$

siendo $\Omega = \{(y_t, x_t) \mid |x_t - x| < \varepsilon\}$. Permitimos que el peso que recibe cada observación de y pueda ser mayor para observaciones asociadas a valores de x_t proximos a x. Esto es un *promedio ponderado local*.

Kernel regression En la regresión kernel se utilizan todas las observaciones muestrales para calcular el valor numérico que debe tomar cada $\hat{y}_j$. Los pesos se determinan a partir de una función kernel, que debe satisfacer:

$$K(u) \geq 0, \int_{-\infty}^{\infty} K(u) du = 1$$

En ocasiones tambien se exige una condición de simetría: $K(-u) = K(u)\ \forall u$. Para ganar flexibilidad, se aplica un factor de escala a la función kermel:

$$K_h(u) = \frac{1}{h} K\left(\frac{u}{h}\right),\ K_h(u) \geq 0,\ \int_{-\infty}^{\infty} K_h(u) du = 1$$

siendo h un parámetro que se conoce como amplitud de banda (*bandwidth*). El ancho de banda es equivalente a la amplitud de los intervalos de un histograma de frecuencias, y el objeto de estimar un kernel es precisamente encontrar el ancho de banda óptimo para esta representación.

La función de pesos se define entonces:

$$w_j(x_j, x) = \frac{K_h(x - x_j)}{\sum_{t=1}^{T} K_h(x - x_j)}$$

Llevando esta expresión a (16) tenemos el estimador de Nadaraya-Watson:

$$\hat{m}(x) = \sum_{t=1}^{T} w_j(x_j, x).y_j = \frac{\sum_{t=1}^{T} K_h(x - x_j).y_j}{\sum_{t=1}^{T} K_h(x - x_j)}$$

La elección más frecuente de funciones kernel es:

El kernel Gaussiano:

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \Rightarrow w_j(x_j, x) = \frac{1}{nh} \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{n} e^{-\frac{1}{2}\left(\frac{x - x_i}{h}\right)^2}$$

consiste en aproximar la densidad por una mixtura de n densidades Normales con esperanza matemática $x_i = x_i$ e igual varianza, h^2.

El kernel de Epanechnikov, denotando por $u = \frac{x-x_i}{h}$, se define:

$$\begin{aligned} K(u) &= \frac{3}{4}(1-u^2), \ si \ -1 \leq u \leq 1 \\ &= 0 \ en \ otro \ caso \end{aligned}$$

que conduce a:

$$w_j(x_j, x) = \frac{1}{nh}\frac{3}{4}\sum_{i=1}^{n}\left[1-\left(\frac{x-x_i}{h}\right)^2\right]1_{\{|\frac{x-x_i}{h}|<1\}}$$

El kernel logistico:

$$K(u) = \frac{1}{e^u + 2 + e^{-u}} = \frac{e^u}{(e^u+1)^2}$$

conduce a:

$$w_j(x_j, x) = \frac{1}{nh}\sum_{i=1}^{n}\left(e^{\frac{x-x_i}{h}} + 2 + e^{-\frac{x-x_i}{h}}\right)^{-1}$$

Cuando el bandwidth tiende a cero, la función $\hat{m}(x)$ tiende a reproducir los datos, es decir, devuelve las observaciones de y_j. Cuando el bandwidth tiende a infinito, la función $\hat{m}(x)$ tiende a asignar a todas las observaciones de x_j la media muestral de y . Es decir, el suavizado es claramente excesivo. A veces se propone obtener un bandwidth optimo minimizando el error cuadrático en media:

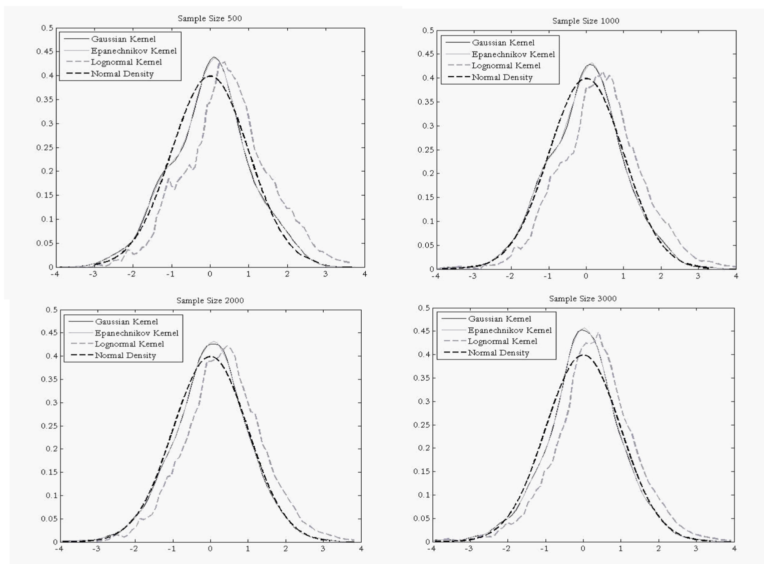
$$E\int_{-\infty}^{\infty}\left[\hat{m}(x) - m(x)\right]^2 dx$$

pero se trata de un problema computacional complicado. Una sugerencia mas sencilla consiste en utilizar:

$$\begin{aligned} h_{opt} &= 1.06\sigma_x T^{-1/5} \text{ para el kernel Gaussiano} \\ h_{opt} &= 2.34\sigma_x T^{-1/5} \text{ para el kernel de Epanechnikov} \end{aligned}$$

Cuando hay varias variables explicativas, puede tomarse como kernel una función de densidad multivariante, como la Normal.

Para el cálculo del VaR, las funciones kernel pueden utilizarse para estimar la función de densidad de las rentabilidades a partir de una muestra relativamente reducida que no nos permitiría construir un histograma de frecuencias suficientemente suavizado. En tal aplicación, la elección de kernel no es excesivamente importante. Puede comenzarse aplicando un ajuste de volatilidad como se ha explicado, a partir de varianzas que pueden estimarse por EWMA o GARCH. Posteriormente, las rentabilidades se estandarizan para tener esperanza 0 y varianza 1 antes de ajustar el Kernel, y luego habria que deshacer la transformacion, ya sea de acuerdo con valores medios muestrales o con los valores de final de muestra. Por último, se calculan los cuantiles de las distribuciones continuas estimadas. El ejercicio se repite con las últimas 500, 1000, 2000 y 3000 observaciones muestrales.



Otros kernel:

$$
\begin{aligned}
\text{Sigmoide:} \quad & K(u) = \frac{2}{\pi}\frac{1}{e^{u}+e^{-u}} \\
\text{Uniforme:} \quad & K(u) = \frac{1}{2} \quad \text{si} \quad -1 \leq u \leq 1 \\
\text{Triangular:} \quad & K(u) = 1-|u| \quad \text{si} \quad -1 \leq u \leq 1 \\
\text{Coseno:} \quad & K(u) = \frac{\pi}{4}\cos\left(\frac{\pi}{2}u\right) \quad \text{si} \quad -1 \leq u \leq 1 \\
\text{Cuártico:} \quad & K(u) = \frac{15}{16}\left(1-u^{2}\right)^{2} \quad \text{si} \quad -1 \leq u \leq 1
\end{aligned}
$$

Regresión local lineal La regresión lineal local es la solución al problema:

$$\min_{(a,b)} L(a,b) = \sum_{t=1}^{T} \left[y_t - a - b(x-x_t)\right]^2 K_h(x-x_t) \tag{17}$$

siendo $K_h(u)$ una función kernel y h su bandwidth. El estimador de $m(x)$ es el parámetro a mientras que b es un estimador de la derivada de $m(x)$ evaluada en $X = x$.

La solución puede expresarse:

$$\hat{m}(x) = \frac{\sum_{t=1}^{T} w_t y_t}{\sum_{t=1}^{T} \left(w_t + \frac{1}{T^2} \right)} \tag{18}$$

donde los pesos se definen:

$$w_t = K_h(x - x_t) \left[s_{T,2} - (x - x_t) s_{T,1} \right]$$

con $s_{T,i} = \sum_{t=1}^{T} (x - x_t)^i . K_h(x - x_t)$. Dichos pesos satisfacen la condición:

$$\sum_{t=1}^{T} (x - x_t) w_t = 0$$

La función (18) se conoce como *local linear regression smoother.*

El estimador de Nadaraya-Watson surge como solución al problema:

$$\min_{(a,b)} L(a,b) = \sum_{t=1}^{T} (y_t - a)^2 K_h(x - x_t)$$

bajo el supuesto de que la función $m(x)$ tiene primera derivada. En general, si $m(x)$ admite k derivadas, entonces podemos sustituir el polinomio lineal de (17) por un polinomio de grado $k - 1$.

9.4.2 Aproximacion de Cornish-Fisher

Una limitación importante de la distribución t-Student en la modelización de las rentabilidades condicionales está originada por su dependencia de un sólo parámetro, el número de grados de libertad, d. No permite asimetrías ni tampoco puede reproducir el elevado grado de curtosis que se observa a menudo en las rentabilidades estandarizadas empíricas. Una alternativa consiste en utilizar la aproximación de Cornish-Fisher, que proporciona aproximaciones a los cuantiles de una distribución de probabilidad de rentabilidades utilizando los momentos de orden superior (en particular, la asimetria y curtosis) de la misma.

La aproximación de Cornish-Fisher se basa en un desarrollo en serie de Taylor de cuarto orden alrededor de la distribución $N(0,1)$. Es una aproximación a la inversa de la función de distribución seguida por las rentabilidades, es decir, es una aproximación al valor numérico del cuantil $\tilde{x}_p$ de la distribución de rentabilidades para un nivel de probabilidad p. La expresión de Cornish-Fisher es:

$$\tilde{x}_p \equiv f(z_p) = z_p + \frac{\tau}{6} \left(z_p^2 - 1 \right) + \frac{\kappa}{24} z_p (z_p^2 - 3) - \frac{\tau^2}{36} z_p (2 z_p^2 - 5)$$

donde $z_p = \Phi^{-1}(p)$ es el p cuantil de la distribución $N(0,1)$, y τ, κ denotan la asimetria y exceso de curtosis de la distribución original de rentabilidades. Nótese que cuando ambos coeficientes (asimetría y curtosis) son cero, tenemos el cuantil $N(0,1)$.

En la estimación del cuantil $\tilde{x}_p$ hemos ignorado la esperanza y varianza de la distribución muestral de rentabilidades. Por tanto, $\tilde{x}_p$ puede interpretarse como una aproximación al p-cuantil de una distribución empírica con esperanza 0 y varianza 1. Si μ y σ denotan la esperanza y desviación típica de la distribución de rentabilidades observadas, sus cuantiles vienen dados por:

$$x_p = \hat{\sigma}\tilde{x}_p + \hat{\mu}$$

Por tanto, una vez obtenida la aproximación Cornish-Fisher, el VaR puede calcularse (notemos que $\tilde{x}_p < 0$) ,

$$VaR_{t+1}^p = -\hat{\sigma}_{t+1}\tilde{x}_p - \hat{\mu}_{t+1}$$

y la ETL (calculada como valor numérico positivo, como habitualmente se presenta):

$$ETL_p = -f\left(ETL_{\Phi,p}\right)\hat{\sigma} - \hat{\mu} = f\left(p^{-1}\varphi\left(\Phi^{-1}(p)\right)\right)\hat{\sigma} - \hat{\mu}$$

donde $f(.)$ es la función que define el cuantil CF, y $ETL_{\Phi,p}$ es la ETL de una distribución $N(0,1)$. Recordemos de la sección 1.7 que si $X \sim N(\mu_h, \sigma_h^2)$, entonces: $ETL_{p,h}(X) = p^{-1}\varphi\left(\Phi^{-1}(p)\right)\sigma_h - \mu_h$, donde φ y Φ denotan las funciones de densidad y distribución de una variable $N(0,1)$.

Una limitación de este enfoque es que puede verse influido numéricamente por una alta frecuencia de rentabilidades estandarizadas en el entorno de cero (es decir, no extremas), lo cual puede resolverse mediante la Teoría de Valores Extremos, que examinamos más adelante, ya que se concentra en la modelización de las rentabilidades en la cola de la distribución.

La aproximación es buena para el cálculo del VaR cuando las rentabilidades no tienen mucha asimetria o curtosis. Para una rentabilidad con distribución t-Student con 50 grados de libertad, la aproximación puede ser buena, subestimando el VaR ligeramente, pero para una distribución con 10 grados de libertad, que tiene exceso de curtosis igual a 1, la subestimación será importante, en torno al 10%.

Ejercicio [ver también EIV.3.4]

Consideremos, por ejemplo, el VaR 1%. Bajo Normalidad, tenemos: $\Phi_{.01}^{-1} = -2.33$. Alternativamente, la aproximación de Cornish-Fisher del cuantil 1% es:

$$\tilde{x}_p = -2,33 + 0,74\tau - 0,24\kappa - 0,38\tau^2$$

Supongamos que la asimetría es: $\tau = -1$ y el exceso de curtosis: $\kappa = 4$. Tendríamos entonces,

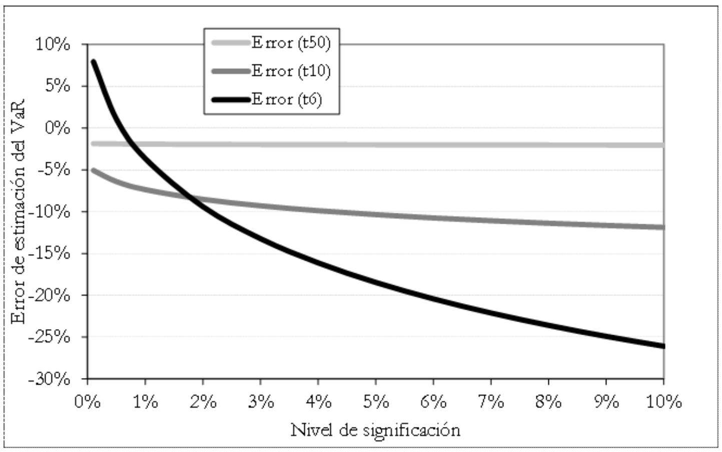
$$\tilde{x}_p = -2,33 - 0,74 - 4(0,24) - 0,38 = -4,41$$

y tenemos el VaR:

$$VaR_{t+1}^p = -4,41\sigma_{t+1}$$

casi el doble de lo que habríamos obtenido suponiendo la Normalidad de los rendimientos estandarizados.

Cuando la distribución de rentabilidades se separa de la Normal la aproximación Cornish-Fisher puede no dar buenos resultados. A modo de ejemplo, el siguiente gráfico muestra el error porcentual de estimación del cuantil cuando las rentabilidades siguen una distribución t-Student con distintos grados de libertad. Puee verse que con $\nu = 50$, Cornish-Fisher subestima el VaR alrededor de un 2%, pero para $\nu = 6$ el error de estimación del VaR puede ser muy alto.



9.4.3 Distribuciones de valor extremo

La aproximación de Cornish-Fisher proporciona estimaciones de los cuantiles de la distribución de rendimientos estandarizados a partir de estimaciones de los coeficientes de asimetría y de exceso de curtosis de dicha distribución. Pero las estimaciones de estos estadísticos pueden estar excesivamente condicionadas por el amplio conjunto de rentabilidades en el entorno de cero, lo que entenderíamos por rentabilidades "estándar". Por esta razón puede ser conveniente un enfoque basado únicamente en los rendimientos más extremos. Además, el mayor riesgo al que se enfrenta una cartera es la ocurrencia repentina de una rentabilidad negativa extremadamente grande, por lo que estimar con precisión la probabilidad de tales sucesos es la esencia de la gestión de riesgos.

El resultado básico sobre el que se basa la EVT es que la cola extrema de una amplia familia de distribuciones F puede describirse aproximadamente por una distribución relativamente sencilla, la llamada distribución de Pareto. La teoría se basa en el supuesto de independencia e idéntica distribución de los rendimientos. Como la dependencia temporal surge en muchos casos debido a la persistencia en volatilidades, es conveniente trabajar con rendimientos estandarizados mediante un modelo de volatilidad condicional previamente estimado:

$$z_{t+1} = r_{t+1}/\sigma_{t+1}$$

que, generalmente, ya podemos suponer $i., i.d.$, con esperanza nula y varianza unitaria.

Por otra parte, los rendimientos en períodos de tiempo relativamente largos se aproximan a la distribución Normal, por lo que la EVT tiene mayor interés para rendimientos observados en datos de alta frecuencia, una vez estandarizados.

Una alternativa es ajustar a los datos disponibles una distribución que se sabe que tiene colas gruesas, como la distribución Generalizada de Valor Extremo (GVE). La distribución GVE se ajusta a la muestra completa, mientras que la distribución generalizada de Pareto (GPD), que veremos luego, se ajusta a las rentabilidades que exceden de cierto umbral predefinido u. Pero necesitamos tener un cierto número de datos por encima de dicho umbral (mínimo 20) para ajustar con precisión la distribución.

La distribución Generalizada de Valor Extremo (GVE) depende de parámetros de posición (μ) y de escala (β), asi como de un parámetro ξ que se conoce como índice de cola (*tail index*). Su inverso, $1/\xi$, se conoce como *shape* (perfil) de la distribución. La función de distribución es:

$$\begin{aligned} F(z) &= \exp\left(-\exp\left(-\frac{z-\mu}{\beta}\right)\right),\ \xi = 0, \\ F(z) &= \exp\left(-\left[1+\xi\frac{z-\mu}{\beta}\right]^{-1/\xi}\right),\ \xi \neq 0,\ \xi\frac{z-\mu}{\beta} > -1 \end{aligned}$$

siendo su función de densidad:

$$\begin{aligned} f(x) &= \beta^{-1}\exp\left(-\frac{z-\mu}{\beta}\right)\exp\left(-\exp\left(-\frac{z-\mu}{\beta}\right)\right),\ \xi = 0, \\ f(x) &= \beta^{-1}\left[1+\xi\frac{z-\mu}{\beta}\right]^{-1/\xi-1}\exp\left(-\left[1+\xi\frac{z-\mu}{\beta}\right]^{-1/\xi}\right),\ \xi \neq 0,\ \xi\frac{z-\mu}{\beta} > -1 \end{aligned}$$

La condición $\xi\frac{z-\mu}{\beta} > -1$ es una condición sobre el rango de valores de x,

$$\begin{aligned} z &> \mu - \beta\xi^{-1}\ si\ \beta\xi > 0 \\ z &< \mu - \beta\xi^{-1}\ si\ \beta\xi < 0 \end{aligned}$$

Hay tres tipos de distribución GVE, según el valor del índice de cola ξ :

- para $\xi = 0$ tenemos la distribución de Gumbel, cuya densidad tiene moda en 0, asimetría positiva y disminuye exponencialmente en las colas

- para $\xi < 0$ tenemos la distribución de Weibull, cuya densidad tiende a degenerar en el punto μ segun tiende ξ a $-\infty$. La cola inferior permanece finita.

- para $\xi > 0$ tenemos la distribución de Fréchet, que también tiende a degenerar en el punto μ, pero esta vez según tiende ξ a ∞. Converge más lentamente que la distribución de Weibull, puesto que en las colas disminuye de acuerdo con la potencia de los valores de la variable. Este resultado aplica a la mayoría de las distribuciones con colas pesadas, como la t de Student.

Para distribuciones con colas ligeras, no muy útiles en Finanzas, el parámetro ξ sería negativo. En series financieras, tendremos habitualmente: $\xi > 0$, una distribución de Fréchet.

9.4.4 Distribución en las colas: La distribución Generalizada de Pareto

Consideremos ahora la distribución de Pareto, que se utiliza para representar valores extremos por encima de un cierto umbral. Supongamos que los rendimientos *estandarizados* z siguen la distribución incondicional $F(z)$, y consideremos la probabilidad de que el rendimiento observado un determinado instante, excediendo de un cierto umbral u, lo haga en menos de una cuantía x. Esto define una función de distribución diferente sobre el conjunto de valores de z superiores a u, que denotamos por F_u y que tiene como argumento a x, el exceso de z sobre u. La función F_u es la distribución de probabilidad de z truncada a la izquierda de u, que podemos calcular:[18]

$$x \geq 0 \Rightarrow F_u(x) = P[z \leq u+x \mid u < z] = \frac{P\left(u < z \leq u + x\right)}{P(z > u)} = \frac{F(u+x) - F(u)}{1 - F(u)}$$

Como vemos, F_u es una función paramétrica del umbral fijado, u, y puede escribirse en función de la distribución de rendimientos estandarizados F.

Para un umbral dado u, suficientemente alto, y para muchas distribuciones de probabilidad F, la distribución F_u pertenece a la familia de distribuciones Generalizada de Pareto (GPD), que está definida por (nótese que el argumento es x, el exceso de z sobre u):

$$\begin{aligned} F_u(x) &= P(z - u < x \mid z > u) \approx 1 - \left(1 + \frac{\xi x}{\beta}\right)^{-1/\xi}, \xi \neq 0 \qquad (19) \\ F_u(x) &\approx 1 - \exp\left(-\frac{x}{\beta}\right),\ \xi = 0 \end{aligned}$$

donde $\beta > 0$ es el parámetro de escala y ξ es el tail index.

La distribución de probabilidad de Pareto habitual tiene funciones de distribución y densidad:

[18] Se dice truncadaa la izquierda porque es la densidad de probabilidad a la derecha del umbral u.

$$\begin{aligned} F(x) &= 1-\left(\frac{\kappa}{\kappa+x}\right)^{\alpha} \\ f(x) &= \frac{\alpha\kappa^{\alpha}}{(\kappa+x)^{\alpha+1}}, \text{ para } \alpha>0, \kappa>0, x\geq 0 \end{aligned}$$

Cuando $\xi > 0$, la distribución Generalizada de Pareto tiene colas pesadas y se reduce entonces a la distribución de Pareto con $\alpha = 1/\xi$, $\kappa = \beta/\xi$, como facilmente puede comprobarse a partir de las expresiones anteriores. La distribución Exponencial es otro caso particular, cuando $\xi = 0$. Cuando $\xi < 0$ tenemos la "Short-tailed" distribución de Pareto de tipo II.

Su función de densidad es:

$$\begin{aligned} f_u(x) &\approx \beta^{-1}\left(1+\frac{\xi x}{\beta}\right)^{-1/\xi-1}, \xi\neq 0 \\ f_u(x) &\approx \beta^{-1}\exp\left(-\frac{x}{\beta}\right),\ \xi=0 \end{aligned}$$

El rango de valores admisibles de x en esta función de densidad está definido por la condición: $1+\xi\frac{x}{\beta}\geq 0$, similar a la que vimos en la distribución GVE. Por tanto, si $\xi > 0$, entonces el rango admisible es $x \geq -\beta/\xi$. Si $\xi < 0$, como debe cumplirse $\xi\frac{x}{\beta} \geq -1$, es decir $x\xi \geq -\beta$, ello implica: $0 \leq x \leq -\beta/\xi$. Cuando $\xi < 0$,el rango de valores efectivos para la función de densidad aumenta al aumentar el parámetro de escala β . Al aumentar el indice de cola ξ aumenta la probabilidad en las colas. En series financieras, ξ suele estar entre 0.1 y 0.4. Esta distribucion tiene momentos finitos de orden k solo para $k < 1/\xi$. Además: $E(X) = \beta/(1-\xi\ )$.

Como la distribución DGP se utiliza sobre rentabilidades estandarizadas, una vez estimado el VaR, hay que deshacer tal normalizacion multiplicando el VaR paramétrico estimado por la desviación tipica de las rentabilidades originales y sumando su media.

Para estimar sus parámetros por máxima verosimilitud, seleccionamos las observaciones z_i superiores al umbral escogido u, y maximizamos:

$$\underset{\beta,\xi}{Max}\sum_{i=1}^{n_u}\ln\left[\frac{1}{\beta}\left(1+\xi\frac{z_i-u}{\beta}\right)^{-1/\xi-1}\right] = \sum_{i=1}^{n_u}\ln\left[\frac{1}{\beta}\left(1+\xi\frac{x_i}{\beta}\right)^{-1/\xi-1}\right] =$$
$$-n_u\ln\beta-\left(1+\frac{1}{\xi}\right)\sum_{i=1}^{n_u}\left(1+\xi\frac{x_i}{\beta}\right), \text{ sujeto a } : \quad \beta>0, 1+\xi\frac{x_i}{\beta}\geq 0,\ \forall i$$

Estimación sencilla de la GPD Si en la expresión: $F_u(x) = \frac{F(u+x)-F(u)}{1-F(u)}$ tenemos en cuenta que: $z = x + u$, tenemos:

$$1-F_u(z-u) = 1-\frac{F(z)-F(u)}{1-F(u)} = \frac{1-F(z)}{1-F(u)} \Rightarrow F(z) = 1-[1-F(u)]\,[1-F_u(z-u)]$$

Si T denota el tamaño muestral total, y T_u denota el número de observaciones que exceden del umbral u, el término $[1-F(u)]$ puede estimarse mediante el cociente T_u/T. Utilizando la aproximación (19) para sustituir en la expresion anterior $F_u(z-u)$, tenemos, para las rentabilidades estandarizadas z que exceden del umbral u, la distribución:

$$F(z) = 1 - \frac{T_u}{T}\left(1+\frac{\xi(z-u)}{\beta}\right)^{-1/\xi} \tag{20}$$

que es la estimación de la *cola superio*r de la distribución de rentabilidades.

Nótese que aunque la GPD es una distribución sobre los excesos de rentabilidad a partir de un determinado umbral, de ella deducimos una distribución sobre rentabilidades z.

Example 10 *(Hull) Supongamos que* $\xi = 0.3232, \beta = 0.0055$, *que elegimos un umbral* $u = 0.02$, *y que entre las* $T = 2256$ *observaciones muestrales, hay* $T_u = 28$ *rentabilidades menores que* -2%. *Si quisiéramos estimar la probabilidad de que la rentabilidad* z *sea inferior a -0,04, tendríamos:*

$$F(0.04) = 1 - \frac{28}{22556}\left(1+0.3232\frac{0.04-0.02}{0.0055}\right)^{-1/0.3232} = 0.9989$$

por lo que: $P(z < -0.04) = 1 - 0.9989 =$ *0.0011, un estimador que, seguramente, será mas preciso que el resultado que hubiesemos obtenido contando observaciones. Del mismo modo, se obtendria que la probabilidad de obtener una rentabilidad inferior a -6% sería 0.0003.*

Vamos a utilizar este resultado para estimar el parámetro ξ , que determina el grosor de la cola de la distribución F, por máxima verosimilitud. Para ello, suponemos que para valores de z superiores al umbral u, es decir, para $z > u$, la función anterior puede aproximarse por:[19]

$$F(z) = 1 - L(z)z^{-1/\xi} \approx 1 - cz^{-1/\xi}$$

con función de densidad:

$$f(z) = \frac{dF(z)}{dz} = \frac{1}{\xi}cz^{-1/\xi-1}$$

donde puede observarse que en la aproximación hemos prescindido del parámetro β.

[19] En esta aproximación, la función $L(z)$ es: $L(z) = \frac{T_u}{T}\left(\frac{\beta}{\xi}\right)^{1/\xi}\left(1+\frac{\beta/\xi-u}{z}\right)^{-1/\xi}$

La aproximación se basa en el hecho de que la función $L(z)$ varía lentamente con z para la mayoría de las distribuciones F, por lo que podemos suponerla constante. De este modo, tenemos en $F(z) \approx 1 - cz^{-1/\xi}$ una expresión que nos permite aproximar el valor numérico de un amplio conjunto de funciones de distribución en su *cola superior*.

Utilizando la definición de distribución condicional tenemos la función de densidad de rendimientos a la derecha del umbral u:

$$f(z/z > u) = \frac{f(z)}{P(z > u)}$$

Recordemos que, esencialmente, una función de densidad truncada se obtiene normalizando la función de densidad original por la probabilidad existente en la región que se considera tras el truncamiento (en este caso, la región a la derecha del umbral u).

Suponiendo independencia de los rendimientos, tenemos la verosimilitud:

$$L = \prod_{i=1}^{T_u} \frac{f(z_i)}{1 - F(u)} = \prod_{i=1}^{T_u} \left(\frac{1}{\xi} \frac{cz_i^{-1/\xi - 1}}{cu^{-1/\xi}} \right)$$

para las observaciones $z_i > u$. Por tanto, el logaritmo de dicha función es:

$$\ln L = \sum_{i=1}^{T_u} \left[-\ln \xi - \left(\frac{1}{\xi} + 1 \right) \ln x_i + \frac{1}{\xi} \ln u \right]$$

Derivando respecto de ξ e igualando a cero, tenemos el estimador de Hill del *parámetro de grosor de cola*:

$$\hat{\xi} = \frac{1}{T_u} \sum_{i=1}^{T_u} \ln \left(\frac{z_i}{u} \right)$$

Ya solo nos falta estimar el parámetro c de la aproximación a la distribución F. Para ello, notamos que:

$$F(u) = 1 - \frac{T_u}{T} = 1 - cu^{-1/\xi}$$

y, puesto que $F(u)$ está estimado mediante T_u/T, esto nos lleva al estimador:

$$\hat{c} = \frac{T_u}{T} u^{1/\xi}$$

por lo que la estimación de la función de distribución para observaciones que *exceden* del umbral u puede aproximarse por:

$$F(z) = 1 - cz^{-1/\xi} = 1 - \frac{T_u}{T} \left(\frac{z}{u} \right)^{-1/\xi} \Rightarrow \quad (21)$$

$$\Rightarrow \quad f(z) = \hat{c} \frac{1}{\xi} z^{-\left(\frac{1}{\xi} + 1\right)} = \frac{T_u}{T} \left(\frac{z}{u} \right)^{-1/\xi} \frac{1}{z\xi} \quad (22)$$

Por tanto, bajo el supuesto que antes hicimos acerca del comportamiento de la función $F(z)$, tenemos estimadores sencillos, sin tener que recurrir a la optimización numérica de la función de verosimilitud que antes vimos.

Nótese que la GPD es una distribución definida para los excesos de rentabilidad sobre un determinado umbral u, y de ella hemos deducido una distribución sobre las propias rentabilidades enla cola, z.

Remark 11 *Es muy importante observar que hemos desarrollado la EVT para la cola derecha de la distribución, al considerar la probabilidad de que el rendimiento* z *observado un determinado instante, excediendo de un cierto umbral* u, *lo haga en menos de una cuantía* x. *Por tanto, para aplicar la teoría de EVT que hemos desarrollado al análisis de la cola izquierda de una distribución de rendimientos, hemos de trabajar con pérdidas estandarizadas:*

$$z_t = -\frac{r_t}{\sigma_t}$$

no con las rentabilidades estandarizadas.

Ejercicio [EIV.3.3] Utilizando las rentabilidades del S&P500 ajustadas de volatilidad que se incluyen en la hoja de cálculo para enero 1950 a marzo 2007, se estiman los parámetros de la distribución GPD y se calcula el VaR bajo esta distribución, comparándolo con el estimado mediante el percentil adecuado de la distribución de rentabilidades ajustadas de volatilidad.

Threshold	n_u	u	ξ	β
20%	1000	-0,693	-0,1194	2,2280
10%	500	-1,1708	-0,1776	3,3405
5%	250	-1,536	-0,2305	4,4352
1%	50	-2,5292	-0,4036	8,2127

Standardized: Ajustado de volatilidad

VaR (Standardized)	GPD VaR
8,8421	5,94%
10,2978	6,91%
11,0486	7,41%
11,7463	7,88%
Empirical VaR	4,86%

Elección del umbral u La elección del umbral u es siempre delicada. Si escogemos un umbral excesivamente pequeño, entonces estaremos trabajando con algunos rendimientos no excesivamente atípicos, y la aproximación funcional a la cola de la distribución en que nos hemos basado puede no ser suficientemente buena para dichos valores numéricos. Si, por el contrario, escogemos un umbral excesivamente elevado, tendremos muy pocas observaciones para estimar los parámetros de la distribución, por lo que tendremos baja precisión en la estimación de dichos parámetros y, consecuentemente, en los cálculos posteriores de Valor en Riesgo, Pérdida Esperada y otros, que veremos a continuación. Una regla relativamente habitual es elegir un umbral que deje un 5% de los datos en la cola de la distribución, aunque en función del número de observaciones de que dispongamos podríamos variar dicho criterio.

VaR y ETL en la teoría de valores extremos Supongamos que hemos ajustado una distribución GPD a las pérdidas en exceso de un umbral u. Dado un nivel de significación p (por ej., 1%), el cuantil q de orden $1-p$ (99%), definido por: $F(q)=1-p$, debe satisfacer (20):

$$\begin{aligned} 1-p &= F(q) = 1 - \frac{T_u}{T}\left(1+\frac{\xi(q_p - u)}{\beta}\right)^{-1/\xi} \Rightarrow p = \frac{T_u}{T}\left(1+\frac{\xi(q_p-u)}{\beta}\right)^{-1/\xi} \\ &\Rightarrow VaR_{1-p} \equiv q_p = u + \frac{\beta}{\xi}\left[\left(\frac{T_u}{Tp}\right)^{\xi} - 1\right] = g(p) \end{aligned}$$

donde T es el tamaño muestral y T_u es el número de observaciones por debajo del umbral u. Como puede verse: a) utilizamos la probabilidad $1-p$ en el cálculo del cuantil, ya que estamos aplicando una teoría desarrollada para valores a la derecha de un determinado umbral u, así que cambiamos signo a las rentabilidades r_t para trabajar con pérdidas y caracterizar la probabilidad de obtener pérdidas a la derecha (es decir, mayores) de un umbral u . El cuantil q para el que $100(1-p)\%$ (con $p=0,01$, por ejemplo) de las observaciones muestrales son pérdidas inferiores a q es, cambiado de signo, el cuantil para el que el $100p\%$ de las rentabilidades estandarizadas (no pérdidas) son inferiores (más negativas que $-q$). La estimación del VaR resultante tendrá ya signo positivo.

Example 12 *En el ejemplo anterior de Hull, si queremos calcular el VaR 99% a 1 día (que interpretamos como VaR 1% cambiado de signo) de una cartera de 1 millón de euros invertida en este activo, tendriamos:*

$$VaR_{0.99} = 0.02 + \frac{0.0055}{0.3232}\left[\left(\frac{2256}{28}\underbrace{(1-0.99)}_{p}\right)^{-0.3232} - 1\right] = 0.0212$$

es decir, 21.200 euros, 02,12% del valor de la cartera.

Si no disponemos de estimaciones del parámetro β, podemos aplicar la expresión aproximada (21):

$$1-p = F(q) \simeq 1 - \frac{T_u}{T}\left(\frac{q}{u}\right)^{-1/\xi}$$

por lo que dicho cuantil sería entonces:

$$q \equiv VaR_{1-p} = F^{-1}(1-p) = u\left(\frac{T_u}{Tp}\right)^{\xi} = 0.0214$$

Definition 13 *La pérdida media en exceso sobre un umbral u se define:*

$$e(u) = E(z - u \mid z > u)$$

y si la distribución en exceso del umbral sigue una distribución GPD, su expresión es:

$$\begin{aligned} e(u) &= \frac{\beta + \xi u}{1 - \xi}, \\ \text{con } u &\geq 0 \text{ si } 0 \leq \xi \text{ y } 0 \leq u \leq -\frac{\beta}{\xi} \text{ si } \xi < 0 \end{aligned}$$

De hecho, puede probarse que la distribución de los excesos sobre otro umbral más elevado que u, el umbral elegido para definir la "cola" de la distribución tienen asimismo una distribución GPD con el mismo parámetro ξ pero con un parámetro de escala β que crece linealmente con el umbral escogido, v, $v > u$. Supuesto que $\xi < 1$, puede probarse que el exceso medio es:

$$e(v) = \frac{\beta + \xi(v - u)}{1 - \xi} = \frac{\xi v}{1 - \xi} + \frac{\beta - \xi u}{1 - \xi} \tag{23}$$

El Expected Tail Loss o VaR condicional puede estimarse como el promedio de los cuantiles desde p hasta 1:

$$\begin{aligned} ES &= \frac{1}{1-p} \int_p^1 q_x dx = \frac{1}{1-p} \int_p^1 \left(u + \frac{\beta}{\xi} \left[\left(\frac{T_u}{T} \right)^{\xi} \frac{1}{x^{\xi}} - 1 \right] \right) dx = \\ &= \frac{1}{1-p} \left[ux + \frac{\beta}{\xi} \left(\frac{T_u^{\xi}}{T^{\xi}} \frac{x^{-\xi+1}}{1-\xi} - x \right) \right]_p^1 = \frac{VaR_{1-p}}{1-\xi} + \frac{\beta - \xi u}{1-\xi} \end{aligned}$$

donde u es el umbral que se haya elegido para definir lo que se entiende por la "cola" de la distribución de rentabilidades. A esta misma expresión se llegaria agregando al VaR_{1-p} el exceso medio (23) para $v \equiv VaR_{1-p} : VaR_{1-p} + \frac{\xi}{1-\xi} VaR_{1-p} + \frac{\beta - \xi u}{1-\xi} = \frac{VaR_{1-p}}{1-\xi} + \frac{\beta-\xi u}{1-\xi}$.

[Christoffersen Chapter4Results.xls, question 6 compara los valores numéricos del VaR estimado en un modelo Normal, un modelo t-Student, la aproximación Cornish-Fisher, y el modelo EVT].

Construcción del QQ-plot bajo la EVT. El QQ-plot se construye utilizando los pares de puntos formados por el cuantil teórico q que acabamos de caracterizar y el cuantil empírico q_i (el cuantil empírico es q_i es la i-ésima pérdida estandarizada, en orden descendente desde la más elevada):

$$\{X_i, Y_i\} = \left\{ u \left(\frac{T_u}{Tp} \right)^{\xi}, q_i \right\} = \left\{ u \left(\frac{T_u}{i - 0,5} \right)^{\xi}, q_i \right\}$$

ya que p se estima mediante: $p = \frac{i-0,5}{T}$. Las coordenadas y_i del QQ-plot son las (T_u/T) mayores pérdidas realmente observadas en la muestra.

[ver Christoffersen Chapter4Results.xls, question 7]

Aplicación práctica de los procedimientos de EVT

1. Comenzamos estimando un modelo de volatilidad para las rentabilidades, que utilizaremos para convertirlas en rentabilidades *estandarizadas*, dividiendo por σ_t.

2. Fijado un umbral q_u para la definición de cola de la distribución (99% ó 95%, por ejemplo) calculamos el umbral u como el percentil $1 - q_u$ en las rentabilidades estandarizadas. El umbral será una rentabilidad negativa.

3. Consideramos ahora la distribución de *pérdidas estandarizadas*, obtenidas cambiando de signo las rentabilidades estandarizadas.

4. Estimamos el parámetro ξ de grosor de cola a partir de las pérdidas estandarizadas.

5. Fijado un nivel de significación p para la estimación del VaR, calculamos $F_{1-p}^{-1} = u\left(\frac{T_u}{T}\frac{1}{p}\right)^{\xi}$ [*denotado como EVT en Christoffersen Chapter4Results.Question 6*] y multiplicamos por la volatilidad de cada día para obtener el VaR. Hemos de utilizar el umbral u cambiado de signo, positivo, puesto que estamos trabajando ahora con la distribución de las pérdidas.

6. Para generar un QQ-plot, representamos las T_u rentabilidades estandarizadas menores (las más negativas, recordemos que estamos modelizando la cola izquierda de la distribución de rentabilidades) frente a los cuantiles de las distribución que queremos utilizar como referencia en el QQ-plot. En el caso de la EVT, los cuantiles están dados por $F_{1-p}^{-1} = u\left(\frac{T_u}{T}\frac{1}{p}\right)^{\xi} = u\left(\frac{T_u}{i-.50}\right)^{\xi}$

Ejemplo: En Chapter4Results.xls, de Christoffersen, se estima un GARCH(1,1) bajo una distribución t-Student para la innovación del proceso, y se dibuja el QQ-plot para los residuos del modelo estimado (Question 4 y Question 5). En Question 6 se utiliza el estimador de Hill para la GPD.

9.4.5 Distribuciones de Johnson

Johnson (1949) introdujo una familia de distribuciones mediante una traslación de la distribución $N(0,1)$.Si $Z \sim N(0,1)$, la variable aleatoria Y definida mediante la transformación:

$$Z = \gamma + \delta.g\left(\frac{Y-\xi}{\lambda}\right)$$

donde: $\xi, \lambda > 0, \gamma, \delta > 0$ son parámetros. La distribución resultante depende de la función g que se utilice. Si $g(u) = u$, tenemos la distribución Normal con $E(Y) = \xi - \gamma\lambda/\delta, Var(Y) = (\lambda/\delta)^2$ (de modo que si $\delta = \lambda$ y $\gamma = \xi$, Y seria $N(0,1)$); si $g(u) = \log\frac{u}{1-u}$ tenemos una distribución acotada, conocida como distribución Johnson $S_B : Y \sim JS_B(\xi, \lambda, \gamma, \delta)$; si $g(u) = \ln(u)$ tenemos la distribución log-normal; si $g(u) = senh^{-1}(u)$ obtenemos una distribución de probabilidad no acotada, conocida como la distribución Johnson $S_U : Y \sim JS_U(\xi, \lambda, \gamma, \delta)$.

Por tanto, una variable aleatoria Y sigue una *distribución S_U de Johnson* no acotada si:

$$\frac{Y-\xi}{\lambda} = senh\left(\frac{Z-\gamma}{\delta}\right)$$

donde $Z \sim N(0,1)$. Los parámetros $(\xi, \lambda, \gamma, \delta)$ determinan, respectivamente, la localización, la escala, la asimetría, y la curtosis de la distribución. Con cuatro parámetros, la distribución es muy flexible y puede recoger comportamientos muy variados.

Si $Y \sim JS_U(\xi, \lambda, \gamma, \delta)$:

$$F_Y(y) = P(Y < y) = P\left[N(0,1) < \gamma + \delta.senh^{-1}\left(\frac{y-\xi}{\lambda}\right)\right] = \Phi\left(\gamma + \delta.senh^{-1}\left(\frac{y-\xi}{\lambda}\right)\right)$$

donde Φ denota la función de distribución de una variable $N(0,1)$. La función de densidad es:

$$f_Y(y) = \frac{\delta}{\lambda\sqrt{1+\left(\frac{y-\xi}{\lambda}\right)^2}}\phi\left(\gamma + \delta.senh^{-1}\left(\frac{y-\xi}{\lambda}\right)\right)$$

donde ϕ denota la función de densidad de una variable $N(0,1)$. Su media y varianza son:

$$\begin{aligned} \mu &= \xi + \lambda\sqrt{\omega}senh\,(\Omega) \\ \sigma^2 &= \frac{\lambda^2}{2}(\omega - 1)\,(\omega.\cosh(2\Omega) + 1) \end{aligned}$$

donde $\omega = \exp(\delta^{-2})$, y $\Omega = \gamma/\delta$.

Cada cuantil x_p de la distribución de Johnson no acotada está relacionado con el cuantil análogo z_p de la distribución $N(0,1)$:

$$q = F_Y(x_p) \Rightarrow \Phi^{-1}(q) = z_p = \gamma + \delta.senh^{-1}\left(\frac{x_p-\xi}{\lambda}\right) \Rightarrow x_p = \lambda.senh\left(\frac{z_p-\gamma}{\delta}\right) + \xi$$

por lo que dado el VaR z_p de una $N(0,1)$ a un determinado nivel de significación y para un horizonte dado, tenemos:

$$VaR_p\left(JS_U\right) = -\left(\lambda.senh\left(\frac{z_p-\gamma}{\delta}\right)+\xi\right)$$

que será un valor numérico de signo positivo.

Es posible estimar los parámetros de la distribución de Johnson a partir de los cuatro primeros momentos muestrales de la distribución de rentabilidades utilizando el Método de Momentos, del siguiente modo (algoritmo de Tuenter (2001)):

Calculamos una cota superior y una inferior para ω :

$$\begin{aligned}
\omega^{\text{sup}} &= \left(-1+\sqrt{2(\kappa+2)}\right)^{1/2} \\
\omega^{\text{inf}} &= \max(\omega_1,\omega_2)
\end{aligned}$$

donde ω_1 es la única raiz positiva de $\omega^4+2\omega^3+3\omega^2-\kappa-6=0$, donde κ es el *exceso* de curtosis muestral, y ω_2 es la única raiz positiva de $(\omega-1)(\omega+2)^2=\tau^2$ siendo τ el coeficiente muestral de asimetría.

A continuación, denotando por $m(\omega)$ la función:

$$m(\omega)=\left(4+2\left[\omega^2-\frac{\kappa+6}{\omega^2+2\omega}+3\right]\right)^{1/2}-2,$$

hallamos el valor numérico ω^*, comprendido entre ω^{inf} y ω^{sup} tal que

$$\left(\omega^*-1-m(\omega^*)\right)\left(\omega^*+2+\frac{m(\omega^*)}{2}\right)^2=\tau^2$$

Finalmente, las estimaciones paramétricas son:

$$\begin{aligned}
\delta &= \frac{1}{\sqrt{\ln\omega^*}} \\
\gamma &= \theta\delta,\ \text{siendo } \theta=-sgn(\tau)senh^{-1}\left(\sqrt{\frac{(\omega^*+1)(\omega^*-1-m(\omega^*))}{2\omega^* m}}\right) \\
\lambda &= \sqrt{\frac{2\sigma^2}{(\omega^*-1)(\omega^*\cosh 2\theta+1)}} \\
\xi &= \mu-sgn(\tau)\frac{\sigma\sqrt{\omega^*-m(\omega^*)-1}}{\omega^*-1}
\end{aligned}$$

donde μ y σ son la media y desviación típica de la muestral de rentabilidades.

La distribución de Johnson da resultados mucho mejores que la aproximación de Cornish-Fisher. Para una distribución t-Student con 30 grados de libertad apenas comete error de aproximación. Incluso con 10 grados de libertad el error es despreciable. Para una distribución t-Student con 6 grados de libertad, la distribución de Johnson subestima ligeramente el VaR al 5% y 10% y lo sobreestima ligeramente para percentiles 0,1% y 1%.

Ejercicio [ver EIV.3.5] En la hoja de cálculo, primero se obtienen, mediante el algoritmo de Tuenter, los valores de ω_1, ω_2, y a continuacion, el valor de ω^*.

La distribución JS_U admite una representación alternativa mediante la función de densidad:

$$f_Y(y) = \frac{\delta}{\lambda c} \frac{1}{\sqrt{r^2+1}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

donde:

$$\begin{aligned}
z &= -\gamma + \delta . senh^{-1}(r) = -\gamma + \delta . \log \left[r + (r^2+1)^{1/2} \right], \ con\ Z \sim N(0,1); \\
r &= \frac{y - \left(\xi + c\lambda \omega^{1/2} senh\,(\Omega) \right)}{c\lambda}; \ \omega = \exp(\delta^{-2}); \ \Omega = -\gamma/\delta; \\
c &= \left[\frac{1}{2} (\omega - 1) \left(\omega . \cosh(2\Omega) + 1 \right) \right]^{-1/2}
\end{aligned}$$

con esta parametrización, los parámetros ξ y λ son la esperanza y la varianza de Y. El parámetro γ determina la asimetria, y tiene el mismo signo que ésta. El parámetro δ debe estar en muchos casos por encima de 1, y determina la curtosis.

9.4.6 Distribución t-Student Generalizada Asimétrica

La distribución SGT (Theodossiou, 1998), es una distribución muy flexible que puede captar tanto el exceso de curtosis como la asimetría presente en las series financieras. Y que engloba como casos particulares un conjunto amplio de distribuciones más sencillas, como la distribución normal o la t-Student.

$$f(z_t \mid \lambda, \eta, k) = C \left(1 + \frac{\mid z_t + \delta \mid^k}{\left(\frac{\eta+1}{k}\right) \left[1 + sign(z_t + \delta)\lambda \right]^k \theta^k} \right)^{-(\eta+1)/k}$$

donde:

$$\begin{aligned}
C &= (0,5)\, k \left(\frac{\eta+1}{k} \right)^{-1/k} \frac{1}{\theta B\left(\frac{\eta}{k}, \frac{1}{k}\right)}, \\
\theta &= \frac{1}{\sqrt{g - \rho^2}}, \quad \delta = \theta\rho, \\
\rho &= 2\lambda B \left(\frac{\eta}{k}, \frac{1}{k} \right) \left(\frac{\eta+1}{k} \right)^{1/k} B \left(\frac{\eta-1}{k}, \frac{2}{k} \right), \\
g &= (1+3\lambda)^2 B \left(\frac{\eta}{k}, \frac{1}{k} \right)^{-1} \left(\frac{\eta+1}{k} \right)^{1/k} B \left(\frac{\eta-1}{k}, \frac{3}{k} \right)
\end{aligned}$$

En esta densidad, λ es el parámetro de simetría ($\mid \lambda \mid < 1$), $\eta > 2$ recoge el grosor de las colas, $k > 0$ recoge la curtosis o grado de apuntamiento, $B(.)$ es la función Beta, y z_t son los rendimientos estandarizados.

La distribución normal surge como caso particular cuando $\lambda = 0, k = 2, \eta = \infty$, mientras que la t-Student con v grados de libertad aparece cuando $\lambda = 0, k = 2, \eta = v$.

9.5 VaR histórico para carteras lineales

Analizamos el cálculo del VaR histórico en carteras lineales sin plantearnos la precisión en la estimación del VaR en cuantiles extremos, lo que podría hacerse utilizando Cornish-Fisher o las distribuciones de Johnson. Por el contrario, en los ejemplos vamos a aplicar la regla de la raiz cuadrada al extrapolar la varianza en el tiempo. Hacemos el ajuste en volatilidad mediante EWMA, aunque podría hacerse mediante un modelo GARCH. Cnviene dejar claro que utilizamos los modelos EWMA y GARCH unicamente por simplicidad.

En toda esta sección podria añadirse la preocupación por la precisión en la estimación del VaR a un alto nivel de confianza.

9.5.1 VaR histórico para secuencia de flujos de caja

[Case Study IV.3.5.1,pestaña EIV.3.5.1]

Despues de proyectar los cash flows sobre los vertices de tipos de interés, se utiliza la representacion habitual de la cuenta P&L: $\Delta PV_t \approx -\theta' \Delta r_t$.Previamente se ha utilizado un esquema EWMA para ajustar volatilidad al final de la muestra. Luego se ajustan las P&L diarias en volatilidad, para hacerlas comparables con la última observación muestral. Se comparan los histogramas de P&L resultantes, y se calculan los VaR con ajuste de volatilidad y sin dicho ajuste.

Start date:	05-Jan-00	7		
End date:	31-dic-07	2023		
Sample size				
2017	Skewness	XS kurtosis		
Unweighted	-0.0288	45.0152		
Volatility Adjusted	-0.0687	-0.2267		
Significance level:	1.0%		0.5%	
Risk horizon (days):	10		10	
lambda	0.9		0.9	
	Unweighted	Volatility Adjusted	Unweighted	Volatility Adjusted
Historical VaR	£207,188	£266,505	£237,590	£278,591
Normal VaR	£194,403	£264,065	£215,475	£292,854
Ratio	107%	101%	110%	95%

9.5.2 VaR total, sistemático y específico para una cartera de acciones

[EIV.3.6] [EIV.3.7], ambos en [Case Study IV.3.5.2].

La rentabilidad de una cartera se construye a partir de las rentabilidades de los activos que la componen mediante:

$$r_t = w_T' x_t,\ t = 1, 2, ..., T$$

donde utilizamos la serie temporal reconstruida de una cartera que se hubiese mantenido históricamente con la misma composición que la que hemos configurado en T. Por otra parte, como el VaR histórico reflejará mejor las actuales condiciones de mercado si se utilizan rentabilidades ajustadas de volatilidad, podemos hacer este ajuste de dos modos distintos: 1) ajustando las rentabilidades de cada activo y aplicando los pesos que definen la cartera, 2) aplicando los pesos para definir la rentabilidad de la cartera y ajustando directamente esta rentabilidad.

Es importante que el ajuste de volatilidad preserve la estructura de correlaciones de las rentabilidades r_t de los activos. Sin embargo, en el primer caso, el ajuste individual por volatilidad modifica las correlaciones, por lo que necesitamos un ajutse multivariante por volatilidad. Para ello, si denotamos por V_t la matriz de covarianzas GARCH o EWMA de las rentabilidades y por Q_t la matriz de Cholesky correspondiente, tenemos:

$$V_t(r_t) = V_t = Q_t Q_t'$$

y hacemos la transformación del vector de rentabilidades r_t:

$$\tilde{r}_t = Q_T Q_t^{-1} r_t,\ t = 1, 2, ..., T$$

tenemos:

$$V_t(\tilde{r}_t) = V_t\left(Q_T Q_t^{-1} r_t\right) = Q_T Q_t^{-1} V_t(r_t)\left(Q_T Q_t^{-1}\right)' = Q_T\left(Q_t^{-1} Q_t\right)\left(Q_t' Q_t'^{-1}\right) Q_T' = Q_T Q_T' = V_T$$

manteniendo por tanto constante la estructura de la matriz de covarianzas.

En el caso de 2 activos, si denotamos:

$$V_t = \begin{pmatrix} \sigma_{1t}^2 & \sigma_{12t} \\ \sigma_{12t} & \sigma_{2t}^2 \end{pmatrix} \Rightarrow Q_t = \begin{pmatrix} a_t & 0 \\ b_t & c_t \end{pmatrix} \Rightarrow Q_t^{-1} = \frac{1}{a_t c_t}\begin{pmatrix} c_t & 0 \\ -b_t & a_t \end{pmatrix}$$

con $a_t = \sqrt{\sigma_{1t}^2}, b_t = \sigma_{12t}/\sqrt{\sigma_{1t}^2}, c_t = \sqrt{\sigma_{2t}^2 - \frac{\sigma_{12t}}{\sigma_{1t}^2}}$, por lo que:

$$Q_T Q_t^{-1} = (a_t c_t)^{-1}\begin{pmatrix} a_T c_t & 0 \\ b_T c_t - b_t c_T & a_t c_T \end{pmatrix}$$

y la transformación es:

$$r_{1t} = r_{1t}\sqrt{\frac{\sigma_{1T}^2}{\sigma_{1t}^2}}\ ;\ r_{2t} = r_{1t}\frac{b_T c_t - b_t c_T}{a_t c_t} + r_{2t}\frac{c_T}{c_t}$$

de modo que la primera rentabilidad solo experimenta el ajuste de volatilidad habitual.

En el Case Study IV.3.5.2 (pestaña EX_IV.3.6) se calcula el VaR ajustado de una cartera formada en un 70% por acciones de Citigroup y en un 30% por acciones de Apple. En la pestaña EX_IV.3.7 se hace la descomposición en VaR sistemático y VaR specífico. La estimación del VaR se realiza de 3 formas distintas: a) transformando las dos rentabilidades de modo que se preserve la correlacion entre ellas, b) generando la serie temporal histórica de la rentabilidad de cada activo de la cartera, ajustándola de volatilidad, y calculando su VaR, c) ignorando el efecto que la transformación tiene sobre la correlacion: ajustamos la rentabilidad de cada activo por separado, y se reconstruye la serie temporal de la rentabilidad de la cartera. En el ejemplo se utiliza un modelo EWMA para generar las series temporales de volatilidades, pero de igual modo podria haberse utilizado un modelo GARCH.

Si se va a calcular el VaR de distintas carteras sobre los mismos activos, es preferible ajustar las rentabilidades de los activos individuales. Entonces, debemos utilizar el primer procedimiento. Sin embargo, el error introducido con el tercer procedimiento no es generalmente muy grande.

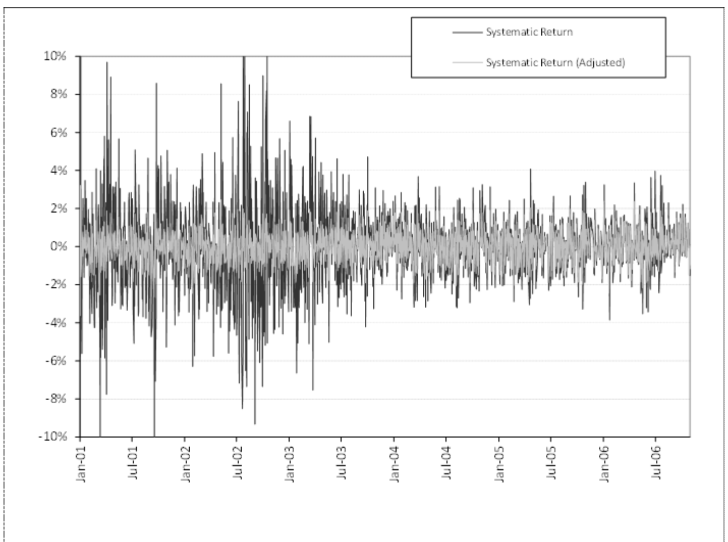
Citibank	0.3		
Apple	0.7		
Lambda EWMA	0.94		
Var Parameters			
Significance Level	0.01	0.05	0.001
Holding Period	10	10	10
Historical VaR (Unadjusted)	17.5%	10.5%	29.3%
Historical VaR Volatility Adjusted - Method (a)	19.2%	13.4%	25.6%
Historical VaR Volatility Adjusted - Method (b)	18.4%	13.1%	25.4%
Historical VaR Volatility Adjusted - Method (c)	18.1%	12.3%	25.2%

En un modelo factorial, podemos descomponer el VaR total en un VaR sistematico, debido a los factores de riesgo, y un VaR especifico. Solo que a diferencia de lo que sucede en el enfoque paramétrico del VaR, en el enfoque histórico y en el de Monte Carlo, esta descomposición no es aditiva. Para ello, comenzamos estimando las betas de las dos acciones [Fig.IV.3.23 en Case Study IV.3.5.2] El calculo del VaR 1% a 10 dias y su descomposición se lleva a cabo en la pestaña IV.3.7 en Case Study IV.3.5.2, para dos fechas diferentes, 21 de abril de 2008 y 30 de octubre de 2006. El componente sistemático de la rentabilidad de la cartera es:

$$r_{sis} = (\omega_{1T}\beta_{1T} + \omega_{2T}\beta_{2T})r_{m_t}$$

La descomposición se lleva a cabo, además, sin ajustar las rentabilidades y ajustándolas por la volatilidad del final de la muestra.

Lambda				
0.94				
	Weights	Betas	Weights	Betas
Citibank	0.3	1.08	0.3	2.43
Apple	0.7	2.43	0.7	1.51
Net		2.02		1.79
Start date:	02/01/2001		02/01/2001	
End date:	30/10/2006		21/04/2008	
Significance Level	1%		1%	
Risk Horizon	10		10	
Historical VaR	Unadjusted	Vol. Adjusted	Unadjusted	Vol. Adjusted
Total	17.4%	5.8%	17.5%	17.1%
Systematic	18.8%	6.0%	16.8%	17.7%
Specific	15.1%	6.1%	13.3%	17.6%
Normal VaR	Unadjusted	Vol. Adjusted	Unadjusted	Vol. Adjusted
Total	16.6%	6.3%	16.3%	18.1%
Systematic	16.9%	6.3%	14.8%	18.1%
Specific	14.2%	6.2%	13.1%	17.7%



9.5.3 VaR Equity y Forex de una cartera de accciones internacionales

[Case Study IV.3.5.3] ¿Cómo debe desagregarse el VaR de una cartera de acciones internacionales en sus componentes Equity y Forex? Sabemos que los cuantiles se transfoman de modo natural al aplicar transformaciones monotonas continuas. También sabemos que si la rentabilidad de la cartera sigue una distribución estable, tambien es posible agregar temporalmente los cuantiles, al igual que hacemos con la varianza. Sin embargo, no existe una manera de caracterizar la transformación de cuantiles dentro de una cartera, es decir, la relacion que existe entre los cuantiles de la cartera y los de los activos que la componen.

Consideremos un inversor UK que tiene una cartera en acciones US y UK a

21 de abril de 2008. Supongamos que el 70% del total de la cartera está invertido en libras esterlinas, en la forma de acciones UK, y 30% está en acciones US. La cartera US tiene una beta de 1.8 en relacion con el S&P500, y la cartera UK tiene una beta de 1.25 respecto del FTSE100.

Puesto que vamos a utilizar una muestra larga de datos diarios, desde 3 de enero de 1996 (3000 datos), utilizaremos *varianzas ajustadas* por un esquema EWMA con $\lambda = 0.94$.

Hemos de tener en cuenta que el precio en libras de una inversión en US$ es:

$$P_t^{£} = P_t^{\$}.X_t^{£/\$}$$

por lo que, en términos de rentabilidades logaritmicas:

$$\ln\left(P_{t+1}^{£}/P_t^{£}\right) = \ln(P_{t+1}^{\$}/P_t^{\$}) + \ln(X_{t+1}^{£/\$}/X_t^{£/\$}) \Rightarrow r_t^{£} = r_t^{\$} + r_t^{£/\$}$$

Si denotamos con subindice 1 la inversión en libras, y con subindice 2 la inversion en US$, el valor de la cartera en libras en cada momento es:

$$P_t^{£} = \omega_1 P_{1t}^{£} + \omega_2 P_{2t}^{£}$$

con $\omega_1 = 0.70$ y $\omega_2 = 0.30$ en este ejemplo. Para continuar, hemos de suponer que las rentabilidades logaritmicas y porcentuales son similares. Si denotamos por $y_{1t}^{£}$ e $y_{2t}^{\$}$ las rentabilidades diarias de los dos indices de mercado, tenemos las rentabilidades diarias de las carteras UK y US:

$$r_{1t}^{£} = \beta_1 y_{1t}^{£};\ r_{2t}^{\$} = \beta_2 y_{2t}^{\$}$$

donde β_1 y β_2 son las betas porcentuales de la cartera en libras respecto de FTSE100 y de la cartera en US$ respecto del S&P500, respectivamente. Por tanto, la rentabilidad de la cartera es:

$$\begin{aligned} r_t^{£} &\approx \omega_1 r_{1t}^{£} + \omega_2 r_{2t}^{£} = \omega_1 r_{1t}^{£} + \omega_2 (r_{2t}^{\$} + r_{2t}^{£/\$}) = \\ &= (\omega_1 \beta_1 y_{1t}^{£} + \omega_2 \beta_2 y_{2t}^{\$}) + \omega_2 r_{2t}^{£/\$} = r_t^E + r_t^F \end{aligned}$$

cuyo cuantil empírico, cambiado de signo, nos daria la estimación del VaR. Los componentes Equity VaR y Forex VaR son los cuantiles de las distribuciones de los dos sumandos que aparecen en la expresión anterior.

Para el cálculo del VaR marginal, al estar utilizando el método histórico, no tenemos expresiones analíticas del VaR en el que tomar derivadas, por lo que estimamos numéricamente las derivadas del VaR respecto a los parámetros del mismo:

$$Descomposición\ del\ VaR_p\ M\arg inal\ = \theta_1 g(\theta_1) + \theta_2 g(\theta_2)$$

Dichos parámetros son los que dependen de la elección del inversor, y estos son en este caso la sensibilidad de la cartera a cada factor (Equity, Forex): $\theta = (\omega_1 + \omega_2, \omega_2) = (1, \omega_2)$, por lo que calculamos las rentabilidades perturbadas:

$$
\begin{aligned}
r_t^{£}(\varepsilon,0) &\approx (1+\varepsilon)(\omega_1\beta_1 y_{1t}^{£}+\omega_2\beta_2 y_{2t}^{\$})+\omega_2 r_{2t}^{£/\$}=(1+\varepsilon)r_t^E+r_t^F \\
r_t^{£}(0,\varepsilon) &\approx (\omega_1\beta_1 y_{1t}^{£}+\omega_2\beta_2 y_{2t}^{\$})+(1+\varepsilon)\omega_2 r_{2t}^{£/\$}=r_t^E+(1+\varepsilon)r_t^F
\end{aligned}
$$

y obtenemos el vector gradiente: $g(\theta)=(g(\theta_1),g(\theta_2))$ de primeras derivadas del VaR:

$$
\begin{aligned}
g(\theta_1) &= \frac{VaR(1+\varepsilon;\omega_2)-VaR(1;\omega_2)}{\varepsilon} \\
g(\theta_2) &= \frac{VaR(1;\omega_2+\varepsilon)-VaR(1;\omega_2)}{\varepsilon}
\end{aligned}
$$

lo que implica generar una serie perturbada de rentabilidades de la cartera mediante $(1+\varepsilon)r_t^E+r_t^F$ y calcular su VaR, para compararlo con el VaR de las rentabilidades originales. Luego hacemos lo mismo con $r_t^E+(1+\varepsilon)r_t^F$.

Para obtener el VaR marginal multiplicamos el vector gradiente por el vector de sensibilidades a los factores, antes de perturbaciones:

$$
\begin{aligned}
VaR_p\ M\arg inal\ Equity &= (\omega_1+\omega_2)g(\theta_1)=g(\theta_1) \\
VaR_p\ M\arg inal\ Forex &= \omega_2 g(\theta_2)
\end{aligned}
$$

Finalmente, repetimos el ejercicio con las rentabilidades ajustadas, para ver el efecto que tiene el ajuste de la varianza muestral a la varianza al final de la muestra.

9.5.4 VaR de tipos de interés y Forex de una cartera de bonos internacionales

[Case Study IV.3.5.4] Un banco UK compra el 21 de abril de 2008 £50 millones de bonos US AA que se acaban de emitir a 5 años, con cupon anual del 4%. Puesto que el banco necesita comprar £50 millones en $US para financiar la compra, la rentabilidad total tendrá un componente de divisa. Los factores de riesgo son, por tanto, la curva swap, con calificacion AA, y el tipo de cambio £/US$. Queremos descomponer el VaR histórico en sus componentes de tipos de interés y Forex.

Calculamos el valor presente de los flujos de pago (4,4,4,4,104) utilizando la curva swap a 21 de abril de 2008. La suma nos da el precio (*fair price*) del bono: 101,1398. Por tanto, los £50 permite comprar 494.365 bonos de nominal 100 libras.

El VaR de tipos de interés puede calcularse de dos modos:

- Calculamos los PV01 de los flujos de caja en libras:

$$PV01_{it}^{£}=T_iC_i(1+R_{it,T})^{-(T+1)}10^{-4}$$

donde los cuatro primeros flujos vienen dados por: $C_i = n(100x) = (N/P)(100x)$, siendo x el cupon, n el numero de bonos, N el nominal, y P el precio del bono en el momento de compra. El último flujo es: $C_i = n100(1+x) = (N/P)100(1+x)$. A continuación, multiplicamos el vector constante $PV01_i$ por Δr_{it} a lo largo de la muestra para obtener los $P\&L_t$ en libras. Calculamos el percentil y multiplicamos por la raiz cuadrada del horizonte de riesgo, en días,

- Seguimos el enfoque histórico de cálculo de la distribución de $P\&L$. Para ello, aplicamos *a la curva swap actual* las variaciones registradas en puntos básicos a lao largo de la muestra, y calculamos el valor del bono dia a dia descontando los flujos de caja a los tipos resultantes. Los $P\&L$ son entonces la diferencia entre el valor de la cartera de bonos cada dia y la cuantía inicial, £50 millones. El valor del bono se puede convertir a $US al tipo de cambio diario, y calcular las $P\&L$ en dolares diariamente, y su VaR. Las variaciones diarias en $P\&L$ se deberán en parte a las variaciones en la curva swap y, en parte, a las variaciones en el tipo de cambio.

El *VaR total* está basado en la distribución de $P\&L$ en $US, porque hemos de recoger simultáneamente el efecto de variaciones en los tipos de interés, y también el efecto del tipo de cambio. El precio en dolares cada dia se obtiene aplicando al precio en libras el tipo de cambio. El VaR Forex se estima a partir de la distribución histórica de rentabilidades del tipo de cambio, y la distribución de $P\&L$ para el VaR total se obtiene convirtiendo el valor de la posición en libras a $US, y recalculando las $P\&L$. Solo que para obtener de acuerdo con el enfoque histórico el tipo de cambio que debe utilizarse cada dia en esta conversión del valor de la cartera de bonos, se aplica el rendimiento logaritmico del tipo de cambio de ese dia *al tipo de cambio actual*, el del dia en que se estima el VaR. Los resultados son casi idénticos.

Este ajuste es similar al que aplicamos en rentabilidades para igualar la varianza que experimentan las mismas a lo largo de la muestra, a la varianza al final de la misma, cuando se calcula el VaR. En este caso, el tipo de cambio ha mostrado una tendencia creciente desde el entorno del 1,5 $/£ hacia los 2,0 $/£. Lo que hacemos entonces es utilizar las rentabilidades diarias del tipo de cambio, pero aplicarlas a su última cotización, el 21 de abril de 2008, para calcular el VaR ese dia.

El VaR Forex es mayor que el VaR de tipos de interés, y el VaR total es inferior a la raiz cuadrada de la suma de los componentes del VaR al cuadrado, lo que indica una dependencia negativa entre el tipo de cambio y los tipos swap.

9.5.5 VaR de un crack spread trader

[Case Study IV.3.5.5] Crack spread is a term used in the oil industry and futures trading for the differential between the price of crude oil and petroleum products extracted from it - that is, the profit margin that an oil refinery can expect to make by "cracking" crude oil (breaking its long-chain hydrocarbons into useful shorter-chain petroleum products). In the futures markets, the "crack spread" is a specific spread trade involving simultaneously buying and selling contracts in crude oil and one or more derivative products, typically gasoline and heating oil. Oil refineries may trade a crack spread to hedge the price risk of their operations, while speculators attempt to profit from a change in the oil/gasoline price differential.

En este ejemplo analizamos una inversión que considera dos spreads: la diferencia entre el precio del futuro del fuel calefacción, y el precio del futuro en petróleo WTI de igual vencimiento, por un lado; la diferencia entre los precios de los futuros sobre gasolina y el futuro en petróleo WTI de igual vencimiento, por otro. Como los diferenciales pueden tomar valores negativos, no tiene sentido calcular rentabilidades porcentuales de estos factores de riesgo. Lo que hacemos es calcular las $P\&L$ diarias sobre el spread, expresado en términos nominales. Cada contrato de futuros es por 1000 barriles, con precio en $US por barril. Queremos calcular el VaR de la cartera y descomponerlo en los VaR de ambos crack spreads. Los factores de riesgo son los futuros con vencimiento constante en 1, 2, 3, 4, 5 y 6 meses. Supongamos que un inversor tiene posiciones en el spread HO_WTI: $(100, 50, 50, 0, 0, 250)$ y en el spread UL_WTI: $(-50, -100, 0, 0, 150, 100)$. Calculamos las volatilidades mediante un esquema EWMA con $\lambda = 0.94$ y ante la apariencia de variabilidad en volatilidad a lo largo de la muestra, aplicamos a las rentabilidades un ajuste en volatilidad. Como la volatilidad es alta hacia el final de la muestra, este ajuste hará que el VaR sobre datos $P\&L$ ajustados, sea superior al VaR calculado sobre datos sin ajustar. Calculamos las $P\&L$ multiplicando el número de contratos a cada vencimiento por la variación en el spread ese día, despues de hacer el ajuste en volatilidad. A continuación, sumamos los $P\&L$ de cada dia para las 6 posiciones en un spread, las 6 posiciones en el otro spread, y las 12 posiciones conjuntamente. Ambos cálculos se hacen con las$P\&L$ ajustadas y sin ajustar de volatilidad. Así podemos calcular el VaR de cada posición, así como el VaR total. Para el cálculo del VaR, multiplicamos el percentil, cambiado de signo, por 1000, que es el número de barriles en cada contrato.

El VaR total es muy inferior a la suma de los VaR de los dos spreads, pero superior a la raiz cuadrada de la suma de sus cuadrados. El ajuste en volatilidad aumenta sensiblemente, en efecto, el VaR.

9.6 Estimación de la ETL en el modelo histórico de VaR

En el modelo lineal paramétrico de VaR nos resultó posible proporcionar expresiones analíticas para el ETL porque haciamos supuestos especificos acerca de la forma funcional de la distribución de rentabilidades. En el modelo histórico de VaR, el ETL debe estimarse directamente, calculando el promedio de todas las pérdidas que han excedido del umbral definido por el VaR. Pero si, a modo de aproximación, hemos ajustado una distribución paramétrica a la distribución histórica de rentabilidades, podemos utilizar la expresión analítica correspondiente para el cálculo de la ETL.

ETL histórico paramétrico A partir de la expresión analítica para la pérdida en exceso en la *distribución Generalizada de Pareto*, tenemos:

$$ETL_p = VaR_p + \frac{\beta + \xi VaR_p}{1-\xi} = \frac{VaR_p + \beta}{1-\xi}$$

En otros dos casos, podemos deducir expresiones analíticas para la ETL a partir de la transformación de variables aleatorias que define dichas distribuciones. Utilizando el hecho de que la ETL al $100\%p$ en el caso de una Normal es $p^{-1}\varphi(Z_p)$, tenemos para la distribución de Johnson SU:

$$ETL_p = \lambda \; senh\left(\frac{p^{-1}\varphi(Z_p) - \gamma}{\delta}\right) + \xi$$

donde $Z_p = \Phi^{-1}(p)$.

Finalmente, la ETL bajo una distribución Cornish-Fisher puede aproximarse por:

$$ETL_p = f\left(p^{-1}\varphi(Z_p)\right)\sigma - \mu$$

donde: $f(x) = x + \frac{\tau}{6}\left(x^2-1\right) + \frac{\kappa}{24}x(x^2-3) - \frac{\tau^2}{36}x(2x^2-5)$.

Resultados empíricos acerca del ETL histórico [Case Study IV.3.6.2 S&P 500] Utilizando rentabilidades ajustadas de volatilidad para el período 4 de enero de 1950 a 9 de marzo de 2007, se estiman la expansión de Cornish-Fisher y la distribución SU de Johnson utilizando los cuatro primeros momentos de la distribución de rentabilidades. Se estima asimismo el ETL empirico, como promedio de la pérdidas en exceso del VaR, así como el ETL bajo el supuesto de Normalidad.

Al haber 14000 datos en la muestra, podemos calcular VaR y ETL con precisión incluso para altos niveles de confianza. Por el contrario, el trabajo con la GPD requiere muchos datos, puesto que consiste en estimar la distribución utilizando las observaciones que exceden de un determinado umbral. Las estimaciones del VaR de Johnson están más proximas a las obtenidas con GPD que las de Cornish-Fisher, el VaR empírico o el VaR Normal. Las estiamciones baj la distribucion de Johnson tienen la ventaja de que, a diferencia de las obtenidas con GPD, no dependen de la elección de un determinado umbral.

Como la distribución empírica de rentabilidades tiene una alta curtosis, con exceso de curtosis de 4,93, y una asimetria negativa muy significativa, el ETL Normal es muy inferior al proporcionado por otros métodos. Tengamos en cuenta que incluso con 14.000 datos, la estimación del ETL al 0,1% se realiza con solo 14 datos, y al 0,05%, con solo 7 datos, por lo que serán estimaciones poco precisas. En general, cuanto menor sea la muestra, mayor será el riesgo de modelo en la estimación del VaR y el ETL al utilizar métodos paramétricos o semi-paramétricos. Trabajando con rentabilidades del S&P500 ajustadas de volatilidad al 0.05%, las estimaciones segun distintos supuestos sobre la distribución de probabilidad de las rentabilidades son:

VaR (Standardized)	GPD VaR	GPD ETL (Standardized)	GPD ETL	Threshold
8.7960	5.81%	9.8112	6.48%	20%
10.1629	6.72%	11.4081	7.54%	10%
10.8413	7.16%	12.3379	8.16%	5%
10.9948	7.27%	13.4301	8.88%	1%
Empirical VaR	3.87%	Empirical ETL	5.56%	
Normal VaR	2.16%	Normal ETL	2.33%	
CF VaR	5.96%	CF ETL	7.31%	
Johnson VaR	5.60%	Johnson ETL	6.34%	

Desagregación del ETL histórico [Case Study IV.3.6.3] (Continúa del Case Study IV.3.5.3). La ETL se puede desagregar igual que el VaR, con la peculiaridad de que es siempre subaditivo. En la hoja de cálculo se estiman las ETL directamente, promediando las pérdidas en exceso del VaR. En este caso particular, el ajuste de volatilidad no tiene apenas efecto debido a que el 21 de abril de 2008 la volatilidad no estaba muy lejos de su promedio histórico. Pero si se cambia la muestra, puede verse que el ajuste puede resultar importante para otros momentos cronológicos.

10 Riesgo de estimación

Entendemos por Riesgo de Estimación la cantidad de incertidumbre que se genera por el hecho de que un determinado parámetro o estadístico, como puede ser la volatilidad de una cartera, o su VaR, hayan de ser estimados, al ser desconocidos. Hay diversas fuentes de riesgo de estimación, que surgen de los distintos componentes que entran a formar parte del proceso de estimación: el modelo estadístico/econométrico escogido, la metodología de estimación, o la muestra de datos seleccionada, por lo que el riesgo de estimación tiene varios componentes: el llamado Riesgo de Modelo, que se genera por la elección de un determinado modelo, entre todos los alternativos, para llevar a cabo la estimación; el riesgo procedente del inevitable error muestral en toda estimación, que depende de la muestra seleccionada para llevar a cabo la estimación, y el Riesgo Paramétrico, que se debe a la ausencia de precisión con que se lleva a cabo la estimación.

Es importante asociar a la estimacion del VaR una medida de la precisión con que se ha obtenido su valor numérico. El modo en que se haga depende del enfoque que se haya seguido en la estimación numérica del VaR, ya sea por simulación o mediante un modelo paramétrico. Cuando se obtiene por simulación, se genera una distribución de probabilidad simulada para la rentabilidad, pero la estimación del VaR puede ser muy sensible al tamaño de la muestra utilizada o al número de simulaciones históricas en que basamos su cálculo. De hecho, el tamaño muestral parece ser un condicionante importante del valor numérico del VaR (Case Study 4.3.3.1). Cuando se utiliza un enfoque paramétrico, suponiendo Normalidad e independencia de las rentabilidades, es sencillo estimar la matriz de covarianzas de los factores de riesgo.

Recordemos (ver capítulo sobre Volatilidad) que cuando se estima la varianza de la rentabilidad mediante la varianza muestral, entonces la desviación típica de la volatilidad es:

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

mientras que si se estima mediante un esquema de media móvil ponderada (EWMA), tenemos:

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

10.1 Distribución del estimador del VaR en modelos lineales paramétricos

10.1.1 Rentabilidades con distribución Normal

Es interesante utilizar la desviación típica asociada a la estimación de la volatilidad para obtener una desviación típica aproximada para el VaR. Tal desviación típica podría utilizarse para obtener intervalos de confianza para el VaR o para contrastar si los valores numéricos obtenidos para el VaR por dos procedimientos distintos son estadisticamente diferentes.

Si se espera que la cartera genere una rentabilidad igual a la del activo sin riesgo, y si dicha rentabilidad puede suponerse Normal, entonces el VaR es proporcional a la volatilidad (desviación típica) de la cartera, siendo sencillo obtener sus propiedades estadísticas. Para un horizonte de h dias, tenemos:

$$VaR_p(h) = \Phi^{-1}_{1-p}\sigma\sqrt{h}$$

Puesto que p y h son constantes, tenemos:

$$DT(VaR_p(h)) \simeq \Phi^{-1}_{1-p}\frac{\hat{\sigma}}{\sqrt{2T}}\sqrt{h} = \frac{VaR_p(h)}{\sqrt{2T}} \qquad (24)$$

Mientras que si la varianza se ha estimado mediante un esquema EWMA, tenemos:

$$DT(VaR_p(h)) \simeq \Phi^{-1}_{1-p}\hat{\sigma}\sqrt{\frac{1-\lambda}{2(1+\lambda)}}\sqrt{h} = VaR_p(h).\sqrt{\frac{1-\lambda}{2(1+\lambda)}}$$

Rentabilidades con distribución t-Student Cuando las rentabilidades son i., i.d., con distribución t-Student con ν grados de libertad, y se espera que dicha cartera proporcione, en media, la rentabilidad del activo sin riesgo, la estimación del VaR a un período es:

$$tStudent\ VaR_{p,\nu} = \sqrt{\frac{\nu-2}{\nu}}t^{-1}_{\nu,1-p}\sigma$$

Para horizontes cortos (h pequeño), la extrapolación lineal en la desviación típica es razonable, y tenemos para h períodos:

$$tStudent\ VaR_{p,\nu}(h) = \sqrt{h\frac{\nu-2}{\nu}}t_{\nu,1-p}^{-1}\sigma$$

En comparación con el supuesto de Normalidad de las rentabilidades, la leptocurtosis de la distribución t-Student hará que el VaR calculado bajo este supuesto sea superior al estimado bajo Normalidad, y que los errores estándar sean asimismo mayores. Las expresiones del VaR pueden utilizarse de nuevo para la estimación del error estándar del VaR de modo similar a como vimos en el caso de la distribución Normal.

$$DT(VaR_{p,v}(h)) \simeq \sqrt{\frac{\nu-2}{\nu}}t_{\nu,1-p}^{-1}\frac{\hat{\sigma}}{\sqrt{2T}}\sqrt{h} = \frac{tStudent\ VaR_{p,v}(h)}{\sqrt{2T}} \tag{25}$$

10.1.2 Distribución no paramétrica del estimador del VaR

Cuando el cálculo del VaR se basa en un modelo factorial, el VaR se obtiene a partir de una forma cuadrática, en la que aparecen tanto la matriz de covarianzas de los factores, como las estimaciones numéricas de las betas, las sensibilidades de la cartera a cada uno de los factores. Calcular expresiones para la desviación típica del VaR incluso con carácter de meras aproximaciones es entonces complejo, y puede ser preferible construir directamente una aproximación a la distribución de probabilidad del VaR, tenendo en cuenta que es un cuantil de la distribución de probabilidad de la rentabilidad de la cartera, como hacemos en esta sección.

Cuando la estimación del VaR se basa en simulación histórica o en simulación Monte Carlo, el error muestral puede ser una fuente importante de error de estimación. Incluso en el modelo de simulación histórica (el modelo con ajuste de volatilidad o filtrado de rentabilidades no paramétrico o semiparamétrico), el error muestral puede introducir bastante imprecisión acerca de la estimación numérica del VaR. El error muestral es más fácil de controlar en la simulación Monte Carlo, pero en ella tenemos el componente de riesgo procedente de la estimación de los parámetros del modelo. En la simulación histórica, el tamaño muestral escogido para las simulaciones es el principal determinante de la precisión (o su ausencia) en la estimación del VaR. En la simulación Monte Carlo, utilizando un elevado número de simulaciones, junto con técnicas de reducción de varianza, podemos conseguir que el riesgo de estimación paramétrica y el riesgo de modelo sean las principales factores explicativos de la falta de precisión (elevada desviación típica) en la estimación del VaR por este método.

Como proporción del valor de la cartera, el $VaR_p(h)$ es igual a $-\alpha$, siendo α el cuantil p de la distribución de rentabilidades de la cartera a horizonte de h días. Pero desconocemos dicha distribución de probabilidad, de la que solo disponemos de una muestra. Por tanto, podemos utilizar dicha muestra para estimar el cuantil p poblacional, pero sería conveniente tener una medida del grado de precisión de dicha estimación.

Aproximación Normal Consideremos un activo cuyas rentabilidades siguen una función de distribución de probabilidad F. Si conociésemos F entonces, dado un nivel de significación p, podríamos calcular el cuantil poblacional $\alpha = F^{-1}(p)$ [por ejemplo, en una población $N(0,1)$, si $p = 1\%$, entonces $\alpha = -2.33$].

Si ahora tomamos una muestra aleatoria de rentabilidades (una serie temporal) de dicha distribución, denotemos por $I_t(p)$ la función indicatriz que toma el valor 1 si la observación t está por debajo de $F^{-1}(p)$, y toma el valor 0 en caso contrario. Sea ahora $X(T,p)$ la suma de dichas funciones a lo largo de toda la muestra: $X(T,p) = \sum_{t=1}^{T} I_t(p)$. La variable aleatoria $I_t(p)$ sigue una distribución de Bernoulli, $B(p)$,[20] mientras que $X(T,p)$ sigue una distribución binomial $B(T,p)$. Por tanto, su esperanza matemática es Tp y su varianza $Tp(1-p)$. Si $T = 1000$ y $p = 0,01$,entonces $X(T,p)$ oscilará en distintas muestras en torno a $Tp = 10$.

Permitiendo tender T a infinito, podemos aplicar la aproximación Normal, y tenemos que:

$$\frac{X(T,p) - Tp}{\sqrt{Tp(1-p)}} \sim N(0,1) \tag{26}$$

A partir de esta expresión podemos generar un intervalo al nivel del 100(1-ε)% de confianza para el número de observaciones muestrales por debajo del cuantil p poblacional, $X(T,p)$:

$$\left[Tp - \Phi^{-1}_{\varepsilon/2}\sqrt{Tp(1-p)}\ ;\ Tp + \Phi^{-1}_{\varepsilon/2}\sqrt{Tp(1-p)}\right]$$

donde Φ denota la función de distribución Normal estándar, $N(0,1)$.

Es importante distinguir entre la probabilidad p que caracteriza el cuantil que buscamos, y el nivel de confianza $1-\varepsilon$ al que construimos el intervalo para dicho cuantil. Por ejemplo, con $p = 5\%, \varepsilon = 1\%$ y una muestra de 500 datos, dicho intervalo sería $(11\ ;\ 37)$. Ordenaríamos la muestra de rentabilidades de menor a mayor y tomaríamos las rentabilidades numéricas que ocupasen los puestos 11 y 37 en dicha ordenación. Supongamos que estas son -1,40% y -1,05%, respectivamente. Nuestra conclusión sería que el cuantil 5% de la distribución de rentabilidades de la cual disponemos de una muestra de tamaño 500, está comprendida en el intervalo (-1,40%;-1,05%) con una probabilidad del 99%. Por supuesto que podemos también calcular exactamente el cuantil 5% muestral, que será la rentabilidad que ocupe el lugar 25 en la muestra ordenada, y estará comprendida entre -1,40% y -1,05%. Supongamos que resulta ser igual a -1,21%, lo que nos daría una estimación del cuantil 5% poblacional, pero no conoceríamos con qué precisión habríamos obtenido dicha estimación. Eso es lo que nos proporciona el intervalo de confianza que hemos construido.

Utilizando un supuesto acerca de la función de densidad Hasta aquí no hemos utilizado ninguna información acerca de F, la distribución muestral de

[20] Pues toma el valor 1 cuando el dato es inferior a α lo que sucede con probabilidad p, y otma el valor cero en caso contrario.

rentabilidades de la cartera de activos. Veamos ahora cómo podemos incorporar tal información, lo cual hace el procedimiento aplicable, por ejemplo, en la estimación del VaR histórico.

Dividiendo en (26) numerador y denominador por T :

$$\frac{\varphi(T,p)-p}{\sqrt{p(1-p)/T}} \sim N(0,1)$$

donde $\varphi(T,p)$ denota el porcentaje de observaciones muestrales por debajo del cuantil poblacional α. Denotemos por $q(T,p)$ el estimador del $p-$cuantil de F en una muestra concreta de rentabilidades. Dicho estimador está en relación biunivoca con el estimador de la proporción de observaciones muestrales que caen por debajo de $q(T,p)$ mediante: $q(T,p) = F^{-1}(\varphi(T,p))$.[21] En consecuencia:

$$F\left(q(T,p)\right) \sim N\left(p, T^{-1}p(1-p)\right) \tag{27}$$

Ahora, supongamos que F es aproximadamente lineal en el rango relevante de la distribución, lo que podemos expresar como:

$$F\left(q(T,p)\right) \simeq q(T,p) f\left(q(T,p)\right) \tag{28}$$

siendo f la función de densidad. Dicho de otro modo, suponemos que la función de densidad es lineal en la región de la cola de la distribución que estamos considerando.[22] No puede negarse que este supuesto, sugerido en Kendall (1940), es bastante estricto y condiciona la validez rigurosa de los resultados que siguen.

Bajo este supuesto, tenemos, por (27) que:

$$q(T,p) f\left(q(T,p)\right) \sim N\left(p, T^{-1}p(1-p)\right)$$

es decir:

$$q(T,p) \sim N\left(\frac{p}{f\left(q(T,p)\right)}, \frac{T^{-1}p(1-p)}{\left[f\left(q(T,p)\right)\right]^2}\right)$$

por lo que la desviación típica aproximada para el estimador muestral del cuantil p es:

$$DT\left(q(T,p)\right) \sim \frac{\sqrt{T^{-1}p(1-p)}}{f\left(q(T,p)\right)} \tag{29}$$

que podemos utilizar para estimar un intervalo de confianza al 100(1-ε)% para el cuantil poblacional utilizando información acerca de la distribución de

[21] De hecho si las estimaciones muestrales d ela proporción y del cuantil coincidiesen con los verdaderos valores poblacionales, tendriamos: $\varphi(T,p) = \alpha$ y también: $q(T,p) = F^{-1}(\alpha)$, cumpliéndose la igualdad.

[22] Si F es lineal en dicha región: $F(x) = kx$, por lo que, tomando derivadas: $f(x) = k$, lo que implica que: $F(x) = xf(x)$, como aparece en (28)

rentabilidades a través de la presencia en esta expresión de la función de densidad f.

El VaR es el cuantil p cambiado de signo, por lo que su desviación típica será misma que la desviación típica del cuantil, que acabamos de deducir. Dado un nivel de significación para el VaR, por ejemplo del 1%, ese será nuestro p. Si partimos de un determinado supuesto acerca de la función de distribución F seguida por las rentabilidades de la cartera, comenzamos invirtiendo dicha función al nivel p. Es decir, encontramos el número real α tal que[23] $F(\alpha) = p$. A continuación, calculamos el valor numérico de la función de densidad[24] de las rentabilidades en dicho punto $f(\alpha)$ y lo llevamos a la expresión (29).

Por ejemplo, bajo una distribución Normal estándar, el cuantil 5% es $\alpha = 1,65$. Con una muestra con $T = 100$, el intervalo de confianza del 95% de confianza sería $(1,23; 2,06)$, que puede considerarse muy amplio. Con $T = 250$, el intervalo sería $(1,35; 1,94)$, mientras que con $T = 1250$, el intervalo se reduciría a $(1,53; 1,76)$. Si trabajamos con el VaR 1%, tenemos para la Normal estándar el cuantil 2,33. Con un año de datos ($T = 250$), el intervalo de confianza sería $(1,81; 2,84)$, notablemente más amplio que el obtenido para el VaR 5%.

Dos comentarios acerca de la aplicación práctica de este resultado: 1) requiere conocer las funciones de distribucion y de densidad f de la rentabilidad de la cartera, y 2) se basa en la validez del supuesto acerca de que la cola de la distribución de rentabilidades es localmente plana, siendo un resultado sólo aproximado en los demás casos.

Comparación de la precisión de distintos métodos en la estimación del VaR Hemos visto dos procedimientos para la estimación del VaR: por un lado, acabamos de ver como utilizar el cuantil de la distribución observada de rentabilidades; por otro lado, antes vimos cómo utilizar la desviación típica de dicha distribución. A diferencia de los cuantiles, la desviación típica utiliza información relativa a toda la distribución de rentabilidades, por lo que cabe esperar que sea más eficiente, generando estimaciones más precisas. Para el VaR al 95% de confianza tendríamos: $DT(\alpha\hat{\sigma}) = \alpha.DT(\hat{\sigma}) = \alpha\hat{\sigma}/\sqrt{2T}$. Con un año de datos tendríamos alrededor del percentil 5%, que es 1,65, un intervalo del 95% (semiamplitud dos desviaciones típicas) igual a $1,65 \pm (1,96)(1,65)\hat{\sigma}/\sqrt{500}$. En una Normal estándar ($\sigma = 1$), esto genera un intervalo de confianza para el VaR de: $(1,51; 1,79)$, frente al intervalo (1,38; 1,97) que obtuvimos estimando el VaR a través del percentil.

Que los errores estándar obtenidos a partir del error estándar de la estimación de la varianza (24) sean menores, sugiere que la estimación del VaR por este

[23] Si suponemos una distribución de rentabilidades Normal, la función NORMINV de Excel, a la que hemos de dar como inputs los valores numéricos de α, junto con la esperanza matemática y la desviación típica de la distribución de rentabilidades, nos proporciona x.

[24] En el caso de una Normal, la función NORMDIST de Excel, con inputs $(x, \mu, \sigma, falso)$ nos proporciona dicha estimación, siendo μ, σ la esperanza y desviación típica de la distribución de rentabilidades. El último input "falso" hace que el resultado sea el valor numérico de la función de densidad en el punto x, y no el valor numérico de la función de distribución, lo que sucedería si escribiésemos "verdadero", en lugar de "falso".

procedimiento es más eficiente (ver *Example* IV.6.4, Ejercicio 4.2). Esto sucede porque dichos errores estandar utilizan la información acerca de que la distribución de rentabilidades es Normal. Si dicho supuesto es correcto, sería preferible utilizar estos errores estándar. Sin embargo, si las rentabilidades no siguen una distribución Normal, los errores estandar que resultan por este procedimiento no serían correctos. Por otra parte, si el supuesto que se ha hecho acerca de que F sea lineal en la cola de la distribución no se cumple, las desviaciones típicas calculadas pueden ser bastante incorrectas, como puede comprobarse si se utilizan ambos métodos para estimar desviaciones tipicas del VaR en el caso de una distribución Normal de rentabilidades.

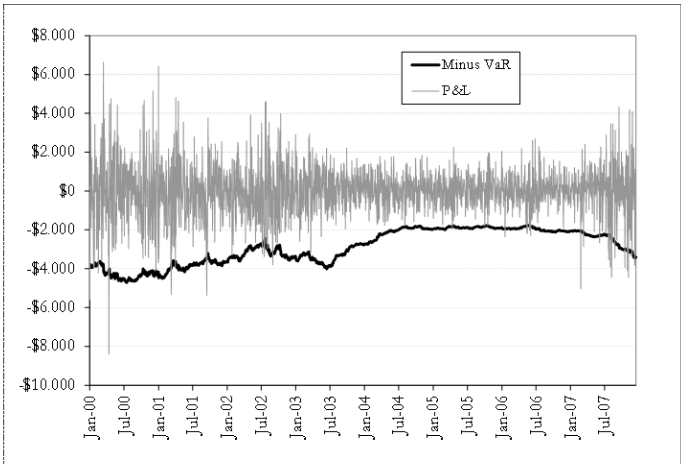
10.2 Validación del model de estimación del VaR

10.2.1 Backtesting

Un procedimiento de backtest se basa en el resultado obtenido por una determinada cartera representativa, en el sentido de tener una exposición típica a los distintos factores de riesgo que se consideren relevantes en el análisis. Hemos de suponer que o bien las posiciones en los distintos activos o los porcentajes invertidos en ellos. Mantener constantes las ponderaciones implica que se hayan estado ajustando continuamente las posiciones, segun iban variado los precios de los activos, por lo que se habla entonces de un VaR dinámico. Si en la cartera se permiten posiciones cortas y largas, entonces no cabe mantener las ponderaciones constantes, sino las posiciones en los activos. Otro criterio consiste en suponer que lo que se ha mantenido constante con el paso del tiempo han sido las betas de la cartera con respecto a cada uno de los factores considerados.

Por supuesto que el resultado del backtest depende no solo del modelo seguido para el cálculo de laVaR, sino de la cartera considerada, por lo que el contraste puede dar resultados contrarios en distintas carteras. Un backtest debe llevarse a acabo con un número de datos suficientemente grande pues, de lo contrario, el contraste carecerá de potencia, no siendo capaz de rechazar un modelo de VaR que sea inadecuado. Asimismo, el backtest ha de realizarse con observaciones no solapadas. Esto significa que si se trata de un VaR a horizonte h, comenzaremos utilizando n observaciones (*período de estimación*) para estimar el VaR a horizonte h, es decir, en el instante $n+h$. Se calcula la rentabilidad observada en la muestra sobre ese periodo de 10 dias y se comprueba si se trata o no de una excepción. A continuación, moveremos la ventana muestral para recoger las observaciones $h+1$ a $n+h$. Con estas, se estima el VaR para el periodo $n+2h$, y se vuelve a desplazar la ventana muestral h observaciones hacia el futuro. Se procede de este modo hasta agotar la muestra. Con ello se tienen aproximadamente $(T-n)/h$ estimaciones del VaR a h periodos, y otras tantas observaciones de la rentabilidad sobre esos mismos h períodos. Ese es el número de ocasiones en que concluirímos si se ha producido o no una excepción. El periodo de estimación es generalmente bastante mas corto si se trata de un VaR paramétrico o se estima el VaR por Monte Carlo, que si se estima mediante simulación histórica.

Figure IV.6.7



La validación del modelo de estimación del VaR se basan en la excepciones que se observen a lo largo de la muestra. Unos contrates se basan en el número de tales excepciones; otros en la ocurrencia temporal de tales excepciones, y otros, en ambos aspectos. Otros contrastes tratan de encontrar factores explicativos de las excepciones, lo que nos llevaria a creer que no suceden de forma puramente aleatoria, como se supone en todo modelo de estimación del VaR. Para definir los contrastes, conviene tener en mente una función indicatriz que toma el valor 1 cuando se produce una excepción, y toma el valor cero cuando no se produce. Es decir, dado un nivel de significación α del VaR, tenemos:

$$\begin{aligned} I_{\alpha,t+1} &= 1 \text{ si } R_{t+1} < -VaR_{1,\alpha,t} \\ I_{\alpha,t+1} &= 0 \text{ en caso contrario} \end{aligned}$$

10.2.2 Contraste del número de excepciones

Si el modelo de estimación del VaR fuese adecuado, $I_{\alpha,t+1}$ seguiria una distribución de Bernoulli con probabilidad α. Si denotamos por $X_{n,\alpha}$ el número de excepciones observadas, esta vriable sigue una distribución Binomial $B(n,\alpha)$. Por tanto, el´número de excepciones esperadas sería $n\alpha$, y su varianza $n\alpha(1-\alpha)$. Cuando n es suficientemente grande, $X_{n,\alpha}$ es aproximadamente Normal, de modo que un intervalo de confianza del 95% para el número de excepciones sería:

$$\left(n\alpha - 1,96\sqrt{n\alpha(1-\alpha)}\ ;\ n\alpha + 1,96\sqrt{n\alpha(1-\alpha)}\right)$$

10.2.3 Guidelines

Para las instituciones financieras, es muy importante disponer de un modelo adecuado de estimación de su VaR. Debe justificar la calidad de dicho modelo mediante pruebas de backtesting, siguiendo las pautas propuestas por los acuerdos de Basilea. SI el modelo de VaR no resulta suficientemente satisfactorio de acuerdo con dichas normas, la institución puede verse obligada a disponer de un capital regulartorio igual a 4 veces su VaR nominal. El backtest que se proponía en Basilea II consideraba un VaR 1% a horizonte de un dia, utilizando 250 observaciones en su estimación. Por tanto, el número esperado de excepciones es de 2,50, con una desviación típica $\sqrt{n\alpha(1-\alpha)}$ =1,5732. El objetivo de los reguladores de riesgos es rechazar modelos que generan un VaR excesivamente bajo. Ello seria conveniente para la institución financiera, que mantendría un capital regulatorio reducido, con un coste bajo, pero tendría un alto número de excepciones. Los reguladores aceptan modelos de VaR con un número de excepciones igual o inferior a 4. TSe dice que tales modelos están en la zona verde, con un capital regulatorio igual a 3 veces su VaR. Si tienen entre 5 y 9 excepciones, el modelo pasa a estar en zona amarilla, y el multiplicador aumenta gradualmente entre 3,40 y 3,85 veces el VaR. La zona roja viene definida por 10 o más excepciones, con un multiplicador de 4, o la institución financiera debe variar su modelo de estimación del VaR.

Sin embargo, cuando el capital regulatorio se calcula utilizando un modelo de VaR interno a la institución financiera, se considera un VaR 1% a horizonte de 10 dias. ¿Por qué plantean entonces los reguladores un backtest con VaR a 1 día? La razon es que sería dificil poder disponer de un elevado número de observaciones del VaR a 10 dias. Por ejemplo, con 10 años de muestra dispondriamos de tan solo 250 observaciones de VaR, que no son pocas para un backtest. Podriamos plantearnos analizar elos resultados que genera un modelo de VaR con horizonte a 10 dias utilizando muestra diarias, pero el problema que tendríamos es que las excepciones ya no serían independientes en el tiempo, pues una caida importante dlemercado en un determinado dia tenderia a hacer que se produjeran excepciones durante los 10 días siguientes. Por tanto, aunque el número esperado de excepciones no variaría, la varianza de dicho número aumentaría considerablemente.

A pesar de ello, las instituciones calculan el VaR a 10 dias cada día. No está justificado efectuar contrastes basados en la independencia de las excepciones, pero se pueden obtener estimaciones de distintos eventos, como el de si el capital regulatorio se va a ver superado cada dia de una determinada semana.

Asimismo, los reguladores permiten en ocasiones a la sinstituciones financieras estimar su VaR a un dia, y multiplicarlo adecuadamente para elevarlo al horizonte de riesgo deseado. Pero, como sabemos, esto está justificado unicamente si las rentabilidades de la cartera son independientes en el tiempo y siguen una distribución Normal, lo que suele rechazarse en los datos. Por otra parte, si se utilizan muestras solapadas en la estimación del VaR a 10 días, podría pensarse en elevar el número de excepciones que define cada una de las zonas que antes describimos. Ello haría que las instituciones tuviesen que preocuparse de como aumentar el VaR diario para equipararlo a un VaR a 10 dias, o tratar de disponer de un buen modelo de estimacion directa del VaR a 10 dias. Asimismo, la muestra de 250 observaciones que los reguladores proponen para estimar el VaR resulta insuficiente, generando estimaciones numéricas poco precisas, lo que dificulta que se llegue a rechazar un modelo de VaR inadecuado.

En la aplicación práctica de backtests, los reguladores distinguen entre los $P\&L$ realizados, que se obtendrían mediante la negociación intradía, incorporando comisiones, reservas, y costes de financiación, y los $P\&L$ teóricos, no realizados, que se obtienen mediante el llamado *mark-to-market*, es decir, valorando la cartera a precios de mercado, sin incorporar los elementos que hemos mencionado.

10.2.4 Contrastes de modelos de VaR

Contraste de cobertura incondicional Los contrastes de cobertura incondicional, propuestos por Kupiec (1995) se basan asimismo en el número de excepciones. La hipótesis nula es que la función indicatriz que antes definimos sigue una distribución de Bernoulli con una probabilidad constante de observar el suceso (una excepción), igual al nivel de significación α del VaR. El estadístico de contraste es:

$$LR_{ci} = \frac{\pi_m^{n_1}(1-\pi_m)^{n_0}}{\hat{\pi}^{n_1}(1-\hat{\pi})^{n_0}}$$

donde n_1 es el número observado de excepciones, y $n_0 = n - n_1$, siendo n el tamaño de la muestra utilizada para la validación del VaR, $\hat{\pi}$ denota el porcentaje observado de excepciones $\hat{\pi} = n_1/n$, π_m es el porcentaje esperado de acuerdo con el modelo teórico: $\pi_m = \alpha$. En muestras grandes, $-2.ln(LR_{ci})$ se distribuye como una chi-cuadrado con 1 grado de libertad, con cuyos valores críticos habrá que comparar el valor numérico que hayamos obtenido para $-2.ln(LR_{ci})$.[EIV.6.5]

Contraste de independencia temporal Suele observarse que las excepciones no están distribuidas de modo aleatorio a lo largo de la muestra, sino que ocurren en mayor frecuencia en determinados intervalos de tiempo. Ello reflejaría que el modelo de VaR que se está utilizando no responde suficientemente a las circunstancias de mercado. En particular, esto sucedería si la estimación de la varianza a lo largo de la muestra no es suficientemente flexible.

Por tanto, un contraste de validación el VaR es el de la independencia temporal en la ocurrencia de las excepciones. Para ello, se tiene en cuenta que si las excepciones no se presentan de modo independiente, la probabilidad de que mañana se observe una excepción es más elevada si hoy se produjo una excepción, que si no se produjo. Si denotamos por n_{ij} el número de observaciones en que el indicador toma el valor i en un período, seguido del valor j el período siguiente. Por ejemplo, n_{10} sería el número de observaciones en que no se produjo excepción, habiéndose producido una excepción en el período anterior. Nótese que $n_1 = n_{11} + n_{01}, n_0 = n_{00} + n_{10}$. Definamos las proporciones:

$$\pi_{01} = \frac{n_{01}}{n_{00}+n_{01}} \; ; \; \pi_{11} = \frac{n_{11}}{n_{10}+n_{11}}$$

que miden la proporcion de excepciones, dado que el dato previo no lo fue, π_{01}, y la proporción de excepciones que han seguido a una excepción, π_{11}.

El estadístico para el contraste de independencia temporal en la ocurrencia de excepciones, propuesto por Christoffersen (998), es:

$$LR_{ind} = \frac{\hat{\pi}^{n_1}(1-\hat{\pi})^{n_0}}{\pi_{01}^{n_{01}}(1-\pi_{01})^{n_{00}}\pi_{11}^{n_{11}}(1-\pi_{11})^{n_{10}}}$$

y se tiene que $-2ln(LR_{ind})$ sigue una distribución chi-cuadrado con 1 grado de libertad.

Este contraste no detectará la dependencia temporal en la ocurrencia de excepciones si se producen rentabilidades "normales" entre excepciones. Esto es debido a que se basa en una cadena de Markov de primer orden, por lo que el contraste funciona bien si las excepciones viene realmente en dias consecutivos.[25]

Nótese que para la aplicación del contraste es preciso que se hayan observado algunos episodios de excepciones consecutivas, de modo que π_{11} no sea igual a cero.[EIV.6.6]

Contraste de cobertura condicional El siguiente estadístico puede utilizarse para contrastar la cobertura incondicional combinada con la independencia de las excepciones, y constituye un contraste de cobertura condicional:

$$LR_{cc} = \frac{\pi_m^{n_1}(1-\pi_m)^{n_0}}{\pi_{01}^{n_{01}}(1-\pi_{01})^{n_{00}}\pi_{11}^{n_{11}}(1-\pi_{11})^{n_{10}}}$$

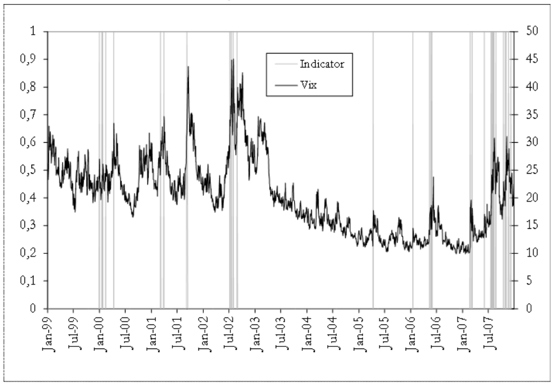
tenéndose que $-2\ln(LR_{cc})$ sigue una distribución chi-cuadrado con dos grados de libertad.

Es evidente que $LR_{cc} = LR_{ci}LR_{ind}$, por lo que $-2\ln(LR_{cc}) = -2\ln(LR_{ci}) - 2\ln(LR_{ind})$.

Contrastes basados en regresión Estos contrastes se basan en tratar de explicar la ocurrencia de excepciones por algun factor que era conocido en el momento de estimar el VaR para el periodo t, como la volatilidad del mercado, el nivel de los tipos de interés, la pendiente de la estructura temporal, u otros. En el caso de la estimación del VaR un período hacia adelante, se trataría de avriables conocidas en $t-1$. Si alguno de estos factores tuviese capacidad explicativa, rechazariamos la ocurrencia puramente aleatoria de las excepciones, lo que implicaría un rechazo del modelo de estimación del VaR. El contraste se basa en el estadístico F de capacidad explicativa global de la regresión. Una alternativa al modelo de regresión lineal es, en este caso, el uso de un modelo logit o probit para explicar la ocurrencia de una excepción.

[25] Para los interesados, esta limitación se debe a que el contraste de independencia se basa en un modelo de cadena de Markov de orden 1, cuando sería conveniente disponer de un contraste de orden superior.

Figure IV.6.9



En ocasiones se observa que la ocurrencia de excepciones está relacionada con la volatilidad de la cartera ese mismo día. Al no tratarse de la volatilidad del dia en que se estimó el VaR, ello no tiene ninguna implicación para el contraste del modelo VaR, pero sugiere que el modelo de volatilidad podria estar mal especificado. Esto puede suceder si, por ejemplo, se han utilizado ventanas móviles para estimar la volatilidad, que no tienen en cuenta los posibles agrupamientos de volatilidad, como hacen tanto el modelo EWMA de volatilidad como los modelos GARCH. [EIV.6.7]

Contrastes basados en la calidad de las predicciones de volatilidad (añadir otros) Cuando se utiliza el modelo lineal Normal, el contrate de validación del VaR se reduce a un contraste de validación de las predicciones de volatilidad, puesto que ambos son proporcionales. Si denotamos por $\hat{\sigma}_t$ la predicción, hecha en t, de la desviación típica de la rentabilidad obtenida entre t y $t+1$, la rentabilidad estandarizada es:

$$z_{t+1} = \frac{R_{t+1}}{\hat{\sigma}_t}$$

Si las rentabilidades son i., i.d. con desviación típica σ, y nuestro modelo predice σ adecuadamente, entonces la desviación típica de z_t debería ser igual a 1.

El *estadístico de sesgo* (*bias statistic*), b, es precisamente igual a la desviación típica muestral de z_t y se trata de contrastar $H_0 : b = 1$, frente a una alternativa unilateral o bilateral, dependiendo de que creamos que el modelo infraestima o sobreestima el VaR, o no tengamos de tal intuición.

Recordemos que el error estándar del estimador de la desviación típica de una variable aletoria es igual a dicha estimación numérica, multiplicada por $\sqrt{\frac{1}{2T}}$siendo T el tamaño de la muestra. Por tanto, como el valor numérico de b, que es una desviación típica será muy próxima a 1, su error estándar será aproximadamente igual a $\sqrt{\frac{1}{2T}}$. Para contrastar $H_0 : b = 1$, podriamos pensar en utilizar el estadístico $\sqrt{2T}\left(\hat{b}-1\right)$. Pero al no tener b una distribución

Normal,dicho estadístico no tendría una distribución $t - Student$.

Si suponemos que las rentabilidades estandarizadas son independientes y siguen una distribución $N(0,1)$, entonces Tb seguirá una distribución chi-cuadrado con T grados de libertad, por ser la suma de cuadrados de variables $N(0,1)$. Su media sería T y su varianza $2T$, por lo que, aplicando el teorema central del limite para muestras grandes,

$$\frac{Tb^2 - T}{\sqrt{2T}} \approx N(0,1)$$

es decir,

$$b^2 \approx N\left(1, \sqrt{\frac{2}{T}}\right)$$

y utilizando la expansión en serie de Taylor del mismo modo que hicimos para obtener los errores estándar del estimador de la desviación típica, tenemos:

$$b \approx N\left(1, \sqrt{\frac{1}{2T}}\right)$$

que será válida en muestras de backtesting grandes, trabajando con rentabilidades independientes, con distribución Normal.A partir de esta distribución, puede construirse un intervalo de confianza y examinar si la estimación de b cae dentro del mismo, no rechazando H_0, fuera del mismo por su izquierda (subestimación del VaR), o por su derecha (sobreestimación del VaR). [EIV.6.10]

Contrastes de la distribución prevista en el backtesting del VaR Este contraste examina si los valores numéricos observados para las rentabilidades futuras se corresponden con la distribución de probabilidad utilizada en el cálculo del VaR. Utiliza por tanto todas las rentabilidades observadas, y no sólo las que caen en las colas de la distribución. Si el backtesting se basa en datos no solapados, y denotamos por F_t la distribución prevista en t para las rentabilidades en $t+h$, tenemos:

$$p_{ht} = F_t(R_{t+h})$$

Si el modelo del VaR es correcto, entonces las probabilidades p_{ht} deberán seguir una distribución Uniforme $U(0,1)$, ya que deberian ser una secuencia de números puramente aleatorios. Dicho de otro modo, si el modelo de riesgo es correcto, no debería ser capaz de predecir la probabilidad de una rentabilidad futura. Si, por el contrario, el modelo subestima el riesgo en las colas de la distribución, tendremos un número relativamente alto de observaciones en las colas (con p_{ht} próximo a 0 o a 1), y tambien en el centro ($p_{ht} = 0,5$) por leptocurtosis. Nos alejariamos asi de la distribución uniforme de dichas probabilidades en el intervalo (0,1).

Para facilitar el contraste, transformamos las probabilidades p_{ht} en variables aleatorias Normales mediante:

$$Z_{ht} = \Phi^{-1}(p_{ht})$$

donde Φ denota la función de distribución de una variable $N(0,1)$. Queremos contrastar la hipótesis nula:

$$H_0 : Z_{ht} \sim i., i.d.\ N(0,1)$$

frente a la alternativa:

$$H_1 : Z_{ht} \sim i., i.d.\ N(\mu_h, \sigma_h^2),\ \mu_h \neq 0, \sigma_h \neq 1$$

El estadístico de razón de verosimilitudes:

$$-2\ln(LR) = -2\ln\left(\frac{L_0}{L_1}\right) \sim \chi_2$$

Es fácil ver que:

$$-2\ln(LR) = \sum_{t=1}^{T} z_{ht}^2 - \sum_{t=1}^{T}\left(\frac{z_{ht}-\hat{\mu}_h}{\hat{\sigma}_h}\right)^2 - T\ln\hat{\sigma}_h^2$$

donde $\hat{\mu}_h, \hat{\sigma}_h$ denotan la estimación de la media y desviación típica de Z_{ht}.

Una alternativa sería aplicar un contraste dinámico, que utilizase una alternativa:

$$H_1 : Z_{ht} \sim i., i.d.\ N(\mu, \sigma_t^2),\ \mu \neq 0, \sigma_t \neq 1$$

donde σ_t proviniese de algun modelo específico, como el suavizado exponencial (EWMA). [EIV.6.11]

Backtesting predicciones de la ETL McNeil y Frey (2000) sugieren utiizar los *residuos en exceso estandarizados*, definidos como:

$$\begin{aligned}
\varepsilon_{t+1} &= \frac{-y_{t+1} - ETL_{1,\alpha,t}}{\hat{\sigma}_t},\ si\ y_{t+1} < -VaR_{1,\alpha,t} \\
\varepsilon_{t+1} &= 0 \text{ en otro caso}
\end{aligned}$$

siendo $\hat{\sigma}_t$ la predicción de la desviación típica de la rentabilidad diaria (o de la $P\&L$) entre t y t+1. es decir, la predicción un periodo hacia adelante, hecha en t.

Si la estructura dinamica del proceso estámodelizada correctamente, y la ETL es una estimación insesgada de la esperanza en la cola a la izquierda del VaR, entonces los residuos en exceso estandarizados deberian conmprtarse como una muestra de observaciones independientes e identicamente distribuidas, con esperanza igual a cero. Por tanto, la hipótesis nula es que su media es cero, frente a la alternativa de que sea positiva, lo que sugeriria que hemos subestimado la

ETL, que es la eventualidad contra la cual nos queremos proteger. El estadistico de contraste es, por tanto:

$$t = \frac{\bar{\varepsilon}}{\hat{\sigma}_{\varepsilon}}$$

Este estadistico no tiene una distribución estándar, por lo que su distribución debes estimarse mediante bootstrapping. El problema es que, al utilizar unicamente las observaciones que exceden del VaR puede ser un numero reducido.

11 Apéndices

11.1 Apéndice 1: Distribuciones no estándar

La distribución generalizada de errores tiene por función de densidad:

$$f(x) = \frac{v}{\lambda 2^{(1+1/v)} \Gamma(1/v)} e^{-\left(0,5\left|\frac{x}{\lambda}\right|^{v}\right)},\ \lambda = \sqrt{2^{-(2/v)} \frac{\Gamma\left(\frac{1}{v}\right)}{\Gamma\left(\frac{3}{v}\right)}}$$

donde v es una parámetro que controla el perfil de la función de densidad. Incluye la densidad Normal cuando $v = 2$, y tiene colas más gruesas que la Normal para v menor que 2. El parámetro λ asegura que la varianza sea igual a 1.

Distribuciones truncadas

La función de densidad truncada que resulta al considerar los valores por debajo de un umbral q en una función de densidad $f(x)$ es:

$$g(x) = \frac{f(x)}{\int_{-\infty}^{q} f(x)dx},\ x < q$$

con esperanza matemática:

$$E\left[g(x)\right] = \frac{\int_{-\infty}^{q} x f(x)}{\int_{-\infty}^{q} f(x)dx}$$

que mide el valor esperado de la variable X cuando $X < q$.

Su varianza es:

$$Var\left[g(x)\right] = \frac{\int_{-\infty}^{q} x^2 f(x)}{\int_{-\infty}^{q} f(x)dx} - \left(E\left[g(x)\right]\right)^2$$

En el caso de una variable $\varepsilon \sim$Normal(0,1) tenemos:

$$E(\varepsilon \mid \varepsilon < -\alpha) = \frac{-\phi(\alpha)}{\Phi(-\alpha)}$$

Por ejemplo, para $\alpha = 0$, tenemos:

$$E(\varepsilon \mid \varepsilon < 0) = \frac{-\frac{1}{\sqrt{2\pi}} e^{0}}{0,5} = -\sqrt{\frac{2}{\pi}}$$

11.2 Apéndice 2: Valoración de opciones en presencia de asimetría y curtosis. El modelo Gram-Charlier.

El precio de una opción call debe ser igual al valor esperado y descontado de su pago a vencimiento, donde la expectativa se calcula de acuerdo con la distribución libre de riesgo:

$$c = e^{-rT} E_t^* \left[Max(S_{t+T} - X, 0) \right]$$

El modelo de Black Scholes Merton supone que las rentabilidades diarias del activo subyacente se distribuyen independientemente en el tiempo, de acuerdo con una distribución Normal con esperanza y varianza constantes, $N(\mu, \sigma^2)$. En tal caso, la rentabilidad sobre el horizonte a vencimiento de la opción seguirá una distribución $N(T\mu, T\sigma^2)$, y el precio del activo subyacente al vencimiento de la opción será: $S_{t+T} = S_t e^{R_{t+1,t+T}}$, siendo $R_{t+1,t+T}$ el tipo libre de riesgo para una inversión entre $t+1$ y $t+T$.

Esto conduce a:

$$\begin{aligned} c &= e^{-rT} \int_{-\infty}^{\infty} Max(S_t e^{x^*} - X, 0) f(x^*) dx^* = \\ &= e^{-rT} \int_{\ln(X/S_t)}^{\infty} S_t e^{x^*} f(x^*) dx^* - \int_{\ln(X/S_t)}^{\infty} X f(x^*) dx^* \end{aligned}$$

donde x^*denota la variable riesgo-neutro correspondiente a la rentabilidad del activo subyacente entre t y $t+T$. La integral anterior resulta:

$$c_{BSM} = e^{-rT} \left[S_t e^{rT} \Phi(d) - X\Phi(d - \sigma\sqrt{T}) \right] = S_t \Phi(d) - Xe^{-rT} \Phi(d - \sigma\sqrt{T})$$

donde Φ denota la función de distribución de la variable $N(0,1)$, y $d = \frac{\ln(S_t/X) + T(r+\sigma^2/2)}{\sigma\sqrt{T}}$.

La paridad put-call es una relación de ausencia de arbitraje, que no precisa de ningún modelo de valoración:

$$S_t + p = c + Xe^{-rT}$$

y, junto con la expresión anterior para el precio de la opción call, conduce al precio de la opción put:

$$c_{BSM} = c + Xe^{-rT} - S_t = Xe^{-rT} \Phi(\sigma\sqrt{T} - d) - S_t \Phi(-d)$$

En el caso en que el activo reparte una tasa de dividiendos (u otro tipo de rentas) constante, anual, igual a q, la expresión para d es: $d = \frac{\ln(S_t/X) + T(r-q+\sigma^2/2)}{\sigma\sqrt{T}}$, puesto que el inversor que tiene la opción en su cartera recibe al vencimiento de la opción tan sólo el activo subyacente, pero no la renta que su posesión ha generado desde que se compró la opción.

En consecuencia, de acuerdo con el modelo BS, el precio de una opción call es una función: $c_{BS} = c(S_t, r, X, T, q; \sigma)$ y, si disponemos de una muestra de n opciones negociadas un determinado día sobre un mismo activo subyacente, la volatilidad de dicho subyacente puede estimarse mediante el problema:

$$\underset{\sigma}{Min}\ MSE_{BSM} = \underset{\sigma}{Min} \frac{1}{n} \sum_{i=1}^{n} \left(c_i^{mkt} - c_{BSM}(S_t, r, X_i, T_i, q; \sigma) \right)^2$$

La volatilidad implícita se define:

$$\sigma_{BSM}^{vi} = c_{BSM}^{-1}(S_t, r, X, T, q; \sigma)$$

que puede calcularse para cada opción por separado. De acuerdo con el modelo BSM, dicha volatilidad debería ser única para cada activo subyacente, con independencia del vencimiento del opción considerada, o de su precio de ejercicio. Sin embargo, se observa que esto no es así, apareciendo sonrisas o muecas de volatilidad. En el primer caso, la curva de volatilidad sobre el grado de Moneyness describe una curva cóncava, indicando la infravaloración de las opciones muy out-of-the-money por parte del modelo BSM, debido a un exceso de curtosis en la distribución de rentabilidades del activo subyacente. La mueca (smirk) refleja una infravaloración de una cola del mercado por parte del modelo BSM, habitualmente la formada por las opciones muy in-the-money. Estadísticamente, se debe a una asimetría en la distribución de rentabilidades del activo subyacente. En consecuencia, las opciones put muy out-of-the-money están asimismo infravaloradas por el modelo BSM.

Consideremos ahora la existencia de asimetría y curtosis en la distribución de rentabilidades del activo subyacente. Por tanto, si definimos la rentabilidad estandarizada:

$$\omega_T = \frac{R_{t+1,t+T} - T\mu}{\sqrt{T}\sigma}$$

tenemos: $R_{t+1,t+T} = T\mu + \sqrt{T}\sigma\omega_T$.

Si suponemos ahora que las rentabilidades estandarizadas siguen una distribución que admite la expansión de Gram-Charlier, tenemos:

$$f(\omega_T) = \phi(\omega_T) - \zeta_{1T} \frac{1}{3!} D^3 \phi(\omega_T) + \zeta_{2T} \frac{1}{4!} D^4 \phi(\omega_T)$$

donde $\phi(.)$ denota la función de densidad de la $N(0,1)$,y D^i denota el operador derivada. Es sencillo ver que los coeficientes de asimetría y curtosis de la rentabilidad sobre un período de longitud T se relacionan con los coeficientes de las rentabilidades diarias mediante: $\zeta_{1T} = \zeta_{11}/\sqrt{T}, \zeta_{2T} = \zeta_{21}/T$.

Tenemos, por tanto:

$$\begin{aligned} D^1\phi(\omega) &= -\omega\phi(\omega);\ D^2\phi(\omega) = (\omega^2 - 1)\phi(\omega); \\ D^3\phi(\omega) &= -(\omega^3 - 3\omega)\phi(\omega);\ D^4\phi(\omega) = (\omega^4 - 6\omega^2 + 3)\phi(\omega); \end{aligned}$$

La función de densidad Gram-Charlier $f(\omega_T)$ es una expansión alrededor de la función de densidad Normal, que permite asimetría y curtosis no nulas, pero que se reduce a la densidad $N(0,1)$ cuando el coeficiente de asimetría y el exceso de curtosis son ambos cero. La expansión de Cornish-Fisher, por el contrario, se aplica a la inversa de la función de distribución de una variable aleatoria.

Para poner precio a opciones Europeas, partimos de nuevo de la fórmula libre de riesgo de valoración de una opción:

$$c = e^{-rT} E_t^* \left[Max(S_{t+T} - X, 0) \right]$$

por lo que debemos resolver:

$$c = e^{-rT} \int_{\ln(X/S_t)}^{\infty} (S_t e^{x^*} - X) f(x^*) dx^*$$

Antes trabajamos con la distribución Normal con esperanza r y varianza σ^2 diariamente. Ahora, en cambio, definimos la rentabilidad estandarizada a horizonte T :

$$\omega_T = \frac{x^* - rT}{\sqrt{T}\sigma}$$

y suponemos que sigue una distribución Gram-Charlier.

Bajo tal supuesto, el precio de la opción call es aproximadamente igual a:

$$\begin{aligned} c_{GC} &\approx S_t \Phi(d) - X e^{-rT} \Phi(d - \sigma\sqrt{T}) + S_t \phi(d) \sqrt{T} \sigma \left[\frac{\zeta_{1T}}{3!} (2\sqrt{T}\sigma - d) - \frac{\zeta_{2T}}{4!} (1 - d^2 + 3d\sqrt{T}\sigma - 3T\sigma^2) \right] = \\ &= S_t \Phi(d) - X e^{-rT} \Phi(d - \sigma\sqrt{T}) + S_t \phi(d) \sigma \left[\frac{\zeta_1}{3!} (2\sqrt{T}\sigma - d) - \frac{\zeta_{21}}{4!} (1 - d^2 + 3d\sqrt{T}\sigma - 3T\sigma^2) \right] \end{aligned}$$

La expresión es aproximada porque hemos prescindido de los términos en σ^3y σ^4, lo que nos permite mantener la misma definición para el parámetro d que en el modelo BSM. De este modo, el modelo Gram-Charlier (GC) es una extensión del modelo BSM para el caso en que hay asimetría y curtosis. La existencia de dividendos o rentas puede ser tenida en cuenta del modo habitual.

El modelo GC tiene tres parámetros desconocidos: $\sigma, \zeta_{11}, \zeta_{21.}$Pueden estimarse por un procedimiento numérico resolviendo el problema de optimización:

$$\underset{\sigma,\zeta_{11},\zeta_{21}}{Min} \; MSE_{GC} = \underset{\sigma,\zeta_{11},\zeta_{21}}{Min} \frac{1}{n} \sum_{i=1}^{n} \left[c_i^{mkt} - c_{GC}(S_t, r, X_i, T_i; \sigma, \zeta_{11}, \zeta_{21}) \right]^2$$

mientras que la volatilidad implícita σ_{GC}^{vi} puede calcularse para cada opción mediante:

$$\sigma_{GC}^{vi} = c_{BSM}^{-1}(S_t, r, X, T; c_{GC})$$

de modo que, una vez que se dispone de valores numéricos para los parámetros $\sigma, \zeta_{11}, \zeta_{12}$, se lleva el precio teórico generado por el modelo GC a la fórmula de valoración del modelo BSM, y se invierte para encontrar así la volatilidad implícita.

Puede utilizarse asimismo la expresión aproximada:

$$\sigma_{GC}^{vi} = c_{BSM}^{-1}(S_t, r, X, T; c_{GC}) \approx \sigma\sqrt{T}\left(1 - \frac{\zeta_{11}/\sqrt{T}}{3!}d - \frac{\zeta_{21}/T}{4!}(1-d^2)\right)$$

que se reduce a la expresión habitual en ausencia de asimetría y curtosis. El modelo CG proporciona una formula de valoración cerrada, en un contexto de asimetria y curtosis, que permite recoger las pautas sistemáticas de volatilidad que se observan en los mercados.

11.3 Apéndice 3: El modelo GARCH de valoración de opciones

El modelo de Gram-Charlier para valorar opciones es capaz de recoger la asimetría y curtosis en volatilidad, pero tiene la desventaja de que supone que ésta es constante en el tiempo, contrariamente a la robusta observación empírica al contrario en todos los mercados. Puede decirse, que mientras que el modelo GC captura la estructura de precios de las opciones a través de los precios de ejercicio, sin embargo no recoge la estructura existente a lo largo de los vencimientos. En esta sección consideramos la formación de precios de opciones cuando la rentabilidad esperada del subyacente sigue un proceso GARCH. La diferencia estriba en que bajo volatilidad constante, la estructura temporal de volatilidades es constante, ya que la varianza de la rentabilidad a un horizonte de T periodos es igual a $T\sigma^2$, siento σ^2 la varianza de la rentabilidad sobre un período.

Suponemos que el proceso GARCH especifica que la rentabilidad esperada es igual a la tasa libre de riesgo, r, más una prima por riesgo de volatilidad λ, así como un término de normalización. Por otro lado, se supone que la rentabilidad observada cada período es igual a la rentabilidad esperada r, más una prima por el riesgo de volatilidad, $\lambda\sigma_{t+1}$, un término de normalización, $-\frac{1}{2}\sigma_{t+1}^2$, más una innovación. Se supone que dicha innovación sigue una distribución condicional $N(0, \sigma_t^2)$,donde σ_t^2 evoluciona de acuerdo con un proceso GARCH(1,1) con apalancamiento, lo que crea asimetría en la distribución de rentabilidades, lo cual es importante para explicar la asimetría observada en los precios de las opciones:

$$\begin{aligned} R_{t+1} &= \ln S_{t+1} - \ln S_t = r + \lambda\sigma_{t+1} - \frac{1}{2}\sigma_{t+1}^2 + \sigma_{t+1}z_{t+1},\ z_{t+1}/\Omega_t \sim N(0,1) \\ \sigma_{t+1}^2 &= \omega + \alpha(\sigma_t z_t - \theta\sigma_t)^2 + \beta\sigma_t^2 \end{aligned}$$

que implican una esperanza y varianza condicional para las rentabilidades:

$$\begin{aligned} E_t R_{t+1} &= r + \lambda\sigma_{t+1} - \frac{1}{2}\sigma_{t+1}^2; \\ V_t R_{t+1} &= \sigma_{t+1}^2 \end{aligned}$$

Utilizando la conocida propiedad: $x \sim N(\mu, \sigma^2) \Rightarrow E(e^x) = e^{\mu+\sigma^2/2}$, tenemos:

$$\begin{aligned} E_t(S_{t+1}/S_t) &= E_t(R_t) = E_t\left[e^{r+\lambda\sigma_{t+1}-\frac{1}{2}\sigma_{t+1}^2+\sigma_{t+1}z_{t+1}}\right] = e^{r+\lambda\sigma_{t+1}-\frac{1}{2}\sigma_{t+1}^2}E_t\left[e^{\sigma_{t+1}z_{t+1}}\right] = \\ &= e^{r+\lambda\sigma_{t+1}-\frac{1}{2}\sigma_{t+1}^2}e^{\frac{1}{2}\sigma_{t+1}^2} = e^{r+\lambda\sigma_{t+1}} \end{aligned}$$

que muestra el papel que juega el parámetro λ como precio del riesgo de volatilidad.

Si partimos nuevamente de la expresión genérica para el precio de una opción call:

$$c = e^{-rT}E_t^*\left[Max(S_{t+T} - X, 0)\right]$$

Bajo neutralidad al riesgo, debemos tener una rentabilidad esperada igual a la tasa libre de riesgo, y una volatilidad esperada igual a la del proceso original:

$$\begin{aligned} E_t^*(S_{t+1}/S_t) &= r \\ V_t^*(R_{t+1}) &= \sigma_{t+1}^2 \end{aligned}$$

Consideremos ahora el proceso:

$$\begin{aligned} R_{t+1} &= \ln S_{t+1} - \ln S_t = r - \frac{1}{2}\sigma_{t+1}^2 + \sigma_{t+1}z_{t+1}^*, \qquad (30) \\ z_{t+1}^*/\Omega_t &\sim N(0,1) \qquad (31) \\ \sigma_{t+1}^2 &= \omega + \alpha(\sigma_t z_t^* - \lambda\sigma_t - \theta\sigma_t)^2 + \beta\sigma_t^2 \end{aligned}$$

cuya esperanza condicional, bajo la distribución de probabilidad libre de riesgo es: $E_t^*(S_{t+1}/S_t) = r$, y cuya varianza condicional bajo esa misma distribución es:

$$\begin{aligned} V_t^*(R_{t+1}) &= E_t^*\sigma_{t+1}^2 = E_t^*\left[\omega + \alpha(\sigma_t z_t^* - \lambda\sigma_t - \theta\sigma_t)^2 + \beta\sigma_t^2\right] = \\ [\text{Por (30)}] &= E_t\left[\omega + \alpha\left(R_t - r - \frac{1}{2}\sigma_t^2 - \lambda\sigma_t - \theta\sigma_t\right)^2 + \beta\sigma_t^2\right] = \\ &= E_t\left[\omega + \alpha(\sigma_t z_t - \theta\sigma_t)^2 + \beta\sigma_t^2\right] = E_t\sigma_{t+1}^2 = \sigma_{t+1}^2 \end{aligned}$$

Por tanto, (30) satisface las dos condiciones que debe satisfacer un proceso libre de riesgo.

La ventaja de este modelo es su flexibilidad, pudiendo ser adaptado a cualquiera de las especificaciones GARCH vistas. Además, ajusta los precios de las opciones con bastante aproximación. La limitación es que no existe una fórmula cerrada para el precio de las opciones,que deben valorarse mediante simulación.

Para ello notemos que podemos eliminar un parametro mediante la especificación:

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

donde $\lambda^* = \lambda + \theta$.

Para llevar a cabo las simulaciones con objeto de valorar una opción, a partir de una observación para σ_{t+1}^2, generamos N observaciones $N(0,1)$ para z_{t+1}^*/Ω_t. Como queremos calcular la esperanza matemática E_t^* utilizando el proceso estocástico libre de riesgo, calculamos ahora la rentabilidad y varianza riesgo-neutro en el período $t+s$ para la simulación j-ésima mediante:

$$\begin{aligned} R_{j,t+s}^* &= r - \frac{1}{2}\sigma_{j,t+s}^2 + \sigma_{j,t+s} z_{j,t+s}^*, \ j = 1,2,...N \\ \sigma_{j,t+s+1}^2 &= \omega + \alpha(\sigma_{j,t+s} z_{j,t+s}^* - \lambda^*\sigma_{j,t+s})^2 + \beta\sigma_{j,t+s}^2, \ s = 1,2,... \end{aligned}$$

Repitiendo el ejercicio de simulación, obtenemos N realizaciones para el horizonte deseado. El precio hipotético del activo a vencimiento bajo la distribución riesgo-neutro puede calcularse, para cada realización:

$$S_{j,t+T}^* = S_t e^{\sum_{s=1}^T R_{j,t+s}^*}, \ j = 1,2,...,N$$

y el precio de la opción se calcula mediante el promedio descontando los pagos hipotéticos a vencimiento:

$$c_{GH} \approx e^{-rT}\frac{1}{N}\sum_{j=1}^N Max\left(S_{j,t+T}^* - X, 0\right)$$

que converge a la esperanza matemática según aumenta el número de simulaciones. $N = 5000$ debería ser suficiente para proporcionar una aproximación suficientemente buena en la mayoría de los casos.

Los parámetros del modelo GARCH deben estimarse previamente, lo que puede hacerse mediante el procedimiento de Máxima Verosimilitud. Alternativamente, si la muestra de opciones disponible para un determinado día es suficientemente amplia, podemos estimar resolviendo el problema de optimización:

$$\underset{\sigma_{t+1}^2,\omega,\alpha,\beta,\lambda^*}{Min} MSE_{GH} = \underset{\sigma_{t+1}^2,\omega,\alpha,\beta,\lambda^*}{Min} \frac{1}{n}\sum_{i=1}^n \left[c_i^{mkt} - c_{GH}(S_t, r, X_i, T_i; \sigma_{t+1}^2, \omega, \alpha, \beta, \lambda^*)\right]^2$$

donde estamos tratando σ_{t+1}^2 como un parámetro desconocido. Debe tenerse en cuenta, sin embargo, que según el algoritmo numérico va buscando en el espacio paramétrico un vector de valores numéricos para $\sigma_{t+1}^2, \omega, \alpha, \beta, \lambda^*$, hay que proceder a la valoración de las opciones mediante simulación, por lo que se trata de un procedimiento bastante exigente desde el punto de vista computacional. Por otra parte, este procedimiento permitiría analizar la variabilidad temporal de los valores numéricos de los parametros del modelo, $\sigma_{t+1}^2, \omega, \alpha, \beta, \lambda^*$.

Existe una especificación GARCH algo más particular que la anterior, que genera una fórmula cerrada para el precio de la opción:

$$
\begin{aligned}
R_{t+1} &= \ln S_{t+1} - \ln S_t = r + \lambda \sigma_{t+1}^2 + \sigma_{t+1} z_{t+1},\ z_{t+1}/\Omega_t \sim N(0,1) \\
\sigma_{t+1}^2 &= \omega + \alpha (z_t - \theta \sigma_t)^2 + \beta \sigma_t^2
\end{aligned}
$$

La persistencia de la varianza en este modelo viene dada por $\alpha\theta^2 + \beta$, y la varianza incondicional es $\frac{\omega+\alpha}{1-\alpha\theta^2-\beta}$.

La versión riesgo-neutro de este proceso es:

$$
\begin{aligned}
R_{t+1} &= \ln S_{t+1} - \ln S_t = r - \frac{1}{2}\sigma_{t+1}^2 + \sigma_{t+1} z_{t+1}^*,\ z_{t+1}^*/\Omega_t \sim N(0,1) \\
\sigma_{t+1}^2 &= \omega + \alpha (z_t^* - \theta^* \sigma_t)^2 + \beta \sigma_t^2
\end{aligned}
$$

siendo sencillo ver que:

$$
\begin{aligned}
E_t^*(S_{t+1}/S_t) &= r \\
V_t^*(R_{t+1}) &= \sigma_{t+1}^2
\end{aligned}
$$

Bajo este proceso GARCH, el precio de una opción call europea es:

$$c_{CFG} = S_t P_1 - X e^{-rT} P_2$$

con:

$$P_1 = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re}\left(\frac{X^{-i\varphi} f^*(i\varphi+1)}{i\varphi f^*(1)} \right) d\varphi;\ P_2 = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re}\left(\frac{X^{i\varphi} f^*(i\varphi)}{i\varphi} \right) d\varphi;$$

donde la función $f(.)$ está definida por:

$$f(\varphi) = S_t^{\varphi} e^{A_{t,t+T}(\varphi) + B_{t,t+T}(\varphi)\sigma_{t+1}^2}$$

con expresiones recursivas:

$$
\begin{aligned}
A_{t,t+T}(\varphi) &= A_{t+1,t+T}(\varphi) + \varphi r + B_{t+1,t+T}(\varphi)\omega - \frac{1}{2}\ln\left(1 - 2\alpha B_{t+1,t+T}(\varphi)\right) \\
A_{t,t+T}(\varphi) &= \varphi(\lambda+\theta) - \frac{1}{2}\theta^2 + \beta B_{t+1,t+T}(\varphi)\omega - \frac{1}{2}\frac{(\varphi-\theta)^2/2}{1 - 2\alpha B_{t+1,t+T}(\varphi)}
\end{aligned}
$$

que pueden resolverse a partir de condiciones terminales:

$$A_{t+T,t+T}(\varphi) = 0;\ B_{t+T,t+T}(\varphi) = 0$$

11.4 Apéndice 4: Teoría de valores extremos (versión 2)

Consideremos una serie de rentabilidades diarias de una cartera: $\{r_1, r_2, ..., r_n\}$, cuyos estadísticos de orden extremos son $r_{(1)}$ y $r_{(n)}$: $r_{(1)} = \min_{1\leq j\leq n}\{r_j\}$, y $r_{(n)} = \max_{1\leq j\leq n}\{r_j\}$. Nos vamos a centrar en las propiedades del mínimo, que son las relevantes para el cálculo del VaR de una posición larga. Sin embargo, la misma teoría es válida para el cálculo de la rentabilidad máxima de la cartera, mediante un cambio de signo:

$$r_{(n)} = -\min_{1\leq j\leq n}\{-r_j\} = -r^c_{(1)}$$

donde $r^c_t = -r_t$.

Supongamos que las rentabilidades son incorrelacionadas e igualmente distribuidas, de acuerdo con $F(x)$, y con un rango $[l, u]$, .donde los extremos pueden ser finitos o no. La función de distribución de $r_{(1)}$, $F_{n,1}(x)$, es:

$$F_{n,1}(x) = 1 - [1 - F(x)]^n$$

que tiende a ser degenerada según $n \to \infty : F_{n,1}(x) \to 0$ si $x \leq l$, y $F_{n,1}(x) \to 1$ si $x > l$.

La Teoría de Valores Extremos se refiere a la posible existencia de sucesiones $\{\alpha_n\}$, (factores de escala) $\{\beta_n\}$, (parámetros de localización), con $\alpha_n > 0$, tales que las distribución de:

$$r_{(1*)} \equiv \frac{r_{(1)} - \beta_n}{\alpha_n}$$

converja a una distribución no degerada cuando $n \to \infty$.

La Teoría de Valores Extremos tiene dos implicaciones importantes:

- la distribución límite del mínimo normalizado, $F_*(x)$, está caracterizada por el comportamiento en las colas de la distribución $F(x)$ de r_t, no por la distribución específica de las rentabilidades, por lo que es aplicable a una gama amplia de distribuciones de rentabilidades. Sin embargo, las sucesiones $\{\alpha_n\}$ y $\{\beta_n\}$ dependerán de la distribución concreta de rentabilidades,

- el índice de cola k, o el parámetro de perfil, no depende del intervalo temporal considerado para las rentabilidades, lo que resulta útil en el cálculo del VaR.

11.4.1 Estimación del modelo

Los parámetros del modelo: k, escala, α_n, perfil, β_n, localización, puedes estimarse por métodos paramétricos (Máxima Verosimilitud o regresión) o por métodos no paramétricos.

Máxima verosimilitud

Método de Regresión

Método no paramétrico