Modelos de Valoración de Activos. Estimación y Contraste

Mª Victoria Esteban González

Departamento de Economía Aplicada III. Econometría y Estadística
Facultad de Ciencias Económicas y Empresariales
Universidad del País Vasco/Euskal Herriko Unibertsitatea

Contenido

Tema 1

Modelo de Valoración de Activos con Cartera de Mercado

1.1. Introducción

Bajo el supuesto de imposibilidad de arbitraje sabemos que todos los activos que prometen ofrecer los mismos pagos futuros deben tener hoy el mismo precio. Debe satisfacerse la ecuación fundamental de valoración, que en forma general se escribe:

$$E_t(M_{t+1}\tilde{R}_{jt+1}) = 1 \qquad j = 1, \ldots, N \tag{1.1}$$

y donde los supuestos alternativos sobre la variable agregada (o factor de descuento) M_{t+1} permiten obtener modelos diferentes de valoración como el CAPM o el APT. Lo que nos interesa en este punto es la forma, que tienen estos dos modelos, de explicar la compensación por el riesgo soportado por los agentes económicos al tomar sus decisiones de cartera. El modelo general de la expresión (1.1) puede resumirse como:

$$E(R_j) = \text{tipo de interés libre de riesgo} + \text{compensación por riesgo} \tag{1.2}$$

siendo la forma explícita que adopta la compensación por riesgo lo que distingue a los diferentes modelos de valoración. La importancia de un modelo u otro dependerá, por tanto, de su habilidad para predecir los rendimientos observados en los mercados, lo más ajustadamente posible. Nos centraremos en estimar y contrastar el CAPM. El mensaje más importante del CAPM es que existe una relación lineal y positiva entre el rendimiento esperado de cualquier activo financiero j y su riesgo beta. Dependiendo de la existencia o no de un activo libre de riesgo en la economía estaremos en una versión u otra y por tanto, en una modelización u otra con sus correspondientes implicaciones.

1.2. CAPM, versiones e implicaciones

1.2.1. CAPM de Sharpe-Lintner

A partir de un trabajo de Markowitz (1959), Sharpe (1964) y Lintner (1965b) muestran que si los inversores tienen expectativas homogéneas y construyen carteras eficientes en el sentido media-varianza, en ausencia de fricciones, la cartera de mercado es una cartera eficiente en el sentido media-varianza. La versión del CAPM de Sharpe-Lintner supone la existencia de un activo libre de riesgo al que prestar y tomar prestado. Bajo esta versión el rendimiento esperado de un activo j, $E(R_j)$ se especifica como[1]:

$$E(R_j) = r + \beta_{jm}\left[E(R_m) - r\right] \qquad j = 1, \ldots, N \tag{1.3}$$

$$\beta_{jm} = \frac{Cov(R_j, R_m)}{Var(R_m)} \tag{1.4}$$

donde:

- r es el tipo de interés libre de riesgo.
- $[E(R_m) - r]$ es la prima de riesgo de la cartera de mercado compuesta por todos los activos de la economía ponderados según su valor de mercado.

Podemos escribir la versión de Sharpe-Lintner de forma más compacta expresándola en términos de rendimientos en exceso del tipo de interés libre de riesgo o *exceso de rendimiento* como:

$$E(r_j) = \beta_{jm} E(r_m) \tag{1.5}$$

$$\beta_{jm} = \frac{Cov(r_j, r_m)}{Var(r_m)} \tag{1.6}$$

donde

$$r_j \equiv R_j - r; \quad r_m \equiv R_m - r$$

Notar que las ecuaciones (1.4) y (1.6) coinciden dado que tratamos al tipo de interés libre de riesgo como una variable no estocástica. En las aplicaciones empíricas el tipo de interés libre de riesgo es estocástico y por tanto, las betas pueden diferir. En las aplicaciones empíricas es habitual usar rendimientos en exceso de un activo seguro y por tanto usar la ecuación (1.5).

A partir de la ecuación (1.5) obtenemos las siguientes implicaciones contrastables:

a) El valor esperado del exceso de rendimiento de cada activo j es proporcional a su beta. Por tanto el término independiente del correspondiente modelo econométrico debe ser cero.

b) Las betas capturan completamente la variación de sección cruzada de los excesos de rendimiento esperados.

c) La prima por riesgo del mercado es positiva $E(r_m) > 0$.

[1] Por el momento trabajaremos con momentos incondicionales.

1.2.2. CAPM de Black

Black (1972) deriva una versión más general ya que no supone la existencia del activo seguro, es decir, no exige que existan oportunidades de pedir prestado y prestar a la tasa libre de riesgo. Admitiendo ventas al descubierto, existe una cartera que tiene beta igual a cero respecto a la cartera de mercado que es, a su vez, una cartera eficiente en el sentido media-varianza:

$$E(R_j) = E(R_{0m}) + \beta_{jm} \left[E(R_m) - E(R_{0m}) \right] \qquad j = 1, \ldots, N \tag{1.7}$$

$$\beta_{jm} = \frac{Cov(R_j, R_m)}{Var(R_m)} \tag{1.8}$$

donde:

- $E(R_m)$ es el rendimiento esperado en la cartera de mercado.
- $E(R_{0m})$ es el rendimiento esperado de la cartera con beta igual a cero respecto de la cartera de mercado, m. Esta cartera se define como la cartera que tiene la varianza más pequeña dentro de la clase de carteras incorrelacionadas con la cartera de mercado, m. Puede haber otra cartera incorrelacionada con m y con el mismo rendimiento esperado, pero tendrá varianza mayor.
- $[E(R_m) - E(R_{0m})]$ es la prima por riesgo del mercado cuando no existe activo seguro.

Dado que en el modelo de Black la riqueza en téminos reales es relevante, en la ecuación (1.8), β_{jm} se define en términos de rendimientos reales.
Las implicaciones de este modelo son:

a) La relación establecida está en términos de rendimientos reales.

b) Las betas capturan completamente la variación de sección cruzada de los rendimientos.

El análisis econométrico de la versión de Black del CAPM trata el rendimiento de la cartera cero-beta como una cantidad no observable, haciendo el análisis más complicado que en la versión de Sharpe-Lintner.

La versión de Black puede contrastarse como una restricción en el modelo de mercado con rendimientos reales. El modelo de mercado para rendimientos reales, en términos esperados es [2]:

$$E(R_j) = \alpha_{jm} + \beta_{jm} E(R_m) \tag{1.9}$$

[2] El subíndice m en los parámetros α y β sirve para enfatizar que la relación a estudiar es la existente entre el rendimiento esperado de cada activo y el rendimiento esperado del mercado.

y *la implicación* de la versión de Black sería:

$$\alpha_{jm} = E(R_{0m})(1 - \beta_{jm}) \quad \forall i$$

Es decir, el modelo de Black restringe al término constante del modelo de mercado con rendimientos reales a ser igual al producto entre el rendimiento esperado de la cartera cero-beta y la diferencia entre la unidad y la beta del activo. Más adelante volveremos sobre esta relación.

Trabajando en (1.7) como en (1.3) podemos escribir el CAPM en su forma más general posible en términos de *exceso de rendimiento* sobre el tipo de interés libre de riesgo. Obtenemos:

$$E(R_j) - r = [E(R_{0m}) - r] + \beta_{jm}\left[[E(R_m) - r] - [E(R_{0m}) - r]\right] \quad j = 1, \ldots, N \quad (1.10)$$

Así en su versión más general el CAPM puede escribirse como:

$$E(r_j) = E(r_{0m}) + \beta_{jm}\left[E(r_m) - E(r_{0m})\right] \qquad j = 1, \ldots, N \qquad (1.11)$$

En este caso las implicaciones del CAPM para los rendimientos de los activos son:

$$[E(r_m) - E(r_{0m})] > 0 \quad \text{y} \quad E(r_{0m}) \geq 0$$

Resumiendo, las implicaciones del CAPM para los rendimientos de los activos que vamos a contrastar a partir de la ecuación (1.11) son:

1. Para el modelo de Sharpe y Lintner en el que existen ilimitadas oportunidades de prestar y pedir prestado a la tasa libre de riesgo, las implicaciones a contrastar son:

   $$E(r_{0m}) = 0 \qquad \text{y} \qquad E(r_m) > 0$$

   de forma que las primas de riesgo de los activos sean proporcionales a sus betas, es decir, $E(r_j) = E(r_m)\beta_{jm}$.

2. Para la versión cero-beta de Black donde se restringe la posibilidad de pedir prestado a la tasa libre de riesgo las implicaciones del modelo a contrastar son menos precisas:

   $$E(r_{0m}) \geq 0 \qquad \text{y} \qquad [E(r_m) - E(r_{0m})] > 0$$

   Para éste caso el modelo viene dado por la expresión (1.7).

En definitiva, el modelo implica una relación *lineal y positiva* entre el rendimiento esperado y riesgo beta de cada activo.

forma más habitual de denotar el cero-beta CAPM de la ecuación (1.7) es:

$$E(R_j) = \gamma_0 + \gamma_1 \beta_{jm} \qquad j = 1, \ldots N \tag{1.12}$$

donde γ_0 y γ_1 son el rendimiento esperado de la cartera cero-beta respecto al mercado y la prima por riesgo del mercado respectivamente. Notar que dado que el CAPM se cumple periodo a periodo, esto implica que en t, γ_0 y γ_1 son comunes a todos los activos.

Nuestro cometido es estimar y contrastar las relaciones (1.3) y (1.7). Estimar la prima por riesgo del mercado y las betas de los activos y contrastar si el CAPM es un buen modelo. Para ello necesitamos proponer estimadores consistentes de los parámetros del modelo y diseñar un contraste para verificar estadísticamente si se cumplen o no las implicaciones derivadas de ambas relaciones. Veamos en primer lugar los problemas que aparecen a simple vista:

a) Las relaciones se han establecido en términos de rendimientos esperados que no son observados en la práctica. Por ello debemos justificar la realización de los contrastes en términos de realizaciones de rendimientos (rendimientos observados) de un modelo con expectativas. En los contrastes que se presentarán existe, de hecho, un supuesto sobre la racionalidad de las expectativas. Si suponemos que éstas en promedio son correctas, sobre periodos de tiempo suficientemente largos, podemos utilizar a los rendimientos observados como una aproximación a los rendimientos esperados.

   También se puede justificar haciendo uso del modelo de mercado donde suponemos que la distribución conjunta de rendimientos de los activos y del mercado es Normal, estacionaria y serialmente independiente[3]:

$$R_{jt} = \alpha_j + \beta_{jm} R_{mt} + \epsilon_{jt} \tag{1.13}$$

   donde $E(\epsilon_{jt}) = 0$, $var(\epsilon_{jt}) = \sigma^2_{\epsilon_j}$ y ϵ_{jt} es independiente de R_{mt}. El rendimiento esperado incondicional del activo j es:

$$E[R_j] = \alpha_j + \beta_{jm} E[R_m]$$

   de forma que

$$E[R_j] - \alpha_j - \beta_{jm} E[R_m] = 0 \tag{1.14}$$

   Restando las expresiones (1.13) y (1.14) obtenemos:

$$R_{jt} = E[R_j] + \beta_{jm}[R_{mt} - E[R_m]] + \epsilon_{jt} \tag{1.15}$$

   bajo la hipótesis de que el CAPM se satisface, es decir,

$$E[R_j] = r + \beta_{jm}[E(R_m) - r]$$

   por lo que operando obtenemos:

$$R_{jt} = r + \beta_{jm}(R_{mt} - r) + \epsilon_{jt} \tag{1.16}$$

   de forma que el modelo con datos observados *ex-post* parece razonable y apropiado.

[3]Se siguen suponiendo expectativas racionales.

b) El CAPM es un modelo para un único periodo, por tanto, las expresiones (1.5) y (1.7) no tienen dimensión temporal. Para el análisis econométrico es necesario añadir un supuesto en relación a la conducta de las series temporales de rentabilidades y estimar el modelo sobre el tiempo. Suponemos que los rendimientos siguen una distribución conjunta normal multivariante, son independientes e idénticamente distribuidos (IID). Los supuestos son muy fuertes pero nos permiten mantener una coherencia teórica con el modelo CAPM que se cumple periodo a periodo. Además son una buena aproximación empírica para observaciones muestrales mensuales. Los supuestos se aplican a los excesos de rendimiento de la versión de Sharpe-Lintner y a los rendimientos reales de la versión de Black. Por otro lado, al suponer que el modelo se satisface periodo a periodo añadimos un supuesto de estacionariedad sobre las primas de riesgo y sobre las betas. Por ello, en los contrastes parece apropiado la utilización de carteras en vez de rendimientos individuales.

c) Por último, necesitamos observar la verdadera cartera de mercado, lo cual resulta imposible pués numerosos activos no se negocian en mercados organizados.

La utilización de una *proxy* para la verdadera cartera de mercado implica que el CAPM como modelo teórico *no es contrastable* en la práctica. Los contrastes habituales suponen que el rendimiento de la verdadera cartera de mercado es una función lineal exacta del rendimiento de una cartera de activos cuyos precios son observables. El uso de dichas aproximaciones no aporta evidencia directa sobre si los contrastes aceptan o rechazan el modelo.

Dado que contrastar el CAPM es equivalente a contrastar que la verdadera cartera de mercado es eficiente, la no observabilidad de la cartera de mercado hace imposible realizar dicho contraste. Cuando utilizamos una proxy para el mercado lo que contrastamos es si la cartera elegida es eficiente en el sentido media-varianza y nunca podremos establecer conclusiones sobre la validez del modelo teórico ya que todas las conclusiones obtenidas han de pasar por la validez de la cartera elegida.
Por tanto, el CAPM como modelo teórico podría ser incorrecto pero la cartera empleada como aproximación podría ser eficiente en media-varianza. De igual forma la cartera usada como proxy podría no ser eficiente y sin embargo la verdadera cartera podría serlo y el CAPM sería un modelo correcto.

La estimación del modelo requiere que tengamos disponibles tres tipos de variables: las betas de los activos, la prima por riesgo del mercado y el rendimiento del activo libre de riesgo. La forma habitual de obtener la beta del activo es mediante una regresión de MCO. Así, el estimador de la pendiente en el modelo de mercado en términos de excesos de rendimiento, es un estimador habitual de la beta. El modelo a estimar es:

$$r_{jt} = \alpha_{jm} + \beta_{jm} r_{mt} + \epsilon_{jt} \qquad t = 1, \ldots, T$$

donde

$$\hat{\beta}_{jm,MCO} = \frac{\sum_t (r_{mt} - \bar{r}_m)(r_{jt} - \bar{r}_j)}{\sum_t (r_{mt} - \bar{r}_m)^2}$$

El subíndice j denota el activo para el cual estimamos su beta mediante el modelo de mercado con rendimientos en exceso. r_{jt} y r_{mt} son los rendimientos en exceso, observables en el momento t, para el activo j y la cartera de mercado respectivamente. Habitualmente como cartera de mercado se utiliza un índice de mercado y como rendimiento del activo sin riesgo se utiliza el rendimiento de las Letras del Tesoro. El periodo muestral para el que se estima la relación anterior es, habitualmente, cinco años ($T = 60$). Este tipo de aplicación sólo es útil si el CAPM es una buena descripción de los datos.

Todos los contrastes del CAPM propuestos se derivan de regresiones de serie temporal o regresiones de sección cruzada. Vamos a ver los más relevantes en la literatura.

1.3. Estimación y contrastes de sección cruzada

1.3.1. Método de Fama y MacBeth (1973)

El CAPM implica una relación lineal entre los rendimientos esperados y la beta del mercado que explica completamente la sección cruzada de los rendimientos esperados. Fama y MacBeth (1973) desarrollaron una estimación de sección cruzada, para contrastar esta implicación, que ha tenido gran influencia histórica.

El método de Fama y MacBeth es un método en etapas. En la primera se estima la relación de sección cruzada. En la segunda etapa se utilizan las estimaciones obtenidas para contrastar las implicaciones del modelo. Veamos como.

Queremos contrastar el modelo teórico recogido por la expresión (1.12):

$$E(R_j) = \gamma_0 + \gamma_1 \beta_j \qquad j = 1, \dots N$$

donde γ_0 es el rendimiento esperado de la cartera cero-beta respecto al mercado, γ_1 es la prima por riesgo del mercado y $\beta_j = Cov(R_j, R_m)/Var(R_m)$. Prescindimos del subíndice m de la beta para reconocer que el contraste se lleva a cabo con una cartera que no es la verdadera cartera de mercado.

Recordemos que el CAPM es un modelo estático. Así, periodo a periodo, se cumple la expresión (1.12). Para estimar el correspondiente modelo econométrico y verificar las implicaciones derivadas del modelo recogido por la ecuación (1.12)
dispondremos de N activos ($j = 1, \dots, N$) con T observaciones para cada activo ($t = 1, \dots, T$).

Dado que la beta del mercado explica completamente la sección cruzada de los rendimientos esperados, la beta será el único regresor del modelo. Además, necesitamos una variable que aproxime el exceso de rendimiento de cada activo en t. Así, el modelo econométrico a estimar se corresponde con el familiar Modelo Lineal Simple, que en nuestra notación correspondiente podemos escribir para el momento t como:

$$R_j = \gamma_0 + \gamma_1 \beta_j + \eta_j \qquad j = 1, \dots, N$$

donde R_j es el rendimiento de cada activo j en el periodo t, (habitualmente un mes). Esta es una variable observable que utilizamos para aproximar la no observable $E(R_j)$.

Los parámetros a estimar son γ_0 y γ_1. $\beta_j \quad j = 1, \ldots, N$ es el valor del regresor (la beta) para cada activo en t. Matricialmente el modelo a estimar, en el momento t, es:

$$\begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_N \end{bmatrix} = \begin{bmatrix} 1 & \beta_1 \\ 1 & \beta_2 \\ \vdots & \vdots \\ 1 & \beta_N \end{bmatrix} \begin{bmatrix} \gamma_0 \\ \gamma_1 \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_N \end{bmatrix} \Longleftrightarrow \underset{(N\times 1)}{R} = \underset{(N\times 2)}{X} \; \underset{(2\times 1)}{\Gamma} + \underset{(N\times 1)}{\eta}$$

Si suponemos que las betas son conocidas podemos obtener estimaciones para γ_0 y γ_1 mediante la regresión de sección cruzada anterior aplicada en cada momento t. Tendremos así estimaciones del término independiente o intercepto y de la prima por riesgo del mercado en t (recordar que ambos son comunes para todos los activos en el momento t). El estimador MCO de los parámetros del modelo, en el momento t, es $\hat{\Gamma} = (X'X)^{-1}(X'R)$. En realidad, como disponemos de T observaciones para cada uno de los N activos podemos realizar la regresión anterior en cada t, es decir, no tendremos una única regresión sino T regresiones de sección cruzada con N activos cada una. De esta forma y para remarcar la importancia del tiempo en el proceso escribiremos la sección cruzada anterior como:

$$R_{jt} = \gamma_{0t} + \gamma_{1t}\beta_{jt} + \eta_{jt} \qquad j = 1, \ldots, N \tag{1.17}$$

Donde el subíndice t indica que corremos la regresión de sección cruzada en cada t. [4]

Estimamos cada vez por MCO y obtenemos estimaciones de los parámetros en cada t, en total tendremos T estimaciones de cada parámetro. Es decir, dispondremos de una serie temporal, de tamaño T, de estimaciones para cada parámetro, que podemos escribir como $\{\hat{\gamma}_{0t}\}_{t=1}^{T}$ y $\{\hat{\gamma}_{1t}\}_{t=1}^{T}$. Posteriormente, con las estimaciones obtenidas podemos diseñar un contraste para las implicaciones que deseemos contrastar.

El problema es que en la regresión propuesta en la ecuación (1.17) las betas son desconocidas y habremos de estimarlas.

Cómo conseguir estimadores de las betas

Nuestro objetivo es estimar en <u>cada momento t</u> la ecuación de sección cruzada (1.17):

$$R_{jt} = \gamma_{0t} + \gamma_{1t}\beta_{jt} + \eta_{jt} \qquad j = 1, \ldots, N$$

Los parámetros a estimar son γ_0 y γ_1. Bajo el supuesto de rendimientos temporales IID y normalmente distribuidos el modelo anterior puede ser estimado por MCO y los estimadores obtenidos serán a su vez normales. El problema más evidente de la ecuación (1.17)

[4] Por ejemplo si $t = 1$, matricialmente:

$$\begin{bmatrix} R_{11} \\ R_{21} \\ \vdots \\ R_{N1} \end{bmatrix} = \begin{bmatrix} 1 & \beta_{11} \\ 1 & \beta_{21} \\ \vdots & \vdots \\ 1 & \beta_{N1} \end{bmatrix} \begin{bmatrix} \gamma_{01} \\ \gamma_{11} \end{bmatrix} + \begin{bmatrix} \eta_{11} \\ \eta_{21} \\ \vdots \\ \eta_{N1} \end{bmatrix}$$

es que las betas, β_{jt}, que seran los regresores del modelo, son desconocidas. Para poder estimar la ecuación (1.17) hemos de conocer el valor de estas betas. Así, hemos de estimar estas betas previamente a la estimación de la ecuación (1.17).
La manera habitual de hacerlo es utilizar el *modelo de mercado* donde suponemos que la distribución conjunta de los rendimientos de los activos y del mercado es normal, estacionaria y serialmente independiente. El modelo de mercado para el activo j podemos escribirlo como:

$$R_{jt} = \alpha_j + \beta_{jm} R_{mt} + \epsilon_{jt} \qquad t = t-1, t-2, \ldots, t-60 \tag{1.18}$$

donde $E(\epsilon_{jt}) = 0$ y ϵ_{jt} es independiente de R_{mt}. Habitualmente R_{mt} es un índice de mercado que utilizamos para aproximar la verdadera cartera de mercado. Suprimiremos el subíndice m de la beta para remarcar que la cartera de mercado utilizada es una aproximación a la verdadera cartera de mercado.

La estimación por MCO de este modelo[5] nos proporciona para cada activo j su beta correspondiente en cada t. Es decir, tenemos una serie $\{\hat{\beta}_{jt}\}_{t=1}^{T}$ para cada j=1, ..., N. Veámoslo.

El modelo a estimar, en forma matricial es ahora para cada j:

$$\begin{bmatrix} R_{j1} \\ R_{j2} \\ \ldots \\ R_{jT} \end{bmatrix} = \begin{bmatrix} 1 & R_{m1} \\ 1 & R_{m2} \\ \ldots & \ldots \\ 1 & R_{mT} \end{bmatrix} \begin{bmatrix} \alpha_j \\ \beta_j \end{bmatrix} + \begin{bmatrix} \epsilon_{j1} \\ \epsilon_{j2} \\ \ldots \\ \epsilon_{jT} \end{bmatrix} \Longleftrightarrow \underset{(T \times 1)}{R^\star} = \underset{(T \times 2)}{X^\star} \underset{(2 \times 1)}{\beta} + \underset{(T \times 1)}{\epsilon}$$

El estimador MCO de la beta correspondiente al activo j lo encontramos en la segunda fila del estimador $\hat{\beta} = (X^{\star\prime} X^\star)^{-1}(X^{\star\prime} R^\star)$. Si corremos esta regresión de serie temporal para cada uno de los N activos tendremos la betas estimadas correspondientes al momento t.
Sin embargo, la regresión (1.17) se corre en una muestra de sección cruzada en t cuando $t = 1, \ldots, T$. Por tanto, para cada activo j necesitamos una serie temporal de betas estimadas, que serán utilizadas en sección cruzada. Por ello, para conseguir esa serie temporal de betas estimadas necesitamos que la regresión (3.10) se realice para cada activo en t cuando $t = 1, \ldots, T$, en total $N \times T$ regresiones de serie temporal con una muestra de tamaño 60. Dispondremos, por tanto, de N series tipo $\{\hat{\beta}_{jt}\}_{t=1}^{T} \quad j = 1, \ldots, N$ que habremos de reordenar en sección cruzada para estimar a continuación la regresión (1.17).

Cómo utilizar la estimaciones de las betas

Una vez estimadas las betas podemos estimar la ecuación (1.17) donde sustituimos las betas originales por su correspondiente estimación. Así, en cada t estimamos por MCO la siguiente regresión con datos de sección cruzada.

$$R_{jt} = \gamma_{0t} + \gamma_{1t} \hat{\beta}_{jt} + \eta_{jt} \qquad j = 1, \ldots, N \tag{1.19}$$

[5] En la literatura, dados los estudios empíricos realizados, es habitual utilizar las 60 observaciones previas a la observación t como periodo muestral que permite un mejor ajuste a las verdaderas betas, de las que no disponemos.

obteniendo $\hat{\gamma}_0$ y $\hat{\gamma}_1$ en t. Si repetimos el ejercicio de regresión T veces es decir desde $t = 1, \ldots, T$ obtenemos, para cada uno de los coeficientes, una serie temporal de tamaño T, $\{\hat{\gamma}_{0t}\}_{t=1}^T$ y $\{\hat{\gamma}_{1t}\}_{t=1}^T$, que nos serán útiles para realizar inferencia.

Contrastes

Definimos $\gamma_0 = E(\gamma_{0t})$ y $\gamma_1 = E(\gamma_{1t})$. Las implicaciones del cero-beta CAPM son que:

$$\gamma_0 > 0 \qquad \text{y} \qquad \gamma_1 > 0$$

Es decir esperamos que el rendimiento esperado de la cartera cero-beta y la prima por riesgo del mercado sean, en promedio, estrictamente positivas.
Dado que hemos supuesto que los rendimientos son normalmente distribuidos y temporalmente IID los parámetros estimados también seran normalmente distribuidos. Por tanto, dadas las series temporales de estimaciones de los parámetros γ_{0t} y γ_{1t} $\;T = 1, \ldots, T$, podemos contrastar estas implicaciones usando el estadístico-t habitual para contrastes sobre la media de una población.

Las hipótesis de contraste podemos escribirlas como:

$$\begin{array}{ll} H_0 : \gamma_i = 0 & \qquad \text{donde} \quad i = 0, 1 \\ H_a : \gamma_i > 0 & \end{array}$$

dadas las propiedades temporales que hemos supuesto para los rendimientos de los activos, para contrastar estas hipótesis podemos utilizar un estadístico t habitual definido:

$$t_c = \frac{\bar{\hat{\gamma}}_i}{\hat{\sigma}_{\hat{\gamma}_i}} \overset{H_0}{\sim} t_{(T-1)}$$

donde

$$\bar{\hat{\gamma}}_i = \frac{1}{T} \sum_{t=1}^{T} \hat{\gamma}_{it}$$

$$\hat{\sigma}^2_{\hat{\gamma}_i} = \frac{1}{T(T-1)} \sum_{t=1}^{T} \left(\hat{\gamma}_{it} - \bar{\hat{\gamma}}_i \right)^2$$

las cuales calculamos utilizando $\{\hat{\gamma}_{it}\}_{t=1}^T$. Asintóticamente este estadístico tiende a una $N(0, 1)$.
Si lo que analizamos es la versión de Sharpe-Lintner estaríamos mirando a la ecuación (1.17) en términos de excesos de rendimiento y las implicaciones a contrastar serían $\gamma_0 = 0$ y $\gamma_1 > 0$. El proceso sería exactamente igual al anterior [6].

[6] Si la hipótesis alternativa es a dos colas, como es el caso para $H_0 : \gamma_0 = 0$, debemos comparar el estadístico t-calculado con el valor en tablas para $\alpha/2$, mientras que si la hipótesis alternativa es a una cola, como en el resto de casos, comparamos con el valor en tablas para α, siendo α el valor de significación de la prueba.

Notar que dado que las varianzas de los activos en la ecuación de sección cruzada (1.19) son heterocedásticas el estimador MCO es ineficiente. Este resultado es irrelevante ya que el estadístico de contraste no depende de la forma explícita de la matriz de varianzas y covarianzas de las perturbaciones en t.

Problemas econométricos

El primer problema que encontramos en relación a la estimación es el hecho de que las betas son desconocidas y nos vemos obligados a estimarlas previamente. La utilización de estimaciones de las betas en (1.19) en vez de las betas verdaderas introduce un problema de errores en variables en la regresión que produce sesgo e inconsistencia en los estimadores MCO de γ_0 y γ_1. Veámos como:

Nosotros no observamos β_{jt} y disponemos de $\hat{\beta}_{jt}$, es decir, la beta verdadera más un error de medida tal que $\hat{\beta}_{jt} = \beta_{jt} + w_{jt}$ donde[7] $E(w_{jt}) = 0$, $E(w_{jt}^2) = \sigma_{w_j}^2$ y además $Cov(\eta_{jt}, w_{jt}) = 0$, es decir, no hay correlación entre el error de medida y las perturbaciones de la regresión, y tampoco con las verdaderas betas, $Cov(\beta_{jt}, w_{jt}) = 0$. El modelo estimable en estas circunstancias, en cada momento t, es:

$$R_{jt} = \gamma_{0t} + \gamma_{1t}(\hat{\beta}_{jt} - w_{jt}) + \eta_{jt} \qquad j = 1, \ldots, N$$

es decir,

$$R_{jt} = \gamma_{0t} + \gamma_{1t}\hat{\beta}_{jt} + (\eta_{jt} - \gamma_{1t}w_{jt}) \qquad j = 1, \ldots, N$$

donde recordemos que hemos supuesto:

$$\begin{array}{ll} \eta_{jt} \sim iid(0, \sigma_{\eta_j}^2) & Cov(\eta_{jt}, w_{jt}) = 0 \\ w_{jt} \sim iid(0, \sigma_{w_j}^2) & Cov(\beta_{jt}, w_{jt}) = 0 \end{array}$$

Si llamamos a la perturbación del modelo estimable $\xi_{jt} = (\eta_{jt} - \gamma_{1t}w_{jt})$ y buscamos sus propiedades tenemos que:

$$\begin{array}{l} E(\xi_{jt}) = 0 \\ E(\xi_{jt}^2) = \sigma_{\eta_j}^2 + \gamma_{1t}^2\sigma_{w_j}^2 \\ Cov(\xi_{jt}, \hat{\beta}_{jt}) = Cov(\eta_{jt} - \gamma_{1t}w_{jt}, \beta_{jt} + w_{jt}) = -\gamma_{1t}\sigma_{w_j}^2 \neq 0 \end{array}$$

Con lo que el estimador de MCO es sesgado. Si buscamos el límite en probabilidad cuando $N \to \infty$

$$plim\hat{\gamma}_{1t} = \gamma_{1t} - \frac{\gamma_{1t}\sigma_{w_j}^2}{\sigma_{\beta_j}^2 + \sigma_{w_j}^2}$$

y el estimador MCO es inconsistente.
El problema de errores en variables puede ser abordado de diferentes formas dependiendo de nuestro objetivo. Habitualmente el objetivo es realizar contraste de hipótesis sobre

[7] El subíndice t indica que la beta corresponde a la regresión de sección cruzada realizada en el momento t, por tanto, es un subíndice fijo, no indica aleatoriedad y desde ese punto de vista es totalmente prescindible para el análisis de errores en variables. Se incluye por consistencia notacional.

un parámetro del modelo, empleando la distribución de dicho parámetro. En este caso se suele proponer utilizar el estimador de Variables Instrumentales que proporciona un estimador consistente con el que podemos hacer inferencia asintótica. Sin embargo, no es el caso aquí, ya que nosotros queremos estimar las betas para luego poder estimar los parámetros γ_{it} y basamos el estadístico de contraste en la serie temporal de estimadores $\hat{\gamma}_{it}$. El estimador de VI sólo proporciona propiedades en muestras grandes, concretamente nos proporciona una matriz de varianzas y covarianzas consistente del estimador de las betas, que no interviene en el estadístico, pero no nos garantiza la insesgadez del estimador. Por ello, estaremos interesados en corregir el sesgo producido por el problema de errores en variables más que en la consistencia del estimador de las betas. En la literatura hay dos formas de enfocar el problema.

- La solución propuesta por Fama y MacBeth (1973) es utilizar carteras en la estimación de las betas lo que incrementa la precisión de los estimadores. Estas carteras han de estar bien definidas. En concreto deben tener máxima dispersión entre las betas de las carteras, ya que si la dispersión es mínima la relación (1.17) no tiene contenido como relación de sección cruzada. Hay varias formas correctas de construirlas, por betas, por capitalización bursátil, por sector industrial, etc. El problema es que esta medida en general no elimina totalmente el problema de errores en variables ya que habitualmente los errores de medida entre las betas de los activos de una misma cartera están correlacionados.

- Litzenberger y Ramaswamy (1979) propusieron una solución en términos de ajustar el estimador de $\sigma_{\hat{\gamma}_i}$ por el sesgo introducido por los errores en variables. El procedimiento fue refinado por Shanken (1992b, 1996). Shanken propuso multiplicar el estimador $\hat{\sigma}^2_{\hat{\gamma}_i}$ por el siguiente factor de ajuste:

  $$\left(1 + \frac{(\bar{R}_m - \bar{\hat{\gamma}}_0)^2}{\hat{\sigma}^2_m}\right)$$

  donde $\bar{\hat{\gamma}}_0 = \frac{1}{T}\sum_{t=1}^{T} \hat{\gamma}_{0t}$, $\quad \bar{R}_m = \frac{1}{T}\sum_{t=1}^{T} R_{mt}$ y $\hat{\sigma}^2_m$ es el estimador de la varianza muestral, para $t = 1, \ldots, T$, del índice bursátil utilizado como aproximación a la verdadera cartera de mercado. Este procedimiento elimina el sesgo producido en el estadístico t por el problema de errores en variables.

Otro problema añadido es el ya mencionado al inicio del tema sobre la cartera de mercado, la utilización de una aproximación a esta verdadera cartera de mercado. Roll y Ross (1994) muestran cómo si la verdadera cartera de mercado es eficiente, la relación de sección cruzada entre los rendimientos esperados y betas es muy sensible a pequeñas desviaciones de la proxy del mercado con respecto a la verdadera cartera. En un trabajo de Kandel y Stambaugh (1995) se muestra que el efecto se reduce en parte utilizando el estimador de Mínimos Cuadrados Generalizados en vez de MCO, pero en este caso hay que conocer la verdadera matriz de varianzas y covarianzas de los rendimientos ya que su estimación implica nuevos problemas.

1.3.2. Ventajas o extensiones

Cuando decimos que las betas explican toda la sección cruzada de los rendimientos esperados lo que queremos decir es que ninguna otra variable debe explicar a éstos. La aproximación de Fama y MacBeth (1973) permite fácilmente contrastar ésta implicación, basta con introducir estas variables en el modelo de sección cruzada y contrastar la significatividad del parámetro asociado con el estadístico de contraste propuesto anteriormente. En este caso el modelo a estimar sería:

$$R_{jt} = \gamma_{0t} + \gamma_{1t}\beta_{jt} + \gamma_{2t}X_{jt} + \eta_{jt} \qquad j = 1, \ldots, N$$

donde la variable X_{jt} es la variable que creemos que explica la sección cruzada de los rendimientos medios. Una vez estimado por el método propuesto por Fama y MacBeth contrastaríamos $H_0 : \gamma_2 = 0$ con el estadístico t_c propuesto anteriormente donde $i = 2$.

Fama y MacBeth en su trabajo propusieron varias variables candidatas para esta regresión:

- La variable β_{jt}^2 permite contrastar la linealidad del modelo.
- La desviación estándar de los residuos del modelo de mercado para cada activo j permite estudiar si el mercado valora el riesgo idiosincrásico.
- El tamaño de la empresa (medido como el logarítmo del valor de mercado) también es una variable candidata a explicar las rentabilidades.

El ajuste propuesto por Shanken para solucionar el problema de errores en variables soluciona el sesgo por este error pero no impide que entren nuevas variables en la relación por el hecho de no observar las verdaderos betas. Por lo que este tipo de extensiones tienen sentido incluso cuando aplicamos la corrección de Shanken.

1.4. Contrastes de eficiencia en media varianza de una determinada cartera. Sección cruzada

Cuando decimos que el CAPM se satisface en realidad decimos que la cartera de mercado es eficiente en el sentido media-varianza. Así podemos pensar en contrastar si la cartera de mercado es eficiente o no. Como la verdadera cartera de mercado no es observable lo que contrastaríamos es si el índice bursátil que utilizamos para aproximar la verdadera cartera de mercado es eficiente o no. La forma más habitual de denotar el cero-beta CAPM se recogía en la ecuación (1.12):

$$E(R_j) = \gamma_0 + \gamma_1\beta_{jm} \qquad j = 1, \ldots N$$

donde γ_0 y γ_1 son el rendimiento esperado de la cartera cero-beta respecto al mercado y la prima por riesgo del mercado respectivamente. Dado que el CAPM se cumple periodo a periodo implica que en t, γ_0 y γ_1 son comunes a todos los activos.

En (1.12) contrastamos si el modelo se satisface respecto a una beta (estimada) con relación al índice bursátil cuya eficiencia se quiere contrastar.

1.4.1. Cómo contrastar cuando suponemos que conocemos las verdaderas betas y la verdadera matriz de varianzas y covarianzas de los rendimientos

En este caso los supuestos son muy fuertes pero es el marco más sencillo de abordar por lo que es un buen principio para lograr el objetivo final de contrastar la eficiencia en media y varianza relajando ambos supuestos.
Al igual que en las secciones anteriores vamos a disponer de una muestra de N activos ($j = 1, \ldots, N$) y para cada uno de ellos dispondremos de una serie temporal de tamaño T de observaciones de rendimientos.

La ecuación (1.12) en notación matricial podemos escribirla como:

$$E = X\Gamma \tag{1.20}$$

donde

$$\begin{bmatrix} E(R_1) \\ E(R_2) \\ \vdots \\ E(R_N) \end{bmatrix} = \begin{bmatrix} 1 & \beta_1 \\ 1 & \beta_2 \\ \vdots & \vdots \\ 1 & \beta_N \end{bmatrix} \begin{bmatrix} \gamma_0 \\ \gamma_1 \end{bmatrix} \Longleftrightarrow \underset{(N\times 1)}{E} = \underset{(N\times 2)}{X} \; \underset{(2\times 1)}{\Gamma}$$

Siendo E el vector de **rendimientos esperados incondicionales**.
Sea R_t el vector de rendimientos observados de los activos en el momento t, de orden $(N \times 1)$. Definimos $\eta_t \equiv R_t - E$ tal que bajo la hipótesis nula dada por (1.20) podemos escribir las regresiones de sección cruzada vistas anteriormente en notación matricial periodo a periodo (es decir, en el momento t) como:

$$R_t = X\Gamma + \eta_t \tag{1.21}$$

donde el subíndice t indica que la regresión (1.21) se corre T veces, $t = 1, \ldots, T$. En realidad hay T regresiones de sección cruzada del tipo:

$$\begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_N \end{bmatrix} = \begin{bmatrix} 1 & \beta_1 \\ 1 & \beta_2 \\ \vdots & \vdots \\ 1 & \beta_N \end{bmatrix} \begin{bmatrix} \gamma_0 \\ \gamma_1 \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_N \end{bmatrix}$$

Supondremos que las perturbaciones de estas regresiones de sección cruzada tienen media cero pero son heterocedásticas, dado que las varianzas de los rendimientos de los activos son distintas. Vamos a suponer que la verdadera matriz de varianzas y covarianzas de los rendimientos, a la que llamaremos V, (matriz de orden $N \times N$), es conocida. Además supondremos que las verdaderas betas son conocidas. También supondremos que la distribución conjunta de los rendimientos de los activos es normal, estacionaria y serialmente independiente. En este contexto donde las betas son conocidas y por tanto X es una matrix conocida y dadas las propiedades recogidas en la matriz de varianzas y covarianzas de la perturbación, V, deberíamos estimar los parámetros de la relación en t, $R_t = X\Gamma + \eta_t$, por Mínimos Cuadrados Generalizados (MCG) de tal forma que en cada momento t el vector de estimadores es:

$$\hat{\Gamma}_t = (X'V^{-1}X)^{-1}X'V^{-1}R_t \tag{1.22}$$

Estimador lineal, insesgado, eficiente, con matriz de varianzas y covarianzas $Var(\hat{\Gamma}_t) = (X'V^{-1}X)^{-1}$ y consistente. El estimador tiene distribución normal dados los supuestos sobre la distribución conjunta de los rendimientos.

Dado el supuesto de estacionariedad e independencia el estimador más eficiente de Γ utilizando **todos** los datos del periodo muestral en una **única** regresión de sección cruzada con N observaciones será:

$$\hat{\Gamma} = (X'V^{-1}X)^{-1}X'V^{-1}\bar{R} \tag{1.23}$$

donde $\bar{R}$ es el vector $N \times 1$ de rendimientos medios muestrales de los N activos. La matriz de varianzas y covarianzas de $\hat{\Gamma}$ es[8]:

$$Var(\hat{\Gamma}) = \frac{1}{T}(X'V^{-1}X)^{-1}$$

Una vez estimado el modelo estamos en disposición de contrastar. Queremos contrastar la eficiencia en media-varianza del mercado. Esta hipótesis está recogida en la ecuación (1.20) e implica que los residuos de la regresión han de ser cero. Por tanto si la hipótesis nula no se rechaza observaremos que los residuos de la regresión oscilan alrededor de cero y si es falsa se desviarán sistemáticamente de cero, esta desviación puede surgir por ejemplo porque estamos intentando ajustar una relación lineal cuando la verdadera relación no lo es, o porque deberíamos incorporar otras variables que expliquen los rendimientos esperados además de su beta. Para contrastar esta hipótesis nula podemos basarnos en la magnitud de los residuos al cuadrado y utilizar el siguiente estadístico:

$$e' \left[\frac{V}{T}\right]^{-1} e = Te'V^{-1}e \overset{H_0}{\sim} \chi^2_{N-2} \tag{1.24}$$

donde $e = \bar{R} - X\hat{\Gamma}$ son los residos de la regresión MCG en (1.23).
Observaciones:

- Intuición del contraste: Una suma de residuos al cuadrado baja indica que los residuos no se separan significativamente de cero. Por tanto, valores pequeños del estadístico nos llevan a no rechazar la hipótesis nula y valores altos del mismo nos llevarían a rechazarla.

- Para pensar en la distribución, pensemos en el modelo lineal general $Y = X\beta + u$ con $E(uu') = \sigma^2\Omega$ donde Ω es conocida y σ^2 es un parámetro desconocido. El estimador de σ^2

  $$\tilde{\sigma}^2_{MCG} = \frac{\tilde{u}'_{MCG}\Omega^{-1}\tilde{u}_{MCG}}{N-k}$$

  es insesgado y tal que

  $$\frac{(N-k)\tilde{\sigma}^2_{MCG}}{\sigma^2} \sim \chi^2_{N-k}$$

[8]Notar que dado $X_t \sim iid(\mu, \sigma^2)$, $E(\bar{X}) = \mu$ y $Var(\bar{X}) = \frac{\sigma^2}{T}$. Entonces para el estimador de la media de $\hat{\Gamma}_t$, $\hat{\Gamma}$, tendremos $Var(\hat{\Gamma}) = \frac{Var(\hat{\Gamma}_t)}{T} = \frac{(X'V^{-1}X)^{-1}}{T}$

$$\frac{\tilde{u}'_{MCG}\Omega^{-1}\tilde{u}_{MCG}}{\sigma^2} \sim \chi^2_{N-k}$$

si

$$\sigma^2 = 1, \quad \tilde{u}'_{MCG}\Omega^{-1}\tilde{u}_{MCG} \sim \chi^2_{N-k}$$

En el modelo con una única regresión de sección cruzada la matriz de varianzas y covarianzas de la perturbación es V/T. Si los residuos se denotan por e obtenemos:

$$e'\left[\frac{V}{T}\right]^{-1} e = Te'V^{-1}e \overset{H_0}{\sim} \chi^2_{N-2}$$

- Otra manera de pensar en el estadístico y su distribución: Sea e_t el vector de orden $N \times 1$ de residuos de la regresión cruzada que corremos en cada t $e_t = R_t - X\hat{\Gamma}_t = M\eta_t$ donde $M = (I - X(X'V^{-1}X)^{-1}X'V^{-1})$ y tal que

$$E(e_t) = 0 \quad E(e_t e_t') = V - X(X'V^{-1}X)^{-1}X'$$

en t. Pensemos en los residuos de la regresión única de sección cruzada $e = \frac{\sum e_t}{T}$, tendremos:

$$E(ee') = \frac{E(e_t e_t')}{T} = \frac{1}{T}[V - X(X'V^{-1}X)^{-1}X'] \qquad \text{matriz de orden } (N \times N)$$

Podríamos utilizar esta matriz de varianzas y covarianzas para construir el estadístico pero podemos utilizar una forma más sencilla para hacerlo. Utilizando la descomposición de Choleski, definimos una matriz C tal que $CC' = V^{-1}$. La matriz de covarianzas de la forma $\sqrt{T}C'e$ es:

$$Cov(\sqrt{T}C'e) = C'\left((CC')^{-1} - X(X'CC'X)^{-1}X'\right)C = I - \delta(\delta'\delta)^{-1}\delta'$$

donde $\delta = C'X$. Al ser los rendimientos normales, también lo serán los residuos y la forma $\sqrt{T}C'e$ será asintóticamente normal, $Cov(\sqrt{T}C'e)$ será una matriz idempotente con rango $N-2$ y por lo tanto $Te'CC'e = Te'V^{-1}e \overset{H_0}{\sim} \chi^2_{N-2}$.

- Los grados de libertad son $N-2$ ya que estimamos dos parámetros γ_0 y γ_1.

1.4.2. ¿Qué ocurre cuando reconocemos que V es desconocida?

En la práctica la matriz de varianzas y covarianzas V no es conocida y por tanto el estimador MCG propuesto en la sección anterior no es factible. El estimador que debemos aplicar sería el estimador de Mínimos Cuadrados Generalizados Factibles (MCGF) donde estimamos previamente la matriz V para a continuación estimar los parámetros del modelo. Además el estimador de V que buscamos debe ser consistente para que a su vez el estimador MCGF lo sea y podamos realizar inferencia en el límite.

Notar que seguimos manteniendo el supuesto de que las betas verdaderas son conocidas y reconocemos que la matriz V es desconocida.

Para encontrar el estimador de V empezaremos *analizando la relación* **condicional** *entre rendimiento esperado y riesgo bajo la hipótesis nula de que el CAPM se satisface*, es decir se cumple la relación incondicional dada por la ecuación (1.20):

$$E = X\Gamma$$

El modelo de mercado en notación matricial para el momento t se escribe:

$$R_t = \alpha + \beta R_{mt} + \epsilon_t \tag{1.25}$$

donde α es el vector de orden $(N \times 1)$ de ordenadas en el origen, β es el vector $(N \times 1)$ de betas y ϵ_t es el vector $(N \times 1)$ de perturbaciones (condicionales dado R_{mt}) del modelo de mercado o componente idiosincrásico del rendimiento (dado R_{mt}). Tal que:

$$E(\epsilon_t) = E(\epsilon_t|R_{mt}) = 0 \qquad V(\epsilon_t\epsilon_t'|R_{mt}) = \Sigma$$

Siendo Σ la matriz de varianzas y covarianzas de las perturbaciones condicionales de la ecuación (1.25) en t. La matriz es de orden $N \times N$.

Si estimamos el modelo de mercado para cada activo en la muestra usando las observaciones temporales, obtenemos una serie temporal de residuos $\{\hat{\epsilon}_t\}_{t=1}^T$ para cada activo j que puede ser útil para estimar la matriz poblacional.

• Llamamos $E(R_m)$ al rendimiento esperado del mercado y tomamos valores esperados **incondicionales** en (1.25)

$$E = \alpha + \beta E(R_m) \tag{1.26}$$

Bajo el supuesto de que el CAPM se satisface, es decir $E = X\Gamma$, podemos escribir:

$$E = X\Gamma = \gamma_0 1_N + \gamma_1\beta + \beta E(R_m) - \beta E(R_m) = \gamma_0 1_N + [\gamma_1 - E(R_m)]\,\beta + \beta E(R_m)$$

si tenemos en cuenta (1.26) $E = X\Gamma$ se satisface si y sólo si:

$$\alpha = \gamma_0 1_N + [\gamma_1 - E(R_m)]\,\beta \tag{1.27}$$

que dada la definición de la prima de mercado, es equivalente a la expresión (1.44).

• Ahora tomamos valores esperados **condicionales** en el modelo de mercado recogido en la ecuación (1.25):

$$E(R_t|R_{mt}) = \alpha + \beta R_{mt} \tag{1.28}$$

sustituyendo el valor de α dado por (1.27) en (1.28) obtenemos:

$$E(R_t|R_{mt}) = \gamma_0 1_N + [\gamma_1 + R_{mt} - E(R_m)]\,\beta \tag{1.29}$$

La ecuación (1.29) implica que bajo la hipótesis de que el CAPM se satisface, es decir, $E = X\Gamma$, la relación condicional (dado R_{mt}) entre rendimiento esperado y riesgo beta es también lineal. Trabajando en la expresión (1.25) tenemos:

$$R_t = a + \beta R_{mt} + \epsilon_t = \gamma_0 1_N + [\gamma_1 - E(R_m)]\,\beta + \beta R_{mt} + \epsilon_t$$

$$R_t = \gamma_0 1_N + \gamma_1 \beta + [R_{mt} - E(R_m)]\, \beta + \epsilon_t$$

Llamamos

$$\Gamma_t = \Gamma + \left[\begin{array}{c} R_{mt} - E(R_m) \end{array}\right] \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$$

con lo que podemos denotar la expresión de la relación rendimiento-riesgo *ex-post* anterior de la siguiente forma matricial:

$$R_t = X\Gamma_t + \epsilon_t \qquad t = 1, \ldots, T \tag{1.30}$$

Lo fundamental de esta nueva relación lineal rendimiento-riesgo condicional es que el parámetro de la pendiente, γ_{1t}, es aleatorio al ser función del rendimiento de mercado en t, R_{mt}.

En resumen, tenemos dos formas alternativas de representar el mismo proceso:

$$R_t = X\Gamma + \eta_t \tag{1.31}$$

$$R_t = X\Gamma_t + \epsilon_t \tag{1.32}$$

Las perturbaciones de ambos modelos tienen la siguiente relación: dado que $\eta_t = R_t - X\Gamma = R_t - E$

$$\eta_t = \epsilon_t + \beta\, [R_{mt} - E(R_m)] \tag{1.33}$$

por tanto

$$V = \Sigma + \beta\beta'\sigma_m^2 \tag{1.34}$$

donde σ_m^2 es la varianza de la cartera de mercado. Shanken (1985) demostró que:

$$(X'V^{-1}X)^{-1}X'V^{-1} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}$$

Esto implica que el estimador MCG de Γ es:

$$\hat{\Gamma} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\bar{R} \tag{1.35}$$

donde Σ puede ser estimada previamente. En realidad al tener que estimar Σ previamente nuestro estimador es Mínimos Cuadrados Generalizados Factibles y en el resto de ecuaciones Σ debe ser sustituida por $\hat{\Sigma}$ para reconocer este hecho.

El estimador de la varianza de $\hat{\Gamma}$ cuando suponíamos que V era conocida era:

$$Var(\hat{\Gamma}) = \frac{(X'V^{-1}X)^{-1}}{T}$$

ahora esta expresión debe de tener en cuenta la relación entre la matriz de varianzas y covarianzas incondicional, V, y la condicional Σ

$$V = \Sigma + \beta\beta'\sigma_m^2$$

y no podemos intercambiar una por otra en la expresión de $Var(\hat{\Gamma})$. Si pensamos en la relación rendimiento-riesgo condicional, ecuación (1.30), vemos que el parámetro γ_{1t} depende ahora del rendimiento de mercado en t, R_{mt}. Es decir el rendimiento de mercado introduce variabilidad en el parámetro γ_1 que debe ser tenida en cuenta. Por ello el estimador de la varianza de $\hat{\Gamma}$ en el contexto de una única regresión de sección cruzada es:

$$Var(\hat{\Gamma}) = \frac{(X'\Sigma^{-1}X)^{-1}}{T} + \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right] \frac{\sigma_m^2}{T} \tag{1.36}$$

donde la variabilidad del rendimiento de mercado afecta únicamente al parámetro γ_1 y no al parámetro γ_0

Notar que este estimador es equivalente al propuesto por Fama y MacBeth. Por tanto podemos realizar T regresiones de sección cruzada y utilizar el estadístico t_c propuesto por Fama y MacBeth o una única regresión de sección cruzada con los rendimientos medios de los activos y las betas de todo el periodo muestral y emplear (1.35) y (1.36).

Para llevar a cabo el contraste de eficiencia media-varianza podemos emplear el estadístico:

$$Te'\hat{\Sigma}^{-1}e \xrightarrow{d,H_0} \chi^2_{N-2} \tag{1.37}$$

con distribución asintótica ya que usamos un estimador Factible cuya única propiedad es la consistencia. Además estamos suponiendo que conocemos las verdaderas betas.

1.4.3. ¿Qué ocurre si reconocemos que las verdaderas betas no son conocidas?

En este caso las betas deben ser estimadas previamente y como ya vimos en secciones anteriores esto introduce un problema de errores en variables en el modelo. Vamos a estudiar ahora los efectos de estos errores de estimación en los estadísticos anteriores. Sea la relación:

$$R_t = X\Gamma + \eta_t$$

tomamos valores medios muestrales y

$$\bar{R} = X\Gamma + \bar{\eta} \tag{1.38}$$

la ecuación (1.33) nos dice que $\eta_t = \epsilon_t + \beta[R_{mt} - E(R_m)]$ de donde:

$$\bar{\eta} = \bar{\epsilon} + \beta[\bar{R}_m - E(R_m)]$$

Cuando tratamos las consecuencias del error de estimación en las betas definíamos $\hat{\beta}_{jt} = \beta_{jt} + w_{jt}$ de donde el error de estimación en las betas en forma matricial es: $w = \hat{\beta} - \beta$. Así trabajando con la expresión (1.38) obtenemos:

$$\bar{R} = \gamma_0 1_N + \gamma_1\beta + \bar{\epsilon} + \beta[\bar{R}_m - E(R_m)]$$
$$\bar{R} = \gamma_0 1_N + \gamma_1\hat{\beta} - \gamma_1 w + \bar{\epsilon} + \beta[\bar{R}_m - E(R_m)]$$
$$\bar{R} = \hat{X}\Gamma + \{(\bar{\epsilon} - \gamma_1 w) + \beta[\bar{R}_m - E(R_m)]\}$$

donde

$\hat{X}$ denota el hecho de que ahora las betas son estimadas.

$(\bar{\epsilon} - \gamma_1 w)$ es el cambio en la perturbación dado el error de medida en las betas.

$\beta[\bar{R}_m - E(R_m)]$ es el cambio en la perturbación debido a la innovación en el rendimiento de mercado como consecuencia de la aleatoriedad de la prima por riesgo del propio mercado.

Definimos la matriz $\hat{A}$ de orden $(2 \times N)$ tal que $\hat{A}\hat{X} = I_2$:

$$\hat{A} = \left(\hat{X}'\hat{\Sigma}^{-1}\hat{X}\right)^{-1} \hat{X}'\hat{\Sigma}^{-1}$$

y

$$\hat{A}\bar{R} = \left(\hat{X}'\hat{\Sigma}^{-1}\hat{X}\right)^{-1} \hat{X}'\hat{\Sigma}^{-1}\bar{R}$$

Por tanto $\hat{\Gamma}_{MCG} = \hat{A}\bar{R}$. Así dado $\hat{A}$ tenemos:

$$\hat{A}\bar{R} = \hat{A}\hat{X}\Gamma + \hat{A}\{(\bar{\epsilon} - \gamma_1 w) + \beta[\bar{R}_m - E(R_m)]\}$$

dado que $\hat{A}\hat{X} = I_2$

$$\hat{A}\bar{R} - \Gamma = \hat{A}\{(\bar{\epsilon} - \gamma_1 w) + \beta[\bar{R}_m - E(R_m)]\}$$

Cuando $T \to \infty$ se satisface:

$$E(\bar{\epsilon}) = 0 \quad E(w) = 0 \quad E[\bar{R}_m - E(R_m)] = 0$$

por tanto

$$plim\{(\bar{\epsilon} - \gamma_1 w) + \beta[\bar{R}_m - E(R_m)]\} = 0$$

y

$$plim(\hat{A}\bar{R} - \Gamma) = 0$$
$$plim(\hat{A}\bar{R}) = \Gamma$$

por tanto el estimador $\hat{A}\bar{R}$ es un estimador consistente de Γ cuando $T \to \infty$. Sin embargo sigue siendo inconsistente cuando $N \to \infty$ dado que el problema de errores en variables no desaparece. Recordar que estamos trabajando simultáneamente con dos dimensiones, T y N y el problema de errores en variables está asociado a la inconsistencia del estimador cuando $N \to \infty$.

Necesitamos analizar la varianza asintótica de $\hat{\Gamma}$ cuando $T \to \infty$. En el modelo de mercado $R_{jt} = \alpha_j + \beta_{jm} R_{mt} + \epsilon_{jt} \qquad t = 1, \ldots, T$ la varianza del estimador de la beta es:

$$Var(\hat{\beta}_j) = \sigma^2_{\hat{\beta}_j} = \frac{\sigma^2_{\epsilon_j}}{\sum_{t=1}^T (R_{mt} - \bar{R}_m)^2} = \frac{\sigma^2_{\epsilon_j}}{T\sigma^2_m}$$

y se puede demostrar que la varianza del error de estimación de las betas es:

$$\sigma^2_w = \left(\frac{1}{\sigma^2_m}\right)\Sigma$$

Además bajo los supuestos realizados de normalidad e independencia de los rendimientos, $\bar{\epsilon}, \bar{R}_m$ y w son independientes. Por tanto, usando esta propiedad de independencia y la expresión de σ_w^2 tenemos la siguiente convergencia en distibución:

$$\sqrt{T}\{(\bar{\epsilon} - \gamma_1 w) + \beta[\bar{R}_m - E(R_m)]\} \xrightarrow{d} N\left(0, \Sigma\left(1 + \frac{\gamma_1^2}{\sigma_m^2}\right) + \beta\beta'\sigma_m^2\right) \tag{1.39}$$

$$\sqrt{T}\hat{A}\{(\bar{\epsilon} - \gamma_1 w) + \beta[\bar{R}_m - E(R_m)]\} \xrightarrow{d} N\left(0, A\Sigma A'\left(1 + \frac{\gamma_1^2}{\sigma_m^2}\right) + \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]\sigma_m^2\right)$$

dado que:

$$\hat{A}\{(\bar{\epsilon} - \gamma_1 w) + \beta[\bar{R}_m - E(R_m)]\} = \hat{A}\bar{R} - \Gamma$$

$$\sqrt{T}(\hat{A}\bar{R} - \Gamma) \xrightarrow{d} N\left(0, A\Sigma A'\left(1 + \frac{\gamma_1^2}{\sigma_m^2}\right) + \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]\sigma_m^2\right) \tag{1.40}$$

al ser $\hat{A}\bar{R} = \hat{\Gamma}$ y $\hat{A} = (\hat{X}'\hat{\Sigma}^{-1}\hat{X})^{-1}\hat{X}'\hat{\Sigma}^{-1}$, la distribución asintótica del estimador de Γ podemos escribirla como:

$$\sqrt{T}(\hat{\Gamma} - \Gamma) \xrightarrow{d} N\left(0, X'\Sigma^{-1}X\left(1 + \frac{\gamma_1^2}{\sigma_m^2}\right) + \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]\sigma_m^2\right)$$

donde $\left(1 + \frac{\gamma_1^2}{\sigma_m^2}\right)$ es el término de ajuste que proviene de la varianza del error de estimación de las betas y que al ser positivo incrementa la varianza. Tenemos una estimación menos precisa de Γ. Notar que cuanto mayor sea la variabilidad del mercado en relación a γ_1^2 (si la volatilidad del mercado es grande con relación a la prima de riesgo) este término será menor. Este término de ajuste es el propuesto por Shanken (1996).

Para este caso en el que las betas verdaderas no son conocidas el estadístico de contraste para la hipótesis de que la cartera de mercado utilizada es eficiente en media y varianza es:

$$\frac{Te'\hat{\Sigma}^{-1}e}{\left(1 + \frac{\gamma_1^2}{\sigma_m^2}\right)} \xrightarrow{d,H_0} \chi_{N-2}^2$$

1.5. Estimación y Contrastes de series temporales

1.5.1. El alfa de Jensen

Vamos a relacionar dos modelos distintos como son el modelo de valoración CAPM y el modelo de mercado. El modelo de mercado es el resultado de imponer una serie de supuestos estadísticos sobre los rendimientos bivariantes de los activos y el mercado. Se supone que la distribución conjunta de rendimientos de los activos y del mercado es normal, estacionaria y serialmente independiente. Este modelo no tiene relación con el CAPM pero podemos relacionar ambos planteamientos. El modelo de mercado en términos esperados es:

$$E(R_j) = \alpha_{jm} + \beta_{jm} E(R_m) \tag{1.41}$$

Imponiendo el CAPM en la expresión (1.41), ($E(R_j) = \gamma_0 + \gamma_1 \beta_{jm}$), obtenemos:

$$\gamma_0 + \gamma_1 \beta_{jm} = \alpha_{jm} + \beta_{jm} E(R_m) \tag{1.42}$$

Bajo el CAPM, $E(R_m) = \gamma_0 + \gamma_1$ dado que $\beta_{mm} = 1$, por lo que podemos escribir la ecuación (1.42) como:

$$\gamma_0 + \gamma_1 \beta_{jm} = \alpha_{jm} + \beta_{jm}(\gamma_0 + \gamma_1) \tag{1.43}$$

Por tanto, bajo la hipótesis nula de que el CAPM se satisface, la ordenada en el origen del modelo de mercado debe ser igual a:

$$\alpha_{jm} = \gamma_0 (1 - \beta_{jm}) \tag{1.44}$$

Este planteamiento sirve de base para un contraste de sección cruzada donde se regresa la ordenada en el origen del modelo de mercado en $(1 - \beta_{jm})$ [9]. Así pensemos en el siguiente modelo de regresión

$$\alpha_{jm} = \delta_1 + \delta_2 (1 - \beta_{jm}) + \zeta_j \qquad j = 1, \ldots, N$$

donde

$$\hat{\delta}_{2,MCO} = \frac{Cov((1 - \beta_{jm}), \alpha_{jm})}{Var(1 - \beta_{jm})}$$

si contrastamos $H_0 : \delta_2 = 0$ y no rechazamos la hipótesis nula para un nivel de significación dado, la variable $(1 - \beta_{jm})$ no explica nada de α_{jm} y no existe la relación $\alpha_{jm} = \gamma_0(1 - \beta_{jm})$ propuesta por el modelo.

Sin embargo, lo que queremos introducir son los contrastes de serie temporal. Para ello veamos una versión distinta del modelo de mercado. Regresamos los excesos de rendimiento de los activos sobre el exceso de rendimiento del mercado. Para cada activo j tendremos el modelo siguiente:

$$R_{jt} - r_t = a_j + b_{jm}(R_{mt} - r_t) + \epsilon_{jt} \qquad t = 1, \ldots, T \tag{1.45}$$

donde r_t es el rendimiento del activo seguro en el periodo t.

$$\hat{b}_{jm,MCO} = \frac{Cov((R_{jt} - r_t), (R_{mt} - r_t))}{Var(R_{mt} - r_t)}$$

dado que la variabilidad del tipo de interés del activo libre de riesgo a lo largo del tiempo en relación a la variabilidad de los rendimientos de los activos y del mercado es irrelevante se cumple que:

$$\hat{b}_{jm,MCO} = \frac{Cov((R_{jt} - r_t), (R_{mt} - r_t))}{Var(R_{mt} - r_t)} = \frac{Cov(R_{jt}, R_{mt})}{Var(R_{mt})} = \hat{\beta}_{jm} \tag{1.46}$$

[9] Esta es la especificación empleada por Black, Jensen y Scholes (1972) en un contraste que ha tenido gran influencia histórica.

Es decir, el coeficiente $\hat{b}$ de la regresión (1.45) será igual (a todos los efectos) al coeficiente beta estimado. Además se cumple:

$$\hat{a}_j = (\bar{R}_j - \bar{r}) - \hat{\beta}_{jm}(\bar{R}_m - \bar{r})$$

donde $\bar{R}_j$ es el rendimiento medio del activo j durante el periodo muestral en el que se ha llevado a cabo la regresión y $\bar{r}$ es el rendimiento medio del activo seguro durante el mismo periodo. A este coeficiente se le denomina **alfa de Jensen**.

- ¿Cuál es el alfa de Jensen, para cualquier activo j, si el CAPM fuera el modelo correcto?
  Si el CAPM se cumple, para cada periodo tenemos que:

  $$E(R_j) - r = \hat{\beta}_{jm}[E(R_m) - r]$$

  Si la media muestral refleja, en promedio, lo que se esperaba en el mercado se satisface que $E(R_m) = \bar{R}_m$ y $E(R_j) = \bar{R}_j$ de donde:

  $$(\bar{R}_j - \bar{r}) = \hat{\beta}_{jm}(\bar{R}_m - \bar{r})$$

  $$(\bar{R}_j - \bar{r}) - \hat{\beta}_{jm}(\bar{R}_m - \bar{r}) = 0$$

  y por tanto el CAPM predice que el alfa de Jensen debe ser cero para todos los activos.
  El correspondiente contraste se llevará a cabo en la sección 6.

- ¿Cómo interpretamos el alfa de Jensen? Dado lo anterior, podemos interpretar éste coeficiente como la diferencia entre el rendimiento medio de un activo sobre el periodo muestral y el rendimiento que predice el CAPM dado el riesgo del activo. Puede ser visto como una medida del rendimiento ajustado por el riesgo beta. Si el alfa estimado por la regresión (1.45) es estadísticamente mayor que cero, el activo tiene un rendimiento superior al sugerido por su riesgo beta. Si el alfa estimado por la regresión (1.45) es estadísticamente menor que cero, el activo tiene un rendimiento inferior al sugerido por su riesgo beta. Por tanto, bajo la hipótesis nula de que el CAPM se satisface, el coeficiente alfa de Jensen permite determinar si un activo, inversión o cartera es en general positiva o negativa, dado su riesgo. El alfa de Jensen es útil como una medida del rendimiento ajustado por el riesgo beta.

1.5.2. Contrastes de series temporales para la eficiencia en media varianza de una determinada cartera. Estadísticos alternativos

Hemos visto que si el CAPM se cumple el alfa de Jensen debe ser cero para todos los activos. Ahora queremos diseñar un contraste conjunto para contrastar que el término independiente en la regresión (1.45) es cero para todos los activos a la vez[10]:

[10] El subíndice m de α_{jm} y β_{jm} indica que el valor de las estimaciones de los parámetros α_{jm} y β_{jm} dependen de la variable utilizada para aproximar la verdadera cartera de mercado.

$$R_{jt} - r_t = \alpha_{jm} + \beta_{jm}(R_{mt} - r_t) + \epsilon_{jt}$$

Gibbons, Ross y Shanken (1989) propusieron el siguiente estadístico de eficiencia de una cartera:

$$W = \frac{\hat{\alpha}'_m \hat{\Sigma}^{-1} \hat{\alpha}_m}{(1 + \hat{\theta}^2_m)} \tag{1.47}$$

donde

$\hat{\alpha}_m$ es un vector $N \times 1$ de estimadores de la constante.

$\hat{\Sigma}$ es el estimador de la matriz de varianzas y covarianzas de los residuos de la regresión en el modelo de mercado con T en el denominador de cada uno de sus componentes.

$\hat{\theta}_m = \frac{\bar{r}_m}{S_m}$ siendo $\bar{r}_m = \bar{R}_m - \bar{r}$ el exceso de rendimiento medio muestral de la cartera de mercado y S_m su desviación estándar con T en el denominador (en lugar de T-1).

Puede demostrarse que bajo normalidad de los rendimientos podemos contrastar:

$$H_0 : \alpha_{jm} = 0 \qquad j = 1, \ldots, N$$

con el estadístico

$$\frac{T(T-N-1)}{NT} \frac{\hat{\alpha}'_m \hat{\Sigma}^{-1} \hat{\alpha}_m}{(1 + \hat{\theta}^2_m)} \overset{H_0}{\sim} F_{(N,T-N-1)}$$

La utilización de este procedimiento evita el problema de errores en variables que aparece en el contexto de sección cruzada.

Existen estadísticos alternativos al anterior que son transformaciones monótonas del estadístico W con distribución asintótica $\chi^2_{(N)}$ como son:

Estadístico de Wald	$Wald = TW$
Razón de verosimilitudes	$LRT = Tln(1 + W)$
Contraste de multiplicador de Lagrange	$LMT = \frac{TW}{(1+W)}$

La utilización alternativa de estos estadísticos implica tener cuidado con los p-value de los mismos, por ejemplo para Wald el p-value es siempre menor mientras que para LRT es siempre mayor.

1.6. Contrastes de series temporales para la eficiencia en media varianza de una determinada cartera. Formalización

A continuación vamos a formalizar los estadísticos de contraste anteriores, para ello realizamos los siguientes supuestos que definen el marco de trabajo:

Disponemos de series de rendimientos para N activos (o carteras) y para la cartera de mercado y los denotamos respectivamente $\{R_{jt}\}_{t=1}^{T} \quad j = 1, \ldots, N$, $\{R_{mt}\}_{t=1}^{T}$. Consideramos que cada serie es un proceso estocástico de forma que las variables aleatorias que lo componen son iid. En cada periodo t la distribución de la variable de dimensión N para R_{jt} y de dimensión 1 para R_{mt} es normal multivariante. Además admitimos que los rendimientos de los activos estén correlados y sean heterocedásticos. Por tanto la matriz de covarianzas de la normal multivariante es del tipo general.

1.6.1. Contrastes en la versión de Sharpe y Lintner

En este caso se puede prestar y pedir prestado a la tasa de interés de un activo libre de riesgo. Por tanto hemos de añadir el supuesto de que existe un activo libre de riesgo cuya tasa de rendimiento está disponible y denotamos por: R_{ft}. Vamos a trabajar con excesos de rendimiento sobre el activo libre de riesgo, así que definimos:

$$Z_{jt} = R_{jt} - R_{ft}$$

$$Z_{mt} = R_{mt} - R_{ft}$$

El CAPM de Sharpe y Lintner es:

$$E(Z_j) = \beta_{jm} E(Z_m) \tag{1.48}$$

donde $\beta_{jm} = \frac{Cov(Z_j, Z_m)}{Var(Z_m)}$

Definimos $\boldsymbol{Z}_t$ como un vector de tamaño $(N \times 1)$ de excesos de rendimiento para los N activos. El exceso de rendimiento para estos N activos queda descrito con el modelo de mercado en términos de excesos de rendimiento.
Recordemos que $\forall j$ el modelo de mercado en términos de excesos de rendimiento lo definimos como (suprimimos el subíndice m, por conveniencia, en α y β):

$$R_{jt} - R_{ft} = \alpha_j + \beta_j [R_{mt} - R_{ft}] + \epsilon_{jt} \qquad t = 1, \ldots, T$$

en la notación actual:

$$\forall j \qquad Z_{jt} = \alpha_j + \beta_j Z_{mt} + \epsilon_{jt} \qquad t = 1, \ldots, T$$

En forma matricial:

$$\boldsymbol{Z}_t = \boldsymbol{\alpha} + \boldsymbol{\beta} Z_{mt} + \boldsymbol{\epsilon}_t \tag{1.49}$$

y se cumple:

$$E(\boldsymbol{\epsilon}_t) = 0$$

$$E(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t') = \boldsymbol{\Sigma}$$

$$E(Z_{mt}) = \mu_m, \qquad E[(Z_{mt} - \mu_m)^2] = \sigma_m^2$$

$$Cov(Z_{mt}, \boldsymbol{\epsilon}_t) = 0$$

donde definimos

$\boldsymbol{\beta}$ es el vector de orden $(N \times 1)$ de betas.

$\boldsymbol{\alpha}$ es el vector de orden $(N \times 1)$ de términos independientes.

Z_{mt} es el exceso de rendimiento de la cartera de mercado en el periodo t.

$\boldsymbol{\epsilon}_t$ es el vector de orden $(N \times 1)$ de perturbaciones.

Además por conveniencia vamos a denotar o redefinir como $\boldsymbol{\mu}$ al rendimiento en exceso esperado.
Objetivo: Queremos contrastar la implicación del CAPM en la versión de Sharpe y Lintner, la implicación es que los términos independientes de todos los activos son cero, $H_0 : \boldsymbol{\alpha} = 0$. Si esta hipótesis se acepta la cartera de mercado es eficiente en el sentido media y varianza. Para contrastar primero debemos estimar el modelo. Supondremos normalidad. En primer lugar vamos a estimar por Máxima Verosimilitud el modelo no restringido, es decir sin imponer la hipótesis nula. Los estimadores maximoverosímiles son consistentes, asintóticamente eficientes y asintóticamente normales, por lo que podemos construir estadísticos de contraste válidos a partir de sus distribuciones. Notar que bajo los supuestos que trabajamos estos estimadores son equivalentes a los obtenidos por MCO en el modelo periodo a periodo.

Dado que asumimos normalidad conjunta de los rendimientos en exceso en cada periodo, la función de densidad conjunta de los excesos de rendimiento condicionada a la muestra es decir, condicionada a los excesos de rendimientos del mercado, es:

$$\begin{aligned} f(\boldsymbol{Z}_t|Z_{mt}) &= (2\pi)^{-\frac{N}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}} \times \\ & exp\left[-\frac{1}{2}(\boldsymbol{Z}_t - \boldsymbol{\alpha} - \boldsymbol{\beta}Z_{mt})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{Z}_t - \boldsymbol{\alpha} - \boldsymbol{\beta}Z_{mt})\right] \end{aligned} \tag{1.50}$$

Nosotros lo que tenemos son series temporales de excesos de rendimientos para N activos, $(\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_T)$, y para el mercado, $(Z_{m1}, Z_{m2}, \ldots, Z_{mT})$ por ello necesitamos la función de densidad conjunta $f(\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_T|Z_{m1}, Z_{m2}, \ldots, Z_{mT})$.

Suponemos a los excesos de rendimiento temporalmente IID, independientes y normales, por lo que la función de densidad conjunta condicionada es el producto de las marginales condicionadas, y por tanto:

$$f(\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_T|Z_{m1}, Z_{m2}, \ldots, Z_{mT})$$

$$= \prod_{t=1}^{T} f(\boldsymbol{Z}_t|Z_{mt}) \tag{1.51}$$

$$\begin{aligned} &= \prod_{t=1}^{T}(2\pi)^{-\frac{N}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}} \times \\ & exp\left[-\frac{1}{2}(\boldsymbol{Z}_t - \boldsymbol{\alpha} - \boldsymbol{\beta}Z_{mt})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{Z}_t - \boldsymbol{\alpha} - \boldsymbol{\beta}Z_{mt})\right] \end{aligned} \tag{1.52}$$

Tomamos logarítmos en la función de verosimilitud y la función a maximizar es:

$$
\begin{aligned}
L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}) &= -\frac{NT}{2} log(2\pi) - \frac{T}{2} log|\boldsymbol{\Sigma}| \\
&\quad -\frac{1}{2} \sum_{t=1}^{T} (\boldsymbol{Z}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} Z_{mt})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{Z}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} Z_{mt})
\end{aligned} \tag{1.53}
$$

Para encontrar los estimadores de $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}$ debemos diferenciar la función con respecto a estos parámetros e igualar a cero estas ecuaciones. Las derivadas parciales son:

$$
\frac{\partial L}{\partial \boldsymbol{\alpha}} = \boldsymbol{\Sigma}^{-1} \left[\sum_{t=1}^{T} (\boldsymbol{Z}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} Z_{mt}) \right] \tag{1.54}
$$

$$
\frac{\partial L}{\partial \boldsymbol{\beta}} = \boldsymbol{\Sigma}^{-1} \left[\sum_{t=1}^{T} (\boldsymbol{Z}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} Z_{mt}) Z_{mt} \right] \tag{1.55}
$$

$$
\begin{aligned}
\frac{\partial L}{\partial \boldsymbol{\Sigma}} &= -\frac{T}{2} \boldsymbol{\Sigma}^{-1} \\
&\quad +\frac{1}{2} \boldsymbol{\Sigma}^{-1} \left[\sum_{t=1}^{T} (\boldsymbol{Z}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} Z_{mt})(\boldsymbol{Z}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} Z_{mt})' \right] \boldsymbol{\Sigma}^{-1}
\end{aligned} \tag{1.56}
$$

Y los estimadores maximoverosímiles son:

$$
\hat{\boldsymbol{\alpha}} = \hat{\boldsymbol{\mu}} - \hat{\boldsymbol{\beta}} \hat{\mu}_m \tag{1.57}
$$

$$
\hat{\boldsymbol{\beta}} = \frac{\sum_{t=1}^{T} (\boldsymbol{Z}_t - \hat{\boldsymbol{\mu}})(Z_{mt} - \hat{\mu}_m)}{\sum_{t=1}^{T} (Z_{mt} - \hat{\mu}_m)^2} \tag{1.58}
$$

$$
\hat{\boldsymbol{\Sigma}} = \frac{1}{T} \sum_{t=1}^{T} (\boldsymbol{Z}_t - \hat{\boldsymbol{\alpha}} - \hat{\boldsymbol{\beta}} Z_{mt})(\boldsymbol{Z}_t - \hat{\boldsymbol{\alpha}} - \hat{\boldsymbol{\beta}} Z_{mt})' \tag{1.59}
$$

donde

$$
\hat{\boldsymbol{\mu}} = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{Z}_t \qquad \hat{\mu}_m = \frac{1}{T} \sum_{t=1}^{T} Z_{mt}
$$

Las distribuciones condicionadas (en el exceso de rendimiento de mercado) que siguen estos estimadores son:

$$
\hat{\boldsymbol{\alpha}} \sim N \left(\boldsymbol{\alpha}, \frac{1}{T} \left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2} \right] \boldsymbol{\Sigma} \right) \tag{1.60}
$$

$$
\hat{\boldsymbol{\beta}} \sim N \left(\boldsymbol{\beta}, \frac{1}{T} \left[\frac{1}{\hat{\sigma}_m^2} \right] \boldsymbol{\Sigma} \right) \tag{1.61}
$$

$$
T \hat{\boldsymbol{\Sigma}} \sim \boldsymbol{W}_N (T - 2, \boldsymbol{\Sigma}) \tag{1.62}
$$

donde

$\hat{\sigma}_m^2 = \frac{1}{T}\sum_{t=1}^{T}(Z_{mt} - \hat{\mu}_m)^2$

$\boldsymbol{W}_N(T-2, \boldsymbol{\Sigma})$ indica distribución de Whishart con $(T-2)$ grados de libertad y matriz de covarianzas $\boldsymbol{\Sigma}$. Esta distribución es una generalización multivariante de la χ^2. (Ver Anderson (1984) y Muirhead (1983) para sus propiedades).

La covarianza entre $\hat{\boldsymbol{\alpha}}$ y $\hat{\boldsymbol{\beta}}$ es:

$$Cov(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}') = -\frac{1}{T}\left[\frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]\boldsymbol{\Sigma} \tag{1.63}$$

y $\hat{\boldsymbol{\Sigma}}$ es independiente de $\hat{\boldsymbol{\alpha}}$ y $\hat{\boldsymbol{\beta}}$.

Contrastes

Vamos a contrastar:

$$H_0 : \boldsymbol{\alpha} = 0$$

$$H_a : \boldsymbol{\alpha} \neq 0$$

Hay diferentes estadísticos de contraste, derivaremos sólo alguno de ellos.

• Contraste de Wald

Podemos contrastar la hipótesis anterior con el estadístico de Wald:

$$\begin{aligned} W &= \hat{\boldsymbol{\alpha}}' \left[Var(\hat{\boldsymbol{\alpha}})\right]^{-1} \hat{\boldsymbol{\alpha}} \\ &= T\left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1} \hat{\boldsymbol{\alpha}}'\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\alpha}} \overset{H_o}{\sim} \chi_N^2 \end{aligned} \tag{1.64}$$

Observaciones:

1. Un test de tipo Wald utiliza el hecho que si X es v.a n-variante tal que $X \sim N(0,1)$, entonces $X'X \sim \chi_N^2$

2. En la expresión del estadístico W la matriz $\boldsymbol{\Sigma}$ es desconocida y debe ser sustituida por su estimador consistente. Para ello podemos utilizar su estimador maximoverosímil que es consistente y viene dado por la ecuación (1.59). La utilización de un estimador para $\boldsymbol{\Sigma}$ hace que la distribución del estadístico sea una distribución asintótica. Así
$$W \xrightarrow{d,H_0} \chi_N^2$$

• Contraste de Gibbons, Ross y Shanken (1989)

En (1989) Gibbons, Ross y Shanken derivaron un estadístico de contraste alternativo al test de Wald sin necesidad de utilizar la teoría asintótica. Este estadístico se basaba en el siguiente resultado de Muirhead (1983):

Teorema: Sea $\boldsymbol{x}$ un vector de dimensión m con distribución $N(0, \boldsymbol{\Omega})$, sea $\boldsymbol{A}$ una matriz de orden $(m \times m)$ que sigue una distribución de Wishart $\boldsymbol{W}_m(n, \boldsymbol{\Omega})$, (con $n \geq m$). Si $\boldsymbol{A}$ y $\boldsymbol{x}$ son independientes se cumple que

$$\frac{(n-m+1)}{m}\boldsymbol{x}'\boldsymbol{A}^{-1}\boldsymbol{x} \sim F_{(m,n-m+1)}$$

Aplicando este resultado a:

$$\boldsymbol{x} = \sqrt{T}\left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1/2} \hat{\boldsymbol{\alpha}}$$

$$\boldsymbol{A} = T\hat{\boldsymbol{\Sigma}}$$

$$m = N$$

$$n = (T-2)$$

Tenemos el siguiente estadístico:

$$GRS = \frac{(T-N-1)}{N}\left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1} \hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}} \overset{H_0}{\sim} F_{(N,T-N-1)} \tag{1.65}$$

• **Contraste de Razón de Verosimilitudes**
Para formar el estadístico de Razón de Verosimilitudes debemos obtener los estimadores en el modelo restringido. Es decir debemos estimar (1.49) imponiendo la restricción $\boldsymbol{\alpha} = 0$. El modelo a estimar en este caso es:

$$\boldsymbol{Z}_t = \boldsymbol{\beta} Z_{mt} + \boldsymbol{\epsilon}_t$$

con lo que

$$\hat{\boldsymbol{\beta}}^\star = \frac{\sum_{t=1}^T \boldsymbol{Z}_t Z_{mt}}{\sum_{t=1}^T Z_{mt}^2} \tag{1.66}$$

$$\hat{\boldsymbol{\Sigma}}^\star = \frac{1}{T}\sum_{t=1}^T (\boldsymbol{Z}_t - \hat{\boldsymbol{\beta}}^\star Z_{mt})(\boldsymbol{Z}_t - \hat{\boldsymbol{\beta}}^\star Z_{mt})' \tag{1.67}$$

Bajo la H_o las distribuciones de estos estimadores son:

$$\hat{\boldsymbol{\beta}}^\star \sim N\left(\boldsymbol{\beta}, \frac{1}{T}\left[\frac{1}{\hat{\mu}_m^2 + \hat{\sigma}_m^2}\right]\boldsymbol{\Sigma}\right) \tag{1.68}$$

$$T\hat{\boldsymbol{\Sigma}}^\star \sim \boldsymbol{W}_N(T-1, \boldsymbol{\Sigma}) \tag{1.69}$$

El test de razón de verosimilitudes se define:

$$RV = -2LR$$

donde LR es la razón del logaritmo de verosimilitudes y ésta es igual a la diferencia entre el valor del logaritmo de la función de verosimilitud en el modelo restringido y el valor del logaritmo de la función de verosimilitud en el modelo no restringido, $LR = L^\star - L$, evaluado en los estimadores máximo verosímiles. Así:

$$LR = L^\star - L = -\frac{T}{2}\left[log|\hat{\boldsymbol{\Sigma}}^\star| - log|\hat{\boldsymbol{\Sigma}}|\right]$$

Así el estadístico de Razón de Verosimilitudes es:

$$RV = -2LR = T\left[log|\hat{\boldsymbol{\Sigma}}^\star| - log|\hat{\boldsymbol{\Sigma}}|\right] \overset{a}{\sim} \chi^2_N \tag{1.70}$$

Además se puede demostrar que los estadísticos de contraste GRS y RV se relacionan de la siguiente manera:

$$GRS = \frac{(T-N-1)}{N}\left(exp\left[\frac{RV}{T}\right] - 1\right) \tag{1.71}$$

con lo cual no es necesario recurrir a la teoría asintótica para poder aplicar el test RV.

En (1982) Jobson y Korkie refinan el estadístico RV para mejorar las propiedades en muestras finitas. Este estadístico es:

$$\begin{aligned} JK &= \frac{(T - \frac{N}{2} - 2)}{T} RV \\ &= (T - \frac{N}{2} - 2)\left[log|\hat{\boldsymbol{\Sigma}}^\star| - log|\hat{\boldsymbol{\Sigma}}|\right] \overset{a}{\sim} \chi^2_N \end{aligned} \tag{1.72}$$

1.6.2. Versión de Black

La versión del CAPM de Black no permite la existencia del activo libre de riesgo. El rendimiento esperado de la cartera cero-beta $E(R_{0m})$ se trata como una variable inobservable y por lo tanto es un parámetro desconocido en el modelo. Se denota por γ. El modelo se especifica:

$$\begin{aligned} E(\boldsymbol{R}_t) &= \boldsymbol{\iota}\gamma + \boldsymbol{\beta}(E(R_{mt}) - \gamma) \\ &= (\boldsymbol{\iota} - \boldsymbol{\beta})\gamma + \boldsymbol{\beta}E(R_{mt}) \end{aligned} \tag{1.73}$$

En el modelo de Black el modelo no restringido es el modelo de mercado en términos de rendimientos reales. Definimos $\boldsymbol{R}_t$ como un vector de tamaño $(N \times 1)$ de rendimientos reales para los N activos. El modelo de mercado en términos de rendimientos reales para los N activos en forma matricial es:

$$\boldsymbol{R}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}R_{mt} + \boldsymbol{\epsilon}_t \tag{1.74}$$

y se cumple:

$$E(\boldsymbol{\epsilon}_t) = 0$$
$$E(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t') = \boldsymbol{\Sigma}$$
$$E(R_{mt}) = \mu_m, \qquad E[(R_{mt} - \mu_m)^2] = \sigma^2_m$$
$$Cov(R_{mt}, \boldsymbol{\epsilon}_t) = 0$$

donde definimos

$\boldsymbol{\beta}$ es el vector de orden $(N \times 1)$ de betas de los activos.

$\boldsymbol{\alpha}$ es el vector de orden $(N \times 1)$ de términos independientes.

$\boldsymbol{\epsilon}_t$ es el vector de orden $(N \times 1)$ de perturbaciones.

La implicación a contrastar en la versión de Black es:

$$\boldsymbol{\alpha} = (\boldsymbol{\iota} - \boldsymbol{\beta})\gamma \tag{1.75}$$

Contrastar esta implicación es complicado dada la relación no lineal ente $\boldsymbol{\beta}$ y γ.

Dado el supuesto de rendimientos IID y conjuntamente normales la versión de Black puede ser estimada y contrastada mediante el estimador Máximo Verosímil. Los estimadores maximoverosímiles del modelo no restringido, el dado por la ecuación (1.74) son idénticos a los estimadores en el modelo de mercado en términos de excesos de rendimientos excepto porque sustituimos el exceso de rendimiento por el rendimiento real. Ahora $\hat{\boldsymbol{\mu}}$ es el vector de medias muestrales de los rendimientos reales. Y los estimadores maximoverosímiles son:

$$\hat{\boldsymbol{\alpha}} = \hat{\boldsymbol{\mu}} - \hat{\boldsymbol{\beta}}\hat{\mu}_m \tag{1.76}$$

$$\hat{\boldsymbol{\beta}} = \frac{\sum_{t=1}^{T}(\boldsymbol{R}_t - \hat{\boldsymbol{\mu}})(R_{mt} - \hat{\mu}_m)}{\sum_{t=1}^{T}(R_{mt} - \hat{\mu}_m)^2} \tag{1.77}$$

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{R}_t - \hat{\boldsymbol{\alpha}} - \hat{\boldsymbol{\beta}}R_{mt})(\boldsymbol{R}_t - \hat{\boldsymbol{\alpha}} - \hat{\boldsymbol{\beta}}R_{mt})' \tag{1.78}$$

donde

$$\hat{\boldsymbol{\mu}} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{R}_t \qquad \hat{\mu}_m = \frac{1}{T}\sum_{t=1}^{T}R_{mt}$$

Las distribuciones condicionadas en el rendimiento real del mercado que siguen estos estimadores son:

$$\hat{\boldsymbol{\alpha}} \sim N\left(\boldsymbol{\alpha}, \frac{1}{T}\left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]\boldsymbol{\Sigma}\right) \tag{1.79}$$

$$\hat{\boldsymbol{\beta}} \sim N\left(\boldsymbol{\beta}, \frac{1}{T}\left[\frac{1}{\hat{\sigma}_m^2}\right]\boldsymbol{\Sigma}\right) \tag{1.80}$$

$$T\hat{\boldsymbol{\Sigma}} \sim \boldsymbol{W}_N(T-2, \boldsymbol{\Sigma}) \tag{1.81}$$

donde

$$\hat{\sigma}_m^2 = \frac{1}{T}\sum_{t=1}^{T}(R_{mt} - \hat{\mu}_m)^2$$

La covarianza entre $\hat{\boldsymbol{\alpha}}$ y $\hat{\boldsymbol{\beta}}$ es:

$$Cov(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}') = -\left[\frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]\boldsymbol{\Sigma} \tag{1.82}$$

para el modelo restringido el logarítmo de la función de verosimilitud es:

$$\begin{aligned} L(\gamma, \boldsymbol{\beta}, \boldsymbol{\Sigma}) &= -\frac{NT}{2}log(2\pi) - \frac{T}{2}log|\boldsymbol{\Sigma}| \\ & -\frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{R}_t - \gamma(\boldsymbol{\iota} - \boldsymbol{\beta}) - \boldsymbol{\beta}R_{mt})'\boldsymbol{\Sigma}^{-1} \\ & \times(\boldsymbol{R}_t - \gamma(\boldsymbol{\iota} - \boldsymbol{\beta}) - \boldsymbol{\beta}R_{mt}) \end{aligned} \tag{1.83}$$

Las derivadas parciales respecto de $\gamma, \boldsymbol{\beta}$ y $\boldsymbol{\Sigma}$ son:

$$\frac{\partial L}{\partial \gamma} = (\boldsymbol{\iota} - \boldsymbol{\beta})'\boldsymbol{\Sigma}^{-1}\left[\sum_{t=1}^{T}(\boldsymbol{R}_t - \gamma(\boldsymbol{\iota} - \boldsymbol{\beta}) - \boldsymbol{\beta}R_{mt})\right] \tag{1.84}$$

$$\frac{\partial L}{\partial \boldsymbol{\beta}} = \boldsymbol{\Sigma}^{-1}\left[\sum_{t=1}^{T}(\boldsymbol{R}_t - \gamma(\boldsymbol{\iota} - \boldsymbol{\beta}) - \boldsymbol{\beta}R_{mt})(R_{mt} - \gamma)\right] \tag{1.85}$$

$$\begin{aligned} \frac{\partial L}{\partial \boldsymbol{\Sigma}} &= -\frac{T}{2}\boldsymbol{\Sigma}^{-1} + \frac{1}{2}\boldsymbol{\Sigma}^{-1} \\ & \left[\sum_{t=1}^{T}(\boldsymbol{R}_t - \gamma(\boldsymbol{\iota} - \boldsymbol{\beta}) - \boldsymbol{\beta}R_{mt})(\boldsymbol{R}_t - \gamma(\boldsymbol{\iota} - \boldsymbol{\beta}) - \boldsymbol{\beta}R_{mt})'\right]\boldsymbol{\Sigma}^{-1} \end{aligned} \tag{1.86}$$

Y los estimadores restringidos son:

$$\hat{\gamma}^{\star} = \frac{(\boldsymbol{\iota} - \hat{\boldsymbol{\beta}}^{\star})'\hat{\boldsymbol{\Sigma}}^{\star -1}(\hat{\boldsymbol{\mu}} - \hat{\boldsymbol{\beta}}^{\star}\hat{\mu}_m)}{(\boldsymbol{\iota} - \hat{\boldsymbol{\beta}}^{\star})'\hat{\boldsymbol{\Sigma}}^{\star -1}(\boldsymbol{\iota} - \hat{\boldsymbol{\beta}}^{\star})} \tag{1.87}$$

$$\hat{\boldsymbol{\beta}}^{\star} = \frac{\sum_{t=1}^{T}(\boldsymbol{R}_t - \hat{\gamma}^{\star}\boldsymbol{\iota})(R_{mt} - \hat{\gamma}^{\star})}{\sum_{t=1}^{T}(R_{mt} - \hat{\gamma}^{\star})^2} \tag{1.88}$$

$$\hat{\boldsymbol{\Sigma}}^{\star} = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{R}_t - \hat{\gamma}^{\star}(\boldsymbol{\iota} - \hat{\boldsymbol{\beta}}^{\star}) - \hat{\boldsymbol{\beta}}^{\star}R_{mt})(\boldsymbol{R}_t - \hat{\gamma}^{\star}(\boldsymbol{\iota} - \hat{\boldsymbol{\beta}}^{\star}) - \hat{\boldsymbol{\beta}}^{\star}R_{mt})' \tag{1.89}$$

La solución de las ecuaciones (1.87), (1.88) y (1.89) se obtiene mediante un procedimiento iterativo que puede iniciarse en los estimadores bajo el modelo no restringido y que finaliza cuando alcanzamos un grado de convergencia, prefijado de antemano, en los estimadores.

Debemos contrastar la siguiente hipótesis:

$$H_0 : \boldsymbol{\alpha} = (\boldsymbol{\iota} - \boldsymbol{\beta})\gamma$$

$$H_0 : \boldsymbol{\alpha} \neq (\boldsymbol{\iota} - \boldsymbol{\beta})\gamma$$

Como estadístico de contraste se puede construir el contraste de Razón de Verosimilitudes tal que:

$$RV_B = T\left[log|\hat{\boldsymbol{\Sigma}}^{\star}| - log|\hat{\boldsymbol{\Sigma}}|\right] \overset{a}{\sim} \chi^2_{N-1} \tag{1.90}$$

Observaciones:

1. Los grados de libertad son ahora $N-1$ ya que estimamos el parámetro γ.

2. Con respecto a los parámetros a estimar en el modelo tenemos:

   - $N(N-1)/2$ parámetros en la matriz de covarianzas de los residuos.
   - En el modelo no restringido tenemos N parámetros en $\boldsymbol{\alpha}$ y N parámetros en $\boldsymbol{\beta}$ más los correspondientes $N(N-1)/2$ parámetros en la matriz de covarianzas de los residuos.
   - En el modelo restringido tenemos N parámetros en $\boldsymbol{\beta}$ más el parámetro γ y más los $N(N-1)/2$ parámetros en la matriz de covarianzas de los residuos.
   - En el modelo no restringido hay $N-1$ parámetros más que en el modelo restringido.

3. Si ajustamos el estadístico para mejorar las propiedades en muestras finitas podemos definirle como:

$$JK = (T - \frac{N}{2} - 2)\left[log|\hat{\boldsymbol{\Sigma}}^{\star}| - log|\hat{\boldsymbol{\Sigma}}|\right] \overset{a}{\sim} \chi^2_{N-1} \tag{1.91}$$

4. El test de contraste propuesto está basado en el comportamiento del estadístico en muestras grandes pero tiene pobres propiedades en muestras finitas. Trabajos de Kandel (1984) y Shanken (1986) muestran como refinar este estadístico para mejorar sus propiedades en muestras finitas.

1.6.3. Propiedades empíricas de los contrastes

Un contraste está sujeto a dos tipos de error:

a) Error Tipo I: En este caso el contraste nos lleva a rechazar la hipótesis nula cuando es cierta.

b) Error Tipo II: En este caso el contraste nos lleva a no rechazar la hipótesis nula siendo ésta falsa.

Definición: Tamaño del contraste
Llamamos tamaño del contraste a la probabilidad de incurrir en el error tipo I. generalmente se designa por α y se le llama nivel de significación.
Definición: Potencia del contraste
La potencia del contraste es la probabilidad de que el contraste nos conduzca correctamente a rechazar la hipótesis nula cuando es falsa. Por tanto Potencia =1-Prob(error tipo II).

- **Tamaño del contraste:**
  Con respecto a los estadísticos de contraste anteriores hay que hacer una seria advertencia con respecto al tamaño muestral para el cual se están aplicando. El contraste de Gibbons, Ross y Shanken tiene una distribución exacta en muestras finitas mientras que el contraste de Wald o la razón de verosimilitudes son test

asintóticos. Cuando el tamaño muestral no es suficientemente grande los resultados obtenidos con estadísticos asintóticos pueden no ser los adecuados y hay que tener cuidado con ello. Es necesario que comprobemos para un nivel de significación $\alpha(\alpha = 5\,\%)$ el porcentaje de veces que se rechaza la hipótesis nula, cuando es cierta.

- **Potencia del contraste:**
  La potencia de un estadístico de contraste es la probabilidad (porcentaje de veces) de rechazar la hipótesis nula dado que la alternativa es cierta. Un test con baja potencia sugiere que no es capaz de distinguir entre la hipótesis nula y la alternativa. El punto contrario, un test con mucha potencia, implica tener un estadístico de contraste muy informativo pero que puede rechazar la hipótesis nula contra alternativas muy cercanas a la misma desde el punto de vista económico. Podemos pensar en rechazos de la nula cuando se producen desviaciones de la nula que no son importantes económicamente.
  Los estudios realizados muestran que la potencia de los estadísticos dependen tanto de la dimensión de sección cruzada N como de la dimensión de serie temporal T. Aumenta cuando aumenta T y también aumenta cuando se reduce N, con respecto a esta dimensión parece que lo ideal es mantener N alrededor del valor 10.

1.7. Rendimientos no IID y no normales

Las secciones anteriores se han desarrollado bajo el supuesto de rendimientos temporalmente IID y normalidad conjunta. Sin el supuesto de normalidad es difícil demostrar las propiedades de los contrastes del modelo de valoración. En la literatura hay suficiente evidencia sobre la no normalidad de los rendimientos mensuales, ver Fama (1965, 1976), Blattberg y Gonedes (1974), Affleck-Graves y MacDonald (1989). Además numerosos estudios muestran que los rendimientos son heterocedásticos y temporalmente dependientes. Por tanto, es importante considerar los efectos de relajar estos supuestos. Para derivar estadísticos de contraste del CAPM robustos a estos supuestos podemos trabajar en el marco del Método Generalizado de Momentos.

Vamos a desarrollar el estimador GMM para los parámetros del modelo CAPM de Sharpe y Lintner. Vamos a suponer que disponemos de una muestra de serie temporal de tamaño T para N activos. Dentro del marco del método GMM la distribución de los rendimientos condicionada al rendimiento de mercado puede ser serialmente dependiente y condicionalmente heterocedástica. Sólo es necesario asumir que el rendimiento en exceso es estacionario y ergódico con momentos de orden cuarto finitos. El conjunto de condiciones de ortogonalidad es de orden $(2N \times 1)$, lo formamos con las N condiciones del vector de residuos. $E(\epsilon_t) = 0$ y las N condiciones del producto del exceso de rendimiento de mercado por los residuos, $E(Z_{mt}\epsilon_t) = 0$. Tenemos por tanto la función:

$$\boldsymbol{h}_t(\boldsymbol{\theta}) = \boldsymbol{z}_t \otimes \boldsymbol{\epsilon}_t$$

tal que:

$$\boldsymbol{z}_t' = \begin{bmatrix} 1 & Z_{mt} \end{bmatrix} \quad \boldsymbol{\theta}' = [\boldsymbol{\alpha}' \ \boldsymbol{\beta}'] \quad \boldsymbol{\epsilon}_t = \boldsymbol{Z}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} Z_{mt}$$

y por tanto podemos leer la condición anterior en forma matricial como:

$$\boldsymbol{h}_t(\boldsymbol{\theta}) = \boldsymbol{z}_t \otimes \boldsymbol{\epsilon}_t = \begin{bmatrix} \epsilon_{1t} \\ \vdots \\ \epsilon_{Nt} \\ Z_{mt}\epsilon_{1t} \\ \vdots \\ Z_{mt}\epsilon_{Nt} \end{bmatrix} = \begin{bmatrix} Z_{1t} - \alpha_1 - \beta_1 Z_{mt} \\ \vdots \\ Z_{Nt} - \alpha_N - \beta_N Z_{mt} \\ Z_{mt}(Z_{1t} - \alpha_1 - \beta_1 Z_{mt}) \\ \vdots \\ Z_{mt}(Z_{Nt} - \alpha_N - \beta_N Z_{mt}) \end{bmatrix}$$

La condición de los momentos es la siguiente:

$$E[\boldsymbol{h}_t(\boldsymbol{\theta}_0)] = 0$$

donde $\boldsymbol{\theta}_0$ es el vector de verdaderos parámetros. Definimos:

$$\begin{aligned} \boldsymbol{g}_T(\boldsymbol{\theta}) &= \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{h}_t(\boldsymbol{\theta}) \\ &= \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{z}_t \otimes \boldsymbol{\epsilon}_t \end{aligned}$$

El estimador GMM de $\boldsymbol{\theta}$ es aquel que minimiza la siguiente forma cuadrática:

$$\boldsymbol{Q}_T(\boldsymbol{\theta}) = \boldsymbol{g}_T(\boldsymbol{\theta})'\boldsymbol{W}^{-1}\boldsymbol{g}_T(\boldsymbol{\theta})$$

donde $\boldsymbol{W}$ es una matriz de pesos definida positiva de orden $(2N \times 2N)$.

Dado que tenemos $2N$ parámetros desconocidos y $2N$ restricciones de momentos el sistema está exactamente identificado y $\hat{\boldsymbol{\theta}}$ se elige de forma que las medias de los momentos muestrales de $\boldsymbol{g}_T(\boldsymbol{\theta})$ sean cero. Además en este caso el estimador GMM no depende de la matriz $\boldsymbol{W}$. Los estimadores GMM de los parámetros coinciden con los estimadores máximo verosímiles dados por (1.57) y (1.58) y que recordemos son.

$$\hat{\boldsymbol{\alpha}} = \hat{\boldsymbol{\mu}} - \hat{\boldsymbol{\beta}}\hat{\mu}_m$$

$$\hat{\boldsymbol{\beta}} = \frac{\sum_{t=1}^{T}(\boldsymbol{Z}_t - \hat{\boldsymbol{\mu}})(Z_{mt} - \hat{\mu}_m)}{\sum_{t=1}^{T}(Z_{mt} - \hat{\mu}_m)^2}$$

La importancia del método GMM consiste en que podemos encontrar una matriz de varianzas y covarianzas de los parámetros robusta a los supuestos del modelo. La matriz de varianzas y covarianzas asintótica de los parámetros estimados por el método GMM es:

$$\boldsymbol{V} = \left[\boldsymbol{G}_0'\boldsymbol{S}_0^{-1}\boldsymbol{G}_0\right]^{-1} \tag{1.92}$$

donde:

$$\boldsymbol{G}_0 = E\left[\frac{\partial \boldsymbol{g}_T(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right] \tag{1.93}$$

y

$$\boldsymbol{S}_0 = \sum_{t=-\infty}^{+\infty} E[\boldsymbol{h}_t(\boldsymbol{\theta})\boldsymbol{h}_{t-l}(\boldsymbol{\theta})'] \tag{1.94}$$

La distribución asintótica del estimador es:

$$\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \overset{a}{\sim} N(0, \boldsymbol{V}) \tag{1.95}$$

Para estimar consistentemente $\boldsymbol{V}$ necesitamos estimadores consistentes de $\boldsymbol{G}_0$ y de $\boldsymbol{S}_0$.

- En este caso $\boldsymbol{G}_0$ es:

  $$\boldsymbol{G}_0 = -\begin{bmatrix} 1 & \mu_m \\ \mu_m & (\sigma_m^2 + \mu_m^2) \end{bmatrix} \otimes I_N$$

  Podemos obtener un estimador consistente de $\boldsymbol{G}_0$, $\boldsymbol{G}_T(\hat{\boldsymbol{\theta}})$ utilizando los estimadores de máxima verosimilitud de μ_m y σ_m^2, ya que son consistentes.

- Para obtener un estimador consistente de $\boldsymbol{S}_0$ necesitamos reducir el sumatorio a un sumatorio finito. Un estimador consistente de $\boldsymbol{S}_0$ podemos obtenerlo estimando $\boldsymbol{W}$ consistentemente por Newey West (1987). El estimador consistente de $\boldsymbol{V}$ sería:

  $$\boldsymbol{V} = (1/T)\left[\hat{\boldsymbol{G}}_T(\hat{\boldsymbol{\theta}})'\widehat{\boldsymbol{W}}_T^{-1}\hat{\boldsymbol{G}}_T(\hat{\boldsymbol{\theta}})\right]^{-1}$$

Una vez obtenidos los estimadores consistentes $\boldsymbol{S}_T = \widehat{\boldsymbol{W}}_T$ y $\boldsymbol{G}_T(\hat{\boldsymbol{\theta}})$, $\widehat{\boldsymbol{V}}$ es un estimador consistente de $\boldsymbol{V}$.

Para contrastar

$$H_0 : \boldsymbol{\alpha} = 0$$

podemos utilizar el estadístico.

$$J_T = T\hat{\boldsymbol{\alpha}}'\left[\boldsymbol{R}\left[\boldsymbol{G}_T(\hat{\boldsymbol{\theta}})'\widehat{\boldsymbol{W}}_T^{-1}\boldsymbol{G}_T(\hat{\boldsymbol{\theta}})\right]^{-1}\boldsymbol{R}'\right]^{-1}\hat{\boldsymbol{\alpha}} \overset{a,H_0}{\sim} \chi_N^2 \tag{1.96}$$

donde:

$$\boldsymbol{R} = \begin{bmatrix} 1 & 0 \end{bmatrix} \otimes I_N$$

y $\frac{1}{T}\boldsymbol{R}\left[\boldsymbol{G}_T(\hat{\boldsymbol{\theta}})'\widehat{\boldsymbol{W}}_T^{-1}\boldsymbol{G}_T(\hat{\boldsymbol{\theta}})\right]^{-1}\boldsymbol{R}'$ es un estimador robusto de $Var(\hat{\boldsymbol{\alpha}})$

Tema 2

Modelos de Valoración Multifactoriales. APT

2.1. Introducción

Bajo el supuesto de ausencia de arbitraje los precios de los activos financieros deben ser tales que satisfagan la expresión:

$$E_t(M_{t+1}\tilde{R}_{t+1}) = 1 \qquad j = 1, \ldots, N \tag{2.1}$$

es decir, el rendimiento esperado ponderado por la variable agregada o factor de descuento, M, es constante e igual a 1 para todos los activos financieros. La ecuación (3.1) dice que todos los activos que prometen ofrecer los mismos pagos futuros deben tener hoy el mismo precio.

Esta expresión nos permite obtener una fórmula de valoración en términos de la prima de riesgo esperada de cualquier activo incierto j:

$$E(R_j - r) = \left[-\frac{1}{E(M)}\right] Cov(M, R_j) \tag{2.2}$$

Supuestos alternativos sobre la variable agregada M permiten obtener diferentes modelos de valoración. En el CAPM se supone que solo existe un factor de riesgo que es la cartera de mercado y se supone que la variable M es lineal en ella. Sin embargo la evidencia empírica sugiere que para caracterizar el comportamiento de los rendimientos en exceso es necesario más de un factor. En esta línea podemos distinguir dos aproximaciones teóricas, la teoría de valoración, APT de Ross (1976) basado en argumentos de arbitraje y el CAPM Intertemporal, ICAPM, de Merton (1973) basado en argumentos de equilibrio.

2.1.1. APT de Ross (1976)

El APT es un modelo más general que el CAPM ya que permite múltiples factores de riesgo. Además no requiere la identificación de la cartera de mercado. Supongamos que existen múltiples factores de riesgo sistemático que no puedan ser recogidos exclusivamente mediante la cartera de mercado. Suponiendo que existan K factores de riesgo sistemático podemos representar la variable agregada M como:

$$M = \delta_0 + \delta_1 f_1 + \ldots + \delta_K f_K \tag{2.3}$$

donde $f_1, f_2, \ldots, f_K$ son los factores de riesgo sistemático de la economía tal que

$$E[\tilde{R}_j(\delta_0 + \delta_1 f_1 + \ldots + \delta_K f_K)] = 1; \quad j = 1, \ldots, N \tag{2.4}$$

Dado que podemos escribir el proceso generador de los rendimientos como

$$R_j = E(R_j) + \text{innovación sistemática} + \underbrace{\epsilon_j}_{\text{innovación idiosincrásica}} \tag{2.5}$$

en términos de los K-factores de riesgo podemos escribirlo como:

$$R_j = E(R_j) + b_{j1} f_1 + b_{j2} f_2 + \ldots + b_{jK} f_K + \epsilon_j \tag{2.6}$$

En definitiva podemos escribir el proceso generador de rendimientos que suponemos que existe en los mercados como

$$R_j = a_j + b_{j1} f_1 + b_{j2} f_2 + \ldots + b_{jK} f_K + \epsilon_j \tag{2.7}$$

donde:

- $f_1, f_2, \ldots, f_K$, son los factores sistemáticos o agregados comunes a todos los activos existentes expresados como innovaciones por lo que tienen valor esperado cero y sus covarianzas entre dos factores cualesquiera también son cero. $E(f_K) = 0 \quad \forall j, K$
- $b_{j1}, b_{j2}, \ldots, b_{jK}$, son las sensibilidades de los rendimientos del activo j a los factores.
- $E(\epsilon_j f_K) = 0 \quad \forall j, K$, los factores de riesgo sistemático son considerados innovaciones y no están correlacionados con el componente idiosincrásico.
- El componente idiosincrásico es también una innovación. No están correlacionados entre sí para las diferentes empresas,
  $E(\epsilon_j) = E(\epsilon_j \epsilon_h) = 0 \quad \forall j, j \neq h$
- $a_j = E(R_j) \quad \forall j$ ya que los factores sistemáticos de riesgo y el componente idiosincrásico son innovaciones cuyo valor esperado es igual a cero.

Matricialmente podemos escribir la ecuación anterior en forma matricial para el activo i como:

$$R_i = a_i + \boldsymbol{b}_i' \boldsymbol{f} + \epsilon_i \tag{2.8}$$

donde

$$E(\epsilon_i | \boldsymbol{f}) = 0$$
$$E(\epsilon_i^2) = \sigma_i^2 \leq \sigma^2 < \infty$$

donde R_t es el rendimiento del activo i, a_i es el término independiente del modelo de factores, $\boldsymbol{b}_i$ es el vector $(K \times 1)$ de sensibilidades para el activo i, $\boldsymbol{f}$ es el vector $(K \times 1)$ de realizaciones de los factores comunes, ϵ_i son las perturbaciones. Para los N activos de la economía, podemos escribir el siguiente modelo matricial:

$$\boldsymbol{R} = \boldsymbol{a} + \boldsymbol{B}\boldsymbol{f} + \boldsymbol{\epsilon} \tag{2.9}$$

$$E[\boldsymbol{\epsilon}_t | \boldsymbol{f}] = 0 \tag{2.10}$$

$$E(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t' | \boldsymbol{f}] = \boldsymbol{\Sigma} \tag{2.11}$$

Siendo

$\boldsymbol{R}$ es un vector $(N \times 1)$ con $\boldsymbol{R} = [R_1 \; R_2 \ldots R_N]'$

$\boldsymbol{a}$ es un vector $(N \times 1)$ con $\boldsymbol{a} = [a_1 \; a_2 \ldots a_N]'$

$\boldsymbol{B}$ es un vector $(N \times K)$ con $\boldsymbol{B} = [\boldsymbol{b}_1 \, \boldsymbol{b}_2 \ldots \boldsymbol{b}_N]'$

$\boldsymbol{\epsilon}$ es un vector $(N \times 1)$ con $\boldsymbol{\epsilon} = [\epsilon_1 \; \boldsymbol{\epsilon}_2 \ldots \boldsymbol{\epsilon}_N]'$

El APT exacto viene dado por la siguiente expresión:

$$E(R_j) = \lambda_0 + \sum_{k=1}^{K} b_{jk} \lambda_k \qquad j = 1, \ldots, N \tag{2.12}$$

donde λ_k son las primas por riesgo esperadas de los factores de riesgo sistemático y representan los incrementos en el rendimiento esperado de los activos para cada unidad adicional de riesgo beta como sensibilidad al factor k.

Las implicaciones empíricas del APT son.

a) El rendimiento esperado de la cartera que tiene betas igual a cero con respecto a todos los factores de riesgo sistemático representa al activo que juega el papel del activo libre de riesgo como tal o simplemente como una cartera de cero-betas.

b) El rendimiento esperado de los activos aumenta linealmente con incrementos en una beta cualquiera dada, b_{jk}.

c) El rendimiento esperado de los activos no está determinado por ninguna otra característica de los activos que no sea una de las betas asociadas a alguno de los factores de riesgo sistemático.

2.1.2. ICAMP

El Modelo Intertemporal de Valoración de Activos con Cartera de Mercado, ICAPM, de Merton (1973) junto con supuestos en la distribución condicionada de los rendimientos desarrolla un modelo factorial. En este modelo la cartera de mercado sirve como factor y variables estado sirven como factores adicionales. Desde el punto de vista econométrico podemos estudiar ambos modelos donde tenemos factores de valoración exactos. Podemos escribir la expresión (2.12) en términos matriciales como:

$$\boldsymbol{\mu} = \boldsymbol{\iota}\lambda_0 + \boldsymbol{B}\boldsymbol{\lambda}_K \tag{2.13}$$

donde $\boldsymbol{\mu}$ es el vector $(N \times 1)$ de rendimientos esperados, λ_0 es el parámetro del modelo cero-beta y que coincide con el rendimiento del activo sin riesgo, si este existe. $\boldsymbol{\lambda}_K$ es el vector $(K \times 1)$ de premios al riesgo de los factores. Esta estructura lineal es común a todos los modelos.

2.2. Estimación y contraste

Para poder estimar y contrastar necesitamos hacer supuestos sobre la conducta de los rendimientos a lo largo del tiempo. Suponemos que los rendimientos condicionados en los factores son IID a través del tiempo y conjuntamente siguen una normal multivariante. Estos supuestos pueden ser relajados estimando por el Método Generalizado de Momentos. Vamos a distinguir cuatro versiones del modelo de valoración:

- Los factores son carteras de activos comercializadas y existe un activo libre de riesgo.
- Los factores son carteras de activos comercializadas y no existe un activo libre de riesgo.
- variables económicas como factores.
- Los factores son carteras de activos y estas carteras generan la frontera media-varianza de los activos.

Estimaremos por Maximaverosimilitud, dado el supuesto de normalidad conjunta para los rendimientos condicionados en los factores podemos construir test de contraste para cada uno de los casos anteriores. El mecanismo de contraste conlleva los siguientes pasos:

a) Estimar el modelo no restringido y estimar $\boldsymbol{\Sigma}$, la matriz de varianzas y covarianzas residual.

b) Estimar el modelo no restringido y estimar $\boldsymbol{\Sigma}^\star$, la matriz de varianzas y covarianzas residual del modelo restringido.

c) Contrastar la hipótesis nula correspondiente con el siguiente estadístico general:

$$J = -\left(T - \frac{N}{2} - K - 1\right)\left[log|\hat{\boldsymbol{\Sigma}}| - log|\hat{\boldsymbol{\Sigma}}^\star|\right] \overset{a}{\sim} \chi^2_{(N)} \tag{2.14}$$

donde N es el número de activos, T el número de observaciones de serie temporal y K el número de factores. El estadístico está reescalado por el término $\left[T - \frac{N}{2} - K - 1\right]$ en vez de la habitual T para mejorar la convergencia de la distribución en muestras finitas, bajo la nula, a la distribución en muestras grandes.

2.2.1. Los factores son carteras de activos comercializadas y existe un activo libre de riesgo

En este caso el modelo no restringido para K-factores se expresa en exceso de rendimientos. Sea $\boldsymbol{Z}_t$ un vector de tamaño $(N \times 1)$ de excesos de rendimiento para los N activos (o carteras de activos). Escribimos el modelo lineal de K-factores en forma matricial compacta como:

$$\boldsymbol{Z}_t = \boldsymbol{a} + \boldsymbol{B}\boldsymbol{Z}_{Kt} + \boldsymbol{\epsilon}_t \tag{2.15}$$

$$E(\boldsymbol{\epsilon}_t) = 0$$

$$E(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t') = \boldsymbol{\Sigma}$$

$$E(\boldsymbol{Z}_{Kt}) = \boldsymbol{\mu}_K, \qquad E[(\boldsymbol{Z}_{Kt} - \boldsymbol{\mu}_K)(\boldsymbol{Z}_{Kt} - \boldsymbol{\mu}_K)'] = \boldsymbol{\Omega}_K$$

$$Cov(\boldsymbol{Z}_{Kt}, \boldsymbol{\epsilon}_t') = \boldsymbol{0}$$

donde definimos

- $\boldsymbol{B}$ es el vector de orden $(N \times K)$ de sensibilidades a los factores.
- $\boldsymbol{a}$ es el vector de orden $(N \times 1)$ de términos independientes.
- $\boldsymbol{Z}_{Kt}$ es vector de orden $(K \times 1)$ de excesos de rendimiento de las carteras que hacen de factores y $\boldsymbol{\Omega}_K$ su matriz de varianzas y covarianzas.
- $\boldsymbol{\epsilon}_t$ es el vector de orden $(N \times 1)$ de perturbaciones y $\boldsymbol{\Sigma}$ su matriz de varianzas y covarianzas.
- $\boldsymbol{0}$ es una matriz de ceros $(K \times N)$

La valoración exacta de los factores implica que $\boldsymbol{a}$ es cero.

Los estimadores máximoverosímiles del modelo no restringido coinciden con los MCO y son:

$$\hat{\boldsymbol{a}} = \hat{\boldsymbol{\mu}} - \hat{\boldsymbol{B}}\hat{\boldsymbol{\mu}}_K \tag{2.16}$$

$$\hat{\boldsymbol{B}} = \left[\sum_{t=1}^{T}(\boldsymbol{Z}_t - \hat{\boldsymbol{\mu}})(\boldsymbol{Z}_{Kt} - \hat{\boldsymbol{\mu}}_K)'\right]\left[\sum_{t=1}^{T}(\boldsymbol{Z}_{Kt} - \hat{\boldsymbol{\mu}}_K)(\boldsymbol{Z}_{Kt} - \hat{\boldsymbol{\mu}}_K)'\right]^{-1} \tag{2.17}$$

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{Z}_t - \hat{\boldsymbol{a}} - \hat{\boldsymbol{B}}\boldsymbol{Z}_{Kt})(\boldsymbol{Z}_t - \hat{\boldsymbol{a}} - \hat{\boldsymbol{B}}\boldsymbol{Z}_{Kt})' \tag{2.18}$$

donde

$$\hat{\boldsymbol{\mu}} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{Z}_t \qquad \hat{\boldsymbol{\mu}}_K = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{Z}_{Kt}$$

El modelo restringido se obtiene imponiendo en (2.15) la restricción $\boldsymbol{a} = 0$. El modelo a estimar en este caso es:

$$\boldsymbol{Z}_t = \boldsymbol{B}\boldsymbol{Z}_{Kt} + \boldsymbol{\epsilon}_t$$

con lo que los correspondientes estimadores maximoverosímiles son:

$$\hat{\boldsymbol{B}}^{\star} = \left[\sum_{t=1}^{T} \boldsymbol{Z}_t \boldsymbol{Z}'_{Kt}\right] \left[\sum_{t=1}^{T} \boldsymbol{Z}_{Kt} \boldsymbol{Z}'_{Kt}\right]^{-1} \tag{2.19}$$

$$\hat{\boldsymbol{\Sigma}}^{\star} = \frac{1}{T} \sum_{t=1}^{T} (\boldsymbol{Z}_t - \hat{\boldsymbol{B}}^{\star} \boldsymbol{Z}_{Kt})(\boldsymbol{Z}_t - \hat{\boldsymbol{B}}^{\star} \boldsymbol{Z}_{Kt})' \tag{2.20}$$

Para contrastar $H_0 : \boldsymbol{a} = 0$ podemos utilizar el estadístico

$$J = -\left(T - \frac{N}{2} - K - 1\right) \left[log|\hat{\boldsymbol{\Sigma}}| - log|\hat{\boldsymbol{\Sigma}}^{\star}|\right] \overset{a}{\sim} \chi^2_{(N)} \tag{2.21}$$

donde $\hat{\boldsymbol{\Sigma}}$ y $\hat{\boldsymbol{\Sigma}}^{\star}$ son los estimadores maximoverosímiles de la matriz de varianzas y covarianzas de los residuos bajo el modelo no restringido y el restringido respectivamente. T es el número de observaciones de serie temporal, N es el número de activos y K el número de factores. Dado que estamos contrastando N restricciones en la hipótesis nula los grados de libertad de la distribución χ^2 son N.

• En este caso podemos construir un estadístico con distribución exacta en muestras finitas:

$$J_1 = \frac{(T - N - K)}{N} \left[1 + \hat{\boldsymbol{\mu}}_K^{'} \hat{\boldsymbol{\Omega}}^{-1} \boldsymbol{\mu}_k\right]^{-1} \hat{\boldsymbol{a}}' \hat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{a}} \tag{2.22}$$

donde $\hat{\boldsymbol{\Omega}}$ es el estimador maximoverosímil de $\boldsymbol{\Omega}_K$,

$$\hat{\boldsymbol{\Omega}} = \frac{1}{T} \sum_{t=1}^{T} (\boldsymbol{Z}_{Kt} - \hat{\boldsymbol{\mu}}_K)(\boldsymbol{Z}_{Kt} - \hat{\boldsymbol{\mu}}_k)'$$

$$J_1 \overset{H_0}{\sim} F_{(N,(T-N-K))}$$

Este test dado que tienen distribución exacta en muestras finitas puede evitar los problemas que acompañan a la utilización de distribuciones asintóticas.

2.2.2. Los factores son carteras de activos comercializadas y no existe un activo libre de riesgo

En ausencia de un activo libre de riesgo hay un modelo cero-beta que es un modelo multifactorial equivalente al CAPM de Black.

En un contexto multifactorial la cartera cero-beta es una cartera con sensibilidad nula a los factores y el rendimiento esperado en exceso del rendimiento cero-beta están linealmente relacionados con las columnas de la matriz de sensibilidades a los factores. Los factores se suponen que son rendimientos de carteras en exceso del rendimiento cero-beta.

Definimos el vector $(N \times 1)$ de rendimientos reales para N activos o carteras de activos, $\boldsymbol{R}_t$. Para el modelo no restringido tenemos:

$$\boldsymbol{R}_t = \boldsymbol{a} + \boldsymbol{B}\boldsymbol{R}_{Kt} + \boldsymbol{\epsilon}_t \tag{2.23}$$

$$
\begin{aligned}
&E(\boldsymbol{\epsilon}_t) = 0 \\
&E(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t') = \boldsymbol{\Sigma} \\
E(\boldsymbol{R}_{Kt}) = \boldsymbol{\mu}_K, \qquad &E[(\boldsymbol{R}_{Kt} - \boldsymbol{\mu}_K)(\boldsymbol{R}_{Kt} - \boldsymbol{\mu}_K)'] = \boldsymbol{\Omega}_K \\
&Cov(\boldsymbol{R}_{Kt}, \boldsymbol{\epsilon}_t') = \mathbf{0}
\end{aligned}
$$

donde definimos

$\boldsymbol{B}$ es el vector de orden $(N \times K)$ de sensibilidades a los factores.

$\boldsymbol{R}_{Kt}$ es vector de orden $(K \times 1)$ de rendimientos reales de las carteras que actúan de factores y $\boldsymbol{\Omega}_K$ su matriz de varianzas y covarianzas.

$\boldsymbol{a}$ es el vector de orden $(N \times 1)$ de términos independientes.

$\boldsymbol{\epsilon}_t$ es el vector de orden $(N \times 1)$ de perturbaciones y $\boldsymbol{\Sigma}$ su matriz de varianzas y covarianzas.

$\mathbf{0}$ es una matriz de ceros $(K \times N)$

Los estimadores máximoverosímiles del modelo no restringido coinciden con los MCO y son:

$$
\hat{\boldsymbol{a}} = \hat{\boldsymbol{\mu}} - \hat{\boldsymbol{B}}\hat{\boldsymbol{\mu}}_K \tag{2.24}
$$

$$
\hat{\boldsymbol{B}} = \left[\sum_{t=1}^{T}(\boldsymbol{R}_t - \hat{\boldsymbol{\mu}})(\boldsymbol{R}_{Kt} - \hat{\boldsymbol{\mu}}_K)'\right]\left[\sum_{t=1}^{T}(\boldsymbol{R}_{Kt} - \hat{\boldsymbol{\mu}}_K)(\boldsymbol{R}_{Kt} - \hat{\boldsymbol{\mu}}_K)'\right]^{-1} \tag{2.25}
$$

$$
\hat{\boldsymbol{\Sigma}} = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{R}_t - \hat{\boldsymbol{a}} - \hat{\boldsymbol{B}}\boldsymbol{R}_{Kt})(\boldsymbol{R}_t - \hat{\boldsymbol{a}} - \hat{\boldsymbol{B}}\boldsymbol{R}_{Kt})' \tag{2.26}
$$

donde

$$
\hat{\boldsymbol{\mu}} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{R}_t \qquad \hat{\boldsymbol{\mu}}_K = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{R}_{Kt}
$$

En el modelo restringido los rendimientos reales entran en forma de excesos respecto del rendimiento esperado de la cartera cero-beta, γ_0. Por tanto cumplen:

$$
\begin{aligned}
\boldsymbol{R}_t &= \boldsymbol{\iota}\gamma_0 + \boldsymbol{B}(\boldsymbol{R}_{Kt} - \boldsymbol{\iota}\gamma_0) + \boldsymbol{\epsilon}_t \\
&= (\boldsymbol{\iota} - \boldsymbol{B}\boldsymbol{\iota})\gamma_0 + \boldsymbol{B}\boldsymbol{R}_{Kt} + \boldsymbol{\epsilon}_t
\end{aligned} \tag{2.27}
$$

Los estimadores del modelo restringido son:

$$
\hat{\boldsymbol{B}}^{\star} = \left[\sum_{t=1}^{T}(\boldsymbol{R}_t - \boldsymbol{\iota}\hat{\gamma}_0)(\boldsymbol{R}_{Kt} - \boldsymbol{\iota}\hat{\gamma}_0)'\right] \times \left[\sum_{t=1}^{T}(\boldsymbol{R}_{Kt} - \boldsymbol{\iota}\hat{\gamma}_0)(\boldsymbol{R}_{Kt} - \boldsymbol{\iota}\hat{\gamma}_0)'\right]^{-1} \tag{2.28}
$$

$$
\hat{\boldsymbol{\Sigma}}^{\star} = \frac{1}{T}\sum_{t=1}^{T}[\boldsymbol{R}_t - \boldsymbol{\iota}\hat{\gamma}_0 - \hat{\boldsymbol{B}}^{\star}(\boldsymbol{R}_{Kt} - \boldsymbol{\iota}\hat{\gamma}_0)] \times [\boldsymbol{R}_t - \boldsymbol{\iota}\hat{\gamma}_0 - \hat{\boldsymbol{B}}^{\star}(\boldsymbol{R}_{Kt} - \boldsymbol{\iota}\hat{\gamma}_0)]' \tag{2.29}
$$

$$\hat{\gamma}_0 = [(\boldsymbol{\iota} - \hat{\boldsymbol{B}}^{\star}\boldsymbol{\iota})'\hat{\boldsymbol{\Sigma}}^{\star -1}(\boldsymbol{\iota} - \hat{\boldsymbol{B}}^{\star}\boldsymbol{\iota})]^{-1} \times [(\boldsymbol{\iota} - \hat{\boldsymbol{B}}^{\star}\boldsymbol{\iota})'\hat{\boldsymbol{\Sigma}}^{\star -1}(\hat{\boldsymbol{\mu}} - \hat{\boldsymbol{B}}^{\star}\hat{\boldsymbol{\mu}}_K)] \tag{2.30}$$

El estimador maximoverosímil no se puede obtener directamente. Debemos aplicar un proceso iterativo de (2.28) a (2.30). $\hat{\boldsymbol{B}}$ en (2.25) y $\hat{\boldsymbol{\Sigma}}$ en (2.26) pueden ser utilizados como valores iniciales para $\boldsymbol{B}$ y $\boldsymbol{\Sigma}$ en (2.30). Contrastamos la hipótesis nula $H_0 : \boldsymbol{a} = (\boldsymbol{\iota} - \boldsymbol{B}\boldsymbol{\iota})\gamma_0$ utilizando el estadístico de ratio de verosimilitudes definido en (2.14) siendo los grados de libertad $N-1$, perdemos un grado de libertad ya es necesario estimar el rendimiento cero-beta esperado, es decir el parámetro γ_0.

$$J = -\left(T - \frac{N}{2} - K - 1\right)\left[log|\hat{\boldsymbol{\Sigma}}| - log|\hat{\boldsymbol{\Sigma}}^{\star}|\right] \overset{a}{\sim} \chi^2_{(N-1)} \tag{2.31}$$

2.2.3. Variables económicas como factores

No es necesario que los factores sean activos comercializados, en algunos casos como factores se incluyen variables económicas como innovaciones en el PNB, cambios en la curva de los tipos de interés de los bonos, o inflación no anticipada. Sea el siguiente modelo no restringido para K-factores:

$$\boldsymbol{R}_t = \boldsymbol{a} + \boldsymbol{B}\boldsymbol{f}_{Kt} + \boldsymbol{\epsilon}_t \tag{2.32}$$

$$E(\boldsymbol{\epsilon}_t) = 0$$
$$E(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t') = \boldsymbol{\Sigma}$$
$$E(\boldsymbol{f}_{Kt}) = \boldsymbol{\mu}_{fK}, \qquad E[(\boldsymbol{f}_{Kt} - \boldsymbol{\mu}_{fK})(\boldsymbol{f}_{Kt} - \boldsymbol{\mu}_{fK})'] = \boldsymbol{\Omega}_K$$
$$Cov(\boldsymbol{f}_{Kt}, \boldsymbol{\epsilon}_t') = \boldsymbol{0}$$

donde definimos

- $\boldsymbol{B}$ es el vector de orden $(N \times K)$ de sensibilidades a los factores.
- $\boldsymbol{f}_{Kt}$ es vector de orden $(K \times 1)$ de realizaciones de los factores y $\boldsymbol{\Omega}_K$ su matriz de varianzas y covarianzas.
- $\boldsymbol{a}$ es el vector de orden $(N \times 1)$ de términos independientes.
- $\boldsymbol{\epsilon}_t$ es el vector de orden $(N \times 1)$ de perturbaciones y $\boldsymbol{\Sigma}$ su matriz de varianzas y covarianzas.
- $\boldsymbol{0}$ es una matriz de ceros $(K \times N)$

Los estimadores máximoverosímiles del modelo no restringido son:

$$\hat{\boldsymbol{a}} = \hat{\boldsymbol{\mu}} - \hat{\boldsymbol{B}}\hat{\boldsymbol{\mu}}_{fK} \tag{2.33}$$

$$\hat{\boldsymbol{B}} = \left[\sum_{t=1}^{T}(\boldsymbol{R}_t - \hat{\boldsymbol{\mu}})(\boldsymbol{f}_{Kt} - \hat{\boldsymbol{\mu}}_{fK})'\right] \times \left[\sum_{t=1}^{T}(\boldsymbol{f}_{Kt} - \hat{\boldsymbol{\mu}}_{fK})(\boldsymbol{f}_{Kt} - \hat{\boldsymbol{\mu}}_{fK})'\right]^{-1} \tag{2.34}$$

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{R}_t - \hat{\boldsymbol{a}} - \hat{\boldsymbol{B}}\boldsymbol{f}_{Kt})(\boldsymbol{R}_t - \hat{\boldsymbol{a}} - \hat{\boldsymbol{B}}\boldsymbol{f}_{Kt})' \tag{2.35}$$

donde

$$\hat{\boldsymbol{\mu}} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{R}_t \qquad \hat{\boldsymbol{\mu}}_{fK} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{f}_{Kt}$$

Para construir el modelo restringido vamos a comparar la esperanza no condicionada de (2.32) con (2.13). La esperanza no condicionada de (2.32) es:

$$\boldsymbol{\mu} = \boldsymbol{a} + \boldsymbol{B}\boldsymbol{\mu}_{fK} \tag{2.36}$$

donde $\boldsymbol{\mu}_{fK} = E[\boldsymbol{f}_{Kt}]$. Igualando el lado derecho de (2.13) y (2.36) tenemos

$$\boldsymbol{a} = \boldsymbol{\iota}\lambda_0 + \boldsymbol{B}(\boldsymbol{\lambda}_K - \boldsymbol{\mu}_{fK}) \tag{2.37}$$

Llamamos $\gamma_0 = \lambda_0$ y $\gamma_1 = (\boldsymbol{\lambda}_K - \boldsymbol{\mu}_{fK})$ donde $\boldsymbol{\lambda}_K$ es el vector $(K \times 1)$ de premios al riesgo de los factores, tenemos el siguiente modelo restringido:

$$\boldsymbol{R}_t = \boldsymbol{\iota}\gamma_0 + \boldsymbol{B}\gamma_1 + \boldsymbol{B}\boldsymbol{f}_{Kt} + \boldsymbol{\epsilon}_t \tag{2.38}$$

Los estimadores del modelo restringido son:

$$\hat{\boldsymbol{B}}^\star = \left[\sum_{t=1}^{T}(\boldsymbol{R}_t - \boldsymbol{\iota}\hat{\gamma}_0)(\boldsymbol{f}_{Kt} + \hat{\gamma}_1)'\right] \times \left[\sum_{t=1}^{T}(\boldsymbol{f}_{Kt} + \hat{\gamma}_1)(\boldsymbol{f}_{Kt} + \hat{\gamma}_1)'\right]^{-1} \tag{2.39}$$

$$\hat{\boldsymbol{\Sigma}}^\star = \frac{1}{T}\sum_{t=1}^{T}[(\boldsymbol{R}_t - \boldsymbol{\iota}\hat{\gamma}_0) - \hat{\boldsymbol{B}}^\star(\boldsymbol{f}_{Kt} + \hat{\gamma}_1)] \times [(\boldsymbol{R}_t - \boldsymbol{\iota}\hat{\gamma}_0) - \hat{\boldsymbol{B}}^\star(\boldsymbol{f}_{Kt} + \hat{\gamma}_1)]' \tag{2.40}$$

$$\hat{\gamma} = [\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{\star-1}\boldsymbol{X}]^{-1}[\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{\star-1}(\hat{\boldsymbol{\mu}} - \hat{\boldsymbol{B}}^\star\hat{\boldsymbol{\mu}}_{fK})] \tag{2.41}$$

donde $\boldsymbol{X} = [\boldsymbol{\iota} \quad \hat{\boldsymbol{B}}^\star]$ y $\gamma = [\gamma_0 \quad \gamma_1']'$

El estimador maximoverosímil no se puede obtener directamente. Debemos aplicar un proceso iterativo de (2.39) a (2.41). $\hat{\boldsymbol{B}}$ en (2.34) y $\hat{\boldsymbol{\Sigma}}$ en (2.35) pueden ser utilizados como valores iniciales para $\boldsymbol{B}$ y $\boldsymbol{\Sigma}$ en (2.41).

Contrastamos la hipótesis nula $H_0 : \boldsymbol{a} = (\boldsymbol{\iota}\gamma_0 + \boldsymbol{B}\gamma_1)$ utilizando el estadístico de ratio de verosimilitudes definido en (2.14) siendo los grados de libertad $N - K - 1$, perdemos un grado de libertad estimando γ_0 y K estimado los K elementos de $\boldsymbol{\lambda}_K$.

$$J = -\left(T - \frac{N}{2} - K - 1\right)\left[log|\hat{\boldsymbol{\Sigma}}| - log|\hat{\boldsymbol{\Sigma}}^\star|\right] \overset{a}{\sim} \chi^2_{(N-K-1)} \tag{2.42}$$

2.2.4. Los factores son carteras de activos y estas carteras generan la frontera media-varianza de los activos

Cuando los factores generan la frontera media-varianza el término independiente de la relación de valoración exacta, λ_0, es cero sin necesidad de un activo libre de riesgo. En el contexto del APT, generación ocurre cuando dos carteras bien diversificadas son de varianza mínima en el límite.

El modelo no restringido se debe expresar en rendimientos reales. Sea $\boldsymbol{R}_t$ el vector de rendimientos reales de orden $(N \times 1)$, definimos el modelo de K-factores como:

$$\boldsymbol{R}_t = \boldsymbol{a} + \boldsymbol{B}\boldsymbol{R}_{Kt} + \boldsymbol{\epsilon}_t \tag{2.43}$$

$$E(\boldsymbol{\epsilon}_t) = 0$$

$$E(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t') = \boldsymbol{\Sigma}$$

$$E(\boldsymbol{R}_{Kt}) = \boldsymbol{\mu}_K, \qquad E[(\boldsymbol{R}_{Kt} - \boldsymbol{\mu}_K)(\boldsymbol{R}_{Kt} - \boldsymbol{\mu}_K)'] = \boldsymbol{\Omega}_K$$

$$Cov(\boldsymbol{R}_{Kt}, \boldsymbol{\epsilon}_t') = \boldsymbol{0}$$

donde definimos

$\boldsymbol{B}$ es el vector de orden $(N \times K)$ de sensibilidades a los factores.

$\boldsymbol{R}_{Kt}$ es vector de orden $(K \times 1)$ de rendimientos reales de las carteras y $\boldsymbol{\Omega}_K$ su matriz de varianzas y covarianzas.

$\boldsymbol{a}$ es el vector de orden $(N \times 1)$ de términos independientes.

$\boldsymbol{\epsilon}_t$ es el vector de orden $(N \times 1)$ de perturbaciones y $\boldsymbol{\Sigma}$ su matriz de varianzas y covarianzas.

$\boldsymbol{0}$ es una matriz de ceros $(K \times N)$

Las restricciones incluidas en (2.43) impuestas por la inclusión de carteras de factores que generan la frontera media-varianza son $\boldsymbol{a} = 0$ y $\boldsymbol{B}\boldsymbol{\iota} = \boldsymbol{\iota}$. Los estimadores máximoverosímiles del modelo no restringido son:

$$\hat{\boldsymbol{a}} = \hat{\boldsymbol{\mu}} - \hat{\boldsymbol{B}}\hat{\boldsymbol{\mu}}_K \tag{2.44}$$

$$\hat{\boldsymbol{B}} = \left[\sum_{t=1}^{T}(\boldsymbol{R}_t - \hat{\boldsymbol{\mu}})(\boldsymbol{R}_{Kt} - \hat{\boldsymbol{\mu}}_K)'\right] \times \left[\sum_{t=1}^{T}(\boldsymbol{R}_{Kt} - \hat{\boldsymbol{\mu}}_K)(\boldsymbol{R}_{Kt} - \hat{\boldsymbol{\mu}}_K)'\right]^{-1} \tag{2.45}$$

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{R}_t - \hat{\boldsymbol{a}} - \hat{\boldsymbol{B}}\boldsymbol{R}_{Kt})(\boldsymbol{R}_t - \hat{\boldsymbol{a}} - \hat{\boldsymbol{B}}\boldsymbol{R}_{Kt})' \tag{2.46}$$

donde

$$\hat{\boldsymbol{\mu}} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{R}_t \qquad \hat{\boldsymbol{\mu}}_K = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{R}_{Kt}$$

Para escribir el modelo restringido consideremos $\boldsymbol{B}$ particionada en un vector columna $\boldsymbol{b}_1$ de orden $(N \times 1)$ y la matriz $\boldsymbol{B}_1$ de orden $(N \times (K-1))$. Igualmente particionamos el vector de factores por filas, la primera fila es el vector $\boldsymbol{R}_{1t}$ y las $(K-1)$ restantes las agrupamos en $\boldsymbol{R}_{K^\star t}$. Con esta partición la restricción $\boldsymbol{B\iota} = \boldsymbol{\iota}$ puede ser escrita como $\boldsymbol{b}_1 + \boldsymbol{B}_1\boldsymbol{\iota} = \boldsymbol{\iota}$. De esta forma hemos escrito el modelo no restringido como

$$\boldsymbol{R}_t = \boldsymbol{a} + \boldsymbol{b}_1\boldsymbol{R}_{1t} + \boldsymbol{B}_1\boldsymbol{R}_{K^\star t} + \boldsymbol{\epsilon}_t \tag{2.47}$$

Sustituyendo $\boldsymbol{a} = 0$ y $\boldsymbol{b}_1 = \boldsymbol{\iota} - \boldsymbol{B}_1\boldsymbol{\iota}$ en (2.47) obtenemos el siguiente modelo restringido:

$$\boldsymbol{R}_t - \boldsymbol{\iota}\boldsymbol{R}_{1t} = \boldsymbol{B}_1(\boldsymbol{R}_{K^\star t} - \boldsymbol{\iota}\boldsymbol{R}_{1t}) + \boldsymbol{\epsilon}_t \tag{2.48}$$

Los estimadores del modelo restringido son:

$$\hat{\boldsymbol{B}}_1^\star = \left[\sum_{t=1}^{T}(\boldsymbol{R}_t - \boldsymbol{\iota}\boldsymbol{R}_{1t})(\boldsymbol{R}_{K^\star t} - \boldsymbol{\iota}\boldsymbol{R}_{1t})'\right] \times \left[\sum_{t=1}^{T}(\boldsymbol{R}_{KT} - \boldsymbol{\iota}\boldsymbol{R}_{1t})(\boldsymbol{R}_{K^\star t} - \boldsymbol{\iota}\boldsymbol{R}_{1t})'\right]^{-1} \tag{2.49}$$

$$\hat{\boldsymbol{b}}_1^\star = \boldsymbol{\iota} - \hat{\boldsymbol{B}}_1^\star\boldsymbol{\iota} \tag{2.50}$$

$$\hat{\boldsymbol{\Sigma}}^\star = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{R}_t - \hat{\boldsymbol{B}}^\star\boldsymbol{R}_{Kt})(\boldsymbol{R}_t - \hat{\boldsymbol{B}}^\star\boldsymbol{R}_{Kt})' \tag{2.51}$$

Contrastamos $H_0 : \boldsymbol{a} = 0 \quad \text{y} \quad \boldsymbol{B\iota} = \boldsymbol{\iota}$ con el estadístico (2.14) con $2N$ grados de libertad. Además podemos construir un estadístico de contraste con distribución exacta:

$$J_2 = \frac{(T-N-K)}{N}\left[\frac{|\hat{\boldsymbol{\Sigma}}^\star|}{|\hat{\boldsymbol{\Sigma}}|} - 1\right] \overset{H_0}{\sim} F_{(2N,2(T-N-K))} \tag{2.52}$$

2.3. Estimación de la prima de riesgo y de los rendimientos esperados

Los modelos de valoración exactos nos permiten estimar el rendimiento esperado de un activo, $\boldsymbol{\mu} = \boldsymbol{\iota}\lambda_0 + \boldsymbol{B}\boldsymbol{\lambda}_k$, para lo cual necesitamos estimadores de la matriz de sensibilidades al riesgo $\boldsymbol{B}$, la tasa de riesgo o rendimiento esperado cero-beta λ_0 y la prima por riesgo de los factores $\boldsymbol{\lambda}_k$. Las dos primeras se estiman directamente y la prima por riesgo de los factores se puede obtener para cada caso:

1. Caso 1: Podemos estimar la prima por riesgo directamente de las medias muestrales de los excesos de rendimiento en las carteras:

$$\hat{\boldsymbol{\lambda}}_K = \hat{\boldsymbol{\mu}}_K = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{Z}_{Kt} \tag{2.53}$$

   un estimador de la matriz de varianzas y covarianzas de $\hat{\boldsymbol{\lambda}}_K$ es:

$$\widehat{Var}[\hat{\boldsymbol{\lambda}}_K] = \frac{1}{T}\hat{\boldsymbol{\Omega}}_K = \frac{1}{T^2}\sum_{t=1}^{T}(\boldsymbol{Z}_{Kt} - \hat{\boldsymbol{\mu}}_K)(\boldsymbol{Z}_{Kt} - \hat{\boldsymbol{\mu}}_K)'$$

2. Caso 2: La prima por riesgo de los factores puede ser estimada usando la diferencia entre la media muestral de las carteras-factores y el rendimiento cero-beta estimado:

$$\hat{\boldsymbol{\lambda}}_K = \hat{\boldsymbol{\mu}}_K - \boldsymbol{\iota}\hat{\gamma}_0 \tag{2.54}$$

   un estimador de la matriz de varianzas y covarianzas de $\hat{\boldsymbol{\lambda}}_K$ es:

$$\widehat{Var}[\hat{\boldsymbol{\lambda}}_K] = \frac{1}{T}\hat{\boldsymbol{\Omega}}_K + \widehat{Var}[\hat{\gamma}_0]\boldsymbol{\iota}\boldsymbol{\iota}'$$

   siendo

$$\widehat{Var}[\hat{\gamma}_0] = \frac{1}{T}\left(1 + (\hat{\boldsymbol{\mu}}_K - \hat{\gamma}_0\boldsymbol{\iota})'\hat{\boldsymbol{\Omega}}_K^{-1}(\hat{\boldsymbol{\mu}}_K - \hat{\gamma}_0\boldsymbol{\iota})\right) \times [(\boldsymbol{\iota} - \hat{\boldsymbol{B}}^{\star}\boldsymbol{\iota})'\hat{\boldsymbol{\Sigma}}^{\star-1}(\boldsymbol{\iota} - \hat{\boldsymbol{B}}^{\star}\boldsymbol{\iota})]^{-1}$$

3. Caso 3: En este caso, donde los factores no son carteras comercializadas, obtenemos la prima por riesgo de los factores como:

$$\hat{\boldsymbol{\lambda}}_K = \hat{\boldsymbol{\mu}}_{fK} + \hat{\gamma}_1 \tag{2.55}$$

   un estimador de la matriz de varianzas y covarianzas de $\hat{\boldsymbol{\lambda}}_K$ es:

$$\widehat{Var}[\hat{\boldsymbol{\lambda}}_K] = \frac{1}{T}\hat{\boldsymbol{\Omega}}_K = \widehat{var}[\hat{\gamma}_1]$$

4. Caso 4: Este caso es el mismo que el primero salvo porque los rendimientos reales son sustituidos por excesos de rendimiento. $\boldsymbol{\lambda}_K$ es el vector de medias muestrales de los factores y λ_0 es cero.

El rendimiento esperado de un activo puede ser estimado sustituyendo los estimadores de $\boldsymbol{B}, \lambda_0, \boldsymbol{\lambda}_K$ en (2.13). Dado que (2.13) no es lineal en parámetros, calcular las desviaciones requiere usar aproximaciones lineales y estimar las covarianzas de los parámetros estimados.

2.3.1. Valoración de los factores

- Podemos contrastar la significatividad conjunta de los factores (contrastamos si los factores son valorados), $H_0 : \boldsymbol{\lambda}_K = 0$, con el estadístico.

$$J_3 = \frac{(T-K)}{TK}\hat{\boldsymbol{\lambda}}_K'\widehat{Var}[\hat{\boldsymbol{\lambda}}_K]^{-1}\hat{\boldsymbol{\lambda}}_K \xrightarrow{d,H_0} T^2_{(K,T-K)} \tag{2.56}$$

  En el caso en que el estimador de $\boldsymbol{\lambda}$ esté basado únicamente en medias muestrales de los factores, el estadístico anterior tiene distribución exacta $F_{(K,T-K)}$

- Para contrastar la significatividad individual de un factor, $H_0 : \lambda_{jK} = 0$ podemos utilizar el estadístico:

$$J_4 = \frac{\hat{\lambda}_{jK}}{\sqrt{v_{jj}}} \overset{a}{\sim} N(0,1) \tag{2.57}$$

siendo $\hat{\lambda}_{jk}$ es el j-ésimo elemento de $\hat{\boldsymbol{\lambda}}$ y v_{jj} es el elemento (j,j) de $\widehat{Var}[\hat{\boldsymbol{\lambda}}_k]$. Notar que en el caso de que los factores sean derivados estadísticamente no tienen una interpretación económica clara, solo una aproximación basada en conjeturas. En esta situación los contrastes anteriores sobre estos factores no son de utilidad.

2.3.2. Estimador alternativo de Shanken (1992b)

Shanken (1992b) muestra que las primas por riesgo de los factores pueden ser estimadas mediante un procedimiento alternativo, en dos etapas. Este procedimiento sigue la misma secuencia de estimación y contraste que el de Fama y MacBeth (1973) para el CAPM.

1. En la primera etapa se estiman las sensibilidades a los factores activo por activo mediante MCO. Estas estimaciones sirven para obtener una aproximación a la matrix $\boldsymbol{B}$, $\tilde{\boldsymbol{B}}$. Este estimador de $\boldsymbol{B}$ es idéntico al estimador maximoverosímil del modelo no restringido obtenido bajo el supuesto de normalidad conjunta y residuos IID.

2. En la segunda etapa se estima el modelo

$$\boldsymbol{Z}_{Kt} = \boldsymbol{\iota}\lambda_{0t} + \tilde{\boldsymbol{B}}\boldsymbol{\lambda}_{Kt} + \boldsymbol{\eta}_t \tag{2.58}$$

Estimando por MCO en (2.58) en cada momento t obtenemos una serie temporal de estimadores consistentes, para la prima por riesgo de los factores $\boldsymbol{\lambda}_{Kt}$ y de la cartera cero-beta λ_0. Esta serie temporal es utilizada para desarrollar las inferencias en términos de las propiedades de la serie temporal. Shanken (1992b) derivó el ajuste necesario para valorar el problema de errores en variables derivado de la estimación de $\boldsymbol{B}$.

Si se utiliza el estimador maximoverosímil este ajuste no es necesario ya que el estimador lo incorpora en la derivación.

2.4. Selección de factores

La estimación desarrollada supone que los factores son conocidos. Podemos distinguir dos tipos de factores: estadísticos y teóricos.

2.4.1. Aproximación estadística

La aproximación estadística implica construir factores a partir de un conjunto de activos mucho mayor que el que se utiliza para estimar y contrastar el modelo, en general activos individuales. Los datos muestrales de estas rentabilidades son utilizadas para construir carteras que representan a los factores. Hay dos procedimientos alternativos con numerosas variantes, el primero es el análisis de factores y el segundo son las componentes principales. Esta metodología utiliza el exceso de rendimiento de una cartera de activos individuales para obtener una proxy para cada factor de riesgo.

Este método fue desarrollado por Connor y Korajczyk (1986, 1988) y Ferson y Korajczyk (1991). Mucha literatura está dirigida a compararlos y bastante a determinar el número de factores suficiente para explicar las rentabilidades.

Connor y Korajczyk (1986) muestran que la aproximación deriva estimadores consistentes de los factores de un APT si el número de activos es suficientemente grande. Prueban además que los errores de estimación para los factores pueden ser pequeños en muestras finitas.

Sea $\boldsymbol{Z}^N$ la matriz de excesos de rendimiento de N activos (individuales) con T observaciones cada uno. En una economía con N activos donde suponemos que las rentabilidades de los títulos siguen un modelo factorial y existe un activo libre de riesgo, podemos escribir:

$$\boldsymbol{Z}^N = \boldsymbol{B}^N \boldsymbol{f}^N + \epsilon^N \tag{2.59}$$

donde $\boldsymbol{Z}^N, \boldsymbol{f}^N$ y ϵ^N tienen la interpretación habitual. Además suponemos que $E(\epsilon^N \epsilon^{N'}) = V$ no necesariamente diagonal.

Chamberlain y Rothschild (1983) demuestran que cuando el número de observaciones de sección cruzada crece, el análisis de los vectores propios es asintóticamente equivalente al análisis de factores. Connor y Korajczyk (1986) utilizaron la teoría de componentes principales para identificar estadísticamente los factores de la matriz $\boldsymbol{f}^N$. Demostraron que los k primeros vectores propios de la matriz $T \times T$, $\Omega^N = 1/N(\boldsymbol{Z}^N \boldsymbol{Z}^{N'})$ se aproximan a una transformación no singular de $\boldsymbol{f}$ cuando N tiende a infinito. Por tanto si $\boldsymbol{G}^N$ de orden $(k \times T)$ es la matriz que recoge a los k primeros vectores propios de Ω^N, puede ser utilizada en vez de $\boldsymbol{f}$ para estimar las sensibilidades a los factores en (2.59).

Connor y Korajczyk (1988) mostraron que una versión iterada del procedimiento anterior mejora la eficiencia en la estimación. En este caso no se utiliza $\boldsymbol{G}^N$ como la matriz de factores sino $\boldsymbol{G}^{N*}$ que sería la matriz de orden $(k \times T)$ que recoge los k primeros vectores propios de la matriz $\Omega^{N*} = 1/N(\boldsymbol{Z}^{N*} \boldsymbol{Z}^{N*'})$. Siendo $\boldsymbol{Z}^{N*}$ la matriz de excesos de rendimiento escalada por la diagonal de la matriz de varianzas y covarianzas de los rendimientos idiosincrásicos:

$$\boldsymbol{Z}^{N*} = diag(V^N)^{-1/2} \boldsymbol{Z}$$

y como estimador de los elementos de la diagonal de V^N se utiliza la varianza residual de la regresión:

$$\boldsymbol{Z}^N = \alpha + \boldsymbol{G}^{N'} \beta + \eta$$

siendo β un vector de parámetros de orden $(k \times 1)$.

La principal ventaja de la versión iterada es la mayor rapidez de convergencia. Sin embargo, implica estimar V lo que puede reducir la eficiencia del procedimiento cuando N no sea muy grande.

Una cuestión inherente al análisis empírico de un modelo de factores es la determinación del número de factores que se deben incluir. Connor y Korajczyk (1993) desarrollan un procedimiento de contraste para la determinación del número de factores en un modelo factorial basado en la idea:

"Si k es el número de factores necesario, la inclusión del factor $k+1$ no debe reducir significativamente la varianza residual del modelo con el nuevo factor". Es decir la variabilidad de los residuos explicada por el factor $k+1$ debe ser asintóticamente igual a cero. El procedimiento implica estimar el modelo con k y $k+1$ factores y comparar su capacidad explicativa.

La utilización de factores determinados por un análisis de componentes principales tiene como ventaja que prescinde de la determinación subjetiva de las variables económicas que actúen como factores. Por contra carecen de contenido económico e interpretación económica.

2.4.2. Aproximación teórica

La aproximación teórica distingue entre dos prácticas alternativas:

- Una aproximación implica especificar factores como variables macroeconómicas y del mercado financiero que se cree capturan riesgo sistemático de la economía. En esta línea dos de los artículos más relevantes son Chen, Roll y Ross (1986) o Fama y French (1993).

- Una aproximación alternativa determina los factores como características específicas de las empresas que probablemente explican diferenciales en sensibilidad al riesgo sistemático, y forman carteras de activos basados en estas características. Las características utilizadas son el valor de mercado, el ratio precio-ganancias, el ratio valor contable-valor de mercado. La literatura muestra que modelos de factores que incluyen la cartera de mercado equiponderada y factores construidos en base a estas características específicas de las empresas explican la sección cruzada de las rentabilidades.

•Factores de Chen, Roll y Ross (CRR) (1986)

Chen, N., Roll, R., y S. Ross (1986), "Economic Forces And the Stock Market", *Journal of Business*, 59, 383-403.

Estos autores utilizan variables macroeconómicas como factores de riesgo agregado. Partiendo de la ecuación fundamental de valoración escrita en términos de los precios

$$P_{jt} = E[M_{t+1} X_{j,t+1}] \qquad j = 1, \ldots, N \tag{2.60}$$

determinan con argumentos intuitivos un conjunto de variables que influyen en la determinación de los precios de los activos.

Siendo $X_{j,t+1}$ los flujos que genera la empresa j en el futuro y M_{t+1} el factor de descuento, las variables macroeconómicas pueden afectar a los flujos futuros de la empresa o al factor de descuento. Las variables que especifican son:

- El cambio mensual porcentual en el índice de producción industrial.

- Una medida de inflación no esperada.

- La variación mensual en la inflación esperada.

- La diferencia entre el tipo de interés de la deuda pública a largo plazo y a corto plazo. Este factor está asociado a cambios en la estructura temporal de los tipos de interés.

- El diferencial de insolvencia financiera medido como la diferencia entre el tipo de interés de la deuda empresarial a largo plazo y el tipo de interés de la deuda pública a largo plazo.

Estiman el siguiente modelo de sección cruzada:

$$R_{jt} = \gamma_{0t} + \gamma_{1t}\hat{b}_{jmt} + \sum_{k=1}^{5} \gamma_{kt}\hat{b}_{jkt} + \eta_{jt} \qquad t = 1, \ldots, T \tag{2.61}$$

donde $\hat{b}_{jmt}$ es estimada previamente en el modelo de mercado habitual y las $\hat{b}_{jkt}$ son estimadas previamente en la apropiada versión del modelo de mercado para cada factor:

$$R_{jt} = \alpha_{jt} + b_{jt}F_{kt} + \epsilon_{jt} \qquad t = 1, \ldots, T \tag{2.62}$$

Los contrastes se realizan con la metodología de Fama y MacBeth y la apropiada corrección para el problema de errores en variables se corresponde con ajustar la varianza de la serie temporal de estimadores para cada factor con el correspondiente término de ajuste:

$$\left(1 + \frac{\hat{\gamma}_k^2}{\sigma_k^2}\right) \qquad k = 1, \ldots, K$$

Siendo k el factor contrastado.

• **Factores de Fama y French (1993, 1996)**

Fama E., y K. French, (1993), "Common Risk Factors in the Returns on Stocks and Bonds", *Journal of Financial Economics*, 33, 3-54.

Fama E., y K. French, (1996), "Multifactor Explanations of Asset Pricing Anomalies", *Journal of Finance*, 51, 55-84.

Fama y French (1993) proponen un modelo de tres factores de riesgo que pueden replicarse mediante unas determinadas carteras de los activos existentes en la economía. El modelo que estiman es.

$$R_{jt} - r_j = \alpha_j + b_{jm}(R_{mt} - r_t) + b_{j,SMB}SMB_t + b_{j,HML}HML_t + \varepsilon_{jt} \quad j = 1, \ldots, 25 \tag{2.63}$$

Siendo:

- El primer factor consiste en replicar el riesgo de mercado mediante una cartera de coste cero con posición larga en la propia cartera de mercado y corta (endeudamiento) en el activo libre de riesgo. La prima asociada a esta cartera es la prima por riesgo del mercado.

- SMB es la cartera que replica al factor de riesgo asociado al tamaño entendido como capitalización bursátil. Esta cartera representa la diferencia entre el rendimiento de las carteras más pequeñas y las más grandes controlando por los efectos potenciales del cociente VC/VM.

- HML es la cartera que replica al factor de riesgo asociado al cociente VC/VM, esta cartera representa la diferencia entre las carteras con más alto y más bajo VC/VM una vez controlado el efecto tamaño.

FF obtienen que el modelo con los tres factores aumenta la capacidad explicativa de un $77,9\,\%$ a un $93,1\,\%$ con respecto a mantener como único factor de riesgo el asociado a la cartera de mercado. Los resultados obtenidos por FF indican que sus dos carteras réplica poseen betas que deberían explicar en sección cruzada los rendimientos medios de los activos.

Tema 3

Modelo de Mercado

3.1. Introducción

Bajo el supuesto de ausencia de arbitraje los precios de los activos financieros deben ser tales que satisfagan la expresión:

$$E_t(M_{t+1}\tilde{R}_{t+1}) = 1 \qquad j = 1, \ldots, N \tag{3.1}$$

es decir, el rendimiento esperado ponderado por la variable agregada o factor de descuento, M, es constante e igual a 1 para todos los activos financieros [1]. La ecuación (3.1) dice que todos los activos que prometen ofrecer los mismos pagos futuros deben tener hoy el mismo precio.

Esta expresión nos permite obtener una fórmula de valoración en términos de la prima de riesgo esperada de cualquier activo incierto j:

$$E(R_j - r) = \left[-\frac{1}{E(M)}\right] Cov(M, R_j) \tag{3.2}$$

Supuestos alternativos sobre la variable agregada M permiten obtener diferentes modelos de valoración.

- Si suponemos que el factor de descuento o M es lineal en el rendimiento de la cartera de mercado,

  $$M = \delta_{0m} + \delta_{1m} R_m$$

  podemos contrastar la ecuación anterior utilizando el modelo de valoración de activos con cartera de mercado, CAPM. En este caso los inversores al escoger como cartera óptima a la cartera de mercado, están soportando exclusivamente riesgo sistemático (el riesgo global de la economía) pero no soportan además riesgos específicos asociados a los activos individuales que componen la cartera de mercado.

[1] En adelante prescindiremos de los subíndices temporales.

- Supongamos que existen múltiples factores de riesgo sistemático que no puedan ser recogidos exclusivamente mediante la cartera de mercado. Suponiendo que existan K factores de riesgo sistemático podemos representar la variable agregada M como:

$$M = \delta_0 + \delta_1 F_1 + \ldots + \delta_K F_K \tag{3.3}$$

donde $F_1, F_2, \ldots, F_K$ son los factores de riesgo sistemático de la economía tal que

$$E[\tilde{R}_j(\delta_0 + \delta_1 F_1 + \ldots + \delta_K F_K)] = 1; \quad j = 1, \ldots, N \tag{3.4}$$

En este caso podemos contrastar la ecuación (3.2) anterior utilizando el modelo de valoración de activos bajo ausencia de arbitraje o APT.

Vamos a pensar en cómo se generan los rendimientos de los activos financieros inciertos:

El rendimiento observado de un activo j se puede descomponer en dos partes, la parte esperada por el inversor $E(R_j)$ y la parte no esperada, componente sorpresa, error o innovación:

$$R_j = E(R_j) + \text{innovación}_j \tag{3.5}$$

El error o componente sorpresa se puede a su vez descomponer en dos partes:

a) Sorpresas en los rendimientos de los activos individuales debidas a la llegada de nueva información de la economía en general que afectan a todas las empresas, por supuesto en diferente grado. Le vamos a denominar innovación sistemática ya que afecta de manera sistemática a todas las empresas. Es un componente del riesgo no diversificable.

b) Sorpresas debidas a la llegada de nueva información que afecta exclusivamente a un activo. Le vamos a denominar innovación idiosincrásica. Es un componente del riesgo diversificable, una adecuada combinación de activos individuales en carteras puede eliminar en parte este componente de riesgo individual.

Por tanto, el rendimiento observado del activo j se genera como la siguiente suma

$$R_j = E(R_j) + \underbrace{\beta_{j1} F_1 + \beta_{j2} F_2 + \ldots + \beta_{jK} F_K}_{\text{innovación sistemática}} + \underbrace{\epsilon_j}_{\text{innovación idiosincrásica}} \tag{3.6}$$

Así podemos escribir el proceso generador de rendimientos que suponemos que existe en los mercados como

$$R_j = \alpha_j + \beta_{j1} F_1 + \beta_{j2} F_2 + \ldots + \beta_{jK} F_K + \epsilon_j \tag{3.7}$$

donde:

- $F_1, F_2, \ldots, F_K$, son los factores sistemáticos o agregados comunes a todos los activos existentes expresados como innovaciones por lo que tienen valor esperado cero y sus covarianzas entre dos factores cualesquiera también son cero. $E(F_K) = 0 \quad \forall j, K$
- $\beta_{j1}, \beta_{j2}, \ldots, \beta_{jK}$, son las sensibilidades de los rendimientos del activo j a los factores.
- $E(\epsilon_j F_K) = 0 \quad \forall j, K$, los factores de riesgo sistemático son considerados innovaciones y no están correlacionados con el componente idiosincrásico.
- El componente idiosincrásico es también una innovación. No están correlacionados entre sí para las diferentes empresas,
  $E(\epsilon_j) = E(\epsilon_j \epsilon_h) = 0 \quad \forall j, j \neq h$
- $a_j = E(R_j) \quad \forall j$ ya que los factores sistemáticos de riesgo y el componente idiosincrásico son innovaciones cuyo valor esperado es igual a cero.

Al proceso generador de rendimientos recogido en (3.7) le llamamos modelo factorial. Si le añadimos el supuesto de ausencia de arbitraje podemos derivar el modelo APT.

3.2. El modelo de mercado

3.2.1. El modelo de mercado como modelo factorial

El modelo de mercado es un caso particular de los modelos factoriales ya que considera un único factor (una única fuente de riesgo sistemático) en el proceso de generación del rendimiento. Este factor es el rendimiento de la cartera de mercado. El modelo de mercado se escribe:

$$R_j = \alpha_j + \beta_{jm} R_m + \epsilon_j \tag{3.8}$$

donde suponemos

$$E(\epsilon_j R_m) = E(\epsilon_j \epsilon_h) = E(\epsilon_j) = 0 \quad \forall j \quad \text{y} \quad \forall j \neq h$$

Dada la ecuación (3.8) debemos notar:

- Puesto que el único factor de riesgo sistemático es el rendimiento de la cartera de mercado estamos suponiendo que todas las noticias macroeconómicas quedan recogidas en ésta.
- Dado que $E(\epsilon_j \epsilon_h) = 0$ la única razón por la que los rendimientos de los activos tienden a moverse de forma conjunta es porque experimentan movimientos en común con la cartera de mercado como única fuente de riesgo.
- El modelo de mercado es un modelo factorial, no hay un razonamiento económico que justifique la cartera de mercado como único factor de riesgo sistemático ni se puede interpretar la perturbación del modelo como componente idiosincrásico del rendimiento de un activo financiero. El verdadero componente idiosincrásico será el

correspondiente a un modelo que incluya todos los factores de riesgo existentes, suponer que sólo hay uno y que éste es el rendimiento de la cartera de mercado es demasiado exigente. Sin embargo es un modelo sencillo y útil como luego se verá.

- Dado que la verdadera cartera de mercado es desconocida, en la práctica, se aproxima por el rendimiento de un índice bursátil por lo cual el factor de riesgo sistemático tiene interpretación económica.

3.2.2. El modelo de mercado bajo supuestos estadísticos

El modelo de mercado también puede ser justificado bajo supuestos estadísticos suponiendo que la distribución conjunta entre el rendimiento del activo j y el de la cartera de mercado es Normal Bivariante.

Recordemos que la media o valor esperado de la distribución de R_j condicionada en R_m es[2]

$$E(R_j|R_m) = \int R_j f(R_j|R_m) dR_j$$

Si la distribución conjunta de R_j y R_m es Normal bivariante tal que

$$(R_j, R_m) \sim N_2(E(R_j), E(R_m), \sigma_j^2, \sigma_m^2, \rho_{jm})$$

entonces tenemos los siguientes resultados

a) Las distribuciones marginales son normales

$$f(R_j) = N(E(R_j), \sigma_j^2)$$

$$f(R_m) = N(E(R_m), \sigma_m^2)$$

b) Las distribuciones condicionales son normales

$$f(R_j|R_m) = N(E(R_j|R_m), Var(R_j|R_m))$$

siendo su media

$$E(R_j|R_m) = \alpha_j + \beta_{jm} R_m$$

donde

$$\beta_{jm} = \frac{Cov(R_j, R_m)}{Var(R_m)}$$

$$\alpha_j = E(R_j) - \beta_{jm} E(R_m)$$

y su varianza

$$Var(R_j|R_m) = \int_{R_j} \left[R_j - E(R_j|R_m)\right]^2 f(R_j|R_m) dR_j = \sigma_j^2 (1 - \rho_{jm}^2)$$

[2]La media o valor esperado es una suma ponderada de todos los posibles valores de R_j. Al tomar el valor esperado condicional, la ponderación asociada a R_j es su función de densidad condicional, $f(R_j|R_m)$, en lugar de la función de densidad marginal $f(R_j)$.

donde ρ_{jm} es el coeficiente de correlación entre el rendimiento del activo j y el rendimiento de la cartera de mercado y que viene dado por

$$\rho_{jm} = \frac{Cov(R_j, R_m)}{\sigma_j \sigma_m}$$

Además $Var(R_j|R_m)$ es constante y menor que $< \sigma_j^2$.

Podemos escribir:

$$f(R_j|R_m) = N(\alpha_j + \beta_{jm}R_m, \sigma_j^2(1 - \rho_{jm}^2))$$

La función de regresión de R_j sobre R_m, que es la función de media condicional $E(R_j|R_m)$ es lineal e igual a:

$$E(R_j|R_m) = \alpha_j + \beta_{jm}R_m$$

c) $Var(R_j|R_m)$ es independiente de R_m ya que su valor es el mismo para todos los valores de R_m. Por tanto:

$$\epsilon_j = R_j - E(R_j|R_m) = R_j - \alpha_j - \beta_{jm}R_m \sim N(0, \sigma_{\epsilon_j})$$
$$E(\epsilon_j|R_m) = E(\epsilon_j) = 0$$
$$Var(\epsilon_j|R_m) = Var(R_j|R_m) = Var(R_j)(1 - \rho_{jm}^2 = Var(\epsilon_j)$$

Por tanto la perturbación ϵ_j tiene la misma distribución condicional Normal para todos los valores de R_m, es decir ϵ_j y R_m son independientes, (su covarianza es cero, su correlación es cero y además suponemos normalidad luego son independientes).

Si reescribimos

$$\epsilon_j = R_j - E(R_j|R_m) = R_j - \alpha_j - \beta_{jm}R_m \longrightarrow R_j = \alpha_j - \beta_{jm}R_m + \epsilon_j$$

obtenemos el modelo de mercado donde

$E(\epsilon_j) = 0, E(\epsilon_j^2) = \sigma_{\epsilon_j}^2,$ y $E(\epsilon_j|R_m) = 0$

El modelo de mercado como modelo factorial expresa el rendimiento $\forall j$ como la suma de dos partes

$$R_j = \underbrace{\alpha_j}_{\text{comp. idiosincrásico constante}} + \underbrace{\underbrace{\beta_{jm}R_m}_{\text{sistemático}} + \underbrace{\epsilon_j}_{\text{comp. idios. aleatorio}}}_{\text{innovación}}$$

Como ϵ_j y R_m son independientes podemos escribir

$$\sigma_j^2 = \beta_{jm}^2\sigma_m^2 + \sigma_{\epsilon_j}^2$$

donde σ_m^2 es la varianza del rendimiento de la cartera de mercado y $\sigma_{\epsilon_j}^2$ es la varianza del componente del rendimiento que se debe a innovaciones o llegadas de nueva información que afectan únicamente a la empresa j.

3.3. Estimación del modelo de mercado

Suponiendo que sólo existe un factor de riesgo agregado en la economía podemos recurrir al modelo de mercado para estimar el riesgo beta de cualquier activo financiero. El modelo de mercado se define como:

$$R_{jt} = \alpha_j + \beta_{jm} R_{mt} + \epsilon_{jt} \qquad t = 1, \ldots, T \quad j = 1, \ldots, N \tag{3.9}$$

donde suponemos que la distribución conjunta de los rendimientos de los activos y del mercado es normal, estacionaria y serialmente independiente:

$$\begin{array}{l} E(\epsilon_{jt}|R_{mt}) = E(\epsilon_{jt}) = 0 \quad \forall t \\ Var(R_{jt}|R_{mt}) = Var(\epsilon_{jt}|R_{mt}) = Var(\epsilon_{jt}) = \sigma^2_{\epsilon_j} \quad \forall t \\ Cov(\epsilon_{jt}, R_{mt}) = Cov(\epsilon_{jt}, \epsilon_{jt-\tau}) = 0 \quad t = 1, \ldots, T; \quad \tau \geq 1 \end{array}$$

El modelo de mercado para el activo j podemos escribirlo como:

$$R_{jt} = \alpha_j + \beta_{jm} R_{mt} + \epsilon_{jt} \qquad t = 1, 2, \ldots, T \tag{3.10}$$

donde $E(\epsilon_{jt}) = 0$ y ϵ_{jt} es independiente de R_{mt}. Habitualmente R_{mt} es un índice de mercado que utilizamos para aproximar la verdadera cartera de mercado.

Estimamos el modelo por MCO. El modelo a estimar, para cada j en forma matricial es:

$$\begin{bmatrix} R_{j1} \\ R_{j2} \\ \vdots \\ R_{jT} \end{bmatrix} = \begin{bmatrix} 1 & R_{m1} \\ 1 & R_{m2} \\ \vdots & \vdots \\ 1 & R_{mT} \end{bmatrix} \begin{bmatrix} \alpha_j \\ \beta_j \end{bmatrix} + \begin{bmatrix} \epsilon_{j1} \\ \epsilon_{j2} \\ \vdots \\ \epsilon_{jT} \end{bmatrix} \Longleftrightarrow \begin{array}{ccccccc} R & = & X & \beta & + & \epsilon \\ (T \times 1) & & (T \times 2) & (2 \times 1) & & (T \times 1) \end{array}$$

Los parámetros desconocidos de este modelos son α_j, β_j y $\sigma^2_{\epsilon_j}$ y su estimación es la siguiente:

$$\hat{\beta}_{MCO} = (X'X)^{-1}(X'R)$$

$$\hat{\sigma}^2_{\epsilon} = \frac{\hat{\epsilon}'\hat{\epsilon}}{T-2}$$

siendo $\hat{\epsilon} = R - X\hat{\beta}$

$$Var(\hat{\beta}_{MCO}) = \sigma^2_{\epsilon}(X'X)^{-1}$$

Si nos fijamos en los parámetros del modelo de mercado el riesgo beta de cualquier activo j viene dado por la expresión:

$$\beta_{jm} = \frac{Cov(R_{jt}, R_{mt})}{Var(R_{mt})}$$

Así el estudio del comportamiento del coeficiente beta como medida de riesgo relevante es importante ya que es una medida de riesgo que va más allá de la propia validez del modelo teórico que la sustenta como medida de riesgo relevante. El coeficiente beta como medida de riesgo es utilizado en la gestión de carteras, en estrategias de cobertura de riesgo con activos derivados, en evaluación de la gestión de carteras etc. También se emplea en modelos de valoración que incorporan más de un factor de riesgo sistemático.

Anexo: Algunas definiciones:

• Sea P_t el precio de un activo en t, (sin pago de dividendos)
• El rendimiento neto entre t y $(t-1)$ es

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

• El rendimiento bruto se define: $1 + R_t$

• El rendimiento bruto sobre los τ periodos más recientes, de $(t-\tau)$ a t, se denota $(1 + R_t^\tau)$ y es igual al producto de los τ rendimientos de un solo periodo desde $t - \tau + 1$ hasta t

$$\begin{aligned} 1 + R_t^\tau &\equiv (1+R_t)(1+R_{t-1})\dots(1+R_{t-\tau+1}) \\ &= \frac{P_t}{P_{t-1}}\frac{P_{t-1}}{P_{t-2}}\frac{P_{t-2}}{P_{t-3}}\cdots\frac{P_{t-\tau+1}}{P_{t-\tau}} \\ &= \frac{P_t}{P_{t-\tau}} \end{aligned}$$

• El rendimiento es una tasa de variación en los precios así que lo correcto es hablar de tasa de rendimiento. Además son rendimientos compuestos.

• El rendimiento continuamente compuesto es el logaritmo natural de su rendimiento bruto:

$$r_t \equiv ln(1 + R_t) = ln\frac{P_t}{P_{t-1}} = p_t - p_{t-1}$$

siendo $p_t = lnP_t$

• En general la evidencia empírica sobre los modelos de valoración cuando se estudian en un contexto de sección cruzada, donde suele ser habitual el empleo de carteras, se basa en rendimientos simples mientras que la evidencia sobre el comportamiento temporal de los activos suele estar basada en rendimientos continuamente compuestos.

• El rendimento de una cartera se define:

$$R_{ct} = \sum_{j=1}^{N} \omega_{jc} R_{jt}$$

es decir es la suma ponderada de los rendimientos de los activos que forman la cartera.

Si las ponderaciones de todos los activos que forman la cartera son iguales se le llama equiponderada. Si las ponderaciones se basan en su capitalización se le dice índice construido por capitalización bursátil.

Un índice bursátil no es más que una cartera donde participan todos los activos que cotizan en el mercado, es por tanto una cartera de mercado.

Podemos distinguir por construcción dos:

I) Indice equiponderado:

$$R_{EW,t} = \sum_{j=1}^{N} \omega_j R_{j,t} = \frac{\sum_{j=1}^{N} R_{j,t}}{N}$$

por tanto $\omega_j = 1/N$

II) Indice por capitalización bursátil

$$R_{VW,t} = \sum_{j=1}^{N} \omega_j R_{j,t}$$

donde

$$\omega_j = \frac{P_t \times \text{n}^{\text{o}} \text{ de acciones desembolsadas}}{\sum_{i=1}^{N} \text{capital de bolsa}}$$

Bibliografía

[1] Affleck-Graves, J., y B. MacDonald, (1989), "Nonnormalities and Tests of Asset Pricing Theories," *Journal of Finance*, 44, 889-908.

[2] Anderson, T., (1984), *An Introduction to Multivariate Statistical Analysis*, Segunda Edición, John Wiley and Sons, New York.

[3] Baesel, J. B. (1974): "On the Assessment of Risk: Some Further Considerations", *Journal of Finance* 29, 1491-1494.

[4] Basarrate, B., y G. Rubio, (1990), "A Note on the Seasonality in the Risk-Return Relation-ship", *Investigaciones Económicas*, 14, 311-318.

[5] Basarrate, B., y G. Rubio, (1994), "La imposición sobre plusvalías y minusvalías: sus efectos sobre el comportamiento estacional del mercado de valores", *Revista Española de Economía*, 11, 247-277.

[6] Black, F., (1972), "Capital Market Equilibrium with Restricted Borrowing," *Journal of Business*, 45, 444-454.

[7] Black, F., Jensen, M., y M. Scholes, (1972), "The Capital Asset Pricing Model: Some Empirical Tests", en *Studies in the Theory of Capital Markets*, M. Jensen editor, Praeger, Nueva York.

[8] Blattberg, R., y N. Gonedes, (1974), " A comparison of stable and Student Distributions as Statistical Models for Stock Prices," *Journal of Business*, 47, 244-280.

[9] Blume, M (1968): The Assessments of Portfolio Performance, PhD Thesis, University of Chicago: Graduate School of Business.

[10] Blume, M (1971): "On the Assessment of Risk", *Journal of Finance* 26, 1-10.

[11] Blume, M. (1975): "Betas and Their Regression Tendencies", *Journal of Finance* 30, 785-795.

[12] Campbell, J., Y., A., W., Lo, y A. C. MacKinlay, (1997), *The Econometrics of Financial Markets*, Princeton University Press, Princeton, New Jersey.

[13] Chamberlain, G. y M. Rothschild, (1983), "Arbitage and Mean Variance Analysis on Large Asset Market", *Econometrica*, vol. 51, 1281-1304.

[14] Chen, N., Roll, R., y S. Ross (1986), "Economic Forces And the Stock Market", *Journal of Business*, 59, 383-403.

[15] Cochrane, J., (2001), *Asset Pricing*, Princeton University Press, Princeton, New Jersey.

[16] Connor, G. y R. Korajczyk, (1986), "Economic Forces and the Stock Market", *Econometrica*, vol. 51, 1281-1304.

[17] Connor, G. y R. Korajczyk, (1988), "Risk and Return in an Equilibrium APT: Application of a New Test Methodology", *Journal of Financial Economics*, 21, 255-289.

[18] Connor, G. y R. Korajczyk, (1993), "A Test for the Number of Factors in an Approximate Factor Model", *The Journal of Finance*, 48, 1263-1291.

[19] Connor, G. y R. Korajczyk, (1995), "The Arbitrage Pricing Theory and Multifactor Models of Asset Returns", en *Handbook in Operations Research and Management Science*, vol. 9, eds. R. Jarrow, V. Maksimovic y W. Ziemba, North-Holland.

[20] Esteban, M. V., (1997), "Variabilidad predecible en los rendimientos de los activos: evidencia e implicaciones", *Investigaciones Económicas. Segunda Epoca.* vol. XXI, 523-542.

[21] Fama, E., (1965), "The Behavior os Stock Markets Prices," *Journal of Business*, 38, 34-105.

[22] Fama, E., (1976), *Foundations of Finance*, Basic Books, New York.

[23] Fama E., y K. French, (1992), "The Cross-Section of Expected Returns", *Journal of Finance*, 47, 427-465.

[24] Fama, E., y J. MacBeth, (1973), "Risk, return, and Equilibrium," *Journal of Political Economy*, 71, 607-636.

[25] Fisher, L. (1970): The Estimation of Systematic Risk: Some New Findings, in "Proceedings of the Seminar on the Analysis of Security Price", University of Chicago, May.

[26] Gibbons, M., S. Ross, y J. Shanken, (1989), "A Test of Efficiency of a Given Portfolio," *Econometrica*, 57, 1121-1152.

[27] Gonedes, N. (1973): "Evidence on the Information Content of Accounting Numbers: Accounting-based and Market-based Estimates of Systematic Risk", *Journal of Financial and Quantitative Analysis* 8, 407-443.

[28] Gordon, J. A. and Chervany, N. L. (1980): "On the Estimation and Stability of Beta", *Journal of Financial and Quantitative Analysis* XV, 123-137.

[29] Hawawini, G. y D. Keim, (995), "On the Predictability of Common Stock Returns : World-Wide Evidence" en *Hadbook in Operations Research and Management Science*, vol. 9, eds. R. Jarrow, V. Maksimovic y W. Ziemba, North-Holland.

[30] Jagannathan, R. y Z. Wang (1996), "The Conditional CAPM and the Cross-Section of Expected Returns", *Journal of Finance*, 51, 3-53.

[31] Jegadeesh, N. y S. Titman, (1993), "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency", *Journal of Finance*, 48, 65-91.

[32] Jobson, D., y R. Korkie, (1982), "Potential Performance and Test of Portfolio Efficiency," *Journal of Financial Economics*, 10, 443-466.

[33] Kandel, S., (1984), "The Likelihood Ratio Test of Mean-Variance Efficiency without a Riskless Asset," *Journal of Financial Economics*, 13, 575-592.

[34] Kandel, S., y R. Stambaugh, (1995), "Portfolio Inefficiency and the Cross-Section of Expected Returns," *Journal of Financial Economics*, 18, 61-90.

[35] Keim, D., (1983), "Size-Related Anomalies and Stock Return Seasonality: Further Empirical evidence", *Journal of Financial Economics*, 12, 13-32.

[36] Kothari, S., Shanken, J. Y R. Sloan, (1995), "Another Look at the Cross-Section of Expected Stock Returns", *Journal of Finance*, 50, 185-224.

[37] Lintner, J,, (1965b), "The Valuation of Risky Asset and the Selection of Risky Investment in Stock Portfolios and Capital Budgets," *Review of Economics ans Statistics*, 47, 13-37.

[38] Litzenberger, R., y K. Ramaswamy, (1979), " The Effect of personal Taxes and Dividends on Capital Asset Prices: Theory and Evidence," *Journal of financial Economics*, 7, 163-196.

[39] Marín, J. M. y Rubio, G., (2001), *Economía Financiera*, Antoni Bosch Ed., Barcelona.

[40] Markowitz, H., (1959), *Portfolio Selection: Efficient Diversification of Investments*, John Wiley, New York. Marín, J. M. y Rubio, G., (2001), *Economía Financiera*, Antoni Bosch Ed., Barcelona.

[41] Merton, R. C. (1973), "An intertemporal Capital Asset Pricing Model", *Econometrica*, vol. 41, 867-887.

[42] Muirhead, R., (1983), *Aspects of Multivariate Statistical Theory*, John Wiley and Sons, New York.

[43] Newey, W., y K. West, (1987), "A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703-708.

[44] Roll, R. (1977), "A Critique of the Asset Pricing Theory's Tests: Part I: On Past and Potential Testability of the Theory", *Journal of Financial Economics*, 4, 129-176.

[45] Roll, R., y S. Ross, (1994), " On the Cross-Sectional Relation between Expected Returns and Betas, " *Journal of Finance*, 49, 101-122.

[46] Ross, S. A. (1976), "The Arbitrage Theory of Capital Asset Pricing", *Journal of Economic Theory*, 13, 341-360.

[47] Rubio, G. (1988), "Further International evidence on asset Pricing: The Case of the Spanish Capital Market", *Journal of Banking and Finance*, 12, 221-242.

[48] Rubio, G., (2001), *Economía Financiera*, Antoni Bosch Ed., Barcelona.

[49] Shanken, J., (1986), "Testing Portfolio Efficiency When the Zero-Beta Rate Is Unknown," *Journal of Finance*, 41, 269-276.

[50] Shanken, J., (1992b), "On the Estimation of Beta-Pricing Models," *Review of Financial Studies*, 5, 1-34.

[51] Shanken, J., (1985), "Multivariate Tests of the Zero-Beta CAPM," *Journal of Financial Economics*, 14, 327-348.

[52] Shanken, J., (1996), "Statistical Methods in Test of Portfolio Efficiency: A Synthesis," en *Handbook of Statistics*, vol. 14, eds, S. Maddala y C. Rao, Elsevier Sciences.

[53] Sharpe, W., (1964), "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance*, 19, 425-442.

[54] White, H., (1980), "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48, 817-838.