miento de emisiones a precios mayores), el riesgo es mayor. Es por esto por lo que normalmente interesa más el pago único que el aplazado.

Como muestra de ello, reseñaremos que de las ampliaciones de capital producidas desde octubre de 1990 a marzo de 1991, el 80 % aproximadamente de las mismas fueron desembolsadas en su totalidad en el momento de la suscripción, constatándose esta que pudiéramos llamar conducta racional del mercado, inspirada en la liquidez del mismo.

El valor resultante de la fórmula mencionada (la prima o precio de mercado del DPS) multiplicado por la relación acciones nuevas/acciones antiguas será el valor del DPS en función de las variables que influyen en él (que ya han sido resaltadas), aunque no olvidemos que el mercado rara vez considera la bondad de esta valoración, puesto que no está corregida por el período en que no se dispone del activo, y que supone claros riesgos para el inversor.

Estos riesgos que soporta el inversor son de dos tipos:

a) Riesgo financiero: asociado a la imposibilidad de revertir el dinero de la inversión durante el tiempo en que ésta no sea líquida. Es conocido, ya que es el coste del dinero.

b) Riesgo económico o «prima por riesgo» que el inversor aplicará al coste del derecho y que le suponga cierta garantía ante una bajada generalizada de las cotizaciones durante el tiempo que no es líquida esta inversión. Esta prima resulta mucho más difícil de estimar, ya que depende tanto de factores propios de la empresa que amplíe capital como del propio mercado. Pudiéramos realizar estimaciones de esta prima a aplicar a través del análisis fundamental, por ejemplo.

A este respecto, Fernández estudia empíricamente estas primas por riesgo (tanto económico como financiero) en el período 1985-1988, aunque sin tener en cuenta el tiempo que las empresas tardaron en incorporar a la Bolsa las nuevas acciones, en el que esas primas se sitúan en torno al 10 % del valor teórico del derecho arrojado por su valoración como opción financiera.

Como consecuencia a todo lo hasta aquí comentado y como juego de la oferta y demanda, en la que como señalábamos con anterioridad compradores y vendedores se encuentran regidos por motivaciones distintas, se formará el valor de mercado del DPS, sobre el que podrán realizarse arbitrajes y operaciones de especulación, descartando nosotros la de cobertura de riesgos, que son las tres utilidades que poseen los instrumentos derivados.

Muestreo por cuotas: Consideraciones desde un modelo de superpoblación

José Miguel Casas Sánchez
Universidad de Alcalá de Henares

Marta Guijarro Garvi
Universidad de Cantabria

1. INTRODUCCIÓN

El muestreo por cuotas, como se sabe, es bastante parecido al muestreo estratificado; así pues, se conoce la distribución de una característica h $(h = 1,...,H)$ que poseen N_h individuos de la población. Para cada característica se seleccionan, por parte del entrevistador, n_h elementos. La fracción de muestreo, $f_h = n_h / N_h$, puede ser diferente para cada característica.

En el muestreo por cuotas, la población de tamaño N se clasifica según las diferentes características $i,j,...,h$ $(i = 1,...,I; j = 1,...,J; ... h = 1,...,H)$, siendo $N_{i,j,...,h}$ el número de elementos de la población correspondientes a cada una de las casillas de una tabla de contingencia múltiple, y $n_{i,j,...,h}$ el tamaño de la respectiva muestra, cuya fracción de muestreo representaremos por $f_{i,j,...,h}$. Emplearemos *cuotas marginales* y adoptaremos una única fracción de muestreo, f, para cada conjunto de cuotas, es decir (1):

$$n_{i\cdot\ldots\cdot} = fN_{i\cdot\ldots\cdot}; n_{\cdot j\ldots\cdot} = fN_{\cdot j\ldots\cdot}; \ldots n_{\cdot\cdot\ldots h} = fN_{\cdot\cdot\ldots h}$$

En este tipo de muestreo, el plan de muestreo es desconocido por el estadístico, es un muestreo no probabilístico y, consecuentemente, no es posible hacer inferencias sin utilizar algún modelo. Para ello, haremos uso de un modelo de superpoblación en el que la población finita es generada como una muestra aleatoria de una superpoblación infinita, es decir, el vector de valores poblaciones de la variable objeto de

(1) La sustitución del subíndice por · significa sumar en todas las modalidades de la característica representada por el subíndice.

estudio es una realización de un vector de variables aleatorias cuya distribución conjunta se puede formular a partir de un modelo de análisis de la varianza.

Nuestro objetivo final será estimar la media poblacional de una cierta variable, mediante un estimador basado en el modelo, estudiando su ξ-insesgadez y su error cuadrático medio con el diseño de muestreo formulado.

2. MUESTREO POR CUOTAS CON MODELOS DE SUPERPOBLACIONES

Consideramos un modelo de superpoblación cuya formulación sea análoga a la de un modelo de análisis de la varianza aditivo y con interacción (2).

Si designamos por Y_{ijk} $(i = 1,...,I; j = ,1,...,J; k = 1,...,N_{ij})$ el valor de la variable correspondiente al k-ésimo elemento de la casilla (i,j) de una tabla de contingencia, podremos escribir:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk} \qquad [1]$$

donde las variables ϵ_{ijk} son independientes y tales que:

$$E_\xi(\varepsilon_{ijk}) = 0$$

$$Var_\xi(\varepsilon_{ijk}) = \sigma_{ij}^2 = \tau_i^2 + \omega_j^2$$

siendo σ_{ij}^2 constante desconocida (3).

Partiendo del modelo anterior, la estimación de la media poblacional se basará en la estimación de los valores de la variable que no están en la muestra, es decir, el *estimador predictivo* (4) adoptará la forma:

$$e = \frac{1}{N}\left(\sum_i\sum_j\sum_{k\in s} Y_{ijk} + \sum_i\sum_j\sum_{k\in r} \hat{Y}_{ijk}\right) \qquad [2]$$

(2) Este enfoque generaliza, en cierto modo, el analizado por Deville (1991), ya que incluye la interacción entre ambos factores o características.

(3) Admitimos que que se verifican las condiciones de identificabilidad:

$$\sum_i N_{ij}\gamma_{ij} = \sum_j N_{ij}\gamma_{ij} = 0 \qquad [3]$$

(4) Algunos autores, Royall (1970), Herson y Royall (1973), denominan a estos estimadores *estimadores basados en el modelo*, ya que se obtienen únicamente a partir de consideraciones que se deducen del modelo de trabajo, sin tener en cuenta el diseño muestral utilizado.

donde s y r indican las unidades de la población que están y no están en la muestra, respectivamente.

La estimación de los parámetros por mínimos cuadrados ordinarios nos lleva a:

$$\hat{\mu} = \bar{y}$$
$$\hat{\alpha}_i = \bar{y}_{i\cdot} - \bar{y} \quad \forall i$$
$$\hat{\beta}_j = \bar{y}_{\cdot j} - \bar{y} \quad \forall j$$
$$\hat{\gamma}_{ij} = \bar{y}_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}$$

donde $\bar{y}$ es la media de los valores muestrales, $\bar{y}_{i\cdot}$ e $\bar{y}_{\cdot j}$ son las medias de los valores de la muestra en las modalidades i-ésima y j-ésima del primer y segundo factor, respectivamente, y, por último, $\bar{y}_{ij}$ es la media de los valores muestrales de la casilla (i,j).

Sustituyendo en la expresión [3] del estimador predictivo, tendremos:

$$\begin{aligned} e &= \frac{1}{N}\sum_i\sum_j n_{ij}\bar{y}_{ij} + \\ &+ \frac{1}{N}\sum_i\sum_j (N_{ij} - n_{ij})(\bar{y} + \bar{y}_{i\cdot} - \bar{y} + \bar{y}_{\cdot j} - \bar{y} + \bar{y}_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}) = \\ &= \frac{1}{N}\sum_i\sum_j N_{ij}\bar{y}_{ij} \end{aligned} \qquad [4]$$

Teorema 2.1

Bajo el modelo de superpoblación [1], el estimador predictivo

$$e = \frac{1}{N}\sum_i\sum_j N_{ij}\bar{Y}_{ij}$$

es ξ-insesgado y su varianza es:

$$Var_\xi(e) = \frac{1}{N^2}\sum_i\sum_j N_{ij}\left(\frac{N_{ij}}{n_{ij}} - 1\right)\sigma_{ij}^2$$

Demostración

La primera parte es inmediata, ya que basta considerar que la media respecto al modelo de los valores de la variable en cada casilla es constante.

Para llegar a obtener la varianza del estimador, nos basaremos en que es ξ-insesgado pudiendo, así, escribir:

$$
\begin{aligned}
& Var_\xi(e) = E_\xi(e-\bar{Y})^2 = \\
&= \frac{1}{N^2}E_\xi\left[\sum_i\sum_j\frac{N_{ij}}{n_{ij}}\sum_{k=1}^{n_{ij}}(Y_{ijk}-E_\xi(Y_{ijk}))-\sum_i\sum_j\sum_{k=1}^{N_{ij}}(Y_{ijk}-E_\xi Y_{ijk})\right]^2 = \\
&= \frac{1}{N^2}E_\xi\left(\sum_i\sum_j\frac{N_{ij}}{n_{ij}}\sum_{k=1}^{n_{ij}}\varepsilon_{ijk}-\sum_i\sum_j\sum_{k=1}^{N_{ij}}\varepsilon_{ijk}\right)^2 = \\
&= \frac{1}{N^2}\left[\sum_i\sum_j\frac{N_{ij}^2}{n_{ij}^2}\sum_{k=1}^{n_{ij}}E_\xi(\varepsilon_{ijk})^2+\sum_i\sum_j\sum_{k=1}^{N_{ij}}E_\xi(\varepsilon_{ijk})^2\right]- \\
&- \frac{2}{N^2}\sum_i\sum_j\frac{N_{ij}}{n_{ij}}\sum_{k=1}^{n_{ij}}E_\xi(\varepsilon_{ijk})^2 = \\
&= \frac{1}{N^2}\left(\sum_i\sum_j\frac{N_{ij}^2}{n_{ij}}\sigma_{ij}^2+\sum_i\sum_j N_{ij}\sigma_{ij}^2-2\sum_i\sum_j N_{ij}\sigma_{ij}^2\right) = \\
&= \frac{1}{N^2}\left[\sum_i\sum_j N_{ij}\left(\frac{N_{ij}}{n_{ij}}-1\right)\sigma_{ij}^2\right] \qquad \text{c.q.d.}
\end{aligned}
$$

Además, como:

$$\hat{\sigma}_{ij}^2 = \frac{1}{N}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij}}(Y_{ij}-\bar{y}_{ij})^2$$

es un estimador ξ-insesgado de σ_{ij}^2 tendremos que

$$\widehat{Var_\xi}(e) = \frac{1}{N^2}\sum_{i=1}^{I}\sum_{j=1}^{J}N_{ij}\left(\frac{N_{ij}}{n_{ij}}-1\right)\hat{\sigma}_{ij}^2$$

es un estimador ξ-insesgado de la varianza del estimador predictivo.

3. DISEÑO MUESTRAL EN EL MUESTREO POR CUOTAS (5)

En el caso de cuotas marginales, el muestreo aleatorio simple selecciona muestras con n_{ij} elementos en diferentes casillas que pueden ser considerados como un vector aleatorio en el espacio R^{IJ}, del tal manera que:

(5) En este apartado seguiremos el trabajo de Deville (1991), en donde se utiliza el muestreo aleatorio simple restringido a las cuotas marginales; sin embargo, la estimación que aquí realizaremos será para la media y no para el total poblacional, como se indica allí.

$$\sum_j n_{ij} = n_{i\cdot} \quad (i=1,\dots,I); \quad \sum_i n_{ij} = n_{\cdot j} \quad (j=1,\dots,J-1)$$

donde la variable n_{ij}, número de elementos en cada casilla, sigue una distribución multinomial de parámetros n y $p_{ij} = N_{ij}/N$.

Según esto, y sin más que sustituir en la expresión [4], el estimador predictivo de la media poblacional se puede escribir como:

$$e = \sum_i\sum_j p_{ij}\bar{y}_{ij}$$

siendo $p_{ij} = N_{ij}/N$ conocido.

En la situación planteada, nuestro objetivo es obtener un estimador de la forma:

$$e = \sum_i\sum_j \hat{p}_{ij}\bar{y}_{ij}$$

donde $\hat{p}_{ij} = \hat{N}_{ij}/N$ es el estimador máximo-verosímil de p_{ij}.

Aplicando el método de estimación de la máxima verosimilitud, llegamos a obtener:

$$\hat{p}_{ij} = \frac{n_{ij}}{n}(a_i+b_j)^{-1}$$

siendo a_i, b_j, $(i = 1,\dots,I; j = 1,\dots,J-1)$, soluciones del sistema de ecuaciones:

$$
\left.
\begin{aligned}
&\sum_j \frac{n_{ij}}{n}(a_i+b_j)^{-1} = \frac{N_{i\cdot}}{N} \quad i=1,\dots,I \\
&\sum_i \frac{n_{ij}}{n}(a_i+b_j)^{-1} = \frac{N_{\cdot j}}{N} \quad j=1,\dots,J-1
\end{aligned}
\right. \qquad [5]
$$

y el estimador predictivo de la media poblacional tendrá la forma:

$$e = \sum_i\sum_j \frac{n_{ij}}{n}(a_i+b_j)^{-1}\bar{y}_{\cdot j}$$

3.1. Estudio de la ξ-insesgadez del estimador

Veamos que el ξ-sesgo del estimador, e, es igual a:

$$\sum_i \sum_j \frac{\hat{N}_{ij}}{N}\gamma_{ij}$$

En efecto, teniendo en cuenta que:

$$\bar{Y} = \sum_i \sum_j \left[\frac{N_{ij}}{N}(\mu + \alpha_i + \beta_j + \gamma_{ij}) + \frac{1}{N}\sum_{k=1}^{N_{ij}} \varepsilon_{ijk} \right]$$

y

$$e = \sum_i \sum_j \left[\frac{\hat{N}_{ij}}{N}(\mu + \alpha_i + \beta_j + \gamma_{ij}) + \frac{\hat{N}_{ij}}{n_{ij}N}\sum_{k=1}^{n_{ij}} \varepsilon_{ijk} \right]$$

bastará tomar esperanzas con respecto al modelo ξ para obtener:

$$E_\xi(e - \bar{Y}) = \sum_i \sum_j \frac{\hat{N}_{ij} - N_{ij}}{N}(\mu + \alpha_i + \beta_j + \gamma_{ij})$$

Ahora bien, puesto que $\hat{N}_{ij} = N\,(n_{ij}/n)\,(a_i + b_j)^{-1}$, las ecuaciones del sistema [5] nos llevan de modo inmediato a las condiciones:

$$\begin{aligned} \sum_i \frac{\hat{N}_{ij}}{N} &= \sum_i \frac{N_{ij}}{N} \\ \sum_j \frac{\hat{N}_{ij}}{N} &= \sum_j \frac{N_{ij}}{N} \end{aligned} \qquad [6]$$

que, junto con la aplicación de las condiciones de identificabilidad [3], nos permiten demostrar el resultado.

Evidentemente, si los valores γ_{ij} son prácticamente nulos, el ξ-sesgo se acerca a cero. Es decir, cuanto menor sea la interacción entre los factores, menor será el ξ-sesgo del estimador, llegando a ser cero cuando no exista interacción.

Cabe señalar que, para un tamaño muestral suficientemente grande, el ξ-sesgo será aproximadamente nulo, puesto que, bajo condiciones de regularidad, es posible asegurar que los estimadores máximo-verosímiles son consistentes (6).

(6) Ver Kendall y Stuart (1961).

3.2. Error cuadrático medio del estimador respecto al modelo

Teorema 3.2.1

Bajo el modelo [1],

$$E_\xi(e - \bar{Y})^2 = [b_\xi(e)]^2 + \frac{1}{N^2}\sum_i \sum_j \hat{N}_{ij}\left(\frac{\hat{N}_{ij}}{n_{ij}} - 1\right)\sigma_{ij}^2$$

con $b_\xi(e)$ igual al ξ-sesgo del estimador e.

Demostración

Dado que:

$$e - \bar{Y} = \sum_i \sum_j \left[\frac{\hat{N}_{ij}}{N}\gamma_{ij} + \frac{1}{N}\left(\frac{\hat{N}_{ij}}{n_{ij}} - 1\right)\sum_{k=1}^{n_{ij}} \varepsilon_{ijk} - \frac{1}{N}\sum_{k=n_{ij}+1}^{N_{ij}} \varepsilon_{ijk} \right]$$

Elevando al cuadrado y tomando esperanzas respecto al modelo resultará:

$$\begin{aligned} E_\xi(e - \bar{Y})^2 &= \left(\sum_i \sum_j \frac{\hat{N}_{ij}}{N}\gamma_{ij}\right)^2 + \\ &+ \frac{1}{N^2}\sum_i \sum_j \left[\left(\frac{\hat{N}_{ij}}{n_{ij}} - 1\right)^2 n_{ij} + (N_{ij} - n_{ij})\right](\tau_i^2 + \omega_j^2) = \\ &= [b_\xi(e)]^2 + \frac{1}{N^2}\sum_i \sum_j \left(\frac{\hat{N}_{ij}^2}{n_{ij}} - 2\hat{N}_{ij} + N_{ij}\right)(\tau_i^2 + \omega_j^2) \end{aligned}$$

Aplicando las condiciones [6], podremos escribir la expresión anterior:

$$E_\xi(e - \bar{Y})^2 = [b_\xi(e)]^2 + \frac{1}{N^2}\sum_i \sum_j \hat{N}_{ij}\left(\frac{\hat{N}_{ij}}{n_{ij}} - 1\right)(\tau_i^2 + \omega_j^2) \qquad [7]$$

c.q.d.

Cuando $b_\xi(e)$ es cero, el error cuadrático medio coincide con la varianza del estimador, que en ese caso sería:

$$Var_\xi(e) = \frac{1}{N^2}\sum_i \sum_j \hat{N}_{ij}\left(\frac{\hat{N}_{ij}}{n_{ij}} - 1\right)\sigma_{ij}^2$$

Lógicamente, la expresión [7] coincide con el resultado obtenido en el teorema 2.1 cuando los valores N_{ij} son conocidos.

BIBLIOGRAFÍA

AZORÍN, F. y SÁNCHEZ-CRESPO, J. L. (1986), *Métodos y aplicaciones del muestreo*, Madrid, Alianza.

DEVILLE, J. (1991), «A theory of quota surveys», *Survey Methodology* 17, n.º 2, 163-181.

GOURIEROUX, C. (1981), «Theory des sondages», *Economica.*

HERSON, J. y ROYALL, R. M. (1973), «Robust estimation in finite populations», *Journal of American Satistical Association* 68, 880-893.

KENDALL, M. G. y STUART, A. (1967), *The advanced theory of statistics*, Nueva York, Hafner.

ROHATGI, V. K. (1976), *An introduction to probability theory and mathematical statistics,* Nueva York, John Wiley.

ROYALL, R. M. (1970), «On finite population sampling theory under certain linear regression models», *Biometrika* 57, 377-387.

TAM, S. M. (1988b), «Some results on robust estimation in finite population sampling», *Journal of American Statistical Association* 83, 242-248.

WOLTER, K. M. (1985), *Introduction to variance estimation,* Nueva York, Springer-Verlag.

Uso y abuso del coeficient de determinación (I

JOSÉ HERNÁNDEZ ALON
Profesor de ES

1. INTRODUCCIÓN

Analizada en la primera parte del estudio las imperfecciones del coeficiente determinación, en cuanto medida para describir relaciones entre variables, ante tructuras particulares de los datos, o si el tamaño de información en que se basa regresión estadística es pequeño, pasamos ahora a evaluar la capacidad del mismo pa medir esta relación en series temporales y en series con transformaciones previas datos.

En el primer caso, la evolución regular creciente en el tiempo, sin cambios de t dencia, que les es característica a las series temporales, suele ser el motivo para q de forma relativamente fácil pueda aproximarse una ecuación de regresión y obten coeficientes de determinación elevados, sin que, por otra parte, sea esto el signo e dente de una relación causal adecuada.

A su vez, en el segundo caso, ciertas transformaciones previas de los datos (p ejemplo la toma de logaritmos), para ajustar mediante la técnica de regresión lin ecuaciones no lineales, pueden implicar, también, un incremento impropio en el c ficiente de determinación y generar errores de interpretación, como destacaremo continuación.

En las páginas que siguen se acude, como en la primera parte del estudio, al c curso de ejemplos numéricos, que invalidan la utilización abusiva de este coeficie ante las nuevas situaciones, y se propondrán, igualmente, medidas alternativas o co plementarias.

En todo caso, ante las limitaciones de interpretación del R^2 e incluso de las m didas alternativas propuestas, se tratará de poner de manifiesto cómo la evaluaci