

SOLUCIONS ALS PROBLEMES PROPOSATS AL VOLUM 11. N° 2

PROBLEMA N° 15

Utilitzarem la següent extensió multivariant del Teorema Central del Límit (TCL) (veure, per exemple, Serfling, 1980, p. 28):

Si $Z_i, i = 1, 2, \dots, n$, són vectors aleatoris idempendents i identicament distribuïts, amb esperança μ i matriu de variàncies i covariàncies Σ , aleshores

$$n^{1/2} \left(n^{-1} \sum_{i=1}^{i=n} Z_i - \mu \right) \xrightarrow{L} N(\mu, \Sigma),$$

on " \xrightarrow{L} " denota convergència en llei (distribució), per n tendint a infinit, i $N(\mu, \Sigma)$ denota la distribució normal multivariant d'esperança μ i matriu de variàncies i covariàncies Σ .

Utilitzarem també els següents resultats típics de teoria asymptòtica (consulteu, per exemple, Serfling, 1980, p. 19):

Si $Z_n \xrightarrow{L} Z$, i $Y_n \xrightarrow{P} c$, on c és una constant i " \xrightarrow{P} " denota convergència en probabilitat, aleshores

- (A) $Z_n + Y_n \xrightarrow{L} Z + c$
- (B) $Z_n Y_n \xrightarrow{L} cZ$

Passem ara a resoldre el problema. Per començar, nota que

$$\begin{aligned} \sum_{i=1}^{i=n} (X_i - \bar{X})(X_i - \bar{X})' &= \\ &= \sum_{i=1}^{i=n} (X_i - \mu_X)(X_i - \mu_X)' + n(\bar{X} - \mu_X)(\bar{X} - \mu_X)', \end{aligned}$$

on μ_X i \bar{X} són els vectors mitjana poblacional i mostra, respectivament, definits a l'enunciat del problema. Per tant, si S és la matriu de variàncies i covariàncies mostra, tindrem que

$$(1) \quad n^{1/2}S = n^{1/2}(n-1)^{-1} \sum_{i=1}^{i=n} (X_i - \mu_X)(X_i - \mu_X)' + n^{1/2}(n-1)^{-1}n(\bar{X} - \mu_X)(\bar{X} - \mu_X)'.$$

Aplicant ara el TCL enunciat anteriorment, pel cas en que $Z_i \equiv X_i$, s'obté

$$(2) \quad n^{1/2}(\bar{X} - \mu_X) \xrightarrow{L} N(0, \Sigma),$$

i, per tant,

$$(3) \quad n^{1/2}(n-1)^{-1}n(\bar{X} - \mu_X)(\bar{X} - \mu_X)' \xrightarrow{P} 0,$$

ja que $n^{1/2}(n-1)^{-1} \rightarrow 0$ i $n(\bar{X} - \mu_X)(\bar{X} - \mu_X)'$ està afitada en probabilitat (degut a (2)).

Definint ara

$$(4) \quad d_i \equiv \text{vec } (X_i - \mu_X)(X_i - \mu_X)',$$

on “vec” és un operador que vectoritza matrius (per exemple, posant una columna sota de l’altra), i aplicant de nou el TCL, ara pel cas en que $Z_i \equiv d_i$, obtenim

$$(5) \quad n^{1/2} \left(n^{-1} \sum_{i=1}^{i=n} d_i - E d_i \right) \xrightarrow{L} N(0, \text{var}(d_i)),$$

on “var” denota matriu de variàncies i covariàncies. Per tant,

$$(6) \quad n^{1/2} \left((n-1)^{-1} \sum_{i=1}^{i=n} \text{vec } (X_i - \mu_X)(X_i - \mu_X)' - \text{vec } \Sigma \right) \\ \xrightarrow{L} N(0, \text{var}(d_i)),$$

que s’obté de (5) després de substituir d_i per la seva expressió (veure (4)), $E d_i$ pel resultat de càlcul següent:

$$E d_i = E \text{vec } (X_i - \mu_X)(X_i - \mu_X)' = \text{vec } E(X_i - \mu_X)(X_i - \mu_X)' = \text{vec } \Sigma,$$

i, finalment, n^{-1} per $(n-1)^{-1}$ (que, obviament, tenint en compte (A) i (B), no altera la distribució límit).

Combinant ara (1), (3) i (6) i aplicant (A), obtenim

$$(7) \quad n^{1/2}(\text{vec } S - \text{vec } \Sigma) \xrightarrow{L} N(0, \text{var}(d_i)).$$

Nota que hem utilitzat (A) amb Y_n igual al terme que apareix a (3) i $c = 0$.)

Per altra banda, és obvi que

$$\text{var } (d_i) = Ed_i d'_i - Ed_i Ed'_i = Ed_i d'_i - (\text{vec } \Sigma)(\text{vec } \Sigma)',$$

i que

$$Ed_i d'_i = E(\text{vec } (X_i - \mu_X)(X_i - \mu_X)') (\text{vec } (X_i - \mu_X)(X_i - \mu_X)')' = \\ E((X_i - \mu_X) \otimes (X_i - \mu_X))((X_i - \mu_X) \otimes (X_i - \mu_X))'.$$

on, en l'última igualtat, hem utilitzat propietats bàsiques de la relació entre producte tensorial i l'operador "vec" (e.g., Magnus i Neudecker 1988). Per tant,

$$(8) \quad \text{var } (d_i) = \Sigma_{XXXX} - (\text{vec } \Sigma)(\text{vec } \Sigma)',$$

on Σ_{XXXX} és la matriu definida l'enunciat del problema. (7) i (8) clouen la resolució del problema.

OBSERVACIONS:

- 1.- Una versió més general d'aquest problema ve resolta a Fuller (1987, Teorema 1.C.1 p. 89-90.)
- 2.- Un estimador consistent de la matriu de variàncies i covariàncies assimptòtica de $\text{vec } S_{XX}$, és a dir, la matriu

$$\Sigma_{XXXX} - (\text{vec } \Sigma)(\text{vec } \Sigma)',$$

és la matriu següent de moments mostrals de quart ordre:

$$H = n^{-1} \sum_{i=1}^{i=n} (b_i - \bar{b})(b_i - \bar{b})',$$

on $b_i \equiv \text{vec } X_i X'_i$ i $\bar{b} = n^{-1} \sum_{i=1}^{i=n} b_i$. (Demostra aquesta propietat de consistència; és un exercici curt.)

- 3.- La matriu anterior de variàncies i covariàncies assimptòtica serà singular, ja que S és una matriu simètrica i, per tant, $\text{vec } S$ conté elements repetits. Si en el desenvolupament anterior substituïm "vec" per "vecs", on vecs és l'operador vectorització simètrica (opera sobre una matriu simètrica,

posem per cas d'ordre $p \times p$, recollint només els $p^* = p(p + 1)/2$ elements distints) aleshores la matriu de variàncies i covariàncies assumptòtica serà de dimensió més reduïda (dimensió $p^* \times p^*$) i no necessàriament singular. L'estimador consistent continuarà essent H , substituint però vec per vecs en la definició de b_i .

- 4.- La matriu H conté informació sobre la variabilitat mostral de S , és a dir, de com varia S al llarg de diferents rèpliques; de la mateixa manera que S informa sobre la variabilitat mostral de \bar{X} . Nota que bona part de l'anàlisi multivariant es dedica a extreure estadístics que són funció (explícita o implícita) de la matriu S (valors propis, correlacions canòniques, estimadors de paràmetres en un model d'equacions estructurals, estadístics de contrast, etc.). Per tant, coneixent la variabilitat mostral de S , podem determinar (via l'anomenat mètode-delta; per exemple, Rao (1973)) la variabilitat mostral d'aquests estadístics. Es a dir, H ens permet calcular l'error estàndard dels estimadors, així com les distribucions assumptòtiques dels estadístics de contrast. Nota que aquesta informació l'obtenim sota hipòtesis distribucionals mínimes: l'única hipòtesi que utilitzem és la de que X té moments de quart-ordre finits. Els errors estàndars i distribucions assumptòtiques obtinguts d'aquesta manera no tenen perquè coincidir amb els més habituals que s'obtenen sota l'hipòtesi que X té distribució normal multivariant. (Per una millor comprensió d'aquest punt, res millor que resoldre el problema que proposarem properament.)

REFERÈNCIES BIBLIOGRÀFIQUES:

- Fuller, W.A.** (1987). "Measurement Error Models". New York, John Wiley.
Serfling, R.J. (1980). "Approximation Theorems of Mathematical Statistics". New York, John Wiley.
Magnus, J.R., & Neudecker, H. (1988). "Matrix Differential Calculus with Applications in Statistics and Econometrics" New York, John Wiley.
Rao, C.R. (1973). "Linear Statistical Inference and its Applications". New York, John Wiley.

Albert Satorra
Universitat de Barcelona

PROBLEMA N° 16

Sean X_1, \dots, X_n las variables aleatorias muestrales correspondientes a una muestra aleatoria simple de tamaño n . Definamos:

$$U = \begin{cases} 1 & \text{si } X_1 = 0 \\ 0 & \text{si } X_1 \neq 0 \end{cases}$$

Entonces:

$$E(U) = P(X_1 = 0) = e^{-\lambda}$$

por tanto U es un estimador insesgado de $\psi(\lambda) = e^{-\lambda}$. Por otra parte $T = \sum_{i=1}^n X_i$ es bien sabido que se trata de un estimador suficiente para λ . Ello permite, aplicando el teorema de Rao-Blackwell, construir un nuevo estimador insesgado de $\psi(\lambda)$, más eficiente. Para ello definamos:

$$\begin{aligned} \varphi(t) &= E(U/T = t) = P(U = 1/T = t) = \frac{P([X_1=0] \cap [T=t])}{P([T=t])} = \\ &= \frac{P(X_1=0) \cdot P(T=t/X_1=0)}{P(T=t)} = \\ &= \frac{e^{-\lambda} e^{-(n-1)\lambda} \frac{\{(n-1)\lambda\}^t}{t!}}{e^{n\lambda} \frac{\{n\lambda\}^t}{t!}} = \left(1 - \frac{1}{n}\right)^t \end{aligned}$$

por tanto el estimador buscado será:

$$W = \varphi(T) = \left(1 - \frac{1}{n}\right)^T = \left(1 - \frac{1}{n}\right)^{\sum_{i=1}^n X_i}$$

W es pues un estimador insesgado de $\psi(\lambda)$ y más eficiente que U . Para demostrar que se trata del estimador de varianza mínima, podremos utilizar el teorema de Lehmann-Scheffé. Para ello sólo tendremos que comprobar que el estadístico suficiente para λ , T , es completo, es decir si $E(f(T)) = 0$ implica $f(T) \equiv 0$. En efecto:

$$E(f(T)) = \sum_{k=0}^{\infty} f(k) \cdot e^{-n\lambda} \frac{(n\lambda)^k}{k!} = e^{-n\lambda} \sum_{k=0}^{\infty} \frac{f(k) n^k}{k!} \lambda^k$$

igualando a cero, resulta:

$$\sum_{k=0}^{\infty} \frac{f(k) n^k}{k!} \lambda^k = 0 \quad \lambda > 0$$

Esto sólo puede verificarse si los coeficientes de dicha serie de potencias, en la variable λ , son identicamente cero:

$$\frac{f(k)n^k}{k!} = 0 \quad k = 0, 1, 2, \dots$$

equivalente a

$$f(k) = 0 \quad k = 0, 1, 2, \dots$$

Por tanto si $E(f(T)) = 0 \Rightarrow f(T) = 0$, con lo que demostramos la completitud de T respecto λ y por tanto al poder aplicar el teorema de Lehmann–Scheffé, concluimos que el estimador W construido anteriormente es el estimador de varianza mínima de $\psi(\lambda)$.

Veamos a continuación si dicho estimador alcanza o no la cota de Cramer–Rao. Calculemos en primer lugar el momento de segundo orden de W :

$$\begin{aligned} E(W^2) &= \sum_{k=0}^{\infty} \left(1 - \frac{1}{n}\right)^{2k} e^{-n\lambda} \frac{(n\lambda)^k}{k!} = \\ &= e^{-n\lambda} \sum_{k=0}^{\infty} \frac{\{(1 - \frac{1}{n})^2 n\lambda\}}{k!} = e^{-n\lambda} e^{(1 - \frac{1}{n})^2 n\lambda} = \\ &= e^{(\frac{1}{n}-2)\lambda} \end{aligned}$$

por tanto, al ser $E(W) = e^{-\lambda}$, resulta:

$$\text{var}(W) = e^{(\frac{1}{n}-2)\lambda} - e^{-2\lambda} = e^{-2\lambda} \left(e^{\frac{\lambda}{n}} - 1\right)$$

Por otra parte, la cota de Cramer–Rao para $\varphi(\lambda)$ es igual a:

$$C = \frac{(\psi'(\lambda))^2}{I_n(\lambda)} = \frac{e^{-2\lambda}}{\frac{n}{\lambda}} = e^{-2\lambda} \frac{\lambda}{n}$$

donde $I_n(\lambda)$ es la cantidad de información de Fisher respecto del parámetro λ . Por tanto:

$$e^{-2\lambda} \frac{\lambda}{n} < e^{-2\lambda} \left(\frac{\lambda}{n} + \sum_{k=2}^{\infty} \frac{1}{k!} \left(\frac{\lambda}{n}\right)^k \right) = e^{-2\lambda} \left(e^{\frac{\lambda}{n}} - 1\right)$$

no alcanzando W la cota de Cramer–Rao, a pesar de ser el estimador insesgado de varianza mínima. Ello no impide la existencia de estimadores, no insesgados, con error cuadrático medio menor.

C.M. Cuadras
J.M. Oller

PROBLEMA N° 17

Sean X_1, \dots, X_n las variables aleatorias muestrales correspondientes a una muestra aleatoria simple de tamaño n .

La media muestral es un estimador insesgado de la media poblacional $\mu = \frac{\alpha+\beta}{2}$ y su error cuadrático medio viene dado por:

$$\epsilon_1 = E((\bar{X}_n - \mu)^2) = \text{var}(\bar{X}_n) = \frac{1}{n} \text{var}(X_i) = \frac{(\beta - \alpha)^2}{12n}$$

Este estimador corresponde obviamente al método de los momentos.

El estimador que se obtiene a partir del método de máxima verosimilitud es:

$$W = (U + V)/2$$

donde

$$U = \max \{X_1, \dots, X_n\} \quad y \quad V = \min \{X_1, \dots, X_n\}$$

Es bien sabido que la densidad conjunta del par (U, V) , en el caso absolutamente continuo en general, viene dada por:

$$p(u, v) = n(n-1)f(u)f(v)\{F(u) - F(v)\}^{n-2} \quad v \leq u$$

donde f es la función de densidad y F la función de distribución de cada variable aleatoria muestral.

Por consiguiente, en nuestro caso, tendremos:

$$p(u, v) = n(n-1) \frac{1}{(\beta-\alpha)^2} \left\{ \frac{u-v}{\beta-\alpha} \right\}^{n-2} \quad \alpha < v \leq u \leq \beta$$

y $p(u, v) = 0$ en caso contrario.

El momento de primer orden de W será

$$E(W) = \int_{\alpha}^{\beta} \int_v^{\beta} \frac{u+v}{2} \cdot n(n-1) \frac{1}{(\beta-\alpha)^2} \left\{ \frac{u-v}{\beta-\alpha} \right\}^{n-2} du dv$$

que con el cambio $y = u+v$, $z = u-v$, resulta:

$$\begin{aligned}
E(W) &= \frac{n(n-1)}{2(\beta-\alpha)^n} \int_0^{\beta-\alpha} \int_{2\alpha+z}^{2\beta-z} yz^{n-2} \frac{1}{2} dy dz = \\
&= \frac{n(n-1)}{4(\beta-\alpha)^n} \int_0^{\beta-\alpha} \left\{ \frac{(2\beta-z)^2}{2} - \frac{(2\alpha+z)^2}{2} \right\} z^{n-2} dz \\
&= \frac{n(n-1)}{2(\beta-\alpha)^n} \int_0^{\beta-\alpha} \left\{ (\beta^2 - \alpha^2) z^{n-2} - (\beta - \alpha) z^{n-1} \right\} dz = \\
&= \frac{n(n-1)}{2(\beta-\alpha)^n} \left\{ (\beta^2 - \alpha^2) \frac{(\beta-\alpha)^{n-1}}{n-1} - (\beta + \alpha) \frac{(\beta-\alpha)^n}{n} \right\} = \frac{\beta + \alpha}{2} = \mu
\end{aligned}$$

Obteniendo que W es un estimador insesgado de μ .

El error cuadrático medio, ϵ_2 , coincidirá con la varianza, por tanto:

$$\begin{aligned}
\epsilon_2 &= \text{var}(W) = E((W - \mu)^2) = \\
&= \int_{\alpha}^{\beta} \int_v^{\beta} \left(\frac{u+v}{2} - \frac{\alpha+\beta}{2} \right)^2 n(n-1) \frac{1}{(\beta-\alpha)^2} \left\{ \frac{u-v}{\beta-\alpha} \right\}^{n-2} du dv
\end{aligned}$$

que con el cambio:

$$y = u + v - (\alpha + \beta) \quad z = u - v$$

resulta:

$$\begin{aligned}
\epsilon_2 &= \frac{n(n-1)}{4(\beta-\alpha)^2} \int_0^{\beta-\alpha} \int_{-\beta+\alpha+z}^{\beta-\alpha-z} y^2 z^{n-2} \frac{1}{2} dy dz = \\
&= \frac{n(n-1)}{8(\beta-\alpha)^n} \int_0^{\beta-\alpha} \left\{ \frac{(\beta-\alpha-z)^3}{3} - \frac{(-\beta+\alpha+z)^3}{3} \right\} z^{n-2} dz = \\
&= \frac{n(n-1)}{12(\beta-\alpha)^n} \int_0^{\beta-\alpha} \left\{ (\beta-\alpha)^3 z^{n-2} - 3(\beta-\alpha)^2 z^{n-1} + 3(\beta-\alpha) z^n - z^{n+1} \right\} dz = \\
&= \frac{n(n-1)}{12} (\beta-\alpha)^2 \left\{ \frac{1}{n-1} - \frac{3}{n} + \frac{3}{n+1} - \frac{1}{n+2} \right\} = \\
&= \frac{(\beta-\alpha)^2}{2(n+1)(n+2)}
\end{aligned}$$

Nótese que para $n = 2$ $W = \bar{X}_n$ y por tanto $\in_1 = \in_2$ pero para $n > 2$ $\in_2 < \in_1$ ya que

$$2(n+1)(n+2) > 12n \quad n > 2$$

Por tanto W es mas eficiente que \bar{X}_n para estimar la media muestral para muestras de tamaño $n > 2$.

Nota: El estadístico W aparece de forma natural al aplicar el teorema de Rao–Blackwell al estimador insesgado de μ , \bar{X}_n , respecto los estadísticos suficientes para α y β , U y V .

Jordi Ocaña