PROBLEMES PROPOSATS

PROBLEMA N° 62

Siguin X, Y dues variables aleatòries $N(0,1)$ independents. Considera el canvi de variables

$$X = R \cdot \cos\theta \qquad Y = R \cdot \sin\theta$$

1) Demostra que R i θ són estocàsticament independents.

2) Dóna una demostració, exclusivament geomètrica, de que la variable suma $X + Y$ és també normal.

C.M. Cuadras

Universitat de Barcelona

PROBLEMA N° 63

Considerem una mostra aleatòria simple $x_1, x_2, \ldots, x_n$ d'una variable aleatòria amb densitat $f(x, \theta)$. Suposem que volem contrastar

$$H_0: \quad \theta \in \omega \qquad H_1: \quad \theta \in \Omega - \omega$$

on $\omega \subset \Omega$ són regions paramètriques. Sigui

$$\lambda = \max_{\theta \in \omega} L(x_1, \ldots, x_n; \theta) / \max_{\theta \in \Omega} L(x_1, \ldots, x_n; \theta)$$

la raó de versemblança i suposem que es compleixen les condicions de regularitat de manera que la distribució asimptòtica de $V = -2\log\lambda$ és χ^2_{2k} sota H_0, on $2k > 0$.

1) Demostrar que, asimptòticament, la distribució de λ és la mateixa que la del producte

$$U = \prod_{i=1}^{k} U_i$$

on cada U_i és uniforme (0,1) i les U_i són independents.

2) Suposem $k = 1$. Especifica les regions d'acceptació de H_0 i de H_1, expressades com $\lambda \geqslant c$ i $\lambda < c$, respectivament.

C.M. Cuadras

Universitat de Barcelona

SOLUCIONS ALS PROBLEMES PROPOSATS AL VOLUM 20. N° 2

PROBLEMA N° 60

El enunciado del problema aparece en el libro de Barnett (1982, p. 139). Para resolverlo, reescribimos el estimador r_u del modo

$$r_u = \frac{1}{n} \sum_{i=1}^{N} \frac{x_i}{y_i} e_i,$$

donde e_i se distribuye binomial $\mathrm{B}(n, y_i/Y)$, y por tanto

$$\mathrm{E}(e_i) = ny_i/Y, \quad \text{y} \quad \mathrm{V}(e_i) = (ny_i/Y)(1 - y_i/Y).$$

a) $\mathrm{E}(r_u) = \dfrac{1}{n} \displaystyle\sum_{i=1}^{N} \frac{x_i}{y_i} \mathrm{E}(e_i) = \frac{X}{Y} = R,$

y por ello, r_u es insesgado para R.

b)

$$\begin{aligned}
\mathrm{V}(r_u) &= \frac{1}{n^2}\left[\sum_{i=1}^{N} \frac{x_i^2}{y_i^2} \mathrm{V}(e_i) + \sum_{i=1}^{N}\sum_{j\neq i}^{N} \frac{x_i x_j}{y_i y_j} \mathrm{Cov}(e_i, e_j)\right] = \\
&= \frac{1}{n^2}\left[\sum_{i=1}^{N} \frac{x_i^2}{y_i^2} \frac{ny_i}{Y}\left(1 - \frac{y_i}{Y}\right) + \sum_{i=1}^{N}\sum_{j\neq i}^{N} \frac{x_i x_j}{y_i y_j}\left(-n\frac{y_i}{Y}\frac{y_j}{Y}\right)\right] = \\
&= \frac{1}{n}\left[\sum_{i=1}^{N} \frac{x_i^2}{y_i}\frac{1}{Y} - \sum_{i=1}^{N} \frac{x_i^2}{Y^2} - \sum_{i=1}^{N}\sum_{j\neq i}^{N} \frac{x_i x_j}{Y^2}\right] = \\
&= \frac{1}{n}\left[\frac{1}{n}\sum_{i=1}^{N} \frac{x_i^2}{y_i^2}\frac{ny_i}{Y} - \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{x_i x_j}{Y^2}\right] = \\
&= \frac{1}{n}\left\{\left[\frac{1}{n}\sum_{i=1}^{N} r_i^2 \mathrm{E}(e_i)\right] - \left[\frac{1}{n}\sum_{i=1}^{N} r_i \mathrm{E}(e_i)\right]^2\right\} = \\
&= \frac{1}{n}\left[\frac{1}{n}\sum_{i=1}^{N} (r_i - R)^2 \mathrm{E}(e_i)\right] = \frac{1}{n}\sum_{i=1}^{N} (r_i - R)^2 \frac{y_i}{Y}
\end{aligned}$$

$$c)\ \mathrm{E}[s^2(r_u)] = \frac{1}{n(n-1)}\,\mathrm{E}\left[\sum_{i=1}^{n}(r_i - r_u)^2\right] =$$

$$= \frac{1}{n(n-1)}\,\mathrm{E}\left[\sum_{i=1}^{N}(r_i^2 - 2r_ir_u + r_u^2)\right] =$$

$$= \frac{1}{n(n-1)}\,\mathrm{E}\left[\sum_{i=1}^{N} r_i^2 e_i - 2\sum_{i=1}^{n} r_i\left(\frac{1}{n}\sum_{j=1}^{N} r_j e_j\right) + nr_u^2\right] =$$

$$= \frac{1}{n(n-1)}\left[\sum_{i=1}^{N} r_i^2 \frac{ny_i}{Y} - 2\mathrm{E}\left(\sum_{i=1}^{n} r_ir_u\right) + n\mathrm{E}(r_u^2)\right] = (*)$$

Pero,

$$\mathrm{E}(r_i\, r_u) = \mathrm{Cov}(r_i, r_u) + \mathrm{E}(r_i)\,\mathrm{E}(r_u) = \frac{1}{n}\,\mathrm{V}(r_i) + R^2 = \mathrm{V}(r_u) + R^2,$$

luego,

$$\mathrm{E}\left(\sum_{i=1}^{n} r_i\, r_u\right) = n\left[\mathrm{V}(r_u) + R^2\right],$$

de donde,

$$(*) = \frac{1}{n-1}\left[\sum_{i=1}^{N} r_i^2 \frac{y_i}{Y} - 2R^2 - 2\mathrm{V}(r_u) + \mathrm{V}(r_u) + R^2\right] =$$

$$= \frac{1}{n-1}\left[\frac{1}{n}\sum_{i=1}^{N} r_i^2\,\mathrm{E}(e_i) - R^2 - \mathrm{V}(r_u)\right] =$$

$$= \frac{1}{n-1}\left[\frac{1}{n}\sum_{i=1}^{N} (r_i - R)^2\,\mathrm{E}(e_i) - \mathrm{V}(r_u)\right] =$$

$$= \frac{1}{n-1}\left[n\mathrm{V}(r_u) - \mathrm{V}(r_u)\right] = \frac{n-1}{n-1}\,\mathrm{V}(r_u) = \mathrm{V}(r_u)$$

Referencia

Barnett, V. (1982). *Elements of Sampling Theory*. Hodder and Stoughton. Londres.

M. Ruiz Espejo

U.N.E.D.

PROBLEMA N° 61

a) Llamamos $p_{i|j}$ a la probabilidad de seleccionar en la segunda etapa a la unidad i, condicionado a que en la primera etapa se seleccionó la unidad j. Por ser esquema IPPS ($n = 2$),

$$\pi_i = \sum_{j\neq i}^{N} \left(p_i\, p_{j|i} + p_j\, p_{i|j}\right) + p_i\, p_{i|i} =$$

$$(1) \qquad = p_i \sum_{j=1}^{N} p_{j|i} + \sum_{j\neq i}^{N} p_j\, p_{i|j} = p_i + \sum_{j\neq i}^{N} p_j\, p_{i|j}$$

Imponiendo que el esquema es IPPS ($n = 2$), tenemos también que

$$\pi_i = 2\,p_i$$

De (1) y (2) deducimos que

$$(2) \qquad \sum_{j\neq i}^{N} p_j\, p_{i|j} = p_i \qquad (i = 1, 2, \ldots, N)$$

por lo que sumando en i, tenemos

$$(3) \qquad \sum_{i=1}^{N}\sum_{j\neq i}^{N} p_j\, p_{i|j} = \sum_{i=1}^{N} p_i = 1,$$

por ser los p_i normalizados, lo que implica que

$$\sum_{i=1}^{N} p_i\, p_{i|i} = 0,$$

y al ser $p_i > 0$ para todo $i = 1, 2, \ldots, N$; deducimos que

$$p_{i|i} = 0 \qquad \text{para todo} \quad i = 1, 2, \ldots, N.$$

Es decir, el esquema IPPS ($n = 2$) es sin reemplazamiento. Además, se dará

$$(4) \qquad \sum_{i=1}^{N} p_{i|j} = 1 \qquad (j = 1, 2, \ldots, N).$$

b) Como consecuencia, nos queda ahora el sistema simplificado, de (3) y (4),

$$(5) \qquad \sum_{j \neq i}^{N} p_j \, p_{i|j} = p_i \qquad (i = 1, 2, \ldots, N)$$

y

$$(6) \qquad \sum_{i \neq j}^{N} p_{i|j} = 1 \qquad (j = 1, 2, \ldots, N)$$

Si los p_i están fijados, el sistema (5) y (6) tiene $2N$ ecuaciones con $N(N-1)$ incógnitas. Como $N(N-1) \geqslant 2N$ (si $N > 2$), con igualdad si y sólo si $N = 3$, deducimos que sólo en el caso $N = 3$ habrá solución única para el esquema muestral, aunque para $N \geqslant 4$ puede haber más de una solución, como lo prueban los esquemas de Rao (1965) y Durbin (1967).

El esquema único IPPS ($n = 2$) para $N = 3$, se obtiene reescribiendo las ecuaciones (5) y (6). El sistema resultante es

$$(7) \qquad \left.\begin{aligned} p_2 \, p_{1|2} + p_3 \, p_{1|3} &= p_1 \\ p_1 \, p_{2|1} + p_3 \, p_{2|3} &= p_2 \\ p_1 \, p_{3|1} + p_2 \, p_{3|2} &= p_3 \\ p_{2|1} + p_{3|1} &= 1 \\ p_{1|2} + p_{3|2} &= 1 \\ p_{1|3} + p_{2|3} &= 1 \end{aligned}\right\}$$

que resolviéndolo tenemos por solución única,

$$(8) \qquad p_{2|1} = (p_2 - p_3)/p_1 \quad (\geqslant 0)$$

$$(9) \qquad p_{1|2} = p_1/p_2 \quad (\leqslant 1)$$

$$(10) \qquad p_{3|1} = (p_1 - p_2 + p_3)/p_1 \quad (\geqslant 0)$$

$$\begin{aligned} p_{1|3} &= 0 \\ p_{2|3} &= 1 \\ p_{3|2} &= 1 - p_1/p_2. \end{aligned}$$

De (8), (9) y (10) podemos asegurar que las condiciones necesarias y suficientes para que exista solución única posible del sistema (7) con $N = 3$, y tengan sentido probabilístico son:

$$(1.a) \qquad p_1 \leqslant p_2.$$

$$(1.b) \qquad p_3 \leqslant p_2 \qquad \text{y}$$

$$(1.c) \qquad p_2 \leqslant p_1 + p_3.$$

La demostración es inmediata por lo expuesto previamente. Ahora queda comprobar que $\pi_i = 2p_i \quad (i = 1, 2, 3)$. En efecto,

$$\begin{aligned} \pi_1 &= p_1 + p_2 p_{1|2} + p_3 p_{1|3} = p_1 + p_2\frac{p_1}{p_2} + p_3 \cdot 0 = 2p_1, \\ \pi_2 &= p_2 + p_1 p_{2|1} + p_3 p_{2|3} = p_2 + p_1\frac{p_2 - p_3}{p_1} + p_3 = 2p_2, \qquad \text{y} \\ \pi_3 &= p_3 + p_1 p_{3|1} + p_2 p_{3|2} = p_3 + p_1\frac{p_1 - p_2 + p_3}{p_1} + p_2\frac{p_2 - p_1}{p_2} = 2p_3, \end{aligned}$$

luego efectivamente, estamos ante un esquema IPPS ($n = 2$). Además es fácil comprobar que para este esquema,

$$\pi_{12} = p_1 + p_2 - p_3$$
$$\pi_{13} = p_1 - p_2 + p_3 \qquad \text{y}$$
$$\pi_{23} = -p_1 + p_2 + p_3.$$

Si $N \geqslant 4$, las soluciones $\{p_{i|j}\}_{1 \leqslant i \neq j \leqslant N}$ pueden no ser únicas y admitir, como ya hemos indicado, otros esquemas muestrales IPPS ($n = 2$). La mera comprobación de que el esquema muestral más simple estudiado aquí no está contenido en los esquemas de Rao (1965) o Durbin (1967), se dejan como ejercicio para el lector.

Referencia

Durbin, J. (1967). «Design of multistage surveys for the estimating of sampling errors». *Appl. Statist.*, **16**, 152–164.

Rao, J.N.K. (1965). «On two simple schemes of unequal probability sampling without replacement». *J. Indian Statist. Assoc.*, **3**, 173–180.

M. Ruiz Espejo

U.N.E.D.