SOLUCIONS ALS PROBLEMES PROPOSATS AL VOLUM 23 N. 1

PROBLEMA N. 74

Llamando

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2, \tag{1}$$

De Vries (1986, p. 8) ha demostrado que

$$S^2 = \frac{n-1}{n^2}\sum_{i=1}^{n} x_i^2 - \frac{2}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} x_i x_j. \tag{2}$$

Veamos ahora que

$$S^2 = \frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(x_i - x_j)^2 \tag{3}$$

por inducción en n.

Para $n = 1$, S^2 no está definida o vale cero, lo que supone un caso obviamente sin interés práctico especial.

Partimos del caso $n = 2$. Según la fórmula (1),

$$S^2 = \frac{1}{2}\sum_{i=1}^{2}\left[x_i - \frac{1}{2}(x_1 + x_2)\right]^2 = \frac{1}{2}\left\{\left[\frac{1}{2}(x_1 - x_2)\right]^2 + \left[\frac{1}{2}(-x_1 + x_2)\right]^2\right\} =$$

$$= \frac{1}{8}\left(2x_1^2 + 2x_2^2 - 4x_1x_2\right) = \frac{1}{2^2}(x_1 - x_2)^2 = \frac{1}{2^2}\sum_{i=1}^{1}\sum_{j=i+1}^{2}(x_i - x_j)^2,$$

que es también la fórmula (3) para $n = 2$. Con esto queda demostrada la fórmula (3) en el caso $n = 2$.

Ahora suponemos que la fórmula (3) es cierta para n ($n \geq 2$), y vamos a demostrar o justificar que la fórmula (3) sigue siendo cierta para $n + 1$. La hipótesis de inducción es que (3) es válida para n, y en lo que resta demostraremos que (3) será válida cuando tenemos un dato más, x_{n+1}, basándonos en la hipótesis de inducción (que consiste en que las relaciones (1) y (3) son iguales para n) y en que la fórmula (2) es igual a la (1), ya demostrado para cualquier valor entero de n $(n \geq 1)$ por De Vries (1986). Usaremos como subíndice en la notación de S^2 ó $\bar{x}$ al número de datos de la variable, por lo que

$$S_{n+1}^2 = \frac{1}{n+1} \sum_{i=1}^{n+1} (x_i - \bar{x}_{n+1})^2 =$$

$$= \frac{1}{n+1} \left[\sum_{i=1}^{n} (x_i - \bar{x}_n + \bar{x}_n - \bar{x}_{n+1})^2 + (x_{n+1} - \bar{x}_{n+1})^2 \right] =$$

$$= \frac{1}{n+1} \left[\sum_{i=1}^{n} (x_i - \bar{x}_n)^2 + n (\bar{x}_n - \bar{x}_{n+1})^2 + (x_{n+1} - \bar{x}_{n+1})^2 \right] =$$

por la fórmula (1),

$$= \frac{1}{n+1} \left[nS_n^2 + \left(n\bar{x}_{n+1}^2 + \bar{x}_{n+1}^2 \right) - \right.$$

$$\left. - \left(2n\bar{x}_n\bar{x}_{n+1} + 2x_{n+1}\bar{x}_{n+1} \right) + \left(n\bar{x}_n^2 + x_{n+1}^2 \right) \right] =$$

y haciendo uso de la hipótesis de inducción para n

$$= \frac{1}{n+1} \left\{ \frac{1}{n} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (x_i - x_j)^2 + \left[(n+1)\, \bar{x}_{n+1}^2 - 2\bar{x}_{n+1} \left(x_{n+1} + n\bar{x}_n \right) \right] + \right.$$

$$\left. + \left[x_{n+1}^2 + \frac{1}{n} \left(\sum_{i=1}^{n} x_i \right)^2 \right] \right\} =$$

$$= \frac{1}{n\,(n+1)} \left[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (x_i - x_j)^2 - n\,(n+1)\, \bar{x}_{n+1}^2 + \right.$$

$$\left. + \left(n x_{n+1}^2 + \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} \sum_{j \neq i}^{n} x_i x_j \right) \right]. \tag{4}$$

Por otro lado, hay que demostrar que (4) coincide con

$$S_{n+1}^2 = \frac{1}{(n+1)^2} \sum_{i=1}^{n} \sum_{j=i+1}^{n+1} (x_i - x_j)^2 =$$

$$= \frac{1}{(n+1)^2} \left[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (x_i - x_j)^2 + \sum_{i=1}^{n} (x_i - x_{n+1})^2 \right]. \tag{5}$$

El resultado buscado será cierto si, y sólo si, los dos últimos términos de (4) y (5) coinciden, y esto se cumple si, y sólo si, se da la siguiente igualdad:

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (x_i - x_j)^2 - n(n+1)^2 \bar{x}_{n+1}^2 + n(n+1) x_{n+1}^2 +$$

$$+ (n+1) \sum_{i=1}^{n} x_i^2 + (n+1) \sum_{i=1}^{n} \sum_{j \neq i}^{n} x_i x_j -$$

$$- \left(n \sum_{i=1}^{n} x_i^2 + n^2 x_{n+1}^2 - 2n^2 x_{n+1} \bar{x}_n \right) = 0,$$

o bien, haciendo uso de la hipótesis de inducción o de la fórmula (3) para n, la condición necesaria y suficiente es que se dé la igualdad:

$$n^2 S_n^2 - n \left(\sum_{i=1}^{n+1} x_i \right)^2 + n x_{n+1}^2 + \sum_{i=1}^{n} x_i^2 +$$

$$+ (n+1) \sum_{i=1}^{n} \sum_{j \neq i}^{n} x_i x_j + \left(2n x_{n+1} \sum_{i=1}^{n} x_i \right) = 0,$$

y simplificando quedará:

$$n^2 S_n^2 - n \left[\left(\sum_{i=1}^{n} x_i \right)^2 + x_{n+1}^2 + 2x_{n+1} \sum_{i=1}^{n} x_i \right] + n x_{n+1}^2 +$$

$$+ \sum_{i=1}^{n} x_i^2 + (n+1) \sum_{i=1}^{n} \sum_{j \neq i}^{n} x_i x_j + 2n x_{n+1} \sum_{i=1}^{n} x_i = 0,$$

o bien:

$$n^2 S_n^2 + \sum_{i=1}^{n} \sum_{j \neq i}^{n} x_i x_j - (n-1) \sum_{i=1}^{n} x_i^2 = 0,$$

es decir si, y sólo si:

$$S_n^2 = \frac{n-1}{n^2} \sum_{i=1}^{n} x_i^2 - \frac{2}{n^2} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} x_i x_j,$$

y ésta es cierta ya que es la fórmula (2) debida a De Vries (1986). De aquí concluimos que para $n \geq 2$ y n número natural, la fórmula (3) es otra expresión alternativa de la varianza $S^2 = S_n^2$ que no se refiere explícitamente a la media $\bar{x} = \bar{x}_n$. La fórmula (1) refiere la varianza a la media, pero las fórmulas (2) y (3) de la varianza no se refieren a la media explícitamente.

REFERENCIA

De Vries, P.J. (1986). *Sampling Theory for Forest Inventory.* Springer-Verlag. Berlín.

M. Ruíz Espejo
UNED

PROBLEMA N. 75

Para el estadístico media muestral $\bar{x}_n$, basado en la muestra aleatoria simple con reemplazamiento $(x_1, x_2, ..., x_n)$,

$$\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i = t_n \text{ (digamos)},$$

de donde

$$\bar{x}_{n-1,i} = \frac{1}{n-1}\sum_{j\neq i}^{n} x_j = t_{n-1,i} \text{ (digamos)}$$

es la media muestral de la muestra aleatoria simple excluyendo la componente i-ésima x_i. También,

$$\begin{aligned}\bar{t}_{n-1} &= \frac{1}{n}\textstyle\sum_{i=1}^{n} t_{n-1,i} = \frac{1}{n}\sum_{i=1}^{n} \bar{x}_{n-1,i} = \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}^{n} x_j = \\ &= \frac{1}{n(n-1)}\textstyle\sum_{i=1}^{n}(n\bar{x}_n - x_i) = \\ &= \frac{1}{n(n-1)}\left(n^2\bar{x}_n - n\bar{x}_n\right) = \bar{x}_n,\end{aligned}$$

por lo que

$$t_n' = nt_n - (n-1)\,\bar{t}_{n-1} = n\bar{x}_n - (n-1)\,\bar{x}_n = \bar{x}_n,$$

y de aquí, el estimador jackknife de la varianza de la media muestral, es

$$\widehat{V}_J(\bar{x}_n) = \frac{n-1}{n}\sum_{i=1}^{n}\left(t_{n-1,i}^2 + \bar{t}_{n-1}^2 - 2t_{n-1,i}\bar{t}_{n-1}\right) =$$

$$= \frac{n-1}{n}\left(\sum_{i=1}^{n}\bar{x}_{n-1,i}^2 + n\bar{x}_n^2 - 2\bar{x}_n\sum_{i=1}^{n}\bar{x}_{n-1,i}\right) = \frac{n-1}{n}\left(\sum_{i=1}^{n}\bar{x}_{n-1,i}^2 - n\bar{x}_n^2\right) =$$

$$= (n-1)\left(\frac{1}{n}\sum_{i=1}^{n}\bar{x}_{n-1,i}^2 - \bar{x}_n^2\right). \tag{1}$$

Y ya que si denotamos por

$$a_k = \frac{1}{n} \sum_{i=1}^{n} x_i^k$$

al momento muestral respecto al origen de orden k, tenemos que

$$\begin{aligned}
\frac{1}{n} \textstyle\sum_{i=1}^{n} \bar{x}_{n-1,i}^2 &= \frac{1}{n} \textstyle\sum_{i=1}^{n} \left(\frac{1}{n-1} \sum_{j \neq i}^{n} x_j \right)^2 = \\
&= \frac{1}{n(n-1)^2} \textstyle\sum_{i=1}^{n} \left(\sum_{j \neq i}^{n} x_j^2 + \sum_{j \neq i}^{n} \sum_{k \neq i,j}^{n} x_j x_k \right) = \\
&= \frac{1}{n(n-1)^2} \left[\textstyle\sum_{i=1}^{n} \left(n a_2 - x_i^2 \right) + \sum_{i=1}^{n} \sum_{j \neq i}^{n} x_j \sum_{k \neq i,j}^{n} x_k \right] = \\
&= \frac{1}{n(n-1)^2} \left[\left(n^2 a_2 - n a_2 \right) + \textstyle\sum_{i=1}^{n} \sum_{j \neq i}^{n} x_j \left(n a_1 - x_i - x_j \right) \right] = \\
&= \frac{1}{n(n-1)^2} \left[n(n-1) a_2 + \textstyle\sum_{i=1}^{n} n a_1 \left(n a_1 - x_i \right) - \right. \\
&- \left. \textstyle\sum_{i=1}^{n} x_i \left(n a_1 - x_i \right) - \sum_{i=1}^{n} \sum_{j \neq i}^{n} x_j^2 \right] =
\end{aligned}$$

$$= \frac{1}{(n-1)^2} \left[n(n-2) a_1^2 + a_2 \right]. \tag{2}$$

Luego, de (1) y (2) en las dos primeras igualdades respectivamente:

$$\begin{aligned}
\widehat{V}_J(\bar{x}_n) &= (n-1) \left(\frac{1}{n} \textstyle\sum_{i=1}^{n} \bar{x}_{n-1,i}^2 - \bar{x}_n^2 \right) = \\
&= (n-1) \left\{ \frac{1}{(n-1)^2} \left[n(n-2) a_1^2 - a_2 \right] - a_1^2 \right\} = \\
&= \frac{1}{n-1} \left[n(n-2) a_1^2 + a_2 - (n-1)^2 a_1^2 \right] = \\
&= \frac{1}{n-1} \left(a_2 - a_1^2 \right) = \frac{1}{n(n-1)} \textstyle\sum_{i=1}^{n} \left(x_i - a_1 \right)^2 .
\end{aligned}$$

M. Ruíz Espejo
UNED

PROBLEMES PROPOSATS

PROBLEMA N. 76

Sea un sistema de espera abierto con *una* estación de servicio, en el que los clientes llegan según una ley de Poison con media λ y en el que la duración del servicio sigue una ley exponencial con media $1/\mu$. En el supuesto de existir régimen permanente, obtener:

1. La distribución de probabilidad del número de clientes: $\Pi_n \equiv p(X = n)$.
2. En número medio de clientes en el sistema, de clientes esperando *y* de *unidades desocupadas*.
3. Determinar $P(X \leq n)$, la probabilidad de espera y el tiempo medio de espera.
4. ¿Qué número de usuarios se encuentra en el sistema con mayor probabilidad ?
5. ¿Para qué valor de la intensidad de tráfico se encuentra la máxima probabilidad de un número de unidades en el sistema dado? Analizar los casos $n = 1$ y $n = 2$.

R. Alonso

PROBLEMA N. 77

En los supuestos del Problema 1, se considera adicionalmente que un cliente que llega al sistema y no encuentra la unidad de servicio vacía rechaza entrar con probabilidad α.

a) Resolver 1), 2) y 3) del Problema 1.

b) Estudiar el supuesto $\alpha = 1$.

R. Alonso

PROBLEMA N. 78

En los supuestos del Problema 1, se considera adicionalmente que se produce abandono del sistema, antes de ser atendido, según una tasa β constante, independiente del número de clientes. Determinar:

1) La distribución de probabilidad del número de clientes.

2) El número medio de clientes en el sistema

R. Alonso

PROBLEMA N. 79

En los supuestos del Problema 1, se considera adicionalmente que se produce abandono del sistema, antes de ser atendido, según una tasa β y que un cliente que llega al sistema y no encuentra la unidad de servicio libre rechaza entrar con probabilidad α (supuestos adicionales de los problemas 2 y 3). Obtener:

a) La distribución de probabilidad del número de clientes

b) El número medio de clientes en el sistema.

R. Alonso

PROBLEMA N. 80

En los supuestos del Problema 1, se considera adicionalmente que un cliente que llega al sistema y no encuentra la unidad de servicio vacía rechaza entrar con probabilidad $b(n) = \frac{n}{n+1}$. Obtener:

a) La distribución de probabilidad del número de clientes (X).

b) El número medio de clientes en el sistema.

R. Alonso

PROBLEMA N. 81

En los supuestos del Problema 1, se considera adicionalmente que el sistema es de capacidad limitada (K) y la tasa de circulación (ψ) es igual a la unidad. Determinar:

a) La distribución de probabilidad del número de clientes.

b) El número medio de clientes en el sistema.

c) El número medio de unidades desocupadas.

d) El número medio de clientes en espera.

R. Alonso