Estadística y análisis de datos

Segundo curso

Profesores: Imanol Mozo, Irati Susperregi

Examen resuelto (20 de mayo de 2019)

Donostiako Ekonomia eta Enpresa Fakultatea EHU

Autor: Beñat Zunzunegi

Resolución extraoficial:
el autor no asume
ninguna responsabilidad
sobre la corrección oficial de los ejercicios.

Gizapedia

gizapedia.hirusta.io

Problema 1 (1 punto)

Un comercial consigue cerrar una venta el 35 % de las veces que establece una relación nueva. Si en una semana consigue establecer 80 nuevas relaciones:

(a) ¿Cuántas ventas cerraría como mínimo el 90 % de las semanas? (0.5 puntos)

(b) ¿Cuál es la probabilidad de que cierre entre 40 y 50 ventas, ambas inclusive? (0.5 puntos)

(a)

El número de ventas se distrbiuye de acuerdo a la distribución binomial. Como n es grande (superior a 30) puede utilizarse la aproximación normal:

$$X : \text{número de ventas} \sim B(n = 80, p = 0{,}35) \to N(\mu = np = 80 \times 0{,}35 = 28, \sigma = \sqrt{npq} = \sqrt{80 \times 0{,}35 \times 0{,}65} = 4{,}26)$$

Formalizamos la pregunta del siguiente modo:

$$P[X \geq x] = P\left[Z > \frac{x-28}{4{,}26}\right] = 0{,}90$$

$$\frac{x-28}{4{,}26} = -1{,}28 \to x = 22{,}54$$

Se cierran pues 23 ventas como mínimo al menos el 90 % de los dias.

(b)

Realizamos la aproximación continua y tipificamos:

$$P[40 \leq X \leq 50] = P[39{,}5 < X < 50{,}5] = P\left[\frac{39{,}5-28}{4{,}26} < Z < \frac{50{,}5-28}{4{,}26}\right] = P[2{,}69 < Z < 5{,}28]$$

Calculamos en tablas:

$$P[2{,}69 < Z < 5{,}28] = P[Z < 5{,}28] - P[Z < 2{,}69] = 1 - 0{,}9964 = 0{,}0036$$

Problema 2 (1 punto)

El peso medio de los pasajeros de un avión es de 75kg y su desviación t'pica 7.5kg. ¿Cuántos pasajeros puede admitir el avión para que la probabilidad de que el peso total de los pasajeros no supere 3000kg sea como mínimo 0.95?

Si denominamos n al número de pasajeros, el peso total (X) se distribuye de esta forma segun el teorema central del límite:

$$X \sim N(\mu = 75n, \sigma = \sqrt{7{,}5^2 n} = 7{,}5\sqrt{n})$$

Planteamos el problema:

$$P[X < 3000] = 0{,}95 \rightarrow P\left[Z < \frac{3000 - 75n}{7{,}5\sqrt{n}}\right] = 0{,}95$$

Buscamos el percentil correspondiente a la probabilidad en tablas:

$$\frac{3000 - 75n}{7{,}5\sqrt{n}} = 1{,}64$$

Reordenamos, operamos y hacemos cambio de variable:

$$3000 - 75n = 7{,}5 \times 1{,}64\sqrt{n} \rightarrow 3000 - 75n = 12{,}3\sqrt{n} \rightarrow$$

$$75n - 12{,}3\sqrt{n} - 3000 = 0 \rightarrow 75x^2 - 12{,}3x - 3000 = 0$$

Resolvemos la ecuación de segundo grado (solución positiva) y deshacemos el cambio de variable: $x = 6{,}40 \rightarrow n = 40{,}96 \rightarrow$ 40 pasajeros (por prudencia redondeamos hacia abajo).

Problema 3 (2 puntos)

Se sabe que la distribución de probabilidad de los beneficios en miles de euros de una inversión determinada, es la siguiente:

- **la probabilidad de obtener beneficios es $P(1) = \frac{1}{k}$.**
- **la probabilidad de no obtener ni beneficios ni pérdidas es $P(0) = \frac{2}{k}$.**
- **la probabilidad de obtener pérdidas es $P(-1) = \frac{k-3}{k}$.**

Para estimar el parámetro k se toma una muestra de 7 inversiones idénticas independientes. En dos de ellas se obtienen beneficios, en una resultan pérdidas y en los otros cuatro casos no hay ni beneficios ni pérdidas.Estima la probabilidad de obtener pérdidas por:

(a) el método de máxima verosimlitud (1 punto).

(b) el método de los momentos (1 punto).

(a)

Primero, construimos la función de verosimilitud:

$$L(\mathbf{X}/k) = \left(\frac{1}{k}\right)^2 \cdot \left(\frac{2}{k}\right)^4 \cdot \left(\frac{k-3}{k}\right) \cdot \frac{7!}{2!1!4!}$$

Se trata de tomar como estimación el valor de k que maximiza la probabilidad, dada por la función de verosimilitud, de obtener la muestra. De ahí estimaremos luego la probabilidad de pérdida.

Tomamos el logaritmo de la función de verosimilitud, que tiene el mismo máximo, para determinar dicha estimación:

$$\ln L(\mathbf{X}/k) = -2\ln k + 4(\ln 2 - \ln k) + [\ln(k-3) - \ln k] + [\ln 7! - \ln(2!1!4!)]$$

Derivamos respecto de k e igualamos a 0 como condición del máximo:

$$\frac{d\ln L(\mathbf{X}/k)}{dk} = -2\frac{1}{k} - 4\frac{1}{k} + \frac{1}{k-3} - \frac{1}{k} = 0 \rightarrow \frac{-7}{k} + \frac{1}{k-3} = 0$$

Simplificamos y despejamos k:

$$\frac{d\ln L(\mathbf{X}/k)}{dk} = \frac{-6k+21}{k(k-3)} = 0 \rightarrow -6k + 21 = 0 \rightarrow \hat{k}_{MV} = \frac{21}{6} = 3{,}5$$

De este resultado, se estima la probabilidad de pérdida:

$$\hat{P}(-1)_{MV} = \frac{3{,}5-3}{3{,}5} = \frac{1}{7}$$

Resultado bastante lógico, ya que si de 7 inversiones una tiene pérdidas, la probabilidad de pérdidas se estima intuitivamente en 1/7.

$\boxed{\text{(b)}}$

Para estimar por momentos, debo calcular el primer momento de la distrbución (su esperanza matemática):

x	$p(x)$	$xp(x)$
1	$\frac{1}{k}$	$\frac{1}{k}$
0	$\frac{2}{k}$	0
-1	$\frac{k-3}{k}$	$\frac{-(k-3)}{k}$
	1	$\alpha_1 = \frac{4-k}{k}$

Ahora calculamos el primer momento muestral, es decir, la media aritmética:

$$\overline{x} = \frac{1 \times (-1) + 4 \times 0 + 2 \times 1}{7} = \frac{1}{7}$$

Igualamos el momento poblacional con el muestral:

$$\frac{4-k}{k} = \frac{1}{7}$$

Y logramos la estimación por momentos, despejando k:

$$28 - 7k = k \rightarrow \hat{k}_{MOM} = \frac{28}{8} = 3{,}5$$

Y entonces la estimación de la probabilidad:

$$\hat{P}(-1)_{MOM} = \frac{3{,}5 - 3}{3{,}5} = \frac{1}{7}$$

Resulta la misma estimación que por el método de máxima verosimolitud, aunque no siempre es así.