Bioestadística

Universidad de la República
Facultad de Ciencias

Nombre:	CI:

Examen 7 de Diciembre de 2022

Ejercicio 1 En dos grandes poblaciones A y B la proporción de individuos infectados por un determinado virus es del 30 y 60 % respectivamente. Se toma al azar una de las 2 poblaciones (las dos son igual de probables), y de ella elegimos a 5 individuos al azar.

(a) Calcular la probabilidad de que hayan dos infectados si la muestra se toma de la población A. Idem para la población B.

Sabiendo que en la muestra hay **dos** individuos infectados,

(b) Calcular la probabilidad de que la muestra corresponda a la población A.

(c) Calcular la probabilidad de que la muestra corresponda a la población B.

(d) ¿A qué población es más probable que pertenezca la muestra? Justifique.

Ejercicio 2 La longitud en milímetros X que se puede estirar un filamento de nylon sin ruptura sigue una distribución exponencial de parámetro λ desconocido. Dicho filamento es sometido a un control de calidad y pasa dicho control si puede estirarse más de 0,18 milímetros.

(a) Calcular en función de λ la probabilidad de que un filamento pase dicho control. A esa probabilidad le llamaremos p.

Se tomó una muestra de 100 filamentos y se obtuvo que solo 72 de ellos pasaron el control de calidad.

(b) A partir del dato anterior, dar una estimación $\hat{p}$ de p.

(c) A partir $\hat{p}$ dar una estimación de λ.

(d) Construir un intervalo de confianza para p con una confianza del 95 %.

(e) El fabricante de los filamentos dice que al menos el 90 % de sus filamentos deberían de pasar el control de calidad. Realizar el siguiente test de hipótesis:

$$\begin{cases} H_0 : p \geq 0{,}90 \\ H_A : p < 0{,}90 \end{cases}$$

y sacar una conclusión acerca de la veracidad de la afirmación del fabricante para $\alpha = 0{,}05$.

Datos que pueden ser de utilidad: `qnorm(0.975)= 1.96` `qnorm(0.95)=1.65` `qnorm(0.9)=1.28`

SOLUCIÓN

Ejercicio 1 Sea X la cantidad de individuos infectados en la muestra, entonces $X \sim \text{Bin}(5, p)$ con

$$p = \begin{cases} 0{,}3, & \text{si la muestra proviene de la población A,} \\ 0{,}6, & \text{si la muestra proviene de la población B} \end{cases}$$

(a) $\mathbb{P}(X = 2|A) = \mathbb{P}(\text{Bin}(5; 0{,}3) = 2) = C^5_2 0{,}3^2(1 - 0{,}3)^3 = 0{,}309$
$\mathbb{P}(X = 2|B) = \mathbb{P}(\text{Bin}(5; 0{,}6) = 2) = C^5_2 0{,}6^2(1 - 0{,}6)^3 = 0{,}230.$

(b) $\mathbb{P}(A|X = 2) = \frac{\mathbb{P}(X=2|A)\mathbb{P}(A)}{\mathbb{P}(X=2)} = \frac{0{,}309 \times 0{,}5}{\mathbb{P}(X=2)}$. Como

$$\mathbb{P}(X = 2) = \mathbb{P}(X = 2|A)\mathbb{P}(A) + \mathbb{P}(X = 2|B)\mathbb{P}(B) = 0{,}309 \times 0{,}5 + 0{,}230 \times 0{,}5 = 0{,}270,$$

entonces $\mathbb{P}(A|X = 2) = \frac{0{,}309 \times 0{,}5}{0{,}270} = 0{,}572$.

(c) $\mathbb{P}(B|X = 2) = \frac{\mathbb{P}(X=2|B)\mathbb{P}(B)}{\mathbb{P}(X=2)} = \frac{0{,}230 \times 0{,}5}{0{,}270} = 0{,}426$

(d) Es más probable que pertenezcan a la muestra A ya que $\mathbb{P}(X = 2|A) > \mathbb{P}(X = 2|B)$.

Ejercicio 2 $X \sim \text{Exp}(\lambda)$ y pasa el control de calidad si $X \geq 0{,}18$.

(a) $p = \mathbb{P}(X \geq 0{,}18) = e^{-0{,}18\lambda}$.

(b) Se tomó una muestra de 100 filamentos y se obtuvo que solo 72 de ellos pasaron el control de calidad. A partir del dato anterior, podemos estimar p por la frecuencia relativa: $\hat{p} = \frac{72}{100} = 0{,}72$.

(c) Planteamos $0{,}72 = e^{-0{,}18\lambda}$, despejamos λ y obtenemos $\hat{\lambda} = -\frac{\log(0{,}72)}{0{,}18} = 1{,}83$.

(d) Queremos un intervalo de confianza para una proporción: como $p = \mathbb{E}(Y)$, siendo Y la variable aleatoria tal que $Y = \begin{cases} 1, & \text{si } X \geq 0{,}18, \\ 0, & \text{si } X < 0{,}18 \end{cases}$ $(Y \sim \text{Ber}(p))$, entonces queremos un intervalo de confianza para la esperanza de Y. Si la llamamos $y_1, y_2, \ldots, y_{100}$ a los resultados de la variable Y para la muestra, entonces $\hat{p} = \bar{y}_{100}$ y

$$I_{1-\alpha}(p) = \left[\hat{p} \pm \frac{\sigma z_{1-\frac{\alpha}{2}}}{\sqrt{100}}\right]$$

y podemos estimar $\sigma = \sqrt{p(1-p)}$ por $\hat{\sigma} = \sqrt{\hat{p}(1-\hat{p})}$. Luego,

$$I_{0{,}95}(p) = \left[\hat{p} \pm \frac{\sqrt{\hat{p}(1-\hat{p})} \times 1{,}96}{10}\right] = [0{,}72 \pm 0{,}09] = [0{,}63,\ 0{,}81].$$

.

(e) Queremos testear:

$$\begin{cases} H_0 : p \geq 0{,}90 \\ H_A : p < 0{,}90 \end{cases}$$

La región crítica para este test es

$$RC = \left\{\bar{y}_{100} < 0{,}90 - \frac{\sigma z_{1-\alpha}}{\sqrt{100}}\right\} \approx \left\{\bar{y}_{100} < 0{,}90 - \frac{\sqrt{\bar{y}_{100}(1-\bar{y}_{100})} \times 1{,}65}{10}\right\}$$

Como $\hat{p} = 0{,}72 < 0{,}90 - \frac{\sqrt{\hat{p}(1-\hat{p})} \times 1{,}65}{10} = 0{,}83$ y $\bar{y}_{100} = \hat{p}$, con una confianza del 95 % rechazamos la hipótesis del fabricante.