corregido

21-2-18

Universidad Tecnológica Nacional Facultad Regional Buenos Aires Examen Final de Probabilidad y Estadística -21 de febrero de 2018

Apellido y Nombre: ....

Legajo: ....

Ej. 1	Ej. 2	Ej. 3	Ej. 4	Teórico 1	Teórico 2	Nota

La condición mínima de aprobación es de tres puntos correctos. Todas las respuestas deberán estar debidamente justificadas.

Ejercicio 1 *El tiempo en horas hasta la falla de cierto tipo de generador eléctrico tiene distribución exponencial. Mediante ensayos previos se determinó que bajo condiciones normales el 20% de los mismos falla antes de 50 horas de operación. Se arma un sistema con dos de estos generadores(cuya duración es independiente) conectados en paralelo.*

1. Hallar la probabilidad de que el sistema funcione por lo menos 150 horas.
2. ¿Cuál es la vida media de cada uno de estos generadores?.

Ejercicio 2 *De un cultivo de una variedad de guayule se seleccionaron 27 plantas al azar. De estas 15 resultaron atípicas(A) y 12 aberrantes(B). Los porcentajes de caucho en las plantas se suponen normales y con igual varianza para ambas variedades. Los resultados obtenidos son:*

A: 6.21 - 5.7 - 6.04 - 4.47 - 5.22 - 4.45 - 4.84 - 5.88 - 5.82 - 6.09 - 5.59 - 6.06 - 5.59 - 6.74 - 5.55

B: 4.28 - 7.71 - 6.48 - 7.71 - 7.37 - 7.2 - 7.06 - 6.4 - 8.93 - 5.91 - 5.51 - 6.36

a. *Decidir si hay diferencias estadísticamente significativas entre las medias de caucho de las dos poblaciones, utilizando un 5% de significación.*

b. *Indicar en el contexto de esta prueba que sería cometer error de tipo I y error de tipo II.*

Ejercicio 3 *Se realizó una encuesta sobre el perfil de la audiencia del sitio web de ESPN con el fin de estimar la proporción poblacional de mujeres entre los usuario. Entre los 400 encuestados, 104 resultaron mujeres. Asumiendo que el tamaño poblacional es muy grande:*

1. Proponga un estimador insesgado para la proporción poblacional de mujeres.
2. Estime mediante un intervalo del 95% de confianza esta proporción.
3. ¿Cuál sería el tamaño muestral necesario para que la longitud del intervalo resulte de a lo sumo 0.06?

Ejercicio 4 *En un negocio hay dos vendedoras Claudia y Patricia. De los clientes que atiende Claudia el 75% efectúa la compra; mientras que de los clientes que atiende Patricia sólo lo hacen el 65%. Los clientes eligen con igual probabilidad a ambas, pero Claudia falta el 12% de los días y Patricia el 10%. Cuando faltan ambas vendedoras, atiende el dueño que vende con seguridad.*

1. Qué porcentaje de los clientes que llegan al negocio compra?
2. Si compró un cliente, qué probabilidad hay de que lo haya atendido Claudia?

Teórico 1 *Se sabe que los estimadores $\hat{\theta}_1$ y $\hat{\theta}_2$ son insesgados para θ. Mostrar que $\hat{\theta}_3 = \alpha\hat{\theta}_1 + (1-\alpha)\hat{\theta}_2$, con $\alpha > 0$ también es un estimador insesgado de θ b) Si $V(\hat{\theta}_1) = a^2$ y $V(\hat{\theta}_2) = 3a^2$, para qué valor de α el estimador $\hat{\theta}_3$ alcanza la menor varianza.*

Teórico 2 *Probar que A y B son eventos independientes de un mismo espacio muestral si y sólo si se verifica que: $P(A \cup B) = P(A) + P(\bar{A})P(B)$*

$P(A\cup B) = P(A) + P(B) - P(A\cap B) = P(A) + 1 - P(\bar{B}) - P(A\cap B)$

$P(A) + P(\bar{A})P(B) = \cancel{P(A)} + (1-P(A))P(B) \overset{P(A)+}{=} P(B) - P(A)P(B)$

HOJA N° 2

FECHA 21-2-18

Final - PyE UTN

① El tiempo en horas hasta la falla de cierto tipo de generador eléctrico tiene distribución exponencial. Mediante ensayos previos se determinó que bajo condiciones normales el 20% de los mismos falla antes de 50 horas de operación.
Se arma un sistema con dos de estos generadores. (cuya duración es independiente) conectados en paralelo.

a) Hallar la probabilidad de que el sistema funcione por lo menos 150 horas

X: "tiempo, en horas, de funcionamiento de un generador"

$X \sim \mathcal{E}(\lambda)$

$$P(X < 50) = 0.20 \overset{\text{x continuidad}}{=} P(X \leq 50) = 1 - e^{-\lambda 50} \rightarrow e^{-\lambda 50} = 0.80$$

$$\rightarrow -\lambda 50 = \ln(0.80) \rightarrow \lambda = \frac{-\ln(0.8)}{50} = 0{,}0045 \rightsquigarrow X \sim \mathcal{E}(0{,}0045)$$

Sistema en paralelo: funciona mientras alguno funcione
no funciona cuando ninguno de los dos funcionan

$$P(X \geq 150) = 1 - P(X < 150) \overset{\text{continuidad}}{=} 1 - P(X \leq 150) = 1 - (1 - e^{-0{,}0045 \times 150}) = 0{,}509$$

Y: "cant. de generadores, de un sistema de 2, que funcionan que duran más de 150 horas"

$Y \sim Bi(2, 0.509)$

$$P(Y \geq 1) = 1 - P(Y < 1) \overset{\text{discreta}}{=} 1 - P(Y = 0) = 1 - \binom{2}{0} 0{,}509^0 (1 - 0.509)^2 = 0{,}759$$

$$\boxed{P\ Y \geq 1 = 0{,}759}$$

b) ¿cuál es la vida media de cada uno de estos generadores?

$$E(X) = \frac{1}{\lambda} = \frac{1}{0{,}0045} = 222$$

$$\boxed{E(X) = 222 \text{ horas}}$$

NOTA

② De un cultivo de una variedad de guayule se seleccionaron 27 plantas al azar. De estas 15 resultaron atípicos (A), y 12 aberrantes (B). Los porcentajes de caucho en las plantas se suponen normales y con igual varianza para ambas variedades.
Los resultados obtenidos son:

A: 6,21 5.7 6.04 4.47 5.22 4.45 4.84 5.88 5.82 6.09
5.59 6.06 5.59 6.74 5.55

B: 4.28 7.71 6.48 7.71 7.37 7.2 7.06 6.4 8.93 5.91 5.51 6.96

a) Decidir si hay diferencias estadísticamente significativas entre las medias de caucho de las dos poblaciones, utilizando un 5% de significación

$\bar{X}_A = 5,617$ $S_A = 0,642$ $n_A = 15$ $\sigma_A = \sigma_B$ (desconocidas)

$\bar{X}_B = 6,793$ $S_B = 1,2$ $n_B = 12$ $\alpha = 0.05$

$$H_0: \mu_A = \mu_B \quad vs \quad H_1: \mu_A \neq \mu_B$$

$$e_m = \frac{(\bar{X}_B - \bar{X}_A) - \Delta_0}{S_p\sqrt{\frac{1}{n_B} + \frac{1}{n_A}}}$$

$$T_{obs} = \frac{\bar{X}_B - \bar{X}_A}{0,864\sqrt{3/20}} \sim t_{25} \text{ (bajo } H_0)$$

$$S_p = \frac{(n_B - 1)S_B^2 + (n_A - 1)S_A^2}{n_A + n_B - 2}$$

Rechazo H_0 si $|T_{obs}| > t_{25, 0.025}$

$$\Rightarrow S_p = \frac{11 \times 1,2^2 + 14 \times 0.642^2}{25} = 0,864$$

$$T_{obs} = \frac{6,793 - 5,617}{0,3346} = 3,5144$$

$$t_{25, 0.025} = 2,060$$

$\Bigg\}$ $T_{obs} > t_{25, 0.025} \Rightarrow$ Rechazo H_0

Hay diferencias estadísticamente significativas entre las medias de las dos poblaciones

b) Indicar en el contexto de esta prueba qué sería cometer error tipo I y error tipo II

Error tipo I: Rechazo H_0 cuando H_0 es verdadera $\rightarrow$ Decidir que $\mu_A \neq \mu_B$ cuando, en realidad, $\mu_A = \mu_B$

Error tipo II: No rechazar H_0 cuando H_0 es falsa $\rightarrow$ Decidir que $\mu_A = \mu_B$ cuando, en realidad, son distintas

NOTA

Final PyE UTN HOJA N° FECHA 21-2-18

(3) Se realizó una encuesta sobre el perfil de la audiencia del sitio web ESPN con el fin de estimar la proporción poblacional de mujeres entre los usuarios. Entre los 400 encuestados, 104 resultaron mujeres. Asumiendo que el tamaño poblacional es muy grande:

a) Proponga un estimador insesgado para la proporción poblacional de mujeres

X_i: "la iésima persona encuestada es mujer" $X_i \sim Be(p)$

$i \in [1, 400]$ $n = 400$, $R(X) = \{0, 1\}$

$$\overbrace{\hat{p} = \bar{X}}^{\text{tomo}} = \frac{\sum_{i=1}^{400} X_i}{400} = \frac{104}{400} = \boxed{0{,}26 = \hat{p}}$$

$X_i \sim Be(p) \rightarrow E(X_i) = p \quad V(X_i) = p(1-p)$

$$\text{sesgo} = E(\hat{p}) - p = E(\bar{X}) - p = E\left(\frac{\sum_{i=1}^{400} X_i}{400}\right) - p = \frac{1}{400} E\left(\sum_{i=1}^{400} X_i\right) - p =$$

$$= \frac{1}{400} \sum_{i=1}^{400} E(X_i) - p = \frac{1}{400} \cdot 400 \underbrace{E(X_i)}_{p} - p = 0 \checkmark$$

b) Estime mediante un intervalo del 95% de confianza de esta proporción

$1 - \alpha = 0.95 \rightarrow \alpha = 0.05$; $\hat{p} = 0.26$; $z_{1-\frac{\alpha}{2}} = z_{0.975} = 1{,}96$; $n = 400$

$$I.C_{095}(p) = \left[\hat{p} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} ; \hat{p} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right] = [0{,}217 ; 0{,}303]$$

$$\boxed{IC_{095}(p) = [0.217 ; 0.303]}$$

c) ¿Cuál sería el tamaño muestral necesario para que la longitud del intervalo resulte de a lo sumo 0.06?

long. intervalo $= 0.06 \rightarrow$ long. error $= 0.03 = z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ n desconocido

$$\rightarrow 0.03 = 1{,}96 \sqrt{\frac{0{,}26(1-0{,}26)}{n}} \rightarrow \frac{0.03}{1{,}96} = \sqrt{\frac{0.1924}{n}} = \frac{\sqrt{0.1924}}{\sqrt{n}}$$

$$\rightarrow \frac{0.03}{1.96 \times \sqrt{0.1924}} = \frac{1}{\sqrt{n}} \rightarrow n = \frac{1{,}96^2 \times 0{,}1924}{0{,}03^2} = 821{,}25 \rightarrow \boxed{n = 822}$$

NOTA

4) En un negocio hay dos vendedoras: Claudia y Patricia. De los clientes que atiende Claudia el 75% efectúa la compra; mientras que de los clientes que atiende Patricia sólo lo hacen el 65%.
Los clientes eligen con igual probabilidad a ambas pero Claudia falta el 12% de los días y Patricia el 10%. Cuando falten ambas vendedoras, atiende el dueño que vende con seguridad.

a) ¿qué porcentaje de los clientes que llegan al negocio compra?

V: "el cliente que entró al local realizó la compra".
Cl: "el cliente que entró al negocio fue atendido por Claudia" $P(V|Cl) = 0.75$
Pa: " " " " " " " " Patricia" $P(V|Pa) = 0.65$
D: " " " " " " " " el Dueño" $P(V|D) = 1$

C presente: "Claudia está en el local" $P(C \text{ presente}) = 0.88$
P presente: "Patricia está en el local" $P(P \text{ presente}) = 0.9$

- $P(Cl) = P(C\text{ presente} \cap \overline{P\text{ presente}}) \cdot 1 + P(C\text{ presente} \cap P\text{ presente}) \cdot 0,5 \overset{\text{independ.}}{=}$
$$= P(C\text{ presente}) \cdot \underbrace{P(\overline{P\text{ presente}})}_{1-P(P\text{ presente})} + P(C\text{ presente}) \cdot P(P\text{ presente}) \cdot 0,5 =$$
$$= 0,88 \cdot 0,10 + (0,88 \cdot 0,9) \cdot 0,5 = \boxed{0,484 = P(Cl)}$$

Similar que con P(Cl) x independencia

- $P(Pa) = \underbrace{P(\overline{C\text{ presente}})}_{0,12} \cdot \underbrace{P(P\text{ presente})}_{0,9} + \underbrace{P(C\text{ presente})}_{0,88} \cdot \underbrace{P(P\text{ presente})}_{0,9} \cdot 0,5 = \boxed{0,504 = P(Pa)}$
- $P(D) = P(\overline{C\text{ presente}} \cap \overline{P\text{ presente}}) \overset{\text{indep}}{=} \underbrace{P(\overline{C\text{ presente}})}_{0,12} \underbrace{P(\overline{P\text{ presente}})}_{0,1} = \boxed{0,012 = P(D)}$

$$\rightarrow P(V) \overset{\text{Prob. total}}{=} P(V|Cl)\,P(Cl) + P(V|Pa)\,P(Pa) + P(V|D)\,P(D) =$$
$$= 0.75 \times 0.484 + 0.65 \times 0.504 + 1 \times 0.012 = \boxed{0,7026 = P(V)}$$

b) ¿Si compró un cliente, qué probabilidad hay de que lo haya atendido Claudia?

$$P(Cl|V) \overset{\text{Bayes}}{=} \frac{P(V|Cl)\,P(Cl)}{P(V)} = \frac{0.75 \times 0.484}{0.7026} = 0,5166$$

$$\boxed{P(Cl|V) = 0,5166}$$

NOTA

HOJA N°

FECHA 21-2-18

Final PyE UTN

Teórico 1

a) Se sabe que los estimadores $\hat{\theta}_1$ y $\hat{\theta}_2$ son insesgados para θ. Mostrar que $\hat{\theta}_3 = \alpha\hat{\theta}_1 + (1-\alpha)\hat{\theta}_2$, con $\alpha > 0$ también es un estimador insesgado de θ

$$\hat{\theta}_1 \text{ y } \hat{\theta}_2 \text{ insesgados} \rightarrow \underbrace{E(\hat{\theta}_1) = \theta}_{\mathrm{I}} \wedge E(\hat{\theta}_2) = \theta \rightarrow \underbrace{E(\hat{\theta}_1) = E(\hat{\theta}_2)}_{\mathrm{II}}$$

- $\hat{\theta}_3 = \alpha\hat{\theta}_1 + (1-\alpha)\hat{\theta}_2, \ \alpha > 0$

$\rightarrow \hat{\theta}_3$ es insesgado si $E(\hat{\theta}_3) = \theta$

$$E(\hat{\theta}_3) = E(\alpha\hat{\theta}_1 + (1-\alpha)\hat{\theta}_2) = E(\alpha\hat{\theta}_1) + E((1-\alpha)\hat{\theta}_2) =$$

$$= \alpha E(\hat{\theta}_1) + (1-\alpha)E(\hat{\theta}_2) = \alpha E(\hat{\theta}_1) + E(\hat{\theta}_2) - \alpha E(\hat{\theta}_2) =$$

$$\overset{\mathrm{II}}{=} \cancel{\alpha E(\hat{\theta}_1)} + E(\hat{\theta}_1) - \cancel{\alpha E(\hat{\theta}_1)} = E(\hat{\theta}_1) \overset{\mathrm{I}}{=} \theta$$

$$\therefore E(\hat{\theta}_3) = \theta \rightarrow \hat{\theta}_3 \text{ es insesgado}$$

b) Si $V(\hat{\theta}_1) = \alpha^2$ y $V(\hat{\theta}_2) = 3\alpha^2$ para qué valor de α el estimador de $\hat{\theta}_3$ alcanza la menor varianza

$$V(\hat{\theta}_3) = V(\alpha\hat{\theta}_1 + (1-\alpha)\hat{\theta}_2) = V(\alpha\hat{\theta}_1) + V((1-\alpha)\hat{\theta}_2) =$$

$$= \alpha^2 V(\hat{\theta}_1) + (1-\alpha)^2 V(\hat{\theta}_2) = \alpha^2 V(\hat{\theta}_1) + (1-2\alpha+\alpha^2)V(\hat{\theta}_2) =$$

$$= \alpha^2 V(\hat{\theta}_1) + V(\hat{\theta}_2) - 2\alpha V(\hat{\theta}_2) + \alpha^2 V(\hat{\theta}_2) =$$

x enunciado

$$= \alpha^2\alpha^2 + 3\alpha^2 - 2\alpha 3\alpha^2 + \alpha^2 3\alpha^2 = \underbrace{\alpha^4 + 3\alpha^2 - 6\alpha^3 + 3\alpha^4}_{4\alpha^4 - 6\alpha^3 + 3\alpha^2} =$$

$$= \alpha^2(4\alpha^2 - 6\alpha + 3)$$

$\alpha > 0$ x enunciado $\rightarrow$ busco α / $4\alpha^2 - 6\alpha + 3$ sea mínimo

Sea $g(\alpha) = 4\alpha^2 - 6\alpha + 3$

busco α tal que $g'(\alpha) = 0 \rightarrow g'(\alpha) = 8\alpha - 6 = 0 \rightarrow \alpha = 6/8 \rightarrow \alpha = 3/4$

analizo si es mínimo $\rightarrow g''(\alpha) = 8 > 0 \rightarrow \alpha = 3/4$ es mínimo

$$\boxed{\alpha = 0.75}$$

NOTA

teórico 2

Probar que A y B son eventos independientes de un mismo espacio muestral si y sólo si se verifica que:

$$P(A\cup B) = P(A) + P(\bar{A})\,P(B)$$

→ 1°: Pruebo que esa ecuación se cumple si $\overbrace{A \text{ y } B \text{ son indep.}}^{\text{hipótesis}}$

$$P(A\cup B) = P(A) + P(B) - P(A\cap B) \overset{A \text{ y } B \text{ indep.}}{=} P(A) + P(B) - P(A).P(B) =$$

$$= P(A) + P(B)\,(1 - P(A)) = P(A) + P(B)\,P(\bar{A}) =$$

$$= \boxed{P(A) + P(\bar{A}).P(B)}\ ✓$$

→ 2°: Pruebo que A y B son indep. si se cumple $\overbrace{\text{esa ecuación}}^{\text{hipótesis}}$

A y B son indep. si $P(A\cap B) = P(A).\,P(B)$

$$\bullet\ P(A\cap B) = 1 - P(\overline{A\cap B}) \overset{\text{De Morgan}}{=} 1 - P(\bar{A}\cup\bar{B}) =$$

$$= 1 - [P(\bar{A}) + P(\bar{B}) - P(\bar{A}\cap\bar{B})] =$$

De Morgan $$= 1 - P(\bar{A}) - P(\bar{B}) + P(\overline{A\cup B}) =$$

$$= \underbrace{1 - P(\bar{A})}_{P(A)} - P(\bar{B}) + [1 - \underbrace{P(A\cup B)}_{}] =$$

hipótesis $$= P(A) - P(\bar{B}) + 1 - [P(A) + P(\bar{A})\,P(B)] =$$

$$= \cancel{P(A)} + \underbrace{1 - P(\bar{B})}_{P(B)} - \cancel{P(A)} - P(\bar{A})\,P(B) =$$

$$= P(B) - P(\bar{A})\,P(B) = P(B)\,(1 - P(\bar{A})) =$$

$$= P(B).P(A) = \boxed{P(A).P(B)}\ ✓$$

$$\therefore \boxed{P(A\cap B) = P(A).P(B)} \rightarrow A \text{ y } B \text{ son indep.}$$

NOTA