

Exercici 3

[2,5 punts en total]

Puntuació total de l'exercici 3	
---------------------------------	--

Una empresa de telecomunicacions considera que una pàgina web és eficient si el seu temps de càrrega és inferior a 3 segons. L'empresa afirma que almenys el 50 % de les pàgines web que gestiona són eficients. Per comprovar aquesta afirmació, seleccionem una mostra aleatòria de $n = 2.500$ pàgines web gestionades per aquesta empresa.

FÓRMULES PER A RESOLDRE L'EXERCICI:

- $Z \sim \text{normal}(0, 1) \rightarrow P(-1,96 \leq Z \leq 1,96) = 0,95$ i $P(-2,58 \leq Z \leq 2,58) = 0,99$

- Intervals de confiança amb un nivell de confiança $\gamma \in (0, 1)$

— per a la proporció (mostres grans):

$$\left[\hat{p} - z_\gamma \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_\gamma \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right]$$

— per a la mitjana (mostres normals amb la variància σ^2 coneguda):

$$\left[\bar{x} - z_\gamma \frac{\sigma}{\sqrt{n}}, \bar{x} + z_\gamma \frac{\sigma}{\sqrt{n}} \right]$$

— per a la mitjana (mostres grans amb la variància σ^2 desconeguda):

$$\left[\bar{x} - z_\gamma \frac{s}{\sqrt{n}}, \bar{x} + z_\gamma \frac{s}{\sqrt{n}} \right]$$

- 3.1.** Suposem que, cada vegada que seleccionem una pàgina web per a la mostra, aquesta pot ser eficient (o no ser-ho) amb una probabilitat $p = 0,5$ i que el fet que una pàgina sigui eficient (o no) és independent del fet que la resta de pàgines ho siguin (o no). Considerem la variable aleatòria X , que compta el nombre de pàgines, d'entre les 2.500 de la mostra, que són eficients.

Quina distribució té la variable X ? Calculeu la probabilitat que com a molt 1.299 pàgines siguin eficients. Per a fer-ho, utilitzeu l'aproximació per la distribució normal sense fer la correcció per continuïtat.

[1,25 punts]

Puntuació de l'apartat 3.1	
----------------------------	--

3.

- 3.1. Anomenem x al nombre de pàgines web eficients d'entre les 2.500 de la mostra. Si les dades de l'empresa són correctes, aquesta variable seguirà una distribució binomial de paràmetres $n = 2500$ i $p = 0,5$.

Per tant, $x \sim \text{Binomial}(n = 2500, p = 0,5)$.

Volem calcular $P(x \leq 1299)$.

La distribució binomial es pot aproximar per una normal quan n és gran i p no és massa propera a 0 o 1. Concretament es pot aproximar per una normal de mitjana np i variància $np(1 - p)$.

Observem que

$$np = 2500 \cdot 0,5 = 1250$$

i

$$np(1 - p) = 2500 \cdot 0,5 \cdot (1 - 0,5) = 625$$

Per tant, la variable x es pot aproximar per una $N(1250, 625)$.

Aleshores, si anomenem Z a la distribució $N(0,1)$, la $P(X \leq 1299)$ es pot aproximar de la següent manera:

$$P(X \leq 1299) \approx P\left(Z \leq \frac{1299 - 1250}{\sqrt{625}}\right) = P(Z \leq 1,96)$$

Sabem que $P(-1,96 \leq Z \leq 1,96) = 0,95$ i, per tant, per la simetria de la distribució normal, $P(Z \leq 1,96) = 0,975$.

Per tant, si les dades de l'empresa són correctes, la probabilitat que com a molt 1.299 pàgines de la mostra siguin eficients és aproximadament de 0,975.

- 3.2. En la mostra de $n = 2.500$ pàgines web, hem obtingut una mitjana mostral del temps de càrrega de $\bar{x} = 2,95$ segons i una desviació típica mostral de $s = 0,38$ segons. Suposeu que el temps de càrrega de les pàgines web segueix una distribució normal

Construïu un interval de confiança del 95 % per al temps mitjà de càrrega de les pàgines web que gestiona l'empresa. A partir de l'interval obtingut, què podem dir sobre l'afirmació de l'empresa que almenys el 50 % de les pàgines web que gestiona són eficients?

[1,25 punts]

Puntuació de l'apartat 3.2	
----------------------------	--

- 3.2. L'interval de confiança per a la mitjana quan les mostres són grans i la variància σ^2 és desconeguda és donat per:

$$\left[\bar{x} - z_{\gamma} \frac{s}{\sqrt{n}}, \bar{x} + z_{\gamma} \frac{s}{\sqrt{n}} \right]$$

En aquest cas tenim les dades següents:

Mida mostral: $n = 2500$

Mitjana mostral: $\bar{x} = 2.95$ s

Desviació típica mostral: $s = 0.38$ s

Nivell de confiança: $\gamma = 0.95 \Rightarrow z_{\gamma} = 1,96$

Per tant,

$$\bar{x} - z_{\gamma} \frac{s}{\sqrt{n}} = 2,95 - 1,96 \cdot \frac{0,38}{\sqrt{2500}} = 2,9351$$

i

$$\bar{x} + z_{\gamma} \frac{s}{\sqrt{n}} = 2,95 + 1,96 \cdot \frac{0,38}{\sqrt{2500}} = 2,9649$$

Així, l'interval de confiança del 95 % per a la mitjana del temps de càrrega de les pàgines web en segons és $[2,9351, 2,9649]$.

Com que tot l'interval és inferior a 3 segons, podem afirmar, amb un nivell de confiança del 95 %, que el temps mitjà de càrrega és inferior a 3 segons.

A més, si suposem que el temps de càrrega segueix una distribució normal, aquesta és simètrica i el valor central de la distribució coincideix amb la mitjana. Per tant, si la mitjana és inferior a 3 segons, més de la meitat de les pàgines web tenen un temps de càrrega inferior a 3 segons.

En conseqüència, les dades obtingudes són coherents amb l'afirmació de l'empresa que almenys el 50 % de les pàgines web que gestiona són eficients