Es. 5.4 PAG. 139

5A

$n = 2003$

$\hat{\pi}$ = stima puntuale di chi ha risposto sì (proporzione campionaria)

54 % sì
42 % NO
4 % NON SO

se = ?

$\hat{\pi} = 0.54 \qquad 1 - \hat{\pi} = 1 - 0.54 = 0.46$

$$se = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}} = \sqrt{\frac{0.54 \cdot 0.46}{2003}} = \sqrt{\frac{0.2484}{2003}} = 0.0111 = 0.01$$

↳ è una stima dell'errore standard (sostituisco nella formula dell'errore standard la proporzione campionaria). $\sigma_{\hat{\pi}} = \sqrt{\frac{\pi(1-\pi)}{n}}$

↳ errore standard

Es. 5.5 PAG. 139

$Sì_D = 51\%$, $n = 1008$ → margine d'errore = 3.1 % = 0.031

⇓ come ottengo ?

Margine d'errore = $z \cdot se$

$\hat{\pi} = 0.51 \qquad 1 - \hat{\pi} = 0.49$

↳ stima puntuale della proporz. di pop. che ha risposto sì.

$$se = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}} = \sqrt{\frac{0.51 \cdot 0.49}{1008}} = \sqrt{\frac{0.2499}{1008}} = 0.0157 = 0.016$$

$0.031 = z \cdot 0.016$

$z = \frac{0.031}{0.016} = 1.94$ → N.B. Influenza degli arrotondamenti

Es. 5.6 PAG. 139

$P(sì)_D = \frac{90}{142} = 0.63$ a) $\hat{\pi}_D = 0.63$

$P(sì)_R = \frac{26}{102} = 0.25$ $\hat{\pi}_R = 0.25$

b) 95 % IC per $\hat{\pi}_D$ ⇒ $\hat{\pi}_D \pm z \cdot se$

↳ lo z-score viene scelto in modo tale che la probab. sottesa alla curva normale entro z errori standard dalla media sia pari al livello di fiducia.

TAVOLA A → probab. sottesa alla coda di dx → 0.025 ←

↳ $z = 1.96$

$se = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}} = \sqrt{\frac{0.63 \cdot 0.37}{142}} = \sqrt{\frac{0.2331}{142}} = 0.04$

95% IC —> $0.63 \pm 1.96 \cdot 0.04$
0.63 ± 0.08 $(0.55 - 0.71)$ *

95% IC per $\hat{\pi}_R$ => $0.25 \pm 1.96 \cdot se$

$se = \sqrt{\frac{0.25 \cdot 0.75}{102}} = \sqrt{\frac{0.1875}{102}} = 0.04$

—> $0.25 \pm 1.96 \cdot 0.04$
0.25 ± 0.08 $(0.17 - 0.33)$ * → stima intervallare con un livello di fiducia pari a 0.95

* IC: Intervallo di valori entro cui ricade il parametro π
La probabilità associata al fatto che l'intervallo contenga il parametro è denominata livello di fiducia.

Es. 5.7 PAG. 139

$P(sì) = 0.36$ —> $\hat{\pi} = 0.36$
$n = 883$

a) se = ? $se = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}} = \sqrt{\frac{0.36 \cdot (1-0.36)}{883}} = \sqrt{\frac{0.2304}{883}} = 0.016$

b) Margine d'errore = z · se
ME = ? per 95% IC

$z = 1.96$
$ME = 1.96 \cdot 0.016 = 0.031$
↳ è un multiplo (uno z-score) dell' errore standard.

c) 95% IC

$\hat{\pi} \pm z \cdot se$

$\left.\begin{array}{l} \hat{\pi} = 0.36 \\ z = 1.96 \\ se = 0.016 \end{array}\right\} = ME = 0.03$

—> 0.36 ± 0.03 $(0.33, 0.39)$

Es. 5.9 PAG. 139

229 sì
$n = 1200$

$\hat{\pi} = \frac{229}{1200} = 0.19$

(i) 95% CI $\longrightarrow$ $\hat{\pi} \pm z \cdot se$ $\qquad se = \sqrt{\frac{0.19 \cdot 0.81}{1200}} = \sqrt{\frac{0.1539}{1200}} = 0.0113$

$0.19 \pm 1.96 \cdot 0.01$
$0.19 \pm 0.02 \quad (0{,}17 , 0.21)$

(ii) 99% CI $\longrightarrow$ $z = 2.58$
$0.19 \pm 2.58 \cdot 0.01$
$0.19 \pm 0.03 \quad (0.16 , 0.22)$

Es. 5.17 PAG 140

40% J
60% S

a) $n = 400$
99% IC $\hat{\pi}_J$ $\quad z = 2.58$

$\hat{\pi} \pm z \cdot se$

$se = \sqrt{\frac{0.40 \cdot 0.60}{400}} = \sqrt{\frac{0.24}{400}} = 0.02$

$0.40 \pm 2.58 \cdot 0.02$
$= 0.40 \pm 0.05$
$(0.35 , 0.45)$

b) $n = 40$
99% IC $\hat{\pi}_J$ $\quad z = 2.58$

$\hat{\pi} \pm z \cdot se$

$se = \sqrt{\frac{0.24}{40}} = 0.08$

$0.40 \pm 2.58 \cdot 0.08$
$= 0.40 \pm 0.21$
$(0.19 ; 0.61)$

$\Longrightarrow$ Saresti disposto a prevedere il vincitore?

a) SÌ
perchè tutti i valori contenuti nell' IC sono sotto 0.5
$\hookrightarrow$ VINCITORE Smith

b) NO
perchè ci sono valori contenuti nell' interv. di confidenza sia sopra che sotto 0.5
$\hookrightarrow$ ESITO INCERTO

NB. La stima puntuale è uguale MA al diminuire di n aumenta se e quindi l'intervallo risulta più largo.

Es. 5.22 PAG. 141

$n = 497$
$me = 2$
$\bar{y} = 3.02$
$s = 1.81$

$gdl = n - 1 = 497 - 1 = 496 \longrightarrow gdl > 30$
$gdl \to \infty \longrightarrow$ t o z è uguale

a) Stima puntuale della media della popolazione
$\bar{y} = 3.02$

b) $se = \frac{s}{\sqrt{n}} = \frac{1.81}{\sqrt{497}} = \frac{1.81}{22.29} = 0.08$

IC = intervallo di valori attorno a stima puntuale costruito per contenere con una prefissata probab. il valore del parametro.

c) 95 % IC $\quad t = z = 1.96$

$\bar{y} \pm t \cdot se = 3.02 \pm 1.96 \cdot 0.08$
$= 3.02 \pm 0.16 \quad (2.9, 3.2)$

d) 99 % IC $\quad t = 2.58 \; (2.576)$

$= 3.02 \pm 2.58 \cdot 0.08$
$= 3.02 \pm 0.21 \quad (2.81, 3.23)$

Poco probabile, non plausibile perchè 99 % IC valore minimo è 2.81 (> 2.0).

Es 5.23 PAG. 141

$n = 397$
$\bar{y} = 2.89$
$s = 1.77$

a) $se = \frac{s}{\sqrt{n}} = \frac{1.77}{\sqrt{397}} = \frac{1.77}{19.92} = 0.089$

b) 95 % IC
gdl = ∞
$\bar{y} \pm t \cdot se = 2.89 \pm 1.96 \cdot 0.089 = 2.89 \pm 0.17 \quad (2.7, 3.1)$

Es 5.24 PAG. 141

$n = 17$

a) $\bar{y} = \frac{\Sigma y_i}{n} = \frac{11 + 11 + \dots + 6 + 11}{17} = \frac{124}{17} = 7.29$

$s = \sqrt{\frac{\Sigma (y_i - \bar{y})^2}{n-1}} = 7.18$

b) $se = \frac{s}{\sqrt{n}} = \frac{7.18}{\sqrt{17}} = \frac{7.18}{4.12} = 1.74$

c) gdl $= n - 1 = 17 - 1 = 16$
95 % IC ? → t-score = TAVOLA B = 2.12

d) 95% IC ?
$\bar{y} \pm t(se) = 7.29 \pm 2.12 \cdot 1.74 = 7.29 \pm 3.69 \quad (3.6, 11)$ (10.98)

Es. 5.35 PAG. 142

5c

$n = ?$ con $M = 0.06$
90 % IC
$\pi = 0.30$

↓

la dimensione campionaria per la stima entro M è:

$n = \pi(1-\pi)\left(\frac{z}{M}\right)^2$

z per probab. $\frac{(1-0.90)}{2} = 0.05$

↳ $z = 1.64$ (TAVOLA A)

$= 0.30 \cdot 0.70 \cdot \left(\frac{1.64}{0.06}\right)^2$

$= 0.21 \cdot (27.33)^2 = 0.21 \cdot 746{,}93 = 156.18 = 157$

Es. 5.37 PAG. 142

$M = 0.02$ $n = ?$
95 % IC

z per probab. sottesa alla coda di dx (TAVOLA A) pari a $(0.05/2 = 0.025)$

a) con $\pi = 0.10$

$n = \pi(1-\pi)\left(\frac{z}{M}\right)^2 = 0.10 \cdot 0.90 \cdot \left(\frac{1.96}{0.02}\right)^2$

$= 0.09 \cdot 9604 = 864{,}36 = 864$

↓

b) con $\pi = 0.50$ → approccio prudenziale

$n = 0.50 \cdot 0.50 \cdot \underbrace{9604}_{\left(\frac{z}{M}\right)^2} = 2401$ → Notevole riduzione della dimensione campionaria se scelgo un valore di π più appropriato

Es. 5.27 PAG. 141

$n = 1415$
$Me = 16$ anni
$\bar{y} = 20.3$
$s = 18.2$

a) Probabilmente no. Media e mediana non coincidono, ma la media è superiore alla mediana → ASIMMETRICA POSITIVA

Inoltre valore ipotetico minimo della distribuzione è 0.

$z = \frac{y_i - \bar{y}}{s} = \frac{0 - 20.3}{18.2} = -1.12$

La media è a sole 1.12 dev. st. SOPRA lo 0 (in una distrib. normale $\mu - 3\sigma$ e $\mu + 3\sigma$ comprende tot. dei valori)

↳ ASIMMETRICA POSITIVA

b) Sì, n è molto grande perciò la distribuzione campionaria è approx normale per il teorema del limite centrale.

99% IC ?

$\bar{y} \pm t \cdot se$ $\quad t = z = 2.576$ perchè $n = 1415 \rightarrow gdl = 1414 = \infty$ TAVOLA B

$= 20.3 \pm 2.576 \cdot 0.484$ $\quad se = \frac{s}{\sqrt{n}} = \frac{18.2}{\sqrt{1415}} = 0.484$

$= 20.3 \pm 1.25 \quad (19.05, 21.55)$

Scanned by CamScanner

BACKUP

5.12 P. 140

a) $\frac{x}{n} = \frac{1885}{2815}$

b) Sì, tutti i valori contenuti nell'intervallo di confidenza sono superiori a 0.50.

c) 95% IC contrari

0.652 - 0.6696 = -0.0176 → 0.0176 margine d'errore

→ $\underbrace{0.3304}_{1 - 0.3304} \pm 0.0176$

5.11 PAG. 139

Z-score ?

a) 0.98 IC

probab. sottesa alle 2 code: 0.02 ($\uparrow$ 1 - 0.98)

probab. sottesa alla coda di dx: 0.02/2 = 0.01 → Z = 2.33 (TAVOLA A) (valore più vicino a probab.)

b) 0.90 IC

(0.10/2 = 0.05)

Z = 1.64 (probab. che contiene valore)

c) 0.50 IC

(0.50/2 = 0.25)

Z = 0.67 (più vicino)

d) 0.9973

(0.0027/2 = 0.00135)

Z = 3.00

Es. 5.10 PAG. 139

95% IC $\hat{\pi}$ P(sì) = (0.87, 0.90)

99% IC più ampio? Sì

perché? M (margine d'errore) maggiore perché Z è più alto

↳ livello di fiducia maggiore: intervallo di confidenza si allunga perché cresce il valore dello Z-score

↳ ho una stima intervallare meno precisa → per essere sicuri di includere il vero valore del parametro, sacrifichiamo la precisione

Scanned by CamScanner