Università degli Studi di Cassino

Esercitazione di Statistica 2 del 22.02.2007

Simona Balzano

Esercizio 1

Un processo automatico di produzione è caratterizzato da una distribuzione normale e varianza pari a 4.

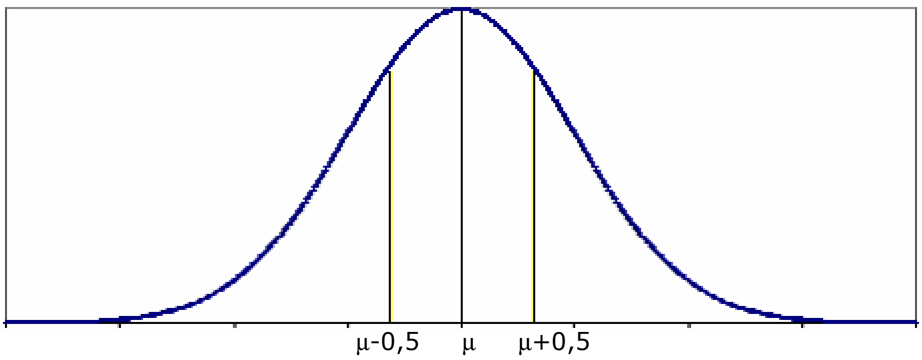
a) Qual è la probabilità di avere una divergenza dalla media di 0,5 in un campione di 25 unità?
b) Quanto deve essere ampio il campione se per tale errore si ammette un rischio pari al 1%?

Soluzione

a)

La probabilità di commettere un errore e = 0,5 intorno alla media equivale alla probabilità associata ad un intervallo di ampiezza 2 × 0,5 = 1 centrato sulla media:

$$\mu \pm \bar{x} = 2 \times 0,5 = 1$$



Tale probabilità può essere calcolata attraverso la v.c. normale standardizzata come probabilità che Z sia esterna all'intervallo [μ - 0,5; μ + 0,5].
In pratica, se:

$$P\left(Z > z_{\alpha/2}\right) = \frac{\alpha}{2}$$

allora:

$$\frac{\alpha}{2} = P\left(Z > \frac{x - \mu}{\sigma/\sqrt{n}}\right)$$

$$\alpha = 2 \times P\left(Z > \frac{0,5}{2/\sqrt{25}}\right) = 2 \times P(Z > 1,25) = 2 \times (1 - F(1,25)) = 2 \times (1 - 0,8944) = \mathbf{0,2112}$$

b)

$\alpha = 0,01$

$$\frac{\alpha}{2} = P\left(z_{\alpha/2} > \frac{x - \mu}{\sigma/\sqrt{n}}\right) = 0,005$$

ciò vuol dire che $\frac{x - \mu}{\sigma/\sqrt{n}} = 2,58$, quindi che:

$$n = \left(\frac{2}{0,5} 2,58\right)^2 = 10,32^2 = 106,5 \cong \mathbf{107}$$

Esercizio 2

Si vuole stimare la lunghezza media delle foglie di una data specie di albero attraverso un campione di ampiezza n = 20 da cui risulta un valore medio pari a 98 mm. Sapendo che la varianza è pari a 9, costruire un intervallo di confidenza per la media al livello di significatività del 95 %.

Soluzione

Intervallo di confidenza per la v.c. <u>media (μ).</u>

<u>Popolazione non nota, varianza nota, n < 30</u>

In questo caso si può costruire l'intervallo di confidenza sfruttando la disuguaglianza di Cebicev:

$\bar{x} = 98$
$\sigma^2 = 9$
$n = 20$
$\alpha = 0,05$

$$P\left(\left|\bar{x} - \mu\right| \leq k\sigma\right) \geq 1 - \frac{1}{k^2}$$

$$1 - \alpha = 0,05 = 1 - \frac{1}{k^2}$$

da cui consegue che:

$$\frac{1}{k^2} = 0,05$$

e quindi che:

$$k = \sqrt{\frac{1}{0,05}} = 4,47$$

Definiamo quindi l'intervallo per la media a partire dalla seguente espressione:

$$P\left(\left|98 - \mu\right| \leq 4,47 \cdot 3\right) \geq 0,05$$

da cui si ricava:

$$P\left(\left|98 - \mu\right| \leq 13,41\right) \geq 0,05; \qquad P\left(98 - 6,705 \leq \mu \leq 98 + 6,705\right) \geq 0,05,$$

quindi:

$$P\left(91,295 \leq \mu \leq 104,705\right) \geq 0,05$$

Esercizio 3

Per verificare l'efficacia di un corso di statistica vengono confrontati i rendimenti medi di due campioni di studenti di ampiezza $n_1 = n_2 = 50$, di cui uno costituito da studenti frequentanti l'altro da studenti non frequentanti. Tali valori risultano essere pari rispettivamente a 28 e 25.
Sapendo che le popolazioni da cui sono stati estratti i campioni sono distribuite normalmente con varianze pari a 16, costruire un intervallo di confidenza al 95% per la differenza tra le medie delle due popolazioni.

Soluzione

Intervallo di confidenza per la v.c. <u>differenza tra medie (μ_1 - μ_2).</u>

<u>Popolazione Normale, varianze note</u>

$\bar{x}_1 = 28 \qquad \bar{x}_2 = 25$

$\sigma_1^2 = \sigma_2^2 = 16$

$n_1 = n_2 = 50$

$\alpha = 0,05$

$$IC_{1-\alpha} = \left|\bar{x}_1 - \bar{x}_2\right| \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$z_{\alpha/2} = 1,96$

$$IC_{0,95} = (28 - 25) \pm 1,96\sqrt{2 \cdot \frac{16}{50}} = 3 \pm 1,568 =$$
$$= \left[\mathbf{1,432;\ 4,568}\right]$$

$$P\left(1,432 \leq \left|\mu_1 - \mu_2\right| \leq 4,568\right) = 0,95$$

L'intervallo di confidenza non contiene lo 0: questo vuol dire che una differenza nulla tra le medie è poco probabile, ossia che esiste una differenza *significativa* (significativamente diversa da 0) tra il rendimenti medio dei frequentanti e quello dei non frequentanti, il che a sua volta vuol dire che il corso è efficace.

Esercizio 4

Per valutare se esiste o meno una differenza tra il reddito medio di uomini e donne si estraggono due campioni di 25 donne e 20 uomini, da cui risulta un reddito medio di 9 (migliaia di euro) con s.q.m. di 2 per le donne e 15 con s.q.m. di 3 per gli uomini. Costruire un intervallo di confidenza per la differenza tra i redditi medi al livello di significatività del 99%

Soluzione

Intervallo di confidenza per la v.c. differenza tra medie (μ_1 - μ_2).

Popolazione non nota, varianze non note, campioni piccoli

$\bar{x}_1 = 9$ $\quad\bar{x}_2 = 15$

$s_1 = 2$ $\quad s_2 = 3$

$n_1 = 25$ $\quad n_2 = 20$

$\alpha = 0{,}01$

Ottenuta la stima della varianza comune, attraverso lo stimatore:

$$\tilde{s}^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

l'intervallo di confidenza è costruito come:

$$IC_{1-\alpha} = \left|\bar{x}_1 - \bar{x}_2\right| \pm t_{\alpha/2;n_1+n_2-2}\,\tilde{s}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$\tilde{s}^2 = \frac{(25-1)\cdot 4 + (20-1)\cdot 9}{25+20-2} = 6,21$$

$$t_{\alpha/2;n_1+n_2-2} = t_{0,005;43} = 2,704$$

$$IC_{0,99} = |9-15| \pm 2,704 \cdot 6,21\sqrt{\frac{1}{25} + \frac{1}{20}} = 6 \pm 5,0376 =$$

$$= \left[\mathbf{0,9624;\ 11,0376}\right]$$

$$P\left(0,9624 \le |\mu_1 - \mu_2| \le 11,0376\right) = 0,99$$

L'intervallo di confidenza non contiene lo 0: questo vuol dire che esiste una differenza *significativa* (significativamente diversa da 0) tra il reddito medio di donne e uomini (il sesso incide sul livello medio di reddito).

Esercizio 5

Per stimare la proporzione π di studenti promossi ad un esame si estrae un campione di 40 studenti, su cui si osserva una percentuale di promossi pari al 25%. Determinare un intervallo di confidenza per π, al livello di significatività

a) del 95%;
b) del 98%.

Soluzione

Intervallo di confidenza per la v.c. <u>proporzione π.</u>

p = 0,25
n = 40

$$IC_{1-\alpha} = \left[p \pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \right]$$

a)

α = 0,05
$Z_{0,025}$ = 1,96

$$IC_{0,95} = \left[0,25 \pm 1,96\sqrt{\frac{0,25 \times 0,75}{40}} \right] = [0,25 \pm 0,13] =$$
$$= \mathbf{[0,12;\ 0,38]}$$

$$P(0,12 \leq \pi \leq 0,38) = 0,95$$

b)

α = 0,01
$Z_{0,005}$ = 2,58

$$IC_{0,95} = \left[0,25 \pm 2,58\sqrt{\frac{0,25 \times 0,75}{40}} \right] = [0,25 \pm 0,17] =$$
$$= \mathbf{[0,08;\ 0,42]}$$

$$P(0,08 \leq \pi \leq 0,42) = 0,99$$

ALLEGATO

Schema per la costruzione di un intervallo di confidenza per la v.c. <u>differenza tra medie</u>, al livello di significatività 1 - α

Varianze note	Varianze non note ma uguali
$IC = \left(\bar{x}_1 - \bar{x}_2\right) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	campioni piccoli $\tilde{s}^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$ $IC = \left(\bar{x}_1 - \bar{x}_2\right) \pm t_{\alpha/2;n_1+n_2-2}\,\tilde{s}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ campioni grandi $IC = \left(\bar{x}_1 - \bar{x}_2\right) \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$