Esercitazione 7 del corso di Statistica 2

Prof. Domenico Vistocco

Dott.ssa Paola Costantini

9 Giugno 2008

Esercizio 1

La distribuzione dei pesi dei pesi pacchetti per confezionare per confezionare le caramelle, in grammi, prodotti da un'azienda, ha una distribuzione Normale con scarto quadratico medio pari a 7. Per stimare il peso medio si estrae un campione di 10 pacchetti ottenendo i pesi seguenti:

170, 180, 172, 171, 183, 181, 175, 178, 185, 184

Si vuole costruire un intervallo di confidenza al 90%.
La stima puntuale della media è $\mu = 177{,}9$.

Soluzione

Per costruire un intervallo di confidenza al livello 1-α è necessario determinare quel valore $z_{\alpha/2}$ tale che la probabilità che z assuma valore nell'intervallo $(-z_{\alpha/2}; z_{\alpha/2})$ sia uguale a 1-α, per cui avremo $P(-z_{\alpha/2} \leq z \leq z_{\alpha/2}) = 0{,}90$.
La funzione di ripartizione di una v.c. Normale standardizzata in $z_{\alpha/2}$ vale $1-\alpha/2$, cioè $\phi(z_{\alpha/2}) = 1-\alpha/2$.
Nel nostro caso $\alpha = 0{,}10$, per cui $\alpha/2 = 0{,}05$.
Il valore $z_{\alpha/2}$ è tale che $\phi(0{,}05) = 1-0{,}05 = 0{,}95$ e dalle tavole risulta che $z_{0{,}05} = 1{,}645$. Di conseguenza lo stimatore per l'intervallo è

$$stimatore\left\{\begin{array}{l} L_1 = \overline{X} - 1{,}645 \cdot 7/\sqrt{10} \\ L_2 = \overline{X} + 1{,}645 \cdot 7/\sqrt{10} \end{array}\right\}$$

Sostituendo il valore osservato della media campionaria $\overline{x} = 177{,}9$ si ottiene la stima dell'intervallo. Gli estremi sono:

$$\left\{\begin{array}{l} l_1 = 177{,}9 - 1{,}645 \cdot 7/\sqrt{10} = 174{,}26 \\ l_2 = 177{,}9 + 1{,}645 \cdot 7/\sqrt{10} = 181{,}54 \end{array}\right.$$

Procedendo in modo analogo è possibile ottenere l'intervallo di confidenza al 95%.
In questo caso $P(-z_{\alpha/2} \leq z \leq z_{\alpha/2}) = 0{,}95$, con $\alpha = 0{,}05$ e $\alpha/2 = 0{,}025$.

$\phi(z_{\alpha/2})$= 1-α/2, cioè ϕ(0,025)= 1-0,025=0,975 e dalle tavole risulta che $z_{0,025}$ = 1,96.

Gli estremi del nostro intervallo saranno:

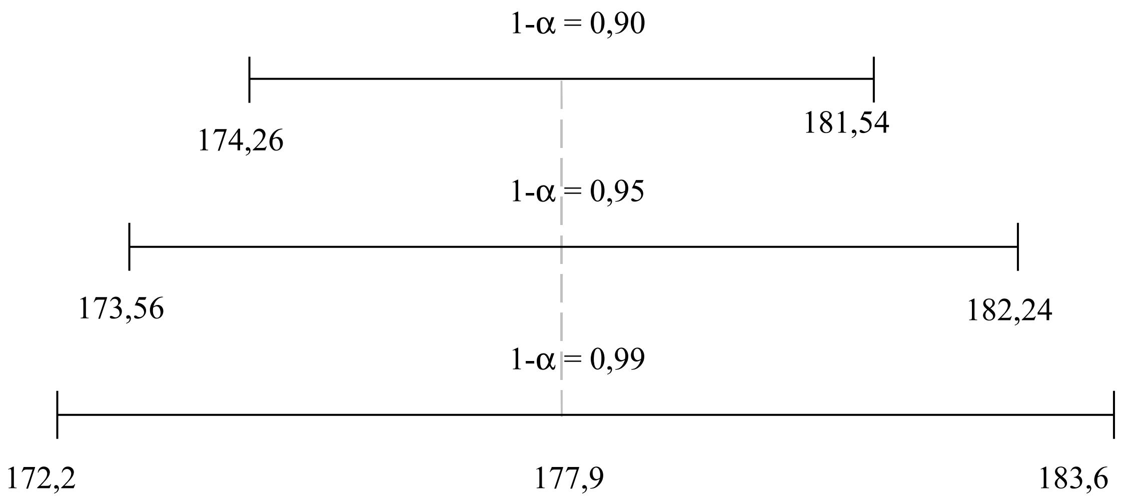
$$\begin{cases} l_1 = 177{,}9 - 1{,}96 \cdot 7/\sqrt{10} = 173{,}56 \\ l_2 = 177{,}9 + 1{,}96 \cdot 7/\sqrt{10} = 182{,}24 \end{cases}$$

Infine determiniamo l'intervallo di confidenza al 99%.

In questo caso P($-z_{\alpha/2} \leq z \leq z_{\alpha/2}$) =0,99, con α= 0,01 e α/2= 0,005.
$\phi(z_{\alpha/2})$= 1-α/2, cioè ϕ(0,005)= 1-0,005=0,995 e dalle tavole risulta che $z_{0,005}$ = 2,575.

Gli estremi del nostro intervallo saranno:

$$\begin{cases} l_1 = 177{,}9 - 2{,}575 \cdot 7/\sqrt{10} = 172{,}2 \\ l_2 = 177{,}9 + 2{,}575 \cdot 7/\sqrt{10} = 183{,}6 \end{cases}$$



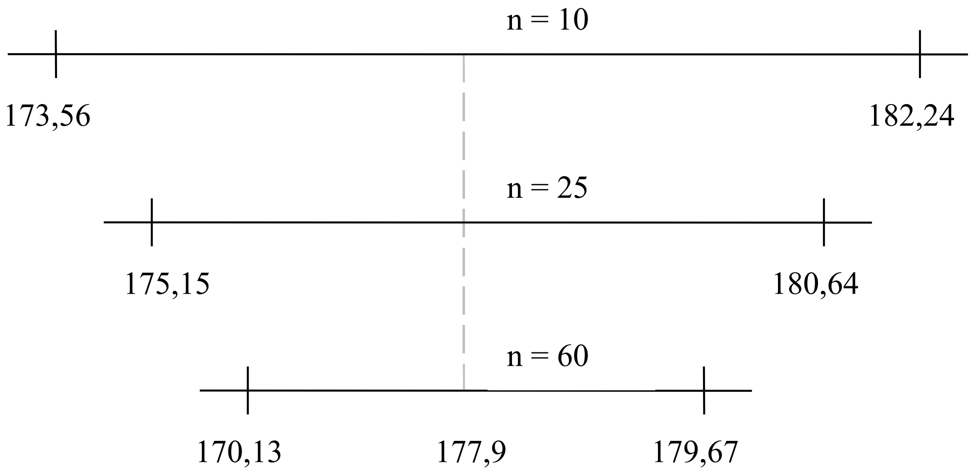
All'aumentare del livello di confidenza aumenta la lunghezza degli intervalli.

Ora supponiamo che dalla stessa popolazione di pesi di sacchetti per confezionare caramelle, si estraggono campioni di numerosità diversa, ad esempio: n=10, n=25, n=60. Assumiamo per semplicità che la stima puntuale della media sia sempre la stessa, $\bar{x} = 177{,}9$ per i diversi valori di n. Per n=10 sappiamo che l'intervallo di confidenza al 95% è (173,56;182,24); per n=25 avremo:

$$\begin{cases} l_1 = 177{,}9 - 1{,}96 \cdot 7/\sqrt{25} = 175{,}156 \\ l_2 = 177{,}9 + 1{,}96 \cdot 7/\sqrt{25} = 180{,}64 \end{cases}$$

Per n = 60 avremo gli estremi:

$$\begin{cases} l_1 = 177{,}9 - 1{,}96 \cdot 7 / \sqrt{60} = 176{,}13 \\ l_2 = 177{,}9 + 1{,}96 \cdot 7 / \sqrt{60} = 179{,}67 \end{cases}$$



All'aumentare della numerosità campionaria, a parità di livello di confidenza, si riduce la lunghezza degli intervalli in quanto vi è un minore grado di incertezza. Aumentando n si raccoglie una maggiore quantità di informazioni e ciò consente una stima più precisa.

Esercizio n. 2

In una città sono svolte periodicamente indagini per verificare qual è la spesa media mensile delle famiglie in generi alimentari. In un'indagine di qualche anno fa lo scarto quadratico medio della spesa alimentare è risultato pari a 120 €. Ora si intende ripetere l'indagine e si vuole determinare n affinché la media campionaria non disti dalla media effettiva di oltre 20 € con probabilità 0,95.

Soluzione

Si può richiedere che lo stimatore $\bar{x}$ non disti dal parametro μ di un ammontare superiore a $\varepsilon>0$. La quantità ε è un errore ritenuto tollerabile per gli scopi dell'indagine e varia in funzione del contesto in cui si opera.

La probabilità con la quale si vuole raggiungere l'obiettivo desiderato è $1 - \alpha = 0{,}95$, per cui avremo che $\alpha = 0{,}05$ e $\alpha/2 = 0{,}025$. Per determinare n dobbiamo individuare il valore $z_{0,025}$, tale che la funzione di ripartizione di una v.c. Normale standardizzata in $z_{0,025}$ sia pari a $\phi(0{,}025)= 1-0{,}025=0{,}975$ e dalle tavole risulta che $z_{0,025} = 1{,}96$.
Calcoliamo n applicando la formula:

$$n = \frac{z_{\alpha/2}^2 \cdot \sigma^2}{\varepsilon^2}$$

Sapendo che ε è la quantità di errore ritenuta tollerabile (20 € nel nostro caso) e utilizzando 120^2 come approssimazione di σ^2 si ottiene la numerosità necessaria per effettuare l'indagine secondo i vincoli previsti

$$n = \frac{1{,}96^2 \cdot 120^2}{20^2} = 138{,}3 \cong 138$$

Per ogni numerosità campionaria non inferiore a 138, la probabilità che la media campionaria non disti dalla media effettiva di oltre 20 euro è almeno 0,95.

Esercizio n. 3

Si vuole stimare il numero medio di km per litro di carburante percorsi da un particolare fuori strada su percorso misto. In un'indagine precedente lo σ è risultato pari a 23 km. Determinare la numerosità campionaria affinché la media campionaria non disti dalla media della popolazione di oltre 0,25 km con probabilità 0,90.

Soluzione

$\alpha = 0{,}10$ e $\alpha/2 = 0{,}05$
$\phi(0{,}05) = 1-0{,}05=0{,}95$ e dalle tavole risulta che $z_{0,05} = 1{,}645$.

$$n = \frac{1{,}645^2 \cdot 2{,}3^2}{0{,}25^2} = 228{,}6 \cong 229$$

Per ogni numerosità campionaria non inferiore a 229, la probabilità che la media campionaria non disti dalla media effettiva di oltre 0,25 km è almeno 0,90.

Esercizio n 4

Volendo stimare la media di una popolazione distribuita in modo normale e con varianza non nota, si estrae da essa un campione di prefissata numerosità e di esso si calcola la media e la varianza campionaria corretta.

Supponiamo che la perdita di peso di n =16 pezzi di metallo, dopo un certo intervallo di tempo sia di 3,42 grammi, con una varianza pari a 0,4624. Costruire un intervallo di confidenza al 99% per la media della perdita di peso di metallo.

Soluzione

Il problema posto consiste nella stima per intervalli della media della popolazione di cui non si conoscer la varianza, sulla base di un campione di piccole dimensioni (n=16).
Lo stimatore per intervallo della media di una popolazione normale, con varianza incognita, al livello di confidenza 1 - α ha estremi:

$$\left\{\begin{array}{l} L_1 = \overline{X} - t_{n-1;\alpha/2} \cdot \hat{\sigma} / \sqrt{n} \\ L_2 = \overline{X} + t_{n-1;\alpha/2} \cdot \hat{\sigma} / \sqrt{n} \end{array}\right\} \text{STIMATORE}$$

Poiché la numerosità campionaria è n=16, la media campionaria studentizzata ha una distribuzione *t* di Student con 15 gradi di libertà. Per costruire l'intervallo di confidenza è necessario determinare il valore $t_{15,0,005}$ tale che la v.c. t_{15} assuma valori maggiori con probabilità α/2= 0,005. Dalle tavole risulta $t_{15,0,005}$ = 2,947, pertanto la stima per intervallo sarà:

$$\left\{\begin{array}{l} l_1 = 3{,}42 - 2{,}947 \cdot 0{,}68 / \sqrt{16} = 2{,}91 \\ l_2 = 3{,}42 + 2{,}947 \cdot 0{,}68 / \sqrt{16} = 3{,}92 \end{array}\right.$$

Il tempo medio per la perdita di peso del metallo è compreso tra 2,91e 3,92 grammi, al livello di confidenza del 99%.

Esercizio n 5

Una società telefonica vuole stimare il tempo medio che intercorre fra il momento nel quale sono segnalati i guasti e quello in cui avviene la riparazione. Si assume che itempi si distribuiscono in modo normale. In un campione casuale di 16 richieste di assistenza, la media è risultata $\bar{x} = 47$ e lo scarto quadratico medio è risultato $\hat{\sigma} = 12$. Si vuole costruire un intervallo di confidenza al 95%.

Soluzione

Lo stimatore per intervallo della media di una popolazione normale, con varianza incognita, al livello di confidenza 1 - α ha estremi:

$$\left\{\begin{array}{l} L_1 = \overline{X} - t_{n-1;\alpha/2} \cdot \hat{\sigma} / \sqrt{n} \\ L_2 = \overline{X} + t_{n-1;\alpha/2} \cdot \hat{\sigma} / \sqrt{n} \end{array}\right\} \text{STIMATORE}$$

Poiché la numerosità campionaria è n=16, la media campionaria studentizzata ha una distribuzione *t* di Student con 15 gradi di libertà. Per costruire l'intervallo di confidenza è necessario determinare il valore $t_{15,0,025}$ tale che la v.c. t_{15} assuma valori maggiori con probabilità $\alpha/2$= 0,025. Dalle tavole risulta $t_{15,0,025}$ = 2,131, pertanto la stima per intervallo sarà:

$$\begin{cases} l_1 = 47 - 2{,}131 \cdot 12/\sqrt{16} = 40{,}61 \\ l_2 = 47 + 2{,}131 \cdot 12/\sqrt{16} = 53{,}39 \end{cases}$$

Il tempo medio per le riparazioni è compreso tra (40,61; 53,39) minuti, al livello di confidenza del 95%.