Esercitazione 6 del corso di Statistica (parte 2)

Dott.ssa Paola Costantini

20 Maggio 2009

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 μ = 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 α= 0,10, per cui $\alpha/2$= 0,05.
Il valore $z_{\alpha/2}$ è tale che ϕ(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

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

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

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

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 α= 0,05 e $\alpha/2$= 0,025.
$\phi(z_{\alpha/2})$= 1-$\alpha/2$, 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:

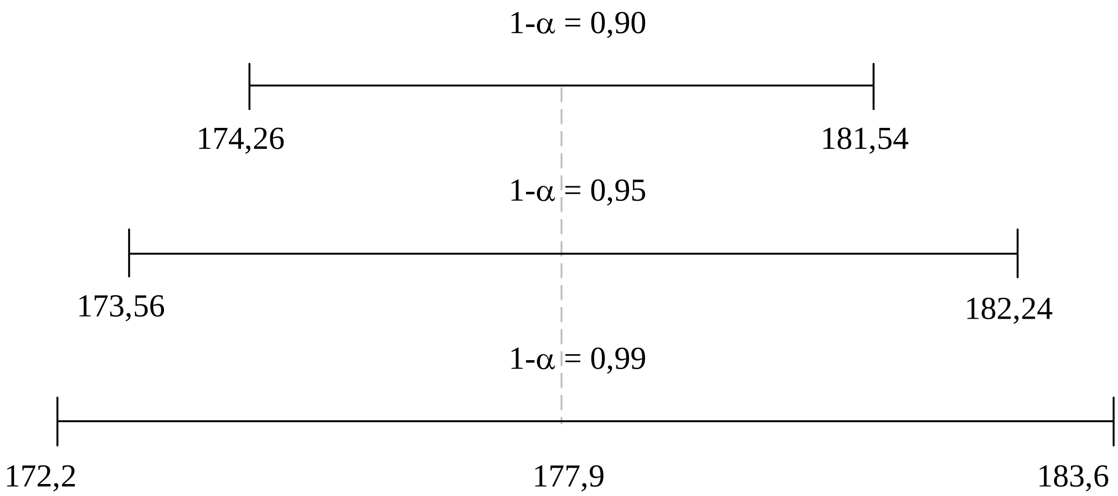
$$\begin{cases} l_1 = 177{,}9 - 1{,}96 \cdot 7/\sqrt{10} = 173{,}56 \\ l_2 = 177{,}9 + 1{,}96 \cdot 7/\sqrt{10} = 182{,}24 \end{cases}$$

Infine determiniamo l'intervallo di confidenza al 99%.

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

Gli estremi del nostro intervallo saranno:

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



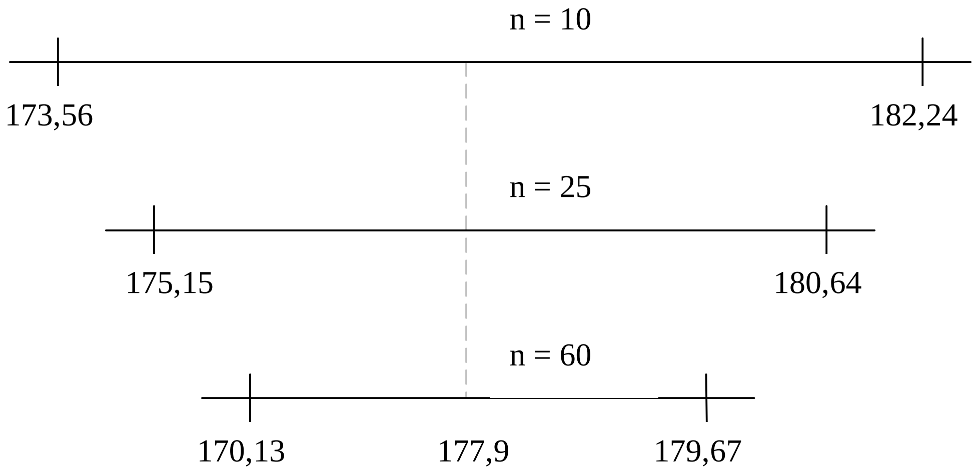
All'aumentare del livello di confidenza aumenta la lunghezza degli intervalli.

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

Si consideri il seguente campione casuale semplice estratto da una popolazione normale di media μ e varianza 16:

33.2, 27.8, 30.8, 39.6, 36.2, 43.4, 36.6, 28.8.

Calcolare l'intervallo di confidenza per μ al livello di confidenza $1 - \alpha = 0.99$.

SVOLGIMENTO

Si richiede di costruire l'intervallo di confidenza per la media avendo osservato un carattere X la cui distribuzione è normale e di cui è nota anche la varianza.
Per costruire l'intervallo di confidenza per μ bisogna considerare la v.c. media campionaria che ha una distribuzione normale con media μ e varianza σ^2/n. La v.c. standardizzata $Z = \left(\bar{X} - \mu\right) / \left(\sigma/\sqrt{n}\right)$ ha una distribuzione normale con media 0 e varianza 1.
E' possibile costruire un intervallo simmetrico intorno a Z individuando due valori $-z_{\alpha/2}$ e $+z_{\alpha/2}$ tali che $P\left(-z_{\alpha/2} \le Z \le +z_{\alpha/2}\right) = 1 - \alpha$. Tale relazione equivale a:

$$P\left(-z_{\alpha/2} \le \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \le +z_{\alpha/2}\right) = P\left(\bar{X} - z_{\alpha/2}\,\sigma/\sqrt{n} \le \mu \le \bar{X} + z_{\alpha/2}\,\sigma/\sqrt{n}\right) = 1 - \alpha$$

Dai dati a disposizione risulta:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{8}(33.2+27.8+30.8+39.6+36.2+43.4+36.6+28.8) = 34.55$$

Poichè $1 - \alpha = 0.99$ si avrà $\alpha/2 = 0.005$ e $|z_{\alpha/2}| = 2.575$. La probabilità di cui sopra può essere riscritta nel seguente modo:

$$P\left(-2.575 \leq \frac{34.55-\mu}{4/\sqrt{8}} \leq +2.575\right) = 0.99.$$

L'intervallo di confidenza per μ al 99% risulta pari a:

$$\left[34.55 - 2.575 \cdot 4/\sqrt{8}, 34.55 + 2.575 \cdot 4/\sqrt{8}\right] = [30.91, 38.19]$$

Esercizio n 3

E' stato scoperto che l'aggiunta di un determinato composto chimico ad un certo fertilizzante permette di aumentare la quantità di fragole ottenibile da ogni singola unità di fertilizzante. Un esperimento viene condotto, introducendo il composto e misurando poi la quantità di fragole prodotta da 5 unità di fertilizzante. I risultati (in grammi di fragole per unità di fertilizzante) sono i seguenti:

229, 255, 280, 203, 229.

Stimare un intervallo per la media della quantità di fragole ottenuta dai 5 fertilizzanti utilizzati, ad un livello di confidenza del 99%.

SVOLGIMENTO

Si richiede di costruire l'intervallo di confidenza per la media avendo osservato un carattere X di cui non è nota la varianza.
Come regola pratica, se la dimensione del campione è minore di 30 non si può invocare il teorema del limite centrale e sfruttare l'approssimazione alla distribuzione normale, ma occorre utilizzare la distribuzione campionaria della media facendo riferimento alla v.c. t di Student con *n-1* gradi di libertà.
In questo caso occorre verificare che la distribuzione della popolazione sia normale. Se questa ipotesi non è soddisfatta il test può non essere significativo.
Nel caso in oggetto per costruire l'intervallo di confidenza per la media bisogna comunque ipotizzare che, per la popolazione, il carattere osservato X abbia una distribuzione normale.
Un intervallo di confidenza è ricavato dalla seguente relazione:

$$P\left(-t_{\alpha/2} \leq \frac{\overline{X}-\mu}{s/\sqrt{n}} \leq +t_{\alpha/2}\right) = P\left(\overline{X} - t_{\alpha/2}\, s/\sqrt{n} \leq \mu \leq \overline{X} + t_{\alpha/2}\, s/\sqrt{n}\right) = 1-\alpha$$

in cui il valore $|t_{\alpha/2}|$ è il quantile di una distribuzione t di Student con *n-1* gradi di libertà.
Dai dati a disposizione risulta:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{5}\left(229 + 255 + 280 + 203 + 229\right) = 239.2$$

$$s = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\overline{x}\right)^2} = \sqrt{\frac{1}{5}\left(229^2 + 255^2 + 280^2 + 203^2 + 229^2\right) - \left(239.2\right)^2} = 29.3$$

Poiché 1 − α = 0.99 si avrà $\alpha/2 = 0.005$ e $\left|t_{\alpha/2}\right| = 4.604$. La probabilità di cui sopra può essere riscritta nel seguente modo:

$$P\left(-4.604 \leq \frac{239.2 - \mu}{29.3/\sqrt{5}} \leq +4.604\right) = 0.99\,.$$

L'intervallo di confidenza per μ al 99%, assumendo che X abbia una distribuzione normale, risulta pari a:

$$\left[239.2 - 4.604 \cdot 29.2/\sqrt{5}\,, 239.2 + 4.604 \cdot 29.2/\sqrt{5}\right] = \left[179.08, 299.32\right]$$