Probabilità e Statistica
Esercitazioni

a.a. 2006/2007

C.d.L.: Ingegneria per l'Ambiente ed il Territorio, Ingegneria Civile, Ingegneria Gestionale, Ingegneria dell'Informazione

C.d.L.S.: Ingegneria Civile

Intervalli di confidenza

Ines Campa e Marco Longhi

Esercizi

Esercizio 1. Un laboratorio analizza una certa quantità di un prodotto farmaceutico per determinare la concentrazione di principio attivo in esso presente. Tali analisi non sono perfettamente precise; se vengono ripetute per altre quantità estratte dal medesimo prodotto i risultati seguono una distribuzione normale con media μ, concentrazione del principio attivo nel prodotto, incognita e deviazione standard, caratteristica della procedura analitica usata, nota e pari a $\sigma = 0.19\frac{g}{l}$. Il laboratorio analizza 4 quantità estratte dal prodotto ottenendo i seguenti risultati: $2.066;\ 2.187;\ 1.893;\ 2.009$ in $\frac{g}{l}$. La casa farmaceutica è interessata ad un intervallo di confidenza per la concentrazione di principio attivo presente nel prodotto al 90%, determinarlo. Calcolare, inoltre, l'intervallo di confidenza per la concentrazione di principio attivo presente nel prodotto al 99%.

Risoluzione. Indichiamo con X la variabile casuale in esame. Per ipotesi si ha: $X \tilde{} N\left(\mu, 0.19^2\right)$. Per determinare **l'intervallo di confidenza per la media di una popolazione normale con varianza nota**, si ricorre alla variabile

$$Z_n = \frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$$

che è distribuita come una normale di media 0 e varianza 1.

Fissato il livello di confidenza $1-\alpha$ e indicato con $z_{\frac{\alpha}{2}}$ il percentile tale che

$$P\left[Z_n \geq z_{\frac{\alpha}{2}}\right] = \frac{\alpha}{2},$$

dall'intervallo di probabilità per la variabile casuale con distribuzione normale standardizzata si ha:

$$1-\alpha = P\left[|Z_n| < z_{\frac{\alpha}{2}}\right] = P\left[\left|\frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}\right| < z_{\frac{\alpha}{2}}\right].$$

Effettuando alcuni passaggi algebrici e sostituendo i valori campionari osservati si ottiene

$$\overline{x}_n - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} < \mu < \overline{x}_n + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}.$$

Essendo $n = 4$, $\overline{x}_n = \overline{x}_4 = \frac{2.066+2.187+1.893+2.009}{4} = 2.03875 \approx 2.039$ e

$$1 - \alpha = 0.90 \Longrightarrow \alpha = 0.10 \Longrightarrow \frac{\alpha}{2} = 0.05 \Longrightarrow z_{0.05} = 1.6449 \approx 1.645.$$

L'intervallo di confidenza per il parametro μ risulta pari a

$$2.039 - 1.645\frac{0,19}{2} < \mu < 2.039 + 1.645\frac{0,19}{2} \Longrightarrow 1.883 < \mu < 2.195.$$

Se il livello di confidenza è al 99%, l'intervallo di confidenza per il parametro μ risulta:

$$\overline{x}_4 - z_{0.005} \cdot \frac{\sigma}{\sqrt{4}} < \mu < \overline{x}_4 + z_{0.005} \cdot \frac{\sigma}{\sqrt{4}}.$$

con $P\left[Z_4 \leq z_{0.005}\right] = 0.995 \Longrightarrow z_{0.005} = 2.5758 \approx 2.576$. Quindi

$$I = (1.794,\ 2.284)\,,\quad l\,(I) = 0.49$$

Esercizio 2. Un segnale radio viene emesso con frequenza distribuita normalmente e con valore atteso μ e deviazione standard $30\ kHz$. Supponendo di osservare la seguente serie di frequenze in kHz:

610	601	578	615	640	630	618	602	613	610	625	585	622	608	59

determinare una stima di μ e la probabilità che la frequenza stia nell'intervallo $[590\ kHz,\ 610\ kHz]$. Determinare poi un intervallo di confidenza per μ al 95 per cento

Risoluzione. Sia X la v.c. che denota la frequenza del segnale radio emesso, si ha $X \tilde{} N(\mu, 900)$. Lo stimatore per μ è $\overline{X}_{15}$, la stima di μ richiesta è

$$\overline{x}_{15} = \frac{9154}{15} \approx 610.27.$$

Determiniamo ora la $P[590 \leq X \leq 610] =$

$$= P\left[\frac{590 - 610.27}{30} \leq Z \leq \frac{610 - 610.27}{30}\right] \approx P[-0,68 \leq Z \leq -0,01].$$

Ne segue che

$$P[590 \leq X \leq 610] = 0.75175 - 0.50399 = 0.24776.$$

Per determinare l'intervallo di confidenza per la media, essendo la distribuzione normale di varianza nota, si ricorre alla variabile

$$Z_{15} = \frac{\overline{X}_{15} - \mu}{\frac{\sigma}{\sqrt{15}}}$$

L'intervallo di confidenza bilaterale al livello $0,95 = 1-\alpha$ per μ è

$$I = \left(\overline{x}_{15} - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}},\ \overline{x}_{15} + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right)$$

con $z_{\frac{\alpha}{2}} = z_{0.025}$, tale che $P\left[Z_{15} \leq z_{0.025}\right] = 0.975$. Consultando le tavole della normale, troviamo che $z_{0.025} = 1.96$, quindi sostituendo in I: $\sigma = 30$, $n = 15$, $\overline{x}_{15} = 610.27$ risulta

$$I = (595.09,\ 625.45)$$

Supponiamo invece di voler ricavare per μ **un intervallo di confidenza unilaterale** al 95 per cento:

Intervallo destro

Posto

$$\alpha = P\left[\frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} > z_\alpha \right] = P\left[\mu < \overline{X}_n - z_\alpha \cdot \frac{\sigma}{\sqrt{n}} \right],$$

si ha

$1 - \alpha = P\left[\mu > \overline{X}_n - z_\alpha \cdot \frac{\sigma}{\sqrt{n}}\right]$, con

$$0.95 = 1 - \alpha = P\left[Z_{15} \le z_\alpha\right] \Longrightarrow z_\alpha = z_{0.05} = 1.6449 \approx 1.65.$$

Sostituendo

$$\left(\overline{x}_{15} - z_{0.05} \cdot \frac{\sigma}{\sqrt{15}}, +\infty\right) = (597.49, +\infty).$$

Intervallo sinistro

$$\left(-\infty,\ \overline{x}_{15} + z_{0.05} \cdot \frac{\sigma}{\sqrt{15}}\right) = (-\infty,\ 623.05).$$

Esercizio 3. Un laboratorio farmaceutico deve calcolare la concentrazione μ di principio attivo in un dato composto chimico. I risultati dell'analisi non sono certi, ma ripetuti possono mostrare che seguono una distribuzione normale. Dato un campione di ampiezza 3: $3.853;\ 3.588;\ 3.954$ in $\frac{g}{l}$, determinare un intervallo di confidenza per la concentrazione μ di principio attivo al 90%.

Risoluzione. Indichiamo con X la variabile casuale in esame. Per ipotesi si ha: $X\tilde{\ }N\left(\mu,\sigma^2\right)$. Per determinare **l'intervallo di confidenza per la media di una popolazione normale con varianza incognita**, si ricorre alla variabile

$$T_{n-1}=\frac{\overline{X}_n-\mu}{\frac{S}{\sqrt{n}}}$$

distribuita come una t_g di Student con $g=n-1$ gradi di libertà. Fissato il livello di confidenza $1-\alpha$ e indicato con $t_{\frac{\alpha}{2};n-1}$ il percentile tale che $F\left(t_{\frac{\alpha}{2};n-1}\right)=1-\frac{\alpha}{2}$, dall'intervallo di probabilità per la variabile casuale t_{n-1} si ha:

$$1-\alpha=P\left[\left|T_{n-1}\right|<t_{\frac{\alpha}{2};n-1}\right]=P\left[\left|\frac{\overline{X}_n-\mu}{\frac{S}{\sqrt{n}}}\right|<t_{\frac{\alpha}{2};n-1}\right].$$

Effettuando alcuni passaggi algebrici all'interno delle parentesi e sostituendo i valori campionari osservati, si ottiene

$$\overline{x}_n - t_{\frac{\alpha}{2};n-1} \cdot \frac{s}{\sqrt{n}} < \mu < \overline{x}_n + t_{\frac{\alpha}{2};n-1} \cdot \frac{s}{\sqrt{n}}.$$

Ricordiamo che

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x}_n)^2 = \frac{1}{n-1} \left[\sum_{i=1}^{n} x_i^2 - n\overline{x}_n^2 \right]$$

e sostituendo i valori: $n = 3$, $\overline{x}_n = \overline{x}_3 = \frac{3.853+3.588+3.954}{3} \approx 3.798$,

$$s^2 = \frac{1}{2} \left[(3.853)^2 + (3.588)^2 + (3.954)^2 - 3(3.798)^2 \right] \approx (0.199)^2,$$

$$1 - \alpha = 0.90 \Longrightarrow \alpha = 0.10 \Longrightarrow \frac{\alpha}{2} = 0.05 \Longrightarrow t_{0.05;2} = 2.91999 \approx 2.920.$$

L'intervallo di confidenza per il parametro μ risulta pari a

$$3.798 - 2.920 \frac{0.199}{\sqrt{3}} < \mu < 3.798 + 2.920 \frac{0,.199}{\sqrt{3}} \Longrightarrow 3.463 < \mu < 4.133.$$

Supponiamo di dover determinare un intervallo di confidenza unilaterale per la μ

Intervallo sinistro

$$\left(-\infty,\ \overline{x}_n + t_{\alpha;n-1} \cdot \frac{s}{\sqrt{n}}\right)$$

con $1-\alpha = P\left[T_{n-1} > -t_{\alpha;n-1}\right] = P\left[T_{n-1} < t_{\alpha;n-1}\right]$. Sostituendo $n=3$, $1-\alpha = 0.90$ e $\alpha = 0.10$ risulta

$$P\left[T_2 < t_{0.10;2}\right] = 0.90 \Longrightarrow t_{0.10;2} = 1.88562 \approx 1.886 \Longrightarrow I = (-\infty,\ 4.004)\,.$$

Intervallo destro

$$\left(\overline{x}_n - t_{\alpha;n-1} \cdot \frac{s}{\sqrt{n}},\ +\infty\right)$$

con $1-\alpha = P\left[T_{n-1} < t_{\alpha;n-1}\right]$. Sostituendo $n=3$ e $\alpha = 0.10$ risulta

$$t_{0.10;2} \approx 1.886 \Longrightarrow I = (3.592,\ +\infty)\,.$$

Esercizio 4. Per una industria di trafilati in alluminio è essenziale, per la qualità del prodotto, che la variabilità dello spessore sia molto bassa. Una nuova apparecchiatura promette una riduzione di tale variabilità; questa viene sperimentata tramite la produzione di un trafilato di spessore $\mu = 3\ mm$. Dato il campione $2.88;\ 2.93;\ 2.98$ in mm, sapendo che la distribuzione dello spessore è normale, determinare l'intervallo di confidenza per la varianza dello spessore del trafilato al 95%.

Risoluzione. Indichiamo con X la variabile casuale in esame. Per ipotesi si ha: $X \tilde{} N\left(3, \sigma^2\right)$. Per determinare **l'intervallo di confidenza per la varianza di una popolazione normale con media nota**, si ricorre alla variabile

$$\sum_{i=1}^{n} \frac{\left(X_i - \mu\right)^2}{\sigma^2}$$

che è distribuita come una variabile casuale Chi-quadro con n gradi di libertà: χ_n^2. Fissato il livello di confidenza $1-\alpha$ e indicato con $\chi_{\alpha;n}^2$ il percentile tale che

$$P\left[\chi \geq \chi_{\alpha;n}^2\right] = \alpha$$

–

Si ha che $a = 1 - \alpha + \frac{\alpha}{2} = 1 - \frac{\alpha}{2}$

–

$$1-\alpha = P\left[\chi^2_{1-\frac{\alpha}{2};n} < \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} < \chi^2_{\frac{\alpha}{2};n}\right]$$

$$= P\left[\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2};n}} < \sigma^2 < \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2};n}}\right]$$

Sostituendo i valori: $n = 3$,

$\sum_{i=1}^{n}(x_i-\mu)^2 = \sum_{i=1}^{3}(x_i-3)^2 = (2.88-3)^2+(2.93-3)^2+(2.98-3)^2 = 0.0197$,

inoltre $1-\alpha = 0.95 \Longrightarrow \alpha = 0.05 \Longrightarrow \frac{\alpha}{2} = 0.025$,

$$P\left[\chi \leq \chi^2_{1-\frac{\alpha}{2};n}\right] = P\left[\chi \leq \chi^2_{0.975;3}\right] = 0.025 \Longrightarrow \chi^2_{0.975;3} = 0.21579 \approx 0.2158$$

$$P\left[\chi \leq \chi^2_{\frac{\alpha}{2};n}\right] = P\left[\chi \leq \chi^2_{0.025;3}\right] = 0.975 \Longrightarrow \chi^2_{0.025;3} = 9.34840 \approx 9.3484$$

Ne segue che l'intervallo di confidenza per la varianza al 95% è

$$\frac{0.0197}{9.3484} < \sigma^2 < \frac{0.0197}{0.2158} \Longrightarrow 0.0021 < \sigma^2 < 0.0913$$

Esercizio 5. Le marmitte catalitiche devono essere sottoposte ad un test per verificare se i livelli di certe sostanze tossiche siano entro limiti precisi. Un campione casuale di ampiezza 3 viene estratto dalla produzione settimanale di una ditta produttrice di marmitte catalitiche. Una prova su strada rileva che i valori per una particolare sostanza nociva prodotti da ciascuna marmitta catalitica sono $885,\ 889,\ 893$, dove l'unità di misura è milligrammi al chilometro, Sapendo che l'emissione di tale sostanza tossica ha distribuzione normale, si determini un intervallo di confidenza per la varianza al 99%.

Risoluzione. Indichiamo con X la variabile casuale in esame. Per ipotesi si ha: $X\tilde{\ } N\left(\mu, \sigma^2\right)$. Per determinare **l'intervallo di confidenza per la varianza di una popolazione normale con media incognita**, si ricorre alla variabile

$$\frac{(n-1)\,S^2}{\sigma^2} = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2}{\sigma^2}$$

che è distribuita come una variabile casuale Chi-quadro con $n-1$ gradi di libertà: χ^2_{n-1}. Fissato il livello di confidenza $1-\alpha$ e indicato con $\chi^2_{\alpha;n-1}$ il percentile tale che

$$P\left[\chi \geq \chi^2_{\alpha;n-1}\right] = \alpha,$$

si ha $P\left[\chi \leq \chi^2_{\frac{\alpha}{2};n-1}\right] = 1-\frac{\alpha}{2}$.

Si ha che $a = 1-\alpha+\frac{\alpha}{2} = 1-\frac{\alpha}{2}$

–

$$1-\alpha = P\left[\chi^2_{1-\frac{\alpha}{2};n-1} < \frac{(n-1)\,S^2}{\sigma^2} < \chi^2_{\frac{\alpha}{2};n-1}\right]$$

$$= P\left[\frac{\sum_{i=1}^{n}\left(x_i-\overline{X}_n\right)^2}{\chi^2_{\frac{\alpha}{2};n-1}} < \sigma^2 < \frac{\sum_{i=1}^{n}\left(x_i-\overline{X}_n\right)^2}{\chi^2_{1-\frac{\alpha}{2};n-1}}\right]$$

Sostituendo i valori: $n=3$, $\overline{x}_n=\overline{x}_3=\frac{885+889+893}{3}=889$,

$$s^2=\frac{1}{2}\left[\sum_{i=1}^{3}x_i^2-3\overline{x}_3^2\right]=\frac{1}{2}\left[(885)^2+(889)^2+(893)^2-3(889)^2\right]=16,$$

inoltre $1-\alpha=0.99 \Longrightarrow \alpha=0.01 \Longrightarrow \frac{\alpha}{2}=0.005$,

$$P\left[\chi\le\chi^2_{1-\frac{\alpha}{2};n-1}\right]=1-\frac{\alpha}{2}\Longrightarrow P\left[\chi\le\chi^2_{0.995;3}\right]=0.005$$

Quindi $\chi^2_{0.995;2} = 0.01002 \approx 0.0100$

$$P\left[\chi \leq \chi^2_{\frac{\alpha}{2};n-1}\right] = P\left[\chi \leq \chi^2_{0.005;2}\right] = 0.995 \Longrightarrow \chi^2_{0.005;3} = 10.5965$$

Ne segue che l'intervallo di confidenza per la varianza al 99% è

$$\frac{32}{10.5965} < \sigma^2 < \frac{32}{0.0100} \Longrightarrow 3.0199 < \sigma^2 < 3200,\ l(I) = 3196.9801$$

con approssimazione alla quarta cifra decimale. Si osservi come tale intervallo sia di utilità dubbia, vista la sua ampiezza, tale ampiezza è dovuta principalmente alla limitatezza del numero di osservazioni, $n = 3$ ed alla alta richiesta di confidenza.

Esercizio 6. Un laboratorio analizza una certa quantità di un prodotto farmaceutico per determinare la concentrazione di principio attivo in esso presente. Tali analisi non sono perfettamente precise; se vengono ripetute per altre quantità estratte dal medesimo prodotto i risultati seguono una distribuzione normale con media μ, concentrazione del principio attivo nel prodotto, incognita e deviazione standard, caratteristica della procedura analitica usata, nota e pari a $\sigma = 0,19\frac{g}{l}$. Si supponga che il laboratorio farmaceutico sia interessato ad analisi molto precise e stabilisca che, per l'intervallo di confidenza per la media al 90%, il margine di errore massimo tollerabile è $0,02\ \frac{g}{l}$, ovvero che l'ampiezza dell'intervallo sia minore o uguale a $2 \cdot 0,02\ \frac{g}{l}$. Quale numerosità deve avere il campione per soddisfare tale richiesta?

Risoluzione. Indichiamo con X la variabile casuale in esame. Per ipotesi si ha: $X\tilde{\ }N\left(\mu, 0,19^2\right)$.

$$P\left[\overline{X}_n - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} < \mu < \overline{X}_n + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right] = 0.90.$$

Sostituendo risulta

$$I = \left(\overline{x}_n - 1.645 \cdot \frac{0.19}{\sqrt{n}}, \ \overline{x}_n + 1.645 \cdot \frac{0.19}{\sqrt{n}}\right),$$

e $l(I) = \overline{x}_n + 1.645 \cdot \frac{0.19}{\sqrt{n}} - \overline{x}_n + 1.645 \cdot \frac{0.19}{\sqrt{n}} = 2 \cdot 1.645 \cdot \frac{0.19}{\sqrt{n}}$. Per ipotesi $l(I) \leq 2 \cdot 0.02$, quindi

$$\frac{0.6251}{\sqrt{n}} \leq 0.04 \Longrightarrow n \geq 244.21,$$

ma n numero naturale, per cui il campione deve avere numerosità maggiore o uguale a 245.

Esercizio 7 (Tema d'esame del 12/10/2004).

Vengono effettuate in tempi diversi 20 misurazioni della concentrazione di un dato elemento in un materiale, e si osserva una media di 1.23 unità ed una varianza di 0.4 unità al quadrato. Nell'ipotesi che questa concentrazione abbia un modello statistico normale con parametri sconosciuti, determinare l'intervallo di confidenza per la media al 95%.

Risoluzione. Indichiamo con X la variabile casuale in esame. Per ipotesi si ha: $X\tilde{\ }N\left(\mu,\sigma^2\right)$. Si tratta di determinare l'intervallo di confidenza per la media di una popolazione normale con varianza incognita, si ricorre alla variabile

$$T_{20-1} = \frac{\overline{X}_{20} - \mu}{\frac{S}{\sqrt{20}}}$$

distribuita come una t_{19} di Student con 19 gradi di libertà.

$$P\left[\overline{x}_{20} - t_{\frac{\alpha}{2};19} \cdot \frac{s}{\sqrt{20}} < \mu < \overline{x}_{20} + t_{\frac{\alpha}{2};19} \cdot \frac{s}{\sqrt{20}}\right] = 0.95.$$

$$1 - \alpha = 0.95 \Longrightarrow \alpha = 0.05 \Longrightarrow \frac{\alpha}{2} = 0.025.$$

Ne segue $P\left[T_{19} \leq t_{\frac{\alpha}{2};19}\right] = 1 - \frac{\alpha}{2} = 0.975 \Longrightarrow t_{\frac{\alpha}{2};19} = t_{0.025;19} = 2.093$
Quindi $I = (0.934,\ 1.526)$

Esercizio 8 (Tema d'esame del 13/12/2005).

Si è misurata 15 volte la temperatura di una stanza ottenendo i seguenti valori

temperatura $^\circ$C	13.5	13.7	14	14.1	14.4	14.8
frequenza	2	3	4	3	2	1

Supponendo che la temperatura sia una v.a. normale con varianza $\sigma^2 = 9$, determinare un intervallo di confidenza della media al 95%.

Quante misure occorre effettuare affinché l'intervallo di confidenza della media al 90% abbia lunghezza minore di 1?

[Risposta $I = (12,48; 15,52)$ $n > 98$]

Esercizio 9 (Tema d'esame del 25/07/2006).

Il diametro delle sfere di cuscinetti costruiti in serie da una macchina automatica ha distribuzione normale con media incognita e varianza cm^2 $(0,048)^2$. Determinare un intervallo di confidenza al 98% per il diametro medio della produzione di sfere, sapendo che le misurazioni dei diametri di un campione casuale di 256 sfere di cuscinetti costruiti dalla macchina danno un diametro medio di cm $0,824$ (scrivere gli estremi dell'intervallo con quattro cifre decimali). Inoltre, qual è l'ampiezza minima del campione affinché l'intervallo di confidenza al 90% abbia ampiezza minore o uguale a cm $0,01$?

[Risposta $\begin{array}{l} I = (0,8170; 0,8310) \\ n = 250 \end{array}$]

Esercizio 10 (Tema d'esame del 12/07/2005).

Si ritiene che i tempi di reazione ad un certo stimolo siano distribuiti normalmente con media μ. Di seguito sono riportati i tempi di reazione, espressi in secondi, di un campione casuale

0.2	0.4	0.2	0.6	0.1

Determinare un intervallo di confidenza al 90% per μ.

[Risposta $I = (0,1093; 0,4907)$]

Esercizio 11 (Tema d'esame del 14/09/2004).

Le misure dei diametri di un campione casuale di 200 sferette da cuscinetto prodotte da una macchina in una settimana hanno una media campionaria 0.824 cm ed una deviazione standard campionaria 0.042 cm. Determinare l'intervallo di confidenza per la media della popolazione con livello di confidenza del 95%.

[Risposta $I = (0,818; 0,830)$]