Corso di Statistica per Ingegneria Fisica
Anno accademico 2020/2021

Esercitazione 7 - Intervalli di confidenza per media, varianza e proporzione

Esercizio 1. La resa di un impianto chimico viene sottoposta a uno studio. In base all'esperienza precedente, si sa che la resa giornaliera dell'impianto è una variabile aleatoria gaussiana con media incognita μ e varianza uguale a 2. Inoltre, gli ultimi cinque giorni di operatività hanno dato le rese seguenti (in %):

$$91.60 \qquad 88.75 \qquad 90.80 \qquad 89.95 \qquad 91.30\,.$$

(a) Costruire un intervallo di confidenza bilatero per il parametro μ al 95%. Quanto vale l'ampiezza dell'intervallo di confidenza? *[$IC_\mu(0.95) = (89.240\,,\, 91.720)$; ampiezza $= 2.480$]*

(b) Al livello di confidenza del 95%, qual è l'errore massimo che si commette stimando μ con la media campionaria delle cinque osservazioni precedenti? *[1.240]*

(c) Dal momento che più è ampio l'intervallo di confidenza per μ, maggiore è il costo che deve sostenere l'azienda, il capo reparto dice agli statistici che l'ampiezza dell'intervallo non deve essere superiore a 2.2. A questo scopo, è disposto a fornire loro ulteriori osservazioni, purché però si mantenga invariato il livello di confidenza del 95%. Quante nuove misure deve fornire il capo reparto affinché la sua richiesta venga rispettata? *[Almeno altre 2 misure.]*

Esercizio 2. Il peso in Kg dei sacchetti di farina prodotti dal mulino del Signor Sabbioso può essere descritto con una variabile aleatoria normale di media μ e deviazione standard pari a 0.5. Vengono pesati 100 sacchetti trovando un peso medio campionario $\bar{x}_n = 15.3\,\mathrm{Kg}$.

(a) Si costruisca un intervallo di confidenza bilatero per μ al livello del 90% e si indichi l'errore massimo commesso stimando μ con $\bar{x}_n$. *[$(15.3 \pm 0.08225)\,\mathrm{Kg}$; $0.08225\,\mathrm{Kg}$]*

(b) Se, mantenendo lo stesso livello di confidenza del 90%, al punto (a) avessimo voluto trovare un intervallo di ampiezza inferiore a 0.1 Kg, come avremmo dovuto scegliere la numerosità del campione? *[$n \geq 271$]*

(c) Quanto può essere confidente il Signor Sabbioso che il peso medio dei suoi sacchetti di farina sia maggiore di 15.2 Kg? *[97.725%]*

(d) In tutte le risposte precedenti, è essenziale l'ipotesi di normalità? *[No]*

Esercizio 3. Sia $X_1, \ldots, X_n$ un campione casuale da una popolazione normale di media μ incognita e varianza $\sigma^2 = 16$.

(a) Sapendo che $n = 9$ e che $(2.387, 7.613)$ è un intervallo di confidenza simmetrico per μ, se ne determini il livello di confidenza. *[$\gamma = 95\%$]*

(b) Si determini la minima numerosità n del campione tale che un intervallo di confidenza per μ di livello 0.95 non risulti più lungo di 4. *[16]*

(c) Nelle situazioni concrete, aumentare la numerosità del campione ha un costo; d'altra parte, costa anche avere stime imprecise. Supponiamo che il costo di ogni misurazione sia di 25 €, e che il costo di ogni unità di ampiezza dell'intervallo sia di 5200 €. Si determini n in modo da rendere minimo il costo totale di un intervallo di confidenza per μ di livello 0.95 *[$n = 139$]*.

Esercizio 4. Sulle scatole di puntine da disegno prodotte dalla PUNT è dichiarato che il contenuto della scatola è di 100 puntine per ogni scatola. Un ispettore di qualità, che deve certificare la PUNT secondo lo standard ISO9002, allo scopo di verificare la veridicità dell'affermazione precedente esamina il contenuto di 100 scatole; ne ottiene una media campionaria di 95.3 puntine per scatola. Si supponga che la quantità di puntine nelle scatole segua una densità incognita, di cui si sa solo che la deviazione standard è $\sigma = 18.25$.

(a) Si determini un intervallo di confidenza di livello 95% per il numero medio di puntine in ciascuna scatola. *[95.3 ± 3.577]*

(b) Quante scatole avrebbe dovuto esaminare l'ispettore se avesse voluto un intervallo di confidenza di livello 95% di lunghezza non superiore a 4? *[$n \geq 320$]*

Esercizio 5. Si consideri la densità di probabilità

$$f(t) = \frac{3}{4}\Big[1 - (t + 1 - \theta)^2\Big]1_{(\theta-2,\theta)}(t),$$

dove θ è un numero reale incognito.

(a) Si disegni il grafico di f.

(b) Si calcoli il valore atteso μ e la varianza σ^2 di una variabile aleatoria con densità f. *[$\mu = \theta - 1$, $\sigma^2 = 1/5$]*

(c) Dato un campione $X_1, \dots, X_n$ di variabili i.i.d. con densità f, si costruisca uno stimatore T_n non distorto per θ. *[$T_n = \overline{X}_n + 1$]*

(d) Si calcoli l'errore quadratico medio dello stimatore T_n del punto precedente. *[$\text{mse}(T_n; \theta) = \frac{1}{5n}$]*

(e) Lo stimatore T_n è consistente in media quadratica? *[Sì]*

(f) Dato un campione di numerosità 50 la cui media vale 15.13, si forniscano una stima puntuale per θ e un intervallo di confidenza per μ di livello 90%. *[$\hat{\theta} = 16.13$; $\mu \in (15.13 \pm 0.10)$]*

Esercizio 6. Si considerino due campioni casuali, indipendenti fra di loro, il primo dei quali, $X_1, \dots, X_n$, viene estratto da una $N(\mu, 1)$, mentre il secondo, $Y_1, \dots, Y_m$, viene estratto da una $N(\mu, 3)$. Le due popolazioni hanno quindi lo stesso valore medio μ, che però è incognito. Si consideri come stimatore di μ la seguente variabile aleatoria:

$$T = \frac{\sum_{i=1}^n X_i + \sum_{j=1}^m Y_j}{n + m}.$$

(a) T è uno stimatore corretto di μ? *[Sì]*

(b) Si determini la densità di T. *[$T \sim N\Big(\mu, \dfrac{n + 3m}{(n + m)^2}\Big)$]*

(c) Per ogni $1/2 < \alpha < 1$, si calcoli il livello del seguente intervallo di confidenza per μ:

$$\mu \in \left(t - z_\alpha\sqrt{\frac{n + 3m}{(n + m)^2}}\,,\, t + z_\alpha\sqrt{\frac{n + 3m}{(n + m)^2}}\right),$$

dove t è la realizzazione della variabile aleatoria T sul risultato delle $n + m$ misure. *[$2\alpha - 1$]*

Esercizio 7. Supponiamo che quando un segnale elettrico di valore μ viene trasmesso dalla sorgente A, il ricevente registri un valore $\mu + N$, dove N, che denota il rumore, è distribuito come una normale di media 0 e varianza 4. Immaginiamo che per ridurre l'errore il segnale sia trasmesso 9volte. I valori registrati in ricezione sono stati

$$5 \quad 8.5 \quad 12 \quad 15 \quad 7 \quad 9 \quad 7.5 \quad 6.5 \quad 10.5\,.$$

(a) Calcolare un intervallo di confidenza simmetrico al 99% per μ, e i due intervalli di confidenza unilaterali dello stesso livello. *[*$\bar{x} = 9$*;* $(7.28; 10.72)$*;* $(7.45; +\infty)$*;* $(-\infty; 10.55)$*.]*

Si supponga ora di non conoscere la varianza di N.

(b) Determinare gli intervalli di confidenza bilaterale e unilaterali al livello 99% per μ. *[*$s^2 = 9.5$*;* $(5.55; 12.45)$*;* $(6.02; \infty)$*;* $(-\infty; 11.98)$*.]*

Esercizio 8. Vengono effettuate 20 misurazioni della concentrazione di un certo enzima nel sangue di diversi individui e si osservano una media campionaria $\bar{x} = 1.23$ ed una varianza campionaria $s^2 = 0.4$.

(a) Supponendo che i valori di questa concentrazione seguano una distribuzione normale (con entrambi i parametri incogniti!), qual è un intervallo di confidenza bilatero al livello 95% per la media della concentrazione? *[*$IC_\mu(95\%) = (0.934\,,\, 1.526)$

(b) Quale sarebbe l'intervallo di cui sopra se la concentrazione a cui si è interessati fosse distribuita come una normale di varianza nota $\sigma^2 = 0.4$? *[IC= [0.953;1.507]]*

(c) Quale tra i due intervalli trovati in (a) e in (b) è più ampio? Il risultato ottenuto sembra ragionevole? *[L'intervallo in (a). Sì; però perchè?]*

Esercizio 9. Avendo osservato il campione, proveniente da una legge normale,

$$2.28 \quad 3.45 \quad 5.62 \quad 6.12 \quad 4.26 \quad 6.30 \quad 7.09 \quad 1.97 \quad 5.77 \quad 7.68 \quad 8.16$$

(a) si fornisca una stima puntuale della media e della varianza, utilizzando stimatori non distorti; *[*$\bar{x}_n = 5.336$*;* $s^2 = 4.374$*]*

(b) si calcoli l'intervallo di confidenza per la media al 90%. *[*(5.336 ± 1.143)*]*

Esercizio 10. Il Centro di Ricerche Spaziali (CRS) vuole determinare la distanza d della stella Vega dalla Terra. Le sue misure X_1 sono affette da errore casuale $\epsilon \sim N(0, \sigma^2)$, dove la deviazione standard σ è incognita. Il CRS sulla base di $n = 15$ misure indipendenti stima d con $(25.09\,,\, 25.35)$, intervallo di confidenza bilatero simmetrico di livello 95%.

Quanto valgono le stime puntuali di d e σ fornite dal campione? *[*$\bar{x}_{10} = \frac{25.09+25.35}{2} = 25.22$ *e* $s_{10} = \dfrac{(25.35 - 25.09)\sqrt{15}}{2 \cdot 2.1448} = 0.2347$*]*

Esercizio 11. In uno studio sull'inquinamento atmosferico effettuato da una certa stazione sperimentale, si sono registrati, su 8 diversi campioni di aria, i seguenti valori di una certa sostanza tossica (in microgrammi per metro cubo):

$$2.2\,;\ 1.8\,;\ 3.1\,;\ 2.0\,;\ 2.4\,;\ 2.0\,;\ 2.1\,;\ 1.2$$

(a) Assumendo che la popolazione campionaria sia normale, si stabilisca un intervallo di confidenza per la media, al livello del 95%. *[*$IC_\mu(0.95) = 2.1 \pm 0.45$*]*

(b) L'ipotesi di normalità della popolazione è essenziale per giustificare il procedimento seguito? *[Sì, perchè la numerosità del campione non è elevata, dunque la densità della sua media campionaria non può essere approssimata con una normale.]*

Esercizio 12. In seguito ad un controllo medico in una scuola elementare durante un'epidemia di morbillo, 9 bambini su 75 risultano aver contratto il virus. Detta p la probabilità che un bambino della stessa scuola sia malato, determinare:

(a) una stima puntuale $\hat{p}$ per p; *[$\hat{p} = 0.12$]*

(b) un intervallo di confidenza per p di livello 0.95 simmetrico rispetto a $\hat{p}$; *[$IC_p(95\%) = (0.046, 0.194)$]*

(c) un intervallo di confidenza per p di livello 0.9 del tipo $(0, a)$. *[$IC_p(90\%) = (0\,,\, 0.168)$]*

Esercizio 13. Il Sindaco Will Piedebianco vuole conoscere la proporzione p di abitanti della Contea contrari al consumo di erba-pipa. Fredegario Bolgeri riceve pertanto l'incarico di intervistare un campione casuale di n abitanti per stimare p sulla base della proporzione $\bar{x}_n$ di contrari al consumo di erba-pipa fra gli intervistati.

(a) Si trovi la minima numerosità n_0 del campione grazie alla quale un intervallo di confidenza di livello 90% simmetrico rispetto a $\bar{x}_n$ ha ampiezza sicuramente inferiore a 0.05. *[$n_0 = 1083$]*

Dopo aver chiesto il parere di n_0 abitanti della Contea scelti a caso, Fredegario Bolgeri trova ben 268 contrari al consumo di erba-pipa! Si stimi p mediante:

(b) uno stimatore puntuale non distorto; *[$\hat{p} = 0.2475$]*

(c) un intervallo di confidenza bilatero di livello 90%; *[$IC_p(90\%) = (0.2475 \pm 0.0216)$]*

(d) un intervallo di confidenza di livello 90% di tipo $(0, a)$. *[$IC_p(90\%) = (0\,,\, 0.2642)$]*

Esercizio 14. L'Assessorato all'Ecologia di Pietraforata vuole conoscere la proporzione p di abitanti favorevoli all'iniziativa delle "domeniche ecologiche". Per questo motivo si decide di intervistare un campione di n abitanti, scelti a caso e giudicati rappresentativi della popolazione, per stimare p sulla base della proporzione $\bar{x}_n$ di risposte favorevoli.

(a) Si stabilisca quale deve essere la numerosità minima n_0 del campione affinchè un intervallo di confidenza bilatero per p di livello 95% abbia ampiezza sicuramente inferiore a 0.05. *[$n_0 = 1537$]*

(b) Eseguita l'indagine, 882 degli n_0 abitanti intervistati si sono dichiarati favorevoli alle "domeniche ecologiche"; si stimi p sia con una stima puntuale sia con un intervallo di confidenza bilatero di livello 95%. *[$\hat{p} = 0.5738; IC_p(95\%) = (0.5491\,,\, 0.5986)$]*

(c) Si supponga ora $p = 0.65$. Con quale probabilità un'indagine condotta su n_0 abitanti fornisce un errore di stima puntuale inferiore a 0.05? *[$\mathbb{P}\left(\left|\bar{X}_{n_0} - p\right| < 0.05\right) = 2\Phi(4.11) - 1 = 99.996\%$]*

Esercizio 15. Per una proiezione dei risultati elettorali, sono state esaminate 3000 schede, estratte a caso tra le schede da scrutinare. Il partito A ha ottenuto 600 voti. Determinare un intervallo di confidenza al 95% (proiezione) per la percentuale di voti ottenuti dal partito A. *[$(20 \pm 1.4)\%$]*

Esercizio 16. Una ditta riceve un lotto molto numeroso di lampade. Per giudicarne la qualità, ne esamina 60 scelte a caso, trovandone 8 difettose.

(a) Fornire un intervallo di confidenza bilatero di livello 90% per la proporzione p di lampade difettose nel lotto. *[$IC_p(90\%) = (0.06114\,,\, 0.20552)$]*

(b) Per una successiva indagine, quanto deve valere la numerosità n del campione affinché l'intervallo di confidenza per p (sempre di livello $0,9$) risulti avere ampiezza inferiore a 0.1? *[$n \geq 271$]*

Esercizio 17. La Società degli Acquedotti del comune di Pietraforata vuole determinare il pH dell'acqua che circola nei suoi impianti. Per questo motivo, incarica un chimico di misurare il pH di 10 campioni provenienti da punti diversi della città, e, in base al risultato delle misure, gli chiede di stimare la media e la varianza della variabile aleatoria

$$X = \text{pH dell'acqua in un punto a caso dell'acquedotto}\,.$$

Il chimico trova nei campioni i seguenti valori di pH:

$$7.94 \quad 8.18 \quad 6.76 \quad 9.43 \quad 8.23 \quad 7.82 \quad 8.01 \quad 8.41 \quad 8.27 \quad 7.16\,.$$

(a) Calcolare un intervallo di confidenza bilatero di livello 95% per la media della variabile aleatoria X, specificando le ipotesi statistiche che si stanno usando. *[Ipotesi: $X \sim N(\mu, \sigma^2)$; $IC_\mu(95\%) = (7.5070\,,\, 8.5350)$]*

(b) Specificando anche qui le ipotesi statistiche che si stanno usando, calcolare un intervallo di confidenza per la varianza della variabile aleatoria X al livello 95% e del tipo seguente:

  (i) bilatero; *[Stessa ipotesi del punto (a); $IC_{\sigma^2}(95\%) = (0.2442\,,\, 1.7205)$]*

  (ii) unilatero del tipo $(0, c)$; *[Stessa ipotesi del punto (a); $IC_{\sigma^2}(95\%) = (0\,,\, 1.3973)$]*

  (iii) unilatero del tipo $(c, +\infty)$. *[Stessa ipotesi del punto (a); $IC_{\sigma^2}(95\%) = (0.2746\,,\, +\infty)$]*

Esercizio 18. Una ditta produce termostati per scaldabagni; i termostati sono costruiti per aprire un circuito a una temperatura t_0 incognita. Un addetto al controllo di qualità esamina uno di questi termostati nel modo seguente: immerge il termostato in un bagno termico a temperatura controllata da un termometro di precisione e fa aumentare la temperatura fino a quando il termostato apre il circuito, poi annota la temperatura alla quale il termostato ha aperto il circuito. Ripetendo questa procedura per 250 volte, l'addetto ottiene i seguenti valori per la media e la varianza campionarie delle 250 misure (espressi in °C):

$$\bar{x}_{250} = 53.375 \qquad\qquad s^2_{250} = 6.908\,.$$

(a) Calcolare un intervallo di confidenza bilatero di livello 95% per la media della variabile aleatoria $X =$ temperatura a cui il termostato apre il circuito, specificando le ipotesi statistiche sulla densità di X che si stanno assumendo. *[Non è necessario fare alcuna ipotesi sulla densità di X. $IC_{\mathbb{E}[X]}(95\%) = (53.049\,,\, 53.701)$]*

(b) Calcolare un intervallo di confidenza bilatero, di livello 95% per la varianza della variabile $X =$ temperatura a cui il termostato apre il circuito, specificando le ipotesi statistiche sulla densità di X che si stanno assumendo. *[Ipotesi: $X \sim N(\mu, \sigma^2)$; $IC_{\mathrm{Var}(X_i)}(95\%) = (5.848\,,\, 8.321)$]*

Esercizio 19. Una ditta produce cinghie di trasmissione per automobili. È noto che la durata di tali cinghie, calcolata in *migliaia* di chilometri percorsi, ha distribuzione normale con media 50 e deviazione standard 5. Denotiamo con X la durata di una cinghia. Successivamente la ditta progetta un congegno che – applicato al sistema di trasmissione – consente di allungare la 'vita' di una cinghia di una quantità $Y \sim N(\mu, \sigma^2)$, con parametri μ e σ^2 entrambi incogniti. In altre parole, la durata delle cinghie in presenza del congegno è $Z = X + Y$ e assumiamo che X e Y siano indipendenti. Un test su 30 automobili provviste del nuovo congegno ha permesso di raccogliere dati relativi ad un campione $Z_1, \ldots, Z_{30}$ che ha fornito $\bar{z}_{30} = 57$, $s_{30} = 6.5$, dove $\bar{Z}_{30}$ indica la media campionaria e S^2_{30} è la varianza campionaria di $Z_1, \ldots, Z_{30}$.

(a) Qual è la distribuzione di Z? *[$Z \sim N(50 + \mu, 25 + \sigma^2)$]*

(b) Sulla base del campione $Z_1, \ldots, Z_{30}$, fornite due stimatori non distorti per μ e σ^2, rispettivamente. Calcolate inoltre le stime corrispondenti. *[$\hat{\mu} = \bar{Z} - 50 = 7$; $\hat{\sigma}^2 = S_Z^2 - 25 = 17.25$]*

(c) Determinare un intervallo di confidenza per μ di livello 0.95. *[$IC_\mu(95\%) = (4.573, 9.427)$]*

(d) Determinare un intervallo di confidenza bilatero per la deviazione standard σ di livello 0.95. *[$IC_\sigma(95\%) = (1.341, 7.166)$]*