

STATISTICA (MODULO II - INFERENZA STATISTICA)
Soluzione esercitazione 5

A.

1. I valori della concentrazione di zinco riscontrati in un campione di $n = 12$ pesci possono essere usati per calcolare la concentrazione media e la varianza campionaria:

$$\begin{aligned}\bar{x} &= \frac{9,89 + 10,05 + \dots + 11,46}{12} = 9,188 \\ s^2 &= \frac{(9,89 - 9,188)^2 + (10,05 - 9,188)^2 + \dots + (11,46 - 9,188)^2}{12 - 1} = 1,386\end{aligned}$$

2. Assumendo che la concentrazione di zinco si distribuisca secondo una v.c. Gaussiana, gli estremi di un intervallo fiduciario al 99% per la concentrazione media di zinco sono dati da $\bar{x} \pm t_{\alpha/2; (n-1)} \sqrt{s^2/n}$, dove $t_{\alpha/2; (n-1)}$ è il quantile della distribuzione t di Student con $n - 1$ gradi di libertà corrispondente ad una probabilità sulla coda destra pari a $\alpha/2$. Nel caso in oggetto $\alpha = 0,01$, da cui $t_{0,005; 11} = 3,106$, per cui i limiti dell'intervallo sono pari a:

$$[9,188 \pm 3,106 \times \sqrt{1,386/12}] = [8,132; 10,246]$$

3. Assumendo che la concentrazione di zinco si distribuisca secondo una v.c. Gaussiana, gli estremi di un intervallo fiduciario al 95% per la varianza sono dati da $[s^2(n-1)/\chi_{\alpha/2, n-1}^2, s^2(n-1)/\chi_{1-\alpha/2, n-1}^2]$. Quindi:

$$[1,386 \times 11/21,92; 1,386 \times 11/3,82] = [0,695; 3,991]$$

dove $\chi_{0,025, 11}^2 = 21,92$ e $\chi_{0,975, 11}^2 = 3,82$.

4. Assumendo di conoscere $\sigma = 1,71$, gli estremi di un intervallo fiduciario al 99% per μ sono dati da $\bar{x} \pm z_{\alpha/2} \sigma / \sqrt{n}$. Quindi:

$$[9,188 \pm 2,576 \times 1,71/\sqrt{12}] = [7,916; 10,460]$$

dove $z_{0,005} = 2,576$.

B.

1. Stimatori non distorti della media e della varianza del consumo mensile di benzina delle famiglie sono rispettivamente:

$$\begin{aligned}\bar{x} &= \frac{34293}{213} = 161 \\ s^2 &= \frac{18062400}{213 - 1} = 85200\end{aligned}$$

2. L'intervallo fiduciario al 95% per μ è pari a

$$\bar{x} \pm z_{\alpha/2} \sqrt{s^2/n} = [161 \pm 1,96 \times \sqrt{85200/213}] = [121,8; 200,2]$$

dove $z_{0,025} = 1,96$.

3. Assumendo che il consumo mensile delle famiglie si distribuisca secondo una v.c. Gaussiana, gli estremi di un intervallo fiduciario al 90% per la varianza sono dati da

$$[s^2(n-1)/\chi_{\alpha/2, n-1}^2, s^2(n-1)/\chi_{1-\alpha/2, n-1}^2] = [85200 \times \frac{213-1}{246,968}; 85200 \times \frac{213-1}{179,305}] = [73136,6; 100735,6]$$

dove $\chi_{0,05; 212}^2 = 246,968$ e $\chi_{0,95; 212}^2 = 179,305$.

C. L'errore massimo di stima ϵ che si può commettere con probabilità del 99% si può scrivere come segue:

$$P(|\bar{X} - \mu| < \epsilon) = 0,99$$

Considerando un campione $n = 100$ sufficiente per poter affermare che $\bar{X} \approx N(\mu, \sigma^2/n)$ si ha:

$$P\left(\frac{|\bar{X} - \mu|}{\sqrt{\sigma^2/n}} < \frac{\epsilon}{\sqrt{\sigma^2/n}}\right) = P\left(|Z| < \frac{\epsilon}{\sqrt{\sigma^2/n}}\right) = 0,99$$

dove $Z \approx N(0,1)$. Sostituendo a σ^2 la sua stima $s^2 = 0,01$ si ottiene

$$P(|Z| < \epsilon/\sqrt{s^2/n}) = 0,99$$

per cui il quantile della normale standardizzata tale $\Phi(z) = 1 - 0,01/2 = 0,995$ è pari a $z = 2,576$. Sostituendo e risolvendo rispetto a ϵ si ottiene $\epsilon = 2,576\sqrt{0,01/100} = 0,02576$.

D. Si definisca la v.c. Y che esprime il numero di possessori di carta di credito che hanno subito un addebito durante l'anno precedente. Tale v.c. ha distribuzione binomiale con $n = 200$ e p incognito. La stima del parametro incognito è data da $\hat{p} = 23/200 = 0,115$, la frazione di possessori che hanno avuto un addebito. Gli estremi di un intervallo fiduciario al 90% per p sono dati da $\hat{p} \pm z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}$. Quindi:

$$[0,115 \pm 1,645\sqrt{0,115(1-0,115)/200}] = [0,0779; 0,1521]$$

dove $z_{0,1/2} = 1,645$.

E.

1. Relativamente ai lavoratori che guadagnano tra i 20 e i 40 mila euro l'anno, la proporzione dei lavoratori dipendenti è pari a $\hat{p} = 58/108 = 0,537$. L'intervallo fiduciario al 95% ha limiti posti a

$$[\hat{p} \pm z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}] = [0,537 \pm 1,96\sqrt{0,537(1-0,537)/108}] = [0,443; 0,631]$$

2. I valori centrali di classe per il reddito sono $y_j = \{15; 30,5; 60,5\}$, mentre le frequenze per la distribuzione dei liberi professionisti sono $n_{2j} = \{22; 50; 54\}$. Il reddito medio e la varianza campionaria sono, rispettivamente, pari a $\bar{y}_2 = 40,65$ e $s_2^2 = 327,22$.

L'intervallo fiduciario al 99% per il reddito medio dei liberi professionisti ha limiti posti a:

$$[\bar{y}_2 \pm z_{\alpha/2}\sqrt{s_2^2/n_{2o}}] = [40,65 \pm 2,576\sqrt{327,22/126}] = [36,5; 44,8]$$

3. La proporzione di lavoratori dipendenti è pari a $\hat{p} = 116/242 = 0,479$. I limiti dell'intervallo fiduciario al 99% sono pari a

$$[0,479 \pm 2,576\sqrt{0,479(1-0,479)/242}] = [0,396; 0,562]$$

Se la forza lavoro fosse equamente suddivisa tra lavoratori dipendenti e liberi professionisti allora $p = 0,5$. Dal momento che tale valore è compreso all'interno dei limiti dell'intervallo fiduciario sopra calcolato, possiamo ritenere plausibile l'affermazione che nella popolazione $p = 0,5$.