Soluzioni Esame di Probabilità e Statistica [3231]

Soluzioni Esame di Calcolo delle Probabilità e Statistica [2959]

Corso di Studi di Ingegneria Gestionale (D.M.270/04) (L)

Dipartimento di Meccanica, Matematica e Management
Politecnico di Bari

Cognome: ______________________	A.A.: 2021/2022
Nome: ______________________	Docente: Gianluca Orlando
Matricola: ______________________	Appello: luglio 2022
Corso di studi: ______________________	Data: 18/07/2022

Tempo massimo: 2 ore.

Esercizio 1. Un gruppo di topi di 5 settimane viene sottoposto a una dose di radiazione di 300 rad. La seguente tabella riporta le frequenze assolute dei giorni di vita dei topi suddivisi in intervalli di classi:

intervallo	frequenza
$[150, 200)$	6
$[200, 300)$	10
$[300, 500)$	12
$[500, 600)$	2
$[600, 900)$	3

1. Rappresentare un istogramma delle densità di frequenze relative.

2. Determinare la classe modale.

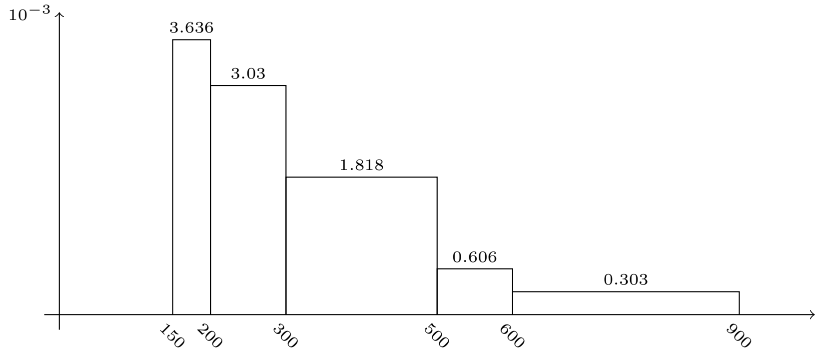
3. Calcolare un'approssimazione della media e della deviazione standard dei dati.

4. Calcolare un'approssimazione della mediana dei dati.

Soluzione. 1. Denotiamo con $I_1, \dots, I_5$ gli intervalli, $f_1, \dots, f_5$ le frequenze assolute. Abbiamo che $n = f_1 + \dots + f_5 = 6 + 10 + 12 + 2 + 3 = 33$. Ricordiamo che le frequenze relative sono date da $p_j = f_j/n$ e le densità di frequenze relative da $d_j = p_j/|I_j|$ dove $|I_j| = b_j - a_j$ se $I_j = [a_j, b_j)$. Completiamo la tabella (scriviamo anche le frequenze assolute cumulate per il punto 4.):

intervallo	f. assolute	f. relative	densità f. rel.	f. cumulate
$[150, 200)$	6	18.18%	$3.636 \cdot 10^{-3}$	6
$[200, 300)$	10	30.30%	$3.03 \cdot 10^{-3}$	16
$[300, 500)$	12	36.36%	$1.818 \cdot 10^{-3}$	28
$[500, 600)$	2	6.06%	$6.06 \cdot 10^{-4}$	30
$[600, 900)$	3	9.1%	$3.03 \cdot 10^{-4}$	33

(nell'ultima percentuale abbiamo approssimato a 9.1 in modo che $18.18+30.30+36.36+6.06+9.1=100$.) Rappresentiamo l'istogramma:



2. La classe modale è l'intervallo $[150,200)$ poiché è l'intervallo con la maggiore densità di frequenza relativa.
3. Ricordando che la media calcolata su un campione di dati $x_1,\dots,x_n$ è

$$\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i,$$

possiamo riscrivere la formula in termini dei valori assunti $v_1,\dots,v_k$ utilizzando le frequenze assolute $f_1,\dots,f_k$ e le frequenze relative $p_1,\dots,p_k$

$$\overline{x}=\frac{1}{n}\sum_{j=1}^{k}f_jv_j=\sum_{j=1}^{k}p_jv_j\,.$$

Per approssimare la media sostituiamo ai valori v_j i valori centrali $\tilde{v}_j$ degli intervalli e le frequenze relative:

$$\overline{x}\simeq 18.18\%\cdot 175+30.30\%\cdot 250+36.36\%\cdot 400+6.06\%\cdot 550+9.1\%\cdot 750=354.6\,.$$

Ricordiamo che la varianza calcolata su un campione di dati $x_1,\dots,x_n$ può essere calcolata mediante la formula:

$$s^2=\frac{1}{n-1}\Big(\sum_{i=1}^{n}x_i^2-n\overline{x}^2\Big).$$

Possiamo riscrivere la formula in termini dei valori assunti $v_1,\dots,v_k$ utilizzando le frequenze assolute $f_1,\dots,f_k$ e le frequenze relative $p_1,\dots,p_k$:

$$s^2=\frac{1}{n-1}\Big(\sum_{j=1}^{k}f_jv_j^2-n\overline{x}^2\Big)=\frac{n}{n-1}\Big(\sum_{j=1}^{k}p_jv_j^2-\overline{x}^2\Big).$$

Per approssimare la varianza sostituiamo ai valori v_j i valori centrali $\tilde{v}_j$ degli intervalli e le frequenze relative e la media approssimata calcolata nel punto precedente:

$$s^2\simeq\frac{33}{32}\Big(18.18\%\cdot 175^2+30.30\%\cdot 250^2+36.36\%\cdot 400^2+6.06\%\cdot 550^2+9.1\%\cdot 750^2-354.6^2\Big)=27286$$

da cui segue che la deviazione standard è approssimata da

$$s\simeq 165.184.$$

4. Per calcolare un'approssimazione della mediana utilizziamo le frequenze cumulate $F_1, \dots, F_k$. La mediana divide l'insieme il campione di dati in due parti, quindi calcoliamo $\frac{n}{2} = \frac{33}{2} = 16.5$ e osserviamo che per l'intervallo $I_3 = [a_3, b_3) = [300, 500)$ si ha che

$$F_2 = 16 < 16.5 < 28 = F_3 \,.$$

La mediana è allora approssimata da

$$Q_2 \simeq a_3 + \lambda_3(b_3 - a_3), \quad \lambda_3 = \frac{\frac{n}{2} - F_2}{F_3 - F_2} = \frac{16.5 - 16}{12} = \frac{1}{24}$$

quindi

$$Q_2 \simeq 300 + \frac{1}{24} 200 \simeq 308.33.$$

Esercizio 2. Un'azienda produce grandi numeri di mobili montabili. In un particolare mobile ci sono due pezzi (che chiameremo A e B) che possono risultare difettosi. In media vengono prodotti ogni giorno 1 pezzo A difettoso e (indipendentemente) 2 pezzi B difettosi. Per entrambi i tipi, il numero di pezzi difettosi è distribuito con una legge di Poisson.

1. Qual è la probabilità che vengano prodotti (strettamente) più di 4 pezzi A difettosi in 5 giorni? (Si assumano i difetti nei diversi giorni indipendenti)

2. Sappiamo che in 5 giorni sono stati prodotti in tutto 12 pezzi difettosi (contando sia tipo A che B). Qual è la probabilità che al più 3 pezzi A siano difettosi?

Soluzione. 1. Il numero di pezzi A difettosi in un giorno è distribuito come una variabile aleatoria $X_1 \sim P(\lambda)$. Poiché $\mathbb{E}(X_1) = \lambda$ e la traccia ci dice che il numero medio di pezzi A difettosi in un giorno è 1, abbiamo che $\lambda = 1$. Quindi

$$\mathbb{P}(\{X_1 = k\}) = \frac{e^{-1}}{k!}.$$

Il numero di pezzi A difettosi in 5 giorni è la somma di 5 variabili aleatorie di Poisson $X_i \sim P(1)$

$$X = X_1 + X_2 + X_3 + X_4 + X_5 \sim P(5\lambda) = P(5)\,,$$

poiché i difetti nei diversi giorni sono indipendenti e la somma di Poisson indipendenti è una Poisson con il parametro dato dalla somma dei parametri. Quindi

$$\mathbb{P}(\{X = k\}) = e^{-5}\frac{5^k}{k!}\,.$$

Possiamo calcolare

$$\begin{aligned}
\mathbb{P}(\{X > 4\}) &= 1 - \mathbb{P}(\{X \le 4\}) \\
&= 1 - \mathbb{P}(\{X = 0\}) - \mathbb{P}(\{X = 1\}) - \mathbb{P}(\{X = 2\}) - \mathbb{P}(\{X = 3\}) - \mathbb{P}(\{X = 4\}) \\
&= 1 - e^{-5} - e^{-5}5 - e^{-5}\frac{5^2}{2} - e^{-5}\frac{5^3}{3!} - e^{-5}\frac{5^4}{4!} \simeq 55.95\%.
\end{aligned}$$

2. Il numero di pezzi B difettosi in un giorno è distribuito come una variabile aleatoria $Y_1 \sim P(\mu)$. Poiché $\mathbb{E}(Y_1) = \mu$ e la traccia ci dice che il numero medio di pezzi B difettosi in un giorno

è 2, abbiamo che $\mu = 2$. Il numero di pezzi B difettosi in 5 giorni è la somma di 5 variabili aleatorie di Poisson $Y_i \sim P(2)$

$$Y = Y_1 + Y_2 + Y_3 + Y_4 + Y_5 \sim P(5\mu) = P(10)\,,$$

poiché i difetti nei diversi giorni sono indipendenti e la somma di Poisson indipendenti è una Poisson con il parametro dato dalla somma dei parametri. Quindi

$$\mathbb{P}(\{Y = k\}) = e^{-10}\frac{10^k}{k!}\,.$$

Il numero di pezzi difettosi sia di tipo A che B è la somma $X + Y$. Queste sono due variabili aleatorie di Poisson indipendenti (lo dice la traccia), quindi $X + Y \sim P(5 + 10) = P(15)$ e

$$\mathbb{P}(\{X + Y = k\}) = e^{-15}\frac{15^k}{k!}.$$

La traccia chiede di calcolare

$$\begin{aligned}
\mathbb{P}(\{X \le 3\}|\{X+Y=12\}) &= \frac{\mathbb{P}(\{X \le 3\}\cap\{X+Y=12\})}{\mathbb{P}(\{X+Y=12\})}\\
= \frac{\mathbb{P}(\{X=0\}\cap\{Y=12\})}{\mathbb{P}(\{X+Y=12\})} &+ \frac{\mathbb{P}(\{X=1\}\cap\{Y=11\})}{\mathbb{P}(\{X+Y=12\})}\\
+ \frac{\mathbb{P}(\{X=2\}\cap\{Y=10\})}{\mathbb{P}(\{X+Y=12\})} &+ \frac{\mathbb{P}(\{X=3\}\cap\{Y=9\})}{\mathbb{P}(\{X+Y=12\})}.
\end{aligned}$$

Utilizzando l'indipendenza di X e Y:

$$\begin{aligned}
&\mathbb{P}(\{X \le 3\}|\{X+Y=12\})\\
&= \frac{\mathbb{P}(\{X=0\})\mathbb{P}(\{Y=12\})}{\mathbb{P}(\{X+Y=12\})} + \frac{\mathbb{P}(\{X=1\})\mathbb{P}(\{Y=11\})}{\mathbb{P}(\{X+Y=12\})}\\
&\quad + \frac{\mathbb{P}(\{X=2\})\mathbb{P}(\{Y=10\})}{\mathbb{P}(\{X+Y=12\})} + \frac{\mathbb{P}(\{X=3\})\mathbb{P}(\{Y=9\})}{\mathbb{P}(\{X+Y=12\})}\\
&= \frac{e^{-5}e^{-10}\frac{10^{12}}{12!}}{e^{-15}\frac{15^{12}}{12!}} + \frac{e^{-5}5e^{-10}\frac{10^{11}}{11!}}{e^{-15}\frac{15^{12}}{12!}} + \frac{e^{-5}\frac{5^2}{2!}e^{-10}\frac{10^{10}}{10!}}{e^{-15}\frac{15^{12}}{12!}} + \frac{e^{-5}\frac{5^3}{3!}e^{-10}\frac{10^9}{9!}}{e^{-15}\frac{15^12}{12!}}\\
&= \frac{1}{\frac{15^{12}}{12!}}\left(\frac{10^{12}}{12!} + 5\frac{10^{11}}{11!} + \frac{5^2}{2!}\frac{10^{10}}{10!} + \frac{5^3}{3!}\frac{10^9}{9!}\right) \simeq 39.31\%.
\end{aligned}$$

Esercizio 3. Sono le 8:30 e sei allo sportello della tua banca per sbrigare una pratica. In media la pratica dura 10 minuti, e la durata in minuti è distribuita come una legge esponenziale. Dopo aver sbrigato la pratica in banca devi andare a lavoro prendendo un bus, che però ha un tempo di arrivo alla fermata (proprio all'uscita della banca) incerto. L'orario di arrivo del bus ha distribuzione uniforme, in media arriva alle 8:40, con una deviazione standard di 10 min.

1. Con che probabilità si verifica il seguente evento: finirai la pratica dopo le 8:50 e il bus arriverà prima delle 8:50?

2. Hai finito la pratica, corri fuori e scopri che il bus è già passato. Non hai guardato l'orologio, quindi non sai quanto tempo è durata la pratica e a che ora è passato il bus. Constatando che non sei riuscito a prendere il bus, è più probabile che la pratica sia durata meno di 10 min oppure che il bus sia arrivato prima delle 8:40? Motivare la risposta. (N.B.: non è richiesto il calcolo esplicito delle due probabilità!)

Soluzione. 1. Denotiamo con $X \sim \mathrm{Exp}(\lambda)$ la durata in minuti della pratica allo sportello della banca. Sappiamo che $\mathbb{E}(X) = \frac{1}{\lambda} = 10$, quindi $\lambda = \frac{1}{10}$ e quindi

$$\mathbb{P}(\{X \geq t\}) = \int_t^{+\infty} \frac{1}{10} e^{-\frac{1}{10}x}\, \mathrm{d}x = \Big[- e^{-\frac{1}{10}x} \Big]_t^{+\infty} = e^{-\frac{1}{10}t}.$$

Denotiamo con Y il minuto di arrivo del bus. Sappiamo che $Y \sim \mathrm{U}(a,b)$ per qualche intervallo $[a,b]$. Calcoliamo a, b utilizzando le informazioni sulla media e sulla varianza. Ricordiamo che

$$\mathbb{E}(Y) = \int_{-\infty}^{+\infty} y f_Y(y)\, \mathrm{d}y = \frac{1}{b-a}\int_a^b y\, \mathrm{d}y = \frac{1}{b-a}\Big[\frac{y^2}{2}\Big]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{2}$$

da cui

$$\frac{b+a}{2} = 40.$$

Inoltre

$$\begin{aligned}
\mathrm{Var}(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 &= \int_{-\infty}^{+\infty} y^2 f_Y(y)\, \mathrm{d}y - \Big(\frac{b+a}{2}\Big)^2 = \frac{1}{b-a}\int_a^b y^2\, \mathrm{d}y - \Big(\frac{b+a}{2}\Big)^2 \\
&= \frac{1}{b-a}\Big[\frac{y^3}{3}\Big]_a^b - \Big(\frac{b+a}{2}\Big)^2 = \frac{b^3-a^3}{3(b-a)} - \frac{b^2+2ab+a^2}{4} = \\
&= \frac{b^2+ab+a^2}{3} - \frac{b^2+2ab+a^2}{4} = \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12}
\end{aligned}$$

da cui

$$\sqrt{\frac{(b-a)^2}{12}} = 10 \implies \frac{b-a}{2} = 5\sqrt{12}.$$

Segue che

$$a = 40 - 5\sqrt{12}\,, \quad b = 40 + 5\sqrt{12}$$

Calcoliamo, utilizzando il fatto che la durata della pratica e il tempo di arrivo del bus sono indipendenti,

$$\begin{aligned}
\mathbb{P}(\{X > 20\} \cap \{Y < 50\}) &= \mathbb{P}(\{X > 20\})\mathbb{P}(\{Y < 50\}) \\
&= \Big(\int_{20}^{+\infty} \frac{1}{10} e^{-\frac{1}{10}x}\, \mathrm{d}x\Big)\Big(\frac{1}{10\sqrt{12}}\int_{40-5\sqrt{12}}^{50} \mathrm{d}y\Big) \\
&= e^{-\frac{20}{10}} \frac{50-40+5\sqrt{12}}{10\sqrt{12}} = e^{-2}\frac{1+\sqrt{3}}{2\sqrt{3}} \simeq 10.67\%
\end{aligned}$$

2. L'evento "perdo il bus" è $\{30 + X > Y\}$. Osserviamo che se $X < 10$ e $X + 30 > Y$, necessariamente $Y < 40$. Quindi

$$\{X < 10\} \cap \{X + 30 > Y\} = \{X < 10\} \cap \{X + 30 > Y\} \cap \{Y < 40\} \subset \{X + 30 > Y\} \cap \{Y < 40\}.$$

Usando la monotonia della probabilità possiamo allora stimare che

$$\begin{aligned}
\mathbb{P}(\{X < 10\} | \{X + 30 > Y\}) &= \frac{\mathbb{P}(\{X < 10\} \cap \{X + 30 > Y\})}{\mathbb{P}(\{X + 30 > Y\})} \\
&\leq \frac{\mathbb{P}(\{Y < 40\} \cap \{X + 30 > Y\})}{\mathbb{P}(\{X + 30 > Y\})} = \mathbb{P}(\{Y < 40\} | \{X + 30 > Y\})
\end{aligned}$$

ovvero, è più probabile che il bus sia arrivato prima delle 8:40.

Osserviamo che

$$\mathbb{P}(\{X < 10\}) = 1 - \mathbb{P}(\{X \geq 10\}) = 1 - e^{-1} \simeq 63.21\%$$

e

$$\mathbb{P}(\{Y < 40\}) = \frac{1}{2},$$

cioè

$$\mathbb{P}(\{X < 10\}) \geq \mathbb{P}(\{Y < 40\}),$$

quindi il fatto di sapere che abbiamo perso il bus cambia la relazione tra le due probabilità.

Esercizio 4. Una riempitrice automatica viene utilizzata per riempire dosatori da 100 ml con gel igienizzante. Viene misurato il volume di riempimento effettivo in un campione casuale di 40 dosatori. La media campionaria e la deviazione standard campionaria calcolate sui dati risultano essere 99.90 ml e 0.55 ml rispettivamente. È possibile affermare con il 5% di significatività che in media la macchina immette meno di 100 ml di gel? Qual è il più piccolo livello di significatività per cui i dati permettono di affermare che la macchina immette meno di 100 ml di gel?
(N.B.: Ricavare le formule!)

Soluzione. Si tratta di un campione casuale $X_1, \dots, X_n$ con $n = 40$ di cui non è nota la distribuzione. I dati forniscono le realizzazioni della media campionaria e della deviazione standard campionaria

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

che sono data rispettivamente da 99.90 ml e 0.55 ml.
La domanda della traccia (se i dati sono abbastanza significativi per concludere che in media la macchina immette meno di 100ml di gel) può essere impostata come un test d'ipotesi sulla media μ della popolazione

$$H_0 : \mu = \mu_0 \quad H_1 : \mu < \mu_0$$

dove $\mu_0 = 100$. (È accettata come soluzione corretta anche la formulazione con $H_0 : \mu \geq \mu_0$, ma in quel caso bisogna stare attenti a usare μ come media vera della popolazione per l'utilizzo della statistica corretta!).
Il livello di significatività del test α è la probabilità di commettere un errore del I tipo, ovvero di rifiutare l'ipotesi nulla quando questa è vera. Supponiamo H_0 vera, ovvero $\mu = \mu_0$. La regione di rifiuto è della forma

$$R_C = \{(x_1, \dots, x_n) \in R(X_1, \dots, X_n) \ : \ \overline{x}_n < \mu_0 - \delta\},$$

dove $\overline{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$. Allora

$$\alpha = \mathbb{P}(\{\overline{X}_n < \mu_0 - \delta\}) = \mathbb{P}(\{\overline{X}_n - \mu_0 < -\delta\}) = \mathbb{P}\Big(\Big\{\frac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}} < -\frac{\delta}{S_n/\sqrt{n}}\Big\}\Big).$$

(Se l'ipotesi nulla è invece $H_0 : \mu \geq \mu_0$, allora si deve usare il seguente fatto: se $\overline{X}_n < \mu_0 - \delta$ allora a maggior ragione $\overline{X}_n < \mu - \delta$ e quindi $\mathbb{P}(\{\overline{X}_n < \mu_0 - \delta\}) \leq \mathbb{P}(\{\overline{X}_n < \mu - \delta\})$ e si può procedere con i calcoli utilizzando la media vera della popolazione μ al posto di μ_0.)
Non è detto che la statistica $\frac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}}$ sia una t-Student con $n-1$ gradi di libertà, poiché non conosciamo la distribuzione della popolazione. Tuttavia il campione è numeroso ($n = 40 > 30$)

e possiamo quindi sfruttare un teorema di approssimazione. Nella fattispecie utilizziamo il Teorema di Slutsky (N.B.: a denominatore c'è S_n, non σ) per affermare che $\frac{\overline{X}_n-\mu_0}{S_n/\sqrt{n}}$ è approssimata da una variabile aleatoria Z distribuita con legge normale standard. Quindi

$$\mathbb{P}\Big(\Big\{\frac{\overline{X}_n-\mu_0}{S_n/\sqrt{n}} < -\frac{\delta}{S_n/\sqrt{n}}\Big\}\Big) \simeq \mathbb{P}\Big(\Big\{Z < -\frac{\delta}{S_n/\sqrt{n}}\Big\}\Big).$$

Scegliamo $\frac{\delta}{S_n/\sqrt{n}} = z_\alpha$, dove z_α è il quantile gaussiano. In questo modo

$$\mathbb{P}\Big(\Big\{Z < -\frac{\delta}{S_n/\sqrt{n}}\Big\}\Big) = \mathbb{P}(\{Z < -z_\alpha\}) = \alpha$$

e la condizione sulla significatività $\mathbb{P}(\{\overline{X}_n < \mu_0 - \delta\}) = \alpha$ è effettivamente verificata. In conclusione, la regione critica è

$$R_C = \Big\{(x_1,\dots,x_n) \in R(X_1,\dots,X_n)\ :\ \overline{x}_n < \mu_0 - \frac{s_n}{\sqrt{n}}z_\alpha\Big\},$$

dove $s_n^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \overline{x}_n)^2$ è la varianza campionaria calcolata sul campione dei dati.
Risolviamo direttamente il secondo quesito, poiché la risposta porterà a rispondere al primo quesito. Per definizione, il p-value è il più piccolo livello di significatività per cui i dati osservati portano a un rifiuto dell'ipotesi nulla, ovvero

$$\begin{aligned}
p\text{-value} &= \inf_\alpha \Big\{\overline{x}_n < \mu_0 - \frac{s_n}{\sqrt{n}}z_\alpha\Big\} = \inf_\alpha \Big\{\frac{\overline{x}_n-\mu_0}{s_n/\sqrt{n}} < -z_\alpha\Big\} = \inf_\alpha \Big\{\Phi\Big(\frac{\overline{x}_n-\mu_0}{s_n/\sqrt{n}}\Big) < \Phi(-z_\alpha)\Big\} \\
&= \inf_\alpha \Big\{\Phi\Big(\frac{\overline{x}_n-\mu_0}{s_n/\sqrt{n}}\Big) < \Phi(-z_\alpha)\Big\} = \inf_\alpha \Big\{\Phi\Big(\frac{\overline{x}_n-\mu_0}{s_n/\sqrt{n}}\Big) < \alpha\Big\} = \Phi\Big(\frac{\overline{x}_n-\mu_0}{s_n/\sqrt{n}}\Big).
\end{aligned}$$

dove abbiamo usato che

$$\Phi(-z_\alpha) = \mathbb{P}(\{Z < -z_\alpha\}) = \alpha.$$

Quindi, utilizzando le tavole,

$$\begin{aligned}
p\text{-value} &= \Phi\Big(\frac{\overline{x}_n-\mu_0}{s_n/\sqrt{n}}\Big) = \Phi\Big(\frac{99.90-100}{0.55/\sqrt{40}}\Big) = \Phi(-1.15) = 1-\Phi(1.15) \\
&\simeq 1 - 0.8749 = 0.1251 = 12.51\%.
\end{aligned}$$

In particolare, $5\% < 12.51\%$, quindi l'ipotesi nulla non è rifiutata con il 5% di significatività.
(Per completezza, mostriamo come controllarlo direttamente. La regione critica è costituita dai dati tali che

$$\overline{x}_n < \mu_0 - \frac{s_n}{\sqrt{n}}z_\alpha.$$

Per calcolare z_α utilizziamo che $\alpha = \mathbb{P}(\{Z > z_\alpha\})$, quindi $0.95 = 1-\alpha = \mathbb{P}(\{Z \le z_\alpha\})$. Utilizzando le tavole:

$$z_\alpha = z_{0.05} \simeq 1.645.$$

Segue che

$$\mu_0 - \frac{s_n}{\sqrt{n}}z_\alpha = 100 - \frac{0.55}{\sqrt{40}}1.645 \simeq 99.86.$$

Poiché $\overline{x}_n = 99.90 > 99.86$, non rifiutiamo l'ipotesi nulla.)