Esame di Probabilità e Statistica [3231]

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: ______________________ Docente: Gianluca Orlando
Nome: ______________________ Appello: novembre 2023
Matricola: ______________________ Data: 06/11/2023

Tempo massimo: 2 ore.

Esercizio 1. (6 punti) È stato misurato il tempo di attesa (in giorni) per un particolare intervento chirurgico su un campione di individui. La distribuzione dei tempi di attesa è riassunta nella seguente tabella:

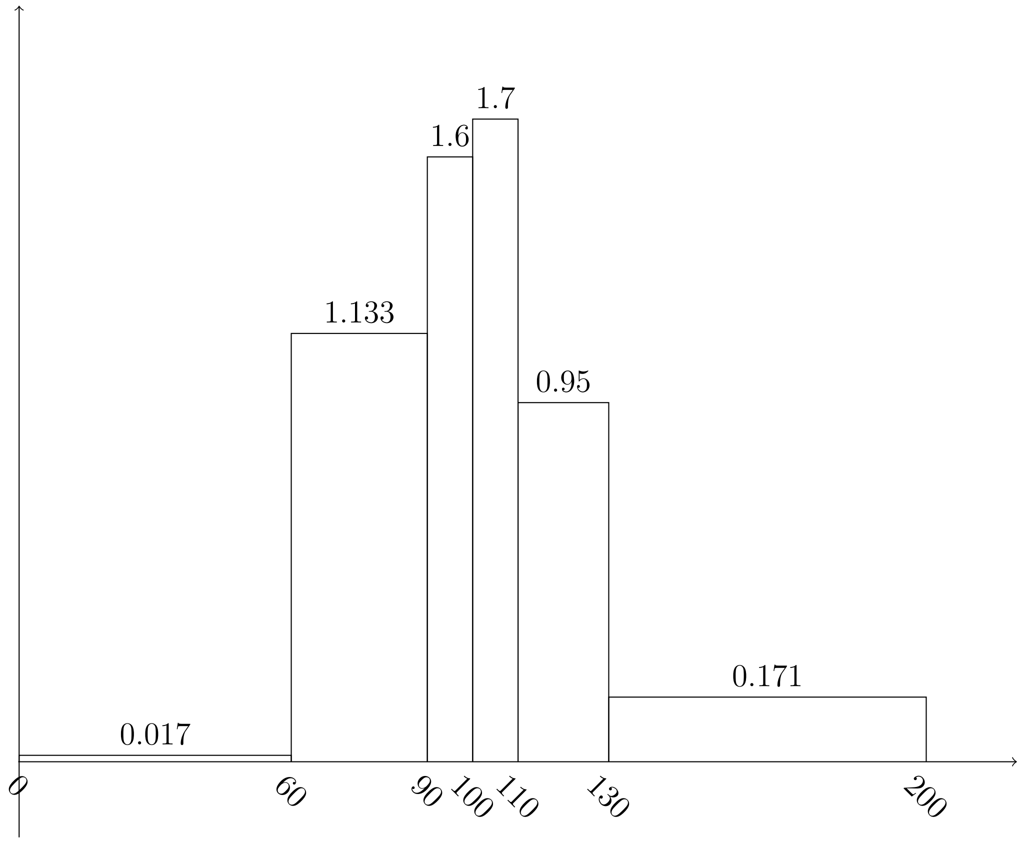
intervalli (giorni)	frequenza assoluta
$[0, 60)$	1
$[60, 90)$	34
$[90, 100)$	16
$[100, 110)$	18
$[110, 130)$	19
$[130, 200)$	12

1. Rappresentare un istogramma delle densità di frequenze assolute.
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. Completiamo la tabella con tutte le informazioni necessarie. Ricordiamo che la densità di frequenza assoluta si ottiene rapportando le frequenze assolute all'ampiezza degli intervalli.

intervallo	freq. assolute	freq. relative	densità di freq. ass.	freq. cumulate
$[0, 60)$	1	1%	0.017	1
$[60, 90)$	34	34%	1.133	35
$[90, 100)$	16	16%	1.6	51
$[100, 110)$	18	18%	1.7	69
$[110, 130)$	19	19%	0.95	88
$[130, 200)$	12	12%	0.171	100

Rappresentiamo le densità di frequenze assolute in un istogramma.



2. La classe modale è quella con maggiore densità di frequenza assoluta, quindi è l'intervallo $[100, 110)$.

3. Per calcolare un'approssimazione della media utilizziamo le frequenze relative ottenute da $p_j = f_j/n$ dove $n = 100$ e i valori centrali $\tilde{v}_j$ degli intervalli

$$\begin{aligned}\overline{x} &= \frac{1}{n}\sum_{i=1}^{n} x_i \simeq \frac{1}{n}\sum_{j=1}^{k} f_j\tilde{v}_j = \sum_{j=1}^{k} p_j\tilde{v}_j \\ &= 1\% \cdot 30 + 34\% \cdot 75 + 16\% \cdot 95 + 18\% \cdot 105 + 19\% \cdot 120 + 12\% \cdot 165 \simeq 102.5\,.\end{aligned}$$

Calcoliamo un'approssimazione della varianza

$$\begin{aligned}s^2 &= \frac{1}{n-1}\Big(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\Big) \simeq \frac{1}{n-1}\Big(\sum_{j=1}^{n} f_j\tilde{v}_j^2 - n\overline{x}^2\Big) = \frac{n}{n-1}\Big(\sum_{j=1}^{n} p_j\tilde{v}_j^2 - \overline{x}^2\Big) \\ &= \frac{100}{99}\Big(1\% \cdot 30^2 + 34\% \cdot 75^2 + 16\% \cdot 95^2 + 18\% \cdot 105^2 + 19\% \cdot 120^2 + 12\% \cdot 165^2 - 102.5^2\Big) \\ &= \frac{100}{99}(11353 - 10506.25) \simeq 855.30\,,\end{aligned}$$

da cui

$$s \simeq 29.25\,.$$

4. Per calcolare un'approssimazione della mediana dei dati, usiamo le frequenze cumulate. Troviamo l'intervallo I_j tale che $F_j \le \frac{n}{2} = 50 < F_{j+1}$. Si tratta dell'intervallo $[90, 100)$. Approssimiamo la mediana con

$$Q_2 \simeq a_j + \lambda_j(b_j - a_j)$$

dove

$$\lambda_j = \frac{n/2 - F_j}{F_{j+1} - F_j} = \frac{50 - 35}{51 - 35} = 0.9375\,.$$

Quindi

$$Q_2 \simeq 90 + 0.9375(100 - 90) = 99.375\,.$$

Esercizio 2. (8 punti) Si consideri un vettore aleatorio discreto (X_1, X_2) con $R(X_1) = \{0,1\}$ e $R(X_2) = \{0,1\}$. Si assuma che $\mathrm{Cov}(X_1, X_2) = 0$.

1. Dimostrare che X_1 e X_2 sono indipendenti. (N.B.: Il punteggio massimo si ottiene se non si utilizzano le informazioni di sotto.)

Si assuma ulteriormente che:

- $\mathrm{Var}(X_1) = \frac{1}{4}$.
- $\mathbb{P}(\{X_2 = 1\}|\{X_1 = 1\}) = \frac{1}{3}$.

Quesiti:

2. Calcolare $\mathbb{P}(\{X_1 \cdot X_2 > 0\})$.
3. Calcolare $\mathbb{P}(\{X_1 + X_2 = 1\})$.
4. Alice e Bob giocano al seguente gioco in cui si susseguono in *lanci*: Alice inizia e *lancia* generando un numero con X_1; Bob segue e *lancia* generando un numero con X_2. Alice e Bob continuano a *lanciare* alternandosi, finché i risultati degli ultimi due *lanci* non sono diversi. In tal caso vince chi dei due ha ottenuto il risultato più grande negli ultimi due lanci. Esempi:

$$X_1 = 1 \text{ (Alice)} \to X_2 = 1 \text{ (Bob)} \to X_1 = 0 \text{ (Alice)} \to \text{Fine: Bob vince.}$$

$$X_1 = 0 \text{ (Alice)} \to X_2 = 0 \text{ (Bob)} \to X_1 = 0 \text{ (Alice)} \to X_2 = 1 \text{ (Bob)} \to \text{Fine: Bob vince.}$$

Che probabilità di vincere ha Alice?

Soluzione. Scriviamo una tabella per la funzione di probabilità congiunta:

X_2 \ X_1	0	1
0	a_{00}	a_{10}
1	a_{01}	a_{11}

1. Poiché $R(X_1) = \{0,1\}$ e $R(X_2) = \{0,1\}$, entrambe sono variabili aleatorie distribuite con legge di Bernoulli, quindi $X_1 \sim \mathrm{Be}(p)$ e $X_2 \sim \mathrm{Be}(q)$ per certi p e q. Ricordando che $\mathbb{E}(X_1) = p$ e $\mathbb{E}(X_2) = q$, usiamo che $\mathrm{Cov}(X_1, X_2) = 0$ per dedurre

$$0 = \mathrm{Cov}(X_1, X_2) = \mathbb{E}(X_1 \cdot X_2) - \mathbb{E}(X_1)\mathbb{E}(X_2) = a_{11} - pq \implies pq = a_{11}\,.$$

Otteniamo che

$$p = \mathbb{P}(\{X_1 = 1\}) = a_{10} + a_{11} \implies a_{10} = p - a_{11} = p - pq = p(1-q)\,.$$

Analogamente

$$q = \mathbb{P}(\{X_2 = 1\}) = a_{01} + a_{11} \implies a_{01} = q - a_{11} = q - pq = q(1-p)\,.$$

Infine,

$$1 - p = \mathbb{P}(\{X_1 = 0\}) = a_{00} + a_{01} = a_{00} + q - pq \implies a_{00} = 1 - p - q + pq = (1-p)(1-q)\,.$$

Possiamo completare la tabella:

X_2 \ X_1	0	1
0	$(1-p)(1-q)$	$p(1-q)$
1	$(1-p)q$	pq

e osserviamo che $\mathbb{P}(\{X_1 = i\} \cap \{X_2 = j\}) = \mathbb{P}(\{X_1 = i\})\mathbb{P}(\{X_2 = j\})$ per ogni $i, j \in \{0, 1\}$.

2. Ricordiamo che per una Bernoulli $\mathrm{Var}(X_1) = p(1-p)$. Quindi

$$p - p^2 = \frac{1}{4} \implies p^2 - p + \frac{1}{4} = 0 \implies \left(p - \frac{1}{2}\right)^2 = 0 \implies p = \frac{1}{2}\,.$$

Usiamo la seconda informazione. Si ha che

$$\frac{1}{3} = \mathbb{P}(\{X_2 = 1\} | \{X_1 = 1\}) = \frac{\mathbb{P}(\{X_2 = 1\} \cap \{X_1 = 1\})}{\mathbb{P}(\{X_1 = 1\})} = \frac{pq}{p} = q\,.$$

(Chi riesce a dimostrare il punto 1., può usarlo qui per dire che $\mathbb{P}(\{X_2 = 1\}|\{X_1 = 1\})) = \mathbb{P}(\{X_2 = 1\}) = q$.) Abbiamo la tabella completa:

X_2 \ X_1	0	1
0	$\frac{1}{3}$	$\frac{1}{3}$
1	$\frac{1}{6}$	$\frac{1}{6}$

Possiamo calcolare

$$\mathbb{P}(\{X_1 \cdot X_2 > 0\}) = \mathbb{P}(\{X_1 \cdot X_2 = 1\}) = \frac{1}{6}\,.$$

3. Si ha che

$$\mathbb{P}(\{X_1 + X_2 = 1\}) = \mathbb{P}(\{X_1 = 1\} \cap \{X_2 = 0\}) + \mathbb{P}(\{X_1 = 0\} \cap \{X_2 = 1\}) = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}\,.$$

4. Consideriamo l'evento

$$A = \text{“Alice vince”}$$

e la sequenza di variabili aleatorie

$$Y_n = \text{“esito dell'}n\text{-esimo } \textit{lancio}\text{”}$$

Distinguiamo i casi in base al risultato iniziale. Supponiamo di sapere che al primo lancio $Y_1 = 1$. Se Bob ottiene $Y_2 = 0$ (con probabilità $\mathbb{P}(\{X_2 = 0\}) = 1 - q = \frac{2}{3}$), Alice vince automaticamente; se Bob ottiene $Y_2 = 1$ (con probabilità $\mathbb{P}(\{X_2 = 1\}) = q = \frac{1}{3}$), tocca nuovamente a Alice. Se Alice ottiene $Y_3 = 0$ (con probabilità $\mathbb{P}(\{X_1 = 0\}) = 1 - p = \frac{1}{2}$), perde;

se Alice ottiene $Y_3 = 1$ (con probabilità $\mathbb{P}(\{X_1 = 1\}) = p = \frac{1}{2}$), il gioco si è resettato e siamo nella situazione precedente! Scriviamo questa relazione ricorsiva:

$$\begin{aligned}
\mathbb{P}(A|\{Y_1 = 1\}) &= (1-q) + q\Big(p\mathbb{P}(A|\{Y_1 = 1\})\Big) \\
&= \frac{2}{3} + \frac{1}{3}\Big(\frac{1}{2}\mathbb{P}(A|\{Y_1 = 1\})\Big) \\
\implies &\frac{5}{6}\mathbb{P}(A|\{Y_1 = 1\}) = \frac{2}{3} \\
\implies &\mathbb{P}(A|\{Y_1 = 1\}) = \frac{4}{5}\,.
\end{aligned}$$

Supponiamo di sapere che al primo lancio $Y_1 = 0$. Se Bob ottiene $Y_2 = 1$ (con probabilità $\mathbb{P}(\{X_2 = 1\}) = q = \frac{1}{3}$), Alice perde automaticamente; se Bob ottiene $Y_2 = 0$ (con probabilità $\mathbb{P}(\{X_2 = 0\}) = 1 - q = \frac{2}{3}$), tocca nuovamente a Alice. Se Alice ottiene $Y_3 = 1$ (con probabilità $\mathbb{P}(\{X_1 = 1\}) = p = \frac{1}{2}$), vince; se Alice ottiene $Y_3 = 0$ (con probabilità $\mathbb{P}(\{X_1 = 0\}) = 1-p = \frac{1}{2}$), il gioco si è resettato e siamo nella situazione precedente! Scriviamo questa relazione ricorsiva:

$$\begin{aligned}
\mathbb{P}(A|\{Y_1 = 0\}) &= (1-q)\Big(p + (1-p)\mathbb{P}(A|\{Y_1 = 0\})\Big) \\
&= \frac{2}{3}\Big(\frac{1}{2} + \frac{1}{2}\mathbb{P}(A|\{Y_1 = 0\})\Big) \\
\implies &\frac{2}{3}\mathbb{P}(A|\{Y_1 = 0\}) = \frac{1}{3} \\
\implies &\mathbb{P}(A|\{Y_1 = 0\}) = \frac{1}{2}\,.
\end{aligned}$$

Utilizziamo la formula della probabilità totale per concludere che

$$\mathbb{P}(A) = \mathbb{P}(A|\{Y_1 = 1\})\mathbb{P}(\{Y_1 = 1\}) + \mathbb{P}(A|\{Y_1 = 0\})\mathbb{P}(\{Y_1 = 0\}) = \frac{4}{5}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = \frac{13}{20}\,.$$

Esercizio 3. (8 punti) Si assuma che il ritardo di un singolo treno di un'azienda di trasporti sia distribuito con legge esponenziale. Sono gestiti due tipi di treni: il 20% dei treni sono treni veloci; il resto sono treni regionali. Si assumano i seguenti fatti:

- il ritardo medio di un treno veloce è 5 minuti;
- la probabilità che un treno regionale ritardi meno di $\ln(3^{10})$ (logaritmo naturale di tre elevato alla 10) minuti è uguale a $\frac{2}{3}$.

Rispondere ai seguenti quesiti:

1. Si consideri un treno veloce. Calcolare la probabilità che il ritardo sia compreso tra 3 e 7 minuti.
2. Calcolare la deviazione standard del ritardo di un treno regionale.
3. Si consideri un treno qualunque. Qual è la probabilità che ritardi più di 8 minuti?
4. (Difficile) Sei su un treno veloce e devi prendere una coincidenza con un treno regionale ad una certa stazione. La tabella di marcia prevede che il treno veloce arrivi e il treno regionale arrivino alla stazione entrambi alle 14:00. Considerando i possibili ritardi, con che probabilità perderai la coincidenza? Supponi che i ritardi siano indipendenti e che il tempo per passare da un treno all'altro sia trascurabile.

Soluzione. Consideriamo tre variabili aleatorie

$$X_{\text{vel}} = \text{“ritardo di un treno veloce”} \sim \text{Exp}(\lambda)\,,$$

$$X_{\text{reg}} = \text{“ritardo di un treno regionale”} \sim \text{Exp}(\mu)\,,$$

$$Y = \text{“tipo di treno”} \sim \text{Be}(20\%)\,,$$

dove $Y = 1$ se il treno è veloce, $Y = 0$ se il treno è regionale.

Per determinare λ, ricordiamo che $\mathbb{E}(X_{\text{vel}}) = \frac{1}{\lambda}$, quindi $\frac{1}{\lambda} = 5$, da cui $\lambda = \frac{1}{5}$.

Per determinare μ, chiamiamo $a = \ln(3^{10})$ e imponiamo

$$\frac{2}{3} = \mathbb{P}(\{X_{\text{reg}} \le a\}) = \int_0^a \mu e^{-\mu x}\,\mathrm{d}x = \Big[-e^{-\mu x}\Big]_{x=0}^{x=a} = 1 - e^{-\mu\,a} \implies e^{-\mu\,a} = 1 - \frac{2}{3} = \frac{1}{3}$$
$$\implies -\mu\,a = \ln\left(\frac{1}{3}\right) \implies \mu = \frac{\ln(3)}{a} = \frac{\ln(3)}{\ln(3^{10})} = \frac{1}{10}\,.$$

1. Calcoliamo

$$\mathbb{P}(\{3 \le X_{\text{vel}} \le 7\}) = \int_3^7 \lambda e^{-\lambda x}\,\mathrm{d}x = \Big[e^{-\lambda x}\Big]_{x=3}^{x=7} = e^{-3\lambda} - e^{-7\lambda} = e^{-\frac{3}{5}} - e^{-\frac{7}{5}} \simeq 30.22\%\,.$$

2. Ricordiamo che $\text{Var}(X_{\text{reg}}) = \frac{1}{\lambda^2}$, quindi la deviazione standard è $\sqrt{\text{Var}(X_{\text{reg}})} = \frac{1}{\lambda} = 5$.
3. Consideriamo la variabile aleatoria

$$X = \text{“ritardo di un treno”}\,.$$

Per rispondere al quesito, utilizziamo il teorema della probabilità totale e utilizziamo il fatto di conoscere la distribuzione per i treni veloci e i treni regionali:

$$\begin{aligned}
\mathbb{P}(\{X \ge 8\}) &= \mathbb{P}(\{X \ge 8\}|\{Y = 1\})\mathbb{P}(\{Y = 1\}) + \mathbb{P}(\{X \ge 8\}|\{Y = 0\})\mathbb{P}(\{Y = 0\}) \\
&= \mathbb{P}(\{X_{\text{vel}} \ge 8\})\mathbb{P}(\{Y = 1\}) + \mathbb{P}(\{X_{\text{reg}} \ge 8\})\mathbb{P}(\{Y = 0\}) \\
&= e^{-8\lambda}\,20\% + e^{-8\mu}\,80\% = e^{-\frac{8}{5}}\,20\% + e^{-\frac{8}{10}}\,80\% \simeq 39.98\%\,.
\end{aligned}$$

4. Si perde la coincidenza se il ritardo del treno veloce è maggiore del ritardo del treno regionale, ovvero $\{X_{\text{vel}} > X_{\text{reg}}\}$. Per l'indipendenza, la densità congiunta è data da $f_{(X_{\text{vel}},X_{\text{reg}})}(x_1, x_2) = \lambda\mu e^{-\lambda x_1} e^{-\mu x_2}$ per $x_1 > 0$ e $x_2 > 0$, zero altrimenti. Allora, integrando sul dominio normale $\{x_1 > x_2,\ x_1 > 0,\ x_2 > 0\}$, otteniamo

$$\begin{aligned}
\mathbb{P}(\{X_{\text{vel}} > X_{\text{reg}}\}) &= \iint_{\{x_1 > x_2\}} f_{(X_{\text{vel}},X_{\text{reg}})}(x_1, x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \\
&= \int_0^{+\infty} \Big(\int_0^{x_1} \lambda\mu e^{-\lambda x_1} e^{-\mu x_2}\,\mathrm{d}x_2\Big)\,\mathrm{d}x_1 \\
&= \int_0^{+\infty} \lambda e^{-\lambda x_1}\Big(\int_0^{x_1} \mu e^{-\mu x_2}\,\mathrm{d}x_2\Big)\,\mathrm{d}x_1 = \\
&= \int_0^{+\infty} \lambda e^{-\lambda x_1}\Big(1 - e^{-\mu x_1}\Big)\,\mathrm{d}x_1 \\
&= \int_0^{+\infty} \lambda e^{-\lambda x_1}\,\mathrm{d}x_1 - \int_0^{+\infty} \lambda e^{-(\lambda+\mu)x_1}\,\mathrm{d}x_1 \\
&= 1 - \frac{\lambda}{\lambda+\mu} = \frac{\mu}{\lambda+\mu} = \frac{\frac{1}{10}}{\frac{1}{5}+\frac{1}{10}} = \frac{1}{3}\,.
\end{aligned}$$

Esercizio 4. (7 punti) Un'azienda farmaceutica produce tranquillanti. Si assuma che la distribuzione della durata della loro efficacia sia normale. Un laboratorio di ricerca di un ospedale ha sperimentato il farmaco su un campione casuale di 6 pazienti. Il periodo di efficacia del tranquillante per ogni paziente (in ore) è stato il seguente:

$$2.7 \quad 2.8 \quad 3.0 \quad 2.3 \quad 2.3 \quad 2.2\,.$$

1. Calcolare sui dati un intervallo di confidenza bilaterale al 90% per la media del periodo di efficacia.

2. Supponiamo che l'esperimento di sopra venga svolto 5 volte (in ogni esperimento vengono misurati i dati di 6 pazienti). Dopo ogni campionamento, viene calcolato un intervallo di confidenza bilaterale al 90% per la media come sopra. Qual è la probabilità che la media vera sia effettivamente nell'intervallo calcolato almeno 4 volte?

Soluzione. 1. Abbiamo un campione casuale $X_1, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ con $n = 6$. Sia μ che σ^2 sono incognite.

Dalla definizione di IC si ha che, ponendo $\beta = 90\%$,

$$\beta = \mathbb{P}(\{U_n \leq \mu \leq V_n\})\,.$$

Per stimare μ sfrutteremo lo stimatore media campionaria $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. Per stimare σ^2 sfrutteremo lo stimatore varianza campionaria $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$. Poiché $X_1, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ sono indipendenti, $T_{n-1} = \frac{\overline{X}_n - \mu}{S_n/\sqrt{n}} \sim t(n-1)$. Allora

$$\begin{aligned}
\beta = \mathbb{P}(\{U_n \leq \mu \leq V_n\}) &= \mathbb{P}\Big(\Big\{\frac{\overline{X}_n - V_n}{S_n/\sqrt{n}} \leq \frac{\overline{X}_n - \mu}{S_n/\sqrt{n}} \leq \frac{\overline{X}_n - U_n}{S_n/\sqrt{n}}\Big\}\Big) \\
&= \mathbb{P}\Big(\Big\{\frac{\overline{X}_n - V_n}{S_n/\sqrt{n}} \leq T_{n-1} \leq \frac{\overline{X}_n - U_n}{S_n/\sqrt{n}}\Big\}\Big) \\
&= 1 - \mathbb{P}\Big(\Big\{T_{n-1} < \frac{\overline{X}_n - V_n}{S_n/\sqrt{n}}\Big\}\Big) - \mathbb{P}\Big(\Big\{T_{n-1} > \frac{\overline{X}_n - U_n}{S_n/\sqrt{n}}\Big\}\Big)\,.
\end{aligned}$$

Segue che

$$\mathbb{P}\Big(\Big\{T_{n-1} < \frac{\overline{X}_n - V_n}{S_n/\sqrt{n}}\Big\}\Big) + \mathbb{P}\Big(\Big\{T_{n-1} > \frac{\overline{X}_n - U_n}{S_n/\sqrt{n}}\Big\}\Big) = 1 - \beta = \alpha = 10\%\,.$$

Decidiamo di equipartire α:

$$\mathbb{P}\Big(\Big\{T_{n-1} < \frac{\overline{X}_n - V_n}{S_n/\sqrt{n}}\Big\}\Big) = \mathbb{P}\Big(\Big\{T_{n-1} < \frac{\overline{X}_n - V_n}{S_n/\sqrt{n}}\Big\}\Big) = \frac{\alpha}{2} = 5\%\,. \tag{1}$$

Definiamo $t_{n-1,\alpha/2}$ come il punto tale che

$$\mathbb{P}(\{T_{n-1} \geq t_{n-1,\alpha/2}) = \frac{\alpha}{2}\,.$$

Scegliendo

$$\begin{aligned}
\frac{\overline{X}_n - U_n}{S_n/\sqrt{n}} &= t_{n-1,\alpha/2} \implies U_n = \overline{X}_n - \frac{S_n}{\sqrt{n}} t_{n-1,\alpha/2}\,, \\
\frac{\overline{X}_n - V_n}{S_n/\sqrt{n}} &= -t_{n-1,\alpha/2} \implies V_n = \overline{X}_n + \frac{S_n}{\sqrt{n}} t_{n-1,\alpha/2}\,,
\end{aligned}$$

si ottiene (1). In conclusione

$$\left[\overline{X}_n - \frac{S_n}{\sqrt{n}} t_{n-1,\alpha/2}, \overline{X}_n + \frac{S_n}{\sqrt{n}} t_{n-1,\alpha/2}\right]$$

è un intervallo di confidenza bilaterale per μ con livello di confidenza $\beta = 1 - \alpha$. Calcoliamolo sui dati. Dalla tavola della t-Student otteniamo che

$$t_{n-1,\alpha/2} = t_{5,0.05} = 2.015\,.$$

Calcoliamo media e varianza sul campione di dati:

$$\overline{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{6}(2.7 + 2.8 + 3.0 + 2.3 + 2.3 + 2.2) = 2.55\,,$$

$$s_n = \sqrt{\frac{1}{n-1}\Big(\sum_{i=1}^{n} x_i^2 - n\overline{x}_n^2\Big)} = \sqrt{\frac{1}{5}\Big(2.7^2 + 2.8^2 + 3.0^2 + 2.3^2 + 2.3^2 + 2.2^2 - 6 \cdot 2.55^2\Big)}$$
$$= \sqrt{0.107} \simeq 0.33\,.$$

Segue che

$$\left[\overline{x}_n - \frac{s_n}{\sqrt{n}} t_{n-1,\alpha/2}, \overline{x}_n + \frac{s_n}{\sqrt{n}} t_{n-1,\alpha/2}\right] \simeq [2.27, 2.82]\,.$$

2. Consideriamo una variabile aleatoria

$$Y = \text{“numero di volte che la media è nell'IC al 90\% su 5 tentativi”} \sim \mathrm{B}(5, 90\%)\,.$$

Allora

$$\mathbb{P}(\{Y \geq 4\}) = \mathbb{P}(\{Y = 4\}) + \mathbb{P}(\{Y = 5\}) = \binom{5}{4}(90\%)^4 10\% + \binom{5}{5}(90\%)^5 = 91.854\%\,.$$