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

A.A.: 2020/2021
Docente: Gianluca Orlando
Appello: VI Appello
Data: 13/09/2021

Esercizio 1. Alice e Bob sono a cena e hanno a disposizione una moneta (con facce “testa” e “croce”) per decidere chi deve pagare il conto. Alice sa però che la moneta *non* è bilanciata: “testa” esce con probabilità $\frac{1}{3}$. Ma è onesta e propone a Bob un gioco equo, con cui i due hanno la stessa probabilità di vincere. Il gioco è basato sulla scommessa sull'esito di due lanci consecutivi della moneta. Per descriverlo, si risolvano i seguenti quesiti:

1. Si consideri l'esperimento che consiste nel lanciare due volte consecutivamente la moneta non bilanciata. Si descriva lo spazio campione Ω e si elenchi la probabilità dei suoi elementi (gli eventi elementari).

2. Ci sono eventi elementari equiprobabili in Ω?

3. Basandosi sulla risposta precedente, descrivere il gioco equo proposto da Alice. (Suggerimento: alcuni risultati dell'esperimento possono essere scartati!)

4. Si consideri ora lo stesso gioco con una moneta non bilanciata per cui “testa” esce con probabilità p, dove $p \in [0, 1]$. Per quali valori di p il gioco nel punto 3 non può funzionare?

Soluzione. Denotiamo con T l'esito “testa” del lancio della moneta e con C l'esito “croce”.

1. Lo spazio campione è dato dalle possibili coppie ordinate

$$\Omega = \{(T, T), (C, C), (T, C), (C, T)\} .$$

Per l'ipotesi sulla moneta abbiamo che $\mathbf{P}(\text{“}T\text{ esce al primo lancio”}) = \mathbf{P}(\{(T, T), (T, C)\}) = \frac{1}{3}$ e analogamente $\mathbf{P}(\text{“}T\text{ esce al secondo lancio”}) = \mathbf{P}(\{(T, T), (C, T)\}) = \frac{1}{3}$. Analogamente per gli esiti su “croce”: $\mathbf{P}(\text{“}C\text{ esce al primo lancio”}) = \mathbf{P}(\{(C, C), (C, T)\}) = \frac{2}{3}$ e analogamente $\mathbf{P}(\text{“}C\text{ esce al secondo lancio”}) = \mathbf{P}(\{(C, C), (T, C)\}) = \frac{2}{3}$. Possiamo supporre che i due lanci siano indipendenti. La probabilità degli eventi elementari è quindi data da

$$\mathbf{P}(\{(T, T)\}) = \mathbf{P}(\text{“}T\text{ esce al primo lancio”})\mathbf{P}(\text{“}T\text{ esce al secondo lancio”}) = \frac{1}{3}\frac{1}{3} = \frac{1}{9} ,$$

$$\mathbf{P}(\{(C, C)\}) = \mathbf{P}(\text{“}C\text{ esce al primo lancio”})\mathbf{P}(\text{“}C\text{ esce al secondo lancio”}) = \frac{2}{3}\frac{2}{3} = \frac{4}{9} ,$$

$$\mathbf{P}(\{(T, C)\}) = \mathbf{P}(\text{“}T\text{ esce al primo lancio”})\mathbf{P}(\text{“}C\text{ esce al secondo lancio”}) = \frac{1}{3}\frac{2}{3} = \frac{2}{9} ,$$

$$\mathbf{P}(\{(C, T)\}) = \mathbf{P}(\text{“}C\text{ esce al primo lancio”})\mathbf{P}(\text{“}T\text{ esce al secondo lancio”}) = \frac{2}{3}\frac{1}{3} = \frac{2}{9} .$$

2. Gli eventi $\{(T,C)\}$ e $\{(C,T)\}$ sono equiprobabili.

3. Alice e Bob possono scommettere sull'esito dei due lanci, ma solo sugli eventi equiprobabili $\{(T,C)\}$ e $\{(C,T)\}$.

- Vengono effettuati i due lanci: il risultato è un elemento di Ω.
- Se il risultato dei due lanci è $\{(T,T)\}$ oppure $\{(C,C)\}$ il risultato viene scartato e si ricomincia il gioco.
- Altrimenti il risultato è $\{(T,C)\}$ oppure $\{(C,T)\}$. Vince chi ha scommesso sul risultato realizzato.

In questo modo Alice e Bob hanno la stessa probabilità di vincere il gioco.

4. Se $\mathbf{P}(\text{“}T \text{ esce al primo lancio”}) = p$, ripetendo i conti del punto 1 si osserva che la probabilità degli eventi equiprobabili è $p(p-1)$. Se $p=0$ oppure $p=1$, gli eventi $\{(T,C)\}$ e $\{(C,T)\}$ sono impossibili, quindi il gioco non termina mai.

Esercizio 2. Siano dati gli eventi A e B di uno spazio campione Ω. È noto che $\mathbf{P}(A\cup B) = \frac{3}{4}$ e $\mathbf{P}(A\cap B) = \frac{1}{4}$.

1. È possibile ricavare i valori numerici di $\mathbf{P}(A)$ e di $\mathbf{P}(B)$? In caso di risposta affermativa, indicarli.
2. Se gli eventi A e B sono indipendenti, è possibile ricavare i valori numerici di $\mathbf{P}(A)$ e di $\mathbf{P}(B)$? In caso di risposta affermativa, indicarli.
3. Cosa si può dire del caso in cui $\mathbf{P}(A|B) = \frac{1}{8}$?

Soluzione. Denotiamo con $a = \mathbf{P}(A)$ e $b = P(B)$ le incognite del problema.
1. Utilizziamo la formula per la probabilità dell'unione di eventi non mutuamente esclusivi:

$$\mathbf{P}(A\cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A\cap B)$$

da cui ricaviamo che

$$\frac{3}{4} = a + b - \frac{1}{4}$$

cioè

$$a + b = \frac{1}{2}\,.$$

Poiché questa è una equazione lineare in due incognite, i valori a e b non sono determinati in modo univoco.

2. Se gli eventi A e B sono indipendenti, per definizione si ha che

$$\mathbf{P}(A\cap B) = \mathbf{P}(A)\mathbf{P}(B)\,,$$

da cui

$$ab = \frac{1}{4}\,.$$

Per determinare a e b risolviamo quindi il sistema

$$\begin{cases} a+b=1 \\ ab=\dfrac{1}{4} \end{cases} \implies \begin{cases} a+\dfrac{1}{4a}=1 \\ b=\dfrac{1}{4a} \end{cases} \implies \begin{cases} 4a^2-4a+1=0 \\ b=\dfrac{1}{4a} \end{cases} \implies \begin{cases} (2a-1)^2=0 \\ b=\dfrac{1}{4a} \end{cases}$$

la cui soluzione è data da $a=\frac{1}{2}$ e $b=\frac{1}{2}$.

3. Dalla definizione di probabilità condizionata si ha che

$$\mathbf{P}(A\cap B)=\mathbf{P}(A|B)\mathbf{P}(B)$$

cioè

$$\frac{1}{4}=\frac{1}{8}b \implies b=2\,.$$

Questo valore non è ammissibile perché $b=\mathbf{P}(B)\in[0,1]$. Pertanto la condizione $\mathbf{P}(A|B)=\frac{1}{8}$ è incompatibile con l'ipotesi $\mathbf{P}(A\cap B)=\frac{1}{4}$.

Esercizio 3. Una variabile aleatoria continua X assume valori nell'intervallo $[-\frac{\pi}{2},\frac{\pi}{2}]$ e ha come Funzione di Distribuzione Cumulativa (FDC) una funzione continua $F\colon\mathbb{R}\to\mathbb{R}$ che verifica

$$F(x)=a\sin(x)+b \quad \text{per ogni } x\in\Big[-\frac{\pi}{2},\frac{\pi}{2}\Big]\,,$$

dove $a,b\in\mathbb{R}$.

1. Si determinino i valori a e b in modo che $F(x)$ soddisfi le proprietà di Funzione di Distribuzione Cumulativa. Si descriva la funzione F anche fuori dall'intervallo $[-\frac{\pi}{2},\frac{\pi}{2}]$.
2. Calcolare $\mathbf{P}(0\le X\le 6)$.
3. Determinare la funzione di densità di probabilità (f.d.p.) $f\colon\mathbb{R}\to\mathbb{R}$ di X.
4. Calcolare $\mathbf{E}(X)$.

Soluzione.
1. La Funzione di Distribuzione Cumulativa è una funzione continua e monotona crescente a valori nell'intervallo $[0,1]$. Poiché X assume valori nell'intervallo $[-\frac{\pi}{2},\frac{\pi}{2}]$, si ha che

$$0=\mathbf{P}\Big(X\le-\frac{\pi}{2}\Big)=F\Big(-\frac{\pi}{2}\Big)=-a+b\,.$$

Inoltre $\mathbf{P}(X>\frac{\pi}{2})=0$, da cui segue che $\mathbf{P}(X\le\frac{\pi}{2})=1-\mathbf{P}(X>\frac{\pi}{2})=1$. Pertanto

$$1=\mathbf{P}\Big(X\le\frac{\pi}{2}\Big)=F\Big(\frac{\pi}{2}\Big)=a+b\,.$$

Possiamo allora risolvere il sistema

$$\begin{cases} -a+b=0 \\ a+b=1 \end{cases} \implies \begin{cases} a=b \\ 2a=1 \end{cases}$$

la cui soluzione è data da $a=b=\frac{1}{2}$. Quindi la FDC è data da

$$F(x)=\begin{cases} 0 & \text{per } x<-\frac{\pi}{2}\,, \\ \frac{1}{2}\sin(x)+\frac{1}{2} & \text{per } -\frac{\pi}{2}\le x\le\frac{\pi}{2}\,, \\ 1 & \text{per } x>\frac{\pi}{2}\,. \end{cases}$$

Osserviamo che la funzione F assume valori nell'intervallo $[0,1]$ ed è monotona crescente.

2. Poiché X assume valori nell'intervallo $[-\frac{\pi}{2}, \frac{\pi}{2}]$, si ha che $\mathbf{P}(X > \frac{\pi}{2}) = 0$. In particolare $\mathbf{P}(\frac{\pi}{2} < X \leq 6) \leq \mathbf{P}(X > \frac{\pi}{2}) = 0$ per monotonia. Poiché gli eventi $\{0 \leq X \leq \frac{\pi}{2}\}$ e $\{\frac{\pi}{2} < X \leq 6\}$ sono mutuamente esclusivi, abbiamo che

$$\begin{aligned}\mathbf{P}(0 \leq X \leq 6) &= \mathbf{P}\Big(\Big\{0 \leq X \leq \frac{\pi}{2}\Big\} \cup \Big\{\frac{\pi}{2} < X \leq 6\Big\}\Big) = \mathbf{P}\Big(0 \leq X \leq \frac{\pi}{2}\Big) + \mathbf{P}\Big(\frac{\pi}{2} < X \leq 6\Big) \\ &= \mathbf{P}\Big(0 \leq X \leq \frac{\pi}{2}\Big) = F\Big(\frac{\pi}{2}\Big) - F(0) = 1 - \frac{1}{2} = \frac{1}{2}\,.\end{aligned}$$

3. La funzione di densità di probabilità è la derivata della Funzione di Distribuzione Cumulativa, $f(x) = F'(x)$. Quindi

$$f(x) = \begin{cases} 0 & \text{per } x < -\frac{\pi}{2}\,, \\ \frac{1}{2}\cos(x) & \text{per } -\frac{\pi}{2} \leq x \leq \frac{\pi}{2}\,, \\ 0 & \text{per } x > \frac{\pi}{2}\,. \end{cases}$$

4. Per calcolare il valore atteso utilizziamo il punto 3 e la formula

$$\mathbf{E}(X) = \int_{-\infty}^{+\infty} x f(x)\,\mathrm{d}x = \frac{1}{2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} x\cos(x)\,\mathrm{d}x = 0\,,$$

poiché $x\cos(x)$ è una funzione dispari.

Esercizio 4. In un certo procedimento chimico, è di fondamentale importanza che il pH di uno dei reagenti sia 8.20. Si sa che il metodo per misurare il pH fornisce valori con distribuzione normale con media pari al valore autentico e deviazione standard 0.02. Supponiamo che 10 misurazioni indipendenti abbiano dato i seguenti valori:

$$8.18 \quad 8.16 \quad 8.17 \quad 8.22 \quad 8.19 \quad 8.17 \quad 8.15 \quad 8.21 \quad 8.16 \quad 8.18\,.$$

1. Verificare l'ipotesi che il pH sia 8.20 con un livello di significatività del 5%.
2. Verificare l'ipotesi che il pH sia 8.20 con un livello di significatività del 10%.

Soluzione. La variabile aleatoria in esame X ha distribuzione normale con media μ e varianza $\sigma^2 = 0.0004$. Viene estratto un campione $(X_1, \ldots, X_n)$ composto da $n = 10$ elementi. Vogliamo verificare la seguente ipotesi:

$$H_0 : \mu = 8.20\,, \quad H_1 : \mu \neq 8.20\,.$$

Consideriamo la media campionaria

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i\,.$$

Poiché ogni X_i ha distribuzione normale, per le proprietà di linearità anche $\overline{X}$ ha distribuzione normale. Utilizzeremo la statistica test

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$$

per verificare il test d'ipotesi.
Denotiamo con z_α il valore di Z per cui l'area sottesa alla curva normale standard a destra di tale valore sia uguale ad α. Osserviamo che, per simmetria,

$$\alpha = \frac{\alpha}{2} + \frac{\alpha}{2} = \mathbf{P}(Z < -z_{\alpha/2}) + \mathbf{P}(Z > z_{\alpha/2}) = \mathbf{P}(|Z| > z_{\alpha/2}) = \mathbf{P}\Big(\Big|\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}\Big| > z_{\alpha/2}\Big).$$

Il test con livello di significatività α accetta l'ipotesi H_0 se la realizzazione di Z ricade nell'intervallo

$$[-z_{\alpha/2}, z_{\alpha/2}].$$

In questo caso, la realizzazione della media campionaria è $\overline{X} = 8.179$, quindi, assumendo H_0 vera, $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} = \frac{8.179-8.20}{0.02/\sqrt{10}} = -3.32$.

1. Se il livello di significatività è $\alpha = 5\% = 0.05$, e $\alpha/2 = 0.025$, si ha che $z_{0.025}$ è il numero tale che

$$1 - 0.025 = 0.975 = \Phi(z_{0.025}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_{0.025}} e^{-x^2/2}\,\mathrm{d}x.$$

Consultando le tabelle si ha che $z_{0.025} = 1.96$. Poiché $Z = -3.32$ non è incluso nell'intervallo $[-1.96, 1.96]$, l'ipotesi H_0 è rifiutata.

2. Se il livello di significatività è $\alpha = 10\% = 0.1$, la regione di accettazione è più piccola rispetto al punto precedente, e quindi l'ipotesi va rifiutata. Infatti, per $\alpha/2 = 0.05$ si ha che $z_{0.05} \sim \frac{1.64+1.65}{2} = 1.645$ e $Z = -3.32$ non appartiene all'intervallo $[-1.645, 1.645]$.

Esercizio 5. Una sostanza metallica viene esposta ad un'atmosfera di ossigeno puro. L'aumento relativo di massa della sostanza viene utilizzato come indicatore della quantità di ossigeno che ha reagito. I dati raccolti sono i seguenti:

ore di esposizione	1.0	2.0	2.5	3.0	3.5	4.0
incremento percentuale	0.02	0.03	0.035	0.042	0.05	0.054

1. Trovare la retta di regressione che interpola i dati e disegnarla.
2. Calcolare il coefficiente di correlazione.

Soluzione. Denotiamo con x_i i valori dati dalle ore di esposizione e con y_i i valori dell'incremento percentuale. Introduciamo le medie

$$\overline{x} = \frac{1}{6}\sum_{i=1}^{6} x_i = 8/3 \sim 2.6667 \qquad \overline{y} = \frac{1}{6}\sum_{i=1}^{6} y_i = 0.0385.$$

1. Per trovare i coefficienti α e β della retta di regressione utilizziamo il metodo dei minimi quadrati, cioè minimizziamo la quantità

$$\sum_{i=1}^{6} |y_i - \alpha x_i - \beta|^2.$$

Imponendo che le derivate rispetto ad α e β si annullino, otteniamo che

$$\begin{cases} \displaystyle\sum_{i=1}^{6}(y_i - \alpha x_i - \beta)x_i = 0 \\ \displaystyle\sum_{i=1}^{6}(y_i - \alpha x_i - \beta) = 0\,. \end{cases}$$

Dalla seconda equazione otteniamo:

$$\hat{\beta} = \overline{y} - \alpha\overline{x}$$

e quindi, dalla prima

$$\sum_{i=1}^{6}(y_i - \hat{\alpha}x_i - \overline{y} + \hat{\alpha}\overline{x})x_i = 0 \implies \hat{\alpha} = \frac{\sum_{i=1}^{6} x_i y_i - 6\,\overline{y}\,\overline{x}}{\sum_{i=1}^{6} x_i^2 - 6\overline{x}^2} = \frac{0.6845 - 0.616}{48.5 - 128/3} \sim 0.0117\,.$$

Pertanto

$$\hat{\beta} = \overline{y} - \hat{\alpha}\overline{x} \sim 0.0073\,.$$

La retta di regressione è

$$0.0117x + 0.0073\,.$$

2. Per calcolare il coefficiente di correlazione lineare utilizziamo la formula

$$\rho = \hat{\alpha}\frac{S_X}{S_Y}\,.$$

Calcoliamo

$$5S_X^2 = \sum_{i=1}^{6} x_i^2 - 6\overline{x}^2 = 48.5 - 128/3$$

$$5S_Y^2 = \sum_{i=1}^{6} y_i^2 - 6\overline{y}^2 = 0.0008115$$

quindi

$$\rho = \hat{\alpha}\frac{S_X}{S_Y} \sim 0.0117\sqrt{\frac{5.8\overline{3}}{0.0008115}} \sim 0.992\,.$$