Calcolo delle Probabilità e Statistica, Ingegneria Civile e A&T e Informatica
Prova scritta del 15 Settembre 2023

Punteggi: **1:** 3×3; **2:** $4 + 4 + 5$; **3:** $4 + 4$.

1. Due roulette regolari vengono azionate più volte; sia T il numero di volte che occorre azionare la prima roulette per ottenere per la prima volta un multiplo di 12, ed S il numero di volte che occorre azionare la seconda roulette per ottenere per la prima volta un multiplo di 7 (si ricordi che una roulette regolare fornisce ad ogni azione un numero $N \in \{0, \ldots, 36\}$ con $P(N = i) = p,\ i = 0, 1, \ldots, 36$).
(i) Trovare le leggi delle v.a. T ed S e calcolare $E(T)$ ed $E(-3T + 5S)$. Si può ritenere che S e T siano v.a. indipendenti?
(ii) Calcolare $P(T - 2S \geq 0)$.
(iii) Si consideri una terza roulette, truccata in modo che sia uguale a $q = 10^{-3}$ la probabilità che esca il numero 0, ogni volta che viene azionata. Se la si aziona 6000 volte e si indica con Z il numero delle volte che esce 0, stimare $P(Z \geq 2)$.

2. Sia (X, Y) un vettore aleatorio tale che Y sia esponenziale di parametro λ mentre la densità condizionale di X dato $\{Y = y\}$ ha la seguente espressione

$$f_{X|Y}(x|y) = ye^{-yx} I_{(0,+\infty)}(x).$$

(i) Calcolare la legge della variabile aleatoria $Z = XY$. Si tratta di una legge nota?
(ii) Calcolare la densità di X.
(iii) Calcolare la densità condizionale di Y dato $\{X = x\}$ e $E(Y|X = 1)$.

3. Una compagnia di assicurazioni vuole condurre una indagine statistica per stimare l'indennizzo medio pagato a seguito di incidenti domestici. Analisi pregresse mostrano che tali importi possono essere assunti gaussiani con media μ incognita e deviazione standard nota e pari a 400 euro. Su un campione casuale di 144 incidenti è stato osservato un indennizzo medio pari a 6230 euro.

(i) Determinare un intervallo di confidenza di livello 95% per il parametro μ.

(ii) Determinare la dimensione minima del campione necessaria affinché l'ampiezza dell'intervallo di confidenza di livello 95% per il parametro μ non superi i 160 euro.

SOLUZIONI della prova scritta del 15-09-2023

1. (i) Siccome i multipli di 12 tra 0 e 36 sono 3 e i multipli di 7 tra 0 e 36 sono 5, T ed S sono istanti di primo successo in serie di prove di Bernoulli indipendenti, in cui la probabilità del successo in ogni prova è $p = \frac{3}{37}$, rispettivamente $q = \frac{5}{37}$. Dunque, T ed S hanno distribuzione geometrica modificata di parametro p e q, rispettivamente. Pertanto:

$$P(T = k) = p(1-p)^{k-1}, \ k = 1, \dots, \ P(S = h) = q(1-q)^{h-1}, \ h = 1, \dots$$

Si ha $E(T) = 1/p = \frac{37}{3}$ e $E(S) = 1/q = \frac{37}{5}$, dunque $E(-3T + 5S) = -3E(T) + 5E(S) = -3 \cdot \frac{37}{3} + 5 \cdot \frac{37}{5} = 0$. Le v.a. S e T possono ritenersi indipendenti.
(ii)

$$P(T - 2S \geq 0) = P(T \geq 2S) = \sum_{k=1}^{\infty} P(T \geq 2k, S = k)$$

siccome S e T sono indipendenti, tale probabilità vale:

$$\sum_{k=1}^{\infty} P(T \geq 2k) P(S = k)$$

Ora $T - 1 = X$ ha distribuzione geometrica di parametro p; ricordando che per una tale X, risulta $P(X \geq h) = (1-p)^h$, si ha $P(T \geq 2k) = P(T - 1 \geq 2k - 1) = P(X \geq 2k - 1) = (1-p)^{2k-1}$. Dunque, riprendendo il calcolo

$$P(T - 2S \geq 0) = \sum_{k=1}^{\infty} (1-p)^{2k-1} P(S = k) = \sum_{k=1}^{\infty} (1-p)^{2k-1} q(1-q)^{k-1} =$$

$$= \frac{q}{(1-p)(1-q)} \sum_{k=1}^{\infty} \left[(1-p)^2(1-q)\right]^k = \frac{q}{(1-p)(1-q)} \left(\frac{1}{1 - (1-p)^2(1-q)} - 1 \right) =$$

$$= \frac{q(1-p)}{1 - (1-p)^2(1-q)}$$

(iii) Per l'approssimazione di Poisson, si ha $P(Z = k) \approx P(W = k)$, ove W ha distribuzione di Poisson di parametro $\lambda = nq = 6000/1000 = 6$. Pertanto $P(Z \geq 2) = 1 - P(Z = 0) - P(Z = 1) \approx 1 - P(W = 0) - P(W = 1) = 1 - e^{-6} - 6e^{-6} = 1 - 7e^{-6}$.

2. È necessario notare che a densità congiunta del vettore (X, Y) ha l'espressione

$$f_{X,Y}(x, y) = \lambda y e^{-\lambda y} e^{-yx} I_{(0,+\infty)}(x) I_{(0,+\infty)}(y).$$

(i) Sia $Z = XY$. Allora, per ogni $t > 0:$

$$F_Z(z) = P(Z \leq z) = P\left(X \leq \frac{z}{Y}\right) =$$

$$= \int_0^{+\infty} dy \int_0^{\frac{t}{y}} dx\ \lambda y e^{-yx} e^{-\lambda y} =$$

$$= \int_0^{+\infty} \lambda e^{-\lambda y} dy \int_0^{\frac{t}{y}} dx\ y e^{-yx} =$$

$$= (1 - e^{-z}) \int_0^{+\infty} \lambda e^{-\lambda y} dy = (1 - e^{-z}).$$

Dunque, $Z \sim Exp(1))$.

(ii) La densità di X è nulla per $x \leq 0$, mentre per $x > 0$ si ha

$$f_X(x) = \int_0^{+\infty} \lambda y e^{-yx} e^{-\lambda y} dy = \frac{\lambda}{(x+\lambda)} \int_0^{+\infty} (x+\lambda) y e^{-(x+\lambda)y} dy =$$

$$= \frac{\lambda}{(x+\lambda)^2},$$

essendo l'integrale al secondo membro l'espressione del valore medio di una variabile aleatoria esponenziale di parametro $(x + \lambda)$.

(iii) La densità condizionale di Y dato $\{X = x\}$, $x > 0$ è nulla per $y \leq 0$, mentre per $y > 0$ si ha

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{\lambda y e^{-\lambda y} e^{-yx}}{\frac{\lambda}{(x+\lambda)^2}} = (x+\lambda)^2 y e^{-(\lambda+x)y},$$

ovvero è una densità $\Gamma(2, (\lambda + x))$. Di conseguenza $E(Y|X = 1) = \frac{2}{\lambda+1}$.

3. (i) Un intervallo di confidenza di livello $1 - \alpha$ per la media incognita μ di una distribuzione gaussiana con varianza σ^2 nota ha l'espressione:

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

dove n è la dimensione del campione, $\phi_{1-\frac{\alpha}{2}}$ è il quantile di ordine $1 - \frac{\alpha}{2}$ della distribuzione gaussiana standard, e $\overline{X}_n$ è la media campionaria.

Sostituendo i dati forniti dal problema si ottiene l'intervallo

$$\left[6230 - \frac{400}{12} \phi_{0.975}, 6230 + \frac{400}{12} \phi_{0.975} \right] = \left[6230 - \frac{196}{3}, 6230 + \frac{196}{3} \right] = [6164.67, 6295.3] \,.$$

(ii) L'ampiezza dell' intervallo di confidenza in esame è $2 \frac{\sigma}{\sqrt{n}} \phi_{1-\frac{\alpha}{2}}$. Pertanto dobbiamo trovare il più piccolo n che soddisfa la disequazione

$$2 \frac{\sigma}{\sqrt{n}} \phi_{0.975} \leq 160 \qquad \sqrt{n} \geq \frac{\sigma}{80} \phi_{0.975} = \frac{400}{80} 1.96 = 9.8$$

la cui soluzione è $n \geq 96.04$. Pertanto è sufficiente considerare un campione di dimensione $n = 97$.