Statistica - CPS
Corso di Laurea in Informatica
Compito del 11-01-2023

Esercizio 1. (10 punti) Ci sono due scatole, contenenti ciascuna 10 HD, dello stesso tipo e usati: la prima scatola contiene 2 HD difettosi e la seconda 5 HD difettosi.

(i) Per effettuare le riparazioni, si prelevano (in blocco) 3 HD dalla prima scatola. Indicando con X il numero di HD funzionanti, calcolare $E[X]$.

(ii) Supponiamo di scegliere a caso, in maniera equiprobabile, una scatola e di prelevare (in blocco) da essa 3 HD che risultano tutti funzionanti: qual è la probabilità che sia stata scelta la prima scatola?

Esercizio 2. (10 punti) Si consideri una v.a. X avente funzione di ripartizione così definita:

$$F(x) = \begin{cases} 0, & \text{se } x < 0; \\ a\,x + b\,x^3, & \text{se } 0 \le x \le 1; \\ 1, & \text{se } x > 1; \end{cases}$$

(dove a e b sono due costanti reali, non necessariamente positive).

(i) Stabilire delle relazioni tra le costanti a e b che garantiscano che la funzione sopra scritta sia la funzione di ripartizione di una variabile con densità.

(ii) Determinare le costanti a e b in modo tale che si abbia $E\big[X\big] = 0{,}5$. Quali momenti finiti ha la v.a. X con funzione di ripartizione $F(x)$?

(iii) Scrivere la densità della v.a. $Y = e^X$.

Esercizio 3. (10 punti) Una grossa ditta vuole lanciare un nuovo tipo di smartphone, prodotto con una tecnologia innovativa: i costi per lanciare su larga scala la produzione però sono tali che il lancio può essere finanziariamente conveniente solo se almeno il 20% dei potenziali acquirenti gradisce il nuovo prodotto. Viene prodotto un prototipo offerto in prova a 10000 persone e di queste la percentuale che dichiara di gradirlo è del 19,4%.

(i) Nonostante questa percentuale sia (leggermente) inferiore al 20%, il capo del progetto sostiene che il gradimento non è inferiore alla soglia desiderata e a tale scopo imposta il test dell'ipotesi $H_0\big)\ p \ge 0,2$ contro $H_1\big)\ p < 0,2$. Dopo aver calcolato il relativo *p-value*, qual è la conclusione a cui si arriva?

(ii) Se 0.194 è la stima della percentuale di gradimento, qual è la *precisione di questa stima* con un livello di fiducia dell'80% ?

Esercizio 1) Indichiamo con B_1 e B_2 le due scatole e con D e F rispettivamente, un HD difettoso e funzionante. Sappiamo che

$$B_1 = \{D, D, F, F, F, F, F, F, F, F\}$$
$$B_2 = \{D, D, D, D, D, F, F, F, F, F\}$$

(i) Se X è la v.a. che conta il numero di HD funzionanti estratti da B_1 prelevandone 3 in blocco, X assume valori interi tra 1 e 3 (non può assumere il valore 0 perché B_1 contiene solo 2 D) e si ha

$$\mathbb{P}(X=h) = \frac{\binom{8}{h}\binom{2}{3-h}}{\binom{10}{3}} \quad \text{se } h=1,2,3 \; ; \quad \mathbb{P}(X=h)=0 \text{ , alt.}$$

quindi

$$\mathbb{P}(X=1) = \frac{\binom{8}{1}\binom{2}{2}}{\binom{10}{3}} = \frac{8 \cdot 3!}{10 \cdot 9 \cdot 8} = \frac{1}{15}$$

$$\mathbb{P}(X=2) = \frac{\binom{8}{2}\binom{2}{1}}{\binom{10}{3}} = \frac{8 \cdot 7 \cdot 2 \cdot 3!}{2! \cdot 10 \cdot 9 \cdot 8} = \frac{7}{15}$$

$$\mathbb{P}(X=3) = \frac{\binom{8}{3}\binom{2}{0}}{\binom{10}{3}} = \frac{8 \cdot 7 \cdot 6 \cdot 3!}{3! \cdot 10 \cdot 9 \cdot 8} = \frac{7}{15}$$

Di conseguenza

$$\boxed{\begin{aligned}\mathbb{E}[X] &= 1 \cdot \mathbb{P}(X=1) + 2 \cdot \mathbb{P}(X=2) + 3 \cdot \mathbb{P}(X=3) = \\ &= \frac{1+14+21}{15} = \frac{12}{5}\end{aligned}}$$

(ii) Sia $A = \{$si prelevano 3 HD funzionanti$\}$, vogliamo calcolare $\mathbb{P}(B_1 | A)$.

Applicando la formula di probabilità delle cause abbiamo

$$\mathbb{P}(B_1 | A) = \frac{\mathbb{P}(A | B_1)\, \mathbb{P}(B_1)}{\mathbb{P}(A | B_1)\mathbb{P}(B_1) + \mathbb{P}(A | B_2)\mathbb{P}(B_2)}$$

Utilizzando

$$\mathbb{P}(B_1) = \mathbb{P}(B_2) = \tfrac{1}{2}$$

$\mathbb{P}(A | B_1) = 7/15$ (dal punto (i))

$$\mathbb{P}(A | B_2) = \frac{\binom{5}{3}\binom{5}{0}}{\binom{10}{3}} = \frac{5 \cdot 4 \cdot 3 \cdot 3!}{3! \cdot 10 \cdot 9 \cdot 8} = \frac{1}{12}$$

si trova

$$\boxed{\mathbb{P}(B_1 / A) = \frac{7/15 \cdot \frac{1}{2}}{\frac{7}{15} \cdot \frac{1}{2} + \frac{1}{12} \cdot \frac{1}{2}} = \frac{28}{28 + 5} = \frac{28}{33}}$$

Esercizio 2 Sia

$$F(x) = \begin{cases} 0, & x < 0 \\ ax + bx^3, & 0 \leq x \leq 1 \\ 1, & x > 1 \end{cases} \qquad \text{con } a, b \in \mathbb{R}$$

(i) Affinché $F(x)$ sia la funzione di ripartizione di una v.a. con densità devono essere verificate le seguenti condizioni:

(a) $\lim_{x \to -\infty} F(x) = 0$; $\lim_{x \to +\infty} F(x) = 1$

(b) F è debolmente crescente

(c) F è continua

Iniziamo con la condizione (c). Deve valere:

$$F(0) = \lim_{x \to 0^-} F(x) \iff 0 = 0 \ \checkmark$$

$$F(1) = \lim_{x \to 1^+} F(x) \iff \boxed{a + b = 1}$$

$F|_{(-\infty, 0)}$, $F|_{(0,1)}$, $F|_{(1, +\infty)}$ è continua $\checkmark$

La condizione (a) è verificata. Perché sia verificata la condizione (b) deve valere:

$F|_{(-\infty,0)}$, $F|_{(1,+\infty)}$ è debolmente crescente ✓

$F|_{(0,1)}$ è debolmente crescente $\Leftrightarrow$ $F'|_{(0,1)} \geq 0$ $\Leftrightarrow$ $a + 3bx^2 \geq 0 \ \forall x \in (0,1)$

da cui si ricava che $\boxed{a \geq 0}$ e $\boxed{a + 3b \geq 0}$

$F(0) \geq F|_{(-\infty,0)}$ $\Leftrightarrow$ $0 \geq 0$ ✓

$F(1) \leq F|_{(1,+\infty)}$ $\Leftrightarrow$ $a+b \leq 1$ ✓ (abbiamo imposto già $a+b=1$).

Mettendo insieme le condizioni trovate, abbiamo che deve valere

$$\boxed{0 \leq a \leq \frac{3}{2} \quad e \quad b = 1-a}$$

(ii) La densità di una v.a. con funzione di ripartizione $F(x)$ è

$$f(x) = \begin{cases} a + 3bx^2, & 0 \leq x \leq 1 \\ 0, & \text{altrimenti} \end{cases}$$

quindi

$$E[X] = \int_{-\infty}^{+\infty} x f(x)\, dx = \int_0^1 x\,(a + 3bx^2)\, dx = \left(\frac{1}{2}ax^2 + \frac{3}{4}bx^4\right)\Big|_0^1 =$$

$$= \frac{1}{2}a + \frac{3}{4}b = \frac{1}{2}a + \frac{3}{4}(1-a) = \frac{3}{4} - \frac{1}{4}a$$

Allora $E[X] = \frac{1}{2}$ $\Leftrightarrow$ $\frac{3}{4} - \frac{1}{4}a = \frac{1}{2}$ $\Leftrightarrow$ $\boxed{a = 1 \quad e \quad b = 0}$

La v.a. X ha momento n-simo finito se $E[X^n] < +\infty$. Poiché X assume valori in $[0,1]$ e la densità f è limitata si ha che tutti i suoi momenti sono finiti. In particolare

$$E[X^n] = \int_{-\infty}^{+\infty} x^n f(x)\, dx = \int_0^1 (ax^n + 3bx^{n+2})\, dx =$$

$$= \left(\frac{a}{n+1}x^{n+1} + \frac{3b}{n+3}x^{n+3}\right)\Big|_0^1 = \frac{a}{n+1} + \frac{3(1-a)}{n+3}$$

(iii) La v.a. $Y = e^X$ si scrive come $Y = h(X)$ con $h(t) = e^t$.

La funzione h ha dominio $\mathbb{R}$ ed è iniettiva e derivabile, con inversa $h^{-1}(s) = \log s$ derivabile. Quindi

$$f_Y(y) = \begin{cases} f(h^{-1}(y)) \left| \frac{d}{dy} h^{-1}(y) \right|, & y \in h(\mathrm{Imm}\, X) \\ 0, & \text{altrimenti} \end{cases}$$

Usando che $\mathrm{Imm}\, X = [0,1]$ e $h([0,1]) = [1,e]$, e che $\frac{d}{dy} h^{-1}(y) = \frac{1}{y}$, si ha quindi che

$$\boxed{f_Y(y) = \begin{cases} \left[a + 3b(\log y)^2\right]/y, & \text{se } 1 \le y \le e \\ 0, & \text{altrimenti} \end{cases}}$$

Esercizio 3 Supponiamo di avere un campione statistico di $n = 10000$ v.a. $X_1, \dots, X_{10000}$ di Bernoulli di parametro p ignoto. Dall'indagine svolta abbiamo la stima $\hat{p} = 0.194$.

(i) Svolgiamo il test unilatero approssimato sulla media del campione con

$H_0)\ p \ge p_0 = 0.2$, $H_1)\ p < p_0 = 0.2$

Il p-value $\bar{\alpha}$ si calcola con la formula

$$\bar{\alpha} = \Phi\left(\frac{\sqrt{n}}{\sqrt{p_0(1-p_0)}} (\hat{p} - p_0) \right) = \Phi\left(\frac{\sqrt{10000}}{\sqrt{0.2 \cdot 0.8}} (0.194 - 0.2) \right) =$$

$$= \Phi(-1.5) = 1 - \Phi(1.5) \sim 1 - 0.93319 \sim 0.067$$

Poiché $\bar{\alpha} < 0.1$, l'ipotesi del capo del progetto va scartata.

(ii) L'intervallo di fiducia dell'80% di un campione statistico di v.a. di Bernoulli ha precisione della stima data da

$$\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}\, q_{1-\alpha/2} \quad \text{dove} \quad 1-\alpha = 0.8 \;, \quad \hat{p} = 0.194 \;, \quad n = 10000$$

quindi

$$\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}\, q_{1-\alpha/2} = \frac{\sqrt{0.194 \cdot 0.806}}{\sqrt{10000}}\, q_{0.9} \sim 0.004 \cdot 1.285 \sim \boxed{5 \cdot 10^{-3}}$$