Probabilità e Statistica

Laurea Triennale in Matematica

17/06/2014

Soluzioni traccia A

Esercizio 2. (Primo appello)

Consideriamo due urne A e B. L'urna A contiene 3 biglie rosse e 2 nere, mentre l'urna B contiene 2 biglie bianche, 2 rosse e 2 nere. Scegliamo un'urna a caso ed estraiamo due biglie (senza reinserimento). Per $i \in \{1, 2\}$ siano B_i, R_i e N_i gli eventi l'i-esima biglia estratta è bianca, rossa, nera. Indichiamo infine con A risp. B gli eventi le biglie sono state estratte dall'urna A risp. B. (`Se possibile esprimere i risultati sotto forma di frazioni`)

(a) Calcolare $P(R_2|A)$.
(b) Calcolare $P(N_1R_2|A)$.
(c) Calcolare $P(R_1)$.
(d) Calcolare $P(R_1 \cap A)$.
(e) Calcolare $P(A|R_1)$.
(f) Calcolare $P(R_1N_2)$.
(g) Calcolare $P(B|R_1N_2)$.
(h) Sia X il numero di biglie rosse estratte, quanto vale $\mathbb{E}[X]$?

Soluzione

(a) $\frac{3}{5}$ (b) $\frac{3}{10}$ (c) $\frac{7}{15}$ (d) $\frac{3}{10}$ (e) $\frac{9}{14}$ (f) $\frac{13}{60}$ (g) $\frac{4}{13}$ (h) $\frac{14}{15}$

Esercizio 2. (Secondo compitino)
Siano X e Y due variabili aleatorie (non necessariamente indipendenti) supponiamo che $X \sim Bin(3, \frac{1}{2})$ e $Y \sim exp(1)$. Sia $W := X + Y$. E' possibile che W abbia distribuzione Poissoniana? (`Sugg. Se è possibile fornire un esempio. Se non è possibile fornire una dimostrazione`)

Soluzione Dimostreremo che le ipotesi $X \sim Bin(3, \frac{1}{2})$, $Y \sim exp(1)$ e $W := X + Y$ implicano che W non può essere una v.a. di Poisson.

Metodo 1 Supponiamo per assurdo che W sia una v.a. di Poisson. Da $W := X + Y$ ricaviamo:

$$Y = W - X$$

osserviamo che

$$P(X \in \mathbb{Z}) = 1, \qquad P(Y \in \mathbb{Z}) = 0 \qquad P(W \in \mathbb{Z}) = 1$$

poiché la differenza di due numeri interi è ancora un numero intero allora vale:

$$P(W - X \in \mathbb{Z}) = 1, \qquad e \qquad P(Y \in \mathbb{Z}) = 0$$

Queste ultime due uguaglianze sono in contraddizione con $Y = W - X$.

Metodo 2 Sempre per assurdo supponiamo che $W \sim Poisson(\lambda)$ e $W := X + Y$. Essendo W sia una v.a. di Poisson vale

$$P(W = 0) = e^{-\lambda} > 0$$

mentre per $P(X + Y = 0)$ abbiamo:

$$P(X + Y = 0) = P(X = 0, Y = 0) \leq P(Y = 0) = 0$$

Le due equazioni $P(W = 0) > 0$ e $P(X + Y = 0) = 0$ sono in contraddizione con $W = X + Y$.

Esercizio 3. (Secondo compitino)
Sia $(X_n)_{n\in\mathbb{N}}$ una successione di variabili aleatorie indipendenti ed identicamente distribuite con distribuzione uniforme sull'intervallo $(1,7)$,
$X_n \sim Unif(1,7)$. Per ogni $n \in \mathbb{N}$ siano definite le seguenti variabili aleatorie:

$$Y_n := \frac{X_1 + \ldots + X_n}{n} \qquad T_n := \frac{Y_1 + \ldots + Y_n}{n}$$

Calcolare media e varianza di Y_n. Le variabili aleatorie $(Y_n)_{n\in\mathbb{N}}$ sono indipendenti? Sono non correlate? Dimostrare che la successione $(Y_n)_{n\in\mathbb{N}}$ soddisfa una legge dei grandi numeri (debole o forte a scelta).

Soluzione

$$\mathbb{E}[X_n] = 4 \qquad Var(X_n) = \frac{(7-1)^2}{12} = 3$$

$$\mathbb{E}[Y_n] = 4 \qquad Var(Y_n) = \frac{3}{n}$$

Calcoliamo ora la covarianza di $Cov(Y_n, Y_m)$, supponiamo $n \geq m$

$$Cov(Y_n, Y_m) = Cov\left(\frac{X_1 + \ldots + X_m}{m}, \frac{X_1 + \ldots + X_n}{n}\right) =$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{Cov(X_i, X_j)}{nm}$$

Poiché per $i \neq j$,le v.a. X_i e X_j sono indipendenti allora si ha:

$$Cov(X_i, X_j) = \begin{cases} 0 & i \neq j \\ Var(X_i) = 3 & i = j \end{cases}$$

dunque

$$Cov(Y_n, Y_m) = \sum_{i=1}^{m} \frac{Var(X_i)}{nm} = \frac{3}{n} \tag{1}$$

Dunque le variabili $(Y_n)_{n\in\mathbb{N}}$ non sono correlate e quindi non sono nemmeno indipendenti.

Legge debole dei grandi numeri per $(Y_n)_{n\in\mathbb{N}}$. Dimostrazione di Roselli Michele.
Bisogna dimostrare che per ogni $\epsilon > 0$ vale $\lim_{n\to\infty} P(|T_n - \mathbb{E}[T_n]| > \epsilon) = 0$
Applichiamo a T_n la disuguaglianza di Chebyschev:

$$\forall n \in \mathbb{N},\ \forall \epsilon > 0,\ \ P(|T_n - \mathbb{E}[T_n]| > \epsilon) \leq \frac{Var(T_n)}{\epsilon^2} = \frac{\frac{1}{n^2}Cov(Y_1 + \ldots + Y_n, Y_1 + \ldots + Y_n)}{\epsilon^2}$$

da cui ricaviamo:

$$P(|T_n - \mathbb{E}[T_n]| > \epsilon) \leq \frac{1}{n^2\epsilon^2} \cdot \sum_{i,j\in\{1,\dots,n\}} Cov(Y_i, Y_j)$$

Dall'equazione (1) ricaviamo $Cov(Y_i, Y_j) = \frac{3}{max(i,j)}$ quindi

$$P(|T_n - \mathbb{E}[T_n]| > \epsilon) \leq \frac{3}{n^2\epsilon^2} \left(\frac{1}{1} + \frac{3}{2} + \frac{5}{3} + \frac{7}{4} + \ldots + \frac{2n-1}{n}\right)$$

$$P(|T_n - \mathbb{E}[T_n]| > \epsilon) \leq \frac{3}{n^2\epsilon^2} \sum_{i=1}^{n} \frac{2n-1}{n} \leq \frac{3}{n^2\epsilon^2} \sum_{i=1}^{n} 2 = \frac{6}{n\epsilon^2}$$

Passando al limite si ha:

$$\limsup_{n\to\infty} P(|T_n - \mathbb{E}[T_n]| > \epsilon) \leq \limsup_{n\to\infty} \frac{6}{n\epsilon^2} = 0$$

Dunque per ogni $\epsilon > 0$ vale

$$\lim_{n\to\infty} P(|T_n - \mathbb{E}[T_n]| > \epsilon) = 0$$

Legge forte dei grandi numeri per $(Y_n)_{n\in\mathbb{N}}$. Bisogna dimostrare che per quasi ogni $\omega \in \Omega$ vale: $\lim_{n\to\infty} T_n(\omega) = \mathbb{E}[Y_n]$. Poiché le v.a. $(X_n)_{n\in\mathbb{N}}$ sono i.i.d. con $\mathbb{E}[X_n] = 4 < \infty$ allora per $(X_n)_{n\in\mathbb{N}}$ vale la legge forte dei grandi numeri quindi per quasi ogni $\omega \in \Omega$ si ha:

$$\lim_{n\to\infty} Y_n(\omega) = 4$$

Sia ora ω fissato tale che $\lim_{n\to\infty} Y_n(\omega) = 4$, per concludere la dimostrazione sarà sufficiente mostrare che $\lim_{n\to\infty} T_n(\omega) = 4$ ovvero che $\lim_{n\to\infty} |T_n(\omega) - 4| = 0$.

Sia $\epsilon > 0$ dalla condizione $\lim_{n\to\infty} Y_n(\omega) = 4$ ricaviamo che esiste N tale che

$$\forall n > N \qquad |Y_n(\omega) - 4| < \epsilon$$

inoltre per ogni n vale $|Y_n(\omega) - 4| \leq 3$. Proviamo adesso ad utilizzare le disuguaglianze precedenti per stimare il limite $\limsup_{n\to\infty} |T_n(\omega) - 4|$.

$$\limsup_{n\to\infty} |T_n(\omega) - 4| = \limsup_{n\to\infty} \left|\sum_{i=1}^{n} \frac{Y_i(\omega) - 4}{n}\right|$$

$$\leq \limsup_{n\to\infty} \left(\sum_{i=1}^{N} \frac{|Y_i(\omega) - 4|}{n} + \sum_{i=N+1}^{n} \frac{|Y_i(\omega) - 4|}{n}\right)$$

utilizzando le maggiorazioni $|Y_n(\omega) - 4| \leq 3$ per i numeratori della prima sommatoria e $|Y_n(\omega) - 4| < \epsilon$ per i numeratori della seconda sommatoria si ottiene

$$\limsup_{n\to\infty} |T_n(\omega) - 4| \leq \limsup_{n\to\infty} \left(\sum_{i=1}^{N} \frac{3}{n} + \sum_{i=N+1}^{n} \frac{\epsilon}{n} \right)$$

$$\limsup_{n\to\infty} |T_n(\omega) - 4| \leq \limsup_{n\to\infty} \left(\frac{3N}{n} + \frac{\epsilon(n-N)}{n} \right) = \epsilon$$

dunque per l'arbitrarietà di $\epsilon > 0$ si ha

$$\lim_{n\to\infty} |T_n(\omega) - 4| = 0$$

Esercizio 4. (Secondo compitino)
Sia $(\Omega, \mathcal{A}, P_\theta)_{\theta\in\Theta}$ un modello statistico con $\Theta = (0,\infty)$. Sia $(X_n)_{n\in\mathbb{N}}$ un campione tale che per ogni $\theta > 0$, le v.a. X_n abbiano distribuzione assolutamente continua con densità f_θ data da:

$$f_\theta(x) = \begin{cases} \alpha_\theta x & x \in (0,\theta) \\ 0 & \text{altrimenti} \end{cases}$$

Sia infine $T_n := \beta\frac{X_1+\ldots+X_n}{n}$ per ogni $n \in \mathbb{N}$.
(a) Calcolare α_θ.
(b) Dimostrare che per $\beta = \frac{3}{2}$ si ha: T_n stimatore corretto per θ.
(c) Sia $\beta = \frac{3}{2}$ dimostrare che per ogni $\theta > 0$ vale:
$P_\theta(\omega \in \Omega | \lim_{n\to\infty} T_n(\omega) = \theta) = 1$

Soluzione (a) Affinché f_θ sia effettivamente una densità è necessario che valgano le relazioni $f_\theta(x) \geq 0$ per ogni x e $\int_{\mathbb{R}} f_\theta(x)dx = 1$. Dalla prima disuguaglinza ricaviamo che deve valere $\alpha_\theta \geq 0$. Per imporre la seconda condizione dobbiamo calcolare $\int_{\mathbb{R}} f_\theta(x)dx$,

$$\int_{\mathbb{R}} f_\theta(x)dx = \int_0^\theta \alpha_\theta x dx = \frac{\alpha_\theta\theta^2}{2}$$

ponendo $\frac{\alpha_\theta\theta^2}{2} = 1$ si ottiene $\boxed{\alpha_\theta = \frac{2}{\theta^2}}$
(b) E' sufficiente verificare che $\mathbb{E}[T_n] = \theta$, innanzitutto calcoliamo il valore atteso di X_n.

$$\mathbb{E}[X_n] = \int_{\mathbb{R}} x f_\theta(x)dx = \int_0^\theta \alpha_\theta x^2 dx = \frac{2}{3}\theta$$

dunque

$$\mathbb{E}[T_n] = \mathbb{E}\left[\beta\frac{X_1+\ldots+X_n}{n}\right] = \beta\mathbb{E}[X_n] = \theta$$

(c) Le variabili aleatorie X_n sono i.i.d. con $\mathbb{E}[X_n] = \frac{2}{3}\theta$ per la legge forte dei grandi numeri vale:

$$\text{per quasi ogni } \omega \in \Omega \qquad \lim_{n\to\infty} \frac{X_1(\omega)+\ldots+X_n(\omega)}{n} = \mathbb{E}[X] = \frac{2}{3}\theta$$

moltiplicando a destra e sinistra per $\frac{3}{2}$ si ha

$$\text{per quasi ogni } \omega \in \Omega \qquad \lim_{n\to\infty} \frac{3}{2}\frac{X_1(\omega)+\ldots+X_n(\omega)}{n} = \theta$$

quindi

$$\text{per quasi ogni } \omega \in \Omega \qquad \lim_{n\to\infty} T_n(\omega) = \theta$$

Esercizio 5. (Secondo compitino)
Una compagnia di assicurazioni auto prevede per i guidatori giovani una polizza più alta, in quanto questo gruppo tende ad avere un numero maggiore di incidenti. La compagnia ha 80000 assicurati divisi in 3 gruppi:
A (sotto i 25 anni, 20% di tutti i suoi assicurati),
B (25-35 anni, 30%),
C (sopra i 35 anni 50%).
I dati mostrano che in media ogni anno le percentuali di assicurati che hanno un incidente sono: 10% per il gruppo A, 5% per il B, 2% per il C. Sia X il numero di assicurati che avranno un incidente il prossimo anno.
(a) Calcolare media e varianza di X.
(b) Scelgo un assicurato a caso (tra gli 80000), qual è la probabilit che abbia un incidente.
(c) Scelgo un assicurato a caso (tra gli 80000), se so che ha avuto un incidente qual è la probabilità che sia del gruppo A?
(d) Stimare (utilizzando l'approssimazione normale) la probabilità che il numero di assicurati che hanno un incidente sia maggiore (strettamente) di 3700.

Soluzione
(a) $\mathbb{E}[X] = 3600$, $Var(X) = 3364$
(b) $P(I) = 0.045 = 4.5\%$
(c) $P(A|I) = \frac{4}{9}$
(d) $P(X > 3700) \simeq 1 - \Phi\left(\frac{3700.5-3600}{58}\right) = 0.04182$