SOLUZIONI APPELLO 4 APRILE 2019 - INGEGNERIA INFORMATICA

Esercizio 1
Un treno sta viaggiando da Roma a Torino, con n passeggeri pi 2 persone di equipaggio. Ad un certo punto, a causa di un guasto il treno si ferma, e si fanno scendere le persone a bordo (sia passeggeri che membri dell'equipaggio) una alla volta ed in un ordine casuale. Si indichi con Y la variabile aleatoria che corrisponde al numero di passeggeri che scendono dal treno prima che scenda il primo membro dell'equipaggio. Si calcolino

(1) la distribuzione di probabilità di Y
(2) il suo valore atteso

Suggerimento: si ricordi la formula seguente

$$\sum_{j=1}^{n} j^2 = \frac{n(n+1)(2n+1)}{6}$$

N.B. Tutte le risposte devono essere opportunamente giustificate
Soluzione

$$\begin{aligned}
\mathbb{P}(Y=k) &= \mathbb{P}(\text{prima scendono} k \text{ passeggeri, il } (k+1)-\text{esimo è un membro dell'equipaggio}) \\
&= \mathbb{P}(\text{il } (k+1)-\text{esimo è un membro dell'equipaggio} \mid \text{prima scendono } k \text{ passeggeri}) \\
&\times \mathbb{P}(\text{prima scendono } k \text{ passeggeri}) \\
&= \frac{\binom{n}{k}\binom{2}{0}}{\binom{n+2}{k}} \frac{2}{n+2-k} \\
&= \frac{2(n+1-k)}{(n+2)(n+1)}.
\end{aligned}$$

Verifica:

$$\sum_{k=0}^{n} \mathbb{P}(Y=k) = \frac{2}{(n+2)(n+1)} \sum_{k=0}^{n} (n+1-k) = \frac{2}{(n+2)(n+1)} \sum_{l=1}^{n+1} l = 1.$$

$$\begin{aligned}
\mathbb{E}(Y) &= \frac{2}{(n+2)(n+1)} \sum_{k=0}^{n} k(n+1-k) = \frac{2}{(n+2)} \sum_{k=0}^{n} k - \frac{2}{(n+2)(n+1)} \sum_{k=0}^{n} k^2 \\
&= \frac{2n(n+1)}{2(n+2)} - \frac{2n(n+1)(2n+1)}{6(n+1)(n+2)} = \frac{n}{3}.
\end{aligned}$$

Esercizio 2
Siano X_1 e X_2 due v.a. esponenziali indipendenti di parametri rispettivamente pari a λ_1 e λ_2.

(1) Si calcoli la distribuzione di probabilità della v.a. $Z = \log(X) - \log(Y)$
(2) Si calcoli la probabilità che Z sia negativa

N.B. Tutte le risposte devono essere opportunamente giustificate
Soluzione $X \sim \text{Exp}(\lambda_1)$, $Y \sim \text{Exp}(\lambda_2)$, dunque $Z = \log(X) - \log(Y)$ e $Z \in (-\infty, \infty)$ q.c.

$$\begin{aligned}
F_Z(z) &= \mathbb{P}(\log(X) - \log(Y) < z) = \mathbb{P}\left(\log\left(\frac{X}{Y}\right) < z\right) = \mathbb{P}\left(\frac{X}{Y} < e^z\right) = \mathbb{P}\left(X < e^z Y\right) \\
&= \lambda_2 \int_0^{\infty} e^{-\lambda_2 y} \lambda_1 \int_0^{e^z y} e^{-\lambda_1 x} dx dy
\end{aligned}$$

$$
\begin{aligned}
&= \lambda_2 \int_0^\infty e^{-\lambda_2 y} dy \left[-e^{-\lambda_1 x}\right]_0^{e^z y} \\
&= \lambda_2 \int_0^\infty e^{-\lambda_2 y} dy - \lambda_2 \int_0^\infty e^{-\lambda_2 y - \lambda_1 e^z y} dy \\
&= 1 - \frac{\lambda_2}{\lambda_2 + \lambda_1 e^z} \left[-e^{-(\lambda_2 + \lambda_1 e^z) y}\right]_0^\infty \\
&= 1 - \frac{\lambda_2}{\lambda_2 + \lambda_1 e^z} = \frac{\lambda_1 e^z}{\lambda_2 + \lambda_1 e^z},
\end{aligned}
$$

per $z \in (-\infty, \infty)$.

La probabilità che Z è negativa è data da

$$
\mathbb{P}(Z < 0) = F_Z(0) = \frac{\lambda_1}{\lambda_1 + \lambda_2}
$$

il che sigignifica che tale probabilità è tanto maggiore quanto maggiore è il valore di λ_1 (rispetto a λ_2). Poichè nell'esponenziale il parametro è l'inverso del valore medio, ciò significa che $\mathbb{P}(Z < 0)$ sarà tanto maggiore quanto minore è il valore medio di X come è confermato dalla definizione di $Z = \log(X) - \log(Y)$.

Esercizio 3 Siano Y e $\{X_n\}_{n\in\mathbb{N}}$ un insieme di variabili aleatorie indipendenti. Supponiamo che Y abbia una distribuzione bernoulliana di parametro $p = \frac{1}{6}$ e che le X_n siano uniformi sull'intervallo $(0, 3)$. Definiamo $Z_n = \min(X_1, X_2, \ldots, X_n)$ e $T_n = \min(Y, X_n)$. Determinare la finzione di ripartizione di Z_n e la funzione di ripartizione di T_n. Studiare le convergenza in distribuzione e la convergenza in probabilità di Z_n.

N.B. Tutte le risposte devono essere opportunamente giustificate

Soluzione

$$
F_{X_n}(x) = \begin{cases} 0 & x < 0, \\ \frac{x}{3} & 0 \le x < 3 \\ 1 & x \ge 3. \end{cases}
$$

$$
\begin{aligned}
F_{Z_n}(t) &= \mathbb{P}(Z_n \le t) \\
&= \mathbb{P}(\min(X_1, X_2, \ldots, X_n) \le t) \\
&= 1 - \mathbb{P}(\min(X_1, X_2, \ldots, X_n) > t) \\
&= 1 - \mathbb{P}(X_1 > t, X_2 > t, \ldots, X_n > t) \\
&= 1 - \mathbb{P}(X_1 > t)\mathbb{P}(X_2 > t) \cdots \mathbb{P}(X_n > t) \\
&= 1 - (1 - F_{X_1}(t))(1 - F_{X_2}(t)) \cdots (1 - F_{X_n}(t)) \\
&= 1 - (1 - F_{X_1}(t))^n,
\end{aligned}
$$

dunque

$$
F_{Z_n}(t) = \begin{cases} 0 & t < 0, \\ 1 - (1 - t/3)^n & 0 \le t < 3, \\ 1 & t \ge 3, \end{cases} \to \begin{cases} 0 & t < 0, \\ 1 & 0 \le t < 3, \\ 1 & t \ge 3, \end{cases}
$$

dunque $Z_n \xrightarrow{d} 0$ e $Z_n \xrightarrow{p} 0$. Inoltre

$$
\begin{aligned}
F_{T_n}(t) &= \mathbb{P}(\min(Y, X_n) \le t) \\
&= 1 - \mathbb{P}(\min(Y, X_n) > t) \\
&= 1 - \mathbb{P}(Y > t, X_n > t) \\
&= 1 - \mathbb{P}(Y > t)\mathbb{P}(X_n > t),
\end{aligned}
$$

e poiché

$$F_Y(t) = \begin{cases} 0 & t < 0 \\ \frac{5}{6} & 0 \leq t < 1 \\ 1 & t \geq 1, \end{cases}$$

la funzione di ripartizione di T_n è data da

$$F_{T_n}(t) = \begin{cases} 0 & t < 0 \\ \frac{5}{6} + \frac{t}{18} & t \in [0, 1) \\ 1 & t \geq 1. \end{cases}$$

Esercizio 3 alternativo

Sia Y la variabile casuale che descrive il numero di veicoli che transitano, in una determinata fascia oraria giornaliera, su un tratto del GRA nel quale è installato un autovelox. Si suppone che Y abbia legge di Poisson di parametro λ e che, per il generico veicolo che transita in quel tratto di GRA in quella fascia oraria, la probabilità di prendere una multa per eccesso di velocità è $\theta/(\lambda+1)$, con $\theta \in (0,1)$, indipendentemente da quanto possa accadere a qualsiasi altro veicolo.

Una operazione di monitoraggio del traffico permette di ottenere, per n giorni, i dati $\{(y_i, x_i),\ i = 1, 2, ..., n)\}$ sul numero di veicoli transitati (nel tratto e nella fascia oraria considerati) e il numero corrispondente di multe. Le osservazioni relative a giorni distinti si assumono indipendenti. Si determini lo stimatore di massima verosimiglianza di θ assumendo che λ sia noto.

N.B. Tutte le risposte devono essere opportunamente giustificate

Soluzione

Per la funzione di verosimiglianza, relativa all'osservazione $(y_1, x_1), (y_2, x_2), \ldots, (y_n, x_n)$, si ha

$$L(\theta) \propto \prod_{i=1}^{n} \lambda^{y_i} e^{-\lambda} \left(\frac{\theta}{\lambda+1}\right)^{x_i} \left(1 - \frac{\theta}{\lambda+1}\right)^{y_i - x_i}.$$

Per λ noto possiamo considerare la funzione

$$g(\theta) = \prod_{i=1}^{n} \left(\frac{\theta}{\lambda+1}\right)^{x_i} \left(1 - \frac{\theta}{\lambda+1}\right)^{y_i - x_i},$$

dove

$$\log g(\theta) = \sum_{i=1}^{n} [x_i \log \theta + (y_i - x_i) \log(\lambda + 1 - \theta)] + \text{const}$$

e

$$\frac{d \log g(\theta)}{d\theta} = \sum_{i=1}^{n} \left[\frac{x_i}{\theta} - \frac{y_i - x_i}{\lambda + 1 - \theta}\right].$$

Uguagliando a zero e risolvendo in θ si ottiene

$$\hat{\theta} = (\lambda + 1)\frac{\bar{x}}{\bar{y}},$$

dove $\bar{x}$ e $\bar{y}$ indicano le medie campionarie dei valori $x_1, x_2, \ldots, x_n$ e $y_1, y_2, \ldots, y_n$, rispettivamente.

D1)

Dimostrare la proprietà di continuità della probabilità per:

- successioni monotone di eventi
- successioni non monotone, ma convergenti di eventi

N.B. Fornire tutti i dettagli delle dimostrazioni

D2) Dare la definizione di varianza di una variabile aleatoria. Enunciare e dimostrare le proprietà della varianza.

N.B. Fornire tutti i dettagli delle dimostrazioni.

D3)

(1) Dare la definizione di uguaglianza, uguaglianza quasi certa, uguaglianza in distribuzione ed indipendenza per due variabili aleatorie X e Y.

(2) Dimostrare che l'uguaglianza di due variabili aleatorie implica l'uguaglianza in distribuzione e che non vale il viceversa, ossia che l'uguaglianza in distribuzione non implica l'uguaglianza.

(3) Data la variabile aleatoria doppia (X, Y) dare la definizione di distribuzione di probabilità di Y condizionata ad $X = x$, nel caso discreto e continuo. Cosa succede se le due variabili aleatorie X e Y sono indipendenti?

N.B. Fornire tutti i dettagli delle dimostrazioni.