Dipartimento di Matematica, Roma Tre
Pietro Caputo

2011-12, II semestre
13 settembre, 2012

CP110 Probabilità: Esame 13 settembre 2012

Testo e soluzione

Nome:____

1. **(6 pts)** Una scatola contiene 10 palline, 8 bianche e due nere. Vengono estratte a caso, senza rimpiazzo, una alla volta. Calcolare:

   (a) la probabilità che k palline bianche precedano la prima pallina nera.

   (b) la probabilità che tra la prima pallina nera e la seconda pallina nera ci siano k palline bianche.

   (c) il numero medio di palline bianche che precedono la prima pallina nera.

Soluzione: Immaginiamo la sequenza come 10 caselle numerate di cui 8 sono bianche e 2 nere, e scegliamo lo spazio campionario come l'insieme delle $\binom{10}{2}$ possibili sequenze distinte. Tutte le sequenze hanno la stessa probabilità $1/\binom{10}{2} = 1/45$.

(a). Sia E_k l'evento che k caselle bianche precedono la prima nera. Dunque $k = 0, \ldots, 8$, altrimenti si ha $P(E_k) = 0$. Se vale E_k allora abbiamo una casella nera nella $k+1$-esima posizione e resta da scegliere una casella nera tra le $10 - (k+1)$ a destra della $k+1$-esima casella. Quindi

$$P(E_k) = \frac{\binom{10-(k+1)}{1}}{\binom{10}{2}} = \frac{9-k}{45}\,, \qquad k = 0, \ldots, 8.$$

(b). Consideriamo tutti i modi di ottenere $k = 0, \ldots, 8$ caselle bianche tra le due caselle nere. Se F_k denota l'evento in questione, e A_j, $j = 1, \ldots, 9$ denota la posizione della prima pallina nera, si ha che $P(F_k \cap A_j) = 0$ ogni volta che $j + k > 9$ (non ci sono abbastanza caselle a destra delle j-esima per realizzare l'evento F_k). Se invece $j + k \leqslant 9$ si ha $P(F_k \cap A_j) = 1/\binom{10}{2}$ (c'è un'unica sequenza che realizza l'evento $F_k \cap A_j$ in questo caso). In conclusione

$$P(F_k) = \sum_{j=1}^{9} P(F_k \cap A_j) = \sum_{j=1}^{9-k} \frac{1}{\binom{10}{2}} = \frac{9-k}{45}\,, \qquad k = 0, \ldots, 8.$$

Dunque si ha la stessa distribuzione del punto (a) sopra.

(c). Siano X_1, X_2, X_3 rispettivamente il numero di palline bianche che precedono la prima nera; il numero di palline bianche tra la prima nera e la seconda nera; il numero di palline bianche che seguono la seconda nera. Abbiamo visto che X_1 e X_2 hanno la stessa distribuzione. Inoltre, per simmetria X_1 e X_3 hanno la stessa distribuzione. Allora $E[X_1] = E[X_2] = E[X_3]$, e quindi usando il fatto che $X_1 + X_2 + X_3 = 8$ si ha

$$E[X_1] = \frac{1}{3}\left(E[X_1] + E[X_2] + E[X_3]\right) = \frac{1}{3}\, E[X_1 + X_2 + X_3] = \frac{8}{3}.$$

Alla stessa risposta si può arrivare scrivendo

$$E[X_1] = \sum_{k=0}^{8} k P(X_1 = k) = \sum_{k=0}^{8} k P(E_k) = \frac{1}{45} \sum_{k=1}^{8} k(9-k) = \frac{120}{45} = \frac{8}{3}$$

Nome:______________________________________

2. **(6 pts)** Supponiamo che la durata in minuti di una partita di baseball sia una variabile uniforme nell'intervallo $[60, 240]$. Due squadre si incontrano per una serie di 5 partite. Assumendo indipendenza tra le partite determinare

   (a) il numero medio e la varianza dei minuti giocati in totale.

   (b) la covarianza tra la durata della prima partita e il numero di minuti giocati in totale

Soluzione: Sia X_i la durata in minuti della partita i-esima e sia $Y = X_1 + \cdots + X_5$ la durata totale.

(a). Per la media si ha

$$E[Y] = 5E[X_1] = 5 \times 150 = 750$$

dove 150 è il valor medio di una uniforme in $[60, 240]$. La varianza, usando l'indipendenza, è data da

$$\mathrm{Var}[Y] = 5\mathrm{Var}[X_1] = 5 \times 180^2/12 = 13500$$

dove $180^2/12 = (240 - 60)^2/12 = 2700$ è la varianza di una uniforme in $[60, 240]$.

(b). Per l'indipendenza si ha

$$\mathrm{Cov}(X_1, Y) = \mathrm{Var}[X_1] + \mathrm{Cov}(X_1, X_2 + \cdots + X_5) = \mathrm{Var}[X_1] = 2700$$

Nome:____

3. **(6 pts)** Le variabili aleatorie continue X, Y hanno densità congiunta

$$f(x,y) = c\, e^{-|x|-y^2}\,, \qquad x \in \mathbb{R},\ y \in \mathbb{R}$$

dove $c > 0$ è una costante.

(a) Dire se X e Y sono indipendenti.

(b) Calcolare il valore della costante c.

(c) Calcolare la varianza di $X + Y$.

Soluzione:

(a). A parte la costante c abbiamo che f è data dal prodotto delle due funzioni $e^{-|x|}$, $x \in \mathbb{R}$ e e^{-y^2}, $y \in \mathbb{R}$. Quindi X e Y sono indipendenti.

(b). Si ha

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = c\int_{-\infty}^{\infty} e^{-|x|}dx \int_{-\infty}^{\infty} e^{-y^2}dy = 2c\sqrt{\pi},$$

dove abbiamo calcolato gli integrali

$$\int_{-\infty}^{\infty} e^{-|x|}dx = 2\int_{0}^{\infty} e^{-x}dx = 2\,, \qquad \int_{-\infty}^{\infty} e^{-y^2}dy = \sqrt{\pi}$$

Allora $c = (2\sqrt{\pi})^{-1}$.

(c). Per l'indipendenza si ha

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y].$$

Notiamo che X è una variabile con densità $\frac{1}{2}e^{-|x|}$, $x \in \mathbb{R}$, e quindi ha media nulla e momento secondo

$$E[X^2] = \frac{1}{2}\int_{-\infty}^{\infty} x^2 e^{-|x|}dx = \int_{0}^{\infty} x^2 e^{-x}dx = 2.$$

Inoltre Y è una normale di media 0 e varianza $\sigma^2 = 1/2$. Quindi

$$\mathrm{Var}[X+Y] = 2 + \frac{1}{2} = \frac{5}{2}.$$

Nome:____

4. **(6 pts)** Un segnale è costituito da una sequenza di $n+1$ bit ottenuti come segue: i primi n bit sono indipendenti e ciascuno vale 1 o 0 con probabilità p e $1-p$ rispettivamente; l'$n+1$-esimo bit vale 1 con probabilità k/n e 0 con probabilità $1-(k/n)$, dove k è il numero di 1 nei primi n bit. Calcolare, in funzione di $n \in \mathbb{N}$ e $p \in [0,1]$:
   (a) il numero medio di 1 nella sequenza
   (b) la covarianza tra l'$n+1$-esimo bit e il numero di 1 nei primi n bit

Soluzione: Il numero di 1 nella sequenza è dato da $N = X + Y$ dove X è una binomiale di parametri n e p, mentre $Y \in \{0,1\}$ soddisfa, per $k = 0, \dots, n$:

$$P(Y=1|X=k) = \frac{k}{n}, \qquad P(Y=0|X=k) = 1 - \frac{k}{n}.$$

(a). Il valor medio di Y è dato da

$$E[Y] = P(Y=1) = \sum_{k=0}^{n} P(X=k)P(Y=1|X=k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}\, \frac{k}{n} = \frac{np}{n} = p\,,$$

dove abbiamo usato il fatto che la media della binomiale è np.

(b). Si ha

$$\text{Cov}(X,Y) = E[XY] - E[X]E[Y] = E[XY] - np^2.$$

Per calcolare $E[XY]$ scriviamo

$$\begin{aligned} E[XY] &= \sum_{k=0}^{n} kP(X=k, Y=1) \\ &= \sum_{k=0}^{n} kP(X=k)P(Y=1|X=k) \\ &= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}\, \frac{k^2}{n} = \frac{1}{n}(np(1-p) + n^2p^2) = p(1-p) + np^2 \end{aligned}$$

dove usiamo il fatto che il momento secondo della binomiale è $np(1-p) + n^2p^2$. Allora

$$\text{Cov}(X,Y) = p(1-p).$$

Nome:____

5. **(6 pts)** Siano X_1, X_2, X_3 tre variabili esponenziali indipendenti di parametro $\lambda = 2$. Calcolare
   (a) $\mathrm{Cov}(X_1 + X_2, X_2 - X_3)$
   (b) $E[3X_1X_2(X_3 - 1)]$
   (c) $E[X_1X_2(3X_3 - 2X_2)]$

Soluzione:

(a). Per la linearità e l'indipendenza:

$$\mathrm{Cov}(X_1 + X_2, X_2 - X_3) = \mathrm{Cov}(X_1, X_2) + \mathrm{Var}[X_2] - \mathrm{Cov}(X_1, X_3) - \mathrm{Cov}(X_2, X_3) = \mathrm{Var}[X_2] = \frac{1}{4}$$

dove usiamo il fatto che la varianza di un'esponenziale di parametro λ è pari a $1/\lambda^2$.

(b). Usando $E[X_i] = 1/2$ si ha

$$E[3X_1X_2(X_3 - 1)] = 3E[X_1]E[X_2](E[X_3] - 1) = -\frac{3}{8}$$

(c). Si ha:

$$E[X_1X_2(3X_3 - 2X_2)] = 3E[X_1]E[X_2]E[X_3] - 2E[X_1]E[X_2^2] = \frac{3}{8} - \frac{1}{2} = -\frac{1}{8}$$

dove usiamo il fatto che il momento secondo di un'esponenziale di parametro λ vale $2/\lambda^2$.

Nome:______________________________________

6. **(6 pts)** Sia Z una variabile aleatoria con distribuzione poissoniana di parametro $\lambda = n$. Fornire un valore approssimato, per valori molto grandi di n, per le seguenti probabilità

   (a) $P(Z \geqslant 1)$

   (b) $P(Z \geqslant n + e^{-n})$

   (c) $P(Z \geqslant n + \sqrt{n})$

Soluzione:
Sappiamo che Z è una somma di n variabili di poisson X_i indipendenti di parametro $\lambda_0 = 1$:

$$Z = \sum_{i=1}^{n} X_i.$$

Inoltre la media e la varianza di Z sono entrambi uguali a n. Dunque per il teorema del limite centrale la variabile

$$\zeta(n) = \frac{\sum_{i=1}^{n}(X_i - 1)}{\sqrt{n}} = \frac{Z}{\sqrt{n}} - \sqrt{n}$$

è approssimativamente normale standard. Possiamo scrivere dunque per ogni $t \geqslant 0$:

$$P(Z \geqslant t) = P(\sqrt{n}\zeta(n) + n \geqslant t) = P(\zeta(n) \geqslant (t-n)/\sqrt{n}) \approx 1 - \Phi((t-n)/\sqrt{n}) \,, \qquad (1)$$

dove Φ sta per la funzione di distribuzione della normale standard.

(a). Usiamo la formula (1) con $t = 1$. Poiché $(1-n)/\sqrt{n} \to -\infty$ si ha $\Phi((1-n)/\sqrt{n}) \to 0$ e quindi otteniamo

$$P(Z \geqslant 1) \approx 1$$

(b). Usiamo la formula (1) con $t = n + e^{-n}$. Poiché $e^{-n}/\sqrt{n} \to 0$, si ha $\Phi((t-n)/\sqrt{n}) \to 1/2$ e pertanto

$$P(Z \geqslant n + e^{-n}) \approx \frac{1}{2}$$

(c). Ragionando come sopra, con $t = n + \sqrt{n}$ abbiamo

$$P(Z \geqslant n + \sqrt{n}) \approx 1 - \Phi(1) \approx 0.16.$$