Dipartimento di Matematica, Roma Tre
Pietro Caputo

2018-19, II semestre
4 giugno, 2019

CP210 Introduzione alla Probabilità: Esonero 2

Cognome	
Nome	
Matricola	
Firma	

Nota:

1. L'unica cosa che si può usare durante l'esame è una penna o una matita. Tutto il resto (calcolatrice, libri, appunti, altri fogli di carta, ...) deve essere messo da parte.
2. Risposte implicite sotto forma di coefficienti binomiali, potenze, esponenziali ecc. sono benvenute. Mostrate in dettaglio il vostro lavoro.
3. Non parlate durante l'esame. Copiare o far copiare non è tollerabile.
4. Scrivete il vostro nome su ogni pagina. In caso di utilizzo di piu' pagine per un singolo esercizio indicare chiaramente l'ordine.
5. Il punteggio massimo per ogni esercizio è indicato nel testo. Notare che già con quattro esercizi risolti correttamente si arriva vicino al trenta.

Buon lavoro!

esercizio	1	2	3	4	5	totale
punti						
su	7	7	7	7	7	35

Nome:____

1. **(7 punti)** Sia $P = (X, Y)$ un punto aleatorio nel piano con distribuzione uniforme nel quadrato $[0,1]^2$. Calcolare la probabilità degli eventi:

   (a) $A = \{X^2 + Y^2 \leqslant 1\}$

   (b) $B = \{Y \leqslant X^2\}$

   (c) $C = \{1/2 \leqslant X/Y \leqslant 2\}$

Soluzione:

1. Notiamo che la regione del piano corrispondente all'evento A ha area $\pi/4$ e poiché l'area del quadrato vale 1 abbiamo $\mathbb{P}(A) = \pi/4$. Allo stesso risultato si può arrivare scrivendo

$$\mathbb{P}(A) = \int_0^1 dx \int_0^{\sqrt{1-x^2}} dy = \int_0^1 \sqrt{1-x^2}dx = \frac{\pi}{4}.$$

Abbiamo usato il fatto che le coordinate X, Y del punto P hanno densità congiunta

$$f_{X,Y}(x,y) = 1_{[0,1]}(x)1_{[0,1]}(y).$$

2. Allo stesso modo

$$\mathbb{P}(B) = \int_0^1 dx \int_0^{x^2} dy = \int_0^1 x^2 dx = \frac{1}{3}.$$

3. Per l'evento C abbiamo $C = \{Y/2 \leqslant X \leqslant 2Y\}$ e dunque

$$\mathbb{P}(C) = \int_0^1 dy \int_{y/2}^{2y} 1_{[0,1]}(x)dx = \int_0^{1/2} dy \int_{y/2}^{2y} dx + \int_{1/2}^1 dy \int_{y/2}^1 dx = \frac{1}{2}.$$

Lo stesso risultato può essere ottenuto piú rapidamente per via geometrica calcolando l'area della regione associata a C.

Nome:____

2. **(7 punti)** Sia X la variabile aleatoria tale che $X = 1$ con probabilità $1/3$ e $X = 2$ con probabilità $2/3$. Siano $X_1, \ldots, X_n$ copie indipendenti di X e sia S_n la loro somma.

   (a) Calcolare il valore atteso e la varianza di S_n per ogni $n \in \mathbb{N}$.

   (b) Scrivere la probabilità dell'evento $S_n = k$, per ogni $k, n \in \mathbb{N}$

   (c) Fornire un valore approssimato, per n grandi, per la probabilità dell'evento

$$3S_n \in \left[5n, 5n + \sqrt{2n}\right].$$

Soluzione: Notiamo che $X_i = 1 + Y_i$ dove Y_i è una Bernoulli di parametro $p = 2/3$. Allora $S_n = n + B_n$ dove B_n è una binomiale di parametri n e p. Dunque abbiamo

$$\mathbb{E}[S_n] = n + np = n(1+p) = \frac{5}{3}n\,, \qquad \mathrm{Var}(S_n) = \mathrm{Var}(B_n) = np(1-p) = \frac{2}{9}n.$$

(b). Inoltre

$$\mathbb{P}(S_n = k) = \mathbb{P}(B_n = k - n) = \binom{n}{k-n} p^{k-n}(1-p)^{n-(k-n)}\mathbf{1}(k - n \in \{0, \ldots, n\})$$

(c). Per il teorema del limite centrale sappiamo che S_n è approssimata da una normale, ossia che la variabile

$$Z_n = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathrm{Var}(S_n)}}$$

converge in distribuzione a una normale standard. Allora per ogni $a < b$ si ha

$$\mathbb{P}\left(S_n \in \left[\frac{5}{3}n + a\sqrt{n}, \frac{5}{3}n + b\sqrt{n}\right]\right) = \mathbb{P}(Z_n \in [3a/\sqrt{2}, 3b/\sqrt{2}])$$

Ponendo $a = 0$ e $b = \sqrt{2}/3$ abbiamo

$$\begin{aligned}\mathbb{P}\left(3S_n \in \left[5\,n, 5\,n + \sqrt{2n}\right]\right) &= \mathbb{P}\left(S_n \in \left[\frac{5}{3}n, \frac{5}{3}n + \frac{1}{3}\sqrt{2n}\right]\right)\\ &= \mathbb{P}(Z_n \in [0,1]) \approx \frac{1}{\sqrt{2\pi}}\int_0^1 e^{-x^2/2}dx = \Phi(1) - \frac{1}{2} \approx 0.34\end{aligned}$$

Nome:____

3. **(7 punti)** Lanciamo 2 dadi. Sia X la somma. Successivamente, i dadi con faccia diversa da 6 vengono lanciati nuovamente mentre quelli con faccia uguale a 6 non vengono toccati. Sia Y la somma al termine di questo secondo lancio (che include eventuali 6 del primo lancio). Calcolare:

   (a) il valore atteso $\mathbb{E}[X]$

   (b) la probabilità dell'evento $\{Y = 7\}$

   (c) il valore atteso $\mathbb{E}[Y]$.

Soluzione: Notiamo che X è la somma di due dadi e dunque

$$\mathbb{E}[X] = 7.$$

Per il calcolo di $\mathbb{P}(Y = 7)$, chiamiamo Z il numero di 6 nei primi due lanci. Abbiamo tre casi: $Z = 0$, con probab. $(5/6)^2 = 25/36$, $Z = 1$ con probab. $2(1/6)(5/6) = 10/36$ e $Z = 2$ con probab. $(1/6)^2 = 1/36$. Allora

$$\mathbb{P}(Y = 7) = \frac{25}{36}\mathbb{P}(Y = 7|Z = 0) + \frac{10}{36}\mathbb{P}(Y = 7|Z = 1) + \frac{1}{36}\mathbb{P}(Y = 7|Z = 2).$$

Se $Z = 0$ lanciamo nuovamente due dadi e dunque $\mathbb{P}(Y = 7|Z = 0) = \mathbb{P}(X = 7) = 1/6$. Se $Z = 1$ lanciamo solo un dado che sommato al 6 precedente deve fare 7 e quindi $\mathbb{P}(Y = 7|Z = 1) = 1/6$. Infine se $Z = 2$ si ha $\mathbb{P}(Y = 7|Z = 2) = 0$ poiché $Y = X = 12$ se $Z = 2$. In conclusione,

$$\mathbb{P}(Y = 7) = \frac{25}{36} \times \frac{1}{6} + \frac{10}{36} \times \frac{1}{6} = \frac{35}{216}.$$

Per il calcolo del valore atteso possiamo ragionare come sopra condizionando al valore di Z. Otteniamo:

$$\begin{aligned}
\mathbb{E}[Y] &= \sum_{j=2}^{12} j\mathbb{P}(Y = j) \\
&= \sum_{j=2}^{12} j\left(\mathbb{P}(Z = 0)\mathbb{P}(Y = j|Z = 0) + \mathbb{P}(Z = 1)\mathbb{P}(Y = j|Z = 1) + \mathbb{P}(Z = 2)\mathbb{P}(Y = j|Z = 2)\right) \\
&= \frac{25}{36}\mathbb{E}[Y|Z = 0] + \frac{10}{36}\mathbb{E}[Y|Z = 1] + \frac{1}{36}\mathbb{E}[Y|Z = 2].
\end{aligned}$$

Ora osserviamo che $\mathbb{E}[Y|Z = 0] = 7$, $\mathbb{E}[Y|Z = 1] = 6 + 7/2 = 19/2$ e $\mathbb{E}[Y|Z = 2] = 6 + 6 = 12$. In conclusione

$$\mathbb{E}[Y] = \frac{175}{36} + \frac{95}{36} + \frac{12}{36} = \frac{282}{36} \approx 7.8$$

Nome:______________________________________

4. **(7 punti)** Siano $X_1, X_2, \dots$ esponenziali di parametro 1 indipendenti. Calcolare:

   (a) $\mathbb{E}[X_1 + X_2 + 2X_3]$

   (b) $\mathrm{Cov}(X_1 + X_2, X_1 + 2X_3)$

   (c) La probabilità dell'evento $A = \{X_1 + X_2 > 2X_3\}$

Soluzione: (a). Le X_i hanno valore atteso $\mathbb{E}[X_i] = 1$. Allora

$$\mathbb{E}[X_1 + X_2 + 2X_3] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + 2\mathbb{E}[X_3] = 4.$$

(b). Per linearità si ha anche

$$\mathrm{Cov}(X_1 + X_2, X_1 + 2X_3) = \mathrm{Cov}(X_1, X_1) + 2\mathrm{Cov}(X_1, X_3) + \mathrm{Cov}(X_2, X_1) + 2\mathrm{Cov}(X_2, X_3).$$

Per l'indipendenza

$$\mathrm{Cov}(X_1, X_3) = \mathrm{Cov}(X_2, X_1) = \mathrm{Cov}(X_2, X_3) = 0.$$

Allora

$$\mathrm{Cov}(X_1 + X_2, X_1 + 2X_3) = \mathrm{Cov}(X_1, X_1) = \mathrm{Var}(X_1) = 1,$$

dove abbiamo usato il fatto che la varianza di un'esponenziale di parametro λ vale $1/\lambda^2$.

(c). Notiamo che $Y = X_1 + X_2$ è una variabile gamma di parametri $\alpha = 2, \lambda = 1$, con densità $f_Y(y) = \frac{ye^{-y}}{\Gamma(2)}\mathbf{1}(y > 0) = ye^{-y}\mathbf{1}(y > 0)$. Inoltre $Z = 2X_3$ è una variabile esponenziale di parametro $\lambda = 1/2$. Essendo Y, Z indipendenti, la loro densità congiunta è

$$f_{Y,Z}(y, z) = f_Y(y) f_Z(z) = ye^{-y}\mathbf{1}_{y>0}\frac{1}{2}e^{-z/2}\mathbf{1}_{z>0}.$$

Allora,

$$\begin{aligned}\mathbb{P}(X_1 + X_2 > 2X_3) &= \mathbb{P}(Y > Z) \\ &= \int_0^\infty dz \int_z^\infty dy f_{Y,Z}(y, z) = \frac{1}{2}\int_0^\infty dz\, e^{-z/2}\int_z^\infty dy\, ye^{-y}\end{aligned}$$

Calcoliamo

$$\int_z^\infty dy\, ye^{-y} = (z + 1)e^{-z}.$$

In conclusione,

$$\mathbb{P}(X_1 + X_2 > 2X_3) = \frac{1}{2}\int_0^\infty dz(z + 1)e^{-3z/2} = \frac{5}{9}.$$

Nome:____

5. **(7 punti)** Consideriamo i lanci di un dado. Sia T il numero di lanci necessario per vedere tutte e 6 le facce del dado.

   (a) Qual è il valore atteso di T ?

   (b) Utilizzando la disuguaglianza di Markov, trovare un valore $t > 0$ tale che l'evento $T \geqslant t$ ha probabilità minore di 1/10.

Soluzione: (a). Come nel problema della raccolta di figurine si ha

$$T = 1 + X_1 + X_2 + X_3 + X_4 + X_5,$$

dove X_i è il numero di lanci necessari per passare dall'avere visto i facce differenti all'avere visto $i+1$ facce differenti. Notiamo che X_i è una geometrica di parametro $p_i = (6-i)/6$. Allora

$$\mathbb{E}[T] = 1 + \sum_{i=1}^{5} \mathbb{E}[X_i] = 1 + \sum_{i=1}^{5} \frac{1}{p_i} = 1 + \frac{6}{5} + \frac{6}{4} + \frac{6}{3} + \frac{6}{2} + 6 = \frac{147}{10}$$

(b). La disuguaglianza di Markov permette di stimare

$$\mathbb{P}(T \geqslant t) \leqslant \frac{\mathbb{E}[T]}{t},$$

per ogni $t > 0$. Allora basta prendere $t = 10\mathbb{E}[T] = 147$ e si ha

$$\mathbb{P}(T \geqslant 147) \leqslant \frac{1}{10}.$$

Osservazione: con un calcolo della varianza, usando Chebyshev invece di Markov possiamo ottenere una stima migliore per t. Infatti, per il calcolo della varianza osserviamo che le X_i sono tutte indipendenti e dunque

$$\mathrm{Var}(T) = \sum_{i=1}^{5} \mathrm{Var}(X_i) = \sum_{i=1}^{5} \frac{1-p_i}{p_i^2} = \sum_{i=1}^{5} \frac{6}{(6-i)^2} = \frac{5269}{600}.$$

La disuguaglianza di Chebyshev stabilisce che

$$\mathbb{P}(|T - \mathbb{E}[T]| \geqslant \delta) \leqslant \frac{1}{\delta^2} \mathrm{Var}(T),$$

per ogni $\delta > 0$. Notiamo che $T \geqslant 25$ implica $|T - \mathbb{E}[T]| \geqslant 10$, e dunque

$$\mathbb{P}(T \geqslant 25) \leqslant \mathbb{P}(|T - \mathbb{E}[T]| \geqslant 10) \leqslant \frac{1}{100} \mathrm{Var}(T) = \frac{1}{100} \frac{5269}{600} \leqslant \frac{1}{10}.$$

Quindi abbiamo ottenuto che $t = 25$ è sufficiente per per avere $\mathbb{P}(T \geqslant t) \leqslant 1/10$. Come regola generale: piu' alti sono i momenti e migliore è il controllo delle probabilità delle deviazioni.

Nome:_______________________________________