Corso di Statistica - Prof. Fabio Zucca	I Prova in itinere - 5 maggio 2017

Nome e cognome:		**Matricola:**	

 12957

Esercizio 1 Si consideri la seguente famiglia di funzioni di parametro $c \in \mathbb{R}$

$$f_c(x) := \begin{cases} c/x & 1 \leq x \leq e^2 \\ 0 & \text{altrimenti} \end{cases}$$

1. Determinare per quale valore del parametro c la funzione f_c rappresenta la densità di una variabile aleatoria assolutamente continua X.
2. Calcolare valore atteso e varianza di X.
3. Calcolare la funzione di ripartizione F di X.
4. Calcolare $\mathbb{P}(\sqrt{e} \leq X \leq e)$ e $\mathbb{P}\left(e \leq X \leq e^{5/2}\right)$.
5. Calcolare per ogni $\alpha \in (0,1)$ il valore $q_\alpha :=\min\{t : F(t) \geq \alpha\}$.

Soluzione.

1. Facilmente
$$1=\int_{\mathbb{R}} f_c(x)\mathrm{d}x=\int_1^{e^2}\frac{c}{x}\mathrm{d}x=\ln(|x|)\Big|_1^{e^2}=2c$$
da cui si ha l'unica soluzione $c=1/2$. In questo caso poichè la funzione ha segno costante (dove è differente da zero), si ha che se l'integrale è uguale a 1 allora la funzione integranda è nonnegativa; ad ogni modo si verifica immediatamente che la funzione f_c è non negativa se e solo se $c\geq 0$.

2. $$E[X]=\int_{\mathbb{R}} x f_{1/2}(x)\mathrm{d}x=\int_1^{e^2}\frac{\mathrm{d}x}{2}=\frac{x}{2}\Big|_1^{e^2}=\frac{e^2-1}{2}$$
Ricordando che $\mathrm{var}(x)=\mathbb{E}[X^2]-[EX]^2$ calcoliamo
$$\mathbb{E}[X^2]=\int_{\mathbb{R}} x^2 f_{1/2}(x)\mathrm{d}x=\int_1^{e^2}\frac{x}{2}\mathrm{d}x=\frac{x^2}{2\cdot 2}\Big|_1^{e^2}=\frac{e^4-1}{4}$$
da cui $\mathrm{var}(X)=\frac{e^4-1}{4}-\left(\frac{e^2-1}{2}\right)^2\approx 3.2$

3. Dalla definizione
$$F_X(t):=\int_{-\infty}^t f_{1/2}(x)\mathrm{d}x=\begin{cases}0 & t<1\\ \int_1^t\frac{\mathrm{d}x}{2x} & 1\leq t<e^2\\ 1 & t>e^2\end{cases}=\begin{cases}0 & t<1\\ \frac{\ln(t)}{2} & 1\leq t<e^2\\ 1 & t>e^2\end{cases}$$

4. Ricordiamo che per una variabile aleatoria assolutamente continua $\mathbb{P}(\sqrt{e}\leq X\leq e)=\mathbb{P}(\sqrt{e}<X\leq e)=F_X(e)-F_X(\sqrt{e})=1/4$. Nel secondo caso avremo $\mathbb{P}\left(e\leq X\leq e^{5/2}\right)=F_X(e^{5/2})-F_X(e)=1-1/2=1/2$.

5. Utilizzando la continuità della F_X e la stretta monotonia su $(1,e^2)$, si ha che F_X è biettiva da $(1,e^2)$ su $(0,1)$ pertanto $F_X(q_\alpha)=\alpha$. In altre parole $t=q_\alpha$ è la soluzione dell'equazione
$$\frac{\ln(t)}{2}=\alpha$$
cioé $q_\alpha=e^{2\alpha}$

Esercizio 2 I giorni di pioggia in agosto nelle zone A e B sono ben descritti da variabili aleatorie $X \sim \mathcal{P}(\lambda_A)$ e $Y \sim \mathcal{P}(\lambda_B)$ (variabili relative a zone o ad anni differenti, si possono considerare indipendenti). Si sa che la probabilità di non avere alcun giorno di pioggia è $\exp(-3)$ nella zona A e $\exp(-1)$ nella zona B.

1. Qual è la probabilità di avere esattamente 4 giorni di pioggia nella zona A? E nella zona B?
2. Qual è la probabilità che vi siano, complessivamente nelle due zone, almeno 2 giorni di pioggia?
3. Qual è la probabilità che in 10 anni differenti vi siano esattamente 2 mesi di agosto senza pioggia nella zona A?
4. Abbiamo 3 fogli contenenti, rispettivamente, il numero di giorni di pioggia di (1) agosto 2015 nella zona A, (2) agosto 2016 nella zona A e (3) agosto 2016 nella zona B. I fogli non sono distinguibili; ne peschiamo uno a caso con ugual probabilità e leggiamo che ci sono stati 2 giorni di pioggia. Qual è la probabilità che si riferisca alla zona A? E alla zona B?
5. Calcolare, ricorrendo ad un'opportuna approssimazione, la probabilità che in 100 anni, nei soli mesi di agosto, vi siano almeno 305 giorni di pioggia nella zona A.

Soluzione. Per una Poissoniana di parametro λ si sa che la probabilità di essere uguale a 0 è $\exp(-\lambda)$ pertanto $\lambda_A = 3$ e $\lambda_B = 1$.

1. Si ha $\mathbb{P}(X = 4) = \exp(-3)\cdot 3^4/4! \approx 0.168031356$. Similmente $\mathbb{P}(Y = 4) = \exp(-1)\cdot 1^4/4! \approx 0.01532831$.

2. Dall'indipendenza si ha $X+Y \sim \mathcal{P}(4)$ da cui $\mathbb{P}(X+Y \geq 2) = 1-\mathbb{P}(X+Y \leq 1) = 1-\exp(-4)(1+4) = 0.908421806$.

   Equivalentemente si sarebbe potuto procedere come segue: $\mathbb{P}(X + Y \geq 2) = 1 - \mathbb{P}(X + Y \leq 1) = 1-\mathbb{P}(X+Y = 0)-\mathbb{P}(X+Y = 1) = 1-\mathbb{P}(X = 0)\mathbb{P}(Y = 0)-\mathbb{P}(X = 1)\mathbb{P}(Y = 0)-\mathbb{P}(X = 0)\mathbb{P}(Y = 1)$.

3. Siano $\{Z_i\}_{i=1}^{10}$ le variabili di Bernoulli che ci dicono se non vi sono stati giorni di pioggia nei vari mesi di agosto nella zona A. Evidentemente $Z_i \sim \mathcal{B}(\exp(-3))$ quindi $Z := \sum_i Z_i \sim \mathcal{B}(10, \exp(-3))$ da cui $\mathbb{P}(Z = 2) = \binom{10}{2} \exp(-3 \cdot 2)(1 - \exp(-3))^8 \approx 0.074133263$.

4. Sia $A :=$"prendo un foglio relativo alla zona A" e $C :=$"vi sono 2 giorni di pioggia sul foglio scelto". Sappiamo che $\mathbb{P}(A) = 2/3$, $\mathbb{P}(C|A) = \exp(-3) \cdot 3^2/2! \approx 0.224041808$ e $\mathbb{P}(C|A^c) = \exp(-1) \cdot 1^2/2! \approx 0.183939721$. Pertanto, dalla formula di Bayes, $\mathbb{P}(A|C) = \mathbb{P}(C|A)\mathbb{P}(A)/(\mathbb{P}(C|A)\mathbb{P}(A)+\mathbb{P}(C|A^c)\mathbb{P}(A^c)) = 0.224041808 \cdot (2/3)/(0.224041808 \cdot 2/3 + 0.183939721/3) = 0.708966884$. Ovviamente $\mathbb{P}(A^c|C) = 1 - \mathbb{P}(A|C)$.

5. Se X_i sono i giorni di pioggia del mese di agosto nell'i-esimo anno, allora $\sum_{i=1}^{100} X_i \sim \mathcal{P}(100 \cdot 3)$; pertanto, essendo $200 \geq 5$, dal TCL si ha che $\sum_{i=1}^{100} X_i \approx Z \sim \mathcal{N}(300, 300)$. Da cui $\mathbb{P}\Big(\sum_{i=1}^{100} X_i \geq 305\Big) = \mathbb{P}\Big(\sum_{i=1}^{100} X_i \geq 305 - 0.5\Big) \approx \mathbb{P}\Big((Z - 300)/\sqrt{300} \geq 4.5/\sqrt{300}\Big) = 1 - \Phi(4.5/\sqrt{300}) = 1 - \Phi(0.2598076211) = 1 - 0.602493914 = 0.397506086$.