II Appello di Probabilità e Statistica	Cognome: ______________________
Laurea in Matematica	Nome: ______________________
2 luglio 2009	Matricola: ______________________

ESERCIZIO 1. Per una certa specie africana di uccelli, i neonati hanno – indipendentemente l'uno dal l'altro – una probabilità di sopravvivere al primo mese pari a 1/7. Quelli che sopravvivono al primo mese hanno una probabilità pari a 1/3 di superare l'anno.

(a) Qual è la probabilità che un neonato sopravviva al primo anno?

(b) Se un neonato muore entro il primo anno, qual è la probabilità che sia sopravvissuto al primo mese?

(c) Si determini approssimativamente il numero minimo n di neonati da monitorare affinché la probabilità che ne sopravvivano almeno 30 dopo un mese sia almeno 0.95.

SOLUZIONE.

(a) Introducendo gli eventi $A :=$ "il neonato sopravvive al primo mese" e $B :=$ "il neonato sopravvive al primo anno", i dati del problema ci dicono che

$$P(A) = \frac{1}{7}, \qquad P(B|A) = \frac{1}{3}.$$

Di conseguenza, per il teorema delle probabilità totali,

$$P(B) = P(B|A)P(A) + P(B|A^c)P(A^c) = P(B|A)P(A) = \frac{1}{3}\frac{1}{7} = \frac{1}{21},$$

dove si è usato il fatto che $P(B|A^c) = 0$, perché $B \subseteq A$.

(b) Per la formula di Bayes

$$P(A|B^c) = \frac{P(B^c|A)P(A)}{P(B^c)} = \frac{(1-\frac{1}{3})\frac{1}{7}}{1-\frac{1}{21}} = \frac{1}{10}.$$

(c) Dati n neonati, il numero di questi che sopravvive dopo un mese è una variabile casuale X con distribuzione $B(n,p)$, con $p = \frac{1}{7}$. Applicando l'approssimazione normale e la correzione di continuità, si ottiene

$$P(X \geq 30) = P(X \geq 29.5) = P\left(\frac{\frac{X}{n}-p}{\sqrt{p(1-p)}}\sqrt{n} \geq \frac{\frac{29.5}{n}-\frac{1}{7}}{\sqrt{\frac{1}{7}(1-\frac{1}{7})}}\sqrt{n}\right) \simeq \Phi\left(\frac{\frac{1}{7}-\frac{29.5}{n}}{\sqrt{6}/7}\sqrt{n}\right).$$

Dato che $\Phi(z) \geq 0.95$ se e solo se $z \geq \Phi^{-1}(0.95) \simeq 1.64$, si ottiene la disequazione

$$\frac{\frac{1}{7}-\frac{29.5}{n}}{\sqrt{6}/7}\sqrt{n} \geq 1.64 \qquad \Longleftrightarrow \qquad n^2 - 1.64\sqrt{6}\sqrt{n} - 29.5 \cdot 7 \geq 0.$$

Le soluzioni positive sono date da $\sqrt{n} \geq 16.52$, cioè $n \geq 273$.

ESERCIZIO 2. Al tavolo di un bistrot alcuni avventori giocano a dadi. Ciascun giocatore ha a disposizione un dado equilibrato a sei facce: se ottiene 5 oppure 6, lancia nuovamente il dado; la prima volta che ottiene 1, 2, 3 oppure 4, passa il turno al giocatore seguente. Il punteggio ottenuto da un giocatore in un turno è pari alla somma degli esiti dei lanci effettuati.

(a) Si determini la probabilità che il turno di un giocatore duri n lanci, per $n \in \{1, 2, 3, \ldots\}$.

(b) Se il turno di un giocatore dura due lanci, qual è la probabilità che il punteggio da lui ottenuto sia dispari?

(c) Qual è la probabilità che un giocatore ottenga un punteggio pari a 5? E pari a 11?

SOLUZIONE.

(a) Il turno di un giocatore dura n lanci se nei primi $n-1$ lanci esce 5 oppure 6, mentre nell'n-esimo lancio esce un numero tra 1 e 4. In ogni lancio, la probabilità di ottenere un numero tra 1 e 4 vale $\frac{2}{3}$, mentre la probabilità di ottenere 5 oppure 6 vale $\frac{1}{3}$. Dato che i lanci sono indipendenti, segue che la probabilità che il turno duri n lanci vale $(\frac{1}{3})^{n-1}\frac{2}{3}$.

(b) Sia $\Omega = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$ lo spazio campionario che descrive gli esiti possibili di due lanci di dadi, munito della probabilità uniforme. Introdotti gli eventi A = "il turno dura due lanci" e B = "il punteggio ottenuto è pari" si ha

$$A = \big\{(i, j) \in \Omega : i \in \{5, 6\} \text{ e } j \in \{1, 2, 3, 4\}\big\},$$

$$A \cap B = \big\{(i, j) \in \Omega : i \in \{5, 6\},\ j \in \{1, 2, 3, 4\} \text{ e } i + j \text{ è pari}\big\}.$$

Chiaramente $|A| = 2 \cdot 4 = 8$, $|A \cap B| = 4$ e dunque

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{|A \cap B|}{|A|} = \frac{1}{2}.$$

(c) Non è possibile che il punteggio totale sia pari a 5 (la probabilità richiesta è dunque pari a zero). Infatti, se il primo lancio dà come esito un numero tra 1 e 4, il turno termina con un punteggio minore di 5; se invece l'esito del primo lancio è 5 oppure 6, bisogna effettuare (almeno) un altro lancio e il punteggio totale è dunque maggiore di 5.

Affinché il punteggio totale sia 11, il primo lancio deve dare come esito 5 oppure 6 (altrimenti il turno finirebbe). Tuttavia, se l'esito del primo lancio fosse 6, non sarebbe possibile ottenere come punteggio totale 11, poiché:

- se il secondo lancio desse come esito un numero tra 1 e 4, il turno terminerebbe senza raggiungere 11;
- se desse come esito 5, il punteggio salirebbe a 11 ma si dovrebbe tirare il dado un'altra volta, superando dunque il valore di 11;
- se desse come esito 6, il valore 11 sarebbe già superato al secondo lancio.

Quindi, per ottenere un punteggio totale pari a 11, il primo lancio deve dare come esito 5. Con analoghi ragionamenti, ci si convince che il secondo lancio deve necessariamente dare come esito 5 e il terzo lancio 1. In altri termini, il punteggio totale è pari a 11 se e soltanto se il primo lancio dà come esito 5, il secondo lancio dà come esito 5 e il terzo lancio dà come esito 1: la probabilità richiesta vale dunque $\frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{216}$.

ESERCIZIO 3. Sia $n \in \mathbb{N}$, $n \geq 3$ e indichiamo con S_n il gruppo delle permutazioni di $\{1, \ldots, n\}$, munito della probabilità P uniforme. Gli elementi di S_n saranno indicati con $\sigma = (\sigma(1), \ldots, \sigma(n))$. Introduciamo le variabili casuali scalari X, Y definite su S_n:

$$X(\sigma) \ := \ \sigma(1) \,, \qquad Y(\sigma) \ := \ \sigma(2) \,.$$

(a) Si mostri che, per ogni $i, j \in \{1, 2, \ldots, n\}$, la densità congiunta di (X, Y) è data da

$$p_{X,Y}(i,j) \ = \ \begin{cases} \dfrac{1}{c_n} & \text{se } i \neq j \\ 0 & \text{se } i = j \end{cases} \,,$$

dove c_n è un'opportuna costante che è richiesto di determinare.

(b) (*) Si determini la densità della variabile $D := Y - X$.
[*Sugg:* basta calcolare $p_D(m)$ per $m > 0$, poiché per simmetria $p_D(-m) = p_D(m)$.]

Indichiamo ora con Z, W due variabili casuali scalari indipendenti, definite su un altro spazio di probabilità $(\Omega, \widetilde{P})$, ciascuna con distribuzione uniforme nell'insieme $\{1, \ldots, n\}$: in altri termini, $\widetilde{P}(Z = i) = \frac{1}{n}$, per ogni $i \in \{1, \ldots, n\}$, e analogamente per W.

(c) Si calcoli $\widetilde{P}(Z \neq W)$.

(d) Si mostri che, per ogni $i, j \in \{1, 2, \ldots, n\}$, si ha che $\widetilde{P}(Z = i, W = j \,|\, Z \neq W) = p_{X,Y}(i, j)$.

SOLUZIONE.

(a) Per definizione di spazio di probabilità uniforme

$$p_{X,Y}(i,j) \ = \ P(X = i, Y = j) \ = \ \frac{|\{\sigma \in S_n : \ \sigma(1) = j, \, \sigma(2) = j\}|}{|S_n|} \,.$$

Chiaramente $p_{X,Y}(i,j) = 0$ se $i = j$, perché le permutazioni sono biunivoche e dunque non si può avere $\sigma(1) = \sigma(2)$. Per $i \neq j$, le permutazioni $\sigma \in S_n$ tali che $\sigma(1) = j$, $\sigma(2) = j$ sono in corrispondenza con le applicazioni biunivoche da $\{3, \ldots, n\}$ a valori in $\{1, \ldots, n\} \setminus \{i, j\}$, di conseguenza $|\{\sigma \in S_n : \ \sigma(1) = j, \, \sigma(2) = j\}| = (n-2)!$ e si ottiene

$$p_{X,Y}(i,j) \ = \ \frac{(n-2)!}{n!} \ = \ \frac{1}{c_n} \,, \qquad \text{dove } \ c_n \ := \ n(n-1) \,.$$

(Si noti che $c_n = |\{(i,j) \in \{1, \ldots, n\} \times \{1, \ldots, n\} : \ i \neq j\}| = |(X,Y)(S_n)|$.)

(b) Chiaramente i valori possibili per la variabile D sono $D(S_n) = \{-(n-1), \ldots, (n-1)\} \setminus \{0\}$. Per $m \in D(S_n)$, con $m > 0$, si ha

$$\{D = m\} \ = \ \bigcup_{k=1}^{n-m} \{X = k, Y = k + m\} \,,$$

e dato che gli eventi che appaiono nell'unione sono disgiunti segue che per ogni $m \in \{1, \ldots, n-1\}$

$$P(D = m) \ = \ \sum_{k=1}^{n-m} P(X = k, Y = k + m) \ = \ \frac{1}{c_n} \, (n - m) \ = \ \frac{n - m}{n(n-1)} \,.$$

Con analoghi argomenti (oppure per simmetria) si ha $P(D = m) = P(D = -m)$ per $m < 0$, per cui la formula generale è

$$P(D = m) \; = \; \frac{n - |m|}{n(n-1)} \, ,$$

per ogni $m \in \{-(n-1), \ldots, (n-1)\} \setminus \{0\}$.

(c)

$$\widetilde{P}(Z \neq W) \; = \; 1 - \widetilde{P}(Z = W) \; = \; 1 - \sum_{i=1}^{n} \widetilde{P}(Z = i, W = i) \; = \; 1 - n \cdot \frac{1}{n^2} \; = \; 1 - \frac{1}{n} \, .$$

(d) Chiaramente $\widetilde{P}(Z = i, W = j \,|\, Z \neq W) = 0$ se $i = j$. Per $i \neq j$, con $i, j \in \{1, \ldots, n\}$, si ha

$$\begin{aligned} \widetilde{P}(Z = i, W = j \,|\, Z \neq W) \; &= \; \frac{\widetilde{P}(Z = i, W = j)}{\widetilde{P}(Z \neq W)} \; = \; \frac{\widetilde{P}(Z = i)\widetilde{P}(W = j)}{\widetilde{P}(Z \neq W)} \; = \; \frac{\frac{1}{n}\frac{1}{n}}{1 - \frac{1}{n}} \\ &= \; \frac{1}{n(n-1)} \; = \; \frac{1}{c_n} \; = \; p_{X,Y}(i,j) \, . \end{aligned}$$

ESERCIZIO 4. Sia $X \sim U(0,1)$. Per $x \in \mathbb{R}$ poniamo $g(x) := 4x(1-x)$ e definiamo $Y := g(X)$.

(a) Si determini la funzione di ripartizione di Y, si deduca che la variabile Y è assolutamente continua e se ne calcoli la densità.

(b) Si calcoli $Cov(X,Y)$.

SOLUZIONE.

(a) Per $y \in \mathbb{R}$, la disequazione $4x(1-x) \leq y$ ha come soluzioni $x \leq \frac{1}{2}(1-\sqrt{1-y})$ oppure $x \geq \frac{1}{2}(1+\sqrt{1-y})$ se $y \leq 1$, mentre ha come soluzione tutta la retta reale se $y > 1$. Di conseguenza,

$$F_Y(y) \;=\; P(Y \leq y) \;=\; \begin{cases} P\left(X \leq \frac{1}{2}(1-\sqrt{1-y})\right) \;+\; P\left(X \geq \frac{1}{2}(1+\sqrt{1-y})\right) & \text{se } y \leq 1 \\ 1 & \text{se } y > 1 \end{cases} .$$

Si noti che $\frac{1}{2}(1+\sqrt{1-y}) > 1$ se $y < 0$, per cui in questo caso $P\left(X \geq \frac{1}{2}(1+\sqrt{1-y})\right) = 0$. Analogamente, per $y < 0$ si ha $P\left(X \leq \frac{1}{2}(1-\sqrt{1-y})\right) = 0$, poiché $\frac{1}{2}(1-\sqrt{1-y}) < 0$. Di conseguenza $F_Y(y) = 0$ per $y < 0$ e possiamo restringere[1] l'attenzione ai valori $y \in [0,1]$.

Dato che per $z \in [0,1]$ si ha $P(X \geq z) = 1 - z$ e $P(X \leq z) = z$, si ottiene

$$\begin{aligned} F_Y(y) \;&=\; P\left(X \leq \frac{1}{2}(1-\sqrt{1-y})\right) \;+\; P\left(X \geq \frac{1}{2}(1+\sqrt{1-y})\right) \\ &=\; 1 \;-\; \frac{1}{2}(1+\sqrt{1-y}) \;+\; \frac{1}{2}(1-\sqrt{1-y}) \;=\; 1 - \sqrt{1-y}\,, \qquad \forall y \in [0,1]\,. \end{aligned}$$

In definitiva,

$$F_Y(y) \;=\; P(Y \leq y) \;=\; \begin{cases} 0 & \text{se } y < 0 \\ 1 - \sqrt{1-y} & \text{se } y \in [0,1] \\ 1 & \text{se } y > 1 \end{cases} .$$

La funzione $F_Y(\cdot)$ è dunque C^1 a tratti, di conseguenza la variabile Y è assolutamente continua. La sua densità è data da

$$f_Y(y) \;=\; F_Y'(y) \;=\; \frac{1}{2}\,\frac{1}{\sqrt{1-y}}\,\mathbf{1}_{[0,1]}(y)\,.$$

(b) Per le proprietà del valor medio

$$E(XY) \;=\; E(X \cdot 4X(1-X)) \;=\; 4E(X^2) - 4E(X^3)\,,$$

$$E(X)E(Y) \;=\; E(X)E(4X(1-X)) \;=\; 4E(X)^2 - 4E(X)E(X^2)\,,$$

quindi

$$Cov(X,Y) \;=\; E(XY) - E(X)E(Y) \;=\; 4E(X^2)(1+E(X)) - 4(E(X^3) + E(X)^2)\,.$$

Dato che $E(X^n) = \int_0^2 x^n \, \mathrm{d}x = \frac{1}{n+1}$, si ha $E(X) = \frac{1}{2}$, $E(X^2) = \frac{1}{3}$ e $E(X^3) = \frac{1}{4}$, da cui

$$Cov(X,Y) \;=\; \frac{4}{3}\left(1 + \frac{1}{2}\right) \;-\; 4\left(\frac{1}{4} + \frac{1}{4}\right) \;=\; 0\,.$$

[1] Questo si sarebbe potuto dedurre subito, considerando che la funzione g manda l'intervallo $[0,1]$ in $[0,1]$.