Prova Scritta di Probabilità e Statistica	Cognome: ____________________
Laurea in Matematica	Nome: ____________________
28 giugno 2012	Matricola: ____________________

ESERCIZIO 1. Sia $(A_n)_{n\geq 2}$ una successione di eventi indipendenti, tali che $P(A_n) \leq 1 - \frac{1}{n}$. Sia $B := \bigcap_{n=2}^{+\infty} A_n$. Mostrare che $P(B) = 0$.

SOLUZIONE. Sia $B_n := \cap_{k=2}^{n} A_k$. Per l'ipotesi di indipendenza,

$$P(B_n) = \prod_{k=2}^{n} P(A_k) \leq \prod_{k=2}^{n} \left(1 - \frac{1}{k}\right) = \prod_{k=2}^{n} \frac{k-1}{k} = \frac{1}{n}.$$

Si noti che (B_n) è una successione decrescente di eventi, e $B = \cap_n B_n$, e quindi $P(B) = \lim_n P(B_n)$. La conclusione segue subito dalla disuguaglianza precedente.

ESERCIZIO 2. Diciamo che una funzione $f : \mathbb{R} \to \mathbb{R}$ è *fortemente positiva* se:

- è derivabile infinite volte;
- $f^{(n)}(x) > 0$ per ogni $x \in \mathbb{R}$ e ogni $n \geq 1$;
- per ogni $x \in \mathbb{R}$
$$f(x) = \sum_{n=0}^{+\infty} f^{(n)}(0)x^n.$$

Sia X una variabile casuale scalare tale che $f(|X|) \in L^1$ per una funzione f fortemente positiva. Mostrare che $X \in L^p$ per ogni $p \geq 1$.

SOLUZIONE. Per ipotesi, per ogni $x \in \mathbb{R}$ e $m \in \mathbb{N}$,

$$f(|x|) - f(0) = \sum_{n=0}^{+\infty} f^{(n)}(0)|x|^n \geq f^{(m)}(0)|x|^m.$$

Perciò
$$|X|^m \leq \frac{f(|X|) - f(0)}{f^{(m)}(0)}.$$

Quest'ultima variabile è in L^1, poiché per ipotesi $f(|X|) \in L^1$. Ma allora $X \in L^m$ per ogni $m \in \mathbb{N}$. Se $p \geq 1$, esiste $m \in \mathbb{N}$ tale che $p < m$, e perciò $L^p \subseteq L^m$. Pertanto $X \in L^p$ per ogni $p \geq 1$.

ESERCIZIO 3. Sia Y una variabile casuale a valori in $[0, +\infty)$, e $f : [0, +\infty) \to [0, +\infty)$ una funzione strettamente crescente. Mostrare che per ogni $\epsilon > 0$

$$P(Y \geq \epsilon) \leq \frac{E[f(Y)]}{f(\epsilon)}.$$

SOLUZIONE. Vale la disuguaglianza

$$f(Y) \geq f(\epsilon)\mathbf{1}_{[\epsilon,+\infty)}(Y),$$

che segue dal fatto che f è crescente. Prendendo il valor medio di ambo i membri (si può fare poiché si tratta di variabili non negative), si ottiene

$$E[f(Y)] \geq f(\epsilon)P(Y \geq \epsilon),$$

da cui si conclude.

ESERCIZIO 4. Un'urna contiene N palline numerate da 1 a N. Si estraggono due palline (senza reimmissione). Sia X il più piccolo fra i due numeri estratti, e Y il più grande. Calcolare $Cov(X, Y)$. (Sugg: ricordare che $\sum_1^N i = \frac{N(N+1)}{2}$ e $\sum_1^N i^2 = \frac{N(N+1)(2N+1)}{6}$)

SOLUZIONE. Anzitutto

$$p_{X,Y}(m,n) = \begin{cases} \frac{1}{\binom{N}{2}} & \text{se } 1 \leq m < n \leq N \\ 0 & \text{altrimenti.} \end{cases}$$

Perciò

$$p_X(m) = \sum_{n=m+1}^{N} \frac{1}{\binom{N}{2}} = \frac{N-m}{\binom{N}{2}}.$$

Quindi

$$E(X) = \frac{1}{\binom{N}{2}} \sum_{m=1}^{N-1} m(N-m) = \frac{1}{\binom{N}{2}} \left[\frac{N^2(N-1)}{2} - \frac{N(N-1)(2N-1)}{6} \right] = \frac{N+1}{3}.$$

Con un calcolo analogo, o usando il fatto che per simmetria $E(Y) = N + 1 - E(X)$, si trova $E(Y) = \frac{2(N+1)}{3}$. Inoltre

$$\begin{aligned} E(XY) = \frac{1}{\binom{N}{2}} \sum_{1 \leq m < n \leq N} mn &= \frac{1}{\binom{N}{2}} \sum_{n=2}^{N} n \sum_{m=1}^{n-1} m \\ &= \frac{1}{\binom{N}{2}} \sum_{n=2}^{N} \frac{n^2(n-1)}{2} = \frac{1}{\binom{N}{2}} \sum_{n=1}^{N} \frac{n^2(n-1)}{2} \\ &= \frac{1}{N(N-1)} \sum_{n=1}^{N} n^3 - \frac{(N+1)(2N+1)}{6(N-1)}. \end{aligned}$$

Mettendo assieme quanto trovato si trova un'espressione per $Cov(X, Y) = E(XY) - E(X)E(Y)$. Un'espressione più semplice si ottiene usando l'identità, che si mostra facilmente per induzione, $\sum_1^N n^3 = \frac{N^2(N+1)^2}{4}$.

ESERCIZIO 5. Da un mazzo di 52 carte da poker viene estratta a caso una carta, ed inserita in un altro mazzo (sempre da poker). Questo secondo mazzo viene poi accuratamente mescolato.

(a) Si gira la prima carta del secondo mazzo. Questa è un asso di cuori. Qual è la probabilità, condizionata a tale informazione, che la carta inserita fosse un asso di cuori?

(b) Si girano le prime due carte del secondo mazzo. Queste sono un asso di cuori e un tre di picche. Qual è la probabilità, condizionata a tale informazione, che la carta inserita non fosse né un asso di cuori né un tre di picche?

SOLUZIONE.

(a) Consideriamo gli eventi A = "la carta inserita è l'asso di cuori", B = "la prima carta girata è l'asso di cuori". Abbiamo:

$$P(A) = \frac{1}{52} \quad P(B|A) = \frac{2}{53} \quad P(B|A^c) = \frac{1}{53}.$$

Perciò

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} = \frac{2}{53}.$$

(b) Consideriamo gli eventi A_1 = "la carta inserita è l'asso di cuori", A_2 = "la carta inserita è il tre di picche" B = "le prime due carte girate sono l'asso di cuori e il tre di picche. Abbiamo:

$$P(A_1) = P(A_2) = \frac{1}{52} \quad P(B|A_1 \cup A_2) = P(B|A_2) = \frac{2}{\binom{53}{2}} \quad P(B|A_1^c) = P(B|A_2^c) = \frac{2}{\binom{53}{2}}.$$

Da ciò si ricava

$$P(B|A_1 \cup A_2) = \frac{P(B \cap A_1) + P(B \cap A_2)}{P(A_1) + P(A_2)} = \frac{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)}{P(A_1) + P(A_2)} = \frac{2}{\binom{53}{2}}.$$

Inoltre

$$P(B|(A_1 \cup A_2)^c) = \frac{1}{\binom{53}{2}}.$$

Perciò

$$P(A_1 \cup A_2|B) = \frac{P(B|A_1 \cup A_2)P(A_1 \cup A_2)}{P(B|A_1 \cup A_2)P(A_1 \cup A_2) + P(B|(A_1 \cup A_2)^c)P((A_1 \cup A_2)^c)} = \frac{2}{27}$$

da cui

$$P((A_1 \cup A_2)^c|B) = \frac{25}{27}.$$

ESERCIZIO 6. Sia X una variabile casuale a valori in $\mathbb{R}$ la cui funzione di ripartizione è data da

$$F_X(x) = \begin{cases} 1 - e^{-x^2/2\sigma^2} & \text{se } x \geq 0 \\ 0 & \text{altrimenti.} \end{cases}$$

Posto $Y := X^2$, si determini media e varianza di Y.

SOLUZIONE. Sia $y > 0$.

$$F_Y(y) = P(X \leq \sqrt{y}) = 1 - e^{-y/2\sigma^2}.$$

Chiaramente $F_Y(y) = 0$ se $y \leq 0$. Dunque F_Y è $\mathcal{C}^1$ a tratti, da cui segue che Y è assolutamente continua, con densità

$$f_Y(y) = F_Y'(y) = \frac{1}{2\sigma^2} e^{-y/2\sigma^2} \mathbf{1}_{(0,+\infty)}(y).$$

Quindi $Y \sim Exp(1/2\sigma^2)$, da cui

$$E(Y) = 2\sigma^2, \quad Var(Y) = 4\sigma^4.$$

ESERCIZIO 7. Sia $\Theta \sim U(0, \pi)$. Su un piano cartesiano Oxy si tracci da O una semiretta che forma un angolo Θ con il semiasse positivo delle x. Si denoti con X l'ascissa del punto di intersezione di tale semiretta con la retta di equazione $y = 1$. Mostrare che X ùna variabile casuale assolutamente continua, determinarne la densità e calcolare, se esiste, $E(X)$.

SOLUZIONE. Si noti che $X = \frac{1}{\tan(\Theta)}$. Si ha allora, per $x > 0$

$$F_X(x) = P(X \leq x) = P(\Theta \geq \arctan(1/x)) = \frac{1}{\pi}[\pi - \arctan(1/x)].$$

Per $x < 0$

$$F_X(x) = P(X \leq x) = P(\Theta \geq \pi - \arctan(-1/x)) = \frac{1}{\pi}[\arctan(-1/x)].$$

Notare che i due "rami" si "attaccano" con continuità in $x = 0$. Quindi F_X è $\mathcal{C}^1$ a tratti, da cui segue che X è assolutamente continua. Derivano F_X si trova che la sua densità è

$$f_X(x) = \frac{1}{\pi} \frac{1}{1 + x^2}.$$

Poiché

$$\int_{-\infty}^{+\infty} \frac{|x|}{1 + x^2} dx = +\infty$$

si ha che $X \notin L^1$, quindi $E(X)$ non è definito.

ESERCIZIO 8. Nel gioco della roulette ci sono 37 numeri, di cui 18 rossi, 18 neri e lo 0 che è verde. Se si gioca un Euro sul rosso si vince un Euro se esce un numero rosso, e lo si perde altrimenti. Supponiamo di giocare ripetutamente un Euro sul rosso.

(a) Quali sono le probabilità di essere in attivo di almeno un Euro dopo 100, 1000 e 10000 ripetizioni? (Usare un'opportuna approssimazione)

(b) Se n è il numero di ripetizioni, per quali valori di n la probabilità di cui sopra è ≤ 0.01?

SOLUZIONE. Sia X_i la variabile che fornisce il bilancio dell'n-esima ripetizione del gioco, per cui $P(X_i = 1) = \frac{18}{37}$, $P(X_i = -1) = \frac{19}{37}$. Pertanto

$$E(X_i) = -\frac{1}{37}, \quad Var(X_i) = 1 - \frac{1}{37^2} =: \sigma^2$$

Si ha dunque, per $Z \sim N(0,1)$,

$$P(X_1 + X_2 + \ldots + X_n \geq 1) = P(X_1 + X_2 + \ldots + X_n \geq 1/2) = P\left(Z_n \geq \frac{\frac{1}{2n} + \frac{1}{37}}{\sigma}\sqrt{n}\right)$$

$$\simeq P\left(Z \geq \frac{\frac{1}{2n} + \frac{1}{37}}{\sigma}\sqrt{n}\right) = 1 - \Phi\left(\frac{\frac{1}{2n} + \frac{1}{37}}{\sigma}\sqrt{n}\right).$$

(a) Calcolando la formula precedente per $n = 100$, $n = 1000$ e $n = 10000$ si trovano, rispettivamente, le probabilità 0.3743, 0.1919 e 0.0034.

(b) Dobbiamo risolvere

$$1 - \Phi\left(\frac{\frac{1}{2n} + \frac{1}{37}}{\sigma}\sqrt{n}\right) \leq 0.01,$$

cioè

$$\frac{\frac{1}{2n} + \frac{1}{37}}{\sigma}\sqrt{n} \geq \Phi^{-1}(0.99) = 2.33.$$

Posto $x = \sqrt{n}$, si ottiene la disequazione

$$2x^2 - 74 \cdot 2.33\sigma x + 37 \geq 0.$$

Tale disequazione è verificata per $x \geq x_2$ e $x \leq x_1$, dove $x_1 < x_2$ sono le radici dell'equazione associata. Poiché $x_1 < 1$, le soluzioni $x \leq x_1$ non sono accettabili. Si ottiene dunque $\sqrt{n} \geq x_2 = 85.96$, da cui $n \geq 7390$.

ESERCIZIO 9. Siano $X \sim Exp(\lambda)$ e $Y \sim Ge(p)$ indipendenti. Definita $Z := \frac{X}{1+Y}$, mostrare che Z è una variabile aleatoria assolutamente continua, e determinarne la densità.

SOLUZIONE. Sia $z > 0$.

$$\begin{aligned}
F_Z(z) = P(Z \le z) = P\left(\frac{X}{1+Y} \le z\right) &= \sum_{n=0}^{+\infty} P\left(\frac{X}{1+Y} \le z, Y = n\right) \\
&= \sum_{n=0}^{+\infty} P\left(\frac{X}{1+n} \le z, Y = n\right) = \sum_{n=0}^{+\infty} P\left(X \le z(1+n), Y = n\right) \\
&= \sum_{n=0}^{+\infty} P\left(X \le z(1+n)\right) P\left(Y = n\right) = \sum_{n=0}^{+\infty} \left[1 - e^{-\lambda z - n\lambda z}\right] p(1-p)^n \\
&= 1 - pe^{-\lambda z} \sum_{n=0}^{+\infty} [(1-p)e^{-\lambda z}]^n = 1 - \frac{pe^{-\lambda z}}{1-(1-p)e^{-\lambda z}}.
\end{aligned}$$

Essendo $F_Z(z) = 0$ per $z \le 0$, si ha che F_Z è $\mathcal{C}^1$ a tratti, quindi Z è assolutamente continua e

$$f_Z(z) = F_Z'(z) = \frac{\lambda p e^{-\lambda z}}{[1-(1-p)e^{-\lambda z}]^2} \mathbf{1}_{(0,+\infty)}(z).$$

ESERCIZIO 10. Sia $(X_n)_{n\geq 1}$ una successione di variabili casuali in L^2, tutte aventi media 0 e varianza 1, ma non necessariamente indipendenti. Dimostrare che per ogni $\epsilon, \delta > 0$

$$\lim_{n\to+\infty} P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n^{1+\delta}}\right| > \epsilon\right) = 0.$$

SOLUZIONE. Per la disuguaglianza di Chebischev,

$$P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n^{1+\delta}}\right| > \epsilon\right) \leq \frac{Var(X_1 + X_2 + \cdots + X_n)}{\epsilon^2 n^{2+2\delta}}.$$

Inoltre:

$$Var(X_1 + X_2 + \cdots + X_n) = \sum_{i,j=1}^{n} Cov(X_i, X_j) \leq n^2,$$

in quanto, per la disuguaglianza di Cauchy-Schwartz, $Cov(X_i, X_j) \leq \sqrt{Var(X_i)Var(X_j)} = 1$. La conclusione segue allora immediatamente.

ESERCIZIO 11. Si estraggono, una dopo l'altra, senza reimmissione, quattro palline da un'urna che contiene 10 palline colorate da 1 a 10.

(a) Sia X il più alto numero estratto. Qual è la distribuzione di X?

(b) Sia A l'evento "tra i numeri delle quattro palline estratte non ce ne sono due di consecutivi". Qual è la probabilità di A?

(c) Calcolare $P(A|X = 7)$.

(d) Qual è la probabilità di A condizionata al fatto che uno dei numeri estratti è il 7?

SOLUZIONE.

(a) Notare anzitutto che $4 \leq X \leq 10$. L'evento $X = n$ si realizza in tanti modi quante sono le scelte di tre numeri tra 1 e $n - 1$. Quindi

$$P(X = n) = \frac{\binom{n-1}{3}}{\binom{10}{4}}.$$

(b) Bisogna contare il numero di scelte di quattro numeri tra 1 e 10 senza coppie consecutive. Si può procedere per enumerazione, ma c'è un modo più astuto. Mostriamo che le scelte di quattro numeri tra 1 e 10 senza coppie consecutive sono in corrispondenza biunivoca con tutte le scelte di 4 numeri tra 1 e 7. Infatti:

- Consideriamo quattro numeri tra 1 e 10 senza coppie consecutive, e rappresentiamo la scelta nella forma seguente: $1, \mathbf{2}, 3, \mathbf{4}, 5, 6, \mathbf{7}, 8, 9, \mathbf{10}$, dove i numeri in grassetto sono quelli scelti. Cancelliamo ora il numero immediatamente a destra dei 3 numeri scelti più piccoli: $1, \mathbf{2}, \mathbf{4}, 6, \mathbf{7}, 9, \mathbf{10}$. Infine ri-enumeriamo i 7 numeri restanti: $1, \mathbf{2}, \mathbf{3}, 4, \mathbf{5}, 6, \mathbf{7}$. Si ottiene un sottoinsieme di 4 numeri tra 1 e 7.
- La mappa è invertibile: partendo da $1, \mathbf{2}, \mathbf{3}, 4, \mathbf{5}, 6, \mathbf{7}$, si aggiunge un "oggetto" a destra dei tre numeri scelti più piccoli $(1, \mathbf{2}, \bullet, \mathbf{3}, \bullet, 4, \mathbf{5}, \bullet, 6, \mathbf{7})$, e infine si ri-enumera $(1, \mathbf{2}, 3, \mathbf{4}, 5, 6, \mathbf{7}, 8, 9, \mathbf{10})$.

Quindi il numero di scelte di quattro numeri tra 1 e 10 senza coppie consecutive è $\binom{7}{3}$. Infine

$$P(A) = \frac{\binom{7}{3}}{\binom{10}{4}}.$$

(c) La probabilità $P(A \cap \{X = 7\})$ vale $k/\binom{10}{4}$, dove k è il numero di scelte di quattro numeri senza coppie consecutive di cui il più grande sia 7. Vi è un'unica tale scelta $(1, 3, 5, 7)$, quindi $k = 1$. Perciò

$$P(A|X = 7) = \frac{P(A \cap \{X = 7\})}{P(X = 7)} = \frac{1}{\binom{6}{3}}.$$

(d) Sia $B = A \cap \{$"uno dei numeri estratti è il 7"$\}$. Il numero di elementi di B è il numero delle quaterne di A che contengono 7. Necessariamente 7 deve essere o il numero più grande, o il secondo numero più grande. Come abbiamo visto, A ha un unico elemento avente 7 come numero più grande. Contiamo ora gli elementi di A che hanno 7 come secondo numero più grande. Tale conteggio può essere fatto con il seguente schema di scelte successive:

- Il numero più grande può essere 9 o 10 (due scelte);

- bisogna poi scegliere due numeri non consecutivi fra 1 e 5. Il numero di tale scelte è 6, che si può ottenere con una semplice enumerazione, oppure con il trucco precedente.

Quindi $P(B) = 12/\binom{10}{4}$. Infine

$$P(A|\{\text{“uno dei numeri estratti è il 7”}\}) = \frac{P(B)}{P(\text{“uno dei numeri estratti è il 7”})} = \frac{12}{\binom{9}{3}},$$

dove abbiamo usato il semplice fatto che

$$P(\text{“uno dei numeri estratti è il 7”}) = \frac{\binom{9}{3}}{\binom{10}{4}}.$$