II Prova Parziale di Probabilità e Statistica	Cognome: ______________
Laurea in Matematica	Nome: ______________
28 giugno 2012	Matricola: ______________

TEMA A

ESERCIZIO 1. 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 2. 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 3. Sia $X \sim Exp(\lambda)$, e definiamo $Y := \left(\frac{X}{\alpha}\right)^{1/\beta}$, dove $\alpha, \beta > 0$. Si determini la distribuzione di Y.

SOLUZIONE. Anzitutto:

$$F_X(x) = \mathbf{1}_{(0,+\infty)}(x)\left[1 - e^{-\lambda x}\right].$$

Allora, per $y > 0$,

$$F_Y(y) = P\left[\left(\frac{X}{\alpha}\right)^{1/\beta} \leq y\right] = P\left(X \leq \alpha y^\beta\right) = \left[1 - e^{-\lambda\alpha y^\beta}\right],$$

e $F_Y(y) = 0$ per $y \leq 0$. Quindi F_Y è $\mathcal{C}^1$ a tratti, Y è assolutamente continua e

$$f_Y(y) = F_Y'(y) = \mathbf{1}_{(0,+\infty)}(y)\lambda\alpha\beta y^{\beta-1}e^{-\lambda\alpha y^\beta}.$$

ESERCIZIO 4. 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 5. Un macchinario produce 1000 pezzi al giorno. Ogni pezzo ha probabilità p di essere difettoso. Per quali valori di p la probabilità che il numero di pezzi difettosi (in un dato giorno) sia più di 100 è minore di 0.01? (Usare un'espressione approssimata della probabilità)

SOLUZIONE. Siano $X_i \sim Be(p)$, $i = 1, 2, \ldots, 1000$ indipendenti. Dobbiamo trovare i valori di p per cui

$$P\left(\sum_{i=1}^{1000} X_i > 100\right) \leq 0.01.$$

Usando l'approssimazione normale e la correzione di continuità, se $Z \sim N(0,1)$ abbiamo:

$$P\left(\sum_{i=1}^{1000} X_i > 100.5\right) = P\left(Z_n > \frac{\frac{100.5}{1000} - p}{\sqrt{p(1-p)}}\sqrt{1000}\right).$$

Pertanto dobbiamo trovare i valori di p per cui

$$\frac{\frac{100.5}{1000} - p}{\sqrt{p(1-p)}}\sqrt{1000} \geq \Phi^{-1}(0.99) = 2.33.$$

Risolvendo la disequazione, si trova $p \leq 0.08$.

ESERCIZIO 6. 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 $z > 0$.

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

Essendo $F_Z(z) = 0$ per $z \leq 0$, 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 7. Un test a risposte multiple è costituito da 100 domande, ognuna con 4 risposte possibili, di cui solo una corretta. Ogni studente ha probabilità 1/2 di conoscere la risposta ad una data domanda, indipendentemente dalle altre domande e dagli altri studenti. Alle domande di cui sanno la risposta gli studenti forniscono (ovviamente) la risposta corretta, mentre per le altre rispondono "a caso", cioè scegliendo una delle risposte possibili con ugual probabilità.

(a) Considerato uno studente scelto a caso, sia X il numero di risposte corrette fornite. Qual è la distribuzione di X?

(b) Il test viene sostenuto da molti studenti; gli esaminatori decidono *a priori* di fissare una "soglia" di promozione (= numero minimo di risposte corrette per essere promossi) in modo tale che la percentuale di promossi sia circa del 60%. Qual è tale valore di soglia? (Sugg: usare l'approssimazione normale)

SOLUZIONE.

(a) Sia A = "lo studente fornisce la risposta esatta", B = "lo studente conosce la risposta esatta". Abbiamo

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c) = \frac{1}{2} + \frac{1}{4}\frac{1}{2} = \frac{5}{8}$$

Dunque $X \sim B(100, 5/8)$.

(b)

(c) Si tratta di trovare un numero intero m per cui $P(X \geq m) \simeq 0.6$. Usando l'approssimazione normale con la correzione di continuità, se $Z \sim N(0,1)$:

$$P(X \geq m) = P(X \geq m - 1/2) \simeq P\left(Z \geq \frac{\frac{m-1/2}{100} - \frac{5}{8}}{\sqrt{\frac{5}{8}\frac{3}{8}}}\sqrt{100} \right) \simeq 0.6,$$

che equivale a

$$\frac{\frac{m-1/2}{100} - \frac{5}{8}}{\sqrt{\frac{5}{8}\frac{3}{8}}}\sqrt{100} = \Phi^{-1}(0.4),$$

che, risolvendo per m, fornisce $m = 61.77 \simeq 62$.

ESERCIZIO 8. Siano $X \sim Po(\lambda)$ e $Y \sim Po(\mu)$ variabili casuali indipendenti. Si definisca $Z := \max(X, Y)$. Ricordando la definizione

$$d(p, q) = \frac{1}{2} \sum_x |p(x) - q(x)|,$$

e usando il metodo dell'accoppiamento, dimostrare che

$$d(p_Z, p_X) \leq 1 - e^{-\mu}.$$

SOLUZIONE. Si noti che $\{X \neq Z\} \subseteq \{Y \neq 0\}$. Perciò

$$d(p_Z, p_X) \leq P(X \neq Z) \leq P(Y \neq 0) = 1 - e^{-\mu}.$$