

|                                                                                                |                                                   |
|------------------------------------------------------------------------------------------------|---------------------------------------------------|
| <b>II Prova Parziale di Probabilità e Statistica</b><br>Laurea in Matematica<br>28 giugno 2012 | Cognome: _____<br>Nome: _____<br>Matricola: _____ |
|------------------------------------------------------------------------------------------------|---------------------------------------------------|

## TEMA B

**ESERCIZIO 1.** Sia  $(\Omega, P)$  uno spazio di probabilità discreto,  $X \in L^1(\Omega, P)$ , e sia  $f$  la funzione definita da  $f(x) := x\mathbf{1}_{[0, +\infty)}(x)$ .

- (a) Mostrare che  $f(X) \in L^1(\Omega, P)$ .
- (b) Mostrare che  $E(X) \leq E[f(X)]$  usando la *definizione* di valor medio.
- (c) Mostrare che  $E(X) \leq E[f(X)]$  usando la disuguaglianza di Jensen.

## SOLUZIONE.

- (a) Poiché  $|f(x)| \leq |x|$ , si ha  $E(|f(X)|) \leq E(|X|) < +\infty$ , e si conclude.

- (b)

$$\begin{aligned}
 E(X) &= \sum_{\omega} X(\omega)P(\{\omega\}) = \sum_{\omega: X(\omega) \geq 0} X(\omega)P(\{\omega\}) + \sum_{\omega: X(\omega) < 0} X(\omega)P(\{\omega\}) \\
 &\leq \sum_{\omega: X(\omega) \geq 0} X(\omega)P(\{\omega\}) = E[f(X)].
 \end{aligned}$$

- (c)  $f$  è convessa, e  $f(x) \geq x$  per ogni  $x \in \mathbb{R}$ . Pertanto, per la disuguaglianza di Jensen:

$$E[f(X)] \geq f(E(X)) \geq E(X).$$

**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  $\epsilon, \delta > 0$

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

**SOLUZIONE.** Per la disuguaglianza di Chebischev,

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

Inoltre:

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

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

**ESERCIZIO 3.** 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  $Y \sim \text{Exp}(1/2\sigma^2)$ , da cui

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

**ESERCIZIO 4.** Sia  $X \sim U(0, 2)$ .  $X$  divide l'intervallo  $[0, 2]$  in due segmenti. Sia  $Y$  il rapporto fra il segmento più lungo e quello più corto. Si determini la distribuzione di  $Y$ .

**SOLUZIONE.** Si ha  $Y = \max\left(\frac{2-X}{X}, \frac{X}{2-X}\right) \geq 1$ . Quindi, per  $y \geq 1$ ,

$$F_Y(y) = P\left(\frac{2-X}{X} \leq y, \frac{X}{2-X} \leq y\right) = P\left(\frac{2}{1+y} \leq X \leq \frac{2y}{1+y}\right) = \frac{y-1}{y+1}.$$

Chiaramente  $F_Y(y) = 0$  per  $y < 1$ . Quindi  $F_Y$  è di classe  $\mathcal{C}^1$  a tratti. Segue che  $Y$  è assolutamente continua, con densità

$$f_Y(y) = F'_Y(y) = \frac{2}{(1+y)^2} \mathbf{1}_{[1,+\infty)}(y).$$

**ESERCIZIO 5.** 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  $\leq 0.01$ ?

**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  $Z \sim N(0, 1)$ ,

$$\begin{aligned} P(X_1 + X_2 + \dots + X_n \geq 1) &= P(X_1 + X_2 + \dots + X_n \geq 1/2) = P\left(Z_n \geq \frac{\frac{1}{2n} + \frac{1}{37}}{\sigma} \sqrt{n}\right) \\ &\simeq P\left(Z \geq \frac{\frac{1}{2n} + \frac{1}{37}}{\sigma} \sqrt{n}\right) = 1 - \Phi\left(\frac{\frac{1}{2n} + \frac{1}{37}}{\sigma} \sqrt{n}\right). \end{aligned}$$

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

**SOLUZIONE.** Sia  $z > 0$ .

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

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

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

**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.**

- (a) Si verifichi l'identità, valida per ogni  $a < b$  e ogni  $x \in \mathbb{R}$

$$|x - a| - |x - b| = (b - a) [2\mathbf{1}_{[b, +\infty)}(x) - 1] + 2(x - a)\mathbf{1}_{(a, b)}(x).$$

- (b) Sia  $X \in L^1$ , e sia  $m$  un valore tale che  $P(X \geq m) = \frac{1}{2}$  (che viene detto *mediana* di  $X$ ).  
Mostrare che per ogni  $a \in \mathbb{R}$

$$E(|X - a|) \geq E(|X - m|).$$

- (c) Sia  $X$  una variabile casuale assolutamente continua, la cui densità  $f_X$  è tale che  $f_X(x) > 0$  per ogni  $x \in \mathbb{R}$ . Mostrare che  $m \in \mathbb{R}$  con la proprietà vista nel punto precedente esiste ed è unico.

**SOLUZIONE.**

- (a) Basta verificarlo separatamente per  $x \geq b$ ,  $x \in (a, b)$  e  $x \leq a$ .

- (b) Sia  $a > m$ . Per il punto precedente

$$\begin{aligned} |X - m| - |X - a| &= (a - m) [2\mathbf{1}_{[a, +\infty)}(X) - 1] + 2(X - m)\mathbf{1}_{(m, a)}(X) \\ &\leq (a - m) [2\mathbf{1}_{[a, +\infty)}(X) - 1] + 2(a - m)\mathbf{1}_{(m, a)}(X) \\ &= (a - m) [2\mathbf{1}_{[m, +\infty)}(X) - 1] \end{aligned}$$

Prendendo il valor medio di ambo i membri:

$$E(|X - m|) - E(|X - a|) \leq (a - m) [2P(X \geq m) - 1] = 0,$$

da cui segue  $E(|X - m|) \leq E(|X - a|)$ . Se invece  $a < m$

$$|X - a| - |X - m| = (m - a) [2\mathbf{1}_{[m, +\infty)}(X) - 1] + 2(X - a)\mathbf{1}_{(a, m)}(X) \geq [2\mathbf{1}_{[m, +\infty)}(X) - 1].$$

Prendendo il valor medio di ambo i membri:

$$E(|X - a|) - E(|X - m|) \geq (m - a) [2P(X \geq m) - 1] = 0,$$

e si conclude.

- (c)  $F_X$  è una funzione continua, strettamente crescente, con  $\lim_{x \rightarrow -\infty} F_X(x) = 0$ ,  $\lim_{x \rightarrow +\infty} F_X(x) = 1$ . Per continuità esiste  $m$  tale che  $F_X(m) = P(X \geq m) = 1/2$ , e tale  $m$  è unico in quanto  $F_X$  è strettamente crescente.