

Prova scritta di Probabilità e Statistica Laurea in Matematica 2 settembre 2011	Cognome: _____ Nome: _____ Matricola: _____
--	---

ESERCIZIO 1. Sia $(X_i)_{i \geq 1}$ una successione di variabili aleatorie i.i.d. con distribuzione di Bernoulli di parametro $p \in (0, 1)$. Sia $q \in (p, 1)$.

(a) Utilizzando la disuguaglianza di Chebychev, si dimostri che per ogni $N \in \mathbb{N}$, ogni $t \geq 0$,

$$\mathbf{P} \left(\frac{1}{N} \sum_{i=1}^N X_i \geq q \right) \leq e^{-tNq} (\mathbf{E} [e^{tX_1}])^N = \psi_N(t),$$

$$\text{ove } \psi_N(t) \doteq e^{-tNq} (e^tp + (1-p))^N.$$

(b) Poniamo $h(q, p) \doteq q \log(q/p) + (1-q) \log((1-q)/(1-p))$. Utilizzando la differenziabilità di $\psi_N(\cdot)$, si calcoli il parametro $t = t_*$ per il quale $\psi_N(t)$ sia minima. Si deduca che

$$\mathbf{P} \left(\frac{1}{N} \sum_{i=1}^N X_i \geq q \right) \leq e^{-Nh(q,p)}.$$

SOLUZIONE. Abbiamo per ogni $N \in \mathbb{N}$, ogni $t \geq 0$,

$$\begin{aligned} \mathbf{P} \left(\frac{1}{N} \sum_{i=1}^N X_i \geq q \right) &= \mathbf{P} \left(t \sum_{i=1}^N X_i \geq tNq \right) \\ &= \mathbf{P} \left(e^{t \sum_{i=1}^N X_i} \geq e^{tNq} \right) \\ &\leq e^{-tNq} \mathbf{E} \left[\prod_{i=1}^N e^{tX_i} \right] \\ &= e^{-tNq} (\mathbf{E} [e^{tX_1}])^N. \end{aligned}$$

Siccome $X_1 \sim Be(p)$, si ha $\mathbf{E} [e^{tX_1}] = e^tp + (1-p)$.

Visto che $\psi_N(\cdot)$ possiede derivate continue e $t = 0$ non è la posizione di un minimo, otteniamo t_* come soluzione di $\frac{d}{dt} \psi_N(t) = 0$. Infatti

$$\frac{d}{dt} \psi_N(t) = -Nq e^{-tNq} (e^tp + (1-p))^N + N e^tp e^{-tNq} (e^tp + (1-p))^{N-1},$$

da cui $t_* = \log(q/p) - \log((1-q)/(1-p))$. Infine si vede che $\psi_N(t_*) = Nh(q, p)$.

ESERCIZIO 2. Sia X una variabile aleatoria assolutamente continua con densità

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

- (a) Posto $Y := \log X^2$, determinare la distribuzione di Y .
- (b) Mostrare che $X \notin L^1$ ma $Y \in L^1$.
- (c) Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ una funzione strettamente crescente e di classe $\mathcal{C}^1$, e sia $Z := f(X)$. Si calcoli, in termini di f e della sua derivata f' , la funzione di ripartizione F_Z di Z .
- (d) Si determini una funzione f , come al punto precedente, tale che $Z \sim U(0, 1)$.

SOLUZIONE.

(a)

$$F_Y(y) = P(\log X^2 \leq y) = P(X^2 \leq e^y) = P(-e^{y/2} \leq X \leq e^{y/2}) = F_X(e^{y/2}) - F_X(-e^{y/2}).$$

Derivando

$$f_Y(y) = F'_Y(y) = \frac{1}{2}e^{y/2}f_X(e^{y/2}) + \frac{1}{2}e^{y/2}f_X(-e^{y/2}) = e^{y/2}\frac{1}{\pi(1+e^y)}.$$

(b)

$$E(|X|) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{|x|}{1+x^2} dx = +\infty,$$

dato che $\frac{|x|}{1+x^2} dx \sim \frac{1}{|x|}$ per $x \rightarrow \pm\infty$.

$$E(|Y|) = \int_{-\infty}^{\infty} |y|e^{y/2}\frac{1}{\pi(1+e^y)} dy < +\infty$$

visto che $e^{y/2}\frac{1}{\pi(1+e^y)} \sim e^{-|y|/2}$ per $y \rightarrow \pm\infty$.

(c) Poiché $f : \mathbb{R} \rightarrow f(\mathbb{R})$ è strettamente crescente, e quindi invertibile. Se $z \in f(\mathbb{R})$, si ha:

$$F_Z(z) = P(f(X) \leq z) = P(X \leq f^{-1}(z)) = F_X(f^{-1}(z)). \quad (1)$$

Essendo $f(\mathbb{R})$ un intervallo, $F_Z(z) = 1$ se z maggiore $f(\mathbb{R})$, e $F_Z(z) = 0$ se z minora $f(\mathbb{R})$.

(d) Una variabile $Z \sim U(0, 1)$ è una variabile a valori in $(0, 1)$ tale che $F_Z(z) = z$ per $z \in (0, 1)$. Vista la formula (1), si prende $f(x) = F_X(x) = \frac{1}{\pi}[\arctan(x) + \frac{\pi}{2}]$.

ESERCIZIO 3. Si consideri l'insieme $\Omega := \{0, 1\}^n \times \{0, 1\}^n$, munito della probabilità uniforme P . Gli elementi di Ω sono dunque coppie (σ, τ) , dove $\sigma, \tau \in \{0, 1\}^n$. Si consideri la variabile aleatoria $X : \Omega \rightarrow \mathbb{R}$ definita da

$$X(\sigma, \tau) := |\{i = 1, 2, \dots, n : \sigma(i) = \tau(i)\}|.$$

Si determini la distribuzione, la media e la varianza di X .

SOLUZIONE. Cominciamo col contare gli elementi dell'evento $\{X = k\}$, dove $k = 0, 1, \dots, n$. Un elemento (σ, τ) di $\{X = k\}$ può essere scelto tramite il seguente schema di scelte successive:

- a) scelgo un generico τ : ho 2^n scelte possibili;
- b) scelgo k componenti in cui σ deve coincidere con τ : $\binom{n}{k}$ scelte possibili. Ovviamente questa scelta identifica σ univocamente.

Infine:

$$P(X = k) = \frac{2^n \binom{n}{k}}{2^{2n}} = \binom{n}{k} \frac{1}{2^n},$$

da cui segue che $X \sim B(n, 1/2)$, e perciò $E(X) = n/2$ e $Var(X) = n/4$.

ESERCIZIO 4. Un gioco consiste nell'estrarre a caso due carte da un mazzo di carte da Poker (52 carte, 4 semi); si vince se nessuna delle carte estratte è di quadri.

- (a) Determinare la probabilità di successo in questo gioco.
- (b) Per $n \geq 1$, sia p_n la probabilità che in $2n$ ripetizioni del gioco il numero di successi sia almeno n . Determinare, approssimativamente, il valore p_{50} .
- (c) Determinare il minimo valore di n per cui $p_n \geq 1 - 10^{-4}$.

SOLUZIONE.

- (a) Essendoci 39 carte non di quadri, la probabilità di successo è

$$\frac{\binom{39}{2}}{\binom{52}{2}} = \frac{19}{34}.$$

- (b) Sia $X_i \text{Be}(19/34)$ la variabile che vale 1 se e solo se la prova i -esima è un successo. Usando l'approssimazione normale, trascurando la correzione di continuità,

$$p_n = P(\bar{X}_{2n} \geq 1/2) = P\left(\frac{\bar{X}_{2n} - \frac{19}{34}}{\sqrt{\frac{19}{34} \frac{15}{34}}} \sqrt{2n} \geq \frac{\frac{1}{2} - \frac{19}{34}}{\sqrt{\frac{19}{34} \frac{15}{34}}} \sqrt{2n}\right) \simeq \Phi\left(\frac{\frac{19}{34} - \frac{1}{2}}{\sqrt{\frac{19}{34} \frac{15}{34}}} \sqrt{2n}\right)$$

Per $n = 50$, si trova

$$p_{50} \simeq 0.8819$$

- (c) Usando la formula al punto precedente, troviamo

$$p_n \geq 1 - 10^{-4} \iff \frac{\frac{19}{34} - \frac{1}{2}}{\sqrt{\frac{19}{34} \frac{15}{34}}} \sqrt{2n} \geq \Phi^{-1}(1 - 10^{-4}) \simeq 3.72$$

che fornisce $n \geq 493$.