Laurea Triennale in Informatica, Università di Roma Tor Vergata
Calcolo delle Probabilità (ed insegnamenti mutuati)
Anno accademico: `2014-2015`. *Titolare del corso:* `Claudio Macci`
Simulazione 1

Esercizio 1. Supponiamo di avere 3 urne, ciascuna delle quali ha 4 palline bianche e 4 nere. Da ogni urna vengono estratte a caso 2 palline, una alla volta e *senza* reinserimento. Sia X la variabile aleatoria che conta il numero di urne dalle quali si estrae la sequenza di colori (nero, bianco).
D1) Trovare la densità discreta di X.
D2) Rispondere alla stessa domanda nel caso di estrazioni *con* reiserimento.

Esercizio 2. Sia X una variabile aleatoria con densità discreta $p_X(h) = (1-q)^{h-1}q$ per $h \geq 1$ intero e $q \in (0,1)$. Si compiono X prove indipendenti con probabilità di successo $p \in (0,1)$ e sia Y la variabile aleatoria che conta il numero di successi ottenuti.
D3) Calcolare $P(Y = 0)$.
D4) Calcolare $P(X = h|Y = 0)$ per $h \geq 1$ intero.

Esercizio 3. Definiamo la seguente densità congiunta: $p_{X_1,X_2}(k, 4-k) = \binom{4}{k}(\frac{1}{2})^4$ per $k \in \{0,1,2,3,4\}$.
D5) Calcolare $P(X_1 \geq X_2)$.
D6) Trovare la densità discreta di $Z = |X_1 - X_2|$.

Esercizio 4. Sia X una variabile aleatoria con densità continua $f_X(t) = \frac{e^t}{e^a-1}1_{(0,a)}(t)$ per $a > 0$.
D7) Calcolare $\mathbb{E}[e^{-X}]$.
D8) Trovare la funzione di distribuzione di $Y = \log(1 + X/a)$.

Esercizio 5.
D9) Sia $N_t = \sum_{n\geq 1} 1_{T_n \leq t}$ (per $t \geq 0$) un processo di Poisson con intensità di $\lambda = 3$. Calcolare $\mathbb{E}[N_6]$ e $\mathbb{E}[T_9]$.
D10) Sia X una variabile aleatoria normale con media $\mu = -6$ e varianza $\sigma^2 = 36$. Calcolare $P(X \geq 0)$.

Esercizio 6. Sia $\{X_n : n \geq 1\}$ una successione di variabili aleatorie i.i.d. (indipendenti e identicamente distribuite).
D11) Dire per quale valore di m si ha

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + \cdots + X_n}{n} - m\right| > \varepsilon\right) = 0 \text{ per ogni } \varepsilon > 0,$$

nel caso in cui le variabili aleatorie $\{X_n : n \geq 1\}$ abbiano densità continua $f(t) = (1 - |t|)1_{(-1,1)}(t)$.
D12) Dire per quale valore di x si ha

$$\lim_{n\to\infty} P\left(\frac{X_1 + \cdots + X_n - 4n}{\sqrt{n}} \leq x\right) = \Phi(1/2)$$

nel caso in cui le variabili aleatorie $\{X_n : n \geq 1\}$ abbiano speranza matematica 4 e varianza 121.

Esercizio 7 (*solo per Lauree Magistrali*). Consideriamo una catena di Markov omogenea $\{X_n : n \geq 0\}$ con spazio degli stati $E = \{1,2,3\}$ e matrice di transizione

$$P = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1 & 0 & 0 \end{pmatrix}.$$

D13) Motivare la possibilità di applicare il Teorema di Markov e, con l'approssimazione fornita da tale teorema, calcolare $P(X_{1000} = 1)$.
D14) Calcolare $P(X_1 \neq 3, X_2 \neq 3|X_0 = 1)$.

Cenno alle soluzioni (*Ogni segnalazione di errori o sviste (sempre possibili) è gradita*)

Esercizio 1.
In entrambi i casi la variabile aleatoria X ha distribuzione binomiale con parametri $n = 3$ e p da determinare in ciascun caso; quindi si ha $p_X(k) = \binom{3}{k} p^k (1-p)^{3-k}$ per $k \in \{0,1,2,3\}$.
`D1`) In questo caso si ha $p = \frac{4}{8}\frac{4}{7} = \frac{2}{7}$ e quindi $p_X(0) = \frac{125}{343}$, $p_X(1) = \frac{150}{343}$, $p_X(2) = \frac{60}{343}$ e $p_X(3) = \frac{8}{343}$.
`D2`) In questo caso si ha $p = \frac{4}{8}\frac{4}{8} = \frac{1}{4}$ e quindi $p_X(0) = \frac{27}{64}$, $p_X(1) = \frac{27}{64}$, $p_X(2) = \frac{9}{64}$ e $p_X(3) = \frac{1}{64}$.

Esercizio 2.
`D3`) Per la formula delle probabilità totali si ha (si osservi che le probabilità condizionate si esprimono con la densità della distribuzione binomiale) $P(Y = 0) = \sum_{h=1}^{\infty} P(Y = 0|X = h)P(X = h) = \sum_{h=1}^{\infty} \binom{h}{0} p^0 (1-p)^{h-0}(1-q)^{h-1}q = \frac{q}{1-q}\sum_{h=1}^{\infty}[(1-p)(1-q)]^h = \frac{q}{1-q}\cdot\frac{(1-p)(1-q)}{1-(1-p)(1-q)} = \frac{q(1-p)}{1-(1-p-q+pq)} = \frac{q(1-p)}{p+q-pq}$.
`D4`) Per la formula di Bayes, e tenendo conto di alcuni valori calcolati prima, si ha $P(X = h|Y = 0) = \frac{P(Y=0|X=h)P(X=h)}{P(Y=0)} = \frac{\binom{h}{0}p^0(1-p)^{h-0}(1-q)^{h-1}q}{\frac{q(1-p)}{1-(1-p)(1-q)}} = ((1-p)(1-q))^{h-1}(1-(1-p)(1-q))$.

Esercizio 3.
`D5`) Si devono sommare i valori della densità congiunta per $k \in \{0,1,2,3,4\}$ tale che $k \geq 4-k$; quindi si ha $P(X_1 \geq X_2) = \sum_{k=2}^{4}\binom{4}{k}(\frac{1}{2})^4 = \frac{6+4+1}{16} = \frac{11}{16}$.
`D6`) Si ha $p_Z(0) = p_{X_1,X_2}(2,2) = \frac{6}{16} = \frac{3}{8}$, $p_Z(2) = p_{X_1,X_2}(3,1) + p_{X_1,X_2}(1,3) = \frac{4+4}{16} = \frac{8}{16} = \frac{4}{8}$ e $p_Z(4) = p_{X_1,X_2}(4,0) + p_{X_1,X_2}(0,4) = \frac{1+1}{16} = \frac{2}{16} = \frac{1}{8}$.

Esercizio 4.
`D7`) Si ha $\mathbb{E}[e^{-X}] = \int_0^a e^{-t}\frac{e^t}{e^a-1}dt = \frac{1}{e^a-1}\int_0^a dt = \frac{1}{e^a-1}[t]_{t=0}^{t=a} = \frac{a}{e^a-1}$.
`D8`) Si vede che $P(0 \leq \log(1+X/a) \leq \log 2) = 1$, da cui $F_Y(y) = 0$ per $y \leq 0$ e $F_Y(y) = 1$ per $y \geq \log 2$. Per $y \in (0, \log 2)$ si ha $F_Y(y) = P(\log(1+X/a) \leq y) = P(1+X/a \leq e^y) = P(X/a \leq e^y - 1) = P(X \leq a(e^y-1)) = \int_0^{a(e^y-1)}\frac{e^t}{e^a-1}dt = \frac{1}{e^a-1}[e^t]_{t=0}^{t=a(e^y-1)} = \frac{e^{a(e^y-1)}-1}{e^a-1}$.

Esercizio 5.
`D9`) Si ha $\mathbb{E}[N_6] = 3\cdot 6 = 18$ e $\mathbb{E}[T_9] = \frac{9}{3} = 3$.
`D10`) Si ha $P(X \geq 0) = P(Z_X \geq \frac{0-(-6)}{\sqrt{36}}) = P(Z_X \geq 1) = 1 - \Phi(1) = 1 - 0.84134 = 0.15866$.

Esercizio 6.
`D11`) Per la legge dei grandi numeri si ha $m = \int_{-1}^{1} t(1-|t|)dt = \int_{-1}^{0} t(1+t)dt + \int_0^1 t(1-t)dt = \int_{-1}^{0} t + t^2 dt + \int_0^1 t - t^2 dt = [\frac{t^2}{2} + \frac{t^3}{3}]_{t=-1}^{t=0} + [\frac{t^2}{2} - \frac{t^3}{3}]_{t=0}^{t=1} = -(\frac{1}{2} - \frac{1}{3}) + (\frac{1}{2} - \frac{1}{3}) = 0$.
`D12`) Indichiamo con Z_{S_n} la standardizzata di $S_n = X_1 + \cdots + X_n$. Allora $\left\{\frac{X_1+\cdots+X_n-4n}{\sqrt{n}} \leq x\right\} = \left\{Z_{S_n} \leq \frac{x}{\sqrt{121}}\right\}$ e, per il teorema limite centrale, si deve avere $\frac{x}{\sqrt{121}} = 1/2$, da cui segue $x = \frac{\sqrt{121}}{2} = \frac{11}{2}$.

Esercizio 7.
`D13`) Possiamo applicare il Teorema di Markov perché la catena è regolare; infatti la catena è irriducibile ed esiste un elemento positivo della diagonale principale della matrice di transizione (questo è sufficiente per la regolarità). Allora esiste un'unica distribuzione stazionaria (π_1, π_2, π_3) e, applicando tale teorema, possiamo dire che

$$\lim_{n\to\infty} P(X_n = j) = \lim_{n\to\infty}\sum_{i=1}^{3} P(X_n = j|X_0 = i)P(X_0 = i) = \pi_j \sum_{i=1}^{3} P(X_0 = i) = \pi_j \text{ (per ogni } j \in \{1,2,3\})$$

qualsiasi sia la distribuzione iniziale. Nel caso specifico si ha il seguente sistema:

$$\begin{cases} \frac{\pi_1}{3} + \frac{\pi_2}{3} + \pi_3 = \pi_1 \\ \frac{\pi_1}{3} + \frac{\pi_2}{3} = \pi_2 \\ \frac{\pi_1}{3} + \frac{\pi_2}{3} = \pi_3 \end{cases}$$

Sostituendo la terza nella seconda si ottiene $\pi_2 = \pi_3$; poi sostituendo la terza nella prima si ottiene si ottiene $\pi_1 = 2\pi_3$. Allora, tenendo conto che $\pi_1 + \pi_2 + \pi_3 = 1$, si ottiene la distribuzione stazionaria è $(1/2, 1/4, 1/4)$. In conclusione, facendo riferimento alla approssimazione fornita dal Teorema di Markov, si ha $P(X_{1000} = 1) = \pi_1 = \frac{1}{2}$.

`D14`) Si ha $P(X_1 \neq 3, X_2 \neq 3 | X_0 = 1) = \sum_{i,j=1}^{2} p_{1i} p_{ij} = \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} = \frac{4}{9}$.

Commenti.

La somma dei valori di ciascuna densità discreta che appare è 1 in accordo con la teoria.

`D3`) Si ha $\frac{q(1-p)}{p+q-pq} \in (0,1)$ come deve essere osservando che il denominatore è maggiore del numeratore, il quale a sua volta è positivo; infatti $p + q - pq = p + q(1-p) > q(1-p) > 0$.

`D4`) Possiamo dire che X ha distribuzione geometrica traslata di parametro q e, condizionatamente all'evento $\{Y = 0\}$, si ha una geometrica traslata di parametro $r = 1 - (1-p)(1-q)$.

`D7`) In altro modo (meno conveniente) si calcola la speranza matematica richiesta dopo aver calcolato la densità continua di $Z = e^{-X}$. Si ha $F_Z(z) = 0$ per $z \leq e^{-a}$, $F_Z(z) = 1$ per $z \geq 1$ e $F_Z(z) = P(e^{-X} \leq z) = P(-X \leq \log z) = P(X \geq -\log z) = \int_{-\log z}^{\infty} f_X(t)dt = \int_{-\log z}^{a} \frac{e^t}{e^a-1} dt = \frac{1}{e^a-1}[e^t]_{t=-\log z}^{t=a} = \frac{e^a-1/z}{e^a-1}$ per $z \in (e^{-a}, 1)$, da cui segue $f_Z(z) = \frac{1}{(e^a-1)z^2} 1_{(e^{-a},1)}(z)$, e quindi $\mathbb{E}[Z] = \int_{e^{-a}}^{1} z \frac{1}{(e^a-1)z^2} dz = \frac{1}{e^a-1} \int_{e^{-a}}^{1} \frac{1}{z} dz = \frac{1}{e^a-1}[\log z]_{z=e^{-a}}^{z=1} = \frac{0-\log e^{-a}}{e^a-1} = \frac{a}{e^a-1}$.

`D11`) L'uguaglianza $\int_{-1}^{1} t(1-|t|)dt = 0$ si deduce senza fare calcoli osservando che si ha l'integrale di una funzione "dispari" su un intervallo simmetrico rispetto all'origine, cioè $[-1, 1]$.