Laurea triennale in MATEMATICA, Corso di **PROBABILITÀ 1**

Prof. L. Bertini - G. Nappo - F. Spizzichino

ESONERO DEL 03.06.2010

N.B. Scrivere le soluzioni degli esercizi su questi fogli giustificando brevemente i passaggi svolti. Non è necessario svolgere tutti i calcoli fino in fondo.

Esercizio 1. In un'elezione vi sono sette elettori e tre candidati (A,B e C) ed indichiamo con X_A, X_B e X_C il numero di voti che vengono rispettivamente attribuiti ad A, B e C.

Supponiamo che ciascun elettore voti, indipendentemente dagli altri, A con probabilità $\frac{1}{2}$, B con probabilità $\frac{1}{4}$ e C con probabilità $\frac{1}{4}$.

i) Calcolare $\alpha_1 = P(X_A = 2, X_B = 3, X_C = 2)$ e $\beta_1 = P(X_A = 2, X_B = 3, X_C = 5)$.

ii) Calcolare $\alpha_2 = P(X_C = 2)$ e $\beta_2 = P(X_C = 5)$.

iii) Calcolare $\alpha_3 = P(X_B = 3, X_C = 2)$ e $\beta_3 = P(X_B = 3, X_C = 5)$.

iv) Supponiamo di sapere che $X_B = 3$. Condizionatamente a tale evento, trovare la distribuzione di probabilità di X_C, il valore atteso e la varianza.

i) **SOLUZIONE** Dalla definizione del meccanismo di elezione, la distribuzione congiunta di X_A, X_B, X_C è una multinomiale. Più precisamente, si ha

$$\alpha_1 = \binom{7}{2\ 3\ 2} \left(\frac{1}{2}\right)^2 \left(\frac{1}{4}\right)^3 \left(\frac{1}{4}\right)^2 \left(= \frac{7!}{2!3!2!}\,\frac{1}{2^{12}} = 7 \cdot 6 \cdot 5\,\frac{1}{2^{12}} \right).$$

Ovviamente deve essere

$$P\left(X_A + X_B + X_C = 7\right) = 1,$$

e dunque si ha

$$\beta_1 = 0.$$

ii) **SOLUZIONE** Dalla definizione del meccanismo di elezione segue anche che le distribuzione marginali di X_A, X_B, X_C sono delle binomiali $bin(7, \frac{1}{2})$, $bin(7, \frac{1}{4}), bin(7, \frac{1}{4})$, rispettivamente. Quindi, in particolare,

$$\alpha_2 = P(X_C = 2) = \binom{7}{2} \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^5$$

$$\beta_2 = P(X_C = 5) = \binom{7}{5} \left(\frac{1}{4}\right)^5 \left(\frac{3}{4}\right)^2 = \binom{7}{2} \left(\frac{1}{4}\right)^5 \left(\frac{3}{4}\right)^2.$$

iii) **SOLUZIONE** In virtù della condizione $P\left(X_A + X_B + X_C = 7\right) = 1$, si ha

$$\alpha_3 = P(X_B = 3, X_C = 2) = P(X_A = 2, X_B = 3, X_C = 2) = \alpha_1;$$

inoltre

$$\beta_3 = P(X_B = 3, X_C = 5) = 0.$$

iv) **SOLUZIONE** La distribuzione di probabilità condizionata di X_C dato l'evento $(X_B = 3)$ deve essere $bin(4, \frac{1}{3})$. Infatti, sapendo che B ha preso 3 voti, C può prendere al massimo 4 voti (e al minimo 0). Inoltre possiamo scrivere, per $k = 0, 1, 2, 4$,

$$\begin{aligned} P(X_C = k | X_B = 3) &= \frac{P(X_C = k, X_B = 3)}{P(X_B = 3)} = \frac{P(X_A = 4 - k, X_B = 3, X_C = k)}{P(X_B = 3)} \\ &= \frac{\binom{7}{4-k\ \ 3\ \ k} \left(\frac{1}{2}\right)^{4-k} \left(\frac{1}{4}\right)^3 \left(\frac{1}{4}\right)^k}{\binom{7}{3} \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^4} = \frac{7!}{(4-k)!3!k!}\,\frac{3!4!}{7!}\,\frac{\left(\frac{1}{4}\right)^k \left(\frac{1}{2}\right)^{4-k}}{\left(\frac{3}{4}\right)^4} \end{aligned}$$

e quindi, considerando che $4 = k + (4-k)$, e che $\frac{4!}{(4-k)!k!} = \binom{4}{k}$,

$$P(X_C = k | X_B = 3) = \binom{4}{k} \frac{\left(\frac{1}{4}\right)^k}{\left(\frac{3}{4}\right)^k} \frac{\left(\frac{1}{2}\right)^{4-k}}{\left(\frac{3}{4}\right)^{4-k}} = \binom{4}{k} \left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{4-k}.$$

A tale risultato si può anche arrivare tramite un ragionamento diretto.

Da quanto sopra [cioè dal fatto che, condizionatamente a $\{X_B = 3\}$, X_C segue una distribuzione $bin(4, \frac{1}{3})$] segue immediatamente che il valore atteso e la varianza di tale distribuzione condizionata sono dati da

$$E(X_C | X_B = 3) = 4\,\frac{1}{3} = \frac{4}{3}$$

e

$$Var(X_C | X_B = 3) = 4\,\frac{1}{3}\,\frac{2}{3} = \frac{8}{9}.$$

Laurea triennale in MATEMATICA, Corso di **PROBABILITÀ 1**

Prof. L. Bertini - G. Nappo - F. Spizzichino

ESONERO DEL 03.06.2010

N.B. Scrivere le soluzioni degli esercizi su questi fogli giustificando brevemente i passaggi svolti. Non è necessario svolgere tutti i calcoli fino in fondo.

Esercizio 2. Supponiamo che, per ciascuno dei numerosi clienti di una compagnia di assicurazione, vi sia uguale probabilità (molto piccola) di denunciare un sinistro nel corso del mese di luglio 2010. Indichiamo con N il numero complessivo di denunce relative a tale periodo. Sulla base di sue informazioni e di precedenti dati statistici, la compagnia stima che il valore atteso di N sia uguale a $7,5$.

i) Esprimere un valore approssimato per $P(N = 0)$

ii) Esprimere un valore approssimato per $P(N < 6)$

iii) Utilizzando la disuguaglianza di Čebišev, dare una limitazione inferiore a $P(4 \le N \le 11)$.

PARTE FACOLTATIVA

Si supponga inoltre che la compagnia abbia introdotto una soglia massima di rimborso di 9000 euro, ossia la compagnia rimborsa al massimo 9000 euro, anche se il danno causato da un sinistro supera tale soglia. Si assuma che la probabilità che il danno (causato da ciascun sinistro) sia maggiore o uguale alla soglia di 9000 euro sia $p = 0,02$, indipendentemente l'uno dall'altro.

Sia Y l'ammontare che la compagnia dovrà rimborsare, limitatamente ai sinistri denunciati nel mese di luglio e che comportano danni maggiori o uguali a 9000 euro.

iv) Calcolare $E(Y)$.

i) **SOLUZIONE** Se conoscessimo il numero esatto n di clienti e la probabilità p della probabilità che un cliente denuci un sinistro nel mese di luglio 2010 [trascurando l'eventualità che un cliente ne denuci più di uno nello stesso mese] potremmo dire che la variabile aleatoria N segue una distribuzione binomiale di parametri n e p. Poiché sappiamo solo che il suo valore atteso è $np = \lambda = 7,5$, n grande e p molto piccola, è ragionevole utilizzare l'approssimazione di Poisson, cioè calcolare approssimativamente le probabilità $P(N = k)$ tramite $\frac{\lambda^k}{k!} e^{-\lambda}$, con $\lambda = 7,5$ (almeno per valori "non troppo grandi" di k).

Quindi possiamo calcolare approssimativamente

$$P(N = 0) \simeq e^{-7,5}$$

ii) **SOLUZIONE** Ovviamente

$$P(N < 6) = P(N \le 5) = \sum_{k=0}^{5} P(N = k),$$

e quindi, per lo stesso motivo del punto *i)*, possiamo approssimare la probabilità cercata con

$$P(N < 6) \simeq \sum_{k=0}^{5} \frac{(7,5)^k}{k!} e^{-7,5}.$$

iii) **SOLUZIONE** Per ottenere, con la disuguaglianza di Čebišev, una limitazione inferiore a

$$P(4 \le N \le 11) = P(|N - 7,5| \le 3,5) = 1 - P(|N - 7,5| > 3,5)$$

basta ottenere, sempre con la disuguaglianza di Čebišev, una maggiorazione per l'insieme complementare, ossia

$$P(|N - 7,5| > 3,5) = P(|N - E(N)| > 3,5) \le \frac{Var(N)}{(3,5)^2} = \frac{np(1-p)}{12,25} \le \frac{np}{12,25} = \frac{7,5}{12,25}.$$

Infatti allora si ottiene

$$P\,(4 \le N \le 11) = P\,(|N-7,5| \le 3,5) = 1 - P\,(|N-7,5| > 3,5) \ge 1 - \frac{7,5}{12,25} = \frac{4,75}{12,25}(\simeq 0,387).$$

NOTA di APPROFONDIMENTO

Uno studente particolarmente attento avrebbe potuto ottenere una limitazione inferiore migliore: basta infatti osservare che, essendo N una variabile aleatoria a valori interi, per ogni $\delta \in (0,1]$, si ha

$$P\,(4 \le N \le 11) = P\,(3+\delta \le N \le 12-\delta) = P\,(|N-7,5| \le 4,5-\delta) = 1 - P\,(|N-7,5| > 4,5-\delta)$$

e quindi ripetendo il ragionamento con $4,5-\delta$ al posto di $3,5$ avrebbe ottenuto che, per ogni $\delta \in (0,1]$,

$$P\,(4 \le N \le 11) = P\,(|N-7,5| \le 4,5-\delta) = 1 - P\,(|N-7,5| > 4,5-\delta) \ge 1 - \frac{7,5}{(4,5-\delta)^2}.$$

Mandando δ a zero si ottiene allora

$$P\,(4 \le N \le 11) \ge 1 - \frac{7,5}{(4,5)^2} = 1 - \frac{7,5}{20,25}(\simeq 0,629).$$

Alla stessa conclusione si può arrivare direttamente anche osservando che vale anche la seguente disuguaglianza (sempre di Čebišev): per ogni variabile aleatoria X con valore atteso μ (e varianza finita) si ha

$$P(|X-\mu| \ge \alpha) \le \frac{Var(X)}{\alpha^2} \qquad \text{qualunque sia } \alpha > 0.$$

iv) **SOLUZIONE** La variabile aleatoria Y si può scrivere come

$$Y = 9000 \sum_{n=1}^{N} \mathbf{1}_{A_i},$$

(con la convenzione che $\sum_{k=1}^{0} a_k = 0$) dove

$$A_i = \{\text{il danno dell'}i\text{-simo sinistro denunciato nel mese di luglio è maggiore o uguale a 9000 euro}\}.$$

Il valore atteso di Y è allora

$$E(Y) = 9000 \cdot E\left(\sum_{i=1}^{N} \mathbf{1}_{A_i}\right) = 9000 \cdot 7,5 \cdot 0,002 = 9000\,\frac{15}{2}\,\frac{2}{100} = 1350.$$

Infatti

$$\begin{aligned} E\left(\sum_{i=1}^{N} \mathbf{1}_{A_i}\right) &= \sum_{n=0}^{\infty} P(N=n)\,E\left(\sum_{i=1}^{N} \mathbf{1}_{A_i}\middle|\{N=n\}\right) \\ &= \sum_{n=0}^{\infty} P(N=n)\,n \cdot 0,02 = 0,02\,E(N) = 0,02 \cdot 7,5. \end{aligned}$$

È importante sottolineare che nella formula precedente si è usato il fatto che

$$E\left(\sum_{i=1}^{N} \mathbf{1}_{A_i}\middle|\{N=n\}\right) = n \cdot 0,02$$

in quanto, condizionatamente a sapere che si è verificato $\{N=n\}$, la variabile aleatoria $\sum_{i=1}^{N} \mathbf{1}_{A_i}$ ha una distribuzione binomiale di parametri n e $p = 0,02$.

Alternativamente si poteva sfruttare il fatto che, se N ha distribuzione di Poisson di parametro λ e gli eventi A_i sono indipendenti fra loro e da N, allora la variabile aleatoria $\sum_{i=1}^{N} \mathbf{1}_{A_i}$ ha distribuzione di Poisson di parametro $\lambda\,p = 7,5\,p$.

Quest'ultima affermazione deriva sempre dal fatto che, condizionatamente a sapere che si è verificato $\{N = n\}$, la variabile aleatoria $\sum_{i=1}^{N} \mathbf{1}_{A_i}$ ha una distribuzione binomiale di paramentri n e $p = 0,02$, e allora

$$P\big(\sum_{i=1}^{N} \mathbf{1}_{A_i} = k\big) = \sum_{n=0}^{\infty} P(N = n)\, P\big(\sum_{i=1}^{N} \mathbf{1}_{A_i} = k|\{N = n\}\big)$$

(tenendo conto che $P\big(\sum_{i=1}^{N} \mathbf{1}_{A_i} = k|\{N = n\}\big) = 0$ se $k > n$)

$$= \sum_{n=k}^{\infty} \frac{\lambda^n}{n!}\, e^{-\lambda} \binom{n}{k} p^k\, (1-p)^{n-k}$$

$$= \lambda^k\, p^k\, e^{-\lambda} \sum_{n=k}^{\infty} \frac{\lambda^{n-k}}{n!} \frac{n!}{k!(n-k)!} (1-p)^{n-k} = \frac{(\lambda\, p)^k}{k!}\, e^{-\lambda} \sum_{n=k}^{\infty} \frac{\big(\lambda\,(1-p)\big)^{n-k}}{(n-k)!}$$

(ponendo $h = n - k$ nella somma della serie, che diventa allora una serie esponenziale)

$$= \frac{(\lambda\, p)^k}{k!}\, e^{-\lambda}\, e^{\lambda\,(1-p)} = \frac{(\lambda\, p)^k}{k!}\, e^{-\lambda\, p}.$$

Laurea triennale in MATEMATICA, Corso di **PROBABILITÀ 1**

Prof. L. Bertini - G. Nappo - F. Spizzichino

Esonero del 03.06.2010

N.B. Scrivere le soluzioni degli esercizi su questi fogli giustificando brevemente i passaggi svolti.
ATTENZIONE: è necessario svolgere tutti i calcoli fino in fondo.

Esercizio 3. La distribuzione di una variabile aleatoria continua X ammette una funzione di densità di probabilità della forma

$$f(x) = \begin{cases} 0 & x \leq -1, \\ a & -1 < x \leq 1, \\ \frac{1}{2}ax & 1 < x \leq 3, \\ 0 & x > 3. \end{cases}$$

i) Dimostrare che $a = \frac{1}{4}$.

ii) Dimostrare che $\mathbb{E}(X) = \frac{13}{12}$.

iii) Calcolare $P(0 < X < 2)$.

i) **SOLUZIONE** La funzione $f(x)$ è una densità (di una variabile aleatoria continua) se e solo se

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1$$

Nel nostro caso

$$\begin{aligned}\int_{-\infty}^{+\infty} f(x)\,dx &= \int_{-\infty}^{-1} f(x)\,dx + \int_{-1}^{1} f(x)\,dx + \int_{1}^{3} f(x)\,dx + \int_{3}^{+\infty} f(x)\,dx \\ &= \int_{-1}^{1} a\,dx + \int_{1}^{3} \frac{1}{2}ax\,dx = 2\,a + \frac{1}{2}\,a\,\frac{1}{2}\,x^2\Big|_1^3 = a\Big(2 + \frac{1}{4}\,(9-1)\Big) = 4\,a,\end{aligned}$$

da cui $a = \frac{1}{4}$ e quindi $f(x) = \begin{cases} 0 & x \leq -1, \\ \frac{1}{4} & -1 < x \leq 1, \\ \frac{1}{8}\,x & 1 < x \leq 3, \\ 0 & x > 3. \end{cases}$

ii) **SOLUZIONE** Per le variabili aleatorie con densità si ha che

$$E(X) = \int_{-\infty}^{+\infty} x\,f(x)\,dx, \qquad \text{purché esista finito} \int_{-\infty}^{+\infty} |x|\,f(x)\,dx.$$

$\big[$In questo caso non è necessario fare questa verifica perché X è una variabile aleatoria limitata in quanto, ad esempio, $P(|X| \leq 3) = 1\big]$

$$\begin{aligned}E(X) &= \int_{-\infty}^{+\infty} x\,f(x)\,dx \left(= \int_{-\infty}^{-1} x\,f(x)\,dx + \int_{-1}^{1} x\,f(x)\,dx + \int_{1}^{3} x\,f(x)\,dx + \int_{3}^{+\infty} x\,f(x)\,dx \right) \\ &= \int_{-1}^{1} x\,\frac{1}{4}\,dx + \int_{1}^{3} x\,\frac{1}{8}\,x\,dx = 0 + \frac{1}{8}\,\frac{1}{3}\,x^3\Big|_1^3 = \frac{1}{24}\,(27-1) = \frac{26}{24} = \frac{13}{12}.\end{aligned}$$

iii) **SOLUZIONE** Essendo X una variabile aleatoria con densità si ha che

$$\begin{aligned}P(0 < X < 2) &= \int_0^2 f(x)\,dx = \int_0^1 f(x)\,dx + \int_1^2 f(x)\,dx = \int_0^1 \frac{1}{4}\,dx + \int_1^2 \frac{1}{8}\,x\,dx \\ &= \frac{1}{4} + \frac{1}{8}\,\frac{1}{2}\,x^2\Big|_1^2 = \frac{1}{4} + \frac{1}{8}\,\frac{1}{2}\,(4-1) = \frac{7}{16}.\end{aligned}$$