Capitolo 3

Principali distribuzioni di probabilità discrete

Vogliamo ora presentare alcune distribuzioni di probabilità notevoli che costituiscono lo schema teorico di fenomeni naturali di vario tipo. In questo capitolo consideriamo alcune distribuzioni di probabilità di tipo discreto; in seguito affronteremo le distribuzioni di probabilità di tipo continuo.

3.1 La distribuzione uniforme

La distribuzione uniforme è la distribuzione di eventi equiprobabili. Diremo che una variabile aleatoria X segue una distribuzione uniforme su $\{1, 2, \ldots, n\}$ se ha funzione di probabilità:

$$p(x) := \begin{cases} \dfrac{1}{n} & \text{per } 1, 2 \ldots, n \\ 0 & \text{altrimenti.} \end{cases}$$

In tal caso diremo che X segue una *distribuzione uniforme* e scriveremo $X \sim U(n)$.

Speranza matematica e Varianza. Tenendo conto che:

$$\begin{aligned} 1 + 2 + \cdots + n &= \frac{n(n+1)}{2} \\ 1^2 + 2^2 + \cdots + n^2 &= \frac{n(n+1)(2n+1)}{6} \end{aligned}$$

possiamo calcolare speranza matematica e varianza:

$$\mathbb{E}(X) = \frac{1}{n}(1 + 2 + \cdots + n) = \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2}$$

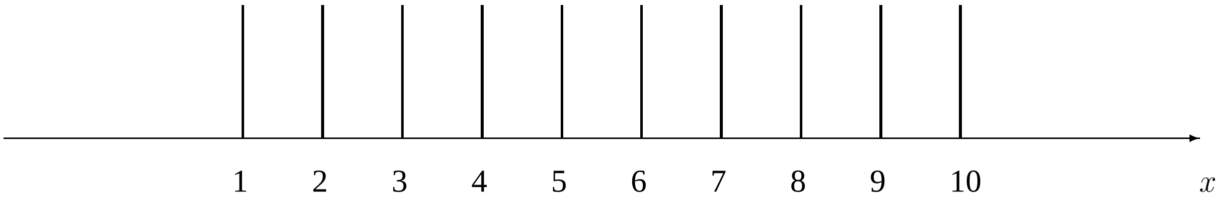


Figura 3.1: Funzione di probabilità di distribuzione uniforme discreta su $\{1, 2, \ldots, 10\}$.

e quindi la varianza di X:

$$\begin{aligned}\text{Var}(X) &= \mathbb{E}(X^2) - \mu^2 = \frac{1}{n}(1^2 + 2^2 + \cdots + n^2) - \left(\frac{n+1}{2}\right)^2 \\ &= \frac{1}{n}\frac{n(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{n+1}{12}[2(2n+1) - 3(n+1)] \\ &= \frac{n^2-1}{12} \ .\end{aligned}$$

3.2 La distribuzione binomiale

La distribuzione binomiale è la distribuzione di probabilità del numero di successi in una successione di n prove ripetute ed indipendenti in ciascuna delle quali vi è probabilità costante di successo. Data una successione di n eventi $A_1, A_2, \ldots, A_n$ indipendenti ed equiprobabili, con $P(A_1) = \cdots = P(A_n) = p$, la variabile aleatoria:

$$X := |A_1| + |A_2| + \cdots + |A_n| \tag{3.1}$$

fornisce il numero di successi ottenuto nelle n prove. Ovviamente il numero di successi che può essere ottenuto nelle n prove è un numero intero compreso fra 0 e n; pertanto X assume valori nell'insieme $\Omega_X = \{0, 1, \ldots, n\}$. Vogliamo calcolare la distribuzione su Ω_X originata dalla v.a. X.

Consideriamo dapprima il caso in cui nelle prime x prove $A_1, \ldots, A_x$ si ottenga un successo e nei successive $n - x$ prove $A_{x+1}, \ldots, A_n$ si ottenga un insuccesso, per $x \in \{0, 1, \ldots, n\}$; cioè l'evento $A_1 \ldots A_x A^c_{x+1} \ldots A^c_n$. Poichè per ipotesi gli eventi $A_1, \ldots, A_n$ sono stocasticamente indipendenti ed equiprobabili, si ha:

$$\begin{aligned}P(A_1, \ldots, A_x, A^c_{x+1}, \ldots, A^c_n) &= P(A_1) \cdots P(A_x) P(A^c_{x+1}) \cdots P(A^c_n) \\ &= \overbrace{p \cdots p}^{x \text{ volte}} \cdot \overbrace{(1-p) \cdots (1-p)}^{n-x \text{ volte}} \\ &= p^x (n-p)^{n-x} \ .\end{aligned}$$

L'evento $\{X = x\}$, cioè l'evento "si sono ottenuti x successi nelle n prove", è dato dall'unione di tutti gli eventi contenenti esattamente x successi e $n - x$ insuccessi (ovviamente l'ordine non ha alcuna importanza); tali eventi sono disgiunti ed equiprobabili in quanto ciascuno di essi ha probabilità $p^x(n-p)^{n-x}$. Il numero di tali eventi è uguale al numero di combinazioni di classe x da un insieme di n elementi. Si ha pertanto:

$$P\{X = x\} \;=\; \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, \ldots, n$$

Possiamo scrivere pertanto:

$$p(x) := \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{se } x = 0, 1, \ldots, n \\ 0 & \text{altrimenti.} \end{cases} \tag{3.2}$$

In tal caso diremo che X segue una *distribuzione binomiale* e scriveremo $X \sim B(n, p)$.

Vogliamo dimostrare che la (3.2) gode della proprietà $\sum_{x \in \Omega_X} p(x) = 1$. In base alla formula del binomio di Newton, si ha che per $a, b \in \mathbb{R}$ e per ogni n:

$$(a+b)^n \;=\; \sum_{k=0}^{n} \binom{n}{x} a^k b^{n-k} \;.$$

In particolare, per $a = p$ e $b = 1 - p$, si ottiene

$$1 \;=\; \sum_{k=0}^{n} \binom{n}{x} p^k (1-p)^{n-k} \;,$$

cioè la nostra tesi.

Esempio 3.1 Sia $X \sim B(10, 0.3)$. Si ha allora:

$$\begin{aligned} P(X=2) &= \binom{10}{2} 0.3^2 \cdot 0.7^8 = \frac{10!}{2!8!} 0.3^2 \cdot 0.7^8 = \frac{10 \cdot 9}{2} 0.3^2 \cdot 0.7^8 = 45 \cdot 0.3^2 \cdot 0.7^8 \\ &= 0.2335 \\ P(X \le 2) &= P(X=0) + P(X=1) + P(X=2) \\ &= \binom{10}{0} 0.3^0 \cdot 0.7^{10} + \binom{10}{1} 0.3^1 \cdot 0.7^9 + \binom{10}{2} 0.3^2 \cdot 0.7^8 \\ &= 0.0282 + 0.1211 + 0.2335 = 0.3828 \\ P(X > 3) &= 1 - P(X \le 2) = 1 - (P(X=0) + P(X=1) + P(X=2)) \\ &= 1 - 0.3828 = 0.6172 \;. \end{aligned}$$

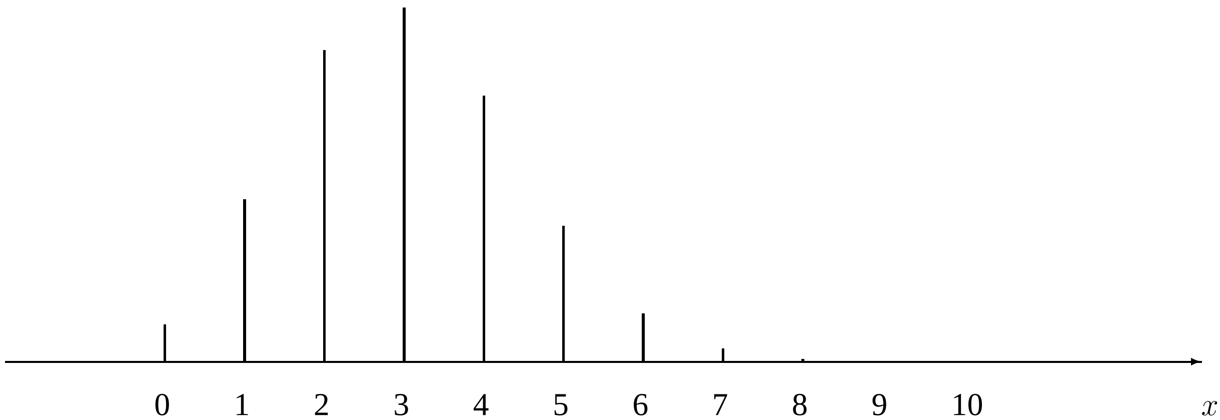


Figura 3.2: Funzione di probabilità distribuzione binomiale di parametri $n = 10$ e $p = 0.3$.

Speranza matematica e Varianza. Per il calcolo della speranza matematica e della varianza di X, premettiamo il calcolo della speranza e della varianza di $|A_i|$:

$$\mathbb{E}(|A_i|) = 1 \cdot p + 0 \cdot (1-p) \; = \; p \tag{3.3}$$

$$\text{Var}(|A_i|) = \mathbb{E}(|A_i|^2) - \{\mathbb{E}(|A_i|)\}^2 \; = \; p - p^2 = p(1-p) \; . \tag{3.4}$$

Pertanto dalla proprietà (2.22) della speranza e dalla (2.40) della varianza, la speranza matematica e la varianza di una variabile aleatoria $X \sim B(n,p)$ sono date rispettivamente da:

$$\begin{aligned} \mathbb{E}(X) &= \mathbb{E}(|A_1|) + \mathbb{E}(|A_2|) + \cdots + \mathbb{E}(|A_n|) \\ &= p + p + \cdots + p \; = \; np \qquad (3.5) \\ \text{Var}(X) &= \text{Var}(|A_1|) + \text{Var}(|A_2|) + \cdots + \text{Var}(|A_n|) \\ &= p(1-p) + p(1-p) + \cdots p(1-p) \; = \; np(1-p) \; . \qquad (3.6) \end{aligned}$$

Proposizione 3.2 Siano $X_1, \ldots, X_k$ v.a. indipendenti, con $X_i \sim B(n_i, p)$ per $i = 1, \ldots, k$. Allora la variabile aleatoria $S_k = X_1 + \cdots + X_k$ ha distribuzione $B(n_1 + \cdots + n_k, p)$.

Dimostrazione. Poichè X_1 segue una distribuzione binomiale di parametri n_1 e p, allora vuol dire che si può scrivere come:

$$X_1 = |A_1| + \cdots + |A_{n_1}| \; ,$$

dove $A_1, \ldots, A_{n_1}$ sono eventi indipendenti ed equiprobabili con $P(A_i) = p$, ($i = 1, \ldots, n_1$). Analogamente per le altre variabili $X_2, \ldots, X_k$ esisteranno eventi rispettivamente

$(A_{n_1+1},\ldots,A_{N-1+n_2}),\ldots,(A_{n_{k-1}+1},\ldots,A_{n_k})$ eventi indipendenti ed equiprobabili con probabilità p. Si ha pertanto:

$$\begin{aligned}
X_1 &= |A_1| + \cdots + |A_{n_1}| \\
X_2 &= |A_{n_1+1}| + \cdots + |A_{n_1+n_2}| \\
\cdots &= \cdots \\
X_k &= |A_{n_1+\cdots+n_{k-1}+1}| + \cdots + |A_{n_k}|
\end{aligned}$$

e quindi considerara la variabile aleatoria:

$$X = X_1 + \cdots + X_k \;=\; |A_1| + \cdots + |A_{n_1+n_2+\cdots+n_k}|$$

segue una distribuzione binomiale di parametri $n = n_1 + \cdots + n_k$ e p. □

Il successivo corollario è segue immediatamente dal caso precedente per $n_1 = \cdots = n_k = n$:

Corollario 3.3 Siano $X_1,\ldots,X_k$ v.a. indipendenti aventi distribuzione $B(n,p)$ per $i = 1,\ldots,k$. Allora la variabile aleatoria $S_k = X_1 + \cdots + X_k$ ha distribuzione $B(nk,p)$.

Esempio 3.4 Ad esempio, se X, Y sono v.a. indipendenti con $X \sim B(10, 0.2)$ e $Y \sim B(7, 0.2)$, allora $X + Y \sim (17, 0.2)$

Esempio 3.5 Sapendo che il 30 % dei passeggeri che hanno prenotato non si presenta alla partenza, una compagnia aerea accetta fino a 28 prenotazioni su un volo con capienza 24 posti. Qual è la probabilità che almeno un passeggero, che ha regolarmente prenotato, resti a terra?

Sia X =”numero di passeggeri che, avendo prenotato un posto sull’aereo, si presentano per l’imbarco”. Evidentemente si ha $X \sim B(28, 0.7)$. Poichè l’aereo dispone di 24 posti, fra coloro che hanno prenotato qualcuno non potrà partire se si presentano almeno 25 passeggeri al *check-in*; la probabilità di tale evento è:

$$P(X \geq 25) \;=\; P(X = 25) + P(X = 26) + P(X = 27) + P(X = 28)\ .$$

Poichè $X \sim B(28, 0.7)$ allora si ha:

$$\begin{aligned}
P(X = 25) &= \binom{28}{25} 0.7^{25}(1 - 0.7)^{28-25} = 0.0119 \\
P(X = 26) &= \binom{28}{26} 0.7^{26}(1 - 0.7)^{28-26} = 0.0032 \\
P(X = 27) &= \binom{28}{27} 0.7^{27}(1 - 0.7)^{28-27} = 0.006 \\
P(X = 28) &= \binom{28}{28} 0.7^{28}(1 - 0.7)^{28-28} = 4.6 \cdot 10^{-5}
\end{aligned}$$

E quindi, la probabilità che al più 24 passeggeri si presentino per l’imbarco è data da

$$P(X \geq 25) = 0.0157 = 1.57\%\ .$$

♣

Esempio 3.6 Analizzando le acque di scarico di un'industria fortemente inquinata, si è osservato che una frazione $\theta = 0.7$ di campioni di tali acque contiene una sostanza fortemente nociva e che tale sostanza di campioni inquinati si riduce a $\theta_D = 0.4$ se l'industria mette in funzione l'impianto di depurazione di cui è dotata (ma che ha un costo di esercizio molto elevato). Le autorità sanitarie effettuano senza preavviso un controllo delle acque di scarico, analizzando $n = 12$ campioni e trovando la sostanza nociva in $h = 7$ di essi. Assumendo che, prima di effettuare il prelievo, la probabilità che l'impianto di depurazione sia attivo o meno sia la stessa (cioè $P(D) = P(D^c) = 0.5$), qual è la probabilità che l'impianto sia attivo dopo aver preso visione dei risultati?

Sia X = "numero di campioni inquinati da un insieme di 12 prelievi". L'insieme di tutti i possibili campioni che si possono prelevare dalla acque di scarico, rispettivamente nei due casi D e D^c (depuratore in funzione o fermo al momento del prelievo), si può identificare con un'urna contenente palline bianche (corrispondente ai campioni contenenti la sostanza nociva) nelle proporzioni $\theta_D = 0.4$ oppure $\theta = 0.7$. Allora ricorrendo all'approssimazione binomiale, si può scrivere:

$$
\begin{aligned}
P(X = 7|D) &= \binom{12}{7} 0.4^7 0.6^5 = 0.101 \\
P(X = 7|D^c) &= \binom{12}{7} 0.7^7 0.3^5 = 0.158 \; .
\end{aligned}
$$

E quindi per il teorema di Bayes, supponendo equiprobabili le due ipotesi in esame, cioè $P(D) = P(D^c) = 1/2$ (ma se si hanno altre informazioni questa scalta non è obbligatoria), si ha:

$$
\begin{aligned}
P(D|X = 7) &= \frac{P(X = 7|D)P(D)}{P(X = 7|D)P(D) + P(X = 7|D^c)P(D^c)} = \frac{0.101}{0.101 + 0.158} = 0.390 \\
P(D^c|X = 7) &= \frac{P(X = 7|D^c)P(D^c)}{P(X = 7|D)P(D) + P(X = 7|D^c)P(D^c)} = \frac{0,158}{0.101 + 0.158} \\
&= 1 - P(D|X = 7) = 0.610 \; .
\end{aligned}
$$

♣

3.2.1 La distribuzione di Bernoulli

Nel caso in cui si abbia $n = 1$, la distribuzione binomiale prende il nome di *distribuzione di Bernoulli*. Pertanto $X \sim B(1, p)$ costituisce la variabile aleatoria concernente una singola estrazione da un'urna avente proporzione p di eventi successo. In base alla Proposizione (3.2, una v.a. $X \sim B(n, p)$ può essere vista come la somma di n v.a. indipendenti $Y_1, \ldots, Y_n$ aventi distribuzione di Bernoulli, cioè $Y_i \sim B(1, p)$ per $i = 1, \ldots, n$.

3.3 La distribuzione ipergeometrica

La distribuzione ipergeometrica segue lo stesso schema della legge binomiale ma, a differenza di quest'ultima, si considerano esperimenti senza reimmissione. Considerata una successione (finita) di n estrazioni senza ripetizione da un'urna contenente N palline, di cui una frazione pN bianche (con $0 < p < 1$) e la restante parte $qN = (1-p)N$ non bianche, sia X il numero di palline bianche X ottenute in una successione di n estrazioni senza reimmissione. In questo caso Ω_X è costituito dall'insieme dei valori interi tali che:

$$\max\{0, n-qN\} \leq x \leq \min\{n, pN\} .$$

Infatti in questo caso si tratta di estrazioni senza reimmissione ed indicano che il numero minimo e massimo di successi ottenibili in n estrazioni dipendono dal numero N di palline e dalla proporzione p di palline bianche contenute nell'urna. Pertanto se il numero di palline estratte n è superiore al numero di palline nere presenti nell'urna, il numero minimo di palline bianche non può essere 0 ma pari a $n-qN$; al contrario non si può estrarre un numero di palline bianche superiore a quello presente nell'urna.

La funzione di probabilità data da:

$$p(x) := \begin{cases} \dfrac{\dbinom{pN}{x}\dbinom{(1-p)N}{n-x}}{\dbinom{N}{n}} & \text{se } \max\{0, n-qN\} \leq x \leq \min\{n, pN\} \\ 0 & \text{altrimenti.} \end{cases}$$

In questo caso diremo che la variabile aleatoria X segue una *distribuzione ipergeometrica* di parametri interi positivi N, p e n, e scriveremo $X \sim \mathcal{H}(N,p,n)$. La verifica della condizione $\sum_x p(x) = 1$ è abbastanza laboriosa e verrà qui omessa.

Si noti che anche in questo caso la v.a. X è data da:

$$X = |A_1| + |A_2| + \cdots + |A_n| \tag{3.7}$$

dove gli eventi $A_1, \ldots, A_n$ sono equiprobabili, con $P(A_i) = p$ per $i = 1, \ldots, n$ e correlati. Si ha infatti:

$$P(A_1) = p$$
$$P(A_2) = P(A_2|A_1)P(A_1) + P(A_2|A_1^c)P(A_1^c) = \frac{pN-1}{N-1}p + \frac{pN}{N-1}(1-p) = p .$$

Iterando il ragionamento per $i = 3, \ldots, N$ si vede che $P(A_i) = p$ per ogni i.

Si può dimostrare che per $N \to \infty$, mantenendo costante la proporzione p di palline bianche, si ha:

$$\lim_{N\to\infty} \frac{\dbinom{pN}{x}\dbinom{(1-p)N}{n-x}}{\dbinom{N}{n}} = \binom{n}{x} p^x (1-p)^{n-x} \tag{3.8}$$

e pertanto la distribuzione binomiale costituisce il limite della distribuzione ipergeometrica per $N \to \infty$.

Speranza matematica e Varianza. In base alla (3.7), la speranza matematica è da:

$$\mathbb{E}(X) = \mathbb{E}(|A_1|) + \mathbb{E}(|A_2|) + \cdots + \mathbb{E}(|A_n|) = np , \tag{3.9}$$

Per quanto riguarda la varianza, dalla (3.7) segue:

$$\text{Var}(X) = \sum_{i=1}^{n} \text{Var}(|A_i|) + \sum_{i \neq j} \text{Cov}(|A_i|, |A_j|) . \tag{3.10}$$

In base a quanto visto in precedenza nella (3.4) si ha Var$(|A_i|) = pq$.

Per il calcolo della covarianza Cov$(|A_i|, |A_j|)$, consideriamo la relazione (2.35) che nel caso qui in esame si scrive:

$$\text{Cov}(|A_i|, |A_j|) = \mathbb{E}(|A_i| \cdot |A_j|) - \mathbb{E}(|A_i|)\mathbb{E}(|A_j|) .$$

Si noti che risulta $|A_i| \cdot |A_j| = |A_i \cap A_j| = 1$ se e solo l'evento $A_i \cap A_j$ è vero; e quindi si ha:

$$\begin{aligned}\mathbb{E}(|A_i| \cdot |A_j|) &= \mathbb{E}(|A_i \cap A_j|) = P(A_i \cap A_j) = P(A_i)P(A_j|A_i) \\ &= p\frac{pN-1}{N-1} .\end{aligned}$$

Segue pertanto

$$\text{Cov}(|A_i|, |A_j|) = \mathbb{E}(|A_i| \cdot |A_j|) - \mathbb{E}(|A_i|)\mathbb{E}(|A_j|) = p\frac{pN-1}{N-1} - p^2$$

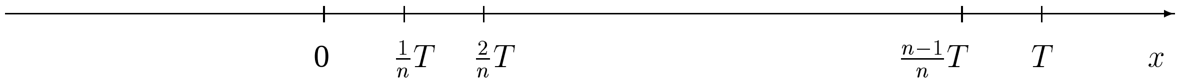
e quindi sostituendo nella (3.10) si ottiene:

$$\begin{aligned}\text{Var}(X) &= \sum_{i=1}^{n} \text{Var}(|A_i|) + \sum_{i \neq j} \text{Cov}(|A_i|, |A_j|) = npq + n(n-1)p\left(\frac{pN-1}{N-1} - p\right) \\ &= npq + n(n-1)p\frac{p-1}{N-1} \\ &= npq\left(1 - \frac{n-1}{N-1}\right) .\end{aligned} \tag{3.11}$$

In particolare, per $n = 1$ la varianza della distribuzione ipergeometrica coincide con quella della distribuzione binomiale in quanto se si effettua un'unica estrazione lo schema è lo stesso nei due casi. Inoltre per valori grandi di N la varianza della distribuzione ipergeometrica tende a quella della distribuzione binomiale, come del resto prevedibile in base alla (3.8).

Esercizio 3.7 Gli archivi di un ufficio sono stati memorizzati in 30 dischi da 100 Mbyte, contenenti ognuno 100 file. Un programma dovrà accedere a 28 di questi file (tutti diversi). Qual è la probabilità che esso non debba usare il primo disco?

La situazione in esame si modellizza con uno schema successo-insuccesso senza rimpiazzo: sono 28 prove ripetute in ognuna delle quali viene scelto un file fra i 3000 possibili. Considereremo "successo" la scelta di un file appartenente al primo disco (sono 100) e insuccesso la



scelta di uno degli altri dischi (2900 files). Non c'è restituzione perchè un file, una volta letto, non verrà più richiamato in seguito, quindi $X \sim \mathcal{H}(3000, \frac{1}{30}, 28)$. La probabilità richiesta è dunque

$$P(X = 0) = \frac{\binom{100}{0}\binom{2900}{28}}{\binom{3000}{28}} = 0.385 \ .$$

Se invece le 28 voci fossero state scelte a caso con possibilità di ripetizione, il numero di volte in cui il programma avrebbe dovuto accedere al disco n.1 sarebbe stata una $B(28, \frac{1}{30})$. La probabilità che il disco non venga consultato è

$$P(X = 0) = \binom{28}{0}\left(\frac{1}{30}\right)^0\left(1 - \frac{1}{30}\right)^{28} = 0.387.$$

Si noti che i due risultati sono molto vicini. Ciò è dovuto al fatto che le varianze delle due distribuzioni binomiale e ipergeometrica sono praticamente uguali, infatti nella (3.11) il rapporto:

$$\frac{n-1}{N-1} = \frac{27}{2999} = 0.009$$

è molto prossimo allo zero. ♣

3.4 La distribuzione di Poisson

La distribuzione di Poisson è certamente una delle leggi di probabilità che trova maggiori applicazioni per descrivere fenomeni del mondo reale. Tale distribuzione è anche detta *distribuzione degli eventi rari* per sottolineare il fatto che in ciascuno dei casi in cui si applica, la probabilità che il singolo evento si verifichi è estremamente bassa. La distribuzione di probabilità di Poisson può quindi servire da modello matematico per fenomeni del tipo: il numero di automobili che si fermano ad una stazione di servizio in un certo intervallo di tempo, numero di utenti che prelevano del denaro da uno sportello bancomat in un certo intervallo di tempo, etc.

La distribuzione di Poisson può essere costruita a partire da una successione di prove indipendenti, secondo lo schema che ha condotto alla legge binomiale.

Si consideri un intervallo di lunghezza $T > 0$ (e per semplicità si pensi a $(0, T)$) suddiviso in n sottointervalli di lunghezza T/n. Si supponga che un certo evento A (successo) possa manifestarsi con la stessa probabilità $p = \alpha(T/n)$ – proporzionale, secondo un coefficiente $\alpha > 0$, all'ampiezza T/n del singolo sottointervallo – ed in modo indipendente in ognuno dei sottontervalli, e sia $1 - p = 1 - \alpha(T/n)$ la probabilità che nel generico sottointervallo non si manifesti l'evento A. Vogliamo ricavare la probabilità che l'evento A si presenti in x degli n sottointervalli consederati.

Indicato con X il numero dei sottointervalli in cui si manifesta l'evento A, cioè il numero di successi, le ipotesi adottate consentono di applicare la distribuzione binomiale, e si ha:

$$p(x) = \binom{n}{x}\left(\frac{\alpha T}{n}\right)^x\left(1 - \frac{\alpha T}{n}\right)^{n-x} \qquad x = 0, 1, 2, \dots, n.$$

All'aumentare del numero n dei sottointervalli di $(0, T)$, si ha:

$$\lim_{n\to\infty}\binom{n}{x}\left(\frac{\alpha T}{n}\right)^x\left(1 - \frac{\alpha T}{n}\right)^{n-x} = \frac{(\alpha T)^x}{x!}\mathrm{e}^{-\alpha T} = \frac{\lambda^x}{x!}\mathrm{e}^{-\lambda}$$

dove si è posto $\lambda = \alpha T > 0$.

L'espressione precedente si ottiene considerando che:

$$\begin{aligned}\binom{n}{x}\left(\frac{\alpha T}{n}\right)^x\left(1 - \frac{\alpha T}{n}\right)^{n-x} &= \frac{n(n-1)\cdots(n-x+1)(n-x)!}{x!(n-x)!}\left(\frac{\alpha T}{n}\right)^x \times \\ &\times \left(1 - \frac{\alpha T}{n}\right)^n\left(1 - \frac{\alpha T}{n}\right)^{-x} \\ &= \frac{1}{x!}\frac{n(n-1)\cdots(n-x+1)}{n^x}(\alpha T)^x\left(1 - \frac{\alpha T}{n}\right)^n\left(1 - \frac{\alpha T}{n}\right)^{-x}\end{aligned}$$

e che per $n \to \infty$ si ha:

$$\begin{aligned}\frac{n(n-1)\cdots(n-x+1)}{n^x} = \frac{n\cdot n(1-\frac{1}{n})\cdots n(1-\frac{x-1}{n})}{n\cdot n\cdots n} &\to 1 \\ \left(1 - \frac{\alpha T}{n}\right)^n &\to e^{-\alpha T} \\ \left(1 - \frac{\alpha T}{n}\right)^{-x} &\to 1\ .\end{aligned}$$

La legge di Poisson appare, in questo modo, come una distribuzione che approssima la distribuzione binomiale nel caso in cui in questa si abbia un numero di prove n molto grande ed una probabilità di successo p in ciascuna prova molto piccola (essendo $np = \lambda$ = costante).

Più in generale, i campi di applicazione della distribuzione di Poisson si evincono meglio dallo schema seguente. Si supponga di essere interessati ad un certo evento (ad esempio, incidenti automobilistico su un tratto di strada). Si indichi con $\eta(s, s+t)$ il numero di volte in cui l'evento in questione si verifica nell'intervallo temporale $(s, s+t)$ e supponiamo che:

1. La probabilità che l'evento si verifichi in un intervallo di tempo di ampiezza $t > 0$ sia uguale a αt, con $\alpha > 0$, indipendentemente dall'origine dell'intervallo e per valori di t sufficientemente piccolo:

$$P(\eta(s, s+t) = 1) \;=\; \alpha t + o(t) \qquad s \geq 0$$

2. la probabilità che in un intervallo di ampiezza h l'evento si verifichi più di una volta sia trascurabile:
$$P(\eta(s, s+t) > 1) = o(t)$$
3. Il numero degli eventi che si verificano in due intervalli disgiunti siano stocasticamente indipendenti, cioè dati $s_1 < s_2 < s_3 < s_4$, allora le v.a. $\eta(s_1, s_2)$ e $\eta(s_3, s_4)$ sono stocasticamente indipendenti.

Sotto le precedenti ipotesi si può dimostrare che:
$$P(\eta(0,t) = x) = \mathrm{e}^{-\alpha t}\frac{(\alpha t)^x}{x!} \qquad\qquad x = 0, 1, 2, \ldots$$
Diremo che una variabile aleatoria X segue una *distribuzione di Poisson* di parametro $\lambda > 0$, e scriveremo $X \sim \mathcal{P}(\lambda)$, se ha la seguente funzione di probabilità:
$$p(x) := \begin{cases} \mathrm{e}^{-\lambda}\dfrac{\lambda^x}{x!} & \text{se } x = 0, 1, 2, \ldots \\ 0 & \text{altrimenti.} \end{cases} \tag{3.12}$$

Per quanto riguarda la verifica della (2.6) $\sum_{x=0}^{\infty} p(x) = 1$, consideriamo lo sviluppo in serie della funzione esponenziale e^{λ}:
$$\mathrm{e}^{\lambda} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$$
da cui, moltiplicando ambo i membri per $\mathrm{e}^{-\lambda}$ si ottiene
$$1 = \sum_{k=0}^{\infty} \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$$
cioè la nostra tesi.

Speranza matematica e Varianza. La speranza matematica è data da:
$$\mathbb{E}(X) := \sum_{x=0}^{\infty} x \cdot \mathrm{e}^{-\lambda}\frac{\lambda^x}{x!} = \mathrm{e}^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda \cdot \lambda^{x-1}}{(x-1)!} = \lambda \mathrm{e}^{-\lambda}\sum_{z=0}^{\infty}\frac{\lambda^z}{z!} = \lambda\ .$$
Per il calcolo della varianza premettiamo il calcolo del momento secondo:
$$\mathbb{E}(X^2) = \sum_{x=0}^{\infty} x^2 \cdot \mathrm{e}^{-\lambda}\frac{\lambda^x}{x!} = \mathrm{e}^{-\lambda}\lambda\sum_{x=1}^{\infty} x\frac{\lambda^{x-1}}{(x-1)!}$$
Posto $z = x - 1$, segue
$$\mathbb{E}(X^2) = \mathrm{e}^{-\lambda}\lambda\sum_{z=0}^{\infty}(z+1)\frac{\lambda^z}{z!} = \mathrm{e}^{-\lambda}\lambda\sum_{z=0}^{\infty}(z+1)\frac{\lambda^z}{z!} = \lambda\mathbb{E}(X+1) = \lambda(\lambda+1)$$
Segue pertanto:
$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mu^2 = \lambda(\lambda+1) - \lambda^2 = \lambda\ .$$
Pertanto, nel caso della distribuzione di Poisson, speranza matematica e varianza coincidono.

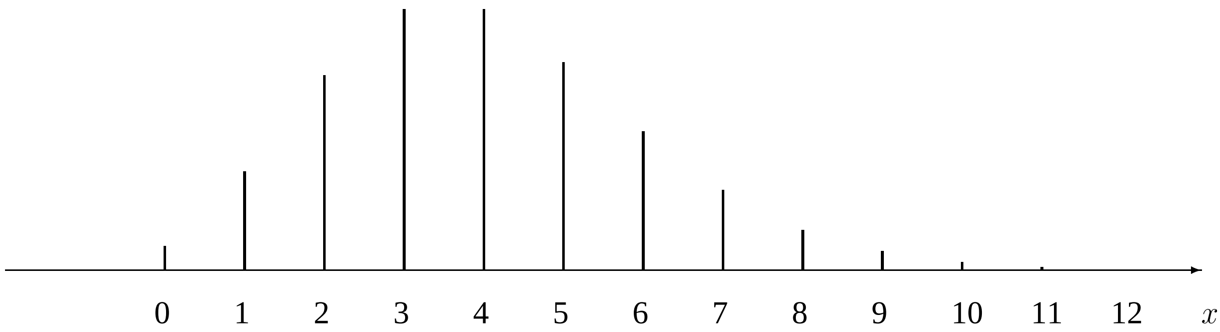


Figura 3.3: Funzione di probabilità distribuzione di Poisson con parametro $\lambda = 4$.

Approssimazione della distribuzione binomiale con quella di Poisson. Per $n > 50$ e $p < 0.1$, si può approssimare una distribuzione Binomiale $B(n, p)$ con una di Poisson $\mathcal{P}(\lambda)$ di parametro $\lambda = np$

Proposizione 3.8 Siano $X_1, \ldots, X_k$ v.a. indipendenti, con $X_i \sim P(\lambda_i)$ per $i = 1, \ldots, k$. Allora la variabile aleatoria $S_k = X_1 + \cdots + X_k$ ha distribuzione $P(\lambda_1 + \cdots + \lambda_k)$.

Esercizio 3.9 All'inizio di ogni mese, un negoziante ordina un certo prodotto per le vendite di tutto il mese. Egli può plausibilmente assumere che la variabile delle unità di prodotto venduto durante il mese segua una legge di Poisson con parametro $\lambda = 4$. Quante unità di prodotto deve avere in magazzino affinchè abbia una probabilità dell'88% di non restare senza prodotto?

In base ai dati del problema, il numero X di unità vendute in un mese segue una distribuzione di Poisson di parametro $\lambda = 4$, cioè $X \sim \mathcal{P}(4)$. Ogni mese possono essere vendute 0 unità di prodotto oppure un'unità di prodotto, etc. Indicato con k il numero di unità di prodotto da ordinare, si richiede di determinare il più piccolo valore di k per cui risulti:

$$P(X \leq k) = \sum_{x=1}^{k} P(X = x) = \sum_{x=1}^{k} \mathrm{e}^{-4} \frac{4^x}{x!} \; .$$

Dalla funzione di probabilità della distribuzione di Poisson $\mathcal{P}(4)$ si ricavano i valori:

x	$P(X = x)$	$P(X \leq x)$
0	0.0183	0.0183
1	0.0733	0.0916
2	0.1465	0.2381
3	0.1954	0.4335
4	0.1954	0.6289
5	0.1563	0.7852
6	0.1042	0.8894
7	0.0595	0.9489
8	0.0298	0.9787

e quindi segue che è necessario ordinare almeno $k = 6$ unità di prodotto. ♣

Esercizio 3.10 In una data succursale bancaria è stata valutata una probabilità $p = 0.001$ che una persona effettui un prelievo superiore a 3000 euro. Calcolare la probabilità che, durante una normale mattina di lavoro in cui si presentano 200 persone, vi siano più di due persone che effettuino un prelievo superiore a 3000 euro.

Il problema richiede il calcolo della probabilità che in una successione di $n = 200$ prove ripetute ed indipendenti, ciascuna delle quali presenta probabilità $p = 0.001$ di successo e $q = 1 - p = 0.999$ di insuccesso, si verifichi un variabile di successi maggiore di due.

Il modello del problema è pertanto costituito da una distribuzione binomiale di parametri $n = 200$ e $p = 0.001$. Tuttavia, essendo la probabilità di successo in ciascuna prova molto piccola, lo stesso problema può essere formalizzato mediante una distribuzione di Poisson di parametro $\lambda = np = 0.2$.

Indichiamo con X il "numero di persone che effettuano un prelievo superiore a 6 milioni in un normale giorno di lavoro", il problema pertanto richiede il calcolo di

$$P\{X > 2\} = \sum_{k=3}^{200} P\{X = k\} \; .$$

Ovviamente il calcolo può essere semplificato tenendo presente che

$$P\{X > 2\} = 1 - P\{X \leq 2\} = 1 - [P\{X = 0\} + P\{X = 1\} + P\{X = 2\}] \; .$$

Utilizziamo dapprima una distribuzione binomiale, in tal caso $X \sim B(200, 0.001)$, e pertanto si ha:

$$\begin{aligned}
P\{X = 0\} &= \binom{200}{0} p^0 \,(1-p)^{200} = 0.999^{200} = 0.8186 \\
P(X = 1) &= \binom{200}{1} p \,(1-p)^{199} = 0.001\, 0.999^{199} = 0.1639 \\
P(X = 2) &= \binom{200}{2} p^2 \,(1-p)^{198} = 0.001^2\, 0.999^{198} = 0.0163
\end{aligned}$$

da cui

$$P\{X > 2\} = 1 - [P\{X = 0\} + P\{X = 1\} + P\{X = 2\}] = 0.0012 \; .$$

Utilizzando una distribuzione di Poisson di parametro $\lambda = np = 0, 2$, cioè $X \sim \mathcal{P}(0.2)$, si ha:

$$\begin{aligned}
P\{X = 0\} &= \frac{\mathrm{e}^{-\lambda}\lambda^0}{0!} = \mathrm{e}^{-0.2} \; = \; 0.8187 \\
P\{X = 1\} &= \frac{\mathrm{e}^{-\lambda}\lambda^1}{1!} = \mathrm{e}^{-0,2}\, 0, 2 = 0.1637 \\
P\{X = 2\} &= \frac{\mathrm{e}^{-\lambda}\lambda^2}{2!} = \frac{\mathrm{e}^{-0.2}\, 0, 2^2}{2} = 0.0164
\end{aligned}$$

e quindi si perviene allo stesso risultato

$$P\{X > 2\} = 0.0012\ .$$

♣

3.5 La distribuzione multinomiale

Una generalizzazione della distribuzione binomiale può condurre alla cosidetta *distribuzione multinomiale* nel modo seguente. Consideriamo una successione di n prove ripetute ed indipendenti ciascuna delle quali ammetta k possibili risultati $A_1, \ldots, A_k$ mutuamente esclusivi. Sia p_i la probabilità che l'esperimento fornisca il risultato A_i, per $i = 1, \ldots, k$.

Siano $x_1, \ldots, x_{k-1}$ interi non negativi tali che $x_1 + \cdots + x_{k-1} \leq n$. La probabilità che esattamente x_i prove abbiano risultato A_i, per $i = 1, \ldots k-1$, e quindi che le restanti $x_k = n - (x_1 + \ldots + x_{k-1})$ prove abbiano risultato A_k, è chiaramente:

$$\binom{n}{x_1\ x_2\ \cdots x_k} p_1^{x_1} \cdots p_k^{x_k}\ ,$$

dove la quantità:

$$\binom{n}{x_1\ x_2\ \cdots x_k} := \frac{n!}{x_1! \cdots x_k!}$$

costituisce il *coefficiente multinomiale* ed è uguale al numero di modi in cui scegliere, fra le n prove, quelle in cui si verifica A_1, quelle in cui si verifica A_2, etc.

Se $(X_1, \ldots, X_k)$ è un vettore aleatorio dove X_j è il numero (aleatorio) di volte in cui l'evento A_j si è verificato, per $x_j = 0, 1, \ldots, n$, allora la funzione di probabilità di $(X_1, \ldots, X_k)$ è data da:

$$P(X_1 = x_1, \ldots, X_k = x_k) \quad = \quad \begin{cases} \dfrac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k} & \text{se } \sum_{i=1}^{k} x_i = n \\ 0 & \text{altrimenti} \end{cases} \tag{3.13}$$

La distribuzione del vettore aleatorio $(X_1, \ldots, X_k)$ si chiama *distribuzione multinomiale* e scriveremo $X \sim M(n, p_1, \ldots, p_k)$.

Per la distribuzione multinomiale si ha:

$$\begin{aligned} \mathbb{E}(X_j) &= np_j\ , \\ \text{Var}(X_j) &= np_j(1 - p_j)\ , \\ \text{Cov}(X_i, X_j) &= E[(X_i - np_i)(X_j - np_j)] \ = \ -np_i p_j\ . \end{aligned}$$

In particolare, ne segue che il coefficiente di correlazione fra X_i e X_j è dato da:

$$\rho_{ij} \quad = \quad -\left(\frac{p_i p_j}{(1 - p_i)(1 - p_j)}\right)^{1/2} \qquad\qquad i, j = 1, \ldots, k-1,\ (i \neq j)\ .$$

3.6 La distribuzione geometrica

La distribuzione geometrica deriva dallo stesso schema della distribuzione binomiale ma nel caso ora in esame, invece di analizzare il numero complessivo di eventi favorevoli conseguiti in n prove, ci si interessa al numero di prove necessario per ottenere il primo successo, cioè il *tempo di attesa del primo successo*.

Data una successione di eventi equiprobabili ed indipendenti $A_1, A_2, \ldots$ con $P(A_i) = p$ per ogni $i = 1, 2, \ldots$, con $0 < p < 1$, denotiamo con X la quantità aleatoria ”indice del primo successo nella sequenza assegnata”. Allora la v.a. X assume valori nell'insieme dei numeri naturali ed ha funzione di probabilità

$$p(x) := \begin{cases} p(1-p)^{x-1} & \text{se } x = 1, 2, \ldots \\ 0 & \text{altrimenti} \end{cases}$$

Diremo in tal caso che X segue una *distribuzione geometrica* di parametro p e scriveremo $X \sim G(p)$.

Per quanto riguarda la verifica della relazione $\sum_{x=1}^{\infty} p(x) = 1$, posto $q = 1 - p$, si può immediatamente verificare che risulta:

$$\sum_{r=0}^{k} q^r = 1 + q + \cdots + q^k = \frac{1 - q^{k+1}}{1 - q} = \frac{1 - (1-p)^{k+1}}{p} .$$

da cui

$$\sum_{r=0}^{\infty} q^r = \lim_{k\to\infty} \sum_{r=0}^{k} q^r = \lim_{k\to\infty} \frac{1 - (1-p)^{k+1}}{p} = \frac{1}{p} .$$

cioè:

$$\sum_{r=0}^{\infty} (1-p)^r = \frac{1}{p} \qquad \text{o, equivalentemente,} \qquad \sum_{r=1}^{\infty} (1-p)^{r-1} = \frac{1}{p} .$$

Da tale relazione segue immediatamente:

$$\sum_{r=1}^{\infty} (1-p)p^{r-1} = 1 , \tag{3.14}$$

e quindi la verifica della proprietà $\sum_x p(x) = 1$.

Nel caso di leggi geometriche, particolare interesse riveste la seguente relazione:

$$\begin{aligned} P\{X \geq x\} &= \sum_{r=x}^{\infty} p(1-p)^{r-1} = \sum_{r=1}^{\infty} p(1-p)^{r-1+(x-1)} = (1-p)^{x-1} \sum_{r=1}^{\infty} p(1-p)^{r-1} \\ &= (1-p)^{x-1} , \end{aligned} \tag{3.15}$$

essendo $\sum_{r=1}^{\infty} p(1-p)^{r-1} = 1$ per la (3.14).

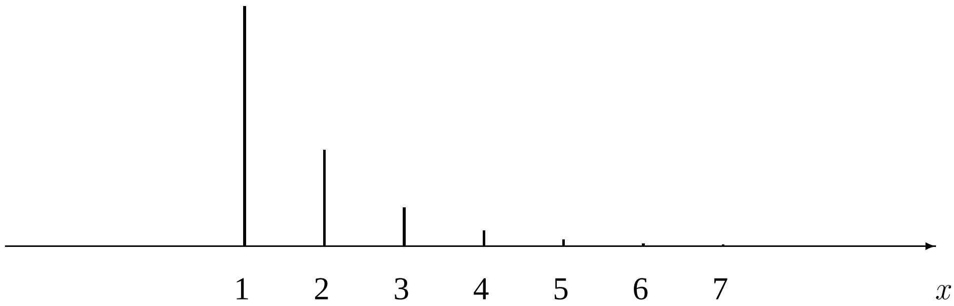


Figura 3.4: Funzione di probabilità distribuzione geometrica con parametro $p = 0,4$.

Speranza matematica e Varianza. Per quanto riguarda la speranza matematica si ha:

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{x=1}^{\infty} xp(1-p)^{x-1} = p\sum_{x=1}^{\infty} x(1-p)^{x-1} = -p\sum_{x=1}^{\infty} \frac{\mathrm{d}}{\mathrm{d}p}(1-p)^x \\
&= -p\frac{\mathrm{d}}{\mathrm{d}p}\sum_{x=1}^{\infty}(1-p)^x = -p\frac{\mathrm{d}}{\mathrm{d}p}\frac{1-p}{1-(1-p)} = -p\frac{\mathrm{d}}{\mathrm{d}p}\frac{1-p}{p} = \frac{1}{p}\ .
\end{aligned}$$

Per il calcolo della varianza, premettiamo il calcolo di $\mathbb{E}(X^2)$:

$$\begin{aligned}
\mathbb{E}(X^2) &= \sum_{x=1}^{\infty} x^2 p(1-p)^{x-1} = \sum_{x=1}^{\infty}(x^2 - x + x)p(1-p)^{x-1} \\
&= \sum_{x=1}^{\infty} x(x-1)p(1-p)^{x-1} + \sum_{x=1}^{\infty} xp(1-p)^{x-1} \\
&= p(1-p)\sum_{x=1}^{\infty} x(x-1)(1-p)^{x-2} + \frac{1}{p} \\
&= p(1-p)\sum_{x=1}^{\infty}\frac{\mathrm{d}^2}{\mathrm{d}p^2}(1-p)^x + \frac{1}{p}0p(1-p)\frac{\mathrm{d}^2}{\mathrm{d}p^2}\frac{1-p}{p} + \frac{1}{p} \\
&= p(1-p)\frac{2}{p^3} + \frac{1}{p} = \frac{2(1-p)}{p^2} + \frac{1}{p}0\frac{2-p}{p^2}\ .
\end{aligned}$$

Segue pertanto:

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mu^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}\ .$$

Proprietà di assenza di memoria. La distribuzione geometrica è caratterizzata dalla seguente proprietà di *assenza di memoria*.

Proposizione 3.11 $X \sim G(p)$ se e solo se

$$P\{X = r + n \mid X > n\} = P\{X = r\}\ ,$$

per qualunque coppia r, n di numeri interi.

Dimostrazione. La proprietà in esame segue da un'applicazione della formula della probabilità condizionale. Si ha infatti:

$$\begin{aligned} P\{X = r + n \mid X > n\} &= \frac{P\{X = r + n, X > n\}}{P\{X > n\}} \\ &= \frac{P\{X = r + n\}}{P\{X > n\}} \, . \end{aligned} \tag{3.16}$$

Poichè, in base alla (3.15), risulta:

$$P\{X > n\} = P\{X \geq n + 1\} = (1 - p)^n$$

allora dalla (3.16) segue per $r \geq 1$:

$$P\{X = r + n \mid X > n\} = \frac{p(1-p)^{n+r-1}}{(1-p)^n} = p(1-p)^{r-1} = P\{X = r\} \, .$$

□

Più in generale, la proprietà di assenza di memoria implica che

$$P\{X \geq r + n \mid X > n\} = P\{X \geq r\} \, .$$

Tale risultato afferma che, se un evento E non si è verificato nelle prime n prove, allora la probabilità che esso non si verifichi nelle successive r prove è la stessa di quella di non manifestarsi nelle prime r prove.

Esempio 3.12 In una certa località balneare la probabilità che piova in un qualunque giorno del mese di agosto è $p = 0,05$. Assumendo l'indipendenza qual è la probabilità che la prima pioggia del mese avvenga il 15 agosto? E prima del 15 agosto?
Si tratta di un'applicazione della distribuzione geometrica. Se X indica la quantità aleatoria "numero di giorni (a partire dal primo agosto) fino alla prima pioggia", allora $X \sim G(0,04)$, pertanto si avrà:

$$\begin{aligned} P\{X = 15\} &= 0.05 \cdot 0.95^{14} \cong 0.024 \, , \\ P\{X \leq 14\} &= 1 - 0.95^{14} \cong 0.5124 \, . \end{aligned}$$

♣

Esempio 3.13 Si consideri un macchinario la cui durata di funzionamento, in un'opportuna unità di misura (mese, anno o altro), sia retta dalla legge geometrica di densità:

$$p(x) \;=\; 0.2 \cdot (0,8)^{x-1} \qquad\qquad x = 1, 2, \ldots$$

Sapendo che il macchinario non ha ancora cessato di funzionare dopo 5 periodi (unità di misura), qual è la probabilità che si guasti all'ottavo periodo?

$$P\{X = 8 | X > 5\} \;=\; \frac{P\{X = 8\}}{P\{X > 5\}} \;=\; P\{X = 3\} \;=\; 0.2 \cdot (0.8)^2 \;=\; 0.128 \, .$$

♣

3.7 La distribuzione binomiale negativa o di Pascal

La distribuzione binomiale negativa è una generalizzazione della distribuzione geometrica. Sia data una successione di prove indipendenti ognuna delle quali produce un successo con probabilità $0 < p < 1$ oppure un insuccesso con probabilità $1 - p$. La distribuzione binomiale negativa è la distribuzione di probabilità del numero di prove per ottenere esattamente r successi. Diremo che una v.a. X segue una *distribuzione binomiale negativa* di parametri interi positivi $r \geq 1$ e $0 < p < 1$, e scriveremo $X \sim BN(r, p)$, se ha la seguente distribuzione di probabilità:

$$p(x) = \begin{cases} \binom{x-1}{r-1} p^r (1-p)^{x-r} & \text{se } x = r, r+1, \ldots \\ 0 & \text{altrimenti.} \end{cases} \tag{3.17}$$

In particolare per $r = 1$ si ottiene la distribuzione geometrica.

Speranza matematica e Varianza. La speranza matematica e la varianza di una v.a. $X \sim BN(r, p)$ sono date rispettivamente da:

$$\begin{aligned} \mathbb{E}(X) &= \frac{r}{p} \\ \text{Var}(X) &= \frac{r(1-p)}{p^2} \end{aligned}$$

Si noti che, dati $X \sim G(p)$ e $Y \sim BN(r, p)$, risulta

$$\begin{aligned} \mathbb{E}(Y) &= r\,\mathbb{E}(X) \\ \text{Var}(Y) &= r\,\text{Var}(X)\ . \end{aligned}$$

Proposizione 3.14 Siano $X_1, \ldots, X_k$ v.a. indipendenti, con $X_i \sim BN(r_i, p)$ per $i = 1, \ldots, k$. Allora la variabile aleatoria $S_k = X_1 + \cdots + X_k$ ha distribuzione $NB(r_1 + \cdots + r_k, p)$.

Corollario 3.15 Siano $X_1, \ldots, X_k$ v.a. indipendenti aventi distribuzione geometrica di parametro p . Allora la variabile aleatoria $S_k = X_1 + \cdots + X_k$ ha distribuzione $BN(k, p)$.

3.8 Altri esercizi

Esercizio 3.16 n commensali si ripartiscono a caso intorno ad una tavola rotonda. Sia X il più piccolo numero (aleatorio) di commensali che separa Pietro e Paolo. Calcolare la legge di X e la speranza matematica di X.

Esercizio 3.17 Si supponga che in una grande azienda si effettuino mediamente quattro interventi di manutenzione del settore informatico in una giornata lavorativa di 8 ore. Si assuma che il numero di interventi di manutenzione nei periodi indicati segua una distribuzione di Poisson.

1. Valutare la probabilità che in due ore non vi sia nessun intervento di manutenzione;
2. Valutare la probabilità che in un'ora vengano effettuati almeno due interventi di manutenzione.

Esercizio 3.18 Data una successione di eventi indipendenti ed equiprobabili con probabilità uguale a 1/100, determinare

1. la probabilità che, osservandone 100, non se ne verifichi alcuno;
2. La probabilità che su 200, se ne verifichino almeno 3;
3. la probabilità che su 50 se ne verifichino più di due;
4. il numero minimo n di prove necessarie affinchè la probabilità di avere almeno un successo sia maggiore di 0.98.

Esercizio 3.19 Un'urna contiene N palline di cui 30 bianche e le restanti nere. Calcolare N sapendo che la probabilità di ottenere 2 palline bianche in 4 estrazioni con restituzione è pari a $3^3/2^7$ e che le palline nere sono più numerose delle palline bianche.

Esercizio 3.20 Da un'urna U, contenente 2 palline bianche e 2 nere, e da un'urna V, contenente 2 palline bianche e tre nere, si estraggono rispettivamente 2 palline e 3 palline senza restituzione. Indicando con X e Y rispettivamente i numeri aleatori di palline bianche estratte rispettivamente da U e da V, calcolare $P(X+Y>2)$, $P(X=1|X+Y>2)$ e $\mathbb{E}(X+Y)$

Esercizio 3.21 Date tre variabili aleatorie, indipendenti e con distribuzione geometrica di parametro p, calcolare:

1. la speranza matematica m della variabile $X+Y-Z$;
2. la probabilità dell'evento $\{2.5 \leq X+Y+Z \leq 4.2\}$;
3. la covarianza delle v.a. $U=X+Y$ e $V=X-Y$.

Esercizio 3.22 Un canale di trasmissione dati può ricevere messaggi binari da due sorgenti diverse A e B con probabilità $\frac{1}{2}$ ciascuna. Ognuna delle due sorgenti produce messaggi in cui i bit successivi sono tra loro indipendenti. Per la sorgente A i bit possono essere 1 oppure 0 con probabilità $\frac{1}{2}$, mentre per la sorgente B, il valore 1 si verifica con probabilità $\frac{1}{4}$ e lo 0 con probabilità $\frac{3}{4}$. Un messaggio di lunghezza n viene ricevuto in cui si osserva una proporzione di 1 pari a 0.4 (cioè in cui si trovano $0.4\,n$ bit uguali a 1).

1. Supponiamo $n=10$. Qual è la probabilità che si tratti della sorgente A? Quale delle due sorgenti è la più probabile? E se invece fosse $n=100$?
2. Rispondere alle stesse domande del punto precedente supponendo che a priori i messaggi provengano con probabilità 30% dalla sorgente A e 70% dalla sorgente B.

Esercizio 3.23 Un'urna U_1 contiene una pallina bianca e 2 nere, mentre un'urna U_2 contiene 2 palline bianche ed una nera. Da U_1 si estraggono due palline che vengono inserite in U_2. Successivamente da U_2 si estrae una pallina che viene inserita in U_1. Considerati gli eventi:

A : le due palline inserite in U_2 sono nere
B : la pallina inserita in U_1 è bianca

calcolare $P(B)$, $P(A|B)$ e $P(A \cup B)$.

Esercizio 3.24 Due squadre di atletica A e B partecipano alla fase finale di un torneo che consiste in una serie di 7 incontri diretti. La squadra vincitrice è quella che si aggiudica per prima 4 vittorie. Assumendo che le due squadre abbiano la stessa probabilità di vincere una qualunque partita e che i risultati siano indipendenti, calcolare:

1. La probabilità che il torneo si concluda in al più 6 partite.

2. La probabilità che il torneo si concluda esattamente dopo 6 incontri dato che la squadra A ha vinto le prime due partite.

Esercizio 3.25 Un'azienda alimentare propone un nuovo tipo di merenda al cioccolato e crema di zabaione, rivolta particolarmente ai ragazzi aventi un'età compresa fra 10 e 12 anni. L'area marketing ha successivamente rilevato che, a livello nazionale, il 30% degli interessati ha visto almeno uno degli spot televisivi predisposti per la campagna pubblicitaria che ha lanciato il prodotto. Considerato un gruppo di 150 potenziali consumatori, ci si chiede:

1. qual è il numero atteso che ha visto almeno uno spot televisivo?

2. qual è la probabilità che ve siano stati almeno 50 che ha visto almeno uno spot?

Esercizio 3.26 Siano X_1 e X_2 due v.a. indipendenti aventi distribuzione di Poisson di parametri rispettivamente $\lambda_1 = 3$ e $\lambda_2 = 1$, cioè $X_1 \sim \mathcal{P}(3)$ e $X_2 \sim \mathcal{P}(1)$. *i*) Calcolare la funzione di ripartizione di X_2 e rappresentarla graficamente; *ii*) considerata la v.a. $Y = 3X_1$ calcolare $\mathbb{E}[Y]$ e $P(Y \leq 6)$. Considerata successivamente la v.a. $W = 2X_1 - X_2$, calcolare: $\mathbb{E}[W]$ e Var(W).

Esercizio 3.27 Siano X e Y due v.a. aventi distribuzione di Poisson di parametro $\lambda = 9$.

i) Calcolare $P(X \geq 2)$ e $P(3X \leq 2)$

ii) Calcolare la covarianza di X e Y sapendo che il coefficiente di correlazione fra X e Y è uguale a $2/9$.

Esercizio 3.28 Una moneta viene tarata in maniera tale che, in una successione di 10 prove ripetute e indipendenti, la probabilità che l'evento T ('testa') si verifichi 4 volte sia uguale 0.2377 e che si verifichi 5 volte sia uguale a 0.1536. Calcolare la probabilità che nelle 10 prove l'evento T si verifichi una sola volta.

Esercizio 3.29 Uno studente, per ottenere informazioni relative ad assegnazioni di borse di studio, si rivolge all'ufficio competente il quale dà al 60% informazioni esatte. Sapendo ciò, lo studente va 3 volte a chiedere sempre la stessa informazione ed ottiene 2 volte la risposta A ed una volta la risposta contraria B. Calcolare la probabilità p che la risposta esatta sia la A. (Si supponga che A e B siano inizialmente equiprobabili e costituiscano una partizione dell'evento certo).

Esercizio 3.30 Il parco macchine di un noto autonoleggio dispone, fra le altre, di 7 auto di gran lusso: 3 *Mercedes Classe S* e 4 *Rolls Royce Silver Seraph*. In base a opportune analisi statistiche, si può assumere che la richesta quotidiana di noleggio di queste autovetture seguono approssimativamente una distribuzione di Poisson di parametro $\lambda_1 = 2$ per le Mercedes e $\lambda_2 = 3$ per le Rolls-Royce; si assuma inoltre che le richieste siano indipendenti. Supponendo che le macchine siano noleggiate per un periodo di una giornata, calcolare la probabilità che, in una giornata particolare: *i*) la compagnia non abbia abbastanza macchine disponibili per soddisfare la domanda di Mercedes; *ii*) la compagnia non abbia abbastanza macchine disponibili per soddisfare la domanda di auto di gran lusso.

Esercizio 3.31 E' noto che i prodotti di una volta erano soggetti a guasti. Si supponga che una macchina per cucire possa essere (con uguale probabilità) di 3 diversi tipi, a seconda dell'anno di produzione: tipo A (1960), tipo B (1970) e tipo C (1990). Il numero di guasti che una macchina per cucire subisce nell'arco di 5 anni segue la distribuzione di Poisson con valore medio diverso a seconda del tipo di macchina: $\lambda = 1$ (tipo A), $\lambda = 2$ (tipo B) e $\lambda = 3$ (tipo C). Sapendo che nei primi cinque anni di vita la macchina per cucire ha subito due guasti, determinare la probabilità che si tratti di una macchina di tipo A.

Esercizio 3.32 Siano X_1 e X_2 due v.a. indipendenti aventi distribuzione binomiale di parametri rispettivamente $X_1 \sim B(20, 0.4)$ e $X_2 \sim (5, 0.4)$.

i) Calcolare la funzione di densità di X_2 e rappresentarla graficamente;

ii) calcolare $P(0.2 \leq X_2 \leq 3.5)$;

iii) considerata la v.a. $Y = X_1 + X_2$ calcolare Var(Y) e $P(11.2 \leq Y < 12)$.

Esercizio 3.33 Una moneta è stata truccata in maniera tale che su 20 lanci l'evento "testa" su presenti 6 volte con probabilità 0.171 e 7 volte con probabilità 0.184. Qual è la probabilità che su 10 lanci l'evento "testa" non si verifichi mai?