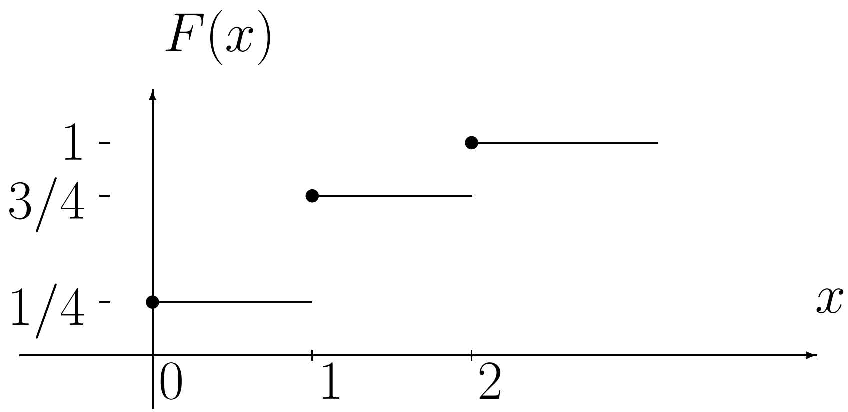
Se X è una variabile aleatoria discreta che assume i valori $x_1, x_2, x_3, \ldots$, dove $x_1 < x_2 < x_3 < \ldots$, allora la sua funzione di distribuzione $F(x)$ è costante negli intervalli $[x_{i-1}, x_i)$, ed in x_i ha un salto di ampiezza pari a $p(x_i)$. Quindi $F(x)$ può essere così espressa in termini della densità discreta:

$$F(x) = \sum_{k\colon x_k \le x} p(x_k) = \begin{cases} 0, & x < x_1 \\ p(x_1), & x_1 \le x < x_2 \\ p(x_1) + p(x_2), & x_2 \le x < x_3 \\ p(x_1) + p(x_2) + p(x_3), & x_3 \le x < x_4 \\ \vdots \end{cases}$$

Ad esempio, se X ha densità discreta $p(0) = \frac{1}{4}$, $p(1) = \frac{1}{2}$, $p(2) = \frac{1}{4}$, si ha



$$F(x) = \begin{cases} 0, & x < 0 \\ \dfrac{1}{4} & 0 \le x < 1 \\ \dfrac{3}{4} & 1 \le x < 2 \\ 1 & x \ge 2. \end{cases}$$

Esempio. In un gioco, in cui si lancia una moneta per 3 volte, si vincono k euro se esce testa per la prima volta al lancio k-esimo ($k = 1, 2, 3$), e si perdono c euro se non esce mai testa. Indicando con X la variabile aleatoria che descrive la vincita, ricavarne la funzione di distribuzione $F(x)$, e determinare il valore di c affinché il gioco sia equo.

Soluzione. Indicando con X la vincita al gioco, e posto $p(k) = P(X = k)$, risulta

$$p(1) = P(T_1) = \frac{1}{2}, \qquad p(2) = P(\overline{T_1} \cap T_2) = \frac{1}{4},$$

$$p(3) = P(\overline{T_1} \cap \overline{T_2} \cap T_3) = \frac{1}{8}, \qquad p(-c) = P(\overline{T_1} \cap \overline{T_2} \cap \overline{T_3}) = \frac{1}{8},$$

Pertanto: $F(x) = 0$ per $x < -c$, $F(x) = 1/8$ per $-c \leq x < 1$, $F(x) = 5/8$ per $1 \leq x < 2$, $F(x) = 7/8$ per $2 \leq x < 3$, $F(x) = 1$ per $x \geq 3$. La vincita media è

$$\sum_k k \cdot p(k) = -c \cdot \frac{1}{8} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} = \frac{11 - c}{8}.$$

Ne segue che il gioco è equo se $c = 11$.

Esercizio. Cosa cambia se si vincono k euro se esce testa k volte? ($k = 1, 2, 3$)

4.4 Valore atteso

Introduciamo uno dei più importanti concetti in calcolo delle probabilità.

Definizione. Se X è una variabile aleatoria discreta con densità discreta $p(x)$, il valore atteso (o valore medio, o speranza matematica) di X è definito da

$$E(X) = \sum_{x\colon p(x)>0} x \cdot p(x).$$

Il valore atteso di X è la media pesata di tutti i possibili valori che X può assumere, ognuno pesato con la probabilità che X lo assuma. Ad esempio, se $p(0) = p(1) = \frac{1}{2}$:

$$E(X) = 0 \cdot p(0) + 1 \cdot p(1) = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}.$$

Se, invece, risulta $p(0) = \frac{1}{3}$, $p(1) = \frac{2}{3}$, allora

$$E(X) = 0 \cdot p(0) + 1 \cdot p(1) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}.$$

Esempio. Calcolare il valore atteso nel lancio di un dado equilibrato.

Soluzione. Essendo $p(k) = \frac{1}{6}$ per $k = 1, 2, \ldots, 6$ otteniamo

$$E(X) = \sum_{k=1}^{6} k \cdot p(k) = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = \frac{7}{2} = 3{,}5.$$

Esempio. Dato un evento A, definiamo la funzione indicatrice I_A di A come

$$I_A(\omega) = \begin{cases} 1 & \text{se } \omega \in A \\ 0 & \text{altrimenti.} \end{cases}$$

Si calcoli $E(I_A)$.

Soluzione. Poiché $p(1) = P(I_A = 1) = P(A)$ e $p(0) = P(I_A = 0) = P(\overline{A})$, abbiamo

$$E(I_A) = 0 \cdot p(0) + 1 \cdot p(1) = p(1) = P(A).$$

Pertanto, il valore atteso della variabile indicatrice di un evento è uguale alla probabilità dell'evento stesso.

Esempio. Il concorrente di un gioco a quiz deve rispondere a due domande, D_1 e D_2, ed è libero di scegliere in che ordine rispondere. Se risponde per prima alla domanda D_i gli sarà consentito di rispondere all'altra domanda solo se avrà risposto correttamente alla prima. Egli riceve V_i euro se risponde correttamente alla domanda D_i, $i = 1, 2$. Se la probabilità di $E_i = \{\text{conosce la risposta di } D_i\}$ è P_i, con E_1 e E_2 indipendenti per ipotesi, a quale domanda dovrà rispondere prima per massimizzare il guadagno atteso?

Soluzione. Sia X_i la vincita del concorrente se risponde prima a D_i $(i = 1, 2)$; si ha

$$\begin{array}{ll} P(X_1 = 0) = 1 - P_1 & \qquad P(X_2 = 0) = 1 - P_2 \\ P(X_1 = V_1) = P_1\,(1 - P_2) & \qquad P(X_2 = V_2) = P_2\,(1 - P_1) \\ P(X_1 = V_1 + V_2) = P_1\,P_2 & \qquad P(X_2 = V_1 + V_2) = P_1\,P_2 \end{array}$$

con guadagno atteso $E[X_1] = V_1\,P_1\,(1 - P_2) + (V_1 + V_2)\,P_1\,P_2$ se risponde prima a D_1, mentre risulta $E[X_2] = V_2\,P_2\,(1 - P_1) + (V_1 + V_2)\,P_1\,P_2$ se risponde prima a D_2.
È più vantaggioso rispondere prima a D_1 se $V_1\,P_1\,(1 - P_2) \geq V_2\,P_2\,(1 - P_1)$, ovvero se

$$\frac{V_1\,P_1}{1 - P_1} \geq \frac{V_2\,P_2}{1 - P_2}.$$

Esempio. Una comitiva di 120 studenti viene condotta in gita in 3 autobus. Nel primo ci sono 36 studenti, nel secondo 40, e nel terzo 44. All'arrivo si sceglie a caso uno studente tra i 120. Se X denota il numero di studenti che hanno viaggiato sull'autobus dello studente scelto a caso, calcolare $E(X)$.

Soluzione. Si ha $P(X = 36) = \frac{36}{120}$, $P(X = 40) = \frac{40}{120}$, $P(X = 44) = \frac{44}{120}$ e quindi

$$E(X) = 36 \cdot \frac{36}{120} + 40 \cdot \frac{40}{120} + 44 \cdot \frac{44}{120} = \frac{4832}{120} = 40{,}2667.$$

Il numero medio di studenti presenti su un autobus è $120/3 = 40$, che quindi è minore di $E(X)$. Questo fenomeno si verifica perché: più studenti sono presenti su un singolo autobus e più sarà probabile che lo studente scelto a caso provenga proprio da quello. Così si assegna peso maggiore agli autobus con più studenti.

Più in generale, se vi sono k autobus con $n_1, n_2, \ldots, n_k$ studenti, si ha

$$E(X) = \sum_{i=1}^{k} n_i \cdot \frac{n_i}{\sum_{j=1}^{k} n_j} = \frac{\sum_{i=1}^{k} n_i^2}{\sum_{j=1}^{k} n_j} \quad \left(\geq \frac{\sum_{i=1}^{k} n_i}{k} \right)$$

Esempio. In seguito a trial clinici risulta che un farmaco sperimentale produce un miglioramento del 50% in pazienti gravi, un miglioramento del 5% in pazienti di media gravità e un peggioramento dell'1% in pazienti lievi. Si supponga che i pazienti affetti dalla patologia specifica siano per il 10% gravi, per il 15% medi e per il 75% lievi. Qual è il miglioramento medio prodotto dal farmaco sperimentale?

Soluzione. Descriviamo il miglioramento in percentuale prodotto dal farmaco sperimentale mediante una variabile aleatoria X discreta tale che

$$P(X = -1) = 0{,}75 \qquad P(X = 5) = 0{,}15 \qquad P(X = 50) = 0{,}10$$

Il miglioramento medio prodotto dal farmaco è del 5%, essendo

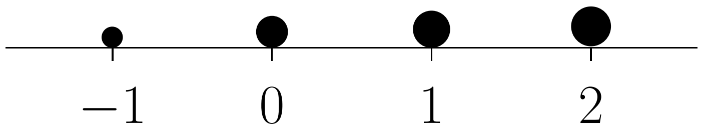
$$E(X) = \sum_x x \cdot P(X = x) = -1 \cdot 0{,}75 + 5 \cdot 0{,}15 + 50 \cdot 0{,}10 = 5.$$

Notiamo che se fosse stabilito che un farmaco si può porre in commercio se produce un miglioramento medio non minore del 5% in pazienti affetti da patologia specifica, quel farmaco sarebbe accettato sebbene causi un peggioramento nel 75% dei pazienti.

Il concetto di valore atteso è analogo al concetto fisico di *baricentro* di una distribuzione di masse. Sia X una variabile aleatoria discreta di densità discreta $p(x_i)$, $i \geq 1$. Immaginiamo una sbarra di peso trascurabile su cui sono poste delle masse

$$p(-1) = 0{,}10 \qquad p(0) = 0{,}25 \qquad p(1) = 0{,}30 \qquad p(2) = 0{,}35$$

nei punti $x_1 = -1$, $x_2 = 0$, $x_3 = 1$, $x_4 = 2$.



Allora il punto m nel quale la sbarra rimane in equilibrio è il centro di gravità, rispetto al quale è nulla la somma dei momenti dei pesi delle singole porzioni di massa, ossia

$$\sum_i [x_i - m]\, p(x_i) = 0 \qquad \Longrightarrow \qquad m = \frac{\sum_i x_i\, p(x_i)}{\sum_i p(x_i)} = \sum_i x_i\, p(x_i) = E(X).$$

Nell'esempio, $E(X) = \sum_{x=-1}^{2} x \cdot p(x) = -1 \cdot 0{,}10 + 0 \cdot 0{,}25 + 1 \cdot 0{,}30 + 2 \cdot 0{,}35 = 0{,}9.$

Esempio. Da un'urna contenente k biglie nere e $n-k$ biglie bianche si effettuano n estrazioni a caso senza reinserimento. Determinare $E(X)$, dove X denota l'estrazione in cui si estrae una biglia nera per la prima volta.

Soluzione. Poiché le estrazioni sono senza reinserimento l'evento $\{X=j\}$ si verifica se le prime $j-1$ biglie estratte sono bianche, la j-esima è nera, e nelle rimanenti $n-j$ estrazioni vi sono $k-1$ biglie nere. Quindi risulta

$$p(j)=P(X=j)=\frac{\binom{j-1}{0}\binom{1}{1}\binom{n-j}{k-1}}{\binom{n}{k}}=\frac{\binom{n-j}{k-1}}{\binom{n}{k}}, \qquad j=1,2,\ldots,n-k+1.$$

Notiamo che

$$\sum_j p(j)=\frac{1}{\binom{n}{k}}\sum_{j=1}^{n-k+1}\binom{n-j}{k-1}=\frac{1}{\binom{n}{k}}\sum_{r=k-1}^{n-1}\binom{r}{k-1}=\frac{1}{\binom{n}{k}}\binom{n}{k}=1,$$

avendo posto $r=n-j$ e fatto uso dell'identità $\sum_{r=k}^{n}\binom{r}{k}=\binom{n+1}{k+1}$.

Si ha dunque

$$E(X) \;=\; \sum_j j\,p(j) = \frac{1}{\binom{n}{k}} \sum_{j=1}^{n-k+1} j\binom{n-j}{k-1} = \frac{1}{\binom{n}{k}} \sum_{j=1}^{n-k+1} \sum_{h=1}^{j} \binom{n-j}{k-1}$$

$$= \frac{1}{\binom{n}{k}} \sum_{h=1}^{n-k+1} \sum_{j=h}^{n-k+1} \binom{n-j}{k-1} = \frac{1}{\binom{n}{k}} \sum_{h=1}^{n-k+1} \sum_{r=k-1}^{n-h} \binom{r}{k-1},$$

avendo posto $r = n - j$. Facendo uso dell'identità $\sum_{r=k}^{n} \binom{r}{k} = \binom{n+1}{k+1}$ segue

$$E(X) = \frac{1}{\binom{n}{k}} \sum_{h=1}^{n-k+1} \binom{n-h+1}{k} = \frac{1}{\binom{n}{k}} \sum_{r=k}^{n} \binom{r}{k} = \frac{1}{\binom{n}{k}} \binom{n+1}{k+1},$$

avendo posto $r = n-h+1$ e usato nuovamente l'identità $\sum_{r=k}^{n} \binom{r}{k} = \binom{n+1}{k+1}$. Pertanto,

$$E(X) = \frac{n+1}{k+1}.$$

Nel caso particolare in cui $k = 1$, risulta

$$p(j) = P(X = j) = \frac{\binom{n-j}{0}}{\binom{n}{1}} = \frac{1}{n}, \qquad j = 1, 2, \ldots, n.$$

In tal caso X si dice avere distribuzione uniforme discreta, e risulta

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

Notiamo che il problema considerato nell'esempio precedente può essere anche interpretato come il problema della ricerca sequenziale, in cui è data una lista di n elementi, k dei quali sono del tipo da individuare e sono distribuiti a caso. L'algoritmo si basa sulla scansione sequenziale della lista, fermandosi in corrispondenza del primo dei k elementi da individuare. La variabile aleatoria X descrive il numero di confronti necessari per individuare tale elemento, e quindi dà una misura della complessità temporale dell'algoritmo. Pertanto la complessità nel caso medio è data da

$$E(X) = \frac{n+1}{k+1}.$$

Esempio. Da un'urna contenente k biglie nere e $n-k$ biglie bianche si effettuano n estrazioni a caso con reinserimento. Determinare $E(Y)$, dove Y denota l'estrazione in cui si estrae una biglia nera per la prima volta.

Soluzione. Poiché le estrazioni sono con reinserimento, per l'indipendenza risulta

$$p(j) = P(Y = j) = p(1-p)^{j-1}, \qquad j = 1, 2, \dots,$$

dove $p = k/n$. Si ha pertanto

$$E(Y) = p\sum_{j=1}^{\infty} j(1-p)^{j-1} = p\sum_{j=1}^{\infty}\sum_{k=1}^{j}(1-p)^{j-1} = p\sum_{k=1}^{\infty}\sum_{j=k}^{\infty}(1-p)^{j-1}.$$

Ponendo $n = j - k$ si ha

$$E(Y) = p\sum_{k=1}^{\infty}(1-p)^{k-1}\sum_{n=0}^{\infty}(1-p)^{n} = p\,\frac{1}{p}\,\frac{1}{p} = \frac{1}{p} = \frac{n}{k}.$$

Notiamo che, per $k \le n$, risulta $E(X) = \dfrac{n+1}{k+1} \le \dfrac{n}{k} = E(Y)$.

4.5 Valore atteso di una funzione di una variabile aleatoria

Sia X una variabile aleatoria discreta di cui è nota la sua densità discreta $p(x_i)$.
Si desidera calcolare il valore atteso di una qualche funzione di X, diciamo $g(X)$.

Essendo $g(X)$ stessa una variabile aleatoria discreta, avrà una densità discreta che possiamo determinare conoscendo quella di X. Ricavata la densità discreta di $g(X)$ possiamo calcolare $E[g(X)]$ utilizzando la definizione di valore atteso.

Esempio. Sia X una variabile aleatoria che assuma i valori -1, 0, 1 con probabilità $P(X = -1) = 0{,}2 \quad P(X = 0) = 0{,}5 \quad P(X = 1) = 0{,}3$. Calcolare $E(X^2)$.

Soluzione. Sia $Y = X^2$; è immediato verificare che la densità discreta di Y è

$$P(Y = 1) = P(X^2 = 1) = P(X = \pm 1) = P(X = 1) + P(X = -1) = 0{,}5$$

$$P(Y = 0) = P(X^2 = 0) = P(X = 0) = 0{,}5$$

Quindi $E(X^2) = E(Y) = \sum_{y=0}^{1} y \cdot P(Y = y) = 0 \cdot 0{,}5 + 1 \cdot 0{,}5 = 0{,}5$.
Si noti che $0{,}5 = E(X^2) \neq [E(X)]^2 = [-1 \cdot 0{,}2 + 0 \cdot 0{,}5 + 1 \cdot 0{,}3]^2 = [0{,}1]^2 = 0{,}01$.

Se si considera che $g(X) = g(x)$ quando $X = x$, appare ragionevole che $E[g(X)]$ sia la media pesata dei valori $g(x)$, assegnando $P(X = x)$ come peso a $g(x)$.

Proposizione. Se X è una variabile aleatoria discreta che assume i valori x_i, $i \geq 1$, con probabilità $p(x_i)$, allora per ogni funzione a valori reali $g(x)$ risulta

$$E[g(X)] = \sum_i g(x_i)\, p(x_i).$$

Dimostrazione. Denotiamo con y_j, $j \geq 1$, i diversi valori di $g(x_i)$, $i \geq 1$. Allora, raggruppando tutti gli x_i che hanno stesso valore, abbiamo

$$\begin{aligned}\sum_i g(x_i)\, p(x_i) &= \sum_j \sum_{i:\, g(x_i)=y_j} g(x_i)\, p(x_i) = \sum_j y_j \sum_{i:\, g(x_i)=y_j} p(x_i)\\ &= \sum_j y_j\, P[g(X) = y_j] = E[g(X)],\end{aligned}$$

avendo fatto uso dell'identità $\boxed{P(X \in B) = \sum_{k:\, x_k \in B} p(x_k)}$ per $B = \{x_i : g(x_i) = y_j\}$.

Esempio. Sia X una variabile aleatoria che assume i valori -1, 0, 1 con probabilità $P(X=-1)=0{,}2 \quad P(X=0)=0{,}5 \quad P(X=1)=0{,}3$. Calcolare $E(X^2)$ facendo uso di $E[g(X)] = \sum_i g(x_i)\, p(x_i)$.

Soluzione. Otteniamo $E(X^2) = \displaystyle\sum_i x_i^2\, p(x_i) = (-1)^2 \cdot 0{,}2 + 0^2 \cdot 0{,}5 + 1^2 \cdot 0{,}3 = 0{,}5.$

Esempio. Sia X una variabile aleatoria che descrive il risultato del lancio di un dado non truccato. Calcolare $E[g(X)]$, con $g(x) = \min\{x, 4\}$.

Soluzione. Si ha

$$E[g(X)] = \sum_i g(x_i)\, p(x_i) = \sum_{k=1}^{6} \min\{k, 4\}\, p(k) = \frac{1}{6}(1+2+3+4+4+4) = 3.$$

Esempio. Sia X la variabile aleatoria tale che $P(X=0) = 1-p$ e $P(X=1) = p$, $0 \le p \le 1$. Calcolare il minimo di $\psi(a) := E[(X-a)^2]$, per $a \in \mathbb{R}$.

Soluzione. Risulta $\psi(a) = E[(X-a)^2] = \sum_{k=0}^{1} (k-a)^2\, p(k) = a^2\,(1-p) + (1-a)^2\, p$ e quindi $\psi'(a) = 2a(1-p) - 2(1-a)p = 0$ per $a = p$. Quindi il minimo di $\psi(a)$ è p.

Proposizione. (Proprietà di linearità) Sia X una variabile aleatoria discreta. Se a e b sono costanti reali, allora

$$E(a\,X + b) = a\,E(X) + b.$$

Dimostrazione. Ricordando che $E[g(X)] = \sum_i g(x_i)\,p(x_i)$, per $g(x) = a\,x + b$ è

$$E[g(X)] = \sum_i (a\,x_i + b)\,p(x_i) = a \sum_i x_i\,p(x_i) + b \sum_i p(x_i) = a\,E(X) + b,$$

avendo fatto uso dell'identità $\sum_i p(x_i) = 1$.

Definizione. Sia X una variabile aleatoria discreta; la quantità $E(X^n)$, $n \geq 1$, è detta *momento di ordine n* di X. Risulta

$$E(X^n) = \sum_{x:\,p(x)>0} x^n\,p(x).$$

Segue che il valore atteso di X è anche il momento di ordine 1 di X.

Esempio. Sia X una variabile aleatoria che assuma i valori -1, 0, 1 con probabilità $P(X=-1)=0{,}2 \quad P(X=0)=0{,}5 \quad P(X=1)=0{,}3$. Calcolare $E(X^n)$.

Soluzione. Il momento di ordine n di X è dato da

$$E(X^n)=\sum_{x=-1}^{1} x^n\,P(X=x)=(-1)^n\,0{,}2+0^n\,0{,}5+1^n\,0{,}3=(-1)^n\,0{,}2+0{,}3$$

quindi $E(X^n)=0{,}5$ per n pari e $E(X^n)=0{,}1$ per n dispari.

Esempio. Determinare il momento di ordine n della variabile aleatoria $Y=aX+b$, con a e b costanti reali.

Soluzione. Facendo uso del teorema del binomio si ha

$$E[Y^n]=E[(aX+b)^n]=E\left[\sum_{k=0}^{n}\binom{n}{k}a^kX^kb^{n-k}\right]$$

e pertanto, per la proprietà di linearità,

$$E[Y^n]=\sum_{k=0}^{n}\binom{n}{k}a^kb^{n-k}E[X^k].$$

4.6 Varianza

Sebbene il valore atteso fornisca una media pesata dei possibili valori di una variabile aleatoria, esso non dà alcuna informazione riguardo alla variabilità, o dispersione, di questi valori. Per esempio, le variabili aleatorie date da

$$W = 0 \ \text{ con prob. } 1, \qquad Y = \begin{cases} -1 \text{ con prob. } 1/2 \\ +1 \text{ con prob. } 1/2, \end{cases} \qquad Z = \begin{cases} -100 \text{ con prob. } 1/2 \\ +100 \text{ con prob. } 1/2 \end{cases}$$

hanno lo stesso valore atteso, pari a 0, ma è presente maggiore dispersione nei valori di Y piuttosto che in quelli di W e nei valori di Z rispetto a Y.

Poiché ci si attende che X assuma valori disposti attorno al suo valore atteso $E(X)$, appare ragionevole misurare la variabilità di X mediante la media della distanza dal valor medio che i possibili valori di X assumono, ad esempio considerando la quantità $E(|X-\mu|)$, dove $\mu = E(X)$. Per superare alcune difficoltà di tipo matematico si preferisce invece adoperare la media delle differenze al quadrato tra X ed il suo valore atteso μ, e ciò conduce alla seguente definizione.

Definizione. Sia X una variabile aleatoria discreta di valore atteso $E(X) = \mu$; la varianza di X, denotata con $\mathrm{Var}(X)$, è così definita:

$$\mathrm{Var}(X) = E[(X-\mu)^2] = \sum_{x:\,p(x)>0} (x-\mu)^2\, p(x).$$

Una formula alternativa per la varianza si ottiene usando la proprietà di linearità:

$$\begin{aligned}\mathrm{Var}(X) &= E[(X-\mu)^2] = E(X^2 - 2\mu\, X + \mu^2) = E(X^2) - 2\mu\, E(X) + \mu^2 \\ &= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2.\end{aligned}$$

La maniera più semplice di valutare la varianza di X consiste quindi nel calcolare la differenza tra il momento del secondo ordine di X e il quadrato del suo valore atteso.

Esempio. Calcolare $\mathrm{Var}(X)$, dove X rappresenta l'esito del lancio di un dado.
Soluzione. Essendo $p(k) = \frac{1}{6}$ per $k = 1, 2, \ldots, 6$ e $E(X) = \frac{7}{2}$ otteniamo

$$\mathrm{Var}(X) = E(X^2) - \mu^2 = \sum_{k=1}^{6} k^2 \cdot p(k) - \left(\frac{7}{2}\right)^2 = \frac{1}{6} \cdot 91 - \frac{49}{4} = \frac{35}{12} \approx 2{,}9167.$$