4 Le variabili casuali

4.1 Variabili casuali discrete e continue

Nei problemi di calcolo delle probabilità si è spesso condotti a considerare delle quantità che sono funzioni del risultato di un certo fenomeno casuale.
La nozione di variabile casuale o aleatoria viene usata per descrivere gli eventi e misura i risultati di una prova, i quali possono assumere differenti valori anche quando le prove vengono eseguite sotto le stesse condizioni.
Sono esempi di variabili aleatorie: il valore numerico degli esiti del lancio di un dado, la durata della vita di una lampadina, il modulo della velocità di una molecola di un gas. Nel primo caso la variabile casuale è *discreta*, cioè assume un insieme finito o numerabile di valori; negli altri due casi è *continua*, cioè è definita su un intervallo dell'asse dei numeri reali.
Definizione: Dato $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità, definiamo variabile casuale o aleatoria un'appli-

cazione

$$X : \Omega \to \mathbb{R}$$

tale che $\forall r \in \mathbb{R}$

$$A_r = \{w \in \Omega : X(w) \leq r\} \in \mathcal{A},$$

cioè A_r è un evento.
Una variabile casuale X è perciò una funzione di $w \in \Omega$ per la quale ha senso calcolare la probabilità:

$$P[\{w \in \Omega : X(w) \leq r\}] \quad \forall r \in \mathbb{R},$$

cioè la probabilità che X assuma valori più piccoli o uguali ad r.

$$A_r = \{w \in \Omega : X(w) \leq r\} \stackrel{\text{def}}{=} (X \leq r).$$

La variabile casuale X fa corrispondere un numero reale ad ogni esito dell'esperimento.

Esempio: Lancio di due dadi.

$$\Omega = \{(i, j) = w : i = 1, ..., 6; j = 1, ..., 6\}$$

Consideriamo la seguente variabile casuale:

- $X(w = (i, j)) = i + j$
  X assegna ad ogni lancio la somma dei punteggi

dei dadi. Quindi X è una variabile casuale che assume i valori $\{2; 3; 4; ...; 12\}$.

4.2 Funzione di ripartizione e funzione di densità

Definizione: Dato $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità, definiamo funzione di ripartizione di una variabile casuale X la funzione:

$$F : \mathbb{R} \to [0, 1]$$

tale che

$$\forall x \in \mathbb{R}\,, \;\; F(x) = P[X \leq x] = P[\{w \in \Omega : X(w) \leq x\}]$$

Esempio: Lancio di due dadi.
$X = i + j \in [2, 12]$.

$$\text{se}\, x < 2 \; (X \leq x) = \varnothing \quad P[X \leq x] = 0 \qquad F(x) = 0$$

$$\text{se}\, 2 \leq x < 3 \; (X \leq x) = \{(1, 1)\} \qquad F(x) = \frac{1}{36}$$

$$\text{se}\, 3 \leq x < 4 \; (X \leq x) = \{(1, 1); (1, 2); (2, 1)\} \; F(x) = \frac{3}{36}$$

$$\vdots$$

$$\text{se}\, 11 \leq x < 12 \; (X \leq x) = \Omega \setminus \{(6, 6)\} \quad F(x) = \frac{35}{36}$$

$$\text{se}\, x \geq 12 \; (X \leq x) = \Omega \qquad F(x) = 1$$

È utile ricorrere alla forma grafica della funzione di ripartizione per verificare che essa sia una funzione monotona non decrescente definita non solo per tutti i valori assunti dalla variabile casuale, ma per l'intero asse dei numeri reali.

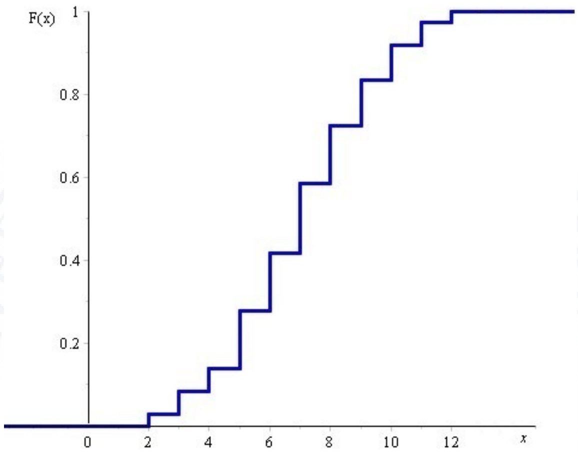

Figura 13: Funzione di ripartizione

Proprietà della funzione di ripartizione F

1. $\lim_{x \to -\infty} F(x) = 0, \quad \lim_{x \to +\infty} F(x) = 1$

2. F è monotona non decrescente:
   se $x_1 < x_2 \to F(x_1) \leq F(x_2) \quad \forall x_1, x_2 \in \mathbb{R}$

3. $P[X > x] = 1 - P[X \leq x] = 1 - F(x)$

4. F è continua a destra:

$$\lim_{h \to 0^+} F(x+h) = F(x)\,, \ \forall x \in \mathbb{R}$$

5. $P[x_1 < X \leq x_2] = F(x_2) - F(x_1)$,
   $\forall x_1, x_2 \in \mathbb{R}, \quad x_1 < x_2$

Nel definire la funzione di ripartizione F non abbiamo fatto distinzione tra variabile casuale discreta e variabile casuale continua. Tale distinzione invece è necessaria quando introduciamo un'altra funzione importante chiamata funzione di densità della variabile casuale X.

Definizione: Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità e X una variabile casuale discreta. Definiamo funzione di densità discreta (o funzione di massa o distribuzione di probabilità) la funzione:

$$f : \mathbb{R} \rightarrow [0, 1]$$

tale che

$$f(x) = P[X = x] = P[\{w \in \Omega : X(w) = x\}], \ \forall x \in \mathbb{R},$$

dove x è detto anche *punto massa*.

Proprietà della funzione di densità discreta f

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$

Osservazione: Assegnata la funzione di ripartizione F di una variabile casuale X è possibile calcolare la funzione di densità f e viceversa.

Esempi

1) lancio di un dado.

X = numero della faccia in alto.

$f(x) = \frac{1}{6}\,, \quad x = x_k\,, \quad k = 1, ..., 6.$

$$
\begin{array}{lll}
\text{se}\, x < 1 & P[X \leq x] = P[\varnothing] = 0 & F(x) = 0 \\
\text{se}\, 1 \leq x < 2 & P[X \leq x] = \sum_{t \leq x} f(t) = f(1) & F(x) = \frac{1}{6} \\
\text{se}\, 2 \leq x < 3 & P[X \leq x] = f(1) + f(2) & F(x) = \frac{2}{6} \\
\text{se}\, 3 \leq x < 4 & P[X \leq x] = f(1) + f(2) + f(3) & F(x) = \frac{3}{6} \\
\text{se}\, 4 \leq x < 5 & P[X \leq x] = f(1) + ... + f(4) & F(x) = \frac{4}{6} \\
\text{se}\, 5 \leq x < 6 & P[X \leq x] = 1 - f(6) & F(x) = \frac{5}{6} \\
\text{se}\, x \geq 6 & P[X \leq x] = P[\Omega] = 1 & F(x) = 1
\end{array}
$$

La funzione di ripartizione è costante tra un punto massa e il successivo ed ha un salto costante ($\sim$ funzione di densità) nei punti massa.

Scelto $x = 2.5$, quanto vale $F(2.5)$?
$F(2.5) = \sum_{t \leq 2.5} f(t) = f(1) + f(2) = \frac{1}{3}$

Scelto $x = 3$, quanto vale $f(3)$?
$f(3) = F(3) - \lim_{h \to 0^+} F(3-h) = \frac{3-2}{6} = \frac{1}{6}$

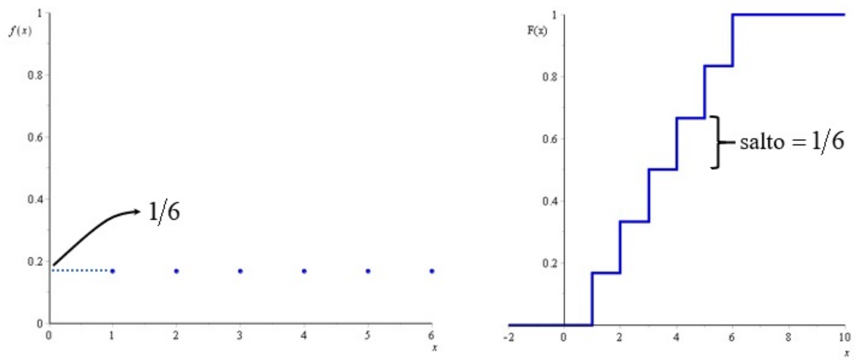

Figura 14: Funzione di densità e di ripartizione

2) Lancio di due dadi.
Abbiamo già costruito la funzione di ripartizione, determiniamo ora la funzione di densità.

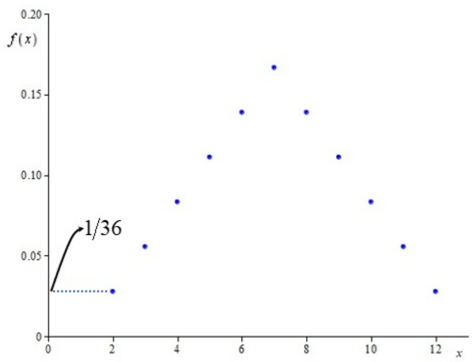

Figura 15: Funzione di densità

se $x = 2$ $P[X = 2] = P[X = \{(1,1)\}] = \dfrac{1}{36}$

se $x = 3$ $P[X = 3] = P[X = \{(1,2);(2,1)\}] = \dfrac{2}{36}$

$\vdots$

$\vdots$

se $x = 11$ $P[X = 11] = P[X = \{(5,6);(6,5)\}] = \dfrac{2}{36}$

se $x = 12$ $P[X = 12] = P[X = \{(6,6)\}] = \dfrac{1}{36}$

Per una variabile casuale X continua non ha senso calcolare la probabilità che essa sia uguale esattamente ad uno dei valori che può assumere, mentre ha senso calcolare la probabilità che la variabile casuale cada all'interno di un determinato intervallo di valori ammissibili.

<u>Definizione</u>: Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità. Una variabile casuale X è detta continua se esiste una funzione

$$f : \mathbb{R} \to [0, +\infty)$$

tale che:

$$\forall x \in \mathbb{R}\,,\; F(x) = P[X \leq x] = \int_{-\infty}^{x} f(t)dt.$$

La funzione f è detta funzione di densità di probabilità continua.

Proprietà della funzione di densità continua f

1. $f(x) \geq 0\,,\; \forall x \in \mathbb{R}$ in cui è definita $F'(x)$.
2. $\int_{-\infty}^{+\infty} f(x)dx = 1$
3. $\forall a, b \in \mathbb{R}, a < b$

$$P[a \leq X \leq b] = \int_{a}^{b} f(x)dx = F(b) - F(a)$$

Esempio: Durata di una conversazione telefonica.

La durata di una conversazione telefonica può essere descritta da una variabile casuale X avente la seguente funzione di ripartizione:

$$F(x) = \begin{cases} 0 & \text{se } x < 0\,, \\ 1 - e^{-\lambda x} & \text{se } x \geq 0, \quad \lambda > 0\,. \end{cases}$$

allora
$f(x) = F'(x) = \lambda e^{-\lambda x}, \quad x \geq 0$, mentre $f(x) = 0$ per $x < 0$, quindi

$$f(x) = \begin{cases} 0 & \text{se } x < 0\,, \\ \lambda e^{-\lambda x} & \text{se } x \geq 0\,. \end{cases}$$

Se la durata della conversazione è espressa in minuti, calcolare la probabilità che la conversazione duri tra i 5 ed i 10 minuti.

$$\begin{aligned} P[5 < X < 10] &= \int_5^{10} \lambda e^{-\lambda x} dx = -e^{-\lambda x}\Big|_5^{10} \\ &= e^{-5\lambda} - e^{-10\lambda} \end{aligned}$$

oppure

$$\begin{aligned} P[5 < X < 10] &= F(10) - F(5) = \\ &= \left(1 - e^{-\lambda x}\right)\Big|_{x=10} - \left(1 - e^{-\lambda x}\right)\Big|_{x=5} \\ &= -e^{-10\lambda} + e^{-5\lambda} \end{aligned}$$

Una variabile casuale continua X con funzione di densità e di ripartizione sopra definite viene detta *esponenziale*.

4.3 Il valore atteso

Un importante indicatore numerico che misura la tendenza centrale di una variabile casuale X è la *media* o *valore atteso* o *speranza matematica.*
Il valore atteso viene calcolato utilizzando la distribuzione di probabilità.
Definizione: Sia X una variabile casuale con distribuzione di probabilità $f(x)$. Chiamiamo valore atteso o media di X il numero μ_X o $E[X]$ definito come:

- $\mu_X = E[X] = \sum_i x_i f(x_i)$ se X è variabile casuale discreta,
- $\mu_X = E[X] = \int_{-\infty}^{+\infty} x f(x) dx$ se X è variabile casuale continua.

La media, indicando dove sono "centrati" i valori di X è anche detta *centro di gravità* o *baricentro* di una densità.

Osservazione
Per una variabile casuale discreta, $E[X]$ rappresenta la media pesata di tutti i possibili valori che X può assumere, ognuno pesato con la probabilità che lo

assuma (ricorda che $f(x_i) = P[X = x_i]$).

Esempi

1) Lancio di due dadi.

$X = i + j$, $i = 1, ..., 6$; $j = 1, ..., 6$. $X \in [2, 12]$

$$E[X] = 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + \ldots + 11 \cdot \frac{2}{36} + 12 \cdot \frac{1}{36} = 7$$

(si veda anche il grafico di $f(x)$)

2) Durata di una conversazione telefonica.

X è una variabile casuale continua avente funzione di densità:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{se } x \geq 0, \quad \lambda > 0\,, \\ 0 & \text{se } x < 0\,. \end{cases}$$

$$E[X] = \int_0^{+\infty} x \lambda e^{-\lambda x} dx$$

Integrando per parti

$u(x) = x$, $v'(x) = -\lambda e^{-\lambda x} \to v = e^{-\lambda x}$

$$
\begin{aligned}
E[X] &= -xe^{-\lambda x}\Big|_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x} dx \\
&= -\lim_{x\to+\infty} \frac{x}{e^{\lambda x}} - \frac{1}{\lambda} e^{-\lambda x}\Big|_0^{+\infty} \\
&= 0 + \frac{1}{\lambda} = \frac{1}{\lambda}
\end{aligned}
$$

dove nel limite abbiamo utilizzato la regola di De l'Hôspital.

4.4 La varianza

Un altro indicatore numerico che è in grado di caratterizzare la variabilità della variabile casuale X attorno alla sua media è la *varianza*.

Definizione: Sia X una variabile casuale con funzione di densità $f(x)$ e media μ_X. Definiamo varianza di X il numero reale positivo $\sigma_X^2 = \text{var}[X]$:

- $\sigma_X^2 = \text{var}[X] = \sum_i (x_i - \mu_X)^2 f(x_i)$ se X è una variabile casuale discreta,
- $\sigma_X^2 = \text{var}[X] = \int_{-\infty}^{+\infty} (x - \mu_X)^2 f(x) dx$ se X è una variabile casuale continua.

Il numero reale positivo:

$$\sigma_X = \sqrt{\text{var}[X]}$$

è detto deviazione standard o scarto quadratico medio di X.

Esempi

1) Lancio di due dadi.
$X = i + j, \qquad X \in [2, 12], \qquad E[X] = 7$

$$\text{var}[X] = (2-7)^2 \cdot \frac{1}{36} + (3-7)^2 \cdot \frac{2}{36} + \ldots (12-7)^2 \cdot \frac{1}{36} = \frac{35}{6}$$

2) Durata di una conversazione telefonica.
X è una variabile casuale continua con funzione di densità:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{se } x \geq 0, \quad \lambda > 0\,, \\ 0 & \text{se } x < 0\,. \end{cases}$$

$E[X] = \dfrac{1}{\lambda}\,.$

Dalla definizione si ha:

$$\mathrm{var}[X] = \int_0^{+\infty} \left(x - \frac{1}{\lambda}\right)^2 \lambda e^{-\lambda x} dx =$$

$$= \underbrace{\int_0^{+\infty} x^2 \lambda e^{-\lambda x} dx}_{1} \underbrace{-2 \int_0^{+\infty} x e^{-\lambda x} dx}_{2} + \underbrace{\frac{1}{\lambda} \int_0^{+\infty} e^{-\lambda x} dx}_{3}$$

Integrando per parti l'integrale 1 con $u(x) = x^2$, $v'(x) = -\lambda e^{-\lambda x}$

$$1 = -x^2 e^{-\lambda x}\Big|_0^{+\infty} + \int_0^{+\infty} 2x e^{-\lambda x} dx$$

e quindi

$$1 + 2 = -x^2 e^{-\lambda x}\Big|_0^{+\infty} = -\lim_{x \to +\infty} \frac{x^2}{e^{\lambda x}} = 0$$

avendo applicato due volte la regola di De l'Hôspital. Rimane l'integrale 3:

$$\mathrm{var}[X] = \frac{1}{\lambda} \int_0^{+\infty} e^{-\lambda x} dx = -\frac{1}{\lambda^2} e^{-\lambda x}\Big|_0^{+\infty} = \frac{1}{\lambda^2}.$$

4.5 Valore atteso di una funzione di variabile casuale

Il concetto di valore atteso viene esteso in maniera naturale anche ad una funzione di variabile casuale. Se X è una variabile casuale e $g : \mathbb{R} \to \mathbb{R}$ una funzione reale di variabile reale, allora $Y = g(X)$ è una variabile casuale.
Definizione; Chiamiamo valore atteso della variabile casuale $g(X)$ la quantità:

- $E[g(X)] = \sum_i g(x_i) f(x_i)$ se X è variabile casuale discreta,
- $E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx$ se X è variabile casuale continua.

Esempi
1) Sia X la variabile casuale che rappresenta il numero di automobili lavate in un autolavaggio tra le 16 e le 17 di un dato giorno. Supponiamo che X abbia la seguente distribuzione di probabilità:

x	4	5	6	7	8	9
$f(x) = P[X=x]$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Figura 16

Sia $g(X) = 2X - 1$ la variabile casuale che rappresenta la quantità, in euro, pagata all'addetto dal gestore. Calcolare il guadagno atteso dall'addetto tra le 16 e le 17.

$$E[g(X)] = E[2X-1] = \sum_{i=4}^{9}(2x_i - 1)f(x_i) =$$

$$= 7\frac{1}{12} + 9\frac{1}{12} + 11\frac{1}{4} + 13\frac{1}{4} + 15\frac{1}{6} + 17\frac{1}{6} = 12.67$$

2) Sia X la variabile casuale (continua) che rappresenta il tempo impiegato per trovare un guasto in un impianto elettrico (misurato in ore).
Supponiamo che X abbia la seguente funzione di densità:

$$f(x) = \begin{cases} 1 & \text{se } 0 < x < 1\,, \\ 0 & \text{altrove}\,. \end{cases}$$

Causa il guasto, l'interruzione di x ore dell'impianto provoca un danno economico $g(X) = X^3$.
Calcolare il valore atteso del danno.

$$E[g(X)] = \int_{-\infty}^{\infty} x^3 f(x)dx = \int_0^1 x^3 dx = \frac{1}{4} = 0.25$$

Proprietà del valore atteso

Elenchiamo ora le proprietà del valore atteso di una funzione di variabile casuale, che possono quindi essere espresse sia per variabili casuali discrete che continue.

1. $E[c] = c$ con c costante $\in \mathbb{R}$.
2. $E[cg(X)] = cE[g(X)]$ con c costante $\in \mathbb{R}$.
3. $E[c_1 g_1(X) + c_2 g_2(X)] = c_1 E[g_1(X)] + c_2 E[g_2(X)]$ con c_1, c_2 costanti $\in \mathbb{R}$.
4. se $g_1(X) \leq g_2(X)\ \forall X \in \mathbb{R}$
   $E[g_1(X)] \leq E[g_2(X)]$
5. Disuguaglianza di Markov
   Data una funzione $g : \mathbb{R} \to \mathbb{R}$ tale che $g(X) \geq 0$

$\forall x \in \mathbb{R}$ per cui esista $E[g(X)]$, allora:

$$P[g(X) \geq r] \leq \frac{1}{r} E[g(X)], \quad \forall r > 0$$

6. Disuguaglianza di Chebyshev
   Data una variabile casuale X con media μ_X e varianza σ_X^2 si ha:

$$P[|X - \mu_X| \geq k\sigma_X] \leq \frac{1}{k^2}, \; \forall k > 0$$

   o equivalentemente

$$P[|X - \mu_X| < k\sigma_X] =$$
$$= P[\mu_X - k\sigma_X < X < \mu_X + k\sigma_X] \geq 1 - \frac{1}{k^2}$$

La disuguaglianza di Chebyshev si ottiene dalla disuguaglianza di Markov tramite un'opportuna scelta di $g(X)$ e di r. Più precisamente $g(X) = (X - \mu_X)^2$ ed $r = k^2\sigma_X^2$ (> 0).

Essa afferma che più la varianza della variabile casuale X è piccola, più è piccola la probabilità che X assuma valori lontani dalla sua media.

La disuguaglianza di Chebyshev afferma che la probabilità che una qualsiasi variabile casuale X assuma un valore all'interno di k deviazioni standard dalla media è almeno $1 - \frac{1}{k^2}$, ad esempio:

- per $k = 2$, X ha una probabilità di almeno $3/4$ di cadere all'interno di 2 deviazioni standard dalla media μ $(\mu \pm 2\sigma)$.
- per $k = 3$, X ha una probabilità di almeno $8/9$ di cadere all'interno di 3 deviazioni standard dalla media μ $(\mu \pm 3\sigma)$.

Il valore determinato con la disuguaglianza di Chebyshev è riferito solo al limite inferiore, cioè si sa che la probabilità di una variabile casuale X di cadere all'interno di 2 deviazioni standard è almeno $3/4$, ma non si sa di quanto potrebbe essere maggiore di $3/4$. Questo accade perchè non è nota la funzione di densità $f(x)$ della variabile casuale X. Perciò l'utilizzo della disuguaglianza di Chebyshev è limitato ai casi in cui non è nota la forma della distribuzione della variabile casuale.
Esempio: Sia X la variabile casuale corrispondente al numero di pezzi prodotti da una fabbrica in una settimana. Si sa che $\mu_X = 50$.

- Calcolare la $P[X \geq 75]$.
- Sapendo che $\sigma_X^2 = 25$, dare un limite inferiore alla probabilità $P[40 < X < 60]$.

La funzione di densità $f(x)$ non è nota, posso però applicare Chebyshev e Markov.

- $P[X \geq 75] \leq \frac{1}{75}E[X] = \frac{1}{75}50 = \frac{2}{3}$.
- $P[40 < X < 60] =$
  $= P[40 - 50 < X - 50 < 60 - 50] =$
  $= P[|X - 50| < 10] = 1 - P[|X - 50| \geq 10]$
  ma
  $P[|X - 50| \geq 10] = P[(X - 50)^2 \geq 100]$,
  quindi, applicando Chebyshev
  $1 - P[40 < X < 60] \leq \frac{1}{100} \underbrace{E[(X - 50)^2]}_{=\mathrm{var}[X]} =$
  $= \frac{\sigma_x^2}{100} = \frac{25}{100} = \frac{1}{4}$
  da cui $P[40 < X < 60] \geq \frac{3}{4}$

Nota bene: se fosse nota la funzione di densità $f(x)$ associata ad X, potremmo calcolare esattamente la $P[X \geq 75]$, poichè
$P[X \geq a] = 1 - P[X < a]$ e

- se X è continua $P[X < a] = \int_{-\infty}^{a} f(x)dx$,

- se X è discreta $P[X < a] = P[X \leq a] - P[X = a] = F(a) - f(a)$.

Proprietà della varianza

1. $\text{var}[X] = E[(X - E[X])^2] = E[(X - \mu_X)^2]$
2. $\text{var}[X] = E[X^2] - E[X]^2 = E[X^2] - \mu_X^2$
3. $\text{var}[aX] = a^2 \, \text{var}[X], \ \forall a \in \mathbb{R}$
4. $\text{var}[X + a] = \text{var}[X], \ \forall a \in \mathbb{R}$

4.6 I momenti

Oltre al valore atteso e alla varianza esistono altre quantità che possono misurare le caratteristiche di una variabile casuale: sono i valori attesi delle potenze della variabile casuale, detti *momenti*.
Definizione: Sia X una variabile casuale con media μ_X. Chiamiamo momento di ordine r ($r \in \mathbb{N}$) la quantità:

$$\mu_r' = E[X^r]$$

dove

- $E[X^r] = \sum_i x_i^r f(x_i)$ se X è discreta,

- $E[X^r] = \int_{-\infty}^{\infty} x^r f(x)dx$ se X è continua.

Si noti che
$\mu'_1 = E[X] = \mu_X$.

Definizione: Sia X una variabile casuale con media μ_X. Chiamiamo momento centrale di ordine r ($r \in \mathbb{N}$) la quantità:

$$\mu_r = E[(X - \mu_X)^r]$$

dove

- $\mu_r = \sum_i (x_i - \mu_X)^r f(x_i)$ se X è discreta,
- $\mu_r = \int_{-\infty}^{\infty} (x - \mu_X)^r f(x)dx$ se X è continua.

Osservazione
Poichè $\mathrm{var}[X] = E[X^2] - E[X]^2$ possiamo anche scrivere, essendo $\mu'_1 = \mu_X$

$$\sigma_X^2 = \mu'_2 - (\mu'_1)^2$$