

Il valore atteso di una v.a. (discreta)

Introduciamo un nuovo concetto.

DEFINIZIONE DI VALORE ATTESO DI UNA V.A. DISCRETA

Sia $X : \Omega \rightarrow \mathbb{R}$ una v.a. discreta avente come immagine in \mathbb{R} l'insieme V . Se la serie $\sum_{v \in V} |v| \cdot f_X(v)$ converge, allora si chiama valore atteso di X il numero reale

$$\mathbb{E}(X) = \sum_{v \in V} v \cdot f_X(v).$$

Il valore atteso è anche chiamato *media*, ma cercheremo di evitare questo termine per non confonderlo con la *media campionaria* della statistica inferenziale, semmai lo chiameremo **media teorica**.

La \mathbb{E} sta per *expectation*.

Valore atteso con V finito

Scrivo $V = \{x_1, x_2, \dots, x_n\}$ (qui V ha n elementi):

$$\mathbb{E}(X) = \sum_{i=1}^n x_i \cdot \mathbb{P}(X = x_i) \quad \mathbb{E}(X) = \sum_{i=1}^n x_i \cdot \mathbb{P}(X = x_i) \cdot \mathbb{E}(X)$$

sommo su tutti i valori possibili: il valore moltiplicato per la sua probabilità.

Si tratta di una **media pesata** dei valori, dove il **peso** di un valore è la sua **probabilità**.

Valori più probabili hanno peso maggiore, valori meno probabili peso inferiore.

Valore atteso con V numerabile

Se V numerabile

Scrivo $V = \{x_i\}_{i=1}^{\infty}$ e la sostanza della formula non cambia:

$$\mathbb{E}(X) = \sum_{i=1}^{\infty} x_i \cdot \mathbb{P}(X = x_i).$$

Unico problema tecnico: in questo caso potrebbe non essere vero che

$\sum_{i=1}^{\infty} |x_i| \cdot \mathbb{P}(X = x_i) < +\infty$, in tal caso si dice che X non ha valore atteso. I modelli che vedremo non hanno questo problema.

La richiesta di *convergenza assoluta* non dipende dall'ordinamento scelto per V .

Il valore atteso della vincita

Torniamo alla roulette e calcoliamo

$$\begin{aligned}\mathbb{E}(X) &= 0 \cdot \frac{9}{37} + 20 \cdot \frac{9}{37} + 30 \cdot \frac{10}{37} + 50 \cdot \frac{8}{37} + 200 \cdot \frac{1}{37} \\ &= \frac{1080}{37} \approx 29.19\end{aligned}$$

Quindi la “vincita attesa” è di 29.19€. Peccato che la giocata costi 30€!!!

Ecco perché non conviene darsi al gioco (a meno che non siate il casinò!). Il costo della giocata è *mediamente* maggiore della vincita.

Proprietà del valore atteso

TEOREMA

- se a, b sono numeri reali, $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$;
- se X e Y sono due v.a. allora $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

NOTA BENE

In generale $\mathbb{E}(X)$ **non** è il valore più probabile! Potrebbe anche non essere un valore assunto dalla v.a.

Un esempio

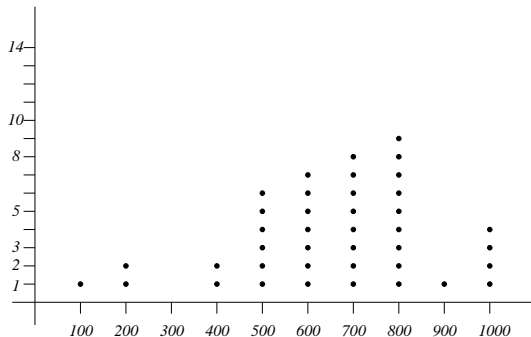
Supponiamo di avere un piccolo paese in cui le case sono disposte su un'unica strada e X sia la variabile "distanza fra la residenza di un abitante e il cartello di inizio paese". X è aleatoria nel senso che **preso un abitante a caso** (nello spazio Ω = abitanti) la distanza è funzione dell'abitante scelto.

Rappresentiamo sull'asse delle x le distanze in metri e i 40 abitanti siano pallini sopra la posizione della loro casa.

Grafici di frequenza

Di fatto se ai pallini sostituissimo delle barre verticali avremmo un grafico di frequenza assoluta (frequenza assoluta di un valore = numero di volte che tale valore è presente).

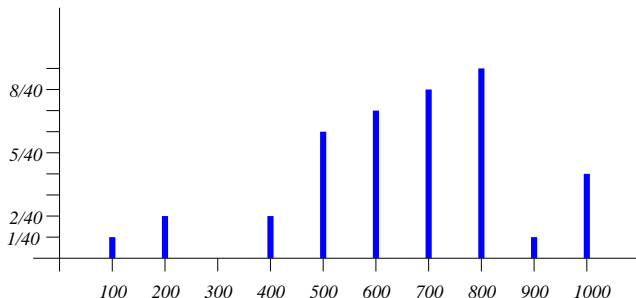
Il grafico



$$V = \{100, 200, 400, 500, 600, 700, 900, 1000\}$$

La frequenza relativa

Rappresentiamo la frequenza relativa (= frequenza assoluta diviso numero totale) di ogni distanza. In pratica riscaldiamo solo l'asse delle y .



In questo caso la frequenza relativa coincide con la densità:

$$f_X(100) = \frac{1}{40}, f_X(200) = \frac{2}{40}, f_X(400) = \frac{2}{40}, f_X(500) = \frac{6}{40}, f_X(600) = \frac{7}{40}, \\ f_X(700) = \frac{8}{40}, f_X(800) = \frac{9}{40}, f_X(900) = \frac{1}{40}, f_X(1000) = \frac{4}{40}.$$

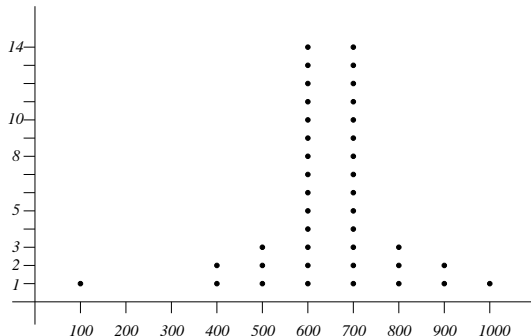
Valore atteso = centro

Nota

In questo caso la densità non è approssimata, bensì esatta: stiamo infatti considerando **tutta** la popolazione.

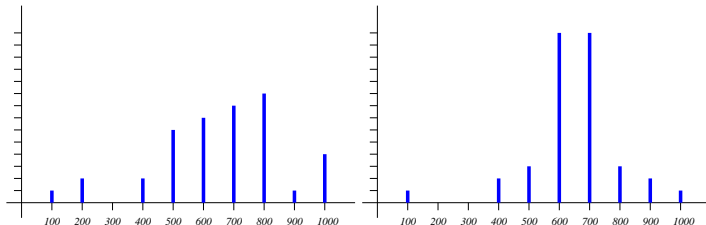
Il centro geografico (pensando che il cartello di uscita paese sia a 1100m da quello di ingresso) è a 550m dall'ingresso paese, ma $\mathbb{E}(X) = 655\text{m}$ rappresenta il “centro democratico”, nel senso che è la media delle distanze pesando maggiormente le distanze più rappresentate (= con più case).

Un altro paese



Qui $\mathbb{E}(X) = 645\text{m}$ ma il paese sembra meno “disperso”.
Confrontiamo i due paesi.

Confronto



Differenze

Anche se i valori più estremi sono gli stessi (100 e 1000), il secondo è meno “disperso”. Inoltre sembrerebbe che se scegliamo un abitante a caso nel secondo paese c'è minor incertezza sull'esito (è assai probabile che pescheremo un abitante alla distanza 600 o 700).

Vogliamo definire un numero che misuri la dispersione (o incertezza) di una v.a.

La varianza di una v.a.

DEFINIZIONE DI VARIANZA DI UNA V.A.

Data una variabile aleatoria X la sua varianza è il numero reale:

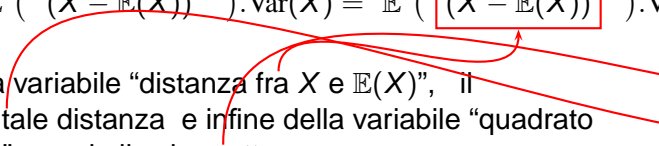
$$\text{Var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2).$$

La radice $\sqrt{\text{Var}(X)}$ è detta deviazione standard o anche scarto quadratico medio.

$\text{Var}(X)$ sarà grande per le X “più sparse sui valori possibili” e piccola per X meno “sparse”.

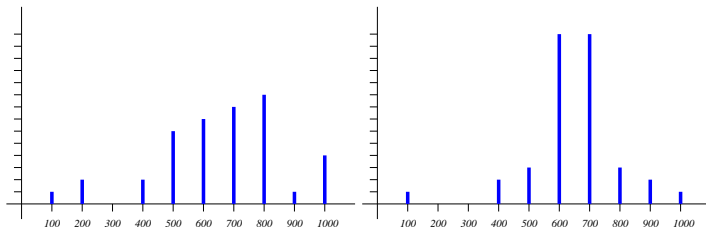
Significato della formula

Ricordiamo che “ \mathbb{E} ” davanti a una v.a. indica il valore atteso – o media pesata dei valori possibili – di tale v.a.

$$\text{Var}(X) = \mathbb{E} \left((X - \mathbb{E}(X))^2 \right).$$


considero la variabile “distanza fra X e $\mathbb{E}(X)$ ”, il quadrato di tale distanza e infine della variabile “quadrato di $X - \mathbb{E}(X)$ ” prendo il valore atteso

La varianza dei due paesi



Il primo paese ha $\text{Var}(X) = 43975m^2$, il secondo $\text{Var}(X) = 21975m^2$.

Calcoliamo per il primo paese:

$$V = \{100, 200, 400, 500, 600, 700, 900, 1000\}$$

$$f_X(100) = \frac{1}{40}, f_X(200) = \frac{2}{40}, f_X(400) = \frac{2}{40}, f_X(500) = \frac{6}{40}, f_X(600) = \frac{7}{40}, \\ f_X(700) = \frac{8}{40}, f_X(800) = \frac{9}{40}, f_X(900) = \frac{1}{40}, f_X(1000) = \frac{4}{40}.$$

$$\mathbb{E}(X) = 100 \cdot \frac{1}{40} + 200 \cdot \frac{2}{40} + \cdots + 1000 \cdot \frac{4}{40} = 655.$$

La variabile $(X - \mathbb{E}(X))^2$ assume valori in

$$\{(100 - 655)^2, (200 - 655)^2, (400 - 655)^2, (500 - 655)^2, (600 - 655)^2, (700 - 655)^2, (900 - 655)^2, (1000 - 655)^2\},$$

rispettivamente con probabilità $\frac{1}{40}, \frac{2}{40}$, etc.

$$\text{Quindi } \text{Var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) =$$

$$(100 - 655)^2 \cdot \frac{1}{40} + (200 - 655)^2 \cdot \frac{2}{40} + \cdots + (1000 - 655)^2 \cdot \frac{4}{40} = 43975.$$

Proprietà della varianza

TEOREMA

- 1 $\text{Var}(X)$ è sempre un numero ≥ 0 ;
- 2 $\text{Var}(X) = 0$ se e solo se X è una v.a. costante*;
- 3 $\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$;
- 4 se a, b sono numeri reali, $\text{Var}(aX + b) = a^2 \text{Var}(X)$;
- 5 se X e Y sono due v.a. allora
$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y) **.$$

* cioè una variabile che può assumere un solo valore.

** con $\text{Cov}(X, Y)$ indichiamo la **covarianza di X e Y** e cioè il numero
$$\text{Cov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

Si dice che

$\mathbb{E}(X)$ è un **indice di posizione**, $\text{Var}(X)$ è un **indice di dispersione**.

Usiamo la (3)

L'uguaglianza

$$\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

è utile per il calcolo: vediamo l'esempio del primo paese.
Sappiamo che $\mathbb{E}(X) = 655m$.

Calcoliamo il valore atteso della v.a. X^2 : assume i valori $\{100^2, 200^2, 400^2, 500^2, 600^2, 700^2, 900^2, 1000^2\}$,
rispettivamente con probabilità $\frac{1}{40}, \frac{2}{40}$, etc.
Allora

$$\mathbb{E}(X^2) = 100^2 \cdot \frac{1}{40} + 200^2 \cdot \frac{2}{40} + \dots + 1000^2 \cdot \frac{4}{40} m^2.$$

Per ottenere la varianza a questo numero sottrarremo $655^2 m^2$.

La formula (3)

$$\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

Per ricordare

Ci sono due operazioni: il quadrato e il valore atteso. Nel primo pezzo si fa prima il quadrato di X e poi il valore atteso, nel secondo prima il valore atteso e poi il quadrato.

Perché è comoda

Si evitano le sottrazioni (valore possibile - $\mathbb{E}(X)$) (che sono tante quante i valori possibili) e se ne fa una sola.

Nell'esempio

Si evitano le sottrazioni $100 - 655$, $200 - 655$, etc.

Dimostrazione delle proprietà

Dimostrazione

- 1 Segue dal fatto che $\text{Var}(X)$ è il valore atteso di una v.a. che assume solo valori ≥ 0 .
- 2 Se $X = c$ con probabilità 1, $\mathbb{E}(X) = c$ e $\text{Var}(X) = \mathbb{E}((c - c)^2) = 0$.
Viceversa se $\text{Var}(X) = 0$ deve essere $X = \mathbb{E}(X)$ ma allora X è costante.
- 3 $\mathbb{E}((X - \mathbb{E}(X))^2) =$

$$\begin{aligned} &= \mathbb{E}(X^2 - 2X \cdot \mathbb{E}(X) + (\mathbb{E}(X))^2) \\ &= \mathbb{E}(X^2) - \mathbb{E}(2X \cdot \mathbb{E}(X)) + \mathbb{E}((\mathbb{E}(X))^2) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X) \cdot \mathbb{E}(X) + (\mathbb{E}(X))^2 = \mathbb{E}(X^2) - (\mathbb{E}(X))^2. \end{aligned}$$

- 4 $\text{Var}(aX + b) = \mathbb{E}((aX + b)^2) - (\mathbb{E}(aX + b))^2$
$$\begin{aligned} &= \mathbb{E}(a^2X^2 + b^2 + 2abX) - (a\mathbb{E}(X) + b)^2 \\ &= a^2\mathbb{E}(X^2) + 2ab\mathbb{E}(X) + b^2 - a^2(\mathbb{E}(X))^2 - b^2 - 2ab\mathbb{E}(X) = a^2\text{Var}(X). \end{aligned}$$

Dimostrazione delle proprietà

Dimostrazione

$$\begin{aligned} \textcircled{5} \quad \text{Var}(X + Y) &= \mathbb{E}((X + Y)^2) - (\mathbb{E}(X + Y))^2 \\ &= \mathbb{E}(X^2 + Y^2 + 2XY) - (\mathbb{E}(X) + \mathbb{E}(Y))^2 \\ &= \mathbb{E}(X^2) + \mathbb{E}(Y^2) + 2\mathbb{E}(XY) - (\mathbb{E}(X))^2 - 2\mathbb{E}(X)\mathbb{E}(Y) - (\mathbb{E}(Y))^2 \\ &= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y). \end{aligned}$$

Un commento su (4)

$$\text{Var}(aX + b) = a^2 \text{Var}(X)$$

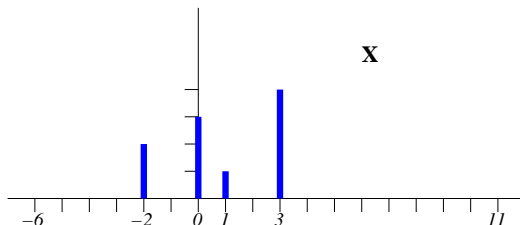
non deve stupire: se V è l'insieme dei valori possibili di X , l'insieme dei valori possibili della v.a. Y dove $Y = aX + b$ non è altro che il risultato di una **dilatazione** (per a) di V , seguita da una **traslazione** di b .

Vediamo un esempio.

Un esempio

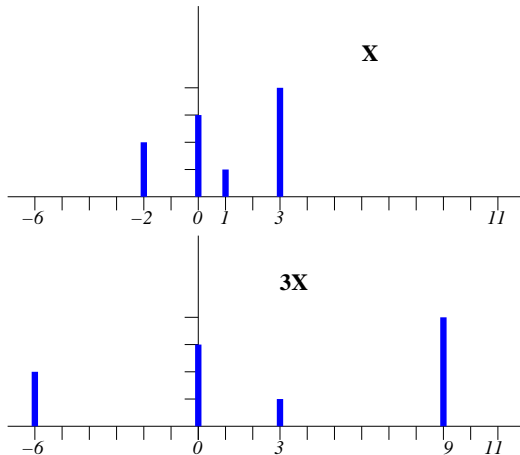
Sia X una v.a. con $V = \{-2, 0, 1, 3\}$ e $f_X(-2) = 0.2$, $f_X(0) = 0.3$, $f_X(1) = 0.1$, $f_X(3) = 0.4$.

Ecco il grafico della densità:



$$Z = 3X$$

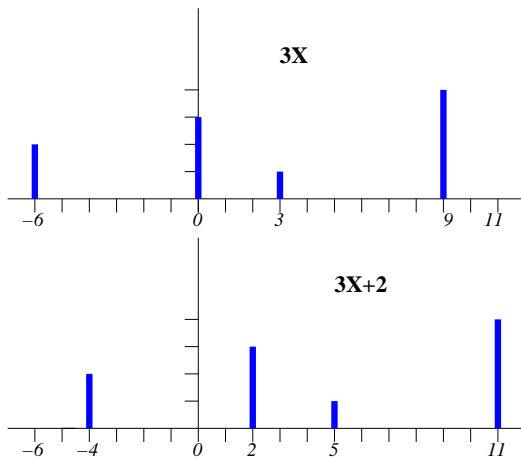
Sia ora $Z = 3X$: confrontiamo le densità:



Vedete come moltiplicare per 3 amplifica la dispersione?

$$Y = Z + 2 = 3X + 2$$

Ora aggiungiamo 2: $Y = Z + 2$. Le densità:



La traslazione di 2 non ha effetto sulla varianza (la dilatazione - moltiplicare per 3 - invece sì).

Multipli negativi

NOTA BENE

Se a è negativo, comunque la varianza non sarà MAI negativa!!!

Ad esempio:

$$\text{Var}(-X) = \text{Var}(X)$$

$$\text{Var}(-3X) = 9\text{Var}(X)$$

$$\text{Var}(-2X + 12) = 4\text{Var}(X).$$