Probabilità e Processi Stocastici (455AA)

Lezione 7

Dario Trevisan – *https://web.dm.unipi.it/trevisan*

13/10/2025

Ross Statistica

Indicatori caratteristici

Come riassumere la legge di una variabile?

Indicatori di *posizione*, di *dispersione* e di *correlazione*.

- Funzione di ripartizione (CDF), funzione di sopravvivenza

Come riassumere la legge di una variabile?

Indicatori di *posizione*, di *dispersione* e di *correlazione*.

- Funzione di ripartizione (CDF), funzione di sopravvivenza
- Funzione quantile, mediana (quartili, decili, ecc.)

Come riassumere la legge di una variabile?

Indicatori di *posizione*, di *dispersione* e di *correlazione*.

- Funzione di ripartizione (CDF), funzione di sopravvivenza
- Funzione quantile, mediana (quartili, decili, ecc.)
- Valor medio

Come riassumere la legge di una variabile?

Indicatori di *posizione*, di *dispersione* e di *correlazione*.

- Funzione di ripartizione (CDF), funzione di sopravvivenza
- Funzione quantile, mediana (quartili, decili, ecc.)
- Valor medio
- Varianza e deviazione standard

Come riassumere la legge di una variabile?

Indicatori di *posizione*, di *dispersione* e di *correlazione*.

- Funzione di ripartizione (CDF), funzione di sopravvivenza
- Funzione quantile, mediana (quartili, decili, ecc.)
- Valor medio
- Varianza e deviazione standard
- Covarianza e correlazione.

Come riassumere la legge di una variabile?

Indicatori di *posizione*, di *dispersione* e di *correlazione*.

- Funzione di ripartizione (CDF), funzione di sopravvivenza
- Funzione quantile, mediana (quartili, decili, ecc.)
- Valor medio
- Varianza e deviazione standard
- Covarianza e correlazione.
- Momenti e funzione generatrice dei momenti

Come riassumere la legge di una variabile?

Indicatori di *posizione*, di *dispersione* e di *correlazione*.

- Funzione di ripartizione (CDF), funzione di sopravvivenza
- Funzione quantile, mediana (quartili, decili, ecc.)
- Valor medio
- Varianza e deviazione standard
- Covarianza e correlazione.
- Momenti e funzione generatrice dei momenti
- Funzione caratteristica

Come riassumere la legge di una variabile?

Indicatori di *posizione*, di *dispersione* e di *correlazione*.

- Funzione di ripartizione (CDF), funzione di sopravvivenza
- Funzione quantile, mediana (quartili, decili, ecc.)
- Valor medio
- Varianza e deviazione standard
- Covarianza e correlazione.
- Momenti e funzione generatrice dei momenti
- Funzione caratteristica
- Cenni all'entropia

Funzione di ripartizione e di sopravvivenza

Data X a valori in $\mathbb{R}$, si è spesso interessati a conoscere la probabilità che X assuma valori “grandi”.

- Esempio: la quantità di denaro che un investitore potrebbe guadagnare (se positiva) o perdere (se negativa) in una fissata data futura, a seconda dell’andamento del mercato

Definiamo

Funzione di ripartizione e di sopravvivenza

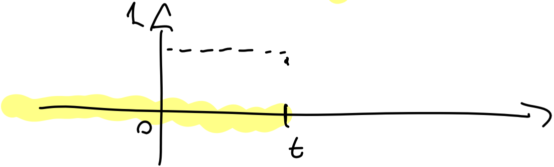
Data X a valori in $\mathbb{R}$, si è spesso interessati a conoscere la probabilità che X assuma valori "grandi".

- Esempio: la quantità di denaro che un investitore potrebbe guadagnare (se positiva) o perdere (se negativa) in una fissata data futura, a seconda dell'andamento del mercato

Definiamo

1. la **funzione di ripartizione** (o funzione cumulativa, in inglese *cumulative distribution function*, CDF,) di X

$$\mathbb{R} \ni \overset{t}{\cancel{x}} \mapsto \mathrm{CDF}_X(\overset{t}{\cancel{x}}) = P(X \leq \overset{t}{\cancel{x}}) \in [0,1]$$



Funzione di ripartizione e di sopravvivenza

Data X a valori in $\mathbb{R}$, si è spesso interessati a conoscere la probabilità che X assuma valori "grandi".

- Esempio: la quantità di denaro che un investitore potrebbe guadagnare (se positiva) o perdere (se negativa) in una fissata data futura, a seconda dell'andamento del mercato

Definiamo

1. la **funzione di ripartizione** (o funzione cumulativa, in inglese *cumulative distribution function*, CDF,) di X

$$\mathbb{R} \ni x \mapsto \mathrm{CDF}_X(x) = P(X \leq x).$$

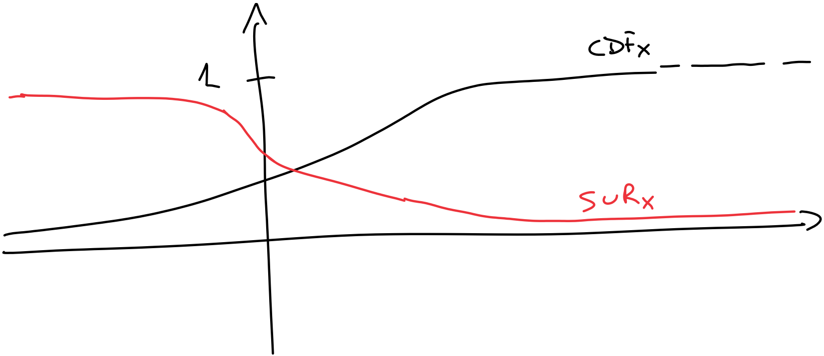
2. la **funzione di sopravvivenza** (in inglese *survival function*) di X,

$$\mathbb{R} \ni x \mapsto \mathrm{SUR}_X(x) = P(X > x).$$

Come calcolarle

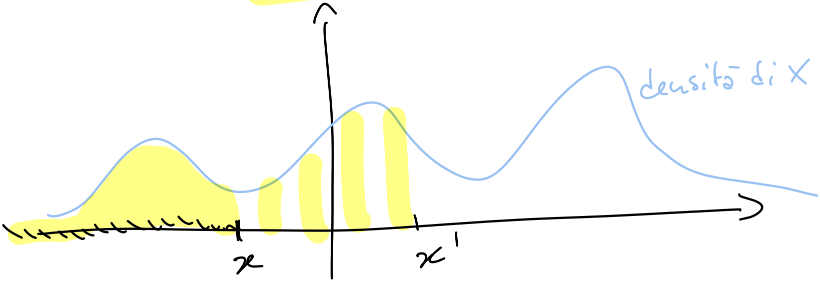
- Essendo $\{X \leq x\}$ e $\{X > x\}$ un sistema di alternative,

$$\mathrm{CDF}_X(x) + \mathrm{SUR}_X(x) = 1. \qquad \forall x$$



Come calcolarle

- Essendo $\{X \leq x\}$ e $\{X > x\}$ un sistema di alternative,
$$\mathrm{CDF}_X(x) + \mathrm{SUR}_X(x) = 1.$$
- Se la densità (discreta o continua) di X è nota, siccome $P(X \leq x) = P(X \in (-\infty, x])$, si trova
$$\mathrm{CDF}_X(x) = \begin{cases} \sum_{z \leq x} P(X = z) & \text{se } X \text{ ha densità discreta,} \\ \int_{-\infty}^{x} f(z)dz & \text{se } X \text{ ha densità continua.} \end{cases}$$



Come calcolarle

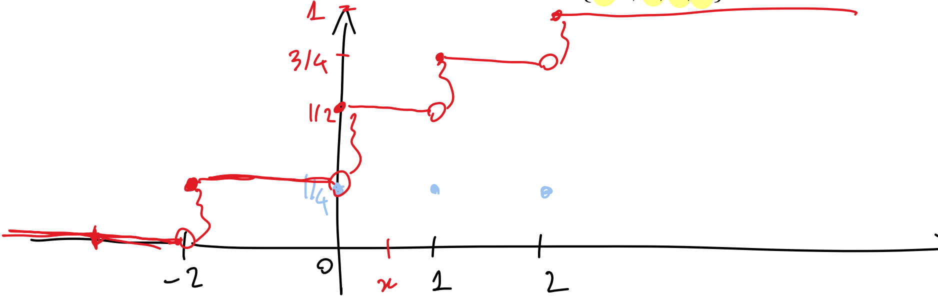
- Essendo $\{X \leq x\}$ e $\{X > x\}$ un sistema di alternative,
$$\mathrm{CDF}_X(x) + \mathrm{SUR}_X(x) = 1.$$
- Se la densità (discreta o continua) di X è nota, siccome $P(X \leq x) = P(X \in (-\infty, x])$, si trova
$$\mathrm{CDF}_X(x) = \begin{cases} \sum_{z \leq x} P(X = z) & \text{se } X \text{ ha densità discreta,} \\ \int_{-\infty}^{x} f(z)dz & \text{se } X \text{ ha densità continua.} \end{cases}$$
- Possiamo quindi intepretare (almeno nel caso continuo) la $\mathrm{CDF}_X(x)$ come l'area del sottografico della densità da $-\infty$ fino ad x.

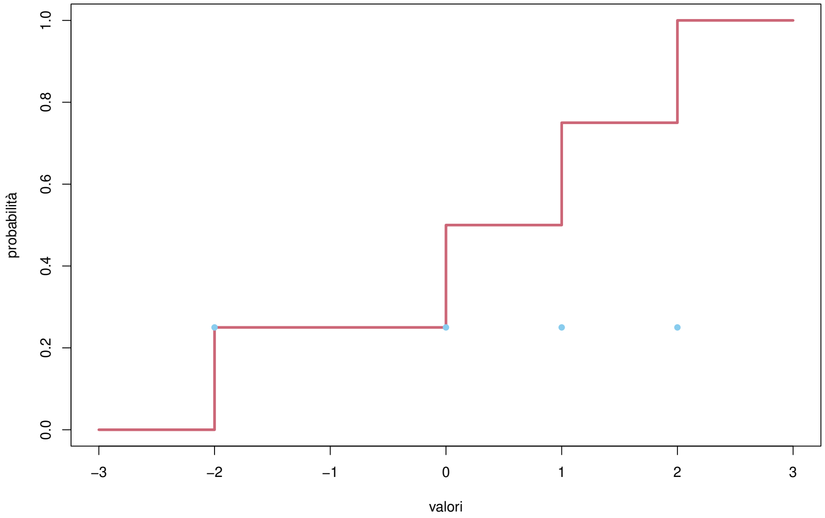
Come calcolarle

- Essendo $\{X \leq x\}$ e $\{X > x\}$ un sistema di alternative,
$$\mathrm{CDF}_X(x) + \mathrm{SUR}_X(x) = 1.$$
- Se la densità (discreta o continua) di X è nota, siccome $P(X \leq x) = P(X \in (-\infty, x])$, si trova
$$\mathrm{CDF}_X(x) = \begin{cases} \sum_{z \leq x} P(X = z) & \text{se } X \text{ ha densità discreta,} \\ \int_{-\infty}^{x} f(z)dz & \text{se } X \text{ ha densità continua.} \end{cases}$$
- Possiamo quindi intepretare (almeno nel caso continuo) la $\mathrm{CDF}_X(x)$ come l'area del sottografico della densità da $-\infty$ fino ad x.
- Per la funzione di sopravvivenza,
$$\mathrm{SUR}_X(x) = \begin{cases} \sum_{z > x} P(X = z) & \text{se } X \text{ ha densità discreta,} \\ \int_{x}^{+\infty} f(z)dz & \text{se } X \text{ ha densità continua,} \end{cases}$$

Esempio nel caso discreto:

- Sia X uniforme discreta sui valori $E = \{-2, 1, 0, 2\}$.



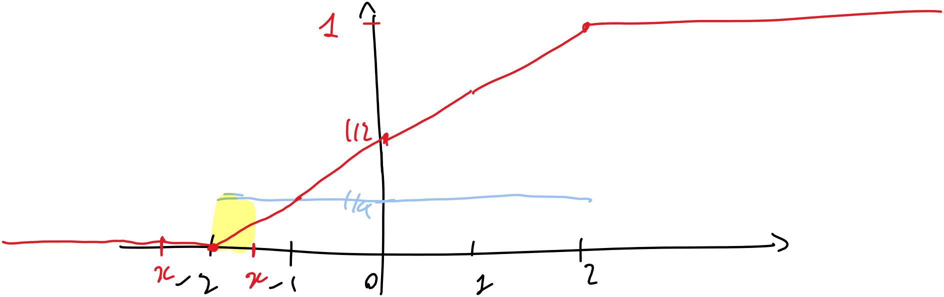


probabilità
0.0
0.2
0.4
0.6
0.8
1.0
−3
−2
−1
0
1
2
3
valori

Esempio nel caso continuo

Si consideri una variabile aleatoria X avente densità uniforme continua nell'intervallo $[-2, 2]$, ossia per $x \in [-2, 2]$,

$$p(X = x) = 1/4.$$



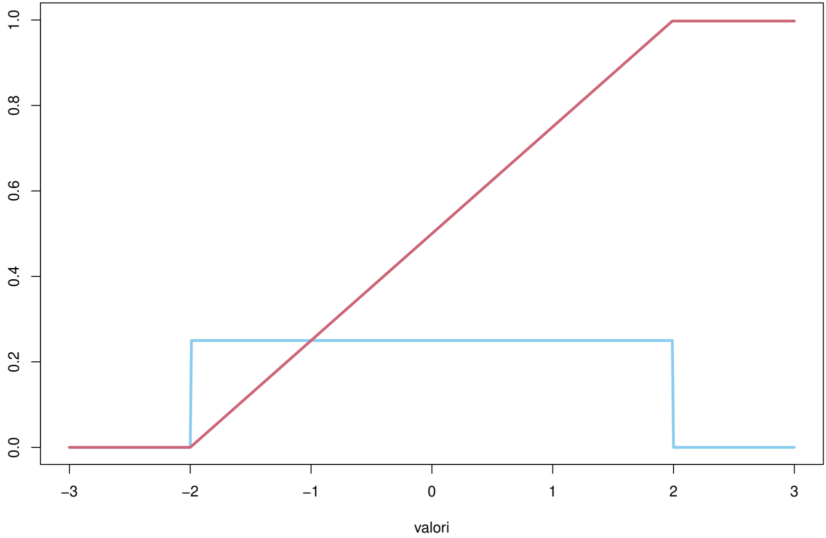


Figure 1: densità uniforme continua e CDF

Proprietà

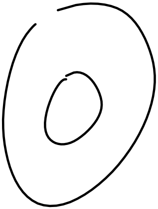
1. vale $\mathrm{CDF}_X(x) \in [0, 1]$, essendo una probabilità.

Proprietà

1. vale $\mathrm{CDF}_X(x) \in [0,1]$, essendo una probabilità.
2. la funzione $x \mapsto \mathrm{CDF}_X(x)$ è crescente (ma non strettamente): se $x < z$, per la monotonia della probabilità,

$$\mathrm{CDF}_X(x) \leq \mathrm{CDF}_X(z)$$

$$\{X \leq x\} \subseteq \{X \leq z\}$$



Proprietà

1. vale $\mathrm{CDF}_X(x) \in [0, 1]$, essendo una probabilità.
2. la funzione $x \mapsto \mathrm{CDF}_X(x)$ è crescente (ma non strettamente): se $x < z$, per la monotonia della probabilità,

$$\mathrm{CDF}_X(x) \leq \mathrm{CDF}_X(z)$$

3. vale $\mathrm{CDF}_X(-\infty) = 0$ e $\mathrm{CDF}_X(+\infty) = 1$ (nel senso di limiti opportuni):

Proprietà

1. vale $\mathrm{CDF}_X(x) \in [0,1]$, essendo una probabilità.
2. la funzione $x \mapsto \mathrm{CDF}_X(x)$ è crescente (ma non strettamente): se $x < z$, per la monotonia della probabilità,
$$\mathrm{CDF}_X(x) \leq \mathrm{CDF}_X(z)$$
3. vale $\mathrm{CDF}_X(-\infty) = 0$ e $\mathrm{CDF}_X(+\infty) = 1$ (nel senso di limiti opportuni):
4. Nel caso di densità discreta, la CDF_X è una funzione costante a tratti, mentre nel caso di variabili con densità continua, la CDF_X è una funzione continua.

Proprietà

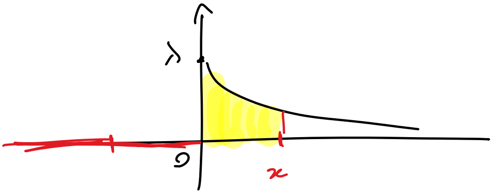
1. vale $\mathrm{CDF}_X(x) \in [0, 1]$, essendo una probabilità.
2. la funzione $x \mapsto \mathrm{CDF}_X(x)$ è crescente (ma non strettamente): se $x < z$, per la monotonia della probabilità,

$$\mathrm{CDF}_X(x) \leq \mathrm{CDF}_X(z)$$

3. vale $\mathrm{CDF}_X(-\infty) = 0$ e $\mathrm{CDF}_X(+\infty) = 1$ (nel senso di limiti opportuni):
4. Nel caso di densità discreta, la CDF_X è una funzione costante a tratti, mentre nel caso di variabili con densità continua, la CDF_X è una funzione continua.
5. Per la funzione SUR, valgono proprietà analoghe: la funzione è decrescente e vale $\mathrm{SUR}_X(-\infty) = 1$ mentre $\mathrm{SUR}_X(+\infty) = 0$.

Un ulteriore esempio

$$\mathcal{E}xp(\lambda)(x) = \begin{cases} \lambda e^{-\lambda x} & \text{se } x \geq 0 \\ 0 & \text{altrimenti} \end{cases}$$



Si consideri una variabile aleatoria X con densità esponenziale di parametro $\lambda > 0$.

Se $x \geq 0$

$$CDF_X(x) = \int_0^x \lambda e^{-\overbrace{\lambda t}^{y}} dt = [y = \lambda t]$$

$$= \int_0^{\lambda x} e^{-y} dy = -e^{-y}\Big|_0^{\lambda x}$$

$$= -e^{-\lambda x} + 1 = 1 - e^{-\lambda x}$$

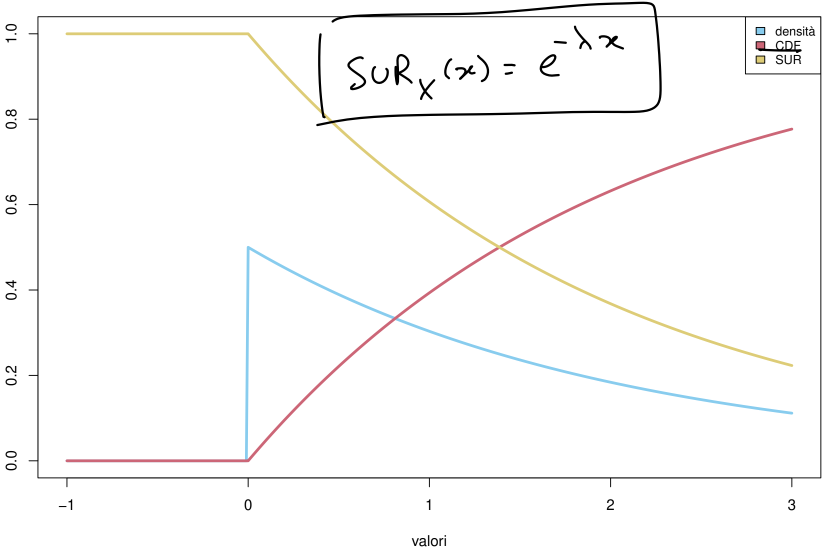
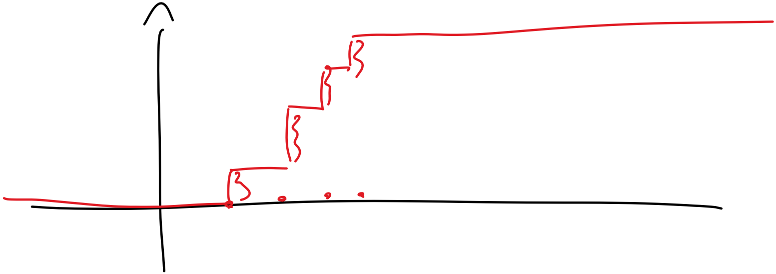


Figure 2: densità, CDF_X e SUR_X di una variabile X con densità esponenziale di parametro $\lambda = 1/2$.

Dalla CDF alla densità

- Nel caso discreto, la densità è non nulla solo nei valori $x \in \mathbb{R}$ in cui la $\mathrm{CDF}_X(x)$ ha un salto, e il valore della densità in quel punto è proprio *l'ampiezza del salto*.



Dalla CDF alla densità

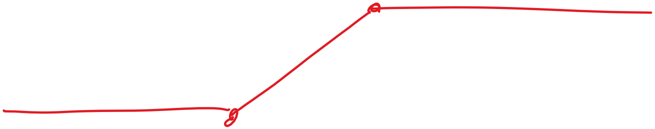
- Nel caso discreto, la densità è non nulla solo nei valori $x \in \mathbb{R}$ in cui la $\mathrm{CDF}_X(x)$ ha un salto, e il valore della densità in quel punto è proprio *l'ampiezza del salto*.
- Nel caso di densità continua, per invertire la formula

$$\int_{-\infty}^{x} p(X = z)dz = \mathrm{CDF}_X(x)$$

è sufficiente applicare il teorema fondamentale del calcolo integrale, e quindi derivare la CDF_X:

$$\frac{d}{dx}\mathrm{CDF}_X(x) = p(X = x).$$

(nei punti in cui CDF_X è derivabile)



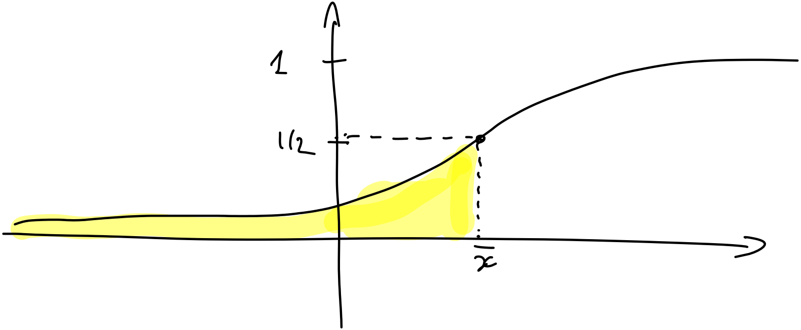
Dalla CDF alla densità

- Nel caso discreto, la densità è non nulla solo nei valori $x \in \mathbb{R}$ in cui la $\mathrm{CDF}_X(x)$ ha un salto, e il valore della densità in quel punto è proprio *l'ampiezza del salto*.
- Nel caso di densità continua, per invertire la formula
$$\int_{-\infty}^{x} p(X = z) dz = \mathrm{CDF}_X(x)$$
è sufficiente applicare il teorema fondamentale del calcolo integrale, e quindi derivare la CDF_X:
$$\frac{d}{dx} \mathrm{CDF}_X(x) = p(X = x).$$
(nei punti in cui CDF_X è derivabile)
- Per la SUR_X basta cambiare di segno: ad esempio (caso continuo)
$$-\frac{d}{dx} \mathrm{SUR}_X(x) = f(x).$$

Mediana (definizione intuitiva)

- La *mediana* di X è definita come un valore $\bar{x} \in \mathbb{R}$, se esiste, tale che

$$\mathrm{CDF}_X(\bar{x}) = \frac{1}{2}.$$



Mediana (definizione intuitiva)

- La *mediana* di X è definita come un valore $\bar{x} \in \mathbb{R}$, se esiste, tale che
$$\mathrm{CDF}_X(\bar{x}) = \frac{1}{2}.$$
- Significa che
$$P(X \leq \bar{x}) = P(X > \bar{x}) = \frac{1}{2},$$
ossia $\bar{x}$ è scelta in modo che le due alternative $\{X \leq \bar{x}\}$ e $\{X > \bar{x}\}$ abbiano la stessa probabilità.

Mediana (definizione intuitiva)

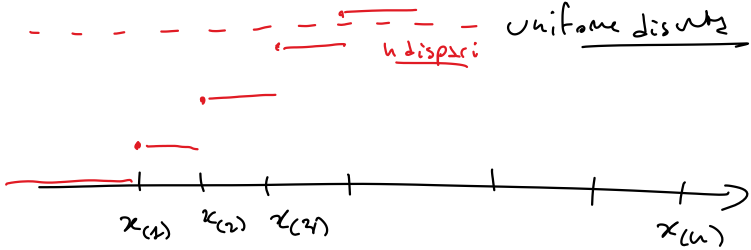
- La *mediana* di X è definita come un valore $\bar{x} \in \mathbb{R}$, se esiste, tale che
$$\mathrm{CDF}_X(\bar{x}) = \frac{1}{2}.$$
- Significa che
$$P(X \leq \bar{x}) = P(X > \bar{x}) = \frac{1}{2},$$
ossia $\bar{x}$ è scelta in modo che le due alternative $\{X \leq \bar{x}\}$ e $\{X > \bar{x}\}$ abbiano la stessa probabilità.
- La mediana è quindi un indicatore di "centralità" per X, alternativo alla moda.

Esempi

- se la densità di X è uniforme (diciamo su un insieme E con un numero pari di valori), significa che metà dei valori possibili di X sono $\leq \bar{x}$, mentre i rimanenti sono $> \bar{x}$. In questo caso invece una moda è uno qualsiasi dei valori di E.

$(x_1, x_2, \dots, x_n)$ si definisce $X = x_I$ dove $I \in \{1, 2, 3, \dots, n\}$

uniforme discreta



Esempi

- se la densità di X è uniforme (diciamo su un insieme E con un numero pari di valori), significa che metà dei valori possibili di X sono $\leq \bar{x}$, mentre i rimanenti sono $> \bar{x}$. In questo caso invece una moda è uno qualsiasi dei valori di E.

- Sia X uniforme su

5/10 6/10

$$E = \{0, 0.1, 0.15, 0.3, 0.4, 0.7, 0.73, 0.9, 0.95, 1.1\}$$

Una mediana per X è $\bar{x} = 0.4$, ma anche un qualsiasi valore compreso tra 0.4 e 0.7 (escluso)

Esempi

- se la densità di X è uniforme (diciamo su un insieme E con un numero pari di valori), significa che metà dei valori possibili di X sono $\leq \bar{x}$, mentre i rimanenti sono $> \bar{x}$. In questo caso invece una moda è uno qualsiasi dei valori di E.

- Sia X uniforme su

$$E = \{0, 0.1, 0.15, 0.3, 0.4, 0.7, 0.73, 0.9, 0.95, 1.1\}$$

ok se n pari

Una mediana per X è $\bar{x} = 0.4$, ma anche un qualsiasi valore compreso tra 0.4 e 0.7 (escluso)

- Nel caso di una variabile esponenziale di parametro λ,

$$CDF_X(\bar{x}) = 1 - e^{-\lambda \bar{x}} = \frac{1}{2} \quad \leftrightarrow \quad \boxed{\lambda \bar{x} = \log_e(2)},$$

ossia

$$\boxed{\bar{x} = \frac{\log_e(2)}{\lambda}.}$$

Funzione quantile

- Problemi della mediana (così definita)

Funzione quantile

- Problemi della mediana (così definita)
  1. non necessariamente la mediana esiste sempre

Funzione quantile

- Problemi della mediana (così definita)
  1. non necessariamente la mediana esiste sempre
  2. non è unicamente determinata

Funzione quantile

- Problemi della mediana (così definita)
  1. non necessariamente la mediana esiste sempre
  2. non è unicamente determinata
  3. non è facile calcolarla

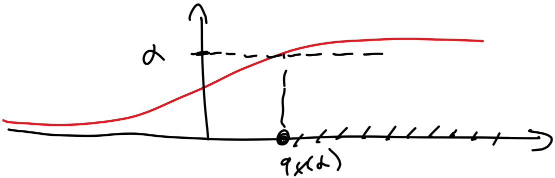
Funzione quantile

- Problemi della mediana (così definita)
  1. non necessariamente la mediana esiste sempre
  2. non è unicamente determinata
  3. non è facile calcolarla
- Per risolvere i primi due, si introduce una inversa generalizzata CDF_X, detta funzione *quantile* di X

$$q_X : (0,1) \to \mathbb{R}$$

che ad ogni *livello* $\alpha \in (0,1)$ associa

$$q_X(\alpha) = \min \{ x \in \mathbb{R} \, : \, \mathrm{CDF}_X(x) \geq \alpha \} \, .$$



Funzione quantile

- Problemi della mediana (così definita)
  1. non necessariamente la mediana esiste sempre
  2. non è unicamente determinata
  3. non è facile calcolarla
- Per risolvere i primi due, si introduce una inversa generalizzata CDF_X, detta funzione *quantile* di X

$$q_X : (0,1) \to \mathbb{R}$$

che ad ogni *livello* $\alpha \in (0,1)$ associa

$$q_X(\alpha) = \min \{ x \in \mathbb{R} \, : \, \mathrm{CDF}_X(x) \geq \alpha \} \, .$$

- se X ha densità continua, allora

se discreta

$$\mathrm{CDF}_X(q_X(\alpha)) = \alpha.$$

Un esempio discreto

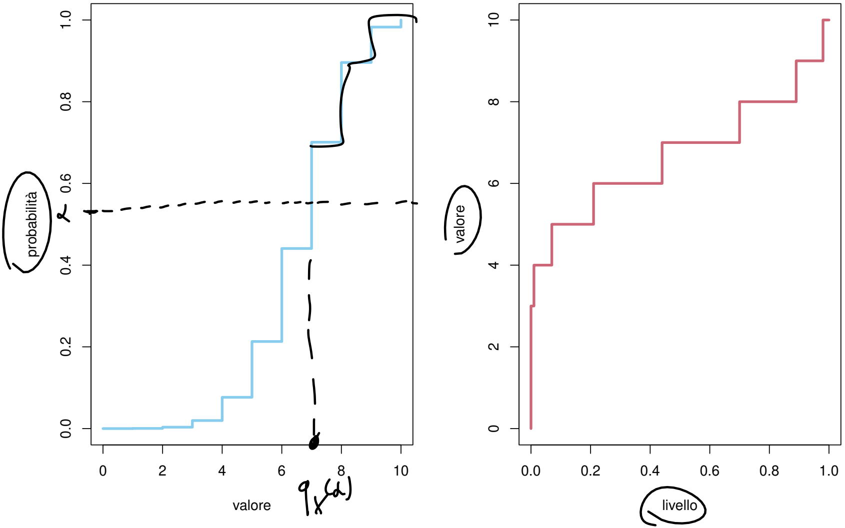


Figure 3: plot di CDF$_X$ per X binomiale (a sinistra) e quantile funzione quantile (a destra)

Mediana, quartili, decili ecc.

Si definiscono quindi

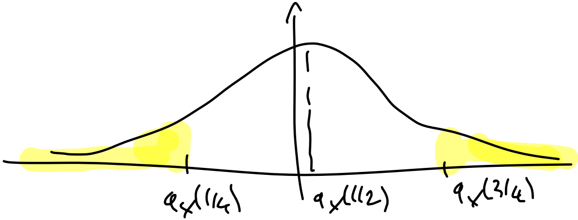
- la mediana di X come il valore $\bar{x} = q_X(1/2)$.

Mediana, quartili, decili ecc.

Si definiscono quindi

- la mediana di X come il valore $\bar{x} = q_X(1/2)$.
- i *quartili*, $q_X(\alpha)$ corrispondenti ad $\alpha \in \{1/4, 2/4, 3/4\}$ (detti il primo, secondo e terzo quartile),

$q_X(1/4)$ $\quad$ $q_X\left(\frac{2}{4}\right) = q_X\left(\frac{1}{2}\right) =$ mediana



Mediana, quartili, decili ecc.

Si definiscono quindi

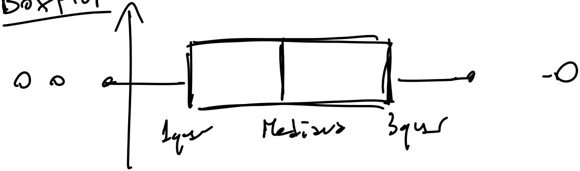
- la mediana di X come il valore $\bar{x} = q_X(1/2)$.
- i *quartili*, $q_X(\alpha)$ corrispondenti ad $\alpha \in \{1/4, 2/4, 3/4\}$ (detti il primo, secondo e terzo quartile),
- i *decili* $q_X(\alpha)$ per $\alpha = k/10$,

Mediana, quartili, decili ecc.

Si definiscono quindi

- la mediana di X come il valore $\bar{x} = q_X(1/2)$.
- i *quartili*, $q_X(\alpha)$ corrispondenti ad $\alpha \in \{1/4, 2/4, 3/4\}$ (detti il primo, secondo e terzo quartile),
- i *decili* $q_X(\alpha)$ per $\alpha = k/10$,
- i *percentili* $q_X(\alpha)$ per $\alpha = k/100$.

Grafico a scatola <u>Boxplot</u>



Valor medio

Definizione

- La mediana, i quartili, decili ecc., sono efficaci ma difficili da calcolare.

Definizione

- La mediana, i quartili, decili ecc., sono efficaci ma difficili da calcolare.
- il valor medio (media, valore atteso o speranza matematica) di X a valori in $\mathbb{R}$, è una *media* dei valori di X *ponderata* tramite la densità (discreta o continua):

$$\mathbb{E}\left[X|I\right] = \begin{cases} \sum_{x \in \mathbb{R}} xP(X = x|I) & \text{se } X \text{ ha densità discreta,} \\ \int_{-\infty}^{\infty} xp(X = x|I)dx & \text{se } X \text{ ha densità continua.} \end{cases}$$

Definizione

- La mediana, i quartili, decili ecc., sono efficaci ma difficili da calcolare.
- il valor medio (media, valore atteso o speranza matematica) di X a valori in $\mathbb{R}$, è una *media* dei valori di X *ponderata* tramite la densità (discreta o continua):

$$\mathbb{E}\left[X|I\right] = \begin{cases} \sum_{x\in\mathbb{R}} xP(X=x|I) & \text{se } X \text{ ha densità discreta,} \\ \int_{-\infty}^{\infty} xp(X=x|I)dx & \text{se } X \text{ ha densità continua.} \end{cases}$$

- La notazione $\mathbb{E}\left[X|I\right]$ ricorda $P(X=x|I)$ (si sottointende a volte I).

Definizione

- La mediana, i quartili, decili ecc., sono efficaci ma difficili da calcolare.
- il valor medio (media, valore atteso o speranza matematica) di X a valori in $\mathbb{R}$, è una *media* dei valori di X *ponderata* tramite la densità (discreta o continua):

$$\mathbb{E}\left[X|I\right] = \begin{cases} \sum_{x\in\mathbb{R}} xP(X = x|I) & \text{se } X \text{ ha densità discreta,} \\ \int_{-\infty}^{\infty} xp(X = x|I)dx & \text{se } X \text{ ha densità continua.} \end{cases}$$

- La notazione $\mathbb{E}\left[X|I\right]$ ricorda $P(X = x|I)$ (si sottointende a volte I).
- Si richiede che serie o integrali convergano assolutamente, ossia

$$\sum_{x\in\mathbb{R}} |x|P(X = x|I) < \infty \text{ oppure } \int_{-\infty}^{\infty} |x|p(X = x|I)dx < \infty.$$

altrimenti il valor medio *non esiste* (può accadere).

Esempi

Sia $X \in \{0,1\}$ la variabile indicatrice di un evento A, ossia $\{X = 1\} = A$.

- La legge di X è Bernoulli di parametro $p = P(X = 1|I) = P(A|I)$

$$\mathbb{E}[X] = 0 \cdot P(X=0) + 1 \cdot P(X=1)$$

$$= P(X=1) = \underline{p} \in [0,1]$$

Esempi

Sia $X \in \{0, 1\}$ la variabile indicatrice di un evento A, ossia $\{X = 1\} = A$.

- La legge di X è Bernoulli di parametro $p = P(X = 1|I) = P(A|I)$
- Per definizione

$$\mathbb{E}[X|I] = 0 \cdot P(X = 0|I) + 1 \cdot P(X = 1|I) = P(X = 1|I) = P(A|I) = p,$$

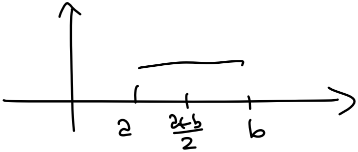
Esempi

Sia $X \in \{0, 1\}$ la variabile indicatrice di un evento A, ossia $\{X = 1\} = A$.

- La legge di X è Bernoulli di parametro $p = P(X = 1|I) = P(A|I)$
- Per definizione

$$\mathbb{E}\,[X|I] = 0 \cdot P(X = 0|I) + 1 \cdot P(X = 1|I) = P(X = 1|I) = P(A|I) = p,$$

- Eccetto i casi limite $p = 0$, oppure $p = 1$, il valor medio non è uno dei possibili valori di X.



Sia X una variabile aleatoria uniforme continua sull'intervallo (a, b).

- Il valor medio di X è

$$\int_a^b x\frac{1}{b-a}dx = \frac{(b^2-a^2)}{2(b-a)} = \frac{a+b}{2},$$

Sia X una variabile aleatoria uniforme continua sull'intervallo (a, b).

- Il valor medio di X è

$$\int_a^b x \frac{1}{b-a} dx = \frac{(b^2 - a^2)}{2(b-a)} = \frac{a+b}{2},$$

- Notiamo che in questo caso il valor medio è uno dei possibili valori e coincide con la mediana.

Sia X una variabile aleatoria con densità binomiale di parametri $n = 10$, $p = 1/3$. Allora

$$\sum_{k=0}^{10} k \underbrace{\binom{10}{k} \frac{1}{3^k} \left(\frac{2}{3}\right)^{10-k}}_{P(X=k)} = 10 \cdot \frac{1}{3}$$

$$\mathbb{E}[X] = n \cdot p$$

- Si può calcolarlo numericamente:

```
## [1] 3.333333
```

Definizione generale

- ci limiteremo al calcolo (esplicito) del valor medio nei casi in cui X ammetta densità discreta o continua.

Definizione generale

- ci limiteremo al calcolo (esplicito) del valor medio nei casi in cui X ammetta densità discreta o continua.
- è possibile dare una definizione generale, che non usa la densità:

$$\mathbb{E}\left[X|I\right] = \int_0^\infty P(X > x|I)dx - \int_{-\infty}^0 P(X < x)dx,$$

La formula di disintegrazione per il valor medio

La proprietà fondamentale è l'analoga della *formula di disintegrazione* per la probabilità, che possiamo scrivere in termini di sistemi di alternative o variabili aleatorie (discrete o continue).

- Sia X una variabile aleatoria reale e sia $Y \in E$ una variabile aleatoria:

$$P(X=x) = \sum_{y \in E} P(X=x \mid Y=y) P(Y=y)$$

$$\sum_{x \in \mathbb{R}} x P(X=x) = \sum_{x \in \mathbb{R}} x \sum_{y} P(X=x \mid Y=y) P(Y=y)$$

$$= \sum_{y \in E} \left(\sum_{x \in \mathbb{R}} x P(X=x \mid Y=y) \right) P(Y=y)$$

$$= \sum_{y} \mathbb{E}[X \mid Y=y] P(Y=y)$$

La formula di disintegrazione per il valor medio

La proprietà fondamentale è l'analoga della *formula di disintegrazione* per la probabilità, che possiamo scrivere in termini di sistemi di alternative o variabili aleatorie (discrete o continue).

- Sia X una variabile aleatoria reale e sia $Y \in E$ una variabile aleatoria:

   1. Se Y ha densità discreta, vale

$$\mathbb{E}[X|I] = \sum_{y \in E} \mathbb{E}[X|I, Y = y] \, P(Y = y|I),$$

La formula di disintegrazione per il valor medio

La proprietà fondamentale è l'analoga della *formula di disintegrazione* per la probabilità, che possiamo scrivere in termini di sistemi di alternative o variabili aleatorie (discrete o continue).

- Sia X una variabile aleatoria reale e sia $Y \in E$ una variabile aleatoria:

  1. Se Y ha densità discreta, vale

  $$\mathbb{E}[X|I] = \sum_{y \in E} \mathbb{E}[X|I, Y = y]\, P(Y = y|I),$$

  2. Se $E = \mathbb{R}^d$ e Y ha densità continua, vale

  $$\mathbb{E}[X|I] = \int_{\mathbb{R}^d} \mathbb{E}[X|I, Y = y]\, p(Y = y|I) dy.$$

Dimostrazione

Fatto

Proprietà del valor medio

Siano X, Y variabili aleatorie reali, e siano a, b, $c \in \mathbb{R}$ (non aleatorie). Allora

1. (linearità) vale $\mathbb{E}[aX|I] = a\mathbb{E}[X|I]$ e $\mathbb{E}[X + Y|I] = \mathbb{E}[X|I] + \mathbb{E}[Y|I]$. ← non vale per moda o mediana

Proprietà del valor medio

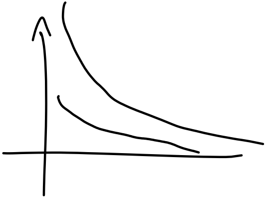
Siano X, Y variabili aleatorie reali, e siano a, b, $c \in \mathbb{R}$ (non aleatorie). Allora

1. (linearità) vale $\mathbb{E}[aX|I] = a\mathbb{E}[X|I]$ e $\mathbb{E}[X + Y|I] = \mathbb{E}[X|I] + \mathbb{E}[Y|I]$.
2. (monotonia) se $P(X \geq Y|I) = 1$, allora $\mathbb{E}[X|I] \geq \mathbb{E}[Y|I]$. In particolare, se $P(X \in [a, b]|I) = 1$, allora $\mathbb{E}[X|I] \in [a, b]$.

Proprietà del valor medio

Siano X, Y variabili aleatorie reali, e siano a, b, $c \in \mathbb{R}$ (non aleatorie). Allora

1. (linearità) vale $\mathbb{E}\left[aX|I\right] = a\mathbb{E}\left[X|I\right]$ e $\mathbb{E}\left[X+Y|I\right] = \mathbb{E}\left[X|I\right] + \mathbb{E}\left[Y|I\right]$.
2. (monotonia) se $P(X \geq Y|I) = 1$, allora $\mathbb{E}\left[X|I\right] \geq \mathbb{E}\left[Y|I\right]$. In particolare, se $P(X \in [a,b]|I) = 1$, allora $\mathbb{E}\left[X|I\right] \in [a,b]$.
3. (diseguaglianza di Markov) se X è a valori non-negativi (rispetto all'informazione I), allora per ogni $c > 0$,



$$P(X > c|I) \leq \frac{\mathbb{E}\left[X|I\right]}{c}.$$

Dimostrazione 1. (linearità)

Dimostrazione 2. (monotonia)

Supponiamo che $\underline{X \geq 0}$ e mostriamo $\mathbb{E}[X] \geq 0$

$$\mathbb{E}[X] = \sum_{\substack{x \in \mathbb{R} \\ x \geq 0}} x\, P(X = x) = \sum (\geq 0)$$

In generale se $X \geq Y$ allora $Z = X - Y \geq 0$

e quindi $\mathbb{E}[Z] \geq 0$

$$\mathbb{E}[X - Y] = \mathbb{E}[X] - \mathbb{E}[Y]$$

<u>linearità</u>

Dimostrazione 3. (Markov)

$$\mathbb{E}[X] = \overset{\geq c}{\mathbb{E}[X \mid X > c]}\; \overset{>0}{P(X > c)} +$$

$$+\; \underset{\geq 0}{\mathbb{E}[X \mid 0 \leq X \leq c]}\; P(X \leq c)$$

$$\geq c \cdot P(X > c)$$

Collegamento tra quantile e valor medio

- Scegliamo $c = q_X(\alpha)$ nella diseguaglianza di Markov.

Collegamento tra quantile e valor medio

- Scegliamo $c = q_X(\alpha)$ nella diseguaglianza di Markov.
- Supponendo che $X \geq 0$ e che abbia densità continua,

$$P(X \geq q_X(\alpha)) = 1 - \mathrm{CDF}_X(q_X(\alpha)) = 1 - \alpha,$$

$$P(X > q_X(\alpha)) = 1 - \alpha \leq \frac{\mathbb{E}[X]}{q_X(\alpha)}$$

Collegamento tra quantile e valor medio

- Scegliamo $c = q_X(\alpha)$ nella diseguaglianza di Markov.
- Supponendo che $X \geq 0$ e che abbia densità continua,

$$P(X \geq q_X(\alpha)) = 1 - \mathrm{CDF}_X(q_X(\alpha)) = 1 - \alpha,$$

- quindi la diseguaglianza implica che

$$1 - \alpha \leq \frac{\mathbb{E}[X]}{q_X(\alpha)}.$$

Collegamento tra quantile e valor medio

- Scegliamo $c = q_X(\alpha)$ nella diseguaglianza di Markov.
- Supponendo che $X \geq 0$ e che abbia densità continua,

$$P(X \geq q_X(\alpha)) = 1 - \mathrm{CDF}_X(q_X(\alpha)) = 1 - \alpha,$$

- quindi la diseguaglianza implica che

$$1 - \alpha \leq \frac{\mathbb{E}[X]}{q_X(\alpha)}.$$

- Ad esempio, con $\alpha = 1/2$ si trova che

$$q_X(1/2) \leq 2\mathbb{E}[X].$$

Valor medio di $g(X)$

Sia $X \in E$ una variabile aleatoria con densità discreta oppure continua (se $E = \mathbb{R}^d$).

- Data $g : E \to \mathbb{R}$, vale

$$\mathbb{E}\left[g(X)|I\right] = \begin{cases} \sum_{x \in E} g(x) P(X = x|I) & \text{se } X \text{ ha densità discreta,} \\ \int_E g(x) p(X = x|I) dx & \text{se } X \text{ ha densità continua.} \end{cases}$$

Valor medio di $g(X)$

Sia $X \in E$ una variabile aleatoria con densità discreta oppure continua (se $E = \mathbb{R}^d$).

- Data $g : E \to \mathbb{R}$, vale

$$\mathbb{E}\left[g(X)|I\right] = \begin{cases} \sum_{x \in E} g(x) P(X = x|I) & \text{se } X \text{ ha densità discreta,} \\ \int_E g(x) p(X = x|I) dx & \text{se } X \text{ ha densità continua.} \end{cases}$$

- Per calcolare il valor medio di una variabile composta $g(X)$, basta conoscere la densità di X (*non* serve calcolare quella di $g(X)$)!

Dimostrazione

Caso discreto

$$\mathbb{E}[g(X)] = \sum_{x \in E} \underbrace{\mathbb{E}[g(X) \mid X = x]}_{= g(x)} P(X = x)$$

$$= \sum_{x \in E} g(x) \, P(X = x)$$

Esempio

Sia X una variabile con densità discreta binomiale di parametri $n = 20$, $p = 1/4$. Si trova

$$\mathbb{E}\left[X^3\right] = \sum_{k=0}^{20} k^3 \binom{20}{k} \frac{1}{4^k} \left(\frac{3}{4}\right)^{20-k}.$$

Possiamo calcolarlo numericamente.

```
## [1] 183.125
```

Valor medio e indipendenza

$$P(X=x, Y=y) = P(X=x)\,P(Y=y)$$

Siano X, Y variabili aleatorie reali indipendenti (rispetto ad una informazione I). Allora

$$\mathbb{E}\left[XY|I\right] = \mathbb{E}\left[X|I\right]\mathbb{E}\left[Y|I\right].$$

- Ricordando che se X, Y sono indipendenti anche eventuali composizioni $g(X)$ e $h(Y)$ lo sono, segue che

$$\mathbb{E}\left[g(X)h(Y)|I\right] = \mathbb{E}\left[g(X)|I\right]\mathbb{E}\left[h(Y)|I\right].$$

(a volte, specie negli esercizi, si è appunto in questa situazione apparentemente più generale)

Dimostrazione

$$P(X,Y) = P(X)P(Y) \iff P(Y|X) = P(Y)$$

$$\mathbb{E}[XY] = \sum_x \mathbb{E}[XY|X=x]\,P(X=x)$$

$$= \sum_x \mathbb{E}[xY|X=x]\,P(X=x)$$

$$= \underbrace{\sum_x x\,P(X=x)\,\mathbb{E}[Y|\cancel{X=x}]}_{\mathbb{E}[X]\,\mathbb{E}[Y]}$$

$$\mathbb{E}[Y|\cancel{X=x}] = \sum_y y\,P(Y=y|\cancel{X=x})$$

Valor medio e indipendenza

Siano X, Y variabili aleatorie reali indipendenti (rispetto ad una informazione I). Allora

$$\mathbb{E}[XY|I] = \mathbb{E}[X|I]\,\mathbb{E}[Y|I].$$

- Ricordando che se X, Y sono indipendenti anche eventuali composizioni $g(X)$ e $h(Y)$ lo sono, segue che

  $$\mathbb{E}[g(X)h(Y)|I] = \mathbb{E}[g(X)|I]\,\mathbb{E}[h(Y)|I].$$

  (a volte, specie negli esercizi, si è appunto in questa situazione apparentemente più generale)

Dimostrazione

Il caso vettoriale

Data una variabile $X = (X_1, \ldots, X_d)$ a valori in $\mathbb{R}^d$, si definisce il *vettore dei valor medi* (o vettore delle medie) di X come il vettore in $\mathbb{R}^d$,

$$\mathbb{E}[X|I] = (\mathbb{E}[X_1|I], \ldots, \mathbb{E}[X_d|I]).$$

Linearità (caso vettoriale)

- Additività:
$$\mathbb{E}\left[X+Y|I\right]=\mathbb{E}\left[X|I\right]+\mathbb{E}\left[Y|I\right]$$
per variabili aleatorie X, Y a valori in $\mathbb{R}^d$.

Linearità (caso vettoriale)

- Additività:
$$\mathbb{E}[X + Y | I] = \mathbb{E}[X | I] + \mathbb{E}[Y | I]$$
per variabili aleatorie X, Y a valori in $\mathbb{R}^d$.
- trasformazioni lineari affini del vettore dei valor medi: data una variabile aleatoria $X \in \mathbb{R}^d$ e posta
$$Y = AX + b \quad \text{ossia} \quad \mathbb{E}[Y_i] = \mathbb{E}\left[\sum_{j=1}^{d} A_{ij} X_j + b_i, \right]$$
dove $A \in \mathbb{R}^{k \times d}$ è una matrice e $b \in \mathbb{R}^k$ è un vettore (noti, ossia costanti rispetto all'informazione I), vale
$$\mathbb{E}[Y | I] = A\mathbb{E}[X | I] + b, \quad \text{ossia} \quad \mathbb{E}[Y_i | I] = \sum_{j=1}^{d} A_{ij} \mathbb{E}[X_j | I] + b_i.$$

Varianza e deviazione standard

.

Varianza

La moda, la mediana ed il valore medio sono tutti indicatori *puntuali*, riassumono tutta la legge con un singolo valore.

Per descrivere in modo più efficace una variabile X si affianca un indicatore della dispersione, ossia della “concentrazione” della sua legge intorno ad un indicatore puntuale.

- Un indicatore di dispersione molto usato è la *deviazione standard* (o scarto quadratico medio), definita come

  $$\sigma_X = \sqrt{\mathrm{Var}\,(X)},$$

  dove la *varianza* è

  $$\mathrm{Var}\,(X) = \mathbb{E}\left[(X - \mathbb{E}\,[X])^2\right].$$

L'unità di misura di σ_X è la stessa di X, mentre $\mathrm{Var}(X)$ ha come unità di misura il quadrato dell'unità di X.

Operativamente:

1. si calcola il valor medio di X,

L'unità di misura di σ_X è la stessa di X, mentre $\mathrm{Var}(X)$ ha come unità di misura il quadrato dell'unità di X.

Operativamente:

1. si calcola il valor medio di X,
2. si considera lo scarto (variabile aleatoria centrata)

$$X - \mathbb{E}[X],$$

L'unità di misura di σ_X è la stessa di X, mentre $\operatorname{Var}(X)$ ha come unità di misura il quadrato dell'unità di X.

Operativamente:

1. si calcola il valor medio di X,
2. si considera lo scarto (variabile aleatoria centrata)

$$X - \mathbb{E}[X],$$

3. se ne prende il quadrato e il valor medio

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

L'unità di misura di σ_X è la stessa di X, mentre $\mathrm{Var}\,(X)$ ha come unità di misura il quadrato dell'unità di X.

Operativamente:

1. si calcola il valor medio di X,
2. si considera lo scarto (variabile aleatoria centrata)
$$X - \mathbb{E}\,[X]\,,$$
3. se ne prende il quadrato e il valor medio
$$\mathrm{Var}\,(X) = \mathbb{E}\left[(X - \mathbb{E}\,[X])^2\right].$$
4. Per la deviazione standard si prende la radice quadrata
$$\sigma_X = \sqrt{\mathbb{E}\,[(X - \mathbb{E}\,[X])^2]}.$$

$$\mathrm{Var}\,(X) = \mathbb{E}\left[(X - \mathbb{E}\,[X])^2\right].$$

Partendo dalla densità di X e usando $g(x) = (x - \mathbb{E}\,[X])^2$, si trova la formula

$$\mathrm{Var}\,(X) = \begin{cases} \sum_{x\in\mathbb{R}}(x - \mathbb{E}\,[X])^2 P(X = x) & \text{se } X \text{ ha densità discreta} \\ \int_{-\infty}^{\infty}(x - \mathbb{E}\,[X])^2 p(X = x)dx & \text{se } X \text{ ha densità continua.} \end{cases}$$

Esempi

Espressione alternativa per la varianza

Vale l'indentità

$$\mathrm{Var}\,(X) = \mathbb{E}\left[X^2\right] - (\mathbb{E}\,[X])^2.$$

Dimostrazione

Un altro esempio

Diseguaglianza di Chebyshev

Sia $X \in \mathbb{R}$ una variabile aleatoria. Allora per ogni costante $k > 0$, si ha

$$P(|X - \mathbb{E}[X]| > k) \leq \frac{\mathrm{Var}(X)}{k^2},$$

o, equivalentemente, per ogni $k \geq 1$,

$$P(\mathbb{E}[X] - k\sigma_X \leq X \leq \mathbb{E}[X] + k\sigma_X) \geq 1 - \frac{1}{k^2}.$$

- Una conseguenza è che, se $\sigma_X = 0$ la variabile X è constante con probabilità 1

Diseguaglianza di Chebyshev

Sia $X \in \mathbb{R}$ una variabile aleatoria. Allora per ogni costante $k > 0$, si ha

$$P(|X - \mathbb{E}[X]| > k) \leq \frac{\mathrm{Var}(X)}{k^2},$$

o, equivalentemente, per ogni $k \geq 1$,

$$P(\mathbb{E}[X] - k\sigma_X \leq X \leq \mathbb{E}[X] + k\sigma_X) \geq 1 - \frac{1}{k^2}.$$

- Una conseguenza è che, se $\sigma_X = 0$ la variabile X è constante con probabilità 1
- Informalmente scriviamo

$$X \approx \mathbb{E}[X] \pm \sigma_X$$

Dimostrazione

Standardizzazione

Data X, la variabile $X - \mathbb{E}[X]$ è centrata, $\mathbb{E}[(X - \mathbb{E}[X])] = 0$,

La *standardizzazione* X è la variabile

$$\hat{X} = \frac{X - \mathbb{E}[X]}{\sigma_X},$$

che è centrata $\mathbb{E}\left[\hat{X}\right] = 0$ e ha deviazione standard $\sigma_{\hat{X}} = 1$.

Esempi