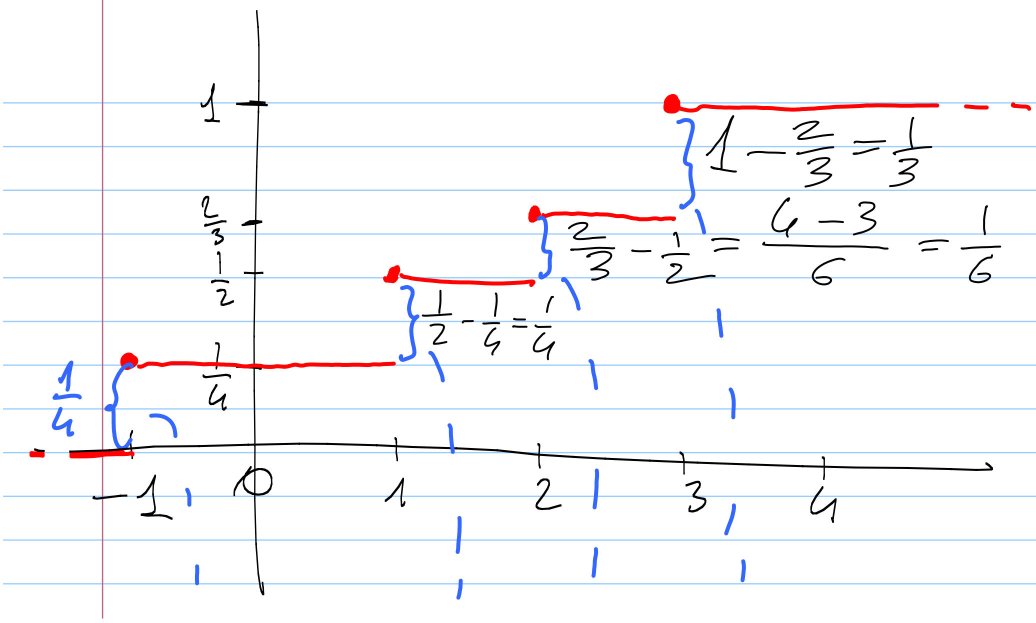
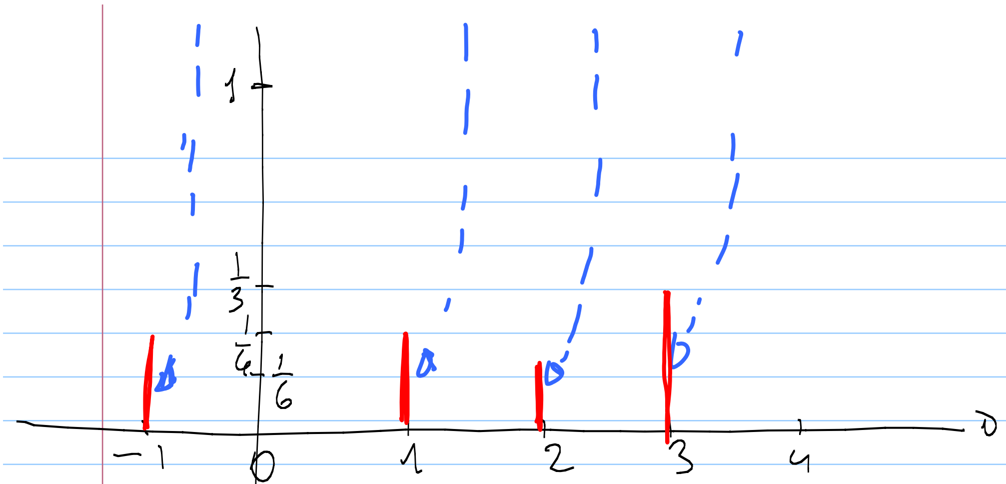
CdLP - TdD - XVIII - 29/10/2008

Titolo nota [redacted]

funzione di distrib. ⇒ densità discreta.

$$F(x) = \begin{cases} 0 & x \quad x < -1 \\ \frac{1}{4} & x \quad -1 \leq x < 1 \\ \frac{1}{2} & x \quad 1 \leq x < 2 \\ \frac{2}{3} & x \quad 2 \leq x < 3 \\ 1 & x \quad x \geq 3 \end{cases}$$





in tabelle

x_i	-1	1	2	3	Σ
$p(x_i)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{3}$	1

Scommessa

Lancio un dado e ricevo tanti euro quanti sono descritti dal dado.

Qual è il prezzo equo per partecipare alla scommessa?

Idea: Ripetiamo la scommessa molte volte e calcoliamo il guadagno medio

$$\text{guadagno medio} = \frac{\text{somma guadagni}}{\#\ \text{scommesse}}$$

$\#$ scommesse $= n$

$\#$ scommesse nelle quali vinco i euro $= n_i$

$i = 1, 2, 3, 4, 5, 6$

$$n_1 + n_2 + n_3 + n_4 + n_5 + n_6 = n$$

somma guadagnata =

$n_1 \cdot 1 + n_2 \cdot 2 + n_3 \cdot 3 + n_4 \cdot 4 + n_5 \cdot 5 + n_6\, 6$

\# scommesse in cui vinco 1 euro

euro vinti ad ogni scommessa

guadagno medio = $\dfrac{\text{somma guadagnata}}{\#\ \text{scommesse}}$ =

$= \dfrac{n_1 1 + n_2 2 + n_3 3 + n_4 \cdot 4 + n_5 5 + n_6\, 6}{n} =$

$= \dfrac{n_1}{n} \cdot 1 + \dfrac{n_2}{n} \cdot 2 + \dfrac{n_3}{n} 3 + \dfrac{n_4}{n} 4 + \dfrac{n_5}{n} 5 + \dfrac{n_6}{n} \cdot 6$

Secondo l'impostazione Frequentista delle Probabilità

$\dfrac{n_1}{n} = \dfrac{\#\ \text{volte che vinco 1}}{\#\ \text{scommesse totali}} \xrightarrow{n \to \infty} P(\text{ottenere 1 col dado})$

$P(1)$ = prob. di ottenere 1 col dado

Similmente per gli altri casi

$$\frac{n_i}{n} \xrightarrow[\text{per } n \to +\infty]{} P(i) = P(\text{ottenere } i \text{ col dado}) \qquad i = 2, 3, 4, 5, 6$$

guadagno medio =

$$\frac{n_1}{n} 1 + \frac{n_2}{n} 2 + \frac{n_3}{n} 3 + \frac{n_4}{n} \cdot 4 + \frac{n_5}{n} 5 + \frac{n_6}{n} 6$$

$$\xrightarrow[n \to +\infty]{} P(1) \cdot 1 + P(2) 2 + P(3) 3 + P(4) 4 + P(5) 5 + P(6) 6$$

Se il dado è regolare

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$$

dunque

$$\text{guadagno medio} = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6 =$$

$$= \frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6} = \frac{7}{2} = 3.5$$

prezzo equo

Il discorso può essere generalizzato a qualsiasi v.a. (discreta)

Definizione

$E(X)$ è il valore atteso di X ed è il prezzo equo da pagare in una scommessa nella quale vinciamo la quantità aleatoria X

Se X ha distribuzione

X	x_1	x_2	...	x_n
$p(x)$	$p(x_1)$	$p(x_2)$	....	$p(x_n)$

il valore atteso di X è

$$E(X) = x_1 p(x_1) + x_2 p(x_2) + \dots + x_n p(x_n)$$

$$= \sum_{i=1}^{n} x_i \, p(x_i)$$

Esempio: lancio di 3 monete

$X = \#$ teste

x	0	1	2	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$$E(X) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \frac{3}{8} + 3 \cdot \frac{1}{8} =$$

$$= \frac{1}{8}(0+3+6+3) = \frac{12}{8} = \frac{3}{2}$$

Esempio 3d. Una classe di 120 studenti viene condotta a teatro suddivisa in 3 autobus. Nel primo autobus ci sono 36 studenti, nel secondo 40 e 44 nel terzo. Quando gli autobus arrivano davanti al teatro, si sceglie a caso uno tra i 120 studenti. Se X denota il numero di studenti che hanno viaggiato sull'autobus dello studente scelto a caso, si calcoli $E[X]$.

tre autobus	b_1	b_2	b_3
capienze	36	40	44

Vogliamo calcolare la capienza media degli autobus

B = capienza di un autobus scelto a caso.

Se si procede mediante estrazione casuale dell'autobus, B ha distribuzione

b	36	40	44
$p(b)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

$$E(B) = \frac{1}{3}\cdot 36 + \frac{1}{3}\cdot 40 + \frac{1}{3}\cdot 44 = \frac{1}{3}(36+40+44)$$
$$= \frac{1}{3}\cdot 120 = 40$$

Tecnica differente.

Prendiamo uno studente (dei 120) a caso e gli/le chiediamo "quanti studenti ci sono nel tuo autobus?"

Questa tecnica produce gli stessi risultati della precedente?

X = risposta data dallo studente

Devo verificare

$$E(X) \overset{?}{=} E(B)$$

$$X \in \{36, 40, 44\}$$

trovo la distribuzione

$$P(X=36) = \frac{\text{studenti nell'autobus } b_1}{\text{totale studenti}}$$

$$= \frac{36}{120} = \frac{3}{10}$$

$$P(X=40) = \frac{\text{studenti in } b_2}{\text{totale}} = \frac{40}{120} = \frac{1}{3}$$

$$P(X=44) = \frac{\text{studenti in } b_3}{\text{tot}} = \frac{44}{120} = \frac{11}{30}$$

X ha distribuzione

x	36	40	44
$p(x)$	$3/10$	$1/3$	$11/30$

$$E(X) = \frac{3}{\cancel{10}_{5}} \cdot \cancel{36}^{18} + \frac{1}{3} \cdot 40 + \frac{11}{\cancel{30}_{15}} \cdot \cancel{44}^{22}$$

$$= \frac{162 + 200 + 242}{15} = \frac{604}{15} \simeq 40.26 \neq 40$$

in conclusione

$$E(X) \neq E(B)$$ (anche se di poco)

Data una v.a. X sappiamo come calcolare $E(X)$

Spesso non siamo interessati direttamente a X, ma a una sua trasformazione

$$Y = f(X)$$

Come possiamo calcolare $E(Y)$?

Esempio test di misurazione della pressione ai pazienti di un ospedale

Il test ha 3 esiti:
- $+1$ pressione alta
- 0 pressione normale
- -1 pressione bassa

X ha dist.

x	-1	0	1
$p(x)$	0.24	0.48	0.28

$$E(X) = -1(0.24) + 0(0.48) + 1(0.28)$$
$$= -0.24 + 0.28 = 0.04$$

Ci interessa la classificazione valori normali/valori anomali

Ricorriamo alla trasformazione

$$Y = X^2$$

$X = -1 \Rightarrow Y = (-1)^2 = 1$ — valori anomali

$X = 0 \Rightarrow Y = (0)^2 = \boxed{0}$ val. normale

$X = 1 \Rightarrow Y = 1^2 = 1$

$E(Y) = ?$

Per calcolare $E(Y)$ dobbiamo conoscere la distribuzione di Y

$P(Y=0) = P(X=0) = 0.48$

$P(Y=1) = P(\{X=-1\} \cup \{X=+1\})$

$= P(X=-1) + P(X=1) = 0.24 + 0.28 = 0.52$

Y ha distribuzione

y	0	1
$P(y)$	0.48	0.52

$E(Y) = 0(0.48) + 1(0.52) = 0.52$

Da quello che sappiamo (finora)

se $Y = f(X)$

per calcolare $E(Y)$ dovremo

1) calcolare la dist. di Y

2) calcolare $E(Y)$ dalla sua distribuzione

In realtà esiste una scorciatoia

Proposizione 4.1

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

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

Esempio

x	-1	0	1
$p(x)$	0.24	0.48	0.28

Per calcolare $E(Y) = E(X^2)$

invece di calcolare la distribuzione di Y e poi $E(Y)$

Posso applicare direttamente la Proposizione 4.1, che ci permette di scrivere

x	-1	0	1
$p(x)$	0.24	0.48	0.28

$$E(Y) = E(X^2) =$$

$$= (-1)^2(0.24) + 0^2(0.48) + 1^2(0.28) =$$

$$1(0.24) + 0(0.48) + 1(0.28) = 0.24 + 0.28 = 0.52$$

Caso particolare: trasformazioni lineari

$$f(X) = aX + b$$

Es. X = temperatura massima di Pescara per il 1° novembre (in gradi centigradi)

Y = stessa temp. in gradi FAHRENHEIT

$$Y = \frac{9}{5}X + 32$$
$$aX + b$$

Corollario 4.1

Se a e b sono costanti, allora

$$E[aX + b] = aE[X] + b$$

In questo caso specifico calcolo $E(Y)$ direttamente da $E(X)$

Esempio se $E(X) = 10$

$$E(Y) = \frac{9}{5}E(X) + 32 = \frac{9}{\cancel{5}}\cancel{10}^{\,2} + 32$$
$$= 18 + 32 = 50$$

Oss Questa scorciatoia vale solo per trasformazioni lineari

Esempio precedente

X	-1	0	1
$p(x)$	0.24	0.48	0.28

se voglio calcolare $E(Y)$ con $Y = X^2$

non posso prendere $E(X)$ e fare il quadrato.

$$E(X) = 0.04$$

$$(E(X))^2 = (0.04)^2 = 0.0016$$

$$\underset{0.52}{E(X^2)} \neq \underset{0.0016}{[E(X)]^2}$$

Titolo nota [redacted]

$E(X)$ valore atteso della v.a. X

dà un'idea della grandezza dei valori assunti dalla v.a. X

$E(X)$ non dà nessuna idea sul fatto che i valori di X siano dispersi o concentrati

Esempio: 3 v.a. con lo stesso valore atteso.

$W = 0$ con probabilità 1.

w	0
$p(w)$	1

$EW = 0 \cdot 1 = 0$

$$Y = \begin{cases} -1 & \text{con prob. } \frac{1}{2} \\ +1 & \text{con prob } \frac{1}{2} \end{cases}$$

y	-1	1
$p(y)$	$\frac{1}{2}$	$\frac{1}{2}$

$$E(Y) = -1\left(\frac{1}{2}\right) + 1\left(\frac{1}{2}\right) = -\frac{1}{2} + \frac{1}{2} = 0$$

$$Z = \begin{cases} -100 & \text{con prob } \frac{1}{2} \\ +100 & \text{con prob } \frac{1}{2} \end{cases}$$

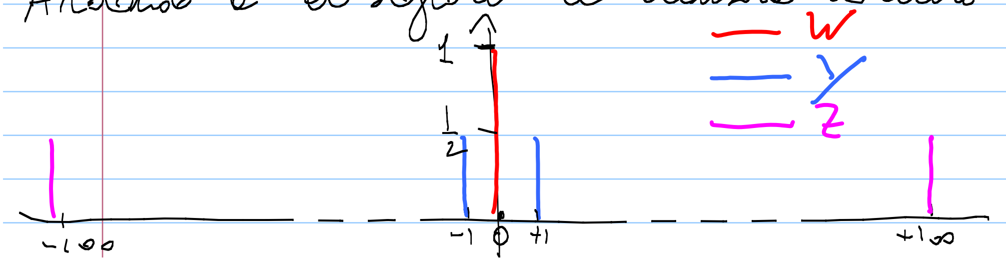
z	-100	100
$p(z)$	$\frac{1}{2}$	$\frac{1}{2}$

$$E(Z) = -100 \cdot \frac{1}{2} + 100 \frac{1}{2} = -50 + 50 = 0$$

quindi

$$E(W) = E(Y) = E(Z) = 0$$

Andando a disegnare le densità discrete



Come possiamo misurare la dispersione dei valori di una v. a.?

IDEA: misura della dispersione attorno ad un valore centrale μ

Come "valore centrale" consideriamo

$$\mu = E(X)$$

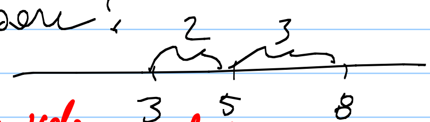
dispersione = distanza fra X e μ $\quad d(X, \mu)$

X v.a $\Rightarrow d(X, \mu)$ v.a

con una media di dispersione $E(d(X,\mu))$

Quale distanza usare?

1^o proposta: differenza **in valore assoluto**



$x = 8 \quad \mu = 5 \qquad d(x,\mu) = d(8,5) = 8-5 = 3$

$x = 3 \quad \mu = 5 \qquad = d(x,\mu) = d(3,5) = |3-5| =$

$|-2| = +2$

1^o proposta $\quad d(a,b) = |a-b|$

Problema: definizione intuitiva.

Tuttavia è di difficile trattazione dal punto di vista matematico.

Più facile considerare

$$d(a,b) = (a-b)^2$$

$d(8,5) = (8-5)^2 = 3^2 = 9 \qquad d(3,5) = (3-5)^2 = (-2)^2 = 4$

idea meno intuitiva, ma di più facile trattazione matematica.

2° proposta (definitiva) $d(a,b) = (a-b)^2$

$$E\left(d(X,\mu)\right) = E\left((X-\mu)^2\right) =$$

$$= E\left((X-E(X))^2\right)$$

la quantità è detta **VARIANZA** di X