Variabili aleatorie discrete / 17-03

$X : S \to X \subseteq \mathbb{R}$

mappa uno spazio campionario in $\mathbb{R}$

→ lancio due dadi e guardo la somma delle facce uscite
- spazio degli eventi sono tutte le coppie
- X e' la funzione che somma i due numeri e restituisce un numero

Si prende X un insieme di cardinalita' N $\quad X = \{x_1, \dots, x_n\}$
X eredita in qualche modo le probabilita' in S
→ discreta perche' comunque il risultato e' un oggetto (come: numeri interi) infiniti ma senza frazioni

Esempi

1) ~~X~~ teste in 3 lanci di una moneta

casi possibili: $\{0, 1, 2, 3\}$ teste

$S = \{T,T,T\}, \{T,T,C\}, \{TCC\}, \{C,T,T\}, \{CTC\}, \{CCT\}$
$\{TCT\}, \{CCC\}$

$\underset{\text{caso zero teste}}{P(X=0)} = \frac{1}{8} \quad P(X=1) = \frac{3}{8} \quad P(X=2) = \frac{3}{8} \quad P(X=3) = \frac{1}{8}$

2) Estraggo 3 palline da 20 numerate da 1 a 20
→ X è il numero più grosso estratto fra le 3 palline

$$P(X=i) = \frac{\binom{i-1}{2}}{\binom{20}{3}}$$

→ ovvero i modi di estrarre le altre due palline rimanenti sotto quel numero (escludendo il numero pescato che è il più grande)

i ↓ numero estratto

Funzione di PMF e CDF

$p(X=i)$ viene indicata $p(i)$ (o $p(x_i)$) con $i = 1, \dots, N$

$$\sum_{i=1}^{N} p(x_i) = 1 \quad \text{dove} \quad 0 \le p(i) \le 1$$

PROBABILITÀ DI MASSA (PMF)

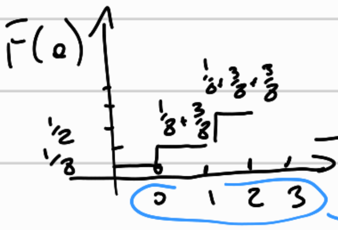
PMF è la probabilità dei valori assunti dalla variabile aleatoria

Una definizione equivalente è la PROBABILITÀ CUMULATA (CDF)

Immaginando di ordinare i valori t.c. $x_1 < x_2 \dots < x_i < \dots$

$$F(a) = \sum_{x_i \le a} p(x_i)$$

fissato un a, ad ogni aggiunta di p salgo di valori (a grande = prendo più valori)



→ a vale 1 (una testa)

→ a vale 0 (zero teste)

→ valori di a (var. aleatorie X)

(es. monete)

Valore atteso di una variabile aleatoria (media var aleatoria)

$$\mu = E[X] = \sum_{i=1}^{\tilde{N}} x_i \cdot p(x_i)$$

valori distinti senza ripetizioni

sommatoria dei valori PESATI secondo la loro probabilita'

(un po' come la media pesata)

Esempio (monete 1))

$$E[X] = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{3}{2} = 1,5$$

in media

media attesa della variabile aleatoria

"dei 3 lanci e' probabile che io ottenga una testa e mezzo"

Valore atteso di una funzione di variabile casuale

$$\{x_1, \dots, x_n\} \rightarrow \{y_1, \dots, y_n\} \qquad y = g(x_i)$$

es. monete 1))

$$g = |x|^2 \qquad \{0, 1, 4, 9\}$$

Se volessi calcolare il valore atteso della funzione?

$$E(g(x)) = \sum_{i=1}^{M} g(x_i) \cdot P(x_i)$$

$$E(g(x)) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 4 \cdot \frac{3}{8} + 9 \cdot \frac{1}{8}$$

più probabili, allora' un valore fra questi due

Varianza

Data la media $E[X] = \sum_{i=1}^{N} x_i p(x_i)$ e se volessi prendere un secondo valore oltre la media?

Non so la 'posizione' originaria, ma posso esprimere la distanza dal primo trovato

→ La VARIANZA (in media quanto i valori si discostano dal valor medio)

→ media delle distanze grande

→ media delle distanze piccola, qui la media esprime meglio la posizione "media" di tutti i valori rispetto alla posizione attuale

→ altrimenti viene zero

→ mi evito il modulo e positivizzo

Varianza: $$\sum_{i=1}^{N} (\underbrace{x}_{\text{valore}} - \underbrace{E[X]}_{\text{rispetto media}})^{2} \cdot \underbrace{p(i)}_{\text{per pesare}}$$

(non ho differenze se tutte le distanze sono al quadrato)

Deviazione standard $$sd(x) = \sqrt{Var(x)}$$

(valore $K\boxed{\pm X}$) di quanto varia

Proprietà

• LINEARITÀ: $g(x) = aX + b$

$p(x_i)$ p relativa a x_i

$$E[aX+b] = \sum_{i=1}^{N} (ax_i + b) \cdot \widehat{p(i)}$$

$$= a \sum_{i=1}^{N} \underbrace{x_i\, p(i)}_{\text{media}} + b \underbrace{\sum_{i=1}^{N} p(i)}_{\text{fa } 1}$$

$$= a\, E[X] + b$$

• riscrittura varianza

$$\sum_{i=1}^{n} (x - \overset{\mu}{E[X]})^2 \cdot p(i) = E[(x-\mu)^2]$$

$$= E[(x^2 - 2\mu X + \mu^2)]$$

poiché sommatorie posso dividerle

$$= E[x^2] + E[\mu^2] - 2E[\mu X]$$

numero μ^2 (come x b sopra) $\hookrightarrow a X$ " $-2\mu \cdot \underbrace{(\mu)}_{E[X]}$

espletazione

$$= E[x^2] - \mu^2 = \boxed{E[x^2] - (E[x])^2}$$

Questo vale per la varianza?

No, perché ho dei prodotti: $Var(aX+b) = a^2 Var(X)$

$$Var(aX+b) = E[(aX+b - E[aX+b])^2]$$

$$= E[(aX+b - aE[X] - b)^2]$$

$$= a^2 E[(X - E[X])^2] = a^2 Var(X)$$

poiché sposto tutto di b, faccio i conti con i nuovi valori, come risultato non ho un b

Esempi var aleatorie discrete

VAR DI BERNOULLI

• lancio moneta: S {0,1} $0 \le p \le 1$ che vengo un numero

$$\boxed{\begin{aligned} p(X=0) &= 1-p \\ p(X=1) &= p \end{aligned}}$$ generico

FORMULA DI Bernoulli

se $p \neq \frac{1}{2}$ allora moneta non onesta

(vale perché $\sum p_i = 1$)

↳ il valore non cambia se scambio, utilizzo valori generici

$$E(X) = 0 \cdot (1-p) + 1 \cdot p = p$$

Binomio di Newton

$(a+b)^2 = a^2 + b^2 + 2ab$

$(a+b)^3 = a^3 + b^3 + 3a^2b + 3ab^2$

$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}$

mi permette di calcolare le variabili di Bernoulli

La binomiale è una variabile i cui valori x_i sono $0, 1, \dots n$

e $p(i) = \binom{n}{i} p^i (1-p)^{n-i}$

$E[X] = np$ (perché $E[\sum X_i] = \sum E[X_i]$)

$Var(X) = np(1-p)$

$E[X_i]$ → Bernoulli $= p$

Per verificare che è valida $\sum p = 1$ e $0 \le p \le 1$

usando

$a = p \quad b = 1-p$ ma allora $(a+b)^n = (\cancel{p} + 1 - \cancel{p})^n = 1^n = 1$

quindi $\sum p = 1$

- serie geometrica $0 < a < 1$

$\sum_{i=0}^{n} a^i = \frac{1}{1-a}$

somma potenze

$\sum_{i=0}^{n} (1-a)^i = \frac{1}{1-(1-a)} = \frac{1}{a}$

- Variabile (distribuzione) geometrica

I valori sono infiniti: $1, 2 \dots \infty$ (no zero, sarebbe no lanci)

$p(i) = (1-p)^{i-1} p$

i lanci di monete, $1-p$ teste, p croci

ma p (croce) viene dopo

p che la prima testa arrivi dopo i lanci (cioè quanti lanci prima di avere testa)

Valida: $0 \le p \le 1$ ✓

$\sum p = 1$? $\longrightarrow$ $p \sum_{i=1}^{\infty} (1-p)^{i-1}$ $= p \sum_{J=0}^{\infty} (1-p)^J$

$i-1 = J$

è un numero, non influenzato dalla sommatoria

↳ quando $i=1$ $J=0$ (val iniziale) riscrivo gli indici

$\longrightarrow \cancel{p} \cdot \frac{1}{\cancel{p}} = 1$ ✓ (usando •)

- $E[X] = \sum_{i=1}^{\infty} i \cdot \underbrace{(1-p)^{i-1} p}_{p(i)} = p \sum_{i=1}^{\infty} (i-1+1)(1-p)^{i-1}$

x_i sono i numeri perchè in questo caso sono i numeri

$= p \sum_{i=1}^{n} (i-1)(1-p)^{i-1} + \underbrace{p \sum_{i=1}^{n} (1-p)^{i-1}}_{=1 \text{ per quanto sopra}}$

$\rightarrow$ rinomino gli indici $p \sum_{J=\cancel{0}\,1}^{\infty} J (1-p)^J + 1$ ma per $J=0$ il termine della sommatoria è zero

quindi faccio partire da 1 — posso togliere un fattore e portarlo fuori

$\longrightarrow (1-p)\, p \sum_{J}^{\infty} J (1-p)^{J-1} + 1$

↳ è la stessa formula iniziale

dunque $\longrightarrow E[X] = (1-p) E[X] + 1$

valore atteso di una geometrica

$\rightarrow \cancel{E[X]} - \cancel{E[X]} + p E[X] = 1 \longrightarrow \boxed{E[X] = 1/p}$

quindi in media lanciando una moneta 10 volte ho 1/10 di probabilità di fare testa

$$Var(X) = \frac{(1-p)}{p^2}$$

Distribuzione di Poisson

$$e^{a} = \sum_{i=0}^{\infty} \frac{a^i}{i!}$$

utile quando cerco
P di eventi successivi
conoscendo l'andamento medio

Variabile casuale di Poisson $\quad i = 0, 1, 2, \dots$

$$p(i) = \frac{\mu^i}{i!} e^{-\mu}$$

non dipende più da p
ma da un numero detto μ

Valida:

$$\sum p(i) = \underbrace{\sum_{i=0}^{\infty} \frac{\mu^i}{i!}}_{e^{\mu}} e^{-\mu} = e^{\mu} \cdot e^{-\mu} = 1 \quad \checkmark$$

$$E[X] = \sum_{\cancel{i=0}\ 1}^{\infty} \cancel{i} \frac{\mu^i}{\cancel{i!}\ \cancel{i}(i-1)!} e^{-\mu}$$

l'obiettivo è renderlo come la distribuzione di Poisson

$$= \sum_{i=1}^{\infty} \frac{\mu^i}{(i-1)!} e^{-\mu}$$

$$= \mu \sum_{i=1}^{\infty} \frac{\mu^{i-1}}{(i-1)!} e^{-\mu}$$

$Var(X)$

$\parallel$

pongo $j = i - 1 \quad \rightarrow \mu \underbrace{\sum_{j=0}^{\infty} \frac{\mu^j}{j!}}_{1} e^{-\mu} \longrightarrow E[X] = \mu$

valore atteso di una poissiana