Estensione delle proprietà viste in precedenza

ESTENSIONE DELLE DEF. E DELLE PROPRIETÀ VISTE IN PRECEDENZA AL CASO DI UN VETTORE DI N VARIABILI CASUALI

$$(X_1, X_2, \dots X_N)$$

Sono variabili casuali associate ad un esperimento (lanci di monete…). Ovviamente possono essere discrete o continue.

V.C. DISCRETE

$$X_K \in \{x_1, x_2 \dots x_{m_K}\} \qquad \begin{array}{l} K = 1, \dots N \\ m_K \in \mathbb{N} \setminus \{0\} \end{array}$$

Appartiene a un certo numero di valori, non sempre uguali per ogni esperimento.

Riconosciamo queste variabili come collegate e definiamo la funzione di massa congiunta.

FUNZIONE DI MASSA CONGIUNTA

$$p(a_1, a_2, \dots a_N) = P(X_1 = a_1, X_2 = a_2 \dots, X_N = a_N)$$

Le virgole rappresentano intersezioni, tutto deve avvenire contemporaneamente.

$$p : \mathbb{R}^N \to [0, 1]$$

Da qui si possono definire le funzioni marginali.

Poi abbiamo la proprietà:

$$\sum_{K_1=1}^{m_1} \sum_{K_2=1}^{m_2} \dots \sum_{K_N=1}^{m_N} p(x_{K_1}, x_{K_2} \dots x_{K_N}) = 1$$

FUNZIONE DI RIPARTIZIONE DI PROB. CONGIUNTA

$$F(a_1, a_2, \dots a_N) = P(X_1 \le a_1, X_2 \le a_2 \dots X_N \le a_N) \quad \forall a_1, a_2 \dots a_N \in \mathbb{R}$$

$$F: \mathbb{R}^N \longrightarrow [0,1]$$

PROPRIETÀ

1) $0 \le F(a_1, a_2, \dots a_N) \le 1$
2) $F(-\infty, -\infty, \dots, -\infty) = 0$
3) $F(+\infty, +\infty, \dots +\infty) = 1$

Ovviamente nel caso discreto la funzione è a tratti, quindi formerà degli iper-gradini negli spazi dimensionali.

VARIABILI CASUALI CONGIUNTAMENTE CONTINUE

Si dice congiuntamente perché il rischio di non dirlo si può pensare alle singole variabili casuali, non pensandole come un insieme. Ci deve essere la funzione congiunta, che riguarda tutte le variabili casuali insieme.

Def. $(X_1, X_2 \dots X_N)$ si dicono v.c. congiuntamente continue se esiste

$f: \mathbb{R}^N \longrightarrow \mathbb{R}^+ \cup \{0\}$ tale che $\forall$

$$\mathcal{H} \subset \mathbb{R}^N \quad P((X_1, X_2 \dots X_N) \in \mathcal{H}) = \int\int_{\mathcal{H}} \dots \int f(s_1, \dots s_N)\, ds_1 \dots ds_N$$

In linea di principio può essere un integrale a 10, 20, n dimensioni, dipende da come introduco l'esperimento e da come associo le variabili casuali a questo esperimento.

Valgono le solite proprietà:

PROPRIETÀ

- $f(a_1, a_2 \dots a_N) \ge 0 \quad \forall (a_1, a_2, \dots a_N) \in \mathbb{R}^N$
- $\displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(s_1, s_2 \dots s_N)\, ds_1 \dots ds_N = 1$

E la funzione di ripartizione congiunta con le solite proprietà.

FUNZIONE DI RIPARTIZIONE CONGIUNTA

$$F: \mathbb{R}^N \to [0,1] \quad F(a_1, a_2 \dots a_N) = P(X_1 \leq a_1, \dots X_N \leq a_N)$$

PROPRIETÀ SIMILI A QUELLE DEL CASO DISCRETO, CON L'UNICA DIFFERENZA

$$\frac{\partial^N F(a_1, a_2 \dots a_N)}{\partial a_1 \, \partial a_2 \dots \partial a_N} = f(a_1, a_2 \dots a_N)$$

Vale anche il caso inverso.

$$F(a_1, a_2 \dots a_N) = \int_{-\infty}^{a_1} \left(\int_{-\infty}^{a_2} \dots \left(\int_{-\infty}^{a_N} f(s_1, s_2 \dots s_N) \, ds_N \right) \dots ds_2 \right) ds_1$$

VARIABILI CASUALI INDIPENDENTI

$(X_1, X_2 \dots X_N)$ si dicono indipendenti se

$$\forall A_1, A_2, \dots A_N \subset \mathbb{R}$$

$$P(X_1 \in A_1, X_2 \in A_2 \dots X_N \in A_N) =$$
$$= P(X_1 \in A_1) \, P(X_2 \in A_2) \dots P(X_N \in A_N)$$

Questo non basta: deve valere l'indipendenza anche per i gruppi da N – 1, N – 2,... 2.

$$= P(X_1 \in A_1) \, P(X_2 \in A_2) \dots P(X_N \in A_N)$$

e lo stesso valga per gruppi di $(N-1)$, $(N-2)$ ..., 2 variabili

Più che sfruttarla per verificare l'indipendenza, si può usare quando viene dato che siano indipendenti. Quindi le probabilità o le funzioni di ripartizione possono essere scritte come prodotti.

Valor medio

È noto con moltissimi termini equivalenti, a seconda dell'autore o della traduzione.

VALORE ATTESO o VALOR MEDIO o MEDIA o MEDIA TEORICA o SPERANZA MATEMATICA

Media teorica perché teorizzata dalle variabili casuali utilizzate.

Tutte queste quantità in realtà sono una sola, il valor medio di una variabile casuale.

Def. Data una v.c. X (discreta o continua) si definisce valore atteso o valor medio, se esiste, la seguente quantità

$$E[X] \equiv \begin{cases} \displaystyle\sum_{k=1}^{m} x_k \, p(x_k) & \text{se } X \text{ v.c. discreta con } X \in \{x_1, \dots x_m\} \\ \displaystyle\int_{-\infty}^{+\infty} x f(x) \, dx & \text{se } X \text{ v.c. continua con } f(x) \text{ funzione di densità} \end{cases}$$

Se ho una variabile casuale discreta sommo il prodotto tra tutti i valori assunti e la relativa funzione di massa, una media pesata.

Nel caso continuo ho un integrale.

È importante il passaggio SE ESISTE, cioè se l'integrale o la serie converge. Se non convergono, non ha senso definire il valor medio (il valor medio non esiste se l'integrale o la serie divergono).

Bernoulliana

es.1 V.C. DI BERNOULLI o BERNOULLIANA

È molto semplice, ma Bernoulli tramite questa variabile ha dimostrato perché la definizione frequentista è compatibile con la definizione assiomatica.

$X \in \{0,1\}$ $\quad X = 1$ se si verifica evento A

$\quad X = 0$ se si " A^c

$$p(1) = P(X=1) = P(A) = p$$

$$p(0) = P(X=0) = P(A^c) = 1 - p = q$$

p e q fanno parte della stessa quantità.

C'è anche una notazione.

$$X \sim Be(p)$$

$\sim$ si comporta come; $Be(p)$ Bernoulliana di parametro p

Una volta che si conosce p, si conosce q e tutto quello che c'è da conoscere della funzione di massa di questa variabile.

Come si calcola il valor medio?

$$E[X] = \sum_{k=1}^{m} x_k\, p(x_k) = \cancel{0\, p(0)} + 1\, p(1) = p$$

Il valor medio vale come la probabilità che si verifichi l'evento A.

Si introduce come contatore di quello che è successo in una singola prova che viene ripetuta.

In generale, il valor medio non viene né 0, né 1. Quindi il valor medio può valere valori che la variabile casuale non assume.

Lo stesso si può vedere nel lancio del dado equilibrato.

Altri esempi

es. 2

$$Y \in \{1,2,3,4,5,6\} \quad \text{con } p(1)=p(2)=p(3)=$$

(Lancio dado equilibrato)

$$=p(4)=p(5)=p(6)=\frac{1}{6}$$

$$E[Y] = \sum_{k=1}^{m} x_k\, p(x_k) = 1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{6} +$$

$$+ 4\cdot\frac{1}{6} + 5\cdot\frac{1}{6} + 6\cdot\frac{1}{6} =$$

$$= \frac{21}{6} = \frac{7}{2} = 3.5$$

Il valor medio non è un valore esatto della variabile casuale, La variabile casuale assume valori con la stessa probabilità e il valor medio è il valore centrale, tra il più piccolo e il più grande.

In meccanica il valor medio può essere associato al baricentro: conoscendolo, non si conosce tutto, ma ci sono buone indicazioni per descrivere il moto. Lo stesso qui, per sapere tutto bisogna sapere la funzione di massa o di densità, ma si può avere un'idea di quello che ci si può aspettare se si ripete tante volte l'esperimento.

es. 3 X v.c. continua con

$$f(x) = \begin{cases} 3x^2 & \text{se } x \in [0,1] \\ 0 & \text{altrove} \end{cases}$$

Sappiamo che il valor medio sarà tra 0 e 1. Va calcolato.

$$E[X] = \int_{-\infty}^{+\infty} x f(x)\,dx = \cancel{\int_{-\infty}^{0} x\cdot 0\,dx} + \int_{0}^{1} x\, 3x^2\,dx +$$

$$+ \cancel{\int_{1}^{+\infty} x\cdot 0\,dx} = \left[3\frac{x^4}{4}\right]_0^1 = \frac{3}{4}$$

es. 4 $\quad X$ v.c. continua con

$$f(x) = \begin{cases} \frac{1}{2} & \text{se } x \in [1, 3] \\ 0 & \text{altrove} \end{cases}$$

Il fatto che la funzione di densità sia costante è l'equivalente dell'equiprobabilità del caso discreto, quindi ogni valore nell'intervallo è indifferente. Il valor medio sarà il punto centrale dell'intervallo.

$$E[X] = \int_{-\infty}^{+\infty} x \cdot f(x)\,dx = \cancel{\int_{-\infty}^{1} x \cdot 0\,dx} + \int_{1}^{3} x \cdot \frac{1}{2}\,dx + \cancel{\int_{3}^{+\infty} x \cdot 0\,dx} = \left[\frac{x^2}{4}\right]_1^3 = \frac{9-1}{4} = \frac{8}{4} = 2$$

Proprietà del valor medio

PROPRIETÀ DI $E[X]$

1) Data una v.c X e definita una seconda v.c.

$$Y = g(X)$$

$$E[Y] = E[g(X)] = \begin{cases} \displaystyle\sum_{k=1}^{m} g(x_k)\,p(x_k) & \text{se } X \text{ v.c. discreta con } X \in \{x_1, \dots, x_m\} \\ \displaystyle\int_{-\infty}^{+\infty} g(x) f(x)\,dx & \text{se } X \text{ v.c. continua con } f(x) \text{ sua funzione di densità di prob.} \end{cases}$$

A cosa serve? Pensiamo a un quadrato il cui lato sia una variabile casuale. L'area del quadrato si fa x al quadrato, quindi o una funzione g. Se dell'area voglio calcolare il valor medio, si usa questo procedimento.

2) Conseguenza di 1) nel caso in cui
$Y = \alpha X + \beta$ con $\alpha, \beta \in \mathbb{R}$ parametri

Perché fare una trasformazione lineare? In fisica per cambiare un'unita di misura (da m/s a km/h, da °C a °F).

Come cambia il valor medio?

$$E[Y] = \alpha E[X] + \beta$$

$\boxed{\text{DIM}}$ CASO DISCRETO

$X \in \{x_1, x_2, \dots x_m\}$ con $p(a) = P(X = a)$

$$E[Y] = E[\alpha X + \beta] = E[g(X)] =$$

$$\underset{1)}{=} \sum_{k=1}^{m} g(x_k)\, p(x_k) = \sum_{k=1}^{m} (\alpha x_k + \beta)\, p(x_k)$$

$$= \alpha \sum_{k=1}^{m} x_k\, p(x_k) + \sum_{k=1}^{m} \beta\, p(x_k) =$$

$$= \alpha \underbrace{\sum_{k=1}^{m} x_k\, p(x_k)}_{= E[X]} + \beta \underbrace{\sum_{k=1}^{m} p(x_k)}_{= 1} =$$

$$= \alpha E[X] + \beta$$

N.B. Se $\alpha = 0$ $E[\beta] = \beta$

3) Date due v.c. X e Y e definita $Z = h(X, Y)$, se esiste $E[Z]$ è

$$E[Z] = E[h(X,Y)] = \begin{cases} \displaystyle\sum_{k=1}^{m} \sum_{j=1}^{n} h(x_k, y_j)\, p(x_k, y_j) & \text{se } X \text{ e } Y \text{ discrete con } X \in \{x_1, \dots x_m\},\ Y \in \{y_1, \dots y_n\} \\ \displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(x,y)\, f(x,y)\, dx\, dy & \text{se } (X,Y) \text{ congiuntamente continue con funzione di densità congiunta } f(x,y) \end{cases}$$

Ovviamente se le le sommatorie o gli integrali convergono.

4) Date due v.c. X e Y (discrete o congiuntamente continue) con valori medi $E[X]$ e $E[Y]$ e definita $Z = X + Y$

$$E[Z] = E[X] + E[Y]$$

Con tutte queste proprietà abbiamo verificato che il valor medio è una funzione lineare.

$\boxed{\text{DIM}}$ CASO CONTINUO (IN MANIERA SIMILE SI DIM CASO DISCRETO)

(X, Y) con $f(x,y)$

$$E[Z] = E[h(X,Y)] = E[X+Y] \overset{3)}{=}$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(x,y)\, f(x,y)\, dx\, dy =$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x+y)\, f(x,y)\, dx\, dy =$$

$$= \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} x\, f(x,y)\, dy\right) dx + \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} y\, f(x,y)\, dx\right) dy$$

Si può fare per arbitrarietà (non cambia integrare prima in x e poi in y o viceversa).

$$= \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} x\, f(x,y)\, dy\right) dx + \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} y\, f(x,y)\, dx\right) dy =$$

$$= \int_{-\infty}^{+\infty} x \underbrace{\left(\int_{-\infty}^{+\infty} f(x,y)\, dy\right)}_{f_X(x)} dx + \int_{-\infty}^{+\infty} y \underbrace{\left(\int_{-\infty}^{+\infty} f(x,y)\, dx\right)}_{f_Y(y)} dy =$$

$$= \underbrace{\int_{-\infty}^{+\infty} x\, f_X(x)\, dx}_{\overset{\parallel}{E[X]}} + \underbrace{\int_{-\infty}^{+\infty} y\, f_Y(y)\, dy}_{\overset{\parallel}{E[Y]}}$$

$$= E[X] + E[Y] \qquad \square$$

N.B. $E[X]$ è lineare nel senso che

$$E[\alpha X + \beta] = \alpha E[X] + \beta$$

$$E[X + Y] = E[X] + E[Y]$$

C'è un'ultima proprietà, ricavata dalla quarta, quindi non serve la dimostrazione.

4bis) Date $X_1, X_2 \dots X_N$ N v.c.

$$E[X_1 + X_2 + \dots + X_N] = E[X_1] + E[X_2] + \dots + E[X_N]$$

Il valor medio della somma di tante variabili casuali è la somma dei valori medi delle variabili casuali.

es. 1

N lettere diverse devono essere spedite a N indirizzi corrispondenti

Le N lettere vengono inviate in maniera casuale spedendo ciascuna ad un indirizzo scelto a caso tra gli N

(anche con ripetizioni)

Quante lettere in media arriveranno all'indirizzo corretto?

È un problema surreale, ma è complicato da risolvere. Pensiamo dal punto di vista della singola lettera: ogni lettera finisce o all'indirizzo giusto o a quello sbagliato, non importa quale, è sbagliato.

Per ogni lettera introduciamo una v.c. di Bernoulli

$$X_K = \begin{cases} 0 & \text{altrimenti} \\ 1 & \text{se la } K\text{-esima lettera viene spedita al giusto indirizzo} \end{cases}$$

$$X_K \sim Be(p) \qquad p = \frac{1}{N}$$

Z = n° di lettere che vengono spedite al giusto indirizzo

$$Z = X_1 + X_2 + \dots + X_N$$

Infatti le variabili di Bernoulli servono da contatori (tutte all'indirizzo giusto Z è 1, tutte all'indirizzo sbagliato Z è 0).

La consegna dice di calcolare il valor medio di Z.

$$E[X_K] = p = \frac{1}{N}$$

$$E[Z] \underset{4bis)}{=} E[X_1] + E[X_2] + \dots + E[X_N] =$$

$$= \underbrace{\frac{1}{N} + \frac{1}{N} + \dots + \frac{1}{N}}_{N \text{ volte}} = 1$$

Indipendentemente dal numero di lettere e indirizzi. Sorprendente.