Variabili aleatorie

Mc128k

2015-11-30

Contenuti

Indice

1 Variabili aleatorie

In molti casi le uscite sperimentali sono fenomeni non rappresentati da numeri reali, fatto che rende più difficile il calcolo probabilistico; una variabile aleatoria fa in modo di prendere in ingresso (come dominio) le uscite sperimentali e associare ad ognuna di esse un numero reale (codominio).

Usando termini brutali, una variabile aleatoria è un modo per tradurre eventi in numeri, dato un evento mi dice quale numero reale corrisponde ad esso.

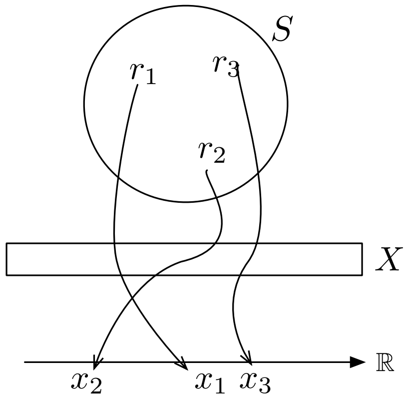


Figura 1: Variabile aleatoria X

$$X(r) = x \tag{1.1}$$

I valori r sono le **uscite sperimentali**, che attraversando la variabile aleatoria X portano alle **realizzazioni** x. Da distinguere X grande da x piccolo. Quando si esegue un esperimento si ha quindi una sola uscita sperimentale, che porta ad uno specifico numero reale. Ovviamente per uno stesso esperimento è possibile definire diverse variabili aleatorie che portano a diversi risultati a seconda delle esigenze.

$$X(r_i) = x_i \tag{1.2}$$

Si indicano tutte le uscite sperimentali per cui una condizione è verificata con:

$$X \leq x \tag{1.3}$$

Mentre si può indicare un evento direttamente con il risultato della variabile aleatoria:

$$\{X = x\} \tag{1.4}$$

1.1 Funzione di distribuzione (CDF)

Detta anche CDF (Cumulative Distribution Function) o funzione di ripartizione, è una funzione della variabile reale x (quella che esce dalla variabile aleatoria) che indica la probabilità "cumulativa" di un possibile evento.

$$F_x(x) := P\{X \leq x\} \tag{1.5}$$

Esempio 1.1.
Una moneta truccata può fare uscire un risultato con più probabilità rispetto all'altro. Lo spazio campione è dato da Testa o Croce:

$$S = \{T, C\}$$

Le probabilità (essendo due eventi mutualmente esclusivi) sono:

$$P\{T\} = p$$
$$P\{C\} = 1 - p$$

Si definisce una variabile aleatoria per "tradurre" testa o croce in numeri reali:

$$X(T) = 1$$
$$X(C) = 0$$

Questo significa che se esce testa la variabile aleatoria varrà 1, altrimenti 0.

Il diagramma di una CDF di una moneta truccata può essere quello raffigurato in fig.2, dove l'ascissa rappresenta la uscita della variabile aleatoria ($P\{X \leq x\}$) e l'ordinata rappresenta la probabilità che un certo evento avvenga.

Per capire come funziona il grafico bisogna notare alcune caratteristiche:

- La probabilità per definizione non può mai essere maggiore di 1, dato che a quel punto dice che "l'evento X considerato succederà sicuramente"
- Se ci si mette nel punto $x = 2$ si nota che la probabilità che la uscita sperimentale sia minore o uguale di x è 1, dato che in questo caso (esce sempre testa oppure croce) la variabile aleatoria è maggiore di 0 (croce) e maggiore di 1 (testa).

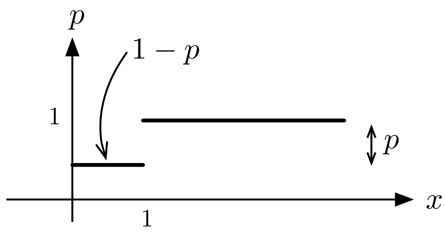


Figura 2: CDF di una moneta

- Considerando un punto X compreso fra 0 e 1 si ha che la probabilità che esca un risultato minore del valore X fissato è data da $1-p$, dove p rappresenta la probabilità di avere testa oppure croce.
- Essendo la moneta truccata si ha che p non è esattamente 0.5 ma risulta maggiore la possibilità che esca testa.
- Il grafico in conclusione indica, dato un punto X sull'asse x, qual è la probabilità che si verifichino eventi che attraversando la VA generano un numero minore o uguale del punto X.
- La variabile aleatoria considerata è detta **discreta** dato che varia con discontinuità (non ci sono punti intermedi fra testa e croce).

Esempio 1.2.
Si prenda in considerazione un dado colorato, la variabile aleatoria scelta assegna ad ogni colore un numero per poter rappresentare le uscite sperimentali sull'asse dei reali:

- Giallo - 10
- Verde - 15
- Rosso - 21
- Blu - 22
- Arancio - 10
- Viola - -12

Non essendo il dado truccato, la probabilità di avere ciascuna uscita sperimentale è uguale a $\frac{1}{6}$, il grafico della CDF risulta quindi a gradini equidistanti, tranne nel punto dove due valori sono uguali a 10 (fig.3).

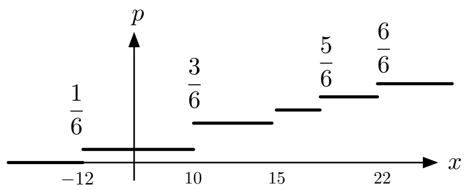


Figura 3: CDF del dado colorato

Esempio 1.3.
Una telefonata arriva a caso fra 0 e T, si ha quindi uno spazio continuo uniforme. Ci si chiede quale sia la probabilità che la telefonata arrivi in un tempo compreso fra t_1 e t_2.

$$P\{\text{telefonata arriva in } (t_1, t_2)\} = \frac{t_2 - t_1}{T}$$

La realizzazione della variabile aleatoria risulta esattamente uguale all'istante di arrivo, dato che si ha già un numero reale non serve trasformarlo:

$$x(t) = t$$

La CDF si compone:

$$F_x(x) := P\{X \leq x\}$$

Essendo uno spazio campione uniforme, ogni valore del tempo t ha una probabilità infinitesima che possa avvenire la chiamata, quindi la CDF non può essere a gradini.

$$P\{X \leq x\} = \frac{x}{T} \quad 0 < x < T$$

La probabilità che la telefonata arrivi fra x_1 e x_2 è data dallo spostamento *verticale* sul grafico, partendo da $f(x_1)$ e arrivando fino a $f(x_2)$.

1.1.1 Proprietà

1. $F_x(+\infty) = \lim_{x \to +\infty} F_x(x) = 1$
2. $F_x(-\infty) = \lim_{x \to +\infty} F_x(x) = 0$
3. F_x è monotona e debolmente crescente in ogni caso

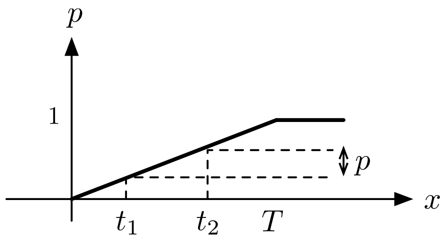


Figura 4: CDF della telefonata

4. $P\{X \leq x\} = 1 - F_x(x)$
5. $P\{X \leq n\} + P\{X > n\} = P\{S\} = 1$
6. $P\{x_1 < x \leq x_2\} = F_x(x_2) - F_x(x_1)$ se $x_1 \leq x_2$
7. $P\{X = x_0\} = F_x(x_0^+) - F_x(x_0^-)$
8. F_x è sempre continua a destra

Inoltre la variabile aleatoria è continua se F_x è **continua** per qualunque x, mentre è **discreta** se la funzione è costante a tratti (gradini). Può anche essere **mista** se presenta entrambe le caratteristiche.

1.2 Densità di probabilità (PDF)

La densità di probabilità (Probability Density Function) è semplicemente data dalla derivata della CDF, indica quindi "come cresce" l'altra funzione.

$$f_x(x) := \frac{d}{dx} F_x(x) \tag{1.6}$$

La PDF fa anche in modo di visualizzare in modo immediato dove sono "condensate" la maggior parte delle probabilità, per esempio il punto di massimo di una curva gaussiana indica che i risultati della V.A. sottostanti sono relativamente più probabili degli altri.

1.2.1 Proprietà

1. $f_x(x) \geq 0$

2. $\int_{x_1}^{x_2} f_x(x)dx = F_x(x_2) - F_x(x_1) = P\{x_1 < X < x_2\}$
3. $\int_{-\infty}^{+\infty} f_x(x)dx = F_x(+\infty) - F_x(-\infty) = 1$ (normalizzazione)
4. Una CDF a tratti ha una PDF definita da delta di Dirac ogni volta che si incontra un gradino; **l'area di ogni delta** è uguale alla altezza della "alzata" di ogni gradino (vedi masse di probabilità).
5. Ad un punto di flesso a tangente verticale nella CDF corrisponde un asintoto verticale nella PDF.

Per la seconda proprietà si può quindi calcolare la probabilità che un evento sia compreso fra due numeri x_1 e x_2, data la PDF (fig.5):

$$P\{x_1 < x \leq x_2\} = \int_{x_1}^{x_2} f_x(x)dx \tag{1.7}$$

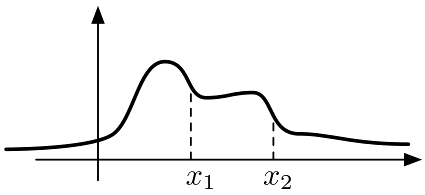


Figura 5: Calcolo probabilità da PDF

In questo modo si approssima la probabilità come se si stesse ripetendo infinite volte l'esperimento (frequenza relativa).

1.3 Funzione di una V.A.

Data una variabile aleatoria X, si può utilizzare una funzione $Y : \mathbb{R} \to \mathbb{R}$ che prende i valori in uscita da essa e produce nuovi valori. Viene quindi creata una nuova variabile aleatoria Y che prende in entrata i valori in uscita da X, $x_1, \dots x_n$ e produce nuovi valori $y_1, \dots y_n$. Ulteriori dettagli nella sezione 8.

$$S\{r_1, \dots r_n\} \to X \to \{x_1, \dots x_n\} \to Y \to \{y_1, \dots y_n\} \tag{1.8}$$

1.4 Massa di probabilità

Una variabile aleatoria potrebbe esprimere la probabilità che un evento avvenga solo se è un certo punto definito, altrimenti non può avvenire. In questo caso si hanno diverse "masse" identificate da delta di Dirac che indicano la probabilità che quell'evento discreto possa verificarsi. Le masse sono concentrate in singoli punti, al di fuori dei quali non si hanno probabilità.

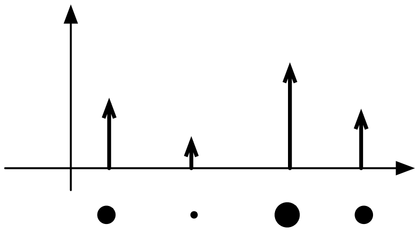


Figura 6: Massa di probabilità

$$f_x(x) : \mathbb{R} \to \mathbb{R}^+ \tag{1.9}$$

$$f_x(x_i) = P(X = x_i) := \begin{cases} P(X = x_i) = p_i, & x_i \in S_x \\ 0, & x_i \notin S_x \end{cases} \tag{1.10}$$

2 Densità notevoli

2.1 Uniforme

$$f_x(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{altrove} \end{cases} \tag{2.1}$$

2.2 Esponenziale negativa

$$f_x(x) = \lambda e^{-\lambda x} u(x), \quad \lambda > 0 \tag{2.2}$$

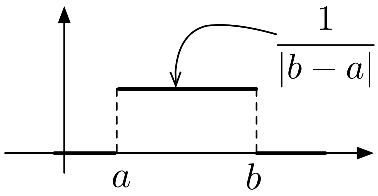


Figura 7: Densità uniforme

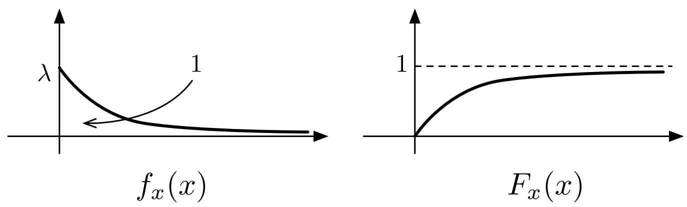


Figura 8: Esponenziale negativa

La esponenziale negativa rappresenta bene le vita di un componente (date certe condizioni, per esempio senza considerare l'usura).

Quando viene dato un valore medio η bisogna porre $\lambda = \frac{1}{\eta}$.

Esempio 2.1.
Si accende una lampadina in un istante t_0, e si vuole sapere fra quanto si brucia. È possibile fare uso di una variabile aleatoria che indica l'istante in cui la lampadina si spegne. La V.A. che si definisce ha lo stesso valore della uscita sperimentale (tempo), e ha **densità** del tipo esponenziale. Per esempio se si prende un valore $\bar{t}$ la probabilità che la lampadina **si spenga** *prima di* $\bar{t}$ è data dall'area sottesa dal grafico da 0 a t, di conseguenza la probabilità che si rompa aumenta asintoticamente man mano che si considera un tempo maggiore.

Potrebbe sembrare che le lampadine "vogliano" bruciarsi presto piuttosto che tardi, ma non è così, dato che bisogna considerare il caso in cui le lampadine non si brucino per usura, ma solo per cause accidentali.

Considerando le prove ripetute si ha che dato un numero finito di lampadine, all'inizio se ne bruciano di più perchè sono in numero maggiore, mentre man mano che si va avanti ce ne sono sempre di meno e se ne bruciano un numero minore.

2.3 Gaussiana

La funzione gaussiana (detta anche variabile aleatoria normale) esprime con precisione numerosi fenomeni di distribuzione naturali. Dipende da due parametri, σ_x e η.

$$f_x(x) = \mathcal{W}(\eta_x; \sigma_x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \cdot e^{-\frac{(x-\eta_x)^2}{2\sigma_x^2}} \tag{2.3}$$

Il parametro σ_x^2 è detto **varianza** della variabile aleatoria, mentre σ_x è la **deviazione standard**, indica quanto è larga la "campana" definita dalla funzione. La deviazione standard indica infatti quanto la distribuzione di probabilità può "deviare" (allontanarsi) dallo "standard" (il valore medio, la punta della campana). η_x indica il **valore medio**, quindi lo spostamento sulla ascissa (fig.9).

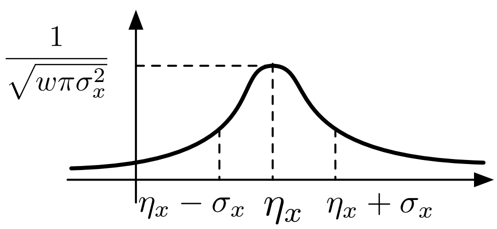


Figura 9: Distribuzione normale

Una particolarità da notare è che se si fa tendere $\sigma_x \to 0$ si ottiene un delta di Dirac centrato in η_x.

Si può costruire una approssimazione di una gaussiana ponendo una variabile aleatoria X con $\mathcal{W}(0, 1)$.

$$Q(x) = \int_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \tag{2.4}$$

Q fa in modo di approssimare bene la "coda" di una gaussiana centrata nello zero, ammesso di prendere un valore $x \geq 3$:

$$F_x(x) = P\{X \leq x\} = 1 - Q\left(\frac{x - \eta_x}{\sigma_x}\right) \tag{2.5}$$

$$Q(x) \lesssim \frac{1}{\sqrt{2\pi}} \frac{e^{-\frac{x^2}{2}}}{x} \tag{2.6}$$

Inoltre i punti di flesso in cui la derivata seconda è nulla ($\frac{d^2}{dx^2} = 0$) sono individuati da $\eta_x - \sigma_x$ e $\eta_x + \sigma_x$.

Esempio 2.2.
Sia X la v.a. che definisce l'istante di morte di un componente. La funzione di distribuzione di probabilità è data da:

$$f_x(t) = \begin{cases} 3 \cdot 10^-9t^2(100 - t)^2, & 0 \leq t \leq 100gg \\ 0, & altri \end{cases}$$

Ci si chiede quale sia la probabilità che la morte sia compresa fra il 60° e il 70° giorno:

$$P\{60 < t < 70\} = \int_{60}^{70} f_x(t)dt \simeq 0.154 = 15\%$$

3 Variabili aleatorie discrete

Indica le probabilità relative ad un evento che può solo verificarsi o non verificarsi.

3.1 Variabile di Bernoulli

$$X = \{0, 1\} \tag{3.1}$$

$$P\{x = 0\} = P_0 \tag{3.2}$$

$$P_1 = 1 - P_0 \tag{3.3}$$

La variabile di Bernoulli può essere utile come indicatore di un evento, quindi dato un evento in uno spazio campione esso si può verificare (P_1) o non verificare (P_0) con una certa probabilità. Le uscite sperimentali dentro l'evento che si vuole verificare hanno valore 1, si può quindi esprimere come $P\{x = 1\} = P\{A\}$.

Considerando ora un esperimento formato da prove ripetute di un evento modellabile con una variabile di Bernoulli, lo si ripete n volte individuando successi e insuccessi. Il numero di successi più probabile su n prove è individuato da:

$$p := P\{successo\} \tag{3.4}$$

$$P\{x = i\} = \binom{n}{i} p^i (1-p)^{n-i}, \quad i = 0, 1, \dots n \tag{3.5}$$

3.2 Geometrica

Sia i il numero delle prove in cui si ha il primo successo, la variabile aleatoria può assumere infiniti valori; la probabilità è data da:

$$P\{x = i\} = P(1-p)^{i-1}, \quad i = 1, 2, \dots \infty \tag{3.6}$$

3.3 Distribuzione di Poisson

$$P\{x = i\} = e^{-\lambda} \cdot \frac{\lambda^i}{i!} \tag{3.7}$$

È una distribuzione di probabilità discrete che esprime la probabilità di avere un risultato sapendo che mediamente è uguale a λ. In pratica si ha una sorta di campana simile alla gaussiana centrata nella media di dimensioni che variano.

4 Teoremi applicabili

4.1 Densità di probabilità condizionata

La formula della probabilità condizionata è data da:

$$P(A|M) = \frac{P(AM)}{P(M)}, \quad P(M) \neq 0 \tag{4.1}$$

Dato l'evento condizionante M si può definire la funzione di distribuzione come:

$$F_x(x|M) := P\{X \leq x|M\} = \frac{P\{X \leq x, M\}}{P(M)} \tag{4.2}$$

Per esempio, dati gli insiemi in fig.10 e la variabile aleatoria associata, si ha che tutti i valori minori o uguali a x sono collegati a r_1, r_2, r_3, ma se si condiziona la scelta a M si ha solo r_1, r_2.

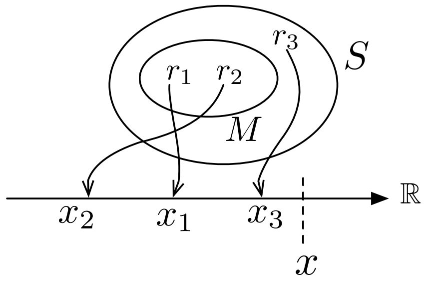


Figura 10: Probabilità condizionata

Di conseguenza vale che:

- $F_x(+\infty|M) = 1$
- $F_x(-\infty|M) = 0$

Volendo applicare il condizionamento alla probabilità di un evento delimitato da estremi:

$$P\{x_1 < x \leq x_2\} = F_x(x_2) - F_x(x_1) \tag{4.3}$$

$$f_x(x|M) := \frac{d}{dx}F_x(x|M) \tag{4.4}$$

$$P\{x_1 < x \leq x_2|M\} = \int_{x_1}^{x_2} f_x(x|M)dx \tag{4.5}$$

L'evento M deve essere espresso tramite la stessa incognita x:

$$M = \{b < X \leq a\} \tag{4.6}$$

$$F_x(x|M) = \frac{P\{X \leq x, b < x \leq a\}}{P\{b < x \leq a\}} \tag{4.7}$$

Se (facendo riferimento a fig.11) si considera l'evento condizionante in un intervallo si può calcolare facilmente la probabilità cercata.

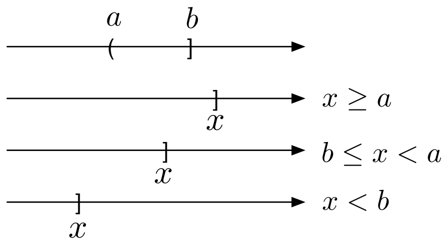


Figura 11: Eventi condizionati

$$x \geq a \implies \frac{P\{X \leq x, b < x \leq a\}}{P\{b < x \leq a\}} = 1 \tag{4.8}$$

$$b \leq x < a \implies \frac{P\{b < X \leq x\}}{P\{b < x \leq a\}} = \frac{F_x(x) - F_x(b)}{F_x(a) - F_x(b)} \tag{4.9}$$

$$x < b \implies \emptyset, P = 0 \tag{4.10}$$

4.2 Probabilità totali

Si vuole estendere l'uso di funzioni di distribuzione al teorema delle probabilità totali. Partendo da una partizione A_i di uno spazio S, la probabilità di un evento B viene costruita sommando i singoli contributi:

$$P(B) = \sum_i P(B|A_i) \cdot P(A_i) \tag{4.11}$$

L'evento viene espresso con una variabile aleatoria, si può quindi trovare la CDF sommando in modo simile:

$$F_x(x) = P\{X \leq x\} = \sum_i P\{X \leq x|A_i\} \cdot P(A_i) \tag{4.12}$$

$$F_x(x) = \sum_i F_x(x|A_i) \cdot P(A_i) \tag{4.13}$$

$$\tag{4.14}$$

Derivando si ottiene la PDF (densità):

$$\frac{d}{dx} F_x(x|A_i) = f_x(x|A_i) \tag{4.15}$$

$$f_x(x) = \sum_i f_x(x|A_i) \cdot P(A_i) \tag{4.16}$$

4.3 Formula di Bayes mista

Dato un evento M e una V.A. X, si può applicare la formula di Bayes. La probabilità di un evento M condizionata dal fatto che una evento è verificato dalla V.A. è quindi:

$$P(M|X \leq x) = \frac{P(X \leq x|M) \cdot P(M)}{P(X \leq x)} = \frac{F_x(x|M)P(M)}{F_x(x)} \tag{4.17}$$

Essendo la variabile aleatoria continua, per dare un significato a $P(M|X = x)$ si passa al limite:

$$P(M|X = x) = \lim_{\varepsilon \to 0} P(M|(x - \varepsilon < X \leq x) \tag{4.18}$$

$$= \lim_{\varepsilon \to 0} \frac{P((x - \varepsilon) < X \leq x|M)P(M)}{P((x - \varepsilon < X \leq x)} \tag{4.19}$$

$$= \lim_{\varepsilon \to 0} \frac{[F_x(x|M) - F)X(x - \varepsilon|M)]P(M)}{F_x(x) - F_x(x - \varepsilon)} \tag{4.20}$$

Dividendo numeratore e denominatore per ε, si riconosce un rapporto incrementale, che identifica quindi la derivata di $F_x(x)$.

$$P(M|X = x) = \frac{f_x(x|M) \cdot P(M)}{f_x(x)} \tag{4.21}$$

È possibile anche ottenere la probabilità totale:

$$P(M|X = x)f_x(x) = f_x(x|M) \cdot P(M) \tag{4.22}$$

$$\int_{-\infty}^{\infty} P(M|X = x) \cdot f_x(x)dx = \int_{-\infty}^{\infty} f_x(x|M) \cdot P(M)dx = P(M) \tag{4.23}$$

$$P(M) = \int_{-\infty}^{\infty} P(M|X = x) \cdot f_x(x)dx \tag{4.24}$$

Esempio 4.1.
Due amici fanno una gara di tiro con l'arco. La distanza della freccia dal centro bersaglio è data da una variabile aleatoria rappresentata da un esponente negativo per entrambi:

$$D_A \to f_{DA}(x) = \frac{1}{10} e^{-\frac{x}{10}} u(x) cm^{-1}$$
$$D_B \to f_{DB}(x) = \frac{1}{20} e^{-\frac{x}{20}} u(x) cm^{-1}$$

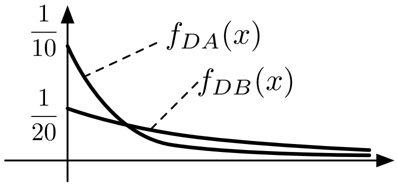


Figura 12: PDF di riferimento

Le variabili aleatorie possono essere espresse nel grafico in fig.12.

Si nota subito che il giocatore più bravo è A, dato che ad una distanza al centro minore corrisponde una probabilità maggiore (area maggiore preso $X \leq x$.

I due amici si alternano regolarmente, ogni freccia viene tolta prima di tirare la successiva. Ad un certo punto passa un terzo amico e trova la freccia piantata a 15cm dal centro, e si chiede quale dei due è più probabile che l'abbia tirata.

Formalizzando il problema si hanno gli evneti:

$$D = \{\text{distanza dal centro}\}$$
$$A = \{\text{ha tirato A}\}$$
$$B = \{\text{ha tirato B}\}$$

Dato che i giocatori si alternavano regolarmente la probabilità che sia stato uno o l'altro è uguale a $\frac{1}{2}$, ma la osservazione viene anche condizionata dal fatto che la variabile aleatoria continua ha assunto un certo valore:

$$P(A|D=15) = \frac{f_D(15|A) \cdot P(A)}{f_D(15)} = \frac{(\frac{1}{10}e^{-\frac{15}{10}}) \cdot \frac{1}{2}}{(\frac{1}{10}e^{-\frac{15}{10}})\frac{1}{2} + (\frac{1}{20}e^{-\frac{15}{20}})\frac{1}{2}}$$

$$f_D(15) = f_D(x)|_{x=15} = f_D(X|A)P(A) + f_D(X|B)P(B)|_{x=15} = f_{DA}(15) \cdot \frac{1}{2} + f_{DB}(15) \cdot \frac{1}{2}$$

5 Parametri statistici di V.A.

Per descrivere le caratteristiche di una variabile aleatoria si utilizzano diversi parametri:

- X: Variabile aleatoria
- $f_x(x)$: Funzione di distribuzione

- η_x: Valore medio $E\{X\}$
- $Var\{X\} = \sigma_x^2$: Varianza

5.1 Valore medio

Intuitivamente si può pensare al valore medio (detto anche **valore atteso**) come il "baricentro" della distribuzione di probabilità.

Per variabili discrete si hanno realizzazioni $X = \{x_1, x_2, \dots x_n\}$, che dopo N esecuzioni dell'esperimento si distribuiscono, creando n_1 realizzazioni di x_1, n_2 realizzazioni per x_2 e così via. La somma di tutte le realizzazioni è uguale al numero di esecuzioni $n_1 + n_2 + \dots + n_n = N$.

La media pesata è data quindi da:

$$\frac{n_1 x_1 + n_2 x_2 + \dots n_n x_n}{N} \tag{5.1}$$

Formalizzando, per **V.A. discrete** si ha che la media, il baricentro di tutte le masse di probabilità, è data dalla media pesata:

$$E\{x\} := \sum_{i=1}^{n} p_i x_i \tag{5.2}$$

Dove p_i è la probabilità di ogni singolo evento $p_i = P\{X = x_i\}, \quad i = 1, \dots n$. La media pesata si esegue sommando le posizioni dei contributi moltiplicati per i loro pesi singolarmente, tutto diviso per il peso totale, quindi 1, che è la somma di tutte le probabilità.

Per **V.A. continue** si generalizza utilizzando gli integrali:

$$E\{x\} = \eta_x := \int_{-\infty}^{+\infty} x \cdot f_x(x) dx \tag{5.3}$$

Dove la probabilità infinitesima è data da $p_i = f_x(x)dx$. La lettera E sta per "Expectation", dato che ci si aspetta che la probabilità di un evento sia vicina al baricentro della distribuzione.

Se la probabilità $f_x(x)$ è definita entro un intervallo $[a, b]$, allora sicuramente la media è dentro il medesimo. Si può calcolare quindi in modo ristretto:

$$\int_a^b x f_x(x) dx \tag{5.4}$$

5.1.1 Linearità del valore medio

Date tre variabili aleatorie, X, Y, Z, sia la prima di esse data dalla combinazione lineare delle altre due:

$$X = a \cdot Y + b \cdot Z \tag{5.5}$$

Si può quindi definire una funzione che restituisce valori di X dati valori di Y e Z:

$$X = g(Y, Z) \tag{5.6}$$

$$x = g(y, z) \text{ (elementi singoli)} \tag{5.7}$$

Allora il valore medio di X è uguale alla combinazione lineare dei valori medi delle altre variabili aleatorie.

5.1.2 Valore medio di una costante

Una costante è una variabile aleatoria che assegna un solo valore c di probabilità unitaria (quindi accade sempre e solo quell'evento). Per la proprietà campionatrice del delta di Dirac si ha che il valor medio è sempre lo stesso valore (ovvio).

$$E\{X = c\} = \int_{-\infty}^{+\infty} x \cdot f_x(x)dx = \int_{-\infty}^{+\infty} x \cdot \delta(x - c)dx = c \tag{5.8}$$

5.1.3 Probabilità totale

Si può applicare il teorema della probabilità totale dato che vale la linearità del valore medio, costruendo una combinazione lineare pesata dalle probabilità.

Data una partizione A_i di S:

$$E\{x\} = E\{X|A_1\} \cdot P(A_1) + E\{X|A_2\} \cdot P(A_2) + \cdots = \sum_x E\{X|A_i\} \cdot P(A_i) \tag{5.9}$$

5.2 Valore quadratico medio

Detto anche **valore efficace**, è il valore medio della variabile al quadrato.

$$E\{x^2\} = \int_{-\infty}^{+\infty} x^2 f_x(x)dx \tag{5.10}$$

Si nota che esso è tanto maggiore quanto la massa dista dallo zero, e anche i valori negativi danno un contributo. Per due distribuzioni uguali ma traslate si ha quindi che quella più lontana dall'origine ha una "importanza" maggiore dell'altra.

Due variabili aleatorie possono per esempio avere lo stesso baricentro ma un valore quadratico medio diverso, come in fig.13.

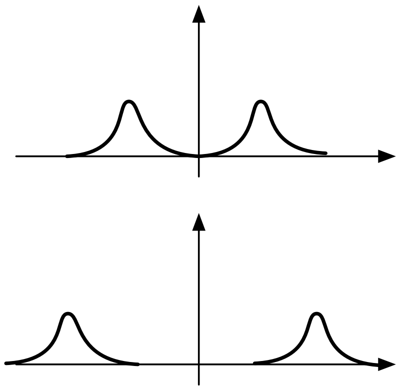


Figura 13: Stesso baricentro diverso val. q.m.

5.3 Varianza

La varianza, indicata con σ_x^2 oppure $Var\{X\}$, indica quanto i valori assunti dalla V.A. si discostano quadraticamente dal valore atteso $E\{X\}$. È il **momento centrale di ordine due** della variabile aleatoria.

Si ottiene semplicemente "traslando" la variabile aleatoria all'origine per poi trovare il valore quadratico medio, che indica quanto sono concentrati oppure "sparpagliati" i valori rispetto allo zero:

$$\sigma_x^2 = E\{(X - \eta_x)^2\} = \int_{-\infty}^{+\infty} (x - \eta_x)^2 f_x(x) dx = \tag{5.11}$$

$$= E\{(x - \eta_x)^2\} = E\{X^2 - 2X \cdot \eta_x + \eta_x^2\} = \tag{5.12}$$

$$= E\{x^2\} - 2\eta_x \cdot E\{x\} + \eta_x^2 = E\{x^2\} - \eta_x^2 \tag{5.13}$$

Di conseguenza la varianza di una singolo punto è zero:

$$X = \{c\}, P\{x = c\} = 1, f_x(x) = \delta(x - c) \tag{5.14}$$

$$\sigma_x^2 = \int_{-\infty}^{+\infty} (x - \eta_x)^2 f_x(x) dx = \int_{-\infty}^{+\infty} (x - c)\delta(x - c) dx = 0 \tag{5.15}$$

La **deviazione standard** si ottiene eseguendo la radice della varianza, ed è utile per avere un valore delle stesse dimensioni della VA.

5.4 Momenti di ordine n di una V.A.

Vengono definiti come segue:

$$E\{x^n\} = \int_{-\infty}^{+\infty} x^n \cdot f_x(x) dx \tag{5.16}$$

Dove per $n = 1$ si ottiene il valore medio, e per $n = 2$ il valore quadratico medio, gli altri non ci interessano.

Un teorema dice che se si conoscono tutti i momenti, quindi noti $E\{x^n\}$ per $\forall n \geq 1$, la $f_x(x)$ è individuata.

5.5 Momenti centrali

Si dicono centrali perché la variabile aleatoria viene "spostata" rispetto al suo centro, dato da $\eta_x = E\{X\}$, e poi si calcola il momento.

$$E\{(x - \eta_x)^n\} = \int_{-\infty}^{+\infty} (x - \eta_x)^n \cdot f_x(x) dx \tag{5.17}$$

Per $n = 1$ si ottiene $E\{x\} - \eta_x = \eta_x - \eta_x = 0$ per qualunque V.A, per $n = 2$ si riconosce la varianza (che quindi è il momento centrale di secondo ordine).

5.6 Valore medio e momenti condizionati

Data una V.A. X, $f_x(x)$ e un evento A:

$$E\{X|A\} = \int_{-\infty}^{+\infty} x \cdot f_x(x|A)dx \tag{5.18}$$

$$E\{X^n|A\} = \int_{-\infty}^{+\infty} x^n \cdot f_x(x|A)dx \tag{5.19}$$

$$E\{g(x)|A\} = \int_{-\infty}^{+\infty} g(x) \cdot f_x(x|A)dx \tag{5.20}$$

5.7 Mediana

La mediana individua un punto tale per cui la massa di probabilità viene divisa in due parti uguali, quindi facendo l'integrale da quel punto a $+\infty$ si ottiene $\frac{1}{2}$.

$$\int_{m}^{+\infty} f_x(x)dx = \int_{-\infty}^{m} f_x(x)dx = \frac{1}{2} \tag{5.21}$$

Nel caso in cui si hanno variabili aleatorie discrete, la mediana si può mettere in un qualunque punto nell'intervallo tra le masse di probabilità che lo contengono, per esempio in fig.14 la mediana si può posizionare in qualunque punto fra p_2 e p_3.

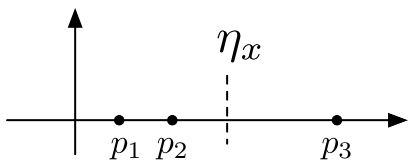


Figura 14: Mediana con V.A. discreta

La mediana serve in casi in cui il valore medio potrebbe perdere di significato a causa di valori molto distanti o lunghe code. Essa indica un valore più vicino alla media effettiva.

5.8 Moda

Si tratta di un altro modo di indicare dove è concentrata la massa di probabilità. È individuata dal punto di massimo nella distribuzione di probabilità:

$$f_x(x)|_{X=M} \tag{5.22}$$

A livello pratico individua il "picco" della funzione. Se ce ne sono più di uno si ha una V.A. **multimodale**.

6 Disuguaglianza di Chebyshev

Per qualunque V.A. e $\forall\varepsilon > 0$, si può sapere che dato il valore atteso η_x e la varianza σ_x^2, si può conoscere la probabilità che la V.A. abbia valori esterni rispetto ad un intervallo centrato in η_x:

$$P\{|x - \eta_x| \geq \varepsilon\} \leq \frac{\sigma_x^2}{\varepsilon^2} \tag{6.1}$$

7 E e σ_x^2 di V.A. notevoli

Segue un elenco di valori medi e varianza delle variabili aleatorie più utilizzate.

7.1 Uniforme

La variabile aleatoria è data da:

$$f_x(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{altrove} \end{cases} \tag{7.1}$$

Il valor medio è dato da:

$$E\{x\} = \int_{-\infty}^{\infty} x f_x(x) dx = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a} \cdot \frac{b^2 - a^2}{2} = \frac{b+a}{2} = \eta_x \tag{7.2}$$

La varianza:

$$\sigma_x^2 = E\{x^2\} - \eta_x^2 = \int_{-\infty}^{+\infty} (x - \eta_x)^2 f_x(x) dx = \frac{(b-a)^2}{12} = \frac{\Delta^2}{12} \tag{7.3}$$

7.2 Esponenziale negativa

Variabile aleatoria:

$$f_x(x) = \lambda e^{-\lambda x} u(x), \quad \lambda > 0 \tag{7.4}$$

Valor medio:

$$E\{x\} = \int_{-\infty}^{+\infty} x \lambda e^{-\lambda x} u(x) dx = \frac{1}{\lambda} \tag{7.5}$$

Varianza:

$$\sigma_x^2 = \frac{1}{\lambda^2} \tag{7.6}$$

7.3 Gaussiana

Variabile aleatoria:

$$f_x(x) = \mathcal{W}(\eta_x, \sigma_x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \cdot e^{-\frac{(1-\eta_x)^2}{2\sigma_x^2}} \tag{7.7}$$

Il valore medio è già definito da $E\{x\} = \eta_x$, stessa cosa per la varianza σ_x^2.

8 V.A. funzioni di altre V.A.

Siano noti $F_x(x)$, $f_x(x)$, $y = g(x)$, si vuole trovare $F_y(y)$, $f_y(y)$.

Si può affrontare il problema in due modi, trovando $F_y(y)$ ed eseguire la sua derivata, oppure applicando il teorema fondamentale.

8.1 Teorema fondamentale

Per calcolare $F_y(y)$ si prende l'asse delle ordinate della funzione $g(x)$ e si fissa un punto, da quel punto in giù si hanno i valori che verificano l'evento:

$$F_y(y) = P\{Y \leq y\} = P\{r \in S : g[x(r)] \leq y\} \tag{8.1}$$

Per $\bar{y}$ le radici $R\bar{y} = \{x_1, x_2, x_3\}$ di $\bar{y} = g(x)$ determinano quali valori prendere per calcolare la probabilità. In altre parole, la funzione $g(x)$ dice quali parti della retta reale selezionare dalla variabile aleatoria X, quindi si passa da $Y \leq y$, si ottengono le parti della retta x che interessano e si applicano alla V.A. originale per selezionare gli eventi.

$$F_y(\bar{y}) = P\{Y \leq \bar{y}\} = P\{x \in R\bar{y}\} \tag{8.2}$$

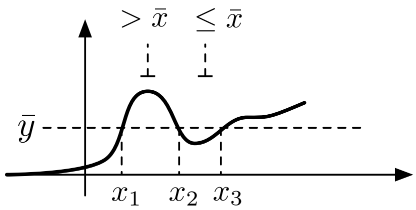


Figura 15: Trovare gli zeri di $g(x)$

Nel caso di fig.15 si ha quindi che:

$$F_y(\bar{y}) = P\{X \leq x_1\} + P\{x_2 < X \leq x_3\} + \dots \tag{8.3}$$

Che data $F_x(x)$ diventa:

$$F_y(\bar{y}) = F_x(x_1) + [F_x(x_3) - F_x(x_2)] + \dots \tag{8.4}$$

Per ottenere la distribuzione si deriva:

$$f_y(y) = \frac{d}{dx} F_y(y) \tag{8.5}$$

Applicando il teorema fondamentale si ha che se $y = g(x)$ non ha radici allora $f_y(y) = 0$, mentre per ogni y per cui $y = g(x)$ ha radici $x_1, \dots x_n$ si ha che:

Se $y = g(x)$ non ha tratti orizzontali né salti allora:

$$P\{y < Y < y + \Delta y\} \simeq f_y(y) \cdot \Delta y \tag{8.6}$$
$$\simeq f_x(x_1)\Delta x_1 + f_x(x+2)\Delta x_2 + \dots \tag{8.7}$$

Volendo estrarre $f_y(y)$ si ottiene:

$$f_y(y) \simeq f_x(x_1)\frac{\Delta x_1}{\Delta y} + f_x(x_2)\frac{\Delta x_2}{\Delta y} + \dots \tag{8.8}$$
$$= \frac{f_x(x_1)}{\left|\frac{\Delta y}{\Delta x_1}\right|} + \frac{f_x(x_2)}{\left|\frac{\Delta y}{\Delta x_1}\right|} + \dots \tag{8.9}$$
$$= \frac{f_x(x_1)}{|g'(x_1)|} + \frac{f_x(x_2)}{|g'(x_2)|} + \dots \quad \text{per altre radici} \tag{8.10}$$

Se invece sono presenti tratti orizzontali in $g(x)$ allora:

$$P\{x_a < X < x_b\} = \int_{x_a}^{x_b} f_x(x)dx \tag{8.11}$$

Per tratti verticali (salti) invece:

$$f_y(y) = 0, \quad y_c < Y < y_d \tag{8.12}$$

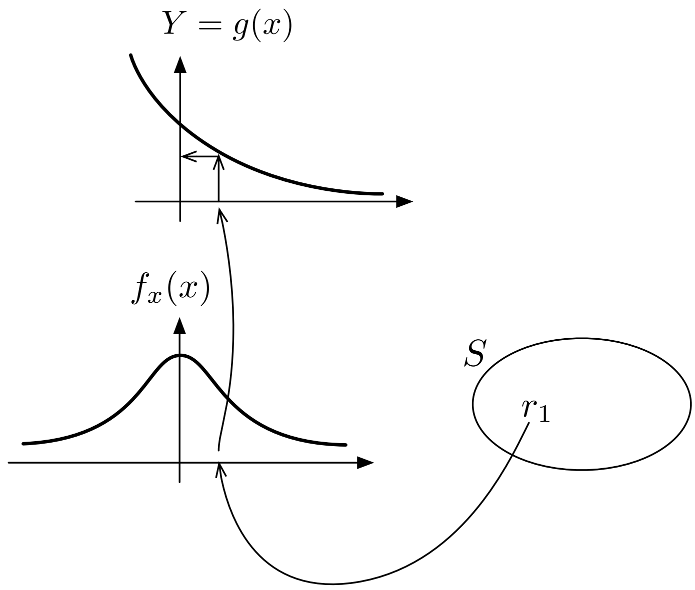


Figura 16: Funzione di VA

Esempio 8.1.
Sia $f_x(x)$ data da una funzione uniforme fra -1 e $+1$ (per normalizzazione ha quindi altezza $1/2$, $y = g(x) = x^2$.

Si calcolano le radici della funzione, quindi si rende esplicita rispetto alla variabile x:

$$x_{1,2} = \pm\sqrt{y}$$

Per calcolare i termini al denominatore si esegue la derivata:

$$g'(x) = 2x \implies |g'(x_1)| = |g'(x_2)| = 2\sqrt{y}$$

Applicando la formula 8.10 per ogni radice si ottiene:

$$\frac{f_x(x_1)}{g'(x_1)} = \frac{f_x(-\sqrt{y})}{|-2\sqrt{y}|} = \begin{cases} \frac{1/2}{|-2\sqrt{y}|}, & -1 < -\sqrt{y} < 1 \\ 0, & \text{altrove} \end{cases}$$

$$\frac{f_x(x_2)}{g'(x_2)} = \frac{f_x(\sqrt{y})}{|2\sqrt{y}|} = \begin{cases} \frac{1/2}{|2\sqrt{y}|}, & -1 < \sqrt{y} < 1 \\ 0, & \text{altrove} \end{cases}$$

La funzione di distribuzione riferita a y è quindi:

$$f_y(y) = \begin{cases} \frac{1}{2\sqrt{y}}, & 0 < y < 1 \\ 0, & \text{altrove} \end{cases}$$

8.1.1 Teorema dell'aspettazione

Sia $Y = g(x)$ una funzione di V.A., si vuole calcolare $E\{Y\}$ ma non si ha la densità di probabilità di Y. Invece che fare i passaggi per ottenerla si può semplicemente moltiplicare la funzione per la densità di X:

$$E(y) = \int_{-\infty}^{+\infty} g(x) \cdot f_x(x)dx \tag{8.13}$$

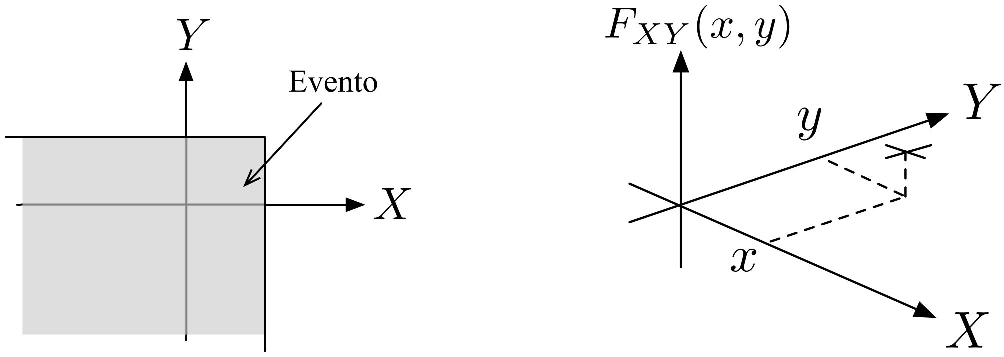
9 Coppie di V.A.

Date due variabili aleatorie X e Y tali che si possano definire due probabilità $P\{X \leq x\}$ e $P\{Y \leq y\}$, esse rappresentano diversi sottoinsiemi dello spazio campione. La probabilità che entrambe siano verificate è data da $P\{X \leq x, Y \leq y\}$ (anche se in generale non si può sapere dato che gli eventi non sono indipendenti).

9.1 Funzione di distribuzione

Si definisce la funzione di distribuzione coniugata come una "generalizzazione in due dimensioni" della funzione che già si conosce:

$$F_{XY}(x,y) := P\{X \leq x, Y \leq y\} \tag{9.1}$$



(a) Evento nel piano (b) Funzione di distribuzione

Figura 17: CDF nel piano

Un evento è quindi rappresentato da un punto in un piano che soddisfa la disuguaglianza, come in fig.17a, la funzione di distribuzione si può rappresentare come una funzione di due variabili nel piano (fig.17b).

Un evento rettangolare si può definire con $A = \{X \leq \bar{x}, y_1 < y \leq y_2\}$, mentre un semipiano è dato da:

$$F_x(x) = P\{X \leq x\} = P\{X \leq x, Y \leq +\infty\} = F_{XY}(x, +\infty) \tag{9.2}$$

Per trovare la funzione di densità in una sola dimensione basta fissarne una, mentre come prima non è possibile avere una probabilità maggiore di 1:

$$F_y(y) = P\{y \leq y\} = F_{XY}(+\infty, y) \tag{9.3}$$
$$F_{XY}(+\infty, +\infty) = 1 \tag{9.4}$$
$$F_{XY}(-\infty, -\infty) = 0 \tag{9.5}$$
$$\tag{9.6}$$

9.2 Coppie di V.A. discrete

Quando le variabili aleatorie sono discrete esse creano "gradini" che si sommano come in fig.18. Per trovare un insieme si agisce per sottrazioni progressive finché non si "ritaglia" la parte interessata.

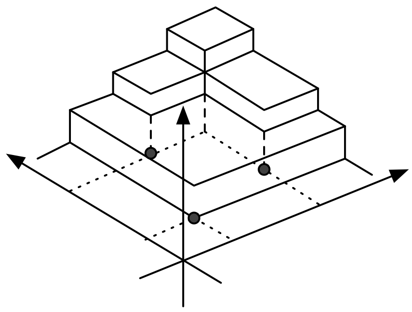


Figura 18: Coppie di V.A. discrete

9.3 Densità di probabilità

Per trovare la densità data la funzione di distribuzione, si eseguono due derivate parziali incrociate:

$$f_{xy}(x,y) = \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y} \tag{9.7}$$

La operazione inversa si fa con un doppio integrale:

$$F_{XY}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(u,v) du, dv \tag{9.8}$$

Si può quindi notare che la probabilità di un evento è data dal volume sotteso dalla funzione di due variabili. Per ricavarla si esegue quindi la integrazione:

$$P\{(x,y) \in D\} = \iint_D f_{XY}(x,y)\, dx\, dy \tag{9.9}$$

Di conseguenza, per la normalizzazione:

$$\iint_{-\infty}^{+\infty} f_{XY}(x,y)\, dx\, dy = 1 \tag{9.10}$$

Inoltre se il dominio è rettangolare si può scambiare l'ordine degli integrali.

9.4 Densità marginali

Data una densità $f_{XY}(x,y)$

$$P\{X_1 < X < X_2\} = \int_{x_1}^{x_2} \int_{-\infty}^{+\infty} f_{XY}(x,y)\,dy\,dx \tag{9.11}$$

$$f_x(x) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\,dy \tag{9.12}$$

$$f_y(y) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\,dx \tag{9.13}$$

Esempio 9.1.
Data la distribuzione (esponenziale negativa):

$$f_{XY}(x,y) = \begin{cases} 2e^{-x}e^{-2y}, & - < x < +\infty \\ 0, & \text{altrove} \end{cases}$$

Si vuole calcolare la probabilità data da:

$$P\{X > 1, Y < 1\}$$
$$P\{X < Y\}$$

$$\begin{aligned} P &= \int_{-\infty}^{+1} \int_{+1}^{+\infty} 2e^{-x}e^{-2y}\,dx\,dy = \\ &= \int_{0}^{+1} 2e^{-2y} \int_{1}^{+\infty} e^{-x}\,dx\,dy = \\ &= 2e^{-1}\left[\frac{e^{-2y}}{-2}\right]_0^1 = \\ &= e^{-1}(1-e^{-2}) \simeq 0.32 \end{aligned}$$

Volendo invece trovare la probabilità data da $P\{X < Y\}$:

$$\begin{aligned} P &= \int_{-\infty}^{+\infty} \int_{-\infty}^{y} f_{XY}(x,y)\,dx\,dy = \\ &= \int_{0}^{+\infty} \int_{0}^{y} 2e^{-x}e^{-2y}\,dx\,dy = \\ &= \int_{0}^{+\infty} 2e^{-2y} \int_{0}^{y} e^{-x}\,dx\,dy = \\ &= \int_{0}^{+\infty} 2e^{-2y}(1-e^{-y})dy = \frac{1}{3} \end{aligned}$$

9.5 PDF e densità condizionata

$$F_X(X|Y) := F_X(X|X=Y) := P\{X \leq x|Y=y\} \tag{9.14}$$

Se la variabile aleatoria Y è continua, sia M un evento:

$$P\{M|Y=y\} = \frac{f_Y(Y|M) \cdot P\{M\}}{F_Y(y)} \tag{9.15}$$

$$P\{X \leq x|Y=y\} = \frac{f_y(y|X \leq x) \cdot P\{X \leq x\}}{f_y(y)} \tag{9.16}$$

$$f_Y(y|M) = \frac{d}{dy} F_y(Y|M) \implies f_Y(Y|X \leq x) = \frac{\partial}{\partial y}\{F_Y(Y|X \leq x)\} \tag{9.17}$$

$$F_X(X|Y) = \frac{\frac{\partial}{\partial y}[F_Y(Y|X \leq x)] \cdot F_x(x)}{f_y(y)} = \tag{9.18}$$

$$= \frac{\frac{\partial}{\partial y}\{F_y(Y|X \leq x) \cdot F_x(x)\}}{f_y(y)} = \tag{9.19}$$

$$= \frac{\frac{\partial}{\partial y}\{P\{Y \leq y, X \leq x\}}{f_y(y)} = \tag{9.20}$$

$$= \frac{\frac{\partial}{\partial y} F_{XY}(x,y)}{f_y(y)} \tag{9.21}$$

$$f_x(X|Y) = \frac{\frac{\partial}{\partial x}\frac{\partial}{\partial y} F_{XY}(x,y)}{f_y(y)} = \frac{f_{XY}(x,y)}{f_y(y)} \tag{9.22}$$

$$f_y(Y|X) = \frac{f_{XY}(x,y)}{f_x(x)} \tag{9.23}$$

$$f_{XY}(x,y) = f_x(x|y) f_y(y) = f_y(y|x) f_x(x) \tag{9.24}$$

$$f_x(x) = \int_{-\infty}^{+\infty} f_{XY}(x,y) dy = \int_{-\infty}^{+\infty} f_x(x|y) f_y(y)\, dy \tag{9.25}$$

$$f_y(y) = \int_{-\infty}^{+\infty} f_y(y|x) f_x(x) dx \tag{9.26}$$

9.6 V.A. indipendenti

Due variabili aleatorie X, Y si dicono indipendenti se lo sono gli eventi che le riguardano, quindi date le probabilità $P\{X \leq x\}$, $P\{Y \leq y\}$, se vale la seguente relazione per qualunque x, y allora le V.A. sono indipendenti:

$$P\{X \leq x, Y \leq y\} = P\{X \leq x\} \cdot P\{Y \leq y\} \tag{9.27}$$

Di conseguenza le funzioni di densità e distribuzione sono date da:

$$F_{XY}(x, y) = F_x(x) \cdot F_y(y) \tag{9.28}$$

$$f_{XY}(x, y) = f_x(x) \cdot f_y(y) \tag{9.29}$$

Per V.A. condizionate se sono indipendenti allora il condizionamento da parte dell'altra è indifferente:

$$f(X|Y) = \frac{f_{XY}(x, y)}{f_y(y)} \to f_{XY}(x, y) = f_x(X|Y) f_y(y) \tag{9.30}$$

$$= f_x(x) \cdot f_y(y) \implies f_x(X|Y) = f_x(x) \tag{9.31}$$

9.7 Densità uniformi

Sono densità che hanno valore costante in un certo dominio, quindi data una porzione di piano xy si ha che la densità di probabilità è uguale per ogni suo punto (cilindro, parallelepipedo, ecc.).

Si tratta semplicemente della estensione della densità uniforme su un segmento in due dimensioni. Ovviamente il volume del solido trovato deve essere unitario.

10 Funzioni di due V.A.

Siano date due variabili aleatorie X, Y e una funzione $z = g(x, y)$. Se si fissa z la funzione diventa una relazione fra x e y, che indica di fatto una curva di livello della distribuzione, trovata tenendo "ferma" una dimensione.

La funzione di distribuzione si ottiene ponendo come punti del dominio tutti quelli che soddisfano la disuguaglianza, quindi rientrano nell'insieme dato:

$$F_z(z) = P\{Z \leq z\} = P\{(X, Y) \in D_z\} = \iint_{D_z} f_{xy}(x, y)\, dx\, dy \tag{10.1}$$

$$f_z(z) = \frac{d}{dz} F_z(z) \tag{10.2}$$

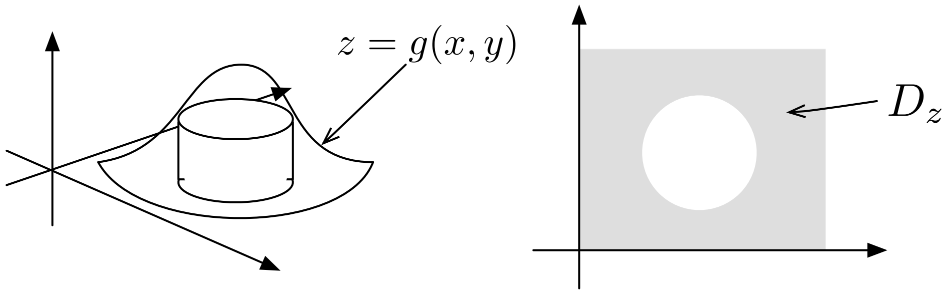


Figura 19: Funzione di due V.A.

Come il teorema fondamentale pone un modo di trovare $f_y(y)$ data $f_x(x)$ e $y = g(x)$ esiste anche una estensione che permette di farlo in due dimensioni (non trattata nel corso).

10.1 Somma di due V.A.

La somma di due variabili aleatorie si può vedere come una funzione che prende due valori e ne ritorna un terzo.

$$Z = X + Y \tag{10.3}$$

$$z = g(x, y) = X + Y \tag{10.4}$$

$$F_z(z) = P\{Z \leq z\} = P\{X + Y \leq z\} = \iint_{D_z} f_{xy}(x, y)\, dx\, dy \tag{10.5}$$

Il dominio si pone come $D_z = \{x, y : x + y \leq z\}$, per la probabilità si ha che i punti da considerare sono dati da $y \leq z - x$. La retta è data dall'intersezione del piano inclinato e quello orizzontale, la curva di livello in questo caso è quindi una retta.

$$F_z(z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{z-y} f_{xy}(x, y)\, dx\, dy \tag{10.6}$$

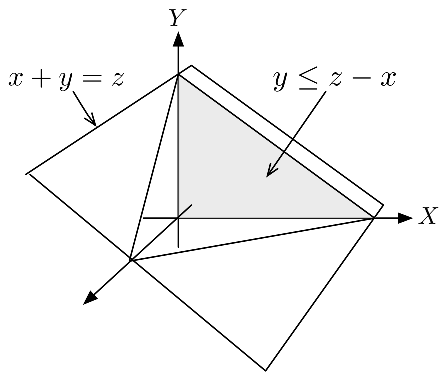


Figura 20: Intersezione di piani

Si ottiene la densità:

$$f_z(z) = \frac{d}{dz}F_z(z) = \tag{10.7}$$
$$= \frac{d}{dz}\int_{-\infty}^{+\infty}\int_{-\infty}^{z-y} f_{xy}(x,y)\,dx\,dy = \tag{10.8}$$
$$= \int_{-\infty}^{+\infty}\frac{d}{dz}\int_{-\infty}^{z-y} f_{xy}(x,y)\,dx\,dy \tag{10.9}$$
$$\frac{d}{dz}(z-y) = 1 \tag{10.10}$$
$$= \int_{-\infty}^{+\infty} f_{xy}(z-y,y)\,dy \tag{10.11}$$

10.2 Somma di V.A. indipendenti

$$Z = X + Y, \quad f_{xy}(x,y) = f_x(x)\cdot f_y(y) \tag{10.12}$$
$$\tag{10.13}$$

Introducendo la indipendenza si ha una operazione di convoluzione:

$$f_z(z) = \int_{-\infty}^{+\infty} f_{xy}(z-y,y)\,dy \tag{10.14}$$
$$= \int_{-\infty}^{+\infty} f_x(z-y)f_y(y)\,dy \tag{10.15}$$
$$= f_x(x) * f_y(y) \tag{10.16}$$

10.3 Minimo fra due V.A.

$$Z = \min(X, Y) \tag{10.17}$$

$$Z = g(X, Y) = \min(X, Y) = \begin{cases} X, & X \leq Y \\ Y, & X > Y \end{cases} \tag{10.18}$$

Esempio 10.1.
Ci sono due sistemi in cascata, ognuno con una vita individuata da una variabile aleatoria:

- X - Vita del primo
- Y - Vita del secondo

Si vuole determinare la vita del sistema complessivo, che quindi sarà determinata dalla vita più breve, dato che se si rompe almeno un sistema il tutto smette di funzionare.

$$D_z = \min(X, Y) \min(X, Y) \leq Z = \begin{cases} X \leq Z, & X \leq Y \\ Y \leq Z, & X > Y \end{cases} \tag{10.19}$$

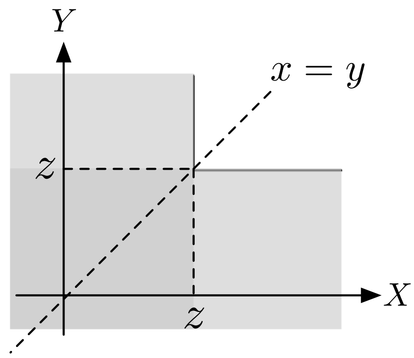


Figura 21: Dominio da considerare

La CDF si ottiene sommando due semipiani e rimuovendo la loro intersezione per non farla comparire due volte:

$$F_z(z) = F_x(z) + F_y(z) - F_{xy}(z, z) \tag{10.20}$$

Se le V.A. sono indipendenti:

$$F_z = F_x(z) + F_y(z) - F_x(z) \cdot F_y(z) \tag{10.21}$$

$$f_z(z) = \frac{d}{dz}F_z(z) = f_x(z) + f_y(z) - f_x(z) \cdot F_y(z) - F_x(z) \cdot f_y(z) \tag{10.22}$$

10.4 Valor medio di due V.A.

Estensione del teorema dell'aspettazione.

$$Z = g(X, Y) \tag{10.23}$$

$$E\{Z\} = E\{g(x, y)\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f_{xy}(x, y)\, dx\, dy \tag{10.24}$$

Nel caso discreto:

$$E\{Z\} = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} g(x_i, y_j) \cdot P\{X = x_i, Y = y_j\} \tag{10.25}$$

11 Parametri sintetici per coppie di V.A.

11.1 Correlazione

Date due V.A. X e Y, la correlazione è data da:

$$R_{XY} := E\{XY\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot y \cdot f_x(x, y)\, dx\, dy \tag{11.1}$$

Due variabili sono incorrelate se il valor medio del prodotto è uguale al prodotto dei valori medi:

$$E\{XY\} = E\{X\} \cdot E\{Y\} \tag{11.2}$$

$$R_{XY} = \eta_x \cdot \eta_y \tag{11.3}$$

Se sono indipendenti allora sono incorrelate, ma non vale il contrario.

La correlazione indica quanto una variabile aleatoria è legata ad un altra (vedi il nome), quindi una correlazione maggiore (in modulo) indica che una variabile è più utile per determinare l'altra.

Per esempio la correlazione fra una variabile e sè stessa è sempre 1, mentre fra due variabili indipendenti è 0. La correlazione positiva è rappresentata da un insieme di punti distribuiti nel primo e il terzo quadrante, viceversa per la negativa (fig.22).

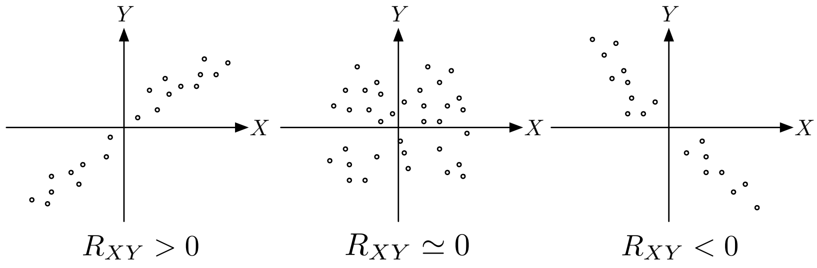


Figura 22: Esempi di correlazioni

11.2 Covarianza

Indica di fatto la correlazione fra due V.A. riferite al valor medio (fig.23).

$$C_{XY} = \text{Cov}(x, y) = E\{(x - \eta_x)(y - \eta_y)\} = \tag{11.4}$$
$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \eta_x)(y - \eta_y) f_{xy}(x, y)\, dx\, dy \tag{11.5}$$
$$C_{XX} = E\{(x - \eta_x)^2\} = \sigma_x^2 \tag{11.6}$$
$$C_{XY} = R_{XY} - \eta_x \eta_y \tag{11.7}$$

Se sono incorrelate, quindi $R_{XY} = \eta_x \eta_y$ la covarianza risulta nulla:

$$C_{XY} = R_{XY} - \eta_x \eta_y = R_{XY} - R_{XY} = 0 \tag{11.8}$$

11.3 Coeff. di correlazione

Indica la pendenza della retta a cui si avvicinano le uscite sperimentali, per esempio in fig.23 a destra il valore si avvicina a 1.

$$\rho_{XY} := \frac{C_{XY}}{\sigma_x \cdot \sigma_y} \tag{11.9}$$
$$-1 < \rho_{XY} < 1 \tag{11.10}$$

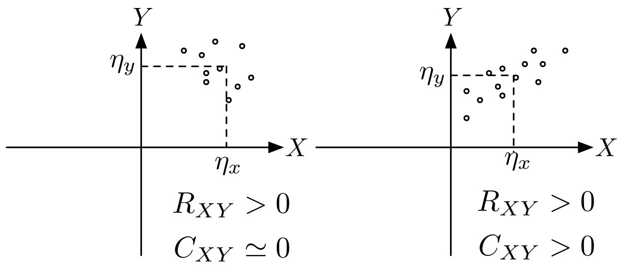


Figura 23: Esempi di covarianze

12 Variabili congiuntamente gaussiane

Segue un formulone di dubbio significato.

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\eta_x)^2}{\sigma_x^2} - 2\rho_{XY}\frac{(x-\eta_x)(y-\eta_y)}{\sigma_x\sigma_y} + \frac{(y-\eta_x)^2}{\sigma_y^2}\right]} \tag{12.1}$$

$$f_x(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x-\eta_x)^2}{2\sigma_x^2}} \tag{12.2}$$

$$f_y(y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{(y-\eta_x)^2}{2\sigma_y^2}} \tag{12.3}$$

Se due variabili sono congiuntamente gaussiane, allora lo è anche la loro combinazione lineare.

13 Vettori di V.A.

Come nel caso di coppie di V.A. si lavora in uno spazio a due dimensioni, si possono generalizzare i teoremi in n dimensioni utilizzando dei vettori che contengono una V.A. in ogni posto.

Un vettore si indica con:

$$\underline{X} = (X_1, X_2, \ldots, X_n) \tag{13.1}$$

La funzione di distribuzione della probabilità congiunta e la densità si scrivono:

$$F_{\underline{x}}(\underline{x}) = F_x(x_1, x_2, \dots, x_n) := P\{X_1 \leq x_1, X_2 \leq x_2, \dots X_n \leq x_n\} \tag{13.2}$$

$$f_{\underline{x}}(\underline{x}) = \frac{\partial^n F_{\underline{x}}(x_1, \dots x_n)}{\partial x_1 \partial x_2 \dots \partial x_n} \tag{13.3}$$

Per coppie di V.A. gli eventi sono descritti da regioni (o sottoinsiemi) del piano in $\mathbb{R}^2$, si estende a tutte le altre dimensioni:

$$\{\underline{X} \in D\}, \quad D \subset \mathbb{R}^n \tag{13.4}$$

$$P\{\underline{X} \in D\} = \int \cdots \int_D f_x(x_1, x_2, \dots x_n)\, dx1\, dx2\, \dots dxn \tag{13.5}$$

È sempre applicata la proprietà di normalizzazione:

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f_{\underline{x}}(x_1, \dots, x_n)\, dx1\, \dots dxn = 1 \tag{13.6}$$

Densità marginale

Si ottiene (volendo rimuovere per esempio le variabili di numero pari, nota i differenziali che determinano quali parti *rimuovere*):

$$f_{x_1 x_3 x_5}(x_1 x_3 x_5) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f_{\underline{x}}(x_1, x_2, \dots x_n)\, dx2\, dx4\, \dots dxn \tag{13.7}$$

$$F_{\underline{x}}(x_1, x_3, x_5) = F_x(x_1, +\infty, x_3, +\infty, x_5, +\infty, +\infty \dots) \tag{13.8}$$

Densità condizionante

$$f_{\underline{x}}(x_1, x_2 - x_k | x_{k+1} \dots x_n) := \frac{f_{\underline{x}}(x_1, x_2, \dots x_n)}{f_{\underline{x}}(x_{k+1} \dots x_n)} \tag{13.9}$$

Regola della catena

$$f(x_1, x_2, x_3 \dots) = f(x_1) \cdot f(x_2|x_1) f(x_3|x_1 x_2) \dots \tag{13.10}$$

Valore medio e teorema dell'aspettazione

$$z = g(X_1, X_2, \dots X_n) \tag{13.11}$$

$$E\{Z\} = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} g(x_1, x_2, \dots x_n) \cdot f_{\underline{x}}(x_1, \dots x_n)\, dx1\, \dots dxn \tag{13.12}$$

Indipendenza

$X_1, \dots X_n$ sono indipendenti se:

$$F_{\underline{x}}(\underline{x}) = F_{X1}(x_1) \cdot F_{X1}(x_2) \dots F_{X1}(x_n) \tag{13.13}$$