 LEZIONE (4) 23/03/09

Esempio (ripresa dell'esempio di pag. 54)
Trasformazione di una Gaussiana attraverso un esponenziale.
La d.d.p. rappresentata a pag. 45 si trasforma nel grafico di pag. 55 che rappresenta la d.d.p. lognormale, definendo la variabile casuale (d'ora in poi v.c.)

$$Y = e^{X}$$

Interessa calcolare la d.d.p. di Y
Partendo da

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

si ottiene

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}\, y} e^{-\frac{(\ln y - \mu)^2}{2\sigma^2}} \quad \text{con } y > 0$$

Si chiama densità lognormale perché applicando il logaritmo si torna alla "normale" (o gaussiana).
Si noti che non ha senso applicare il logaritmo ad una gaussiana, poiché questa si estende da $-\infty$ a $+\infty$ e, in $\mathbb{R}$, non esiste il logaritmo di numeri negativi.
La v.c. lognormale è molto usata per modellizzare fenomeni che per loro natura non sono negativi, come ad esempio la portata di un corso d'acqua.

Esempio Data una v.c. X sia $Y = aX + b$ un'altra v.c.

Determinare $f_Y(y)$

Ricordiamo il teorema di pag. 54.
Y si ottiene con $g(x) = ax + b$.

È richiesto di determinare la d.d.f. di Y a partire dalla conoscenza della d.d.f. di X

Si ottiene: $g'(x) = a$

Inoltre, $y = ax + b$, da cui

$$x_1 = \frac{y-b}{a} \quad (\text{con } a \neq 0)$$

x_1 è l'unica soluzione.

Applicando il teorema, poiché la $g(x)$ ammette sempre la soluzione x_1, il risultato è

$$f_Y(y) = \frac{f_X\left(\frac{y-b}{a}\right)}{|a|}$$

La forma della $f_Y(y)$ dipende da quella della v.c. X.

<u>Osservazione</u>: dimostriamo con esempi che una trasformazione lineare della v.c. X permette di ottenere una v.c. Y la cui d.d.f. è dello stesso tipo della d.d.f. della X.

Caso a: v.c. X con d.d.f. a forma triangolare

Si osservi che l'area sottesa da f_X deve valere 1.

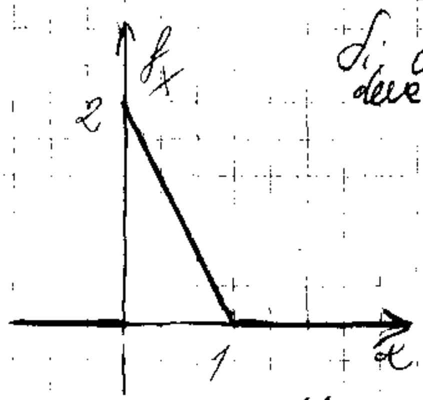
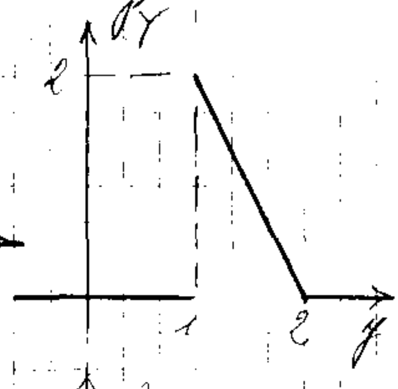


DIAGRAMMA 57

Trasformazione a1

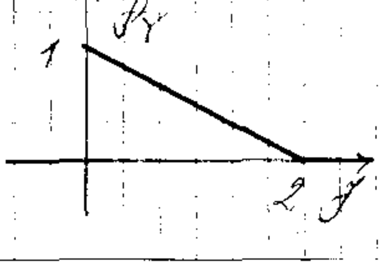
$Y = X + 1$

La trasformazione fa traslare di una unità a destra la f_X.



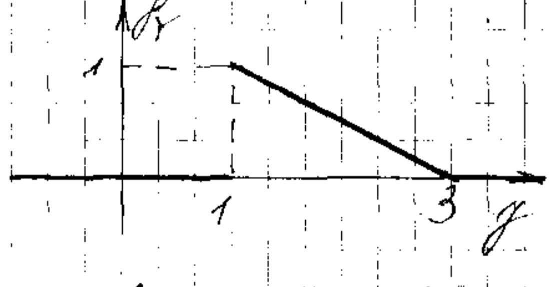
Trasformazione a2

Sempre a partire dal diagramma 57, la trasformazione $Y = 2X$ fa ottenere il grafico rappresentato a destra



58 I valori possibili della X stanno tra 0 e 1.
Quelli possibili della Y stanno tra 0 e 2.
Si ottiene ancora un triangolo, ma con base allargata ed altezza dimezzata.

Trasformazione n 3 $\qquad Y = 2X + 1$

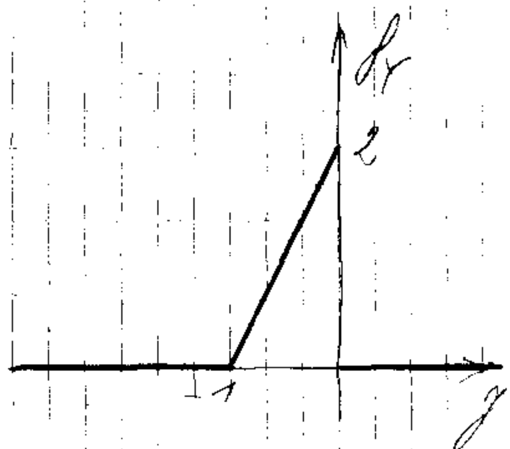


A partire dal grafico rappresentato nell diagramma 57 si ottiene una traslazione a destra di un'unità ed una dilatazione della base.

Trasformazione n 4

$Y = -X$

Si fa un ribaltamento rispetto al diagramma 57.



Conclusione: Operando trasformazioni lineari di v.c. la variabile che si ottiene ha una d.d.f. che è dello stesso tipo di quella da cui si è partiti.

La forma non cambia (si passa da triangolo a triangolo, da rettangolo a rettangolo, ecc.) con eventuale traslazione, ribaltamento, dilatazione o compressione.

Se si partisse da una d.d.f. uniforme si otterrebbe una d.d.f. ancora uniforme.

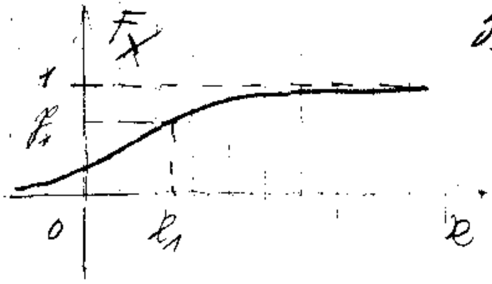
Esempio particolare: data una v.c. X, trasformarla con

$$Y = F_X(X)$$

Determinare la d.d.f. di Y nell'ipotesi che $F_X(x)$ sia continua e strettamente crescente.

In questo caso è $g(x) = F_X(x)$

Si nota che, se $0 \le y_1 \le 1$ l'equazione $Y = F_X(x)$ ammette l'unica soluzione x_1.



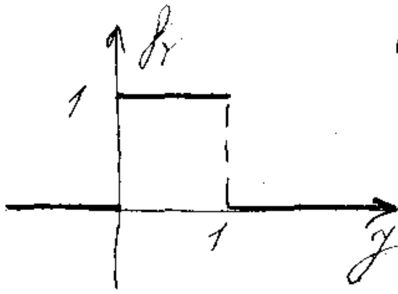
x_1 è tale che

$$F_X(x_1) = y_1$$

ed è la soluzione dell'equazione.

Applicando il teorema di pag. 34:

$$f_Y(y) = \frac{f_X(x_1)}{\left|\frac{d F_X(x_1)}{dx}\right|} = \frac{f_X(x_1)}{|F'_X(x_1)|} = \frac{f_X(x_1)}{f_X(x_1)} = 1$$



Il risultato è rappresentato e unitario e si riferisce a tutte le v.c. trasformate con la loro f.d.d.

Il risultato è 1 solo per i valori che può assumere la Y, che sono quelli compresi tra 0 e 1 ed afferma che qualsiasi v.c. trasformata con la sua f.d.d. diventa uniforme.

Non è interessante generare variabili uniformi, bensì è importante il passaggio contrario.

Può interessare partire da una v.c. uniforme e generare una v.c. con determinate caratteristiche.

Si usa il risultato ma alla rovescia: anziché trasformare con F_X, si trasforma con la sua inversa

$$F_X^{-1}(x).$$

In pratica si vuol affrontare il seguente problema.

Data una v.c. uniforme U in $[0, 1]$, tro=

vare $g(\cdot)$ tale che $X = g(U)$ abbia la f.d.d. F_X desiderata.

Partendo dall'ultimo esempio analizzato, è sufficiente considerare

$$X = F_X^{-1}(U)$$

Esempio: determinare X, partendo da

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{se } x \geq 0 \\ 0, & \text{se } x < 0 \end{cases}$$

Ricordiamo che (pag. 48)

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Partendo da $u = F_X(x)$ si ricava x come funzione di u.

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

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

Passando ai logaritmi

$$-\lambda x = \ln(1-u)$$

$$x = -\frac{\ln(1-u)}{\lambda}$$

La soluzione è

$$X = -\frac{\ln(1-U)}{\lambda}$$

N.B. per procedere nel modo indicato la funzione deve essere in forma chiusa; la funzione di distribuzione gaussiana è nota attraverso valori tabellati.

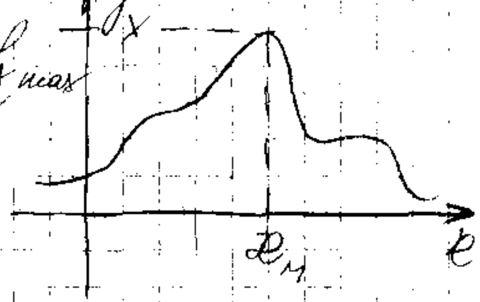
•) Parametri che caratterizzano le d.d.f.

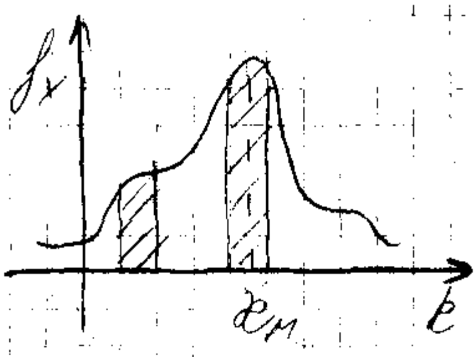
Per caratterizzare una v.c. sono sufficienti la d.d.f. e la f.d.d., ma con queste funzioni la descrizione può essere complicata.

Cerchiamo di condensare una v.c. con la definizione di alcuni parametri.

-) Moda

Essa è la coordinata x in corrispondenza della quale la $f_X(x)$ è massima.



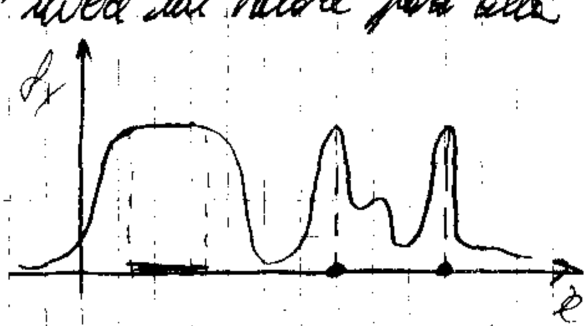


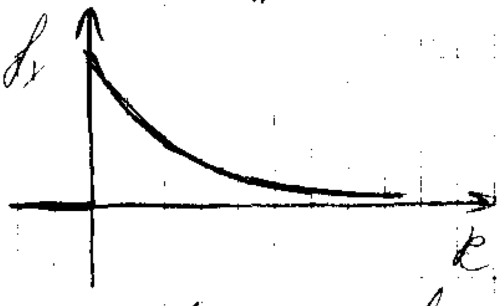
In un certo senso essa rappresenta il valore più probabile.

Considerando dei bins cioè intervalli di larghezza uguale, la probabilità di avere valori prossimi alla moda è maggiore di quella che si ha negli altri intervalli.

La probabilità però, di avere un valore pari alla moda è zero.

La moda può non essere unica, come si vede a destra, in cui una è l'unione di un intervallo e di due punti.





A volte la moda è poco significativa, come nel caso della v.c. esponenziale.

Nel caso a sinistra, che rappresenta tipicamente un tempo d'attesa, la moda è zero, ma tale valore è poco significativo, poiché il grafico mostra che i tempi possono essere molto lunghi.

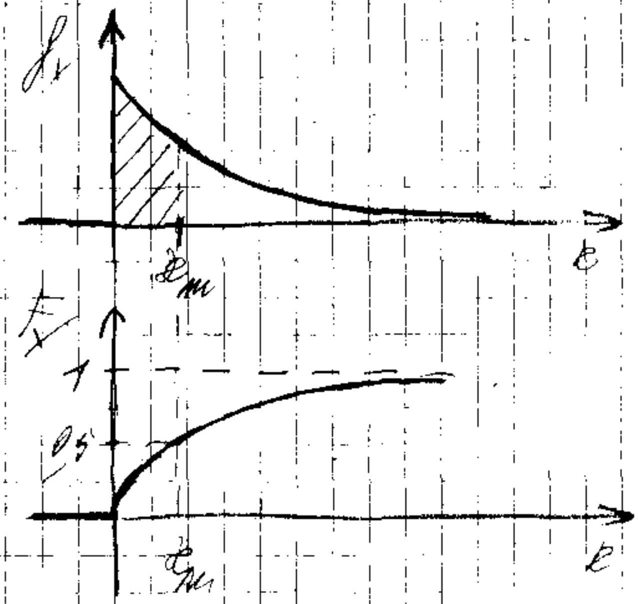
La moda può essere determinata, per lo più, annullando la derivata prima della f_X.

-) Mediana

È un parametro più significativo della media.

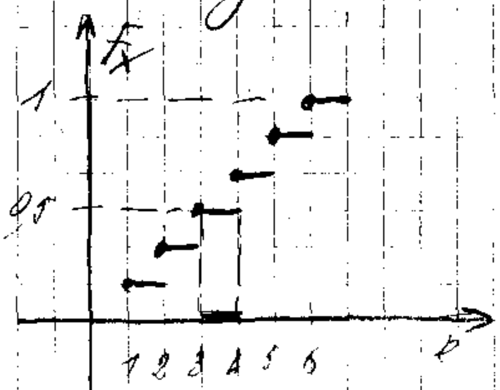
La mediana è il valore x_m che rende

$$F_X(x_m) = 0{,}5$$



La mediana ha la proprietà di dividere l'area sottesa da f_X in due parti uguali pari a 0,5 ciascuna.

La probabilità di avere un valore superiore alla mediana è pari alla probabilità di avere un valore inferiore ad essa.



Anche la mediana può non essere unica, come nel caso del dado, in cui essa è l'intervallo compreso tra 3 e 4.

-) Momenti di ordine k

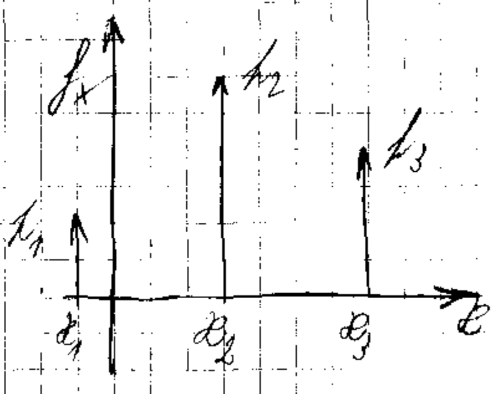
È definito così:

$$m_k = \int_{-\infty}^{+\infty} x^k f_X(x)\,dx$$

·) Variabili casuali discrete

La f_X è una funzione di impulsi, cioè è sempre nulla tranne in alcuni punti in cui si trovano delle delta di Dirac con area p_1, p_2, p_3.

Deve essere $\sum_i p_i = 1$



I momenti sono definiti come sommatorie.

$$m_k = \sum_i x_i^k p_i$$

-) **Momenti centrali di ordine k**

$$\mu_k = \int_{-\infty}^{+\infty} (x - m_1)^k f_X(x)\, dx$$

Per variabili casuali discrete

$$\mu_k = \sum_i (x_i - m_1)^k p_i$$

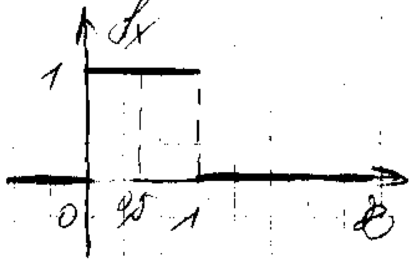
•) **Media o valor atteso**

È il momento del primo ordine e si può indicare in vari modi

$$m_1 \qquad E[X] \qquad m_X$$

La E è l'iniziale di Expectation (aspettativa)

Esempio di v.c. uniforme in $[0, 1]$



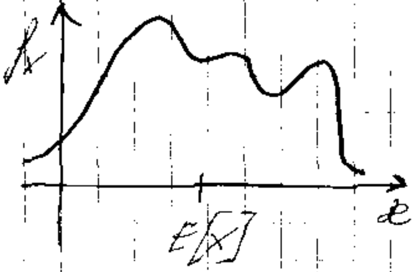
$$E[X] = \int_{-\infty}^{+\infty} x f_X(x)\, dx = \int_{-\infty}^{+\infty} x\, dx = \int_0^1 x\, dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2} = 0.5$$

Interpretazione fisica

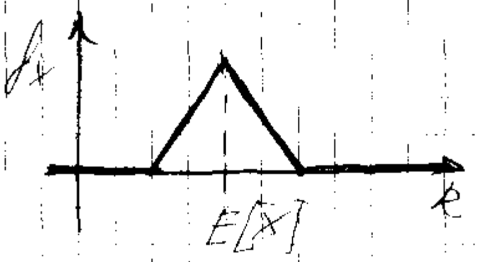
$$E[X] = \int_{-\infty}^{+\infty} x f_X(x)\, dx = \frac{\int_{-\infty}^{+\infty} x f_X(x)\, dx}{\int_{-\infty}^{+\infty} f_X(x)\, dx}$$

È la formula che permette di calcolare il baricentro di una distribuzione di masse f_X, come si è visto nel caso di Fisica 1A.

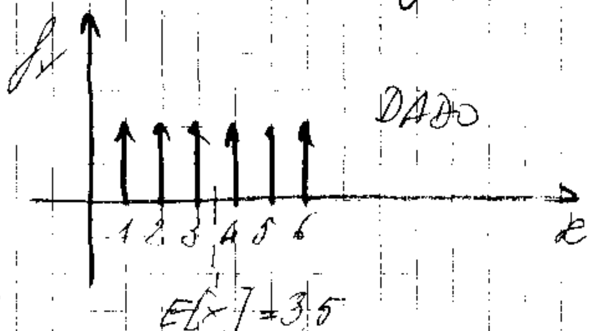
Nel nostro caso $E[X]$ è il baricentro della densità di probabilità.



nota: nel caso di una d.d.p. in cui è presente una simmetria di distribuzione $E[X]$ si trova immediatamente, come nel caso di una funzione pari.



FUNZIONE PARI



Nel caso del dado si può pensare di avere sei masse uguali puntiformi.

·) <u>Stima di $E[X]$</u>

$E[X]$ si stima mediante la media campionaria.
A partire da $x_1, x_2, \dots, x_n$ realizzazioni indipendenti della v.c. X, si ha

$$M_1 = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Nel caso non si possa calcolare la media tramite l'integrale poiché non si conosce la f_X, si possono prelevare le misure di alcuni campioni.

Un modo per stimare $E[X]$ consiste nel calcolare la <u>media aritmetica</u> dei dati sperimentali.

N.B. non bisogna scambiare la media campionaria con la media.

La media è deterministica, poiché si ricava da un integrale.

La media aritmetica è uno stimatore di $E[X]$, al quale si cerca di avvicinarsi il più possibile.

È la stessa differenza che c'è tra istogramma e la d.d.f. vera.

La media campionaria è una quantità casuale. Cambiando i valori dei campioni si trova un valore diverso.

Eseguendo le medie campionarie su un numero di valori che tende all'infinito, ci si avvicina quanto si vuole alla media (teorica).

•) <u>Valor atteso condizionato</u>

$$E[X|M] = \int_{-\infty}^{+\infty} x \, f_{X|M}(x|M) \, dx$$

•) <u>Teorema</u>: data una v.c. X consideriamo

$Y = g(X)$, funzione di v.c.,

risulta che

$$E[Y] = E[g(X)] = \int_{-\infty}^{+\infty} g(x) \, f_X(x) \, dx$$

Il vantaggio consiste nel fatto che si può evitare di calcolare $f_Y(y)$.

•) <u>Conseguenze del teorema</u>

→ $m_2 = E[X^2]$

Infatti, per definizione è: $m_2 = \int_{-\infty}^{+\infty} x^2 f_X(x) \, dx$

Inoltre, ponendo $g(x) = x^2$, si ottiene

$$E[X^2] = E[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x) \, dx = \int_{-\infty}^{+\infty} x^2 f_X(x) \, dx$$

Più in generale:

$$m_k = E[X^k]$$

Tutti i momenti possono essere ricondotti a momenti del 1° ordine essendo momenti del primo ordine di variabili modificate.

Il momento del 2° ordine è il momento del 1° ordine di x^2.

-) se $Y = aX + b$, allora, con $g(x) = ax + b$:

$$E[Y] = \int_{-\infty}^{+\infty} (ax+b)\, f_X(x)\, dx =$$

$$= \int_{-\infty}^{+\infty} a\, x\, f_X(x)\, dx + \int_{-\infty}^{+\infty} b\, f_X(x)\, dx =$$

$$= a\, E[X] + b \cdot 1 = a\, E[X] + b$$

Il valor atteso di una trasformazione lineare è la trasformazione lineare del valor atteso.

Questa è la linearità dell'operatore $E(\cdot)$.

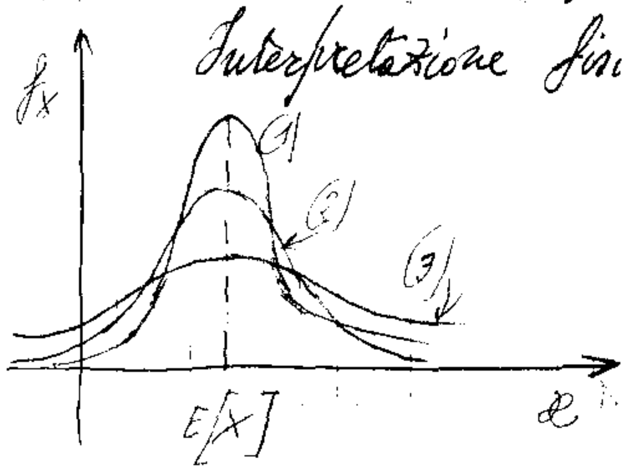
•) <u>Varianza</u>

È un parametro che dà l'idea della variabilità di una v.c.

Definizione:

$$Var[X] = \sigma_X^2 = \mu_2 = \int_{-\infty}^{+\infty} (x - E[X])^2 f_X(x)\, dx$$

Interpretazione fisica: nel grafico sono rappresentate tre v.c. con la stessa media.



La varianza è l'integrale degli scarti al quadrato, pesati mediante f_X.

La varianza è maggiore nel caso (3), dove la dispersione è maggiore e minore laddove i valori sono più concentrati, come nel caso (1).

L'espressione della varianza è formalmente uguale a quella che serve per calcolare il momento d'inerzia intorno ad un asse baricentrico.

L'inerzia alla rotazione di un corpo è tanto maggiore quanto più le masse sono lontane dal centro ed è misurata da μ_2.

La d.d.p. (1) è la meno dispersa e la sua varianza è la minore, riducendo maggior concentrazione attorno al baricentro.

Deviazione standard

$$\sigma_X := \sqrt{Var[X]}$$

I dati riferiti alla deviazione standard sono più comprensibili, poiché dimensionalmente congruenti con la v.c. X.

La varianza può essere comoda ~~nello studio~~ nell'esprimere di determinate proprietà.

Osservazione: ponendo $(x - E[X])^2 = g(x)$, si ottiene

$$Var[X] = \int_{-\infty}^{+\infty} (x - E[X])^2 f_X(x)\,dx =$$
$$= \int_{-\infty}^{+\infty} g(x) f_X(x)\,dx = E[g(x)] =$$
$$= E[(x - E[X])^2]$$

Questa volta ci riconduciamo al concetto di baricentro, cioè ad un momento del 1° ordine.

La varianza può cioè essere intesa come il baricentro della variabile $g(x)$ che rappresenta gli scarti al quadrato rispetto al valor atteso.

Proprietà 1: vale la seguente relazione

$$Var[X] = E[X^2] - E[X]^2,$$ dove

$E[X]^2$ è la media al quadrato (m_1^2).
$E[X^2]$ è il valore quadratico medio, ed è un momento del 2° ordine.

Dimostrazione:

$$Var[X] = E[(X - E[X])^2] = E[(X - m_1)^2] =$$
$$= E[X^2 - 2m_1 X + m_1^2] =$$ (vedi proprietà 2 pag. 61)
$$= E[X^2] - 2m_1 E[X] + E[m_1^2] =$$
$$= E[X^2] - 2E[X]^2 + E[X]^2 =$$
$$= E[X^2] - E[X]^2$$

Proprietà 2:

$$E[g_1(X)+g_2(X)] = E[g_1(x)] + E[g_2(x)]$$

con $g_1(X)$ e $g_2(X)$ funzioni della v.c. X

La proprietà afferma che il valor atteso della somma di due funzioni di X è la somma dei valori attesi di ciascuna funzione.

Esempio: data la v.c. $Y := aX + b$, si determini la

$$Var[Y]$$

$Var[Y] = Var[aX+b]$ = (sfruttando l'osservazione di pag. 64)

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

$$= E[(aX+\cancel{b}-aE[X]-\cancel{b})^2] =$$

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

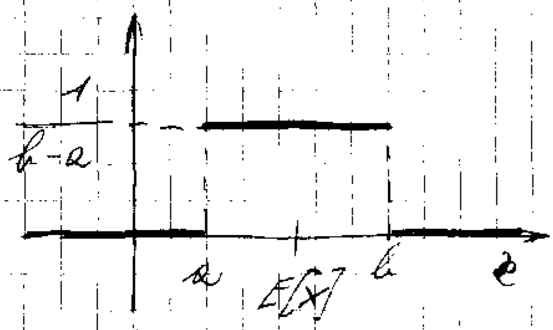
$$= a^2 Var[X]$$

In conclusione:

$$Var[aX+b] = a^2 Var[X]$$

Medie e varianza di alcune v.c. notevoli

-) Uniforme su $[a, b]$



$$E[X] = \frac{a+b}{2}$$

$$Var[X] = \frac{(b-a)^2}{12}$$

-) Gaussiana (pag. 65)

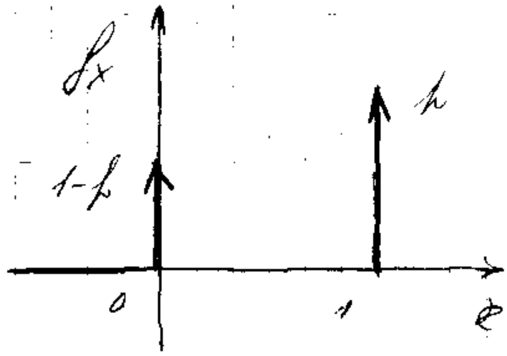
$E[X] = m$, poiché il baricentro è sull'asse di simmetria

$$Var[X] = \sigma^2$$

- <u>Esponenziale</u> (pag. 48)

$$E[X] = \frac{1}{\lambda} \qquad Var[X] = \frac{1}{\lambda^2}$$

- <u>V.C. di Bernoulli</u>



Può assumere solo due valori 0 e 1 ed è rappresentata da due delta di Dirac di area p e $(1-p)$.

Descrive l'esito di una singola prova di Bernoulli: il successo è rappresentato dal valore 1, mentre 0 rappresenta l'insuccesso.

$$f_x(x) = (1-p)\,\delta(x) + p\,\delta(x-1)$$

$$E[X] = p \qquad \underline{Var[X] = p(1-p)}$$

Dimostrazione:

$$E[X] = \sum_{i=1}^{2} x_i p_i = 0(1-p) + 1\cdot p = p$$

$$Var[X] = E[X^2] - E[X]^2 =$$

$$\cancel{= \sum_{i=1}^{2} x_i^2 p_i^2 - p^2 = 0(1-p)^2 + 1\cdot p^2 - p^2}$$

$$= \sum_{i=1}^{2} x_i^2 p_i - p^2 = 0^2\cancel{(1-p)} + 1^2 p - p^2 =$$

$$= p - p^2 = p(1-p) = pq \qquad (q = 1-p)$$

- <u>Binomiale</u> (pag. 47)

$$E[X] = np \qquad Var[X] = np(1-p)$$

La v.c. binomiale può essere considerata come la somma di n v.c. di Bernoulli, essendo la prima è il numero di successi in n prove di Bernoulli.

Ogni successo è una variabile binaria $(0,1)$.

Assumere per la Binomiale $n = 8$ equivale ad eseguire la somma di 8 v.c. di Bernoulli.

Se la v.c. di Bernoulli ha media p, eseguendo n prove di Bernoulli, la media diventa pari ad np.

Esempio: Data una v.c. con $E[X]$ e $Var[X]$ note determinare i valori di a e b tali che, ponendo

$$Y := aX + b$$

si abbia:

$$E[Y] = 0 \quad , \quad Var[Y] = 1$$

In gergo si dice che la v.c. Y è standardizzata.

Occorre determinare una trasformazione lineare che standardizzi la v.c. X.

Si è visto (pag. 58) che una trasformazione lineare presenta la forma.

Intendiamo ottenere una v.c. trasformata che abbia il baricentro nell'origine e con varianza pari ad 1.

Imponiamo che $E[Y] = 0$

$$- E[Y] = E[aX + b] = a\,E[X] + b = 0$$

Imponiamo che $Var[Y] = 1$

$$- Var[Y] = Var[aX + b] = a^2 Var[X] = 1 \quad \text{(pag. 66)}$$

Dall'ultima equazione si ricava a:

$$a = \frac{1}{\sqrt{Var[X]}} = \frac{1}{\sigma_X}$$

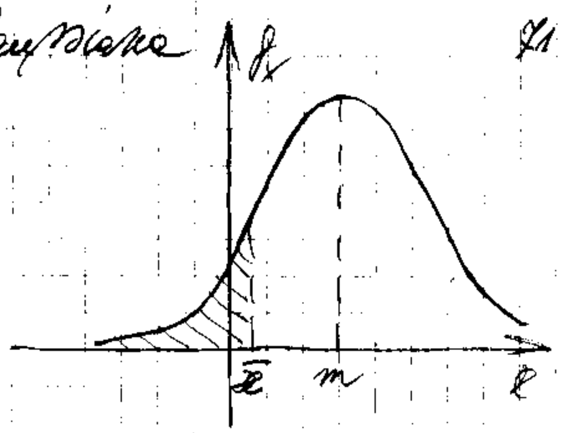
Dalla prima equazione:

$$b = -a\,E[X] = \frac{-E[X]}{\sigma_X}$$

La trasformazione è (trasf. di standardizzazione)

$$Y = aX + b = \frac{1}{\sigma_X} X - \frac{E[X]}{\sigma_X} = \frac{X - E[X]}{\sigma_X}$$

Applicazione al caso della Gaussiana



Occorre calcolare

$$F_X(\bar{x}) = P(X \le \bar{x})$$

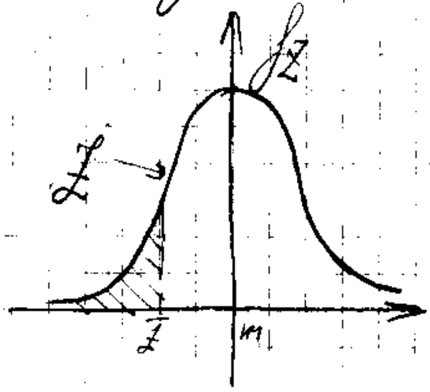
non si conosce la f.d.d. in forma chiusa della Gaussiana i cui valori sono raccolti in tabelle.

Le Gaussiane sono infinite, poiché dipendono da σ e da m, ma le tabelle si riferiscono alla Gaussiana standard.

Le tabelle riportano i valori assunti dalla f.d.d.

$$F_Z(z)$$

della Gaussiana standard Z



$$f_Z(z) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{z^2}{2}}$$

$$E[Z] = 0 \quad Var[Z] = 1$$

$$F_X(\bar{x}) = P(X \le \bar{x}) =$$

$$= P\left(\frac{X - E[X]}{\sigma_X} \le \frac{\bar{x} - E[X]}{\sigma_X}\right) =$$

(ad entrambi i membri è stata sottratta $E[X]$ e i due membri sono stati divisi per σ_X)

$$= P\left(Z \le \frac{\bar{x} - E[X]}{\sigma_X}\right) =$$

(ponendo $Z := \frac{X - E[X]}{\sigma_X}$ e $\bar{z} := \frac{\bar{x} - E[X]}{\sigma_X}$)

$$= P(Z \le \bar{z}) = F_Z(\bar{z})$$

Ora rimane da cercare sulle tabelle della Z il valore $F_Z(\bar{z})$

Si noti che le aree evidenziate sui diagrammi sono uguali.

72 In pratica, si standardizza $\bar{x}$, ottenendo $\bar{z}$

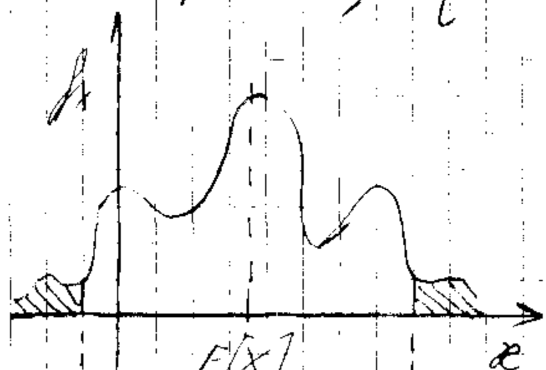
$$\bar{z} = \frac{\bar{x} - E[X]}{\sigma_{\bar{x}}}$$

e consultando poi la tabella della Gaussiana standard per valutare ~~quanto~~ la $F_Z(\bar{z})$ in corrispondenza di $\bar{z}$.

•) <u>Disuguaglianza di Čebyšëv</u> (in italiano anche Cebiceff)

(1821 - 1894 matematico e statistico russo)

$$\forall \varepsilon > 0, \quad P(|X - E[X]| \geq \varepsilon) \leq \left(\frac{\sigma_X}{\varepsilon}\right)^2$$



La relazione vale per ogni v.c. ed riguarda la probabilità che un valore cada lontano dal baricentro più di ε.

Per calcolare la probabilità sarebbe necessario calcolare le aree evidenziate mediante integrazione di f_X.

La disuguaglianza afferma che a prescindere dalla forma della f_X, l'area evidenziata è, comunque, minore di

$$\left(\frac{\sigma_X}{\varepsilon}\right)^2$$

In altre parole se il rapporto ε / σ_X è sufficientemente grande, la probabilità che cada all'esterno dell'intervallo

$$[E[X] - \varepsilon, \ E[X] + \varepsilon]$$

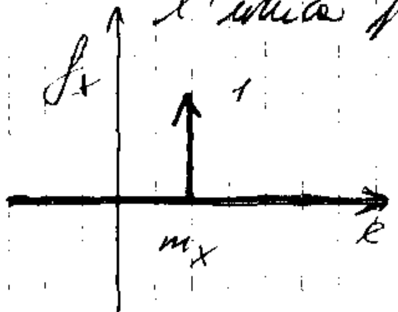
è trascurabile.

<u>Osservazione</u>: se $\sigma_X = 0$ la disuguaglianza si traduce in

$$P(|X - E[X]| \geq \varepsilon) \leq 0 \Rightarrow P(|\cdots|) = 0 \quad \forall \varepsilon$$

Per qualsiasi ε anche arbitrariamente piccolo la probabilità di essere fuori dall'intervallo indicato è nulla.

Rendendo ε sempre più piccolo si riscontra che l'unica possibilità che si ha è che



V.C. DEGENERE

$$X = E[X] = m_X$$

Ne consegue che

$$P(X = E[X]) = 1$$

Se la varianza è nulla la d.d.p. diventa degenere e la v.c. è costretta ad assumere con probabilità 1 valori uguali ad $E[X]$.

Di conseguenza, ogni costante può essere considerata come v.c. con varianza nulla.

Soluzione dell'esercizio di pag. 29.

Si chiede di calcolare $P(A)$ a partire dai suggerimenti proposti.

$A = C_1 \cdot B$ Per andare da α a β devono funzionare il collegamento 1 e quello da γ a β: i collegamenti devono essere simultaneamente veri.

$B = C_2 + C_{43} = C_2 + C_4 C_3$

L'evento B è vero se si può passare per 2 oppure da 4 e 3.

L'obiettivo mira ad esprimere tutto in funzione di C_i.

I nodi funzionano tutti in maniera indipendente, con l'ipotesi che $p_i = P(C_i)$ (C_i indipendenti)

con p_i uguale alla probabilità di funzionamento del ~~negato~~ singolo collegamento.

$$P(A) = P(C_1 \cdot (C_2 + C_4 \cdot C_3)) = \text{(applicando la proprietà distributiva)}$$

$$= P(C_1 C_2 + C_1 C_3 C_4) = \text{(corollario 5 di pag. 13)}$$

$$= P(C_1 C_2) + P(C_1 C_3 C_4) - P(C_1 C_2 C_3 C_4)$$

71 $P(C_1 C_2 C_3 C_4)$ è la probabilità dell'intersezione in cui l'elemento C_1 è contato una sola volta.
Poiché gli eventi sono tutti indipendenti si può scrivere

$$P(A) = P(C_1)P(C_2) + P(C_1)P(C_3)P(C_4) - P(C_1)P(C_2)P(C_3)P(C_4) =$$
$$= p_1 p_2 + p_1 p_3 p_4 - p_1 p_2 p_3 p_4 =$$
$$= p_1 (p_2 + p_3 p_4 - p_2 p_3 p_4)$$

Esercizio 1 del foglio distribuito (Cavalier de Méré) (pag. 71/A)
Definiamo gli eventi:

$A = \{$ almeno un "1" in 4 lanci di dado $\}$

$B = \{$ almeno un doppio "1" in 24 lanci di una coppia di dadi $\}$

Negli esempi precedenti riguardanti la Binomiale si richiedeva che si verificasse esattamente un successo e non almeno un successo, oppure che si verificassero k successi.

"Almeno" un successo su 4 lanci significa che sono accettabili 1 successo oppure 2, 3 o 4.

~~tot~~ Con la Binomiale l'evento favorevole è l'unione di 4 eventi: 1 successo, 2 successi, 3 successi, 4 successi.

Occorrerebbe sommare le 4 probabilità dei 4 eventi.

La strada non sarebbe affatto agevole nel caso B, in quanto dovrei unire 24 eventi: 1 successo, 2 successi, ..., 24 successi.

Meglio considerare la probabilità di non avere nemmeno un successo ~~(A)~~.

La probabilità di avere almeno un successo è data da

1 − prob. di totale insuccesso

Per avere almeno un successo è sufficiente escludere che non ci sia nemmeno un successo.

$$P(\text{almeno un successo}) = 1 - P(\text{nemmeno un successo}) =$$
$$= 1 - q^n$$

q è la probabilità di insuccesso.

La probabilità di avere n insuccessi di seguito in n prove di Bernoulli, in cui vale l'indipendenza è data dal prodotto delle probabilità

$$\underbrace{q \cdot q \cdot \ldots \cdot q}_{n \text{ volte}}$$

Perciò:

$$P(A) = 1 - \left(\frac{5}{6}\right)^4 = 0,5177$$

$$P(B) = 1 - \left(\frac{35}{36}\right)^{24} = 0,4914$$

Variante dell'esercizio precedente

Supponiamo che si richieda di calcolare la probabilità di avere 1 solo successo nel primo e nel secondo caso.

Usando la Binomiale (pag. 31)

$$p_4(1) = \binom{4}{1} p^1 q^{4-1} =$$

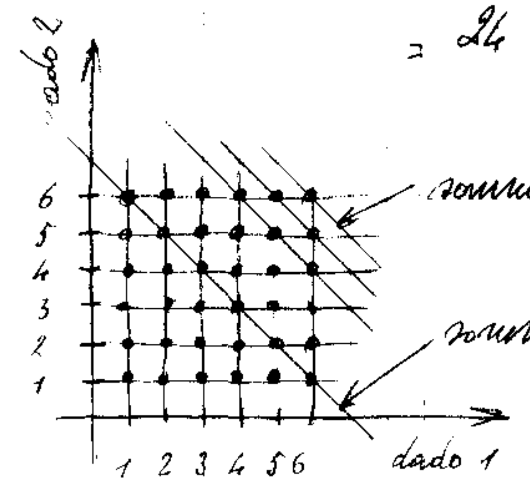
$$= \binom{4}{1} p\, q^3 = 4 \, \frac{1}{6} \left(\frac{5}{6}\right)^3 = 0,3858$$

nel caso di lanciare 1 dado solo.

Nel caso di avere 2 dadi

$$p_{24}(1) = \binom{24}{1} p^1 q^{24-1} =$$

$$= 24 \, \frac{1}{36} \left(\frac{35}{36}\right)^{23} = 0,3391$$



Dal grafico si vede che lanciando 2 dadi, 7 è l'evento con 6 possibilità su 36.

75/4

Identificazione dei Modelli e Analisi dei Dati

Esercizi

1. Nel XVII secolo il Cavalier de Méré, un nobiluomo dell'epoca, sosteneva di aver dimostrato che i seguenti eventi A e B avessero la stessa probabilità:

   A: lanciando un dado 4 volte esce "1" almeno una volta;
   B: lanciando una coppia di dadi 24 volte esce una coppia di "1" almeno una volta.

   Infatti, risulta $4 \times 1/6 = 24 \times 1/36 = 2/3$. L'esperienza mostrava, tuttavia, che il primo evento si verificava più frequentemente. Il Cavalier de Méré chiese di risolvere la questione a Blaise Pascal che si fece aiutare dall'amico Pierre de Fermat. Voi, come rispondereste al Cavalier de Méré?

2. Un automobilista viola quotidianamente un divieto di svolta al fine di accorciare l'itinerario di ritorno a casa dal lavoro. Infatti, ritiene che sia molto bassa la probabilità p di essere multato da un vigile quando si commette tale infrazione.
   Si indichi con x_k la probabilità di farla franca k giorni di seguito. Suppponendo $p = 0.001$, calcolare per quali valori di k risulta $x_k = 0.5$ e $x_k = 0.05$.

3. L'emofilia è una malattia genetica che colpisce gli uomini, mentre le donne possono solo essere "portatrici sane" del gene che causa la malattia. Un figlio di una donna portatrice del gene e di un uomo sano ha il 50% di probabilità di essere ammalato di emofilia. Un figlio di una donna non portatrice del gene e di un uomo sano non può essere ammalato. Una figlia di una donna portatrice del gene e di un uomo sano ha il 50% di probabilità di divenire portatrice del gene.
   Si supponga che la signora X, pur avendo il padre sano, abbia un fratello emofiliaco (pertanto è certo che la madre di X è portatrice del gene). La signora X, avendo sposato un uomo sano, ha avuto due figli maschi sani.

3.a Calcolare la probabilità che la signora X sia portatrice del gene dell'emofilia.

3.b Calcolare la probabilità che il prossimo figlio della signora X sia ammalato di emofilia.

4. Date n persone, quale è la probabilità che non ve ne siano 2 nate lo stesso giorno?

5. Siete nel braccio della morte con altri due prigionieri. Sapete che all'alba uno di voi tre, scelto a caso, verrà giustiziato. Il secondino, che sa già chi è il condannato, dichiara che farà il seguente gioco: mediante il lancio di una moneta sceglierà a caso uno dei vostri due compagni di sventura. Se il compagno scelto è colui che morirà all'alba, il secondino vi informerà che l'altro compagno è destinato a salvarsi. Se invece il compagno scelto a caso è destinato a salvarsi, il secondino vi renderà noto questo fatto. In entrambi i casi il secondino vi indicherà il nome di un compagno destinato a sfuggire all'esecuzione capitale. Dopo che il secondino ha eseguito il suo sadico gioco, quante probabilità avete di essere giustiziati all'alba?