Definizione di probabilità

• ***Prima definizione*** di ***probabilità*** di un evento E, indicata con P(E): rapporto dei casi favorevoli a quelli possibili

$$P(E)= (\text{casi favorevoli})/(\text{casi possibili})$$

Evidentemente $P(E) \leq 1$

Esempio 1: lancio della moneta: con quale probabilità esce testa o croce? Ci sono un caso favorevole e due possibili

$$P(E)=1/2$$

Esempio 2: lancio di un dado : con quale probabilità esce un numero fissato per esempio il 3? Ci sono un caso favorevole e sei possibili

$$P(E)=1/6$$

Esempio 3: lancio di due monete : con quale probabilità esce almeno una testa?
I casi possibili sono TT, TC, CT, CC, i casi favorevoli TT, TC, CT

$$P(E)=3/4$$

Critica alla definizione: appena la situazione si complica a questa definizione bisogna aggiungere la clausola: “purchè siano ugualmente possibili”, che non è molo logica

Definizione di probabilità

• ***Seconda definizione:*** operativa

$$P(E) = \lim_{N \to \infty} (n/N)$$

dove n = numero di volte che si verifica l'evento E, N numero totale di volte che si ripete l'esperimento, n/N è la *frequenza relativa statistica* dell'evento E.

Pochè $n \leq N$ si ha ancora $P(E) \leq 1$

Nota: questa definizione è buona per le applicazioni ma non per una rigorosa trattazione matematica

• ***Terza definizione***: usando la teoria della misura. Chiamo S = {insieme degli eventi E}. Si considerino tanti sottoinsiemi s di S , $(s \subseteq S)$ Sia $\mathcal{E}$ 'insieme dei sottoinsiemi s:

$$\mathcal{E} = \{s : s \subseteq S\}$$

Sia P(s) una funzione definita in $\mathcal{E}$ che associa ad ogni insieme s un numero reale non negativo. P(s) è una probabilità se valgono le seguenti proprietà

- $P(S) = 1,$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \forall A, B \in \mathcal{E},$
- $P(\emptyset) = 0,$

Il simbolo '$\emptyset$' indica l'insieme vuoto. Si ha ancora $P \leq 1$

Variabili casuali e funzioni di distribuzione

*Richiami su proprietà e leggi della probabilità sono date in appendice

•Lanciamo due dadi e calcoliamo la somma dei risultati S: essa è funzione del lancio dei dadi ed è quindi una *variabile casuale.*

•Per ogni valore di S si può calcolare la probabilità che si ottenga proprio quel valore

S	2	3	4	5	6	7	8	9	10	11	12
Casi favorevoli	1	2	3	4	5	6	5	4	3	2	1
P(S)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/3	1/36

•L'insieme dei valori P(S) costituisce la *funzione di distribuzione* della variabile S

•Conoscere la funzione di distribuzione significa conoscere la *probabilità associata a ciascun valore* della variabile casuale

•Il caso considerato è quello di una *variabile casuale discreta* in un intervallo finito. Esistono casi di variabili continue su intervalli finiti o infiniti:

•*Esempio* : supponiamo di sparare con una carabina su un bersaglio: La distanza dal centro a cui arriva il colpo è una variabile casuale continua su un intervallo finito.

Per i casi discreti la *distribuzione è una funzione cha associa a qualunque valore della variabile x la sua probabilità*

L. Andreozzi
<andreozz@df.unipi.it>

•Per il caso di una *variabile continua* la definizione non è applicabile.

•La probabilità di un valore esattamente definito e zero.

•Bisogna chiedersi qual è la probabilità che si abbia un valore in un intervallo assegnato

•Dividiamio l'intervallo in cui può variare la variabile *x* in intervallini Δx.

•Possiamo calcolare la probabilità che la x vada a cadere in ciascun intervallino, cioè la probabilità che la variabile x assuma un valore compreso tra $x - \Delta x$ e $x + \Delta x$.

•Sia $P(x, \Delta x)$

•Per avere una quantità indipendente dall'intervallino Δx si calcola il rapporto $P(x, \Delta x) / \Delta x$

•Considerando il limite:

$$\lim_{(\Delta x \to 0)} P(x, \Delta x) / \Delta x$$

Si ottiene la *densità di probabilità f(x)* che è per definizione la distribuzione della variabile casuale x

•La densità di probabilità è una funzione che non ha ne' le dimensioni ne' il significato di probabilità, ma il cui valore integrato in un certo intervallo della variabile aleatoria, fornisce la probabilità che la variabile x esca in quella zona.

$$p(a < x < b) = \int_a^b f(x)dx$$

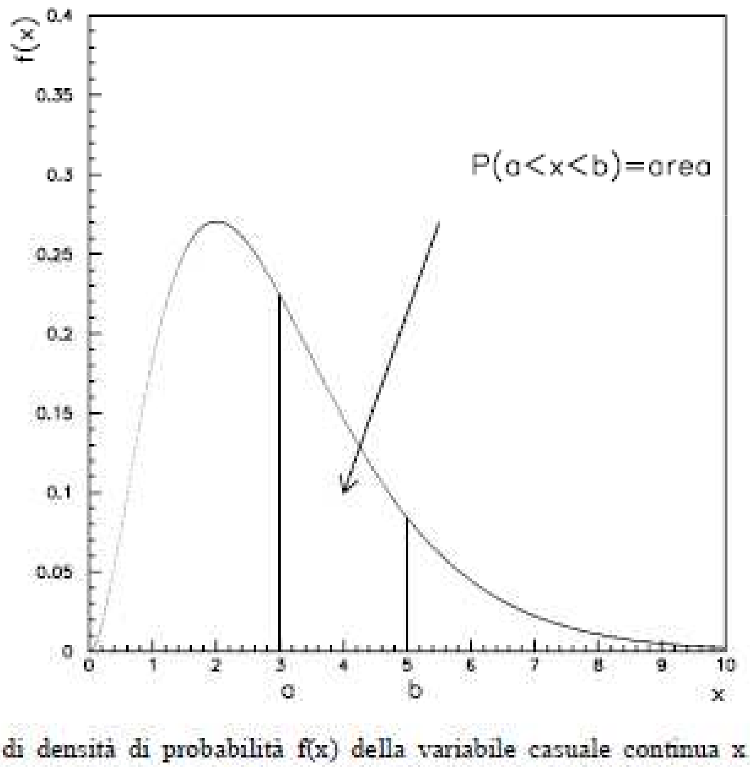


Fig.2.3 Esempio di densità di probabilità f(x) della variabile casuale continua x definita nell'intervallo (0,10). L'area complessiva sottesa alla curva é pari a 1. Per determinare la probabilità che x sia compresa tra i 2 valori a e b, (pari rispettivamente a 3 e 5 in questo caso) si deve valutare l'area indicata.

•*Esempio*: Un’operazione chirurgica consiste nel colpire con un raggio certe cellule maligne. Se si colpisce entro il raggio di 1 mm dal nucleo della cellula, questa muore.

Una equipe medica sostiene di essere in grado di colpire ad una certa distanza x una cellula maligna , dove la variabile casuale x ha una densità di probabilità f(x)

$$f(x) = \begin{cases} (3+x)/9 & \text{per } -3 \le x \le 0 \\ (3-x)/9 & \text{per } 0 \le x \le 3 \\ 0 & \text{altrove} \end{cases}$$

L’equipe medica è affidabile?

R: La variabile casuale è continua e prende valori nell’intervallo [-3,3]. La densità f(x) deve soddisfare la condizione di normalizzazione ad 1, cioè:

$$\int_{-\infty}^{+\infty} f_x(x)\,dx = 1 = P(x \le +\infty).$$

Nel caso specifico $\displaystyle\int_{-\infty}^{+\infty} f_x(x)\,dx = \int_{-3}^{0} \frac{3+x}{9}\,dx + \int_{0}^{+3} \frac{3-x}{9}\,dx =$

$$= \frac{1}{9}\left[\left(3x + \frac{1}{2}x^2\right)\Bigg|_{-3}^{0} + \left(3x - \frac{1}{2}x^2\right)\Bigg|_{0}^{+3}\right] = \frac{1}{9}\left[\left(+9 - \frac{1}{2}9\right) + \left(9 - \frac{9}{2}\right)\right] = 1$$

•La probabilità di rendere innocua una cellula è

$$P(-1 < x < 1) = \int_{-1}^{1} f_X(x)\, dx =$$

$$= \int_{-1}^{0} \frac{3+x}{9}\, dx + \int_{0}^{1} \frac{3-x}{9}\, dx = \frac{1}{9}\left[3 - \frac{1}{2} + 3 - \frac{1}{2}\right] = \frac{5}{9} = 56\%$$

Meglio farsi operare altrove!

Rappresentazione di fenomeni casuali

•Una volta che si disponga di un insieme di più misure della stessa grandezza (un *campione* di misure) i risultati possono essere rappresentati in una tabella o organizzati in modo che il loro significato risulti a colpo d'occhio evidente.

• Caso somma dei dadi.

Sull'asse delle ascisse c'è la variabile S, in corrispondenza ad ogni S sono tracciati segmenti verticali di lunghezza proporzionale al numero n di volte che si è ottenuta la variabile S su N lanci.

Tale rappresentazione è la *rappresentazione di frequenza*

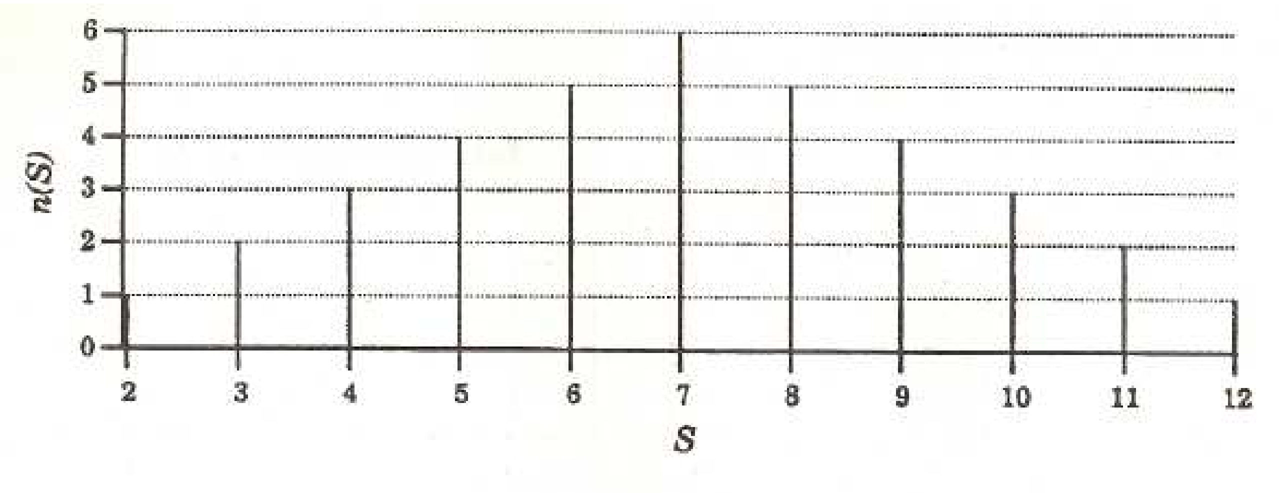


Figura 2.1: $n(S)$ è il numero di volte in cui si è ottenuto S lanciando due dadi.

•*Nota*: la somma delle ordinate è uguale a N

•Talvolta è più utile rappresentare il rapporto
$n(S)/N = f(S)$

•Questa rappresentazione è la *rappresentazione di frequenza relativa* (o di probabilità se N è molto grande), utile per variabili che possono assumere valori discreti e in numero non molto grande

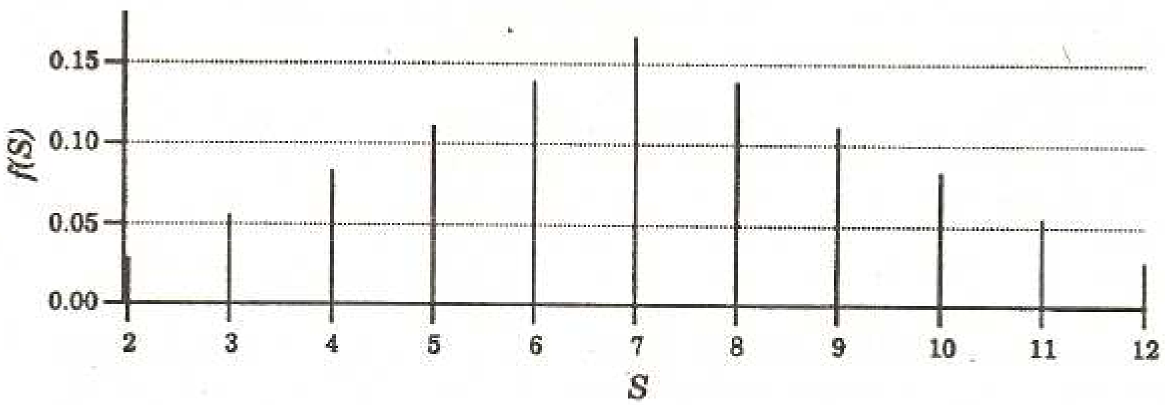


Figura 2.2: $f(S)$ è la frequenza con cui si è ottenuto S lanciando due dadi.

•*Nota*: Per come è stato costruito il grafico la somma delle lunghezza dei segmenti è uguale a 1

•Un altro tipo di rappresentazione è detta ad *istogramma* ed è indispensabile nel caso delle variabili continue

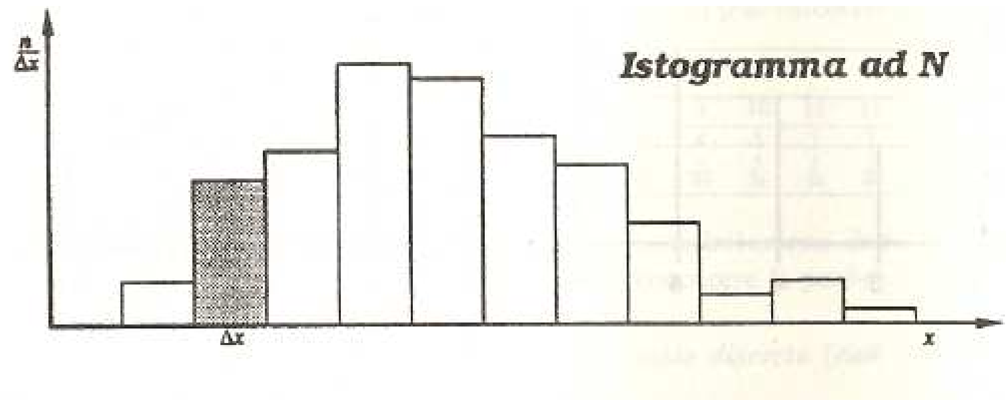


Figura 2.3: Esempio di istogramma normalizzato ad N

•Un istogramma si costruisce in questo modo:

1) si pone sull'asse delle ascisse la variabile x in esame
2) si suddividono i valori che questa variabile può assumere in intervalli di uguale ampiezza
3) si assume ogni intervallo come base di un rettangolo di area proporzionale al numero n di misure

Nota: se si vuole che l'area sia un numero, la grandezza riportata sulle ascisse deve essere sempre l'inverso delle dimensioni delle grandezze riportate in ordinata

•Istogrammi normalizzati a n : area uguale al numero totale N di prove effettuate

•Normalizzati ad 1: si ottengono dividendo l'altezza di ogni intervallo per N. L'area della figura che rappresenta i dati è allora uguale ad 1

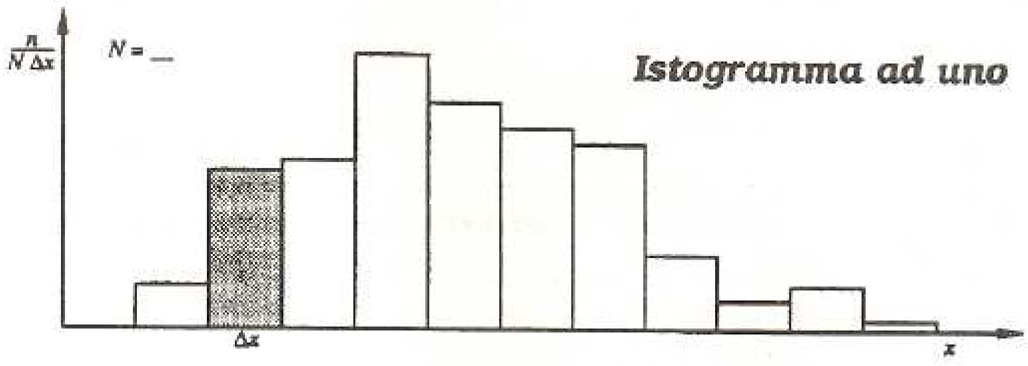


Figura 2.4: Esempio di istogramma normalizzato ad uno

•Un istogramma è molto utile per una prima analisi di un fenomeno è evidente che *viene persa l'informazione* relativa alla distribuzione dei dati dentro ciascun intervallo.

1)Non conviene rimpicciolire troppo l'intervallo per non evidenziare le *fluttuazioni*, perdendo in chiarezza

2) Se l'intervallo è troppo grande si perdono informazioni

3) Regola empirica (da giustificare poi):

$$\Delta x \approx \frac{x_{max} - x_{min}}{\sqrt{N}}$$

•*Elaborazione dati della esercitazione della SOMMA ANGOLI INTERNI DI UN TRIANGOLO (1)*:

disegnare l’istogramma delle 30 misure ottenute

•*Esercizio:* Un gruppo di studenti misura ad occhio e indipendentemente la lunghezza della lavagna dell’aula. Vengono registrati i seguenti risultati (in metri)

6	4.8	6.5	5.62	3	7.5	5.5	4	5.5	5	4.5	5
4.75	4	5.8	5	5	5.5	4.6	4.7	4.5	5.3	4.75	2.80
4.50	5	6.1	4	5.5	4.6	6.60	5.80	5.4	5.5	7.5	5.9
10	5.8	3.7	5.4	5.5	5	5.5	5.8	5	5.5	5.7	5
6.2	6.1	3.45	7.5	5.40	3.5	6.2	5.18	4.5	4.80	5.30	5.5
6	5.3	6	5	6.20	6	5	5	6.5	5.5	5	5.5
5.35	5.2	5.8	4.5	5.45	6.5	6	5	6	5.75	5.25	5.6
5.6	5.9	5	6.5	8	5.40	4	3.2	5.60	6.05	5.5	6.5

Rappresentare questi dati mediante un istogramma.

Alcune definizioni

•Si premettono alcune definizioni. Definiamo

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + a_3 + \cdots + a_n.$$

Proprietà formali:

- $\sum_i a_i = \sum_\ell a_\ell = \sum_\flat a_\flat = \sum_\varsigma a_\varsigma$ cioè l'indice è un indice muto,
- $\sum_i c \cdot a_i = c \cdot \sum_i a_i$,
- $\left(\sum_i a_i\right)^2 = \sum_i \sum_j a_i a_j = \sum_{ij} a_i a_j$.

❑ **Esempio:** $\left(\sum_{i=1}^{3} a_i\right)^2 = (a_1 + a_2 + a_3)^2 = a_1a_1 + a_2a_1 + a_3a_1 + a_1a_2 + a_2a_2 + a_3a_2 + a_1a_3 + a_2a_3 + a_3a_3.$

❍ **Esercizio:** $\sum_{i=1}^{n} k = nk.$

❍ **Esercizio:** $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}.$

❍ **Esercizio:** $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}.$

Caratteristiche comuni alle distribuzioni

Sia x una variabile casuale e P(x) la funzione di distribuzione associata

- *Proprietà di normalizzazione*: la somma su tutti I casi possibili della probabilità sia 1. Matematicamente:

$$\sum_{i=1}^{N} P(x_i) = 1, \qquad \int_a^b P(x)\, dx = 1.$$

dove a e b, estremi di integrazione sono i valori estremi che può assumere la x, eventualmente uguali a infinito. La funzione P è la probabilità nel caso discreto e la densità di probabilità nel caso continuo

- *Media*: si definisce la media μ (o anche m) come

$$\mu = \sum_{i=1}^{N} x_i \cdot P(x_i), \qquad \mu = \int_a^b x \cdot P(x)\, dx.$$

Se le x sono equiprobabili, la media è la media aritmetica.

Infatti se $P(x_1) = P(x_2) = \ldots = P(x_N)$ poichè $\sum_i P(x_i) = 1,$

si ha $P(x_i) = 1/N$ che inserita nella formula che definisce

la media, fornisce la media aritmetica $\sum x_i/N$

- *Mediana*: si definisce la mediana $\mu_{1/2}$ come quel valore tale che

$$P(x_i \leq \mu_{1/2}) = P(x_i \geq \mu_{1/2}), \qquad \int_a^{\mu_{1/2}} P(x)\,dx = \int_{\mu_{1/2}}^b P(x)\,dx = \frac{1}{2}.$$

Nel caso discreto non è detto che tale valore esista

- *Valore più probabile*: è il valore per cui la funzione di distribuzione ha un massimo

$$\mu_{max} : P(x_i) \leq P(\mu_{max})\,\forall i, \qquad \mu_{max} : P(x) \leq P(\mu_{max})\,\forall x \in [a \ldots b].$$

○ **Esercizio**: Supponiamo di lanciare due dadi. Vogliamo trovare la funzione di distribuzione per la variabile S, somma delle uscite dei due dadi. Esaminiamo perciò la figura 2.1, a pagina 11.

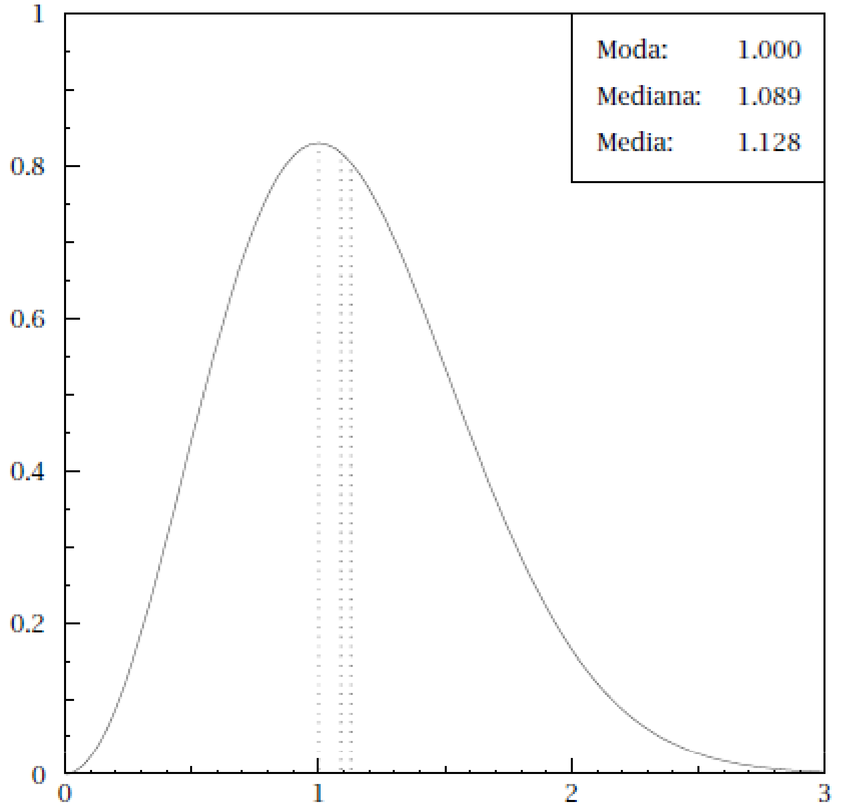
Si vede subito che il numero dei casi favorevoli è dato dalla funzione $f(S) = S - 1$ per S che varia da 2 a 7, e dalla funzione $f(S) = 13 - S$ per S variabile da 8 a 12. Per avere la probabilità di ottenere un certo valore S bisogna dividere il numero dei casi favorevoli per quello dei casi possibili, che sono in tutto $6 \times 6 = 36$. Quindi:

$$P(S) = \begin{cases} \dfrac{S-1}{36} & 2 \leq S \leq 7 \\ \dfrac{13-S}{36} & 7 \leq S \leq 12 \end{cases}$$

Si osservi che il fattore $1/36$ è la costante di normalizzazione per la funzione $f(S)$. Infatti se cerchiamo il fattore moltiplicativo A tale che $\sum_S A \cdot f(S) = 1$ si ottiene:

$$A\sum_{S=2}^{7}(S-1) + A\sum_{S=8}^{12}(13-S) = 1,$$
$$A \cdot (1+2+3+4+5+6+5+4+3+2+1) = 1$$
$$A \cdot 36 = 1, \qquad A = 1/36.$$

Determinare il valor medio e la varianza di S.



Funzione di distribuzione di Maxwell - Boltzmann

$$y = f(v) = \frac{4}{\sqrt{\pi}} \alpha^{\frac{3}{2}} v^2 e^{-\alpha v^2}$$

L. Andreozzi
<andreozz@df.unipi.it>

Misure di dispersione intorno alla media

•Definizione ***del valore di aspettazione*** o di previsione

Consideriamo una funzione f(x) della variabile casuale x:

E[f(x)]

Si chiama *valore di aspettazione* della f ed è:

$$E[f] = \sum_{i=1}^{N} f(x_i) \cdot P(x_i), \qquad E[f] = \int_a^b f(x) \cdot P(x)\, dx.$$

Se f(x) = x, allora

$$E[f] = E[x] = \mu.$$

Perciò E[f] si chiama anche *valor medio di f*

•Proprietà dell'operatore E:

1) E[costante] = costante
2) E[af] = a E[f], dove a è una costante
3) E[f+g] = E[f] + E[g]

Esercizio: dimostrare le proprietà 1,2,3 (basta tener presente la definizione di E e sfruttare la linearità della sommatoria e dell'integrale

Esempio: lancio un dado equo e guadagno 1 euro per ogni punto della faccia che appare: qual è il guadagno aspettato? Se il dado non fosse equo con P(6)= 0.2 e la rimanente distribuzione uniforme, qual è il guadagno aspettato?

a) $E[x] = 1 * (1/6) + 2 * (1/6) + 3 * (1/6) + 4 * (1/6) + 5 * (1/6) + 6 * (1/6) = 21/6 = 3.5$ euro

b) $P(x<6) = 1 - 0.2 = 0.8$
$P(x_i, i = 1\ldots5) = 0.8/5 = 0.16$
$E[x] = 1 * (0.8/5) + 2 * (0.8/5) + 3 * (0.8/5) + 4 * (0.8/5) + 5 * (0.8/5) + 6 * (0.2) = 3.6$ euro

Esempio: sia X una variabile aleatoria e descritta da una densità di probabilità esponenziale di parametro λ. definita da

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

Qual è il suo valore di aspettazione e quello di x^2?

- verifica che è una densità di probabilità

$$\int_0^{+\infty} \lambda e^{-\lambda x}\, dx = -e^{-\lambda x}\Big|_0^{+\infty} = 1$$

$$E[x] = \int_0^{+\infty} x\lambda\, e^{-\lambda x}\, dx = \Big[-x\, e^{-\lambda x}\Big]_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx =$$

$$= \Big\{ 0 + \frac{e^{-\lambda x}}{\lambda} \Big]_0^{\infty} = \frac{1}{\lambda}$$

$$E[x^2] = \int^{\infty} x^2\, \lambda\, e^{-\lambda x}\, dx = -x^2 e^{-\lambda x}\Big]_0^{\infty} + \int 2x\, e^{-\lambda x} dx =$$

$$= -\frac{2}{\lambda} x\, e^{-\lambda x}\Big]_0^{\infty} + \frac{2}{\lambda}\int_0^{\infty} e^{-\lambda x}\, dx = -\frac{2}{\lambda^2}\, e^{-\lambda x}\Big]_0^{\infty} = \frac{2}{\lambda^2}$$

• Misure di *dispersione intorno alla media*

-Si pone il problema di trovare una funzione f tale che E[f] indichi quanto sono lontane le misure dalla media μ

-Prendendo $f(x) = x - \mu$, per cui $E[x - \mu]$ è lo scarto dalla media , si ha:

$$E[x - \mu] = E[x] - E[\mu] = \mu - \mu = 0$$

Tale funzione non dice nulla poiché, come da attendersi, gli errori per eccesso compensano quelli per difetto

-Proviamo con $f(x) = (x - \mu)^2$ e calcoliamo $E[(x - \mu)^2]$. Tale quantità si chiama ***varianza*** e si indica con σ^2.

$$\sigma^2 = E[(x-\mu)^2] = \begin{cases} \sum_{i=1}^{N} P(x_i)\ (x_i - \mu)^2 \\ \int_a^b P(x)\ (x-\mu)^2 dx \end{cases}$$

-Si definisce ***deviazione standard*** σ la radice quadrata della varianza

$$\sigma = \{E[(x-\mu)^2]\}^{1/2}$$

-Si ha che $\sigma^2 = E[x^2] - \mu^2$

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

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

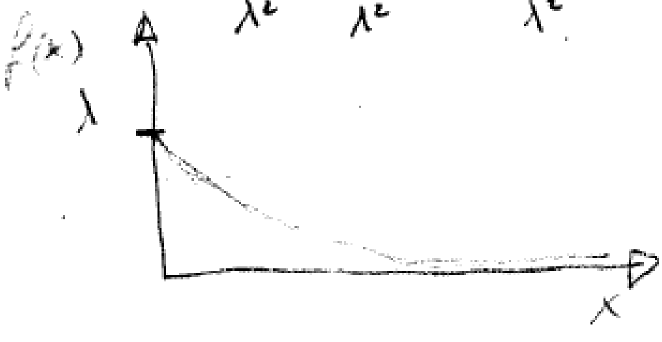
Esempio: sia X una variabile aleatoria e descritta da una densità di probabilità esponenziale di parametro λ. definita da

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

Calcolare varianza e deviazione standard.

Si ha $\sigma^2 = E[x^2] - E^2[x] = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$

$\sigma = \frac{1}{\lambda}$



Discussione e cenno al teorema di Tschebyscheff

- Chiariamo perchè ci interessano la media μ e la varianza σ^2 della distribuzione.
- Quando si hanno misure i cui risultati fluttuano, il ***risultato della misura è una variabile casuale,*** funzione delle condizioni sperimentali.
- *La media* della variabile casuale è il valore più significativo del risultato.
  La *deviazione standard* misura l'incertezza da attribuire al risultato
- Lo dimostra il ***teorema di Tschebyscheff:***
  Sia x una variabile casuale tale che esistono finiti μ e σ.
  Detto k un numero positivo si ha:

$$\boxed{P(|x - \mu| \geq k\sigma) \leq \frac{1}{k^2}.}$$

- Cioè: ***e' poco probabile avere grandi deviazioni dalla media***

Discussione e cenno al teorema di Tschebyscheff (2)

- Dimostrazione:
  se nella definizione di σ^2:

$$\sigma^2 = \int_a^b P(x)(x-\mu)^2\, dx$$

mettiamo al posto di $(x - \mu)^2$

0 se $(x - \mu)^2 \leq k^2 \sigma^2$

$k^2 \sigma^2$ negli altri casi

si ottiene

$$\sigma^2 \geq k^2\sigma^2 \int_{|x-\mu|\geq k\sigma} P(x)\, dx = k^2\sigma^2 \cdot P(|x-\mu| \geq k\sigma)$$

che è esattamente la relazione $\boxed{P(|x-\mu| \geq k\sigma) \leq \frac{1}{k^2}.}$

- *Esempio 1*: tutte le monete da 100 lire sono tra loro differenti. Sia x la v.a. che indica il diametro di una moneta scelta a caso.
  La macchina che le produce alla zecca di stato conia con una σ= 0.3 mm.

  Non avendo a disposizione alcuna informazione sulla densità di probabilità di x, *calcolare la probabilità* che il diametro di x si discosti dal suo valor medio per più di 0.7 mm

si ha σ= 0.3 mm e kσ = 0.7 mm
la disuguaglianza di Tschebyscheff fornisce:

$$\boxed{P(|x-\mu| \geq k\sigma) \leq \frac{1}{k^2}.}$$

$$P(|x-\mu| > 0.7) \leq \frac{(0.3)^2}{(0.7)^2} = 0.18$$

- *Esempio 2*: Sia x la v.a. con densità

$$P(x=-2)=\frac{1}{8} \qquad P(x=0)=\frac{3}{4} \qquad P(x=2)=\frac{1}{8}$$

Calcolare la probabilità che la v. a. si discosti dal suo valor medio per più di 2 incluso e confrontare con il risultato fornito dalla disuguaglianza di Tschebyscheff

È una densità , infatti P(-2)+ P(0)+ P(2)=1

Si ha E[x] =μ= 0. Infatti (1/8)*P(-2)+ 0*P(0)+(1/8)* P(2)=0

Quindi $P(|X| \geq 2) = P(x=-2) + P(x=2) = \frac{1}{4}$

Questo è un conto esatto

Invece con Tschebyscheff:
si ha $k\sigma = 2$

Calcolo di σ usando la : $\sigma^2 = E[x^2] - \mu^2$

$$E[x^2] = \frac{4}{8} + \frac{4}{8} = 1 \qquad \sigma^2 = 1-0 = 1 \qquad \sigma = 1$$

$$P(|x| \geq 2) \leq \frac{1}{4}$$

La disuguaglianza di Tschebyscheff fornisce il valore esatto! Non esiste una maggiorazione piu' stretta della probabilita' di discostarsi da μ

Cenno alla teoria dei campioni

•L'insieme delle misure sperimentali (che è una piccola parte di tutte le misure effetuabili) è un ***campione*** di tutte le possibili misure

- ***Si postula l'esistenza*** di una distribuzione dei risultati di tutte le misure effettuabili: questa viene chiamata ***distribuzione generatrice (parent distribution)***
- Dal campione dobbiamo approssimare la media μ e la varianza σ^2 della distribuzione generatrice
- Vediamo le principali distribuzioni

Appendice sulla probabilità

2.1 Definizione di probabilità

Prima definizione: Probabilità di un evento E, indicata con $P(E)$ è uguale al rapporto tra il numero dei casi favorevoli e quello dei casi possibili:

$$P(E) = \frac{\text{casi favorevoli}}{\text{casi possibili}}. \tag{2.1}$$

Questa viene anche chiamata *probabilità a priori di un evento*. Evidentemente $P \leq 1$.

Questa definizione è sufficiente per i casi più elementari.

❑ **Esempio:** (a) Lancio di una moneta: qual'è la probabilità che esca testa (o croce)? Ci sono un caso favorevole e due possibili, quindi

$$P(T) = {}^1\!/_2.$$

(b) Lancio di un dado: qual'è la probabilità che esca un certo numero fissato, ad esempio il 3? Ci sono un caso favorevole e 6 possibili, quindi:

$$P(3) = {}^1\!/_6.$$

(c) Lancio di due monete: qual'è la probabilità che esca *almeno* una testa. I casi possibili sono TT, TC, CT, CC, i casi favorevoli TT, TC, CT, quindi:

$$P(1T) = {}^3\!/_4.$$

A questa definizione può essere posta la seguente critica: non appena la situazione si complica diventa necessario tenere conto che alcuni casi possono essere più o meno "favorevoli". Quindi si aggiunge alla definizione la clausola: "purchè siano ugualmente possibili" che non è chiaramente molto logica.

Seconda definizione: Operativa:

$$P(E) = \lim_{N\to\infty} \frac{n}{N} \tag{2.2}$$

dove n = numero di volte in cui si verifica l'evento E; N = numero totale di volte in cui si ripete l'esperimento, n/N è la frequenza relativa statistica dell'evento E.

Poichè $n \leq N$ si ha ancora $P \leq 1$.

Questo limite significa che compiendo una serie di prove, con N sempre più grande, il rapporto n/N tende a stabilizzarsi intorno ad un certo valore, con oscillazioni sempre più piccole al crescere di N. Provate a verificarlo sperimentalmente, ad esempio per l'evento 'testa' nel lancio di una moneta.

Questa definizione è buona per le applicazioni ma non per una rigorosa costruzione matematica.

Terza definizione: Usando la teoria della misura: chiamiamo S = {insieme degli eventi E}. Si considerino tanti sottoinsiemi s di S, ($s \subseteq S$). Sia $\mathcal{E}$ l'insieme dei sottoinsiemi s:

$$\mathcal{E} = \{s : s \subseteq S\}$$

Sia $P(s)$ una funzione definita nell'insieme $\mathcal{E}$ che associa ad ogni insieme s un numero reale non negativo. $P(s)$ *è una probabilità se valgono le seguenti proprietà*:

- $P(S) = 1$,
- $P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \forall A, B \in \mathcal{E}$,
- $P(\emptyset) = 0$,

Dove il simbolo '$\emptyset$' indica l'insieme vuoto. Segue che si ha ancora $P \leq 1$.

2.2 Proprietà e leggi della probabilità

Addizione: Se A e B sono due eventi *incompatibili*, cioè che non possono verificarsi contemporaneamente nello stesso esperimento, od anche i sottoinsiemi che li rappresentano hanno intersezione vuota, la probabilità che si verifichi l'evento A o l'evento B è:

$$P(A \text{ o } B) = P(A) + P(B) \qquad \text{Eventi incompatibili.}$$

□ **Esempio**: Si lanci un dado, la probabilità che esca il 2 o il 3 è:

$$P(2 \text{ o } 3) = P(2) + P(3) = 1/6 + 1/6 = 2/6.$$

Se A e B sono compatibili, e cioè non sono disgiunti, possono presentarsi dei casi favorevoli ad entrambi gli eventi, contati quindi sia nel calcolo di $P(A)$ che in quello di $P(B)$. Considerando la somma $P(A) + P(B)$ tali casi sarebbero contati due volte, perciò la legge dell'addizione deve essere così generalizzata:

$$\boxed{P(A \text{ o } B) = P(A) + P(B) - P(A \text{ e } B).} \tag{2.3}$$

□ **Esempio**: Estrarre una carta da un mazzo di 52 carte:

A = la carta è di fiori $\quad P(A) = 13/52$
B = la carta è un re $\quad P(B) = 4/52$
C = la carta è un re o una fiori $\quad P(C) = 16/52 = \frac{\text{casi favorevoli}}{\text{casi possibili}} =$
$= P(A) + P(B) - P(A \cap B)$

In questo caso *non* vale la legge dell'addizione per eventi incompatibili, in quanto gli eventi *non* sono incompatibili.

○ **Esercizio**: Sia data una scatola con 6 palline rosse, 3 blu e 5 bianche. Calcolare la probabilità che esca una pallina blu o bianca.

Moltiplicazione: Consideriamo due eventi A e B, e l'evento $C = (A \text{ e } B)$. Quanto vale $P(C)$? Supponiamo prima che il fatto che si verifichi A non influenzi in nessun modo il fatto che si verifichi B, cioè che tra A e B non ci sia nessuna connessione, A e B sono *indipendenti*. Allora, la legge della moltiplicazione è:

$$P(A \text{ e } B) = P(A) \cdot P(B) \qquad \text{Eventi indipendenti.}$$

Infatti:

$$P(A) = \frac{n_A}{N_A}, \qquad P(B) = \frac{n_B}{N_B}.$$

I casi possibili dell'evento A e B sono $N_A \cdot N_B$, i casi favorevoli $n_A \cdot n_B$,

$$P(A \text{ e } B) = \frac{n_A \cdot n_B}{N_A \cdot N_B} = P(A) \cdot P(B)$$

❑ **Esempio:** Si lanci un dado ed una moneta, la probabilità che escano il 2 e la testa è:

$$P(2 \text{ e testa}) = 1/6 \times 1/2 = 1/12.$$

❑ **Esempio**: Lancio di un dado, $C = 2$ e 3: $P(C) = 0$.

❑ **Esempio**: Lancio di due dadi, $C = 2$ e 3: $P(C) = 1/36$.

❑ **Controesempio**: Estrarre una carta da un mazzo di 52 carte ed un jolly.

A = la carta è di fiori	$P(A) = 13/53$
B = la carta è un re	$P(B) = 4/53$
C = la carta è un re ed una fiori = K♣	$P(C) = 1/53 = \frac{\text{casi favorevoli}}{\text{casi possibili}} \neq P(A) \cdot P(B)$.

In questo caso *non* vale la legge della moltiplicazione per eventi indipendenti, in quanto gli eventi *non* sono indipendenti (se esce una carta di fiori questo esclude che sia un jolly, quindi modifica la probabilità che sia un re).

Generalizziamo la legge della moltiplicazione al caso di eventi dipendenti. Definiamo la *probabilità dell'evento B, condizionata dall'evento A*, e la indichiamo con $P(B/A)$, come la probabilità che si verifichi B una volta che si è verificato A. Allora

$$\boxed{P(A \text{ e } B) = P(A) \cdot P(B/A).} \tag{2.4}$$

Nell'esempio precedente si ha $P(\text{re di fiori}) = 4/53 \times 1/4 = 1/53$. Per eventi A e B indipendenti è ovviamente $P(A) = P(A/B)$.