Costruzione di macchine

Modulo di:

Progettazione probabilistica e affidabilità

Marco Beghini e Leonardo Bertini

Lezione 2:

Probabilità condizionata e variabili casuali

Probabilità condizionata (*P(A/B)*):

La probabilità che si verifichi un evento *A*, assumendo o sapendo che un altro evento *B* dello spazio campionario si è verificato.

Esempio 2.1:

Supponiamo che la probabilità di vincere lo scudetto sia la stessa per ogni squadra della massima divisione, determinare la probabilità :

1) che vinca l'Inter (*A*)

2) che vinca l'Inter sapendo che (oppure 'se', o 'nell'ipotesi che') lo scudetto sia vinto da una milanense (*B*)

$$P(A) = \frac{1}{20}$$

La probabilità (a priori) che vinca l'Inter

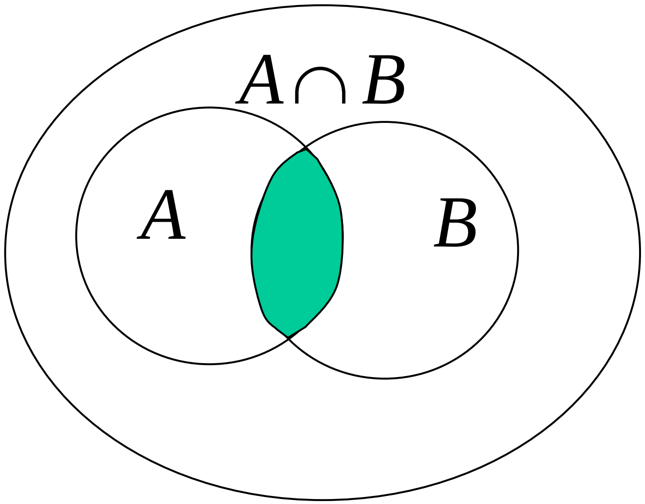
$$P(B) = \frac{2}{20} = \frac{1}{10}$$

La probabilità (a priori) che vinca una milanese

$$P(A|B) = \frac{1}{2}$$

La probabilità che vinca l'Inter se vince una milanese

Nel caso generale



$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Esempio 2.2

Nel lancio di una coppia di dadi, determinare la probabilità di avere almeno un 2 sapendo che la somma è 6.

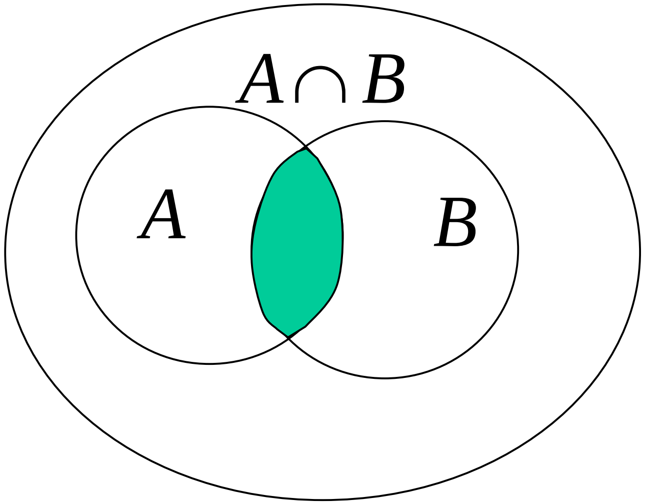
S	**1**	**2**	**3**	**4**	**5**	**6**
1	**O**	**O**	**O**	**O**	**O**	**O**
2	**O**	**O**	**O**	**O**	**O**	**O**
3	**O**	**O**	**O**	**O**	**O**	**O**
4	**O**	**O**	**O**	**O**	**O**	**O**
5	**O**	**O**	**O**	**O**	**O**	**O**
6	**O**	**O**	**O**	**O**	**O**	**O**

A= avere almeno un 2

$$A = \left\{ \begin{matrix} (2,1),(2,2),(2,3),(2,4),(2,5), \\ (2,6),(1,2),(3,2),(4,2),(5,2),(6,2) \end{matrix} \right\}$$

$$P(A) = \frac{11}{36}$$

Nel caso generale



$$P\left(A\middle|B\right)=\frac{P\left(A\cap B\right)}{P\left(B\right)}$$

Esempio 2.2

Nel lancio di una coppia di dadi, determinare la probabilità di avere almeno un 2 sapendo che la somma è 6.

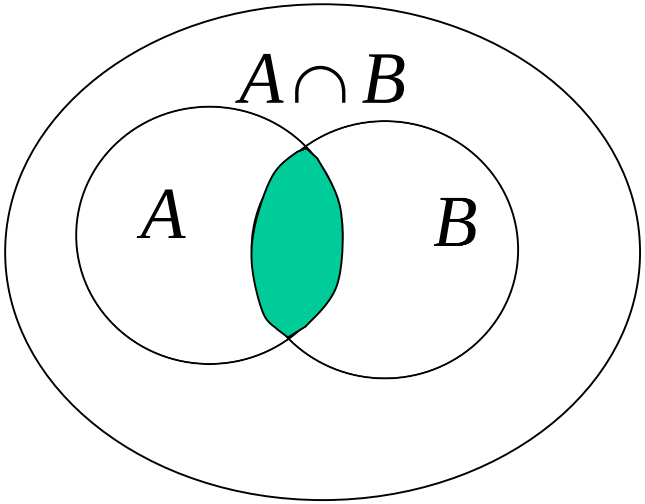
	1	**2**	**3**	**4**	**5**	**6**
1	O	O	O	O	O	O
2	O	O	O	O	O	O
3	O	O	O	O	O	O
4	O	O	O	O	O	O
5	O	O	O	O	O	O
6	O	O	O	O	O	O

B = la somma è 6

$$B=\{(1,5),(2,4),(3,3),(4,2),(5,1)\}$$

$$P(B)=\frac{5}{36}$$

Nel caso generale



$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}}$$

Esempio 2.2

Nel lancio di una coppia di dadi, determinare la probabilità di avere almeno un 2 sapendo che la somma è 6.

	1	**2**	**3**	**4**	**5**	**6**
1	O	O	O	O	O	O
2	O	O	O	O	O	O
3	O	O	O	O	O	O
4	O	O	O	O	O	O
5	O	O	O	O	O	O
6	O	O	O	O	O	O

$$A \cap B = \{(2,4),(4,2)\}$$

$$P(A \cap B) = \frac{2}{36}$$

$$P(A/B) = \frac{\frac{2}{36}}{\frac{5}{36}} = \frac{2}{5}$$

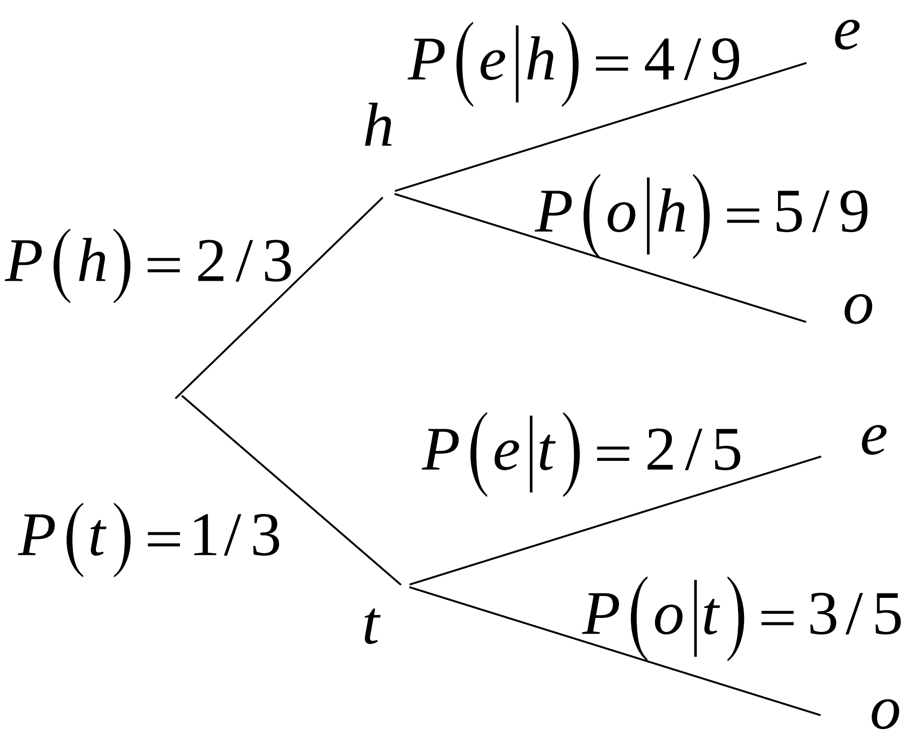
Esempio 2.3

Una moneta truccata (*unfair coin*) è tale per cui

$$P(h) = 2/3; \quad P(t) = 1/3$$

Lanciata la moneta, si sceglie 'a caso' un numero tra 1 e 9 se esce testa (*head*) e un numero tra 1 e 5 se esce croce (*tail*): qual è la probabilità di avere alla fine un numero pari (*even*)?

Processo stocastico e diagramma ad albero

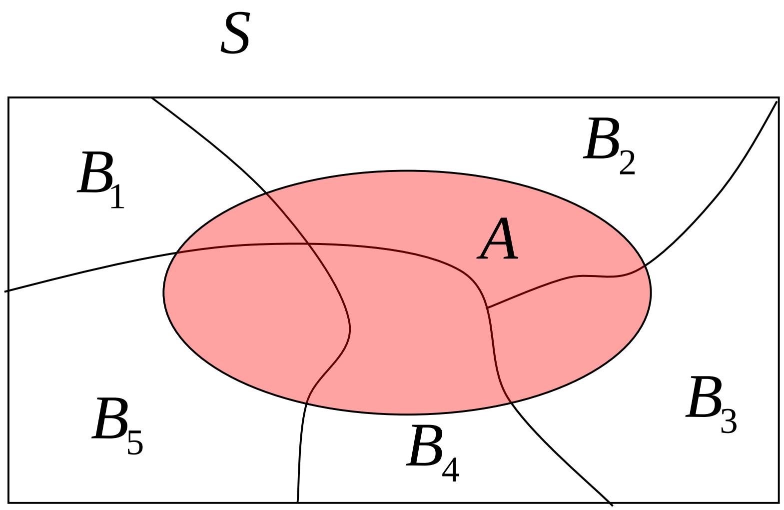


$$P(e) =$$

$$P(h) \cdot P(e|h) + P(t) \cdot P(e|t) =$$

$$\frac{2}{3}\frac{4}{9} + \frac{1}{3}\frac{2}{5} = \frac{58}{135}$$

Partizioni e probabilità condizionate



$$P\left(B_i \cap B_j\right) = \varnothing \quad se \;\; i \neq j$$

$$\bigcup_i B_i = S$$

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \dots. =$$

$$= P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2) + \dots.$$

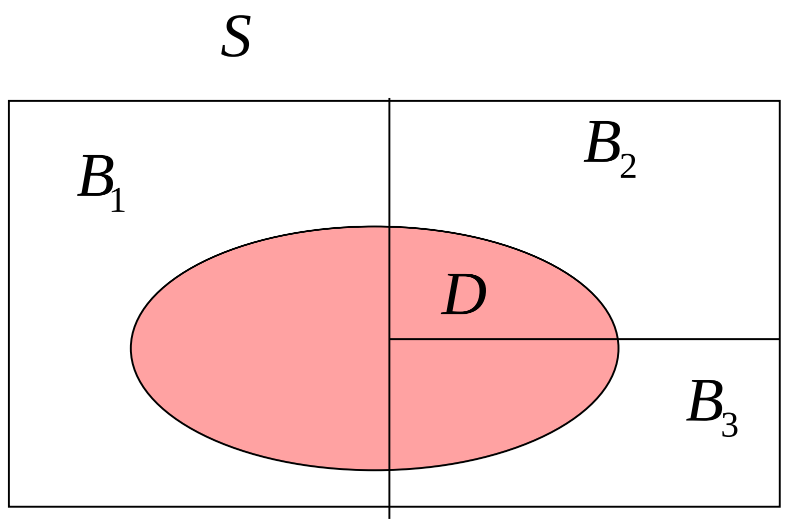
Esempio 2.5

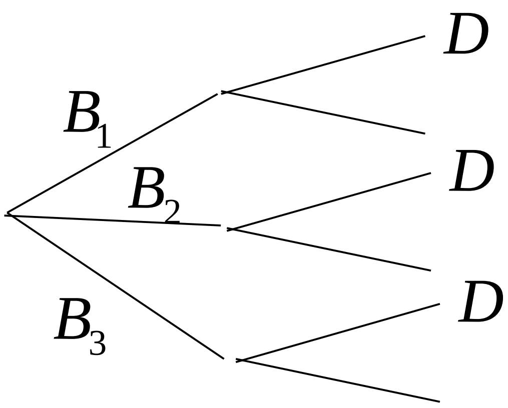
Tre macchine producono rispettivamente il 50%, 30% e il 20% della produzione, le percentuali di pezzi difettosi di ognuna di esse sono rispettivamente: 3%, 4% e 5%. Determinare la percentuale di pezzi difettosi della produzione.

$$P(B_1)=0.5;\ P(B_2)=0.3;\ P(B_3)=0.2$$

$$P(D/B_1)=0.03 \quad P(D/B_2)=0.04$$

$$P(D/B_3)=0.05$$

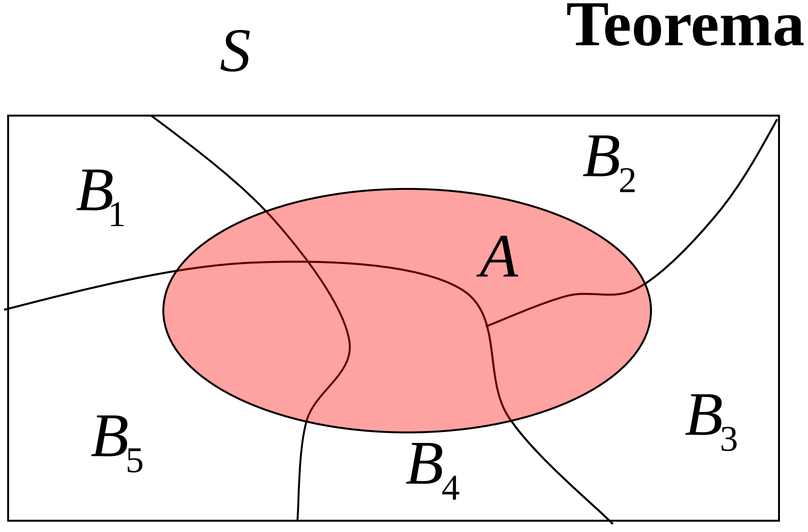




$$P(D)= P(D|B_1)\cdot P(B_1)+ P(D|B_2)\cdot P(B_2)+P(D|B_3)\cdot P(B_3)=$$

$$=0.03\cdot 0.5+0.04\cdot 0.3+0.05\cdot 0.2=0.037=3.7\%$$

Teorema di Bayes



$$P\left(B_1 \middle| A\right) = ?$$

$$P\left(B_1 \middle| A\right) = \frac{P\left(B_1 \cap A\right)}{P\left(A\right)} = \frac{P\left(A \cap B_1\right)}{P\left(A\right)} = \frac{P\left(A \middle| B_1\right) \cdot P\left(B_1\right)}{P\left(A\right)}$$

$$P\left(A\right) = P\left(A \middle| B_1\right) P\left(B_1\right) + P\left(A \middle| B_2\right) P\left(B_2\right) + \ldots\ldots\ldots$$

$$P\left(B_1 \middle| A\right) = \frac{P\left(A \middle| B_1\right) \cdot P\left(B_1\right)}{P\left(A \middle| B_1\right) P\left(B_1\right) + P\left(A \middle| B_2\right) P\left(B_2\right) + \ldots\ldots\ldots}$$

Esempio 2.6

Nell’esempio 2.5, determinare la probabilità che un pezzo difettoso sia stato prodotto dalla prima macchina

$$P\left(B_1 \middle| D\right) = ?$$

$$P\left(B_1 \middle| D\right) = \frac{P\left(D \middle| B_1\right) \cdot P\left(B_1\right)}{P\left(D\right)} = \frac{0.03 \cdot 0.5}{0.037} = 0.405 = 40.5\%$$

$$P\left(B_2 \middle| D\right) = 0.324$$

$$P\left(B_3 \middle| D\right) = 0.27$$

Eventi indipendenti

Definizione: *A* e *B* sono indipendenti se il fatto che si verifichi l'uno non modifica la probabilità che si verifichi l'altro:

$$P\left(A\middle|B\right) = P\left(A\right)$$

ma

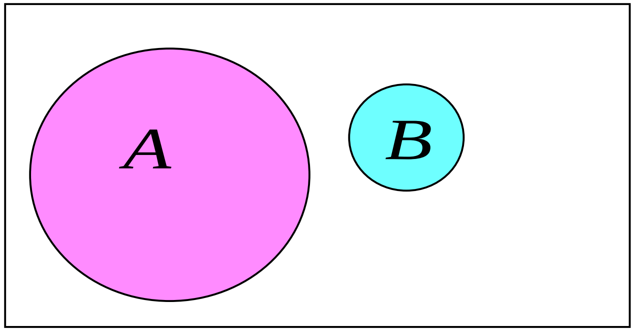
$$P\left(A\middle|B\right) \cdot P\left(B\right) = P\left(A \cap B\right)$$

Quindi sono indipendenti se e solo se:

$$P\left(A \cap B\right) = P\left(A\right) \cdot P\left(B\right)$$

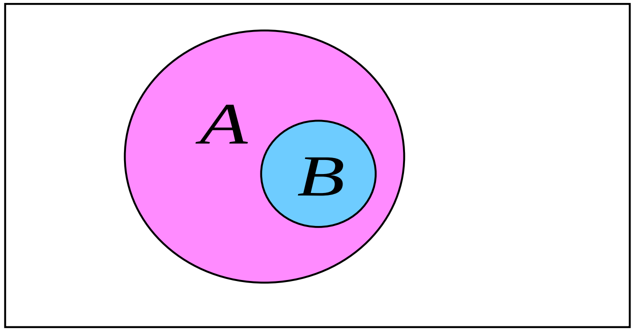
Esercizio 2.1

Verificare che due eventi A e B non impossibili e mutuamente esclusivi **non sono** indipendenti

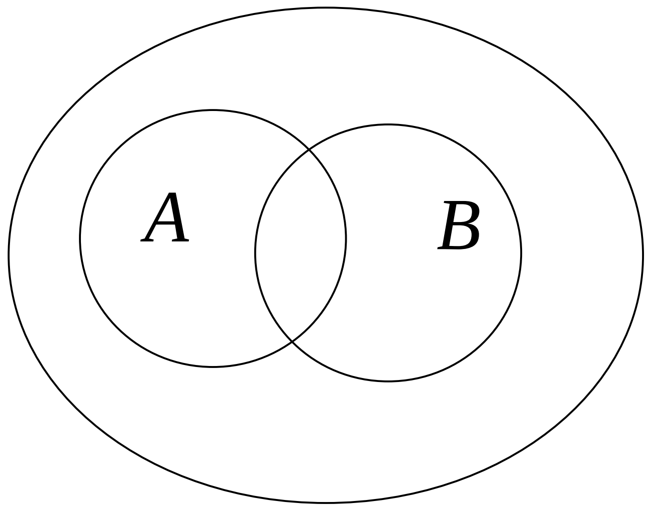


Esercizio 2.2

Verificare che due eventi A e B con $B \subset A$ e $B \neq \varnothing$ (non impossibili e inclusivi) **non sono** indipendenti



Ne consegue che:
due eventi *A* e *B* non impossibili per essere indipendenti devono essere come rappresentato:



$$P(A/B) = \frac{P(A \cap B)}{P(B)} = P(A)$$

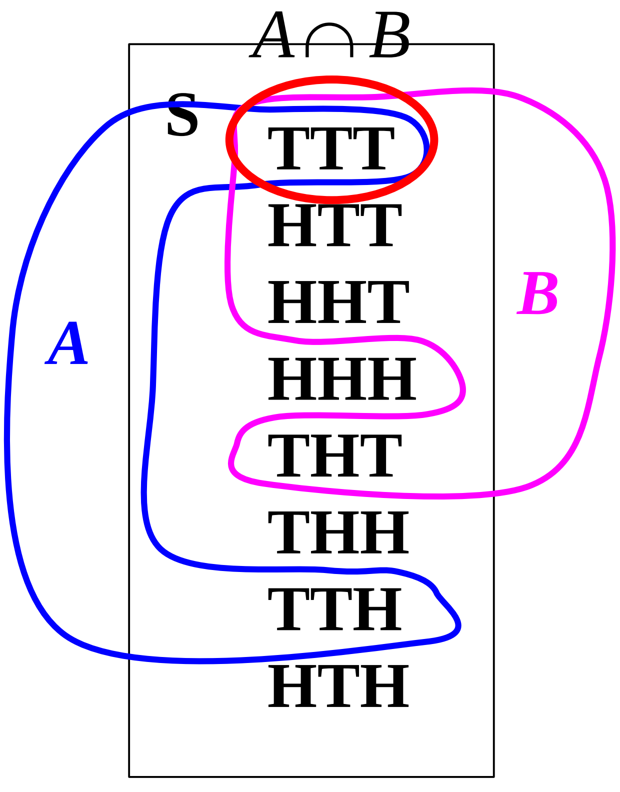
La probabilità che ha un elemento di *B* di appartenere ad *A* è la stessa di un elemento qualsiasi di *S*

Probabilità semplice: spazi campionari finiti

Lancio tre volte una moneta e definisco i seguenti eventi:

- A=i primi due lanci hanno come risultato "tie"
- B= l'ultimo lancio ha come risultato "tie"

A e B sono indipendenti?



$$P(A) = \frac{2}{8} = \frac{1}{4}$$

$$P(B) = \frac{4}{8} = \frac{1}{2}$$

$$P(A \cap B) = \frac{1}{8}$$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{8}}{\frac{1}{2}} = \frac{1}{4} = P(A)$$

A e B sono indipendenti

Esempio 2.7

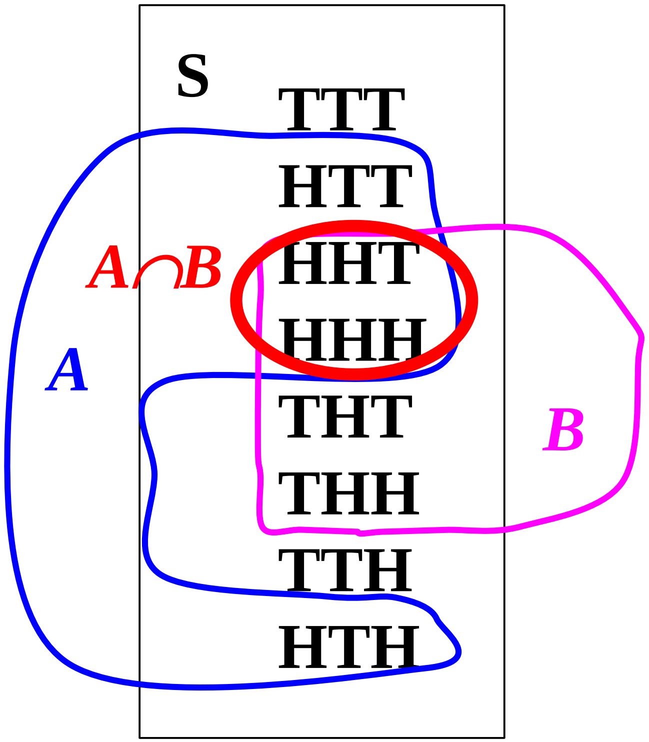
Consideriamo un esperimento consistente in tre lanci di una moneta, (*fair coin*) valutare la dipendenza degli eventi:

A: il primo lancio produce testa

B: il secondo lancio produce testa

C: testa si presenta due sole volte

$$P(A) = P(B) = \frac{4}{8} = \frac{1}{2}$$



$$P(A \cap B) = \frac{2}{8} = \frac{1}{4}$$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2} = P(A)$$

A e *B* indipendenti

Esempio 2.7

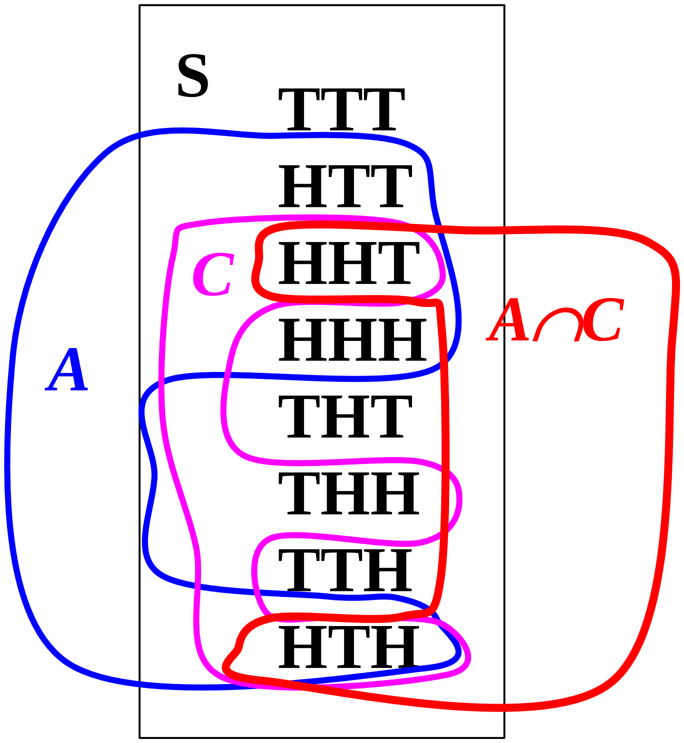
Consideriamo un esperimento consistente in tre lanci di una moneta, (*fair coin*) valutare la dipendenza degli eventi:

A: il primo lancio produce testa

B: il secondo lancio produce testa

C: testa si presenta due sole volte

$$P(A) = P(B) = \frac{4}{8} = \frac{1}{2}$$



$$P(C) = \frac{3}{8}$$

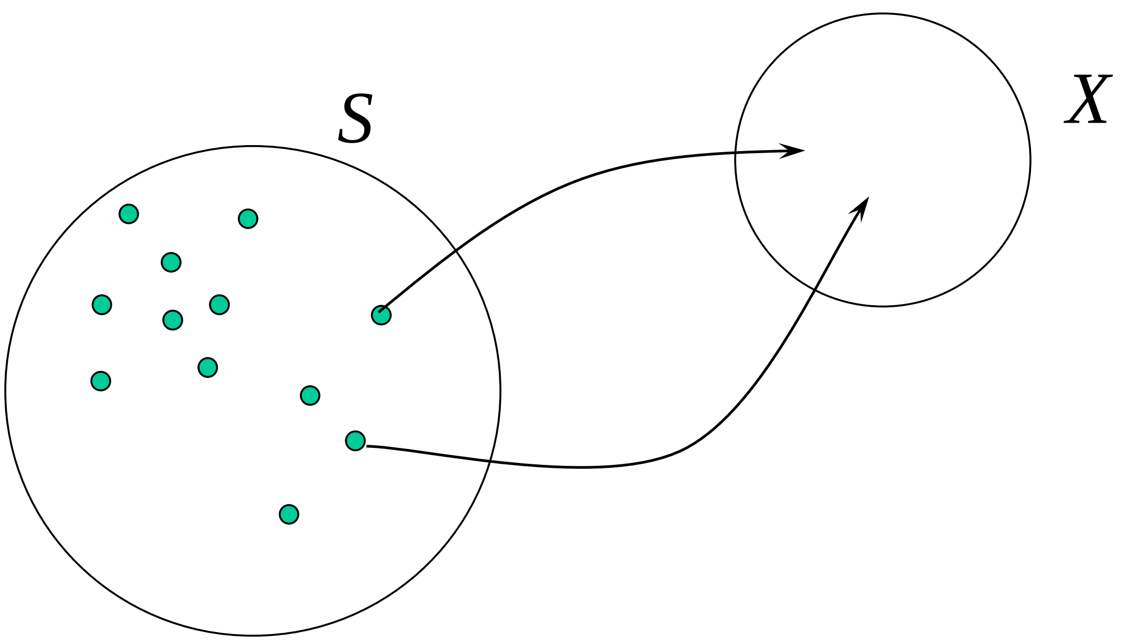
$$P(A \cap C) = \frac{2}{8} = \frac{1}{4}$$

$$P(A/C) = \frac{P(A \cap C)}{P(C)} = \frac{\frac{1}{4}}{\frac{3}{8}} = \frac{8}{12} = \frac{2}{3} \neq P(A)$$

A e C non indipendenti

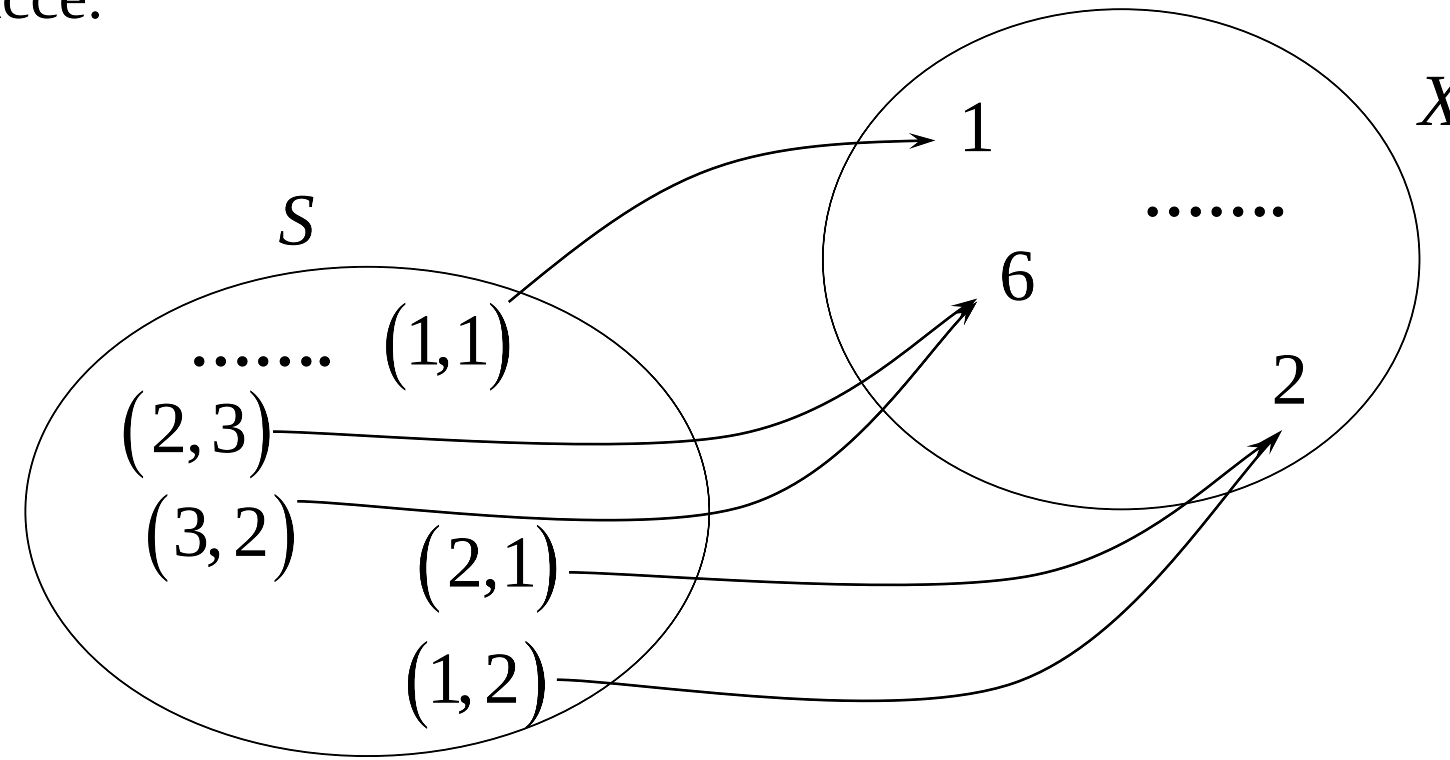
Variabile casuale (*random variable*)

Insieme numerico



Esempio 2.8

Lancio due dadi e la variabile casuale è il prodotto dei valori delle facce.



$$X = \{1; 2; 3; 4; 5; 6; 8; 9; 10; 12; 15; 16; 18; 20; 24; 25; 30; 36\}$$

Variabile casuale (*random variable*): definizione

Una VC (RV) o Variabile Aleatoria (VA) su uno spazio campionario S è una trasformazione da S in $\mathbb{R}$ tale che ogni evento di S sia rappresentato come valore numerico appartenente all'insieme: $\boldsymbol{X} \subseteq \mathbb{R}$

Variabile casuale discreta

$\boldsymbol{X} \subseteq \mathbb{R}$ è un insieme **numerabile** di elementi (finito o infinito) a ognuno dei quali può essere associata una probabilità.

Esempio di variabile casuale discreta non finita:
X contiene il numero di lanci di un dado prima che si presenti un numero primo:

$$\boldsymbol{X}=\{1;2;3;……..\}=\mathbb{N}$$

Variabile casuale discreta

$$x_1 \qquad x_2 \qquad x_3 \qquad x_4 \qquad x_5 \quad \ldots\ldots\ldots\ldots \quad x_n$$

$$x_i \in X$$

$$p(x_1) \quad p(x_2) \quad p(x_3) \quad p(x_4) \quad p(x_5) \ldots\ldots\ldots\ldots \quad p(x_n)$$

Funzione o distribuzione di probabilità discreta: $P_i = p(x_i)$

Proprietà:

$$P_i = p(x_i) \geq 0 \quad \forall x_i$$

$$\sum_{i=1}^{n} P_i = 1$$

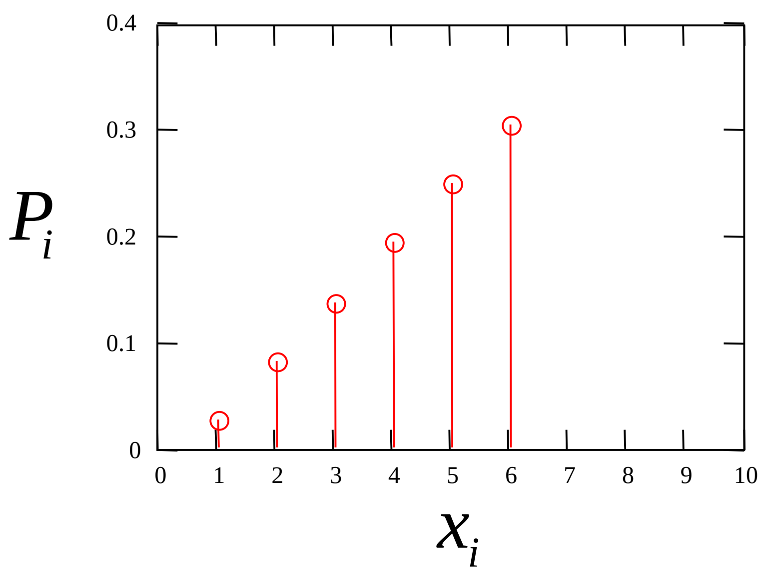
Esempio 2.9

Si lanciano due dadi e si considera la variabile aleatoria x il massimo valore: determinare il dominio della VA e la sua distribuzione

Evento	x_i	P_i
(1,1)	1	1/36
(1,2),(2,1),(2,2)	2	3/36
………		

$$X = \{1; 2; 3; 4; 5; 6\}$$

x_i	1	2	3	4	5	6
P_i	1/36	3/36	5/36	7/36	9/36	11/36

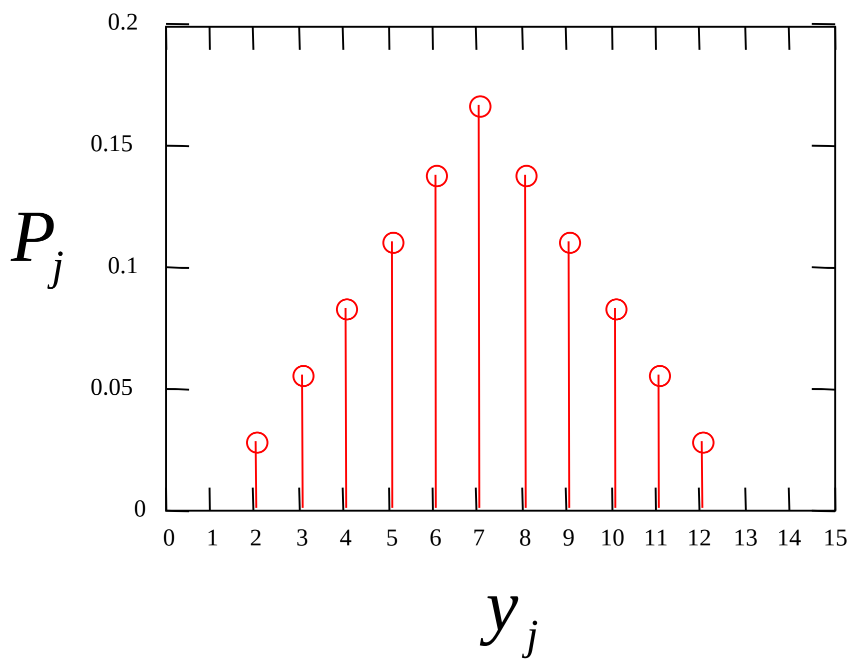


Esempio 2.10

Si lanciano due dadi e si considera la variabile casuale *y* come la somma dei valori: determinare dominio e distribuzione

$$Y = \{y_j\} = \{2;3;4;5;6;7;8;9;10;11;12\}$$

$$P_3 = P(y=4) = P\{(1,3);(3,1);(2,2)\} = \frac{3}{36}$$



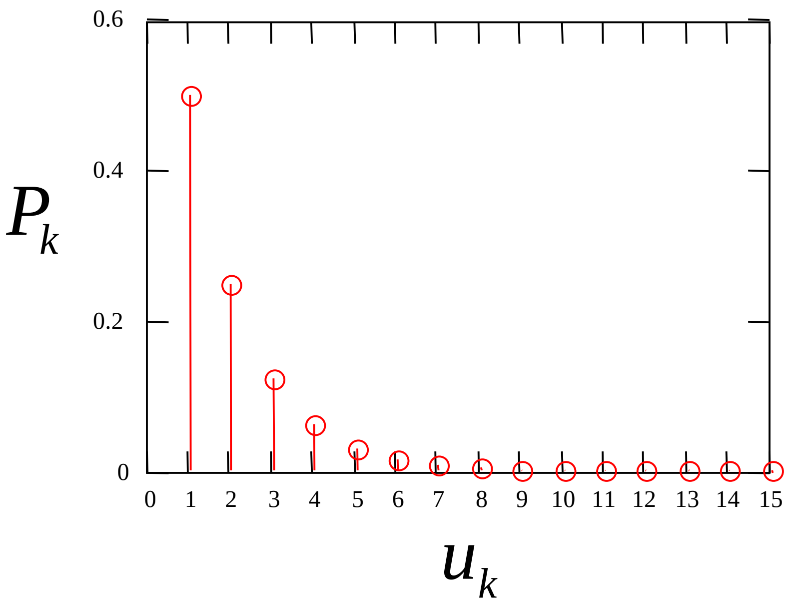
Esempio 2.11

Si lancia una moneta finché non compare testa, la variabile casuale u è il numero di lanci

$$\boldsymbol{U}=\{u_k\}=\{1;2;3;\ldots\ldots\}=\mathbb{N}$$

$$P_1 = P\{u=1\} = \frac{1}{2} \qquad P_2 = P\{(t),(h)\} = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{2^2}$$

$$P_5 = P\{(t),(t),(t),(t),(h)\} = \frac{1}{2^5} \qquad P_k = \frac{1}{2^k}$$



$$\sum_{k=1}^{\infty}\frac{1}{2^k} = ?$$

$$0.5+0.25+0.125+0.0625+\ldots\ldots=1$$

Variabile casuale discreta: proprietà della distribuzione

Valore centrale, valor medio, speranza matematica (sinonimi)
(*central*, *mean*, *expected value*)

$$\mu = \sum_{i=1}^{n} x_i P_i \qquad\qquad \mu = E\left(x_i\right) = E\left(X\right)$$

Significati:

1) Media dei risultati degli esperimenti ponderata sulle probabilità:

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

Variabile casuale discreta: significati della media

2) Se si effettua l’esperimento tante volte e si esegue la media aritmetica semplice dei risultati, **al limite** si ottiene μ

3) Se si ripete l’esperimento *m* volte (con $m>>1$) e si sommano i risultati, **il valore più probabile** è:

$$m \cdot \mu$$

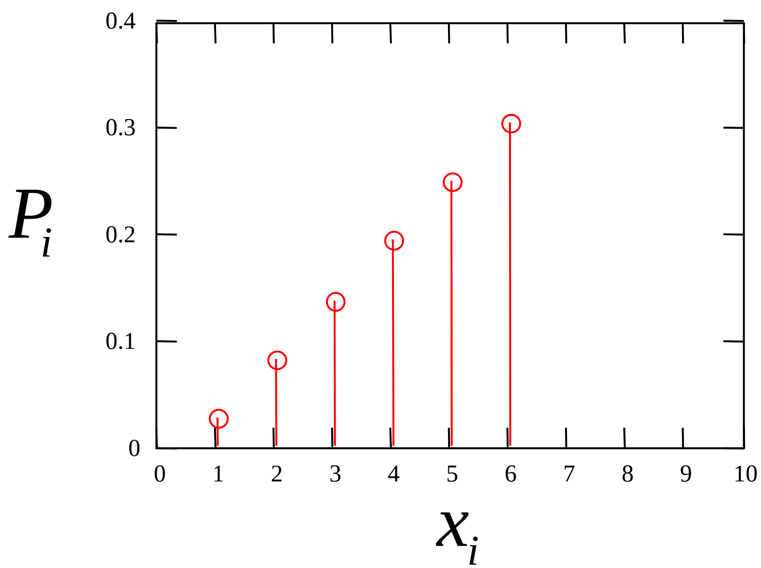
4) Analogia meccanica:
Se associo il valore della variabile casuale alla posizione di un punto disposto sull’asse *x* e la relativa probabilità alla massa del punto, il valor medio della variabile casuale individua la posizione del centro di massa (baricentro) della distribuzione

Esempio 2.9a

Si lanciano due dadi e si considera la variabile aleatoria *x* il massimo valore: determinare il valor medio della distribuzione

y_j	Esiti fav.	P_j	y_jP_j
1	1	0.027778	0.027778
2	3	0.083333	0.166667
3	5	0.138889	0.416667
4	7	0.194444	0.777778
5	9	0.25	1.25
6	11	0.305556	1.833333
		μ	4.472222

Dadi	Possibili esiti					
1\2	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	2	3	4	5	6
3	3	3	3	4	5	6
4	4	4	4	4	5	6
5	5	5	5	5	5	6
6	6	6	6	6	6	6

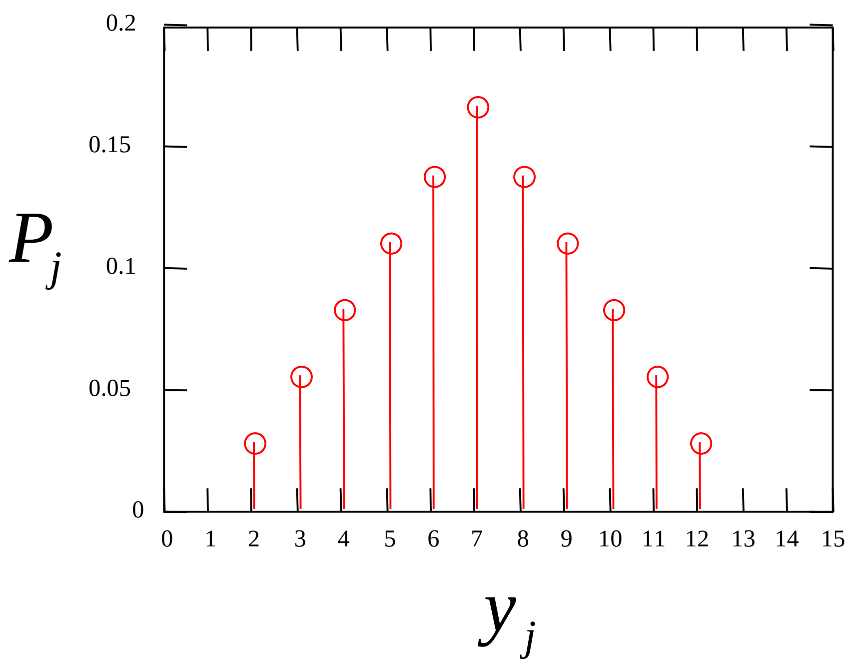


Esempio 2.10a

Si lanciano due dadi e si considera la variabile casuale *y* come la somma dei valori: determinare la media della distribuzione

y_j	Esiti fav.	P_j	y_jP_j
2	1	0.027778	0.055556
3	2	0.055556	0.166667
4	3	0.083333	0.333333
5	4	0.111111	0.555556
6	5	0.138889	0.833333
7	6	0.166667	1.166667
8	5	0.138889	1.111111
9	4	0.111111	1
10	3	0.083333	0.833333
11	2	0.055556	0.611111
12	1	0.027778	0.333333
		μ	7

Dadi	Possibili esiti					
1\2	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12



Esempio 2.12 Giochi equi.

Un gioco di sorte consiste nel lancio di un dado corretto (*fair dice*), se l'esito è un numero primo il giocatore riceve una vincita equivalente in euro, se l'esito non è primo il giocatore paga al banco l'equivalente in euro. Determinare la distribuzione della probabilità di vincita e valutarne la media.

x_i	2	3	5	-1	-4	-6
P_i	1/6	1/6	1/6	1/6	1/6	1/6

$$\mu = \sum_{i=1}^{n} x_i P_i = -\frac{1}{6}$$

Note:

La media può non essere nel dominio della VA (in questo caso non è nemmeno intera)

Mediamente il giocatore perde 1/6 di euro per ogni giocata

Esempio 2.13

Determinare il valor medio della distribuzione:

$$P_k = \frac{1}{2^k}$$

$$\mu_k = \sum_{k=1}^{\infty} \frac{k}{2^k} = 2$$

Nota:

Lanciando tante volte una moneta, il valor medio di lanci tra l'apparizione di due teste è 2

Variabile casuale discreta: proprietà della distribuzione

Misura di dispersione, larghezza (*scatter*): varianza (*variance*)

$$\sigma^2 = VAR(X) = \sum_{i=1}^{n} (x_i - \mu)^2 P_i \quad \text{scritto anche } \sigma_X^2$$

Considerando la VA:

$$d_i = x_i - \mu$$

$$\sigma^2 = VAR(X) = E\left(d_i^2\right)$$

Significati:

Media ponderata (con la pr.) dei quadrati delle distanze dalla media.

Analogia meccanica: momento d'inerzia baricentrico.

Attenzione alle dimensioni!

Esercizio 2.3

Verificare che $\sigma^2 = E\left(x_i^2\right) - \mu^2$

$$\sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 P_i = \sum_{i=1}^{n} \left(x_i^2 - 2x_i\mu + \mu^2\right)^2 P_i =$$

$$= \sum_{i=1}^{n} x_i^2 P_i - 2x_i\mu P_i + \mu^2 P_i = E\left(x_i^2\right) - 2\mu \sum_{i=1}^{n} x_i P_i + \mu^2 \sum_{i=1}^{n} P_i =$$

$$E\left(x_i^2\right) - 2\mu^2 + \mu^2 = E\left(x_i^2\right) - \mu^2$$

Scarto quadratico medio o deviazione standard (*standard deviation*)

$$\sigma = \sqrt{VAR(X)} = \sqrt{\sum_{i=1}^{n} (x_i - \mu)^2 P_i}$$

Significati:

- Distanza quadratica media ponderata dei valori dal centro
- Dimensionalmente omogenea con *x*
- Analogia meccanica: raggio d'inerzia

Esercizio 2.4

Verificare che:

1. $k \in \Re;\ VAR(kX) = k^2 VAR(X);\ \sigma_{kX} = |k| \sigma_X$
2. c costante omogena a x: $\sigma_{X+c} = \sigma_X$

Definizione

Variabile casuale standardizzata:

$$X^* = \frac{X - \mu}{\sigma}$$

Esercizio 2.5

Verificare che:

1. $E\left(X^*\right) = 0$
2. $VAR\left(X^*\right) = 1$

$$k \in \Re;\ \ VAR(kX) = k^2 VAR(X);\ \ \sigma_{kX} = |k|\sigma_X$$

$$\sigma_{kx}^2 = \sum_{i=1}^{n} \left(kx_i - k\mu\right)^2 P_i = k^2 \sum_{i=1}^{n} \left(x_i - \mu\right)^2 P_i = k^2 \sigma^2$$

c costante omogena a x: $\quad \sigma_{X+c} = \sigma_X$

$$\sigma_{x+c}^2 = \sum_{i=1}^{n} \left(\left(x_i + c\right) - \left(\mu + c\right)\right)^2 P_i = \sum_{i=1}^{n} \left(x_i - \mu\right)^2 P_i = \sigma^2$$

$$X^* = \frac{X-\mu}{\sigma}$$

$$E\left(X^*\right) = 0$$

$$\mu_{X^*} = \sum_{i=1}^{n} \frac{\left(x_i - \mu\right)}{\sigma} P_i = \frac{1}{\sigma}\left(\sum_{i=1}^{n} x_i P_i - \mu \sum_{i=1}^{n} P_i\right) = \frac{1}{\sigma}\left(\mu - \mu\right) = 0$$

$$VAR\left(X^*\right) = 1$$

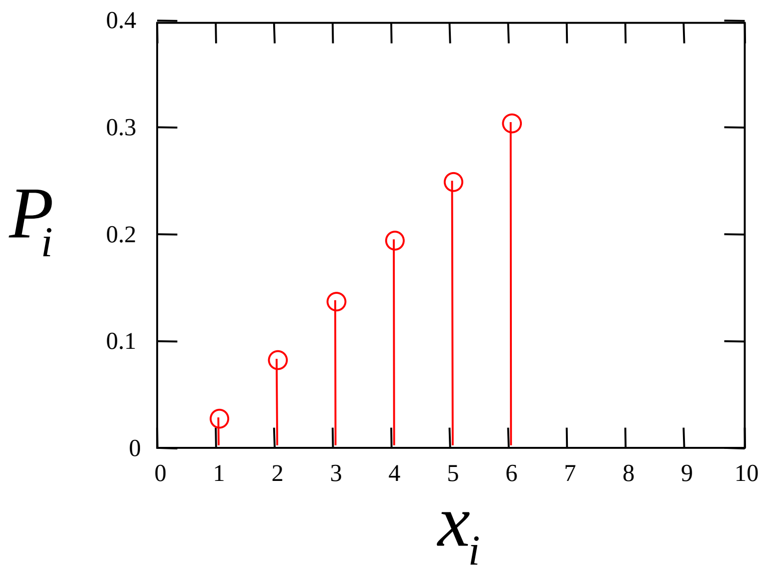
$$\sigma_{X^*}^2 = \sum_{i=1}^{n}\left(\frac{\left(x_i - \mu\right)}{\sigma}\right)^2 P_i = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(x_i - \mu\right)^2 P_i = \frac{\sigma^2}{\sigma^2} = 1$$

Esempio 2.9b

Si lanciano due dadi e si considera la variabile aleatoria *x* il massimo valore: determinare la varianza e la deviazione standard della distribuzione

y_j	Esiti fav.	P_j	$(y_j-\mu)^2P_j$
1	1	0.027778	0.334898
2	3	0.083333	0.509324
3	5	0.138889	0.301033
4	7	0.194444	0.04336
5	9	0.25	0.069637
6	11	0.305556	0.713199
		μ	4.472222
		σ^2	1.971451
		σ	1.404084

Dadi	Possibili esiti					
1\2	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	2	3	4	5	6
3	3	3	3	4	5	6
4	4	4	4	4	5	6
5	5	5	5	5	5	6
6	6	6	6	6	6	6

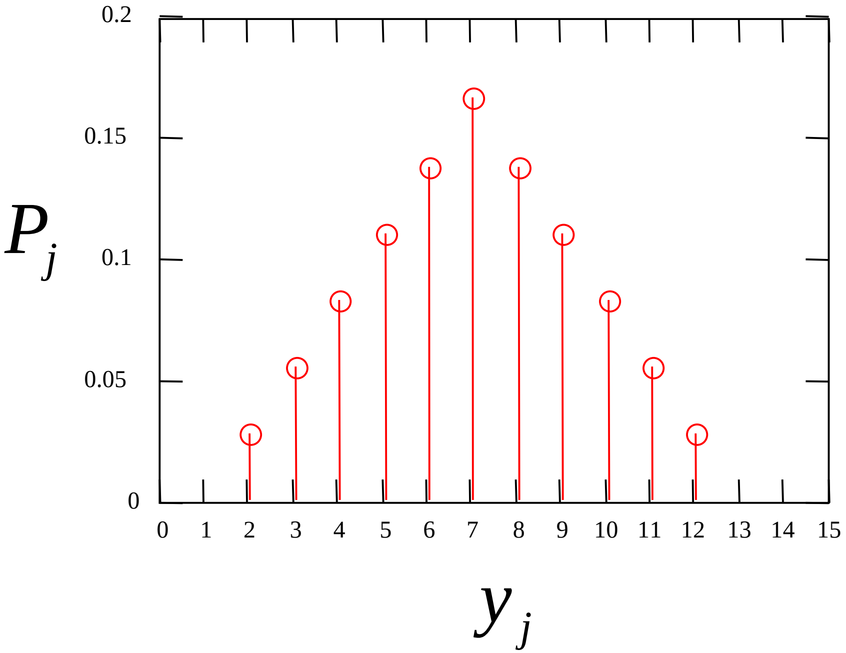


Esempio 2.10b

Si lanciano due dadi e si considera la variabile casuale *y* come la somma dei valori: determinare la varianza e la distribuzione standard della distribuzione

y_j	Esiti fav.	P_j	$(y_j-\mu)^2 P_j$
2	1	0.027778	0.694444
3	2	0.055556	0.888889
4	3	0.083333	0.75
5	4	0.111111	0.444444
6	5	0.138889	0.138889
7	6	0.166667	0
8	5	0.138889	0.138889
9	4	0.111111	0.444444
10	3	0.083333	0.75
11	2	0.055556	0.888889
12	1	0.027778	0.694444
		μ	7
		σ^2	5.833333
		σ	2.415229

Dadi	Possibili esiti					
1\2	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12



Variabili casuali discrete: distribuzione congiunta (*joint distribution*)

Date due VA: *X* e *Y* definite sullo stesso *S*

$$X = \left\{ x_1, x_2, \ldots\ldots, x_n \right\} \quad \text{e} \quad Y = \left\{ y_1, y_2, \ldots, y_m \right\}$$

Il prodotto cartesiano è l’insieme di tutte le *n·m* coppie:

$$X \times Y = \left\{ \left(x_1, y_1 \right), \left(x_1, y_2 \right), \ldots\ldots\ldots, \left(x_n, y_m \right) \right\}$$

Definiamo la probabilità congiunta:

$$P_{ij} = P\left\{ \left(x = x_i \right) \cap \left(y = y_j \right) \right\}$$

Matrice riassuntiva

$X \backslash Y$	y_1	y_2	y_3	.....	y_m	somma
x_1	P_{11}	P_{12}	P_{13}	.....	P_{1m}	Px_1
x_2	P_{21}	P_{22}	P_{23}	.....	P_{2m}	Px_2
x_3	P_{31}	P_{32}	P_{33}	.....	P_{3m}	Px_3
			....			
x_n	P_{n1}	P_{n2}	P_{n3}	.....	P_{nm}	Px_n
somma	Py_1	Py_2	Py_3	.....	Py_m	

Esempio 2.14

Matrice delle probabilità congiunte per le variabili degli esempi 2.9 e 2.10 (Valore massimo e somma nel lancio di due dadi)

	2	3	4	5	6	7	8	9	10	11	12	
1	1	0	0	0	0	0	0	0	0	0	0	1
2	0	2	1	0	0	0	0	0	0	0	0	3
3	0	0	2	2	1	0	0	0	0	0	0	5
4	0	0	0	2	2	2	1	0	0	0	0	7
5	0	0	0	0	2	2	2	2	1	0	0	9
6	0	0	0	0	0	2	2	2	2	2	1	11
	1	2	3	4	5	6	5	4	3	2	1	

Tutti i valori devono essere divisi per 36

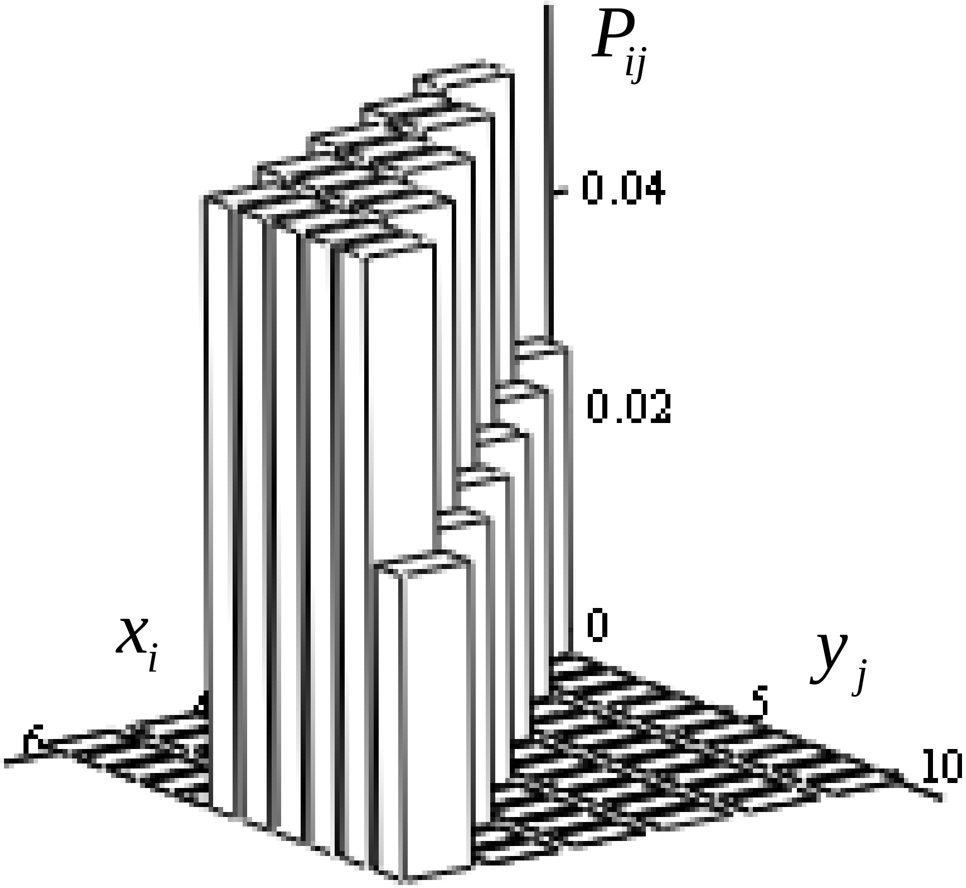
Esempio 2.14

Matrice delle probabilità congiunte per le variabili degli esempi 2.9 e 2.10 (Valore massimo e somma nel lancio di due dadi)

	2	3	4	5	6	7	8	9	10	11	12	
1	1	0	0	0	0	0	0	0	0	0	0	1
2	0	2	1	0	0	0	0	0	0	0	0	3
3	0	0	2	2	1	0	0	0	0	0	0	5
4	0	0	0	2	2	2	1	0	0	0	0	7
5	0	0	0	0	2	2	2	2	1	0	0	9
6	0	0	0	0	0	2	2	2	2	2	1	11
	1	2	3	4	5	6	5	4	3	2	1	

2 casi su 36 (3+1)
2 casi su 36 (3+2)
1 caso su 36 (3+3)

Rappresentazione della matrice delle probabilità congiunte per le variabili degli esempi 2.9 e 2.10



Proprietà:

$$P_{ij} \geq 0 \quad \forall \left(x_i , y_j \right) \qquad\qquad \sum_{i=1}^{n} \sum_{j=1}^{m} P_{ij} = 1$$

Definizione di covarianza di *X* e *Y*

$$COV\left(X , Y \right) = \sum_{i=1}^{n} \sum_{j=1}^{m} \left(x_i - \mu_x \right)\left(y_j - \mu_y \right) P_{ij} = E\left[\left(X - \mu_x \right)\left(Y - \mu_y \right) \right]$$

Significati della covarianza

- Dimensionalmente omogenea alla varianza
- Può essere positiva negativa o nulla
- Analogia meccanica: momento centrale centrifugo o misto

Definizione di correlazione (lineare) di *X* e *Y*

$$\rho\left(X , Y \right) = \frac{COV\left(X , Y \right)}{\sigma_X \sigma_Y}$$

Esercizio 2.6

Per i dati della tabella precedente verificare che:

$$COV\left(X,Y\right)=2.9$$

$$\rho\left(X,Y\right)=0.86$$

Dadi	Possibili esiti somma						μ	7
1\2	1	2	3	4	5	6	σ^2	5.833333
1	2	3	4	5	6	7	σ	2.415229
2	3	4	5	6	7	8		
3	4	5	6	7	8	9		
4	5	6	7	8	9	10		
5	6	7	8	9	10	11		
6	7	8	9	10	11	12		

Dadi	Possibili esiti Valore massimo						μ	4.472222
1\2	1	2	3	4	5	6	σ^2	1.971451
1	1	2	3	4	5	6	σ	1.404084
2	2	2	3	4	5	6		
3	3	3	3	4	5	6		
4	4	4	4	4	5	6		
5	5	5	5	5	5	6		
6	6	6	6	6	6	6		

	Tabella probabilità congiunte										
	2	3	4	5	6	7	8	9	10	11	12
1	1	0	0	0	0	0	0	0	0	0	0
2	0	2	1	0	0	0	0	0	0	0	0
3	0	0	2	2	1	0	0	0	0	0	0
4	0	0	0	2	2	2	1	0	0	0	0
5	0	0	0	0	2	2	2	2	1	0	0
6	0	0	0	0	0	2	2	2	2	2	1

	Tabella contributi covarianza										
	2	3	4	5	6	7	8	9	10	11	12
1	0.482253	0	0	0	0	0	0	0	0	0	0
2	0	0.549383	0.206019	0	0	0	0	0	0	0	0
3	0	0	0.24537	0.16358	0.040895	0	0	0	0	0	0
4	0	0	0	0.052469	0.026235	0	-0.01312	0	0	0	0
5	0	0	0	0	-0.02932	0	0.029321	0.058642	0.043981	0	0
6	0	0	0	0	0	0	0.084877	0.169753	0.25463	0.339506	0.212191

COV	2.916667
ρ	0.860073

NB : in Excel esistono apposite funzioni

Esercizio 2.7

Dimostrare che in generale:

$$COV(Y,X) = COV(X,Y)$$

$$COV(X,Y) = E(XY) - \mu_x \mu_y$$

Esercizio 2.7

Dimostrare che in generale:

$$COV(Y,X) = COV(X,Y)$$

$$COV(X,Y) = E(XY) - \mu_x \mu_y$$

$$COV(X,Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} (x_i - \mu_x)(y_j - \mu_y) P_{ij} =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} (x_i y_j - x_i \mu_y - y_j \mu_x + \mu_x \mu_y) P_{ij} =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j P_{ij} - \mu_y \sum_{i=1}^{n} \sum_{j=1}^{m} x_i P_{ij} - \mu_x \sum_{i=1}^{n} \sum_{j=1}^{m} y_j P_{ij} + \mu_x \mu_y \sum_{i=1}^{n} \sum_{j=1}^{m} P_{ij} =$$

$$= E(x_i y_j) - \mu_x \mu_y - \mu_x \mu_y + \mu_x \mu_y = E(x_i y_j) - \mu_x \mu_y$$

Proprietà della correlazione

$$\rho\left(X,X\right)=1;\quad \rho\left(X,-X\right)=-1$$

$$-1\leq\rho\left(X,Y\right)\leq 1$$

$$\rho\left(aX+b,cY+d\right)=\rho\left(X,Y\right)$$

Se due variabili aleatorie sono correlate positivamente: $\rho>0$, quando una è sopra la propria media anche l'altra tende a esserlo

Se due variabili aleatorie sono correlate negativamente: $\rho<0$, quando una è sopra la propria media l'altra tende a essere sotto

Due variabili aleatorie con $\rho=0$ (o comunque vicino a zero) si dicono scorrelate (linearmente)

Variabili aleatorie indipendenti

X e Y sono indipendenti se:

$$P\left\{\left(X = x_i\right)\cap\left(Y = y_j\right)\right\} = P\left(X = x_i\right)\cdot P\left(Y = y_j\right)$$

$$P_{ij} = PX_i \cdot PY_j$$

La matrice delle probabilità congiunte è ottenuta dal prodotto delle distribuzioni marginali

Esercizio 2.8
Verificare che le variabili della tabella precedente non sono indipendenti

y_j	Esiti fav.	P_j	y_jP_j
2	1	0.027778	0.055556
3	2	0.055556	0.166667
4	3	0.083333	0.333333
5	4	0.111111	0.555556
6	5	0.138889	0.833333
7	6	0.166667	1.166667
8	5	0.138889	1.111111
9	4	0.111111	1
10	3	0.083333	0.833333
11	2	0.055556	0.611111
12	1	0.027778	0.333333
		μ	7

Dadi	Possibili esiti					
1\2	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

y_j	Esiti fav.	P_j	y_jP_j
1	1	0.027778	0.027778
2	3	0.083333	0.166667
3	5	0.138889	0.416667
4	7	0.194444	0.777778
5	9	0.25	1.25
6	11	0.305556	1.833333
		μ	4.472222

Dadi	Possibili esiti					
1\2	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	2	3	4	5	6
3	3	3	3	4	5	6
4	4	4	4	4	5	6
5	5	5	5	5	5	6
6	6	6	6	6	6	6

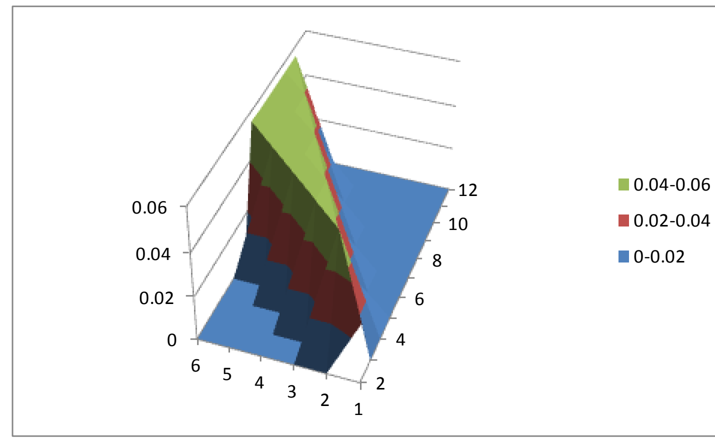


	Tabella probabilità congiunte										
	2	3	4	5	6	7	8	9	10	11	12
1	0.027778	0	0	0	0	0	0	0	0	0	0
2	0	0.055556	0.027778	0	0	0	0	0	0	0	0
3	0	0	0.055556	0.055556	0.027778	0	0	0	0	0	0
4	0	0	0	0.055556	0.055556	0.055556	0.027778	0	0	0	0
5	0	0	0	0	0.055556	0.055556	0.055556	0.055556	0.027778	0	0
6	0	0	0	0	0	0.055556	0.055556	0.055556	0.055556	0.055556	0.027778

	Tabella probabilità marginali										
	2	3	4	5	6	7	8	9	10	11	12
1	0.000772	0.001543	0.002315	0.003086	0.003858	0.00463	0.003858	0.003086	0.002315	0.001543	0.000772
2	0.002315	0.00463	0.006944	0.009259	0.011574	0.013889	0.011574	0.009259	0.006944	0.00463	0.002315
3	0.003858	0.007716	0.011574	0.015432	0.01929	0.023148	0.01929	0.015432	0.011574	0.007716	0.003858
4	0.005401	0.010802	0.016204	0.021605	0.027006	0.032407	0.027006	0.021605	0.016204	0.010802	0.005401
5	0.006944	0.013889	0.020833	0.027778	0.034722	0.041667	0.034722	0.027778	0.020833	0.013889	0.006944
6	0.008488	0.016975	0.025463	0.033951	0.042438	0.050926	0.042438	0.033951	0.025463	0.016975	0.008488

Proprietà delle variabili aleatorie indipendenti

Se X e Y sono VA indipendenti:

$$E(XY) = E(X) \cdot E(Y)$$

$$VAR(X+Y) = VAR(X) + VAR(Y)$$

Relazione generalizzabile per *n* variabili indipendenti:

Se $X_1, X_2, ..., X_n$ sono VA indipendenti:

$$VAR(X_1 + X_2 + ... + X_n) = VAR(X_1) + VAR(X_2) + ... + VAR(X_n)$$

$$COV(X,Y) = 0$$

Due variabili indipendenti sono scorrelate (implicazione semplice)