$$\min_{a,b} \sum_{i=1}^{m} \left[y_i - (a x_i - b)\right]^2$$

TROVIAMO $a, b \in \mathbb{R}$ MINIMIZZANTI:

$$\sum_i \left[(y_i - a x_i) - b\right]^2 \longrightarrow$$ IL VALORE MINIMO DI b È $b = \bar{y} - a\bar{x}$

$\Longrightarrow$ CERCO a ... $\sum_i \left[y_i - a x_i - (\bar{y} - a\bar{x})\right]^2$

$$= \sum_i \left[y_i - a x_i - \bar{y} + a\bar{x}\right]^2 = \sum_i \left[(y_i - \bar{y}) - a(x_i - \bar{x})\right]^2$$

CHIAMIAMO QUESTA FUNZIONE $S(a)$ E NE FACCIAMO LA DERIVATA PER OTTENERE IL MINIMO

$$S(a) = \sum_i \left[(y_i - \bar{y}) - a(x_i - \bar{x})\right]^2$$

$$S'(a) = \sum_i 2\left[(y_i - \bar{y}) - a(x_i - \bar{x})\right] \cdot \left[-(x_i - \bar{x})\right]$$

$$= -2\sum_i \left[(y_i - \bar{y}) - a(x_i - \bar{x})\right] \cdot (x_i - \bar{x})$$

$$= -2\sum_i (y_i - \bar{y})(x_i - \bar{x}) + 2a\sum_i (x_i - \bar{x})^2 \geq 0$$

$$\text{per } \sum_i (x_i - \bar{x})^2 \geq \cancel{2} \sum_i (y_i - \bar{y})(x_i - \bar{x})$$

$$a \geq \frac{\sum_i (y_i - \bar{y})(x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2} = \frac{\sigma_{xy}}{\sigma_x^2}$$

RICAPITOLANDO, ABBIAMO OTTENUTO CHE SE $\rho_{xy} \neq 0$, LA RETTA DI REGRESSIONE LINEARE HA EQ. $Y = aX + b$, DOVE $a = \frac{\sigma_{xy}}{\sigma_x^2}$, $b = \bar{y} - a\bar{x}$.

ESEMPI: 1) CALCOLIAMO LA RETTA DI REGRESS. NELL' ES. PRECEDENTE

$$\{(1,1), (5,9), (4,7), (10,19), (2,3)\}$$

$Y = aX + b$ CON $a = \frac{\sigma_{xy}}{\sigma_x^2}$, $b = \bar{y} - a\bar{x}$

AVEVAMO GIÁ CALCOLATO $\bar{x} = 4.4$, $\bar{y} = 7.8$

$\sigma_x^2 = 9,84$, $\sigma_{xy} = \frac{98.4}{5} = 19.68$

$$\Rightarrow a = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{19.68}{9.84} = 2$$

$$b = \bar{y} - a\bar{x} = 7.8 - 2 \cdot 4.4 = 7.8 - 8.8 = -1$$

OSSIA LA RETTA $Y = 2X - 1$.

2)

ETÀ (anni)	25	30	42	55	55	63	70
PRESSIONE (mmHg)	120	125	135	140	145	140	160

ESERCIZIO: CALCOLARE LA RETTA DI REGRESSIONE DELLA PRESSIONE (Y) RISP. ALL'ETÀ (X) DOPO AVER VERIFICATO CHE $\rho_{xy} = 0.94$

$$Y = 0.73\,X + 102.20$$

3) IN UN GRUPPO DI 5 ADULTI, LA SOMMINISTRAZIONE DI DOSI DIVERSE DI UN FARMACO HA COMPORTATO LE SEGUENTI DIMINUZIONI DELLA PRESSIONE ARTERIOSA:

DOSE (mg)	DIMINUZ. PRESS. (mmHg)
7	10
12	18
15	20
20	25
22	25

ESERCIZI: a) SCRIVERE L'EQ. DELLA RETTA DI

ESERCIZI: a) SCRIVERE L'EQ. DELLA RETTA DI REGRESSIONE (X = DOSE, Y = PRESSIONE)

b) CALCOLARE LA DOSE OTTIMALE PER OTTENERE UNA DIMINUZIONE DELLA PRESSIONE PARI A 15 mmHg ($\Rightarrow$ SOSTITUIRE $Y = 15$ NELLA EQ. DELLA RETTA E CALCOLARE X)

NOTA: ABBIAMO VISTO CHE X, Y SONO TANTO PIÙ ALLINEATE RISPETTO ALLA RETTA DI REGRESSIONE QUANTO PIÙ $|\rho_{xy}| \longrightarrow 1$.

NELLA PRATICA, SI DISTINGUONO:

- $0 < |\rho_{xy}| \leq 0.3$ SI PARLA DI <u>CORRELAZIONE DEBOLE</u>
- $0.3 < |\rho_{xy}| \leq 0.7$ SI PARLA DI <u>CORRELAZIONE MODERATA</u>
- $0.7 < |\rho_{xy}| \leq 1$ SI PARLA DI <u>CORRELAZIONE FORTE</u>

<u>PROBABILITÀ</u>. È LA MISURA DI QUANTO LO ACCADIMENTO DI UN EVENTO SIA PLAUSIBILE TRA TUTTI QUELLI POSSIBILI.

AD ESEMPIO NEL LANCIO DI UNA MONETA

VI SONO 2 EVENTI POSSIBILI: "ESCE TESTA"

AD ESEMPIO NEL LANCIO DI UNA MONETA VI SONO 2 EVENTI POSSIBILI: "ESCE TESTA" O "ESCE CROCE". CIASCUNO DI QUESTI 2 ESITI VIENE CHIAMATO <u>EVENTO ALEATORIO</u> (OSSIA "CASUALE", DAL LATINO ALEA - DADO DA GIOCO).

NELLA PRATICA, LA PROBABILITÁ CI SERVE PER INTRODURRE LA NOZIONE DI <u>INDIPENDENZA</u> TRA EVENTI.

DENOTEREMO CON Ω L'INSIEME DI TUTTI GLI EVENTI POSSIBILI DI UN FENOMENO.

ESEMPI:

- LANCIO DI UNA MONETA $\Omega = \{T, C\}$
- LANCIO DI UN DADO $\Omega = \{1, 6, 2, 3, 4, 5\}$
- LANCIO DI 2 DADI:

$$\Omega = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), \dots\}$$

$$(6,1), \ldots\ldots (6,6)\}$$

DOVE OGNI COPPIA DENOTA (1° DADO, 2° DADO).

DEF: LA <u>PROBABILITÀ</u> (O <u>MISURA DI PROBABILITÀ</u>) È UNA FUNZIONE

$$P : \{\text{SOTTOINSIEMI DI } \Omega\} \longrightarrow [0,1]$$

CHE VERIFICA LE PROPRIETÀ:

1) $P(\Omega) = 1$ $(= 100\%)$

2) $P(\emptyset) = 0$ $(= 0\%)$

3) SE A, B SONO 2 SOTTOINSIEMI DI Ω DISGIUNTI, OSSIA $A \cap B = \emptyset$ $\Rightarrow$

$$P(A \cup B) = P(A) + P(B)$$

QUANDO NON DIVERSAMENTE SPECIFICATO, USEREMO LA PIÙ NATURALE MISURA DI PROBABILITÀ, DETTA <u>UNIFORME</u>: SE $\Omega = \{x_1, \ldots, x_m\}$ ALLORA DATO $A \subseteq \Omega$ SI HA

$$P(A) = \frac{\text{eventi favorevoli}}{\text{eventi totali}} = \frac{|A|}{|\Omega|} = \frac{\text{n° di elem. di } A}{\text{n° di elem. di } \Omega}$$

ESEMPI: 1) LANCIO DI UNA MONETA: $\Omega = \{T, C\}$

$$P(T) = \frac{|\{T\}|}{|\Omega|} = \frac{1}{2} = 0.5 \quad (= 50\%)$$

$$P(\emptyset) = \frac{0}{2} = 0, \qquad P(T, C) = \frac{2}{2} = 1$$

2) LANCIO DI UN DADO: $\Omega = \{1, 3, 5, 4, 6, 2\}$

$$P(2) = \frac{1}{6} = P(1) = P(3) = P(4) = P(5) = P(6)$$

$$P(1,3,4) = \frac{3}{6} = \frac{1}{2}$$

3) DADO TRUCCATO: $\Omega = \{1, 2, 3, 4, 5, 6\}$

$$P(1) = P(2) = P(3) = P(4) = 0.1 \quad (= 10\%)$$

$$P(5) = P(6) = 0.3 \quad (= 30\%)$$

QUAL È LA PROB. DI OTTENERE 2 o 3?

$$P(2,3) \overset{(3)}{=} P(2) + P(3) = 0.1 + 0.1 = 0.2 \quad (20\%)$$

QUAL È LA PROB. DI OTTENERE 2 o 6?

$$P(2,6) \overset{(3)}{=} P(2) + P(6) = 0.1 + 0.3 = 0.4 \quad (40\%)$$

$P(2,6) \overset{(3)}{=} P(2) + P(6) = 0.1 + 0.3 = 0.4$ (40%)

4) LANCIO DI 2 DADI :

$$\Omega = \{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),$$
$$(2,1),(2,2), \ldots$$
$$\vdots \qquad \ddots \qquad \vdots$$
$$(6,1), \ldots \ldots (6,6)\}$$

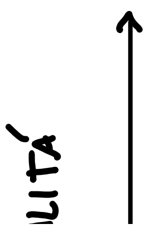
CI SONO IN TUTTO 36 COPPIE.
OGNI COPPIA HA PROBABILITÀ UNIFORME DI ACCADERE, OSSIA $\frac{1}{36}$.

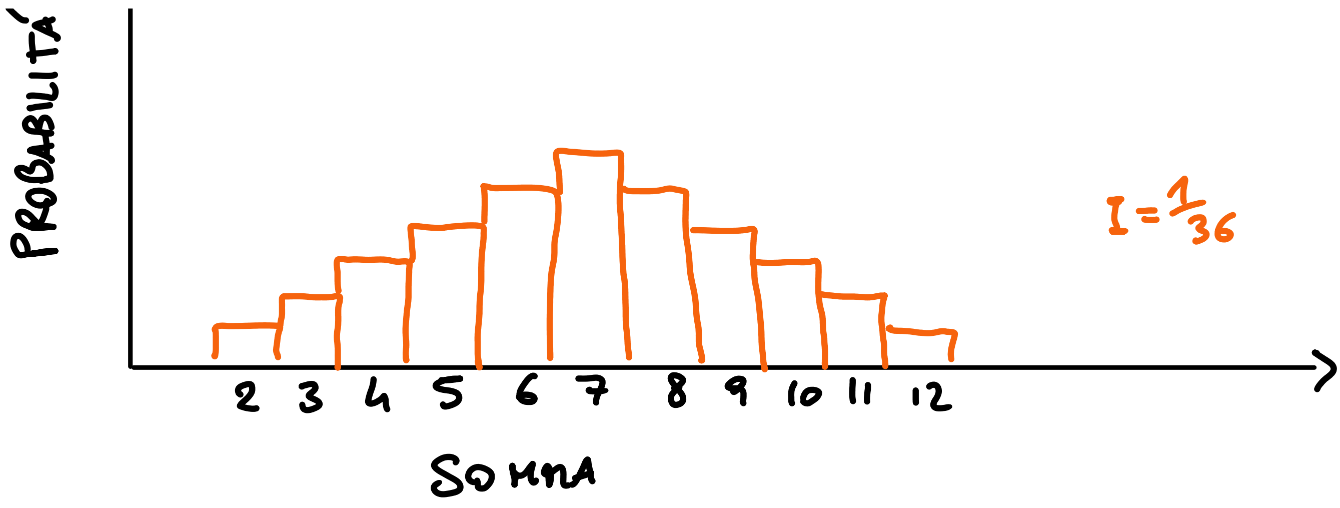
QUAL È LA PROB. DI AVERE SOMMA 10?

$A = \{\text{CASI FAVOREVOLI}\} = \{(4,6),(5,5),(6,4)\}$

$$\Rightarrow \quad P(A) = \frac{3}{36} = \frac{1}{12}$$

RIPORTIAMO IN UN ISTOGRAMMA LE PROB. DI CIASCUNA SOMMA :





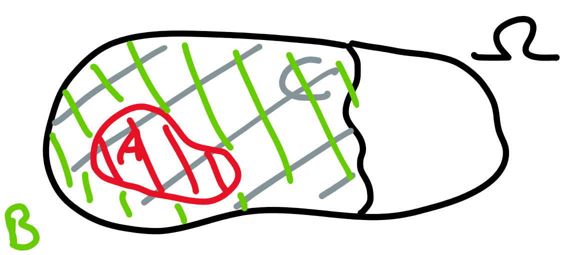
(DISTRIBUZIONE BINOMIALE)

$\boxed{\text{PROPOSIZIONE:}}$ SIA Ω SPAZIO DI PROBABILITÁ CON MISURA P. ALLORA:

a) SE A, C SONO SOTTOINSIEMI DI Ω CON $A \subseteq C \implies P(A) \leq P(C)$

b) SE A^c É IL COMPLEMENTARE DI $A \implies P(A^c) = 1 - P(A)$.

DIMOSTRAZIONE:

a) 

$(A \subseteq C)$

CHIAMIAMO $B = A^c \cap C$

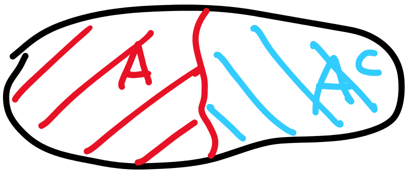
$\Rightarrow A \cap B = \emptyset$ E

$A \cup B = C$

$$\implies P(C) = P(A \cup B) \overset{(3)}{=} P(A) + P(B) \geq P(A)$$

OSSIA $P(C) \geq P(A)$

b)



Ω $\quad A \cap A^c = \phi$ E

$A \cup A^c = \Omega$

$$\Rightarrow 1 \overset{(1)}{=} P(\Omega) = P(A \cup A^c) \overset{(3)}{=} P(A) + P(A^c)$$

OSSIA $1 = P(A) + P(A^c) \Rightarrow P(A^c) = 1 - P(A).$

$\square$

ESEMPIO: LA PROB. DI OTTENERE UN NUMERO DIVERSO DA 2 CON UN DADO É

$$P(2^c) = 1 - P(2) = 1 - \frac{1}{6} = \frac{5}{6}.$$

PROBABILITÁ CONDIZIONATA E INDIPENDENZA DI EVENTI.

ASSUMIAMO CHE A, B SIANO DUE EVENTI (SOTTOINSIEMI DI UNO SPAZIO DI PROBABILITÁ Ω) E SUPPONIAMO CHE B AVVENGA. VOGLIAMO CALCOLARE LA PROBABILITÁ DI A SAPENDO B.

DEF: SI DEFINISCE PROBABILITÁ DI A CONDIZIONATO A B IL NUMERO

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \qquad (P(B) \neq 0)$$

SCRIVEREMO TALVOLTA $P(A \text{ dato } B)$.

ESEMPI: 1) QUAL È LA PROBABILITÀ CHE ESCA 4 LANCIANDO UN DADO, SAPENDO CHE È USCITO UN PARI?

$A = \{4\}$ $\quad B = \{2, 4, 6\} \longrightarrow A \cap B = \{4\}$

$$P(A \text{ dato } B) = P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{3/6} =$$

$$= \frac{1}{6} \cdot \frac{6}{3} = \frac{1}{3}$$

2) QUAL È LA PROB. CHE ESCA 4 SAPENDO CHE È USCITO UN DISPARI?

$A = \{4\}$ $\quad B = \{1, 3, 5\} \rightarrow A \cap B = \emptyset$

$$P(A \text{ dato } B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\emptyset)}{P(B)} = \frac{0}{3/6} = 0$$

3) ESERCIZIO: QUAL È LA PROBABILITÀ CHE LANCIANDO 2 DADI LA SOMMA DIA 10, SAPENDO CHE COL PRIMO DADO ESCE 4?

DEF: DUE EVENTI A, B SI DICONO <u>INDIPENDENTI</u> SE $P(A \cap B) = P(A) \cdot P(B)$

OSSERVAZIONE: SE A, B SONO INDIPENDENTI $\Rightarrow$

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot \cancel{P(B)}}{\cancel{P(B)}} = P(A)$$

OSSIA IL FATTO CHE B AVVENGA NON CAMBIA LA PROBABILITA' DI A!

TEOREMA (FORMULA DI BAYES):

$$P(A \mid B) = \frac{P(A) \cdot P(B \mid A)}{P(B)}$$

ESEMPI: 1) SIANO A, B EVENTI ALEATORI CON

$P(A) = 0.3$, $P(B) = 0.5$, $P(B \mid A) = 0.8$

CALCOLIAMO $P(A \mid B) = \dfrac{P(A) \cdot P(B \mid A)}{P(B)} =$

$$= \frac{0.3 \cdot 0.8}{0.5} = \frac{0.24}{0.5} = 0.48$$

2) DA UN RILEVAMENTO E' NOTO CHE UNA CERTA POPOLAZIONE É COMPOSTA PER IL 40% DA FUMATORI ABITUALI. É NOTO INOLTRE CHE IL 5% DEI DECESSI AVVIENE A CAUSA DI UN CERTO TIPO DI TUMORE. INFINE SI É CONSTATATO CHE TRA QUANTI SONO

DECEDUTI A CAUSA DI QUEL TUMORE, IL 60% ERANO FUMATORI ABITUALI.

CALCOLARE LA PROBABILITÁ, PER I FUMATORI ABITUALI, DI MORIRE PER QUEL TUMORE.

I DUE EVENTI SONO TRA LORO INDIPENDENTI?

$$\Omega = \{ \text{DECESSI DELLA POPOLAZIONE} \}$$

A = FUMATORE $\quad$ B = MALATO

$P(A) = 0.4$ $\quad$ $P(B) = 0.05$

$P(A \text{ dato } B) = 0.6$

$P(B \text{ dato } A) = ?$

$$P(B|A) = \frac{P(B) \cdot P(A|B)}{P(A)} = \frac{0.05 \cdot 0.6}{0.4} = 0.075$$

OSSIA IL 7.5 %.

SONO INDIPENDENTI? (A, B SONO INDIP. SE $P(A \cap B) = P(A) \cdot P(B)$)

CALCOLIAMO $P(A \cap B)$ RICORDANDO CHE

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \Rightarrow \boxed{P(A \cap B)} = P(B) \cdot P(A \mid B)$$

$$= 0.05 \cdot 0.6$$

$$= \boxed{0.03}$$

INVECE $P(A) \cdot P(B) = 0.4 \cdot 0.05 = 0.02 \neq 0.03$

OSSIA A, B SONO DIPENDENTI TRA LORO.