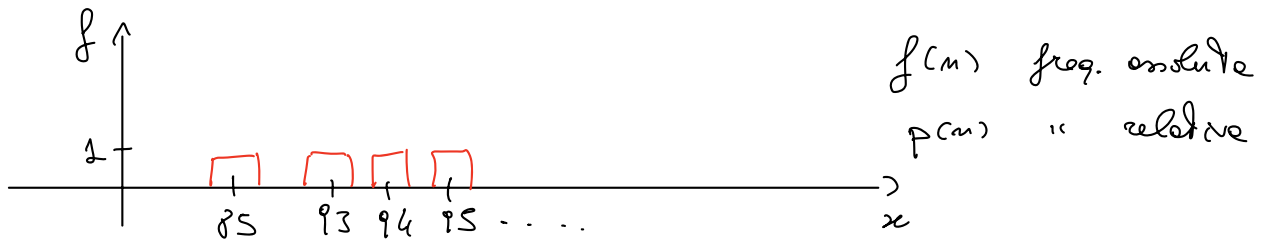


Statistiche descrittive

$$\underline{ES} \quad \underline{x} = \{ \overset{x_1}{110}, \overset{x_2}{103}, \overset{x_3}{102}, \overset{x_4}{97}, 85, 94, 95, 99, 122, 93, \overset{x_{11}}{101} \}, N=11$$



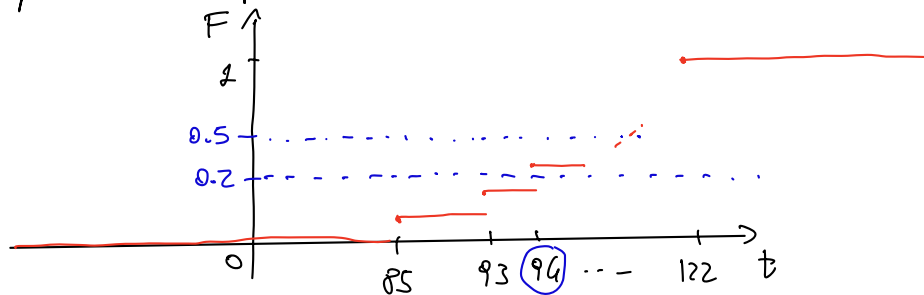
$$\text{media } \bar{x} = \frac{1}{N} \sum_{i=1}^{11} x_i = \sum_{m=0}^{+\infty} p(m) m = \frac{1}{11} (110 + 103 + 102 + \dots) = 100,09$$

$$\begin{aligned} \text{varianza } var(x) &= \frac{1}{N} \sum_{i=1}^{11} (x_i - \bar{x})^2 = \frac{1}{N} \sum_{i=1}^{11} x_i^2 - \bar{x}^2 = \sum_{m=0}^{+\infty} p(m) m^2 - \bar{x}^2 \\ &= \frac{1}{11} (110^2 + 103^2 + 102^2 + \dots) - (100,09)^2 \sim 85,93 \end{aligned}$$

scarto quadratico medio
deviazione standard

$$\sigma(x) = \sqrt{var(x)} \sim 9,27$$

funzione di ripartizione empirica $F: \mathbb{R} \rightarrow [0, 1], F(t) = \frac{\#\{x_i : x_i \leq t\}}{N}$



$\beta \in (0, 1)$ β -quantile $\leftrightarrow$ $100 \cdot \beta$ -esimo percentile ($\beta = 0.2 \leftrightarrow$ 20-esimo perc.)

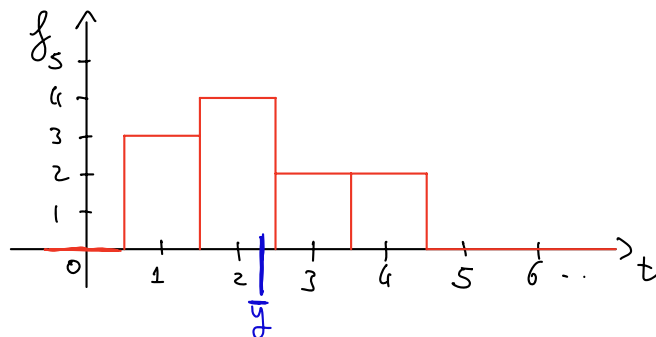
94 $\hat{=}$ il 20-esimo percentile di x

ovvero $\frac{20}{100}$ dei dati sono ≤ 94 e $\left[\sim 27\% \text{ dei dati } \leq 94 \right]$
 ovvero $\frac{80}{100}$ dei dati sono ≥ 94 $\left[\sim 82\% \text{ dei dati } \geq 94 \right]$

mediana $\leftrightarrow$ 0.5-quantile $\leftrightarrow$ 50-esimo percentile $\hat{=}$ 101

ES

$$\underline{y} = \{ 2, 2, 2, 3, 1, 1, 2, 3, 4, 1, 4 \}, \quad N=11$$



$$\bar{y} = \frac{1}{11} (2+2+2+3+1+1+2+3+4+1+4) = \frac{25}{11}$$

$$= 1 \cdot \frac{3}{11} + 2 \cdot \frac{4}{11} + 3 \cdot \frac{2}{11} + 4 \cdot \frac{2}{11} = \frac{25}{11}$$

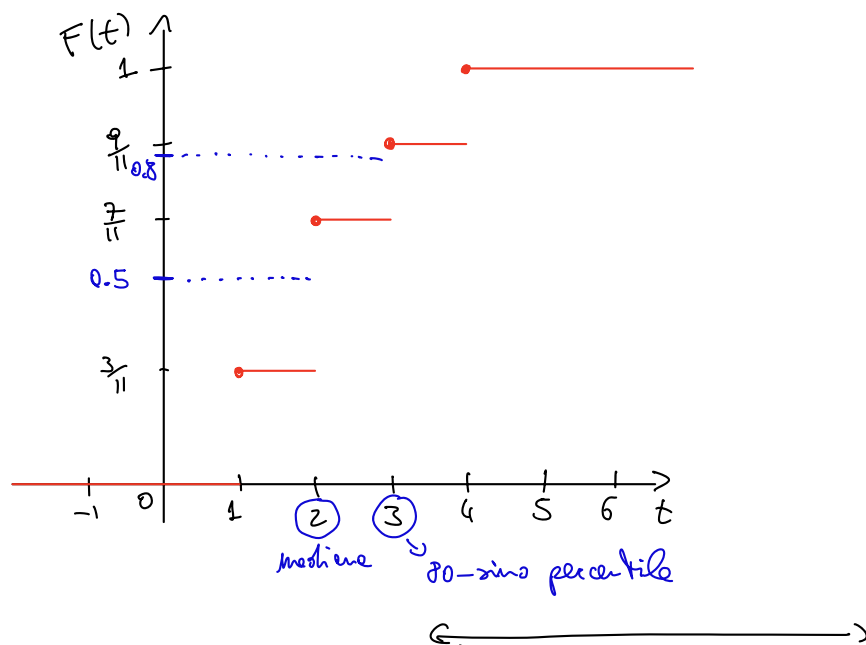
$$= 2,27 \sim 2,27$$

$$\text{var}(\underline{y}) = \frac{1}{11} (2^2+2^2+2^2+3^2+1^2+1^2+2^2+3^2+4^2+1^2+4^2) - \left(\frac{25}{11}\right)^2 =$$

$$= 1^2 \cdot \frac{3}{11} + 2^2 \cdot \frac{4}{11} + 3^2 \cdot \frac{2}{11} + 4^2 \cdot \frac{2}{11} - \left(\frac{25}{11}\right)^2 =$$

$$= \frac{69}{11} - \frac{625}{121} = \frac{759-625}{121} = \frac{134}{121} \sim 1,11$$

$$\sigma(\underline{y}) = \sqrt{\frac{134}{121}} \sim 1,05$$



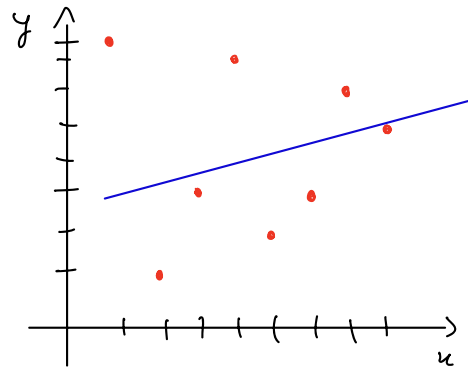
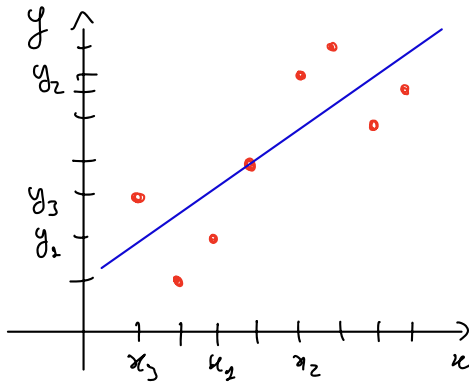
Dati due insiemi di dati: $\underline{x} = \{x_1, x_2, \dots, x_N\}$ e $\underline{y} = \{y_1, y_2, \dots, y_N\}$,
con la stessa numerosità N .

Def Si chiama covarianza di $\underline{x}$ e $\underline{y}$ il numero

$$\text{cov}(\underline{x}, \underline{y}) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$

Si chiama coefficiente di correlazione tra $\underline{x}$ e $\underline{y}$ il numero

$$\rho(\underline{x}, \underline{y}) = \frac{\text{cov}(\underline{x}, \underline{y})}{\sigma(\underline{x})\sigma(\underline{y})} \in [-1, 1].$$



Si chiama retta di regressione di $\underline{x}$ e $\underline{y}$ la retta

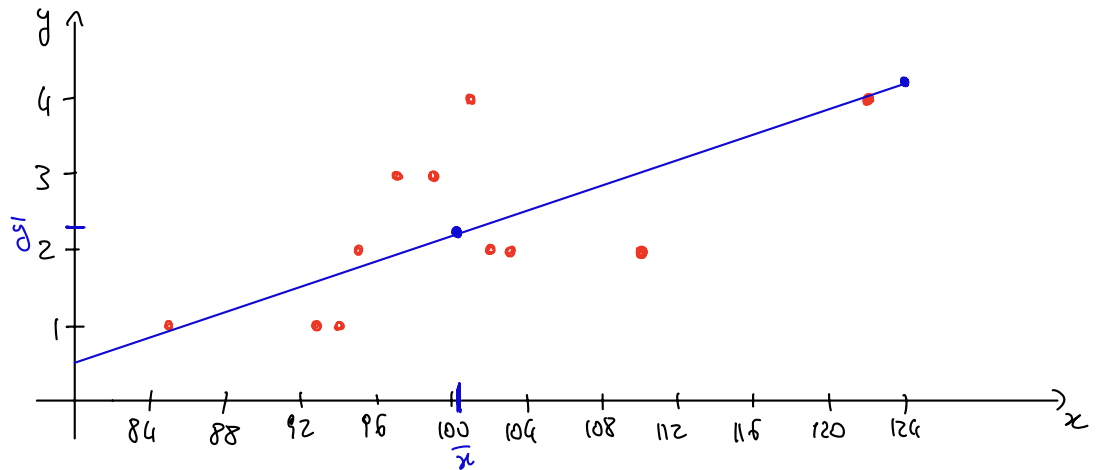
$$y = ax + b$$

dove $a = \frac{\text{cov}(\underline{x}, \underline{y})}{\text{var}(\underline{x})}$, $b = \bar{y} - a\bar{x}$

ES

$$\underline{x} = \{110, 103, 102, 97, 85, 94, 95, 99, 122, 93, 101\}$$

$$\underline{y} = \{2, 2, 2, 3, 1, 1, 2, 3, 4, 1, 4\}$$



$$\text{cov}(\underline{x}, \underline{y}) = \frac{1}{11} \left[(110 - 100,99)(2 - 2,27) + (103 - 100,99)(2 - 2,27) + \dots \right] \sim 6,62$$

$$\rho(\underline{x}, \underline{y}) = 0,67$$

$$a = \frac{\text{cov}(\underline{x}, \underline{y})}{\text{var}(\underline{x})} \sim \frac{6,62}{85,93} \sim 0,08$$

$$y = 0,08x - 5,74$$

$$b = \bar{y} - 0.08\bar{x} = 2.77 - 0.08 \cdot 100.09 \sim -5.74$$

Probabilità

Variabile aleatoria. $X: \underset{\text{universo}}{\Omega} \longrightarrow \mathbb{R}$, $\omega \in \underset{\text{eventi}}{\Omega}$
 $A \subset \Omega$

Variabile aleatoria discreta se $\text{Immagine}(X) = \{x_1, x_2, x_3, \dots\}$ finito o numerabile

Si chiama legge di X la distribuzione di probabilità dei suoi valori
 $\{p(x_i)\}$ $p(x_i) = \text{probabilità che } X \text{ valga } x_i.$
 $\sum p(x_i) = 1$

ES $\Omega = \{\text{esito di lancio di una moneta}\} = \{T, C\}$
 $X: \Omega \rightarrow \mathbb{R}, \quad X(T) = +1, \quad X(C) = -1.$

$$\text{Immag}(X) = \{-1, +1\}$$

Legge di X, $p(-1) = \frac{1}{2}, \quad p(+1) = \frac{1}{2}.$

ES $\Omega = \{\text{esito di tre lanci di una moneta}\} =$
 $= \{TTT, TTC, TCT, TCC, CTT, CTC, CCT, CCC\}$

$X: \Omega \rightarrow \mathbb{R}, \quad X = \text{somma di } +1 \text{ per ogni } T \text{ e di } -1 \text{ per ogni } C.$

$$\text{Immag}(X) = \{-3, -1, +1, +3\}$$

$$p(-3) = \frac{1}{8}, \quad p(-1) = \frac{3}{8}, \quad p(+1) = \frac{3}{8}, \quad p(+3) = \frac{1}{8}$$

ES $\Omega = \{\text{esito di } n \text{ lanci di una moneta}\}$

$X: \Omega \rightarrow \mathbb{R}, \quad X = \text{quante volte esce } T \text{ in } n \text{ lanci}$

$$\text{Immag } X = \{0, 1, 2, \dots, n\}$$

$k=0, \dots, n$

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

legge binomiale

se p la prob. del successo

$$\binom{n}{k} = \frac{n!}{k! (n-k)!}, \quad h! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot h$$

