LEGENDA

1. STATISTICA DESCRITTIVA

- Popolazione
- Campione
- Bias
- Variabile
- Frequenza
- Diagramma a barre
- Diagramma a torta
- Dati raggruppati in classi
- Valori di sintesi
    - Moda
    - Media
    - Varianza
    - Quartili
    - Percentili
    - Quantili
    - Retta di regressione lineare
    - Covarianza
    - Coeff. di correlazione lin.
    - Coeff. di determinazione

2. PROBABILITÀ

- Spazio degli eventi elementari
- Evento
- Collezione di tutti gli eventi
- Proprietà delle collezioni
- Misura di probabilità
- Principio di inclusione/esclusione
- Calcolo combinatorio
    - Disposizione, ripetizioni
    - Disposizione, no ripetizioni
    - Permutazione
    - Combinazione semplice
- Paradosso del compleanno
- Probabilità condizionata
- Th. della probabilità totale
- Th. di Bayes
- Problema di Monty Hall
- Eventi indipendenti
- Variabili aleatorie
- Legge di una var. aleatoria
- Valore atteso
- LOTUS
- Varianza
- Covarianza
- Vettori aleatori
- Funzioni di vettori aleatori
- Leggi note
- Leggi discrete
    - Bernoulli $Be(p)$
    - Binomiale $B(n,p)$
    - Poisson $\mathcal{P}(\lambda)$
    - Geometrica $Geo(p)$
- Leggi continue
    - Uniforme $U(a,b)$
    - Esponenziale $Exp(\lambda)$
    - Gamma $Gamma(\alpha,\lambda)$
    - Normale $\mathcal{N}(\mu,\sigma^2)$
- Vettori aleatori continui
- Variabili aleatorie indipendenti

3. STATISTICA INFERENZIALE

- Definizioni
- Obiettivi
- Legge dei Grandi Numeri
- Th. del Limite Centrale
- Intervalli di Confidenza I.C.
- Leggi continue per analisi inf.
    - Chi-Quadro $\chi^2(n)$
    - tStudent $t(n)$
- I.C. unilaterali
- Test di ipotesi
- p-value

STATISTICA DESCRITTIVA

- POPOLAZIONE: l'insieme contenente tutte le unità statistiche che saranno esaminate (in parte) per l'analisi statistica.
- CAMPIONE: sottoinsieme della popolazione contenente tutte le unità statistiche esaminate.

- BIAS: è un errore che si commette effettuando un campionamento non rappresentativo della popolazione.

<u>Es</u> Survivorship bias

- VARIABILE: l'oggetto di studio dell'analisi statistica.

Può essere:

- qualitativa < DESCRITTIVA / APPARTENENZA AD UNA CLASSE
- quantitativa < DISCRETA / CONTINUA

- FREQUENZA

Siano $x_1, x_2 \dots x_n$ dati del campione

con $v_1, v_2, \dots v_k$ valori assunti dalla var. (range)

f_j = num. assunto da x_i quando $x_i = v_j$

$f_j = \#\{i = 1..n \; : \; x_i = v_j\}$ <u>Oss</u> $\sum_{j=1}^{k} f_j = n$

$p_j = \frac{f_j}{n}$ ← FREQ. RELATIVA   <u>Oss</u> $\sum_{j=1}^{k} p_j = 1$

FREQUENZA ASSOLUTA

Avendo le freq. di variabili qualitative o quantitative discrete possono prodursi dei grafici a barre o a torta.

Diagramma a barre delle frequenze assolute / relative

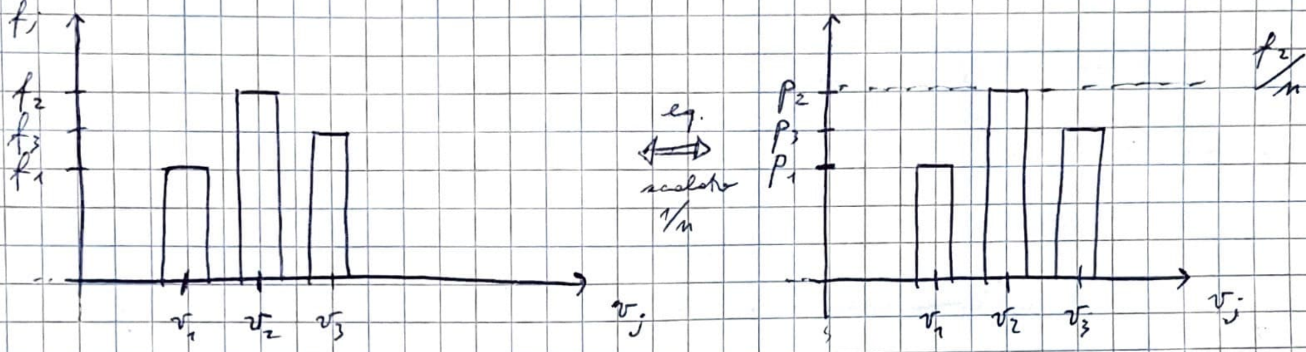
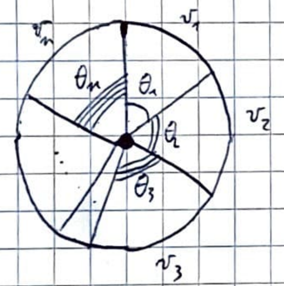


Diagramma a torta

NON CONTA L'ORDINE



$$\theta_i = p_i \cdot 2\pi$$

$$\sum_{j=1}^{n} \theta_j = 2\pi$$

Dati raggruppati in classi

Un esempio di classe di dati è un intervallo.

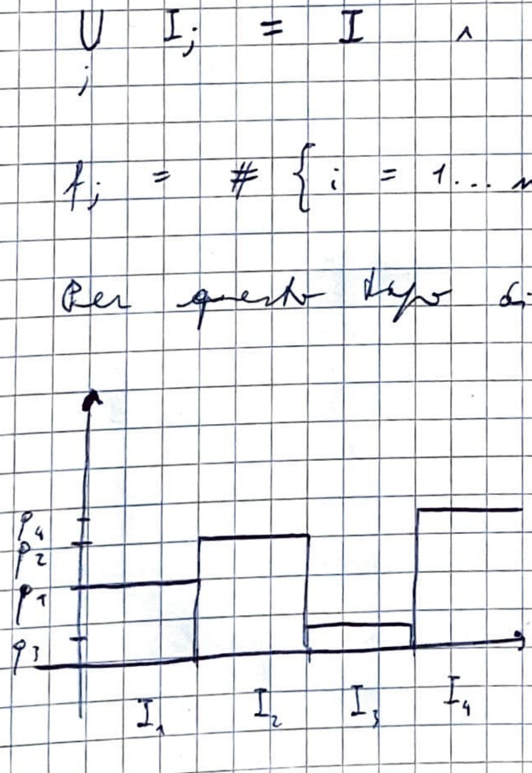
$$I_j = [a_j, b_j) \qquad \text{con} \quad b_j = a_{j+1}$$
$$I_{j+1} = [a_{j+1}, b_{j+1})$$

$$\bigcup_j I_j = I \quad \wedge \quad \bigcap_j I_j = \emptyset$$

$$f_j = \# \{ i = 1 \dots n \ \ni' \ x_i \in I_j \}$$

Per questo tipo di variabile uso l'istogramma.

istogramma delle freq. rel.



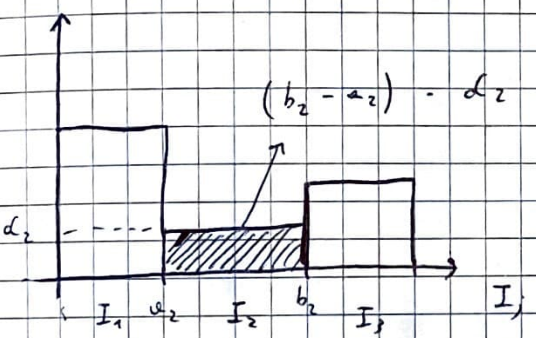
Non è regolamentata l'ampiezza dell'intervallo che può essere scelta a piacere.

Per valutare il miglior intervallo di diverse ampiezze usiamo la densità di frequenza ass. / rel.

$$\frac{f_j}{b_j - a_j} = \frac{f_j}{|I_j|}$$ densità di freq. assoluta

$$\frac{p_j}{b_j - a_j} = \frac{p_j}{|I_j|}$$ densità di freq. relativa

È possibile rappresentare anche l'istogramma delle ~~[illegible]~~ densità di frequenze relative / assolute.



$$(b_2 - a_2) \cdot d_2 = \frac{p_2}{b_2 - a_2} \cdot (b_2 - a_2) = p_2$$

L'area del rettangolo j rappresenta p_j

VALORI DI SINTESI

Sono valori che caratterizzano i dati

- MODA: il valore più frequente nel dataset.

v_j è modale se $f_j \geq f_i \quad \forall\, i = 1..k$

$p_j \geq p_i \quad \forall\, i = 1..k$

Se $f_j > f_i \quad \forall\, i = 1..k$ allora v_j è l'unico
$(p_j) \quad (p_i)$
valore modale $\Rightarrow$ v_j è la moda.

È utilizzata per var. qualitative / quant. discrete.

Nei dati raggrupp. contenuti al posto ~~delle~~ dei valori modali c'è / ci sono la classe MODA o le classi modali che sono quelle con densità di freq. maggiore.

I_j è modale se $d_{j_A} \geq d_{i_A}$ $\forall\, i = 1..k$
(d_{j_R}) (d_{i_R})

con $d_{j_A} = \dfrac{f_j}{b_j - a_j}$ $\wedge$ $d_{j_R} = \dfrac{p_j}{b_j - a_j}$ $\wedge$ $I_j = [a_j, b_j)$

- MEDIA : $\boxed{\bar{x}_n = \dfrac{1}{n} \sum_{i=1}^{n} x_i}$ $\Longrightarrow \begin{cases} \bar{x}_n = \dfrac{1}{n} \sum_{i=1}^{k} v_i f_i & \text{FREQ. ASS.} \\ \bar{x}_n = \sum_{i=1}^{k} v_i p_i & \text{FREQ. REL.} \end{cases}$

Equivalente a trovare il baricentro di un'asta con $m_i = \frac{1}{n}$ $\wedge$ $x_i = x_i$

Nella media tutti i valori hanno un peso.
La media è sensibile a dati anomali.

Media pesata (ponderata)

$$\bar{x} = \sum_{i=1}^{n} \lambda_i x_i \quad \text{con } \lambda_i \text{ peso } \in [0, 1] \wedge \sum_{i=1} \lambda_i = 1$$

Per i dati raggruppati in classe la ~~media~~ media è approssimabile come

$$\bar{x}_n \approx \sum_{i=1}^{k} p_i \underbrace{\frac{a_i + b_i}{2}}_{\text{VALOR MEDIO DI } I_i}$$

Th La media segue le trasformazioni affini

Sia $\bar{x}$ media di un campione di dati x_i

$\Rightarrow$ $\bar{y} = a\bar{x} + b$ sarà la media di un campione di dati x_i soggetti ad una trasformazione affine $y_i = a x_i + b$

Dim $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \Longrightarrow \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$

$$\Longrightarrow \bar{y} = \frac{1}{n}\sum_{i=1}^{n}(a x_i + b) = \frac{1}{n}\left[\sum_{i=1}^{n} a x_i + n b\right]$$

$$\Longrightarrow \bar{y} = a\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) + b = a\bar{x} + b$$

Es $\bar{T}_{°C} = \frac{5}{9}(\bar{T}_{°F} - 32)$

- **VARIANZA**: Indica la dispersione dei dati attorno alla media

$$\boxed{\begin{aligned} s_n^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \\ \sigma^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \end{aligned}}$$

s_n^2 ← conoscendo un campione

σ^2 ← conoscendo tutta la popolazione

$$\Longrightarrow s_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad \leftarrow \text{deviazione standard}$$

$$\left(\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)$$

Th La ~~varianza~~ dev. std. regge le trasformazioni lineari (1) e non è sensibile alle traslazioni. (2)

Dim (1) $y'_i = a x_i \Longrightarrow \bar{y}' = a\bar{x}$

$$s_{y'}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y'_i - \bar{y}')^2 = \frac{1}{n-1}\sum_{i=1}^{n}(a x_i - a\bar{x})^2$$

$$\Longrightarrow s_{y'}^2 = a^2\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \Longrightarrow s_{y'}^2 = a^2 s_x^2$$

$$\Longrightarrow s_{y'} = |a|\, s_x \Longrightarrow s_{y'} = a\, s_x, \quad a > 0$$

(2) $y''_i = x_i + b \Longrightarrow \bar{y}'' = \bar{x} + b$

$$s_{y''}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y''_i - \bar{y}'')^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i + \cancel{b} - \bar{x} - \cancel{b})^2$$

$$\Longrightarrow s_{y''}^2 = s_x^2 \Longrightarrow s_{y''} = s_x$$

Oss $s_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i^2 - 2x_i\bar{x} + \bar{x}^2) =$

$= \frac{1}{n-1} \left(\sum_{i=1}^{n} x_i^2 - 2\bar{x} \sum_{i=1}^{n} x_i + n\bar{x}^2 \right) =$

$= \frac{1}{n-1} \left(\sum_{i=1}^{n} x_i^2 - 2\bar{x}\, n\bar{x} + n\bar{x}^2 \right) =$

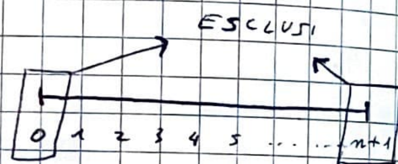
$= \boxed{\frac{1}{n-1} \left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2 \right)}$

- **QUARTILI**: Si basano sulla posizione dei dati in ordine crescente.

Siano $x_1, x_2 \dots x_n$ dati non ordinati

Rinominiamo i dati $x_1 \le x_2 \le \dots \le x_n$

I quartili possono essere < ESCLUSIVI / INCLUSIVI



$\boxed{Q_i = x_{k_i}(1 - r_i) + x_{(k_i)+1}\, r_i}$ con $i\,\frac{n+1}{4} = \underset{\in \mathbb{N}}{k_i} + \underset{\in [0,1)}{r_i}$

ESCLUSIVI

$r_i \in \{0, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}\}$

Q_2 è la <u>mediana</u> $M = x_{k_2}(1 - r_2) + x_{k_2+1}\, r_2$

La mediana non è sensibile a dati anomali.

$\boxed{Q_i = x_{k_i+1}(1 - r_i) + x_{k_i+2}\, r_i}$ con $i\,\frac{n-1}{4} = \underset{\in \mathbb{N}}{k_i} + \underset{\in [0,1)}{r_i}$

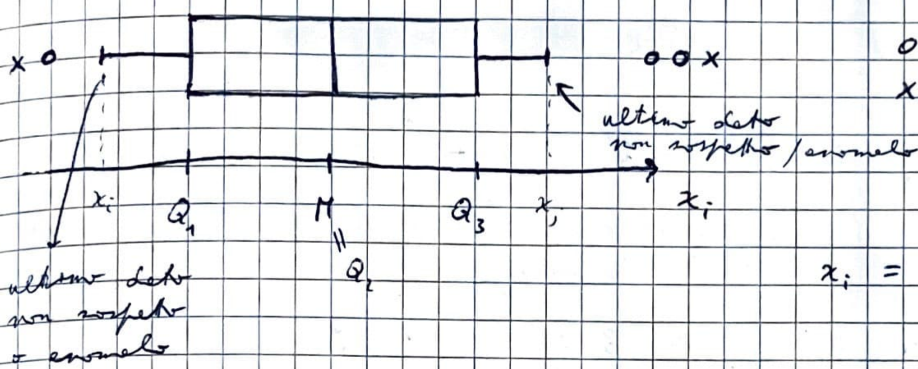
I quartili sono usati per identificare dati anomali o sospetti.

$\boxed{IQR = Q_3 - Q_1} \ge 0$ inter-quartile range

Dato <u>ANOMALO</u>: $x_i \le Q_1 - 3\,IQR \lor x_i \ge Q_3 + 3\,IQR$

Dato <u>SOSPETTO</u>: $x_i \le Q_1 - \frac{3}{2}\,IQR \lor x_i \ge Q_3 + \frac{3}{2}\,IQR$

Box-plot (diagramma scatola / e baffi)



o SOSPETTO

x ANOMALO

$$x_i = \min_{x \in X} x \quad \text{non anomalo, non sospetto}$$

$$x_j = \max_{x \in X} x \quad \text{non anomalo, non sospetto}$$

- **PERCENTILI**: stesso concetto dei quartili dove l'intervallo è diviso in 100 sottointervalli e non in 4.

$$\boxed{P_i = x_{k_i}(1 - r_i) + x_{k_i+1}\, r_i}, \quad k_i + r_i = i\,\frac{n+1}{100}$$

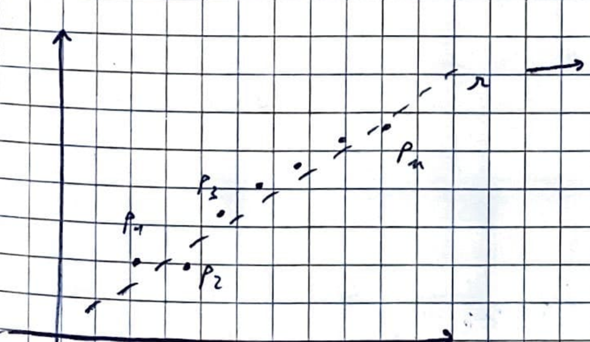
$$P_{25} = Q_1, \quad P_{50} = Q_2 = M, \quad P_{75} = Q_3$$

$$IQR = P_{75} - P_{25}$$

- **QUANTILI** $\quad k + r = (n+1)\,q, \quad q_q = x_k(1-r) + x_{k+1}\, r$

Scatter-plot (diagramma dispersione)

Si usa su campioni con dati bivariati



$\longrightarrow$ retta di regressione lineare

$$P_i = (x_i, y_i)$$

Retta di regressione lineare: È la retta che minimizza l'errore $|E|^2$.

$$e(\alpha,\beta) = \sum_{i=1}^{n} [y_i - (\alpha x_i + \beta)]^2 = \sum_{i=1}^{n} [y_i^2 - 2y_i(\alpha x_i + \beta) + (\alpha x_i + \beta)^2]$$

Minimizziamo α e β $\Rightarrow$ $\nabla e = \vec{0}$

$$\Longrightarrow \begin{cases} \dfrac{\partial e}{\partial \alpha} = 0 \\ \dfrac{\partial e}{\partial \beta} = 0 \end{cases} \Longrightarrow \begin{cases} \sum_{i=1}^{n} (-\cancel{2} x_i y_i + \cancel{2} x_i(\alpha x_i + \beta)) = 0 \\ \sum_{i=1}^{n} (-\cancel{2} y_i + \cancel{2}(\alpha x_i + \beta)) = 0 \end{cases}$$

$$\Longrightarrow \begin{cases} -\sum_{i=1}^{n} x_i y_i + \alpha \sum_{i=1}^{n} x_i^2 + \beta \sum_{i=1}^{n} x_i = 0 \\ -\sum_{i=1}^{n} y_i + \alpha \sum_{i=1}^{n} x_i + \beta n = 0 \end{cases} \Longrightarrow$$

$$\Longrightarrow \begin{cases} \cancel{-Q_{xy}} + \alpha Q_{xx} + \beta n \bar{x} = \cancel{0}\, Q_{xy} \\ \cancel{-\bar{y}\, n} + \alpha \bar{x} \cancel{n} + \beta \cancel{n} = \cancel{0}\, \bar{y} \end{cases}$$

$$\Longrightarrow \begin{cases} Q_{xx}\alpha = \dfrac{Q_{xy} - \beta n \bar{x}}{\cancel{Q_{xx}}} \\ \bar{y} = \alpha \bar{x} + \beta \end{cases} \Longrightarrow \boxed{\begin{cases} \alpha = \dfrac{Q_{xy} - n\bar{x}\bar{y}}{Q_{xx} - n\bar{x}^2} \\ \beta = \bar{y} - \alpha \bar{x} \end{cases}}$$

$$\Longrightarrow \beta = \bar{y} - \frac{Q_{xy} - n\bar{x}\bar{y}}{Q_{xx} - n\bar{x}^2}\,\bar{x} = \frac{Q_{xx}\bar{y} - \cancel{n\bar{x}^2\bar{y}} - Q_{xy}\bar{x} \cancel{+ n\bar{x}^2\bar{y}}}{Q_{xx} - n\bar{x}^2} =$$

$$= \frac{Q_{xx}\bar{y} - Q_{xy}\bar{x}}{Q_{xx} - n\bar{x}^2}$$

$$\boxed{\alpha = \frac{\operatorname{cov}(x,y)}{\operatorname{var}(x)}} = \frac{\operatorname{cov}_{x,y}}{s_x^2} = \frac{\sum_i x_i y_i - \bar{x}\bar{y}n}{\sum_i x_i^2 - \bar{x}^2 n}$$

- COVARIANZA Esprime la relazione lineare tra due variabili statistiche.

La varianza è la covarianza tra le stesse var.

$$\operatorname{cov}_{x,x} = s_x^2, \quad \operatorname{cov}_{y,y} = s_y^2$$

$$\boxed{\operatorname{cov}_{x,y} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}$$

- COEFF. DI CORRELAZIONE LINEARE

$$\boxed{\rho = \frac{cov_{x,y}}{s_x \, s_y}} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sqrt{(\sum x_i^2 - n\bar{x}^2)(\sum y_i^2 - n\bar{y}^2)}}$$

Descrive quanto variano i dati x_i, y_i in coppia rispetto a quanto variano singolarmente x_i e y_i.

$$\alpha = \frac{cov_{x,y}}{s_x^2} = \frac{cov_{x,y}}{s_x^2} \cdot \frac{s_x}{s_y} \cdot \frac{s_y}{s_x} = \boxed{\rho \cdot \frac{s_y}{s_x}}$$

Oss $\rho \in [-1, 1]$

$|\rho|$ indica la correlazione $|\rho| = 0$ NESSUNA CORR.
$|\rho| = 1$ CORR. PERFETTA

$sign(\rho)$ indica se la correlazione è crescente $(\rho > 0)$ o decrescente $(\rho < 0)$.

$$\bar{w} = (x_1 - \bar{x} \quad x_2 - \bar{x} \quad \dots \quad x_n - \bar{x})^T$$
$$\bar{v} = (y_1 - \bar{y} \quad y_2 - \bar{y} \quad \dots \quad y_n - \bar{y})$$

$$\bar{w} \cdot \bar{v} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = cov_{x,y}(n-1)$$

$$\bar{w} \cdot \bar{v} = |\bar{w}| \cdot |\bar{v}| \cdot \cos\theta = \sqrt{n-1}\, s_x \sqrt{n-1}\, s_y \cos\theta \Rightarrow$$

$$\Rightarrow (n-1) s_x s_y \cos\theta = cov_{x,y}(n-1) \Rightarrow$$
$$\Rightarrow \cos\theta = \rho_{x,y} \in [-1, 1]$$

Se $\rho = 0 \Rightarrow \cos\theta = 0 \Rightarrow \theta = \pi/2 + k\pi \Rightarrow$

$\bar{w} \perp \bar{v}$

Se $\rho = 1 \Rightarrow \cos\theta = 1 \Rightarrow \theta = 0$

$\Rightarrow \bar{w} = \lambda \bar{v}$ con $\lambda > 0$

$$\Rightarrow y_i - \bar{y} = \lambda (x_i - \bar{x}) \Rightarrow y_i = \underbrace{\lambda}_{\alpha} x_i + \underbrace{(\bar{y} - \lambda\bar{x})}_{\beta}$$

Se $\rho = -1 \Rightarrow y_i = \overset{\alpha}{\lambda} x_i + \underbrace{(\bar{y} - \lambda \bar{x})}_{\beta}$, $\lambda < 0$

$\text{sign}(\rho) = \text{sign}(\alpha)$

$\Rightarrow$ sign (ρ) indica la pendenza della retta

($\rho > 0$ crescente , $\rho < 0$ decrescente)

- COEFF. DI DETERMINAZIONE

$$R^2 = 1 - \frac{\sum_{i=1}^{n} [y_i - (\alpha x_i + \beta)]^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} = \boxed{1 - \frac{e(\alpha, \beta)}{s_y^2}}$$

$$\boxed{R^2 = \rho^2}$$

Se $R^2 = 0 \Rightarrow \frac{\sum [y_i - (\alpha x_i + \beta)]^2}{s_y^2} = 1 \Rightarrow e(\alpha, \beta) = s_y^2$

Se $\rho = 0 \Rightarrow \alpha = 0 \wedge \beta = \bar{y}$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} [y_i - \beta]^2}{s_y^2} = 1 - \frac{\overbrace{\sum (y_i - \bar{y})^2}^{s_y^2}}{s_y^2} = 0$$

Se $\rho = 1 \Rightarrow e = 0 \Rightarrow R^2 = 1 - 0 = 1$

PROBABILITÀ

- Spazio degli eventi elementari (o campionario)

$\Omega \neq \emptyset$, $\omega \in \Omega$ è un evento elementare e rappresenta una possibile osservazione dell'evento.

Es. $\Omega = \{T, C\}$ lancio della moneta

$\Omega = \{1, 2, 3, 4, 5, 6\}$ lancio di un dado

- Evento è un sottoinsieme $A \subseteq \Omega$
- Collezione di tutti gli eventi $\mathcal{F}$

$\{T\} \cup \{C\}$

$\uparrow$

Ω

$\mathcal{F} = \{A \text{ s.t. } A \subseteq \Omega\}$

Es $\Omega = \{T, C\}$, $\mathcal{F}_\Omega = \{\emptyset, \{T\}, \{C\}, \cancel{\quad}\}$

$\emptyset = \{T\} \cap \{C\}$ (esce sia T che C nello stesso lancio)

Proprietà delle collezioni

1) $\Omega \in \mathcal{F}$
2) Se $A \in \mathcal{F} \Longrightarrow \Omega \setminus A \in \mathcal{F}$
3) $A_1, A_2 \dots A_n$ successione di elementi di $\mathcal{F}$

$\Longrightarrow \bigcup_{i \in \mathbb{N}} A_i \in \mathcal{F}$

$\mathcal{F}$ ha una struttura algebrica σ. (σ-algebra)

Oss $\emptyset = \Omega \setminus \Omega \Longrightarrow \emptyset \in \mathcal{F}$ (4)

$\Omega \setminus \bigcap_{i \in \mathbb{N}} A_i = \bigcup_{i \in \mathbb{N}} (\Omega \setminus A_i) \in \mathcal{F}$

$\Longrightarrow \bigcap_{i \in \mathbb{N}} A_i \in \mathcal{F}$ (5)

- Misura di probabilità $\mathbb{P}$

$\mathbb{P} : \mathcal{F} \to [0, 1]$

1) $\mathbb{P}(\Omega) = 1$
2) $\mathbb{P}(A_i \cup A_j) = \mathbb{P}(A_i) + \mathbb{P}(A_j)$, $A_i \cap A_j = \emptyset$
3) $\mathbb{P}(A_i \cup A_j) = \mathbb{P}(A_i) + \mathbb{P}(A_j) - \mathbb{P}(A_i \cap A_j)$

$\uparrow$

principio di inclusione / esclusione

Oss (2) e (3) viene fuori $\mathbb{P}(\emptyset) = 0$

Prop $\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(B)$, $B \subset A$

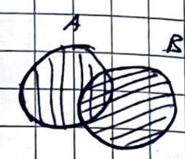
Dim $A = (A \setminus B) \cup B$

$(A \setminus B) \cap B = \emptyset$

$\mathbb{P}(A \setminus B \cup B) = \mathbb{P}(A \setminus B) + \mathbb{P}(B) = \mathbb{P}(A)$

$\Longrightarrow \mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(B)$

Prop $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$



$A \cup B = A \cup (B / A)$

$A \cap (B \setminus A) = \emptyset$

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \overbrace{\mathbb{P}(B \setminus A)}^{\mathbb{P}(B \setminus (A \cap B))} = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

caso del principio di inclusione / esclusione.

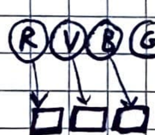
CALCOLO COMBINATORIO

- disposizioni con ripetizioni

Es PIN di 4 cifre $10^4 = \#A = \boxed{n^k}$

- disposizioni senza ripetizioni

Es



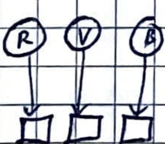
Metto $\overset{n}{4}$ palline in $\overset{k}{3}$ scatole

$$\#A = \frac{4!}{1!} = \boxed{\frac{n!}{(n-k)!}}$$

- permutazioni

$(n = k)$

Es

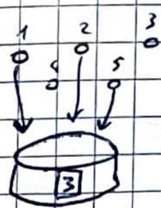


Metto $\overset{n}{3}$ palline in $\overset{k}{3}$ scatole

$\#A = 3! = \boxed{n!}$

• combinazione semplice

Es:



Metto $\overset{k}{3}$ palline su $\overset{n}{5}$ in una scatola (l'ordine non importa)

$$\#A = \boxed{\binom{n}{k}} = \frac{n!}{k!\,(n-k)!} = \frac{5!}{3!\;2!} = 10$$

Da dove viene questa formula?
Prendo ogni disposizione senza ripetizioni e mi accorgo che ci sono casi uguali (nell'es: stesse palline ma in ordine diverso) per il nostro calcolo.
Per ogni caso ho $k!$ permutazioni.
Divido $\#A_{\text{CON ORDINE}}$ per $k!$ $\Longrightarrow$

$$\#A_{\text{SENZA ORDINE}} = \#A_{\text{CON ORDINE}} \,/\, k! = \frac{n!}{(n-k)!} \cdot \frac{1}{k!} = \binom{n}{k}$$

Paradosso del compleanno

A = " almeno 2 persone condividono il compleanno "
$\Omega \setminus A$ = " tutti hanno un compleanno diverso "

$$\mathbb{P}(A) = 1 - \mathbb{P}(\Omega \setminus A) = 1 - \frac{\#(\Omega \setminus A)}{\#\Omega} = 1 - \frac{\frac{365!}{(365-k)!}}{365^k}$$

$\mathbb{P}(A) \approx 0.5037 > 50\%$ per $k = 23$.

P.S. $\mathbb{P}(X) = \frac{\#X}{\#\Omega} \iff \mathbb{P}(x_1) = \mathbb{P}(x_2) = \dots = \mathbb{P}(x_n)$

La famosa formula "casi favorevoli su casi totali"

Proprietà coeff. bin.

TRIANGOLO DI TARTAGLIA (PASCAL)

1) $\binom{n}{k} = \binom{n}{n-k}$ 2) $\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}$

ESPANSIONE POLINOMIALE DI NEWTON

3) $\sum_{i=0}^{n} \binom{n}{i} = 2^n$ 4) $(a+b)^n = \sum_{i=0}^{n} a^i\, b^{n-i} \binom{n}{i}$

PROBABILITÀ CONDIZIONATA

$$\boxed{\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}} \quad , \quad A, B \in \mathcal{F} \quad , \quad \mathbb{P}(B) > 0$$

$$\Longrightarrow \quad \mathbb{P}(A \cap B) = \mathbb{P}(A \mid B)\, \mathbb{P}(B)$$

Se $A \subseteq B \implies \mathbb{P}(A \mid B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = 1$

$\mathbb{P}(\cdot \mid B) : \mathcal{F} \to [0,1]$ è una misura di probabilità.

1) $\mathbb{P}(\Omega) = 1 \quad \wedge \quad \mathbb{P}(\emptyset) = 0$

2) $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cup A_2)$

Th delle probabilità totale

$$\boxed{\mathbb{P}(A) = \sum_{i \in N} \mathbb{P}(A \mid B_i)\, \mathbb{P}(B_i)} \quad \text{con } A \in \mathcal{F} \ \wedge$$

$(B_i)_{i \in N}$ una partizione di Ω $\left(\bigcup_i B_i = \Omega, \ \bigcap_{i \in N} B_i = \emptyset \ , \ N \subset \mathbb{N} \ , \ \#N \geq 2 \right)$

Dim $\quad A = A \cap \Omega = A \cap \bigcup_{i \in N} B_i = \bigcup_{i \in N} (A \cap B_i) \implies$

$$\Longrightarrow \mathbb{P}(A) = \mathbb{P}\left(\bigcup_{i \in N} (A \cap B_i) \right) = \sum_{i \in N} \mathbb{P}(A \cap B_i) =$$

$$= \sum_{i \in N} \mathbb{P}(A \mid B_i)\, \mathbb{P}(B_i)$$

$$(A \cap B_i) \cap (A \cap B_j) = B_i \cap B_j = \emptyset$$

Th. di Bayes

Siano $A, B \in \mathcal{F}$ con $\mathbb{P}(A) > 0$

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

Dim $\quad \mathbb{P}(B \mid A) = \dfrac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} \quad , \quad \mathbb{P}(A \mid B) = \dfrac{\mathbb{P}(B \cap A)}{\mathbb{P}(B)}$

$$\Longrightarrow \mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \mid B)\, \mathbb{P}(B)}{\mathbb{P}(A)}$$

Fondamentale per l'intera probabilità e statistica.
Usato nei test di screening per stabilire la prob. della presenza di una malattia dopo l'esito di un test.
(note le sue SENSIBILITÀ e la sua SPECIFICITÀ)

Sensibilità: probabilità di essere ~~malato~~ positivo se ~~il tampone è positivo~~ nei malati $\mathbb{P}(+ | m) = SE$

Specificità: probabilità di essere ~~sano~~ negativo se ~~il tampone risulta negativo~~ nei sani $\mathbb{P}(- | s) = SP$

$$\mathbb{P}(m | +) = \frac{\mathbb{P}(+|m)\,\mathbb{P}(m)}{\mathbb{P}(+)} , \quad \mathbb{P}(+) = \mathbb{P}(+|m)\,\mathbb{P}(m) + \mathbb{P}(+|s)\,\mathbb{P}(s)$$

sapendo quanto vale $\mathbb{P}(m) = 1 - \mathbb{P}(s) = m$

$$\mathbb{P}(m | +) = \frac{m \cdot \mathbb{P}(+|m)}{\mathbb{P}(+|m)\cdot m + (1-m)(1-\mathbb{P}(-|s))} = \frac{m\,SE}{m\,SE + (1-m)(1-SP)}$$

Problema di Monty Hall

Hai davanti a te 3 porte. Ne scegli una a caso e ti verrà ~~~~ aperta ed esclusa una porta. Una delle 2 porte rimaste contiene una macchina. Cambi o no la porta?

Sol ~~Da~~ All'inizio ho $1/3$ di possibilità di scegliere la porta giusta. $\Rightarrow$ Ho $2/3$ di possibilità di aver scelto quella sbagliata.
Se ho scelto quella sbagliata ($2/3$) il conduttore dovrà eliminare per forza la porta non contenente la macchina tra le 2 porte rimaste $\Rightarrow$ il conduttore tra le porte A e B lascerà quella con dietro la macchina. Se cambi porta hai vinto.

Se quindi scegli la porta sbagliata all'inizio vinci.
($p = 2/3 \approx 0.67$)

Se invece scegli la porta giusta, cambiando porta perderai ma questo succede solo $1/3$ delle volte.

Possiamo risolvere anche con il th. di Bayes

A_n = " ~~la~~ macchina è dietro la porta n"

$$\mathbb{P}(A_1) = \mathbb{P}(A_2) = \mathbb{P}(A_3) = 1/3$$

B_n = " il conduttore apre la porta n"

$$\left.\begin{aligned} \mathbb{P}(B_2 \mid A_1) &= 1/2 \\ \mathbb{P}(B_2 \mid A_2) &= 0 \\ \mathbb{P}(B_2 \mid A_3) &= 1 \end{aligned}\right\} \text{ ipotizzo che scelgo la porta 1}$$

$$\mathbb{P}(A_1 \mid B_2) = \frac{\mathbb{P}(B_2 \mid A_1)\,\mathbb{P}(A_1)}{\mathbb{P}(B_2)} = \frac{1/2 \cdot 1/3}{\mathbb{P}(B_2)} = 1/3$$

$$\mathbb{P}(B_2) = \sum_{i=1}^{3} \mathbb{P}(B_2 \mid A_i)\,\mathbb{P}(A_i) = 1/3 \sum_{i=1}^{3} \mathbb{P}(B_2 \mid A_i) = 1/2$$

$$\mathbb{P}(A_2 \mid B_2) = 0 \ \wedge \ \mathbb{P}(A_1 \mid B_2) = 1/3 \implies$$

$$\implies \mathbb{P}(A_3 \mid B_2) = 2/3$$

__Nota__ ricorda che $\sum_i \mathbb{P}(A_i \mid X) = 1$ con $A_i \cap A_j = \emptyset$ $\forall\, i,j : i \neq j$

__Eventi indipendenti__

Siano A, B eventi $\implies$ $\mathbb{P}(A \mid B) = \mathbb{P}(A)$ $\wedge$
$\mathbb{P}(B \mid A) = \mathbb{P}(B)$

~~diciamo~~ diciamo che se A e B sono INDIPENDENTI

$$\mathbb{P}(A \mid B) = \mathbb{P}(A \cap B)\,\frac{1}{\mathbb{P}(B)} = \mathbb{P}(A) \Rightarrow$$

$$\Rightarrow \boxed{\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)}$$

2 eventi sono ind. se $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$

2 eventi sono DISGIUNTI se $A \cap B = \emptyset$

2 eventi ~~indipen~~ DISGIUNTI non potranno mai essere INDIPENDENTI

DISGIUNTI $\Rightarrow$ DIPENDENTI

Def Eventi $A_1, A_2 \dots A_n \in \mathcal{F}$ sono indipendenti se

$$\forall S \subset \{1 \dots n\} : \mathbb{P}\left(\bigcap_{i \in S} A_i\right) = \prod_{i \in S} \mathbb{P}(A_i)$$

Variabili aleatorie

$X : \Omega \to \mathbb{R}$ è una funzione che verifica la proprietà

- Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità.

Gli insiemi del tipo $\{\omega \in \Omega \text{ t.c. } X(\omega) \in I\}$ sono eventi di $\mathcal{F}$ $\Rightarrow$ ha sempre senso $\mathbb{P}(\{\omega \in \Omega \text{ t.c. } X(\omega) \in I\})$

(I intervallo aperto / chiuso, limitato / illimitato)

Variabili aleatorie equivalenti

Siano X, Y var. aleatorie

X ed Y sono EQUIVALENTI se $\forall k \in \mathbb{R} : \mathbb{P}(\{X = k\}) = \mathbb{P}(\{Y = k\})$

Legge di una var. aleatoria

Sia $X : \Omega \to \mathbb{R}$ una variabile aleatoria

$$\mathbb{P}_X(I) = \mathbb{P}(\{X \in I\}) \quad \forall I \subset \mathbb{R}$$

P_X si dice "legge" (o "distribuzione") di X

(X è distribuita con legge P_X)

Def Il range di una variabile aleatoria X è l'immagine di X $(R(X) \equiv Im(X))$

Def X e Y sono variabili aleatorie IDENTICAMENTE DISTRIBUITE se hanno la stessa legge

$$\mathbb{P}(\{X \in I\}) = \mathbb{P}(\{Y \in I\}) \quad \forall I \subset \mathbb{R}$$

Def X e Y sono indipendenti se

$$\mathbb{P}(\{X_1 \in I_1\} \cap \{X_2 \in I_2\}) = \mathbb{P}(\{X_1 \in I_1\})\,\mathbb{P}(\{X_2 \in I_2\})$$

n var. aleatorie sono INDIPENDENTI se

$$\mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j) \quad \forall J \subset \{1 .. n\}$$

Def X e Y sono i.i.d. se sono INDIPENDENTI E IDENTICAMENTE DISTRIBUITE.

Def Una variabile aleatoria X si dice DISCRETA se il suo range è ~~finito~~ numerabile.

$$R(X) = \{x_1, x_2 \dots x_n\}$$

Se X è discreta per conoscere la sua legge è sufficiente conoscere $\mathbb{P}(\{X = x_i\})$ per $i = 1 .. n ..$

Valore atteso

Sia X una var. aleatoria discreta.

Il valore atteso $\mathbb{E}(X)$ di X è

$$\boxed{\mathbb{E}(X) = \sum_{x \in R(X)} x\, \mathbb{P}(\{X = x\})}$$

LOTUS (Law of the Unconscious Statistician)

Sia X una variabile aleatoria $\Omega \to \mathbb{R}$ discreta

sia $H : \mathbb{R} \to \mathbb{R}$ una trasformazione

$$\boxed{\mathbb{E}(H(X)) = \sum_{x \in R(X)} H(x) \, \mathbb{P}(\{X = x\})}$$

Th il valore atteso è lineare

Dim $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$

$$\mathbb{E}(aX + bY) = \sum_{(x,y) \in R(X,Y)} H(x,y) \, \mathbb{P}(\{(X,Y) = (x,y)\}) =$$

$$= \sum_{xy} (ax + by) \cdot \mathbb{P}(\{X = x\} \cap \{Y = y\}) =$$

$$= a \sum_{x \in R(X)} \sum_{y \in R(Y)} x \, \mathbb{P}(\{X = x\} \cap \{Y = y\}) + b \sum_{y \in R(Y)} \sum_{x \in R(X)} y \, \mathbb{P}(\ldots) =$$

$$= a \sum_{x \in R(X)} x \, \mathbb{P}(\{X = x\}) + b \sum_{y \in R(Y)} y \, \mathbb{P}(\{Y = y\}) =$$

$$= a\mathbb{E}(X) + b\mathbb{E}(Y)$$

Varianza

Sia $X : \Omega \to \mathbb{R}$ una var. aleatoria discreta

La varianza di X è $\text{Var}(X) = \sigma^2(X)$

$$\boxed{\sigma^2(X) = \mathbb{E}((X - \mathbb{E}(X))^2)}$$

$$\sigma(X) = \sqrt{\sigma^2(X)}$$ DEV. STD.

$$\mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2 - \overbrace{2}^{\text{cost.}} X \overbrace{\mathbb{E}(X)}^{\text{cost.}} + \overbrace{\mathbb{E}(X)^2}^{\text{cost}}) =$$

$$= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \cancel{\mathbb{E}(}\mathbb{E}(X)^2\cancel{)} =$$

$$= \boxed{\mathbb{E}(X^2) - \mathbb{E}(X)^2}$$

Oss $\text{Var}(X) = 0 \Longrightarrow X = \mathbb{E}(X)$ cost.

Th $Var(aX+b) = a^2 Var(X)$

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

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

$= a^2 Var(X)$

Covarianza

Siano $X, Y : \Omega \to \mathbb{R}$ var. aleatorie

La covarianza di X e Y è $Cov(X,Y)$

$$\boxed{Cov(X,Y) = E((X-E(X))(Y-E(Y)))} =$$

$$= E(XY - XE(Y) - YE(X) + E(X)E(Y)) =$$

$$= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) =$$

$$= \boxed{E(XY) - E(X)E(Y)}$$

Th $\boxed{Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)}$

Dim $E([X+Y-E(X+Y)]^2) =$

$= E([X+Y-E(X)-E(Y)]^2) =$

$= E([X-E(X)]^2 + [Y-E(Y)]^2 + 2(X-E(X))(Y-E(Y))) =$

$= E([X-E(X)]^2) + E([Y-E(Y)]^2) + 2E((X-E(X))(Y-E(Y))) =$

$= Var(X) + Var(Y) + 2Cov(X,Y)$

Th X e Y sono indipendenti $\overset{\not\Leftarrow}{\Longrightarrow}$ $Cov(X,Y) = 0$

Dim $Cov(X,Y) = E(XY) - E(X)E(Y) = 0$

$$E(XY) = \sum_{(x,y)\in R(X,Y)} xy\, P(\{(X,Y) = (x,y)\}) =$$

$$= \sum_{(x,y)\in R(X,Y)} x\, y\, P(\{X=x\})P(\{Y=y\}) =$$

$$= \sum_{x\in R(X)} x\, P(\{X=x\}) \sum_{y\in R(Y)} y\, P(\{Y=y\}) = E(X)E(Y)$$

Vettori aleatori

È una mappa $X = (X_1, X_2 \dots X_n) : \Omega \to \mathbb{R}^n$ dove $X_1, X_2 \dots X_n$ sono var. aleatorie.

La legge CONGIUNTA è la misura di probabilità su $\mathbb{R}^n$ data da $\mathbb{P}_X(E) = \mathbb{P}(\{X \in E\})$, $E \subset \mathbb{R}^n$

Le leggi $\mathbb{P}_{X_1}, \mathbb{P}_{X_2} \dots \mathbb{P}_{X_n}$ delle componenti $X_1, X_2 \dots X_n$ sono dette leggi MARGINALI.

Def Un vettore aleatorio è DISCRETO se ogni sua componente X_i è discreta.

$X_2 \backslash X_1$	0	1	2	3	... ← $R(X_1)$
0	P_{00}	P_{10}	P_{20}	P_{30}	
1	P_{01}	P_{11}	P_{21}	P_{31}	
2	P_{02}	P_{12}	P_{22}	P_{32}	
3	P_{03}	P_{13}	P_{23}	P_{33}	
↑ $R(X_2)$					

prodotto cart.

$$R(X) = R(X_1) \times R(X_2)$$

con

$$P_{xy} = \mathbb{P}(\{(X_1, X_2) = (x, y)\})$$

$$\mathbb{P}(\{X_1 = \bar{x}\}) = \sum_{y \in R(X_2)} \mathbb{P}(\{X_1 = \bar{x} \mid X_2 = y\}) \cdot \mathbb{P}(\{X_2 = y\}) \Rightarrow$$

$$\Rightarrow \mathbb{P}(\{X_1 = \bar{x}\}) = \sum_{y \in R(X_2)} \mathbb{P}(\{(X_1, X_2) = (\bar{x}, y)\}) \Rightarrow$$

$$\Rightarrow \mathbb{P}(\{X_1 = \bar{x}\}) = \sum_{y \in R(X_2)} P_{\bar{x}y}$$

Def X e Y sono vettori aleatori indipendenti se

$$\mathbb{P}(\{X \in E\} \cap \{Y \in F\}) = \mathbb{P}(\{X \in E\})\, \mathbb{P}(\{Y \in F\})$$

$$\forall E, F \subset \mathbb{R}^n$$

n vettori aleatori sono ind. se $\mathbb{P}\left(\bigcap_{i=1}^{n} \{X_i \in E_i\}\right) =$

$$= \prod_{i=1}^{n} \mathbb{P}(\{X_i \in E_i\}) \quad \forall E_i \subset \mathbb{R}^n$$

(per considerare ~~[illegible]~~ gruppi non globali prendo $E_i = \mathbb{R}^n$)

es. $X_1, X_3, X_7 \Rightarrow E_2, E_4, E_5, E_6 = \mathbb{R}^n$

Prop Siano $X_1, X_2 \dots X_n$, $Y_1, Y_2 \dots Y_m$ ind.

$\Longrightarrow X = (X_1, X_2 \dots X_n)$ e $Y = (Y_1, Y_2 \dots Y_m)$ sono IND.

Dim $E = I_1 \times I_2 \times \dots \times I_n$, $F = J_1 \times J_2 \dots \times J_m$

$$\mathbb{P}(\{X \in E\} \cap \{Y \in F\}) = \mathbb{P}(\{X_1 \in I_1\} \cap \{X_2 \in I_2\} \dots \cap$$
$$\cap \{X_n \in I_n\} \cap \{Y_1 \in J_1\} \cap \{Y_2 \in J_2\} \dots \cap \{Y_m \in J_m\}) =$$
$$= \mathbb{P}(\{X_1 \in I_1\} \cap \dots \cap \{X_n \in I_n\}) \cdot \mathbb{P}(\{Y_1 \in J_1\} \cap \dots \{Y_m \in J_m\}) =$$
$$= \mathbb{P}(\{X \in E\}) \, \mathbb{P}(\{Y \in F\})$$

Funzioni di vettori aleatori

$X = (X_1, X_2 \dots X_n) : \Omega \to \mathbb{R}^n$

Consideriamo $H : \mathbb{R}^n \to \mathbb{R}$, $H(X_1, X_2 \dots X_n)$ è una variabile aleatoria.

Prop Se $X = (X_1, X_2 \dots X_n)$ e $Y = (Y_1, Y_2 \dots Y_m)$ sono vettori aleatori indipendenti

$H : \mathbb{R}^n \to \mathbb{R}$ e $G : \mathbb{R}^m \to \mathbb{R}$ funzioni di vettori aleatori.

$H(X_1, X_2 \dots X_n)$ e $G(Y_1, Y_2 \dots Y_m)$ sono INDIPENDENTI.

Dim $\mathbb{P}(\{H(X) \in I\} \cap \{G(Y) \in J\}) =$

$$= \mathbb{P}(\{X \in H^{-1}(I)\} \cap \{Y \in G^{-1}(J)\}) =$$
$$= \mathbb{P}(\{X \in H^{-1}(I)\}) \, \mathbb{P}(\{Y \in G^{-1}(J)\}) =$$
$$= \mathbb{P}(\{H(X) \in I\}) \, \mathbb{P}(\{G(Y) \in J\})$$

Leggi note

Ci sono varie leggi probabilistiche note.

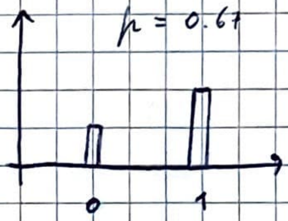
Appartengono a 2 classi principali:

1. DISCRETE
2. CONTINUE

LEGGI DISCRETE

- **Legge di Bernoulli** $X \sim Be(p)$, $p \in (0,1)$

Descrive il successo / insuccesso



$$\boxed{\begin{aligned} \mathbb{P}(\{X=0\}) &= 1-p \\ \mathbb{P}(\{X=1\}) &= p \end{aligned}}$$

$$\boxed{\begin{aligned} \mathbb{E}(X) &= p \\ Var(X) &= p(1-p) \end{aligned}}$$

Dim $\mathbb{E}(X) = \sum_{x \in R(X)} x\,\mathbb{P}(\{X=x\}) = p$

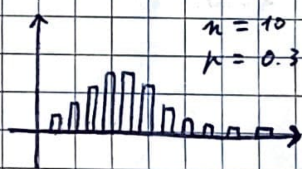
$Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \sum_{x \in R(X)} x^2\,\mathbb{P}(\{X=x\}) = p - p^2$

- **Legge Binomiale** $X \sim B(n,p)$, $n > 0$, $p \in (0,1)$

Descrive il numero di successi k in n tentativi.

È la somma di $X_i \sim Be(p)$ iid $\forall\, i \in \{1 \dots n\}$

$$\boxed{\mathbb{P}(\{X=k\}) = \binom{n}{k} p^k (1-p)^{n-k}} \quad \forall\, k \in [0,n]$$



$$\boxed{\begin{aligned} \mathbb{E}(X) &= np \\ Var(X) &= np(1-p) \end{aligned}}$$

X_i ind $\Rightarrow$ cov $= 0$

Dim $\mathbb{E}(X) = \sum_{i \in \{1..n\}} \mathbb{E}(\underbrace{X_i}_{\sim Be(p)}) = np$

$Var(X) = \sum_{i=1}^{n} Var(X_i) = n p (1-p)$

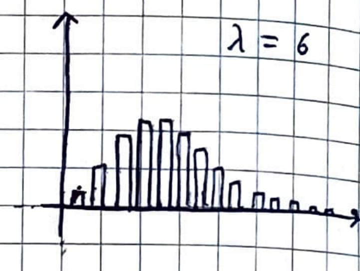
Siano $X \sim B(n,p) \wedge Y \sim B(m,p)$

$\Rightarrow X + Y \sim B(n+m, p)$

Dim $X + Y = \sum_{i=1}^{n} X_i + \sum_{j=1}^{m} Y_j = \sum_{i=1}^{n+m} X_i \sim B(n+m, p)$, $X_i \sim Be(p)$

- **Legge di Poisson** $X \sim \mathcal{P}(\lambda)$, $\lambda > 0$

Descrive il num. di successi in "tanti" tentativi indipend. con una media fissata.



$$\boxed{\mathbb{P}(\{X = k\}) = \frac{\lambda^k}{k!} e^{-\lambda}} \quad , \quad \lambda = n \cdot p$$

$$\boxed{\begin{aligned} \mathbb{E}(X) &= \lambda \\ \mathrm{Var}(X) &= \lambda \end{aligned}}$$

Dem $\lim_{n \to +\infty} \mathbb{P}(\{X_n = k\}) = \frac{\lambda^k}{k!} e^{-\lambda}$, $X_n \sim B(n, p)$

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

$$= \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{\lambda^k}{n^k} \frac{n(n-1)\dots(n-k+1)}{k!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} =$$

$$= \frac{\lambda^k}{k!} \frac{n(n-1)\dots(n-k+1)}{\underbrace{n \cdot n \cdots n}_{k \text{ volte}}} \left(1 - \frac{\lambda}{n}\right)^{-k} \left(1 - \frac{\lambda}{n}\right)^n \longrightarrow$$

$$\longrightarrow \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$\left(\lim_{n \to +\infty} \left(1 - \frac{\lambda}{n}\right)^n \overset{y = -\frac{n}{\lambda}}{=\!=} \lim_{y \to -\infty} \left(1 + \frac{1}{y}\right)^{-\lambda y} = e^{-\lambda} \right)$$

$$\mathbb{E}(X) = \sum_{k=0}^{+\infty} k \, \mathbb{P}(\{X = k\}) = \sum_{k=0}^{+\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = \sum_{k=1}^{+\infty} k e^{-\lambda} \frac{\lambda^k}{k!} =$$

$$= e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

$$\left(\sum_{k=0}^{+\infty} \frac{a^k}{k!} = e^a \right)$$

$$\mathrm{Var}(X) = np(1-p) \Longrightarrow \lambda \left(1 - \frac{\lambda}{n}\right) \xrightarrow{n \to +\infty} \lambda$$

$$\mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 =$$

$$= \sum_{k=0}^{+\infty} k^2 e^{-\lambda} \frac{\lambda^k}{k!} - \lambda^2 = \lambda e^{-\lambda} \sum_{k=1}^{+\infty} k \frac{\lambda^{k-1}}{(k-1)!} - \lambda^2 =$$

$$= \left(\sum_{h=0}^{+\infty} (h+1) \frac{\lambda^h}{h!} \right) \lambda e^{-\lambda} - \lambda^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Th $X \sim \mathcal{P}(\lambda)$, $Y \sim \mathcal{P}(\mu)$ $^{(\text{ind})}$ $\Rightarrow X + Y \sim \mathcal{P}(\lambda + \mu)$.

Dim $\mathbb{P}(\{X+Y=k\}) = \sum_{h=0}^{k} \mathbb{P}(\{X=h\} \cap \{Y=k-h\}) =$

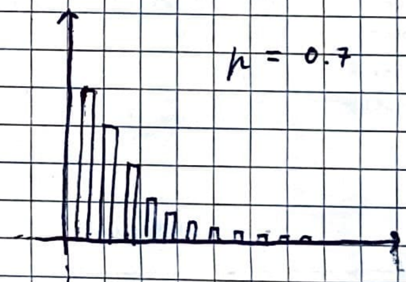
$$= \sum_{h=0}^{k} \mathbb{P}(\{X=h\})\,\mathbb{P}(\{Y=k-h\}) =$$

$$= \sum_{h=0}^{k} e^{-\lambda}\frac{\lambda^h}{h!}\, e^{-\mu}\frac{\mu^{k-h}}{(k-h)!} = \sum_{h=0}^{k} e^{-(\lambda+\mu)}\frac{\lambda^h \mu^{k-h}}{h!\,(k-h)!}\cdot\frac{k!}{k!} =$$

$$= e^{-(\lambda+\mu)}\frac{1}{k!}\underbrace{\sum_{h=0}^{k}\binom{k}{h}\lambda^h\mu^{k-h}}_{\text{POL. DI NETWON}} = e^{-(\lambda+\mu)}\frac{(\lambda+\mu)^k}{k!}$$

- **Legge geometrica** $X \sim Geo(p)$, $p \in (0,1)$

Descrive il primo successo in n tentativi ind. con probabilità p.



$$\boxed{\mathbb{P}(\{X=k\}) = (1-p)^{k-1}\,p}$$

$$\boxed{\begin{aligned} \mathbb{E}(X) &= 1/p \\ Var(X) &= \frac{1-p}{p^2}\end{aligned}}$$

Dim $\mathbb{E}(X) = \sum_{k=1}^{+\infty} k\,\mathbb{P}(\{X=k\}) = p\sum_{k=1}^{+\infty} k(1-p)^{k-1}$ $\quad q = 1-p$

$$= p\sum_{k=1}^{+\infty} k\,q^{k-1} = p\sum_{k=1}^{+\infty}\frac{d}{dq}q^k = p\frac{d}{dq}\sum_{k=0}^{+\infty}q^k =$$

$q^0 = 1$ cost $\left(\frac{d}{dq}1 = 0\right)$

$$= p\frac{d}{dq}\left(\frac{1}{1-q}\right) = p\frac{1}{(1-q)^2} = 1/p$$

$$Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \sum_{k=1}^{+\infty}k^2\,\mathbb{P}(\{X=k\}) - 1/p^2 =$$

$$= p\sum_{k=1}^{+\infty}k^2 q^{k-1} - 1/p^2 = p\frac{d}{dq}\sum_{k=1}^{+\infty}k\,q^k - 1/p^2 = p\frac{d}{dq}\,q\sum_{k=0}^{+\infty}\frac{d}{dq}q^k - 1/p^2 =$$

$$= p\frac{d}{dq}\frac{q}{(1-q)^2} - \frac{1}{p^2} = p\,\frac{(1-q)^{\cancel{2}} + 2q\cancel{(1-q)}}{(1-q)^{\cancel{4}\,3}} - \frac{1}{(1-q)^2} = $$

$$= \cancel{p}\,\frac{p + 2(1-p)}{p^2} - \frac{1}{p^2} = \frac{\cancel{p} + 2 - 2p - \cancel{1}}{p^2} = \frac{1-p}{p^2}$$

La legge geometrica gode di <u>assenza di memoria</u>
(è l'unica legge discreta con questa proprietà)

$$\boxed{\mathbb{P}(\{X > m+n\} \mid \{X > n\}) = \mathbb{P}(\{X > m\})}$$

<u>Dim</u> $\mathbb{P}(\{X > m+n\} \mid \{X > n\}) = \dfrac{\mathbb{P}(\{X > m+n\} \cap \{X > n\})}{\mathbb{P}(\{X > n\})} =$

$$= \frac{\mathbb{P}(\{X > m+n\})}{\mathbb{P}(\{X > n\})} = \frac{(1-p)^{m+n}}{(1-p)^{n}} = (1-p)^{m} =$$

$$= \mathbb{P}(\{X > m\})$$

<u>Prop</u> $\mathbb{P}(\{X > k\}) = (1-p)^k$

<u>Dim</u> $\mathbb{P}(\{X > k\}) = \sum_{h=k+1}^{+\infty} \mathbb{P}(\{X = h\}) = p \sum_{h=k+1}^{+\infty} (1-p)^{h-1} \stackrel{j=h-k-1}{=} p \sum_{j=0}^{+\infty} (1-p)^{j+k} = p(1-p)^k \sum_{j=0}^{+\infty} (1-p)^j = (1-p)^k$

<u>Leggi continue</u>

Una var. aleatoria $X : \Omega \to \mathbb{R}$ ha funzione di densità di probabilità (<u>f.d.p.</u>) $f_X : \mathbb{R} \to [0, +\infty)$ se

$$\mathbb{P}(\{X \in I\}) = \int_I f_X(x)\,dx$$

<u>Oss</u> $\displaystyle\int_{\mathbb{R}} f_X(x)\,dx = 1 \quad \wedge \quad \int_a^a f_X(x)\,dx = 0 = \mathbb{P}(\{X = a\})$

La funzione di distribuzione cumulativa (<u>F.D.C</u>) è

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \mathbb{P}(\{X \le x\})$$

<u>Oss</u> $F_X(x) \xrightarrow{x \to -\infty} 0$

$F_X(x) \xrightarrow{x \to +\infty} 1$

$F_X(x_1) \le F_X(x_2) \iff x_1 \le x_2$

Se X ha f.d.p. $f_X \in \mathcal{C}^0(\mathbb{R})$ [a tratti]

$$\Rightarrow \boxed{F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt} \text{ è derivabile } \wedge \ F_X'(x) = f_X(x)$$

per il Th. fondamentale del calcolo integrale.

Valore atteso

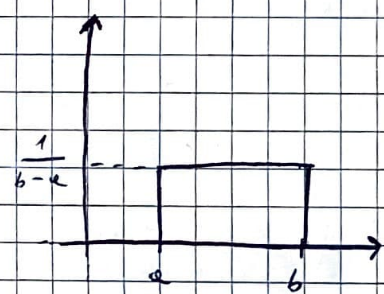
$$\boxed{\mathbb{E}(X) = \int_{\mathbb{R}} x\, f_X(x)\,dx} \ , \quad f_X(x) \text{ f.d.p. di } X : \Omega \to \mathbb{R}$$

Varianza

$$\boxed{Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2}$$

LEGGI CONTINUE

- **Legge uniforme** $X \sim U(a,b)$ $\quad a < b$



$$\boxed{f_X(x) = \frac{1}{\# I} = \frac{1}{b-a}} \ , \quad x \in (a,b)$$

$$\boxed{\begin{aligned} \mathbb{E}(X) &= \frac{a+b}{2} \\ Var(X) &= \frac{(b-a)^2}{12} \end{aligned}}$$

Dim

$$\mathbb{E}(X) = \int_a^b \frac{x}{b-a}\,dx = \frac{b^2-a^2}{b-a} \cdot \frac{1}{2} = \frac{a+b}{2}$$

$$Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \int_a^b \frac{x^2}{b-a}\,dx - \frac{(a+b)^2}{4} =$$

$$= \frac{1}{3}\,\frac{b^3-a^3}{b-a} - \frac{1}{4}(a^2+2ab+b^2) = \frac{1}{3}(b^2+a^2+ab) -$$

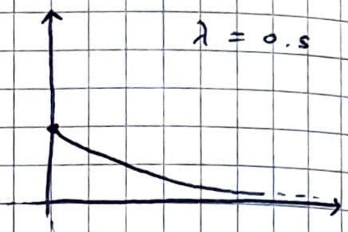
$$- \frac{1}{4}(a^2+2ab+b^2) = \frac{1}{12}\big(4b^2+4a^2+4ab-3a^2-3b^2 - 6ab\big) =$$

$$= \frac{1}{12}(a^2 - 2ab + b^2) = \frac{(b-a)^2}{12}$$

- **Legge esponenziale** $X \sim Exp(\lambda)$, $\lambda > 0$

Descrive i tempi di attesa con proprietà particolari.

$$\boxed{f_X(x) = \lambda e^{-\lambda x}}, \quad x > 0$$



$$\boxed{\begin{aligned} \mathbb{E}(X) &= 1/\lambda \\ Var(X) &= 1/\lambda^2 \end{aligned}}$$

Dim $E(X) = \int_0^{+\infty} x f_X(x)\, dx = \int_0^{+\infty} x \lambda e^{-\lambda x} dx =$

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

$$Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = 1/\lambda^2$$

$$\mathbb{E}(X^2) = \int_0^{+\infty} x^2 f_X(x)\, dx = -x^2 e^{-\lambda x}\Big|_0^{+\infty} + \int_0^{+\infty} 2x e^{-\lambda x} dx =$$

$$= 2\left(x\left(-\frac{1}{\lambda}\right) e^{-\lambda x}\Big|_0^{+\infty} + \int_0^{+\infty} \frac{1}{\lambda} e^{-\lambda x} dx \right) = \frac{2}{\lambda^2}$$

Oss $\mathbb{P}(\{X > t\}) = \int_t^{+\infty} \lambda e^{-\lambda x} dx = e^{-\lambda x}\Big|_{+\infty}^{t} = e^{-\lambda t}$

Oss $\mathbb{P}(\{X > t+s\} \mid \{X > s\}) = \mathbb{P}(\{X > t\})$

La legge esponenziale è l'unica legge continua che gode di assenza di memoria.

Dim $\mathbb{P}(\{X > t+s\} \mid \{X > s\}) = \dfrac{\mathbb{P}(\{X > t+s\} \cap \{X > s\})}{\mathbb{P}(\{X > s\})} =$

$$= \frac{\mathbb{P}(\{X > t+s\})}{\mathbb{P}(\{X > s\})} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t} = \mathbb{P}(\{X > t\})$$

Vettori aleatori continui

Un vettore aleatorio ha f.d.p. **congiunta** $f_{X,Y}(x,y)$ se

$$\mathbb{P}(\{(X,Y)\in E\}) = \iint_E f(x,y)\,dx\,dy$$

f_X e f_Y si chiamano <u>MARGINALI</u>

$$f_X = \int_{-\infty}^{+\infty} f_{(X,Y)}(x,y)\,dy \quad \wedge \quad F_X(x) = \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f_{(X,Y)}(t,y)\,dy\,dt$$

<u>Variabili aleatorie indipendenti</u>

X e Y sono indipendenti se

$$\boxed{f_{(X,Y)}(x,y) = f_X(x)\,f_Y(y)}$$

<u>Dim</u> $\mathbb{P}(\{(X,Y)\in(-\infty,x]\times(-\infty,y]\}) =$

$= \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(t,z)\,dz\,dt =$

$= \mathbb{P}(\{X\le x\}\cap\{Y\le y\}) =$

$= \int_{-\infty}^{x} f_X(t)\,dt \int_{-\infty}^{y} f_Y(z)\,dz =$

$= \mathbb{P}(\{X\le x\})\,\mathbb{P}(\{Y\le y\})$

Derivando in x e y otteniamo $f_{X,Y}(x,y) = f_X(x)\,f_Y(y)$

<u>Th</u> Siano X e Y ind. con f.d.p. $f_X(x)$ e $f_Y(y)$

$\Longrightarrow F_{X+Y}(z) = \mathbb{P}(\{X+Y\le z\}) = \int_{-\infty}^{z} f_{X+Y}(t)\,dt =$

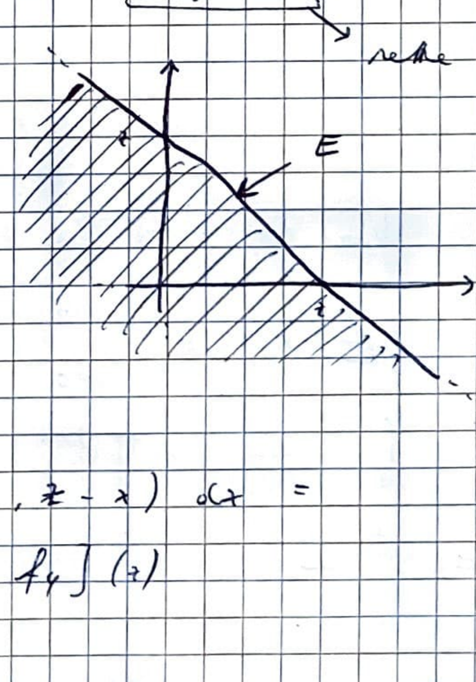
$= \mathbb{P}(\{(X,Y)\in E\})$, $E = \{(x,y)\in\mathbb{R}^2 \ni' \underbrace{x+y\le z}_{\text{retta}}\}$

$= \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{z-x} f_{(X,Y)}(x,y)\,dy\right)dx \overset{z=y+x}{=}$

$= \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{z} f_{(X,Y)}(x,z-x)\,dz\right)dx =$

$= \int_{-\infty}^{z}\underbrace{\int_{-\infty}^{+\infty} f_{X,Y}(x,z-x)\,dx}\,dz =$

$= \int_{-\infty}^{z} f_{X+Y}(t)\,dt$



$F'_{X+Y}(z) = \boxed{f_{X+Y}(z)} = \int_{-\infty}^{+\infty} f_{X,Y}(x,z-x)\,dx =$

$= \boxed{\int_{-\infty}^{+\infty} f_X(x)\,f_Y(z-x)\,dx} = [f_X * f_Y](z)$

$\boxed{f_X * f_Y}$ è la <u>CONVOLUZIONE</u>

Convoluzione $f_X * f_Y(z) = \int_{-\infty}^{+\infty} f_X(x)\, f_Y(z-x)\, dx$

es: Calcoliamo con la convoluzione $X_1 + X_2$ con $X_1, X_2 \sim Exp(\lambda)$ ind.

$$f_{X_1+X_2}(z) = \int_{-\infty}^{+\infty} f_{X_1}(x)\, f_{X_2}(z-x)\, dx =$$

$$= \int_0^z \lambda e^{-\lambda x} \lambda e^{-\lambda(z-x)}\, dx = \int_0^z \lambda^2 e^{-\lambda z}\, dx = z \lambda^2 e^{-\lambda z}$$

$$f_{X_1+X_2+X_3}(z) = f_{X_1+X_2} * f_{X_3}(z) = \dots = \lambda^3 \frac{z^2}{2} e^{-\lambda z}$$

- **Legge Gamma** $X \sim Gamma(\alpha, \lambda)$, $\alpha > 0$, $\lambda > 0$

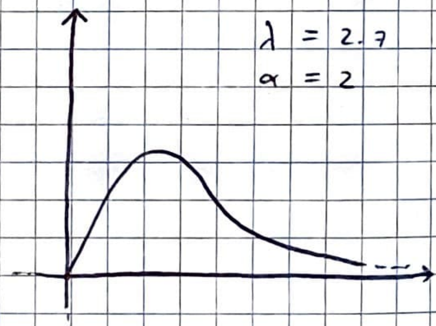
È la somma di α esponenziali $X_i \sim Exp(\lambda)$

$$\sum_{i=1}^{\alpha} X_i \sim Exp(\lambda) = X \sim Gamma(\alpha, \lambda)$$

$$\boxed{f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}}$$

$$\Gamma(\alpha) = \int_0^{+\infty} y^{\alpha-1} e^{-y}\, dy = (\alpha-1)! \quad \forall \alpha \in \mathbb{N}$$

$$\boxed{\mathbb{E}(X) = \frac{\alpha}{\lambda} \qquad Var(X) = \frac{\alpha}{\lambda^2}}$$



Dim ① $\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^{\alpha} X_i \sim Exp(\lambda)\right) = \sum_{i=1}^{\alpha} \mathbb{E}(X_i) = \frac{\alpha}{\lambda}$

$Var(X) = Var\left(\sum_{i=1}^{\alpha} X_i \sim Exp(\lambda)\right) = \sum_{i=1}^{\alpha} Var(X_i) = \frac{\alpha}{\lambda^2}$

② $\mathbb{E}(X) = \int_0^{+\infty} x \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}\, dx =$

$$= \int_0^{+\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{(\alpha-1)+1} e^{-\lambda x}\, dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{\lambda^{\alpha+1}} \underbrace{\int_0^{+\infty} \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} x^{(\alpha+1)-1} e^{-\lambda x}\, dx}_{= 1 \; \left(\int_{-\infty}^{+\infty} f_X(x)\, dx = 1\right)} =$$

$$= \frac{\alpha}{\lambda}$$

$$Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{\alpha^2 + \alpha}{\lambda^2} - \frac{\alpha^2}{\lambda^2} = \frac{\alpha}{\lambda^2}$$

$$E(X^2) = \int_0^{+\infty} x^2 \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}\,dx =$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+2)}{\lambda^{\alpha+2}} \int_0^{+\infty} \frac{\lambda^{\alpha+2}}{\Gamma(\alpha+2)} x^{(\alpha+1)-1} e^{-\lambda x}\,dx =$$

$$= \frac{\alpha(\alpha+1)}{\lambda^2}$$

Th $X \sim \text{Gamma}(\alpha, \lambda) \wedge Y \sim \text{Gamma}(\beta, \lambda) \Rightarrow$

$\Rightarrow X + Y \sim \text{Gamma}(\alpha + \beta, \lambda)$

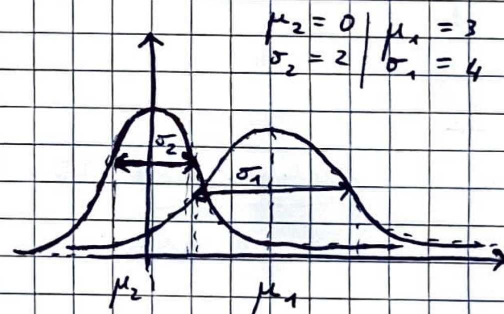
- **Legge Normale** $X \sim N(\mu, \sigma^2)$, $\mu \in R$, $\sigma^2 \geq 0$

La somma di $n > 30$ X_i iid porta ad una legge normale. (TH. LIMITE CENTRALE)

$$\boxed{f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}} \sim N(\mu, \sigma^2)$$

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \sim N(0,1)$$

$$\left.\boxed{\begin{array}{l} E(X) = \mu \\ Var(X) = \sigma^2 \end{array}}\right\} \text{per definizione}$$



$F_X(x) = \Phi(x)$ per $X \sim N(0,1)$

z-value : $z = \frac{x - \mu}{\sigma}$ (quante dev. std. dista x da μ)

Proprietà $X \sim N(\mu, \sigma^2)$, $Y \sim N(\nu, \tau^2)$

$\Rightarrow X + Y \sim N(\mu + \nu, \sigma^2 + \tau^2)$

$X + k \sim N(\mu + k, \sigma^2)$

$X \cdot a \sim N(a\mu, a^2\sigma^2)$

$aX + k \sim N(a\mu + k, a^2\sigma^2)$

STATISTICA INFERENZIALE

Def Un vettore aleatorio $X' = (X_1, X_2 \dots X_n)$ composto da var. aleatorie iid si chiama CAMPIONE CASUALE

Def I dati di un campione casuale $(X_1, X_2 \dots X_n)$ sono $(x_1, x_2 \dots x_n) \in R(X_1, X_2 \dots X_n)$

X è definita da parametri $\theta \in \Theta$

Es $X \sim Be(p) \Rightarrow \theta = p \quad \Theta = [0,1]$

$X \sim B(n,p) \Rightarrow \theta = (n,p) \quad \Theta = \mathbb{N} \times (0,1)$

Obiettivi della statistica inferenziale

Preso un campione casuale $(X_1, X_2 \dots X_n)$ dalla popolazione X definita dai parametri $\theta \in \Theta$ l'obiettivo della statistica inf. è quello di ottenere informazioni sulla popolazione a partire dai dati $(x_1 \dots x_n)$ (questo processo è detto "di INFERENZA")

Def Sia $(X_1 \dots X_n)$ un campione casuale.
Una statistica costruita a partire dal campione è una var. aleatoria $H(X_1 \dots X_n)$

Es La media campionaria è la statistica $\overline{X}_n$

$$\boxed{\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i}$$

$\overline{x}_n$ è la realizzazione di $\overline{X}_n$ calcolata sui dati

$$\overline{x}_n = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Se la popolazione X è definita da parametri $\theta \in \Theta$, $h(\theta)$ il parametro da stimare e $H(X_1 \dots X_n)$ ha lo scopo di stimare $h(\theta) \Rightarrow H(X_1 \dots X_n)$ STIMATORE di $h(\theta)$

Def Uno stimatore si dice CORRETTO se $H(X_1 \dots X_n)$ ha come valore atteso $h(\theta)$

$$\mathbb{E}(H(X_1 \dots X_n)) = h(\theta)$$

$H(x_1 \dots x_n)$ si dice STIMA PUNTUALE di $h(\theta)$

Ex La varianza campionaria è la statistica S_n^2

$$\boxed{S_n^2 = \frac{1}{n-1} \sum_i (X_i - \bar{X}_n)^2}$$

$s_n^2 = \frac{1}{n-1} \sum_i (x_i - \bar{x}_n)^2$ è la realizzazione sui dati

Oss $\boxed{\mathbb{E}(\bar{X}_n) = \mu}$

$\boxed{\mathrm{Var}(\bar{X}_n) = \frac{\sigma^2}{n}} \neq \sigma^2$ legge dei grandi numeri

Con $n \to +\infty$ $\mathrm{Var}(\bar{X}_n) \to 0$ (sempre più certo)

LEGGE DEI GRANDI NUMERI

Siano $X_1 \dots X_n$ iid con $\mathbb{E}(X_i) = \mu$ e $\mathrm{Var}(X_i) = \sigma^2$

$$\Longrightarrow \forall \varepsilon > 0 \quad \lim_{n \to +\infty} \mathbb{P}(\{|\bar{X}_n - \mu| < \varepsilon\}) \Rightarrow 1$$

Dim $\mathbb{P}(\{|\bar{X}_n - \mu| > \varepsilon\}) = 1 - \mathbb{P}(\{|\bar{X}_n - \mu| < \varepsilon\}) \xrightarrow{n \to +\infty} 0$

$$\mathbb{P}(\{|\bar{X}_n - \mu| > \varepsilon\}) = \mathbb{P}(\{\overbrace{(\bar{X}_n - \mu)^2}^{Y} > \varepsilon^2\}) =$$

$$= \int_{\varepsilon^2}^{+\infty} f_Y(y)\, dy = \int_{\varepsilon^2}^{+\infty} \frac{\varepsilon^2}{\varepsilon^2} f_Y(y)\, dy \leq \frac{1}{\varepsilon^2} \int_{\varepsilon^2}^{+\infty} y f_Y(y)\, dy \leq$$

$$\leq \frac{1}{\varepsilon^2} \int_0^{+\infty} y f_Y(y)\, dy = \frac{1}{\varepsilon^2} \mathbb{E}(Y) = \frac{1}{\varepsilon^2} \mathbb{E}((\bar{X}_n - \mu)^2) =$$

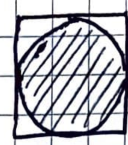
$$= \frac{1}{\varepsilon^2} \left(\mathrm{Var}(\bar{X}_n - \mu) + \mathbb{E}(\bar{X}_n - \mu)^2 \right) =$$

$$= \frac{1}{\varepsilon^2} \left(\frac{\sigma^2}{n} + 0^2 \right) = \frac{\sigma^2}{\varepsilon^2 n} \xrightarrow{n \to +\infty} 0$$

Dalla legge dei grandi numeri nasce la ~~...~~ SIMULAZIONE DI MONTECARLO.

Sapendo che con $n \to +\infty$ tentativi $\bar{X}_n \to \mu$ possiamo usare un calcolatore per simulare n tentativi e trovare μ

<u>Es</u> Sim. Montecarlo per π



$$\pi = \frac{\#\, succ}{\#\, tentativi} \cdot 4$$

$$\#\, successi = A_{\circ} = \pi\left(\frac{1}{2}\right)^2 = \frac{\pi}{4}$$

$$\#\, tentativi = A_{\square} = 1^2 = 1$$

$$\mu \longrightarrow \frac{\pi}{4}$$

TH. DEL LIMITE CENTRALE

Sia X popolazione con $\mathbb{E}(X) = \mu$ e $Var(X) = \sigma^2$

$X_1, X_2 \dots X_n$ iid

$$\Longrightarrow \lim_{n \to +\infty} \mathbb{P}\left(\left\{\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \le z\right\}\right) = \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt$$

Applicabile per $n \gtrsim 30$ senza aver bisogno di conoscere la legge di X_i

INTERVALLI DI CONFIDENZA (I.C.)

Un I.C. con livello di confidenza β è un intervallo aleatorio $[U_n, V_n]$ dove $U_n = u_n(X_1 \dots X_n)$ e $V_n = v_n(X_1 \dots X_n)$ sono statistiche ;

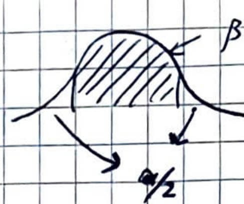
$$\boxed{\beta = \mathbb{P}(\{U_n \le h(\theta) \le V_n\})}$$

<u>Es</u> $h(\theta) = \mu$ (noto σ)

$$\beta = \mathbb{P}(\{U_n \le \mu \le V_n\}) = \mathbb{P}(\{-V_n \le \mu \le U_n\}) =$$

$$= \mathbb{P}\left(\left\{\frac{\bar{X}_n - V_n}{\sigma/\sqrt{n}} \le \underbrace{\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}}_{Z \sim N(0,1)} \le \frac{\bar{X}_n - U_n}{\sigma/\sqrt{n}}\right\}\right) =$$

$$= \mathbb{P}\left(\left\{ \frac{\bar X_n - V_n}{\sigma/\sqrt{n}} \le \cancel{\frac{\bar X_n - \mu}{\sigma/\sqrt{n}}}\; Z \le \frac{\bar X_n - U_n}{\sigma/\sqrt{n}} \right\}\right) =$$



$$= 1 - \cancel{2}\, \mathbb{P}\left(\left\{ Z \le \frac{\bar X_n - V_n}{\sigma/\sqrt{n}} \right\} \cup \left\{ Z \ge \frac{\bar X_n - U_n}{\sigma/\sqrt{n}} \right\}\right) =$$

$$= 1 - 2\, \underbrace{\mathbb{P}(\{ Z \ge z_{\alpha/2} \})}_{\alpha/2} \quad , \quad \alpha = 1 - \beta$$

$$\left(1 - 2\frac{\alpha}{2} = 1 - \alpha = \beta\right)$$

$$\frac{\bar X_n - V_n}{\sigma/\sqrt{n}} = -z_{\alpha/2} \implies V_n = \bar X_n + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}$$

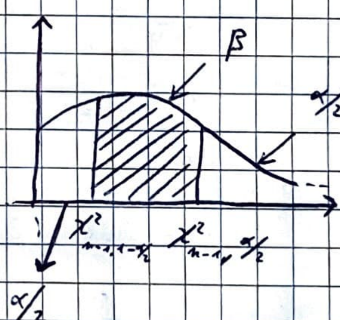
$$\frac{\bar X_n - U_n}{\sigma/\sqrt{n}} = z_{\alpha/2} \implies U_n = \bar X_n - \frac{\sigma}{\sqrt{n}} z_{\alpha/2}$$

$$\boxed{I.C. = \left[\bar X_n - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \, , \, \bar X_n + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right]}$$

$$z_{\alpha/2} = \Phi^{-1}\left(1 - \tfrac{\alpha}{2}\right) \qquad \left[z_x = \Phi^{-1}(1 - x) \right]$$

$$\left[\Phi(x) + \Phi(-x) = 1 \right]$$

Es $\quad h(\theta) = \sigma^2$



$$\beta = \mathbb{P}(\{ U_n \le \sigma^2 \le V_n \}) =$$

$$= \mathbb{P}\left(\left\{ \frac{1}{V_n} \le \frac{1}{\sigma^2} \le \frac{1}{U_n} \right\}\right) =$$

$$= \mathbb{P}\left(\left\{ \frac{(n-1) S_n^2}{V_n} \le \underbrace{\frac{(n-1) S_n^2}{\sigma^2}}_{Z \sim \chi^2(n-1)} \le \frac{(n-1) S_n^2}{U_n} \right\}\right)$$

$$\frac{(n-1) S_n^2}{V_n} = \chi^2_{n-1, 1-\alpha/2} \implies V_n = \frac{(n-1) S_n^2}{\chi^2_{n-1, 1-\alpha/2}}$$

$$\frac{(n-1) S_n^2}{U_n} = \chi^2_{n-1, \alpha/2} \implies U_n = \frac{(n-1) S_n^2}{\chi^2_{n-1, \alpha/2}}$$

$$\boxed{IC = \left[\frac{(n-1) S_n^2}{\chi^2_{n-1, \alpha/2}} \, , \, \frac{(n-1) S_n^2}{\chi^2_{n-1, 1-\alpha/2}} \right]}$$

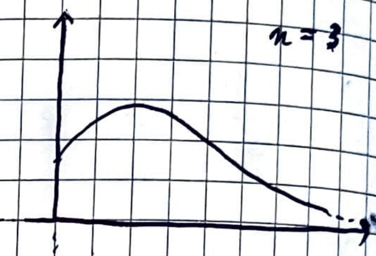
__Es__ $h(\theta) = \mu$ (σ non noto)

$$\beta = \mathbb{P}(\{ U_n \le \mu \le V_n \})$$

$$\beta = \mathbb{P}\left(\left\{ \frac{\bar{X}_n - V_n}{S_n/\sqrt{n}} \le \underbrace{\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}}}_{\sim\, t(n-1)} \le \frac{\bar{X}_n - U_n}{S_n/\sqrt{n}} \right\}\right)$$

__LEGGI CONTINUE PER ANALISI INFERENZIALE__

- __Chi - Quadro__ $X \sim \chi^2(n)$



$$\chi^2(n) = \text{Gamma}\left(\tfrac{n}{2}, \tfrac{1}{2}\right)$$

$$\boxed{f_X(x) = \frac{\lambda^{n/2}}{\Gamma(n/2)} x^{n/2 - 1} e^{-\frac{1}{2}x}}$$

__Oss__ $\sum_{i=1}^{n} X_i^2 = Z \sim \text{Gamma}\left(\tfrac{n}{2}, \tfrac{1}{2}\right)$, $X_i \sim N(0,1)$

__Dim__ $f_{Z^2}(x) = ?$

$$\mathbb{P}(\{Z^2 \le x\}) = 0 \quad \forall x \le 0$$

$$\mathbb{P}(\{Z^2 \le x\}) = \mathbb{P}(\{-\sqrt{x} \le z \le \sqrt{x}\}) = 2\,\mathbb{P}(\{0 \le z \le \sqrt{x}\}) =$$

$$= 2 \int_0^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \overset{s = t^2}{=\!=}$$

$$= \cancel{2} \int_0^{x} \frac{1}{\sqrt{2\pi}} \frac{1}{\cancel{2}\sqrt{s}} e^{-s/2} ds = F_{Z^2}(x)$$

$$F'_{Z^2}(x) = f_{Z^2}(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{x}} e^{-\frac{x}{2}} = \frac{(1/2)^{1/2}}{\underbrace{\Gamma(1/2)}_{\sqrt{\pi}}} x^{1/2 - 1} e^{-\frac{1}{2}x}$$

Perché si usa?

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left[(X_i - \mu)^2 + (\bar{X}_n - \mu)^2 -\right.$$

$$\left.- 2(X_i - \mu)(\bar{X}_n - \mu)\right] = \frac{1}{n-1}\left[\sum_{i=1}^{n}(X_i - \mu)^2 + n(\bar{X}_n - \mu)^2\right.$$

$$\left. - 2(\bar{X}_n - \mu)(n\bar{X}_n - n\mu)\right] =$$

$$= \frac{1}{n-1} \left[\sum_{i=1}^{n} (X_i - \mu)^2 - n (\bar{X}_n - \mu)^2 \right]$$

$$\Longrightarrow (n-1) S_n^2 = \sum_{i=1}^{n} (X_i - \mu)^2 - n (\bar{X}_n - \mu)^2$$

$$\Longrightarrow (n-1) S_n^2 \frac{1}{\sigma^2} + n \left(\frac{\bar{X}_n - \mu}{\sigma} \right)^2 = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2$$

$$\Longrightarrow \frac{(n-1) S_n^2}{\sigma^2} + \underbrace{\left(\frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}} \right)^2}_{\sim N(0,1)} = \sum_{i=1}^{n} \underbrace{\left(\frac{X_i - \mu}{\sigma} \right)^2}_{\sim N(0,1)}$$

$$\Longrightarrow \frac{(n-1) S_n^2}{\sigma^2} + Z_0^2 = \sum_{i=1}^{n} Z_i^2 \quad , \quad Z_i \sim N(0,1)$$

$$Z_0^2 \sim \chi^2(1) = \text{Gamma}(1/2, 1/2)$$

$$Z_i^2 \sim \chi^2(1) = \text{Gamma}(1/2, 1/2)$$

$$\sum Z_i^2 \sim \chi^2(n) = \text{Gamma}(n/2, 1/2)$$

$$\Longrightarrow \frac{(n-1) S_n^2}{\sigma^2} \sim \text{Gamma}\left(\frac{n-1}{2}, 1/2 \right) = \chi^2(n-1)$$

- **t - Student** $\quad X \sim t(n)$

$$X = \frac{Z}{\sqrt{\frac{Q_n}{n}}} \quad , \quad Z \sim N(0,1) \ \wedge \ Q_n \sim \chi^2(n)$$

$$\boxed{f_X(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\ \Gamma(n/2)} \frac{1}{\left(1 + \frac{x^2}{n}\right)^{\frac{n+1}{2}}}}$$

$$n \to +\infty \Longrightarrow X \sim t(n) \approx N(0,1)$$

Perché si usa?

$$\frac{\bar{X}_n - \mu}{S_n / \sqrt{n}} = \frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}} \cdot \frac{\sigma}{S_n} = \overset{\sim N(0,1)}{Z} \cdot \frac{1}{\sqrt{\frac{S_n^2}{\sigma^2}}} \cdot \frac{\sqrt{n-1}}{\sqrt{n-1}} =$$

$$= Z \cdot \frac{\sqrt{n-1}}{\sqrt{\underbrace{\frac{(n-1) S_n^2}{\sigma^2}}_{\sim \chi^2(n-1)}}} \overset{\frac{(n-1)S_n^2}{\sigma^2} = Q_{n-1}}{=\!=\!=} \frac{Z}{\sqrt{\frac{Q_{n-1}}{n-1}}} \sim t(n-1)$$

INTERVALLO DI CONFIDENZA UNILATERALE

$$\boxed{\begin{array}{l} \beta = \mathbb{P}(\{U_n \le h(\theta)\}) \\ \beta = \mathbb{P}(\{h(\theta) \le V_n\}) \end{array}}$$

lim. inf. di confidenza U_n

lim. sup. di confidenza V_n

Nota: In questo caso viene preso z_α e non $z_{\alpha/2}$

TEST DI IPOTESI

Si parte da un'ipotesi nulla $H_0 : h(\theta) \in E_0$
e si verifica che l'ipotesi alternativa $H_1 : h(\theta) \in E_1$
con $E_0 \cap E_1 = \emptyset$ è rifiutabile o accettabile secondo
il livello di significatività $\alpha \in (0,1)$

Errori

Hp. vera / Hp. accettata	H_0	$\bar{H}_0$
H_0	✓	II
$\bar{H}_0$	I	✓

I tipo: $\overbrace{\mu = 3}^{H_0}$ ma pensavo che $\mu > 3$ $\}$ H_1

II tipo: $\mu > 3$ $\}$ H_1 ma pensavo che $\mu = 3$ $\}$ H_0

Il livello di significatività α è la prob. di commettere un errore del I tipo.

$$\boxed{\begin{array}{l} R_c = \{(x_1 \dots x_n) \in R(X_1 \dots X_n) \text{ t.c. } \bar{x}_n > \mu_0 + \delta\} \\ \alpha = \mathbb{P}(\{(X_1 \dots X_n) \in R_c\}) \end{array}}$$

assumendo vera $\mu = \mu_0$

$$= \mathbb{P}(\{\bar{X}_n > \mu_0 + \delta\}) = \mathbb{P}\left(\left\{\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} > \underbrace{\frac{\mu_0 - \mu}{\sigma/\sqrt{n}}}_{= 0 \ (\mu = \mu_0)} + \frac{\delta}{\sigma/\sqrt{n}}\right\}\right)$$

$$\Longrightarrow \alpha = \mathbb{P}\left(\left\{\underset{\sim N(0,1)}{Z} > \frac{\delta}{\sigma/\sqrt{n}}\right\}\right)$$

$$\frac{\delta}{\sigma/\sqrt{n}} = z_\alpha \Longrightarrow \delta = \frac{\sigma}{\sqrt{n}} z_\alpha$$

Se $\bar{x}_n > \mu_0 + \delta = \mu_0 + \frac{\sigma}{\sqrt{n}} z_\alpha \Longrightarrow H_0$ rifiutata

Oss Se H_0 è rifiutata con liv. di significatività α'

$\Rightarrow$ H_0 è rifiutata con $\alpha > \alpha'$

P-VALUE

Il p-value è il più piccolo livello di significatività α per cui i dati permettono di rifiutare l'ipotesi nulla H_0.

$\forall \; \alpha > p\text{-value} \; : \; H_0$ è rifiutata

$\forall \; \alpha < p\text{-value} \; : \; H_0$ è accettata

$$\boxed{p\text{-value} = \inf \left\{ \alpha \in (0,1) \ni' \ \bar{x}_n > \mu_0 + \frac{\sigma}{\sqrt{n}} z_\alpha \right\}} =$$

$$= \inf \left\{ \alpha \in (0,1) \ni' \ \underbrace{\frac{\bar{x}_n - \mu_0}{\sigma/\sqrt{n}}}_{\text{Z-SCORE}} > z_\alpha \right\} \Longrightarrow$$

$$\Longrightarrow p\text{-value} = \inf_{\alpha \in (0,1)} \left\{ \phi(z_\alpha) < \phi\left(\frac{\bar{x}_n - \mu_0}{\sigma/\sqrt{n}} \right) \right\} =$$

$$= \inf_{\alpha \in (0,1)} \left\{ \phi(z_\alpha) < \phi\left(\frac{\bar{x}_n - \mu_0}{\sigma/\sqrt{n}} \right) \right\} =$$

$$= \inf_{\alpha \in (0,1)} \left\{ 1 - \alpha < \phi\left(\frac{\bar{x}_n - \mu_0}{\sigma/\sqrt{n}} \right) \right\} =$$

$$= \inf_{\alpha \in (0,1)} \left\{ \alpha > 1 - \phi\left(\frac{\bar{x}_n - \mu_0}{\sigma/\sqrt{n}} \right) \right\} = \overbrace{1 - \phi\left(\frac{\bar{x}_n - \mu_0}{\sigma/\sqrt{n}} \right)}^{\phi(-\overset{\parallel z}{\text{z-value}})}$$