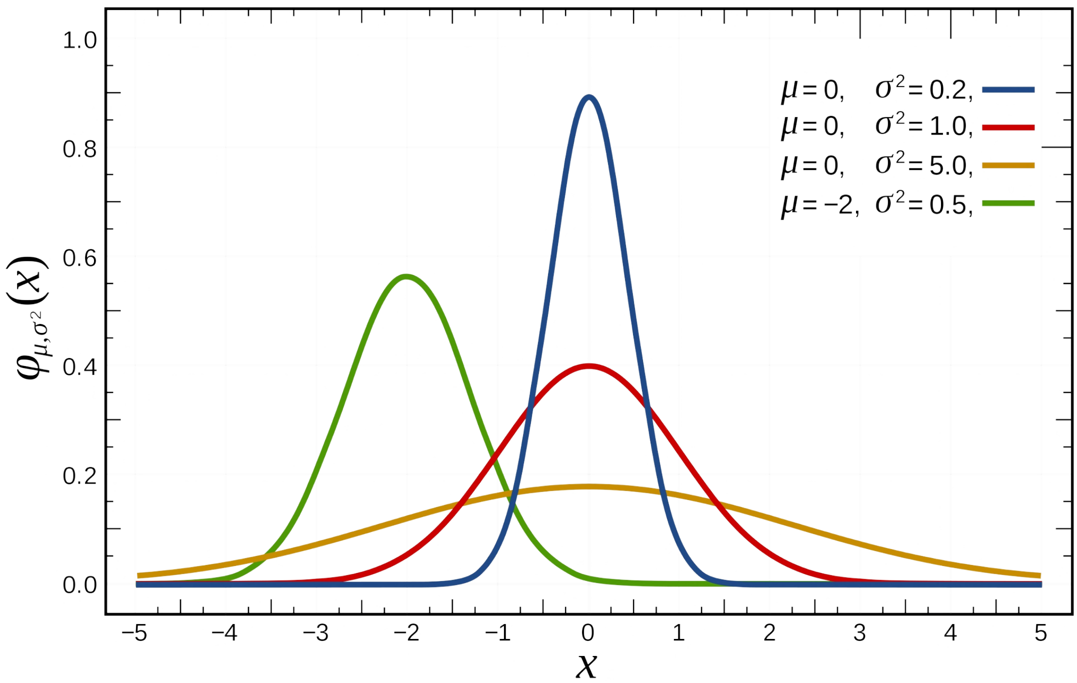
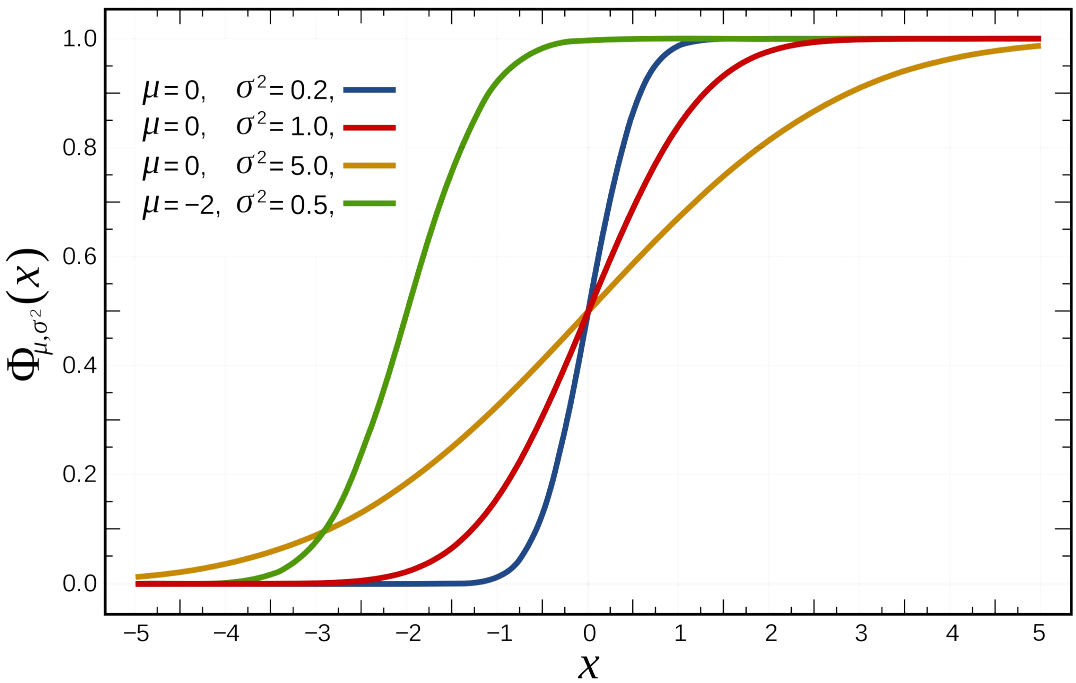


$\mu = 0, \quad \sigma^2 = 0.2,$
$\mu = 0, \quad \sigma^2 = 1.0,$
$\mu = 0, \quad \sigma^2 = 5.0,$
$\mu = -2, \quad \sigma^2 = 0.5,$
$\varphi_{\mu,\sigma^2}(x)$
x



$\mu = 0, \quad \sigma^2 = 0.2,$
$\mu = 0, \quad \sigma^2 = 1.0,$
$\mu = 0, \quad \sigma^2 = 5.0,$
$\mu = -2, \quad \sigma^2 = 0.5,$
$\Phi_{\mu,\sigma^2}(x)$
x

STATISTICA

Statistica Descrittiva

- CONCETTI BASE

popolazione e campione } Esempio: P = italiani che votano; C = sondaggio elettorale

fornisce stime sulla popolazione → inferenza statistica non fornisce mai verità

frequenza assoluta → insieme dei dati

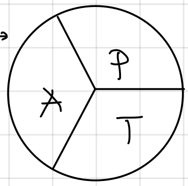
frequenza relativa → % dato

Esempio: elezioni piccolo comune: votanti 6342

candidato A 1234 → frequenza assoluta

$\frac{1234}{6342} = 0.194$ (= 19.4%): frequenza relativa

diagrammi: torta →

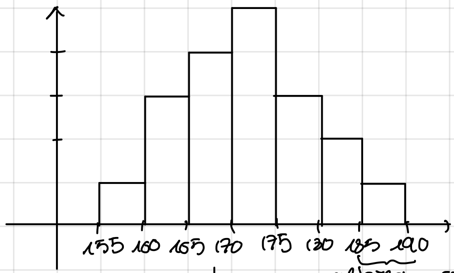


ove P, T, A ad esempio sono: P = matricole di Pisa, T = toscana fuori Pisa, A = altro

↓ regola generale

Frequenza proporzionale ad area

istogrammi (più preciso) → A canne o bastoni



altezza tra quanti valori compresa

↓

Istogramma normale: ha una punta e sale / scende più o meno allo stesso modo

↓ Altrimenti

Left skewed: scende più bruscamente

Right skewed: sale più bruscamente

Bimodale: abbiamo 2 punte

- ANALISI di DATI NUMERICI

Insieme di misurazioni diverse: $x = (x_1, \dots, x_n) \longrightarrow$ vettore di dati

↓

Media (empirica o campionaria): $\frac{x_1 + \dots + x_n}{n} = \bar{x}$

Varianza → Campionaria: $var(x) = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}$: più adatta per campione

→ Empirica: $Var_e(x) = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n}$: più adatta per popolazione

↓

Misura la dispersione di dati: la media può essere uguale per due gruppi diversi di dati ma allo stesso tempo la varianza può essere diversa se i dati sono più distanti

↓ Caso limite

0: $x_1 = x_2 = \dots = x_n$

Uguaglianza di base: $\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$

↓ Dimostrazione

$\sum_{i=1}^{n} (x_i^2 - 2x_i\bar{x} + \bar{x}^2) = \sum_{i=1}^{n} x_i^2 + n\bar{x}^2 - 2\bar{x}\underbrace{\sum_{i=1}^{n} x_i}_{n\bar{x}} \longrightarrow$ svolgendo i calcoli otteniamo l'uguaglianza

Scarto quadratico medio (deviazione standard): $\sigma(x) = \sqrt{\text{var}(x)}$: scarto quadratico medio campionario

$\sigma_e(x) = \sqrt{\text{var}_e(x)}$: scarto quadratico medio empirico

Diseguaglianza di Chebyshev: $\dfrac{\#\{x_i : |x_i - \bar{x}| > d\}}{n} \leq \dfrac{\text{var}_e(x)}{d^2}$ ove $d > 0$

↓ percentuale numeri tali che $|x_i - \bar{x}| > d$

↓ Dimostrazione

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 \geq \sum_{\{x_i : |x_i - \bar{x}| > d\}} (x_i - \bar{x})^2 \geq \sum_{\{x_i : |x_i - \bar{x}| > d\}} d^2 = d^2 \cdot \#\{x_i : |x_i - \bar{x}| > d\}$$ e poi basta dividere per n

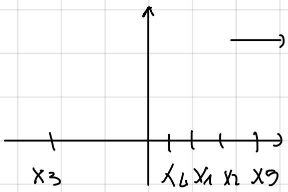
REGOLA EMPIRICA sulla CONCENTRAZIONE dei DATI

Supponiamo di avere $X = (x_1, \dots, x_n)$ e che l'istogramma sia abbastanza normale
- → Circa il 68% dei dati cade nell'intervallo $[\bar{x} - \sigma(x), \bar{x} + \sigma(x)]$
- → Circa il 95% cade in $[\bar{x} - 2\sigma(x), \bar{x} + 2\sigma(x)]$
- → Circa il 99,7% cade in $[\bar{x} - 3\sigma(x), \bar{x} + 3\sigma(x)]$

- **FUNZIONE di RIPARTIZIONE EMPIRICA**

$F_e(\cdot) : \mathbb{R} \longrightarrow [0, 1]$

in pratica ho $x_1, \dots, x_n$



→ Si disegnano su $\mathbb{R}$ e si prende una funzione che fa un salto di $\frac{1}{n}$ in corrispondenza dei dati

proprietà: è debolmente crescente, quindi se $s < t$, $F_e(s) \leq F_e(t)$

$\lim_{t \to -\infty} F_e(t) = 0$ e $\lim_{t \to +\infty} F_e(t) = 1$

Se ci sono due dati uguali il salto si ripete

↓

⚠ Conoscere la F_e equivale a conoscere i numeri ma non il loro ordine

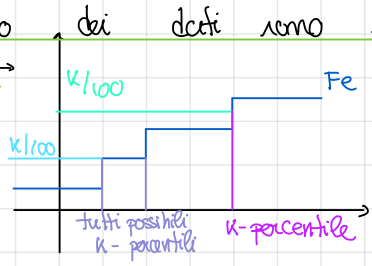
DEFINIZIONE: $\boxed{F_e(t) = \dfrac{\#\{x_i : x_i \leq t\}}{n}}$

- **PERCENTILI e QUANTILI**

K-percentile [$0 < k < 100$, di solito intero] → $X = (x_1, \dots, x_n)$, intuitivamente $k/100$ dei dati sono inferiori a tale percentile

↓ Definizione

Numero t tale che almeno $k/100$ dei dati sono inferiori a t e almeno $1 - k/100$ dei dati sono superiori a t



In statistica si prende $\beta = \frac{k}{100}$ ove $0 < \beta < 1$. Il k-esimo percentile si indicherà quindi come β-quantile

primo quartile: 0.25 quantile

terzo quartile: 0.75 quantile

mediana: 0.50 quantile ⟶ ⚠ Mediana ≠ Media

∘ DATI MULTIPLI (Bivariati)

$(x,y) = ((x_1,y_1),\dots,(x_n,y_n))$

Covarianza campionaria: $cov(x,y) = \sum_{i=1}^{n} \frac{(x_i-\bar{x})(y_i-\bar{y})}{n-1}$

Dall'uguaglianza precedente

covarianza empirica: $cov_e(x,y) = \sum_{i=1}^{n} \frac{(x_i-\bar{x})(y_i-\bar{y})}{n} = \boxed{\sum_{i=1}^{n} \frac{x_i y_i}{n} - \bar{x}\bar{y}}$

coefficiente di correlazione: $r(x,y) = \frac{cov(x,y)}{\sigma(x)\cdot\sigma(y)}$ se $\sigma(x) \wedge \sigma(y) \neq 0$

proprietà

$$= \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2}\cdot\sqrt{\sum(y_i-\bar{y})^2}}$$

$|r(x,y)| \leq 1$ per Schwartz $\left(\left|\sum_{i=1}^{n} a_i b_i\right| \leq \sum_{i=1}^{n}|a_i b_i| \leq \sqrt{\sum_{i=1}^{n} a_i^2}\sqrt{\sum_{i=1}^{n} b_i^2}\right)$

↓ Dimostrazione

$\sum_{i=1}^{n}|x_i-\bar{x}||y_i-\bar{y}| \leq \sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})}$ da cui segue la proprietà

Indica il legame tra i dati ⟶ RETTA di REGRESSIONE

Cerchiamo la retta che minimizza queste distanze al quadrato

↓

$\min_{a,b} \sum (y_i - a - bx_i)^2$

Risultato: la retta di regressione esiste ed è la retta $y = a^* + b^* x$ ove $b^* = \frac{cov(x,y)}{var(x)}$ e $a^* = -b^*\bar{x} + \bar{y}$

↓

Si ha $\min_{a,b} \sum_{i=1}^{n}(y_i - a - bx_i)^2 = \sum_{i=1}^{n}(y_i - a^* - b^* x_i)^2 = \sum_{i=1}^{n}|y_i-\bar{y}|^2(1 - r(x,y)^2)$ e

quindi se $|r(x,y)| = 1$ tutto fa 0 ⟶ punti allineati

$|r(x,y)| \sim 1$ i punti sono abbastanza allineati

$|r(x,y)| \sim 0$ i punti sono dispersi

DIMOSTRAZIONE DISUGUAGLIANZE (facoltativo)

Esempio: $f: \mathbb{R} \longrightarrow [0,+\infty[$ C^1 e supponiamo $\lim_{x\to-\infty} f(x) = \lim_{x\to+\infty} f(x) = +\infty$. Cerchiamo il minimo (che esiste sicuramente)

↓

se $f'(x) = 0$ solo per $x = x_0$, x_0 è il punto di minimo

Chiamiamo $Q(a,b) = \sum_{i=1}^{n}(y_i - a - bx_i)^2 \longrightarrow \lim_{|a|,|b|\to+\infty} Q(a,b) = +\infty$. Per cercare il minimo si annullano le derivate parziali

↓

Entrambe si annullano solo in a^* e b^*. Quindi $\min_{a,b\in\mathbb{R}^2} Q(a,b) = Q(a^*,b^*)$ da cui si arriva a ciò che volevamo dimostrare

Richiami di calcolo combinatorio

- NOTAZIONE
  un insieme di n elementi si indica con $\{1, 2, \dots, n\}$
  ↓
  Applicazioni di $\{1, \dots, n\}$ in $\{1, \dots, n\}$ $= k^n$
- PERMUTAZIONE
  In quanti modi si possono ordinare n elementi: $n!$
- COEFFICIENTI BINOMIALI
  Quanti sono i sottoinsiemi di k elementi da un insieme di n elementi $(0 \le k \le n)$.
  Sono $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. Si usa nel $\underbrace{\text{binomio di Newton}}$

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

sottinsiemi ordinati: $\underbrace{\binom{n}{k}}_{\text{senza ordine}} \cdot \underbrace{k!}_{\text{ordinamenti possibili}} = \frac{n!}{(n-k)!}$

sottinsiemi di stessa cardinalità: $k_1, \dots, k_j$ con $k_1 + \dots + k_j = n$. In quanti modi posso scegliere j sottinsiemi di cardinalità $k_1, \dots k_j$?
↓

$$\frac{n!}{k_1! \dots k_j!} = \binom{n}{k_1}\binom{n-1}{k_2} \dots \binom{n-k_1-k_{j-1}}{k_j}$$

- PROBABILITÀ COMBINATORIA

Definizione: $\frac{\#\text{casi favorevoli}}{\#\text{casi possibili}}$ ⚠ caso particolare

Esempio: Estrazione lotto a Firenze sabato prossimo. Qual è la probabilità che i 5 numeri siano tutti minori di 20
↓

$$\frac{\binom{20}{5}}{\binom{90}{5}} = \frac{20 \cdot 19 \cdot 18 \cdot 17 \dots}{90 \cdot 89 \cdot 88 \cdot 87 \dots}$$

Probabilità che 3 numeri estratti siano inferiori a 20

a. Esattamente 3: $\frac{\binom{20}{3}\binom{70}{2}}{\binom{90}{5}}$

b. Almeno 3: $\frac{\binom{20}{3}\binom{70}{2} + \binom{20}{4} 70 + \binom{20}{5}}{\binom{90}{5}}$

Si scelgono 3 cifre tra 0 e 9. Qual è la probabilità che siano tutte diverse: $\frac{10 \cdot 9 \cdot 8}{10^3} = \frac{9 \cdot 2}{5^2} = \frac{18}{25}$

Mazzo di poker 52 carte. In una presa di poker, qual è la probabilità di avere 2 assi?

a. Esattamente 2: $\frac{\binom{4}{2}\binom{48}{3}}{\binom{52}{5}}$

b. Almeno 2: $\frac{\binom{4}{2}\binom{48}{3} + \binom{4}{3}\binom{48}{2} + \binom{4}{4} \cdot \binom{48}{1}}{\binom{52}{5}}$

Probabilità

- CONCETTI PRIMITIVI

Evento: Avvenimento → Evento stocastico (casuale)

↓

1. Come si rappresentano gli eventi?

Affermazioni legati da connettivi logici "e" "o" "non": Ω (spazio fondamentale degli eventi) ⟶ tutte le possibili realizzazioni, eventi sottoinsiemi di Ω

↓

Esempio: Lanciamo un dado. $\Omega = \{1, 2, \dots, 6\}$

"numero pari" $A = \{2, 4, 6\}$ ⟶ "non pari": A^C

"numero > 3" $B = \{4, 5, 6\}$

↓

"pari o > 3" $= A \cup B$

"pari e > 3" $= A \cap B$

↓ Quindi

"o": $\cup$, "e": $\cap$, "non" = complemento

Decidiamo che $\emptyset$, Ω sono eventi (si indicano con $\mathcal{A}$)

Se A è un evento, anche A^c è un evento ⟶ Se $A \in \mathcal{A} \rightarrow A^c \in \mathcal{A}$

Se $A, B \in \mathcal{A} \rightarrow (A \cup B) \in \mathcal{A}$ e di conseguenza $(A \cap B) \in \mathcal{A}$. Questo perché $(A \cap B)^c = A^c \cup B^c$

↓ Avendo queste proprietà

Algebra di parti di Ω

PROBABILITÀ: $P: \mathcal{A} \longrightarrow [0, 1]$

2. Quali proprietà deve avere?
   a. $P(\Omega) = 1$
   b. se $A \cap B = \emptyset \longrightarrow P(A \cup B) = P(A) + P(B)$ (funzione additiva di insiemi)

      ↓ Conseguente

      $P(A^c) = 1 - P(A)$ in quanto $\Omega = A \cup A^c$ e $A \cap A^c = \emptyset \longrightarrow 1 = P(\Omega) = P(A) + P(A^c)$

      $\emptyset = \Omega^c$ e quindi $P(\emptyset) = 1 - P(\Omega) = 0$

      Se $P(A) = 0$, A è detto **trascurabile**

      Se $P(A) = 1$, A è detto **quasi certo**
   c. se $B \subseteq A$, $P(A/B) = P(A) - P(B)$

      $A/B = A \cap B^c$

      ↓

      $A = B \cup (A/B) \longrightarrow P(A) = P(B) + P(A/B)$
   d. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

      ↓

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

      $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) - P(A \cap B \cap C)$

Osservazione: A volte passando al complementare il problema diventa più semplice

Esempio: Qual è la probabilità che lanciando 5 volte un dado esca almeno una volta il numero 6?

↓

A^c = il numero 6 non esce mai

$P(A^c) = \dfrac{5^5}{6^5} = \left(\dfrac{5}{6}\right)^5$ e quindi $P(A) = 1 - \left(\dfrac{5}{6}\right)^5$

⚠ Le proprietà non permettono di passare al limite

↓

Passiamo da additività ad additività numerabile: $A_1, A_2, \dots$ con $A_i \cap A_j = \emptyset$

Allora $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) = \lim_{n \to +\infty} \left(\sum_{k=1}^{n} P(A_k) \right)$ ⟶ Completa additività

NUOVE DEFINIZIONI: $(\Omega, \mathcal{A}, P)$ = spazio di probabilità

Ω = insieme

$\mathcal{A}$ è una σ-algebra di parti:
1. $\emptyset \in \mathcal{A}$, $\Omega \in \mathcal{A}$
2. $A \in \mathcal{A} \to A \in \mathcal{A}^c$
3. Se $\forall n$, $A_n \in \mathcal{A} \to (\bigcup_{n=1}^{\infty} A_n) \in \mathcal{A}$

↓ Conseguenza

$(\bigcap_{n=1}^{\infty} A_n) \in \mathcal{A}$

PROBABILITÀ: $P: \mathcal{A} \to [0,1]$ con proprietà:
1. $P(\Omega) = 1$
2. se $A_i \cap A_j = \emptyset$, $P(\bigcup_{n=1}^{+\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$

Perché non sono eventi tutti i sottinsiemi? Se Ω è finito o numerabile non ci sono problemi. Se Ω non è numerabile possono esserci problemi e quindi in certe situazioni non è possibile definire P su tutti i sottinsiemi

Ma cosa significa **andare al limite**? $A_1 \subseteq A_2 \subseteq A_3 \dots$

↓ Definiamo

$\lim_{n\to+\infty} A_n = \bigcup_{n=1}^{\infty} A_n$

↓ **Da cui**

$P(\lim_{n\to+\infty} A_n) = \lim_{n\to+\infty} P(A_n)$

Se $A = \bigcup_{n=1}^{\infty} A_n = A_1 \cup \underbrace{(A_2 \setminus A_1) \cup (A_3 \setminus A_2) \dots}_{\text{insiemi disgiunti}}$

↓

$P(A) = \lim_{n\to+\infty} (P(A_1) + P(A_2 \setminus A_1) + \dots + P(A_n \setminus A_{n-1})) = \lim_{n\to+\infty} (P(\cancel{A_1}) + P(\cancel{A_2}) - P(\cancel{A_1}) + \dots + P(A_n) - \cancel{P(A_{n-1})})$

$= \lim_{n\to+\infty} P(A_n)$ (sottointendiamo che se $A \subseteq B$, $P(A) \leq P(B)$)

Se invece $B_1 \supseteq B_2 \supseteq B_3 \dots \to \bigcap_{n=1}^{\infty} B_n = \lim_{n\to+\infty} B_n$

↓

$P(\lim_{n\to+\infty} B_n) = \lim_{n\to+\infty} P(B_n)$

Si può dimostrare che P è σ-additiva $\Longleftrightarrow$ P è additiva
P va al limite

Casi particolari: $\boxed{\Omega \text{ finito}} \to \Omega = \{w_1, \dots, w_n\}$, eventi tutti i sottinsiemi

$p_i = P(\{w_i\}) = P(w_i)$ ove $p_i \geq 0$ e $p_1 + \dots + p_n = 1$

↓

$P(A) = \sum_{w_i \in A} P(w_i) = \sum_{w_i \in A} p_i \to 1 = P(\Omega) = \sum_{i=1}^{n} p_i$

Tutti i punti sono equiprobabili $\to p_i = p_j$ e quindi $p_i = 1/n$

↓

$\boxed{P(A) = \frac{\# A}{\# \Omega} \longrightarrow \text{Probabilità uniforme su } \Omega}$

$\boxed{\Omega \text{ numerabile}} \to \Omega = \{w_1, w_2, \dots\}$ $\mathcal{A}$ tutti i sottinsiemi. Basta conoscere $p_i = P(w_i)$, $\sum_{i=1}^{\infty} p_i = 1$.

↓

$P(A) = \sum_{w_i \in A} p_i$ che può essere finita o infinita

↓

Se Ω è infinito non ha senso la probabilità uniforme: avremo, se $p_i \geq 0$, che $P(\Omega) = +\infty$, mentre se $p = 0$, $P(\Omega) = 0$

ESERCIZI:

1. n persone. Qual è la probabilità che almeno 2 siano nati lo stesso giorno?

Facciamo il complementare: $P(A^c)$ = "tutti compleanni diversi": $\frac{\binom{365}{n} \cdot n!}{365^n}$ = voglio sottoins. ordinati

$$= \frac{365!}{(365-n)!\ 365^n}$$

$$\downarrow$$

$$P(A) = 1 - \frac{365!}{(365-n)! \cdot 365^n}$$

Per $n = 23 \longrightarrow P(A) = 1/2$

Per $n = 40 \longrightarrow P(A) = 0.9$

⚠ Due serie da ricordare

1. **Serie geometrica**: $|a| < 1 \longrightarrow \sum_{n=0}^{\infty} a^n = \frac{1}{1-a}$

↓ Dim.

$$1 + a + a^2 + \dots + a^n = \frac{a^{n+1}-1}{a-1} \longrightarrow \sum_{n=0}^{+\infty} a^n = \lim_{n\to+\infty}(1+\dots+a^n) = \lim_{n\to+\infty} \frac{a^{n+1}-1}{a-1} = \frac{-1}{a-1}$$

2. **Serie esponenziale**: $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$

- PARADOSSO di MONTY-HALL

A: "scelto porta giusta" $\longrightarrow P(A) = 1/3$

B: "scelto porta sbagliata" $\longrightarrow P(B) = 2/3$

↓

Conservo porta: P(vincita auto) = 1/3

Cambio porta: P(vincita auto) = 2/3 (il presentatore mi ha aperto una porta sbagliata)

- PROBABILITÀ CONDIZIONATA

Esempio: Ho puntato il 18 alla roulette

$A = \{18\}$ $\qquad P(A) = 1/37$

È uscito un numero pari, B = "numero pari"

↓

$$P(A|B) = \frac{1}{18} = \frac{1/37}{18/37} = \frac{P(A\cap B)}{P(B)}$$

DEFINIZIONE: $\boxed{P(A|B) = \frac{P(A\cap B)}{P(B)}}$ ove B non deve essere non trascurabile

$A \xrightarrow{\text{funzione}} P(A|B)$ è una probabilità: 1. $P(\Omega|B) = \frac{P(\Omega\cap B)}{P(B)} = 1$

2. $P(A^c|B) = \frac{P(A^c\cap B)}{P(B)} = \frac{P(B-(A\cap B))}{P(B)} =$

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

$P(A_1 \cup A_2 | B) = P(A_1|B) + P(A_2|B)$

Formule: $P(A\cap B) = P(A|B)\cdot P(B) = P(B|A)\cdot P(A)$ [A, B non trascurabile]

$\boxed{P(B|A) = \frac{P(A|B)\cdot P(B)}{P(A)}}$ formula di Bayes

Condizionamento ripetuto: $P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|(A_1\cap A_2)) \cdot$
$\cdot\, P(A_n | (A_1 \cap \dots \cap A_{n-1}))$

↓

Esempio: estrazione di 3 carte da poker. Qual è la possibilità che siano 3 cuori?

soluzione 1: combinatoria

$$P = \frac{\binom{12}{3}}{\binom{52}{3}} = \frac{12 \cdot 11 \cdot 10}{52 \cdot 51 \cdot 50}$$

Soluzione 2: condizionamento ripetuto

$A_1, A_2, A_3 \longrightarrow$ uscito un cuore al 1°, 2°, 3°

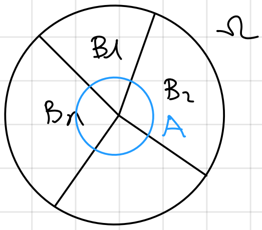
$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2 | A_1) \cdot P(A_3 | A_1 \cap A_2) =$$
$$= \frac{12}{52} \cdot \frac{11}{51} \cdot \frac{10}{50}$$

Problema dei compleanni

Complementare: n persone, compleanni tutti diversi $(A_1, A_2, \dots, A_n)$

$$P(A_1 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2 | A_1) \cdot P(A_n | A_1 \cap \dots \cap A_{n-1}) =$$
$$= \frac{365}{365} \cdot \frac{365-1}{365} \cdot \dots \cdot \frac{365-n}{365} = \frac{\binom{365}{n} \cdot n!}{365^n}$$

Fattorizzazione:



$B_1, \dots, B_n$ partizioni di Ω possibili alternative

$\downarrow$

$$P(A) = \sum_{i=1}^{n} (P(A | B_i) \cdot P(B_i))$$

Dim $\downarrow$

$$A = \bigcup_{i=1}^{n} (A \cap B_i) \longrightarrow P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A | B_i) \cdot P(B_i)$$

$\boxed{\text{Probabilità delle cause (Bayes)}}$:

$$P(\underbrace{B_i}_{\text{cause}} | \overbrace{A}^{\text{effetto}}) = \frac{P(A | B_i) \cdot P(B_i)}{P(A)} = \frac{P(A | B_i) \cdot P(B_i)}{\sum_{j=1}^{n} P(A | B_j) \cdot P(B_j)}$$

$\downarrow$

Esempi: 1. compagnia assicurazione incidenti di cui i clienti sono:

30% giovani (18-30) $\longrightarrow$ Prob. incidenti 0.07

40% adulti (31-65) $\longrightarrow$ Prob. incidenti 0.02

25% anziani (> 65) $\longrightarrow$ Prob. incidenti 0.12

1. Qual è la probabilità che un assicurato abbia un incidente?

2. Arriva la denuncia di un incidente. Qual è la probabilità che arrivi da un giovane?

$\downarrow$

"Cause" B_1, B_2, B_3.
$P(B_1) = 0.3$
$P(B_2) = 0.45$
$P(B_3) = 0.25$

A = "incidente". $P(A | B_1) = 0.07$ $P(A | B_2) = 0.02$ $P(A | B_3) = 0.12$

1. $P(A) = P(A | B_1) \cdot P(B_1) + P(A | B_2) \cdot P(B_2) + P(A | B_3) \cdot P(B_3) =$
$= 0.07 \cdot 0.3 + 0.02 \cdot 0.45 + 0.12 \cdot 0.25$

2. $P(B_1 | A) = \dfrac{P(A | B_1) \cdot P(B_1)}{P(A)} = \dfrac{0.07 \cdot 0.3}{}$

2. Problema di Monty-Hall

B_1 = "porta giusta" $P(B_1) = 1/3$

$B_2 = B_1^c$ $P(B_2) = 2/3$

A = "viene vinta l'auto"

$\downarrow$

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

Strategia conservativa

$$P(A) = 1 \cdot \frac{1}{3} + 0 \cdot \frac{2}{3} = \frac{1}{3}$$

strategia innovativa

$P(A) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}$

- PROBABILITÀ di INDIPENDENZA di EVENTI (STOCASTICA)

$(\Omega, \mathcal{F}, P)$. La conoscenza di B non influenza la probabilità di A e vice-versa.

$P(A|B) = P(A) \iff P(B|A) = P(B)$ che sono equivalenti se

$P(A \cap B) = P(A) \cdot P(B)$ che ha senso anche se A o B sono trascurabili

DEFINIZIONE di INDIPENDENZA STOCASTICA: A, B indipendenti se $P(A \cap B) = P(A) \cdot P(B)$

↓ Conseguenze

1. A, B indipendenti allora A^c e B indipendenti
2. Se $P(A) = 0 \vee P(A) = 1$, A è indipendente da ogni altro B
3. Se A e B sono incompatibili ($A \cap B = \emptyset$) non possono essere indipendenti almeno che uno dei due non sia trascurabile.

↓ Dim

1. $P(A^c \cap B) = P(B) - P(A \cap B) = P(B) - P(A) \cdot P(B) = P(B) \cdot (1 - P(A)) = P(B) \cdot P(A^c)$
2. $P(A) = 0 \longrightarrow \underset{0}{0 \leq P(A \cap B)} \leq \underset{0}{P(A)} \longrightarrow \underset{0}{P(A \cap B)} = \underset{0}{P(A) \cdot P(B)}$

3. A, B, C indipendenti

$P(A \cap B) = P(A) \cdot P(B)$, $P(A \cap C) = P(A) \cdot P(C)$, $P(B \cap C) = P(B) \cdot P(C)$,

$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$

↓ Esempio

$\Omega = \{1, 2, 3, 4\}$ prob. uniforme

$A = \{1, 2\}$, $B = \{1, 3\}$, $C = \{2, 3\}$

$P(A \cap B) = P(A) \cdot P(B)$ $\quad P(A \cap B) = \frac{1}{4}$ e $P(A) = \frac{1}{2}$, $P(B)\ \frac{1}{2}$

$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$? NO: $0 \neq \frac{1}{8}$

↓ Generalizziamo

$A_1, \dots, A_n$ sono indipendenti se $\forall$ scelta di $2 \leq k \leq n$ e $1 \leq i_1 < \dots < i_k \leq n$ si ha $P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot \dots \cdot P(A_{i_k})$

↓ Caso tipico

Prove ripetute nelle medesime condizioni (estrazioni del lotto in settimane successive)

ESERCIZI:

1. **Esercizio 1.2.** In un paese nel quale vi sono 5 alberghi arrivano 3 turisti, questi si sono prenotati in modo del tutto casuale.
a) Qual è la probabilità che finiscano tutti in uno stesso albergo?
b) Qual è la probabilità che finiscano in 3 alberghi differenti?

a. $\frac{5}{5^3} = \frac{1}{25}$ ma posso vederlo come $P(A_1 \cap A_2 \cap A_3) = 1 \cdot \frac{1}{5} \cdot \frac{1}{5} = \frac{1}{25}$

b. $\frac{\binom{5}{3} \cdot 3!}{5^3} = \frac{\cancel{5} \cdot 4 \cdot \cancel{3} \cdot 3 \cdot 2}{\cancel{3} \cdot 2 \cdot 5^{3-2}} = \frac{12}{25}$ oppure $P(B_1 \cap B_2 \cap B_3) = 1 \cdot \frac{4}{5} \cdot \frac{3}{5}$

voglio ordine

2. **Esercizio 1.8.** Supponiamo di avere due scatole, la prima con 5 palline rosse e 7 blu, la seconda con 8 palline rosse e 3 blu. Si sceglie a caso una scatola e si estraggono due palline che risultano essere entrambe rosse: qual è la probabilità che sia stata scelta la prima scatola?

B_1, B_2 ove $P(B_1) = P(B_2) = \frac{1}{2}$

$P(A|B_1) = \frac{\binom{5}{2}}{\binom{12}{2}} = \frac{5 \cdot \cancel{4}}{3 \cancel{12} \cdot 11} = \frac{5}{33}$

$P(A|B_2) = \frac{\binom{8}{2}}{\binom{11}{2}} = \frac{8 \cdot 7}{11 \cdot 10} = \frac{28}{55}$

$$P(B_1|A) = \frac{\frac{5}{33}\cdot\frac{1}{2}}{\frac{5}{33}\cdot\frac{1}{2} + \frac{22}{55}\cdot\frac{1}{2}} < \frac{1}{2}$$

$$P(B_2|A) = 1 - P(B_1|A) > \frac{1}{2}$$

3. variabile binomiale

$\boxed{\textbf{Esercizio 1.9.}}$ Ripetiamo 4 volte un esperimento che ha probabilità di successo eguale a 2/3: calcolare la probabilità che
   a) si abbia almeno un successo
   b) si abbia esattamente un successo.

b. i s i i
i i s i
⋮
4 possibilità

$$\longrightarrow \left(\frac{1}{3}\cdot\frac{2}{3}\cdot\frac{1}{3}\cdot\frac{1}{3}\right)\cdot 4 = 4\cdot\frac{2}{3^4}$$

$$\boxed{\binom{4}{1}\left(\frac{2}{3}\right)^1\cdot\left(1-\frac{2}{3}\right)^{4-1}}$$

n prove di un esperimento con probabilità p di successo
Qual è la probabilità di avere h $(0 \le h \le n)$ successi?

s i i s s ... → h volte, $n-h$ volte

$$\longrightarrow \boxed{p^h}\cdot\boxed{(1-p)^{n-h}}\cdot\boxed{\binom{n}{h}}$$

possibili sequenze di n prove con h successi

a. $A_0, A_1, A_2, A_3, A_4 \longrightarrow 0, 1, \dots$ successi

$$P(A_1) + P(A_2) + P(A_3) + P(A_4) = 1 - P(A_0) = 1 - (1-p)^4 = 1 - \frac{1}{3^4}$$

probabilità insuccesso

4. **Esercizio 1.3.** Il re (di una monarchia che ammette solo regnanti maschi) non è figlio unico: qual è la probabilità che abbia una sorella?

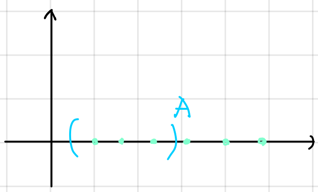
$B_1 = \{M, M\}$ $\quad B_2 = \{M, F \text{ o } F, M\}$ $\quad B_3 = \{F, F\}$.

$P(B_1) = P(B_3) = \frac{1}{4}$ $\qquad P(B_2) = \frac{1}{2}$

$$P(B_2 \mid B_1 \cup B_2) = \frac{P(B_2)}{P(B_1) + P(B_2)} = \frac{1/2}{1/4 + 1/2} = \frac{2}{3}$$

- PROBABILITÀ su $\mathbb{R}$

probabilità discreta: concentrata su un insieme finito o numerabile di punti
$(x_1, x_2, \dots, x_n)$ o $(x_1, x_2, x_3, \dots)$



$p(x_i) = P(\{x_i\}) = P(x_i)$, $\quad p(x_i) \ge 0$

$p(x_1) + p(x_2) + \dots = 1$

$P(A) = \sum_{x_i \in A} p(x_i)$

$$\boxed{p(x) = \begin{cases} p(x_i) & x = x_i \\ 0 & \text{altrimenti} \end{cases}} \longrightarrow \text{funzione di massa / densità discreta}$$

Conosco anche P

ESEMPIO: $p(x) = \begin{cases} \frac{1}{2^k} & k = 1, 2, 3, \dots \\ 0 & \text{altrove} \end{cases} \longrightarrow \sum p(x_i) = 1$ (serie geometrica)

$P([0,3]) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = p(1) + p(2) + p(3)$

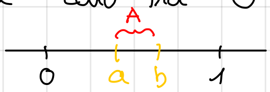
$P(]0,3[) = \frac{1}{2} + \frac{1}{4} = p(1) + p(2)$

$P(]-\infty, 0]) = 0$

probabilità definite da densità

ESEMPIO PRELIMINARE: Un numero a caso tra 0 e 1

$\Omega = [0,1]$



$$P([a,b]) = \frac{\text{lunghezza } [a,b]}{\text{lunghezza } [0,1]} = \frac{b-a}{1-0} = b-a$$

Mi aspetto di prolungarlo a tutti i sottinsiemi di $[0,1]$. ⚠ NON è possibile farlo almeno che l'insieme non sia misurabile

Differenza tra trascurabile e impossibile

trascurabile: probabilità 0; impossibile: $\emptyset$

$x \in [0,1]$ $\{x\}$ $A_n = [x-\frac{1}{n}, x+\frac{1}{n}] \longrightarrow 0 \leq P(x) \leq P(A_n) = \frac{2}{n} \quad \forall n$

$\frac{2}{n} \xrightarrow{n \to +\infty} 0$

e quindi $P(\{x\}) = 0$: trascurabile

Definizione (prob. definita da densità)

$$\boxed{f: \mathbb{R} \longrightarrow [0, +\infty[\text{ integrabile } / \int_{-\infty}^{+\infty} f(x)\,dx = 1}$$

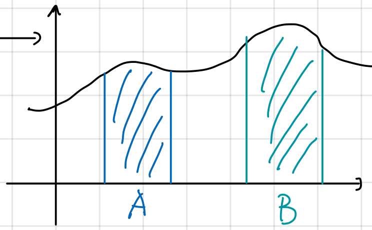
$$P(A) = \int_A f(x)\,dx$$

⚠ Solo su insiemi misurabili

Perché è una probabilità?

1. $P(\mathbb{R}) = \int_{\mathbb{R}} f(x)\,dx = 1$

2. $A \cap B = \emptyset \longrightarrow$



$\longrightarrow P(A \cup B) = \int_{A \cup B} f(x)dx = \int_A f(x)dx + \int_B f(x)dx$ e quindi additiva

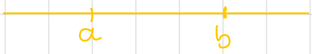
Ma è anche σ-additiva? Si dimostra che va anche al limite e quindi è anche σ-additiva

↓ Tornando all'esempio

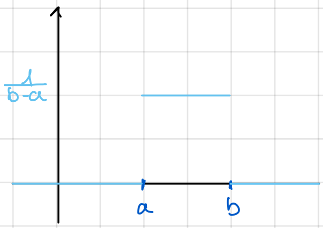
$P([a,b]) = b-a = \int_{[a,b]} 1\,dx \longrightarrow$ **Densità uniforme su $[0,1]$**: $\quad 0 \leq a < b \leq 1$

$$\boxed{f(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{altrove} \end{cases}}$$

Densità uniforme su $[a,b]$:



$$f(x) = \begin{cases} \text{costante su } [a,b] \\ 0 \quad \text{altrove} \end{cases}$$



$$\boxed{f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{altrove} \end{cases}}$$

Densità esponenziale (parametro $\lambda > 0$):

$$\boxed{f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x < 0 \end{cases}}$$

↓ verifica prop. 1

$$\int_{\mathbb{R}} f(x)\,dx = \int_0^{+\infty} \lambda e^{-\lambda x}\,dx = \lim_{M \to +\infty} \int_0^M \lambda e^{-\lambda x}\,dx =$$

$$= \lim_{M \to +\infty} -e^{-\lambda x}\Big|_0^M = \lim_{M \to +\infty} (1 - e^{-\lambda M}) = 1$$

ESEMPI: 1. $P([-2,2]) = \int_{-2}^{2} f(x)dx = \cancel{\int_{-2}^{0} 0\,dx} + \int_0^2 \lambda e^{-\lambda x}\,dx = 1 - e^{-2\lambda}$

2. $f(x) = \begin{cases} c(3-x) & 0 \leq x \leq 3 \\ 0 & \text{altrove} \end{cases}$

a. per quale c è una densità?
b. $P([1,4])$, $P([0,2])$

a. $f(x) \geq 0$ $\qquad \int_{\mathbb{R}} f(x)\,dx = 1$

$$\int_0^3 (3-x)\,dx = 9 - \frac{9}{2} = \frac{9}{2} \longrightarrow c = \frac{2}{9}$$

b. $P([1,4]) = \int_1^4 f(x)\,dx = \frac{2}{9}\int_1^3 (3-x)\,dx$

$P([0,2]) = \int_0^2 \frac{2}{9}(3-x)\,dx$

- VARIABILI ALEATORIE (CASUALI)

ESEMPIO: puntata 1€ sul 18 / puntata 1€ sul dispari / puntata 1€ sul pari / puntata 1€ sullo 0

Qual è la probabilità di perdere? E quella di vincere più di 10€?

$\Omega = \{0, 1, 2, \dots, 36\}$ probabilità uniforme

$X: \Omega \longrightarrow \mathbb{R}$ "vittoria netta" $\longrightarrow X(\omega) = \begin{cases} 34 & \omega = 18 \\ 32 & \omega = 0 \\ -2 & \omega \text{ pari},\ \omega \neq 18 \\ -2 & \omega = \text{dispari} \end{cases}$

Somme vinte — puntata corrispondente

Probabilità di perdere: $P\{\omega \in \Omega \mid X(\omega) < 0\} = P\{X^{-1}(]-\infty, 0[)\} = \frac{35}{37}$

Probabilità di vincere più di 10€: $P\{\omega \in \Omega \mid X(\omega) > 0\} = P\{X^{-1}(]10, +\infty[)\} = P\{0, 18\} = \frac{2}{37}$

Abbiamo introdotto una funzione $X: \Omega \to \mathbb{R}$ variabile casuale (aleatoria). Abbiamo inoltre trasportato la probabilità dai sottoinsiemi di Ω ai sottoinsiemi di $\mathbb{R}$ mediante la formula $P_X(A) = P(X^{-1}(A))$

variabile aleatoria: $X: \Omega \longrightarrow \mathbb{R}$ (su Ω è definita la probabilità P)

legge di probabilità della variabile X: $P_X(A) = P(X^{-1}(A))$

verifichiamo

1. $P_X(\mathbb{R}) = P(X^{-1}(\mathbb{R})) = P(\Omega) = 1$
2. $A_1, A_2, \dots$ sottoinsiemi di $\mathbb{R}$ / $A_i \cap A_j = \emptyset$

$$X^{-1}\left(\bigcup_{n=1}^{\infty} A_n\right) = \bigcup_{n=1}^{\infty} \underbrace{X^{-1}(A_n)}_{}$$

insiemi disgiunti $\longrightarrow P_X\left(\bigcup_{n=1}^{\infty} A_n\right) = P\left(X^{-1}\left(\bigcup_{n=1}^{\infty} A_n\right)\right) = P\left(\bigcup_{n=1}^{\infty} X^{-1}(A_n)\right) =$

$$= \sum_{n=1}^{\infty} P(X^{-1}(A_n)) = \sum_{n=1}^{\infty} P_X(A_n)$$

⚠ In molti esempi non compare $(\Omega, \mathcal{F}, P)$ e $X: \Omega \to \mathbb{R}$ ma solo P_X.

Si dimostra che, data una probabilità Q su $\mathbb{R}$, si possono costruire $(\Omega, \mathcal{F}, P)$ e $X: \Omega \to \mathbb{R}$ tali che $Q = P_X \longrightarrow Q(A) = P_X(A) = P(X^{-1}(A))$.
Anche le variabili aleatorie possono essere discrete, con densità e più generali corrispondenti alla loro legge di probabilità

Siano X, Y variabili aleatorie. Se $P_X = P_Y$, X e Y sono dette equidistribuite.
notazione: $X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\} = \{X \in A\}$

ESERCIZIO:

Esercizio 1.9. Ripetiamo 4 volte un esperimento che ha probabilità di successo eguale a 2/3: calcolare la probabilità che

a) si abbia almeno un successo

b) si abbia esattamente un successo.

$\boxed{\text{Variabile binomiale di parametri } n \text{ e } p}$: ripetiamo n volte un esperimento che ha p probabilità di successo. X conta il numero dei successi. I valori possibili sono $0, 1, \dots, n$. $\longrightarrow$ $\boxed{p(h) = P\{X=h\} = \binom{n}{h} p^h (1-p)^{n-h}}$: variabile aleatoria discreta

$\downarrow$ verifichiamo $p(0)+p(1)+\dots+p(n)=1$

$$\sum_{h=0}^{n} \binom{n}{h} p^h (1-p)^{n-h} = [p+(1-p)]^n = 1 \quad \checkmark$$

Ripetiamo l'esperimento fino al successo: X conta i tentativi che è stato necessario fare $\longrightarrow$ valori possibili: $1, 2, 3, \dots$

$$\underbrace{p(h)}_{\underbrace{1, 1, 1, \dots, 1}_{h}} = (1-p)^{h-1} \cdot p, \text{ ma } p(1)+p(2)+\dots = 1? \quad \sum_{h=1}^{\infty} p(1-p)^{h-1} = p \sum_{n=0}^{+\infty} (1-p)^n = \frac{p}{1-(1-p)} = 1$$

Può capitare che il successo non arrivi mai? $X = +\infty$ se non c'è mai successo

$$\Omega = \bigcup_{h=1}^{\infty} \{X=h\} \cup \{X=+\infty\}$$

$$1 = P(\Omega) = \sum_{h=1}^{\infty} P\{X=h\} + P\{X=+\infty\} = 1 + \cancel{P\{X=+\infty\}}$$

0 = non è possibile

$\boxed{\text{Variabili discrete}}$

X è discreta se prende un insieme finito o numerabile di valori $x_1, x_2, \dots$

Per conoscerla si devono sapere i valori (immagine di X) ed i numeri

$p_X(x_i) = p(x_i) = P\{X=x_i\} \longrightarrow p(x_i) \geq 0$ e $p(x_1)+p(x_2)+\dots+p(x_n) = 1$

$\downarrow$ $\qquad p(x_n)+\dots = \sum_{i=1}^{\infty} p(x_i) = 1$

$\boxed{\text{Binomiale di parametri } n \text{ e } p \quad B(n,p)}$: valori $h = 0, 1, \dots, n$ e $p(h) = P\{X=h\} = \binom{n}{h} p^h (1-p)^{n-h}$. Se $n=1$ è la variabile di $\boxed{\text{Bernoulli}}$ di parametro p e può avere i valori
$$\begin{cases} 0 & p(0) = 1-p \\ 1 & p(1) = P(X=1) = p \end{cases}$$

$\boxed{\text{Geometrica di parametro } p \ (0<p<1)}$: valori possibili $1, 2, \dots$; $p(h) = P\{X=h\} = (1-p)^{h-1} \cdot p$

$\boxed{\text{Variabili con densità}}$: $\exists$ una densità $f: \mathbb{R} \longrightarrow [0, +\infty[$ con $\int_{-\infty}^{+\infty} f(x)\,dx = 1$ tale che $P_X(A) = P\{X \in A\} = \int_A f(x)\,dx$. Per conoscerla mi basta conoscere la densità

$\downarrow$

densità uniforme su $[a,b]$: $f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{altrove} \end{cases}$

densità esponenziale di parametro $\lambda > 0$: $f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$

Riassunto densità

⚠ Densità alta = valori presi con probabilità più alta (area rettangolo). I valori della densità esponenziale sono sempre positivi

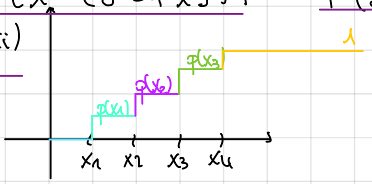
$\downarrow$ Utilizzi

esponenziale: tempo decadimento isotopi instabili

$\boxed{\text{Funzione di ripartizione (cdf)}}$: Data X, $F_X: \mathbb{R} \longrightarrow [0,1]$ / $F_X = P\{X \leq x\} = P\{X^{-1}(]-\infty, x])\} = P(]-\infty, x])$

$\downarrow$ In pratica

X discreta, quindi $P\{X \leq x\} = \sum_{x_i \leq x} p(x_i)$



$\boxed{X \text{ binomiale}}$ allora l'ampiezza del gradino x_i è $p(x_i)$, mentre se X è geometrica allora non arriva mai ad 1.

$\boxed{X \text{ con densità } f} \longrightarrow F(x) = P\{X \le x\} = \int_{-\infty}^{x} f(\tau)\,d\tau$

X $\boxed{\text{densità uniforme}}$ per $[a,b]$: $f(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{altrove} \end{cases} \longrightarrow \boxed{F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a < x < b \\ 1 & x > b \end{cases}}$

Se X ha densità, $F(X) = \int_{-\infty}^{x} f(\tau)\,d\tau$ è continua ⚠ Ma non vice-versa

X $\boxed{\text{esponenziale}} \longrightarrow f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \le 0 \end{cases} \longrightarrow \boxed{F(x) = \begin{cases} 1 - e^{-\lambda x} & x > 0 \\ 0 & x \le 0 \end{cases}}$

proprietà: $F: \mathbb{R} \longrightarrow [0,1]$

a. $x < y \longrightarrow F(x) \le F(y)$ (è debolmente crescente)

b. $\lim_{x \to -\infty} F(x) = 0$ $\qquad \lim_{x \to +\infty} F(x) = 1$

c. F è continua da destra: $F(x) = \lim_{h \to 0^+} F(x+h)$

Dimostrazione: b. e c. sono conseguenza della continuità della probabilità

b. $\lim_{x \to \infty} F(x) = 0$. Prendiamo $x_1 > x_2 > x_3$, quindi $x_n \longrightarrow -\infty$

↓ Dimostriamo

$\lim_{n \to \infty} F(x_n) = 0$

$F(x_n) = P(\underbrace{X^{-1}(]-\infty, x_n])}_{A_n}) \longrightarrow A_n \supseteq A_{n+1} \supseteq A_{n+2}$

$\bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} X^{-1}(]-\infty, x_n]) = X^{-1}(\bigcap_{n=1}^{\infty}]-\infty, x_n]) = X^{-1}(\emptyset) = 0$. Allora

$\lim_{n \to \infty} F(x_n) = \lim_{n \to \infty} P(A_n) = P(\bigcap_{n=1}^{\infty} A_n) = P(\emptyset) = 0$

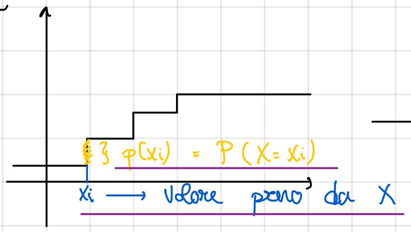
⚠ Presa una funzione $F: \mathbb{R} \to [0,1]$ con le proprietà a, b, c. allora $\exists!$ probabilità Q su $\mathbb{R}$ tale che $F(x) = Q(]-\infty, x])$

↓ Conseguenze

1. Nota F_X, posso ricostruire P_X
2. Se $F_X = F_Y$, X e Y sono equidistribuite

↓ Applicazione

Variabili discrete:



$\longrightarrow$ lo faccio per tutti e ricostruisco

variabili con densità: $F(x) = \int_{-\infty}^{x} f(\tau)\,d\tau$ ma sappiamo $f(x) = \frac{dF(x)}{dx}$ in tutti i punti ove la densità è continua. F è derivabile in tutti i punti eccetto a e b e $\frac{dF(x)}{dx} = f(x)$ eccetto che a e b

ma questo non da problemi

$\boxed{\text{funzione di ripartizione empirica}}$: $x_1, x_2, \dots, x_n$ in ordine. La funzione ogni volta fa un salto di $1/n$. Questa è la cdf di una variabile aleatoria discreta i cui valori sono $x_1, \dots, x_n$ e $p(x_i)$ sono tutti eguali

nozione di quantile: X variabile aleatoria, $0<\beta<1$ $\quad x_\beta = \beta$-quantile

Intuitivamente, x_β è tale che $P\{X \le x_\beta\} = F(x_\beta) = \beta$. Possono esserci problemi quando F fa salti o se F è piatta $\rightarrow$ x_β è un β-quantile se $P(X \le x_\beta) \ge \beta$ $\wedge$ $P(X \ge x_\beta) \ge \beta$

↓ In pratica

Supponiamo X discreta: Nel caso di F piatta si prende il punto di mezzo

Supponiamo X con densità: Non può capitare il salto (continuità) e spesso è anche invertibile $\longrightarrow$ Risolvo $F(x_\beta) = \beta$

ESERCIZIO: Calcolare i β-quantili per densità uniforme ed esponenziale

↳ Quantili variabili con densità

UNIFORME su $[a,b]$: $F(x) = \begin{cases} 0 & x \le a \\ \frac{x-a}{b-a} & a<x<b \\ 1 & x \ge b \end{cases}$ $\longrightarrow F(x_\beta) = \frac{x_\beta - a}{b-a} = \beta \longrightarrow x_\beta = a + \beta(b-a)$

ESPONENZIALE par λ: $F(x) = \begin{cases} 0 & x \le 0 \\ 1 - e^{-\lambda x} & x > 0 \end{cases}$ $\longrightarrow F(x) = 1 - e^{-\lambda x} = \beta$ da cui $e^{-\lambda x} = 1-\beta$

da cui $x_\beta = \frac{-\log(1-\beta)}{\lambda}$

PROPRIETÀ

$F(x) = P\{X \le x\}$. Supponiamo che $x<y \longrightarrow F(y) - F(x) = P\{x < X \le y\}$

↓ Motivazione

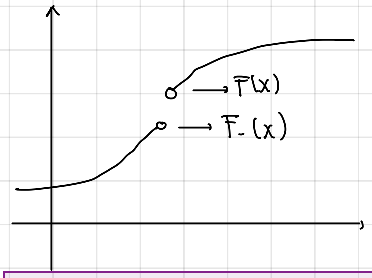
Sappiamo che se $B \subseteq A$, $P(A-B) = P(A) - P(B)$. Da ciò scriviamo $\{x<X\le y\} = \{X \le y\} \setminus \{X \le x\}$. Da ciò si ha $P(x<X\le y) = F(y) - F(x)$

↓ Se la funzione fa un salto

$F(x) - F_-(x) = P\{X = x\}$

↓ Conseguenza

Se X ha densità: $P\{X=x\} = \int_{\{x\}} f(t)\,dt = 0$



variabile di Poisson di parametro $\lambda > 0$

Valori possibili: $0, 1, 2, \dots$

$\boxed{p(h) = P\{X = h\} = e^{-\lambda} \cdot \frac{\lambda^h}{h!}}$

↓ Verifica prop.

$p(0) + p(1) + \dots = \sum_{h=0}^{\infty} p(h) = \sum_{h=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^h}{h!} = e^{-\lambda} \cdot \sum_{h=0}^{\infty} \frac{\lambda^h}{h!} \underset{\text{serie } e^\lambda}{=} e^{-\lambda} \cdot e^{\lambda} = 1$

proprietà: a. È senza memoria: $P\{X = n+h \mid X > n\} = P\{X = h\}$

↓ Dim.

$\frac{P\{X=n+h, X>n\}}{P\{X>n\}} = \frac{P\{X=n+h\}}{P\{X>n\}} = \frac{(1-p)^{n+h-1} \cdot p}{(1-p)^n} = (1-p)^{h-1} \cdot p = P\{X = h\}$

$P\{X>n\} = \sum_{h=n+1}^{+\infty} p(h) = 1 - P\{X \le n\} = 1 - \sum_{h=0}^{n} p(h)$. Ma cosa vuol dire $\{X>n\}$? Ho fallito n volte: $\boxed{P\{X>n\} = (1-p)^n}$

↓

Anche la variabile esponenziale è senza memoria: presi $s, T > 0$ $P\{X \le s+T \mid X > T\} = P\{X \le s\}$. Tipicamente X_{esp} rappresenta un tempo di attesa (rimane uguale anche se aspetto T): guasti di apparecchiature che non invecchiano

$$P\{X \le s\} = \int_0^s \lambda e^{-\lambda x}\,dx = 1 - e^{-\lambda s}$$

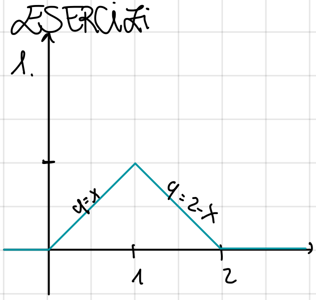
$$P\{X \le s+T \mid X > T\} = \frac{\boxed{P\{T < X \le T+s\}}}{\boxed{P\{X > T\}}} \longrightarrow \int_T^{T+s} \lambda e^{-\lambda x}\,dx = e^{-\lambda T} - e^{-\lambda (T+s)}$$

$e^{-\lambda T}$

$$= \frac{e^{-\lambda T} - e^{-\lambda (T+s)}}{e^{-\lambda T}} = 1 - e^{-\lambda s}$$

ESERCIZI

1.



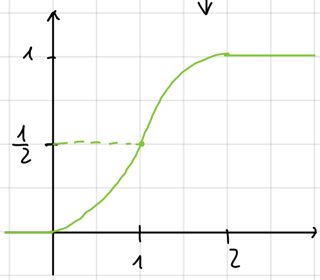
1. Densità

$$f(x) = \begin{cases} 0 & x \le 0 \\ x & 0 < x \le 1 \\ 2-x & 1 \le x < 2 \\ 0 & x > 2 \end{cases}$$

2. Funzione di ripartizione

$$F(x) = \begin{cases} 0 & x \le 0 \\ \frac{x^2}{2} & 0 < x \le 1 \\ 2x - \frac{x^2}{2} - 1 & 1 < x \le 2 \\ 1 & x \ge 2 \end{cases}$$

$$\longrightarrow \frac{1}{2} + \int_1^x (2-\tau)\,d\tau = \frac{1}{2} + 2(x-1) - \frac{x^2}{2} + \frac{1}{2}$$



3. β - quantile

x_β risolve l'equazione $F(x_\beta) = \beta$

⚠ La funzione ha diverse espressioni a seconda delle x considerate

$0 < \beta \le \frac{1}{2} \longrightarrow \frac{x^2}{2} = \beta \longrightarrow x_\beta = \cancel{\pm}\sqrt{2\beta}$ positivo

$\frac{1}{2} < \beta < 1 \longrightarrow 2x - \frac{x^2}{2} - 1 = \beta \longrightarrow \frac{x^2}{2} - 2x + (1+\beta) = 0$

$x = 2 \cancel{\pm} \sqrt{4 - 2(1+\beta)}$

$\boxed{x_\beta \text{ deve essere compreso tra } 1 \text{ e } 2}$

$x = 2 - \sqrt{4 - 2(1+\beta)}$

2. **Esercizio 2.4** Sia X una variabile di Poisson di parametro λ: dire qual è il valore più probabile.
Più in generale dire per quali interi n, $\mathbf{P}\{X = n+1\} \ge \mathbf{P}\{X = n\}$.

$p(n) = e^{-\lambda} \frac{\lambda^n}{n!}$. Su quale intero n la funzione $p(n)$ raggiunge il suo massimo?

⚠ Non posso annullare la derivata sugli interi

$\sum p(n) = 1 < +\infty$ e $\lim_{n \to +\infty} p(n) = 0$ ∴ il massimo lo cerco vicino

Quando $p(n+1) > p(n)$? $\quad e^{-\lambda} \frac{\lambda^{n+1}}{(n+1)!} > e^{-\lambda} \frac{\lambda^n}{n!}$ cioè $\lambda > (n+1)$

Se $(n+1) < \lambda$, $p(n+1) > p(n)$

Se $(n+1) = \lambda$, $p(n+1) = p(n)$ $\longrightarrow$ Massimo qui: sale e poi scende

Se $(n+1) > \lambda$, $p(n+1) < p(n)$

Supponiamo λ intero, $\lambda = m$ $m \in \mathbb{N}$. In questo caso il massimo è raggiunto in due punti: $m-1$ e m

↓ Esempio X esp per 5

$p(4) \overset{deve\ essere}{=} p(5)$: $e^{-5} \frac{5^4}{4!} = e^{-5} \frac{5^5}{5!}$ ✓

Se λ non è intero il massimo si ottiene in $[\lambda]$ (parte intera)

problema: X ha densità $f(\cdot)$. $Y = aX+b$ $(a>0)$. Y ha densità? Come è fatta?

$F_Y(y) = P\{Y \le y\} = P\{aX+b \le y\} = P\{X \le \frac{y-b}{a}\} = F_X\left(\frac{y-b}{a}\right)$

Quindi $f_Y(y) = \frac{d\, F_Y(y)}{dy} = f_X\left(\frac{y-b}{a}\right) \cdot \frac{1}{a}$

↓ Più in generale

X con densità f_X, $Y = h(x)$. Y ha densità? Non sempre

C'è una formula esplicita? Non sempre

↓ Come scoprirlo

Si calcola $F_Y(y) = P\{h(X) \le y\}$ e si guarda se è derivabile. Se lo è, $f_Y(y) = \frac{d}{dy}F_Y(y)$.

C'è però un caso particolare: X ha densità $f_X(\cdot)$ diversa da 0 su un "intervallo" A limitato o illimitato. Supponiamo di avere $h: A \subseteq \mathbb{R} \to B$ intervallo di $\mathbb{R}$ biunivoca, derivabile, con inversa derivabile. $Y = h(X)$ ha densità $\boxed{f_Y(y) = \begin{cases} f_X(h^{-1}(y))\left(\frac{d h^{-1}(y)}{dy}\right) & y \in B \\ 0 & y \notin B \end{cases}}$

⚠ Promemoria: $f_Y(y) = f_X(x(y))\left(\frac{d\,x(y)}{dy}\right)$

ESEMPI: 1. X densità uniforme su $[-1,1]$, $Y = X^2$. Posso usare la formula diretta? $X \geq 0$

$f_X(x) = \begin{cases} \frac{1}{2} & -1 < x < 1 \\ 0 & \text{altrove} \end{cases}$ $\qquad X \to x^2$ non biunivoca su $]-1,1[$

↓

X valori tra -1 e 1

$Y = X^2$ valori tra 0 e 1 $\longrightarrow$ $\boxed{F_Y(y) = \begin{cases} 0 & y < 0 \\ \sqrt{y} & 0 < y < 1 \\ 1 & y \geq 1 \end{cases}}$

$\longrightarrow F_Y(y) = P\{X^2 \le y\} = P\{-\sqrt{y} \le x \le \sqrt{y}\} = \sqrt{y}$

↓

$\boxed{f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}} & 0 < y < 1 \\ 0 & y \le 0 \ \vee \ y \geq 1 \end{cases}}$

2. X esponenziale per λ $Y = X^2$. Posso usare la formula diretta? Sì perché $f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \le 0 \end{cases}$ quindi $f_X(\cdot)$ è diversa da 0 su $]0,+\infty[$

↓

$X \to X^2$ è biunivoca su $]0,+\infty[\to]0,+\infty[$

$h(x) = x^2$, $h^{-1}(y) = \sqrt{y}$

$\boxed{f_Y(y) = \begin{cases} \lambda e^{-\lambda\sqrt{y}} \cdot \frac{1}{2\sqrt{y}} & y > 0 \\ 0 & y \le 0 \end{cases}}$

3. **Esercizio 2.12.** Sia T una variabile con densità esponenziale di parametro λ, e sia $S = \log T$.

a) Per quale t vale l'eguaglianza $\mathbf{P}\{T > t\} = 0.1$?

b) Esaminare se la variabile S ha densità e in caso affermativo calcolarla, ed esaminare se S^2 ha densità.

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & t>0 \\ 0 & t\le 0 \end{cases}$$

$0.1 = P\{T>t\} = \int_t^{+\infty} \lambda e^{-\lambda s} ds = -e^{-\lambda s}\Big|_t^{+\infty} = e^{-\lambda t}$ da cui

$e^{-\lambda t} = 0.1 \longrightarrow t = -\frac{\log(0.1)}{\lambda} > 0$

$S = \log T$. T ha densità $\neq 0$ su $]0,+\infty[$. $t \longrightarrow s = \log t$ è biunivoca da $]0,+\infty[$ a $]-\infty,+\infty[$. Posso applicare la formula diretta ponendo $t = e^s$.

$$f_S(s) = f_T(e^s)\cdot e^s = \lambda e^s \cdot e^{-\lambda e^s} \neq 0 \text{ su tutto } \mathbb{R}$$

Se considero una nuova variabile S^2 non posso applicare la formula diretta

↓ Caso più generale

Sia X variabile aleatoria con densità $f_X(\cdot)$ e $Y = X^2$. Vogliamo una formula per la densità di Y.

$$F_Y(y) = P\{Y\le y\} = P\{X^2 \le y\} = \begin{cases} 0 & y<0 \\ \end{cases}$$

$y \geq 0 \longrightarrow P\{-\sqrt{y} \le X \le \sqrt{y}\} = \int_{-\sqrt{y}}^{\sqrt{y}} f(x)dx$. Sappiamo che $\frac{d}{dy}\int_a^y g(x)dx = g(y)$. Sappiamo
continua

per la derivata di f composte che $\frac{d}{dy}\int_a^{\sqrt{y}} g(x)dx = g(\sqrt{y})\cdot\frac{1}{2\sqrt{y}}$. Sappiamo inoltre che $\frac{d}{dy}\int_{-\sqrt{y}}^{b} g(x)dx = g(-\sqrt{y})\cdot\frac{1}{2\sqrt{y}}$

$$\frac{d}{dy}\int_{-\sqrt{y}}^{\sqrt{y}} f(x)dx = \boxed{f(\sqrt{y})\cdot\frac{1}{2\sqrt{y}} + f(-\sqrt{y})\cdot\frac{1}{2\sqrt{y}}}$$

4. **Esercizio 2.5.** Si lanciano due dadi equilibrati, e consideriamo i due eventi A *"la somma dei due numeri usciti è 7"* , e B *"il numero comparso nel primo dado è 4"*.

1) Provare che gli eventi A e B sono indipendenti.

2) Più in generale, indicando rispettivamente con le v.a. X ed Y il numero uscito nel primo e nel secondo dado, si considerino gli eventi $\{X+Y=n\}$ e $\{X=k\}$: esaminare per quali valori di n e di k questi due eventi rusultano indipendenti.

X, Y risultati 1° e 2° lancio. Valori $1,2,\dots,6$ con $P = \frac{1}{6}$.
$A = \{X+Y=7\}$ $B = \{X=4\}$. È vero che $P(A\cap B) = P(A)\cdot P(B)$?
$P(B) = \frac{1}{6}$ $A\cap B = \{X+Y=7, X=4\} = \{X=4, Y=3\} \longrightarrow P(A\cap B) = \frac{1}{6}\cdot\frac{1}{6} = \frac{1}{36}$
Deve valere $P(A) = \frac{1}{6} \longrightarrow 6$ possibilità ciascuna con $P = \frac{1}{36}$ da cui $6\cdot\frac{1}{36} = \frac{1}{6} \longrightarrow$ OK

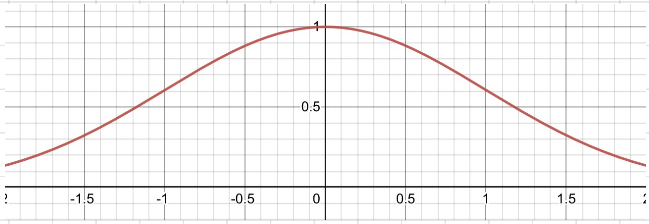
↓ Vogliamo verificarlo generalmente

$A = \{X+Y=n\}$, $B = \{X=k\}$. Per quali n, k sono indip.?
$A\cap B = \{X=k, Y=n-k\}$. Se $1\le k\le 6$ e anche $1\le n-k\le 6$ allora $P(A\cap B) = 1/36$ e $P(B) = 1/6$ e quindi si ha indipendenza se $P\{X+Y=n\} = \frac{1}{6}$. Prendiamo $n=6 \longrightarrow P\{X+Y=6\} = \frac{5}{36}$ (possibili casi) $< \frac{1}{6}$ $\forall n<7$. Se $n=8$ si ha $P\{X+Y=8\} = \frac{5}{36} \rightarrow$ primo val. 2 $< \frac{1}{6}$ e quindi

non può esserci indipendenza. Gli eventi possono essere indipendenti se k è qualsiasi $(1 \leq k \leq 6)$ e $n=7$

• LA DENSITÀ GAUSSIANA

Prendiamo $f(x) = e^{-\frac{x^2}{2}}$



$\longrightarrow \int_{-\infty}^{+\infty} e^{-x^2/2}\,dx < +\infty$

Non si può scrivere la primitiva con funzioni elementari: $\int_a^b e^{-x^2/2}\,dx$ si può solo approssimare. Laplace ha dimostrato che $\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} = \sqrt{2\pi}$. La conseguenza è che $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ è una densità ($N(0,1)$) e $\phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-T^2/2}\,dT$

è la funzione di ripartizione della variabile $\boxed{N(0,1)}$ densità gauss normale $\longrightarrow$ Tavola della variabile $N(0,1)$ per comportamento. Questa fornisce i valori di $\phi(x)$ per $0 \le x \le 4$ (poi circa 1)

Tavola della funzione di ripartizione della distribuzione N(0,1)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91308	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.7	0.99989	0.99990	0.99990	0.99990	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997
4.0	0.99997	0.99997	0.99997	0.99997	0.99997	0.99997	0.99998	0.99998	0.99998	0.99998
4.1	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99999	0.99999
4.2	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
4.3	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
4.4	0.99999	0.99999	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000

Per $x < 0$ si ha $\phi(-x) = 1 - \phi(x)$. ϕ è pari quindi $\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-T^2/2}\,dT =$
$= \int_{x}^{+\infty} \frac{e^{-T^2/2}}{\sqrt{2\pi}}\,dT = 1 - \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-T^2/2}\,dT$

E i quantili? q_α è il quantile di $N(0,1)$ e si trovano usando la tavola all'inverso. Per trovarlo devo fare $\phi(q_\alpha) = \alpha$

$q_{0.9}$ = cerco il valore più vicino a $0.9 \longrightarrow 1.1289$. Si trovano quindi i q_α per $\frac{1}{2} < \alpha < 1$. Se $0 < \alpha < \frac{1}{2}$: $q_{1-\alpha} = -q_\alpha \longrightarrow q_{0.32} = -q_{0.68}$. La relazione è data da $\int_{q_\alpha}^{+\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\,dx = 1-\alpha$ che corrisponde all'integrale tra $-\infty$ e $-q_\alpha$.

alcuni calcoli: X gaussiana standard.

$P\{-1 \le X \le 1\} = \phi(1) - \phi(-1) = 2\phi(1) - 1 \simeq 0.68$

$P\{-2 \le X \le 2\} = 0.94$

$P\{-3 \le X \le 3\} = 0.997$

$\boxed{\text{densità gaussiana generale}}$: $N(m, \sigma^2)$ $\quad m \in \mathbb{R}$, $\sigma > 0$

X gaussiana standard, $Y = m + \sigma X \longrightarrow$ la densità di Y è $N(m, \sigma^2)$.

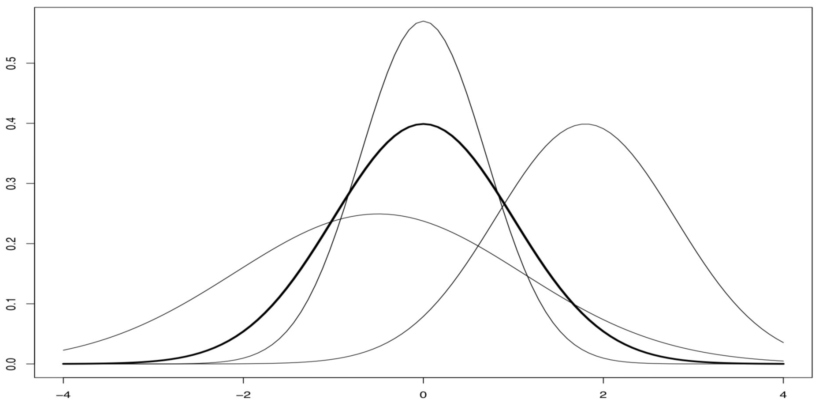
Avremo $g(y) = f\left(\frac{y-m}{\sigma}\right) \cdot \frac{1}{\sigma}$

$$g(y) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2} \xrightarrow{\text{quindi}} Y \sim N(m, \sigma^2) \Longrightarrow \frac{Y-m}{\sigma} \text{ è normale standard}$$

Una variabile $N(m, \sigma^2)$ la si considera della forma $(\sigma X + m)$ con $X \sim N(0,1)$.
Ma quanto vale $P\{m - \sigma \le Y \le m + \sigma\}$? $\longrightarrow P\left\{-1 \le \frac{Y-m}{\sigma} \le 1\right\} = 0.68$. Per la stessa ragione, $P\{m - 2\sigma \le Y \le m + 2\sigma\} = 0.94$ e $P\{m - 3\sigma \le Y \le m + 3\sigma\} = 0.997$.
Se abbiamo dei dati $x = (x_1, \dots, x_n)$ ove n è grande ($n \ge 100$). Se l'istogramma è normale allora circa il 68% dei dati è compreso tra $\bar{x} - \sigma(x)$ e $\bar{x} + \sigma(x) \longrightarrow$ Ma cosa significa che l'istogramma è normale? Assomiglia ad una densità gaussiana.

confrontiamo alcune densità: $\varphi(x) = N(0,1)$ con densità $N(0, 0.3)$ e con $N(0,4)$.

σ grande = valori più concentrati (va a 0 molto velocemente)



ESEMPI

1. $X \sim N(2, 4)$ vogliamo $P\{1 \le X \le 4\}$
$\downarrow$ $\boxed{\text{Mi porto a std (tolgo } m\text{, divido per } \sigma)}$

$$\frac{X-2}{2} = P\left\{\frac{1-2}{2} \le \frac{X-2}{2} \le \frac{4-2}{2}\right\} = P\left\{-\frac{1}{2} \le Z \le 1\right\} = \phi(1) - \phi(0.5) =$$

$$= \phi(1) + \phi(0.5) - 1$$

2. Calcolare, usando la tavola $N(0,1)$ l'integrale $\int_1^3 e^{-x^2}\,dx$

$e^{-x^2} \xrightarrow{\text{cambio variabile}} e^{-\frac{t^2}{2}}$, quindi $x = t/\sqrt{2}$ da cui $t = \sqrt{2}x$

$\downarrow$ Integrale diventa

$$\int_{\sqrt{2}}^{3\sqrt{2}} e^{-\frac{t^2}{2}} \cdot dt/\sqrt{2} = \frac{\sqrt{\pi}}{\sqrt{2\pi}} \int_{\sqrt{2}}^{3\sqrt{2}} e^{-\frac{t^2}{2}}\,dt = \sqrt{\pi}\left(\phi(3\sqrt{2}) - \phi(\sqrt{2})\right)$$

- VARIABILI ALEATORIE DOPPIE

$(X,Y): \Omega \longrightarrow \mathbb{R}^2$. È associata una probabilità sui sottoinsiemi di $\mathbb{R}^2$ $A \subseteq \mathbb{R}^2$. $P_{X,Y}(A) = P((X,Y)^{-1}(A)) = P\{(X,Y) \in A\}$. Anche in questo caso possono essere discrete, con densità o più generali

variabili discrete: l'immagine è finita o numerabile. Si conosce se si conosce l'immagine e la funzione $p(x_i, y_j) = P\{X = x_i, Y = y_j\}$ $(, = \wedge)$

$p(x_i, y_j) \geq 0$
$\sum_{x_i, y_j} p(x_i, y_j) = 1$
$\forall A \subseteq \mathbb{R}^2 \rightarrow P_{X,Y}(A) = \sum_{(x,y) \in A} p(x_i, y_j)$

} proprietà probabilità

variabili con densità: $f: \mathbb{R}^2 \longrightarrow [0, +\infty[$ $(f(x,y))$ tale che $\iint_{\mathbb{R}^2} f(x,y)\,dx\,dy = 1$ e $P_{X,Y}(A) = P((X,Y) \in A) = \iint_A f(x,y)\,dx\,dy$

↓

Prendiamo (X,Y) discreta identificata da $p(x_i, y_j) = P\{X = x_i, Y = y_j\}$. X è una variabile discreta a valori reali. Y è una variabile discreta a valori reali quindi $p_X(x_i) = P\{X = x_i\}$ e vale lo stesso per Y.

↓

Nota $p(x_i, y_j)$ posso ricostruire $p_X(x_i)$ e $p_Y(y_j)$ ma non vice-versa. Per farlo si usa la formula $p_X(x_i) = \sum_{y_j} p(x_i, y_j)$ e $p_Y(y_j) = \sum_{x_i} p(x_i, y_j)$.

↓ Dimostrazione

$p_X(x_i) = P\{X = x_i\}$ e $p(x_i, y_j) = P\{X = x_i, Y = y_j\}$. $\{X = x_i\} = \boxed{\bigcup_{y_j} \{X = x_i, Y = y_j\}}$ disgiunti

$P\{X = x_i\} = \boxed{\sum_{y_j}} P\{X = x_i, Y = y_j\}$

Formule analoghe per variabili con densità: (X,Y) con densità $f(x,y)$. Anche X ha densità data da $f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy$ e $f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx$. Se conosco $f_X(\cdot)$ e $f_Y(\cdot)$ non posso ricostruire $f(x,y)$ tranne nel caso in cui siano **indipendenti**

↓

$(\Omega, \mathcal{F}, P)$, $A, B \subseteq \Omega$ sono indipendenti se $P(A \cap B) = P(A) \cdot P(B)$. $X, Y: \Omega \longrightarrow \mathbb{R}^2$. X e Y sono indipendenti se scelti $A, B \subseteq \mathbb{R}$ gli eventi $X^{-1}(A)$ e $Y^{-1}(B)$ sono indipendenti. $\longrightarrow P(X^{-1}(A) \cap Y^{-1}(B)) = P(X^{-1}(A)) \cdot P(Y^{-1}(B))$ per ogni scelta di A e B.

↓ Caratterizzazione

X, Y discrete sono indipendenti se e solo se $p(x_i, y_j) = p_X(x_i) \cdot p_Y(y_j)$.
X, Y con densità sono indipendenti se e solo se $f(x,y) = f_X(x) \cdot f_Y(y)$

↓ **Dimostrazione caso discreto**

$P\{X \in A, Y \in B\} = P\{X \in A\} \cdot P\{X \in B\}$. X, Y discrete e consideriamo $A = \{x_i\}$ e $B = \{y_j\} \longrightarrow P\{X = x_i, Y = y_j\} = P\{X = x_i\} \cdot P\{Y = y_j\}$ che equivale a $p(x_i, y_j) = p_X(x_i) \cdot p_Y(y_j)$

Vice-versa: $P\{X \in A, Y \in B\} = P\{(X,Y)^{-1}(A \times B)\} = \sum_{x_i, y_j \in A \times B} p(x_i, y_j) = \sum_{\substack{x_i \in A \\ y_j \in B}} p(x_i, y_j) =$

$= \sum_{\substack{x_i \in A \\ y_j \in B}} p_X(x_i) \cdot p_Y(y_j) = \left(\sum_{x_i \in A} p_X(x_i)\right)\left(\sum_{y_j \in B} p_Y(y_j)\right) = P\{X \in A\} \cdot P\{X \in B\}$

⚠ Funzioni di variabili indipendenti sono indipendenti: X, Y indipendenti allora $h(X)$ e $K(Y)$ sono indipendenti (X^2 e $\sqrt{1+|Y|}$ sono indipendenti)

↓ Dim.

$h(X)^{-1}(A) = X^{-1}(h^{-1}(A))$
$K(Y)^{-1}(B) = Y^{-1}(K^{-1}(B))$
} indipendenti $\longrightarrow$ vale anche per più variabili a patto che la stessa variabile non compaia in più funzioni

ESEMPIO (somma v.a. indipendenti)

X binomiale $B(n,p)$ } indipendenti
Y binomiale $B(m,p)$

n e m possono essere $\neq$ ma p deve essere lo stesso. Mi aspetto che $X+Y$ sia binomiale $B(n+m, p)$

↓ Prova

$X \sim B(n,p)$ $\quad Y$ Bernoulli par. $p \longrightarrow \{X+Y=h\} = \{X=h, Y=0\} \cup \{X=h-1, Y=1\}$

$$P\{X+Y=h\} = P\{X=h\}(1-p) + P\{X=h-1\}\cdot p = \binom{n}{h}p^h(1-p)^{n-h}(1-p) + \binom{n}{h-1}p^{h-1}(1-p)^{n-h+1}p =$$
$$= p^h(1-p)^{n+1-h}\left[\binom{n}{h}+\binom{n}{h-1}\right] = p^h(1-p)^{n+1-h}\binom{n+1}{h}$$

↓

Conseguenza: una variabile binomiale $B(n,p)$ può essere vista nella forma $X = X_1 + \ldots + X_n$ ove $X_1, X_2, \ldots, X_n$ sono indipendenti Bernoulli di parametro p.

Per variabili di $\boxed{\text{Poisson}}$: X Poisson par. λ } INDIP
Y Poisson par. μ

↓

Si dimostra che $(X+Y)$ è di Poisson di parametro $(\lambda+\mu)$. Chiamiamo $Z = X+Y$

$$\{Z=n\} = \bigcup_{h=0}^{n}\{X=h, Y=n-h\} \longrightarrow P\{Z=n\} = \sum_{h=0}^{n} e^{-\lambda}\frac{\lambda^h}{h!}\cdot e^{-\mu}\frac{\mu^{n-h}}{(n-h)!} =$$
$$= \frac{e^{-(\lambda+\mu)}}{n!}\sum_{h=0}^{n}\frac{n!}{h!\,(n-h)!}\lambda^h\mu^{n-h} = \frac{e^{-(\lambda+\mu)}}{n!}(\lambda+\mu)^n \longrightarrow \text{formula della}$$

Convoluzione discreta: X, Y indipendenti a valori discreti di $\mathbb{N}$

↓

$$P\{X+Y=n\} = \sum_{h=0}^{n} P\{X=h\}\cdot P\{Y=n-h\}$$

↓ con densità

X, Y con densità indipendenti. $Z = X+Y$ con densità $f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x)\,dx =$
$$= \int_{-\infty}^{+\infty} f_Y(y) f_X(z-y)\,dy$$

Con variabili gaussiane: $X \sim N(m_1, \sigma_1^2)$ } indipendenti
$Y \sim N(m_2, \sigma_2^2)$

↓

$(X+Y)$ è gaussiana $N(m_1+m_2, \sigma_1^2+\sigma_2^2)$

ESERCIZI

Esercizio 2.6. Vi sono due monete, indistinguibili al tatto, delle quali una è perfettamente equilibrata e l'altra è truccata: per quest'ultima è $\frac{2}{3}$ la probabilità che esca testa. Se ne sceglie una a caso e la si lancia 2 volte di seguito, e siano X_1 e X_2 i risultati dei due lanci (X_i vale 1 oppure 0 a seconda che all'i–mo lancio sia uscita testa oppure croce).

1) Le variabili X_1 e X_2 sono indipendenti?

2) Se è uscita testa in entrambi i lanci, qual è la probabilità che sia stata scelta la moneta truccata?

⚠ Non nelle stesse condizioni

B_1: moneta equilibrata $\quad P(B_1) = P(B_2) = \frac{1}{2}$
B_2: moneta truccata

$$P\{X_1=1, X_2=1\} \overset{?}{=} P\{X_1=1\}\cdot P\{X_2=1\}$$
$$P\{X_1=1\} = P\{X_1=1 \mid B_1\}\cdot P(B_1) + P\{X_1=1 \mid B_2\}\cdot P(B_2) = \frac{1}{2}\cdot\frac{1}{2} + \frac{2}{3}\cdot\frac{1}{2} = \frac{7}{12}$$

$$P\{X_2=1\} = \frac{7}{12}$$

$$P\{X_1=1, X_2=1\} = P\{X_1=1, X_2=1 \mid B_1\}\cdot P(B_1) + P\{X_1=1, X_2=1 \mid B_2\}\cdot P(B_2) = \left(\frac{1}{2}\right)^2\frac{1}{2} + \left(\frac{2}{3}\right)^2\frac{1}{2}$$

$\neq \left(\frac{7}{12}\right)^2$: non indipendenti

A = testa 2 volte

$$P(B_1 | A) = \frac{P(A|B_1) \cdot P(B_1)}{P(X)} = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{3}{4}} < \frac{1}{2}$$: ragionevole perché è più probabile che io abbia scelto la seconda

Esercizio 2.7. Sia (X, Y) una variabile doppia avente densità

$$f(x,y) = \begin{cases} x + y & 0 < x < 1,\ 0 < y < 1 \\ 0 & \text{altrove} \end{cases} \longrightarrow \text{quadrato unitario}$$

Calcolare le densità marginali delle componenti X e Y, calcolare valore atteso e varianza della variabile X.

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy \longrightarrow f_X = 0 \text{ se } x < 0 \vee x > 1$$

$$\int_0^1 (x+y)\,dy = x + \frac{1}{2} \quad 0 < x < 1$$

$\downarrow$ verifica

$\int_0^1 (x + \frac{1}{2})\,dx = 1$? sì

$$f_Y(y) = \begin{cases} y + \frac{1}{2} & 0 \le y \le 1 \\ 0 & \text{altrove} \end{cases}$$

VALORI ATTESI e MOMENTI

$X = (X_1, \dots, X_n) \longrightarrow \bar{X} = \sum_{i=1}^{n} \frac{x_i}{n} = \sum_{i=1}^{n} x_i \cdot \frac{1}{n}$. Supponiamo che $x_1, \dots, x_n$ siano i valori di una variabile aleatoria X discreta e $p(x_i) = P\{X = x_i\}$. Si considera come media: $\sum_{i=1}^{n} x_i \cdot p(x_i) \longrightarrow$ "media" naturale della v.a. X

↓ Definizione

Sia $\boxed{X \text{ v.a. discreta}}$: si dice che X possiede valore atteso se $\sum_i |x_i| \cdot p(x_i) < +\infty$ e in tal caso si chiama valore atteso $E[X] = \sum_i x_i \cdot p(x_i)$

Sia $\boxed{X \text{ v.a. con densità } f}$: X ha valore atteso se $\int_{\mathbb{R}} |x| f(x)\, dx < +\infty$ e $E[X] = \int_{-\infty}^{+\infty} x f(x)\, dx$

expectation

↓

Perché si chiede $\sum |x_i| \cdot p(x_i) < +\infty$? Facciamo dei richiami su serie numeriche: abbiamo $x_1, x_2, x_3, \dots \longrightarrow \sum_{n=1}^{+\infty} x_n = \lim_{n \to +\infty} \sum_{i=1}^{n} x_i$: se $\exists$, converge

Se $a_n \geq 0$ ha sempre senso $\sum_{n=1}^{\infty} a_n = \lim_{n \to +\infty} \sum_{k=1}^{n} a_k \in [0, +\infty[$

x_n generici, $\sum_{n=1}^{+\infty} |x_n| \in [0, +\infty[$ se $\sum_{n=1}^{+\infty} |x_n| < +\infty$ la serie converge assolutamente

Se una serie converge assolutamente allora converge ma non vice-versa. Se $(x_n)_{n=1}^{+\infty}$ converge assolutamente si può cambiare l'ordine della somma $\longrightarrow$ prendendo $\nu : \mathbb{N} \longrightarrow \mathbb{N}$ biunivoca, $\sum_{n=1}^{+\infty} x_n = \sum_{n=1}^{+\infty} x_{\nu(n)}$. Si possono inoltre raggruppare i termini. Consideriamo $A_1, A_2, A_3, \dots$ partizioni di $\mathbb{N}$, allora $\sum_{n=1}^{+\infty} x_n = \sum_{n=1}^{+\infty} \left(\sum_{k \in A_n} x_k \right)$

↓

Risultato: supponiamo che $\exists \sum_{n=1}^{+\infty} x_n = \ell$ però $\sum_{n=1}^{+\infty} |x_n| = +\infty$. Preso qualunque $x \in]-\infty, +\infty[$, $\exists\ \nu : \mathbb{N} \longrightarrow \mathbb{N} \,/\, \sum x_{\nu(n)} = x$

↓

Supponiamo X binomiale $B(n,p) \longrightarrow E[X] = \sum_{h=0}^{n} h \binom{n}{p} p^h (1-p)^{n-h}$ esiste!

X geometrica par. $p \longrightarrow E[X] = \sum_{h=1}^{+\infty} h (1-p)^{h-1} \cdot p$ esiste?

X esponenziale par. $\lambda \longrightarrow E[X] = \int_0^{+\infty} x \lambda e^{-\lambda x}\, dx$ esiste?

↓

Supponiamo che X abbia solo valori positivi. Ha sempre senso scrivere $E[X] = \sum_i x_i \cdot p(x_i)$ (discreta) $\in [0, +\infty[$ e $E[X] = \int_0^{+\infty} x f(x)\, dx$ (densità) $[0, +\infty[$

↓

X ha valore atteso se e solo se $E[|X|] < +\infty$ e in tal caso $|E[X]| \leq E[|X|]$

↓

Osservazione: $Y = g(X)$ [X, Y v.a.]

↓

$\boxed{X \text{ discreta}}$: $Y = g(X)$ ha valore atteso se $\sum |g(x_i)| p(x_i) < +\infty$ e in tal caso $E[g(X)] = \sum_{i=1}^{n} g(x_i) \cdot p(x_i)$

↓ giustificazione

Y assume i valori $g(x_1), g(x_2), \dots$ con probabilità rispettive $p(x_1), p(x_2), \dots$

$\boxed{X \text{ con densità}}$: $g(X)$ ha valore atteso se $\int |g(x)| f(x)\, dx < +\infty$ e in tal caso $E[g(X)] = \int g(x) f(x)\, dx$

↓ Osservazione

X ha densità f $Y = g(X)$ ha densità? Supponiamo che sia vero: $E[g(X)] = \int_{\mathbb{R}} g(x) f(x)\, dx$ e, dato che $Y = g(X)$, $E[Y] = E[g(X)] = \int_{\mathbb{R}} y f_Y(y)\, dy$

sono $=$

Chiamiamo momento n-esimo $\mathbb{E}[X^n]$ se esiste

proprietà del valore atteso

1. $\boxed{\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]}$ } regola della somma
$\mathbb{E}[aX+b] = a\,\mathbb{E}[X] + b$

2. Se $X \geq 0$, $\mathbb{E}[X] \geq 0 \longrightarrow X \geq Y$ allora $\mathbb{E}[X] \geq \mathbb{E}[Y]$

Se X e Y hanno valore atteso, anche $X+Y$ ha valore atteso e $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

3. Supponiamo X e Y indipendenti. Allora $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$. Se non sono indipendenti non è detto che $X \cdot Y$ abbia valore atteso

↓ Dim 1 caso discreto

$\sum |x_i| \cdot p_X(x_i) < +\infty$ e $\sum |y_j| \cdot p_Y(y_j) < +\infty$. Devo provare che $\sum |x_i + y_j|\, p(x_i, y_j) < +\infty$

$$\sum_{x_i, y_j} |x_i + y_j|\, p(x_i, y_j) \leq \sum_{x_i, y_j} (|x_i| + |y_j|)\, p(x_i, y_j) = \sum_{x_i, y_j} |x_i|\, p(x_i, y_j) + \sum_{x_i, y_j} |y_j|\, p(x_i, y_j) =$$

$$= \underbrace{\sum_{x_i} |x_i| \underbrace{\sum_{y_j} p(x_i, y_j)}_{p_X(x_i)}}_{<+\infty} + \underbrace{\sum_{y_j} |y_j| \sum_{x_i} p(x_i, y_j)}_{<+\infty} \longrightarrow (X+Y) \text{ ha valore atteso}$$

togliendo i valori assoluti $(\leq \to =)$ e svolgendo i calcoli si arriva alla formula

varianza: $\mathrm{var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2 - 2X\mathbb{E}[X] + \mathbb{E}[X]^2] = \mathbb{E}[X^2] - 2\,\mathbb{E}[X]^2 + \mathbb{E}[X]^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

ESERCIZI:

Esercizio 2.11. Sia X una v.a. con densità uniforme sull'intervallo $[-1, 1]$: presa $Y = X^2$, calcolare densità (se esiste), media e varianza della variabile Y.

2 modi: 1. $\mathbb{E}[Y] = \mathbb{E}[X^2] = \int x^2 f_X(x)\,dx$
2. $\mathbb{E}[Y] = \int y\, f_Y(y)\,dy$

1. $\mathbb{E}[Y] = \mathbb{E}[X^2] = \int_{-1}^{1} \frac{x^2}{2}\,dx = \int_0^1 x^2\,dx = \frac{1}{3}$

$\mathbb{E}[Y^2] = \mathbb{E}[X^4] = \int_{-1}^{1} \frac{x^4}{2}\,dx = \int_0^1 x^4 = \frac{1}{5}$

$\mathrm{Var}(Y) = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = \frac{1}{5} - \frac{1}{9} = \frac{4}{45}$

2. $Y = X^2$ densità di Y

$$f_Y(y) = \begin{cases} 0 & y \leq 0 \\ f_X(\sqrt{y}) \cdot \frac{1}{2\sqrt{y}} + f_X(-\sqrt{y}) \cdot \frac{1}{2\sqrt{y}} & y > 0 \end{cases} = \begin{cases} 0 & y \leq 0 \\ \frac{1}{2\sqrt{y}} & 0 < y < 1 \\ 0 & y \geq 1 \end{cases}$$

$\mathbb{E}[Y] = \int y\, f_Y(y)\,dy = \int_0^1 \frac{y}{2\sqrt{y}}\,dy = \frac{1}{2}\int_0^1 y^{\frac{1}{2}}\,dy = \frac{1}{3}$

$\mathbb{E}[Y^2] = \int_0^1 \frac{y^2}{2\sqrt{y}}\,dy = \frac{1}{2}\int_0^1 y^{\frac{3}{2}}\,dy = \frac{y^{5/2}}{5}\Big|_0^1 = \frac{1}{5}$

Esercizio 2.1. Sia X una v.a. discreta che prende i valori 1, 2 e 3 con probabilità rispettive p_1, p_2 e p_3: sotto la condizione $\mathbf{E}[X] = 2$, per quali valori delle probabilità p_i la varianza è rispettivamente minima o massima?

$p_1 + p_2 + p_3 = 1 \longrightarrow \mathbb{E}[X] = p_1 + 2p_2 + 3p_3 = 2$

$$\begin{cases} p_1 + p_2 + p_3 = 1 \\ p_1 + 2p_2 + 3p_3 = 2 \end{cases} \longrightarrow -p_1 + p_3 = 0 : p_1 = p_3 = p : 0 \leq p \leq \frac{1}{2}$$

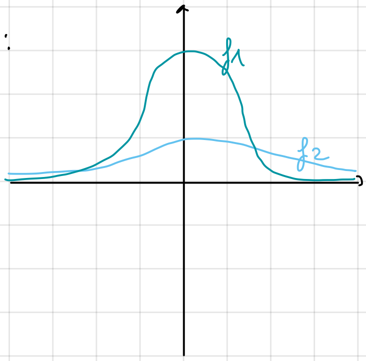
$p_2 = 1 - 2p$

$Var(X) = E[X^2] - E[X]^2 = E[X^2] - 4 \longrightarrow$ massimo della varianza = massimo del momento secondo

$$E[X^2] = 1p_1 + 4p_2 + 9p_3 = p + 4(1-2p) + 9p = 2p+4$$ e quindi $Var(X) = 2p$.

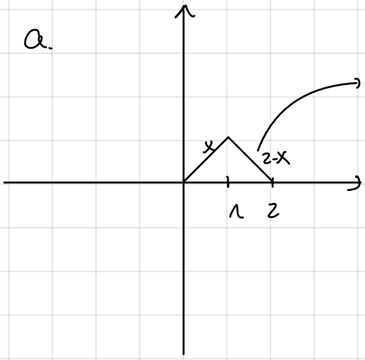
Il massimo sarà $p = \frac{1}{2}$ mentre il minimo con $p=0$

A livello intuitivo:



La densità f_2 ha varianza più grande

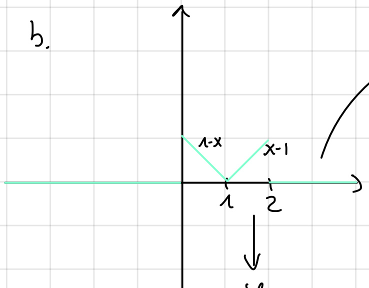
a.



intuitivamente $E[X]=1$

$$E[X] = \int_0^1 x^2 dx + \int_1^2 x(2-x)dx = \frac{1}{3} - \left(\frac{8}{3} - \frac{1}{3}\right) + 4 - 1 = 3 - 2 = 1$$

b.



$E[X] \underset{\text{intuit.}}{=} 1$

La varianza è più grande di a. in quanto i valori sono più dispersi

X Binomiale $B(n,p)$, esiste $E[X] = \sum_{h=0}^{n} h \binom{n}{h} p^h (1-p)^{n-h}$. Ma posso vedere $X = X_1 + \dots + X_n$ ove X_i Bernoulli per p. indipendenti. $E[X] = E[X_1] + \dots + E[X_n] =$
$= n\, E[X_1] = n \cdot p$

$1 \qquad 0$
$p \qquad 1-p$

Esercizio 2.3. In un concorso la prova scritta consiste in un test composto di 5 domande: per ciascuna di esse il test richiede di scegliere una tra 4 possibili risposte, delle quali una sola è corretta. Il test viene valutato attribuendo punteggio 1 ad ogni risposta esatta e $-\frac{1}{4}$ ad ogni risposta errata.

Si consideri un concorrente che risponda a caso ad ogni domanda: quale punteggio ottiene in media?

X "numero risposte esatte": $B(5, \frac{1}{4})$

punteggio $Y = X - \frac{1}{4}(5-X) = \frac{5}{4}X - \frac{5}{4}$

$$E[Y] = \frac{5}{4}E[X] - \frac{5}{4} = \frac{25}{16} - \frac{5}{4} = \frac{5}{16}$$

- MOMENTI, VARIANZA, COVARIANZA

Definizione: momenti $1 \le n < +\infty$ intero, allora $\exists$ il momento n-esimo se $\mathbb{E}[|X|^n] < +\infty$ e in tal caso il momento è $\mathbb{E}[X^n]$
$\downarrow$ 2 casi X

$\boxed{\text{Discreta}}$: $\exists \longrightarrow \sum_{x_i} |x_i|^n p(x_i) < +\infty$
$m \longrightarrow \sum_{x_i} x_i^n p(x_i)$

$\boxed{\text{Densità}}$: $\exists \longrightarrow \int_{-\infty}^{+\infty} |x|^n f(x) dx < +\infty$
$m \longrightarrow \int_{\mathbb{R}} x^n f(x) dx$

proprietà: $1 \le m < n$. Se X ha momento n-esimo, ha anche momento di ordine m. Se $\mathbb{E}[|X|^n] < +\infty$, allora $\mathbb{E}[|X|^m] < +\infty$. Si può dimostrare con la disuguaglianza di Jensen: $\mathbb{E}[|X|^m]^{1/m} \le \mathbb{E}[|X|^n]^{1/n}$ oppure si può usare un modo più rozzo
$\downarrow$
$x \in \mathbb{R}$, $|X|^m \le |X|^n + 1$. Quando l'esponente sale il numero cresce se $x > 1$, viceversa decresce. Se $|X| > 1$, $|X|^m < |X|^n$. Se $|X| < 1$, $|X|^m < 1$ e quindi vale la maggiorazione. Per le variabili con densità: $\mathbb{E}[|X|^m] = \int_{\mathbb{R}} |x|^m f(x) dx \le \underbrace{\int_{\mathbb{R}} |x|^n f(x) dx}_{< +\infty \text{ per Hp}} + \underbrace{\int_{\mathbb{R}} 1 \cdot f(x) dx}_{=1}$

esempi

1. $\boxed{\text{Densità di Cauchy}}$: $f(x) = \dfrac{1}{\pi(1+x^2)}$ è una densità. La v.a. X che ha questa densità non ha momento primo.
$\downarrow$ Dim. densità

$$\int_{-\infty}^{+\infty} \frac{1}{1+x^2} dx = \lim_{M \to +\infty} (\tan^{-1} M - \tan^{-1}(-M)) = \pi$$

$$\mathbb{E}[|X|] = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{|x|}{1+x^2} dx = \frac{2}{\pi} \int_0^{+\infty} \frac{x}{1+x^2} dx = \frac{1}{\pi} \log(1+x^2) \Big|_0^{+\infty} = +\infty$$

2. La funzione $f(x) = \begin{cases} 0 & x \le 1 \\ \frac{2}{x^3} & x > 1 \end{cases}$ è una densità. La v.a. X con questa densità ha momento primo ma non momento secondo.
$\downarrow$

$$a > 0: \int_1^{+\infty} \frac{1}{x^a} dx = \begin{cases} +\infty & 0 < a \le 1 \\ \frac{1}{a-1} & a > 1 \longrightarrow \text{primitiva } x^{-a} = (-a+1) \cdot x^{-a+1} \end{cases}$$

$$\mathbb{E}[|X|] = \int_1^{+\infty} \frac{2|x|}{|x|^3} dx = 2$$

$$\mathbb{E}[X^2] = \int_1^{+\infty} \frac{2x^2}{x^3} dx = +\infty$$

Osservazione (comoda): Se X discreta ha un numero finito di valori sicuramente ha tutti i momenti. Se X ha densità $f(x)$ nulla fuori da un intervallo sicuramente ha tutti i momenti.
$\downarrow$ Dim densità

$$\mathbb{E}[|X|^n] = \int_{\mathbb{R}} |x|^n f(x) dx = \underbrace{\int_a^b |x|^n f(x) dx}_{\emptyset \text{ altrove}} \longrightarrow a \le x \le b: |X|^n \le |a|^n + |b|^n$$ e quindi

$$\int \le (|a|^n + |b|^n) \int_a^b f(x) dx < +\infty$$

Disuguaglianze importanti

1. $\boxed{\text{Markov}}$: Sia Y a valori positivi, preso $a > 0$, $a \cdot P\{Y > a\} \leq \mathcal{E}[Y]$.

↓ Dim.

Definiamo $Z(\omega) = \begin{cases} a & \text{se } Y(\omega) > a \\ 0 & \text{altrimenti} \end{cases} \longrightarrow Z \leq Y \Rightarrow \mathcal{E}[Z] \leq \mathcal{E}[Y]$.

Avremo $\mathcal{E}[Z] = \cancel{0 \cdot P\{Y \leq a\}} + a \cdot P\{Y > a\} \leq \mathcal{E}[Y]$

2. $\boxed{\text{Schwartz}}$: In $\mathbb{R}^n$, presi $X = (x_1, \dots, x_n)$ e $Y = (y_1, \dots, y_n)$, $|X \cdot Y| \leq \|X\| \cdot \|Y\|$.

In sostanza, $\left| \sum_{i=1}^{n} x_i \cdot y_i \right| \leq \sqrt{\sum_{i=1}^{n} x_i^2} \sqrt{\sum_{i=1}^{n} y_i^2}$

↓ Si dimostra

$\mathcal{E}[|X| \cdot |Y|] \leq \sqrt{\mathcal{E}[X^2]} \cdot \sqrt{\mathcal{E}[Y^2]} \xrightarrow{\text{Conseguenza}}$ Se X e Y hanno momento secondo, $X \cdot Y$ ha momento primo

↓

Supponiamo che X, Y, Z abbiano momento

varianza: $\mathcal{E}[(X - \mathcal{E}[X])^2] = \mathcal{E}[X^2] - \mathcal{E}[X]^2$

↓ varianza empirica

$X = (x_1, \dots, x_n)$ e $\text{Var}_e(X) = \sum_{i=1}^{n} (x_i - \bar{x})^2 \cdot \frac{1}{n}$ che è la varianza di una v.a. che prende i valori $x_1, x_2, \dots, x_n$ con $p = \frac{1}{n}$.

Abbiamo inoltre che $\sqrt{\text{Var}(X)} = \sigma(X)$ [scarto quad. medio] e quindi $\text{Var}(X) = \sigma^2(X)$

↓ **Disuguaglianza di Chebyschev**

X v.a., $d > 0$: $P\{|X - \mathcal{E}[X]| > d\} \leq \frac{\text{Var}(X)}{d^2}$

↓ Dimostrazione

$Y = (X - E[X]^2)$, $a = d^2$ e utilizziamo Markov $\longrightarrow \mathcal{E}[Y] = \text{Var}(X) \geq a \cdot P\{Y > a\} = d^2 P\{\underbrace{(X - \mathcal{E}[X])^2 > d^2}_{\Updownarrow \text{ equivalenti}}\}$

$|X - \mathcal{E}[X]| > d$

Caso limite: $\text{Var}(X) = 0 \iff X \equiv c$

covarianza: $\text{Cov}(X, Y) = \mathcal{E}[(X - \mathcal{E}[X])(Y - E[Y])] = \mathcal{E}[X \cdot Y] - \mathcal{E}[X] \cdot \mathcal{E}[Y]$

↓ Conseguenze

1. $\text{Var}(X) = \text{Cov}(X, X)$
2. $\text{Cov}(aX + bY + c, Z) = a\,\text{Cov}(X, Z) + b\,\text{Cov}(Y, Z)$
3. $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
4. $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$

Sono equivalenti:

1. $\mathcal{E}[X \cdot Y] = \mathcal{E}[X] \cdot \mathcal{E}[Y]$
2. $\text{Cov}(X, Y) = 0$
3. $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$

} X, Y sono dette incorrelate

Variabili indipendenti sono incorrelate ma non vice-versa

↓

Coefficiente di correlazione: $\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X) \cdot \sigma(Y)} \longrightarrow$ occorre che siano $\neq 0$

↓ Utilità

$|\rho(X, Y)| \leq 1$ (da Schwartz): legame di natura lineare $\longrightarrow$ come per la retta di regressione: $\min_{(a,b) \in \mathbb{R}^2} \mathcal{E}[(Y - a - bX)^2] = \text{Var}(Y)[1 - \rho^2(X, Y)]$

ESEMPIO: Variabili incorrelate non indipendenti

X uniforme su $[-1, 1]$, $Z = 1 \vee -1$ con $P = \frac{1}{2}$ indip. da X, $Y = X \cdot Z$.

X e Y sono incorrelate $\longrightarrow \mathcal{E}[X] = 0$, $\mathcal{E}[Y] = \mathcal{E}[X \cdot Z] = \mathcal{E}[X] \cdot \mathcal{E}[Z] = 0$, $\mathcal{E}[XY] = \mathcal{E}[X^2 Z] = \mathcal{E}[X^2] \cdot \mathcal{E}[Z] = \frac{1}{3} \cdot \mathcal{E}[Z] = 0$

Non sono però indipendenti $\longrightarrow |X| = |Y|$: non possono essere indipendenti.

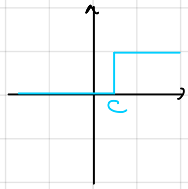
Solo le costanti sono indipendenti da se stesse. (Se so che la costante è 7 non aggiungo altro)

↓ Dim.

Sia X indip. da sé stessa. $F_X(t) = P(X\le t) = P(X\le t, X\le t) \underset{indip}{=} P(X\le t)\cdot P(X\le t) = F_X^2(t)$

$$\downarrow$$

$$\forall t,\quad F_X(t) = \begin{cases}0\\1\end{cases} \longrightarrow$$



è la cdf della costante c

Momenti di una $\boxed{\text{variabile gaussiana } N(0,1)}$ $\longrightarrow$ $X\sim N(0,1)$. X ha tutti i momenti $\xrightarrow{\text{Dim.}}$ $\forall n,\ \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}|x|^n e^{-x^2/2}\,dx < +\infty$: ✓

Se n dispari $(n=2h+1)$, $E[X^n]=0 \longrightarrow E[X^{2h+1}] = \int_{-\infty}^{+\infty} x^{(2h+1)} e^{-x^2/2}\,dx$: essendo dispari e integrabile, l'integrale è $=0$

$$\downarrow$$

$$E[X^2] = \text{var}(X) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2}}\,dx = \frac{1}{\sqrt{2\pi}}\left[\cancel{-x e^{-\frac{x^2}{2}}\Big|_{-\infty}^{+\infty}} + \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\,dx\right] = 1$$

Quindi $\longrightarrow$ $X\sim N(0,1)$ ha $E[X]=0$ e $\text{Var}(X)=1$

Prendiamo $Y\sim N(m,\sigma^2)$, $Y = m+\sigma X$ con X gaussiana standard. Otteniamo che $E[Y] = m+\sigma E[X] = m$ e $\text{Var}(Y) = \sigma^2\,\text{Var}(X) = \sigma^2$

Per n pari: $E[X^{n+2}] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} x^{n+2} e^{-\frac{x^2}{2}}\,dx = \frac{1}{\sqrt{2\pi}}\left[-e^{-\frac{x^2}{2}}\cdot x^{n+1}\Big|_{-\infty}^{+\infty} + (n+1)\int_{-\infty}^{+\infty} x^n e^{-\frac{x^2}{2}}\,dx\right] =$

$= (n+1)\,E[X^n]$

↓ Applicazione

$E[X^4] = 3E[X^3] = 3$

ESERCIZI

Esercizio 2.10. Sia X con densità gaussiana $N(3,4)$: calcolare $\mathbf{E}[X^4]$.

Possiamo scrivere $X = 3+2Z$ con Z gaussiana standard

$$\downarrow$$

$$E[(3+2Z)^4] = 81 + \cancel{4\cdot 27\cdot 2\cdot \underset{=0}{E[Z]}} + 6\cdot 3^2\cdot 2^2\,\underset{=1}{E[Z^2]} + \cancel{\ldots \underset{=0}{E[Z^3]}} + 16\,\underset{=3}{E[Z^4]}$$

∘ FUNZIONE GENERATRICE dei MOMENTI

Prendiamo X v.a. ⟶ funzione generatrice dei momenti $G_X: \mathbb{R} \to [0, +\infty[$:

$G_X(t) = E[e^{tX}]$ — X discreta: $\sum_{x_i} e^{tx_i} p(x_i)$; X densità: $\int_{-\infty}^{+\infty} e^{tx} f(x)\,dx$

Dominio di $G_X = \{ t \in \mathbb{R} \,/\, G_X(t) < +\infty \}$. Qualunque sia X, $G_X(0) = 1 \to E[e^{0X}] = E[1] = 1$.

Alcune volte $G_X(t)$ ⟶ 1 $t=0$; ⟶ $+\infty$ $t \neq 0$ } f.g.m. non serve a nulla

Esempio: densità di Cauchy

$f(x) = \frac{1}{\pi(1+x^2)}$

$t>0$: $G_X(t) = \int_{-\infty}^{+\infty} \frac{e^{tx}}{\pi(1+x^2)}dx \geq \int_0^{+\infty} \frac{\boxed{e^{tx}}}{\pi(1+x^2)}dx \geq \frac{t}{\pi}\int_0^{+\infty} \frac{x}{1+x^2}dx = +\infty$ ($1+tx+\ldots \geq tx$)

$t<0$: $G_X(t) = \int_{-\infty}^{+\infty} \frac{e^{tx}}{\pi(1+x^2)}dx \geq \int_{-\infty}^{0} \frac{\boxed{e^{tx}}}{\pi(1+x^2)}dx \geq \frac{|t|}{\pi}\int_{-\infty}^{0} \frac{|x|}{1+x^2}dx = +\infty$ ($e^{tx} \geq tx = |t|\cdot|x|$)

$\boxed{\text{Ipotesi: Dominio di } G_X(\cdot) \supsetneq \{0\}}$ (Altrimenti G_X inutile)

↓ Risultato

$G_X(\cdot) = G_Y(\cdot) \Longleftrightarrow F_X(\cdot) = F_Y(\cdot)$ cioè X e Y sono equidistribuite (a patto che il dominio non sia $\{0\}$)

↓

$\boxed{\text{Facili regole di calcolo}}$: 1. $G_{aX+b}(t) = G_X(at)\cdot e^{tb}$

2. Se X e Y sono indipendenti, $G_{X+Y}(t) = G_X(t)\cdot G_Y(t)$

↓ Verifica

1. $G_{aX+b}(t) = E[e^{t(aX+b)}] = e^{tb} \underbrace{E[e^{taX}]}_{G_X(at)}$

2. X e Y indipendenti, quindi e^{tX} ed e^{tY} sono indipendenti ⟶ $G_{X+Y}(t) = E[e^{t(X+Y)}] =$
$= E[e^{tX}\cdot e^{tY}] = E[e^{tX}]\cdot E[e^{tY}] = G_X(t)\cdot G_Y(t)$

Teorema: Supponiamo che il dominio di $G_X(t)$ contenga un intervallo $]-\varepsilon, \varepsilon[$. Allora X ha tutti i momenti e si ha $E[X^n] = \frac{d^n}{dx^n} G_X(t)\Big|_{t=0}$ (Di solito non conviene utilizzarla)

↓ Dim per $n=1$

$\frac{d}{dt} G_X(t) = \frac{d}{dt} E[e^{tX}] \underset{\text{anche se non dimostrato}}{=} E\left[\frac{d}{dt} e^{tX}\right] = E[Xe^{tX}]$ e per $t=0$ diventa $E[X]$.

Esempio: $\boxed{\text{variabile geometrica}}$ parametro p.

valori $1, 2, 3, \ldots$; $p(h) = (1-p)^{h-1} p$

$G_X(t) = \sum_{x_i} e^{tx_i} p(x_i) = \sum_{h=1}^{\infty} e^{th}(1-p)^{h-1} p = pe^t \sum_{h=1}^{\infty} [e^t(1-p)]^{h-1} = pe^t \sum_{k=0}^{\infty} ((1-p)e^t)^k$

converge se $(1-p)e^t < 1$ cioè $e^t < \frac{1}{1-p}$ cioè $t < \underbrace{-\log(1-p)}_{>0}$

⟶ $G_X(t) = \frac{pe^t}{1-e^t(1-p)}$ se $t \in]-\infty, -\log(1-p)[$ $(]-\varepsilon, \varepsilon[)$. X ha quindi tutti i momenti e quindi $E[X] = \frac{d}{dt} G_X(t)\Big|_{t=0} = \frac{1}{p}$

Avevamo detto: $X \sim N(m_1, \sigma_1^2)$ e $Y \sim N(m_2, \sigma_2^2)$ indipendenti ⟶ $(X+Y)$ è gaussiana $N(m_1+m_2, \sigma_1^2+\sigma_2^2)$

↓ Dimostriamo con fgm

$Z \sim N(0,1) \longrightarrow \boxed{G_Z(t) = e^{t^2/2}}$ e possiamo scrivere $X = m_1 + \sigma_1 Z$

↓

$G_X(t) = e^{tm_1}\cdot e^{\frac{\sigma_1^2 t^2}{2}} = \exp\left(tm_1 + \frac{\sigma_1^2 t^2}{2}\right)$

$G_Y(t) = \exp\left(tm_2 + \frac{\sigma_2^2 t^2}{2}\right)$

$\downarrow$ 2ª regola calcolo

$G_{X+Y}(t) = \exp\left(t(m_1+m_2) + \frac{(\sigma_1^2+\sigma_2^2)t^2}{2}\right)$ quindi per teorema di prima $(X+Y)$ è $N(m_1+m_2, \sigma_1^2+\sigma_2^2)$ [stessa fgm]

Calcoliamo • $\longrightarrow$ $G_Z(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{tx} e^{-\frac{x^2}{2}} dx = \frac{e^{\frac{t^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{\left(-\frac{t^2}{2} + tx - \frac{x^2}{2}\right)} dx =$

$= \frac{e^{\frac{t^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x-t)^2} dx \underset{x-t=y}{=} \frac{e^{\frac{t^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy = e^{\frac{t^2}{2}}$

Allo stesso modo se X di poisson per λ e Y di poisson per μ, $(X+Y)$ è di poisson di per $(\lambda+\mu)$

$\downarrow$ fgm

$G_X(t) = \sum_{h=0}^{\infty} e^{-\lambda} \frac{\lambda^h}{h!} e^{th} = e^{-\lambda} \sum_{h=0}^{+\infty} \frac{(\lambda e^t)^h}{h!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t-1)}$. X ha tutti i momenti (il dominio è $\mathbb{R}$).

$G_X(t)\, G_Y(t) = e^{\lambda(e^t-1)} e^{\mu(e^t-1)} = e^{(\lambda+\mu)(e^t-1)}$

$\downarrow$ Deriviamo

$G_X'(t) = e^{\lambda(e^t-1)} \lambda e^t \longrightarrow \left.\begin{array}{l} G_X'(0) = \lambda = \mathbb{E}[X] \\ G_X''(0) = \lambda+\lambda^2 = \mathbb{E}[X^2] \end{array}\right\}$ $Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \lambda+\lambda^2-\lambda^2 = \lambda$

Inoltre, X binomiale $B(n,p)$ [non conviene usare $G_X(\cdot)$]. Si considera $X = X_1 + \dots + X_n$ Bernoulli per p indipendenti

$\downarrow$

$\mathbb{E}[X_i] = p$, $Var(X_i) = p(1-p)$

$\mathbb{E}[X] = np$, $Var(X) = np(1-p)$

$\mathbb{E}[X^2] = Var(X) + \mathbb{E}[X]^2 = np(1-p) + n^2p^2$

Consideriamo inoltre X gaussiana standard $N(0,1)$

$\downarrow$

$\left.\begin{array}{l} n \text{ dispari},\ n = 2h+1 : \mathbb{E}[X^{2h+1}] = 0 \\ n \text{ pari},\ n = 2h+2 : \mathbb{E}[X^{2h+2}] = (2h+1)\,\mathbb{E}[X^{2h}] \end{array}\right\}$ $\mathbb{E}[X^2] = 1$, $\mathbb{E}[X^4] = 3$, $\mathbb{E}[X^6] = 5$

Per questa variabile non conviene usare la fgm per calcolare i momenti. Ci serve invece per dimostrare che la somma di due gaussiane indipendenti è gaussiana.

$\downarrow$ Applichiamo

X uniforme su $[a,b]$: $f(x) = \begin{cases} \frac{1}{b-a} & a<x<b \\ 0 & \text{altrove} \end{cases}$

X ha tutti i momenti : $\int_a^b \frac{x}{b-a} dx = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2} = \mathbb{E}[X]$

$Var(X) = \frac{(b-a)^2}{12} \longrightarrow \mathbb{E}[X^2] = \frac{1}{b-a} \int_a^b x^2 dx = \frac{b^3-a^3}{3(b-a)}$ e completando i conti si arriva a $Var(X)$

Esempio: $f(x) = \begin{cases} 0 & x \le 1 \\ \frac{2}{x^3} & x > 1 \end{cases}$. X ha momento primo ma non momento secondo

$\downarrow$ Chi è il dominio di $G_X(\cdot)$?

$G_X(t) = 2\int_1^{+\infty} \frac{e^{tx}}{x^3} dx = \begin{cases} +\infty & t>0 \\ 1 & t=0 \\ <+\infty & t<0 \end{cases}$ $\Big\}$ $D =]-\infty, 0]$ $]-\varepsilon, \varepsilon[$

ESERCIZI: **Esercizio 3.2.** Siano X_1, X_2 indipendenti, con distribuzione di Poisson di parametro λ,
a) Determinare per quale valore di λ si ha $\mathbf{E}\big[(X_1 - X_2)^2\big] = 1$,
b) Utilizzando il valore di λ determinato precedentemente, calcolare $\mathbf{P}\{X_1 + X_2 \geq 2\}$.

a. $1 = E[(X_1 - X_2)^2] = E[X_1^2 + X_2^2 - 2X_1X_2] = E[X_1^2] + E[X_2^2] - 2E[X_1]E[X_2] =$
$= 2(\lambda + \lambda^2) - 2\lambda^2 = 2\lambda \longrightarrow \lambda = 1/2$

b. $(X_1 + X_2)$ di poisson per 1. $Y = X_1 + X_2$

$P(Y \geq 2) = \sum_{h=2} P\{Y = h\} = \sum_{h=2}^{+\infty} e^{-1} \frac{\lambda^h}{h!} = 1 - P(Y<2) = 1 - P(Y=0) - P(Y=1) =$

$= 1 - \frac{1}{e} - \frac{1}{e} = 1 - \frac{2}{e}$

Esercizio 3.3. Siano X e Y indipendenti, con X binomiale B(2,1/5) ed Y con densità uniforme su $[0,2]$: posto $Z = \dfrac{X(Y-1)^2}{X+1}$, calcolare $\mathbf{E}[Z]$.

$E\left[\frac{X(Y-1)^2}{X+1}\right] = E\left[\frac{X}{X+1}\right] \cdot E[(Y-1)^2]$

Densità di $Y = \begin{cases} \frac{1}{2} & 0<y<2 \\ 0 & \text{altrove} \end{cases} \longrightarrow E[(Y-1)^2] = \frac{1}{2}\int_0^2 (y-1)^2\,dy = \frac{1}{2}\int_{-1}^{1} z^2\,dz = \frac{1}{3}$

$E\left[\frac{X}{X+1}\right] = 1 - E\left[\frac{1}{X+1}\right]$. X binomiale $(2, 0.2)$, valori $0 \quad 1 \quad 2$
$p(0) = 0.8^2 = 0.64$ $\downarrow \frac{1}{X+1}$
$p(1) = 2 \cdot 0.2 \cdot 0.8 = 0.32$ valori $1 \quad \frac{1}{2} \quad \frac{1}{3}$
$p(2) = 0.2^2 = 0.04$

$E\left[\frac{1}{X+1}\right] = 0.64 + \underset{(\frac{1}{2}\cdot 0.32)}{0.16} + \frac{0.04}{3}$

Osservazione: Se Z è uniforme su $[a,b] \longrightarrow \mathrm{Var}(Z) = E\left[\left(Z - \frac{a+b}{2}\right)\right] = \frac{(b-a)^2}{12}$

Y uniforme su $[0,2] \longrightarrow E[(Y-1)^2] = \mathrm{var}(Y) = \frac{4}{12} = \frac{1}{3}$

- DUE TEOREMI LIMITE FONDAMENTALI

Legge dei grandi numeri: $X_1, X_2, X_3, \dots$ v.a. indipendenti, equidistribuite con momento secondo. $\mu = E[X_i]$ (valore atteso comune) $\forall \varepsilon > 0$, $\lim_{n\to+\infty} P\left\{\left|\frac{X_1+\dots+X_n}{n} - \mu\right| > \varepsilon\right\} = 0$. La media aritmetica delle prime n variabili $(n \to +\infty)$ converge in probabilità al comune valore atteso

Esempio: Lanciamo una moneta con prob. p che esca testa. $X_i(\omega) = 1$ testa i-mo lancio, 0 croce i-mo lancio

Avremo quindi $\frac{X_1+\dots+X_n}{n} = \frac{\#\text{ teste in } n \text{ lanci}}{\#\text{ lanci}} \xrightarrow{\text{converge}}$ prob. p che esca testa

Teorema limite Centrale: X_1, X_2, X_3 v.a. indipendenti equidistribuite con momento secondo. $\mu = E[X_i]$, $\sigma^2 = Var(X_i) > 0$ $-\infty \le a < b \le +\infty$, allora $\lim_{n\to+\infty} P\left\{a \le \frac{X_1+\dots+X_n - n\mu}{\sigma\sqrt{n}} \le b\right\} =$

$= \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx = \phi(b) - \phi(a)$: convergenza in distribuzione

↓ in pratica

$X_1, \dots, X_n$ equidistribuite, n grande (>0), indipendenti $\longrightarrow \frac{X_1+\dots+X_n - n\mu}{\sigma\sqrt{n}}$ è approssimativamente gaussiana standard

Torniamo alla legge dei grandi numeri. Prendiamo $X_1, X_2, X_3, \dots$ **Definizione** Si dice che $(X_n)_{n=1}^{+\infty}$ converge in probabilità ad X se, $\forall \varepsilon > 0$, $\lim_{n\to+\infty} P\{|X_n - X| > \varepsilon\} = 0$

↓

Proposizione (Criterio sufficiente): supponiamo che $\lim_{n\to+\infty} E[X_n] = c$ e $\lim_{n\to+\infty} Var(X_n) = 0$, allora $(X_n)_{n\ge1}$ converge a c in probabilità

↓ Dim.

Chebyshef: $P\{|X - E[X]| > \varepsilon\} \le \frac{Var(X)}{\varepsilon^2} = \frac{E[(X-E[X])^2]}{\varepsilon^2}$ (definizione varianza con momenti)

$P\{|X_n - c| > \varepsilon\} \le \frac{\boxed{E[(X_n-c)^2]}}{\varepsilon^2}$

$E[(X_n-c)^2] = E[(X_n - E[X_n] + E[X_n] - c)^2] = \underbrace{E[(X_n - E[X_n])^2]}_{\to 0 \,:\, Var(X_n)} + \underbrace{(E[X_n]-c)^2}_{\to 0 \text{ Hp}} + \underbrace{\cancel{2(E[X_n]-c)E[X_n - E[X_n]]}}_{0}$

Vogliamo dimostrare $\frac{X_1+\dots+X_n}{n} = \overline{X_n}$ converge a μ in probabilità (Legge dei grandi numeri)

↓

$E[\overline{X_n}] = \frac{n\mu}{n} = \mu$

$Var(\overline{X_n}) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \longrightarrow 0$

} Per criterio prec. siamo apposto

⚠ Basta che siano incorrelate (varianza somma = somma varianze)
Non è necessario che siano equidistribuite: Basta che $E[X_i] = \mu$ e $Var(X_i) \le M$ $\forall i$.

↓ ⚠

Nel teorema limite centrale è essenziale che le variabili siano indipendenti ed equidistribuite

Applicazione: X binomiale $B(n,p)$ con n grande. $X = X_1 + \dots + X_n$ X_i Bernoulli per p $(0<p<1) \to E[X_i] = p$, $Var(X_i) = p(1-p)$

↓

$\frac{X - np}{\sqrt{np(1-p)}}$ è approssimativamente gaussiana standard

Quanto deve essere grande n?

$np(1-p) \geq 15$: approssimazione buona
$np(1-p) \geq 20$: approssimazione ottima $\longrightarrow$ p vicino a 1/2 e quindi $n \geq 80$

ESERCIZI:

Esercizio 3.6. Gli aerei della compagnia XYZ hanno 180 posti, ma la compagnia sa che in media quelli che si sono prenotati si presentano alla partenza con probabilità 0.9 : per questo la compagnia mette in vendita 195 biglietti (*overbooking*). Qual è la probabilità che qualche cliente rimanga a terra (e di conseguenza la compagnia sia costretta a un risarcimento)?

X binomiale $B(195, 0.9)$. Rimango quando $X \geq 181$

Calcoliamo $P\{X \geq 181\} \xrightarrow{\text{sol. esatta}} P\{X \geq 181\} = \sum_{h=181}^{195} \binom{195}{h} 0.9^h \, 0.1^{195-h}$ non praticabile

Approssimazione gaussiana: $\dfrac{X - 195 \cdot 0.9}{\sqrt{195 \cdot 0.9 \cdot 0.1}} = \dfrac{X - 175.5}{4.18} \sim N(0,1)$

Quindi $P(X \geq 181) = P\left(\dfrac{X-175.5}{4.18} \geq \dfrac{181-175.5}{4.18}\right) \simeq P(\underset{N(0,1)}{Z} \geq 1.31) = 1 - \phi(1.31) = 0.095$

Concentrazione di una variabile binomiale intorno ad un valore atteso

Lanciamo 100 volte una moneta equilibrata. Qual è la probabilità che il numero di teste uscite sia compreso tra 40 e 60?

$X \sim B(100, 0.5) \longrightarrow E[X] = 50$, $Var(X) = 25$

$P(40 \leq X \leq 60) = 1 - \boxed{P(|X-50| > 10)}$

$\longrightarrow$ Chebyshev: $P(|X - E[X]| > 10) \leq \dfrac{25}{100} = 0.25$ (⚠ solo maggiorazione)

Teorema limite centrale: $\dfrac{X-50}{5} \simeq$ gaussiana standard

$P(|X - E[X]| > 10) = P\left(\dfrac{|X - E[X]|}{5} > 2\right) \simeq P(|Z| > 2) =$

$= 2\int_2^{+\infty} \dfrac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 2(1-\phi(2)) = 2(1-0.97) = 0.06$

ESERCIZI:

Esercizio 3.7. Siano $X_1, X_2, \ldots, X_{50}$ indipendenti con densità esponenziale di parametro 2.

a) Stimare la probabilità $\mathbf{P}\{X_1 + \cdots + X_{50} \geq 32\}$

b) Calcolare $\mathbf{E}\left[X_1 \, e^{X_2}\right]$.

b. X_1, X_2 indip. esponenziali per $2 \longrightarrow E[X_1 \cdot e^{X_2}] = \boxed{E[X_1]} \cdot \boxed{E[e^{X_2}]}$

$\longrightarrow$ X esponenziale per λ: $E[X] = \int_0^{+\infty} x \, \lambda e^{-\lambda x} dx = \dfrac{1}{\lambda} \int_0^{+\infty} \lambda x \, e^{-\lambda x} \lambda dx_{\lambda x = t}$

$= \dfrac{1}{\lambda} \int_0^{+\infty} t e^{-t} dt = \dfrac{1}{\lambda}$

$\longrightarrow$ X_2 esponenziale per 2 : $E[e^{X_2}] = \int_{-\infty}^{+\infty} e^x f(x) dx = 2\int_0^{+\infty} e^x e^{-2x} dx = 2\int_0^{+\infty} e^{-x} dx = 2$

unendo

$E[X_1 e^{X_2}] = \dfrac{1}{2} \cdot 2 = 1$

a. Valutare $P\{X_1 + \ldots + X_{50} \geq 32\}$

X esp. per $\lambda \longrightarrow E[X] = \dfrac{1}{\lambda}$, $Var(X) = \dfrac{1}{\lambda^2}$

$\dfrac{X_1 + \ldots + X_{50} - 25}{\sqrt{12.5}} \sim$ gaussiana standard

$P(X_1 + \ldots + X_{50} \geq 32) = P\left(\dfrac{X_1 + \ldots + X_{50} - 25}{\sqrt{12.5}} \geq \underbrace{\dfrac{32-25}{\sqrt{12.5}}}_{\sim 1.98}\right) \simeq P(\underset{N(0,1)}{Z} > 1.98) = 1 - \phi(1.98) = 0.024$

Esercizio 3.9. Siano $X_1, \dots, X_{60}$ $\boxed{\text{i.i.d.}}$ con densità uniforme su $[0,2]$ e sia $\overline{X} = \frac{X_1+\dots+X_n}{n}$ la media empirica. indipendenti equidistribuite

a) Stimare la probabilità $\mathbf{P}\{\frac{5}{6} \le \overline{X} \le \frac{7}{6}\}$.

b) Stimare il più piccolo x per il quale si abbia $\mathbf{P}\{\overline{X} > x\} \le 0.05$.

$$\overline{X} = \frac{X_1 + \dots + X_{60}}{60} \qquad \mathbb{E}[\overline{X}] = 1$$

Y uniforme su $[a,b] \longrightarrow \operatorname{Var}(Y) = \frac{(b-a)^2}{12}$

$\operatorname{Var}(X_1) = \frac{4}{12} = \frac{1}{3} \longrightarrow P\left(\frac{5}{6} \le \overline{X} \le \frac{7}{6}\right) = P\left(|\overline{X} - \mathbb{E}[\overline{X}]| \le \frac{1}{6}\right) = 1 - P(|\overline{X} - \mathbb{E}[\overline{X}]| > \frac{1}{6})$.

Chebyshev: $P(|\overline{X} - \mathbb{E}[\overline{X}]| > \frac{1}{6}) \le \frac{\operatorname{Var}(\overline{X})}{6^2} = \frac{\frac{1}{180}}{36} =$

Approx. gaussiana: $\frac{X_1 + \dots + X_n - n\mathbb{E}[X_i]}{\sigma\sqrt{n}} = \frac{X_1 + \dots + X_{60} - 60}{\sqrt{60 \cdot \frac{1}{180}}} = \sqrt{3}(X_1 + \dots + X_{60} - 60)$ è gaussiana standard

∘ ALCUNE DENSITÀ di RILEVANTE INTERESSE in STATISTICA

Legge degli eventi rari: X_n binomiale $B(n, \frac{\lambda}{n})$. Allora $\lim_{n\to\infty} P\{X_n = h\} = e^{-\lambda}\frac{\lambda^h}{h!} =$

$= P\{X = h\}$ (poisson).

↓

X binomiale n "grande" e p "piccolo" è approssimativamente Poisson per parametro $np = \lambda$

Densità gamma (r,λ) $r>0$, $\lambda>0$ $[\Gamma(r,\lambda)]$: Introduciamo la funzione gamma di Eulero. Per $r>0$, $\Gamma = \int_0^{+\infty} x^{r-1} e^{-x} dx \longrightarrow \Gamma(1) = 1$ integrabile anche in 0 in quanto $r-1 > 1$ (esponente x). L'integrale ha una proprietà interessante: se $\Gamma(r) = (r-1)\Gamma(r-1)$

↓ Dim Integrazione per parti

$$\Gamma(r) = \int_0^{+\infty} x^{r-1} e^{-x} dx = \cancel{-e^{-x} x^{r-1}\Big|_0^{+\infty}} + (r-1)\int_0^{+\infty} x^{r-2} e^{-x} dx$$

↓ Conseguenze

n intero: $\Gamma(n) = (n-1)\Gamma(n-1) = (n-1)!$

Densità:

$$\boxed{f(x) = \begin{cases} \frac{1}{\Gamma(r)} \lambda^r x^{r-1} e^{-\lambda x} & x > 0 \\ 0 & x \le 0 \end{cases}}$$

$\xrightarrow{\text{Prova}} \int_0^{+\infty} \lambda^r x^{r-1} e^{-\lambda x} dx = \int_0^{+\infty} (\lambda x)^{r-1} e^{-\lambda x} \lambda dx = \int_0^{+\infty} t^{r-1} e^{-t} dt = \Gamma(r)$

⚠ $\Gamma(1,\lambda)$ è la densità esponenziale di parametro λ

Formula per i momenti: $X \sim \Gamma(r,\lambda)$ e $\beta > 0 \longrightarrow \boxed{\mathbb{E}[X^\beta] = \frac{\Gamma(r+\beta)}{\Gamma(r)\lambda^\beta}}$

↓ Applicazioni

$$\mathbb{E}[X] = \frac{\Gamma(r+1)}{\Gamma(r)\lambda} = \frac{r\,\cancel{\Gamma(r)}}{\cancel{\Gamma(r)}\lambda} = \frac{r}{\lambda}$$

$$\mathbb{E}[X^2] = \frac{\Gamma(r+2)}{\Gamma(r)\lambda^2} = \frac{(r+1)r}{\lambda^2}$$

$$\operatorname{Var}(X) = \frac{(r+1)r}{\lambda^2} - \frac{r^2}{\lambda^2} = \frac{r}{\lambda^2}$$

↓ Dimostrazione formula

$$\mathbb{E}[X^\beta] = \frac{1}{\Gamma(r)}\int_0^{+\infty} x^\beta \lambda^r x^{r-1} e^{-\lambda x} dx = \frac{1}{\Gamma(r)\lambda^\beta}\int_0^{+\infty} \lambda^{r+\beta} x^{r+\beta-1} e^{-\lambda x} dx = \frac{\Gamma(r+\beta)}{\Gamma(r)\lambda^\beta}$$

Funzione generatrice dei momenti: $G_X(t) = \begin{cases} \left(\frac{\lambda}{\lambda-t}\right)^r & t<\lambda \\ +\infty & t\geq\lambda \end{cases}$

Dim ↓

$$G_X(t) = \frac{1}{\Gamma(r)}\int_0^{+\infty} e^{tx}\lambda^r x^{r-1} e^{-\lambda x}dx = \frac{1}{\Gamma(r)}\int_0^{+\infty}\lambda^r x^{r-1} e^{-(\lambda-t)x}dx$$

→ $+\infty$ $t\geq\lambda$
→ converge $t<\lambda$

$$\longrightarrow \frac{\lambda^r}{(\lambda-t)^r}\cdot\frac{1}{\Gamma(r)}\underbrace{\int_0^{+\infty}(\lambda-t)^r x^{r-1}e^{-(\lambda-t)x}dx}_{\Gamma(r)} = \boxed{\left(\frac{\lambda}{\lambda-t}\right)^r}$$

↓ Applicazione

$X\sim\Gamma(r_1,\lambda)$ $Y\sim\Gamma(r_2,\lambda)$ indipendenti $\Rightarrow (X+Y)\sim\Gamma(r_1+r_2,\lambda)$

↓ verifica

$$G_{X+Y}(t) = G_X(t)\,G_Y(t) = \left(\frac{\lambda}{\lambda-t}\right)^{r_1}\left(\frac{\lambda}{\lambda-t}\right)^{r_2} = \left(\frac{\lambda}{\lambda-t}\right)^{r_1+r_2}$$

Densità chi-quadro a n gradi di libertà: chi-quadro (n) / $\chi^2(n)$. È la densità di $(X_1^2+\dots+X_n^2)$ ove $X_1,\dots,X_n$ sono gaussiane standard indipendenti $\longrightarrow C_n = X_1^2+\dots+X_n^2$.

↓

La densità chi-quadro (n) è in realtà la densità $\Gamma\left(\frac{n}{2},\frac{1}{2}\right)$. Questo è vero perché se $\boxed{X \text{ è } N(0,1),\ X^2 \text{ è } \Gamma\left(\frac12,\frac12\right)}$ quindi $X_1^2+\dots+X_n^2$ è $\Gamma\left(\frac{n}{2},\frac12\right)$

↓ Verifico

$X\sim N(0,1)$ $\quad Y = X^2 \longrightarrow$

$$\boxed{f_Y(y) = \begin{cases}0 & y<0\\ \frac{f_X(\sqrt y)}{2\sqrt y}+\frac{f_X(-\sqrt y)}{2\sqrt y} = \boxed{\frac{1}{\sqrt{2\pi}}}\, y^{-\frac12}e^{-\frac{y}{2}} & y>0\end{cases}}$$

Densità

$$\boxed{\Gamma\left(\tfrac12,\tfrac12\right) = \begin{cases}0 & x<0\\ \boxed{\frac{1}{\Gamma(\frac12)}}\cdot\frac{1}{\sqrt2}\cdot x^{-\frac12}e^{-\frac{x}{2}} & x>0\end{cases}}$$

↓

Necessariamente le costanti devono essere uguali perché le funzioni sono uguali e sono densità $\longrightarrow C_n(n)$ ha densità $\Gamma\left(\frac{n}{2},\frac12\right)$

↓

$C_n = X_1^2+\dots+X_n^2$ e supponiamo n grande. X_i^2 indipendenti, $E[X_i^2]=1$ $E[X_i^4]=3$, $var(X_i^2)=2$

$\xrightarrow{\text{grandi numeri}}$ $\frac{C_n}{n} = \frac{X_1^2+\dots+X_n^2}{n}\sim 1$

$\xrightarrow{\text{limite centrale}}$ $\frac{C_n-n}{\sqrt{2n}}\sim$ gaussiana standard

Densità di Student a n gradi di libertà T_n: È la densità di $T_n = \frac{X}{\sqrt{\frac{C_n}{n}}} = \sqrt{n}\cdot\frac{X}{\sqrt{C_n}}$

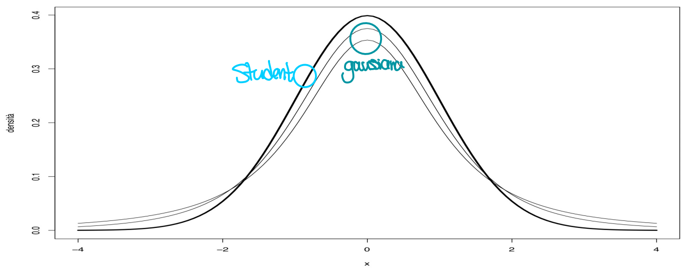
$X \to$ gaussiana standard
↕ indipendenti
$\sqrt{C_n} \to \chi^2(n)$

↓

È una funzione pari ($f(x)=f(-x)$). Supponiamo X con densità $f(x)$ e $-X$ con densità $f(-x)$. f è pari se e solo se X e $-X$ sono equidistribuite. Prendiamo $T_n = \frac{X}{\sqrt{C_n}}\cdot\sqrt n$ e $-T_n = \frac{-X}{\sqrt{C_n}}\cdot\sqrt n$

equidistribuite

n "grande"



$\frac{c_n}{n} \simeq 1 \longrightarrow T_n = \frac{X}{\sqrt{\frac{c_n}{n}}}$ è approssimativamente gaussiana standard

Avendo poi, indicando con $F_n(\cdot)$ la cdf di student (n), $F_n(-x) = 1 - F_n(x)$.
Avremo $\tau(\alpha, n)$ α-quantile variabile $T_n \longrightarrow \tau(\alpha, n) = -\tau(1-\alpha, n)$

ESERCIZI:

Esercizio 3.9. Siano $X_1, \ldots, X_{60}$ i.i.d. con densità uniforme su $[0,2]$ e sia $\overline{X} = \frac{X_1+\cdots+X_n}{n}$ la media empirica.
a) Stimare la probabilità $\mathbf{P}\{\frac{5}{6} \leq \overline{X} \leq \frac{7}{6}\}$.
b) Stimare il più piccolo x per il quale si abbia $\mathbf{P}\{\overline{X} > x\} \leq 0.05$.

a. $E[X_i] = 1 \qquad Var(X_i) = \frac{1}{3}$

↓ Teorema limite centrale

$\frac{X_1 + \ldots + X_n - n}{\sqrt{\frac{n}{3}}} \approx N(0,1)$. Dividiamo sopra e sotto per n

$\frac{\overline{X} - 1}{\sqrt{n/3/n}} = (\overline{X}-1)\sqrt{3n} = (\overline{X}-1)\sqrt{120} \approx N(0,1)$

$\left\{\frac{5}{6} \leq \overline{X} \leq \frac{7}{6}\right\} = \left\{-\frac{1}{6} \leq \overline{X}-1 \leq \frac{1}{6}\right\} = \left\{\frac{-\sqrt{120}}{6} \leq (\overline{X}-1)\sqrt{120} \leq \frac{\sqrt{120}}{6}\right\} =$

$= \{-\sqrt{5} \leq \underbrace{\boxed{(\overline{X}-1)\sqrt{120}}}_{N(0,1)} \leq \sqrt{5}\} \longleftrightarrow P\{-\sqrt{5} \leq Z \leq \sqrt{5}\} = \phi(\sqrt{5}) - \phi(-\sqrt{5}) = 2\phi(\sqrt{5}) - 1 =$
$= 2\phi(2.23) - 1 = 2 \cdot 0.987 - 1 = 0.974$

b. Se x cresce, $P\{\overline{X} > x\}$ diminuisce. Cerco quindi x tale che $P\{\overline{X} > x\} \simeq 0.05$.
Per il limite centrale, $(\overline{X}-1)\sqrt{120} \approx$ gaussiana standard $\longrightarrow P\{\overline{X} > x\} = \{(\overline{X}-1)\sqrt{120} > (x-1)\sqrt{120}\} \approx P\{Z > (x-1)\sqrt{120}\} = 0.05 \longrightarrow P\{Z \leq (x-1)\sqrt{120}\} = 0.95$ e quindi
$(x-1)\sqrt{120} = q_{0.95} \approx 1.645 \longrightarrow (x-1)\sqrt{120} \approx 1.645 \longrightarrow x = \frac{1.645}{\sqrt{120}} + 1 \approx 1.122$

Legge degli eventi rari: $X_n \sim B\left(n, \frac{\lambda}{n}\right) \qquad \lim_{n\to+\infty} P\{X_n = h\} = e^{-\lambda}\frac{\lambda^h}{h!}$

$P\{X_n = h\} = \binom{n}{h}\left(\frac{\lambda}{n}\right)^h\left(1-\frac{\lambda}{n}\right)^{n-h} = \frac{\lambda^h}{h!} \cdot \frac{n(n-1)\ldots(n-h+1)}{n \cdot n \cdot \ldots \cdot n} \cdot \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-h}$

↓ convergenza

$\frac{\lambda^h}{h!} \qquad 1 \qquad e^{-\lambda} \qquad 1 = e^{-\lambda} \cdot \frac{\lambda^h}{h!}$

ESERCIZI:

Esercizio 3.5. Calcolare i momenti primo e secondo di una variabile Binomiale di parametri n e p utilizzando la *funzione generatrice dei momenti*.

$X \sim B(n,p) \qquad E[X] = np \qquad Var(X) = np(1-p)$
$E[X^2] = np(1-p) + n^2p^2$

$G_X(t) = E[e^{tX}] = \sum_{h=0}^{n} e^{th}\binom{n}{h}p^h(1-p)^{n-h} = \sum_{h=0}^{n}\binom{n}{h}(pe^t)^h(1-p)^{n-h} =$
$= [pe^t + (1-p)]^n$

Esercizio 3.4. Si chiama variabile *lognormale* di parametri m e σ^2 una variabile Y della forma $Y = e^X$, dove X è gaussiana $N(m, \sigma^2)$.
a) Calcolare la funzione di ripartizione (c.d.f.) e la densità della variabile Y.
b) Calcolare il valore atteso $\mathbf{E}[Y]$ se $m = 0$ e $\sigma^2 = 3$.

$Y = e^X = e^{m+\sigma Z} \qquad Z \sim N(0,1)$

$F_Y(y) = \begin{cases} \phi\left(\frac{\log y - m}{\sigma}\right) & y > 0 \\ 0 & y \leq 0 \end{cases}$

$\longrightarrow P\{Y \leq y\} = P\{e^{m+\sigma Z} \leq y\} = P\{m + \sigma Z \leq \log y\} = P\left\{Z \leq \frac{\log y - m}{\sigma}\right\} =$

$$= \phi\left(\frac{\log y - m}{\sigma}\right)$$

$$f_Y(y) = \begin{cases} 0 & y \le 0 \\ \frac{1}{\sigma y} \cdot \varphi\left(\frac{\log y - m}{\sigma}\right) & y > 0 \end{cases}$$

$$f_Y(y) = \frac{d}{dy} \phi\left(\frac{\log y - m}{\sigma}\right) \qquad \text{ma} \quad \frac{d}{dz}\phi(z) = \varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

$$= \varphi\left(\frac{\log y - m}{\sigma}\right) \cdot \frac{1}{\sigma y}$$

b. $E[Y]$ se $m = 0$ e $\sigma = \sqrt{3}$

$Y = e^{\sqrt{3}X}$ $\quad X \sim N(0,1)$

$$\downarrow$$

$$E[Y] = E[e^{\sqrt{3}X}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{\sqrt{3}x} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2} + \sqrt{3}x} dx =$$

Vogliamo scrivere : $\frac{-x^2}{2} + \sqrt{3}x - \dots = -\frac{1}{2}(x-a)^2 = -\frac{x^2}{2} - \frac{a^2}{2} + ax \longrightarrow a = \sqrt{3}$

$$= \frac{e^{\frac{3}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2} - \frac{3}{2} + \sqrt{3}x} dx = \frac{e^{\frac{3}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x-\sqrt{3})^2} dx \underset{x-\sqrt{3}=t}{=} e^{\frac{3}{2}}$$

Inferenza statistica

Si parte da un campione per ricostruire informazioni sulla popolazione (ad esempio sondaggio elettorale) → Non sono mai la verità (anche ad esempio il controllo qualità)

↓ ipotesi

C'è un'implicita distribuzione di probabilità nella popolazione (sconosciuta o parzialmente sconosciuta) e le osservazioni $x_1, \dots, x_n$ sono la realizzazione di n variabili indipendenti $X_1, \dots, X_n$ con quella distribuzione di probabilità

Campione statistico: n variabili aleatorie $X_1, \dots, X_n$ con funzione di ripartizione F (sconosciuta)

↓

Abbiamo di dati concreti $x_1, \dots, x_n$ (realtà) ed un modello $X_1, \dots, X_n$ i.i.d. (mondo delle idee). I dati si interpretano come $X_1(\omega), \dots, X_n(\omega)$ ove $\omega \in \Omega$ che rappresenta tutti i possibili risultati.

- **STIMA**

Ipotesi di lavoro: l'altezza di uno studente di Informatica è una v.a. X gaussiana $N(m, \sigma^2)$ → Quale μ, σ^2?

Risultato preliminare: $X_1, \dots, X_n$ i.i.d. con momento secondo, $\overline{X} = \frac{X_1 + \dots + X_n}{n}$. Chiamiamo $S^2 = \sum_{i=1}^{m} \frac{(X_i - \overline{X})^2}{n-1}$: varianza campionaria. Supponiamo che le X_i abbiano media μ e varianza σ^2. Allora $E[\overline{X}] = \mu$, $E[S^2] = E\left[\sum_{i=1}^{m} \frac{(X_i - \overline{X})^2}{n-1}\right] = \sigma^2$. Media e varianza campionarie sono stime corrette di media e varianza.

↓ Verifica

Sappiamo $\overline{X}$, $E[\overline{X}] = \mu$ e $Var(\overline{X}) = \frac{\sigma^2}{n}$

↓ Uguaglianza fond.

$$\sum_{i=1}^{n} (X_i - \overline{X})^2 = \sum_{i=1}^{n} X_i^2 - n\overline{X}^2$$

$$E[X_i^2] = Var(X_i) + E[X_i]^2 = \sigma^2 + \mu^2$$

$$E[(\overline{X})^2] = \frac{\sigma^2}{n} + \mu^2$$

$$E[\sum (X_i - \overline{X})^2] = E[\sum X_i^2] - nE[(\overline{X})^2] = n(\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right) = (n-1)\sigma^2$$

↓

Avendo $X_1, \dots, X_n$, $g(X_1, \dots, X_n)$ è chiamata statistica campionaria.

- **PROBLEMI di STIMA PARAMETRICA**

Il campione $(X_1, \dots, X_n)$ funzione con cdf dipendenti da uno o più parametri θ.

ESEMPIO: Controllo qualità

n variabili di Bernoulli con parametro θ $(0 < \theta < 1)$

Affidabilità: la durata di vita di certe apparecchiature è esponenziale di parametro θ (>0)

2 **metodi di stima**: Di massima verosimiglianza e con metodo dei momenti

Stima di massima verosimiglianza $\hat{\theta}$

Le n variabili hanno densità $f(\theta, x)$. La verosimiglianza è data da $L(\theta; x_1, \dots x_n) = \prod_{i=1}^{n} f(\theta, x_i)$. La stima $\hat{\theta}$ (se esiste) è il punto di massimo della funzione $L \to L(\theta, x_1, \dots, x_n) \Longrightarrow L(\hat{\theta}, x_1, \dots, x_n) = \max_{\theta \in \Theta} L(\theta, x_1, \dots, x_n)$

Stima con metodo dei momenti $\tilde{\theta}$

Si scrivono i momenti teorici in dipendenza del parametro $\theta \longrightarrow E_\theta[X^k] = \int x^k f(\theta, x)dx$.

Si eguagliano con i momenti empirici $E_\theta[X^k] = \sum_{i=1}^{n} \frac{X_i^k}{n}$. Si esamina poi se da questa equazione si ricostruisce il parametro.

ESEMPIO: Campione con densità esponenziale parametro θ

1 $f(\theta, x) = \begin{cases} \theta e^{-\theta x} & x > 0 \\ 0 & x \leq 0 \end{cases}$ → Durata di vita: $x_1, \dots, x_n$ positivi

$\mathcal{L}(\theta, x_1, \dots, x_n) = \prod_{i=1}^{n} \theta e^{-\theta x_i} = \theta^n e^{-\theta(\sum_{i=1}^n x_i)}$

Analizziamo $\theta \to \mathcal{L}(\theta)$: valori positivi, regolare, tende a 0 a $(\mp\infty)$. Per il massimo annulliamo la derivata

$$0 = \frac{d}{d\theta}\mathcal{L}(\theta, x_1, \dots, x_n) = n\theta^{n-1} e^{-\theta(\sum x_i)} - \sum x_i)\theta^n e^{-\theta(\sum x_i)} =$$

$$= n - (\sum x_i)\theta = 0 \longrightarrow \hat{\theta} = \frac{n}{\sum_{i=1}^n x_i} = \frac{\cancel{n}}{\cancel{n}\bar{X}} = \frac{1}{\bar{X}}$$

$\hat{\theta} = (\bar{X})^{-1}$

Con il metodo dei momenti: $E_\theta[X_i] = \frac{1}{\theta} = \frac{\sum x_i}{n} \longrightarrow \tilde{\theta} = \frac{1}{\bar{X}} = (\bar{X})^{-1}$

ESEMPIO: 2

$X_1, \dots, X_n$ campione con densità uniforme su $[0, \theta]$, $0 < \theta < +\infty$.

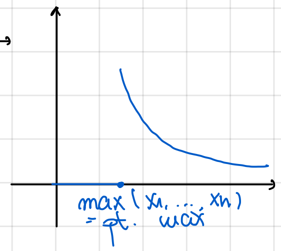
$f(\theta, x) = \begin{cases} \frac{1}{\theta} & 0 < x < \theta \\ 0 & \text{altrove} \end{cases} \longrightarrow X_1, \dots X_n$ positivi: devono stare nell'intervallo

Metodo dei momenti: $E_\theta[X_i] = \frac{\theta}{2} \longrightarrow \frac{\theta}{2} = \sum_{i=1}^n \frac{x_i}{n} \longrightarrow \tilde{\theta} = 2\bar{X}$

Verosimiglianza: $\mathcal{L}(\theta, x_1, \dots, x_n) = \begin{cases} \theta^{-n} & \text{se ogni } x_i \le \theta \\ 0 & \text{se } \exists\ x_i > \theta \end{cases}$

$\mathcal{L}(\theta, x_1, \dots, x_n) = \begin{cases} \theta^{-n} & \text{se } \theta \ge \max(x_1, \dots, x_n) \\ 0 & \text{altrimenti} \end{cases} \longrightarrow$

$\hat{\theta} = \max(x_1, \dots x_n)$



ESERCIZI:

Esercizio 4.1. Consideriamo, per $\theta \in \mathbb{R}$, la funzione

$$f_\theta(x) = \begin{cases} C_\theta(2x + \theta) & 0 < x < 1 \\ 0 & \text{altrove} \end{cases}$$

a) Per quali valori di θ e C_θ la funzione sopra scritta è una densità di probabilità?
b) Scrivere esplicitamente la c.d.f., media e varianza di una v.a. che ha la densità sopra scritta.

a. $\theta \ge 0$ per essere sempre positiva oppure $\theta \le -2$ (sempre negativa)

$\int_0^1 2x + \theta \, dx = 1 + \theta \longrightarrow C_\theta = \frac{1}{1+\theta}$

b. c.d.f. $\longrightarrow F_\theta(x) = \int_{-\infty}^{x} f_\theta(t)\,dt$ ma dal grafico della densità si vede $F_\theta(t) = \begin{cases} 0 & x \le 0 \\ \bullet & \longrightarrow \frac{x^2 + \theta x}{1+\theta} \quad 0 < x < 1 \\ 1 & x \ge 1 \end{cases}$

$\longrightarrow F_\theta(X) = \frac{1}{1+\theta}\int_0^x (2t+\theta)\,dt = \frac{x^2 + \theta x}{1+\theta}$

$E_\theta[X] = \frac{1}{1+\theta}\int_0^1 x(2x+\theta)\,dx = \left(\frac{2}{3} + \frac{\theta}{2}\right)\cdot\frac{1}{1+\theta}$

$E_\theta[X^2] = \frac{1}{1+\theta}\int_0^1 x^2(2x+\theta)\,dx = \frac{1}{1+\theta}\left(\frac{1}{2} + \frac{\theta}{3}\right)$

$Var_\theta(X) = E_\theta[X^2] - E_\theta[X]^2$

Esercizio 4.2. Scrivere esplicitamente la c.d.f. e calcolare il momento primo di una variabile che abbia densità

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

al variare di θ in $\mathbb{R}$.

Considerato poi un campione $X_1, \dots, X_n$ avente la densità sopra scritta, indagare sulla stima di θ col metodo della massima verosimiglianza e col metodo dei momenti.

Campione: n durate di vita con inizio sconosciuto

Verifica densità: $\int_{-\infty}^{+\infty} f_\theta(x)\,dx = \int_\theta^{+\infty} e^{-(x-\theta)}\,dx \underset{x-\theta=t}{=} \int_0^{+\infty} e^{-t}\,dt = 1$ ✓

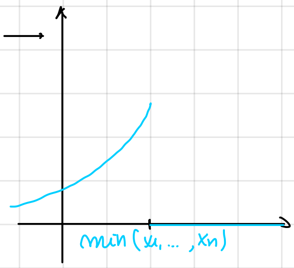
$$F_\theta(x) = \begin{cases} 0 & x \leq \theta \\ 1 - e^{-(x-\theta)} & x > \theta \end{cases}$$

$\rightarrow \int_\theta^x e^{-(t-\theta)}\,dt = e^\theta \int_\theta^x e^{-t}\,dt = e^\theta(-e^{-x} + e^{-\theta}) = 1 - e^{-(x-\theta)}$

$E_\theta[X] = \int_\theta^{+\infty} x e^{-(x-\theta)}\,dx \underset{x-\theta=t}{=} \int_0^{+\infty} e^{-t}(t+\theta)\,dt = 1+\theta$

Stima di $\tilde{\theta}$: $(1+\theta) = \sum_{i=1}^n \frac{x_i}{n} = \overline{X} \longrightarrow \tilde{\theta} = \overline{X} + 1$

$$\mathcal{L}(\theta; x_1, \dots, x_n) = \begin{cases} e^{-\sum x_i + n\theta} & \boxed{\text{se ogni } x_i \text{ è } \geq \theta} \\ 0 & \text{altrimenti} \end{cases} \longrightarrow$$



$\Longrightarrow \theta \leq \min(x_1, \dots, x_n)$

$\hat{\theta} = \min(x_1, \dots, x_n)$

$\boxed{\text{proprietà dei campioni gaussiani}}$

$X_1, \dots, X_n$ indipendenti gaussiane

Caso 1: $\boxed{\text{Gaussiane standard}}$, $X_i \sim N(0,1)$

Risultato principale:
a. $\overline{X}$ e $\sum_{i=1}^n (X_i - \overline{X})^2$ sono indipendenti (solo le gaussiane)
b. $\overline{X}$ ha densità $N(0, \frac{1}{n})$ e $\sum_{i=1}^n (X_i - \overline{X})^2$ ha densità chi-quadro $(n-1)$
c. $T = \sqrt{n}\,\frac{\overline{X}}{S}$ ha densità di Student $T(n-1)$.

⚠ Perché si parla di gradi di libertà? $X_1, \dots X_n$ $\sum (X_i - \overline{X})^2$ chi-quadro $n-1$.
Se $\overline{X}$ è noto, $X_1, \dots, X_{n-1}$ liberi e X_n vincolato

Dimostrazione: Bisogna provare $\overline{X}$ e $\sum_{i=1}^n (X_i - \overline{X})^2$ indipendenti e $\sum_{i=1}^n (X_i - \overline{X})^2$ chi-quadro $(n-1)$ $\longrightarrow$ Non lo facciamo
+
T di Student $T(n) \longrightarrow T = \frac{X}{\sqrt{C_n/n}}$ C_n chi-quadro (n). Prendiamo $\sqrt{n}\,\frac{\overline{X}}{S} =$

$= \dfrac{(\sqrt{n}\,\overline{X})}{\sqrt{\dfrac{\sum (X_i - \overline{X})^2}{n-1}}} \Big\} \sqrt{\dfrac{C_{n-1}}{n-1}}$ $\quad \sqrt{n}\,\overline{X}$ è $N(0,1)$ $\longrightarrow$ Risultato utile per caso generale

Caso 2: $\boxed{\text{Gaussiane generali}}$ $N(m, \sigma^2)$

$X_i = m + \sigma Y_i$, Y_i gaussiane standard indipendenti

Risultato generale:
a. $\overline{X}$ e $\sum (X_i - \overline{X})^2$ sono indipendenti
b. $\overline{X}$ ha densità $N\left(m, \frac{\sigma^2}{n}\right)$ e $\sum_{i=1}^n \frac{(X_i - \overline{X})^2}{\sigma^2}$ ha densità

Chi-quadro $(n-1)$

c. $T = \sqrt{n}\,\dfrac{(\overline{X}-m)}{S}$ ha densità di Student $T(n-1)$

⚠ σ è sparito dalla formula

Si vede subito che il risultato segue dal precedente: $X_i = m + \sigma Y_i$. $\overline{X} = m + \sigma \overline{Y}$. $\sum (X_i - \overline{X})^2 = \sigma^2 \sum (Y_i - \overline{Y})^2$. Infine se prendo $T = \sqrt{n}\,\dfrac{(\overline{X}-m)}{S} =$

$$= \frac{\sqrt{n}\,\sigma\,\overline{Y}}{\sqrt{\dfrac{\sum (Y_i-\overline{Y})^2\sigma^2}{n-1}}} = \frac{\sqrt{n}\,\overline{Y}}{\sqrt{\dfrac{\sum (Y_i-\overline{Y})^2}{n-1}}}$$

ESERCIZI:

Esercizio 4.3. Preso $\theta > 0$, si consideri la funzione

$$f_\theta(x) = \begin{cases} c_\theta\, x^\theta & 0 < x < \theta \\ 0 & \text{altrimenti} \end{cases}$$

a) Dopo aver calcolato la costante c_θ che rende la funzione sopra scritta una densità, calcolare, per una v.a. X che ha quella densità, $\mathbf{P}\{X > \theta/2\}$.

b) Se $\theta = 2$, e $X_1, \dots, X_{100}$ sono indipendenti aventi quella densità, stimare $\mathbf{P}\{X_1 + \cdots + X_{100} \geq 157\}$.

c) Determinare una stima di θ col *metodo dei momenti.*

a. $\displaystyle\int_0^\theta x^\theta\,dx = \left.\frac{x^{\theta+1}}{\theta+1}\right|_0^\theta = \frac{\theta^{\theta+1}}{\theta+1} \longrightarrow c_\theta = \frac{\theta+1}{\theta^{\theta+1}}$

$$P\left(X > \frac{\theta}{2}\right) = \frac{\theta+1}{\theta^{\theta+1}} \int_{\theta/2}^{\theta} x^\theta\,dx = \frac{\cancel{\theta+1}}{\theta^{\theta+1}} \cdot \left.\frac{x^{\theta+1}}{\cancel{\theta+1}}\right|_{\theta/2}^{\theta} = \frac{\theta^{\theta+1}}{\theta^{\theta+1}} - \frac{\left(\frac{\theta}{2}\right)^{\theta+1}}{\theta^{\theta+1}} =$$
$$= 1 - \frac{1}{2^{\theta+1}}$$

$$E_\theta[X] = \frac{\theta+1}{\theta^{\theta+1}} \int_0^\theta x^{\theta+1}\,dx = \frac{\theta+1}{\theta^{\theta+1}} \cdot \frac{\theta^{\theta+2}}{\theta+2} = \frac{\theta(\theta+1)}{\theta+2}$$

b. $\theta = 2 \longrightarrow f(x) = \begin{cases} \frac{3}{8}x^2 & 0 \le x < 2 \\ 0 & \text{altrove} \end{cases}$

$$E[X] = \frac{3}{8}\int_0^2 x^3\,dx = \frac{3}{8} \cdot \frac{16}{4} = \frac{3}{2}$$

$$E[X^2] = \frac{3}{8}\int_0^2 x^4\,dx = \frac{3}{8} \cdot \frac{32}{5} = \frac{12}{5}$$

$$\mathrm{Var}(X) = \frac{12}{5} - \frac{9}{4} = \frac{3}{20}$$

$P\{X_1 + \dots + X_{100} \geq 157\} \longrightarrow X_i$ indip. $E[X_i] = \frac{3}{2}$, $\mathrm{var}(X_i) = \frac{3}{20}$

$$\frac{X_1 + \dots + X_{100} - 150}{\sqrt{\frac{3}{20}\cdot 100}} = \frac{X_1 + \dots + X_{100} - 150}{\sqrt{15}} \sim N(0,1)$$

$$X_1 + \dots + X_{100} \geq 157 \iff \frac{X_1 + \dots + X_{100} - 150}{\sqrt{15}} \geq \frac{157-150}{\sqrt{15}} = \frac{7}{\sqrt{15}} \approx 1{,}807$$

$$P\{X_1 + \dots + X_{100} \geq 157\} \approx P\{Z \geq 1.807\} = 1 - \phi(1.807) = 1 - 0.964 = 0.036$$

c. $E_\theta[X] = \sum_{i=1}^{n} \frac{x_i}{n}$ $\qquad E_\theta(x) = \frac{\theta(\theta+1)}{\theta+2} = \overline{X} \longrightarrow \overline{X} = \frac{x_1 + \dots + x_n}{n} > 0$

$\theta \geq \overline{X}$ perché ogni valore x_i ha valore $[0,\theta]$

$\theta^2 + \theta = \theta\overline{X} + 2\overline{X} \longrightarrow \theta^2 + \theta(1-\overline{X}) - 2\overline{X} = 0 \longrightarrow \theta_{1,2} = \frac{(\overline{X}-1) \pm \sqrt{(1-\overline{X})^2 + 8\overline{X}}}{2} =$

$= \frac{(\overline{X}-1) \pm \sqrt{\overline{X}^2 + 6\overline{X} + 1}}{2} \longrightarrow \tilde{\theta}$ deve essere $> \overline{X}$, quindi scarto la soluzione con il $-$

$> \sqrt{(\overline{X}+1)^2} = \overline{X}+1$

$\downarrow$

$\tilde{\theta} = \frac{(\overline{X}-1) + \boxed{\sqrt{\overline{X}^2 + 6\overline{X} + 1}}}{2} > \frac{\overline{X} - 1 + \overline{X} + 1}{2} = \overline{X}$ e

quindi è ammissibile con il metodo dei momenti

Esercizio 4.6. Consideriamo, per $\boxed{\theta > -1}$, la funzione

garantisce positività

$$F_\theta(x) = \begin{cases} 0 & x \leq 0 \\ \frac{x^2+\theta x}{1+\theta} & 0 < x < 1 \\ 1 & x \geq 1 \end{cases}$$

a) Determinare per quali valori di θ questa è la funzione di ripartizione di una variabile con densità.
b) Sono stati osservati i seguenti dati numerici: 0.35, 0.40, 0.31, 0.85, 0.59, 0.60. Immaginando che questi dati siano l'osservazione di un campione statistico con la c.d.f. sopra scritta, dare una stima di θ col metodo dei momenti.

$F_\theta(\cdot)$ deve essere crescente, o $f_\theta(x) = F'_\theta(x)$ deve essere a valori positivi

$\downarrow$

$f_\theta(x) = \frac{2x+\theta}{1+\theta}$ $\quad 0 \leq x \leq 1$ e quindi deve essere $\theta \geq 0$

$E_\theta[X] = \frac{1}{1+\theta} \int_0^1 x(2x+\theta)\,dx = \frac{1}{1+\theta}\left(\frac{2}{3} + \frac{\theta}{2}\right) = \frac{4+3\theta}{6(1+\theta)}$

$E_\theta[X^2] = \frac{1}{1+\theta} \int_0^1 x^2(2x+\theta)\,dx = \frac{1}{1+\theta}\left(\frac{1}{2} + \frac{\theta}{3}\right) = \frac{3+2\theta}{6(1+\theta)}$

Avremo $\overline{X} = 0.516$. Stimiamo θ col metodo dei momenti

$\overline{X} = \frac{4+3\theta}{6(1+\theta)} \longrightarrow \tilde{\theta} = \frac{4-6\overline{X}}{6\overline{X}-3} \cong 9.41$

- INTERVALLI di FIDUCIA (confidenza)

Proiezioni elettorali $\longrightarrow$ puntiamo sul partito A:
Alle 22 mi danno $21.7 \pm 2.7\%$
Alle 23 mi danno $21.6 \pm 1.7\%$
Alle 24 mi danno $21.8 \pm \underbrace{0.4\%}_{\text{forchetta di previsione}}$

] intervalli di fiducia

Ω = tutti i possibili esiti, Θ = parametri del modello, P_Θ = prob. su Ω.
Supponiamo $\Theta \subseteq \mathbb{R}$

Definizione: Fissiamo $0<\alpha<1$. Si dice che I è un intervallo di fiducia per il parametro θ al livello $1-\alpha$ se, $\forall \theta$: $P_\theta\{\theta \in I\} \geq 1-\alpha$ (oppure $P_\theta\{\theta \notin I\} \leq \alpha$)

$\downarrow$

$I = I(\omega)$: risultato indagine statistica

Esempio: intervallo di fiducia per la media di un campione gaussiano

Caso 1: varianza nota

$X_1, \dots, X_n$ indipendenti $\sim (m, \sigma^2)$ ove σ^2 noto e m da trovare

$\overline{X} = \frac{X_1 + \dots + X_n}{n}$: stima della media

Ci aspettiamo $I(\omega) = [\overline{X}(\omega) - d, \overline{X}(\omega) + d]$. Quale d? Ideale d più piccolo possibile mantenendo alta la fiducia (contrapposizione)

Dobbiamo provare $P_m\{\boxed{m \in I}\} \geq (1-\alpha)$ con d più piccolo possibile

$\longrightarrow m \in I \Longleftrightarrow |X(\omega) - m| \leq d \longrightarrow P_m\{|\overline{X}-m| \leq d\} \geq (1-\alpha)$. Supponiamo

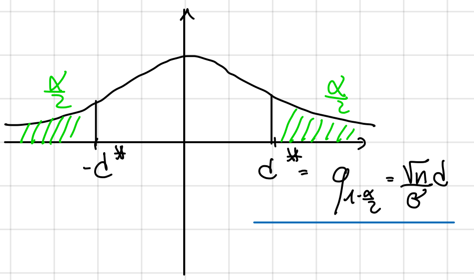
$P_m\{|\overline{X}-m| \leq d\} \cong (1-\alpha)$ (per avere d più piccolo possibile)

$\downarrow$

$\sqrt{n}\,\frac{(\overline{X}-m)}{\sigma}$: gaussiana standard $N(0,1)$

$\downarrow$ l'uguaglianza diventa

$P_m\{\underbrace{\frac{\sqrt{n}}{\sigma}|\overline{X}-m|}_{N(0,1)} \leq \frac{\sqrt{n}}{\sigma} \cdot d\} \cong (1-\alpha) \longrightarrow$



$\phi(d) = \int_{-\infty}^{d^*} \varrho(x)\,dx = 1-\frac{\alpha}{2}$

$d = \frac{\sigma}{\sqrt{n}}\, q_{1-\frac{\alpha}{2}}$

$\overline{X}(\omega) \pm \frac{\sigma}{\sqrt{n}}\, q_{1-\frac{\alpha}{2}}$ (intervallo) $\longrightarrow$ Più aumentano i dati, più si restringe. Al contrario, se σ si alza l'intervallo si ingrandisce

$\downarrow$

$\boxed{\frac{\sigma}{\sqrt{n}}\, q_{1-\frac{\alpha}{2}} \text{ : precisione della stima}}$

$\boxed{\frac{\frac{\sigma}{\sqrt{n}}\, q_{1-\frac{\alpha}{2}}}{|\overline{X}(\omega)|} \text{ : precisione relativa}}$

Caso 2: varianza sconosciuta

$\downarrow$

Student sostituisce σ^2 con $S^2 = \sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{n-1}$: stima corretta varianza

Intervallo di fiducia: $\bar{X}(\omega) \pm \frac{S(\omega)}{\sqrt{n}} \, T(1-\frac{\alpha}{2}, n-1)$ Quantile variabile $T(n-1)$

Quando n "grande" $T(n) \approx$ gaussiana standard e quindi
$T(1-\frac{\alpha}{2}, n-1) \approx q_{1-\frac{\alpha}{2}}$

Il livello di fiducia più usuale è al $95\% = 0.95 = 1-\alpha \longrightarrow \alpha = 0.05$
Sarò quindi guardare $q_{0.975} = 1.96$: Ricetta per l'intervallo di fiducia: $\bar{X}(\omega) \pm \frac{1.96}{\sqrt{n}} S(\omega)$

Ma cosa facciamo con la formula $\bar{X}(\omega) \pm \frac{S(\omega)}{\sqrt{n}} T_{(1-\frac{\alpha}{2}, n-1)}$? In concreto $(x_1, \dots, x_n) = \omega$ sono le osservazioni (rilevazioni statistiche). Si prende $\bar{x} = \frac{x_1 + \dots + x_n}{n}$

e $s = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n-1}}$ e otteniamo $\bar{x} + \frac{s}{\sqrt{n}} T_{(1-\frac{\alpha}{2}, n-1)}$. Utilizziamo le tavole per i quantili della variabile di Student con $\beta \longrightarrow 1$. Questo perché $T_{(\beta, n)} = -T_{(1-\beta, n)}$

Intervalli di fiducia unilateri $( ]-\infty, \dots], \quad [\dots, +\infty[\,)$ la media non è troppo alta

Per media: sinistro $]-\infty, \bar{X}(\omega) + d]$
destro $[X(\omega) - d, +\infty[$

Ottenimento: Partiamo da σ nota al livello $1-\alpha$

sinistro $= \left] -\infty, \bar{X}(\omega) + \frac{\sigma}{\sqrt{n}} q_{1-\alpha} \right]$

destro $= \left[\bar{X}(\omega) - \frac{\sigma}{\sqrt{n}} q_{1-\alpha}, +\infty \right[= \left[\bar{X}(n) + \frac{\sigma}{\sqrt{n}} q_{\alpha}, +\infty \right[$

Analoghi con varianza sconosciuta

Intervallo di fiducia approssimato per il parametro di un campione di Bernoulli

$X_1, \dots, X_n$ n "grande" Bernoulli per p

$\frac{X_1 + \dots + X_n - np}{\sqrt{np(1-p)}} = \frac{\sqrt{n}(\bar{X} - p)}{\sqrt{p(1-p)}}$ approx. gaussiana standard ⚠ Non conosciamo $p \longrightarrow$ mettiamo una stima dipendente dalle osservazioni: $\hat{p} = \bar{X}$: percentuale osservata

Sostituendo

Intervallo di fiducia bilatero: $\bar{X} \pm \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \cdot q_{1-\frac{\alpha}{2}} = \hat{p} \pm \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \cdot q_{1-\frac{\alpha}{2}}$

sostituzione $\hat{p} = \bar{X}$

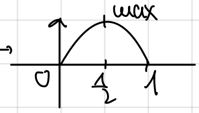
Avremo $\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} q_{1-\frac{\alpha}{2}}$: precisione (approx.) della stima

Esempio iniziale: $21.7 \pm \boxed{2.9}\%$ da cui $\hat{p} = 0.217\%$: stima % partito A
$0.029 = \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} q_{1-\frac{\alpha}{2}}$

ESEMPIO: Voglio calcolare la % di gradimento del governo con una precisione inferiore all'1% con fiducia al 95%. Quante telefonate almeno devo fare?

$1-\alpha = 0.95$. Precisione $\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} q_{0.975} \le 0.01$ da cui $\sqrt{n} \ge \frac{\sqrt{\hat{p}(1-\hat{p})} \cdot 1.96}{0.01}$

Non conosco $\hat{p}$: consideriamo $\max_{0<p<1} p(1-p) \longrightarrow$



Quindi $n \geq \left(\frac{1.96}{0.02}\right)^2 = 9604$: telefonate necessarie

intervallo unilatero per la varianza

risultato: $\sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{\sigma^2} = (n-1)\frac{S^2}{\sigma^2}$ ha densità $\chi^2(n-1)$

↓ Intervallo unilatero sinistro

$$\left]0, \frac{\sum (X_i-\bar{X})^2}{a}\right] = I \quad \{\sigma^2 \in I\} = \left\{\sigma^2 \leq \frac{\sum (X_i-\bar{X})^2}{a}\right\} = \left\{a \leq \frac{\sum (X_i-\bar{X})^2}{\sigma^2}\right\}$$

a → c.te da determinare

notiamo che $P_\sigma\left\{a \leq \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma^2}\right\}$ non dipende da σ

↓

vogliamo che sia $\geq 1-\alpha$ o meglio, per rendere l'intervallo più piccolo, $= 1-\alpha$. Si ottiene se $a = \chi^2(\alpha, n-1)$ α-quantile della variabile $\chi^2(n-1)$

↓ concludendo

sinistro: $\left]0, \frac{\sum (X_i-\bar{X})^2}{\chi^2(\alpha, n-1)}\right]$ destro: $\left[\frac{\sum (X_i-\bar{X})^2}{\chi^2(1-\alpha, n-1)}, +\infty\right[$

- **TEST o VERIFICA di IPOTESI**

Significa formulare un'ipotesi e pianificare un esperimento per decidere se accettare l'ipotesi (con un certo grado di fiducia).

↓

Formalizzazione ipotesi $\Theta = \Theta_0 \cup \Theta_1$

Θ_0: parametri Hp; Θ_1: parametri alternativa

ESEMPIO. La % di pezzi difettosi è inferiore al 2%

$\Theta =]0,1[$ ove θ è la prob. che il pezzo sia difettoso

$\Theta_0 =]0, 0.02]$

$\Theta_1 =]0.02, 1[$

↓

Si usa dire $\underbrace{H_0)\ \theta \leq 0.02}_{\text{ipotesi (ipotesi nulla)}}$ contro $\underbrace{H_1)\ \theta > 0.02}_{\text{alternativa (ipotesi alternativa)}}$

Serve poi pianificare l'esperimento: scegliere risultati che portano a rifiutare l'ipotesi ⟶ Scegliere

$C \subseteq \Omega$: regione critica

$A = C^c$: regione di accettazione

↓ Nel caso degli int. di fiducia

$X_1 \dots X_n$ Bernoulli par θ

$C = \{\omega \in \Omega \mid \bar{X}(\omega) > d\} = \{\bar{X} > d\}$ → dipende dal grado di fiducia

$A = \{\omega \in \Omega \mid \bar{X}(\omega) \leq d\} = \{\bar{X} \leq d\}$

Abbiamo 2 tipi di errore: 1. di prima specie: Rifiuto hp quando è corretta
2. seconda specie: Accetto hp quando è falsa

↓ **Definizione**

il test è di livello α se $\forall \theta \in \Theta_0$, $P_\theta(C) \leq \alpha$ (tipici 0.05, 0.02, 0.01)

Definizione: Si chiama potenza del test la funzione, definita su Θ_1, $\theta \to P_\theta(C)$ per $\theta \in \Theta_1$. È quindi la capacità di accorgersi che l'ipotesi non è soddisfatta.

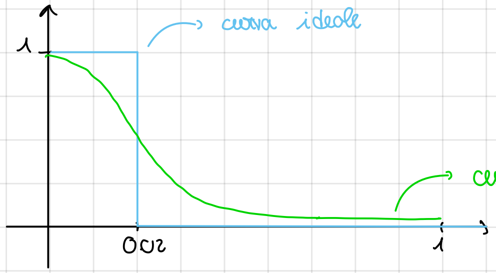
↓

Ideale: livello basso e potenza alta. In teoria dell'affidabilità si usa il linguaggio della $\underbrace{\text{curva operativa}}$

Funzione, definita su tutto Θ, $\theta \longrightarrow P_\theta(A) = 1 - P_\theta(C)$. Essa dovrebbe essere molto vicina a 1 su Θ_0 e molto vicina a 0 su Θ_1

↓ esempio

$H_0)\ \theta \leq 0.02$ $H_1)\ \theta > 0.02$



→ Regione critica, curva operativa etc. si calcolano prima di rilevare i dati

p-value

Considerazione: se si abbassa il livello la regione critica diventa più piccola ed è più facile accettare l'ipotesi

Definizione: Dopo aver osservato i dati, si chiama p-value il numero $\bar{\alpha}$ tale che se $\alpha < \bar{\alpha}$ l'ipotesi è accettata al livello α
se $\alpha > \bar{\alpha}$ l'ipotesi è rifiutata al livello α

↓

L'elemento discriminante misura la plausibilità dell'ipotesi. Se p-value è basso (0.17... 0.09) l'Ip è poco plausibile mentre se il p-value è alto (0.32...0.51) l'Ip è molto plausibile

↓ Osservazioni

1. Il p-value non ha senso da solo ma relativamente ad un'ipotesi
2. Il p-value si calcola dopo aver fatto l'esperimento

↓ **Definizione alternativa**

Probabilità che (alla luce dei dati) il rifiuto sia dovuto al caso, non "strutturale"

ESEMPI: 1. Test sulla media di un campione gaussiano con varianza nota (z-test)

↓

Test bilatero: $H_0) m = m_0$ e $H_1) m \neq m_0$ su $X_1, \dots, X_n$ gaussiane $N(m, \underbrace{\sigma^2}_{\text{fissa e conosciuta}})$

↓

Regione critica: $C = \{|\bar{X} - m_0| > d\}$. A livello α: $P_{m_0}\{|\bar{X} - m_0| > d\} \leq \alpha$
Vogliamo aumentare la potenza, quindi C più grande possibile purché di livello α ⟶ $P_{m_0}\{|\bar{X} - m_0| > d\} \cong \alpha$

↓ Ris. hanno

$\frac{\sqrt{n}}{\sigma}(\bar{X} - m_0)$ gaussiana standard

$$\alpha = P_{m_0}\left\{\frac{\sqrt{n}}{\sigma}|\bar{X} - m_0| > \frac{\sqrt{n}}{\sigma} d\right\} = P\left\{|Z| > \frac{\sqrt{n}}{\sigma} d\right\} = \alpha$$

Quindi $\frac{\sqrt{n}}{\sigma} d = q_{1-\frac{\alpha}{2}}$ da cui arriviamo a $C = \left\{|\bar{X} - m_0| > \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}}\right\}$

che è la regione critica al livello α.

⚠ Ricordiamo la formula dell'intervallo di fiducia per la media al livello $1-\alpha$ ⟶ $\bar{X} \pm \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}}$

↓

Conseguenza: L'ipotesi $H_0) m = m_0$ è accettata al livello α se m_0 appartiene all'intervallo di fiducia al livello $1-\alpha$.

↓

In concreto abbiamo $x_1, \dots, x_n$ (dati) e $\bar{x} = \frac{x_1 + \dots + x_n}{n}$. Rifiutiamo l'Ip se $|\bar{x} - m_0| > \frac{\sigma}{n} q_{1-\frac{\alpha}{2}}$ o se $\frac{\sqrt{n}}{\sigma}|\bar{x} - m_0| > q_{1-\frac{\alpha}{2}} = P\{|Z| > \alpha\}$

Calcolo del p-value: $P_{m_0}\left\{\frac{\sqrt{n}}{\sigma}|\bar{X} - m_0| > \frac{\sqrt{n}}{\sigma}|\bar{x} - m_0|\right\} \leq P_{m_0}\left\{\frac{\sqrt{n}}{\sigma}|\bar{X} - m_0| > q_{1-\frac{\alpha}{2}}\right\} = \alpha$

Si rifiuta al livello α se $\boxed{\alpha > P_{m_0}\left\{\underbrace{\frac{\sqrt{n}}{\sigma}|\bar{X}-m_0|}_{Z \text{ del modello}} > \underbrace{\frac{\sqrt{n}}{\sigma}|\bar{x}-m_0|}_{z \text{ osservato}}\right\}}$

Si accetta al livello α se $\boxed{\alpha < P_{m_0}\left\{\frac{\sqrt{n}}{\sigma}|\bar{X}-m_0| > \frac{\sqrt{n}}{\sigma}|\bar{x}-m_0|\right\}}$

$\bar{\alpha}$ elemento discriminante: $\bar{\alpha} = P_{m_0}\left\{\frac{\sqrt{n}}{\sigma}|\bar{X}-m_0| > \frac{\sqrt{n}}{\sigma}|\bar{x}-m_0|\right\}$ = Prob. che il rifiuto sia stato fatto per errore

Calcolo esplicito

$$\bar{\alpha} = P\left\{|Z| > \frac{\sqrt{n}}{\sigma}|\bar{x}-m_0|\right\} \underset{\text{dal disegno}}{=} 2\left(1-\phi\left(\frac{\sqrt{n}}{\sigma}|\bar{x}-m_0|\right)\right)$$

⚠ Promemoria: **regione critica** $= \{|\bar{X}-m_0| > d\}$

p-value $= P_{m_0}\{|\bar{X}-m_0| > |\bar{x}-m_0|\} = P_{m_0}\left\{\frac{\sqrt{n}}{\sigma}|\bar{X}-m_0| > \frac{\sqrt{n}}{\sigma}|\bar{x}-m_0|\right\}$

$$= P\left\{|Z| > \underbrace{\frac{\sqrt{n}}{\sigma}|\bar{x}-m_0|}_{\text{dati osservati}}\right\}$$

⚠ dati osservati ⚠

Calcolo della curva operativa (al livello α): $\beta(m) = P_m\{A\} = P_m\left\{\frac{\sqrt{n}}{\sigma}|\bar{X}-m_0| \le q_{1-\frac{\alpha}{2}}\right\} =$

$$= P_m\left\{-q_{1-\frac{\alpha}{2}} \le \frac{\sqrt{n}}{\sigma}(x-m_0) \le q_{1-\frac{\alpha}{2}}\right\} =$$

$$= P_m\left\{\frac{\sqrt{n}(m_0-m)}{\sigma} - q_{1-\frac{\alpha}{2}} \le \frac{\sqrt{n}}{\sigma}(\bar{X}-m) \le \frac{(m_0-m)\sqrt{n}}{\sigma} + q_{1-\frac{\alpha}{2}}\right\} =$$

$$= \phi\left(\frac{m_0-m}{\sigma}\sqrt{n} + q_{1-\frac{\alpha}{2}}\right) - \phi\left(\frac{m_0-m}{\sigma}\sqrt{n} - q_{1-\frac{\alpha}{2}}\right)$$

2. Test unilatero

$H_0)\ m \le m_0$ $\qquad H_1)\ m > m_0$

$C = \{(\bar{X}-m_0) > d\}$ di livello α

$\forall m \le m_0,\ P_m\{(\bar{X}-m_0) > d\} \le \alpha$ cerco con m: $P_m\{\bar{X}-m > (m_0-m)+d\}$. Basta avere $P_{m_0}\{(\bar{X}-m_0) > d\} \approx \alpha \rightarrow P\left\{\frac{\sqrt{n}}{\sigma}(\bar{X}-m_0) > \frac{\sqrt{n}}{\sigma}d\right\} \approx \alpha$ da cui $\frac{\sqrt{n}}{\sigma}d = q_{1-\alpha}$.

$C = \left\{\frac{\sqrt{n}}{\sigma}(\bar{X}-m_0) > q_{1-\alpha}\right\}$ di livello α

Calcolo del p-value: Regione critica $\{(\bar{X}-m_0) > d\}$ e quindi il p-value è $P_{m_0}\left\{\frac{\sqrt{n}}{\sigma}(\bar{X}-m_0) \ge \frac{\sqrt{n}}{\sigma}(\bar{x}-m_0)\right\} = 1-\phi\left(\frac{\sqrt{n}}{\sigma}(\bar{x}-m_0)\right)$

Calcolo della curva operativa: $\beta(m) = P_m\left\{\frac{\sqrt{n}}{\sigma}(\bar{X}-m_0) \le q_{1-\alpha}\right\} =$

$$= P_m\left\{\frac{\sqrt{n}}{\sigma}(\bar{x}-m) \le \frac{(m_0-m)}{\sigma}\sqrt{n} + q_{1-\alpha}\right\} =$$

$$= \phi\left(\frac{m_0-m}{\sigma}\sqrt{n} + q_{1-\alpha}\right)$$

Se l'hp fosse $H_0)\ m \ge m_0$ $\quad H_1)\ m < m_0$ si ribaltano $>$ e $\le$ e $q_{1-\alpha}$ diventa q_α

3. $\boxed{\text{Test sulla media di un campione gaussiano con varianza sconosciuta}}$
(T- test)

H0) $m = m_0$, σ^2 qualsiasi, H1) $m \neq m_0$, σ^2 qualsiasi

Avremo $\frac{\sqrt{n}\,(\bar{X} - m)}{S}$ di student $T(n-1)$

$S \longrightarrow \sqrt{S^2} = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}}$

Regione critica di livello α: $C = \left\{ |\bar{X} - m_0| > \frac{S}{\sqrt{n}}\, \tau\left(1-\frac{\alpha}{2}, n-1\right) \right\}$

p-value: $P_{m_0}\left\{ \underbrace{\frac{\sqrt{n}\,|\bar{X} - m_0|}{S}}_{T\text{ modello}} > \underbrace{\frac{\sqrt{n}\,|\bar{x} - m_0|}{S}}_{T\text{ osservata}} \right\} = 2\left[1 - \underbrace{F_{n-1}\left(\frac{\sqrt{n}}{s}|\bar{x} - m_0|\right)}_{\text{c.d.f. variabile } T(n-1)}\right]$

Quando n è "grande" ($n > 70$) $F_n(\cdot) \simeq \phi(\cdot)$

Curva operativa: Non ha senso

$\beta_m = P_m\{A\}$. Se $m = m_0$, $\frac{\sqrt{n}\,(\bar{X} - m_0)}{S}$ è di Student

ma se $m \neq m_0$, $\frac{\sqrt{n}\,(\bar{X} - m_0)}{S}$ ha una densità che

dipende da m e σ e quindi non ha senso

4. $\boxed{\text{Test approssimato su un campione di Bernoulli}}$

$X_1, \dots, X_n$ par. p. H0) $p = p_0$ e H1) $p \neq p_0$.

Se $p = p_0$, $\frac{\sqrt{n}\,(\bar{X} - p_0)}{\sqrt{p_0(1-p_0)}}$ è approssimativamente gaussiana standard

Regione critica livello α: $C = \left\{ \frac{\sqrt{n}\,|\bar{X} - p_0|}{\sqrt{p_0(1-p_0)}} > q_{1-\frac{\alpha}{2}} \right\}$

p-value: $P_{p_0}\left\{ \frac{\sqrt{n}\,|\bar{X} - p_0|}{\sqrt{p_0(1-p_0)}} > \frac{\sqrt{n}\,|\hat{p} - p_0|}{\sqrt{p_0(1-p_0)}} \right\} \simeq 2\left[1 - \phi\left(\frac{\sqrt{n}\,|\hat{p} - p_0|}{\sqrt{p_0(1-p_0)}}\right)\right]$

$\hat{p} = \frac{x_1 + \dots + x_n}{n}$: proporzione osservata

Curva operativa: Si può calcolare ma è una grande complicazione

5. $\boxed{\text{Test unilatero sulla varianza di un campione gaussiano}}$

$X_1, \dots, X_n$ indipendenti $N(\underbrace{m}_{\text{sconosciute}}, \sigma^2)$. $\sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{\sigma^2}$ è $\chi^2(n-1)$ qualunque

sia la media.

H0) $\sigma^2 \leq \sigma_0^2$ e H1) $\sigma^2 > \sigma_0^2$ al livello α

Regione critica: $C = \left\{ \sum (X_i - \bar{X})^2 > d \right\}$

cond. sul livello

$\forall\, \sigma \leq \sigma_0$, $P_\sigma\left\{ \sum (X_i - \bar{X})^2 > d \right\} \leq \alpha$ ovvero con α perché

$P_\sigma\left\{ \frac{\sum (X_i - \bar{X})^2}{\sigma^2} > \frac{d}{\sigma^2} \right\}$ e quindi basta

verificare: $P_{\sigma_0}\left\{ \underbrace{\frac{\sum (X_i - \bar{X})^2}{\sigma_0^2}}_{\chi^2(n-1)} > \frac{d}{\sigma_0^2} \right\} \simeq \alpha$

$\frac{d}{\sigma_0^2} = \chi^2(1-\alpha, n-1)$ da cui $C = \left\{ \frac{\sum (X_i - \bar{X})^2}{\sigma_0^2} > \chi^2(1-\alpha, n-1) \right\}$

p-value : $P_{\sigma_0}\left\{ \frac{\sum (X_i - \bar{X})^2}{\sigma_0^2} > \frac{\sum (x_i - \bar{x})^2}{\sigma_0^2} \right\} = 1 - \underbrace{G_{n-1}}_{\text{c.d.f. } \chi^2(n-1)}\left(\frac{\sum (x_i - \bar{x})^2}{\sigma_0^2} \right) =$

$= 1 - G_{n-1}\left(\frac{var(x)(n-1)}{\sigma_0^2}\right)$

Curva operativa: livello α $\beta(\sigma) = G_{n-1}\left(\frac{\sigma_0^2}{\sigma^2} \cdot \chi^2(1-\alpha, n-1)\right)$

- CONFRONTO TRA DUE CAMPIONI

campioni accoppiati

Esercizio 6.3.1. Viene propagandata una cura dimagrante che promette *7 chili in 7 giorni*: vengono misurati *"prima e dopo"* 7 pazienti ottenendo i valori che seguono

a) 72 89 94 77 86 91 83

b) 68 83 90 71 80 88 74

Si può accettare l'affermazione dell'istituto che offre la cura oppure si deve concludere che *hanno esagerato*? Quale sarebbe stata la conclusione se invece di promettere 7 chili in 7 giorni avessero promesso solo 6 chili in 7 giorni?

$X_i \sim N(m_1, \sigma_1^2)$ $i = 1, \dots, 7$
$Y_i \sim N(m_2, \sigma_2^2)$

$Z_i = X_i - Y_i \sim N(m_1 - m_2, \sigma_1^2 + \sigma_2^2)$

$H_0)$ $(m_1 - m_2) \geq 7$ e $H_1)$ $m_1 - m_2 < 7$

$Z_i \to$ campione con media e varianza sconosciuta e facciamo test $H_0)$ $m \geq 7$ e $H_1)$ $m < 7$ con varianza sconosciuta: test di Student.

$T = \frac{\sqrt{n}\,(\bar{Z} - 7)}{S}$. Calcoliamo il p-value di hp. La regione critica è della forma $\{T < d\}$ e quindi il p-value sarà $P_7\{T < \underset{-2.09}{t_{oss}}\}$. T è di student a 6 gradi di libertà e quindi il p-value sarà $F_6(t) = 0.04$: hp da scartare. Alternativa

$H_0)$ $m \geq 6$ $H_1)$ $m < 6$. $T = \frac{\sqrt{7}\,(\bar{X} - 6)}{S}$ $t_{oss} = -0.76$ e quindi

$F_6(t) = 0.23$: regione rilevante plausibile

Art

$H_0)$ $m \geq 5$ $H_1)$ $m < 5$ $\longrightarrow$ $t_{oss} = 0.57$, $F_6(t) = 0.7$: hp da accettare

⚠ problemi di calcolo con le variabili di Student

Quando n è grande ($n \geq 60$ o meglio $n \geq 30$) si fa l'approssimazione gaussiana: $T(n) \approx Z \sim N(0,1)$.

Quando n piccolo? Non possono esistere tavole per $F_n(t)$ cdf per $T(n)$. Si possono calcolare solo con un software. Esistono le tavole dei quantili per la variabile di Student. Che ci danno $T_{(\beta, n)}$ per alcuni valori di n per alcuni valori di β vicini a 1 (⚠ vale la proprietà $T_{(1-\beta, n)} = -T_{(\beta, n)}$)

Esempi

$T_{(0.95, 6)} = 1.94$ e quindi $T_{(0.05, 6)} = -1.94$

Si rifiuta l'ipotesi al livello 0.05 se $T_{oss} < T_{(0.05, 6)} = -1.94$

Tornando all'esempio

Caso $H_0)$ $m \geq 7$ $\longrightarrow$ $t_{oss} = -2.09$ quindi si rifiuta. p-value = 0.04 (livello a cui si accetta)

Caso $H_0)$ $m \geq 6$ $\longrightarrow$ $t_{oss} = -0.76$ quindi si accetta l'ipotesi al livello 0.05. p-value = 0.23 e quindi concorda con il fatto che si accetta campioni indipendenti

Esercizio 6.3.2. Vengono misurate le lunghezze delle tibie di uomini adulti provenienti da reperti di tombe etrusche: dal sito di Cerveteri vengono effettuate 13 misurazioni ottenendo un valore medio di 47.2 ed una varianza campionaria di 7.98 , mentre dal sito di Ladispoli si ottengono 8 misurazioni con un valore medio di 44.9 ed una varianza campionaria di 8.85.

Si può affermare che la differenza sia una semplice fluttuazione statistica oppure si deve concludere che gli abitanti di Cerveteri erano veramente più alti?

Risultato teorico: $X_1, \dots, X_n \sim N(m_1, \sigma^2)$, $Y_1, \dots, Y_n \sim N(m_2, \sigma^2)$ } La varianza è sconosciuta ma eguale per entrambi i campioni

$$T_{n,k} = \frac{(\bar{X}-\bar{Y})-(m_1-m_2)}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2+\sum_{j=1}^{k}(Y_j-\bar{Y})^2}} \cdot \frac{\sqrt{n+k-2}}{\sqrt{\frac{1}{n}+\frac{1}{k}}}$$ è di Student $T(n+k-2)$

Es. Si considera questa e si applica il test di student

Cerveteri: 13 tibie media 47.2 varianza 7.98 } Dovrei supporre che le varianze siano uguali

Ladispoli: 8 tibie media 44.9 varianza 8.85

Guardando i dati empiricamente concludiamo che i dati sono uguali

$m = m_1 - m_2 \longrightarrow H_0)\ m=0$ avevano in media la stessa altezza $H_1)\ m>0$ Gli abitanti di Cerveteri erano più alti

$$T = \frac{\bar{X}-\bar{Y}-m}{\sqrt{\sum(X_i-\bar{X})^2+\sum(Y_j-\bar{Y})^2}} \cdot \frac{\sqrt{19}}{\sqrt{\frac{1}{13}+\frac{1}{8}}}$$ supponendo $m=0$

$$t_{oss} = \frac{47.2-44.9}{\sqrt{7.98\cdot 12+8.85\cdot 7}} \cdot \frac{\sqrt{19}}{\sqrt{\frac{1}{13}+\frac{1}{8}}} = 1.776$$

p-value: regione critica $= \{T>d\} \longrightarrow$ p-value $= P_0\{T>t_{oss}\} = P\{\underset{\text{Student } T(19)}{T}>1.776\} =$
$= 1-F_{19}(1.776) = 0.045$: hp molto poco plausibile

Quelli di Cerveteri erano veramente più alti

Cosa ci aspettiamo a livello 0.05? Ci aspettiamo rifiuto dell'ipotesi.
E al livello 0.025? Accettazione

Si rifiuta l'ipotesi α se t_{oss} è maggiore di $T(1-\alpha, 19)$. $T(0.95, 19) = 1.72$. $t_{oss} > 1.72$ quindi si rifiuta. Invece, $T(0.975, 19) = 2.09$ e $t_{oss} \le T$ quindi si accetta

Si concluderebbe che il p-value è compreso tra 0.05 e 0.025

ESERCIZI:

Esercizio 5.2. Una barra di metallo di produzione industriale deve essere lunga 40 cm, tuttavia sono state effettuate 400 misurazioni e si è trovato un valore medio di 39.6 con deviazione standard di 4: ci sono seri motivi per ritenere che la produzione sia di cattiva qualità (cioè che non si possa assumere che la lunghezza media sia effettivamente 40 cm)?

È possibile risolvere questo problema senza l'uso di strumenti di calcolo ma con la sola tavola della variabile $N(0,1)$?

$H_0)\ m=40$, σ qualsiasi $H_1)\ m\neq 40$, σ qualsiasi

n "molto alto" quindi ho approssimazione gaussiana p-value: $C = \{|\bar{X} - 40| > d\} \longrightarrow P_{40}\{|\bar{X} - 40| > |\bar{x} - 40|\} = P\left\{\frac{\sqrt{100}}{4}|\bar{X} - 40| > \underbrace{\frac{\sqrt{100}}{4} 0.04}_{2}\right\} =$

gaussiana standard

$= P\{|Z| > 2\}$

Risultato "esatto" $= 2(1 - F_{T_{99}}(2))$
Risultato approssimato $= 2(1 - \phi(2)) = 2(1 - 0.977) = \underbrace{0.046}_{\text{cd esatte}}$

hanno: hp non plausibile e quindi la produzione è di cattiva qualità

Esercizio 5.3. Si osserva un campione di 10000 neonati e si trova che di essi 5106 sono femmine: testare l'ipotesi che il sesso dei neonati sia equamente distribuito, trovare il *p-value* e commentare il risultato.

p: prob che un neonato sia femmina

$H_0)\ p = \frac{1}{2}$ $\quad H_1)\ p \neq \frac{1}{2}$, calcolare il p-value. Potrei anche fare il test $H_0)\ p = \frac{1}{2}$ e $H_1)\ p > \frac{1}{2}$.

Facciamo il primo: $X_1, \dots, X_{10000}$ Bernoulli par. p

$\bar{X} = \frac{X_1 + \dots + X_{10000}}{10000} = \frac{\#\text{ femmine}}{\#\text{ neonati}}$

Regione critica $C = \{|\bar{X} - \frac{1}{2}| > d\} \longrightarrow$ p-value $= P_{0.5}\{|\bar{X} - \frac{1}{2}| > \underbrace{|\bar{x} - \frac{1}{2}|}_{\text{osservati}}\}$

Sappiamo $\frac{X_1 + \dots + X_n - n\overset{E[X_i]}{p}}{\sqrt{np(1-p)}} = \frac{\sqrt{n}(\bar{X} - p)}{\sqrt{p(1-p)}} \approx N(0,1)$

$= 2 \cdot 100 (\bar{X} - \frac{1}{2}) \approx N(0,1)$

p-value $= P_{0.5}(200|\bar{X} - \frac{1}{2}| > 200 \cdot 0.0106\} \approx P\{|Z| > 2.12\} = 2(1 - \phi(2.12)) =$
$= 2(1 - 0.982) = 0.036$: valore basso, hp poco plausibile e quindi è poco più probabile che sia femmina

Facciamo il secondo: regione critica $\{(\bar{X} - \frac{1}{2}) > d\}$

p-value $= P_{0.5}\{200(\bar{X} - \frac{1}{2}) > 200(\bar{x} - \frac{1}{2})\} \approx$

$\approx P(Z > 2.12) = 1 - \phi(2.12) \cong 0.017$ che è da scartare

Esercizio 5.6. Nel mese di gennaio è stata misurata per ogni giorno la temperatura più bassa, ottenendo una media di -12 (gradi centigradi) con deviazione standard 7.2.

a) Qual è una ragionevole stima della minima temperatura più bassa rilevabile, con un'incertezza sulla media con fiducia al 98%? ⟶ ~~estremo~~ estremo più basso int. fiducia livello 1-0.02

b) Determinare un intervallo di fiducia sinistro al 99% per la varianza della temperatura più bassa. ⟶ Int. di fiducia sinistro

c) Consideriamo l'ipotesi che in realtà la temperatura più bassa sia superiore a -10: scegliere il test opportuno, scrivere il *p-value* e la forma della regione critica al livello 0.04.

d) Se nel successivo mese di febbraio è stata rilevata una media di -14.3 con deviazione standard 6.4, si può accettare l'ipotesi che in realtà le temperature minime abbiano la stessa media (dando per scontato per semplicità che le varianze siano eguali)? Quale test è opportuno pianificare, quale risulta la formula per il *p-value*?

$X_1, \dots, X_{31} \sim N(\underbrace{m, \sigma^2}_{\text{sconosciute}})$

$\bar{x} = -12 \qquad s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{30}} = 7.2$

a. Formula generale: $\bar{X}(\omega) \pm \frac{S(\omega)}{\sqrt{n}} \underbrace{T\left(1-\frac{\alpha}{2}, n-1\right)}$

$T(0.99, 30) = 2.457$

$\bar{x} + \frac{s}{\sqrt{31}} \cdot 2.457 = -12 \pm \frac{7.2}{\sqrt{31}} \cdot 2.457 = -12 \pm 3.229$

da cui l'estremo inferiore è -15.229 (temp. minima fiducia al 98%)

b. Intervallo di fiducia sinistro per la varianza al livello $0.99 = 1 - 0.01$

$\left] 0, \frac{\sum (X_i - \bar{X})^2}{\chi^2_{(\alpha, n-1)}} \right] = \left] 0, \frac{(n-1) S^2}{\underbrace{\chi^2_{(\alpha, n-1)}}} \right] = \left] 0, \frac{30 \cdot 7.2^2}{14.95} \right] =]0, 117.85]$

$\chi^2_{(0.01, 30)} = 14.95$

$(n-1) s^2 = 30 \cdot 7.2^2$

Con prob. 0.99 la varianza non supera 117.85

c. $H_0)\ m \geq -10 \qquad H_1)\ m < -10 \qquad$ varianza qualsiasi (t-test)

p value: $m_0 = -10 \longrightarrow$ Regione critica $\{\bar{X} + 10 \leq d\}$

p-value: $P_{m_0}\{\bar{X} + 10 \leq \bar{x} + 10\}$

$\frac{\sqrt{n}(\bar{X} - m)}{S}$ è di Student $T(n-1)$ se la media è m

$P_{m_0}\left\{ \frac{\sqrt{31}(\bar{X}+10)}{S} < \underbrace{\frac{\sqrt{31}(-2)}{7.2}}_{-1.52} \right\} =$

$= F_{30}(-1.52)$: non la possiamo calcolare senza sw

↓ Approssimazione gaussiana

$\approx P(Z \leq -1.52) = 1 - \Phi(1.52) \cong 0.065$: piuttosto basso (ip. debole)

Regione critica a livello 0.04 : esaminare se il test è accettato al livello 0.04 ⟶ ci aspettiamo che l'ipotesi non sia rifiutata ma non siamo sicuri perché approx. troppo sommaria

$C = \left\{ \frac{\sqrt{n}(\bar{X} - m_0)}{S} < \underbrace{T_{(0.04, 30)}}_{\cong -1.69} \right\} : \quad \frac{\sqrt{31}(-12+10)}{7.2} \approx -1.52$ non è inferiore a -1.69 quindi si accetta al livello 0.04.

d. Base teorica $X_1, \dots, X_n$ $N(m_1, \sigma^2)$ } stessa varianza
$Y_1, \dots, Y_k$ $N(m_2, \sigma^2)$ }

se $m_1 = m_2$, $T_{n,k} = \dfrac{(\overline{X} - \overline{Y})}{\sqrt{\Sigma(x_i - \overline{x})^2 + \Sigma(y_i - \overline{y})^2}} \cdot \dfrac{\sqrt{n+k-2}}{\sqrt{\frac{1}{n} + \frac{1}{k}}}$ è di Student $T(n+k-2)$

e nel nostro caso $n+k-2 = 57$ e quindi approx $T(57)$ con Z
g. s. è abbastanza buona
↓
p-value: $P\{|T| > |t_{on}|\}$

$t_{on} = \dfrac{-12 + 14.2}{\sqrt{7.2^2 \cdot 30 + 6.4^2 \cdot 27}} \cdot \dfrac{\sqrt{57}}{\sqrt{\frac{1}{31} + \frac{1}{28}}} \cong 1.23$

$P\{|T| > |t_{on}|\} \simeq P\{|Z| > 1.23\} = 2(1 - \phi(1.23)) = 2(1 - 0.89) =$
$= 0.22$: abbastanza alto, hp ragionevolmente plausibile

Esercizio 5.7. La temperatura ottimale di estrusione di una fibra artificiale è tra i 280 ed i 290 gradi centigradi: una serie di 64 controlli ha permesso di ricavare una temperatura media (empirica) di 277.6 con deviazione standard 10.3.

a) Dare un valore massimo per la stima della temperatura di estrusione a un livello di fiducia del 98%. ⟶ Estremo sup. int. fiducia

b) Qual è il test più opportuno per valutare se la temperatura ottimale è inferiore a 275? Scrivere la formula per il *p-value* e la forma della regione critica (o di rifiuto) con fiducia al 95% (cioè al livello 0.05).

c) Per valutare se possiamo affermare con ragionevole sicurezza che la deviazione standard non supera 15, considerare il test dell'ipotesi $\mathcal{H}_0)\ \sigma \geq 15$: calcolare il *p-value* e scrivere la formula della *potenza* del test al livello 0.025.

$n = 64$ $\overline{X} = 277.6$ $s = 10.3$
modello $X_1, \dots, X_n$ gaussiane $N(\underbrace{m, \sigma^2}_{\text{sconosciute}})$.

a. Estremo superiore int. fiducia per m al livello $1 - 0.02$
↓
$\overline{X}(\omega) \pm \dfrac{S(\omega)}{\sqrt{n}} \underbrace{T(0.99, 63)}_{2.39 \text{ (circa)}}$ $\xrightarrow{\text{dati osservati}}$ $\overline{x} \pm \dfrac{s}{\sqrt{n}} \cdot 2.39 = 277.6 \pm \dfrac{10.3}{8} \cdot 2.39 = 277.6 \pm 3.07$.

L'estremo inferiore è 281.3

b. $H_0)\ m \leq 275$ $H_1)\ m > 275$ σ^2 qualsiasi (T-test)
↓
$C = \{(\overline{X} - 275) > d\}$
$(\overline{X} - 275) \cdot \dfrac{8}{S}$ è $T(63)$ e quindi circa $N(0,1)$

p-value: $P_{275} \left\{ \dfrac{8(\overline{X} - 275)}{S} > \dfrac{8}{10.3}(277.6 - 275) \right\}$
↓
valore esatto: $P\{T > 2.01\} = 1 - F_{63}(2.01)$
valore approssimato: $P\{Z > 2.01\} = 1 - \phi(2.01) = 1 - 0.982 = 0.018$
(abbastanza buono)

e quindi valore molto basso, hp molto poco plausibile (da scartare) c Test $H_0) \sigma \geq 15$ $H_1) \sigma < 15$ al livello 0.01

regione critica: $\left\{ \underbrace{\frac{\sum (x_i - \bar{x})^2}{\sigma_0^2}}_{\text{dati ottenuti}} \leq \underbrace{\chi^2_{(\alpha, n-1)}}_{} \right\}$

dati ottenuti $\chi^2_{(0.01, 60)} \simeq 37$

29.70

↳ Si rifiuta l'ipotesi al livello 0.01 e quindi il p-value è < 0.01 e quindi l'ipotesi è da scartare

Calcolo delle Probabilità e Statistica
Corso di Laurea in Informatica
Compito del 11-04-2022

Esercizio 1. (8 punti)

È data una moneta non equilibrata, per la quale la probabilità che esca testa è p (con $0 < p < 1$). Sappiamo che se vengono effettuati due lanci consecutivi con questa moneta, la probabilità di ottenere due teste è uguale alla probabilità di ottenere due risultati diversi.

a) Ricavare la probabilità p.

b) Lanciando la moneta 450 volte, trovare un valore approssimato per la probabilità di ottenere testa almeno 290 volte.

Esercizio 2. (10 punti)

Dato un parametro $\theta > 0$, si consideri la funzione

$$F_\theta(x) = \begin{cases} 0\,, & \text{per } x \leq 0; \\ c_\theta\, x^2\,, & \text{per } 0 < x < \sqrt{\theta}; \\ 1\,, & \text{per } x \geq \sqrt{\theta}; \end{cases}$$

per un'opportuna costante c_θ.

a) Si determini il valore della costante c_θ affinché la funzione $F_\theta(x)$ sia la funzione di ripartizione di una v.a. con densità. Calcolare media e varianza di una tale v.a. in funzione di θ.

b) Consideriamo ora un campione $X_1, \dots, X_n$ di v.a. con densità aventi la funzione di ripartizione $F_\theta(x)$. Esaminare la stima del parametro θ con il metodo dei momenti e con il metodo di massima verosimiglianza per un insieme di dati $x_1, \dots, x_n$ tali che $x_k > 0$ per ogni $k = 1, \dots, n$.

Esercizio 3. (12 punti)

In una catena di produzione di una ditta viene condotta un'analisi sulla quantità di pezzi difettosi prodotti. Analizzando un campione di 200 pezzi, si trovano 8 pezzi difettosi.

a) Con quale livello di fiducia si ottiene una precisione della stima di 0.02 per la probabilità che un prodotto risulti difettoso? Determinare l'intervallo bilatero al livello di fiducia ottenuto.

b) Il responsabile della ditta sostiene che la probabilità di prodotti difettosi non sia superiore al 2%. Dire, giustificando la risposta, se l'ipotesi del responsabile è plausibile alla luce del campione esaminato.

c) Sui 200 prodotti analizzati, qual è il numero massimo di prodotti difettosi che avremmo dovuto trovare per concludere che l'ipotesi del responsabile fosse molto plausibile (p-value maggiore o uguale a 0.3)?

Esercizio 3 per programma precedente. (12 punti)

Si consideri la catena di Markov con stati $S = \{1, 2, 3, 4\}$ associata alla seguente matrice di transizione:

$$P = \begin{pmatrix} 2/3 & 0 & 1/3 & 0 \\ 3/4 & 0 & 0 & 1/4 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & \lambda & 1-\lambda \end{pmatrix}$$

dove $\lambda \in [0, 1]$.

a) Determinare le probabilit invarianti della catena al variare di λ.

b) Determinare per quale valore di λ si ottiene la massima probabilit di trovarsi in 3 al tempo 2, partendo dallo stato 4 al tempo 0.

c) Partendo dallo stato 1 al tempo 0, determinare al variare di λ un tempo T per cui la probabilità di trovarsi in 4 al tempo T è positiva.

1\. a. $p^2 = 2p(1-p) \longrightarrow 3p^2 = 2p$ da cui $p = \frac{2}{3}$

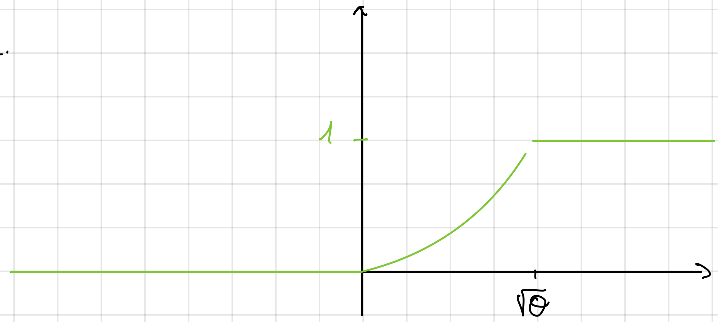
b. X_i i-esimo lancio per $\frac{2}{3}$

$$\frac{X_1 + \ldots + X_{450} - 450 \cdot \frac{2}{3}}{\sqrt{450 \cdot \frac{2}{3} \cdot \frac{1}{3}}} = \frac{X_1 + \ldots + X_{450} - 300}{10} \approx N(0,1)$$

$$\{X_1 + \ldots + X_{450} \geq 290\} = \left\{ \frac{X_1 + \ldots + X_{450} - 300}{10} \geq \frac{290 - 300}{10} \right\}$$

$$P\{X_1 + \ldots + X_{450} \geq 290\} \approx \{Z \geq -1\} = \phi(1) = 0.841$$

2\. a.



perché sia una cdf basta che si abbia $C(\theta) \cdot \theta \leq 1$. Perché sia cdf di una variabile con densità si deve avere $C(\theta) \cdot \theta = 1$ (si deve raccordare) cioè $C(\theta) = \frac{1}{\theta}$

$\downarrow$

Densità associata: $f_\theta(x) = \begin{cases} \frac{2x}{\theta} & 0 < x < \sqrt{\theta} \\ 0 & \text{altrove} \end{cases}$

$$E_\theta[X] = \frac{2}{\theta} \int_0^{\sqrt{\theta}} x \cdot x \, dx = \frac{2}{\theta} \frac{\theta\sqrt{\theta}}{3} = \frac{2\sqrt{\theta}}{3}$$

$$E_\theta[X^2] = \frac{2}{\theta} \int_0^{\sqrt{\theta}} x^2 \cdot x \, dx = \frac{2}{\theta} \cdot \frac{\theta^2}{4} = \frac{\theta}{2} \longrightarrow \mathrm{Var}_\theta(X) = \frac{\theta}{2} - \frac{4}{9}\theta = \frac{\theta}{18}$$

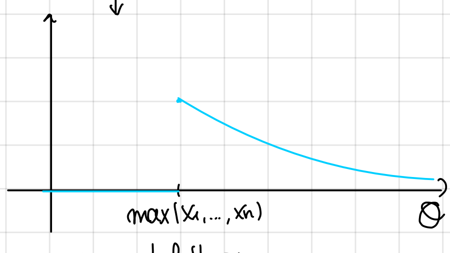
b. Metodo dei momenti

$$E_\theta[X] = \sum_{i=1}^{n} \frac{x_i}{n} = \overline{X} \longrightarrow \frac{2}{3}\sqrt{\tilde{\theta}} = \overline{X} \longrightarrow \tilde{\theta} = \frac{9}{4} \cdot \overline{X}^2 = \frac{9}{4}\left(\sum_{i=1}^{n} \frac{x_i}{n}\right)^2$$

Massima verosimiglianza

$$L(\theta, x_1, \ldots, x_n) = \begin{cases} 2^n \left(\prod_{i=1}^{n} x_i\right) \theta^{-n} & \text{se } x_i < \sqrt{\theta} \ \forall i \longrightarrow x_i \leq \sqrt{\theta} \ \forall i \iff \max(x_1, \ldots, x_n) \leq \sqrt{\theta} \\ 0 & \text{altrimenti} \end{cases}$$

$\downarrow$



θ più piccolo possibile, però $\sqrt{\theta} \geq \max(x_1, \ldots, x_n)$ da cui $\theta = [\max(x_1, \ldots, x_n)]^2$

3\. 8 pezzi difettosi su 200. p = prob. che un pezzo sia difettoso. $\hat{p} = 0.04$

a. Int. fiducia livello $(1-\alpha)$: $\hat{p} \pm \underbrace{\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} q_{1-\frac{\alpha}{2}}}$

precisione stima $= 0.02 \longrightarrow \frac{\sqrt{0.04 \cdot 0.96}}{\sqrt{200}} q_{1-\frac{\alpha}{2}} = 0.02$ da cui $q_{1-\frac{\alpha}{2}} \cong 1.44$

$$\phi(q_\beta) = \beta \longrightarrow 1 - \frac{\alpha}{2} = \phi(1.44) \cong 0.925 \quad \text{da cui} \quad \alpha \cong 2(1 - 0.925) \cong 0.15$$

È quindi l'intervallo di fiducia di livello 0.85

b. $H_0)\ p \le 0.02$ $\qquad H_1)\ p > 0.02$

$C = \{\underbrace{\bar{X} - 0.02}\ > d\}$

$$\frac{\sqrt{200}\,(\bar{X}-0.02)}{\sqrt{0.02\cdot 0.98}} \approx \text{gaussiana standard se } p = 0.02$$

Quindi il p-value è $P_{0.02}\left\{\frac{\sqrt{200}\,(\bar{X}-0.02)}{\sqrt{0.02\cdot 0.98}} > \underbrace{\frac{\sqrt{200}\,(0.04-0.02)}{\sqrt{0.02\cdot 0.98}}}_{2.02}\right\} \approx P\{Z > 2.02\} =$

$= 1-\phi(2.02) = 1-0.978 = 0.022$ molto basso: hp da scartare

c. Si torna al p-value sopra calcolato ma con $\bar{X}(\hat{p})$ generico

$P_{0.02}\left\{\frac{\sqrt{200}\,(\bar{X}-0.02)}{\sqrt{0.02\cdot 0.98}} > \underbrace{\frac{\sqrt{200}\,(\bar{X}-0.02)}{\sqrt{0.02\cdot 0.98}}}_{a}\right\} \underset{g.s.}{\approx} 1-\phi(a) \ge 0.3$ cioè $\phi(a) \le 0.7$

cioè $a \le q_{0.7} = 0.53$

$\frac{\sqrt{200}\,(\bar{X}-0.02)}{\sqrt{0.02\cdot 0.98}} \le 0.53 \longrightarrow$ chiamiamo k il numero di pezzi difettosi trovati.

Si avrà $\frac{k-4}{200} \le \frac{0.53\cdot\sqrt{0.02\cdot 0.98}}{\sqrt{200}} \longrightarrow k \le 4 + 0.53\cdot\sqrt{0.02\cdot 0.98}\cdot\sqrt{200} \longrightarrow k \le 5.049$ e

dato che deve essere intero, $k \le 5$. Si conclude che con al massimo 5 pezzi difettosi accettiamo con fiducia (p-value ≥ 0.3) l'ipotesi $p < 0.02$

ESERCIZI CONCLUSIVI:

Esercizio 5.1. I risultati dei compiti di Statistica hanno media 21 e varianza 7: quanti studenti almeno devono fare il prossimo compito affinché, con probabilità (approssimativa) di almeno 0.99, la media dei loro voti sia compresa tra 19 e 23?

n candidati con risultati $X_1, \dots, X_n$ $\quad N(21, 7)$

$\bar{X} = N\left(21, \frac{7}{n}\right) \longrightarrow \bar{X} = 21 + \sqrt{\frac{7}{n}}\, Z \qquad Z \sim N(0,1)$

Con quale n si ha $P\{19 \le \bar{X} \le 23\} \geq 0.99$ quindi

$P\{|\bar{X} - 21| \le 2\} = P\left\{\sqrt{\frac{7}{n}}\,|Z| \le 2\right\} = P\left\{|Z| \le \underbrace{\frac{2\sqrt{n}}{\sqrt{7}}}_{\text{dal disegno}}\right\} \geq 0.99$

$\frac{2\sqrt{n}}{7} \geq q_{0.995} \approx 2.58$

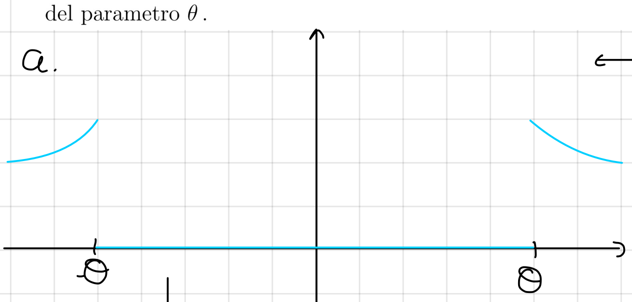
da cui $n \geq 12$

Esercizio 4.4. Consideriamo, per $\theta > 0$, la funzione

$$f_\theta(x) = C_\theta \begin{cases} e^{x} & x < -\theta \\ 0 & -\theta \le x \le \theta \\ e^{-x} & x > \theta \end{cases}$$

a) Determinare la costante che rende la funzione sopra scritta una densità di probabilità e tracciare un grafico approssimativo di $f_\theta(.)$.
b) Tracciare un grafico approssimativo della funzione di ripartizione e scriverla esplicitamente.
c) Considerato un campione con la densità sopra scritta, indagare sulla stima del parametro θ.

a.



$\longleftarrow$ densità : qualunque sia θ, $E_\theta[X] = 0$. per la stima col metodo dei momenti dobbiamo passare al momento secondo

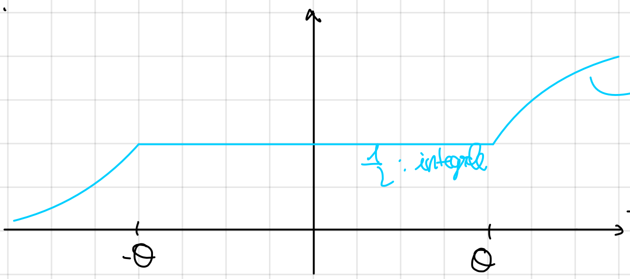
$\int_{\mathbb{R}} f_\theta(x)\,dx = C(\theta)\left[\int_{-\infty}^{-\theta} e^{x}dx + \int_{\theta}^{+\infty} e^{-x}dx\right] = 2C(\theta)\int_{\theta}^{+\infty} e^{-x}dx = 2C(\theta)e^{-\theta} = 1$ da cui $C(\theta) = \frac{e^{\theta}}{2}$

$E_\theta[X] = \frac{e^{\theta}}{2}\left[\int_{-\infty}^{-\theta} x e^{-x}dx + \int_{\theta}^{+\infty} x e^{-x}dx\right] = 0$

$E_\theta[X^2] = \frac{e^{\theta}}{2}\left[\int_{-\infty}^{-\theta} x^2 e^{-x}dx + \int_{\theta}^{+\infty} x^2 e^{-x}dx\right] = e^{\theta}\int_{\theta}^{+\infty} x^2 e^{-x}dx$: integro per parti

Di stima metodo momenti : $e^{\theta}\int_{\theta}^{+\infty} x^2 e^{-x}dx = \sum_{i=1}^{n} \frac{x_i^2}{2}$

b.



andamento derivata

$\rightarrow F_\theta(x) = \begin{cases} \frac{e^{\theta}}{2}\int_{-\infty}^{x} e^{t}dt = \frac{e^{\theta + x}}{2} & x < -\theta \\ \frac{1}{2} & -\theta \le x \le \theta \\ \frac{1}{2} + \frac{e^{\theta}}{2}\int_{\theta}^{x} e^{-t}dt = \frac{1}{2} + \frac{e^{\theta}}{2}(e^{-\theta} - e^{-x}) & x > \theta \end{cases}$

Corso Statistica A, Informatica, anno 2021-22

Esercizi conclusivi

Esercizio 1

Consideriamo la funzione così definita:

$$f(x) = \begin{cases} c\,x^2 & -1 < x < 1 \\ 0 & \text{altrove} \end{cases}$$

a) Calcolare la costante c che rende la funzione sopra scritta una densità di probabilità, e scrivere la funzione di ripartizione di una variabile che abbia quella densità;

b) preso $0 < \beta < 1$, dare una formula per il β-quantile e calcolare il momento n-simo $\mathbf{E}[X^n]$;

c) posto $Y_n = (2 + X^n)$ esaminare se la successione $(Y_n)_{n\geq 1}$ converge in probabilità ed eventualmente a quale limite.

Esercizio 2

Un ragazzo deve riempire un album di 10 figurine. Egli acquista le figurine in busta chiusa; ciascuna busta contiene una sola figurina, e si suppone che le figurine contenute nelle buste siano del tutto casuali ed indipendenti l'una dall'altra. Ovviamente la prima busta che acquista contiene una figurina che egli metterà sicuramente nell'album.

Sia X il numero di buste che deve acquistare, dopo la prima, per trovare la prima figurina diversa da quella già inserita nell'album; sia poi Y il numero di buste che deve acquistare successivamente per trovare la prima figurina diversa dalle prime due già inserite.

a) Calcolare la densità discreta (o funzione di massa) della variabile aleatoria X. È una distribuzione nota?

b) Calcolare la densità discreta della variabile aleatoria Y.

c) Quanto vale $\mathbf{P}\{X + Y = 3\}$?

Esercizio 3

Uno stilista commissiona delle scarpe ad una ditta artigianale e chiede che la varianza delle lunghezze delle scarpe di un dato numero, misurata in centimetri, non superi 0.1 . Vengono misurate accuratamente 40 scarpe prodotte e si ottiene una *varianza campionaria* eguale a 0.135 .

Supponendo che le lunghezze delle scarpe possano essere rappresentate con variabili aleatorie gaussiane, si può accettare al livello 0.05 l'ipotesi

$$H\big)_0\ \sigma^2 \leq 0.1 \qquad \text{contro} \qquad H\big)_1\ \sigma^2 > 0.1\ ?$$

Rispondere alla stessa domanda supponendo però che il numero 0.135 sia stato ottenuto come varianza campionaria delle misurazioni di un campione di 60 scarpe.

Esercizio 4

Consideriamo la funzione così definita:

$$f(x) = \begin{cases} c\,x\,\mathbf{e}^{-\frac{x^2}{2}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

a) Calcolare la costante c che rende la funzione sopra scritta una densità di probabilità, e preso $0 < \beta < 1$, dare una formula per il β-quantile di una v.a. X che abbia quella densità;

b) calcolare la funzione generatrice dei momenti della v.a. X^2 .

Esercizio 5

Il punto di fusione dello stagno allo stato puro è di 231.06 gradi Celsius. Sono stati prelevati 81 campioni di stagno proveniente da una miniera (e quindi contenente delle impurità) ed è stata fatta una accurata misurazione delle loro temperature di fusione ottenendo una temperatura media di 235.44 gradi e una deviazione standard campionaria 3.6 .

a) Il tecnico che ha condotto queste misurazioni ha calcolato che la precisione delle misurazioni è 0.82 : con quale livello di fiducia ha (approssimativamente) calcolato questa precisione?

b) Commentando i risultati, il tecnico afferma che in realtà la varianza di queste misurazioni non supera 10: indicare quale test si deve predisporre per verificare questa affermazione ed esaminare se l'ipotesi può essere accettata ai livelli 0.05 e 0.025. Che cosa si può dire sul *p-value* di questo test?

Esercizio 6

Consideriamo la funzione così definita:

$$f(x) = \begin{cases} x\,\mathbf{e}^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Sia X una v.a. avente quella densità (che è poi la densità Gamma(2,1)) e sia $Y = X^{-1}$.

a) Calcolare (se esiste) la covarianza $Cov(X, Y)$;

b) esaminare quali momenti possiede la v.a. Y ;

c) esaminare quale relazione esiste tra i quantili delle variabili X e Y .

1. a. $\int_{-1}^{1} x^2 dx = 2\int_0^1 x^2\, dx = \frac{2}{3} \longrightarrow c = \frac{3}{2}$

$$F(x) = \begin{cases} 0 & x < -1 \\ \frac{x^3}{2}+\frac{1}{2} & -1 \le x \le 1 \\ 1 & x > 1 \end{cases} \longrightarrow \int_{-\infty}^{x} f(t)dt = \frac{3}{2}\int_{-1}^{x} t^2 dt = \frac{t^3}{2}\Big|_{-1}^{x} = \frac{x^3}{2}+\frac{1}{2}$$

b. $T_\beta \longrightarrow F(T_\beta) = \beta \longrightarrow \frac{T^3}{2} + \frac{1}{2} = \beta$ da cui $T_\beta = \sqrt{2\beta - 1}$

$\mathcal{E}[X^n] = \frac{3}{2}\int_{-1}^{1} x^n \cdot x^2 dx = \frac{3}{2}\int_{-1}^{1} x^{n+2} dx \longrightarrow$ n dispari : $\mathcal{E}[X^n] = 0$
n pari : $\mathcal{E}[X^n] = 3\int_0^1 x^{n+2} dx = \frac{3}{3+n}$

$\lim_{n\to+\infty} \mathcal{E}[X^n] = 0$

c. $Y_n = (2+X^n)$ converge in probabilità? Utilizziamo il criterio sufficiente

$\mathcal{E}[Y_n] = \mathcal{E}[2+X^n] = 2+\mathcal{E}[X^n]$ e, da prima, $\lim_{n\to+\infty} \mathcal{E}[Y_n] = 2$

$Var(Y_n) = Var(2+X^n) = Var(X^n) = \mathcal{E}[X^{2n}] - \mathcal{E}[X^n]^2$ da cui $\lim_{n\to+\infty} Var(Y_n) = 0$

$\downarrow$

Si ha quindi convergenza in probabilità

2. a. X valori $1, 2, 3, \dots$
$P(X=h) = \left(\frac{1}{10}\right)^{h-1} \cdot \frac{9}{10}$ geometrica per $\frac{9}{10}$

b. Y valori $1, 2, 3, \dots$
$P(Y=h) = \left(\frac{2}{10}\right)^{h-1} \frac{8}{10} = \left(\frac{1}{5}\right)^{h-1} \cdot \frac{4}{5}$ geometrica per $\frac{4}{5}$

c. $P\{X+Y=3\} = P\{X=1, Y=2\} + P\{X=2, Y=1\} = P\{X=1\}\cdot P\{Y=2\} + P\{X=2\}\cdot P\{Y=1\}$

3. Test $H_0)\ \sigma^2 \le \sigma_0^2$ $H_1)\ \sigma^2 > \sigma_0^2$ al livello α

$C = \left\{ \frac{\Sigma(x_i - \bar{X})^2}{\sigma_0^2} > \chi^2_{(1-\alpha, n-1)} \right\}$ $\alpha = 0.05$
$1-\alpha = 0.95$

$\downarrow$

$\chi^2_{(0.95, 39)} = 55.75$
$\chi^2_{(0.99, 59)} = 79.08$

Varianza campionaria: $\frac{\Sigma(x_i - x^2)^2}{n-1}$

Per $n = 40$ si rifiuta se $\frac{\Sigma(x_i-x^2)^2}{0.1} > 55.75 \longrightarrow 52.65 < 55.75$ quindi si accetta "a fatica" e mi aspetto un p-value poco superiore a 0.05

Per $n = 60$: $79.65 > 79.08$ quindi si rifiuta per poco l'ipotesi e mi aspetto un p-value poco più basso di 0.05

4. a. $\int_0^{+\infty} x e^{-\frac{x^2}{2}} dx = -e^{-\frac{x^2}{2}} \Big|_0^{\infty} = 1$

funzione di ripartizione : $F(x) = \begin{cases} 0 & x \le 0 \\ 1-e^{-\frac{x^2}{2}} & x > 0 \longrightarrow \int_0^x t e^{-\frac{t^2}{2}} dt = e^{-\frac{t^2}{2}} \Big|_0^x = 1-e^{-\frac{x^2}{2}} \end{cases}$

$\downarrow$ quantile

$F(r_\beta) = \beta \longrightarrow 1-e^{-\frac{r^2}{2}} = \beta$ da cui $r_\beta = \sqrt{-2\log(1-\beta)}$

b. $G_{X^2}(t) = \mathcal{E}[e^{tX^2}] = \int_0^{+\infty} e^{tx^2} x e^{-\frac{x^2}{2}} dx = \int_0^{+\infty} x e^{x^2(t-\frac{1}{2})} dx$

$t \geq \frac{1}{2} \longrightarrow G_X(t) = +\infty$

$t < \frac{1}{2} \longrightarrow G_X(t) < +\infty$

} dominio della f. generatrice: $]-\infty, \frac{1}{2}]$ $(]-\varepsilon, \varepsilon[)$ quindi X^2 possiede tutti i momenti

$G_{X^2}(t) = \int_0^{+\infty} x e^{x^2(t-\frac{1}{2})} dx = \frac{1}{2t-1} e^{x^2(t-\frac{1}{2})} \Big|_0^{+\infty} = \frac{1}{1-2t}$

$\frac{d}{dx} e^{x^2(t-\frac{1}{2})} = x e^{x^2(t-\frac{1}{2})}(2t-1)$

$G_{X^2}(t) = \begin{cases} +\infty & t \geq \frac{1}{2} \\ \frac{1}{1-2t} & t < \frac{1}{2} \end{cases}$

5. a. Bisognerebbe usare la variabile $T(80)$ quindi l'approx. gaussiana è buona

$\overline{X} = 235.44$ $\quad s = 3.6$

Intervallo di fiducia (approx.) al livello $1-\alpha$: $\overline{x} \pm \frac{s}{\sqrt{80}} T_{(1-\frac{\alpha}{2}, 80)} \simeq \overline{x} \pm \frac{s}{\sqrt{80}} q_{1-\frac{\alpha}{2}}$

$0.82 = \frac{s}{80} q_{1-\frac{\alpha}{2}}$ l'incognita è α

$q_{1-\frac{\alpha}{2}} = \frac{0.82 \cdot \sqrt{80}}{3.6} \simeq 2.03$, chi è α? $\phi(q_{1-\frac{\alpha}{2}}) = 1-\frac{\alpha}{2}$ da cui $1-\frac{\alpha}{2} = \phi(2.03) = 0.978$

da cui $\alpha = 2(1-0.978) = 0.044$

Il livello scelto è $1-0.044 = 0.956 = 95.6\%$

b. Test sulla varianza $H_0)\ \sigma^2 \le 10$ $\quad H_1)\ \sigma^2 > 10$

$C = \left\{ \frac{\Sigma(x_i-\overline{x})^2}{\sigma_0^2} > \chi^2_{(1-\alpha, n-1)} \right\}$

$\chi^2_{(0.95, 80)} = 101.87$

$\chi^2_{(0.975, 80)} = 106.62$

$\frac{\Sigma(x_i-\overline{x})^2}{10} = \frac{(3.6)^2 \cdot 80}{10} = 103.68$: si rifiuta l'ipotesi al livello 0.05 e si accetta al livello 0.025. Il p-value è compreso tra 0.05 e 0.025 più o meno a "metà strada".

p-value: $P\left\{ \frac{\Sigma(X_i-\overline{X})^2}{\sigma_0^2} > \frac{\Sigma(x_i-\overline{x})^2}{\sigma_0^2} \right\} = P\{\underset{\text{chi-quadro}(80)}{C} > 103.68\} = 1-G_{80}(103.68) = 1-\text{pchisq}(103.68, 80) = 0.038$

6. a. $\text{cov}(X,Y) = \mathcal{E}[XY] - \mathcal{E}[X]\cdot\mathcal{E}[Y] = 1 - \underset{2}{\underline{\mathcal{E}[X]}} \cdot \underline{\mathcal{E}[Y]} = 1-2\cdot 1 = -1$

$\mathcal{E}[\frac{1}{X}] = \int_0^{+\infty} \frac{1}{x} f(x) dx = \int_0^{+\infty} \frac{1}{x} \cdot x e^{-x} dx = 1$

b. Y ha momento primo? $\mathcal{E}[Y^n] = \mathcal{E}\left[\frac{1}{X^n}\right] = \int_0^{+\infty} \frac{1}{x^n} x e^{-x} dx$: potrebbe $\exists n$ che lo rende $+\infty$

$\mathcal{E}[Y^2] = \int_0^{+\infty} \frac{1}{x} e^{-x} dx = +\infty$: non ha momento secondo e quindi non ha nemmeno i momenti

c. x_β e τ_β quantili di c di Y.

$P\{X \leq x_\beta\} = \beta$

$P\{Y \leq \tau_\beta\} = \boxed{P\{X \geq \frac{1}{\tau_\beta}\}} = \beta$

$P\{X \leq \frac{1}{\tau_\beta}\} = 1 - P\{X > \frac{1}{\tau_\beta}\} = 1-\beta = P\{X \leq x_{1-\beta}\}$.

Si avrà $\frac{1}{\tau_\beta} = x_{1-\beta}$