**SPAZI DI PROBABILITÀ**

**Def.** (Algebra).
Sia $\Omega$ un insieme, un'algebra $\mathcal{A}$ è una classe di sottoinsiemi di $\Omega$ tale che:

- $\Omega \in \mathcal{A}$
- se $A \in \mathcal{A} \Longrightarrow A^C \in \mathcal{A} \quad \forall A \in \mathcal{A}$
- se $A, B \in \mathcal{A} \Longrightarrow A \cap B \in \mathcal{A}$

**Def.** ($\sigma$-Algebra).
Una famiglia $\mathcal{A}$ di parti di un insieme $\Omega$ si dice una $\sigma$-algebra se:

- $\Omega \in \mathcal{A}$
- se $A \in \mathcal{A} \Rightarrow A^C \in \mathcal{A}$
- se $\{A_i\}_{i\in\mathbb{N}} \subset \mathcal{A} \Longrightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$

**Def.** (Misura di Probabilità).
La funzione $\mathbb{P}: \mathcal{A} \longrightarrow [0,1]$ è detta misura di probabilità se:

- $\mathbb{P}(\Omega) = 1$
- se $\{A_i\}_{i\in\mathbb{N}}$ è una successione di elementi di $\mathcal{A}$ a due a due disgiunti, allora vale la proprietà di *$\sigma$-additività*: $\mathbb{P}(\bigcup_{i\in\mathbb{N}} A_i) = \sum_{i\in\mathbb{N}} \mathbb{P}(A_i)$

**Def.** (Spazio di Probabilità).
Uno Spazio di Probabilità è una terna $(\Omega, \mathcal{A}, \mathbb{P})$ dove $\Omega$ è un insieme, $\mathcal{A}$ è una $\sigma$-algebra e $\mathbb{P}$ è una misura di probabilità su $\Omega$.

Indichiamo con $\mathbb{1}_A(x) = \begin{cases} 1 & se\ x \in A \\ 0 & altrimenti \end{cases}$
la FUNZIONE INDICATRICE DI UN EVENTO.
La funzione che associa un sottoinsieme $A$ di $X$ alla sua funzione indicatrice $\mathbb{1}_A$ è iniettiva; il suo codominio è l'insieme delle funzioni $f: X \to \{0, 1\}$.
Se $A$ e $B$ sono due sottoinsiemi di $X$ allora:

∘ $\mathbb{1}_{A\cap B} = \min\{\mathbb{1}_A, \mathbb{1}_B\} = \mathbb{1}_A \mathbb{1}_B$

∘ $\mathbb{1}_{A\cup B} = \max\{\mathbb{1}_A, \mathbb{1}_B\} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A\mathbb{1}_B$

**Proposizione.** Dato $\Omega$ un insieme e $\mathcal{A}, \mathcal{B}$ due $\sigma$-algebre, allora $\mathcal{F} := \mathcal{A} \cap \mathcal{B}$ è una $\sigma$-algebra, mentre in generale $\mathcal{G} := \mathcal{A} \cup \mathcal{B}$ non è una $\sigma$-algebra.

**Proprietà.** (Spazi di Probabilità).

i. $\mathbb{P}(A^C) = 1 - \mathbb{P}(A)$
ii. Se $A \subset B \Longrightarrow \mathbb{P}(A) \le \mathbb{P}(B)$ (*Monotonia*)
iii. $\mathbb{P}(\bigcup_{i\in\mathbb{N}} A_i) = 1 - \mathbb{P}(\bigcap_{i\in\mathbb{N}} {A_i}^c)$
iv. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

**Def.** (Probabilità condizionata).
Siano $A, B \in \mathcal{A}$, con $\mathbb{P}(A) > 0$. Si chiama Probabilità condizionata la quantità $\mathbb{P}_A(B)\mathbb{P}(A) = \mathbb{P}(A \cap B)$.
La nozione di probabilità condizionata è quindi legata al calcolo della probabilità quando si viene a sapere che si sono verificati certi eventi.

**Formula delle Probabilità Totali.**
$$\mathbb{P}(B) = \sum_{i\in\mathbb{N}} \mathbb{P}(A_i)\mathbb{P}_{A_i}(B)$$

**Formula di Bayes.**
Siano $A_1, \dots, A_n$ eventi disgiunti tali che $A_1 \cup \dots \cup A_n = \Omega$ (ossia gli eventi formano una partizione di $\Omega$). Vale allora la formula di Bayes
$$\mathbb{P}_B(A_k) = \frac{\mathbb{P}_{A_k}(B)\mathbb{P}(A_k)}{\sum_{j=1}^{n} \mathbb{P}_{A_j}(B)\mathbb{P}(A_j)} = \frac{\mathbb{P}_{A_k}(B)\mathbb{P}(A_k)}{\mathbb{P}(B)}$$

**Def.** (Indipendenza)
Si dice che $A, B \in \mathcal{A}$ sono indipendenti (in simboli, $A \perp\!\!\!\perp B$), se e solo se $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

**Osservazione.** Se $A \perp\!\!\!\perp B \Rightarrow A^C \perp\!\!\!\perp B^C,\ A^C \perp\!\!\!\perp B, A \perp\!\!\!\perp B^C$.

**VARIABILI ALEATORIE DISCRETE**

**Def.** (Variabile Aleatoria).
Dato uno spazio di probabilità $(\Omega, \mathcal{A}, \mathbb{P})$ si dice variabile aleatoria un'applicazione $X : \Omega \longrightarrow \mathbb{R}$
tale che, per ogni $t \in \mathbb{R}$, l'insieme $\{\omega; X(\omega) \le t\}$ appartenga a $\mathcal{A}$.

**Def.** (Variabile Aleatoria Discreta).
Si definisce v.a. discreta una variabile che possiede al più una infinità numerabile di valori $\{x_1, \dots, x_k, \dots\}$
Data una v.a. discreta X, consideriamo la funzione $p : \mathbb{R} \longrightarrow \mathbb{R}^+$ definita da $p(x) = \mathbb{P}(X = x)$. Questa gode quindi delle proprietà:

a) $p(\mathrm{x}) = 0$ tranne al più per un numero o un'infinità numerabile di valori $x_1, x_2, \dots$
b) $\sum_n p(x_n) = 1$

Chiameremo *densità discreta* una funzione $p$ che soddisfi alle condizioni sopra descritte

**Classificazione v.a. discrete**

- *Legge binomiale*

Descrive eventi del tipo:
successione di prove indipendenti la cui $\mathbb{P}$ di successo è costante ed è *p.*
Si chiama legge binomiale di parametri $n$ e $p$ quella individuata dalla densità
$$p(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & k = 0,1,\dots,n \\ 0 & altrimenti \end{cases}$$
Che verrà indicata nel seguito con il simbolo di B (n, p).
Per n=1 si avrà B (1, p) ovvero la *legge di Bernoulli* di parametro *p.*

- *Densità geometrica*

$$p(k) = \begin{cases} p(1-p)^k & k = 0,1,\dots,n \\ 0 & altrimenti \end{cases}$$
Formula utile:
$P(X \ge k) = \sum_{i=k}^{\infty} p(1-p)^i = (1-p)^k$

- *Densità geometrica modificata*

$$q(k) = \begin{cases} p(1-p)^{k-1} & k = 0,1,\dots,n \\ 0 & altrimenti \end{cases}$$
Notiamo che se Z è una v.a. geometrica modificata, allora Z-1 segue una legge geometrica.

- *Legge di Poisson di parametro $\lambda$*

Descrive eventi del tipo:
in un intervallo di tempo *t possono* verificarsi eventi simili.
Si chiama legge di Poisson di parametro $\lambda$ quella individuata dalla densità
$$q(k) = \begin{cases} e^{-\lambda}\dfrac{\lambda^k}{k!} & k = 0,1,\dots,n \\ 0 & altrimenti \end{cases}$$
La legge di Poisson viene utilizzata in modo naturale per approssimare v.a. con legge di Bernoulli ponendo $\lambda = np$.

APPROSSIMAZIONE DI STIRLING
$$k! \approx k^k e^{-k}\sqrt{2\pi k}$$

**Funzioni di ripartizione per v.a. discrete**

**Def.** (funzione di ripartizione f.r)
Data una v.a. X, discreta o no, si chiama funzione di ripartizione di X la funzione $F_X : \mathbb{R} \longrightarrow [0,1]$ definita da $F_X(t) = \mathbb{P}(X \leq t)$.

**Proprietà** (f.r.)
∘ $0 \leq F_X(t) \leq 1$
∘ $\lim_{t \longrightarrow -\infty} F_X(t) = 0$ e $\lim_{t \longrightarrow +\infty} F_X(t) = 1$.
∘ $F_X(t)$ è monotona non decrescente.
∘ $F_X(t)$ è continua a destra.
∘ $F_X(t)$ ha una discontinuità di salto $\mathbb{P}(X = t)$ a sinistra.

**Osservazioni**
∘ Una f.r. è sempre una funzione non decrescente, poiché se $t$ cresce l'evento {X ≤ t} diventa più grande.
∘ La f.r è importate perché la sua conoscenza è equivalente a quella della distribuzione di X.
Infatti, ricordando che $\mathbb{P}(X \in \mathcal{A}) = \sum_{x_i \in \mathcal{A}} p(x_i)$
Allora si ha che $F_X(\mathrm{t}) = \sum_{x \leq t} p(x)$
che esprime la f.r. in termini della densità.
∘ Talvolta per calcolare la densità di una v.a. discreta può essere più facile calcolare prima la f.r $F_X$ e poi da questa ricavare la densità tramite

$$F_X(k) - F_X(k-1) = p(k)$$

∘ Infine, per una v.a. X geometrica di parametro $p$ la f.r. vale $F_X(k) = 1 - (1-p)^{k+1}$ k=0,1, …
Mentre per le densità binomiali o di Poisson non ci sono formule semplici.

**f.d.r. per v.a. comuni**

- *Geometrica*
  $F(k) = 1 - (1-p)^{k+1}$
- *Esponenziale*
  $F(x) = \begin{cases} 1 - e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$
- *Gamma*
  $F(x) = \begin{cases} 1 - e^{-\lambda x} \sum_{k=0}^{m-1} \frac{(\lambda x)^k}{k!} & x > 0 \\ 0 & x \leq 0 \end{cases}$

**Vettori aleatori, leggi congiunte, indipendenza**

**Def.** (Vettore aleatorio discreto)
Una v.a. m-dimensionale discreta (oppure vettore aleatorio discreto) è un'applicazione
$X = (X_1, \dots, X_m) : \Omega \longrightarrow \mathbb{R}^m$ tale che le applicazioni $X_1, \dots, X_m$ siano delle v.a. reali discrete.

**Def.** (Densità congiunta, Densità marginali)
Se $X_1, \dots, X_m$ sono v.a. reali, la densità $p$ del v.a. $X = (X_1, \dots, X_m)$ si chiama densità congiunta delle v.a. $X_1, \dots, X_m$. Viceversa, se $X = (X_1, \dots, X_m)$ è un v.a. si chiamano densità marginali le densità $p_1, \dots, p_m$ delle v.a. $X_1, \dots, X_m$.

**Osservazione** (D.c. ⟶ D.m.)
Se la densità congiunta è nota, è facile calcolare le densità marginali. Supponiamo m=2, allora
$p_1(z) = \sum_i p(z, x_2^{(i)})$ e $p_2(z) = \sum_i p(x_1^{(i)}, z)$.
Non è invece possibile, in generale, conoscendo le sole densità marginali ricostruire la densità congiunta.

**Def.** (indipendenza)
Diremo che le v.a. $X_1, \dots, X_n$ sono indipendenti ⇔
$\forall\, m > 0$ risultano tra loro indipendenti $X_1, \dots, X_m$.

**Proprietà**
In particolare, nel caso discreto e per m=2 avremo che

$$\mathbb{P}(X_1 = x_1, X_2 = x_2) = \mathbb{P}(X_1 = x_1)\mathbb{P}(X_2 = x_2)$$

E di conseguenza se x= $(x_1, x_2)$:

$$p(x) = p_1(x_1)p_2(x_2)$$

∘ Quindi tale equazione che lega densità marginali e congiunta è una condizione equivalente all'indipendenza. Ovvero nel caso di v.a. indipendenti posso calcolare la densità congiunta a partire dalle marginali.
∘ Altra applicazione è la verifica dell'indipendenza delle v.a. tramite tale equazione.

**Def.** (Densità condizionale)
Date due v.a. X e Y aventi densità congiunta $p$, si chiama densità condizionata di X dato Y=y la quantità $p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)}$ se $p_Y(y) > 0$ e $p_{X|Y}(x|y) = 0$ altrimenti.

**Densità congiunte per v.a. comuni**

- *Binomiale*
  Possiamo definire la densità B (n, p) come la densità congiunta della somma di n v.a. di Bernoulli indipendenti di parametro p.
  Inoltre, la somma di v.a. con legge binomiale è ancora una v.a. con legge binomiale
  B (n=$n_1 + \cdots + n_m, p$).
- *Geometrica*
  Consideriamo U, V v.a. di legge geometrica di parametro r, allora U+V ha densità
  g(k)=(k+1)$r^2(1-r)^k$
- *Poisson*
  Consideriamo U, V v.a. di Poisson di parametri λ e μ rispettivamente, allora U+V è una Poisson di parametro λ+μ.

**VARIABILI ALEATORIE CONTINUE**

**Def.** (Variabili aleatorie continue)
Una v.a. con f.r. F si dice continua se F soddisfa le ipotesi del Teorema Fondamentale del Calcolo Integrale, ossia se
$\exists\ f: \mathbb{R} \longrightarrow \mathbb{R}\ t.c\ F(t) = \int_{-\infty}^{t} f(x)dx\ \ \forall t \in \mathbb{R}$
La funzione f è detta DENSITÀ DI PROBABILITÀ e soddisfa

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

**Osservazioni.**

i. Se X è una v.a. continua allora $\mathbb{P}(X = t) = 0\ \forall t$, in quanto la f.r F è continua. Ciò implica che, ad esempio, $\mathbb{P}(X \leq t) = \mathbb{P}(X < t)$.
ii. Vale che $\mathbb{P}(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(x)dx$ : dunque il calcolo di eventi di questo tipo è ricondotto al calcolo di integrali.
iii. La definizione di v.a. continua permette di calcolare la f.r F se si conosce la densità f (o meglio, una densità f). la definizione infatti non fa che affermare che F è la funzione integrale di f. Grazie al TFC, se F è una funzione derivabile con derivata continua su tutto ℝ (tranne al più in un numero finito di punti) allora F è la funzione integrale della sua derivata F'. Dunque, f=F' può essere scelta come densità per F.

**Teorema.**
Se $F: \mathbb{R} \longrightarrow [0,1]$ soddisfa

i. È continua a destra
ii. È monotona non decrescente
iii. $\lim_{t \to -\infty} F_X(t) = 0$ e $\lim_{t \to +\infty} F_X(t) = 1$

Allora esiste $(\Omega, \mathcal{A}, \mathbb{P})$ spazio di probabilità e una v.a. $X: \Omega \longrightarrow \mathbb{R}$ t.c. $\mathbb{P}(X \le t) = F(t)$.

**Classificazione v.a. CONTINUE**

- *Distribuzione uniforme*

$$X \sim Unif(a,b), X(\omega) = \omega$$

La distribuzione uniforme è *uniforme* su un insieme, ovvero attribuisce la stessa probabilità a tutti i punti appartenenti ad un dato intervallo [a,b] contenuto nell'insieme.
**Legge** di una v.a. uniforme: $\mu_X = \frac{1}{b-a}\mathbb{1}_{(a,b)}(x)$

- *Distribuzione esponenziale*

**Densità** di una distribuzione esponenziale

$$f(x) = e^{-x}\mathbb{1}_{(0,+\infty)}(x)$$

**Densità** di una distribuzione esponenziale di parametro $\lambda$

$$f(x) = \lambda e^{-\lambda x}\mathbb{1}_{(0,+\infty)}(x)$$

- *Legge Beta di parametro $\alpha$ e $\beta$*

$$X \sim Beta(\alpha, \beta)$$

**Densità** di una distribuzione con legge Beta

$$f(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & se\ 0 < x < 1 \\ 0 & altrimenti \end{cases}$$

Con $\Gamma(k) \coloneqq \int_0^\infty t^{k-1} e^{-t} dt$, ma
se $k \in \mathbb{N} \Rightarrow \Gamma(k) = (k-1)!$

- *Legge Gamma di parametro $\alpha$ e $\lambda$*

$$X \sim \Gamma(\alpha, \lambda)$$

**Densità** di una distribuzione con legge Gamma

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

Con $\Gamma(k) \coloneqq \int_0^\infty t^{k-1} e^{-t} dt$, ma
se $k \in \mathbb{N} \Rightarrow \Gamma(k) = (k-1)!$
Notiamo che se $\alpha$=1, allora questa è la legge esponenziale con parametro $\lambda$.

**Teorema.**
Siano $X_1, \dots, X_n$ v.a. i.i.d. con densità esponenziale di parametro $\lambda \Rightarrow S = X_1 + \dots + X_n \sim \Gamma(n, \lambda)$

**Teorema.** (Approssimazione di esponenziali con Poisson)
Consideriamo gli eventi in successione $X_k$ denominati "tempi di attesa" che sono v.a. i.i.d. con legge esponenziale di parametro $\lambda$ e $T_n$ denominati "tempi di arrivo" tale che $T_n = X_1 + \dots + X_n$ e dunque ha legge associata Gamma di parametri $n$ e $\lambda$.
Consideriamo infine il numero di eventi al tempo $t$: $N_t$ tale che $\mathbb{P}(N_t \ge n) = \mathbb{P}(T_n \le t)$. Allora
$\mathbb{P}(N_t = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$, ovvero $N_t \sim Pois(\lambda t)$

**Def.** (Vettore Aleatorio Continuo)
Sia $(\Omega, \mathcal{A}, \mathbb{P})$ uno spazio di probabilità e $X: \Omega \longrightarrow \mathbb{R}$ un'applicazione. Diremo che X è una v.a. m-dimensionale (o vettore aleatorio) se le sue componenti $X_1, \dots, X_m$ sono delle v.a. Si supporrà m=2 e indicheremo con X,Y le componenti del v.a. bidimensionale Z=(X,Y).

Poiché supponiamo che X e Y siano v.a., sappiamo che gli insiemi $(X \le x)$ e $(Y \le y)$ sono degli eventi. Quindi è un evento anche l'insieme $(X \le x, Y \le y)$ che ne è l'intersezione. Esso si può anche esprimere con $(Z \in A_{x,y})$ dove $A_{x,y}$ è il sottoinsieme di $\mathbb{R}^2$ $A_{x,y} = \{(u,v); u \le x, v \le y\}$.

**Def.** (f.r. congiunta)
Indichiamo con F la funzione di ripartizione congiunta di X e Y definita da
$F_Z(x,y) = \mathbb{P}(X \le x, Y \le y) = \mathbb{P}(Z \in A_{x,y})$.
Anche in questo caso parleremo di DENSITÀ DI PROBABILITÀ CONGIUNTA: se essa esiste, allora deve essere ≥0, integrabile e tale che $\int_{\mathbb{R}^2} f(x,y)dxdy = 1$.

**Formula utile.**

$$\mathbb{P}\left((X,Y) \in A_{x,y}\right) = \mathbb{P}(Z \in A) = \int_A f(u,v)dudv$$

**Def.** (legge congiunta e leggi marginali)

$$f_Z(x,y) = \frac{\partial^2}{\partial x \partial y} F_Z(x,y)$$

$$f_X(x) = \int_{\mathbb{R}} f_Z(x,y)dy$$

$$f_Y(y) = \int_{\mathbb{R}} f_Z(x,y)dx$$

**Osservazione.**
Date due v.a. X e Y la v.a. X+Y ha densità data dal "prodotto di convoluzione" delle densità $f_X(x)$ e $f_Y(y)$ cioè

$$f_{X+Y}(t) = (f_X * f_Y)(t) \coloneqq \int_{\mathbb{R}} f_X(x) f_Y(t-x)dx$$

**Def.** (Densità condizionale)
Siano X e Y due v.a. di densità congiunta $f$ e indichiamo con $f_X$ e $f_Y$ le rispettive densità marginali. Si chiama densità condizionale di X dato Y=y la quantità

$$f_{X|Y=y}(x) = \frac{f(x,y)}{f_Y(y)}$$

**LEGGI NORMALI (gaussiane)**

◦ Indichiamo con $Z \sim \mathcal{N}(0,1)$ la v.a. con distribuzione gaussiana standard.
**Densità** di una v.a. gaussiana:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

essa è ≥0 ed integrata su tutto $\mathbb{R}$ dà come risultato 1.
**F.r** di una v.a. gaussiana:

$$\Phi(x) = \int_{-\infty}^x \varphi(x)dx$$

Questa non è esprimibile algebricamente.
**Speranza matematica** di una v.a. gaussiana standard

$$\mathbb{E}[Z] = 0$$

**Varianza** di una v.a. gaussiana standard

$$Var(Z) = 1$$

Osserviamo allora come le quantità 0 e 1 presenti in $\mathcal{N}(0,1)$ siano rispettivamente la media e la varianza di Z.

**Osservazione.**

i. $\mathbb{E}[Z^{2n+1}] = 0$
ii. $\mathbb{E}[Z^{2n}] = (2n-1)!!$

∘ Consideriamo $\boldsymbol{X = \mu + \sigma Z}$. Allora X è una v.a. gaussiana di parametri $\mu$ e $\sigma^2$, cioè $X \sim \mathcal{N}(\mu, \sigma^2)$.
**Densità** di X:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$$

**Speranza matematica** di X:

$\mathbb{E}[X] = \mu$

**Varianza** di X:

$Var(X) = \sigma^2$.

Il grafico di una densità $X \sim \mathcal{N}(\mu, \sigma^2)$ ha l'andamento a campana; $\mu$ è il punto di massimo della densità; inoltre a valori di $\sigma^2$ bassi corrispondono campane strette e alte, a valori grandi campane aperte e appiattite.

**Osservazione.**
Y=-Z è $\mathcal{N}(0,1)$, il che si esprime anche dicendo che la legge $\mathcal{N}(0,1)$ è simmetrica.

**Teorema.**
Se $X \sim \mathcal{N}(\mu, \sigma^2)$ e $Y \sim \mathcal{N}(\nu, \vartheta^2)$ e $X \perp\!\!\!\perp Y$, allora $X + Y \sim \mathcal{N}(\mu + \nu, \sigma^2 + \vartheta^2)$ .

**Teorema.**
Se X e Y sono leggi gaussiane centrate, ovvero hanno speranza matematica finita, allora
sono indipendenti ⇔ non sono correlate ⇔ $\mathbb{E}$[XY]=0

**Def. (**Funzione Gamma)
Si chiama funzione Gamma la funzione $\Gamma: \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ definita da: $\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1}e^{-x}\,dx$, tale integrale non è risolvibile analiticamente.
Tuttavia, per α=1 $\Gamma(1) = 1$ e se α>0 $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$,
da cui si ha che $\forall n \in \mathbb{N}^*$ $\Gamma(n) = (n-1)!$ .

**SPERANZA MATEMATICA (Media)**

**Def.** (Speranza matematica per v.a. discrete)
Diremo che X ha speranza matematica finita se

$\sum_i |x_i| p(x_i) < +\infty$

In questo caso si chiama speranza matematica di X la quantità

$\mathbb{E}[X] = \sum_i x_i p(x_i) = \sum_i x_i \mathbb{P}(X = x_i)$

Se $\mathbb{E}$[X]=0 si dice anche che X è centrata.

**Def.** (Speranza matematica per v.a. continue)
Sia X una v.a. di densità continua *f.*
Si dice che X ha speranza matematica finita ⇔

$\int_{-\infty}^{\infty} |x| f(x)\,dx < +\infty$

Se X ha speranza matematica finita si chiama speranza matematica di X la quantità

$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)\,dx$

**Def.** (Speranza condizionale)
Date due v.a. X e Y aventi densità congiunta *f*, si chiama speranza condizionale di X dato Y=y la media (se esiste) della densità condizionale di X dato Y=y.
Questa definizione vale sia che la densità di X e Y sia continua, discreta o mista.
Se supponiamo che *f* sia continua si avrà che la speranza condizionale di X dato Y=y è definita da

$\mathbb{E}[X|Y = y] = \int x f_{X|Y}(x|y)\,dx$

a condizione che l'integrale converga assolutamente.

**Teorema**
Siano $X = (X_1, \dots, X_m)$ un v.a. discreto e $\phi : \mathbb{R}^m \longrightarrow \mathbb{R}$ una funzione. Poniamo Z= $\phi(X)$ e indichiamo con $x_1, \dots, x_m$ i valori assunti da X. Infine, con $p$ e $g$ rispettivamente le densità di X e Z. Allora Z ha speranza matematica finita ⇔

⇔ $\sum_i |\phi(x^{(i)})| p(x^{(i)}) < +\infty$

e in questo caso
$\mathbb{E}[Z] = \sum_i \phi(x^{(i)}) p(x^{(i)}) = \sum_i \phi(x^{(i)}) P(X = x^{(i)})$

**Proposizione**
L'applicazione $X \longrightarrow \mathbb{E}[X]$ è lineare.
Siano X e Y aventi speranza matematica finita. Allora

a. Per ogni c ∈ ℝ la v.a. cX ha speranza matematica finita e si ha
$\mathbb{E}[cX] = c\mathbb{E}[X]$.
b. X + Y ha speranza matematica finita e
$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

**Proposizione**
Supponiamo che le v.a. X e Y abbiano speranza matematica finita.

a. Se $\mathbb{P}(X \geq Y) = 1$ allora $\mathbb{E}[X] \geq \mathbb{E}[Y]$ e l'uguaglianza è possibile ⇔ $\mathbb{P}(X=Y) = 1$.
b. $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$.

**Speranza matematica per v.a. comuni**

- *Bernoulli*

X~B (1, p): $\mathbb{E}$[X]=p

- *Indicatrice*

Sia A ∈ $\mathcal{A}$ e consideriamo la v.a. indicatrice $\mathbb{1}_A$ ⇒
$\mathbb{E}[\mathbb{1}_A] = \mathbb{P}(A)$

- *Binomiale*

X~B (n, p): $\mathbb{E}$[X]=np

- *Poisson*

X di Poisson di parametro λ:
$\mathbb{E}$[X]=λ

- *Geometrica*

X geometrica di parametro *p:*
$\mathbb{E}[X]=\frac{1-p}{p}$
Se invece Y fosse geometrica modificata:
$\mathbb{E}[X]=\frac{1}{p}$

- *Distribuzione uniforme su [a,b]*

Sia X~Unif([a,b])

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x\mathbb{1}_{[a,b]}(x)\,dx = \frac{a+b}{2}$$

- *Esponenziale*

Sia X una v.a. continua con densità esponenziale di parametro λ

$$\mathbb{E}[X] = \lambda \int_0^{\infty} x\, e^{-\lambda x}\,dx = \frac{1}{\lambda}$$

- *Gamma*

Sia X una v.a. continua con legge Gamma associata di parametri associata di parametri α e λ
$\mathbb{E}[X] = \frac{\alpha}{\lambda}$

- *Beta*

Sia X una v.a. continua con legge Beta associata di parametri associata di parametri α e β
$\mathbb{E}[X] = \frac{\alpha}{\alpha+\beta}$

**Proposizione**
Se X e Y sono v.a. indipendenti ed hanno speranza matematica finita, allora anche la v.a. prodotto XY ha speranza matematica finita e si ha $\mathbb{E}$ [XY]=$\mathbb{E}$[X]$\mathbb{E}$[Y]

## MOMENTI, VARIANZA, COVARIANZA

**Def.** (Momento)
Diremo che una v.a. X ha momento di ordine $k$ finito k=1, 2, … se la v.a. $X^k$ ha media finita. In questo caso la quantità $\mathbb{E}\,[X^k]$ si chiama il momento di ordine *k*.
Analogamente se la v.a. $(X - \mathbb{E}[X])^k$ ha media finita diremo che X ha momento centrato di ordine $k$ finito e chiameremo momento centrato di ordine $k$ la quantità $\mathbb{E}\,[(X - \mathbb{E}[X])^k]$.
Si ha che per v.a. discrete

$$\mathbb{E}[X^k] = \sum_i x_i^k p(x_i)$$
$$\mathbb{E}[(X-\mu)^k] = \sum_i (x_i - \mu)^k p(x_i)$$

Si ha invece che per v.a. continua

$$\mathbb{E}[X^k] = \int x^k f_X(x)\, dx$$
$$\mathbb{E}[(X-\mathbb{E}[X])^k] = \int (x - \mathbb{E}[X])^k f_X(x)\, dx$$

**Def.** (Varianza)
Se X ha momento di secondo ordine finito, si chiama varianza il suo momento centrato del secondo ordine e si scrive

$$\mathrm{Var}(X) = \mathbb{E}[(X-\mathbb{E}[X])^2].$$

Si vede che Var(X) deve essere necessariamente $\geq 0$.

**Def.** (Covarianza)
Si definisce covarianza di X e Y la quantità

$$Cov(X,Y) = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]$$

Da cui

$$Cov(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

**Osservazioni** (Varianza)
◦ La varianza è una "misura" della dispersione di X attorno alla sua media. Infatti, più Var(X) è piccola e più è piccola probabilità che X prenda valori lontani dalla sua media.
◦ Sviluppando il quadrato si ha:

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

◦ Sono utili le seguenti proprietà

$$Var(aX) = a^2 Var(X)$$
$$Var(a+X) = Var(X)$$
$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$

in particolare, se X e Y sono indipendenti allora vale che

$$Cov(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0$$

E dunque

$$Var(X+Y) = Var(X) + Var(Y)$$

**Varianza per v.a. comuni**

- *Bernoulli*
  Se X∼B (1, p) allora
  $$Var(X) = p(1-p)$$
- *Binomiale*
  Se X∼B (n, p) allora
  $$Var(X) = np(1-p)$$
- *Geometrica*
  Se X è una v.a. con densità geometrica allora
  $$Var(X) = \frac{1-p}{p^2}$$
- Poisson
  Sia X una Poisson di parametro λ allora
  $$\mathbb{E}\,[X^2] = \lambda(\lambda+1)$$
  da cui
  $$Var(X) = \lambda$$
- *Distribuzione uniforme su [a,b]*
  Sia X∼Unif([a,b]) allora
  $$Var(X) = \frac{1}{12}(b-a)^2$$
- *Esponenziale*
  Sia X una v.a. continua con densità esponenziale di parametro λ
  $$Var(X) = \frac{1}{\lambda^2}$$
- *Gamma*
  Sia X una v.a. continua con legge Gamma associata di parametri associata di parametri α e λ
  $$V(X) = \frac{\alpha}{\lambda^2}$$
- *Beta*
  Sia X una v.a. continua con legge Beta associata di parametri associata di parametri α e β
  $$V(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

**Osservazione** (Covarianza)
La covarianza viene spesso usata come una misura di indipendenza di sue v.a. se la covarianza è prossima a zero le v.a. sono considerate "quasi" indipendenti, mentre per valori grandi (in valore assoluto) della covarianza fanno pensare a una "forte" dipendenza.
Se $Cov(X,Y) = 0$ si dice che le v.a. X e Y sono *non correlate.*

**Def.** (Coefficiente di correlazione)
Come "misura d'indipendenza" è però meglio usare il coefficiente di correlazione

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

Dove $\sigma_X$ e $\sigma_Y$ sono rispettivamente la deviazione standard di X e Y.
Si ha inoltre che esso è invariante rispetto ai cambiamenti di scala e vale sempre che

$$-1 \leq \rho_{X,Y} \leq 1$$

**Def.** (matrice di covarianza)
Dato un v.a., continuo o discreto, X si chiama matrice di covarianza di X la matrice $C = (c_{ij})_{ij}$ i cui elementi sono i numeri $c_{ij} = Cov(X_i, X_j)$.
Si tratta di una matrice simmetrica m×m, semi-definita positiva.
In particolare, gli autovalori di C sono tutti $\geq 0$.

## INFORMAZIONE ED ENTROPIA

L'informazione riguarda quanto l'accadere di un evento influenza il nostro stato di conoscenza del mondo.

**Def.** (Informazione)
Per $E \in \mathcal{A}$, l'Informazione portata dall'evento E è definita come $I(E) := -log_2(\mathbb{P}(E))$. L'unità di misura è [bit].

**Proprietà.**
a) $I(E) \geq 0$
b) Se E e F sono indipendenti $\Rightarrow I(E \cap F) = I(E) + I(F)$

**N.B.** Se $\mathbb{P}(E) = 0 \Rightarrow log_2(\mathbb{P}(E)) = +\infty$, che si interpreta dicendo che NON è possibile ottenere informazioni da un evento impossibile. Se $\mathbb{P}(E) = 1 \Rightarrow I(E) = 0$, che si interpreta dicendo che un evento certo non dà informazione.

**Def.** (Entropia)
Consideriamo l'informazione media: si chiama Entropia del sistema la quantità

$$H(X) = -\sum_k p_k log_2 p_k$$

Assumendo che $0 \cdot log_2 0 = 0$

**Teorema.**
1. $H(X) \geq 0, H(X) = 0 \Leftrightarrow \mathbb{P}(X = c) = 1$
2. Se X assume n valori $\Rightarrow H(X) \leq log_2(n)$,
   $H(X) = log_2(n) \Leftrightarrow X \sim Unif\{x_1, \dots, x_n\}$.

**Def.** (Entropia condizionata)
Definiamo Entropia condizionata la quantità

$$H_X(Y) = \sum_{k=1}^{+\infty} \mathbb{P}(X = x_k) H_{X=x_k}(Y)$$

con

$$H_{X=x_k}(Y) = -\sum_i \mathbb{P}_{X=x_k}(Y = y_i) log_2 \left(\mathbb{P}_{X=x_k}\left(Y = y_j\right)\right).$$

**Teorema.** (Entropia congiunta)
Sia Z=(X,Y) un vettore aleatorio.

1. L'Entropia congiunta H(X,Y) è data dalla somma dell'entropia di X con l'entropia condizionata di Y dato X: $H(X,Y) = H(X) + H_X(Y)$.
2. Se $X \perp\!\!\!\perp Y \Longrightarrow H(X,Y) = H(X) + H(Y)$.

**Teorema.**
Se $X \sim \{(x_k, p_k)\}$ e se $p_1 \neq p_2$ a meno di riordino, definisco

$$Y \sim \left\{(x_k, p_k^Y) \,\middle|\, P_k^Y = P_k \; se \; k \geq 3, P_1^Y = P_2^Y = \frac{1}{2}(P_1 + P_2)\right\}$$

Allora H(Y)>H(X).

**Principio di massima Entropia**
Consideriamo la v.a. discreta $X \sim \{(x_k, p_k)\}$. Richiediamo che essa soddisfi

$$\begin{cases} \max \{H(X) = -\sum_k p_k log_2 p_k\} \\ condizioni\ note, \quad \sum_k p_k - 1 = 0 \end{cases}$$

Vogliamo cioè determinare i valori $p_k$ affinché l'entropia risulti massima. Il problema si risolve con il metodo dei moltiplicatori di Lagrange: definiamo cioè la funzione

$$\mathcal{L}(p_1, \dots, p_n, \lambda, \mu, \dots)$$

dove i valori $\lambda, \mu, \dots$ moltiplicano, nell'espressione di $\mathcal{L}$, i vincoli imposti. Si dovrà poi risolvere il sistema $\nabla\mathcal{L} = 0$ , per ricavare i punti stazionari.
Ad esempio, qualora la v.a. X assuma tre valori con rispettive probabilità e oltre alla solita
$\sum_k p_k - 1 = 0$ venga imposto un altro vincolo, il sistema da risolvere sarà di 5 equazioni in 5 incognite.

**Def.** (Mutua informazione)
Siano X e Y v.a. allora la muta informazione di X e Y è data da:

$$\mathrm{I(X, Y) = H(Y)} - H_X(Y)$$

che è sempre ≥0 e si può intendere intuitivamente come la quantità d'informazioni in Y dipendente da X.

**Teorema.**
I(X,Y)=I(Y,X)

**Teorema.** (Disuguaglianza di GIBBS)
Date due distribuzioni di probabilità $\{p_k\}$ e $\{q_k\}$ sullo stesso alfabeto vale la seguente diseguaglianza di Gibbs:

$$-\sum_k p_k log_2(p_k) \leq -\sum_k p_k \, log_2(q_k)$$

**FUNZIONE GENERATRICE DEI MOMENTI**

Si chiama *"Funzione generatrice dei momenti"* (o "Trasformata di Laplace") di una v.a. X la funzione $M_X$ definita per $t \in \mathbb{R}$ da

$$M_X(t) := \mathbb{E}[e^{tX}].$$

Il dominio di definizione della f.g.m. è
$D = \{t \in \mathbb{R}: \mathbb{E}[e^{tX}] < +\infty\}$.
Naturalmente, si ha che:

∘ $M_X(t) = \sum_k e^{tx_k} p_k$ se $X$ è una v.a. discreta
∘ $M_X(t) = \int_{\mathbb{R}} e^{tx} f(x) dx$ se $X$ è una v.a. continua

**Osservazione**
La f.g.m. può non essere definita per alcune v.a.
Notiamo che la f.g.m. è sempre definita per $t = 0$ (dunque $D \neq \emptyset$) e ovviamente $M_X(0) = 1$. Tuttavia, può succedere che il valore 0 sia l'unico in cui la f.g.m. è definita.

**Proposizione**
Se $X, Y$sono due v.a. tali che $X \perp\!\!\!\perp Y$ aventi f.g.m. rispettivamente $M_X(t)$ e $M_Y(t)$, allora, per $\alpha, \beta \in \mathbb{R}$,
$M_{\alpha X + \beta Y}(t) = M_X(\alpha t) M_Y(\beta t)$.

**Teorema**
Se esiste $t_0 > 0$ t.c. $M_X(t)$ è definita in $t \in (-t_0, t_0)$ allora

$$\mathbb{E}[X^n] = \frac{d^n}{dt^n} M_X(0)$$

**F.g.m. note**

- $X \sim Unif(0,1)$: $M_X(t) = \frac{e^t - 1}{t} \;\; t \in \mathbb{R} \backslash \{0\}$
- $X \sim \Gamma(\alpha, \lambda)$: $M_X(t) = (\frac{\lambda}{\lambda - t})^\alpha \;\; t < \lambda$
- $X \sim \mathcal{N}(0,1)$: $M_X(t) = e^{-t^2/2} \;\; t \in \mathbb{R}$
- $Y = \mu + \sigma X$: $M_Y(t) = e^{\mu t} e^{-t^2\sigma^2/2}$

**FUNZIONE CARATTERISTICA**
**Def.** (V.a. complessa, speranza matematica di una v.a. complessa)
Diremo che una applicazione $Z = Z_1 + iZ_2$ definita sullo spazio di probabilità $(\Omega, \mathcal{A}, \mathbb{P})$ a valori complessi è una v.a. complessa se e solo se entrambe le applicazioni $Z_1$ e $Z_2$ sono v.a. reali.
Data una v.a. a valori complessi diremo che essa ha speranza matematica finita se ciò è vero per $Z_1$ e $Z_2$ e in questo caso porremo $\mathbb{E}[Z] = \mathbb{E}[Z_1] + i\mathbb{E}[Z_2]$

Sia $i = \sqrt{-1}$, e sia $X$ una v.a. Si chiama *"Funzione caratteristica"* (o "Trasformata di Fourier") di $X$ la funzione $\varphi_X$ a valori complessi definita da

$$\varphi_X(t) := \mathbb{E}[e^{itX}] = \mathbb{E}[\cos(tX)] + i\mathbb{E}[\sin(tX)].$$

Naturalmente, si ha che:

∘ $\varphi_X(t) = \sum_k e^{itx_k} p_k$ se $X$ è una v.a. discreta
∘ $\varphi_X(t) = \int_{\mathbb{R}} e^{itx} f(x) dx$ se $X$ è una v.a. continua

**Osservazione**
La f.c. è sempre definita perché, qualunque sia $t \in \mathbb{R}$ le v.a. $\cos(tX)$ e $\sin(tX)$ sono sempre limitate, anzi, essendo $\mathbb{E}[X]$ un operatore lineare ed intendendo con $|\cdot|$ il modulo complesso vale che $|\varphi_X(t)| = |\mathbb{E}[e^{itX}]| \leq \mathbb{E}[|e^{itX}|] = 1$.

**Proposizione**
Se $X, Y$sono due v.a. tali che $X \perp\!\!\!\perp Y$ aventi f.c. rispettivamente $\varphi_X(t)$ e $\varphi_Y(t)$, allora $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$.

**Proposizione**

1. $\varphi_X(t)$ è una funzione continua in $t$.
2. $\mathbb{E}[X^n]$ esiste finito $\Leftrightarrow \frac{d^n}{dt^n}\varphi_X(t)$ è una funzione continua in $t$.

**Proposizione**
Se la f.g.m. $M_X(t)$ esiste finita in un intorno bilatero di $t = 0$ allora anche la f.c. $\varphi_X(t)$ si esprime come serie di potenze

$$\varphi_X(t) = \sum_{n=0}^{+\infty} \frac{i^n \mathbb{E}[X^n]}{n!} t^n$$

Dunque, per ogni $n \in \mathbb{N}$ $\varphi_X(t)$ ammette derivata $n$-esima finita.

**F.c. note**

- $X \equiv a$: $\varphi_X(t) = e^{ita}$
- $\varphi_{a+bX}(t) = e^{ita}\varphi_X(bt)$
- $X \sim B(p)$: $\varphi_X(t) = (1-p) + pe^{it}$
- $X \sim Bin(n, p)$: $\varphi_X(t) = (1 - p + pe^{it})^n$
- $X \sim Geom(\lambda)$: $\varphi_X(t) = \frac{p}{1-(1-p)e^{it}}$

- $X \sim Pois(\lambda)$: $\varphi_X(t) = e^{\lambda(e^{it}-1)}$
- $X \sim Unif(0,1)$: $\varphi_X(t) = \frac{\sin(t)}{t}$
- $Y \sim Unif(a,b)$: $\varphi_X(t) = \frac{e^{ibt}-e^{iat}}{ibt-iat}$
- $X \sim \exp(\lambda)$: $\varphi_X(t) = \frac{\lambda}{\lambda - it}$
- $X \sim \mathcal{N}(0,1)$: $\varphi_X(t) = e^{-t^2/2}$
- $Y = \mu + \sigma X$: $\varphi_Y(t) = e^{i\mu t} e^{-t^2\sigma^2/2}$

**Teorema (Formula di inversione)**

1. Sia $X$ una v.a. a valori interi. Vale che
$$p_k = \int_{-\pi}^{\pi} e^{-itk}\,\varphi_X(t)\,dt$$
2. Sia $X$ una v.a. continua. Vale che
$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}\varphi_X(t)\,dt$$

**Teorema (Unicità)**
Due v.a. con la stessa f.c. hanno la stessa f.d.r.

TIPI DI CONVERGENZA

**Def.** (Convergenza in probabilità)
Sia $(\Omega, \mathcal{A}, \mathbb{P})$ uno spazio di probabilità, $\{X_n\}$ una successione di variabili aleatorie e X una v.a.
Allora se $\mathbb{P}(|X_n - X| > \varepsilon) \longrightarrow 0 \; per \; n \longrightarrow \infty,\; X_n \longrightarrow^{\mathbb{P}} X$, ovvero $X_n$ converge in probabilità in X.

**Def.** (Convergenza quasi certa)
Se l'insieme degli $\omega \in \Omega$ tali che $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ ha probabilità 1. Allora diremo che $X_n \longrightarrow^{q.c.} X$, ovvero $X_n$ converge quasi certamente a X.

**Def.** (Convergenza in distribuzione o in legge)
Sia $\{X_n\}$ è una successione di v.a. con $F_n$ le rispettive f.d.r.
Diremo che $X_n \longrightarrow^{\mathcal{L}} X$, ovvero $X_n$ converge a X in legge se e solo se $\lim_{n\to\infty} F_n(x) = F(x)$ per ogni punto $x \in \mathbb{R}$ di continuità di F.

**Teorema (di continuità di P. Lévy)**
Sia $\{X_n\}_n$ una successione di v.a. reali, e indichiamo con $\varphi_{X_i}(t)$ le rispettive funzioni caratteristiche. Allora $X_n \longrightarrow^{\mathcal{D}} X$ se e solo se $\varphi_{X_n}(t) \longrightarrow \varphi_X(t)\;\; \forall t \in \mathbb{R}$.

DISUGUAGLIANZE

**Proposizione.** (Disuguaglianza di Markov)
Sia X≥0 una v.a. il cui valore atteso esiste e sia t>0 un numero reale. Allora possiamo scrivere:
$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}$$

**Proposizione.** (Disuguaglianza di Čebyšëv)
Sia X una v.a. con media μ e varianza $\sigma^2$ finite. Allora possiamo scrivere:
$$\mathbb{P}(|X-\mu| \geq t) = \mathbb{P}((X-\mu)^2 \geq t^2) \leq \frac{V(X)}{t^2}$$
Inoltre, vale che:
$$\mathbb{P}(|X-\mu| \geq t\sigma) \leq \frac{1}{t^2}$$

APPROSSIMAZIONI

**Teorema (Legge debole dei grandi numeri – 1)**
Sia $\{X_n\}_n$ una successione di v.a. i.i.d., di media $\mu$ e varianza $\sigma^2$ finite. Allora $\forall\, \delta > 0 \; \lim_{n\to\infty} \mathbb{P}(|\overline{X_n} - \mu| > \delta) = 0$,
in altre parole $\bar{X}_n \longrightarrow^{\mathbb{P}} \mu$

**Teorema (Legge debole dei grandi numeri - 2)**
Se $\{X_n\}_n$ è una successione di v.a. i.i.d., con media $\mu$ finita, allora $\bar{X}_n \longrightarrow^{\mathcal{D}} Y \equiv \mu$. In particolare, $\bar{X}_n \longrightarrow^{\mathbb{P}} Y$.

**Teorema (Limite Centrale)**
Sia $\{X_n\}_n$ una successione di v.a. i.i.d., di media $\mu$ e varianza $\sigma^2$ finite. Sia $S_n = \sum_{k=1}^n X_k$, dunque $\mathbb{E}[S_n] = n\mu$ e $V(S_n) = n\sigma^2$.
Sia poi $S_n^* := \frac{S_n - n\mu}{\sqrt{n\sigma^2}}$ la "standardizzata di $S_n$", dunque $\mathbb{E}[S_n^*] = 0$ e $V(S_n^*) = 1$. Allora $\forall x \in \mathbb{R}$,
$$\lim_{n\to+\infty} \mathbb{P}(S_n^* \leq x) = \Phi(x)$$
dove $\Phi$ è la f.d.r. di $Z \sim \mathcal{N}(0,1)$; in simboli: $S_n^* \longrightarrow^{\mathcal{D}} Z$

**Corollario (Teorema di approssimazione normale)**
$\mathbb{P}(a < S_n^* \leq b) =$
$$= \mathbb{P}\left(\frac{a-n\mu}{\sqrt{n\sigma^2}} < S_n^* \leq \frac{b-n\mu}{\sqrt{n\sigma^2}}\right) \approx \Phi\left(\frac{b-n\mu}{\sqrt{n\sigma^2}}\right) - \Phi\left(\frac{a-n\mu}{\sqrt{n\sigma^2}}\right)$$

**Osservazione (Correzione di continuità)**
Se $X_n$ sono v.a. discrete a valori interi, allora $S_n$ è discreta a valori interi e, per $a \in \mathbb{Z}$
$\mathbb{P}(S_n \leq a) = \mathbb{P}\left(S_n \leq a + \frac{1}{2}\right) =$
$$= \mathbb{P}\left(S_n^* \leq \frac{a+\frac{1}{2}-n\mu}{\sqrt{n\sigma^2}}\right) \approx \Phi\left(\frac{a+\frac{1}{2}-n\mu}{\sqrt{n\sigma^2}}\right)$$
Il coefficiente $\frac{1}{2}$ è detto *correttore di continuità*.

**Teorema (di approssimazione di Weierstrass)**
Sia $f \in C^0([a,b])$. Allora esistono polinomi $f_n$ di grado $n$ che convergono a $f$ uniformemente in $x \in [0,1]$ per $n \to \infty$

STATISTICA

**Elementi di Statistica descrittiva**

- Un *Istogramma* è una rappresentazione grafica dei dati in cui ogni classe è rappresentata da un rettangolo, le basi sono allineate su un asse cartesiano, e la frequenza è proporzionale all'area del rettangolo.
- **Indici di posizione**
  ◦ Media campionaria: $\bar{X} = \frac{1}{n}\sum_{i=1}^n x_i$
  Per i dati raggruppati in classi, si ha $\bar{X} = \frac{1}{n}\sum_{i=1}^k c_i f_i$, dove $c_i$ è il centro dell'intervallo i-esimo, $f_i$ è la frequenza dell'elemento i-esimo.
  ◦ Mediana: è il valore che divide a metà le osservazioni ORDINATE!
  ◦ Quantile $q_\alpha$: è il valore che corrisponde al primo $\alpha\%$ dei dati. I più importanti sono i quartili.
- **Indici di dispersione**
  ◦ Scarto interquartile:
  $IQR = q_{75\%} - q_{25\%}$. È l'ampiezza dell'intervallo che contiene il 50% centrale dei dati.
  ◦ Deviazione standard:
  $$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2}$$

STIME

**Def.** (Popolazione)
Si definisce popolazione la caratteristica numerica descritta tramite v.a. con distribuzioni "note" $X \sim f(\cdot\,|\theta),\; \theta \in \Theta \subset \mathbb{R}^n$.
Dove con $f(\cdot\,|\theta)$ intendiamo che " $\cdot$ " è il parametro di $f$ e $\theta$ è l'insieme dei parametri della legge assegnata

**Def.**
Se $X_1, \ldots, X_n$ sono v.a. i.i.d. ⇒
$\underline{X} = (X_1, \ldots, X_n) \sim f_n(\bar{x}|\theta) = \prod_{k=1}^n f_n(x_k|\theta) \;\; per \; \bar{x} = (x_1, \ldots, x_n)$

**Def.** (Statistica)
Sia $\underline{\boldsymbol{X}} = (X_1, \dots, X_n)$ un vettore delle osservazioni e sia $t$ una funzione misurabile, $t: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ $m \geq 1$, allora si definisca statistica la v.a. $T$ tale che $T = t(X_1, \dots, X_n)$

**Def.** (Stimatore)
Sia $\psi(\theta): \Theta \longrightarrow \mathbb{R}^m$, funzione del parametro $\theta$.
Allora uno stimatore per $\psi$ è una statistica $T$ a valori in $\mathbb{R}^m$, tale che date le osservazioni $X_1, \dots, X_n$ si usa $T = t(X_1, \dots, X_n)$ per stimare $\psi(\theta)$

**Def.** (Stimatore asintoticamente normale)
Sia $T_n$ una famiglia di stimatori consistenti del parametro $\psi(\theta)$. Diremo che $T_n$ è asintoticamente normale se
$\sqrt{n}\big(T_n - \psi(\theta)\big) \longrightarrow^{\mathcal{D}} \mathcal{N}(0, \sigma^2(\theta))$.

**Proprietà di uno stimatore**
∘ *Corretto*
Uno stimatore si dice corretto se $\mathbb{E}^\theta[T] = \psi(\theta)$
∘ *Asintoticamente corretto*
Si dice che uno stimatore è asintoticamente corretto se
$\mathbb{E}^\theta[T_n] \longrightarrow \psi(\theta) \;\; per\ n \longrightarrow \infty$
∘ *Consistente*
Uno stimatore si dice consistente se al tendere a infinito della numerosità del campione, esso converge in probabilità al valore del parametro, ovvero
$T_n \longrightarrow^{\mathbb{P}^\theta} \psi(\theta) \;\; per\ n \longrightarrow \infty$ .
Inoltre, condizioni sufficienti affinché uno stimatore sia consistente sono che:
-sia asintoticamente corretto
- $\lim_{n\to\infty} Var(T_n) = 0$

**Def.** (Rischio quadratico)
$R_T(\theta) = \mathbb{E}^\theta\left[\big(T - \psi(\theta)\big)^2\right] =$
$= \mathbb{E}^\theta\left[\big(T - \mathbb{E}^\theta[T]\big)^2\right]\big(\mathbb{E}^\theta[T] - \psi(\theta)\big)^2$
Se $\forall\theta\ R_T(\theta) \leq R_S(\theta)$ dove $S$ è un altro stimatore, allora $T$ è *preferibile* a $S$, inoltre diremo che $S$ *non è ammissibile*.

**Metodo dei momenti**

Se $\psi(\theta) = \mathbb{E}^\theta[X^k] = m_k(\theta)$ siamo nel caso in cui cerchiamo di stimare il momento n-esimo. Alternativamente $\psi(\theta) = g\big(m_1(\theta), \dots, m_n(\theta)\big)$ è una funzione dei momenti, ad esempio la varianza. Allora useremo i momenti campionari $\widehat{m_p}$ per ottenere come stimatore di $\psi(\theta)$ la v.a.
T= $g\big(\widehat{m_1}(\theta), \dots, \widehat{m_n}(\theta)\big)$

**Def.** (Momento p-esimo campionario)
$\overline{X_p} = \frac{1}{n}\sum_{k=1}^n (X_k)^p$
$\mathbb{E}\big[\,\overline{X_p}\,\big] = \mathbb{E}^\theta[X^p]$

**Stime di massima verosimiglianza**
Indichiamo con $f_n = L(\theta|\underline{x})$ la densità congiunta delle osservazioni. Ottenuti i dati $\underline{x} = (x_1, \dots, x_n)$, consideriamo come stima di $\theta$ il valore $\hat{\theta}$ che rende massima la densità nel punto $\underline{x}$, in simboli $\hat{\theta} = \underset{\theta\in\Theta}{argmax}\, L(\theta|\underline{x})$.
Dunque, se esiste un unico punto di massimo della funzione $L(\theta|\underline{x})$, possiamo cercare di annullare la sua derivata. Anziché di $L(\theta|\underline{x})$, considereremo la funzione $l\big(\theta|\underline{x}\big) = ln(L(\theta|\underline{x}))$, la quale ha gli stessi massimi di $L(\theta|\underline{x})$. Il valore $\hat{\theta}$ deve soddisfare la seguente **equazione di verosimiglianza**

$$\sum_{k=1}^n \frac{\partial}{\partial\theta}\ln\big(f(x_k|\theta)\big) = 0$$

**Teorema**
Sotto opportune ipotesi – che non sempre verifichiamo -
$\exists!\ \hat{\theta} = MLE(\theta)$ stimatore di massima verosimiglianza.

**STIME PER INTERVALLI**

Sia $X \sim \mathcal{N}(\mu, \sigma^2)$ . Allora per ogni $\alpha \in (0,1)$ esiste $\varepsilon > 0$ tale che $\mathbb{P}\big(X \in (\mu - \varepsilon\sigma,\ \mu + \varepsilon\sigma)\big) = \alpha$. Dato che $\frac{X-\mu}{\sigma} = Z \sim \mathcal{N}(0,1)$, $\alpha = \mathbb{P}\big((-\varepsilon, \varepsilon) \ni X\big) \Leftrightarrow \varepsilon = \phi_{\frac{1+\alpha}{2}}$.

**Def.** (Intervallo di Confidenza)
Siano $(X_1, \dots, X_n)$ le osservazioni e $\psi(\theta)$ il parametro da stimare. Un intervallo di confidenza $I(X_1, \dots, X_n)$ al livello $\alpha$ è un intervallo (<u>aleatorio</u>) tale che $\mathbb{P}\big(I(X_1, \dots, X_n) \ni \psi(\theta)\big) = \alpha$

**Stima di media e varianza per campioni gaussiani**

**Primo caso.** Vogliamo stimare la media $\mu$ noto il valore della varianza $\sigma^2$.
Supponiamo che le osservazioni abbiano ciascuna una legge $\mathcal{N}(\mu, \sigma^2)$, dove $\sigma^2$ è un numero fissato e conosciuto. Si ha che $\mathbb{P}\left(\left(\bar{X} - \frac{\sigma}{\sqrt{n}}\phi_{\frac{1+\alpha}{2}} < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}}\phi_{\frac{1+\alpha}{2}}\right)\right) = \alpha$, cioè otteniamo i 3 intervalli di fiducia

- $\mathbb{P}\left(\left(\bar{X} - \frac{\sigma}{\sqrt{n}}\phi_{\frac{1+\alpha}{2}},\ \bar{X} + \frac{\sigma}{\sqrt{n}}\phi_{\frac{1+\alpha}{2}}\right) \ni \mu\right) = \alpha$ bilatero
- $\mathbb{P}\left(\left(\bar{X} - \frac{\sigma}{\sqrt{n}}\phi_\alpha,\ +\infty\right) \ni \mu\right) = \alpha$ unilatero sinistro
- $\mathbb{P}\left(\left(-\infty,\ \bar{X} + \frac{\sigma}{\sqrt{n}}\phi_\alpha\right) \ni \mu\right) = \alpha$ unilatero destro

Notiamo che $\frac{\sigma}{\sqrt{n}}$ è intesa come $\sqrt{\bar{\sigma}^2}$ ovvero della varianza "normalizzata"

**Secondo caso.** ***Leggi $\chi^2$ chi-quadro***
Vogliamo stimare la varianza usando solo media e varianza campionarie.

**Def.**
La distribuzione $\chi_k^2$ è la distribuzione di probabilità della v.a. definita come $\chi_k^2 = \sum_{i=1}^k x_i^2$, dove $x_i$ per *i=1, … ,k* sono v.a. indipendenti con distribuzione normale standard $\mathcal{N}(0,1)$. Il parametro $k$ è detto numero di gradi di libertà.

**Caratteristiche**
∘ Una generalizzazione della distribuzione di $\chi^2$ è la distribuzione Gamma: $\Gamma\left(\frac{k}{2}, \frac{1}{2}\right) = \chi^2(k)$
∘ *media:* $\mathbb{E}[x] = k$
∘ *varianza:* $V(x) = 2k$
∘ Per il teorema centrale del limite la distribuzione $\chi^2(k)$ converge ad una distribuzione normale $\mathcal{N}$ per $k$ che tende all'infinito

**Teorema**
Siano $\{X_i\}_{i=1}^k$ v.a. i.i.d. $\mathcal{N}(\mu, \sigma^2)$

a. $\sum_{i=1}^k \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(k)$
b. $\sum_{i=1}^k \left(\frac{X_i - \bar{X}}{\sigma}\right)^2 \sim \chi^2(k-1)$
c. $S^2 = \frac{1}{n-1}\sum_{i=1}^k (X_i - \bar{X})^2$ è indipendente da $\bar{X}$ e $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

**Intervalli di confidenza per la varianza di leggi $\mathcal{N}(\mu, \sigma^2)$**
La probabilità che $\sigma^2$ vi appartenga è pari a $\alpha$

- Intervallo unilatero sinistro
  $\left(\frac{(n-1)S^2}{\chi^2_{n-1,\alpha}}\ ,\ +\infty\right)$
- Intervallo unilatero destro
  $\left(0\ ,\frac{(n-1)S^2}{\chi^2_{n-1,1-\alpha}}\right)$
- Intervallo bilatero
  $\left(\frac{(n-1)S^2}{\chi^2_{n-1,\frac{1+\alpha}{2}}}\ ,\ \frac{(n-1)S^2}{\chi^2_{n-1,\frac{1-\alpha}{2}}}\right)$

**Terzo caso. Legge t di Student**
Vogliamo stimare la media $\mu$ usando solo varianza e media campionarie

**Def.**
Se $X \sim \mathcal{N}(0,1)$ e $Y \sim \chi^2(n)$ sono indipendenti, allora la v.a.
$T = \frac{X}{\sqrt{Y}}\sqrt{n}$ si dice avere *legge t di Student a n gradi di libertà.*
Per $n \to \infty$, essa converge ad una gaussiana.

**Teorema**
Sia $\{X_n\}_{i=1}^n$ una successione di v.a. i.i.d. con distribuzione $\mathcal{N}(\mu, \sigma^2)$, allora $\overline{X} \sim \mathcal{N}(\mu, {\sigma^2}/{n})$ e $Y$ sono indipendenti e

$T = \frac{\frac{\overline{X}-\mu}{\sqrt{\sigma^2/n}}}{\sqrt{Y/n-1}} \sim t(n-1)$ o, equivalentemente

$T = \frac{\overline{X}-\mu}{\sqrt{S^2/n}} \sim t(n-1)$

Da $\mathbb{P}\left(-t_{n-1,\frac{1+\alpha}{2}} < \frac{\overline{X}-\mu}{\sqrt{S^2/n}} < t_{n-1,\frac{1+\alpha}{2}}\right) = \alpha$ ricaviamo il seguente intervallo di confidenza bilatero per $\mu$:

- $\mu \in \left(\overline{X} - t_{n-1,\frac{1+\alpha}{2}}\sqrt{\frac{S^2}{n}}\ ,\ \overline{X} + t_{n-1,\frac{1+\alpha}{2}}\sqrt{\frac{S^2}{n}}\right)$

**REGRESSIONE LINEARE**
Un problema statistico molto frequente è quello in cui si considera una variabile $y$ che è funzione di altre variabili $x_1, .., x_n$ più una perturbazione aleatoria.
Prendiamo in considerazione il caso in cui tale funzione sia lineare. Parleremo di *regressione lineare*. Ciò significa che assumiamo che la variabile $y$, detta *dipendente*, si possa esprimere come

$$y = \beta_0 + \beta_1 \cdot x + \omega$$

dove $\beta_0, \beta_1$ sono parametri da determinare e ω è una perturbazione stocastica con distribuzione $\mathcal{N}(0,1)$ . Se abbiamo più osservazioni per la variabile $y$, ottenute rispetto a diversi valori di $x$, indicheremo con $y_i$ e $x_i$ tali valori. Se l'assunzione che la dipendenza sia lineare è plausibile, ci aspettiamo che per le varie osservazioni valga:

$$y_i = \beta_0 + \beta_1 \cdot x_i + \omega_i,\ \ i = 1, \dots, n$$

con $\omega_i$ indipendenti e tutti con distribuzione $\mathcal{N}(0, \sigma^2)$, con $\sigma^2$ che non dipende da $i$.
Il problema di ricavare degli stimatori per $\beta_0, \beta_1, \sigma^2$ ha diversi modi di risoluzione, tra i quali risolvere equazioni di massima verosimiglianza.

**Appendice A. Calcolo combinatorio.**

- *<u>Disposizioni</u> di n elementi distinti di classe k*
  Sono tutti i possibili gruppi che si possono formare con k elementi, presi fra gli n, tali che ogni gruppo è diverso dagli altri per gli elementi contenuti o per il loro ordine
  $D_k^n = \frac{n!}{(n-k)!}$

- *<u>Permutazioni</u> di n elementi distinti*
  Sono tutti i gruppi formati dagli n elementi che differiscono per il loro ordine
  $P^n = n!$

- *<u>Disposizioni con ripetizione</u> di n elementi distinti*
  Sono tutti i possibili gruppi che si possono formare con k elementi, anche ripetuti, presi fra gli n, tali che ogni gruppo è diverso dagli altri per gli elementi contenuti o per il loro ordine
  $D_{n,k} = n^k$

- *<u>Permutazioni con ripetizione</u>*
  Sono permutazioni di n elementi di cui $k_1, \dots, k_n$ ripetuti, dunque gruppi formati dagli n elementi che differiscono per l'ordine degli elementi distinti e il posto occupato dagli elementi ripetuti
  $P_n^{k_1,\dots,k_n} = \frac{n!}{k_1!\cdots k_n!}$

- *<u>Combinazione semplice</u> di n elementi distinti di classe k (con k ≤ n)*
  Sono tutti i possibili gruppi che si possono formare con k elementi, presi fra gli n, e tali che ogni gruppo è diverso dagli altri per almeno un elemento contenuto.
  $C_{n,k} = \binom{n}{k} = \frac{n!}{(n-k)!k!}$

- *<u>Combinazioni con ripetizione</u> di n elementi distinti di classe k (con k ⋛ n)*
  Sono tutti i possibili gruppi che si possono formare con k elementi, presi fra gli n; ogni elemento di un gruppo può essere ripetuto fino a k volte, non interessa l'ordine in cui gli elementi si presentano e in ciascun gruppo è diverso il numero delle volte in cui un elemento compare.
  $C'_{n,k} = C_{n+k-1,k} = \binom{n+k-1}{k} = \binom{n+k-1}{n-1}$

**Appendice B .**
***Somme geometriche***
Ricordiamo

$$1 + x + \ldots + x^n = \frac{1-x^{n+1}}{1-x},$$

da cui si ricava subito per $|x| < 1$ il valore della somma della serie

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

e derivando a destra e sinistra rispetto a x si ottiene

$$\sum_{n=0}^{\infty}(n+1)x^n = \sum_{n=1}^{\infty} nx^{n-1} = \frac{1}{(1-x)^2}$$

***Serie per Poisson***
$\sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} = e^{\lambda}$

Tavola 1 *(segue)*: Funzione di ripartizione della Variabile Casuale Normale Standardizzata

| $z$ | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| 0.1 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| 0.2 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| 0.3 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| 0.4 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| 0.5 | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |
| 0.6 | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | 0.7422 | 0.7454 | 0.7486 | 0.7517 | 0.7549 |
| 0.7 | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7704 | 0.7734 | 0.7764 | 0.7794 | 0.7823 | 0.7852 |
| 0.8 | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | 0.8023 | 0.8051 | 0.8078 | 0.8106 | 0.8133 |
| 0.9 | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | 0.8289 | 0.8315 | 0.8340 | 0.8365 | 0.8389 |
| 1.0 | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.8554 | 0.8577 | 0.8599 | 0.8621 |
| 1.1 | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8810 | 0.8830 |
| 1.2 | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | 0.8944 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| 1.3 | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| 1.4 | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |
| 1.5 | 0.9332 | 0.9345 | 0.9357 | 0.9370 | 0.9382 | 0.9394 | 0.9406 | 0.9418 | 0.9429 | 0.9441 |
| 1.6 | 0.9452 | 0.9463 | 0.9474 | 0.9484 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| 1.7 | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| 1.8 | 0.9641 | 0.9649 | 0.9656 | 0.9664 | 0.9671 | 0.9678 | 0.9686 | 0.9693 | 0.9699 | 0.9706 |
| 1.9 | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9738 | 0.9744 | 0.9750 | 0.9756 | 0.9761 | 0.9767 |
| 2.0 | 0.9772 | 0.9778 | 0.9783 | 0.9788 | 0.9793 | 0.9798 | 0.9803 | 0.9808 | 0.9812 | 0.9817 |
| 2.1 | 0.9821 | 0.9826 | 0.9830 | 0.9834 | 0.9838 | 0.9842 | 0.9846 | 0.9850 | 0.9854 | 0.9857 |
| 2.2 | 0.9861 | 0.9864 | 0.9868 | 0.9871 | 0.9875 | 0.9878 | 0.9881 | 0.9884 | 0.9887 | 0.9890 |
| 2.3 | 0.9893 | 0.9896 | 0.9898 | 0.9901 | 0.9904 | 0.9906 | 0.9909 | 0.9911 | 0.9913 | 0.9916 |
| 2.4 | 0.9918 | 0.9920 | 0.9922 | 0.9925 | 0.9927 | 0.9929 | 0.9931 | 0.9932 | 0.9934 | 0.9936 |
| 2.5 | 0.9938 | 0.9940 | 0.9941 | 0.9943 | 0.9945 | 0.9946 | 0.9948 | 0.9949 | 0.9951 | 0.9952 |
| 2.6 | 0.9953 | 0.9955 | 0.9956 | 0.9957 | 0.9959 | 0.9960 | 0.9961 | 0.9962 | 0.9963 | 0.9964 |
| 2.7 | 0.9965 | 0.9966 | 0.9967 | 0.9968 | 0.9969 | 0.9970 | 0.9971 | 0.9972 | 0.9973 | 0.9974 |
| 2.8 | 0.9974 | 0.9975 | 0.9976 | 0.9977 | 0.9977 | 0.9978 | 0.9979 | 0.9979 | 0.9980 | 0.9981 |
| 2.9 | 0.9981 | 0.9982 | 0.9982 | 0.9983 | 0.9984 | 0.9984 | 0.9985 | 0.9985 | 0.9986 | 0.9986 |
| 3.0 | 0.9987 | 0.9987 | 0.9987 | 0.9988 | 0.9988 | 0.9989 | 0.9989 | 0.9989 | 0.9990 | 0.9990 |
| 3.1 | 0.9990 | 0.9991 | 0.9991 | 0.9991 | 0.9992 | 0.9992 | 0.9992 | 0.9992 | 0.9993 | 0.9993 |
| 3.2 | 0.9993 | 0.9993 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9995 | 0.9995 | 0.9995 |
| 3.3 | 0.9995 | 0.9995 | 0.9995 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9997 |
| 3.4 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9998 |

Tavola 1a: Valori critici della Variabile Casuale Normale Standardizzata. $P(Z \geq z_\alpha) = \alpha$.

| $\alpha$ | 0.10 | 0.05 | 0.025 | 0.01 | 0.005 | 0.001 | 0.0005 | 0.0001 | |
|---|---|---|---|---|---|---|---|---|---|
| $z_\alpha$ | 1.2816 | 1.6449 | 1.9600 | 2.3263 | 2.5758 | 3.0902 | 3.2905 | 3.7190 | |
| $\alpha$ | 0.00009 | 0.00008 | 0.00007 | 0.00006 | 0.00005 | 0.00004 | 0.00003 | 0.00002 | 0.00001 |
| $z_\alpha$ | 3.7455 | 3.7750 | 3.8082 | 3.8461 | 3.8906 | 3.9444 | 4.0128 | 4.1075 | 4.2649 |

# B – Distribuzione Chi-quadrato

Quantili della distribuzione $\chi^2_\nu$

| | Probabilità di valori minori | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $\nu$ | 0.005 | 0.01 | 0.05 | 0.1 | 0.9 | 0.95 | 0.975 | 0.99 | 0.995 |
| 1 | 0.00 | 0.00 | 0.00 | 0.02 | 2.71 | 3.84 | 5.02 | 6.63 | 7.88 |
| 2 | 0.01 | 0.02 | 0.10 | 0.21 | 4.61 | 5.99 | 7.38 | 9.21 | 10.60 |
| 3 | 0.07 | 0.11 | 0.35 | 0.58 | 6.25 | 7.81 | 9.35 | 11.34 | 12.84 |
| 4 | 0.21 | 0.30 | 0.71 | 1.06 | 7.78 | 9.49 | 11.14 | 13.28 | 14.86 |
| 5 | 0.41 | 0.55 | 1.15 | 1.61 | 9.24 | 11.07 | 12.83 | 15.09 | 16.75 |
| 6 | 0.68 | 0.87 | 1.64 | 2.20 | 10.64 | 12.59 | 14.45 | 16.81 | 18.55 |
| 7 | 0.99 | 1.24 | 2.17 | 2.83 | 12.02 | 14.07 | 16.01 | 18.48 | 20.28 |
| 8 | 1.34 | 1.65 | 2.73 | 3.49 | 13.36 | 15.51 | 17.53 | 20.09 | 21.95 |
| 9 | 1.73 | 2.09 | 3.33 | 4.17 | 14.68 | 16.92 | 19.02 | 21.67 | 23.59 |
| 10 | 2.16 | 2.56 | 3.94 | 4.87 | 15.99 | 18.31 | 20.48 | 23.21 | 25.19 |
| 11 | 2.60 | 3.05 | 4.57 | 5.58 | 17.28 | 19.68 | 21.92 | 24.72 | 26.76 |
| 12 | 3.07 | 3.57 | 5.23 | 6.30 | 18.55 | 21.03 | 23.34 | 26.22 | 28.30 |
| 13 | 3.57 | 4.11 | 5.89 | 7.04 | 19.81 | 22.36 | 24.74 | 27.69 | 29.82 |
| 14 | 4.07 | 4.66 | 6.57 | 7.79 | 21.06 | 23.68 | 26.12 | 29.14 | 31.32 |
| 15 | 4.60 | 5.23 | 7.26 | 8.55 | 22.31 | 25.00 | 27.49 | 30.58 | 32.80 |
| 16 | 5.14 | 5.81 | 7.96 | 9.31 | 23.54 | 26.30 | 28.85 | 32.00 | 34.27 |
| 17 | 5.70 | 6.41 | 8.67 | 10.09 | 24.77 | 27.59 | 30.19 | 33.41 | 35.72 |
| 18 | 6.26 | 7.01 | 9.39 | 10.86 | 25.99 | 28.87 | 31.53 | 34.81 | 37.16 |
| 19 | 6.84 | 7.63 | 10.12 | 11.65 | 27.20 | 30.14 | 32.85 | 36.19 | 38.58 |
| 20 | 7.43 | 8.26 | 10.85 | 12.44 | 28.41 | 31.41 | 34.17 | 37.57 | 40.00 |
| 21 | 8.03 | 8.90 | 11.59 | 13.24 | 29.62 | 32.67 | 35.48 | 38.93 | 41.40 |
| 22 | 8.64 | 9.54 | 12.34 | 14.04 | 30.81 | 33.92 | 36.78 | 40.29 | 42.80 |
| 23 | 9.26 | 10.20 | 13.09 | 14.85 | 32.01 | 35.17 | 38.08 | 41.64 | 44.18 |
| 24 | 9.89 | 10.86 | 13.85 | 15.66 | 33.20 | 36.42 | 39.36 | 42.98 | 45.56 |
| 25 | 10.52 | 11.52 | 14.61 | 16.47 | 34.38 | 37.65 | 40.65 | 44.31 | 46.93 |
| 26 | 11.16 | 12.20 | 15.38 | 17.29 | 35.56 | 38.89 | 41.92 | 45.64 | 48.29 |
| 27 | 11.81 | 12.88 | 16.15 | 18.11 | 36.74 | 40.11 | 43.19 | 46.96 | 49.64 |
| 28 | 12.46 | 13.56 | 16.93 | 18.94 | 37.92 | 41.34 | 44.46 | 48.28 | 50.99 |
| 29 | 13.12 | 14.26 | 17.71 | 19.77 | 39.09 | 42.56 | 45.72 | 49.59 | 52.34 |
| 30 | 13.79 | 14.95 | 18.49 | 20.60 | 40.26 | 43.77 | 46.98 | 50.89 | 53.67 |
| 31 | 14.46 | 15.66 | 19.28 | 21.43 | 41.42 | 44.99 | 48.23 | 52.19 | 55.00 |
| 32 | 15.13 | 16.36 | 20.07 | 22.27 | 42.58 | 46.19 | 49.48 | 53.49 | 56.33 |
| 33 | 15.82 | 17.07 | 20.87 | 23.11 | 43.75 | 47.40 | 50.73 | 54.78 | 57.65 |
| 34 | 16.50 | 17.79 | 21.66 | 23.95 | 44.90 | 48.60 | 51.97 | 56.06 | 58.96 |
| 35 | 17.19 | 18.51 | 22.47 | 24.80 | 46.06 | 49.80 | 53.20 | 57.34 | 60.27 |
| 36 | 17.89 | 19.23 | 23.27 | 25.64 | 47.21 | 51.00 | 54.44 | 58.62 | 61.58 |
| 37 | 18.59 | 19.96 | 24.07 | 26.49 | 48.36 | 52.19 | 55.67 | 59.89 | 62.88 |
| 38 | 19.29 | 20.69 | 24.88 | 27.34 | 49.51 | 53.38 | 56.90 | 61.16 | 64.18 |
| 39 | 20.00 | 21.43 | 25.70 | 28.20 | 50.66 | 54.57 | 58.12 | 62.43 | 65.48 |
| 40 | 20.71 | 22.16 | 26.51 | 29.05 | 51.81 | 55.76 | 59.34 | 63.69 | 66.77 |
| 45 | 24.31 | 25.90 | 30.61 | 33.35 | 57.51 | 61.66 | 65.41 | 69.96 | 73.17 |
| 50 | 27.99 | 29.71 | 34.76 | 37.69 | 63.17 | 67.50 | 71.42 | 76.15 | 79.49 |
| 60 | 35.53 | 37.48 | 43.19 | 46.46 | 74.40 | 79.08 | 83.30 | 88.38 | 91.95 |
| 70 | 43.28 | 45.44 | 51.74 | 55.33 | 85.53 | 90.53 | 95.02 | 100.43 | 104.21 |
| 80 | 51.17 | 53.54 | 60.39 | 64.28 | 96.58 | 101.88 | 106.63 | 112.33 | 116.32 |
| 90 | 59.20 | 61.75 | 69.13 | 73.29 | 107.57 | 113.15 | 118.14 | 124.12 | 128.30 |
| 100 | 67.33 | 70.06 | 77.93 | 82.36 | 118.50 | 124.34 | 129.56 | 135.81 | 140.17 |

**Tavola dei quantili $t_\alpha(n)$ della legge $t$ di Student.**

$$P(T < t_\alpha(n)) = \alpha \quad \text{con } T \sim t(n)$$

| | 0.9 | 0.95 | 0.975 | 0.99 | 0.995 |
|---|---|---|---|---|---|
| 1 | 3.0777 | 6.3138 | 12.7062 | 31.8205 | 63.6567 |
| 2 | 1.8856 | 2.9200 | 4.3027 | 6.9646 | 9.9248 |
| 3 | 1.6377 | 2.3534 | 3.1824 | 4.5407 | 5.8409 |
| 4 | 1.5332 | 2.1318 | 2.7764 | 3.7469 | 4.6041 |
| 5 | 1.4759 | 2.0150 | 2.5706 | 3.3649 | 4.0321 |
| 6 | 1.4398 | 1.9432 | 2.4469 | 3.1427 | 3.7074 |
| 7 | 1.4149 | 1.8946 | 2.3646 | 2.9980 | 3.4995 |
| 8 | 1.3968 | 1.8595 | 2.3060 | 2.8965 | 3.3554 |
| 9 | 1.3830 | 1.8331 | 2.2622 | 2.8214 | 3.2498 |
| 10 | 1.3722 | 1.8125 | 2.2281 | 2.7638 | 3.1693 |
| 11 | 1.3634 | 1.7959 | 2.2010 | 2.7181 | 3.1058 |
| 12 | 1.3562 | 1.7823 | 2.1788 | 2.6810 | 3.0545 |
| 13 | 1.3502 | 1.7709 | 2.1604 | 2.6503 | 3.0123 |
| 14 | 1.3450 | 1.7613 | 2.1448 | 2.6245 | 2.9768 |
| 15 | 1.3406 | 1.7531 | 2.1314 | 2.6025 | 2.9467 |
| 16 | 1.3368 | 1.7459 | 2.1199 | 2.5835 | 2.9208 |
| 17 | 1.3334 | 1.7396 | 2.1098 | 2.5669 | 2.8982 |
| 18 | 1.3304 | 1.7341 | 2.1009 | 2.5524 | 2.8784 |
| 19 | 1.3277 | 1.7291 | 2.0930 | 2.5395 | 2.8609 |
| 20 | 1.3253 | 1.7247 | 2.0860 | 2.5280 | 2.8453 |
| 21 | 1.3232 | 1.7207 | 2.0796 | 2.5176 | 2.8314 |
| 22 | 1.3212 | 1.7171 | 2.0739 | 2.5083 | 2.8188 |
| 23 | 1.3195 | 1.7139 | 2.0687 | 2.4999 | 2.8073 |
| 24 | 1.3178 | 1.7109 | 2.0639 | 2.4922 | 2.7969 |
| 25 | 1.3163 | 1.7081 | 2.0595 | 2.4851 | 2.7874 |
| 26 | 1.3150 | 1.7056 | 2.0555 | 2.4786 | 2.7787 |
| 27 | 1.3137 | 1.7033 | 2.0518 | 2.4727 | 2.7707 |
| 28 | 1.3125 | 1.7011 | 2.0484 | 2.4671 | 2.7633 |
| 29 | 1.3114 | 1.6991 | 2.0452 | 2.4620 | 2.7564 |
| 30 | 1.3104 | 1.6973 | 2.0423 | 2.4573 | 2.7500 |
| 40 | 1.3031 | 1.6839 | 2.0211 | 2.4233 | 2.7045 |
| 50 | 1.2987 | 1.6759 | 2.0086 | 2.4033 | 2.6778 |
| 60 | 1.2958 | 1.6706 | 2.0003 | 2.3901 | 2.6603 |
| 70 | 1.2938 | 1.6669 | 1.9944 | 2.3808 | 2.6479 |
| 80 | 1.2922 | 1.6641 | 1.9901 | 2.3739 | 2.6387 |
| 90 | 1.2910 | 1.6620 | 1.9867 | 2.3685 | 2.6316 |
| 100 | 1.2901 | 1.6602 | 1.9840 | 2.3642 | 2.6259 |
| 110 | 1.2893 | 1.6588 | 1.9818 | 2.3607 | 2.6213 |
| 120 | 1.2886 | 1.6577 | 1.9799 | 2.3578 | 2.6174 |