



UniCT – DMI

Metodi Matematici Statistici

ANNO ACCADEMICO

2021/22

Autori:

Alessio Tudisco

Sommario

Statistica Descrittiva.....	6
Introduzione (Definizione di dati e caratteristiche dei dati).....	6
Carattere dei dati (Carattere quantitativi discreto e continuo, Carattere non numerico).....	6
Carattere quantitativi discreto.....	6
Carattere quantitativi continuo (compatibile con carattere discreto con numerosità delle modalità elevato)	6
Carattere qualitativo	7
Rappresentazioni grafiche dei dati	7
Rappresentazione grafiche per caratteri qualitativi	7
Rappresentazione grafiche per caratteri quantitativi.....	7
Indici di tendenza centrale.....	8
Media (\bar{x}) e Media pesata \bar{x}_p	8
Mediana (x).....	8
Moda	8
Quantili, Percentili e Quartili.....	8
Calcolo dei Quartili.....	8
Indici di variabilità	9
Varianza (s^2)	9
Scarto quadratico medio (Deviazione standard)	9
Scarto medio assoluto.....	9
Coefficiente di variazione.....	10
Dati raggruppati per classi	10
Istogramma normalizzato	10
Media, Varianza, Mediana e Quartili per le classi.....	10
Indici di forma	10
Asimmetria.....	10
Indice di asimmetria (sk)	10
Curtosi	11
Indice di curtosi (K)	11
La correlazione tra due serie di Dati	11
Scatterplot.....	11
Analisi di regressione per una serie di dati	11
Metodo dei minimi quadrati.....	11
Regressione lineare (e coefficiente di correlazione lineare).....	12
Breve Introduzione al Calcolo Combinatorio	12
Fattoriale	12
Coefficiente Binomiale.....	12
Permutazioni	12
Permutazioni semplici (o senza ripetizione)	13

Permutazioni con ripetizione	13
Permutazioni circolari	13
Disposizioni	13
Disposizioni semplici (o senza ripetizione).....	13
Disposizioni con ripetizione	13
Combinazioni	14
Combinazioni semplici (o senza ripetizione).....	14
Combinazioni con ripetizione.....	14
Elementi di Probabilità.....	14
Spazio Campione ed Eventi (e relazione con le operazioni di insiemistica)	14
Famiglia di eventi e sigma algebra (<i>σalgebra</i>).....	14
Probabilità secondo Kolmogorov	15
Spazi di probabilità.....	15
Spazi di probabilità finiti e Probabilità Classica.....	15
Spazi di probabilità infiniti	15
Spazi di probabilità infiniti numerabili	15
Spazi di probabilità infiniti continui	16
Probabilità della somma logica di eventi $PE \cup F$	16
Probabilità condizionata $PA B$ e rapporto di indipendenza fra eventi	16
Probabilità del prodotto logico di eventi $PA \cap B$	17
Teorema della Catena	17
Eventi stocasticamente indipendenti	17
Partizione dello spazio di probabilità Ω	17
Probabilità totale (Legge delle alternative)	17
Teorema di Bayes.....	17
Variabili aleatorie	18
Variabile aleatoria discreta	18
Distribuzione discreta di probabilità (e densità delle variabili aleatorie discrete)	18
Funzione di ripartizione	18
Variabili aleatorie continue, densità e probabilità	19
Variabile Aleatoria in funzione di un'altra variabile aleatoria	19
Indici di tendenza centrale e variabilità per le variabili aleatorie.....	19
Proprietà degli indici di tendenza centrale	19
Quantili per le variabili aleatorie	20
Disuguaglianze di Markov e di Cebicev.....	20
Momenti di una variabile aleatoria.....	20
Funzione generatrice di momenti.....	20
Funzione caratteristica di variabile aleatoria.....	20
Vettori aleatori.....	20

Densità congiunte e marginali per variabili aleatorie discrete	21
Funzione di ripartizione congiunta	21
Funzione di ripartizione marginale	21
Densità congiunte e marginali per vettori discreti	21
Densità congiunte e marginali per vettori continui	21
Variabili aleatorie indipendenti	21
Densità della somma di variabili aleatorie	22
Valore Atteso per un vettore aleatorie	22
Studio dell'andamento di due variabili aleatorie (covarianza e coefficiente di correlazione)	22
Distribuzioni Notevoli	23
Distribuzione di Bernulli ($X \sim Berp$) e Schema di Bernulli	23
Schema di Bernulli	23
Distribuzione Binomiale ($X \sim Bin(n, p)$)	23
Distribuzione Geometrica ($X \sim Geo(p)$)	23
Distribuzione Pascal o Binomiale Negativa ($X \sim Bin(n, p)$)	23
Distribuzione Poisson ($X \sim Pois(\lambda)$), Schema e Processo di Poisson	24
Schema di Poisson	24
Processo di Poisson	24
Distribuzione Esponenziale ($X \sim Exp(\lambda)$)	24
Distribuzione Gamma ($X \sim \Gamma(n, \lambda)$)	25
Distribuzione di Weibul ($X \sim Wei(n, \lambda)$) e Funzione Rischio	25
Distribuzione di Gauss o Normale ($X \sim Norm(\mu, \sigma^2)$) e Normale Standardizzata	25
Distribuzione Uniforme ($X \sim U(a, b)$)	25
Distribuzione X^2 ($X \sim X^2(n)$)	26
Distribuzione T di Student ($X \sim t(n)$)	26
Rapporto fra le distribuzioni	26
Convergenze	26
Convergenza di una distribuzione	26
Convergenza di in probabilità	27
Convergenza quasi certa	27
Convergenza in media quadratica	27
Rapporto fra le convergenze	27
Legge debole dei grandi numeri	27
Legge forte dei grandi numeri	27
Legge centrale del limite	27
Stima dei Parametri	28
Campione casuale	28
Media Campionaria (e Valore atteso e varianza della media campionaria)	28
Momento Campionario di ordine k	28

Varianza Campionaria (e valore atteso della varianza campionaria).....	28
Distribuzione della media campionaria con popolazione distribuita secondo normale	28
Statistica.....	29
Stimatore	29
Correttezza e Correttezza Asintotica di uno stimatore.....	29
Consistenza in media quadratica e in probabilità di uno stimatore	29
Asintotica normalità.....	29
Costruzione di una stima: Metodo della massima verosomiglianza.....	29
Verifica di Ipotesi	29
Struttura di un test d'ipotesi.....	29
Test parametrici riguardanti una sola popolazione	30
Test riguardante la media di una normale.....	30
Z-Test (Caso $\sigma^2 = v \equiv \text{noto}$)	30
T-Test (Caso $\sigma^2 = ? \equiv \text{incognito}$)	31
Test riguardante la varianza di una normale	31
Caso $\mu = m \equiv \text{noto}$	32
Caso $\mu = ? \equiv \text{incognito}$	32
Test Asintotico su p (probabilità di successo di un processo di Bernulli)	33
Test parametrici di confronto fra due popolazioni	33
Caso σ_x^2 e σ_y^2 note.....	33
Caso σ_x^2 e σ_y^2 note e uguali	34
Formulario.....	35
Mathlab Compendium	39
Esercizi Compendium.....	40
Distribuzioni	40
Distribuzione Binomiale	40
Distribuzione Geometria.....	41
Distribuzione di Poisson.....	41
Stima dei parametri	41
Individuare intervallo (limiti superiori e inferiori)	41
Regressione Lineare	41

Statistica Descrittiva

Introduzione (Definizione di dati e caratteristiche dei dati)

Quando si raccolgono dei dati su una popolazione o su un campione, i valori ottenuti si presentano come un insieme di dati disordinati. I dati che non sono stati organizzati o elaborati sono chiamati *dati grezzi*.

Con il termine *statistica descrittiva* si intende una raccolta di metodi e strumenti matematici atti a organizzare i dati al fine di descrivere in maniera comprensibile le informazioni dagli stessi dati.

Le tecniche di organizzazione dei dati variano a seconda delle caratteristiche degli elementi su cui è svolta l'indagine:

- Si parla di *caratteri qualitativi* quando i dati sono di natura non numerica;
- Si parla di *caratteri quantitativi* quando i dati sono di natura numerica, che si distinguono in:
 - *Discreti* se assumono un numero limitato di valori;
 - *Continui* se assumono qualsiasi valore reale in un certo intervallo;

Carattere dei dati (Carattere quantitativi discreto e continuo, Carattere non numerico)

Si definisce:

$$\text{insieme di dati di numerosità } n \equiv E = \{x_1, x_2, \dots, x_n\}$$

Carattere quantitativi discreto

Quando i dati hanno un carattere quantitativo discreto, la tecnica consiste nel raggruppare i dati considerando l'insieme di tutti i valori assumibili, chiamati *modalità del carattere*.

Sia N il numero dei valori assumibili dai dati dell'insieme E , allora si definisce:

$$\text{insieme delle modalità} \equiv S = \{s_1, s_2, \dots, s_N\}$$

Definiamo:

- *Frequenza assoluta* $\equiv f_j = n^\circ \text{ di elementi di } E \text{ con valore } s_j \rightarrow j = 1, 2, \dots, N$;
- *Frequenza relativa* $\equiv p_j = \frac{f_j}{n} \rightarrow n = \text{numerosità di } E$;
- *Frequenza cumulata assoluta* $\equiv F_j = \sum_k f_k \rightarrow k \text{ tale che } s_k \leq s_j \text{ e } j = 1, 2, \dots, N$;
- *Frequenza cumulata relativa* $\equiv P_j = \sum_k p_k \rightarrow k \text{ tale che } s_k \leq s_j \text{ e } j = 1, 2, \dots, N$;

Le frequenze cumulate permettono di raggruppare i dati in modo da capire quando essi sono *minori di o maggiori di*.

Carattere quantitativi continuo (compatibile con carattere discreto con numerosità delle modalità elevato)

Quando i dati hanno carattere quantitativo continuo, o discreto con un numero elevato di modalità, la tecnica consiste nel considerare dei sottoinsiemi di dell'*insieme delle modalità* S . Definiamo:

- *Classe* C un qualunque sottoinsieme di S : $C_i \subseteq S \rightarrow i = 1, 2, \dots, k \in \mathbb{N}$
- *Partizione di S* ogni famiglia di classi tra loro disgiunte la cui unione sia S :

$$C_i \cap C_j = \emptyset \rightarrow \forall i \neq j; \text{ tale che: } \bigcup_{i=1}^k C_i = S$$

Ad ogni classe si associano diverse grandezze che le caratterizzano:

- I *confini superiore e inferiore*, che sono i valori di massimo e minimo della classe;
- L'*ampiezza*, che è la differenza tra il confine superiore e inferiore;
- Il *valore centrale*, che è la media tra i due confini;
- La *frequenza assoluta*, che è il numero di elementi che appartengono alla classe;
- La *frequenza relativa*, che è la percentuale di elementi che appartengono alla classe;

Per la suddivisione in classi si seguono alcune regole pratiche, poiché un numero elevato di classi potrebbe rendere illeggibile il risultato:

- Il numero di classi dovrebbe idealmente essere compreso fra 5 e 15, ma definisce un numero di classi $k \simeq \sqrt{n}$, dove n è il numero dei dati;

- L'ampiezza delle classi deve essere $a = \frac{R}{k}$, dove R è il *campo di variazione dei dati*, ovvero la differenza fra il valore massimo e minimo assunto dai dati;

Carattere qualitativo

Quando i dati hanno carattere qualitativo, la tecnica consiste nel raggruppare i dati in classi che non sono insiemi numerici ma che forniscono una partizione dell'insieme.

Un esempio può essere che in uno stabilimento vi sono dei malfunzionamenti relativi a una specifica macchina, potremmo quindi organizzare le classi raggruppando gli eventi a seconda delle possibili cause.

Rappresentazioni grafiche dei dati

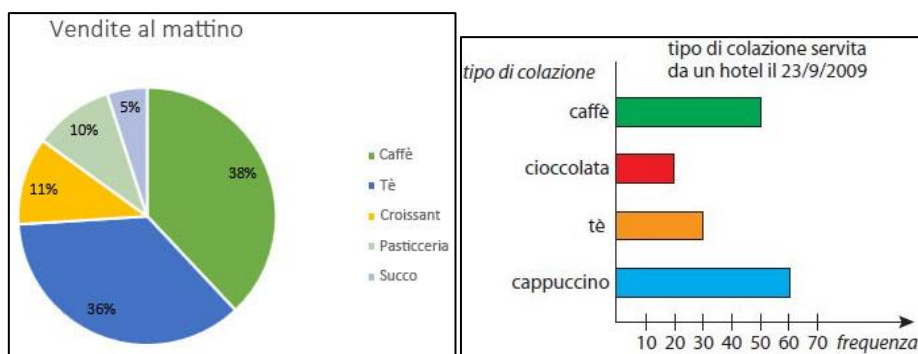
Le rappresentazioni grafiche hanno lo scopo di fornire in maniera rapida le caratteristiche essenziali del fenomeno studiato, permettendo a volte di individuare eventuali relazioni fra le caratteristiche di distribuzione.

Le rappresentazioni grafiche sono basate su una proporzionalità fra le frequenze (assolute o relative) e grandezze geometriche (aree o lunghezze) che vengono utilizzate per rappresentare il fenomeno.

Non vi sono regole fisse per la scelta di una specifica rappresentazione grafica, l'importante è garantire una rapida percezione del fenomeno studiato.

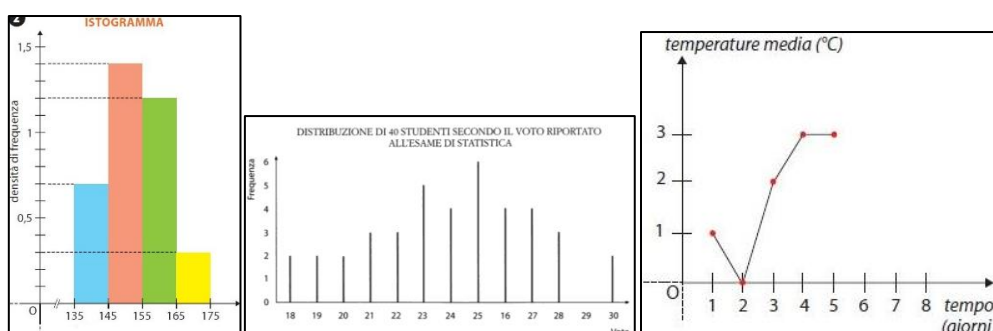
Rappresentazione grafiche per caratteri qualitativi

- *Diagramma a torta*: in cui a ciascuna modalità x_i si associa un settore circolare avente area proporzionale alla sua frequenza f_i ;
- *Diagramma a barre*: in cui a ciascuna modalità x_i si associa un rettangolo avente base costante ed un'altezza proporzionale alla frequenza f_i ;



Rappresentazione grafiche per caratteri quantitativi

- *Istogramma (caratteri continui)*: un plurirettangolo avente basi proporzionale all'ampiezza delle classi e aree proporzionali alla frequenza. Poiché nell'istogramma le aree dei singoli rettangoli sono proporzionali alle frequenze delle rispettive classi, l'altezza h_i del rettangolo della classe i -esima deve essere proporzionale al rapporto fra la frequenza della classe e la corrispondente ampiezza: $h_i = \frac{f_i}{a}$;
- *Diagrammi a segmenti (caratteri discreti)*: un grafico cartesiano in cui, in corrispondenza di ciascuna modalità x_i , si riporta un segmento di altezza proporzionale alla corrispondente frequenza relativa oppure frequenza assoluta;
- *Diagrammi cartesiani*: dove si esprime la dipendenza temporale di un fenomeno quantitativo (asse delle ordinate) in funzione del tempo (asse delle ascisse);



Indici di tendenza centrale

Gli indici di tendenza centrale sono delle misure in grado di sintetizzare i valori assunti dai dati in un solo valore numerico.

Media (\bar{x}) e Media pesata (\bar{x}_p)

$$media \equiv \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \quad media \text{ pesata} \equiv \bar{x}_p = \frac{1}{p} \sum_{i=1}^n p_i x_i \rightarrow p = \sum_{i=1}^n p_i$$

La *media* è quel punto che dista di meno da tutti i punti della serie di dati $\{x_i\}$. La *media pesata* è una variante della media in cui ogni punto ha un'influenza differente nel calcolo.

La media rende minima la funzione: $f(x) = \sum_{i=1}^n (x - x_i)^2$, dimostrazione:

- Dalla derivata prima si ottiene:

$$f'(x) = 2 \sum_{i=1}^n (x - x_i) = 0 \rightarrow \sum_{i=1}^n x - \sum_{i=1}^n x_i = 0 \rightarrow nx = \sum_{i=1}^n x_i \rightarrow x = \frac{1}{n} \sum_{i=1}^n x_i$$

- Tramite il *teorema di Fermat* dimostriamo che in \bar{x} $f(x)$ è minima:

$$f''(x) = 2 \sum_{i=1}^n 1 = 2n > 0$$

Si ottiene che \bar{x} è punto minimo di $f(x)$.

Mediana (\hat{x})

Data una serie di dati $\{x_i\}$ *ordinata in ordine crescente*, si ha che la mediana, indicata con \hat{x} , è *l'elemento che sta in mezzo*:

- Se la numerosità è dispari, la mediana è l'elemento di posto: $\frac{n+1}{2}$;
- Se la numerosità è pari, la mediana non è univocamente determinata:
 - una delle principali approssimazioni è la *media aritmetica dei due valori centrali aventi posto*:
$$\frac{n}{2} e \frac{n}{2} + 1 \rightarrow \hat{x} = \frac{\left(\left[\frac{n}{2}\right] + \left[\frac{n}{2} + 1\right]\right)}{2};$$
 - un'altra approssimazione si ottiene usando l'interpolazione lineare (calcolo basato sulla retta passante fra due punti), a partire da due valori centrali;

Moda

Data una serie di dati $\{x_i\}$ *la moda* è quel valore che si ripete più volte in essa. La moda fornisce un andamento qualitativo dei dati e non è garantito che sia un unico numero (*così come accade nella distribuzione bi-multi modale*).

Quantili, Percentili e Quartili

Data una serie di dati $\{x_i\}$ *ordinata in ordine crescente*, definiamo:

- Quantile di ordine* $a \in [0,1] \equiv q_a$, il valore per cui alla sua sinistra compare il $100 * a\%$ delle osservazioni e alla sua destra il restante $100 * (1 - a)\%$;
- Percentile* $\equiv q_a$, qualora l'ordine del quantile fosse espresso in percentuale;

Definiamo inoltre *primo quartile*, *secondo quartile* e *terzo quartile* (indicati con Q_1 , Q_2 e Q_3) i quantili aventi rispettivamente ordine di 0.25, 0.5, 0.75 tali che:

$$Q_1 \equiv q_{0.25} \quad , \quad Q_2 \equiv q_{0.5} \quad , \quad Q_3 \equiv q_{0.75}$$

Il secondo quartile coincide con la mediana.

Calcolo dei Quartili

Analogamente al concetto di mediana, anche i quartili ed i percentili non sono univocamente determinati. Per calcolarli vi sono 2 tecniche:

- Media aritmetica, tecnica analoga al calcolo della mediana:*
 - Fissiamo $a = 0.25, 0.5, 0.75$ e calcoliamo $a(n + 1) \rightarrow n \equiv \text{numerosità eventi}$;

- Se $a(n+1) \equiv m \in N$ (numero naturale) allora $Q_a = x_m$;
 - Se $a(n+1) \equiv m \notin N$ (numero naturale) allora si prende la parte intera, indicata formalmente con $[m]$ ma che chiameremo m , e si calcola la media aritmetica: $Q_a = \frac{x_m + x_{m+1}}{2}$;
2. Interpolazione lineare: supponendo di avere due punti $P_1(x_1, y_1)$ e $P_2(x_2, y_2)$ si può stimare il valore di y^* che corrisponde a $x^* \in]x_1, x_2[$ tramite la retta passante, detta *retta interpolante*, fra i due punti P_1 e P_2 :

$$y^* = y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x^* - x_1)$$

- Fissiamo $a = 0.25, 0.5, 0.75$ e calcoliamo $a(n+1) \rightarrow n \equiv$ *numerosità eventi*;
- Se $a(n+1) \equiv m \in N$ (numero naturale) allora $Q_a = x_m$;
- Se $x^* \equiv a(n+1) \notin N$ (numero naturale) allora si utilizza la retta interpolante;
- Si prende la parte intera, indicata formalmente con $[x^*]$ ma che chiameremo m , e la parte frazionaria, indicata con $\beta = x^* - m$, tale che $P_1(m, x_m)$ e $P_2(m+1, x_{m+1})$ e ottenere:

$$Q_a = x_m + \frac{(x_{m+1} - x_m)}{m+1-m}(x^* - m) \equiv x_m + (x_{m+1} - x_m)\beta$$

Indici di variabilità

Può accadere che due serie di dati abbiano la stessa media e/o mediana nonostante le due serie siano molto diverse fra loro.

$$E_1 = \{0.5, 0.8, 2.0, 2.7, 4.0\} \quad \bar{x}_1 = \hat{x}_1 = 2$$

Le due serie a lato hanno la stessa media e mediana, nonostante i

dati della serie E_2 risultino più omogenei, ovvero più vicini fra loro.

$$E_2 = \{1.4, 1.7, 2.0, 2.1, 2.8\} \quad \bar{x}_2 = \hat{x}_2 = 2$$

Si introducono quindi degli indici che misurano il grado di *variabilità o dispersione*.

Varianza (s^2)

La *varianza* è definita come:

$$varianza \equiv s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Più è grande la varianza più i dati sono *distanti* dalla media. Una varianza piccola denota una *omogeneità* dei dati della distribuzione.

La varianza può anche essere espressa come *la differenza fra la media dei quadrati dei dati e il quadrato della media*:

$$s^2 = \overline{x^2} - \bar{x}^2$$

Dimostrazione:

$$\begin{aligned} s^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \rightarrow \frac{1}{n} \sum_{i=1}^n x_i^2 + \bar{x}^2 - 2x_i\bar{x} \rightarrow \frac{1}{n} \left[\sum_{i=1}^n x_i^2 + \sum_{i=1}^n \bar{x}^2 - 2 \sum_{i=1}^n x_i\bar{x} \right] \rightarrow \frac{1}{n} \left[\sum_{i=1}^n x_i^2 + n\bar{x}^2 - 2\bar{x} \sum_{i=1}^n x_i \right] \\ &\rightarrow \overline{x^2} + \bar{x}^2 - 2\bar{x}^2 = \overline{x^2} - \bar{x}^2 \end{aligned}$$

Scarto quadratico medio (Deviazione standard)

Lo scarto quadratico medio è definito come la *radice della varianza*:

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

Scarto medio assoluto

Lo scarto medio assoluto è definito come il rapporto fra la numerosità della distribuzione e la somma per ogni evento del valore assoluto della differenza fra il valore i -esimo e la media:

$$s.a. = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|$$

Coefficiente di variazione

Nel caso in cui le due serie di dati in esame abbiano la stessa varianza è necessario utilizzare il *coefficiente di variazione* (*c. v.*):

$$c. v. = \frac{s}{\bar{x}}$$

A valori maggiori del coefficiente di variazione corrisponde una maggiore variabilità dei dati.

Dati raggruppati per classi

Istogramma normalizzato

Per rappresentare tramite grafico un insieme di classi (ognuna avente una propria frequenza assoluta, frequenza relativa o percentuale e ampiezza) si può costruire l'*istogramma normalizzato*.

L'*istogramma normalizzato* consiste in un insieme di rettangoli adiacenti (ognuno relativo ad una classe) aventi la base sull'asse x con punto medio nel valore centrale della classe e l'altezza proporzionale alla frequenza della classe, in modo che l'area del rettangolo sia pari alla frequenza relativa della classe. Quindi l'altezza del rettangolo si ottiene dividendo la frequenza relativa per l'ampiezza della classe. In questo modo se si sommano tutte le aree dei rettangoli si otterrà 100.

Media, Varianza, Mediana e Quartili per le classi

Supponendo che gli n dati siano raggruppati in classi C_i ($i = 1, 2, \dots, k$) e definendo m_i come il *valore medio della classe* C_i , si ha che *la media e la varianza sono così definite*:

$$\text{media} \equiv \bar{x} = \frac{1}{n} \sum_{i=1}^k m_i f_i \quad \text{e} \quad \text{varianza} \equiv s^2 = \frac{1}{n} \sum_{i=1}^k (m_i - \bar{x})^2 f_i$$

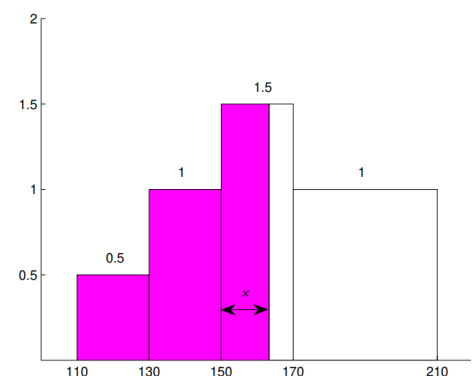
in cui $f_i \equiv$ frequenza assoluta della classe.

Il calcolo dei quartili e della mediana, invece, può essere svolto attraverso l'utilizzo dell'*istogramma normalizzato*. Per calcolare il secondo quartile, ovvero la mediana, occorre trovare quel valore dell'asse x tale che divida esattamente a metà l'area definita dall'istogramma: in genere si sommano le aree dei rettangoli fino ad ottenere una combinazione consecutiva di aree la cui somma è pari a 50.

Nell'esempio a lato, si ha che:

- Il primo rettangolo ha area pari a: $\text{base} * h = 20 * 0.5 = 10$;
- Il secondo rettangolo ha area pari a: $\text{base} * h = 20 * 1 = 20$;
- L'area interessata del terzo rettangolo pari a: $x * h = 1.5x$
- Date che la mediana divide a metà l'area poniamo:
$$10 + 20 + 1.5x = 50$$
$$x = 13.3$$

Otteniamo che $\text{mediana} \equiv Q_2 = 150 + 13.3 = 163.3$.



Indici di forma

Asimmetria

L'*asimmetria* è una misura dello scostamento di una distribuzione dalla simmetria. Se la curva di frequenza di una distribuzione ha una coda più lunga a destra del massimo centrale, piuttosto che a sinistra, la distribuzione si dice *positivamente asimmetrica*, si dirà *negativamente asimmetrica* se il contrario.

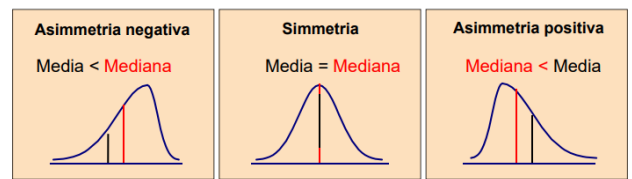
Indice di asimmetria (*sk*)

Date n osservazioni $\{x_1, x_2, \dots, x_n\}$ si definisce *indice di asimmetria* (*sk*) la quantità:

$$\text{indice di asimmetria} \equiv sk = \frac{m_3}{\sqrt{m_2^3}}, m_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k \quad \text{con } k = 2, 3, 4, \dots$$

In cui:

- Gli m_k sono i momenti centrati di ordine k ;
- Questo indice indica se la distribuzione del campione è simmetrica rispetto alla media ($sk = 0$):
 - Se $sk > 0$, la distribuzione sarà più concentrata a sinistra, con una coda più lunga a destra;
 - Se $sk < 0$, la distribuzione sarà più concentrata a destra, con una coda più lunga a sinistra;



Curtosi

La *curtosi* è una misura dell'appiattimento di una distribuzione rispetto alla distribuzione normale (gaussiana).

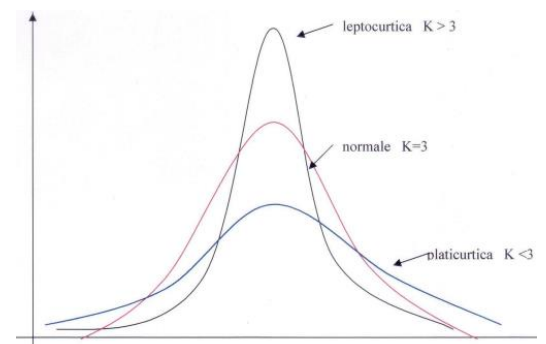
Indice di curtosi (K)

Date n osservazioni $\{x_1, x_2, \dots, x_n\}$ si definisce *indice di curtosi (K)* la quantità:

$$\text{indice di curtosi} \equiv K = \frac{m_4}{m_2^2}, m_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k \text{ con } k = 2, 3, 4, \dots$$

In cui:

- Gli m_k sono i momenti centrati di ordine k ;
- A seconda dell'indice K si ha che:
 - Se $K = 3$, la distribuzione ha l'altezza di una normale;
 - Se $K > 3$, la distribuzione, detta *leptocurtica*, è più appuntita di una normale e avrà code più lunghe;
 - Se $K < 3$, la distribuzione, detta *platicurtica*, sarà più appiattita della normale e avrà code più corte;

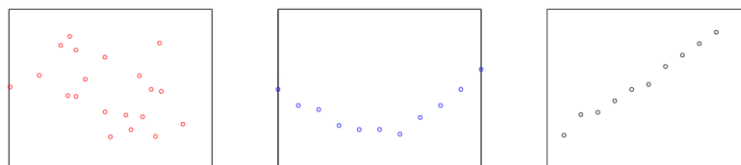


La correlazione tra due serie di Dati

Talvolta più caratteri vengono misurati per ogni evento osservato e ci possiamo chiedere se esiste una certa relazione tra questi dati, ovvero se sono tra loro indipendenti. Consideriamo di avere due caratteri quantitativi X e Y di un evento e supponiamo che i dati siano sotto forma di coppie $(\{x_i\}, \{y_i\})$ di numerosità n , in cui la prima coordinata rappresenta il primo carattere X e la seconda quello Y . Ogni coppia di dati è relativa allo stesso record dell'evento.

Scatterplot

In un primo approccio grafico si possono disegnare sul piano tutti punti (x_i, y_i) e vedere se essi tendono a disporsi secondo un andamento regolare (*scatterplot*).



Analisi di regressione per una serie di dati

Assegnato un insieme E di coppie di dati $\{x_i\}, \{y_i\}$ di numerosità n , ci domandiamo se esiste un legame funzionale del tipo $y = f(x)$ che descriva bene la relazione tra i dati.

Un'analisi di questo tipo si chiama *analisi di regressione*, introdurremo 4 tecniche.

Metodo dei minimi quadrati

Si cerca $f(x)$ tale che sia minima la funzione *residuo*:

$$g(f) = \sum_{i=1}^n [f(x_i) - y_i]^2$$

Questa funzione rappresenta la somma dei quadrati delle distanze fra i dati sperimentali y_i e quelli calcolati con la funzione $f(x_i)$, ovvero la somma degli "errori".

Regressione lineare (e coefficiente di correlazione lineare)

In questo caso la funzione f è una retta:

$$f(x) = mx + q$$

$$g(m, q) = \sum_{i=1}^n [mx_i + q - y_i]^2$$

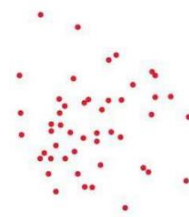
In questa tecnica le incognite sono: m e q .

Il metodo applicato fornisce la retta che meglio approssima i dati, ma non il grado di approssimazione. Per questo motivo si introduce il coefficiente di *correlazione lineare* (o di Pearson):

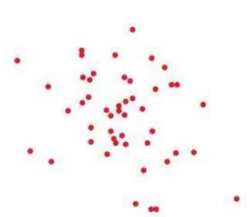
$$r_{xy} = \frac{c_{xy}}{s_x s_y}$$

Tale che:

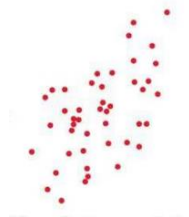
- $r_{xy} \in [-1, 1]$;
- Se $r_{xy} = \pm 1$, i dati (x_i, y_i) sono perfettamente allineati con la retta di regressione;
- Se $r_{xy} > 0$ la retta è ascendente, altrimenti è discendente;
- Se $|r_{xy}| < 0.9$, i dati si allontanano dall'andamento rettilineo;



Correlation $r = 0$



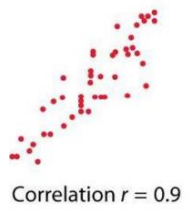
Correlation $r = -0.3$



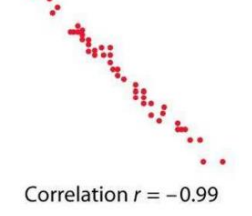
Correlation $r = 0.5$



Correlation $r = -0.7$



Correlation $r = 0.9$



Correlation $r = -0.99$

Breve Introduzione al Calcolo Combinatorio

Fattoriale

Il *fattoriale* di un numero $n \in \mathbb{N}$ si indica con $n!$ ed è uguale al prodotto di tutti i numeri interi positivi minori o uguali a n .

$$n! = \begin{cases} 0, & \text{se } n = 0 \\ \prod_{k=1}^n k, & \text{se } n \in \mathbb{N}, n \neq 0 \end{cases}$$

Coefficiente Binomiale

Il *coefficiente binomiale* è un numero naturale definito a partire da una coppia di numeri naturali, solitamente indicati con n e k , e rappresenta il numero di sottoinsiemi di k elementi distinti che si possono estrarre da un insieme di n elementi distinti.

$$c.b \equiv \binom{n}{k} = \begin{cases} \frac{n!}{k!(n-k)!}, & \text{se } n, k \in \mathbb{N}, \quad 0 \leq k \leq n \\ 0, & \text{se } n, k \in \mathbb{N}, \quad 0 \leq n < k \end{cases}$$

Permutazioni

Dati n elementi di qualsiasi tipo, si dicono *permutazioni* di tali elementi tutti i possibili raggruppamenti formati in modo che:

- Ogni raggruppamento contenga tutti gli n elementi;

- Ogni raggruppamento differisca dagli altri per l'ordine con cui gli elementi si susseguono;

$$A = \{a, b, c\} \rightarrow \{abc, acb, bac, bca, cab, cba\}$$

Si distinguono tre tipi di permutazioni: le *permutazioni semplici*, le *permutazioni con ripetizione* e le *permutazioni circolari*.

Permutazioni semplici (o senza ripetizione)

Si definiscono *permutazioni semplici* i possibili modi per ordinare totalmente n elementi distinti. Il numero di permutazioni semplici si indica con P_n :

$$P_n = n!$$

Permutazioni con ripetizione

Si definiscono *permutazioni con ripetizione* i possibili modi per ordinare totalmente n elementi di cui alcuni sono ripetuti due o più volte.

Dati n elementi, di cui n_1 sono di un tipo, n_2 di un altro tipo, e così via fino a n_k , tali che $n_1 + n_2 + \dots + n_k = n$, il numero di permutazioni con ripetizione si indica con $P_n^{n_1, n_2, \dots, n_k}$:

$$P_n^{n_1, n_2, \dots, n_k} = \frac{n!}{n_1! n_2! \dots n_k!}$$

Permutazioni circolari

Si definiscono *permutazioni circolari* i possibili modi per ordinare totalmente n elementi distinti, da disporre in modo circolare e in cui non si può distinguere la prima dall'ultima posizione. Il numero di permutazioni semplici si indica con P_n^c :

$$P_n^c = (n - 1)!$$

Disposizioni

Dati n elementi distinti e $k \in \mathbb{N}$, si dicono *disposizioni di classe k* i raggruppamenti formati a partire dagli n elementi in modo che:

- Ogni raggruppamento contenga k elementi;
- Ogni raggruppamento differisca dagli altri per almeno un elemento oppure per l'ordine con cui gli elementi si susseguono;

Nelle disposizioni potrebbero essere ammesse o non ammesse ripetizioni degli elementi:

$$\text{senza ripetizioni} \equiv A_{k=2} = \{a, b, c\} \rightarrow \{ab, ac, ba, bc, ca, cb\}$$

$$\text{con ripetizioni} \equiv A_{k=2}^R = \{a, b, c\} \rightarrow \{aa, ab, ac, bb, ba, bc, cc, ca, cb\}$$

Disposizioni semplici (o senza ripetizione)

Si definiscono *disposizioni semplici* le sequenze ordinate di k elementi distinti estratti fra gli n elementi distinti, con $n \geq k$. Il numero di disposizioni semplici si indica con $D_{n,k}$:

$$D_{n,k} = \frac{n!}{(n - k)!}$$

Disposizioni con ripetizione

Si definiscono *disposizioni con ripetizione* le sequenze ordinate di k elementi estratti fra gli n elementi distinti, con la possibilità che ogni elemento possa ripetersi fino a k volte nella stessa sequenza.

Il numero di disposizioni con ripetizione si indica con $D'_{n,k}$:

$$D'_{n,k} = n^k$$

Combinazioni

Dati n elementi distinti e $k \in \mathbb{N}$, si dicono *combinazioni di classe k* i raggruppamenti formati a partire dagli n elementi in modo che:

- Ogni raggruppamento contenga k elementi;
- Ogni raggruppamento differisca dagli altri per almeno un elemento;

Nelle combinazioni potrebbero essere ammesse o non ammesse ripetizioni degli elementi:

$$\text{senza ripetizioni} \equiv A_{k=2} = \{a, b, c\} \rightarrow \{ab, ac, bc\}$$

$$\text{con ripetizioni} \equiv A_{k=2}^R = \{a, b, c\} \rightarrow \{aa, ab, ac, bb, bc, cc\}$$

Combinazioni semplici (o senza ripetizione)

Si definiscono *combinazioni semplici* le sequenze di k elementi distinti estratti fra gli n elementi distinti indipendentemente dall'ordine, con $n \geq k$. Il numero di combinazioni semplici si indica con $C_{n,k}$:

$$C_{n,k} = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Combinazioni con ripetizione

Si definiscono *combinazioni con ripetizione* le sequenze di k elementi estratti fra gli n elementi distinti indipendentemente dall'ordine, con la possibilità che ogni elemento possa ripetersi fino a k volte nella stessa sequenza.

Il numero di combinazioni con ripetizione si indica con $C'_{n,k}$:

$$C'_{n,k} = \binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-1)!}$$

Elementi di Probabilità

I fenomeni casuali o aleatori, ovvero i fenomeni di cui non si può prevedere l'esito, sono oggetto di studio del calcolo delle probabilità, ovvero fare previsione su quello che accadrà.

Spazio Campione ed Eventi (e relazione con le operazioni di insiemistica)

Si definisce *spazio campione* Ω l'insieme di tutti i possibili **esiti** di un esperimento. Da ciò è possibile definire che:

- Un evento A è un sottoinsieme di Ω che può essere *elementare o composto*;
- Due eventi A e B , si dicono *incompatibili* se sono due *sottoinsiemi di Ω disgiunti* (ovvero non hanno parti in comune);

Questa identificazione tra eventi e sottoinsiemi di Ω ci permette di combinare gli eventi per formarne degli altri. In particolare è possibile estendere le nozioni di insiemistica al calcolo delle probabilità:

- $A = \Omega$, è definito come un *evento certo*, poiché *si verifica sempre*;
- $A = \emptyset$, è definito come un *evento impossibile*, poiché *non si verificherà mai*;
- $A \cap B$, è detto *prodotto logico* e corrisponde all'evento: *gli eventi A e B si verificano entrambi*;
 - Se $A \cap B = \emptyset$ si ha che gli eventi sono incompatibili, ovvero mutualmente esclusivi;
- $A \cup B$, è detto *somma logica* e corrisponde all'evento: *almeno uno dei due eventi si verifica*;
 - Segue che $A \cup B \neq \emptyset$;
- \bar{A} , corrisponde all'evento: *l'evento A non si verifica*;
- $A \setminus B$, corrisponde all'evento: *si verifica A ma non B* ;

Famiglia di eventi e sigma algebra ($\sigma_{algebra}$)

Una buona famiglia di eventi deve essere tale da garantire che tutti gli insiemi ottenuti componendo eventi (tramite le operazioni di insiemistica) sia ancora un evento.

Una famiglia \mathcal{F} di sottoinsiemi di Ω si chiama $\sigma_{algebra}$, se soddisfa le proprietà:

- $\Omega \in \mathcal{F}$;
- dato $A \in \mathcal{F}$ allora $\bar{A} \in \mathcal{F}$;
- dati $A_1, A_2, \dots \in \mathcal{F}$ allora $A_1 \cap A_2 \cap \dots \in \mathcal{F}$;
- dati $A_1, A_2, \dots \in \mathcal{F}$ allora $A_1 \cup A_2 \cup \dots \in \mathcal{F}$;

Possiamo notare che:

- Qualunque $\mathcal{F} \subseteq \mathcal{P}(\Omega) \equiv$ insieme delle parti di Ω ;
- Dato Ω , la più piccola $\sigma_{algebra}$ è $\{\emptyset, \Omega\}$ mentre la più grande è $\mathcal{P}(\Omega)$;
- Dati Ω e A_1, A_2, \dots, A_n si descrive la $\sigma_{algebra}$ generata dagli A_i la più piccola $\sigma_{algebra}$ in cui tutti gli A_i fanno parte;

Probabilità secondo Kolmogorov

Sia Ω uno spazio campionario e \mathcal{F} una famiglia di eventi che sia una σ -algebra, una misura di probabilità su Ω è una funzione:

$$P(A) : \mathcal{F} \rightarrow [0, 1]$$

Tale che:

- $P(\Omega) = 1$;
- Per ogni evento $A \in \mathcal{F} \rightarrow P(A) \in R$ con $0 \leq P(A) \leq 1$;
 - Per ogni evento A si ha che $A \cup \bar{A} = \Omega$ e $A \cap \bar{A} = \emptyset$. Quindi:
$$1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A}) \rightarrow P(\bar{A}) = 1 - P(A)$$
- Data una successione di eventi incompatibili $\{A_i, i = 1, \dots, N\} \rightarrow P(\cup_i A_i) = \sum_i P(A_i)$

Spazi di probabilità

Definiamo *spazio di probabilità la terna* (Ω, \mathcal{F}, P) . Distinguiamo spazi finiti e infiniti.

Spazi di probabilità finiti e Probabilità Classica

Sia Ω uno *spazio campione finito*, cioè $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$, la σ -algebra \mathcal{F} è formata da tutti i possibili sottoinsiemi di Ω , ovvero *l'insieme delle parti*.

Per definire la probabilità è sufficiente assegnare una probabilità agli eventi elementari, ovvero:

$$P(\omega_i) = p_i \quad \forall i = 1, 2, \dots, n; \text{ tale che } 0 \leq p_i \leq 1 \text{ e } \sum_{i=1}^n p_i = 1$$

In questo modo si ha ben definita la $P(A)$ per ogni evento, che si ottiene sommando le probabilità degli eventi elementari che compongono A .

Nel caso in cui tutti gli eventi hanno la stessa probabilità, ovvero si ha uno spazio equiprobabile, si ha che:

$$P(\omega_i) = \frac{1}{n} \quad \forall i = 1, 2, \dots, n; \text{ tale che } P(A) = \frac{r}{n} \rightarrow r \equiv \text{cardinalità di } A$$

In questo specifico caso denota la *Probabilità Classica*, secondo cui: $P = \frac{n.casi \text{ favorevoli}}{n.casi \text{ possibili}}$ o $P = \frac{\text{frequenza successi}}{\text{numero di prove}}$

Spazi di probabilità infiniti

Gli spazi di probabilità infiniti si possono suddividere in *numerabili e continui*.

Spazi di probabilità infiniti numerabili

Gli spazi di probabilità infiniti numerabili sono quegli insiemi in cui è possibile definire una corrispondenza biunivoca fra gli elementi dello spazio e numeri naturali:

$$\Omega = \{\omega_1, \omega_2, \dots\} \rightarrow \text{si enumerano gli elementi}$$

In questo caso, possiamo trattare lo spazio di probabilità infinito come una *generalizzazione degli spazi finiti*:

- La $\sigma_{algebra}$ è formata dall'insieme delle parti di Ω ;
- Per definire la probabilità è sufficiente assegnare una probabilità agli eventi elementari, ovvero:

$$P(\omega_i) = p_i \quad \forall i = 1, 2, \dots, n; \text{ tale che } 0 \leq p_i \leq 1 \text{ e } \sum_{i=1}^{\infty} p_i = 1$$

Spazi di probabilità infiniti continui

Uno spazio di probabilità infinito continuo è definito come: $\Omega = (a, b) \subseteq \mathbb{R}$.

In questo caso bisogna assegnare una probabilità definita su un intervallo tramite una funzione di *distribuzione* $f(x) \geq 0$, con

$$\int_a^b f(x)dx = 1 \rightarrow \forall I(x_1, x_2) \subseteq (a, b) \text{ si ha che } P(x_1, x_2) = \int_{x_1}^{x_2} f(x)dx$$

La $\sigma_{algebra}^F$ è costituita dagli intervalli in \mathbb{R} : contiene infatti tutti gli intervalli in \mathbb{R} , tutti i complementari, tutte le intersezioni/unioni numerabili.

Probabilità della somma logica di eventi ($P(E \cup F)$)

L'evento somma logica $E \cup F$ consiste nel fatto che almeno uno dei due eventi si verifica.

Possiamo definire la probabilità della somma logica come:

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

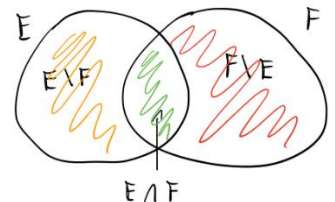
[Dimostrazione] Dati due eventi E e F compatibili ($E \cap F \neq \emptyset$)

Specifichiamo che: $P(E) = P(E \setminus F) + P(E \cap F) \rightarrow P(E \setminus F) = P(E) - P(E \cap F)$

$$P(E \cup F) = P(E \setminus F) + P(E \cap F) + P(F \setminus E) = P(E) - P(E \cap F) + P(E \cap F) + P(F) - P(E \cap F)$$

Si denota il *caso particolare* in cui i due eventi siano *incompatibili*, ovvero $P(E \cap F) = \emptyset = 0$:

$$P(E \cup F) = P(E) + P(F) - 0 = P(E \cup F) = P(E) + P(F)$$



Probabilità condizionata ($P(A|B)$) e rapporto di indipendenza fra eventi

Supponiamo di avere un evento detto B che si verifica, questa informazione dovrebbe alterare le probabilità che assegniamo agli altri eventi. Supponiamo di avere un altro evento detto A , si ha che A si verifica se e solo se si verificano sia A che B , ovvero la sua probabilità dovrebbe essere proporzionale a $P(A \cap B)$.

Siano Ω uno spazio campione, P una sua misura di probabilità, A e B due eventi con $P(A), P(B) > 0$, definiamo la *probabilità di A condizionata dall'evento B* la quantità:

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

Tale che:

- $P(A \cap B)$ è la probabilità che A e B si verifichino contemporaneamente;
- La *probabilità condizionata* ha senso solo se l'evento B che si può verificare, non è un evento impossibile;
- La *probabilità condizionata* di due insiemi disgiunti, ovvero due eventi incompatibili, è pari a 0:
 $P(A \cap B) = \emptyset = 0$;
- Dati due eventi A e B compatibili, si ha che:
 - Se $P(A|B) > P(A)$, diremo che l'evento B è favorevole all'evento A ;
 - Se $P(A|B) < P(A)$, diremo che l'evento B è sfavorevole all'evento A ;
 - Se $P(A|B) = P(A)$, diremo che l'evento B è indipendente dall'evento A ;

- Se A è indipendente da B allora B è indipendente da A

$$P(A|B) = P(A) \leftrightarrow P(B|A) = P(B)$$

- Se $A \subset B$ si ha che $P(A \cap B) = P(A) \rightarrow P(A|B) = \frac{P(A)}{P(B)}$ (che nel caso $B \equiv \Omega$ costituisce la nozione di *probabilità classica*;

Nel caso in cui $A, B \subset \Omega$ (*finito*) possiamo definire la *probabilità condizionata* come:

$$P(A|B) = \frac{\text{elementi di } (A \cap B)}{\text{elementi di } (B)} \rightarrow \text{numero di elementi} \equiv \text{cardinalità dell'insieme}$$

Probabilità del prodotto logico di eventi ($P(A \cap B)$)

L'evento prodotto logico $A \cap B$ consiste nel fatto che i due eventi A e B si verificano entrambi.

Possiamo definire la probabilità del prodotto logico attraverso il teorema della *probabilità composta*:

$$P(A \cap B) = P(B)P(A|B) = P(A)P(B|A)$$

La *probabilità composta* è una conseguenza della *probabilità condizionata*: la probabilità che due eventi si verifichino contemporaneamente è pari alla probabilità di uno dei due eventi moltiplicata per la probabilità dell'altro evento condizionato dal verificarsi del primo.

Teorema della Catena

Dati n eventi E_1, \dots, E_n la cui intersezione ha probabilità positiva ($E_1 \cap \dots \cap E_n > 0$) si ha che:

$$P(E_1 \cap \dots \cap E_n) = P(E_1) * P(E_2|E_1) * P(E_n|E_1 \cap \dots \cap E_{n-1})$$

Eventi stocasticamente indipendenti

Due eventi A e B si definiscono *stocasticamente indipendenti* se:

$$P(A|B) = P(A) \text{ e } P(B|A) = P(B) \leftrightarrow P(A \cap B) = P(A)P(B)$$

Ovvero il presentarsi dell'evento B non influenza la probabilità dell'evento A, e viceversa.

Partizione dello spazio di probabilità Ω

Si definisce partizione di Ω un insieme di eventi A_i (finito e numerabile) tali che:

- $P(A_i) \neq 0 \forall A_i$;
- $P(A_i \cap A_j) = 0 \forall i \neq j$;
- $\bigcup_{i=1}^N A_i = \Omega$;

Probabilità totale (Legge delle alternative)

Sia $\{A_i\}$ una partizione di Ω e B un qualsiasi evento di Ω , si ha che:

$$P(B) = \sum_{i=1} P(A_i \cap B) = \sum_{i=1} P(A_i)P(B|A_i)$$

[Dimostrazione]:

$$F = F \cap \Omega = F \cap \left(\bigcup_i A_i \right) = \bigcup_i (F \cap A_i)$$

$$P\left(\bigcup_i (F \cap A_i)\right) = \sum_{i=1} P(A_i \cap B) = \sum_{i=1} P(A_i)P(B|A_i)$$

Sia $\{A_i\}$ una partizione di Ω , E e F due qualsiasi eventi di Ω , si ha che:

$$P(F|E) = \sum_{i=1} P(F|A_i \cap E)P(A_i|E)$$

Teorema di Bayes

Sia $\{A_i\}$ una partizione di Ω e B un qualsiasi evento di Ω , si ha che:

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^N P(A_i)P(B|A_i)}$$

Dimostrazione:

- Dalle definizioni di probabilità condizionata e probabilità composta si ottiene:

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)P(B|A_i)}{P(B)}$$

- Da teorema della *probabilità totale* si ottiene:

$$\frac{P(A_i)P(B|A_i)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^N P(A_i)P(B|A_i)}$$

Variabili aleatorie

Nei problemi di calcolo delle probabilità si considerano delle quantità che sono funzioni del risultato di un fenomeno casuale dette *variabili aleatorie*, in quanto non conosciamo esattamente il loro valore. Però se riesco ad associare una probabilità ai valori che essa assume, potrò avere delle informazioni sul suo comportamento.

Assegnato uno spazio di probabilità (Ω, \mathcal{F}, P) , si chiama *variabile aleatoria* (*v. a*) un'applicazione che ad ogni elemento ω di uno spazio campione Ω associa un numero reale:

$$X(\omega): \omega \in \Omega \rightarrow S \subseteq \mathbb{R} \rightarrow \forall a \in \mathbb{R} \text{ l'insieme } \{\omega; X(\omega) \leq a\} \text{ sia in } \mathcal{F}.$$

Quest'ultima affermazione equivale a dire che ha senso calcolare la probabilità che X assuma valori più piccoli di un assegnato a .

L'insieme S dei valori assunti da X si chiama *supporto* della variabile aleatoria.

Variabile aleatoria discreta

Una variabile aleatoria si dice *discreta* se l'insieme dei valori s da essa assumibili, ovvero il suo *supporto*, è un insieme finito o numerabile:

$$S = \{s_1, s_2, s_3, \dots, s_n\}$$

Distribuzione discreta di probabilità (e densità delle variabili aleatorie discrete)

Sia S il *supporto* della variabile aleatoria discreta X . Allora definisco *distribuzione discreta di probabilità* la funzione:

$$P(s) = \begin{cases} P(X = s) & \forall s \in S \\ 0 & \text{altrimenti} \end{cases} \leftrightarrow \begin{cases} P(s) \geq 0 \quad \forall s \\ \sum_{s \in S} P(s) = 1 \end{cases} \equiv \text{densità v. a discreta}$$

Funzione di ripartizione

Sia X una *variabile aleatoria*, chiamiamo *funzione di ripartizione* la funzione $F_X: \mathbb{R} \rightarrow [0,1]$ così definita:

$$F_X(a) = P(X \leq a) = P(E_a), \forall a \in \mathbb{R}$$

La funzione $F_X(a)$ rappresenta la probabilità che la variabile aleatorie X assuma un valore inferiore o uguale ad a .

Nel caso in cui la *variabile aleatoria fosse discreta*, si ha che:

$$F_X(a) = \sum_{\substack{s \in S \\ s \leq a}} P(s), \forall a \in \mathbb{R}$$

La funzione di ripartizione ha le seguenti proprietà:

- F_X è una *funzione monotona non decrescente*, ovvero $F_X(a_1) \leq F_X(a_2) \leftrightarrow a_1 \leq a_2$;
- Se $E_{a_1} \subseteq E_{a_2}$ allora $P(E_{a_1}) \leq P(E_{a_2})$;
- $\lim_{a \rightarrow -\infty} F_X(a) = 0$ e $\lim_{a \rightarrow \infty} F_X(a) = 1$;
- $\forall x_0 \in \mathbb{R}$ si ha che $P(X = x_0) = \lim_{a \rightarrow x_0^+} F_X(a) - \lim_{a \rightarrow x_0^-} F_X(a)$;

- F_X è continua a destra: $F_X(x_0) = \lim_{a \rightarrow x_0^+} F_X(a)$;
- Se la variabile aleatoria X è discreta, la funzione di ripartizione sarà costante a tratti;

Variabili aleatorie continue, densità e probabilità

Se una variabile aleatoria X assume i valori di un intervallo $I \subset \mathbb{R}$, la corrispondente funzione di ripartizione $F_X(a)$ è continua e di conseguenza la variabile aleatoria sarà detta *continua*.

Per definire la probabilità per una variabile aleatoria X continua si utilizza una *funzione di densità* f_x , tale che:

$$f_x(u): \mathbb{R} \rightarrow \mathbb{R}_+ \text{ tale che } F_X(a) = P(X \leq a) = \int_{-\infty}^a f_x(u) du, \forall a \in \mathbb{R}$$

$$P(a \leq X \leq b) = F_X(b) - F_X(a) = \int_a^b f_x(u) du$$

Se esiste una funzione di densità f_x tale che:

$$P(a < X \leq b) = \int_a^b f_x(u) du, \forall a, b \in \mathbb{R}$$

La funzione $f_x(u)$ è detta *funzione di densità di probabilità*.

Si consideri:

- $f_x(u) = 0$ se $u \notin I$;
- $P(a < x \leq b) = P(a < x < b)$;
- $\int_{-\infty}^{\infty} f_x(u) du = 1$;

Variabile Aleatoria in funzione di un'altra variabile aleatoria

Quando si ha una variabile aleatoria Y in funzione di un'altra variabile aleatoria X bisogna distinguere il caso in cui le variabili aleatorie siano *discrete* o *continue*:

- Nel caso in cui le v.a siano discrete, si applica la funzione sugli elementi del supporto di X per ottenere il supporto di Y . La probabilità di Y è data dalla somma delle probabilità dei valori di X che hanno come valore Y ;
- Nel caso continuo, in cui $Y = g(X)$, se g è invertibile nel supporto di X e h la sua inversa si ha che:

$$f_y = f_x(h(y)) |h'(y)|$$

Indici di tendenza centrale e variabilità per le variabili aleatorie

Anche per le variabili aleatorie si definiscono indici di tendenza centrale, la cui formula varia a seconda della tipologia di variabile aleatoria.

	Variabile Aleatoria Discreta	Variabile Aleatoria Continua
Valore atteso (valore medio pesato su s)	$\mu = E[X] = \sum_i x_i * P(X = x_i)$	$\mu = E[X] = \int_{-\infty}^{\infty} x * f_x(x) dx$
Moda	Sia X una variabile aleatorie discreta o continua, la <i>moda</i> è quel valore, se esiste, per cui la <i>densità</i> è massima.	
Quantile	Sia X una variabile aleatorie, dato $a \in [0,1]$ si definisce <i>quantile di ordine a</i> il più piccolo numero x_a tale che: $P(X < x_a) \leq a \leq P(X \leq x_a)$	
Mediana	Sia X una variabile aleatorie, si definisce mediana il <i>quantile di ordine a</i> = 0.5	
Varianza	$Var[X] = E[(x - \mu)^2]$	

Proprietà degli indici di tendenza centrale

- $\mu = E[aX + b] = aE[X] + b \forall a, b \in \mathbb{R}$;
- $Var[aX + b] = a^2 Var[X] \forall a, b \in \mathbb{R}$;

- $Var[X] = E[X^2] - (E[X])^2 \quad \forall X;$

Quantili per le variabili aleatorie

Si definisce *quantile di ordine* $a \in [0,1]$, di una variabile aleatoria X , quel valore q_a per cui in quel punto sull'asse delle x , l'area della funzione densità di probabilità compresa fra $]-\infty, x]$ è pari a a :

$$\mathbb{P} = (X \leq q_a) = a$$

$$\text{Se } X \text{ è continua} \rightarrow \mathbb{P} = (X \leq q_a) = \int_{-\infty}^{q_a} f_x(u) du = a$$

Disuguaglianze di Markov e di Cebicev

Sia X una variabile aleatoria e $g: \mathbb{R} \rightarrow \mathbb{R}$ tale che $g(x) \geq 0 \quad \forall x \in \mathbb{R}$, allora, se esiste $E[g(x)]$ si definisce la disuguaglianza di Markov:

$$P(g(X) \geq k) \leq \frac{E[g(X)]}{k} \quad \forall k > 0$$

La disuguaglianza di Cebicev è la seguente:

$$P(|X - E[X]| \geq \epsilon) \leq \frac{Var[X]}{\epsilon^2} \quad \forall \epsilon > 0$$

Momenti di una variabile aleatoria

Si dice *momento N-esimo* di una variabile aleatoria X , indicato con μ_n , la seguente quantità:

$$\mu_n = E[X^n]$$

	Variabile Aleatoria Discreta	Variabile Aleatoria Continua
Momento N-esimo	$\mu_n = E[X^n] = \sum_i x_i^n * P(X = x_i)$	$\mu_n = E[X^n] = \int_{-\infty}^{\infty} x^n * f_x(x) dx$

Osservazioni:

- $Var[X] = \mu_2 - \mu_1^2$

Funzione generatrice di momenti

Data una variabile aleatoria X , si definisce *funzione generatrice di momenti*:

$$m_X(t) = E[e^{tX}]$$

Si osservi che:

- $\mu_1 = \left[\frac{d}{dt} m_X(t) \right]_{t=0};$
- Siano X e Y due variabili aleatorie, se $m_X(t) = m_Y(t) \quad \forall t \in [-t_0, t_0]$ allora X e Y hanno la stessa densità;
- Se X ha una funzione generatrice di momenti e $Y = aX + b$ allora $m_Y = e^{tb} m_X(at);$

Funzione caratteristica di variabile aleatoria

Sia X una variabile aleatoria, si dice *funzione caratteristica di X* : $H_X(t) = E[e^{itX}]$ in cui $i \equiv$ unità immaginaria.

Vettori aleatori

Si definisce *vettore aleatore*, o *variabile aleatoria multi-dimensionale*, un vettore le cui componenti sono *variabili aleatorie*.

$$X \equiv \{X_1, \dots, X_n\}$$

Si osservi che:

$$P(X_1 \cap X_2 \cap \dots \cap X_n) = P(X_1, X_2, \dots, X_n)$$

Densità congiunte e marginali per variabili aleatorie discrete

Date due variabili aleatorie X e Y discrete si ha che:

- $\sum_i \sum_j P(X = i, Y = j) = 1$, i valori $P(X = i, Y = j)$ formano la *densità discreta congiunta del vettore* (X, Y) ;
- $\sum_i P(X = i) = 1$ e $\sum_j P(Y = j) = 1$, le densità $P(X = i)$ e $P(Y = j)$ sono dette *densità marginali*;
- È sempre possibile ricavare le *probabilità marginali* dalle *probabilità congiunte*, ma non è vero il viceversa;

Funzione di ripartizione congiunta

Dato un vettore aleatorio (X, Y) , è detta *funzione di ripartizione congiunta* la funzione:

$$F_{XY}(x, y) = P(X \leq x, Y \leq y) \equiv P((X \leq x) \cap (Y \leq y))$$

Funzione di ripartizione marginale

Dato una funzione di ripartizione congiunta F_{XY} , è definita *funzione di ripartizione marginale* la funzione:

$$F_X(x) = P(X \leq x) = P(X \leq x, Y < +\infty) = \lim_{y \rightarrow +\infty} F_{XY}(x, y)$$

Si osservi:

- $\forall y \in \mathbb{R}, \lim_{x \rightarrow -\infty} F_{XY}(x, y) = 0, \quad \lim_{x \rightarrow \infty} F_{XY}(x, y) = F_Y(y);$
- $\forall x \in \mathbb{R}, \lim_{y \rightarrow -\infty} F_{XY}(x, y) = 0, \quad \lim_{y \rightarrow \infty} F_{XY}(x, y) = F_X(x);$

Densità congiunte e marginali per vettori discreti

Si definisce *densità congiunta* del vettore aleatorio discreto (X, Y) , indicati con P_{XY} , l'insieme dei valori:

$$P_{XY} = (x_i, y_j) = P(X = x_i, Y = y_j)$$

Si definiscono *densità marginali* del vettore aleatorio discreto (X, Y) :

- $P_X(x_i) = \sum_j P_{XY}(x_i, y_j);$
- $P_Y(y_j) = \sum_i P_{XY}(x_i, y_j)$

Densità congiunte e marginali per vettori continui

Si definisce *densità congiunta* del vettore aleatorio continuo (X, Y) la funzione indicata con f_{XY} , per la quale per ogni $a, b, c, d \in \mathbb{R}$ si ha:

$$P(a < x < b, c < y < d) = \int_a^b dx \int_c^d dy f_{XY}(x, y)$$

Si può dimostrare che:

$$F_{XY}(s, t) = P(x < s, y < t) = \int_{-\infty}^s dx \int_{-\infty}^t dy f_{XY}(x, y)$$

Si definiscono *densità marginali* del vettore aleatorio discreto (X, Y) le funzioni:

- $f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dy;$
- $f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dx;$

Variabili aleatorie indipendenti

Date n variabili aleatorie X_i , con $i = 1, 2, \dots, n$, esse si diranno *indipendenti*, se per ogni scelta di intervalli A_i si ha che:

$$P(X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n) = P(X_1 \in A_1) * P(X_2 \in A_2) \dots P(X_n \in A_n)$$

Date due variabili aleatorie indipendenti, si definiscono:

- $P_{XY}(x_i, y_j) = P_X(x_i) * P_Y(y_j)$;
- $f_{XY}(x_i, y_j) = f_X(x_i) * f_Y(y_j)$;
- $F_{XY}(x, y) = F_X(x) * F_Y(y)$;

Densità della somma di variabili aleatorie

Dato vettore aleatorio X, Y continuo o discreto, la somma $U = X + Y$ delle sue componenti ha densità:

$$\begin{aligned} \text{caso discreto} &\rightarrow P_U(u) = \sum_k P_{XY}(k, u - k) \\ \text{caso continuo} &\rightarrow f_U(u) = \int_{-\infty}^{+\infty} f_{XY}(t, u - t) dt \end{aligned}$$

Valore Atteso per un vettore aleatorie

Sia (X, Y) un vettore aleatorie e $g(X, Y)$ una funzione, $E[g(X, Y)]$, se esiste è dato da:

$$\begin{aligned} \text{caso discreto} &\rightarrow E[g(X, Y)] = \sum_i \sum_j g(x_i, y_j) P_{XY}(x_i, y_j) \\ \text{caso continuo} &\rightarrow E[g(X, Y)] = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy g(x, y) f_{XY}(x, y) \end{aligned}$$

Si osservi:

- Per ogni coppia di variabili aleatorie X, Y , il valore atteso della somma è uguale alla somma dei valori attesi:
 $E[X + Y] = E[X] + E[Y]$

Studio dell'andamento di due variabili aleatorie (covarianza e coefficiente di correlazione)

Date due variabili aleatorie X e Y , si definisce *covarianza*, $Cov \in \mathbb{R}$, il valore:

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

Date X, Y, V, W variabili aleatorie e $a, b \in \mathbb{R}$, si osservi che:

- $Cov[X, a] = 0$;
- $Cov[X, Y] = Cov[Y, X]$;
- $Cov[X, X] = Var[X]$;
- $Cov[aX, bY] = abCov[X, Y]$;
- $Cov[X + Y, V + W] = Cov[X, V] + Cov[X, W] + Cov[Y, V] + Cov[Y, W]$;
- Se le due variabili aleatorie sono fortemente correlate, la covarianza sarà grande in valore assoluto;
- Se $Cov[X, Y] > 0$ le due variabili aleatorie si dicono *correlate positivamente* (all'aumentare di una aumenta anche l'altra);
- Se $Cov[X, Y] < 0$ le due variabili aleatorie si dicono *correlate negativamente* (al diminuire di una diminuisce anche l'altra);
- Se $Cov[X, Y] = 0$ le due variabili aleatorie si dicono *incorrelate*;
- Se X e Y sono indipendenti, allora sono anche incorrelate;

Date X e Y due variabili aleatorie, si definisce *coefficiente di correlazione* il valore:

$$\rho_{XY} = \frac{Cov[X, Y]}{\sqrt{Var[X]Var[Y]}} \text{ tale che } -1 \leq \rho_{XY} \leq 1$$

Tale che:

- Se $|\rho_{XY}| = 1$, le due variabili aleatorie sono correlate da $Y = aX + b$;

Distribuzioni Notevoli

Distribuzione di Bernulli ($X \sim \text{Ber}(p)$) e Schema di Bernulli

Consideriamo un esperimento che si può concludere solo in due modi, chiamati convenzionalmente “successo” e “insuccesso”.

Un esperimento di questo tipo è detto di *Bernulli*: sia p la probabilità che si verifichi il successo e q quella che si verifichi l'insuccesso si ha $q = 1 - p$.

La *variabile aleatorie* X che vale 1 se la *prova di Bernulli* si è conclusa con il successo e 0 altrimenti è detta *variabile aleatoria di Bernulli* di parametro p ed è denotata con $X \sim \text{Ber}(p)$.

Si denoti:

- La densità della distribuzione è $P(X = k) = p^k q^{1-k}$;
- La funzione generatrice di momenti è $m_X(t) = pe^t + q$;
- Il valore medio è $E[X] = \frac{d}{dt} [pe^t + q]_{t=0} = p$;
- La varianza è $\text{Var}[X] = pq$;

Schema di Bernulli

Una sequenza di prove di Bernulli, fra loro indipendenti, per le quali la probabilità del successo si mantiene invariato, è detto *schema di Bernulli*.

Distribuzione Binomiale ($X \sim \text{Bin}(n, p)$)

La *variabile aleatorie* X che conta il numero di successi in n prove di uno *schema di Bernulli* è detta *binomiale di parametro p* , dove il parametro p è la probabilità di successo, ed è denotata con $X \sim \text{Bin}(n, p)$.

Si denoti:

- La densità della distribuzione è $P(X = k) = \binom{n}{k} p^k q^{n-k}$;
- La funzione generatrice di momenti è $m_X(t) = (q + pe^t)^n$;
- Il valore medio è $E[X] = \frac{d}{dt} [(q + pe^t)^n]_{t=0} = np$;
- La varianza è $\text{Var}[X] = npq$;

Distribuzione Geometrica ($X \sim \text{Geo}(p)$)

La *variabile aleatorie* X che conta il numero prove di uno *schema di Bernulli* fino alla comparsa del primo successo è detta *geometrica di parametro p* , dove il parametro p è la probabilità del singolo successo, ed è denotata con $X \sim \text{Geo}(p)$

Si denoti:

- La densità della distribuzione è $P(X = k) = pq^{k-1}$;
- La funzione generatrice di momenti è $m_X(t) = \frac{pe^t}{1-qe^t}$ per $t < \ln(\frac{1}{q})$;
- Il valore medio è $E[X] = \frac{d}{dt} \left[\frac{pe^t}{1-qe^t} \right]_{t=0} = \frac{1}{p}$;
- La varianza è $\text{Var}[X] = \frac{q}{p^2}$;
- La distribuzione geometrica gode della proprietà di non memoria, ovvero $P(X = i + j | X > i) = P(X = j)$;

Distribuzione Pascal o Binomiale Negativa ($X \sim \overline{\text{Bin}}(n, p)$)

La *variabile aleatorie* X che conta il numero di prove di uno *schema di Bernulli* necessarie per ottenere l' n -esimo successo è detta *binomiale negativa di parametro p* , dove il parametro p è la probabilità di successo, ed è denotata con $X \sim \overline{\text{Bin}}(n, p)$.

Si denoti:

- La densità della distribuzione è $P(X = k) = \binom{-n}{k} p^n (-q)^k$;

- La funzione generatrice di momenti è $m_X(t) = \left(\frac{pe^t}{1-qe^t}\right)^n$ per $t < \ln\left(\frac{1}{q}\right)$;
- Il valore medio è $E[X] = \frac{d}{dt} \left[\left(\frac{pe^t}{1-qe^t}\right)^n \right]_{t=0} = \frac{n}{p}$;
- La varianza è $Var[X] = n\left(\frac{q}{p^2}\right)$;

Distribuzione Poisson ($X \sim Pois(\lambda)$), Schema e Processo di Poisson

La distribuzione di Poisson può essere vista come un caso limite della distribuzione binomiale, nello specifico quando:

- $n \rightarrow \infty$, il numero di prove tende all'infinito;
- $p \rightarrow 0$, la possibilità di successo è così bassa, tendente allo 0, tale che il valore medio $\lambda = np = \text{costante}$;

La variabile aleatorie X così definita è denominata *variabile di Poisson di parametro λ* ed è denotata con $X \sim Pois(\lambda)$.

Si denoti:

- La densità della distribuzione è $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$;
- La funzione generatrice di momenti è $m_X(t) = e^{\lambda(e^t - 1)}$;
- Il valore medio è $E[X] = \frac{d}{dt} [e^{\lambda(e^t - 1)}]_{t=0} = \lambda$;
- La varianza è $Var[X] = \lambda$;

Schema di Poisson

Per schema di Poisson si descrive la situazione in cui in un certo intervallo di tempo, o spazio, si manifestano degli eventi, chiamati convenzionalmente "*arrivi*" che soddisfano le seguenti condizioni:

- Non vi sono più *arrivi simultanei*;
- Il numero medio di arrivi nell'unità di tempo, o spazio, è costante;
- Ciascun arrivo avviene in modo indipendente dagli altri;

Processo di Poisson

La variabile aleatorie X che conta, in un determinato intervallo di tempo o spazio, il numero di arrivo di uno schema di Poisson di intensità λ è detta *processo di Poisson*.

Si denoti:

- $N_0 = 0$, si inizia il conteggio da 0;
- $N_{t+s} = N_t + N_s$, il numero di arrivi nell'intervallo N_{t+s} è uguale alla somma degli arrivi dei rispettivi intervalli;
- Dato $\Delta t \rightarrow 0$ si ha:
 - $P(N_{t+\Delta t} - N_t = 1) = \lambda \Delta t + o(\Delta t)$;
 - $P(N_{t+\Delta t} - N_t \geq 1) = o(\Delta t)$;
 - $P(N_{t+\Delta t} - N_t = 0) = 1 - \lambda \Delta t + o(\Delta t)$;
- Dato un processo di Poisson di intensità λ , allora $N_t \sim Pois(\lambda t)$

Distribuzione Esponenziale ($X \sim Exp(\lambda)$)

La variabile aleatorie X che descrive il "*tempo di attesa del primo arrivo di un processo di Poisson di intensità λ* " è detta *esponenziale di variabile λ* è denotata con $X \sim Exp(\lambda)$.

Si denoti:

- La densità della distribuzione è $P(x) = \lambda e^{-\lambda x}$ se $x > 0$; 0 se $x \leq 0$;
- La funzione generatrice di momenti è $m_X(t) = \frac{\lambda}{\lambda - t}$ per $t < \lambda$;
- Il valore medio è $E[X] = \frac{d}{dt} \left[\frac{\lambda}{\lambda - t} \right]_{t=0} = \frac{1}{\lambda}$;

- La varianza è $Var[X] = \frac{1}{\lambda^2}$;
- La distribuzione esponenziale gode della proprietà di memoria, ovvero $P(X > t + s | X > t) = P(X > t + s)$;

Distribuzione Gamma ($X \sim \Gamma(n, \lambda)$)

La variabile aleatoria X che descrive il "tempo di attesa del n -esimo arrivo di un processo di Poisson di intensità λ " è detta gamma di parametri n e λ è denotata con $X \sim \Gamma(n, \lambda)$.

Si denoti:

- La densità della distribuzione è $P(x) = \frac{1}{(n-1)!} \lambda^n x^{n-1} e^{-\lambda x}$ se $x > 0$; 0 se $x \leq 0$;

Distribuzione di Weibul ($X \sim \text{Wei}(n, \lambda)$) e Funzione Rischio

Sia $X \sim \text{Exp}(\lambda)$ e $\lambda > 0$, definiamo variabile di Weibul la variabile aleatoria $X = Y^{\frac{1}{a}}$ di parametri a e λ ed è denotata con $X \sim \text{Wei}(n, \lambda)$.

Si denoti:

- La densità della distribuzione è $P(x) = \lambda a x^{a-1} e^{-\lambda x^a}$ se $x > 0$; 0 se $x \leq 0$;
- Il valore medio è $E[X] = a^{-\frac{1}{a}} \Gamma(1 + \frac{1}{a})$;
- La varianza è $Var[X] = a^{-\frac{2}{a}} \left(\Gamma(1 + \frac{2}{a}) - \Gamma(1 + \frac{1}{a})^2 \right)$;

La funzione rischio è la variabile aleatoria T che descrive il "tempo di funzionamento", ovvero la probabilità, condizionatamente al fatto che fino all'istante f l'apparecchio fosse in funzione, esso si guasti inevitabilmente dopo, ossia all'istante $f + \Delta t$:

$$P(T \leq f + \Delta t | T > f)$$

Distribuzione di Gauss o Normale ($X \sim \text{Norm}(\mu, \sigma^2)$) e Normale Standardizzata

Una variabile aleatoria X è detta secondo una normale (o gaussiana) di parametri $\mu \in \mathbb{R}$ e $\sigma^2 \in \mathbb{R}^+$ ed è denotata con $X \sim \text{Norm}(\mu, \sigma^2)$ se ha come densità di probabilità:

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

Si denoti:

- La funzione f_x è simmetrica rispetto alla retta $x = \mu$;
- Ha un massimo assoluto in $(\mu, \frac{1}{\sigma\sqrt{2\pi}})$;
- Ha un asintoto orizzontale $y = 0$;
- Ha due flessi rispettivamente in $(\mu - \sigma, \frac{1}{\sigma\sqrt{2\pi}e})$ e $(\mu + \sigma, \frac{1}{\sigma\sqrt{2\pi}e})$;
- All'aumentare di σ il massimo *diminuisce*;
- Se $\mu = 0$ allora è simmetrica rispetto all'asse delle y ;
- Il valore medio è $E[X] = \mu$;
- La varianza è $Var[X] = \sigma^2$;

In passato a causa dei calcolatori meno potenti e per semplificare i calcoli si faceva uso di una versione standardizzata della normale, la cui densità di probabilità è:

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

Distribuzione Uniforme ($X \sim U(a, b)$)

Una variabile aleatoria X ha distribuzione uniforme in un intervallo $I \equiv [a, b] \subset \mathbb{R}$, ed è denotata con $X \sim U(a, b)$ se è continua e mantiene valore costante:

Si denoti:

- La densità della distribuzione è $f(x) = \frac{1}{b-a}$ se $x > 0$; 0 se $x \leq 0$;
- Il valore medio è $E[X] = \frac{a+b}{2}$;
- La varianza è $Var[X] = \frac{(b-a)^2}{12}$;

Distribuzione X^2 ($X \sim X^2(n)$)

Siano X_i con $i = 1, 2, \dots, n$ variabili aleatorie secondo una $Norm(0,1)$ indipendenti fra loro.

Sia X la variabile aleatorie $X = X_1^2 + X_2^2 + \dots + X_n^2$, diremo che la variabile aleatoria X è distribuita secondo uno chi-quadro con n gradi di libertà, e dotata con $X \sim X^2(n)$, se è assolutamente continua con densità di probabilità:

$$f_x(t) = \frac{1}{\Gamma\left(\frac{n}{2}\right)} \left(\frac{1}{2}\right)^{\frac{n}{2}} t^{\frac{n}{2}-2} e^{\left(-\frac{1}{2}\right)t} \text{ se } t \geq 0, \quad 0 \text{ altrimenti}$$

$$\text{Gamma di Eulero} \equiv \Gamma(a) = \int_0^{+\infty} e^{-x} x^{a-1} dx$$

Si denoti:

- Il valore medio è $E[X] = n$;
- La varianza è $Var[X] = 2n$;

Distribuzione T di Student ($X \sim t(n)$)

Siano $Z \sim Norm(0,1)$ e $Y \sim X^2(n)$ indipendenti fra loro.

Sia X la variabile aleatorie $X = \frac{Z}{\sqrt{\frac{Y}{n}}}$, diremo che la variabile aleatoria X è distribuita secondo una T di Student con n gradi di libertà, e dotata con $X \sim t(n)$, se è assolutamente continua con densità di probabilità:

$$f_x(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} * \left(\frac{1}{\sqrt{n\pi}}\right) * \left(1 + \frac{t^2}{n}\right)^{-\frac{1}{2}(n+1)}$$

$$\text{Gamma di Eulero} \equiv \Gamma(a) = \int_0^{+\infty} e^{-x} x^{a-1} dx$$

Si denoti:

- Il valore medio è $E[X] = 0$ se $n > 1$;
- La varianza è $Var[X] = \frac{n}{n-2}$ se $n > 2$;

Rapporto fra le distribuzioni

- $Bin(n, p)$ si può approssimare a $Norm(\mu = np, \sigma^2 = np)$ se e solo se $n \geq 30, np \geq 5, nq \geq 5$;
- $\overline{Bin}(n, p)$ si può approssimare a $Norm(\mu = \frac{n}{p}, \sigma^2 = \frac{nq}{p^2})$ se e solo se $n \geq 30$;
- $Pois(\lambda)$ si può approssimare a $Norm(\mu = \lambda, \sigma^2 = \lambda)$ se e solo se $\lambda \geq 30$;
- $Ber(p)$ si può approssimare a $Norm(\mu = p, \sigma^2 = \frac{pq}{n})$ se n è molto grande;
- $t(n)$ si può approssimare una $Norm(\mu = 0, \sigma^2 = 1)$ se e solo se $n \geq 30$;

Convergenze

Convergenza di una distribuzione

Si dice che la successione di variabili aleatorie $\{X_n\}$ converge in una distribuzione alla variabile aleatoria X , indicata con $X_n \rightarrow^d X$, se:

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x)$$

Per tutti i valori di x per cui $F_X(x)$ è continua.

Convergenza di in probabilità

Si dice che la successione di variabili aleatorie $\{X_n\}$ converge in probabilità alla variabile aleatoria X , indicata con $X_n \xrightarrow{P} X$, se:

$$\forall \epsilon > 0, \quad \lim_{n \rightarrow \infty} P(|X_n - X| < \epsilon) = 1$$

Convergenza quasi certa

Si dice che la successione di variabili aleatorie $\{X_n\}$ converge quasi certamente alla variabile aleatoria X , indicata con $X_n \xrightarrow{q.c.} X$, se:

$$\forall \epsilon > 0, \quad P\left(\lim_{n \rightarrow \infty} (|X_n - X| < \epsilon)\right) = 1$$

Convergenza in media quadratica

Si dice che la successione di variabili aleatorie $\{X_n\}$ converge in media quadratica alla variabile aleatoria X , indicata con $X_n \xrightarrow{m.q.} X$, se:

$$\lim_{n \rightarrow \infty} E[|X_n - X|^2] = 0$$

Rapporto fra le convergenze

La convergenza *quasi certa* e la convergenza *media quadratica* risultano le *convergenze* più forti poiché da essere è possibile ottenere la *convergenza di probabilità*, da cui è possibile ottenere la *convergenza in una distribuzione*.

Legge debole dei grandi numeri

Sia $\{X_n\}$ una successione di variabili aleatorie identicamente distribuite provenienti da una popolazione X dotata di media ($E[X] = m$), allora:

$$\forall \epsilon > 0, \quad \lim_{n \rightarrow \infty} P(|\overline{X}_n - m| < \epsilon) = 1$$

Ovvero, la successione delle medie campionarie converge in probabilità ad m , il valore atteso di tutte le X_n .

Legge forte dei grandi numeri

Sia $\{X_n\}$ una successione di variabili aleatorie identicamente distribuite provenienti da una popolazione X dotata di media ($E[X] = m$), allora:

$$\forall \epsilon > 0, \quad P\left(\lim_{n \rightarrow \infty} |\overline{X}_n - m| < \epsilon\right) = 1$$

Ovvero, la probabilità che la successione delle medie campionarie converga in probabilità ad m .

Legge centrale del limite

Sia $\{X_n\}$ una successione di variabili aleatorie identicamente distribuite provenienti da una popolazione X dotata di media e varianza ($E[X] = m, Var[X] = v$), allora:

$$\forall x \in \mathbb{R}, \quad \lim_{n \rightarrow \infty} P\left(\frac{\overline{X}_n - m}{\sqrt{\frac{v}{n}}} \leq x\right) = \text{DistribuzioneNormale}(x)$$

Ovvero, la successione delle medie standardizzate converge in legge ad una $Norm(\mu = 0, \sigma^2 = 1)$:

$$\overline{X}_n \sim Norm(\mu = m, \sigma^2 = \frac{v}{n})$$

Stima dei Parametri

Campione casuale

Data una variabile aleatoria X , chiameremo *campione casuale*, o *campioni*, di ampiezza n , una n -tupla di variabili aleatorie fra di loro indipendenti, ciascuna della quali distribuita come la popolazione X .

$$X_1, X_2, \dots, X_n (n_{tuple} \text{ di variabò}) \rightarrow x_1, x_2, \dots, x_n (n_{tuple} \text{ di valori})$$

Media Campionaria (e Valore atteso e varianza della media campionaria)

La *media campionaria* di un *campione casuale* X_1, X_2, \dots, X_n di ampiezza n è dato da:

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

Se la popolazione X è dotata di una media e di una varianza ($E[X] = m$ e $Var[X] = v$), allora si ha:

- Il valore atteso della media campionaria è $E[\bar{X}_n] = m$;
- La varianza della media campionaria è $Var[\bar{X}_n] = \frac{v}{n}$;

Momento Campionario di ordine k

$$M_n^k = \frac{1}{n} \sum_{i=1}^n X_i^k$$

Varianza Campionaria (e valore atteso della varianza campionaria)

La *variazione campionaria* è data a seconda se il *valore atteso della popolazione* è noto o meno:

$$\text{valore atteso noto} \equiv E[X] = m \rightarrow S_o^2 = \frac{1}{n} \sum_{i=1}^n (X_i - m)^2$$

$$\text{valore atteso incognito} \equiv E[X] = ? \rightarrow S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

Se la popolazione X è dotata di una media e di una varianza ($E[X] = m$ e $Var[X] = v$), allora si ha:

- Il valore atteso della media campionaria è $E[S_o^2] = v$;
- La varianza della media campionaria è $Var[S_n^2] = v$;

Si osservi:

- Se la media della popolazione X è nota, allora: $S_o^2 = M_n^2 - 2m\bar{X}_n + m^2$;
- Se la media della popolazione X è incognita, allora: $S_n^2 = \frac{1}{n-1} (nM_n^2 - n\bar{X}_n^2)$;

Distribuzione della media campionaria con popolazione distribuita secondo normale

Sia X_1, X_2, \dots, X_n un campione estratto da una popolazione $X \sim Norm(\mu = m, \sigma^2 = v)$, si ha che:

- La media campionaria è distribuita secondo normale: $\bar{X}_n \sim Norm\left(\mu = m, \sigma^2 = \frac{v}{n}\right)$;
- La varianza campionaria con media nota è distribuita secondo chi-quadro: $S_o^2 \sim \frac{v}{n} X^2 (v = n)$;
- La varianza campionaria con media incognita è distribuita secondo chi-quadro: $S_n^2 \sim \frac{v}{n-1} X^2 (v = n - 1)$;
- La media campionaria \bar{X}_n è indipendente dalla varianza campionaria a media incognita S_n^2 ;
- Il valore: $\frac{\bar{X}_n - m}{\sqrt{\frac{v}{n}}} \sim N(\mu = m, \sigma^2 = 1)$;
- Il valore: $\frac{\bar{X}_n - m}{\sqrt{\frac{S_n^2}{n}}} \sim T(v = n - 1)$

Statistica

Si definisce *statistica* di una variabile aleatoria T , una funzione nota delle variabili aleatorie X_1, \dots, X_n

$$T(X_1, \dots, X_n) \equiv T(X)$$

Stimatore

Data una popolazione dipendente da un parametro ν , e un *campione casuale* X_1, \dots, X_n , si definisce *stimatore* una statistica T che non dipende da ν e che serve per stimare ν .

Correttezza e Correttezza Asintotica di uno stimatore

Uno stimatore si dice *corretto* (e non distorto) se il suo valore atteso coincide con la quantità da stimare:

$$\text{corretto} \equiv E[T_n] = \nu$$

Inoltre, si dice che è *corretto asintoticamente* se con n tendente all'infinito il valore atteso tende a ν :

$$\text{corretto asintoticamente} \equiv E[T_n] \xrightarrow{n \rightarrow \infty} \nu$$

Consistenza in media quadratica e in probabilità di uno stimatore

Uno stimatore si dice *consistente in media quadratica* se:

$$\text{consistente m. q.} \equiv E[(T_n - \nu)^2] \xrightarrow{n \rightarrow \infty} 0$$

Inoltre, si dice che è *consistente in probabilità* se:

$$\text{consistente in probabilità} \equiv \forall \epsilon > 0, \quad P(|T_n - \nu| < \epsilon) \xrightarrow{n \rightarrow \infty} 1$$

Asintotica normalità

$$\frac{T_n - E[T_n]}{\sqrt{\text{Var}[T_n]}} \xrightarrow{n \rightarrow \infty} \text{Norm}(0,1)$$

Costruzione di una stima: Metodo della massima verosomiglianza

Si consideri di avere una successione di variabili aleatorie X_1, \dots, X_n e di conoscere le densità sia per il caso discreto che continuo, si definisce la funzione:

$$\mathcal{L}(x_1, \dots, x_n, \nu) = P_X(x_1, \nu) * \dots * P_X(x_n, \nu) = \prod_{i=1}^n P_X(x_i, \nu)$$

L'obiettivo del metodo della verosomiglianza è scegliere il parametro ν che massimizza la probabilità congiunta.

Si pone la derivata di $\mathcal{L}(\nu) = 0$ e si individua quanto deve valere ν .

Verifica di Ipotesi

Struttura di un test d'ipotesi

Un test di ipotesi consiste nelle seguenti parti:

- *Ipotesi*:
 - L'ipotesi di partenza è detta *ipotesi nulla* ed è indicata con H_0 , questa può essere *semplice* ($H_0 = 5$) o composta ($H_0 \geq 5$ oppure $H_0 \leq 5$);
 - Il contrario dell'ipotesi nulla è l'*ipotesi alternativa* ed è indicata con H_1 ed è quello che dovremmo accettare per vero nel caso in cui rifiutassimo l'ipotesi nulla;
- *Statistica test (U)*:
 - Si tratta di una variabile aleatoria della quale, almeno approssimativamente, possiamo specificare la distribuzione. Indichiamo con u il valore assunto da U ;
- *Regione Critica (C)*:
 - Si rifiuta H_0 se $u \in C$;
 - Non si rifiuta H_0 se $u \notin C$;
- *Errori Possibili*:

- Gli errori che si possono commettere sono i seguenti:

$$\begin{array}{l}
 H_0 \text{ è vera } \begin{cases} \text{se non si rifiuta } H_0, & \text{non c'è errore} \\ \text{se si rifiuta } H_0, & \text{errore 1° tipo} \end{cases} \\
 H_0 \text{ è falsa } \begin{cases} \text{se non si rifiuta } H_0, & \text{errore 2° tipo} \\ \text{se si rifiuta } H_0, & \text{non c'è errore} \end{cases}
 \end{array}$$

- Livello di *significatività* (α):
 - Livello di soglia α , tale che se la probabilità di commettere l'errore del 1° tipo non sia superiore ad α ;
- Funzione potenza ($\pi(\nu)$):
 - Se il test è un test parametrico, ovvero riguarda un parametro, la probabilità di rifiutare l'ipotesi dipende dal valore incognito del parametro ν . Tale probabilità è chiamata funzione potenza del test:

$$\pi(\nu) = P_\nu(u \in C)$$
- P-value:
 - Ci si può chiedere a quale livello di significatività corrisponde una regione critica la cui frontiera fosse costituita proprio da u , cioè per quale livello di significatività il valore asserito da u valica la frontiera di C . Questo valore prende il nome di p-value ed è definito come:

$$p_{value} = \sup\{a \text{ tale che } u \notin C\} = \inf\{a \text{ tale che } u \in C\}$$
 Dato un certo test, rifiuteremo H_0 per tutti i livelli maggiori del p-value e non lo rifiuteremo per tutti i livelli inferiori;

Test parametrici riguardanti una sola popolazione

Test riguardante la media di una normale

Supponiamo di avere una variabile aleatoria distribuita secondo normale $X \sim Norm(\mu, \sigma^2)$ e che n sia il numero di osservazioni di tale variabile aleatoria. Sul parametro μ possiamo dare le seguenti ipotesi:

$$\text{Tipo 1 } \begin{cases} H_0, & \text{se } \mu = \mu_0 \\ H_1, & \text{se } \mu \neq \mu_0 \end{cases} \mid \text{Tipo 2 } \begin{cases} H_0, & \text{se } \mu = \mu_0 \text{ o } \mu \leq \mu_0 \\ H_1, & \text{se } \mu > \mu_0 \end{cases} \mid \text{Tipo 3 } \begin{cases} H_0, & \text{se } \mu = \mu_0 \text{ o } \mu \geq \mu_0 \\ H_1, & \text{se } \mu < \mu_0 \end{cases}$$

La statistica U , la regione critica e il p-value cambiano a seconda se il parametro σ^2 è noto o incognito.

Z-Test (Caso $\sigma^2 = v \equiv \text{noto}$)

Avendo la varianza nota, sappiamo che la media campionaria è pari a $\bar{X}_n \sim Norm(\mu, \sigma^2 = v)$.

Ciò equivale a dire che: $\frac{\bar{X}_n - \mu}{\sqrt{\frac{v}{n}}} \sim N(0,1)$.

Sceglieremo come statistica la standardizzazione che sarebbe corretto eseguire, se fosse vero H_0 , ossia scegliamo come per μ il valore μ_0 .

Avremo $U = \frac{\bar{X}_n - \mu_0}{\sqrt{\frac{v}{n}}}$, diremo quindi che $U = \frac{\bar{X}_n - \mu_0}{\sqrt{\frac{v}{n}}} \sim_{H_0} Norm(0,1)$.

A seconda della tipologia di ipotesi, si procede diversamente:

- Nel caso fossimo nell'ipotesi di tipo 1:
 - Propenderemo per H_0 se il valore asserito da u di U sarà "vicino allo 0", mentre propenderemo per H_1 se u si troverà in una delle due code della distribuzione. Si deve stabilire un valore di \bar{z} in corrispondenza del quale si inizierà a rifiutare H_0 ; poiché il test ha valore di significatività α , vogliamo che $\alpha = \max P(|U| > \bar{z} | H_0 \text{ sia vera})$, cioè α deve essere il valore massimo per cui u finisce in una delle due code. Il valore \bar{z} è quindi $\bar{z} = z_{1-\frac{\alpha}{2}}$ e rifiuteremo H_0 se $|u| > z_{1-\frac{\alpha}{2}}$.
- Nel caso fossimo nell'ipotesi di tipo 2:
 - Propenderemo per H_0 se il valore asserito da u di U sarà "vicino allo 0", mentre usando un valore "troppo piccolo" per la media ci farà propendere per H_1 se u si troverà in nella coda destra. Il valore \bar{z} è quindi $\bar{z} = z_{1-\alpha}$ e rifiuteremo H_0 se $u > z_{1-\alpha}$.
- Nel caso fossimo nell'ipotesi di tipo 3:

- Propenderemo per H_0 se il valore asserito da u di U sarà “vicino allo 0”, mentre usando un valore “troppo grande” per la media ci farà propendere per H_1 se u si troverà in nella coda sinistra. Il valore \bar{z} è quindi $\bar{z} = -z_{1-\alpha}$ e rifiuteremo H_0 se $u < -z_{1-\alpha}$ ovvero se $-u > z_{1-\alpha}$.

Tabella Riassuntiva dei risultati Z-Test	
Tipo	Condizione di Rifiuto
1° Tipo	$ u > z_{1-\frac{\alpha}{2}}$
2° Tipo	$u > z_{1-\alpha}$
3° Tipo	$u < -z_{1-\alpha}$ oppure $-u > z_{1-\alpha}$

T-Test (Caso $\sigma^2 = ? \equiv$ incognito)

Avendo la varianza incognita, sappiamo che se H_0 è vera il valore $\frac{\bar{X}_n - \mu_0}{\sqrt{\frac{S_n^2}{n}}} \sim_{H_0} T(v = n - 1)$

Sceghieremo come statistica la standardizzazione che sarebbe corretto eseguire, se fosse vero H_0 , ossia scegliamo come per μ il valore μ_0 .

Avremo $U = \frac{\bar{X}_n - \mu_0}{\sqrt{\frac{S_n^2}{n}}}$, diremo quindi che $U = \frac{\bar{X}_n - \mu_0}{\sqrt{\frac{S_n^2}{n}}} \sim_{H_0} T(v = n - 1)$.

A seconda della tipologia di ipotesi, si procede diversamente:

- Nel caso fossimo nell'ipotesi di tipo 1:
 - Propenderemo per H_0 se il valore asserito da u di U sarà “vicino allo 0”, mentre propenderemo per H_1 se u si troverà in una delle due code della distribuzione. Si deve stabilire un valore di \bar{t} in corrispondenza del quale si inizierà a rifiutare H_0 ; poiché il test ha valore di significatività α , vogliamo che $\alpha = \max P(|U| > \bar{t} | H_0 \text{ sia vera})$, cioè α deve essere il valore massimo per cui u finisce in una delle due code. Il valore \bar{t} è quindi $\bar{t} = t_{1-\frac{\alpha}{2}}$ e rifiuteremo H_0 se $|u| > t_{1-\frac{\alpha}{2}}$.
- Nel caso fossimo nell'ipotesi di tipo 2:
 - Propenderemo per H_0 se il valore asserito da u di U sarà “vicino allo 0”, mentre usando un valore “troppo piccolo” per la media ci farà propendere per H_1 se u si troverà in nella coda destra. Il valore \bar{t} è quindi $\bar{t} = -t_{1-\alpha}$ e rifiuteremo H_0 se $-u > t_{1-\alpha}$.
- Nel caso fossimo nell'ipotesi di tipo 3:
 - Propenderemo per H_0 se il valore asserito da u di U sarà “vicino allo 0”, mentre usando un valore “troppo grande” per la media ci farà propendere per H_1 se u si troverà in nella coda sinistra. Il valore \bar{t} è quindi $\bar{t} = t_{1-\alpha}$ e rifiuteremo H_0 se $u < -t_{1-\alpha}$ ovvero se $-u > t_{1-\alpha}$.

Tabella Riassuntiva dei risultati T-Test	
Tipo	Condizione di Rifiuto
1° Tipo	$ u > t_{1-\frac{\alpha}{2}}$
2° Tipo	$u > t_{1-\alpha}$
3° Tipo	$u < -t_{1-\alpha}$ oppure $-u > t_{1-\alpha}$

Test riguardante la varianza di una normale

Supponiamo di avere una variabile aleatoria distribuite secondo normale $X \sim \text{Norm}(\mu, \sigma^2)$ e che n sia il numero di osservazioni di tale variabile aleatoria. Sul parametro σ^2 possiamo dare le seguenti ipotesi:

$$\text{Tipo 1 } \left\{ \begin{array}{l} H_0, \text{ se } \sigma^2 = \sigma_0^2 \\ H_1, \text{ se } \sigma^2 \neq \sigma_0^2 \end{array} \right. \mid \text{Tipo 2 } \left\{ \begin{array}{l} H_0, \text{ se } \mu = \sigma^2 \text{ o } \mu \leq \sigma_0^2 \\ H_1, \text{ se } \sigma^2 > \sigma_0^2 \end{array} \right. \mid \text{Tipo 3 } \left\{ \begin{array}{l} H_0, \text{ se } \sigma^2 = \sigma^2 \text{ o } \mu \geq \sigma_0^2 \\ H_1, \text{ se } \sigma^2 < \sigma_0^2 \end{array} \right.$$

La statistica U , la regione critica e il p-value cambiano a seconda se il parametro μ è noto o incognito.

Caso $\mu = m \equiv \text{noto}$

Avendo la media nota, sappiamo che la varianza campionaria è pari a $S_o^2 = \frac{1}{n} \sum_{i=1}^n (X_i - m)^2$.

Ciò equivale a dire che: $S_o^2 \sim \frac{\sigma^2}{n} X^2(v = n)$ da cui segue $\frac{n}{\sigma^2} S_o^2 \sim X^2(v = n)$;

Sceglieremo come statistica la standardizzazione che sarebbe corretto eseguire, se fosse vero H_o , ossia scegliamo come per σ^2 il valore σ_o^2 .

Avremo $U = \frac{n}{\sigma_o^2} S_o^2$, diremo quindi che $U = \frac{n}{\sigma^2} S_o^2 \sim^{H_o} X^2(v = n)$.

A seconda della tipologia di ipotesi, si procede diversamente:

- Nel caso fossimo nell'ipotesi di tipo 1:
 - Propenderemo per H_o se il valore asserito da u di U sarà in corrispondenza della regione dove la densità di X^2 è maggiore, mentre propenderemo per H_1 se u si troverà in una delle due code. Si devono stabilire i valori di \bar{c}_1 e \bar{c}_2 in corrispondenza del quale si inizierà a rifiutare H_o ; I valori sono: $\bar{c}_1 = X_{\frac{\alpha}{2}}^2$, $\bar{c}_2 = X_{1-\frac{\alpha}{2}}^2$ e rifiuteremo H_o se $u < X_{\frac{\alpha}{2}}^2$ oppure $u > X_{1-\frac{\alpha}{2}}^2$;
- Nel caso fossimo nell'ipotesi di tipo 2:
 - Propenderemo per H_o se il valore asserito da u di U sarà in corrispondenza della regione dove la densità di X^2 è maggiore, mentre propenderemo per H_1 se u si troverà nella coda destra. Il valore \bar{c} è quindi $\bar{c} = X_{1-\alpha}^2$ e rifiuteremo H_o se $u > X_{1-\alpha}^2$.
- Nel caso fossimo nell'ipotesi di tipo 3:
 - Propenderemo per H_o se il valore asserito da u di U sarà in corrispondenza della regione dove la densità di X^2 è maggiore, mentre propenderemo per H_1 se u si troverà in nella coda sinistra. Il valore \bar{c} è quindi $\bar{c} = X_{\alpha}^2$ e rifiuteremo H_o se $u < X_{\alpha}^2$.

Tabella Riassuntiva del test sulla varianza con media nota	
Tipo	Condizione di Rifiuto
1° Tipo	$u < X_{\frac{\alpha}{2}}^2$ oppure $u > X_{1-\frac{\alpha}{2}}^2$
2° Tipo	$u > X_{1-\alpha}^2$
3° Tipo	$u < X_{\alpha}^2$

Caso $\mu = ? \equiv \text{incognito}$

Avendo la media incognita, sappiamo che la varianza campionaria è pari a $S_n^2 = \frac{1}{n-1} (nM_n^2 - n\bar{X}_n^2)$.

Ciò equivale a dire che: $S_n^2 = \frac{v}{n-1} X^2(v = n - 1)$ da cui segue $\frac{n-1}{\sigma^2} S_n^2 = X^2(v = n - 1)$;

Sceglieremo come statistica la standardizzazione che sarebbe corretto eseguire, se fosse vero H_o , ossia scegliamo come per σ^2 il valore σ_o^2 .

Avremo $U = \frac{n-1}{\sigma_o^2} S_n^2$, diremo quindi che $U = \frac{n-1}{\sigma^2} S_n^2 = X^2(v = n - 1)$;

A seconda della tipologia di ipotesi, si procede diversamente:

- Nel caso fossimo nell'ipotesi di tipo 1:
 - Propenderemo per H_o se il valore asserito da u di U sarà in corrispondenza della regione dove la densità di X^2 è maggiore, mentre propenderemo per H_1 se u si troverà in una delle due code. Si devono stabilire i valori di \bar{c}_1 e \bar{c}_2 in corrispondenza del quale si inizierà a rifiutare H_o ; I valori sono: $\bar{c}_1 = X_{\frac{\alpha}{2}}^2$, $\bar{c}_2 = X_{1-\frac{\alpha}{2}}^2$ e rifiuteremo H_o se $u < X_{\frac{\alpha}{2}}^2$ oppure $u > X_{1-\frac{\alpha}{2}}^2$;
- Nel caso fossimo nell'ipotesi di tipo 2:
 - Propenderemo per H_o se il valore asserito da u di U sarà in corrispondenza della regione dove la densità di X^2 è maggiore, mentre propenderemo per H_1 se u si troverà nella coda destra. Il valore \bar{c} è quindi $\bar{c} = X_{1-\alpha}^2$ e rifiuteremo H_o se $u > X_{1-\alpha}^2$.

- Nel caso fossimo nell'ipotesi di tipo 3:
 - Propenderemo per H_0 se il valore asserito da u di U sarà in corrispondenza della regione dove la densità di X^2 è maggiore, mentre propenderemo per H_1 se u si troverà in nella coda sinistra. Il valore \bar{c} è quindi $\bar{c} = X_a^2$ e rifiuteremo H_0 se $u < X_a^2$.

Tabella Riassuntiva del test sulla varianza con media incognita	
Tipo	Condizione di Rifiuto
1° Tipo	$u < X_{\frac{\alpha}{2}}^2$ oppure $u > X_{1-\frac{\alpha}{2}}^2$
2° Tipo	$u > X_{1-\alpha}^2$
3° Tipo	$u < X_{\alpha}^2$

Test Asintotico su p (probabilità di successo di un processo di Bernulli)

Dato che una $Ber(p)$ si può approssimare a $Norm(\mu = p, \sigma^2 = \frac{pq}{n})$ se n molto grandi su può applicare la stessa logica del test sulla media di una normale, ponendo come statistica: $U = \frac{\bar{X}_n - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim^{H_0} Norm(0,1)$.

Test parametrici di confronto fra due popolazioni

Supponiamo di avere due popolazioni: X_1, \dots, X_n e Y_1, \dots, Y_m con $n, m \in \mathbb{N}$.

Supponiamo di avere le seguenti variabili aleatorie: $X \sim Norm(\mu = \mu_x, \sigma^2 = \sigma_x^2)$ e $Y \sim Norm(\mu = \mu_y, \sigma^2 = \sigma_y^2)$. Sul parametro μ possiamo dare le seguenti ipotesi:

$$Tipo\ 1 \begin{cases} H_0, \text{ se } \mu_x = \mu_y \\ H_1, \text{ se } \mu_x \neq \mu_y \end{cases} \mid Tipo\ 2 \begin{cases} H_0, \text{ se } \mu_x = \mu_y \text{ o } \mu_x \leq \mu_y \\ H_1, \text{ se } \mu_x > \mu_y \end{cases} \mid Tipo\ 3 \begin{cases} H_0, \text{ se } \mu_x = \mu_y \text{ o } \mu_x \geq \mu_y \\ H_1, \text{ se } \mu_x < \mu_y \end{cases}$$

Si ha $\bar{X}_n \sim Norm(\mu = \mu_x, \sigma^2 = \frac{\sigma_x^2}{n})$ e $\bar{Y}_m \sim Norm(\mu = \mu_y, \sigma^2 = \frac{\sigma_y^2}{m})$ da cui è possibile ricavare che:

$$\bar{X}_n - \bar{Y}_m \sim Norm(\mu = \mu_x - \mu_y, \sigma^2 = \frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m})$$

La statistica U , la regione critica e il p-value cambiano a seconda se il parametro σ^2 sono noti, uguali o incogniti.

Caso σ_x^2 e σ_y^2 note

Avendo le varianze note, sappiamo che la media campionaria è $\frac{(\bar{X}_n - \bar{Y}_m) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}} \sim Norm(0,1)$.

Sceghieremo come statistica la standardizzazione che sarebbe corretto eseguire, se fosse vero H_0 , ossia scegliamo come per μ il valore μ_0 .

$$\text{Avremo } U = \frac{(\bar{X}_n - \bar{Y}_m) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}} \text{ diremo quindi che } U = \frac{(\bar{X}_n - \bar{Y}_m) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}} \sim^{H_0} Norm(0,1).$$

A seconda della tipologia di ipotesi, si procede diversamente:

- Nel caso fossimo nell'ipotesi di tipo 1:
 - Propenderemo per H_0 se il valore asserito da u di U sarà "vicino allo 0", mentre propenderemo per H_1 se u si troverà in una delle due code della distribuzione. Si deve stabilire un valore di \bar{z} in corrispondenza del quale si inizierà a rifiutare H_0 ; poiché il test ha valore di significatività α , vogliamo che $\alpha = \max P(|U| > \bar{z} | H_0 \text{ sia vera})$, cioè α deve essere il valore massimo per cui u finisce in una delle due code. Il valore \bar{z} è quindi $\bar{z} = z_{1-\frac{\alpha}{2}}$ e rifiuteremo H_0 se $|u| > z_{1-\frac{\alpha}{2}}$.
- Nel caso fossimo nell'ipotesi di tipo 2:

- Propenderemo per H_0 se il valore asserito da u di U sarà “vicino allo 0”, mentre usando un valore “troppo piccolo” per la media ci farà propendere per H_1 se u si troverà in nella coda destra. Il valore \bar{z} è quindi $\bar{z} = z_{1-\alpha}$ e rifiuteremo H_0 se $u > z_{1-\alpha}$.
- Nel caso fossimo nell’ipotesi di tipo 3:
 - Propenderemo per H_0 se il valore asserito da u di U sarà “vicino allo 0”, mentre usando un valore “troppo grande” per la media ci farà propendere per H_1 se u si troverà in nella coda sinistra. Il valore \bar{z} è quindi $\bar{z} = -z_{1-\alpha}$ e rifiuteremo H_0 se $u < -z_{1-\alpha}$ ovvero se $-u > z_{1-\alpha}$.

Tabella Riassuntiva dei risultati Z-Test	
Tipo	Condizione di Rifiuto
1° Tipo	$ u > z_{1-\frac{\alpha}{2}}$
2° Tipo	$u > z_{1-\alpha}$
3° Tipo	$u < -z_{1-\alpha}$ oppure $-u > z_{1-\alpha}$

Caso σ_x^2 e σ_y^2 note e uguali

Avendo le varianze note e uguali si ha che:

$$S_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X}_n)^2, S_y^2 = \frac{1}{m-1} \sum_{i=1}^m (y_i - \bar{Y}_m)^2, S_p^2 = \frac{(n-1)S_x^2 + (m-1)S_y^2}{m+n-2}$$

Avremo $U = \frac{(\bar{X}_n - \bar{Y}_m)}{\sqrt{S_p^2(\frac{1}{n} + \frac{1}{m})}}$ diremo quindi che $U = \frac{(\bar{X}_n - \bar{Y}_m)}{\sqrt{S_p^2(\frac{1}{n} + \frac{1}{m})}} \sim^{H_0} t(v = n + m - 2)$.

A seconda della tipologia di ipotesi, si procede diversamente:

- Nel caso fossimo nell’ipotesi di tipo 1:
 - Propenderemo per H_0 se il valore asserito da u di U sarà “vicino allo 0”, mentre propenderemo per H_1 se u si troverà in una delle due code della distribuzione. Si deve stabilire un valore di \bar{t} in corrispondenza del quale si inizierà a rifiutare H_0 ; poiché il test ha valore di significatività α , vogliamo che $\alpha = \max P(|U| > \bar{t} | H_0 \text{ sia vera})$, cioè α deve essere il valore massimo per cui u finisce in una delle due code. Il valore \bar{t} è quindi $\bar{t} = t_{1-\frac{\alpha}{2}}$ e rifiuteremo H_0 se $|u| > t_{1-\frac{\alpha}{2}}$.
- Nel caso fossimo nell’ipotesi di tipo 2:
 - Propenderemo per H_0 se il valore asserito da u di U sarà “vicino allo 0”, mentre usando un valore “troppo piccolo” per la media ci farà propendere per H_1 se u si troverà in nella coda destra. Il valore \bar{t} è quindi $\bar{t} = -t_{1-\alpha}$ e rifiuteremo H_0 se $-u > t_{1-\alpha}$.
- Nel caso fossimo nell’ipotesi di tipo 3:
 - Propenderemo per H_0 se il valore asserito da u di U sarà “vicino allo 0”, mentre usando un valore “troppo grande” per la media ci farà propendere per H_1 se u si troverà in nella coda sinistra. Il valore \bar{t} è quindi $\bar{t} = t_{1-\alpha}$ e rifiuteremo H_0 se $u < -t_{1-\alpha}$ ovvero se $-u > t_{1-\alpha}$.

Tabella Riassuntiva dei risultati T-Test	
Tipo	Condizione di Rifiuto
1° Tipo	$ u > t_{1-\frac{\alpha}{2}}$
2° Tipo	$u > t_{1-\alpha}$
3° Tipo	$u < -t_{1-\alpha}$ oppure $-u > t_{1-\alpha}$

Statistica Descrittiva		
Con E si indica l'insieme dei dati	Con n si indica la numerosità di E	Con S si indica l'insieme delle modalità
Frequenza Assoluta: f_j $= n^\circ \text{ di elementi di } E \text{ di con valore } s_j$	Frequenza Relativa: $p_j = \frac{f_j}{n}$	Frequenza cumula assoluta: $F_j = \sum_k f_k$ $k \text{ tale che } s_k \leq s_j \text{ e } j = 1, 2, \dots, N$
Frequenza cumula relativa: $P_j = \sum_k p_k$ $k \text{ tale che } s_k \leq s_j \text{ e } j = 1, 2, \dots, N$	Numero ottimale di Classi: $k \cong \sqrt{n}$	Ampiezza ottimale per Classi: $a = \frac{R}{k}, R \equiv \max - \min$
Media (\bar{x}) $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$	Media pesata (\bar{x}_p) $\bar{x}_p = \frac{1}{p} \sum_{i=1}^n p_i x_i$	Mediana (\hat{x}) Se N è dispari: <i>elemento</i> $\hat{x} = \text{dati} \left[\frac{n+1}{2} \right]$ Se N è pari: $\hat{x} = \frac{\left(\frac{n}{2} \right) + \left(\frac{n}{2} + 1 \right)}{2}$
Varianza (s^2) $s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ oppure $s^2 = \overline{x^2} - \bar{x}^2$		Deviazione standard $s = \sqrt{s^2}$
Scarto medio assoluto $s.a = \frac{1}{n} \sum_{i=1}^n x_i - \bar{x} $		Coefficiente di variazione $c.v = \frac{s}{\bar{x}}$
Indice di asimmetria (sk) $sk = \frac{m_3}{\sqrt{m_2^3}}, m_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k \text{ con } k = 2, 3, 4, \dots$		
Indice di curtosi (K) $K = \frac{m_4}{m_2^2}, m_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k \text{ con } k = 2, 3, 4, \dots$		
Cenni del Calcolo Combinatorio		
Fattoriale $n! = \begin{cases} 0, & \text{se } n = 0 \\ \prod_{k=1}^n k, & \text{se } n \in \mathbb{N}, n \neq 0 \end{cases}$		
Coefficiente Binomiale $c.b \equiv \binom{n}{k} = \begin{cases} \frac{n!}{k!(n-k)!}, & \text{se } n, k \in \mathbb{N}, \quad 0 \leq k \leq n \\ 0, & \text{se } n, k \in \mathbb{N}, \quad 0 \leq n < k \end{cases}$		
Permutazioni senza ripetizione $P_n = n!$	Permutazioni con ripetizione $P_n^{n_1, n_2, \dots, n_k} = \frac{n!}{n_1! n_2! \dots n_k!}$	Permutazioni cicliche $P_n^c = (n-1)!$
Disposizioni senza ripetizione $D_{n,k} = \frac{n!}{(n-k)!}$		Disposizioni con ripetizione $D'_{n,k} = n^k$
Combinazioni senza ripetizione $C_{n,k} = \binom{n}{k} = \frac{n!}{k!(n-k)!}$		Combinazioni con ripetizione $C'_{n,k} = \binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-1)!}$

Elementi di Probabilità	
Probabilità secondo Kolmogorov $P(A) : \mathcal{F} \rightarrow [0, 1]$	
Probabilità Classica (Applicabile negli spazi di probabilità finiti) $P = \frac{n. \text{ casi favorevoli}}{n. \text{ casi possibili}} \mid P = \frac{\text{frequenza successi}}{\text{numero di prove}}$	
Probabilità della somma logica ($P(E \cup F)$) $P(E \cup F) = P(E) + P(F) - P(E \cap F)$	
Probabilità condizionata ($P(A B)$) $P(A B) = \frac{P(A \cap B)}{P(B)}$	
Probabilità del prodotto logico ($P(E \cap F)$) $P(A \cap B) = P(B)P(A B) = P(A)P(B A)$	Teorema della catena $P(E_1 \cap \dots \cap E_n)$ $= P(E_1)P(E_2 E_1) \dots P(E_n E_1 \cap \dots \cap E_{n-1})$
Probabilità Totale (Legge delle alternative) $P(B) = \sum_{i=1} P(A_i \cap B) = \sum_{i=1} P(A_i)P(B A_i) \text{ oppure } P(F E) = \sum_{i=1} P(F A_i \cap E)P(A_i E)$	
Teorema di Bayes $P(A_i B) = \frac{P(A_i)P(B A_i)}{P(B)} = \frac{P(A_i)P(B A_i)}{\sum_{i=1} P(A_i)P(B A_i)}$	
Valore Atteso per X discreta $\mu = E[X] = \sum_i x_i * P(X = x_i)$	Valore Atteso per X continua $\mu = E[X] = \int_{-\infty}^{\infty} x * f_x(x) dx$
Varianza di X $Var[X] = E[(x - \mu)^2]$ oppure $Var[X] = E[X^2] - (E[X])^2$ oppure $Var[X] = \mu_2 - \mu_1^2$	
Momento N-esimo di X discreta $\mu_n = E[X^n] = \sum_i x_i^n * P(X = x_i)$	Momento N-esimo di X continua $\mu_n = E[X^n] = \int_{-\infty}^{\infty} x^n * f_x(x) dx$
Densità congiunta per vettori discreti $P_{XY} = (x_i, y_j) = P(X = x_i, Y = y_j)$	Densità marginali per vettori discreti <ul style="list-style-type: none"> $P_X(x_i) = \sum_j P_{XY}(x_i, y_j);$ $P_Y(y_j) = \sum_i P_{XY}(x_i, y_j);$
Densità congiunta per vettori continui $P(a < x < b, c < y < d) = \int_a^b dx \int_c^d dy f_{XY}(x, y)$	Densità marginali per vettori continui <ul style="list-style-type: none"> $f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dy;$ $f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dx;$
Densità della somma di variabili aleatorie <i>caso discreto</i> $\rightarrow P_U(u) = \sum_{k=-\infty}^{+\infty} P_{XY}(k, u - k)$ <i>caso continuo</i> $\rightarrow f_U(u) = \int_{-\infty}^{+\infty} f_{XY}(t, u - t) dt$	
Valore atteso per Vettore aleatorio <i>caso discreto</i> $\rightarrow E[g(X, Y)] = \sum_i \sum_j g(x_i, y_j) P_{XY}(x_i, y_j)$	

$$\text{caso continuo} \rightarrow E[g(X,Y)] = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy g(x,y) f_{XY}(x,y)$$

Covarianza

$$Cov[X,Y] = E[(X - E[X]) * (Y - E[Y])]$$

oppure

$$Cov[X,Y] = E[XY] - E[X] * E[Y]$$

Coefficiente di Correlazione

$$\rho_{XY} = \frac{Cov[X,Y]}{\sqrt{Var[X]Var[Y]}}$$

Distribuzione Notevoli

$p = \text{probabilità di successo}, q = 1 - p \equiv \text{insuccesso}$

Distribuzione di Bernulli ($X \sim Ber(p)$)

- La densità della distribuzione è $P(X = k) = p^k q^{1-k}$;
- La funzione generatrice di momenti è $m_X(t) = pe^t + q$;
- Il valore medio è $E[X] = \frac{d}{dt}[pe^t + q]_{t=0} = p$;
- La varianza è $Var[X] = pq$;

Distribuzione Binomiale ($X \sim Bin(n,p)$)

- La densità della distribuzione è $P(X = k) = \binom{n}{k} p^k q^{n-k}$;
- La funzione generatrice di momenti è $m_X(t) = (q + pe^t)^n$;
- Il valore medio è $E[X] = \frac{d}{dt}[(q + pe^t)^n]_{t=0} = np$;
- La varianza è $Var[X] = npq$;

Distribuzione Geometrica ($X \sim Geo(p)$)

- La densità della distribuzione è $P(X = k) = pq^{k-1}$;
- La funzione generatrice di momenti è $m_X(t) = \frac{pe^t}{1-qe^t}$ per $t < \ln(\frac{1}{q})$;
- Il valore medio è $E[X] = \frac{d}{dt}\left[\frac{pe^t}{1-qe^t}\right]_{t=0} = \frac{1}{p}$;
- La varianza è $Var[X] = \frac{q}{p^2}$;

Distribuzione Pascal o Binomiale Negativa ($X \sim \overline{Bin}(n,p)$)

- La densità della distribuzione è $P(X = k) = \binom{-n}{k} p^n (-q)^k$;
- La funzione generatrice di momenti è $m_X(t) = \left(\frac{pe^t}{1-qe^t}\right)^n$ per $t < \ln(\frac{1}{q})$;
- Il valore medio è $E[X] = \frac{d}{dt}\left[\left(\frac{pe^t}{1-qe^t}\right)^n\right]_{t=0} = \frac{n}{p}$;
- La varianza è $Var[X] = n(\frac{q}{p^2})$;

Distribuzione Poisson ($X \sim Pois(\lambda)$)

- La densità della distribuzione è $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$;
- La funzione generatrice di momenti è $m_X(t) = e^{\lambda(e^t-1)}$;
- Il valore medio è $E[X] = \frac{d}{dt}[e^{\lambda(e^t-1)}]_{t=0} = \lambda$;
- La varianza è $Var[X] = \lambda$;

Distribuzione Esponenziale ($X \sim Exp(\lambda)$)

- La densità della distribuzione è $P(x) = \lambda e^{-\lambda x}$ se $x > 0$; 0 se $x \leq 0$;
- La funzione generatrice di momenti è $m_X(t) = \frac{\lambda}{\lambda - t}$ per $t < \lambda$;
- Il valore medio è $E[X] = \frac{d}{dt}\left[\frac{\lambda}{\lambda - t}\right]_{t=0} = \frac{1}{\lambda}$;
- La varianza è $Var[X] = \frac{1}{\lambda^2}$;

Distribuzione Gamma ($X \sim \Gamma(n, \lambda)$)

- La densità della distribuzione è $P(x) = \frac{1}{(n-1)!} \lambda^n x^{n-1} e^{-\lambda x}$ se $x > 0$; 0 se $x \leq 0$;

Distribuzione di Weibul ($X \sim Wei(n, \lambda)$)

- La densità della distribuzione è $P(x) = \lambda a x^{a-1} e^{-\lambda x^a}$ se $x > 0$; 0 se $x \leq 0$;

- Il valore medio è $E[X] = a^{-\frac{1}{\lambda}} \Gamma(1 + \frac{1}{\lambda})$;
- La varianza è $Var[X] = a^{-\frac{2}{\lambda}} \left(\Gamma(1 + \frac{2}{\lambda}) - \Gamma(1 + \frac{1}{\lambda})^2 \right)$;

Distribuzione di Gauss o Normale ($X \sim N(\mu, \sigma)$)

- La funzione f_x è simmetrica rispetto alla retta $x = \mu$;
- Ha un massimo assoluto in $(\mu, \frac{1}{\sigma\sqrt{2\pi}})$;
- Ha un asintoto orizzontale $y = 0$;
- Ha due flessi rispettivamente in $(\mu - \sigma, \frac{1}{\sigma\sqrt{2\pi}e})$ e $(\mu + \sigma, \frac{1}{\sigma\sqrt{2\pi}e})$;
- All'aumentare di σ il massimo *diminuisce*;
- Se $\mu = 0$ allora è simmetrica rispetto all'asse delle y;
- Il valore medio è $E[X] = \mu$;
- La varianza è $Var[X] = \sigma^2$;
- La funzione di densità è: $f_x(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Distribuzione Uniforme ($X \sim U(a, b)$)

- La densità della distribuzione è $f(x) = \frac{1}{b-a}$ se $x > 0$; 0 se $x \leq 0$;
- Il valore medio è $E[X] = \frac{a+b}{2}$;
- La varianza è $Var[X] = \frac{(b-a)^2}{12}$;

Distribuzione X^2 ($X \sim X^2(n)$)

- La densità di distribuzione è

$$f_x(t) = \frac{1}{\Gamma(\frac{n}{2})} \left(\frac{1}{2}\right)^{\frac{n}{2}} t^{\frac{n}{2}-2} e^{-\frac{1}{2}t} \text{ se } t \geq 0, \quad 0 \text{ altrimenti}$$

- Il valore medio è $E[X] = n$;
- La varianza è $Var[X] = 2n$;

Distribuzione T di Student ($X \sim t(n)$)

- La densità di distribuzione è:

$$f_x(t) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} * \left(\frac{1}{\sqrt{n\pi}}\right) * \left(1 + \frac{t^2}{n}\right)^{-\frac{1}{2}(n+1)}$$

- Il valore medio è $E[X] = 0$ se $n > 1$;
- La varianza è $Var[X] = \frac{n}{n-2}$ se $n > 2$;

Stima dei Parametri

Media Campionaria

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

Se la popolazione X è dotata di una media e di una varianza ($E[X] = m$ e $Var[X] = v$), allora si ha:

- Il valore atteso della media campionaria è $E[\bar{X}_n] = m$;
- La varianza della media campionaria è $Var[\bar{X}_n] = \frac{v}{n}$;

Varianza Campionaria

$$\text{valore atteso noto} \equiv E[X] = m \rightarrow S_o^2 = \frac{1}{n} \sum_{i=1}^n (X_i - m)^2$$

$$\text{valore atteso incognito} \equiv E[X] = ? \rightarrow S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

Se la popolazione X è dotata di una media e di una varianza ($E[X] = m$ e $Var[X] = v$), allora si ha:

- Il valore atteso della media campionaria è $E[S_o^2] = v$;
- La varianza della media campionaria è $Var[S_n^2] = v$;

Statistica $T(X_1, \dots, X_n) \equiv T(X)$	Asintotica normalità $\frac{T_n - E[T_n]}{\sqrt{Var[T_n]}} \sim_{n \rightarrow \infty} Norm(0,1)$	

Mathlab Compendium

Metodo/Azione	Descrizione
Assegnamento variabili	Analogo ai programmi di programmazione
Carattere “;” a fine istruzione	Rende l’esecuzione dell’istruzione silenziosa, non sarà stampato l’output informativo
<i>who/whos</i>	Informa su tutte le variabili presenti nel workspace
<i>clear [var]</i>	Cancella tutte o la variabile fornita dal workspace
<i>clc</i>	Pulisce lo standard output
<i>disp(x)</i>	Stampa sullo standard output il contenuto della variabile fornita
Creazione Array	Analogo a Python: <i>arr = [1, 2, 3, ...]</i> Si possono creare array inizializzati a saltelli con: <i>arr = start: step: end</i> Vi sono funzioni helper per la creazione rapida: $arr = zeros(rows, cols)$ e $ones(rows, cols)$ La trasposta di una matrice si ottiene con: $t = arr'$ L’inversa di una matrice si ottiene con: $i = inv(arr)$
Accesso agli array	Analogo ai linguaggi di programmazione ma con indici a partire da 1 e parentesi tonde: $val = arr(1,1)$ Usare come indice il carattere “:” indica tutto
Creazione cella	Analoghe agli <i>object</i> di JS, contengono dati di vario tipo: $\{data1, data2, \dots\}$ L’accesso è analogo agli array.
Creazione Struct	Le struct posso essere visti come array associativi.
Creare funzioni	Una funzione segue la seguente definizione: $function < return_{var} > = nome(argomenti)$ \dots end Si possono creare funzioni inline con: $nome = @(parametri \dots) operazione$
Cicli for	Un ciclo for ha la seguente definizione: $for i = assegnamento$ \dots end
<i>size(arr)</i>	Restituisce il numero di colonne e di righe dell’array fornito
<i>readtable(filepath)</i>	Carica nel workspace un oggetto di tipo tabella dal file fornito
<i>table2array(table array)</i>	Dato in input una tabella o una porzione di essa, restituisce un array
<i>save(filename, vars ...)</i>	Salva in un file matlab denominato a piacimento le variabili fornite
<i>load(file)</i>	Carica i dati contenuti in un file matlab nel workspace
<i>find(condition)</i>	Restituisce gli indici degli elementi che soddisfano la condizione fornita. La ricerca viene eseguita sulla struttura in cui viene applicata la condizione.
<i>plot(...)</i>	Crea un grafico cartesiano 2D con i dati e caratteristiche forniti

<i>surf(...)</i>	Crea un grafico cartesiano 3D con i dati e caratteristiche forniti
<i>histogram(...)</i>	Crea un istogramma con i dati e caratteristiche forniti
<i>histfit(...)</i>	Ricava una distribuzione normale dai dati forniti e crea un istogramma
<i>xlabel e ylabel</i>	Se eseguite dopo una chiama a un grafico, daranno il nome alle relative assi
<i>hold</i>	Durante l'esecuzione di uno script interrompe il runtime fino alla chiusura del grafico attualmente aperto
<i>sqrt(x)</i>	Calcola la radice quadrata
<i>exp(x)</i>	Restituisce il valore e^x
<i>integral(f, down, up)</i>	Calcola l'integrale della funzione nell'intervallo fornito
<i>mean(arr)</i>	Calcola la media aritmetica degli elementi
<i>median(arr)</i>	Calcola la mediana degli elementi
<i>moda(arr)</i>	Calcola la moda degli elementi
<i>min(arr)</i>	Restituisce il valore minore fra gli elementi
<i>max(arr)</i>	Restituisce il valore maggiore fra gli elementi
<i>range(arr)</i>	Restituisce la differenza fra massimo e minimo degli elementi
<i>var(arr)</i>	Calcola la varianza degli elementi
<i>std(arr)</i>	Calcola la deviazione standard degli elementi
<i>cov(arr)</i>	Calcola la covarianza dei dati forniti
<i>corrcoef(X,Y)</i>	Ritorna una matrice di coefficienti di correlazione fra i dati forniti
<i>quantile(arr,n)</i>	Restituisce il quantile di ordine n dei dati forniti
<i>fix()</i>	Ci permette di ottenere la parte intera di un double
Distribuzioni	<p>Le principali distribuzioni che usiamo sono: <i>norm</i>, <i>chi2</i>, <i>t</i> Mathlab mette a disposizione numerose funzionalità che si ottengono concatenando il nome della distribuzione con la funzionalità:</p> <ul style="list-style-type: none"> • <i>normpdf</i>: ci permette di ottenere la relativa distribuzione • <i>normcdf</i>: ci permette di calcolare la funzione di ripartizione • <i>norminv</i>: ci permette di calcolare i valori z, c, t dei test
Regressione Lineare	<p>Mathlab offre la funzione <i>polyfit</i> che tramite il metodo dei minimi quadrati calcola polinomio di grado n che approssima i dati. (usiamo solo $n = 1$) Restituisce un vettore contenente B_1 e B_0 della formula $y = B_1x + B_0$</p>

Esercizi Compendium

Distribuzioni

Distribuzione Binomiale

Viene fornito un esperimento vero falso con la probabilità di successo o fallimento e un numero di prove o quantità.
Viene richiesto di calcolare le probabilità per un dato valore o un range di valori.

La risoluzione di questi si basa sull'utilizzo delle funzionalità di *binopdf* e *binocdf*.

Distribuzione Geometria

Viene fornito una probabilità, ovvero accade una cosa ogni tot. Viene richiesto quale sia la probabilità che l'evento accada per la prima volta dopo una determinata iterazione.

La risoluzione di questi si basa sull'utilizzo delle funzionalità di *geo_pmf*.

Distribuzione di Poisson

Viene fornito un intervallo in cui si verifica un determinato numero di volte un evento. Viene richiesto di calcolare la probabilità che l'evento accada un certo numero *w* di volte in un altro intervallo *c*.

1. Si calcola la frequenza dell'evento: $f = \frac{n_{evento}}{intervallo}$;
 - a. Se l'intervallo è compatibile con quello originale si può eseguire il rapporto sopracitato, nel caso non fosse compatibile bisogna portarlo nei termini minimi e usare quello come frequenza;
2. Si esegue il calcolo della probabilità è pari a *poispmf(w,f*c)*;

Stima dei parametri

Viene fornite una serie di dati come una quantità, una deviazione o una varianza e viene richiesto di trovare intervalli di confidenza, limiti o verosomiglianza.

Individuare intervallo (limiti superiori e inferiori)

- Controllare se si è in possesso di deviazione σ o varianza σ^2 (*varianza = deviazione²*);
 - Se non si ha una varianza ma una lista di campioni si calcola la varianza campionaria secondo formulario;
 - Se si hanno più popolazioni passa al punto successivo;
- Calcolare il valore $\bar{\sigma} = \frac{\sigma}{\sqrt{n}}$ oppure $\bar{\sigma} = \sqrt{\frac{\sigma^2}{n}}$;
 - Se si hanno più popolazioni $\bar{\sigma} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$;
- Un intervallo è individuato da un limite superiore e un limite inferiore:
 - Il limite inferiore è dato da $d_{inv}(1 - intervallo_{fornito}, media, \bar{\sigma})$
 - Il limite superiore è dato da $1 - d_{inv}(1 - (1 - intervallo_{fornito}), media, \bar{\sigma})$

Regressione Lineare

Data una collezione di dati, potrebbe essere chiesto di la retta di regressione, il coefficiente di correlazione, detta di Pearson e infine stimare/stabilire un valore della variabile di risposta.

La variabile di risposta è quella che deve essere predetta (*y*), mentre la variabile spiegata *x* è quella fornita.

Il procedimento è quanto segue:

1. Caricare gli dati negli appositi vettori *X* e *Y*, facendo attenzione a quale sia la variabile di risposta;

- Calcolare la media dei vettori e la loro lunghezza;
- Calcolare i due valori B e definire la retta $y = B_0x + B_1$;
- SE RICHIESTO: calcolare le varianze di x, y , la covarianza e i coefficienti di correlazione e determinazione;
- Stimare il valore richiesto inserendo il corrispondente valore spiegato nella formula della retta;

Commento sui dati:

- Covarianza: indica se i dati sono hanno simile varianza, 0 indica che sono incorrelati, un valore positivo o negativo indica rispettivamente che sono positivamente correlati o negativamente correlati, ovvero all'aumentare o diminuire di uno l'altro aumento o diminuisce di conseguenza. Più grande è il valore assoluto della covarianza, più è forte il loro legame.
- Il coefficiente di correlazione indica che la correlazione è negativa o positiva e quanto i dati siano correlati, visualmente se i punti seguono o meno la retta;
- I coefficienti di determinazione indicano la bontà del modello di regressione, se è pari a 1 non introduce dispersione nei dati, se è pari a 0 non è un buon modello;

```
clear
clc
t=readtable("lavatrici.txt", "VariableNamingRule", "preserve");
x=table2array(t(:,1)); %Variabile Spiegata
y=table2array(t(:,2)); %Variabile Risposta
n=length(x); %Lunghezza vettori
xbar=mean(x); %Media
ybar=mean(y); %Media
scatter(x,y); %Grafico per i punti
p=polyfit(x,y,1); %Vettore dei coefficienti B
m=p(1); %B0
q=p(2); %B1
line(x,(m*x)+q,"Color","blue"); %Retta di regressione
sxx=0.0;
syy=0.0;
sxy=0.0;
for i=1:n
    sxx=sxx+(x(i)-xbar)^2;
    syy=syy+(y(i)-ybar)^2;
    sxy=sxy+((x(i)-xbar)*(y(i)-ybar));
end
ylabel("Spiegata") %Nome asse X
xlabel("Risposta") %Nome asse Y
sxx=sxx/n; %Varianza di x
syy=syy/n; %Varianza di y
sxy=sxy/n; %Covarianza di x e y
ssres=(sxx*syy-sxy^2)/sxx; %Varianza residua
pxy=sxy/sqrt(sxx*syy); %Coef. di correlazione
r2=pxy^2; %Coefficiente di determinazione
r2corr=1-((ssres/(n-2))/(syy/(n-1))); %Coeff. det. corretto
stima=m*40+q; %Stima del valore
```