DEFINIZIONI INIZIALI

Calcolo della probabilità: logicamente autosufficiente in quanto si basa su poche leggi
Statistica: raccolta di dati e fatti in uno stato
la statistica deve inferire con i dati
- Descrittiva: illustra e sintetizza i dati
- Inferenziale: inferisce conclusioni a partire da dati numerici

STATISTICA DESCRITTIVA:

Popolazione: insieme di tutti i possibili individui (unità, oggetti)
Campione: sottoinsieme della popolazione preso in considerazione
└ deve essere rappresentativo ossia informativo
↳ deve essere scelto nel modo più casuale possibile
Variabile: caratteristica che varia da individuo ad individuo
- quantitativa {continue, discrete}
- qualitativa {ordinali, nominali}

Classe: se le variabili sono continue o possono assumere troppi valori, si dividono in classi

Frequenza assoluta: la frequenza assoluta di un valore x è il numero di individui per cui la variabile assume tale valore
$f_a(x) = |\{i : x_i = x\}|$

Frequenza relativa: la f_a, rapportata al numero totale di individui
$f_r(x) = \frac{f_a(x)}{n}$

Funzione cumulativa empirica: associa alla frazione x la frazione di unità nel campione che sono minori o uguali a x
$F(x) = \frac{|\{i : x_i \leq x\}|}{n} = \boxed{\sum_{z \leq x} f_r(z)}$ → cumula i valori di frequenza relativa per ogni z minore o uguale a x

Indici statistici di sintesi:
- MISURE di TENDENZA CENTRALE
- MISURE di DISPERSIONE
- MISURE di FORMA

MISURE di TENDENZA CENTRALE:

1. **Media campionaria**: la media campionaria è un indice lineare: prese due costanti a, b e un nuovo insieme di dati costituito da $\{y_0, y_1, \dots, y_n\}$ con $y_i = ax_i + b$, la media campionaria del nuovo insieme è: $\bar{y} = a\bar{x} + b$
$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$
$\bar{x} = \sum_{v \in X} v \cdot f_r(v)$ — v: valori assunti

2. **Mediana campionaria**: si ordini l'insieme dei dati crescentemente, la mediana è il valore centrale (o la media tra i due valori centrali) della successione di dati ordinata

3. **Moda campionaria**: è, se esiste, l'unico valore con frequenza massima. Se non vi è un unico valore con frequenza massima, ciascuno di essi è detto valore modale

MISURE di DISPERSIONE: Descrivono quanto siano concentrati rispetto alla media i valori in $\{x_1, \dots x_n\}$

1. **Varianza campionaria**: $s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

non è un indice lineare: prese due costanti a,b e un nuovo insieme di dati $\{y_0, \dots, y_n\}$, la varianza campionaria è $s_y^2 = a^2 s^2$

Maggiore è la varianza, maggiore è la dispersione dei dati

1b. **Deviazione standard**:
$s_x = \sqrt{s_x^2}$ [unità di misura della variabile]

DISEGUAGLIANZA di CHEBYSHEV (versione campionaria)

Sia $\{x_1, \dots, x_n\}$ un insieme di dati con media $\bar{x}$ e deviazione standard campionaria $s > 0$

Si consideri $S_k = \{\underbrace{i : |x_i - \bar{x}| < k s_x}_{\text{di quante deviazioni standard sono lontano dalla media}}\}$ insieme con cardinalità $|S_k|$

$$\frac{|S_k|}{n} \geq 1 - \frac{n-1}{nk^2} > 1 - \frac{1}{k^2}$$

Fornisce un limite inferiore alla frazione di dati di un campione che si allontanano dalla loro media campionaria più di un multiplo della deviazione standard

⇒ STIMA per difetto del numero di elementi che distano dalla media almeno ks_x

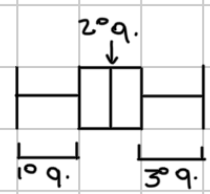
$|S_k|$

PERCENTILI

In un insieme di dati numerici, per ogni valore k esiste un dato che è contemporaneamente maggiore o uguale di almeno il k percento dei dati e minore o uguale di almeno il 100-k percento dei dati

Calcolo del percentile k-esimo:
- $p = k/100$
- $\lceil np \rceil$ dati sono $\leq$ al percentile k-esimo
- $\lceil n(1-p) \rceil$ dati sono $\geq$ al percentile k-esimo

Quartili:
- Il 25 esimo percentile si dice primo quartile
- Il 50 esimo percentile si dice secondo quartile
- Il 75 esimo percentile si dice terzo quartile



BOXPLOT

MISURE di FORMA

1. Distribuzione normale: circa il 68% dei dati cade in $\bar{x} \pm s$, il 95% in $\bar{x} \pm 2s$ ed il 99.7% in $\bar{x} \pm 3s$

miglioriamo la regola di Chebychev

2. Skeweness: assimmetria

3. Dati bivariati: se siamo interessati alla descrizione di coppie di variabili che hanno tra loro qualche associazione, ogni coppia è da considerarsi un'osservazione

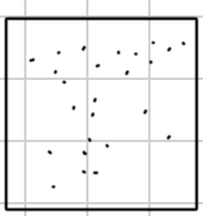


diagramma di dispersione

Covarianza campionaria: assume un valore positivo (negativo) se mediamente le due variabili subiscono oscillazioni concordi (discordi)

$$c = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

coefficiente di correlazione:

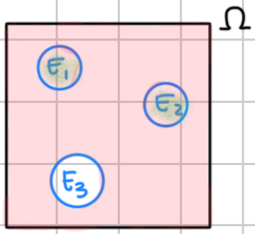
$$r = \frac{c}{s_x s_y}$$

PROBABILITA'

Esito di un esperimento con un certo grado di incertezza
└ procedura ripetibile che ha un certo numero di esiti possibili e ben definiti

Spazio degli eventi/esiti [Ω] : insieme di tutti i possibili esiti
un elemento o un sottoinsieme di Ω

RAPPRESENTAZIONE GRAFICA :



$E_4 = E_1 \cup E_2$
E_3^c
...

ESEMPIO 1: lancio di una moneta
$\Omega = \{Testa, Croce\}$ testa è un evento

ESEMPIO 2: lancio di due monete
$\Omega = \{TT, CC, TC, CT\}$
evento E_1: TT, {at least one T} = {TT, TC, CT}
evento E_2: {first is T} = {TT, TC}

LEGGI DI DE MORGAN:

1. $(E_1 \cup E_2)^c = E_1^c \cap E_2^c$
2. $(E_1 \cap E_2)^c = E_1^c \cup E_2^c$

FUNZIONE DI PROBABILITA' Una funzione che associa ad ogni evento in Ω un valore di probabilità

$\forall E \in \Omega \ \exists \mathbb{P}(E)$

$\mathbb{P}(E)$ frequenza relativa di E (con E ripetuto molte volte)

ASSIOMI:

1. Ogni valore di probabilità è compreso tra 0 e 1
$0 \leq \mathbb{P}(E) \leq 1$
2. $\mathbb{P}(\Omega) = 1$
3. Avendo un insieme finito o numerabile di eventi mutualmente esclusivi
└ non in comune alcun evento ← $E_1 \cap E_2 = \emptyset$

$$\mathbb{P}\Big(\bigcup_{i=0}^{n} E_i\Big) = \sum_{i=0}^{n} \mathbb{P}(E_i)$$
↑ unione degli eventi E_i

PROPRIETA':

1. $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$
2. $\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2) - \mathbb{P}(E_1 \cap E_2)$ principio di inclusione/esclusione

SPAZIO DI EVENTI EQUIPROBABILI

Spazio composto da n eventi equiprobabili $E_1, \dots, E_n$
└ $\mathbb{P}(E_1) = \mathbb{P}(E_2) = \dots = \mathbb{P}(E_n) = p$

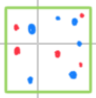
$\mathbb{P}(\Omega) = 1 = np$
→ $\mathbb{P}(E_J)$ $p = 1/n$ $\forall J = 1, \dots, n$ se E_J evento singolo
→ $\mathbb{P}(E_J)$ $p = |E|/n$ $\forall J = 1, \dots, n$ altrimenti

PRINCIPIO di ENUMERAZIONE:

Eseguo R esperimenti, ognuno con n_i esiti possibili

↳ Esiti complementari: $\prod_{i=0}^{R} n_i$ ← esiti possibili

Esempio 1: Da un urna estraggo due palline $\mathbb{P} = \{\bullet, \bullet\}$,



nell'urna ce ne sono 6 blu e 5 rosse

Possibili esiti 1° estrazione: 11
Possibili esiti 2° estrazione: 10 ⇒ 110 possibili esiti

In quanti dei 110 possibili esiti ho una pallina blu e una rossa?

1. 1° estrazione: blu
   2° estrazione: rosso → $6 \times 5 = 30$
2. 1° estrazione: rossa
   2° estrazione: blu → $5 \times 6 = 30$

→ la probabilità è 60 $\mathbb{P}\{\bullet, \bullet\} = \frac{60}{110} = 0{,}54$

PRINCIPIO di ENUMERAZIONE GENERALIZZATO

Eseguo R esperimenti, il primo esperimento ammette n_1 esiti, per ogni esito del primo il secondo esperimento ne ammette n_2

...

Per ogni esito del R−1 esimo esperimento mi restano n_R esiti possibili per l'R esimo esperimento

il numero di esiti totali dell'esperimento sarà $\prod_{i=1}^{R} n_i$

{a, b, c} 3 oggetti da ordinare

abc bac cab
acb bca cba

ci sono 3 possibili posizioni per la prima lettera → $n_1 = 3$
2 possibili posizioni per la 2ª lettera → $n_2 = 2$
la terza ha posizione obbligata → $n_3 = 1$

esiti totali: $3 \times 2 \times 1 = 6$

↳ PERMUTAZIONE di n oggetti → $n!$, K-PERMUTAZIONE: $\frac{n!}{(n-k)!}$

gruppi ordinati

ESERCIZIO: 59 persone su classroom, qual è la probabilità che non ve ne siano 2 che compiono gli anni lo stesso giorno

$\mathbb{P}$ {tutti compleanni diversi} = $\mathbb{P}(E)$?

eventi possibili $S_1 = 365$

eventi possibili $S_2 = 364$

...

eventi possibili complementari: $365! - 306!$

eventi totali: 365^{59}

$\mathbb{P}(E) = \frac{365! - 306!}{365^{59}}$

gruppi NON ordinati

E_2 {almeno 2 persone compiono gli anni lo stesso giorno}

$\mathbb{P}(E_2) = 1 - \frac{365! - 306!}{365^{59}}$

PRINCIPIO di ENUMERAZIONE e COMBINAZIONE

n oggetti, voglio formare gruppi di R oggetti

└ combinazione di R oggetti a partire da n

E {numero gruppi di 3 lettere formabili con gli elementi}

elementi: {a, b, c, d, e} 5 scelte per la prima delle 3 lettere, 4 per la 2ª, 3 per l'ultima

$5 \times 4 \times 3$ modi per scegliere 3 lettere MA tenendo conto dell'ordine

il numero di permutazioni di ogni tripletta è $3!$

$E = 6!/3!$ ← non ordinate

$$\frac{[N \times (N-1) \times \cdots \times N-(R-1)]}{R!} = \frac{N!}{R!(N-R)!} = \binom{N}{R}$$

← COEFFICIENTE BINOMIALE
quante combinazioni diverse di R oggetti posso ottenere da N oggetti

PROBABILITÀ CONDIZIONATA:

Come cambia la probabilità che si verifichi un evento A se ottengo ulteriori informazioni

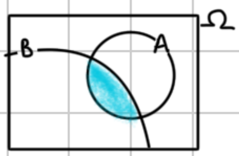
ESEMPIO 1:
lancio una moneta 3 volte
$\mathbb{P}\{HHH\}$
Per il principio di enumerazione se ci sono k esperimenti, ognuno con n esiti possibili gli esiti totali sono: $\prod n_i$
$\hookrightarrow 2 \cdot 2 \cdot 2 = 8$ esiti totali
Poiché il numero di eventi favorevoli è 1
$\mathbb{P}\{HHH\} = \frac{1}{8}$
MA se H è il risulato del primo lancio è H, $\mathbb{P}\{HHH\} = \frac{1}{4}$

ESERCIZIO: Formare un gruppo di 5 studenti in una classe con 6 uomini e 9 donne
casi totali: $\binom{15}{5}$

Gruppo con 2 donne e 3 uomini
Raggruppamento uomini: $\binom{6}{2}$
Raggruppamento donne: $\binom{9}{3}$
Combinazione totale: $\binom{6}{2} \times \binom{9}{3}$
$\hookrightarrow$ casi favorevoli

Probabilità: $\dfrac{\binom{6}{2} \times \binom{9}{3}}{\binom{15}{5}}$

A: evento di cui vogliamo calcolare la probabilità
B: evento che sappiamo essersi verificato
$\mathbb{P}\{A \text{ quando so che si è verificato } B\} = \mathbb{P}\{A|B\}$



$$\mathbb{P}\{A|B\} = \boxed{\frac{\mathbb{P}\{A \cap B\}}{\mathbb{P}\{B\}}}$$

la probabilità che si verifichino entrambi gli eventi rispetto agli eventi B

Es:
Carta da mazzo francese (52 c)
$E_1 = \{$la prima carta è cuori$\}$
$E_2 = \{$la seconda carta è cuori$\}$
13 carte cuori
$\mathbb{P}\{E_2|E_1\}$
metodo 1) eventi favorevoli: 12
eventi totali: 51
$\mathbb{P}\{E_2|E_1\} = \frac{12}{51}$
metodo 2) $\mathbb{P}\{E_2|E_1\} = \dfrac{\mathbb{P}\{E_2 \cap E_1\}}{\mathbb{P}(E_1)}$
$\mathbb{P}\{E_2 \cap E_1\} = \mathbb{P}\{$pesco 2 carte ed entrambe sono cuori$\}$
$= \frac{13}{52} \cdot \frac{12}{51} = \frac{3}{51}$ ← principio di enumerazione
$\mathbb{P}\{E_1\} = \frac{13}{52}$
$\rightarrow \mathbb{P}\{E_2|E_1\} = \frac{3}{51} \cdot \frac{52}{13} = \frac{12}{51}$

Probabilità dell'intersezione tra 2 eventi dipendenti

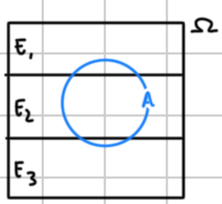
$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$$

se le informazioni fornite dall'evento B vanno ad alterare le probabilità del verificarsi del successivo evento A

probabilità dell'evento A condizionata all'evento B

ALTRIMENTI A e B indipendenti $P(A \cap B) = P(A)\,P(B)$

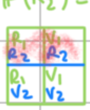
LEGGE di PROBABILITÀ TOTALE:



$\Omega = E_1 \cup E_2 \cup E_3 \quad E_i \cap E_j = \emptyset \quad \forall i,j = 1,\dots,3$
E_1, E_2 ed E_3 sono partizioni di Ω

$$P(A) = \mathbb{P}(A \cap E_1) + \mathbb{P}(A \cap E_2) + \mathbb{P}(A \cap E_3)$$
$$= \sum_{i=1}^{n} \mathbb{P}(A|E_i)\,\mathbb{P}(E_i)$$

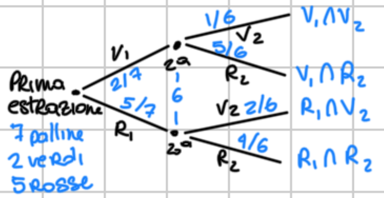
ESEMPIO: 5 palline rosse e 2 verdi, pesco 2 palline
$P(R_2) = \{$la seconda pallina che pesco è rossa$\}$



$\mathbb{P}(R_2) = \mathbb{P}(R_2|R_1)\mathbb{P}(R_1) + \mathbb{P}(R_2|V_1)\mathbb{P}(V_1)$
$\mathbb{P}(R_2|R_1) = \frac{4}{6} \quad \mathbb{P}(R_1) = \frac{5}{7} \quad \mathbb{P}(R_2|V_1) = \frac{5}{6} \quad \mathbb{P}(V_1) = \frac{2}{7}$
$\mathbb{P}(R_2) = \frac{4}{6}\,\frac{5}{7} + \frac{5}{6}\,\frac{2}{7} = \frac{5}{7}$

RAPPRESENTAZIONE TRAMITE ALBERI:

Un nodo diverso per ogni esperimento



$$\Rightarrow \mathbb{P}(R_2) = \mathbb{P}(R_2|R_1)\mathbb{P}(R_1) + \mathbb{P}(R_2|V_1)\mathbb{P}(V_1) = \frac{5}{7}\cdot\frac{4}{6} + \frac{2}{7}\cdot\frac{5}{6} = \frac{5}{7}$$

INDIPENDENZA:

Dati due eventi A e B

A e B si dicono indipendenti se conoscere l'esito di uno non cambia la probabilità dell'altro $\Rightarrow \mathbb{P}(A|B) = \mathbb{P}(A)$

$\overset{\text{def}}{\Rightarrow} \mathbb{P}(A\cap B) = \mathbb{P}(A|B)\,\mathbb{P}(B)$

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

ESEMPIO: lancio di una moneta

H_1: primo lancio H_2: secondo lancio

H_1 e H_2 sono indipendenti

$\Rightarrow \mathbb{P}(H_1 \cap H_2) = \mathbb{P}(H_1)\mathbb{P}(H_2)$

⚠ TEOREMA di BAYES:

Dati due eventi A e B

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B)\,\mathbb{P}(B)}{\mathbb{P}(A)} \xleftarrow{\text{in quanto}} \begin{array}{c}\mathbb{P}(B\cap A) = \mathbb{P}(A\cap B)\\ \mathbb{P}(B|A)\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B)\end{array}$$

ESEMPIO: Base Rate fallacy

Test per verificare se ho una malattia con incidenza (base Rate) = 0.005

con accuratezza alta: 5% falsi positivi

10% falsi negativi

E = {test risulta negativo} avvenuto qual è la probabilità che io abbia la malattia?

M = {ho la malattia} S = {non ho la malattia}

$\mathbb{P}(M) = 0.005$ $\mathbb{P}(S) = 1 - \mathbb{P}(M) = 0.995$

T_n = {test negativo} T_p = {test positivo}

$\Rightarrow$ voglio sapere $\mathbb{P}(M|T_n)$ $\leftarrow$ P di essere malato se il test negativo

$\mathbb{P}\{T_p|S\}$ = {probabilità di un falso positivo} $\leftarrow$ P di avere un test positivo se sono sano

5% falsi positivi $\rightarrow \mathbb{P}(T_p|S) = 0.05$

10% falsi negativi $\rightarrow \mathbb{P}(T_n|M) = 0.1$

$\mathbb{P}(T_n|S)$ è complementare di $\mathbb{P}(T_p|S)$ ossia $\mathbb{P}(T_n|S) = 1 - \mathbb{P}(T_p|S)$

$\mathbb{P}(T_n|S) = 0.95$

$\mathbb{P}(T_p|M) = 0.9$

$\rightarrow \mathbb{P}(M|T_n) \overset{\text{Bayes}}{=} \dfrac{\mathbb{P}(T_n|M)\,\mathbb{P}(M)}{\mathbb{P}(T_n)}$

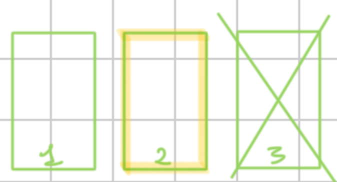
$\mathbb{P}(T_n) = \mathbb{P}(T_n\cap S) + \mathbb{P}(T_n\cap M) = \mathbb{P}(T_n|S)\mathbb{P}(S) + \mathbb{P}(T_n|M)\mathbb{P}(M) = 0.95\cdot 0.995 + 0.1\cdot 0.005 = 0.9457$

$\mathbb{P}(M|T_n) = \dfrac{0.1\cdot 0.005}{0.9457} = 0.0005$ $\longleftarrow$ RICONTROLLARE, LEZIONE 8

ESEMPIO: Gioco a premi

3 porte, 2 con dietro una capra e una una macchina

$\mathbb{P}(A_i) = \mathbb{P}$(c'è un auto dietro alla porta i) $= 1/3$



Scegli la porta 2, il conduttore allora apre la porta 3, mostra che dietro c'è una capra e ti chiede se vuoi cambiare

C_3: il conduttore ha aperto la porta 3

$\mathbb{P}(A_1|C_3) \lesseqgtr \mathbb{P}(A_2|C_3)$?

$\mathbb{P}(A_1|C_3) = 1 - \mathbb{P}(A_2|C_3)$

$\mathbb{P}(C_3|A_1)$ = {probabilità che il conduttore apra la porta 3 dato che l'auto è dietro A_1}

$\mathbb{P}(C_3) = 1/2$ $\leftarrow$ non può aprire la porta scelta dal concorrente

$\mathbb{P}(C_3|A_2) = 1/2$

$\mathbb{P}(A_2|C_3) = \dfrac{\mathbb{P}(C_3|A_2)\,\mathbb{P}(A_2)}{\mathbb{P}(C_3)} = \dfrac{1/2 \cdot 1/3}{1/2} = 1/3 \longrightarrow \mathbb{P}(A_1|C_3) = 1 - \mathbb{P}(A_2|C_3) = 2/3$

VARIABILI ALEATORIE DISCRETE:

- Ω spazio degli eventi
  $\Omega = \{w_1, \dots, w_n\}$ ← spazio finito o numerabile
- esito w_i, $\mathbb{P}\{w_i\}$
- $E \subset \Omega$ $\mathbb{P}\{E\}$, se $w_i \cap w_j = \emptyset$ (disgiunti) $\forall w_i, w_j \in E \rightarrow \mathbb{P}\{E\} = \sum_{w \in E} \mathbb{P}\{w\}$

ESEMPIO: lancio di due dadi
lanciamo più volte due dadi e registriamo gli esiti come coppia (i, j) dove
i: esito primo dado
j: esito secondo dado
$\Omega = \{(1,1), (1,2), \dots, (6,1), \dots, (6,6)\} = \{(i,j) \mid i,j = 1, \dots, 6\}$

$\mathbb{P}\{(3,2)\} = 1/36$

500 guadagno se lanciando 2 dadi ottengo 7 come somma dei due esiti, altrimenti ne perdo 100.
Ricompensa X $X(i,j) = \begin{cases} 500 & \text{if } i+j=7 \\ -100 & \text{otherwise} \end{cases}$

Una variabile aleatoria discreta è una funzione
$X: \Omega \rightarrow \mathbb{R}$ $X(w) = x$ x è realizzazione di X
e può assumere un numero discreto di valori e si chiama aleatoria perché il suo valore dipende dall'esito di un esperimento (casuale) con delle probabilità assegnate agli esiti

$x = 500$ $w = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$
Abbiamo 6 possibili esiti in omega per cui $x = 500$
$\mathbb{P}(X = 500) = 6/36 = 1/6$
L lettera maiuscola perché l'evento non si è ancora realizzato
$\mathbb{P}(X = 127) = 0$

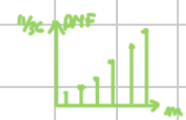
⚠ PROBABILITY MASS FUNCTION (PMF)

$p_X(x) = f_X(x) = \mathbb{P}(X = x)$ $f_X(x): \underset{x}{\mathbb{R}} \rightarrow \underset{p_X(x)}{\mathbb{R}}$

PROPRIETÀ della PMF:
1. $0 \leq p_X(x) \leq 1$
2. x è un valore che X non assume mai
   $p_X(x) = 0 = f_X(x)$

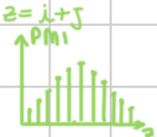
ESEMPIO: lancio di due dadi
$M(i,j) = \max(i,j)$ $M(3,5) = 5$

m:	1	2	3	4	5	6
$p_M(m)$	1/36	3/36	5/36	7/36	9/36	11/36
	(1,1)	(1,2) (2,2) (2,1)	(1,3) (2,3) (3,3) (3,2) (3,1)			



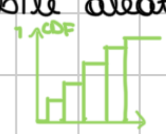
ESEMPIO 2: lancio di due dadi $z = i + j$

z	2	3	4	5	6	7	8	9	10	11	12
	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36



⚠ CUMULATIVE DISTRIBUTION FUNCTION

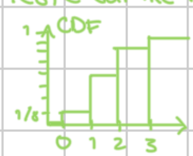
~ funzione cumulativa empirica
la CDF di una variabile aleatoria discreta X è:
$F_X(x) = \mathbb{P}\{X \leq x\}$



PROPRIETÀ:
1. $0 \leq F_X(x) \leq 1$
2. se $a \leq b$ $F_X(a) \leq F_X(b)$ → funzione non decrescente
3. $\lim_{x \to -\infty} F_X(x) = 0$; $\lim_{x \to \infty} F_X(x) = 1$

ESEMPIO: lancio di tre monete
X descrive il numero di teste sui tre lanci

x	0 (ccc)	1	2	3 (HHH)
$f_X(x)$	1/8	3/8	3/8	1/8
CDF	1/8	4/8	7/8	1



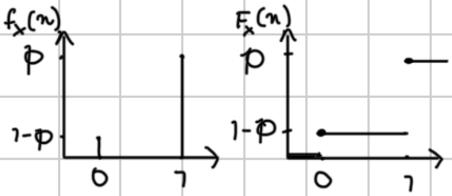
$F_X(y) = \sum_{\substack{y \leq x_i \\ x \in \mathcal{X}}} f_X(y)$ dove $\mathcal{X}$ è l'insieme di tutti i valori x per cui $p_X(x) \neq 0$
$\mathcal{X} = \{x_1, x_2, \dots, x_n\}$
$\mathcal{X}$ è detto supporto di $f_X(x)$

- PMF e CDF caratterizzano la variabile aleatoria

VARIABILE di BERNOULLI

$X \quad f_X(x) = p_X(x) \quad p$: parametro

Una singola realizzazione di un esperimento che ha due possibili esiti
es. lancio di una moneta



- variabile binaria $\begin{cases} 0 & \text{fallimento} \\ 1 & \text{successo} \end{cases}$ croce / testa
- $\mathbb{P}(x=1) = p$
- $\mathbb{P}(x=0) = 1-p$
- $f_X(x;p) = \begin{cases} p & x=1 \\ 1-p & x=0 \end{cases}$; $F_X(x) = \begin{cases} 1-p & x=0 \\ 1 & x=1 \end{cases}$

$p = 0.5$ ← moneta classica equa (in questo caso)

$X \sim \text{Bernoulli}(p)$ ← X è descritta da una distribuzione di Bernoulli con parametro p

X: moneta equa
$X \sim \text{Bernoulli}(0.5)$

ESEMPIO: Referendum PRO/CONTRO – ref costituzionale
$p = 0.7$ ← lo 0.7 degli elettori è a favore
il voto del singolo elettore: $X \sim \text{Bernoulli}(0.7)$

DISTRIBUZIONE BINOMIALE

Modella il numero di successi in n esperimenti indipendenti ogni esperimento segue Bernoulli (p) → gli n esperimenti devono quindi avere ognuno due possibili esiti
↓
il singolo esperimento avrà probabilità o p o $1-p$

$X \sim \text{Bin}(n,p)$
quando $n=1$ si dice che la binomiale degenera in una Bernoulli
$Z \sim \text{Bin}(1,p) = Z \sim \text{Ber}(p)$

$$\boxed{f_Y(r) = \binom{n}{r} p^r (1-p)^{n-r}}$$

$$F_Y(y) = \sum_{r \le y} f_Y(r) = \sum_{r \le y} \binom{n}{r} p^r (1-p)^{n-r}$$

$\mathbb{E}\{X\} = np \qquad \mathbb{V}\{X\} = np(1-p)$

ESEMPIO: $Y \sim \text{Bin}(5, p)$

y: 0 1 2 3 4 5 ← l'esperimento può essere un successo 0,1,2,3,4,5 volte
$f_Y(y)$: non lo so ancora perché il valore di p non è specificato → potrebbero non essere equiprobabili

Singolo esperimento $\mathbb{P}(\text{CROCE}) = 1-p$
$\mathbb{P}(\text{CROCE}, \text{CROCE}) = \mathbb{P}(\text{CROCE}_1 \cap \text{CROCE}_2)$
poiché CROCE_1 e CROCE_2 sono indipendenti → $\mathbb{P}(\text{CROCE}_1 \cap \text{CROCE}_2) = \mathbb{P}(\text{CROCE}_1)\mathbb{P}(\text{CROCE}_2)$
$\mathbb{P}(\text{CROCE}_1, \text{CROCE}_2) = (1-p)(1-p) = (1-p)^2$
↓
$\mathbb{P}(c,c,c,c,c) = (1-p)^5 = \mathbb{P}(y=0)$

Probabilità di 2 teste e 3 croci?
$\mathbb{P}(\text{2 teste e 3 croci}) = f_Y(2)$

Se gli eventi fossero equiprobabili ($p=0.5$), mi basterebbe fare:
$\mathbb{P}(\text{TOT}) = \dfrac{\text{eventi favorevoli}}{\text{eventi possibili}}$

ma in generale $\mathbb{P}(\text{TOT}) = \mathbb{P}$(intersezione tra gli E favorevoli)
poiché $|E_f| = \binom{3}{2} = 3$ e E_f indipende
$\mathbb{P}(2T) = \mathbb{P}(E_1 \cap E_2 \cap E_3) = \mathbb{P}(E_1)\mathbb{P}(E_2)\mathbb{P}(E_3)$
$\mathbb{P}(2T) = \mathbb{P}(HTT)\mathbb{P}(THT)\mathbb{P}(TTH) = p(1-p)^2 \cdot (1-p)p(1-p) \cdot (1-p)^2 p = 3p(1-p)^2$

OPERAZIONI ALGEBRICHE:

$X_j \sim Ber(p)$

$Y = \sum_1^n {}_i \, x_i$ sto sommando delle variabili ognuna delle quali può valere 0 o 1

$x_i = \begin{cases} 0 \\ 1 \end{cases} \rightarrow 0 \le Y \le 5$

$\sim Bin(n, p)$

$Z \sim Bin(m, p)$
$U \sim Bin(\ell, p)$
$Z + U = Bin(m + \ell, p)$

⚠ VALORE ATTESO – MEDIA

Il valore atteso di una variabile aleatoria X, se esiste, è:

$\mathbb{E}\{X\} = \sum_1^{|X|} {}_i \, x_i \, f_X(x_i) = \mu$ ← media pesata sulla probabilità dei singoli valori che X può assumere.

↳ ASPETTAZIONE $\quad = np$

PROPRIETÀ:

① $\mathbb{E}\{Z\} = \mathbb{E}\{X+Y\} = \mathbb{E}\{X\} + \mathbb{E}\{Y\}$ con $Z = X+Y$

② $\mathbb{E}\{aX+b\} = \mathbb{E}\{W\} = a\mathbb{E}\{X\} + b$ con $W = aX+Y$

③ $\mathbb{E}\{XY\} = \mathbb{E}\{X\}\,\mathbb{E}\{Y\}$

- non per forza appartiene al supporto
- è una media (pesata)
- è una statistica che riassume la distribuzione della nostra variabile con un solo numero

DIMOSTRAZIONE $\mathbb{E}\{X\} = np$

Z_i è una $Z_i \sim Ber(p)$ $i = 1, \dots, n$

$X \sim Bin(n, p)$ $\quad X = \sum_1^n {}_i \, Z_i \xrightarrow{\text{PROP. 2}} \mathbb{E}\{X\} = \sum_1^n {}_i \, \mathbb{E}\{Z_i\}$

Poiché Z_i è Bernulliana i valori possibili di $z_i = \{0, 1\} \longrightarrow \mathbb{E}\{Z_i\} = 0(1-p) + 1(p) = p$

$\mathbb{E}\{X\} = \sum_1^n {}_i \, \mathbb{E}\{Z_i\} = \sum_1^n {}_i \, p = np$

CAMBIO di VARIABILE:

$X \quad \mathcal{X} = \{x_1, x_2, \dots, x_n\}$ definiamo $Y = h(X)$ funzione di variabile aleatoria

$\mathbb{E}\{h(X)\} = \sum_1^n {}_i \, h(x_i) f_X(x_i)$

MA $\mathbb{E}\{h(X)\} \neq h(\mathbb{E}\{X\})$

ESEMPIO: X: {Risultato del lancio di un dado}
$Y = h(X) = X^2$
$\mathcal{X} = \{1, 2, 3, 4, 5, 6\}$ $\quad \mathcal{Y} = \{1, 4, 9, 16, 25, 36\}$
$f_X(x) = \{1/6, 1/6, 1/6, 1/6, 1/6, 1/6\}$
$\mathbb{E}(h(X)) = 1 \cdot 1/6 + 4 \cdot 1/6 + 9 \cdot 1/6 \cdots = 15{,}6$
$h(\mathbb{E}(X)) = 3{,}5^2 = 12{,}5$

COPPIE di VARIABILI ALEATORIE

Osservare in maniera congiunta due variabili aleatorie, per scoprire ad esempio il loro effetto congiunto su un dato fenomeno

X, Y **CDF CONGIUNTA** $F_{X,Y}(x, y) = \mathbb{P}\{\{X \le x\} \cap \{Y \le y\}\}$

↳ probabilità dell'intersezione dei due eventi

PMF CONGIUNTA $f_{X,Y}\{x, y\} = \mathbb{P}\{\{X = x\} \cap \{Y = y\}\}$

MARGINALIZZAZIONE:
$f_X(x) = \mathbb{P}\{X = x\} = \sum_1^n {}_i \, f_{X,Y}(x, y_i)$

Mi permette di tornare alle PMF e CDF singole da quelle congiunte

ESEMPIO: lancio di 2 dadi
uno blu: X
uno rosa: Y
$\mathbb{P}\{X = 3\} = \sum \mathbb{P}\{\{X=3\} \cap \{Y = y_i\}\}$

$f_{X,Y}$ (x \ y)	1	2	3	4	5	6
1	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	⋮	⋮			
3	1/36	⋮	⋮			
4	1/36					
5	1/36					
6	1/36					

$X = 1 \cup Y = 1 \cap$
$X = 1 \cup Y = 2 \cap$
$X = 1 \cup Y = 3 \cap$
⋮

$\frac{1}{6} \;\llcorner\; \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36}$

VARIABILI ALEATORIE INDIPENDENTI

Prendo un insieme A contenuto nel supporto di X ed un insieme B contenuto nel supporto di Y:

A e B indipendenti se:
- $\mathbb{P}(A|B) = A$
- $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$

$A \in S_X,\ B \in S_Y$ con X, Y variabili indipendenti

$\mathbb{P}\{\{X \in A\} \cap \{Y \in B\}\} = \mathbb{P}\{X \in A\}\,\mathbb{P}\{Y \in B\}$

$F_{X,Y}(x,y) = F(x)F(y)$

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

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

FUNZIONE di PROBABILITÀ CONDIZIONATA:

X, Y variabili aleatorie dipendenti:

$\mathbb{P}\{X \cap Y\} = \mathbb{P}\{X|Y\}\,\mathbb{P}\{Y\}$

$f_{X|Y}(x|y) = \dfrac{f_{X,Y}(x,y)}{f_Y(y)} = \mathbb{P}\{X = x \mid Y = y\}$

se X, Y indipendenti:
$f_{X|Y}(x|y) = f_X(x)$

VARIANZA: X v.a. discreta $\mathbb{E}\{X\} = \mu$

$\mathrm{Var}\{X\} = \mathbb{E}\{(X-\mu)^2\} = \sigma^2$

PROPRIETÀ:

1. $\mathrm{Var}\{X+Y\} = \mathrm{Var}\{X\} + \mathrm{Var}\{Y\}$ se X Y indipendenti
2. $\mathrm{Var}\{aX+b\} = a^2\,\mathrm{Var}\{X\}$
3. $\mathrm{Var}\{X\} = \mathbb{E}\{X^2\} - \mathbb{E}^2\{X\}$

VARIABILE GEOMETRICA:

$X \sim G(p)$ quanti esperimenti indipendenti sono osservati prima di avere successo.

$\mathbb{P}\{X = k\} = (1-p)^{k-1}p = f_X(k)$ probabilità che alla k-esima prova ottengo il primo successo

$\mathbb{E}\{X\} = 1/p \quad \mathbb{V}\{X\} = \dfrac{1-p}{p^2}$

VARIABILE di POISSON

$X \sim \mathbb{P}(\lambda)$ In un determinato intervallo di tempo quanti eventi si verificano successivamente, sapendo che mediamente se ne verificano λ

$f_X(x) = \dfrac{\lambda^x}{x!}e^{-\lambda} \quad x \in \mathbb{N}$

$\mathbb{E}\{X\} = \lambda \qquad \mathbb{V}\{X\} = \lambda$

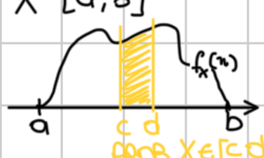
VARIABILI ALEATORIE CONTINUE

Assumono valori su un intervallo continuo $\Omega \to \mathbb{R}$
Una variabile aleatoria X è continua se esiste una funzione $f_X(x)$ tale che:

$$\mathbb{P}\{c < X \leq d\} = \int_c^d \underbrace{f_X(x)}_{\text{funzione di densità di probabilità PDF}} dx \qquad \forall c \leq d$$

ossia, la probabilità che X sia compreso tra c e d si ottiene integrando $f_X(x)$ tra gli estremi dell'intervallo

X [a,b]



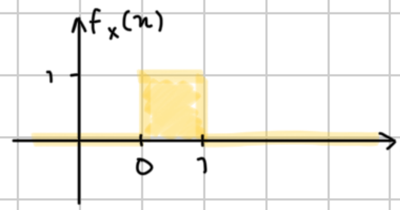
PMF (v.a discrete) $f_X(x) = \mathbb{P}\{X = x\}$

PDF (v.a continue) NON è una funzione di probabilità ma mi permette di calcolare $\mathbb{P}\{X \in [c, c+\delta]\}$

⚠ PROPRIETÀ:

① $f_X(x) \geq 0 \leftarrow$ la PDF può essere > 1, l'importante è che $0 \leq \int_{-\infty}^{+\infty} f_X(x)dx \leq 1$

② $\int_{-\infty}^{+\infty} f_X(x)\, dx = 1 \underset{\text{in quanto}}{=} \mathbb{P}\{-\infty < X \leq +\infty\}$



ESERCIZIO: v.a. continua $X\,[0,1]$ $f_X(x) = c\,x^2$ $c = ?$

Per la prima proprietà della PDF: $f_X(x) \geq 0 \to c \geq 0$

Per la seconda proprietà della PDF: $\int_{-\infty}^{+\infty} f_X(x)\, dx = 1$

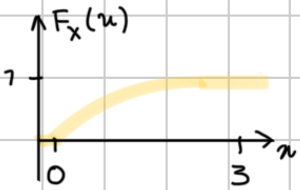
$\int_{-\infty}^{+\infty} f_X(x)\, dx = \int_0^1 c \cdot x^2\, dx = c \int_0^1 x^2 dx = c \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3} c$

$\to \frac{1}{3} c = 1 \to c = 3$

FUNZIONE DI DISTRIBUZIONE CONTINUA: CDF

$F_X(b) = \mathbb{P}\{X \leq b\} = \mathbb{P}\{-\infty \leq X \leq b\}$

$F_X(b) = \int_{-\infty}^{b} f_X(x)\, dx$



PROPRIETÀ:

① $0 \leq F_X(x) \leq 1$ perché la CDF è una funzione di probabilità

② $\lim_{x \to +\infty} F_X(x) = 1$ e $\lim_{x \to -\infty} F_X(x) = 0$

③ $\mathbb{P}\{a \leq X \leq b\} = F_X(b) - F_X(a)$

④ $\frac{dF_X(x)}{dx} = f_X(x)$

$\mathbb{E}\{X - Y\} = 0$
$\mathbb{V}\{X - Y\} = \frac{2}{12}$

ESERCIZIO: $f_X(x) = 3x^2$
$X\,[0,1]$ $\mathbb{P}(X \leq 1/2) = ?$
$\mathbb{P}(X < 0) = 0$
$\mathbb{P}(0 \leq X \leq 1/2) = \int_0^{1/2} 3x^2 dx = \frac{1}{8}$

VARIABILE ALEATORIA UNIFORME

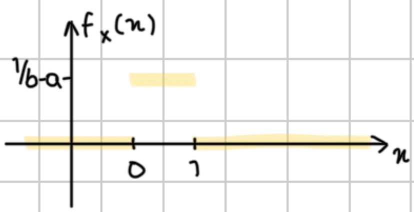
Una v.a X è detta uniforme (uniformemente distribuita) sull'intervallo [a,b] se la sua PDF

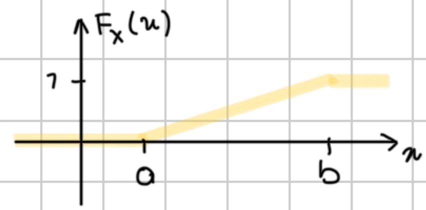
$$f_X(x) = \begin{cases} 1/(b-a) & \forall x \in [a,b] \\ 0 & \text{altrove} \end{cases}$$

$\leftarrow$ i valori nell'intervallo sono tutti equiprobabili

per indicare che X è una v.a. uniforme in [a,b]: $X \sim U[a,b]$

$$F_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases}$$





$\mathbb{E}\{X\} = \frac{a+b}{2}$ $\quad \mathbb{V}\{X\} = \frac{(b-a)^2}{12}$

ESERCIZIO: Autobus intervalli di 15 minuti a partire dalle 7
Una persona arriva tra le 7:00 e le 7:30, quale è la probabilità che aspetti meno di 5 minuti oppure più di 10?
X rappresenta il minuto a cui arriva il passeggero
$X \sim U[0,30]$

1. $\mathbb{P}\{$meno di 5 minuti$\}$: 2 autobus 7:15 e 7:30
$\mathbb{P}\{$meno di 5 m$\} = \mathbb{P}\{10 \le X \le 15\} + \mathbb{P}\{25 \le X \le 30\}$

$f_x(x) = \begin{cases} 1/30 & x \in [0,30] \\ 0 & x \notin [0,30] \end{cases} \longrightarrow \mathbb{P}\{10 < X < 15\} + \mathbb{P}\{25 < X < 30\} = \int_{10}^{15} 1/30 \, dx + \int_{25}^{30} 1/30 \, dx$

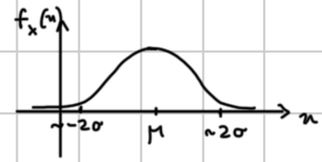
$\mathbb{P}\{$meno di 5 minuti$\} = 1/6 + 1/6 = 1/3$

2. $\mathbb{P}\{$più di 10$\} = \mathbb{P}\{0 \le X \le 5\} + \mathbb{P}\{15 \le X \le 20\} = \int_0^5 1/30 \, dx + \int_{15}^{20} 1/30 \, dx = 1/3$

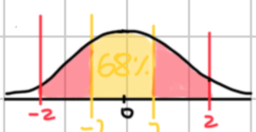
DISTRIBUZIONE GAUSSIANA (-NORMALE)

X v.a. continua è detta normale di parametri μ e σ $\quad X \sim N(\mu, \sigma^2)$
se

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



Gaussiana $\mu = 0$ $\sigma = 1 \longrightarrow$ Distribuzione normale standard
$Z \sim N(0,1)$ $\quad f_z(z) = \varphi(z) = \frac{1}{2\pi} e^{\frac{-z^2}{2}}$ $\quad F_z(z) = \Phi(z) \leftarrow$ si calcola con software (pnorm), per approssimazione oppure con una tavola



$\mathbb{P}\{-1 \le X \le 1\} = 0.68$
$\mathbb{P}\{-2 \le X \le 2\} = 0.95$
$\mathbb{P}\{-3 \le X \le 3\} = 0.99$

z può assumere qualsiasi valore tra $-\infty$ e $+\infty$, ma il 99% dei valori è compreso tra 3 e -3

$X \sim N(\mu, \sigma^2)$, applico ad X
$Z = h(X) = \frac{X - \mu}{\sigma}$ $\quad \longleftarrow$ operazione di standardizzazione
Z è ancora una variabile aleatoria continua gaussiana, ma con parametri differenti:
$$Z \sim N(0,1)$$
Poiché una v.a. normale standard è simmetrica $\varphi(-z) = \varphi(z)$

VARIABILE ALEATORIA ESPONENZIALE

$X \sim esp(\lambda)$ quando: $\quad \lambda \in [0,1]$

$f_x(x) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases}$ con $\lambda > 0$ $\longrightarrow F_x(a) = 1 - e^{-\lambda a}$ $\quad X \in [0, +\infty]$

Possiamo usare una v.a esponenziale per calcolare il tempo di attesa SENZA MEMORIA

$\mathbb{P}\{X > s+t \mid X > s\} = \mathbb{P}\{X > t\} = e^{-\lambda t}$ $\leftarrow$ indipendentemente da quanto io abbia aspettato, la $\mathbb{P}$ che io debba aspettare ancora non cambia

VALORE ATTESO

$X \in [a,b]$

$\mathbb{E}\{X\} = \int_a^b x f_x(x)\, dx$

Esponenziale: $\mathbb{E}\{X\} = \int_0^\infty x e^{-\lambda x} dx \overset{\dots}{=} \frac{1}{\lambda}$ $\quad \mathbb{E}\{X^2\} = \frac{2}{\lambda^2}$

Gaussiana normale standard: $\mathbb{E}\{X\} = \int_{-\infty}^{+\infty} x \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 0$

Normale: $\mathbb{E}\{X\} = \mu$

PROPRIETÀ:

1. $\mathbb{E}\{X+Y\} = \mathbb{E}\{X\} + \mathbb{E}\{Y\}$
2. $\mathbb{E}\{aX+b\} = a\mathbb{E}\{X\} + b$
3. $Y = h(X)$ $\quad \mathbb{E}\{Y\} = \int_a^b h(x) f_x(x)\, dx$
   $X \in [a,b]$

DIMOSTRAZIONE $\mathbb{E}\{X\} = \mu$ $\quad X \sim \mathcal{N}(\mu, \sigma^2)$

$Z \sim \mathcal{N}(0,1)$ $\quad Z = \frac{X-\mu}{\sigma} \longrightarrow X = Z\sigma + \mu$

$\mathbb{E}\{X\} = \mathbb{E}\{\sigma Z + \mu\} = \sigma\, \underbrace{\mathbb{E}\{Z\}}_{=0} + \mu = \mu$

VARIANZA:

$\mathbb{V}\{X\} = \mathbb{E}\{\underbrace{(X-\mu)^2}_{h(X)}\}$

PROPRIETÀ:

1. $\mathbb{V}\{X+Y\} = \mathbb{V}\{X\} + \mathbb{V}\{Y\}$ se X e Y indipendenti
2. $\mathbb{V}\{aX+b\} = a^2\mathbb{V}\{X\}$
3. $X \sim U(0,1)$ $\quad \mathbb{V}\{X\} = 1/12$

Teorema:

$\mathbb{V}\{X\} = \mathbb{E}\{X^2\} - \mathbb{E}\{X\}^2 = \mathbb{E}\{X^2\} - \mu^2$

V.a continue		$\mathbb{E}\{X\}$	$\mathbb{V}\{X\}$
Normale	$\mathcal{N}(\mu, \sigma^2)$	μ	σ^2
Standard	$\mathcal{N}(0,1)$	0	1
Uniforme	$U(a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Esponenziale	$\exp(\lambda)$	$1/\lambda$	$1/\lambda^2$

- statistica descrittiva: come descrivere un insieme molto grande di oggetti
popolazione $\curvearrowright$ campione
quantità misurabili
- STATISTICA INFERENZIALE
analizzare il campione
trarre delle conclusioni valide per la popolazione nel suo insieme

popolazione: modelliamo tramite variabili aleatorie alcune delle sue caratteristiche

Estraiamo un campione casuale che le variabili che stiamo osservando siano realizzazioni di v.a. indipendenti e identicamente distribuite (IID)
↓
con la stessa distribuzione di probabilità

INFERENZIALE: { la distribuzione del campione che stiamo osservando non è mai completamente nota

Inferire: estrapolare quali sono gli elementi non noti (es. parametri) osservando una sequenza abbastanza lunga
- inferenza parametrica: distribuzione nota, a meno di un insieme di parametri
- inferenza non parametrica: la distribuzione non è nota

DISEGUAGLIANZA di CHEBYCHEV

Y v.a. qualsiasi $a>0$

$\mathbb{P}\{|Y-\mathbb{E}\{Y\}|\geq a\}\leq \frac{1}{a^2}\mathbb{V}\{Y\}$ ← la v.a. si discosta dalla media sempre più all'aumentare della varianza

⚠ LEGGE dei GRANDI NUMERI (media)

Una statistica: variabile aleatoria che è funzione dei dati di un campione

$\overline{X}_N=\frac{1}{N}\sum_{1}^{N}X_i$ variabile aleatoria funzione del campione → la media campionaria è una statistica

$\mathbb{E}\{X_i\}=\mu \quad \forall i \qquad \mathbb{V}\{X_i\}=\sigma^2$

$\mathbb{E}\{X_N\}=\mathbb{E}\{\frac{1}{N}\sum_{1}^{N}{}_i X_i\}=\frac{1}{N}\sum_{1}^{N}{}_i\mathbb{E}\{X_i\}=\frac{1}{N}N\cdot\mu=\mu$

v.a. $X\sim BER(p)$

$X_i=\begin{cases}1 & p=p\\ 0 & p=0\end{cases} \rightarrow \mathbb{E}\{\overline{X}_N\}=p \qquad |\overline{X}_N-\mu|<\varepsilon$

la legge dei grandi numeri (o di Bernoulli) dice che la probabilità

$\mathbb{P}\{|\overline{X}_N-\mu|<\varepsilon\}$ cresce al crescere di N

Dimostrazione:

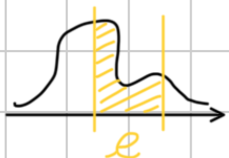
X_i variabili aleatorie, tutte con $\mu,\sigma^2 \quad \forall i=1,\dots,N$

$\mathbb{E}\{\overline{X}_N\}=\mu \quad \mathbb{V}\{\overline{X}_N\}=\frac{1}{N^{\cancel{2}}}\cancel{N}\sigma^2=\frac{\sigma^2}{N}$

$\mathbb{P}\{|\overline{X}_N-\mu|\geq\varepsilon\}\leq\frac{1}{\varepsilon^2}\frac{\sigma^2}{N} \rightarrow \mathbb{P}\{|\overline{X}_N-\mu|\geq\varepsilon\}\leq\frac{\sigma^2}{N\varepsilon^2}$

$\lim_{N\to\infty}\mathbb{P}\{|\overline{X}_N-\mu|\geq\varepsilon\}=\emptyset$

⚠ LEGGE dei GRANDI NUMERI (densità di probabilità)



$\phi = \mathbb{P}\{X \in \mathcal{e}\}$ genero $x_1, x_2, \dots, x_N$ con N grande

se $X_i \in \mathcal{e} \rightarrow Y_i = 1$, altrimenti $Y_i = 0$ Y_i è un indicatore

$\mathbb{E}\{Y_i\} = \phi$

$$\lim_{N \to \infty} \mathbb{P}\{|\overline{Y}_N - \phi| \geq \varepsilon\} = 0$$

Per calcolare con che probabilità una certa variabile aleatoria X_i cade in un intervallo $\mathcal{e}$ senza sapere la sua PDF, possiamo osservare molte volte questa variabile e utilizzare una v.a indicatorea sommando tutte le volte in cui $Y = 1$ e dividendo per N otteniamo una stima di ϕ

$\hookrightarrow \overline{Y}_N \simeq \phi = \int_{\mathcal{e}} f_X(x)\, dx$

$\mathcal{e} = [a - R, a + R]$

$\overline{Y}_N \simeq f_X(a)\, 2R \longrightarrow f_X(a) = \dfrac{\overline{Y}_N}{2R}$

⚠ TEOREMA del LIMITE CENTRALE

Dati $x_1, x_2, \dots, x_N$ se IID $\mathbb{E}\{X_i\} = \mu$ $\mathbb{V}\{X_i\} = \sigma^2$

$\overline{X}_N = \frac{1}{N}\sum_1^N i\, X_i$ $\mathbb{E}\{\overline{X}_N\} = \mu$ $\mathbb{V}\{\overline{X}_N\} = \frac{\sigma^2}{N}$

la media di un grande numero di v.a. IID è circa normale

standardizzo $\overline{X}_N$: $Z = \dfrac{\overline{X}_N - \mu}{\sigma / \sqrt{N}}$ $\mathbb{E}\{Z\} = 0$ $\mathbb{V}\{Z\} = 1$ $\lim_{N \to \infty} F_{Z_N}(z) = \Phi(z)$

Partendo da qualsiasi distribuzione (che non conosciamo) e sapendo che abbiamo realizzazioni IID, quando N è abbastanza grande possiamo approssimare la media campionaria con una v.a gausiana in μ, σ^2/N

$\overline{X}_N \sim \mathcal{N}(\mu, \sigma^2/N)$

STIMA PARAMETRICA:

$Ber(\phi)$ $Bin(n, \phi)$ $\exp(\lambda)$ $\mathcal{N}(\mu, \sigma^2)$

↓ ↓ ↓ ↓

parametrizzare la distribuzione ⟶ distribuzioni parametriche

Ho il campione — conosco la famiglia da cui deriva la sua distribuzione — Non conosco il parametro

↙ stimatori puntuali

↘ stimatori ad intervalli

Uno stimatore è una v.a., è una statistica che fornisce una stima del parametro

↳ è il risultato dello stimatore, è un valore deterministico

STIMATORE di MASSIMA VEROSIMIGLIANZA (MLE)

$\{x_i\}_1^n \quad \theta?$ qual è il valore puntuale di θ per cui è più probabile ottenere questo campione?

Lo stimatore MLE fornisce una stima $\hat{\theta}$ che massimizza la funzione di likelihood

$$\hat{\theta} = \arg\max_{\theta} \mathbb{P}\{x_1, x_2, \dots, x_n; \theta\}$$

$$= \arg\max_{\theta} \Lambda(\theta)$$

poiché $= \arg\max_{\theta} \prod_{i=1}^{n} \mathbb{P}\{x_i\}$ con x_i indipendenti

per il calcolo a volte è più semplice ricondursi al logaritmo naturale in quanto

$$\log\left(\prod_{i=1}^{n} \mathbb{P}_i\right) = \sum_{i=1}^{n} \log \mathbb{P}_i$$

log-likelihood:

$$\hat{\theta} = \arg\max_{\theta} \ln \Lambda(\theta)$$

ESEMPIO: 100 lanci della stessa moneta

numero di teste $X \sim Bin(n, p)$ (n = 100, p = ?)

$x_1 = 55$ teste

$\hat{p}$? MLE

qual è il valore che massimizza la probabilità di avere 55 teste?

$\max_{p} \mathbb{P}_p\{55\}$ ← questa probabilità dipende da p

poiché $X \sim Bin(100, p)$

$$\max_{p} \binom{100}{55} p^{55} (1-p)^{100-55} = \max_{p} f_X(x_1)$$

$$\max_{p} \mathcal{L}(p)$$

$\hat{p} = \arg\max \mathcal{L}(p)$ FUNZIONE di MASSIMA VEROSIMIGLIANZA

↳ per le v.a discrete la funzione di likelihood passa per la PMF congiunta

v.a. continue:

se $X \sim esp(\lambda)$ $\quad \{x_i\}_1^n$

$$\hat{\lambda} = \arg\max_{\lambda} \lambda^n e^{-\lambda \sum_1^n i\, x_i}$$

se $X \sim \mathcal{N}(\mu, \sigma^2)$ $\quad \{x_i\}_1^n$ conoscendo uno dei 2 parametri (σ^2)

$$\hat{\mu} = \arg\max_{\mu} \frac{1}{\sqrt{2\pi}\,\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

ESEMPIO: tempo di vita lampadine

$X \sim exp(\lambda)$

5 lampadine $\quad x_1 = 2$ mesi, $x_2 = 1$ mese, 20gg, $\dots$, x_5 $\quad \{x_i\}_1^5$

x_i indipendenti

$$f_X(x) = \lambda e^{-\lambda x} \rightarrow f_X(x_i) = \prod_1^5 \lambda e^{-\lambda x_i} = \lambda^5 e^{-\lambda \sum_1^5 x_i}$$

argmax

Binomiale: $\hat{p} = \frac{1}{n} \sum_1^n i\, x_i = \frac{k}{n}$ (k: numero successi $\sum x_i$) ← media campionaria su diverse realizzazioni dell'esperimento

Normale: $\hat{\mu} = \frac{1}{n} \sum_1^n i\, x_i$ ← media campionaria

$\hat{\sigma}^2 = \frac{1}{n-1} \sum_1^n i\,(x_i - \hat{\mu})^2$ ← varianza campionaria

Esponenziale: $\hat{\lambda} = \frac{1}{n} \sum_1^n i\, x_i$ ← media campionaria

CONCETTO MLE:

Passiamo dalla PDF in funzione di x alla PDF in funzione del parametro. Questa nuova funzione la chiamiamo di massima verosimiglianza e otteniamo il valore del parametro che massimizza la probabilità di ottenere esattamente quel campione

STIMATORE INTERVALLARE - INTERVALLI di CONFIDENZA

$\theta \, [\ell_\theta ; u_\theta] \quad \{x_i\}_1^n$

gli estremi dell'intervallo variano al variare del campione

↳ la randomicità del campione si riflette sull'intervallo di confidenza

(ℓ_θ, u_θ) $B\%$ confidenza

al variare dei campioni, se calcolassi questo intervallo, il $B\%$ dei campioni produrrebbero un intervallo che contiene il valore vero di θ

$\{x_i\}_1^n$ estratto da n v.a normali IID

$\hat{\mu} = \frac{1}{n} \sum_1^n {}_i x_i = \bar{X}_n$

standardizzazione $\frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$

$P\left\{-u < \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq u\right\} = B$ ← fisso la probabilità

$P\{-u < Y \leq u\} = B \rightarrow \int_{-u}^{u} \varphi(y)\, dy = B$ ← CALCOLARE con TABELLA (lezione 21 1:25:00)

$P\left\{\bar{X}_n - \frac{\sigma u}{\sqrt{n}} > \mu \geq \bar{X}_n + \frac{\sigma u}{\sqrt{n}}\right\}$

$\Phi(-u<x<u) = \Phi(u) - \Phi(-u)$

$\Phi(-u) = 1 - \Phi(u)$ perché code simmetriche

$\Phi(-u<x<u) = \Phi(u) - (1 - \Phi(u)) = 2\Phi(u) - 1$

$2\Phi(u) - 1 = B \rightarrow \Phi(u) = \frac{B+1}{2}$

PRIMA di osservare i dati l'intervallo conterrà θ con $B\%$ di probabilità

DOPO: l'intervallo trovato contiene θ con il $B\%$ di confidenza

probabilità → confidenza

ES: $B = 0.95$

$2\Phi(u) - 1 = 0.95 \rightarrow \Phi(u) = \frac{0.95+1}{2}$

$\xrightarrow{\text{TABELLA}} F_Y(u) = \frac{1.95}{2} = 0.975$

$\xrightarrow{\text{TABELLA}} u = 1.96$

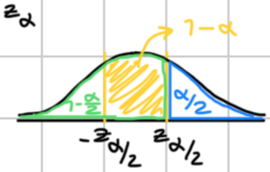
$P\left\{-1.96 \leq \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} < 1.96\right\} = 0.95$

$P\left\{\bar{X}_n + \frac{1.96\sigma}{\sqrt{n}} > \mu \geq \bar{X}_n - \frac{1.96\sigma}{\sqrt{n}}\right\} = 0.95$

FORMULE - Intervallo di confidenza: $1-\alpha$ → z_α si ottiene dalla tavola

Intervallo unilaterale:
- $\left(\bar{X}_n - z_\alpha \frac{\sigma}{\sqrt{n}}, +\infty\right)$
- $\left(-\infty, \bar{X}_n + z_\alpha \frac{\sigma}{\sqrt{n}}\right)$

Intervallo bilaterale $\left(\bar{X}_n - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{X}_n + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right)$



CHI QUADRATO X_k^2 è una variabile aleatoria continua

k v.a. normali standard IID

X_i i-esima v.a $\quad X_i \sim N(0,1) \; \forall i = 1, \dots, k$

$Y = \sum_{i=1}^{k} X_i^2 = C_k$ ← unico parametro

↳ gradi di libertà

T di Student:

v.a. continue $\quad Z \sim N(0,1) \quad C_k \sim X_k^2$

$T_k = \frac{Z}{\sqrt{X_k^2 / k}}$ ← unico parametro → gradi di libertà

$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \quad S_n = \sum_1^n \frac{(X_i - \bar{X}_n)^2}{n-1}$ assumiamo $\hat{\sigma} = S_N$ deviazione CAMPIONARIA

$\frac{\bar{X}_n - \mu}{\hat{\sigma}/\sqrt{n}} \sim T_{N-1} \xrightarrow{\text{STIMATORE}}$

$\left(\bar{X}_n - (S_N \cdot T_{N-1,\alpha/2})/\sqrt{N}, \bar{X}_n + (S_N \cdot T_{N-1,\alpha/2})/\sqrt{N}\right)$

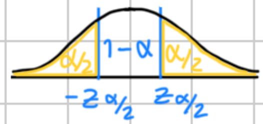
$\left(\bar{X}_n - (S_N \cdot T_{N-1,\alpha})/\sqrt{N}, +\infty\right), \left(-\infty, \bar{X}_n + (S_N \cdot T_{N-1,\alpha/2})/\sqrt{N}\right)$

Intervallo $\mu_1-\mu_2$:

$\{X_i\}_1^N \quad \{Y_i\}_1^M \quad \sigma_1,\sigma_2$ note $\quad \mu_1,\mu_2$ non note

$\mu_1,\mu_2 \in (l,u)$ [caso bilaterale]

$$Z=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\underbrace{\sqrt{\sigma_1^2/n+\sigma_2^2/m}}_{\sigma_d^2}}\sim N(0,1)$$



$1-\alpha$ livello di confidenza

$$\mu_1-\mu_2\in\left(\overline{X}-\overline{Y}-z_{\alpha/2}\sqrt{\sigma_1^2/n+\sigma_2^2/m}\;,\;\overline{X}-\overline{Y}+z_{\alpha/2}\sqrt{\sigma_1^2/n+\sigma_2^2/m}\right)$$

Intervallo per ϕ Bernulli (=intervallo per una proporzione)

Ad ogni elemento di una popolazione

$x_i\begin{cases}1 & \phi\\ 0 & 1-\phi\end{cases} \quad X_i\sim Ber(\phi) \qquad n$ IID

$$\overline{X}=\frac{\sum x_i}{n}=\frac{Y}{n}=\hat{\phi}$$

i-esimo elemento della popolazione $\{x_i\}_1^N \longrightarrow \phi\in(l,u)$?

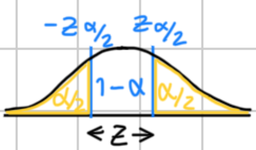
quanti degli N campioni possiedono questa caratteristica ($x_i=1$)?

$Y=\sum_1^N x_i\sim Bin(N,\phi)$

$\mathbb{E}\{Y\}=N\phi \quad \mathbb{V}\{Y\}=N\phi(1-\phi)$

per il teorema del limite centrale: $Y\sim N(N\phi, N\phi(1-\phi))$

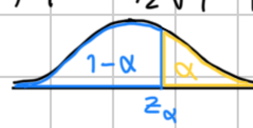
standardizzo Y: $Z=\dfrac{Y-N\phi}{\sqrt{N\phi(1-\phi)}}\sim N(0,1)$



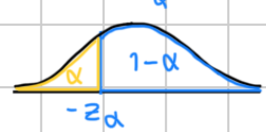
BILATERALE: $\phi\in\left(\hat{\phi}-z_{\alpha/2}\sqrt{\frac{\hat{\phi}(1-\hat{\phi})}{N}}\;,\;\hat{\phi}+z_{\alpha/2}\sqrt{\frac{\hat{\phi}(1-\hat{\phi})}{N}}\right)$

UNILATERALE:

- $\phi\in\left(\hat{\phi}-z_{\alpha}\sqrt{\frac{\hat{\phi}(1-\hat{\phi})}{N}}\;,\;+\infty\right)$



- $\phi\in\left(-\infty\;,\;\hat{\phi}+z_{\alpha}\sqrt{\frac{\hat{\phi}(1-\hat{\phi})}{N}}\right)$



ESERCIZIO: sondaggio, CONF = 0.95

$N=18 \quad X_i\begin{cases}1 & \phi\\ 0 & 1-\phi\end{cases} \quad Y=\sum x_i=17 \leftarrow 17$ SI

$\phi\in\left(\hat{\phi}-z_{\alpha/2}\sqrt{\frac{\hat{\phi}(1-\hat{\phi})}{N}}\;,\;\hat{\phi}+z_{\alpha/2}\sqrt{\frac{\hat{\phi}(1-\hat{\phi})}{N}}\right)$

$\hat{\phi}=\frac{Y}{N}=\frac{17}{18}=0.94$

$1-\alpha=0.95\rightarrow\alpha=0.05\xrightarrow{TAV} z_{\alpha/2}=1.96$

$\phi\in(0.83,1)$

VERIFICA delle IPOTESI

Campione $\{x_i\}_1^N$ estratto da una distribuzione nota a meno di uno o più parametri
└── non siamo interessati alla stima

vogliamo verificare un'ipotesi che dipende da questi parametri

$\begin{cases} H_0 & \text{ipotesi nulla - cauta} \\ H_1/H_a & \text{ipotesi alternativa} \end{cases}$

Rifiuto/Accetto H_0 a seconda di quale ipotesi spiega meglio il risultato

A partire dal dato, scelgo una statistica TEST X v.a

se assumo $H_0 \longrightarrow f_X(x|H_0)$
└── x: realizzazione di X a partire dal dato

$\mathcal{C}$: regione critica

se $x \in \mathcal{C} \rightarrow$ Rifiuto H_0

ipotesi
- SEMPLICI ← possiamo specificare la distribuzione corrispondente $f_x(x|H_0)$ ← singolo parametro con valore specifico in H_0
- COMPOSTE

- ERRORE di TIPO 1 - PRIMA specie: Rifiuto H_0, quando H_0 è vera
- ERRORE di TIPO 2 - SECONDA specie: Non Rifiuto H_0, quando H_1 è vera

livello di significatività del test:

$\mathbb{P}\{\text{Rifiuto } H_0 | H_0\} = \mathbb{P}\{\text{ERRORE I}\} = \alpha$

Potenza del test:

$\mathbb{P}\{\text{NON Rifiuto } H_0 | H_1\} = 1 - \mathbb{P}\{\text{ERRORE II}\} = 1 - \beta$

1. def. H_0
2. def. H_1
3. Statistica Test X
   $f_X(x|H_0)$ $f_X(x|H_1)$
4. Definisco $\mathcal{C}$
5. Decisione
   - $x \in \mathcal{C}$ Rifiuto H_0
   - $x \notin \mathcal{C}$ NON Rifiuto H_0

VERIFICA delle IPOTESI sulla MEDIA di una POPOLAZIONE NORMALE

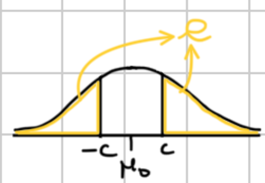
$\{x_i\}_1^N$ $X_i \overset{IID}{\sim} \mathcal{N}(\mu, \sigma^2)$ σ^2 nota

① $H_0: \mu = \mu_0$ ← ipotesi semplice ② $H_1: \mu \neq \mu_0$

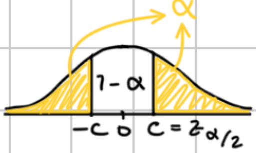
③ Scelgo statistica test: $\bar{X}$ stimatore per μ: $|\bar{X} - \mu|$

④ Regione critica: $\mathcal{C}: \{x_1, \dots, x_N : |\bar{X} - \mu| < c\}$

$\mathbb{P}\{|\bar{X} - \mu_0| > c \,|\, H_0\} = \alpha$ ← ERRORE di tipo I



Standardizzo.



$\left|\frac{\bar{X} - \mu}{\sigma/\sqrt{N}}\right| > z_{\alpha/2} \longrightarrow$ Rifiuto H_0

⑤

$\left|\frac{\bar{X} - \mu}{\sigma/\sqrt{N}}\right| \leq z_{\alpha/2} \longrightarrow$ non Rifiuto H_0

⑥ Calcolo la potenza del test: $z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{N}} \not\sim \mathcal{N}(0,1)$ perchè μ_0 approssima μ

$\frac{\bar{X} - \mu}{\sigma/\sqrt{N}} \sim \mathcal{N}(0,1)$

$\beta(\mu) = \mathbb{P}\left\{\left|\frac{\bar{X} - \mu_0}{\sigma/\sqrt{N}}\right| \leq z_{\alpha/2} \,;\, \mu\right\}$
└── curva operativa caratteristica

$1 - \beta(\mu)$: funzione di potenza del test

ESEMPIO: telefonia - messaggistica
invio 5 sms con valore 8 e ricevo
$\{x_i\}_1^5$ $\bar{X} = 9.5$ $\sigma = 2$
$H_0: \mu_0 = 8$ $H_1: \mu_0 \neq 8$ ← sms diversi
scelgo α (quanto posso sbagliare?)
$\alpha = 0.1$ $\alpha/2 = 0.05 \longrightarrow$ tabella $\rightarrow z_{\alpha/2} = 1.64$
$z = 9.5 - 8/2\sqrt{5} \cong 0.559$ $z < z_{\alpha/2} \rightarrow z \notin \mathcal{C} \rightarrow H_0!$
$\beta(\mu)$

P-VALUE

Probabilità, assumendo H_0, di ottenere un campione di dati estremo come quello osservato

Quando H_0 è vera, il p-value indica la probabilità di ottenere risultati uguali o meno probabili del test

avendo z_a

$$\mathbb{P}\{|\overset{\text{incognita}}{z}| \geq z_a\} = \text{p-value}$$

Se p-value alto non rifiutiamo H_0

↑ utile con σ non noto

ESEMPIO: ELEZIONI

1200 cittadini, 631 rieleggono il sindaco
Il sindaco uscente avrà la maggioranza?

$X_i \begin{cases} \text{vota sindaco} & 1 \\ \text{non vota sindaco} & 0 \end{cases}$

$H_0: p \leq 0.5 \quad H_1: p > 0.5$

Statistica test: $\hat{p} = \bar{X} = \frac{631}{1200} = 0.52$

$z = \frac{\bar{X} - p_0}{\sqrt{p_0(1-p_0)/N}} = 1.79$

Calcolo il p-value: $\mathbb{P}\{|z| > 1.79\} = 0.04$

→ RIFIUTO H_0

VERIFICA delle IPOTESI su una PROPORZIONE

$x_1, \dots, x_N \overset{iid}{\sim} \text{Ber}(p)$

$H_0: p = p_0 \quad H_1: p \neq p_0$

$Y = \sum_{i=1}^{N} X_i \sim \text{Bin}(N, p) \qquad \bar{X} = \frac{Y}{N} = \hat{p}$

$|Y - Np_0| > c \longrightarrow \frac{\bar{X} - p}{\sqrt{p(1-p)/N}} \sim N(0,1) \longrightarrow \frac{\bar{X} - p_0}{\sqrt{p_0(1-p_0)/N}} \sim N(0,1)$

Rifiutare H_0: $\left| \frac{\bar{X} - p_0}{p_0(1-p_0)/N} \right| \geq z_{\alpha/2}$

CAMPIONE di DATI

- QUANTITATIVA
  - $H_0: \mu = \mu_0$
  - $H_a: \mu \neq \mu_0, \mu < \mu_0, \mu \geq \mu_0$
  - σ noto → z-test: $\frac{\bar{X} - \mu_0}{\sigma/\sqrt{N}}$
  - σ non noto → t test: $\frac{\bar{X} - \mu_0}{S_N/\sqrt{N}}$
- QUALITATIVA
  - $H_0: p = p_0$
  - $H_1: p \neq p_0; p < p_0, p \geq p_0$
  - → z test: $\frac{\bar{X} - p_0}{\sqrt{p_0(1-p_0)/N}}$

Relazione tra verifica delle ipotesi e intervalli di confidenza

Intervallo bilaterale a livello $1-\alpha$ per la media di $X \sim N(\mu, \sigma^2)$ è dato da:

$$\mu \in \left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \quad \mathbb{P}\left\{\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu < \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right\} = 1-\alpha$$

Eseguendo una verifica sulle ipotesi bilaterale con $H_0: \mu = \mu_0$ e $H_1: \mu \neq \mu_0$ con significatività α, accettiamo H_0 quando

$$\mu_0 \in \left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \quad \mathbb{P}\left\{\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu < \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right\} = 1-\alpha$$

Di fatto, possiamo costruire un test delle ipotesi partendo dall'intervallo di confidenza con livello $1-\alpha$.
Accettiamo l'ipotesi nulla se il parametro rientra nell'intervallo

CONFRONTO TRA DUE MEDIE DI DUE POPOLAZIONI NORMALI

con $d_i = x_i - y_i$,
$d_i \sim N(\mu_d, \sigma_d^2)$

Due campioni di n e m dati

- Quantitativi
  - indipendenti
    - σ note → z test
      $$z = \frac{\bar{X} - \bar{Y} - \mu_d}{\sqrt{\sigma_1^2/n + \sigma_2^2/m}}$$
    - σ ignote, uguali → t test
      $$t = \frac{\bar{X} + \bar{Y} + \mu_d}{\sqrt{\sigma^2(1/n + 1/m)}}$$
  - non indipendenti
    - σ_d noto → z test
      $$z = \frac{\bar{d} - \mu_d}{\sigma_d / \sqrt{n}}$$
    - σ_d ignoto → t test
      $$t = \frac{\bar{d} - \mu_d}{S_d / \sqrt{n}}$$
- Qualitativi → z test
  $$z = \frac{\hat{p}_1 - \hat{p}_2}{\bar{p}(1-\bar{p})(1/n + 1/m)}$$

CONFRONTO TRA DUE MEDIE di DUE POPOLAZIONI NORMALI

con $d_i = x_i - y_i$,
$d_i \sim N(\mu_d, \sigma_d^2)$

Due campioni di n e m dati

- Quantitativi
  - indipendenti
    - σ note → z test
    - σ ignote, uguali → t test
  - non indipendenti
    - σ_d noto → z test
    - σ_d ignoto → t test
- Qualitativi → z test

$$z = \frac{\bar{X} - \bar{Y} - \mu_d}{\sqrt{\sigma_1^2/n + \sigma_2^2/m}}$$

$$t = \frac{\bar{X} + \bar{Y} + \mu_d}{\sqrt{\sigma^2(1/n + 1/m)}}$$

$$z = \frac{\bar{d} - \mu_d}{\sigma_d / \sqrt{n}}$$

$$t = \frac{\bar{d} - \mu_d}{S_d / \sqrt{n}}$$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\bar{p}(1-\bar{p})(1/n + 1/m)}$$