ESERCIZI: 1) UN FASCIO DI PROTONI IRRADIA DUE STRATI DI TESSUTO. LA PROBABILITÁ CHE UN PROTONE VENGA ASSORBITO (= NON ATTRAVERSI) DAL PRIMO TESSUTO É DEL 12%, MENTRE LA PROB. CHE, DOPO ESSERE PASSATO ATTRAVERSO IL PRIMO, VENGA ASSORBITO DAL SECONDO É DEL 6%. QUAL É LA PROBABILITÁ CHE UN PROTONE ATTRAVERSI ENTRAMBI I TESSUTI?

2) SE P É LA PROB. DI AVERE UN FIGLIO MASCHIO, VERIFICA CHE, AVENDO DUE FIGLI, LA PROB. CHE SIANO ENTRAMBI MASCHI SAPENDO CHE UNO É MASCHIO É PARI A $\frac{p}{2-p}$, MENTRE LA PROB. CHE SIANO ENTRAMBI MASCHI SAPENDO CHE IL PRIMO É MASCHIO É PARI A $p$.

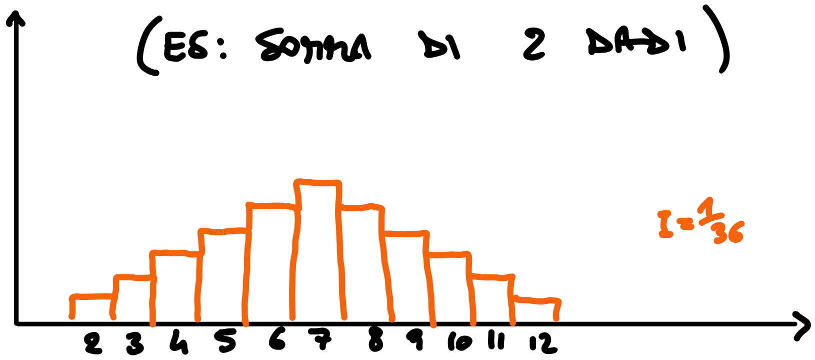
3) [GENOTIPO E FENOTIPO] CALCOLA LA PROB.

3) [GENOTIPO E FENOTIPO] CALCOLA LA PROB. CONDIZIONATA CHE LEONARDO E MARTINA ABBIANO UN FIGLIO CON FENOTIPO $Rh^-$ SAPENDO CHE LEONARDO E MARTINA HANNO GENOTIPO ETEROZIGOTE ($Rh^+Rh^-$).

---

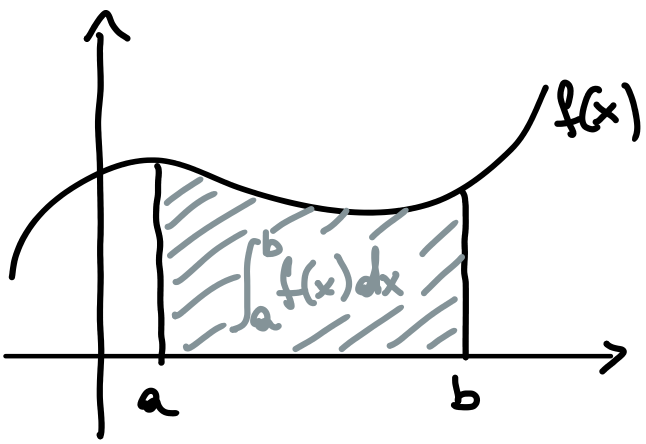
## MISURE DI PROB. CONTINUE.

RICORDIAMO COM'È FATTO UN ISTOGRAMMA: RAPPRESENTIAMO OGNI MODALITÀ DI UN CERTO CARATTERE $X$




DISEGNANDO UN RETTANGOLO AD ESSA ASSOCIATO AVENTE AREA PARI ALLA SUA FREQ. RELATIVA.

PER DEF. DI PROBABILITÀ INOLTRE SI HA CHE LA SOMMA DELLE AREE DI TUTTI I RETTANGOLI È PARI A $P(\Omega) = 1$ (= 100%).




CON GLI STRUMENTI VISTI IN QUESTO CORSO, SIAMO IN GRADO DI CALCOLARE AREE ANCHE IN CASI PIÙ GENERALI DEL SEMPLICE RETTANGOLI. IN PARTICOLARE TRAMITE GLI

INTEGRALI DEFINITI SIAMO CAPACI DI CALCOLARE L'AREA TRA IL GRAFICO DI UNA FUNZIONE $f(x)$ POSITIVA E L'ASSE $x$:

$$\text{Area}\left(\underset{[a,b]}{\underset{\text{IN}}{f(x)}}\right) = \int_a^b f(x)\,dx$$

SUPPORREMO D'ORA IN POI CHE $\Omega = \mathbb{R}$ (O UN SUO SOTTOINSIEME).

DEF: Sia $f: \Omega \longrightarrow \mathbb{R}$ UNA FUNZIONE TALE CHE:

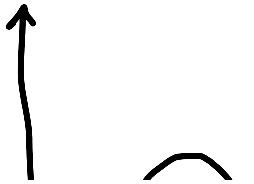
a) $\int_\Omega f(x)\,dx = 1$

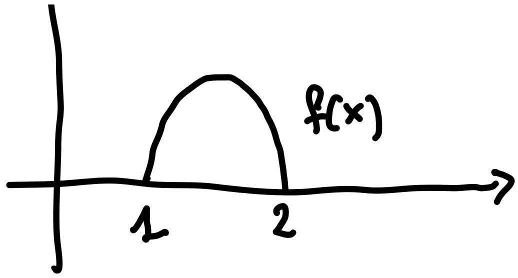
b) $f(x) \geq 0 \quad \forall x \in \Omega$

ALLORA $P(a < X < b) = \int_a^b f(x)\,dx$ É UNA MISURA DI PROBABILITÁ SU $\Omega$. IN PARTICOLARE CIASCUN PUNTO $x \in \Omega$ É UN SINGOLO EVENTO. LA FUNZIONE $f(x)$ É CHIAMATA <u>DENSITÁ DI PROBABILITÁ</u>.

ESEMPI: 1) $f(x) = 6(-x^2+3x-2)$, $\Omega = [1,2]$

$f$ É UNA DENSITÁ DI P.?






a) $\int_1^2 f(x)\,dx = \int_1^2 6(-x^2+3x-2)\,dx$

$= 6\int_1^2 (-x^2+3x-2)\,dx$

$$= 6\left[-\frac{x^3}{3}+3\frac{x^2}{2}-2x\right]_1^2 = 6\left[\left(-\frac{8}{3}+6-4\right)-\left(-\frac{1}{3}+\frac{3}{2}-2\right)\right]$$

$$= 6\left[-\frac{8}{3}+2+\frac{1}{3}-\frac{3}{2}+2\right] = 6\cdot\left[\frac{-16+12+2-9+12}{6}\right] = 6\cdot\frac{1}{6} = \boxed{1}$$

b) $f(x) \geq 0 \quad \forall x \in [1,2] = \Omega \Rightarrow$ ESERCIZIO.

QUINDI EFFETTIVAMENTE, f(x) È UNA DENSITÀ DI PROBABILITÀ. AD ESEMPIO, f(x) POTREBBE ESSERE L'ALTEZZA (IN m) DELLA POPOLAZIONE ITALIANA. LA PROB. CHE UN INDIVIDUO SIA ALTO TRA 1 m E 1,5 m È

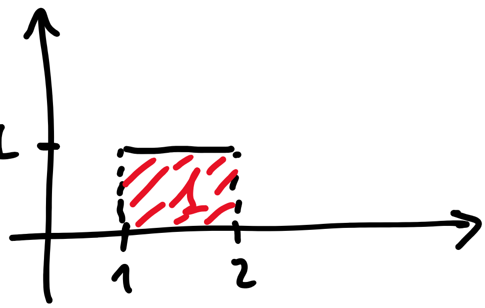
$$P(1 < X < 1.5) = P\left(1 < x < \frac{3}{2}\right) = \int_1^{3/2} f(x)\,dx$$

$$= \int_1^{3/2} 6(-x^2+3x-2)\,dx = \ldots = 0.5$$

$$P\left(1 < x < \frac{7}{4}\right) = \int_1^{7/4} f(x)\,dx = 6\left[-\frac{x^3}{3}+3\frac{x^2}{2}-2x\right]_1^{7/4} =$$

$$= \frac{27}{32} \simeq 0.84 \quad (= 84\,\%)$$

2) $f(x) = \begin{cases} 1 & \text{SE } x \in [1,2] \\ 0 & \text{altrimenti} \end{cases}$




È UNA DENSITÀ DI P.?

a) $\int_{\Omega = \mathbb{R}} f(x)\,dx = \int_1^2 1\,dx = [x]_1^2 = 2 - 1 = 1$

b) $f(x) \geq 0$ SEMPRE? OVVIO.

[OGNI VOLTA CHE PRENDIAMO UN INTERVALLO $[a,b] \subseteq \Omega$, SIAMO IN GRADO DI "CALCOLARE L'AREA" OSSIA IL NUMERO $P(a < X < b)$.
QUINDI ALL'INTERVALLO $[a,b]$ POSSIAMO ASSOCIARE LA MISURA $\int_a^b f(x)\,dx = P(a < X < b)$.
NELL'ESEMPIO (1) ABBIAMO CALCOLATO CHE $[1, \frac{7}{4}]$ CORRISPONDE A UN'AREA DI 0.84.
ALLO STESSO MODO SI POTREBBE CALCOLARE CHE ALL'INTERVALLO $[0, 1.95]$ NELL'ES. (2) CORRISPONDE A UN'AREA DI 0.95.

DEF: UN <u>INTERVALLO DI CONFIDENZA DI LIVELLO $\alpha$</u> (INGLESE: CONFIDENCE INTERVAL)

LIVELLO $\alpha$ (INGLESE: CONFIDENCE INTERVAL) PER LA DISTRIBUZIONE $X$ É UN QUALSIASI INTERVALLO $[a, b] \subset \Omega$ T.C.

$$P(a < X < b) = 1 - \alpha$$

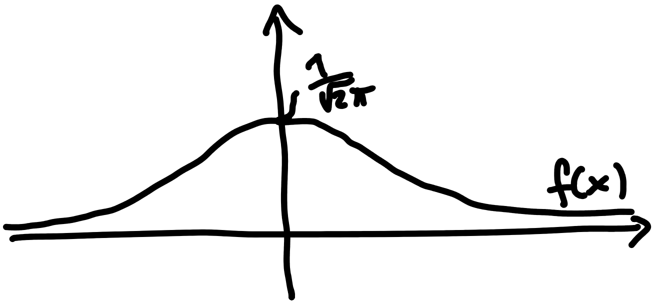
QUINDI, NELL' ES. (1), $\left[1, \frac{7}{4}\right]$ ERA T.C. $P\left(1 < X < \frac{7}{4}\right) = 0.84 \Rightarrow$ ESSO É UN INTERVALLO DI CONFIDENZA DI LIVELLO $\alpha = 0.16$ (O $\alpha = 16\%$). L'INTERVALLO $[1, 1.5]$ AVEVA PROB. $P(1 < X < 1.5) = 0.5$, QUINDI É UN INTERV. DI CONF. DI LIVELLO $\alpha = 0.5$ (50%).

NELL' ES. (2), ABBIAMO VISTO CHE $[0, 1.95]$ HA PROB. $P(0 < X < 1.95) = 0.95$, OSSIA É INTERV. DI CONF. DI LIVELLO $\alpha = 5\%$.

SI NOTI CHE, FISSATO $\alpha$, L'INTERVALLO DI CONFIDENZA DI LIVELLO $\alpha$ PER $X$ IN GENERALE NON É UNICO. AD ESEMPIO NELL' ES (1) ANCHE $\left[\frac{5}{4}, 2\right]$ HA LIVELLO 16%, MENTRE NELL' ES (2) L'INTERVALLO $[1.05, 2]$ HA ANCH'ESSO LIVELLO $\alpha = 0.05$.

$- \frac{x^2}{2}$

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




É UNA DENSITÁ DI PROBABILITÁ CON $\Omega = \mathbb{R}$ CHIAMATA <u>GAUSSIANA</u> O <u>NORMALE STANDARD</u>.

DENOTEREMO $X = N(0,1)$ UNA DISTRIBUZIONE AVENTE TALE DENSITÁ.

PIÚ IN GENERALE PRESI 2 VALORI $\mu, \sigma \in \mathbb{R}$ LA FUNZIONE

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

É ANCORA UNA DENSITÁ DI PROBABILITÁ. LA DISTRIBUZIONE VIENE DENOTATA CON $N(\mu, \sigma^2)$.

L'IMPORTANZA DELLA GAUSSIANA RISIEDE NEL:

**TEOREMA DEL LIMITE CENTRALE:**

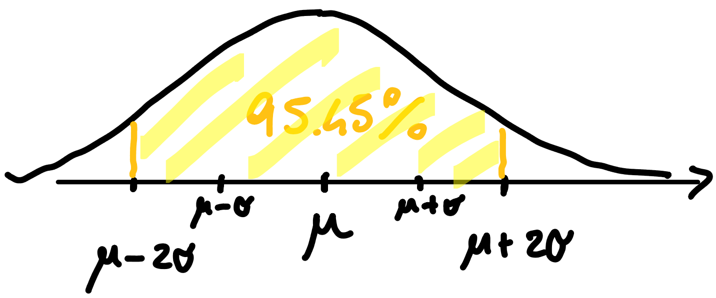
SIANO $X_1, \dots, X_n$ VARIABILI ALEATORIE INDIPENDENTI, LIMITATE E CON VARIANZA NON TENDENTE A ZERO. ALLORA LA LORO "MEDIA"

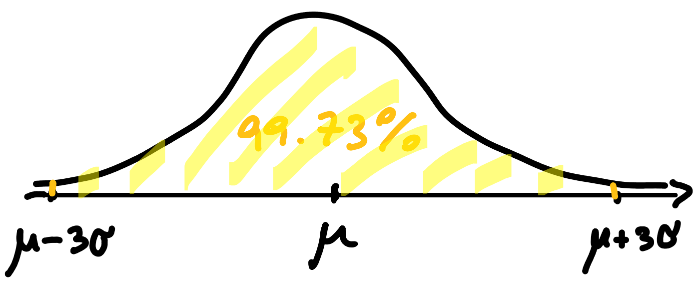
$$\frac{1}{n} \cdot \sum_{i=1}^{n} X_i \xrightarrow[n \to \infty]{} N\left(\mu, \frac{\sigma^2}{n}\right)$$

DOVE $\mu, \sigma^2$ SONO MEDIA E VARIANZA DI TUTTE LE $X_i$.

CONOSCIAMO ALCUNI INTERVALLI DI CONFIDENZA PER $N(0,1)$? CI RIFERIAMO A DATI NUMERICI CHE HANNO DATO I SEGUENTI RISULTATI, VALIDI PIÚ IN GENERALE PER LA NORMALE $N(\mu, \sigma^2)$:

- $[\mu - 2\sigma, \mu + 2\sigma]$ É UN INTERVALLO DI CONFIDENZA AL 95.45% (LIVELLO $\alpha = 4.55\%$)
- $[\mu - 3\sigma, \mu + 3\sigma]$ É UN INTERVALLO DI CONFIDENZA AL 99.73% (LIVELLO $\alpha = 0.27\%$)

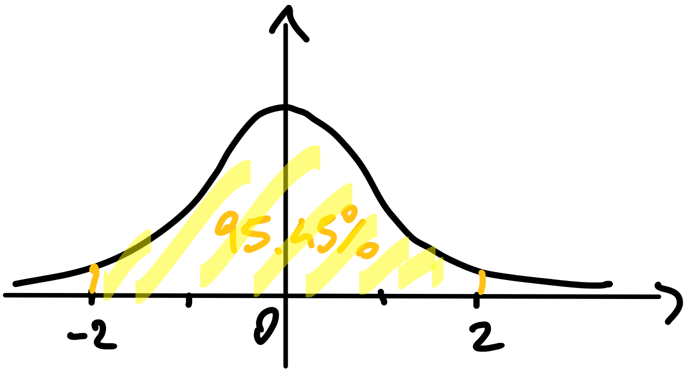


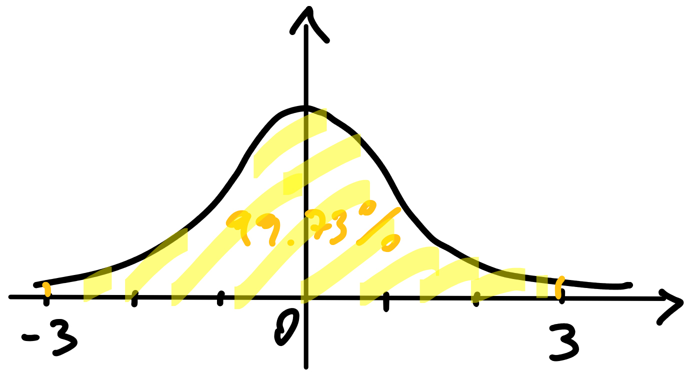




PER LA NORMALE STANDARD $N(0,1) \rightsquigarrow \mu = 0$, $\sigma^2 = 1$, ALLORA ...

$$\bullet\ [0-2\cdot 1,\ 0+2\cdot 1] = [-2,\ 2] \qquad 95.45\%$$

$$\bullet\ [0-3\cdot 1,\ 0+3\cdot 1] = [-3,\ 3] \qquad 99.73\%$$




NELLA PRATICA FAREMO:

- CALCOLO DI MEDIA E VARIANZA (CAMPIONARIA) SU UN CAMPIONE CON CARATTERISTICA $X$ (DISCRETA);
- PER NUMEROSITÀ ABBASTANZA GRANDI (ES. EXCEL $n=100$) IPOTIZZEREMO PER $X$ UN'APPROSSIMAZIONE CON UNA DISTRIBUZIONE NORMALE $N\left(\mu, \frac{\sigma^2}{n}\right)$ SFRUTTANDO IL TEOREMA DEL LIMITE CENTRALE;
- CALCOLEREMO INTERVALLI DI CONFIDENZA
  - AL 95.45% $\rightsquigarrow$ $I = \left[\mu - 2\sqrt{\frac{s^2}{n}},\ \mu + 2\sqrt{\frac{s^2}{n}}\right]$
  - AL 99.73% $\rightsquigarrow$ $I = \left[\mu - 3\sqrt{\frac{s^2}{n}},\ \mu + 3\sqrt{\frac{s^2}{n}}\right]$

4) DISTRIBUZIONE $\chi^2(n) = \sum^{n} X_i^2$ (SI LEGGE

4) DISTRIBUZIONE $\chi^2(n) = \sum_{i=1}^{n} X_i^2$ (SI LEGGE "CHI-QUADRO" DOVE LE $X_i$ SONO VARIABILI ALEATORIE INDIPENDENTI, CIASCUNA CON DISTR. $N(0,1)$. LA SUA DENSITÀ DI PROB. È

$$f(x) = \frac{1}{2^{n/2}\,\Gamma\left(\frac{n}{2}\right)}\, x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \qquad (x \geq 0)$$

DOVE $\Gamma$ È UNA FUNZIONE CHIAMATA GAMMA DI EULERO, DIPENDENTE DALL' INDICE $n$. CHIAMEREMO IL NUMERO $n$ "GRADI DI LIBERTÀ" PER LA DISTRIBUZIONE $\chi^2(n)$.

---

## TEST D'IPOTESI E TEST $\chi^2$.

DEF: UN' IPOTESI STATISTICA È UN'AFFERMAZIONE RIGUARDANTE LA DISTRIBUZIONE DI UNA O PIÙ VARIABILI ALEATORIE.

ESEMPI: 1) LANCIO UNA MONETA DIVERSE VOLTE. FACCIAMO L'IPOTESI CHE ESSA SIA EQUILIBRATA $\left(P(T) = P(C) = \frac{1}{2}\right)$. CHIAMEREMO

$H_0$ L'IPOTESI "MONETA EQUILIBRATA", E
$H_1 = H_0^C$ L'ALTERNATIVA "MONETA NON EQUILIBRATA"

DEF: $H_0$ É CHIAMATA IPOTESI NULLA.
$H_1$ É CHIAMATA IPOTESI ALTERNATIVA.

2) UN IMPUTATO VIENE ACCUSATO DI UN CERTO CRIMINE E DEVE ESSERE GIUDICATO. SE PONIAMO $H_0$ = "IMPUTATO INNOCENTE" E $H_1 = H_0^C$ = "IMPUTATO COLPEVOLE", IN SEGUITO AL GIUDIZIO SI PUÓ ESSERE IN UNO DEI 4 CASI RIASSUNTI IN TABELLA

| | $H_0$ VERA (INNOCENTE) | $H_0$ FALSA (COLPEVOLE) |
|---|---|---|
| ACCETTO $H_0$ (GIUDIZIO INNOCENTE) | OK | ERRORE DEL 2° TIPO |
| RIFIUTARE $H_0$ (GIUDIZIO COLPEVOLE) | ERRORE DEL 1° TIPO | OK |

<u>APPROCCIO IN STATISTICA</u>: PER MINIMIZZARE LO ERRORE SI TENDE AD AVERE UN APPROCCIO "PRUDENTE", OSSIA SI CERCA DI AVERE UN'IPOTESI NULLA $H_0$ CHE SIA VERA (E NEL CASO DI RIFIUTO COMMETTERE ERRORE DEL 1° TIPO) PIUTTOSTO CHE AVERNE UNA FALSA (E IN CASO DI ACCETTAZIONE COMMETTERNE UNO DEL 2° TIPO).

NELL'ES. DEL GIUDIZIO, IN UNO STATO GARANTISTA, SI É INNOCENTI COME IPOTESI ($H_0$) MINIMIZZANDO LA PROPRIETÀ DI COMMETTERE ERRORE DEL 1° TIPO (OSSIA DI GIUDICARE COLPEVOLE UN INNOCENTE). QUESTO PERCHÉ LO ERRORE DI 1° TIPO É CONSIDERATO PIÚ GRAVE DI QUELLO DEL 2° TIPO (LASCIARE IN LIBERTÀ UN COLPEVOLE).

3) VIENE STUDIATA L'EFFICACIA DI UN FARMACO PER CURARE UNA CERTA MALATTIA. SI OTTIENE, SU $n = 100$ PAZIENTI, IL REPORT IN TABELLA:

| | FARMACO | PLACEBO | |
|---|---|---|---|
| GUARITI | 35 | 12 | 47 |
| NON GUARITI | 15 | 38 | 53 |
| | 50 | 50 | 100 |

RICHIESTA: EFFETTUARE UN TEST PER VERIFICARE L'IPOTESI

$H_0$: "TRATTAMENTO E RISULTATO SONO INDIPENDENTI"
$H_1$: "TRATTAMENTO E RISULTATO SONO DIPENDENTI".