# Insiemi

### regole di De Morgan

> - $(A \cup B)^c = A^c \cap B^c$
> - $(A \cap B)^c = A^c \cup B^c$

### algebra

> Un insieme $\mathcal{A}$ è un'algebra di $\Omega$ se
>
> - $\{\emptyset, \Omega\} \subset \mathcal{A}$
> - $\forall A \in \mathcal{A} \to A^c \in \mathcal{A}$
> - L'unione finita di elementi di $\mathcal{A}$ continua a stare in $\mathcal{A}$

# Probabilità

### probabilità dell'unione

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

# Calcolo combinatorio

# Serie geometriche

### distribuzione geometrica

> Siano $n$ eventi $E_1, \ldots, E_n$ che possono verificarsi ciascuno con probabilità $p_1, \ldots, p_n$, in successione e a rotazione
>
> Sia $X_i$ l'evento "si verifica $E_i$ per primo"
>
> Chiamiamo $x$ la probabilità che non si verifichi nessuno degli $n$ eventi $E$ in un certo ciclo, ovvero la ragione della serie geometrica ($x < 1$)

$$x = \prod_{j=0}^{n} (1 - p_j)$$

$$\Pr(X_i) = \sum_{k=0}^{\infty} \left( \prod_{j=0}^{i-1} (1 - p_j) \cdot p_i \cdot x^k \right) = \frac{\prod_{j=0}^{i-1} (1 - p_j) \cdot p_i}{1 - x}$$

# Funzioni

### funzione di ripartizione

> $F(x)$ è una funzione di ripartizione, infatti soddisfa
>
> - non-decrescenza, infatti la derivata $\forall x, \quad F'(x) = \lambda e^{-\lambda x} > 0$
> - continua da destra e anche da sinistra (quindi in particolare limitata)
> - nulla per $x \to -\infty$ e tendente a uno per $x \to \infty$
>
> Grazie all'identità $P((a, b]) = F(b) - F(a)$ possiamo attribuire una probabilità agli intervalli $(a, b]$
>
> Grazie alla continuità di $F$ i punti hanno probabilità nulla, quindi tutti gli intervalli $(a, b], [a, b), [a, b], (a, b)$ sono equiprobabili

# Distribuzioni

### quantile

> La funzione `qbinom` fornisce il più piccolo valore $q$ tale che $\Pr(X \leq q) \geq x$

### minimo di esponenziali

> $X$ e $Y$ sono definite su $(0, +\infty)$, per cui anche il loro minimo avrà valori in $(0, +\infty)$, estremi esclusi
>
> $$F_U(u) = \Pr(U \leq u) = 1 - \Pr(U > u)$$
>
> $$\min\{X, Y\} > u \leftrightarrow X > u \wedge Y > u$$
>
> $$\begin{gathered}\Pr(U > u) = \Pr(\min\{X, Y\} > u) = \Pr(X > u, Y > u) = \\ = \Pr(X > u)\Pr(Y > u) = \exp\{-\lambda_1 u\}\exp\{-\lambda_2 u\} = \exp\{-(\lambda_1 + \lambda_2)u\}\end{gathered}$$
>
> $$F_U(u) = 1 - \exp\{-(\lambda_1 + \lambda_2)u\} \;\to\; U \sim \operatorname{Exp}(\lambda_1 + \lambda_2)$$
>
> $$\lambda = \lambda_1 + \lambda_2$$

### trasformazione non lineare

> $$F_U(u) = \Pr(U < u) = \Pr(T < u^2) = 1 - e^{-\lambda u^2}$$
>
> $$f_U(u) = d(1 - e^{-\lambda u^2}) = -e^{-\lambda u^2} \cdot (-2\lambda u) = 2\lambda u e^{-\lambda u^2}$$

# Variabili aleatorie

### speranza

> Per definizione, la speranza di una v.a. con densità $f$ a supporto in $(a, b)$ è
>
> $$\mathrm{E}[X] = \int_a^b t \cdot f(t)\, dt$$

### varianza

> Per definizione, la varianza di una variabile aleatoria è
>
> $$\mathrm{Var}[X] = \mathrm{E}[(X - \mathrm{E}[X])^2] = \mathrm{E}[X^2] - \mathrm{E}[X]^2 = \int_a^b t^2 f(t)\, dt - \mathrm{E}[X]^2$$

### disuguaglianza di Markov

> Disuguaglianza di Markov per una variabile aleatoria non negativa $T$
>
> $$\Pr(T \geq a) \leq \frac{\mathrm{E}[T]}{a}$$

### poissoniana condizionata

> La distribuzione di una poissoniana $X$ data la somma di due poissoniane $X + Y = n$ è una binomiale di parametri $n$ e $p = \dfrac{\lambda_X}{\lambda_X + \lambda_Y}$

### disuguaglianza di Chebychev

> Disuguaglianza di Chebychev per una variabile aleatoria non negativa $T$
>
> $$\Pr(|T - \mu| \geq a) \leq \frac{\sigma^2}{a^2}$$

### legge della speranza totale

> $$\mathrm{E}[T] = \mathrm{E}_S[\mathrm{E}[T|S]] = \sum_{i=1}^{n} \mathrm{E}[T|S = i] \cdot \Pr(S = i)$$

### legge della varianza totale

$$\mathrm{Var}(X) = \mathrm{E}_Y(\mathrm{Var}(X|Y)) + \mathrm{Var}_Y(\mathrm{E}(X|Y))$$

### approssimazione normale con TLC

Approssimazione normale con teorema del limite centrale e correzione di continuità

$$\mathrm{P}(X < a) = \mathrm{Pr}\left(\frac{X - \mathrm{E}[X]}{\sigma} \leq \frac{(a + \frac{1}{2}) - \mathrm{E}[X]}{\sigma}\right) \approx \Phi\left(\frac{(a + \frac{1}{2}) - \mathrm{E}[X]}{\sigma}\right)$$

# V.C. congiunte

### densità di probabilità congiunta

Affinchè $f(x, y)$ sia una densità di probabilità, deve valere $f(x, y) \geq 0$ per ogni $x, y \in \mathbb{R}^2$ e il suo integrale su $\mathbb{R}^2$ deve essere uguale a $1$

$$\int\int_{\mathbb{R}^2} f(x, y)\, dx\, dy = 1$$

### v.a. indipendenti congiunte

Per determinare se $X$ e $Y$ sono indipendenti calcoliamo le rispettive densità marginali $f_X$ e $f_Y$ e vediamo se il loro prodotto è uguale alla densità congiunta. Ricordiamo infatti che $X$ e $Y$ sono variabili aleatorie indipendenti se e solo se $f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$

$$x, y \in \mathbb{R} \qquad f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x, y)\, dy \qquad f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x, y)\, dx$$

### media di densità marginali

Le medie equivalgono agli integrali di $x \cdot f_X(x)$ e $y \cdot f_Y(y)$

$$\mathrm{E}[X] = \int_{\mathbb{R}} x \cdot f_X(x)\, dx$$

### valore atteso

Per linearità $\mathrm{E}[0.334 - XY] = 0.334 - \mathrm{E}[XY]$

Si ha il vettore aleatorio $(X, Y)$ e $Z = g(X, Y) = XY$, il valore atteso è

$$\mathrm{E}[Z] = \int\int_{\mathbb{R}^2} g(x, y) \cdot f_{X,Y}(x, y)\, dx\, dy$$

# Statistica

## stimatore non distorto di Var

- **Con µ non noto**

  Non conoscendo la media teorica della distribuzione, occorre uno stimatore non distorto della varianza che usi solamente i dati osservati

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

- **Con µ noto**

  Conoscendo la media teorica della distribuzione si può utilizzare

  $$S_*^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$

Dove $\bar{S}^2$ è uno stimatore distorto

## mediana

La mediana di una v.a. è il valore $m_X$ per cui

$$\Pr(X \leq m_X) = \Pr(X \geq m_X) = 0.5$$

## stimatore non distorto

Uno stimatore $\hat{\theta}$ di un parametro $\theta$ si dice corretto o non distorto se

$$\mathrm{E}[\hat{\theta}] = \theta$$

## errore quadratico medio di uno stimatore

L'errore quadratico medio di uno stimatore $\hat{\theta}$ di $\theta$ è

$$\mathrm{EQM}(\hat{\theta}) = \mathrm{E}\left[(\hat{\theta} - \theta)^2\right] = \mathrm{Var}[\hat{\theta}] + \mathrm{bias}^2$$

## stimatore di massima verosimiglianza

- **Funzione di verosimiglianza**

  Per l'indipendenza delle $X_i$, la funzione di verosimiglianza è data dal prodotto delle densità $f_{X_i}$

$$L(\lambda|x_1,\ldots,x_n)=\prod_{i=1}^{n} f_X(x_i)$$

- **log-verosimiglianza**

  Logaritmo naturale della funzione di verosimiglianza

$$l(\lambda|x_1,\ldots,x_n)=\ln(L(\lambda))$$

- **Score function**

  Derivata prima della log-verosimiglianza rispetto al parametro

$$S(\lambda)=\frac{dl(\lambda)}{d\lambda}$$

- **Stimatore di massima verosimiglianza**

  Lo stimatore si ottiene ponendo la $S(\lambda)=0$ e ricavando $\lambda$

$$\hat{\lambda}_{MLE}=S^{-1}(0)$$

## intervalli di confidenza per differenza di medie

$$Z=\overline{X}-\overline{Y}$$

$$\mathrm{E}[X-Y]=\mathrm{E}[X]-\mathrm{E}[Y]$$

- **Varianze note**

$$\mathrm{Var}[X-Y]=\mathrm{Var}[X]+\mathrm{Var}[Y]=\sigma_X^2+\sigma_Y^2$$

$$\overline{X}\sim\mathcal{N}\left(\mu_X,\frac{\sigma_X^2}{n_X}\right)\qquad \overline{Y}\sim\mathcal{N}\left(\mu_Y,\frac{\sigma_Y^2}{n_Y}\right)$$

$$Z=\frac{(\overline{X}-\overline{Y})-(\mu_X-\mu_Y)}{\sqrt{\frac{\sigma_X^2}{n_X}+\frac{\sigma_Y^2}{n_Y}}}$$

$$\left[(\overline{X}-\overline{Y})-F^{-1}_{t(n-2)}\left(1-\frac{\alpha}{2}\right)\cdot\sqrt{\frac{\sigma_X^2}{n_X}+\frac{\sigma_Y^2}{n_Y}}\right.,$$
$$\left.(\overline{X}-\overline{Y})+F^{-1}_{t(n-2)}\left(1-\frac{\alpha}{2}\right)\cdot\sqrt{\frac{\sigma_X^2}{n_X}+\frac{\sigma_Y^2}{n_Y}}\right]$$

- **Varianze ignote**

$$S^2=\frac{(n_X-1)S_X^2+(n_Y-1)S_Y^2}{n_X-n_Y-2}$$

$$Z=\frac{(\overline{X}-\overline{Y})-(\mu_X-\mu_Y)}{\sqrt{S^2\left(\frac{1}{n_X}+\frac{1}{n_Y}\right)}}$$

$$\left[(\overline{X}-\overline{Y})-F^{-1}_{t(n-2)}\left(1-\frac{\alpha}{2}\right)\cdot\sqrt{S^2\left(\frac{1}{n_X}+\frac{1}{n_Y}\right)}\right.,$$
$$\left.(\overline{X}-\overline{Y})+F^{-1}_{t(n-2)}\left(1-\frac{\alpha}{2}\right)\cdot\sqrt{S^2\left(\frac{1}{n_X}+\frac{1}{n_Y}\right)}\right]$$

## p-value

Il p-dei-dati è la probabilità di ottenere risultati uguali o meno probabili di quello osservato durante il test, supposta vera l'ipotesi nulla

$$\text{p-value} = 2 \cdot \min\big(\Pr(X \le z|H),\ \Pr(X \ge z|H)\big)$$

## regione di accettazione

I dati possono

- non contraddire l'ipotesi nulla $H_0$
- sostenere l'ipotesi alternativa $H_1$

Le regioni di accettazione per per le ipotesi sono date dagli intervalli

- $\text{p-value} \ge \alpha \ \to \ H_0$
- $\text{p-value} < \alpha \ \to \ H_1$

## metodo dei momenti

Sia $\hat{\mu}_k$ il $k$-esimo momento campionario, cioè la v.a.

$$\hat{\mu}_k := \frac{1}{n}\sum_{i=1}^{n} X_i^k = \overline{X}$$

Sia $\mu_k$ il $k$-esimo momento della popolazione, cioè il numero

$$\mu_k(\theta) = \mu_k := \mathrm{E}[X_1^k] = \int_{\mathbb{R}} x^k \cdot f(x|\theta)dx$$

La statistica $\hat{\mu}_k$ è uno stimatore corretto di $\mu_k$

## stimatore con metodo dei momenti

Si dice stimatore con metodo dei momenti del parametro $\theta$ la soluzione (se esiste) $\hat{\theta}_{mom}$ dell'equazione

$$\hat{\theta}_{mom} = \mu_1(\hat{\theta}_{mom}) = \hat{\mu}_1$$

## funzione ancillare con TLC

$$\frac{\overline{X}_n - \lambda}{\sqrt{\frac{\lambda}{n}}} \sim \mathcal{N}(0,1)$$

# Regressione lineare

## modello lineare

- Coefficiente di grado 1

$$\hat{\alpha}_1 = \frac{(\sum_i x_i y_i) - n\bar{x}\bar{y}}{(\sum_i x_i^2) - n\bar{x}^2}$$

- Termine noto

$$\hat{\alpha}_0 = \bar{y} - \hat{\alpha}_1 \bar{x}$$

- Modello lineare

$$y = \hat{\alpha}_1 \cdot x + \hat{\alpha}_0$$

## Root Sum Squared

$$\text{RSS} = \sum_i (y_i - \hat{\alpha}_0 - \hat{\alpha}_1 \cdot x_i)^2$$

## funzione ancillare

$$(\hat{\alpha}_1 - \alpha_1)\sqrt{\frac{(n-2)\sum_i (x_i^2 - \bar{x}^2)}{\text{RSS}}} \sim t(n-2)$$

## intervalli di confidenza

Bilaterale

$$\Bigg[\hat{\alpha}_1 - F^{-1}_{t(n-2)}\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\frac{\text{RSS}}{(n-2)\sum_i (x_i^2 - \bar{x}^2)}},$$
$$\hat{\alpha}_1 + F^{-1}_{t(n-2)}\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\frac{\text{RSS}}{(n-2)\sum_i (x_i^2 - \bar{x}^2)}}\Bigg]$$

Unilaterale sinistro

$$\left(-\infty, \quad \hat{\alpha}_1 + F^{-1}_{t(n-2)}(1-\alpha) \cdot \sqrt{\frac{\text{RSS}}{(n-2)\sum_i (x_i^2 - \bar{x}^2)}}\right]$$

## test statistico su $\alpha_1$

Dire che “i dati sostengono” significa fare un test positivo (maggiore, $H_1$)

$$H_0 : \alpha_1 \leq z \qquad H_1 : \alpha_1 > z$$

La statistica test è

$$S_t = (\hat{\alpha}_1 - z)\sqrt{\frac{(n-2)\sum_i (x_i^2 - \bar{x}^2)}{\mathrm{RSS}}}$$

## p-value per $\alpha_1$

Probabilità che una $t$ a $n-2$ gradi di libertà sia maggiore del valore $z$ della statistica test, cosa si può calcolare con

```
pt (St, n-2, lower = FALSE)
```