

Contents

1	Introduzione	3
2	MLE	4
2.1	MLE di una Bernoulliana	5
2.2	MLE di una Poisson	6
2.3	MLE distribuzione Uniforme	7
2.4	MLE distribuzione Normale	7
3	Intervalli di confidenza	8
3.1	μ incognita e varianza σ^2 nota	8
3.2	μ incognita e varianza σ^2 incognita	9
3.3	Metodo Montecarlo	10
4	Intervalli di predizione	11
4.1	μ incognita e varianza σ^2 incognita	11
4.2	Intervalli di confidenza per la varianza	12
4.3	Stime per la differenza tra le medie di due popolazioni normali	13
5	Intervalli di confidenza	16
5.1	Intervalli approssimati per Bernoulli	16
5.2	Qualità ed efficienza degli stimatori	17
5.2.1	Bias e Polarizzazione	17
5.2.2	Combinazioni di stimatori corretti	18
5.3	Stimatore della media di una distribuzione uniforme	20
6	Stimatori Bayesiani	22
6.1	Stimatore di θ per Bernoulli	24
6.2	Stimatore di θ per una Normale	25
6.3	Stimatore di θ per Uniformi	26

7	Verifica delle ipotesi	26
7.1	Livelli di significatività	26
7.2	Verifica di ipotesi sulla media di una popolazione normale	27
7.3	Test unilaterali	32
7.4	Il test t	34
7.5	Verifica se due popolazioni hanno la stessa media	39
7.6	Intervalli di confidenza (Unilaterali)	41
7.7	Esempio:	41
7.8	Intervallo di confidenza	41
7.9	Intervallo di confidenza nella varianza	42
7.10	Intervallo di confidenza	42
7.11	Intervallo di previsione	43
7.12	Qualità di uno stimatore	44
7.13	Proprietà di uno stimatore	45
7.14	Stimatore unbaiese	45
7.15	Valutazione di uno stimatore	45
7.16	Esempio:	46
8	Test di ipotesi	47
8.1	Metodologia alternativa	49
8.2	Test di H_p unilaterale	50
8.3	Test di ipotesi	50
8.4	Uguaglianza media di due popolazioni	51
8.5	Modelli previsionali	53
8.5.1	Modelli di regressione previsionale	53
8.5.2	Regressione lineare	54
8.5.3	Regressione Lineare (e non)	56

1 Introduzione

In probabilità quello che facciamo noi è quello di supporre che le nostre distribuzioni siano **note**.

in statistica facciamo il contrario, ossia dire qualcosa (anche detto *fare dell'inferenza*) su **parametri sconosciuti**.

Dato che i parametri sono sconosciuti il massimo che possiamo fare è quello di ottenere *una stima* dei parametri *incogniti*.

Codesti signorini sono chiamati **stimatori puntuali** e sono indicati con il simbolo $\hat{\theta}$ (in questo caso stiamo parlando di uno stimatore del parametro incognito θ)

Esistono anche gli *stimatori non puntuali*, noti come **intervalli di confidenza**, ossia un intervallo di valori in cui può essere contenuto il *dato incognito*.

Esempio $\hat{\theta}$? Altezza della popolazione

$$X_1 = 1.7$$

$$X_2 = 1.82$$

$$X_3 = 1.73$$

$$X_4 = 1.7$$

$$X_5 = 1.8$$

Possibile soluzione :

$$\hat{\theta}_a = \frac{1}{n} \sum_{i=1}^5 x_i = \frac{1.7 + 1.82 + 1.73 + 1.7 + 1.8}{5} = \frac{8.75}{5} = 1.75$$

$$\hat{\theta}_b = \frac{\min(x_i) + \max(x_i)}{2} = \frac{1.7 + 1.82}{2} = 1.76$$

$$\hat{\theta}_c = \frac{1}{3} \sum_{i=2}^4 x_i = \frac{1}{3}(1.82 + 1.73 + 1.7) = \frac{5.25}{3} = 1.75$$

Scartiamo il più *piccolo* e il *massimo*, calcolando poi la **media** dei rimanenti

2 MLE

Definizione: Stima a Massima Verosomiglianza (Maximum Likelihood Estimation)

Questa classe di stimatori sono molto usati in statistica, servono per determinare i migliori parametri del modello che si adattano ai dati e comparare molteplici modelli per *determinare* quello che si adatta di più ai dati.

Ad esempio la stima di massima verosomiglianza $\hat{\theta}$ è definita come il valore di θ che rende massima $f(x_1, x_2, \dots, x_n | \theta) \rightarrow$ anche detta *funzione di likelihood*

Likelihood: avendo dei dati quale è la probabilità che un certo modello descriva al meglio la natura dei nostri dati

$$\hat{\theta} = \operatorname{argmax} L(\theta) = \operatorname{argmax} [f(X_1 \dots X_n / \theta)]$$

Stima parametrica (Point) Parametric Estimation

Ipotesi: - Esiste un parametro θ incognito n dati a disposizione $\{X_1, X_2, X_n\}$

Legge di probabilità che descrive il fenomeno che ha generato i dati

Formula generica: Bayes

$$P(\theta / X_1 \dots X_n) = \frac{P(X_1 \dots X_n / \theta) P(\theta)}{P(X_1 \dots X_n)}$$

Verosomiglianza (likelihood)

2.1 MLE di una Bernoulliana

Vengono realizzate n prove indipendenti con probabilità p di successo

$$X_i = \begin{cases} 1 & \text{se la prova } i\text{-esima ha successo} \\ 0 & \text{altrimenti} \end{cases}$$

La distribuzione dell X_i è la seguente:

$$P(X_i = k) = p^k(1-p)^{1-k}, \quad k \in \{0, 1\}$$

La likelihood (ossia la *funzione di massa congiunta*) è:

$$\begin{aligned} f(x_1, x_2, \dots, x_n | p) &:= P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n | p) \\ &= p^{x_1}(1-p)^{1-x_1} \dots p^{x_n}(1-p)^{1-x_n} \\ &= p^{\sum_i x_i} (1-p)^{n-\sum_i x_i} \quad x_i = 0, 1 \quad i = 1, \dots, n \end{aligned}$$

Possiamo derivare rispetto a p :

$$\frac{d}{dp} \log f(x_1, x_2, \dots, x_n | p) = \frac{1}{p} \sum_{i=1}^n x_i - \frac{1}{1-p} \left(n - \sum_{i=1}^n x_i \right)$$

Da questo bro possiamo ottenere un'espressione per la stima \hat{p} :

$$\hat{p} = \frac{1}{n} \sum_{i=1}^n x_i$$

2.2 MLE di una Poisson

La funzione di *likelihood* è data da:

$$\begin{aligned} f(x_1, x_2 \dots x_n / \lambda) &= \frac{\lambda^{x_1} e^{-\lambda}}{x_1!} \dots \frac{\lambda^{x_n} e^{-\lambda}}{x_n!} \\ &= \frac{\lambda^{\sum_i x_i} e^{-n\lambda}}{x_1! \dots x_n!} \end{aligned}$$

Come sempre deriviamo e otteniamo:

$$\frac{d}{d\lambda} \log f(x_1, x_2, \dots, x_n | \lambda) = \frac{1}{\lambda} \sum_{i=1}^n x_i - n$$

Da questo bro possiamo ottenere un'espressione per la stima $\hat{\lambda}$:

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n x_i$$

La stessa formula può essere applicata al campione X_1, X_2, \dots, X_n :

$$P\{X_i = 1\} = 1 - P\{X_i = 0\}$$

Esempio Numero di incidenti stradali in 10 giornate senza pioggia

Dataset: $\{ 4 \ 0 \ 6 \ 5 \ 2 \ 1 \ 2 \ 0 \ 4 \ 3 \}$

Si vuole stimare per quell'anno la frazione di giornate senza pioggia con *2 incidenti o meno*

$$\bar{X} = \frac{1}{10} \sum_{i=1}^{10} X_i = \mathbf{2.7}$$

Così otteniamo che la media della poissoniana è 2.7, la stima desiderata è data da:

$$(1 + 2.7 + (2.7)^2 / 2) e^{-2.7} \approx 0.4936$$

2.3 MLE distribuzione Uniforme

$$f(X_1, \dots, X_n | \theta) = \begin{cases} \frac{1}{\theta} & 0 < x_1 < \theta \\ 0 & \text{altrimenti} \end{cases}$$

La formula per la stima di θ

$$\hat{\theta} = \max\{X_1, \dots, X_n\}$$

2.4 MLE distribuzione Normale

TODD AGGIUNGERE THETA MAX THETA / 2

Definizione: La distribuzione normale ha media μ e dev. st. σ **incognite**
La densità congiunta (la likelihood) è data da:

$$f(x_1, x_2, \dots, x_n | \mu, \sigma) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x_i - \mu)^2}{2\sigma^2} \right\}$$

La log-likelihood (metodo semplificato per migliorarci la vita che è già una merda) è data da:

$$\log f(x_1, x_2, \dots, x_n | \mu, \sigma) = -\frac{n}{2} \log(2\pi) - n \log \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

La risoluzione (che lasciamo al libro) ci porta alle formule per le stime:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2}$$

TODO TEORIA DEL LIMITE CENTRALE

3 Intervalli di confidenza

3.1 μ incognita e varianza σ^2 nota

Sia X_1, X_2, \dots, X_n un campione di una popolazione normale con μ *incognita* e varianza σ^2 *nota*

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

Chiedo aiuto alla regia, non so cosa stia sta roba ma comunque:

$$P\left(\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}} \approx 0.95\right)$$

Il 95% circa delle volte μ starà a una distanza non superiore a $1.96 \sigma/\sqrt{n}$ dalla media aritmetica dei dati. Se osserviamo il campione, e registriamo che $\bar{X} = \bar{x}$, allora possiamo dire che "con il 95% di confidenza"

$$\left(\bar{x} - 1.96 \frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96 \frac{\sigma}{\sqrt{n}}\right)$$

Questo intervallo è detto *intervallo di confidenza* ad un livello del 95%

Esempio segnale elettrico di valore μ

i valori registrati sono i seguenti: 5 8.5 12 15 7 9 7.5 6.5 10.5

Otteniamo \bar{x} :

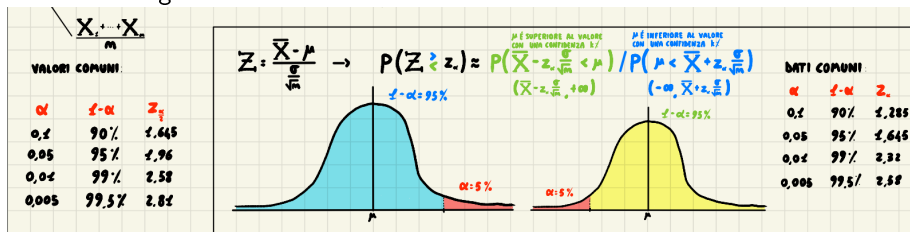
$$\bar{x} = \frac{81}{9} = 9$$

Un intervallo di confidenza al 95% per μ è

$$\left(9 - 1.96 \frac{2}{3}, \quad 9 + 1.96 \frac{2}{3}\right) = (7.69, 10.31)$$

Otteniamo quindi il 95% di fiducia che il messaggio fosse **compreso** tra 7.69 e 10.31

Figure 1: TODO CAPIRE CHE SFACCIAMM È STA ROBA



3.2 μ incognita e varianza σ^2 incognita

Dato che tutti i nostri parametri sono ignoti, non possiamo basarci sul fatto che $\sqrt{n}(\bar{X} - \mu)/\sigma$ è una *normale standard*, dobbiamo quindi ricorrere a una varianza campionaria come segue:

$$S^2 := \frac{1}{n-1} \sum_i (X_i - \bar{X})^2 \rightarrow \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$$

Alla fine otteniamo una variabile aleatoria di tipo t con $n-1$ gradi di libertà

Per Bilaterale

$$P \left\{ \bar{X} - t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}} \right\} = 1 - \alpha$$

Per Unilaterale

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

3.3 Metodo Montecarlo

supponendo di avere una funzione f da \mathbb{R}^r in \mathbb{R} e vogliamo stimare la quantità θ :

$$\theta := \int_0^1 \int_0^1 \cdots \int_0^1 f(y_1, y_2, \dots, y_n) dy_1 dy_2 \dots dy_n$$

Possiamo notare che U_1, U_2, \dots, U_r sono var. al. *uniformi* su $0,1$ quindi:

$$\mathbb{E}[f(U_1, U_2, \dots, U_r)] = \theta$$

Se produciamo un numero casuale distribuito come la funzione e lo ripetiamo n volte, possiamo stimare θ

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n X_i$$

Esempio pensiamo alla stima di questo integrale:

$$\theta := \int_0^1 \sqrt{1-y^2} dy = \mathbb{E}[\sqrt{1-U^2}]$$

Se U_1, U_2, \dots, U_{100} sono variabili aleatorie con tale distribuzione e *indipendenti* ponendo

$$X_i := \sqrt{1 - U_i^2} \quad i = 1, 2, \dots, 100$$

Otteniamo un campione di **100** variabili aleatorie di media θ . Calcoliamo ora la *media campionaria*:

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n X_i = 0.786$$

e successivamente la *deviazione standard campionaria*:

$$S = 0.23$$

dato che $t_{0.025,99} \approx 1.985$ otteniamo che un intervallo di confidenza al 95% per θ è il seguente:

$$0.786 \pm 1.985 \cdot 0.023$$

Quindi il valore è compreso tra 0.740 e 0.832

4 Intervalli di predizione

4.1 μ incognita e varianza σ^2 incognita

Supponiamo che $X_1, X_2, \dots, X_n, X_{n+1}$ sia un campione normale di media μ e varianza σ^2 entrambe *incognite*

$$\mu = \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

Per la riproducibilità

$$X_{n+1} - \bar{X}_n \sim \mathcal{N}(0, \sigma^2 + \frac{\sigma^2}{n}) \longrightarrow \frac{X_{n+1} - \bar{X}_n}{\sigma \sqrt{1 + 1/n}}$$

Dato che σ è incognita dobbiamo sostituirla col suo stimatore (scegliendo la *deviazione standard campionaria* quindi poniamo:

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

Questa grandezza è *indipendente* da \bar{X}_n quindi otteniamo

$$\frac{X_{n+1} - \bar{X}_n}{S_n \sqrt{1 + 1/n}} \sim t_n - 1$$

Esempio prendiamo in campione i valori rilevati da un contapassi negli ultimi 7 giorni

Dataset: 6822 5333 7420 6252 7005 6752

Si trovi l'intervallo di predizione al 95% di confidenza

Risoluzione: le statistiche del campione sono:

$$\bar{X}_7 \approx 6716.57 \quad S_7 \approx 733.97$$

Dalle tabelle ricaviamo che $t_{0.025,6} \approx 2.447$ (+ altri passaggi) concludiamo col dire che il 95% di confidenza che X_8 cadrà nell'intervallo $[4796, 8637]$

4.2 Intervalli di confidenza per la varianza

Se X_1, X_2, \dots, X_n è un campione di una distribuzione *normale* con parametri μ σ^2 **incogniti** ci possiamo basare sul fatto che

Formula generica:

$$(n-1) \frac{S^2}{\sigma^2} \sim \chi_{n-1}^2$$

Per caso Bilaterale :

$$\left(\frac{(n-1)s^2}{\chi_{\frac{\alpha}{2}, n-1}^2}, \frac{(n-1)s^2}{\chi_{1-\frac{\alpha}{2}, n-1}^2} \right) \quad (1)$$

Per caso Unilaterale :

$$P \left(0 < \sigma^2 < \frac{(n-1)S^2}{\chi_{1-\alpha, n-1}^2} \right) / P \left(\frac{(n-1)S^2}{\chi_{\alpha, n-1}^2} < \sigma^2 \right) \quad (2)$$

4.3 Stime per la differenza tra le medie di due popolazioni normali

Siano X_1, X_2, \dots, X_n e Y_1, Y_2, \dots, Y_m due campioni normali e differenti e denotiamo con μ_1 e σ_1^2 e con μ_2 e σ_2^2

$\bar{X} - \bar{Y}$ è lo stimatore di massima verosomiglianza $\mu_1 - \mu_2$

Tabella 7.1 Intervalli con livello di confidenza $1 - \alpha$ per campioni normali.

$X_1, X_2, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ $\bar{X} := \frac{1}{n} \sum_{i=1}^n X_i, \quad S := \left(\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \right)^{1/2}$				
Ipotesi	θ	Intervallo bilaterale	Intervallo sinistro	Intervallo destro
σ^2 nota	μ	$\bar{X} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$	$\left(-\infty, \bar{X} + z_{\alpha} \frac{\sigma}{\sqrt{n}} \right)$	$\left(\bar{X} - z_{\alpha} \frac{\sigma}{\sqrt{n}}, \infty \right)$
σ^2 non nota	μ	$\bar{X} \pm t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}}$	$\left(-\infty, \bar{X} + t_{\alpha, n-1} \frac{S}{\sqrt{n}} \right)$	$\left(\bar{X} - t_{\alpha, n-1} \frac{S}{\sqrt{n}}, \infty \right)$
μ non nota	σ^2	$\left(\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}, n-1}^2}, \frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}, n-1}^2} \right)$	$\left(0, \frac{(n-1)S^2}{\chi_{1-\alpha, n-1}^2} \right)$	$\left(\frac{(n-1)S^2}{\chi_{\alpha, n-1}^2}, \infty \right)$

Per ottenere uno *stimatore non puntuale*, dobbiamo conoscere la distribuzione di $\bar{X} - \bar{Y}$ poiche:

$$\bar{X} \sim \mathcal{N}\left(\mu_1, \frac{\sigma_1^2}{n}\right) \quad e \quad \bar{Y} \sim \mathcal{N}\left(\mu_2, \frac{\sigma_2^2}{m}\right)$$

Possiamo dedurre che:

$$\bar{X} - \bar{Y} \sim \mathcal{N}\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right)$$

Ipotizzando di conoscere σ_1^2 e σ_2^2 abbiamo che:

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n + \sigma_2^2/m}} \sim \mathcal{N}(0, 1)$$

e possiamo dedurre, con i passaggi che ci sono ormai familiari, che

Per caso Bilaterale

$$\begin{aligned} 1 - \alpha &= P \left(-z_{\frac{\alpha}{2}} < \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n + \sigma_2^2/m}} < z_{\frac{\alpha}{2}} \right) \\ &= P \left(\bar{X} - \bar{Y} - z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} < \mu_1 - \mu_2 < \bar{X} - \bar{Y} + z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \right) \end{aligned}$$

Per caso Unilaterale

$$\begin{aligned} 1 - \alpha &= P \left(\bar{X} - \bar{Y} - z_{\alpha} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} < \mu_1 - \mu_2 \right) / P \\ &= \left(\mu_1 - \mu_2 < \bar{X} - \bar{Y} - z_{\alpha} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \right) \end{aligned}$$

5 Intervalli di confidenza

5.1 Intervalli approssimati per Bernoulli

Nel caso avessimo n oggetti con una quantita X di oggetti che soddisfano i requisiti, possiamo dire che X ha distribuzione *binomiale* di parametri n e p

$$\frac{X - np}{\sqrt{np(1-p)}} \sim \mathcal{N}(0, 1)$$

Per ottenere un intervallo per p denotiamo con $\hat{p} := X/n$ la frazione degli oggetti del campione che soddisfano i requisiti, quindi:

$$\frac{X - n\hat{p}}{\sqrt{n\hat{p}(1-\hat{p})}} \sim \mathcal{N}(0, 1)$$

Da questa formula possiamo ottenere così un intervallo di confidenza

Per caso Bilaterale

$$1 - \alpha = P\left(\hat{p} - z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1-\hat{p})/n} < p < \hat{p} + z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1-\hat{p})/n}\right)$$

Per caso Unilaterale

$$P\left(\hat{p} - z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1-\hat{p})/n} < p\right) / P\left(p < \hat{p} + z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1-\hat{p})/n}\right)$$

Esempio Un campione di 100 transistor viene testato. 80 pezzi sono *adeguati*. Volendo trovare un intervallo del 95% per la percentuale p scriviamo:

$$\left(0.8 - 1.96\sqrt{0.8 \cdot 0.2/100}, \quad 0.8 + 1.96\sqrt{0.8 \cdot 0.2/100}\right) = (0.7216, \quad 0.8784)$$

Possiamo dire quindi con il 95% di confidenza che sarà *accettabile* una percentuale compresa tra il **72.16%** e il **87.84%**

Tipo di intervallo	Intervallo di confidenza
Bilaterale	$\hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1-\hat{p})/n}$
Unilaterale sinistro	$(-\infty, \hat{p} + z_{\alpha} \sqrt{\hat{p}(1-\hat{p})/n})$
Unilaterale destro	$(\hat{p} - z_{\alpha} \sqrt{\hat{p}(1-\hat{p})/n}, \infty)$

5.2 Qualità ed efficienza degli stimatori

Sia $X := (X_1, X_2, \dots, X_n)$ un campionario di una distribuzione *nota* tranne per il parametro θ che è incognito e $d(X)$ uno stimatore di θ

Come possiamo valutare la sua efficacia? un criterio può essere quello dell'*errore quadratico medio* ossia:

$$r(d, \theta) := \mathbb{E}[(d(X) - \theta)^2]$$

e sarà questo il nostro indicatore del valore di d come stimatore di θ

5.2.1 Bias e Polarizzazione

Definizione: Sia $d = d(X)$ uno stimatore del parametro θ allora:

$$b_{\theta}(d) := \mathbb{E}[d(X)] - \theta$$

Questo viene detto *bias* di d come stimatore di θ

Se il bias è nullo, si dice che è uno stimatore *corretto* o *non distorto*

Esempio Sia X_1, \dots, X_n un campione con media *incognita* θ quindi:

$$d_1(X_1, X_2, \dots, X_n) = X_1$$

$$d_2(X_1, X_2, \dots, X_n) = \frac{X_1 + X_2 + \dots + X_n}{n}$$

sono entrambi *stimatori non distorti* di θ

verifichiamo:

$$\mathbb{E}[X_1] = \mathbb{E}\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] = \theta$$

Se $d = d(X)$ è uno *stimatore corretto*, il suo errore quadratico medio è

$$\begin{aligned} r(d, \theta) &= \mathbb{E}[(d - \theta)^2] \\ &= \mathbb{E}[(d - \mathbb{E}[d])^2] \\ &= \text{Var}(d) \end{aligned}$$

Questo è lo stimatore migliore con Varianza minima

Regola generale

$$d_3(X_1, X_2, \dots, X_n) := \sum_{i=1}^n \lambda_i X_i \text{ è corretto se } \sum_{i=1}^n \lambda_i = 1$$

5.2.2 Combinazioni di stimatori corretti

Consideriamo due stimatori *corretti* e *indipendenti* di parametro θ (denotati con d_1 e d_2) con varianze rispettivamente σ_1^2 σ_2^2

$$\mathbb{E}[d_i] = \theta \quad \text{Var}(d_i) = \sigma_i^2 \quad i = 1, 2$$

uno stimatore corretto di θ è il seguente

$$d := \lambda d_1 + (1 - \lambda) d_2$$

Successivamente vogliamo trovare anche il valore di λ che produce lo stimatore d con il *minore errore quadratico medio*

$$\begin{aligned} r(d, \theta) &= \text{Var}(d) \\ &= \lambda^2 \text{Var}(d_1) + (1 - \lambda)^2 \text{Var}(d_2) \quad \text{per l'indipendenza di } d_1 \text{ e } d_2 \\ &= \lambda^2 \sigma_1^2 + (1 - \lambda)^2 \sigma_2^2 \end{aligned}$$

ayo bro what's this shit, le me calculate the derivata with latti:

$$\frac{d}{d\lambda} r(d, \theta) = 2\lambda \sigma_1^2 - 2(1 - \lambda) \sigma_2^2$$

e belin lo studiamo sto segno o no? denotiamo con $\hat{\lambda}$ il valore di θ che produce il minimo

$$2\hat{\lambda} \sigma_1^2 - 2(1 - \hat{\lambda}) \sigma_2^2 = 0$$

da cui otteniamo:

$$\hat{\lambda} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} = \frac{1/\sigma_1^2}{1/\sigma_1^2 + 1/\sigma_2^2}$$

il peso ottimale da dare a uno stimatore deve essere **inversamente** proporzionale alla sua varianza

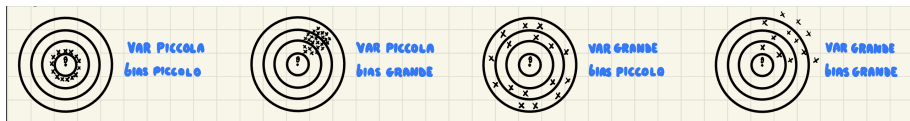
La migliore combinazione lineare delle d_i per l'errore quadratico medio è:

$$\begin{aligned} r(d, \theta) &= \text{Var}(d) \\ &= \left(\frac{1}{\sum_{i=1}^n 1/\sigma_i^2} \right)^2 \sum_{i=1}^n \left(\frac{1}{\sigma_i^2} \sigma_i^2 \right) \\ &= \frac{1}{\sum_{i=1}^n 1/\sigma_i^2} \end{aligned}$$

$$r(d, \theta) = \text{Var}(d)$$

Bias/Polarizzazione Se $d(X)$ è **distorto**:

$$\begin{aligned} r(d, \theta) &= \mathbb{E}[(d - \theta)^2] \\ &= \text{Var}(d) + 0 + \mathbb{E}[b_\theta(d)^2] \\ &= \text{Var}(d) + b_\theta(d)^2 \end{aligned}$$



5.3 Stimatore della media di una distribuzione uniforme

Siano X_1, X_2, \dots, X_n un campione estratto da una popolazione con distribuzione *uniforme* su $(0, \theta)$ dove θ è un parametro incognito.

Dato che (non si sa come) $\mathbb{E}[X_i] = \theta/2$ è uno stimatore naturale per θ è dato da

$$d_1 = d_1(X) := 2\overline{X} := \frac{2}{n} \sum_{i=1}^n X_i$$

Siccome $\mathbb{E}[d_1] = \theta$, otteniamo che:

$$\begin{aligned}r(d_1, \theta) &= Var(d_1) \\&= \frac{4}{n} Var(X_i) \\&= \frac{4}{n} \cdot \frac{\theta^2}{12} \\&= \frac{\theta^2}{3n}\end{aligned}$$

Un secondo stimatore che abbiamo è quello di **massima verosomiglianza** (d_2):

$$d_2 = d_2(X) = MLE := \max(X_i)$$

Per trovare l'errore quadratico medio di d_2 dobbiamo prima conoscere la sua *media* e la sua *varianza*

$$\mathbb{E}[d_2] = \int_0^\theta x \frac{nx^{n-1}}{\theta^n} dx = \frac{n}{n+1} \theta$$

$$\mathbb{E}[d_2^2] = \int_0^\theta x^2 \frac{nx^{n-1}}{\theta^n} dx = \frac{n}{n+2} \theta^2$$

$$Var(d_2) = \mathbb{E}[d_2^2] - \mathbb{E}[d_2]^2 = \frac{n\theta^2}{(n+2)(n+1)^2}$$

Quindi ora calcoliamo la $r(d_2, \theta)$:

$$\begin{aligned}
 r(d_2, \theta) &= \text{Var}(d_2) + (E[d_2] - \theta)^2 \\
 &= \frac{n\theta^2}{(n+2)(n+1)^2} + \frac{\theta^2}{(n+1)^2} \\
 &= \frac{\theta^2}{(n+1)^2} \left[\frac{n}{n+2} + 1 \right] \\
 &= \frac{2\theta^2}{(n+1)(n+2)}
 \end{aligned} \tag{3}$$

Confrontando gli errori quadratici medi notiamo che d_2 è **migliore** di d_1 per θ

$$\frac{2\theta^2}{(n+1)(n+2)} \leq \frac{\theta^2}{3n} \quad d_2 \text{ migliore}$$

* SI DEPOLARIZZA $d_1 \rightarrow d_1 = \frac{m+\varepsilon}{m} d$, CORRETTO $\rightarrow r(d_1, \theta) = \text{VAR}(d_1) = \frac{(m+\varepsilon)^2}{m^2} \text{VAR}(d) = \frac{\theta^2}{m^2+2m}$
 $\frac{\theta^2}{m^2+2m} < \frac{\varepsilon^2 \theta^2}{m^2+3m+2} \Rightarrow d_1 \text{ MIGLIORE}$

* ESISTE UNO MIGLIORE DI $d_1 \rightarrow d_1 = c \cdot d$, $\rightarrow r(d_1, \theta) = \text{VAR}(d_1) + (E[d_1] - \theta)^2 = c^2 \text{VAR}(d) + (c E[d] - \theta)^2 = \frac{2c m \theta^2}{(m+2)(m+2)} + \theta^2 (c \frac{m}{m+2} - \varepsilon)^2$

(MINOR ERRORE: $\frac{d^2 (c \frac{m}{m+2} - \varepsilon)^2}{d^2} = \frac{2c m \theta^2}{(m+2)(m+2)} + \theta^2 (c \frac{m}{m+2} - \varepsilon)^2 \rightarrow \frac{d^2}{d^2} \rightarrow \frac{c}{m+2} + c m \cdot (m+2) \rightarrow c \frac{(m+2)^2}{m+2} = c(m+2)$)
 $r(d_1, \theta) = \frac{\theta^2}{m^2+2m+2}$
 $\frac{\theta^2}{m^2+2m+2} < \frac{\theta^2}{m^2+2m} \Rightarrow d_1 \text{ MIGLIORE}$

6 Stimatori Bayesiani

Definizione: Quando il parametro incognito θ possiamo considerarlo come una variabile aleatoria, questo approccio viene detto *bayesiano*.

Se abbiamo delle informazioni su quelli che possono essere assunti i valori da θ ed esse assumono la forma di distribuzione di probabilità si dice che abbiamo **una**

distribuzione a priori per θ

Se i valori che osserviamo sono $X_i = x_i \quad i = 1, 2, \dots, n$ la *densità di probabilità condizionale di θ* è data da:

$$\begin{aligned} f(\theta|x_1, x_2, \dots, x_n) &= \frac{f(x_1, x_2, \dots, x_n, \theta)}{f(x_1, x_2, \dots, x_n)} \\ &= \frac{f(x_1, x_2, \dots, x_n|\theta)p(\theta)}{\int f(x_1, x_2, \dots, x_n|\theta')p(\theta')d\theta'} \end{aligned}$$

Dove:

- $f(\theta|x_1, x_2, \dots, x_n)$ Viene detta *probabilità a posteriori*
- $f(x_1, x_2, \dots, x_n)$ è la *MLE Marginale*
- $f(x_1, x_2, \dots, x_n|\theta)$ è la *MLE*
- $p(\theta)$ è la *distribuzione a priori*

Una buona stima per θ può essere la **media** perciò:

$$\mathbb{E}[\theta|X_1 = x_1, \dots, X_n = x_n] = \int_{-\infty}^{\infty} \theta f(\theta|x_1, x_2, \dots, x_n) d\theta \quad \text{nel caso continuo}$$

6.1 Stimatore di θ per Bernoulli

Se abbiamo X_1, X_2, \dots, X_n Bernoulliane, con massa di probabilità:

$$f(x|\theta) = \theta^x(1-\theta)^{1-x} \quad x = 0, 1$$

Dove θ è un parametro sconosciuto

Supponiamo quindi che la *distribuzioni a priori* di θ sia uniforme su $(0,1)$, denotiamo con p la densità a propri di θ

$$p(\theta) = 1 \quad 0 < \theta < 1$$

La densità condizionale di θ date x_1, x_2, \dots, x_n è

$$\begin{aligned} f(\theta | x_1, x_2, \dots, x_n) &= \frac{f(x_1, x_2, \dots, x_n, \theta)}{f(x_1, x_2, \dots, x_n)} \\ &= \frac{f(x_1, x_2, \dots, x_n | \theta) p(\theta)}{\int_0^1 f(x_1, x_2, \dots, x_n | \vartheta) p(\vartheta) d\vartheta} \\ &= \frac{\theta^{\sum_i x_i} (1-\theta)^{n-\sum_i x_i}}{\int_0^1 \vartheta^{\sum_i x_i} (1-\vartheta)^{n-\sum_i x_i} d\vartheta} \end{aligned} \quad (4)$$

Non è difficile provare (e invece lo è) integrando per parti un certo numero di volte che per ogni valore di m e r :

$$\int_0^1 \theta^m (1-\theta)^r d\theta = \frac{m!r!}{(m+r+1)!}$$

poniamo ora $x := \sum_{i=1}^n x_i$

$$f(\theta|x_1, x_2, \dots, x_n) = \frac{(n+1)!}{x!(n-x)!} \theta^x (1-\theta)^{n-x} \quad 0 < \theta < 1$$

Ora siamo in grado di calcolare *la stima bayesiana*

$$\begin{aligned} E[\theta \mid x_1, x_2, \dots, x_n] &= \frac{(n+1)!}{x!(n-x)!} \int_0^1 \theta^{1+x} (1-\theta)^{n-x} d\theta \\ &= \frac{(n+1)!}{x!(n-x)!} \frac{(1+x)!(n-x)!}{(n+2)!} \\ &= \frac{x+1}{n+2} \\ &= \frac{1 + \sum_{i=1}^n X_i}{n+2} \end{aligned} \tag{5}$$

Esempio Se raccogliamo un campione di 10 *bernoulliane* e trovassimo **6 successi**, lo stimatore bayesiano di θ fornirebbe un valore di **7 / 12**

Lo stimatore di massima verosomiglianza vale invece **6 / 10**

6.2 Stimatore di θ per una Normale

Supponiamo che X_1, X_2, \dots, X_n sia una distribuzione normale con media θ *incognita* e varianza σ_0^2 **nota**

Calcoliamo la *densità condizionale* di θ :

$$f(\theta \mid x_1, x_2, \dots, x_n) = \frac{f(x_1, x_2, \dots, x_n \mid \theta) p(\theta)}{f(x_1, x_2, \dots, x_n)}$$

Ora calcoliamo la *media*:

$$\mathbb{E}[\theta \mid X_1, X_2, \dots, X_n] = \frac{n/\sigma_0^2}{n/\sigma_0^2 + 1/\sigma^2} \bar{X} + \frac{1/\sigma^2}{n/\sigma_0^2 + 1/\sigma^2} \mu$$

e successivamente la *varianza*:

$$\text{Var}(\theta|X_1, X_2, \dots, X_n) = \frac{1}{n/\sigma_0^2 + 1/\sigma^2}$$

6.3 Stimatore di θ per Uniformi

Avendo una funzione di likelihood $f(x_1, x_2, \dots, x_n|\theta)$ e sapendo che la distribuzione è *uniforme*

$$\theta \in [a, b]$$

$$d_b = d_{MLE}$$

7 Verifica delle ipotesi

Se prima noi cercavamo di stimare determinati parametri ora cerchiamo invece (avendo un campione raccolto) di trovare *le ipotesi che li coinvolga*

7.1 Livelli di significatività

Consideriamo una popolazione con distribuzione F_θ che dipende da θ incognito e vogliamo verificare una qualche ipotesi su θ , una distribuzione normale con media θ e varianza 1 abbiamo due ipotesi

1. $H_0 : \theta = 1 \rightarrow$ *ipotesi nulla semplice*
2. $H_0 : \theta \leq n \rightarrow$ *ipotesi nulla composta*

La prima ipotesi afferma che la popolazione ha come distribuzione $\mathcal{N}(1, 1)$ mentre la seconda sostiene che è normale con *varianza* 1 e *media non* superiore a 1. Queste due ipotesi si possono verificare su un campione aleatorio X_1, X_2, \dots, X_n

Esiste una regione critica **C** per cui se il campione aleatorio vi appartiene l'ipotesi *non viene accettata*. Esiste un livello di tolleranza specificato all'interno della regione critica per cui un'ipotesi può essere *ancora accettata*. Questa tolleranza è definita dal **livello di significatività**, ovvero viene definito α tale che se l'ipotesi è vera la probabilità di rifiutarla non superi α

accetta H_0 se $(X_1, X_2, \dots, X_n) \notin C$

e

rifiuta H_0 se $(X_1, X_2, \dots, X_n) \in C$

Esistono **due tipi di errori**:

1. **Prima Specie**: Si rifiuta H_0 anche se è vera
2. **Seconda Specie**: si accetta H_0 anche se è falsa

7.2 Verifica di ipotesi sulla media di una popolazione normale

Supponiamo di avere X_1, X_2, \dots, X_n sia un campione aleatorio di una popolazione normale di parametri μ, σ^2 con *varianza* nota e *media* incognita, vogliamo verificare le seguenti ipotesi:

$$H_0 : \mu = \mu_0$$

e l'ipotesi *alternativa*

$$H_1 : \mu \neq \mu_0$$

Lo **stimatore puntuale** per μ è:

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

La **regione critica del test** invece è:

$$C := \{(X_1, X_2, \dots, X_n) : |\bar{X} - \mu_0| > c\}$$

Dove c rappresenta la *tolleranza*

Quando $\mu = \mu_0$ sappiamo che \bar{X} ha distribuzione **normale** con media μ_0 e varianza σ^2/n allora:

$$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim Z$$

Dove la relazione \sim è **condizionata** all'ipotesi $H_0 : \mu = \mu_0$

c deve soddisfare la seguente relazione:

$$\alpha = P(\text{errore di I specie}) = P_{\mu_0}(|\bar{X} - \mu_0| > c)$$

Possiamo scrivere l'equazione di sopra in questo modo:

$$\alpha = 2P\left(Z > \frac{c\sqrt{n}}{\sigma}\right)$$

Per $P(Z > c\sqrt{n}/\sigma)$ per la definizione $z_{\frac{\alpha}{2}}$ vale:

$$P(Z > z_{\frac{\alpha}{2}}) = \frac{\alpha}{2} \longrightarrow c = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

Il test con livello di significatività ha **due esiti**:

$$\text{si rifiuta } H_0 \text{ se } \left| \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \right| > z_{\frac{\alpha}{2}}$$

$$\text{si accetta } H_0 \text{ se } \left| \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \right| \leq z_{\frac{\alpha}{2}}$$

$$p \text{ dei dati} = \sqrt{n} \frac{\bar{X} - \mu_0}{\sigma}$$

H_0 si **accetta** se $2P(Z > z_{\frac{\alpha}{2}})$ è elevata

H_0 si **rifiuta** se $2P(Z > z_{\frac{\alpha}{2}})$ è bassa

Perché se la probabilità che Z sia $> z_{\frac{\alpha}{2}}$ è *alta* allora il mio valore sarà vicino al mezzo e va bene. Se è basso allora è *lontano* dal mezzo e non va bene.

Esempio supponiamo una media campionaria dei 5 segnali ricevuti fosse 8.5:

$$\frac{\sqrt{n}}{\sigma} |\bar{X} - \mu_0| = \frac{\sqrt{5}}{2} \cdot 0.5 \approx 0.559$$

Dato che:

$$P(|Z| > 0.559) = 2P(Z > 0.559) \approx 2 \cdot 0.288 = 0.576$$

Otteniamo che il *p-dei-dati* è 0.576 e quindi l'ipotesi nulla che il segnale inviato fosse **8**, che viene accettata per ogni $\alpha < 0.576$

Se avessimo ottenuto che $\bar{X} = 11.5$ il valore del *p-dei-dati* sarebbe:

$$P\left(|Z| > \frac{\sqrt{5}}{2} \cdot 3.5\right) \approx 2P(Z > 3.913) \approx 0.00005$$

Con un valore così piccolo, l'ipotesi che il messaggio fosse stato 8, **va rifiutata**.

Riprendendo il discorso degli errori di specie andiamo a vedere ora *gli errori di seconda specie*.

Rinfreschiamo la memoria, l'errore di seconda specie è quando *si accetta H_0 anche se è falsa*, quindi:

$$\beta(\mu) := P_{\mu}(\text{accettare } H_0) = P_{\mu} \left(-z_{\frac{\alpha}{2}} \leq \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq z_{\frac{\alpha}{2}} \right)$$

La funzione $\beta(\mu)$ è detta **curva OC** (*curva operativa caratteristica*) e rappresenta la probabilità di accettare H_0 quando la media reale è μ .

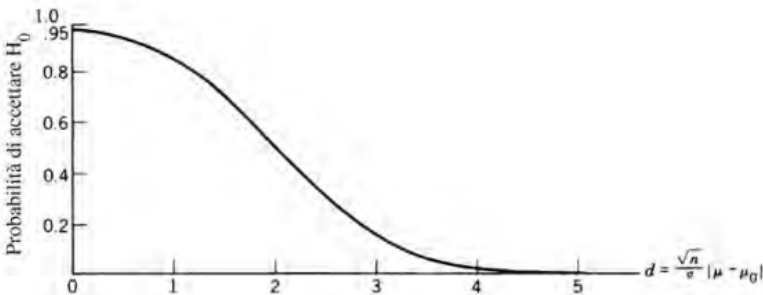


Figure 2: Curva OC di un test *bilaterale* per la media di una popolazione normale, con $\alpha = 0.05$

Per calcolare la probabilità ricordiamoci il fatto che $\bar{X} \sim \mathcal{N}(\mu, \sigma^2/n)$:

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

Quindi:

$$\begin{aligned}\beta(\mu) &= P_{\mu} \left(-z_{\frac{\alpha}{2}} \leq \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq z_{\frac{\alpha}{2}} \right) \\ &= \Phi \left(\frac{\mu_0 - \mu}{\sigma/\sqrt{n}} + z_{\frac{\alpha}{2}} \right) - \Phi \left(\frac{\mu_0 - \mu}{\sigma/\sqrt{n}} - z_{\frac{\alpha}{2}} \right) = 1 - \alpha\end{aligned}$$

Dove Φ indica la *funzione di ripartizione* della distribuzione normale standard

Esempio quanto vale la probabilità di accettare $\mu = 8$ quando in realtà $\mu = 10$:

$$\frac{\sqrt{n}}{\sigma}(\mu_0 - \mu) = \frac{\sqrt{5}}{2}(-2) = -\sqrt{5}$$

Dato che $z_{0.025} \approx 1.96$ ricaviamo la probabilità cercata:

$$\begin{aligned}\beta(10) &\approx \Phi(-\sqrt{5} + 1.96) - \Phi(-\sqrt{5} - 1.96) \\ &= 1 - \Phi(0.276) - 1 + \Phi(4.196) \\ &\approx -0.609 + 1 = \mathbf{0.391}\end{aligned}$$

Riprendendo il discorso della curva OC, ci permette di dimensionare il campione in modo che l'errore di seconda specie soddisfi le condizioni specifiche.

Come facciamo a trovare n tale che la probabilità di accettare $H_0 : \mu = \mu_0$ quando il vero valore è μ_1 sia un valore fissato β per n tale che $\beta(\mu_1) \approx \beta$

$$n \approx \left[\frac{(z_{\frac{\alpha}{2}} + z_{\beta})\sigma}{\mu_1 - \mu_0} \right]^2$$

Notiamo che anche nel caso in cui $\mu_1 < \mu_0$ troviamo sempre la stessa formula

Esempio Quante volte è necessario inviare il segnale con verifica dell'ipotesi $H_0 : \mu = 8$ con livello di significatività 0.05 con almeno il 75% di probabilità di rifiutare l'ipotesi nulla quando $\mu = 9.2$

Dato che $z_{0.025} \approx 1.96$ e $z_{0.25} \approx 0.67$

$$n \approx \left(\frac{1.96 + 0.67}{1.2} \right)^2 \cdot 4 \approx 19.21$$

Come vediamo per il risultato è necessario un *campione di 20 segnali*, quindi con $n = 20$

$$\begin{aligned} \beta(9.2) &\approx \Phi \left(-\frac{1.2\sqrt{20}}{2} + 1.96 \right) - \Phi \left(-\frac{1.2\sqrt{20}}{2} - 1.96 \right) \\ &\approx \Phi(-0.723) - \Phi(-4.643) \\ &\approx 1 - \Phi(0.723) \approx \mathbf{0.235} \end{aligned}$$

Quindi ricapitolando, se il segnale viene trasmesso 20 volte c'è il 76.5% di probabilità che l'ipotesi nulla $\mu = 8$ sia **rifiutata** se la media reale è **9.2**

7.3 Test unilaterali

Introduzione bla bla bla

Verifichiamo due ipotesi:

$$H_0 : \mu = \mu_0 \quad \text{contro} \quad H_1 : \mu > \mu_0$$

Dovremmo rifiutare l'ipotesi nulla quando lo stimatore di μ è molto più grande di μ_0 , la regione critica è quindi:

$$C := \{(X_1, X_2, \dots, X_n) : \bar{X} - \mu_0 > c\}$$

la probabilità di rifiuto dovrebbe essere α quando H_0 è vera, occorre però che c soddisfi la relazione:

$$P_{\mu_0}(\bar{X} - \mu_0 > c) = \alpha$$

Il test *con livello di significatività* α dovrà rifiutare H_0 se $\bar{X} - \mu_0 > z_\alpha \cdot \sigma / \sqrt{n}$

$$\text{si rifiuta } H_0 \text{ se } \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} > z_\alpha$$

$$\text{si accetta } H_0 \text{ se } \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \leq z_\alpha$$

Quella trovata è detta *regione critica* **unilaterale** o a una coda, quindi il problema di verificare le ipotesi alternative

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu > \mu_0$$

Si dice problema di **test unilaterale**

poniamo $Z := \sqrt{n}(\bar{X} - \mu) / \sigma$ questa statistica è una *normale standard* quindi:

$$\begin{aligned} \beta(\mu) &:= P_\mu(\text{accettare } H_0) \\ &= P_\mu\left(Z \leq \frac{\mu_0 - \mu}{\sigma / \sqrt{n}} + z_\alpha\right) \\ &= \Phi\left(\frac{\mu_0 - \mu}{\sigma / \sqrt{n}} + z_\alpha\right) = 1 - \alpha \end{aligned}$$

Dato che Φ in quanto *funzione di ripartizione* è **crescente** però $\beta(\mu)$ è una funzione **decrecente**

L'ipotesi *unilaterale*

$$H_0 : \mu \leq \mu_0$$

contro l'alternativa

$$H_1 : \mu > \mu_0$$

Per accertarci che il *livello di significatività* sia rimasto α

Al variare di μ la probabilità di rifiuto è data da $1 - \beta(\mu)$

Dobbiamo verificare che per ogni μ compatibile con H_0 per ogni $\mu \leq \mu_0$

$$1 - \beta(\mu) \leq \alpha, \quad \text{per ogni } \mu \leq \mu_0$$

Quindi:

$$\beta(\mu) \geq 1 - \alpha, \quad \text{per ogni } \mu \leq \mu_0$$

Osservazione è possibile verificare l'ipotesi

$$H_0 : \mu = \mu_0$$

contro l'ipotesi *alternativa*

$$H_1 : \mu < \mu_0$$

ad un livello di significatività α , decidendo che:

$$\text{si rifiuta } H_0 \text{ se } \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq -z_\alpha$$

$$\text{si accetta } H_0 \text{ se } \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq -z_\alpha$$

7.4 Il test t

Fino ad ora abbiamo supposto che l'unico parametro incognito fosse la *media*, in questo caso la nostra varianza σ^2 **non è nota**

In questa situazione consideriamo che si possa verificare l'ipotesi nulla che μ sia uguale ad un valore assegnato μ_0 contro l'ipotesi alternativa $\mu \neq \mu_0$

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

(Cambiare poi con desc più corta) Come in precedenza, sembra ragionevole rifiutare l'ipotesi nulla quando \bar{X} cade lontano da μ_0 tuttavia la distanza a cui deve essere da μ_0 per giustificare questo rifiuto, dipende dalla deviazione standard σ che in quella sede era nota; in particolare $|\bar{X} - \mu_0|$ doveva essere maggiore di $z_{\frac{\alpha}{2}} \cdot \sigma / \sqrt{n}$ o equivalentemente

$$\left[\frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \right] > z_{\frac{\alpha}{2}}$$

Qui σ non è più conosciuta, sostituiamola quindi con la *deviazione standard campionaria* S

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

rifiutando l'ipotesi nulla quando

$$\left| \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \right|$$

Quindi alla fine noi dobbiamo ottenere una distribuzione t

$$t_{n-1} \sim \frac{\bar{X} - \mu}{S / \sqrt{n}}$$

Se si denota con T la statistica di questo test, ovvero

$$T := \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$$

allora quando H_0 è vera (visto che $\mu = \mu_0$) ha distribuzione t **con $n - 1$ gradi di libertà**.

$$P_{\mu_0} \left(-c \leq \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \leq c \right) = 1 - \alpha$$

Se vogliamo ricavare c :

$$\begin{aligned} \alpha &= 1 - P(-c \leq T < c) \\ &= P(T \leq -c) + P(T \geq c) \\ &= 2P(T \geq c) \end{aligned}$$

Per cui $P(T > c) = \frac{\alpha}{2}$, e quindi deve valere $c = t_{\frac{\alpha}{2}, n-1}$, quindi in fin dei conti:

$$\text{si rifiuta } H_0 \text{ se } \left| \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \right| > t_{\frac{\alpha}{2}, n-1}$$

$$\text{si accetta } H_0 \text{ se } \left| \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \right| \leq t_{\frac{\alpha}{2}, n-1}$$

Vedere tabella sotto per tutt'e cose

Figure 3: X_1, X_2, \dots, X_n è un campionario estratto da una popolazione $\mathcal{N}(\mu, \sigma^2)$

$$\sigma^2 \text{ nota} \quad \bar{X} := \frac{1}{n} \sum_{i=1}^n X_i$$

H_0	H_1	Statistica del test, X_{ts}	Si rifiuta H_0 con livello di significatività α se...	p -dei-dati se $X_{ts} = t$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	$\dots X_{ts} > z_{\frac{\alpha}{2}}$	$2P(Z > t)$
$\mu \leq \mu_0$	$\mu > \mu_0$	$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	$\dots X_{ts} > z_{\alpha}$	$P(Z > t)$
$\mu \geq \mu_0$	$\mu < \mu_0$	$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	$\dots X_{ts} < -z_{\alpha}$	$P(Z < t)$

Esempio Vogliamo verificare l'ipotesi che il consumo *medio* di acqua sia 350 galloni al giorno.

Si misurano i consumi medi di un campione di 20 *campioni* che seguono:

340 356 332 362 318 344 386 402 322 360
362 354 340 372 338 375 364 355 324 370

Dobbiamo verificare le due ipotesi seguenti:

$$H_0 : \mu = 350 \quad \text{contro} \quad H_1 : \mu \neq 350$$

Calcoliamo ora la **media** e la **deviazione standard campionaria**

$$\bar{X} = 353.8 \quad S \approx 21.85$$

troviamo ora il valore della statistica del test:

$$T \approx \frac{\sqrt{20} \cdot 3.8}{21.85} \approx 0.778$$

il valore che abbiamo trovato è minore di $t_{0.05,19} \approx 1.729$ l'ipotesi nulla è *accettata*
ad un livello del 5% TODO FINIRE PAGINA PDF 324 / 342 TOTALE

Figure 4: X_1, X_2, \dots, X_n è un campionario estratto da una popolazione $\mathcal{N}(\mu, \sigma^2)$
e σ^2 non è nota

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

H_0	H_1	Statistica del test, X_{ts}	Si rifiuta H_0 con livello di significatività α se...	p -dei-dati se $X_{ts} = t$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\frac{\bar{X} - \mu_0}{S/\sqrt{n}}$	$\dots X_{ts} > t_{\frac{\alpha}{2}, n-1}$	$2P(T_{n-1} > t)$
$\mu \leq \mu_0$	$\mu > \mu_0$	$\frac{\bar{X} - \mu_0}{S/\sqrt{n}}$	$\dots X_{ts} > t_{\alpha, n-1}$	$P(T_{n-1} > t)$
$\mu \geq \mu_0$	$\mu < \mu_0$	$\frac{\bar{X} - \mu_0}{S/\sqrt{n}}$	$\dots X_{ts} < -t_{\alpha, n-1}$	$P(T_{n-1} < t)$

Nota: T_{n-1} ha distribuzione t con $n - 1$ gradi di libertà. Inoltre $P(T_{n-1} > t_{\alpha, n-1}) = \alpha$.

7.5 Verifica se due popolazioni hanno la stessa media

Esercizio 1 Probabilità che Oneto dia 30L (Lode)

$$n = 120$$

$$\sum_i^{120} x_i = 18$$

$$\hat{P} = \frac{18}{120} = 0.15 \rightarrow 15\%$$

Esercizio 2 N studenti da 30 e lode

$$n_1 = 18 \leftarrow \text{Oneto}$$

$$n_2 = 20 \leftarrow \text{Anguita}$$

$$n_{1,2} = 10 \leftarrow 30L \text{ sia con Oneto che con Anguita}$$

$N = ?$ Studenti da **30 e Lode**

$$\hat{P}_1 \approx \frac{n_1}{n_2}$$

$$\hat{P}_1 \approx \frac{n_1}{N}$$

$$\frac{n_{1,2}}{n_2} = \frac{n_1}{N}$$

$$\Rightarrow N = \frac{n_1 \cdot n_2}{n_{1,2}} \rightarrow \frac{18 \cdot 20}{10} = 36$$

Esercizio 3 Stima del numero di incidenti medio in auto $n = 10$

$$x_1 = \{4, 0, 6, 5, 2, 1, 2, 0, 4, 3\}$$

$$\hat{\lambda} = \frac{\sum_i x_i}{n} = \frac{27}{10} = 2.7$$

$$P\{x \leq 2\} = e^{-2.7} \left(\frac{2.7^0}{0!} + \frac{2.7^1}{1!} + \frac{2.7^2}{2!} \right) \approx .4936 \rightarrow 49.36\%$$

Probabilità che non ci siano più di **2 incidenti**

Esercizio gaussiana primo

$$x_1 = 1.7$$

$$x_2 = 1.82$$

$$x_3 = 1.73$$

$$x_4 = 1.7$$

$$x_5 = 1.8$$

$$\hat{\mu} = \frac{\sum_i x_i}{n} = \frac{1.7 + 1.82 + 1.73 + 1.7 + 1.8}{5} = 1.75$$

$$\hat{\sigma} = \sqrt{\frac{0.05^2 + 0.07^2 + 0.02^2 + 0.05^2 + 0.05^2}{5}} \approx 0.051$$

Esempio: Sistema di comunicazione $\sigma^2 = 4$ $n = 9$

$$x_1 = \{5.85, 12, 15, 7, 9, 7.5, 6, 5, 10.5\}$$

$$\hat{\mu} = \frac{1}{n} \sum_i^n x_i = \frac{1}{9} \sum_i^n x_i = \frac{81}{9} = 9$$

$$P\left(9 - 1.96 \frac{\sigma}{\sqrt{m}} < \mu < 9 + 1.96 \frac{\sigma}{\sqrt{m}}\right) = 0.95$$

$$p\left(9 - 1.96 \frac{2}{3} < \mu < 9 + 1.96 \frac{2}{3}\right) = 0.95$$

$\rightarrow [7.693, 10.31] \rightarrow \mu$ si trova tra 7.693 e 10.31

In generale Prob = $1 - \alpha$

$$(\bar{x} - z_\alpha \frac{\sigma}{\sqrt{n}}, \bar{x} + z_\alpha \frac{\sigma}{\sqrt{n}}) \rightarrow \text{Si rileva dalle tavole}$$

7.6 Intervalli di confidenza (Unilaterali)

7.7 Esempio:

Pesca stagionale dei salmoni (*Fisso intervallo* -> *trovo n*)

Ad ogni stagione il peso medio dei salmoni è diverso ma $\sigma = 0.3$ Kg

Intervallo di confidenza al 95%, quindi $\alpha = 0.05$

$$(\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}})$$

$$1.96 \frac{\sigma}{\sqrt{n}} \geq 0.1 \quad \sqrt{n} \geq \frac{1.96}{0.1} \sigma$$

$$n \geq \left(\frac{1.96}{0.1} 0.3 \right)^2 = 5.88^2 \approx 34.6 \leftarrow \text{salmoni}$$

7.8 Intervallo di confidenza

con *media* e *varianza* **incognite**

Esempio: Trasimittente (μ) e ricevitore ($\mu + \text{rumore}$)

$$95\%(7.69, 10.31) \quad \hat{\mu} = 9, \sigma^2 = 4$$

$$X_i \{5, 8.5, 12, 15, 7, 9, 7.5, 6.5, 10.5\}$$

$$\hat{\mu} = \bar{X} = \frac{1}{9} \sum_i^n X_i = \frac{81}{9} = 9$$

$$s^2 = \frac{1}{8} \sum_i (X_i^2 - 9.81) \approx 9.5 \quad s = 3.082$$

$$\mu \in \left(9 - 2.306 \frac{3.082}{3}, 9 + 2.306 \frac{3.082}{3} \right) = (6.63, 11.37)$$

Si può dimostrare che $T_{\frac{\alpha}{2}, n-1} \mathbb{E}[S] \geq z_{\alpha} \sigma$

7.9 Intervallo di confidenza nella varianza

Esempio: Laminatoio $n = 4$ $X_i = \{0.123, 0.124, 0.126, 0.12\}$ spessore in **mm**

Svolgimento

$$\frac{1}{4} \sum_i^4 X_i = \frac{0.493}{4} = 0.12325$$

$$\frac{1}{4-1} \sum_i^4 (X_i - 0.12325)^2 = 1.875 \cdot 10^{-5}$$

$$\sigma^2 \in \left(\frac{s^2(n-1)}{9.348}, \frac{s^2(n-1)}{0.216} \right)$$

Dove **9.348** e **0.216** sono ricavati dalle tabelle

Facciamo la radice:

$$\sigma \in (0.0014, 0.0093) \rightarrow 95\%$$

7.10 Intervallo di confidenza

della differenza di due medie:

N campioni

$$X_i \sim \mathcal{N}(\mu_1, \sigma_1^2)$$

$$\bar{X} = \frac{1}{n} \sum_i^n X_i$$

M campioni

$$Y_i \sim \mathcal{N}(\mu_2, \sigma_2^2)$$

$$\bar{Y} = \frac{1}{m} \sum_i^m Y_i$$

$$\bar{X} - \bar{Y} \sim \mathcal{N}\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right)$$

$$\mathcal{N}(0,1) \sim \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}}$$

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

Se σ_1^2, σ_2^2 non sono note:

$$S_1^2 = \frac{1}{n-1} \sum_i^n (X_i - \bar{X})$$

$$s_2^2 = \frac{1}{m-1} \sum_i^m (X_i - \bar{Y})$$

$$(n-1) \frac{S_1^2}{\sigma_1^2} \sim \chi_{n-1}^2$$

$$(n-1) \frac{s_2^2}{\sigma_2^2} \sim \chi_{n-1}^2$$

Possiamo andare avanti solo se $\sigma_1^2 = \sigma_2^2 = \sigma^2$

$$(n-1) \frac{s_1^2}{\sigma^2} + (n-1) \frac{s_2^2}{\sigma^2} \sim \chi_{n+m-2}^2$$

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\sigma^2(\frac{1}{n} + \frac{1}{m})}} \rightarrow \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{S_p(\frac{1}{n} + \frac{1}{m})}} \\ \sim \mathcal{N}(0,1) \qquad \qquad \qquad \sim T_{n+m-2}$$

$$S_p = \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}$$

Se σ sono ignote ma uguali

$$\mu_1 - \mu_2 \in (\bar{X} - \bar{Y} - T_{\frac{\alpha}{2}, n+m-2} \sqrt{s^2(\frac{1}{n} + \frac{1}{m})})$$

$$\bar{X} - \bar{Y} + T_{\frac{\alpha}{2}, n+m-2} \sqrt{s^2(\frac{1}{n} + \frac{1}{m})}$$

7.11 Intervallo di previsione

Esempio smartwatch contapassi $n = 7$

<i>LUN</i>	6922	X_1	<i>GIO</i>	7432	X_4
<i>MAR</i>	5333	X_2	<i>VEN</i>	6252	X_5
<i>MER</i>	7420	X_3	<i>SAB</i>	7005	X_6

$$DOM \quad 6752 \quad X_7$$

$$\bar{X}_n = \frac{1}{n} \sum_i^m X_i = \frac{47016}{7} \approx 6717$$

$$1 - \alpha = 95\% \quad \alpha = 5\%$$

$$t_{0.0025,6} = 2.997$$

$$S_n = \sqrt{S_n^2} = 7.333.8$$

$$x_{n+1} \in (6717 - 2.447 \cdot 733397 \sqrt{1 + \frac{1}{7}}, 6717 + 2.447 \cdot 733397 \sqrt{1 + \frac{1}{7}})$$

$$X_{n+1} \in (9796, 8637) \mu \in (6037, 7396)$$

7.12 Qualità di uno stimatore

$$X = X_1 \dots X_n \quad \theta \leftarrow \text{parametro} \quad d(x) \leftarrow \text{stimatore di } \theta$$

$$(d(x) - \theta)^2 \quad \mathbb{E}[(d(x) - \theta)^2]$$

Errore Quadratico (*misura della qualità*) Errore Quadratico Medio (*M.S.E*)

Rischio $r(d, \theta) = \mathbb{E}[(d - \theta)^2]$ Lo stimatore "ottimo" sarà quello con il rischio minimo

-> d con r minimo θ

Esempio $d^*(x) = 4$ se $\theta = 4 \Rightarrow d^* = \text{stimatore ottimo}$ (per tutti gli altri valori non va)

7.13 Proprietà di uno stimatore

Def: $b_\theta(d) = \mathbb{E}[d] - \theta \rightarrow$ bias o polarizzazione Uno stimatore non è **polarizzato** se $b_\theta(d) = 0$

Esempio : $X_1 \dots X_n$ θ media

$$d_1(X_1 \dots X_n) = X_1$$

$$d_2(X_1 \dots X_n) = \frac{X_1 + X_2}{2}$$

$$d_3(X_1 \dots X_n) = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Tutti questi sono **unbiased**

7.14 Stimatore unbaieseio

$$r(d, \theta) = \mathbb{E}[(d(x) - \theta)^2] = \mathbb{E}[(d(x) - \mathbb{E}[d(x)])^2] = \text{Var}(d)$$

tra gli stimatori non polarizzati di ottimo è quello con la varianza minima

7.15 Valutazione di uno stimatore

$$X = X_1 \dots X_n \quad \theta = ?$$

Dove θ è un *parametro* e $d(x)$ è uno *stimatore* di θ

$$r(d, \theta) (\text{mse}) \text{ rischio} \quad b_\theta(d) = \mathbb{E}[d] - \theta$$

$$\text{se } b_\theta(d) = 0 \Rightarrow r(d, \theta) = \text{Var}(d)$$

$$\text{se } b_\theta(d) \neq 0 ? \quad r(d, \theta) = ?$$

$$r(d, \theta) = \mathbb{E}[(d(x) - \theta)^2] = \mathbb{E}[(d(x) - \mathbb{E}[d] + \mathbb{E}[d] - \theta)^2]$$

$$= \mathbb{E}[(d - \mathbb{E}[d])^2 + (\mathbb{E}[d] - \theta)^2 - 2(d - \mathbb{E}[d])(\mathbb{E}[d] - \theta)]$$

$$= \mathbb{E}[(d - \mathbb{E}[d])^2] + \mathbb{E}[(\mathbb{E}[d] - \theta)^2] - 2(\mathbb{E}[d] - \theta) \cdot \mathbb{E}[(d - \mathbb{E}[d])]$$

$$r(d, \theta) = \mathbb{E}[(d - \mathbb{E}[d])^2] + \mathbb{E}[(\mathbb{E}[d] - \theta)^2]$$

$$= \text{Var}(d) + b_\theta(d)^2 \leftarrow \text{bias}^2$$

7.16 Esempio:

Stimatore della media di una *distribuzione uniforme*

$$\begin{array}{ll} \mathbb{E}[X_i] = \theta/2 & d_1 = 2\frac{1}{n} \sum_i^n X_i \\ X_1, X_2 \dots X_n & d_2 = \max X_i \end{array}$$

$$\begin{aligned} d_1 : \mathbb{E}[d_1] &= \frac{2}{n} \sum_i \mathbb{E}[X_i] = \frac{2}{n} n \frac{\theta}{2} = \theta \\ r(d_1, \theta) &= \text{Var}(d_1) = \frac{4}{n^2} n \text{Var}(X_i) = \frac{4}{n} \frac{\theta^2}{12} = \frac{\theta^2}{3n} \Leftarrow \text{Unbiased} \end{aligned}$$

$$\begin{aligned} F_2(x) &= P_r\{d_2(x) \leq x\} = P_r\{\max X_1 \leq x\} \\ &= P_r\{X_1 \leq x, \forall i \in 1\} = \prod_{i=1}^n P_r\{X_i \leq x\} = \left(\frac{x}{\theta}\right)^n \\ f_2(x) &= \frac{d}{dx} F_2(x) = n \frac{x^{n-1}}{\theta^n} \quad x \leq \theta \end{aligned}$$

$$\mathbb{E}[d_2] = \int_0^\theta x f_x(x) dx = \int_0^\theta \frac{n}{\theta^n} x^n dx = \frac{n}{\theta^n} \left[\frac{x^{n+1}}{n+1} \Big|_0^\theta \right] = \frac{n}{n+1} \theta$$

$$\mathbb{E}[d_2^2] = \frac{n}{\theta^n} \int_0^\theta x^2 f(x) dx = \frac{n}{\theta^n} \int_0^\theta x^{n+1} dx = \frac{n}{\theta^n} \left[\frac{x^{n+2}}{n+2} \Big|_0^\theta \right] = \frac{n}{n+2} \theta^2$$

$$\text{Var}(d_2) = \mathbb{E}[d_2^2] - \mathbb{E}[d_2]^2 = \frac{n}{n+2} \theta^2 - \frac{n^2}{(n+1)^2} \theta^2 = \frac{n}{(n+2)(n+1)^2} \theta^2$$

$$r(d_2, \theta) = \text{Var}(d_2) + (\mathbb{E}[d_2] - \theta)^2 = \frac{2 \cdot \theta^2}{(n+1)(n+2)}$$

$$n \geq 4 \quad r(d_2, \theta) < r(d_1, \theta) \quad d_3 = \frac{n+1}{n} d_2$$

In sintesi

$$r(d_1, \theta) = \frac{\theta^2}{3n} \Leftarrow \text{Unbiased}$$

$$r(d_2, \theta) = \frac{2\theta^2}{(n+1)(n+2)} \Leftarrow \text{Biased}$$

$$r(d_3, \theta) = \frac{\theta^2}{n^2 + 2n} \Leftarrow \text{Unbiased}$$

$$r(d_4, \theta) = \frac{\theta^2}{(n+1)^2} \Leftarrow \text{Biased}$$

8 Test di ipotesi

Ipotesi: Affermazione rispetto a uno o più parametri di una distribuzione Ipotesi da confutare: H_0 (ipotesi nulla)

Esempio :

$$\begin{aligned} X_1 \dots X_n &\sim \mathcal{N}(\mu, \sigma^2) \\ H_0 &: \mu = 0 \\ H_a &: \mu \neq 0 \end{aligned} \tag{6}$$

Diamo per scontato che l'ipotesi sia **vera**
Dobbiamo cercare di *confutarla*

Definizione Regione critica tale che:

$(X_1 \dots X_n) \in C \rightarrow H_0$ è rifiutata

$(X_1 \dots X_n) \notin C \rightarrow H_0$ è accettata

α = Livello di **significatività** del test ($\alpha = 10\%, 5\% \dots$)

Procedimento

- Fisso α
- Suppongo che α sia vera
- calcolo stima di μ
- verifico che non sia "*troppo distante*"

$$X_1 \dots X_n \sim \mathcal{N}(\mu, \sigma^2)$$

$$H_0 : \mu = \mu_0 \quad H_a : \mu \neq \mu_0$$

$$\bar{X} = \frac{1}{n} \sum_i X_i$$

Regione critica $\{(X_1 \dots X_n) : |\bar{X} - \mu_0| > c\}$

$$P_{r_{\mu_0}} \{|\bar{X} - \mu_0| > c\} = \alpha$$

$$P_{r_{\mu_0}} \left\{ \frac{|\bar{X} - \mu_0|}{\frac{\sigma}{\sqrt{n}}} > \frac{c}{\frac{\sigma}{\sqrt{n}}} \right\} = \alpha$$

$$P_{r_{\mu_0}} \{|z| > z_\alpha\} = \alpha$$

Esempio (5 transmissioni)

$$n = 5$$

$$\bar{X} = 9.5$$

$$H_0 : \mu = 8$$

$$\alpha = 5\%$$

Ipotizzando che H_0 sia vera:

$$\frac{|\bar{X} - \mu|}{\frac{\sigma}{\sqrt{n}}} = \frac{|9.5 - 8|}{\frac{2}{\sqrt{5}}} \approx 1.68$$

Se:

$$\alpha = P_r(\text{rifiuto } H_0 / H_0 \text{ vera})$$

$\alpha \uparrow$ più "*facile*" rifiutare l'ipotesi

$\alpha \downarrow$ più "*difficile*" rifiutare l'ipotesi

8.1 Metodologia alternativa

$$Ts = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \rightarrow \text{Statistica di test}$$

P -value = **probabilità** di ottenere un valore più "anomalo" di quello osservato

Esempio: $X_i \sim \mathcal{N}(\mu, 4)$

$$n = 5$$

$$H_0 : \mu = 8$$

$$\bar{X} = 8.5$$

$$H_a : \mu \neq 8$$

$$\frac{|\bar{X} - \mu_0|}{\frac{\sigma}{\sqrt{n}}} = \frac{|8.5 - 8|}{\frac{2}{\sqrt{5}}} = \frac{\sqrt{5}}{2} 0.5 \approx 0.559$$

$$P\{|z| > 0.559\} = 2P\{z > 0.559\} \approx 2 \cdot 0.288 = 0.579 \rightarrow P\text{-value}$$

Se $\bar{X} = 11.5$:

$$\frac{|\bar{X} - \mu_0|}{\frac{\sigma}{\sqrt{n}}} = \frac{|11.5 - 8|}{\frac{2}{\sqrt{5}}} \approx 3.913$$

$$P\{|z| > 3.913\} = 2P\{z > 3.913\} \leq 0.00005 \rightarrow \underline{\text{Rifiuto ipotesi } H_0}$$

8.2 Test di Hp unilaterale

$$H_0 : \mu = \mu_0 (\mu \leq \mu_0)$$

$$H_a : \mu > \mu_0$$

$$C = \{(X_1 \dots n) \cdot \bar{X} - \mu_0 > c\}$$

$$P_{r_{\mu_0}}\{\bar{X} - \mu_0 > c\} = P_r\left\{\frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} > \frac{c}{\frac{\sigma}{\sqrt{n}}}\right\} = P_{r_{\mu_0}}\{z > z_a\} = \alpha$$

$$\text{Statistica test } \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \leq z_\alpha \text{ accetto}$$

8.3 Test di ipotesi

H_0	H_a	TS	Livello α	P - Value
$\mu = \mu_0$	$\mu \neq \mu_0$	$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{1}}$	Rifiuto H_0 se $TS > \frac{z_\alpha}{2}$	$2P(z \geq TS)$

Altre ipotesi :

H_0	H_a	TS	Livello α	P - Value
$\mu < \mu_0$	$\mu > \mu_0$	$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{2}}$	$H_0 \quad z_\alpha > TS$	$P(z \geq TS)$
$\mu \geq \mu_0$	$\mu < \mu_0$	//	$H_0 \quad z_\alpha < -TS$	$P(z \leq TS)$

8.4 Uguaglianza media di due popolazioni

$$\begin{array}{ll} X_1 \dots X_n \sim \mathcal{N}(\mu_1, \sigma_2^2) & Y_i \dots Y_m \sim \mathcal{N}(\mu_2, \sigma_2^2) \\ \bar{X} = \frac{1}{n} \sum_i X_i & \bar{Y} = \frac{1}{m} \sum_i Y_i \\ S_x^2 = \frac{1}{n-1} \sum_i (X_i - \bar{X})^2 & S_y^2 = \frac{1}{m-1} \sum_i (Y_i - \bar{Y})^2 \end{array}$$

$$S_p^2 = \frac{(n-1)S_x^2 + (m-1)S_y^2}{n+m}$$

H_0	H_a	TS	Livello α	P - Value
$\mu_1 = \mu_2$	$\mu \neq \mu_2$	$\frac{\bar{X} - \bar{Y}}{\sqrt{\sigma_1^2/n + \sigma_2^2/m}}$		
$\mu_1 = \mu_2$	$\mu \neq \mu_2$	$\frac{\bar{X} - \bar{Y}}{\sqrt{S_p^2(\frac{1}{n} + \frac{1}{m})}}$	ref. $ TS > z_{\frac{\alpha}{2}}$	$2P(z \geq TS)$
$\mu_1 = \mu_2$	$\mu \neq \mu_2$	$S_i \in \text{T-student}$		

4) T-test per coppie di dati Se X_1 e X_2 **NON** sono indipendenti

$$W_i = X_i - Y_i \quad \begin{array}{l} X_i \sim \mathcal{N}(\mu_1, \sigma_1^2) \\ Y_i \sim \mathcal{N}(\mu_2, \sigma_2^2) \end{array}$$

ES Manutenzione (n guasti) tagliand

$$H_0 : \mu_a - \mu_b \geq 0 \quad \bar{W} = \frac{1}{5}(-7.5 + 2.5 - 2.5 - 3.5 - 1.5) = -2.5$$

$$S_W^2 = \frac{1}{4}(W_i - \bar{W})^2 = 13$$

$$Ts = \frac{\frac{\bar{W}}{S_W}}{\sqrt{n}} = \frac{\frac{-2.5}{\sqrt{13}}}{\sqrt{5}} = 1.55$$

$$Pr\{T_{n-1} \leq Ts\} = \{T_4 \leq Ts\}$$

5) Test sulla varianza

$$H_0 : \sigma^2 = \sigma_0^2 \quad H_a : \sigma^2 \neq \sigma_0^2$$

$$\frac{(n-1)}{\sigma_0^2} \sim \chi_{n-1}^2$$

$$Pr\{\chi_{1-\frac{\alpha}{2}, n-1}^2 \leq \frac{(n-1)s^2}{\sigma_0^2} \leq \chi_{\frac{\alpha}{2}, n-1}^2\} = 1 - \alpha$$

Uguaglianza di varianza :

$$X_1 \dots X_n \quad Y_1 \dots Y_n \quad H_0 : \sigma_x^2 = \sigma_y^2 \quad H_a : \sigma_x^2 \neq \sigma_y^2$$

$$S_x^2 - S_y^2 \quad Ts = \frac{\frac{S_x^2}{\sigma_x^2}}{\frac{S_y^2}{\sigma_y^2}} = \frac{S_x^2}{S_y^2}$$

$$\frac{S_x^2}{S_y^2} \sim F_{n-1, m-1} \quad Pr\{F_{1-\frac{\alpha}{2}, n-1, m-1} \leq \frac{S_x^2}{S_y^2} \leq F_{\frac{\alpha}{2}, n-1, m-1}\}$$

Non rifiuto se soddisfa la disuguaglianza

Test parametro Bernoulli (Var discrete.)

$$H_0 : p \leq p_0 \quad H_a : p > p_0$$

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad n \text{ campioni (Bernoulli)}$$

Binomiale \sim **Gaussiana** (quando n è grande)

X n eventi favorevoli

$$\mathbb{E}[X] = np \quad \text{Var}(X) = np(1-p) \quad \mathcal{N}(np, np(1-p))$$

Esempio Difetti di fabbricazione:

$$n = 300 \quad H_0: p \leq p_0 \quad p_0 = 2\%$$

$X = 10$ n difetti

$$\frac{X - np}{\sqrt{np_0(1-p_0)}} = \frac{10 - 300 \cdot 0.02}{\sqrt{300 \cdot 0.02 \cdot 0.98}} = 1.65$$

$$\Pr\{z > 1.65\} = 0.0495$$

8.5 Modelli previsionali

8.5.1 Modelli di regressione previsionale

$$Y_i = \alpha + \beta x_i + e_i \quad e_i \sim \mathcal{N}(0, \sigma^2)$$

Problema $\{x_i, y_i\}_{i=1}^n \quad \alpha, \beta = ?$

Sum of square $\rightarrow SS$

$$SS = \sum_i^n (y_i - \alpha - \beta x_i)^2 \text{ Dove } B \text{ e } A \rightarrow \text{var aleatoria}$$

$$\begin{cases} \frac{dSS}{dA} = -2 \sum_i^n (y_i - A - Bx_i) = 0 \\ \frac{dSS}{dB} = -2 \sum_i^n (y_i - A - Bx_i)^2 x_i = 0 \end{cases}$$

$$\begin{cases} \sum_i^n y_i = nA + B \sum_i x_i \\ \sum_i^n x_i y_i = n \sum_i^n x_i + B \sum_i^n x_i^2 \end{cases}$$

$$A = \frac{1}{n} \sum_i y_i - B \frac{1}{n} \sum_i x_i = \bar{y} - \beta \bar{x}$$

8.5.2 Regressione lineare

$$y = \alpha + \beta x \quad e \sim (0, 1) \quad y_i = A + \beta x_i$$

$$\mathbb{E}[B] = \beta \quad \mathbb{E}[A] = \alpha$$

$$Var[B] = \frac{\sigma^2}{\sum_i x_i^2 - n\bar{x}^2} \quad Var[A] = \frac{\sigma^2 \sum_i x_i^2}{n(\sum_i x_i^2 - n\bar{x}^2)}$$

$$SS_R = \sum_i (y_i - (A + \beta x_i))^2 \quad (\text{Somma dei quadrati dei residui})$$

$$\frac{SS_R}{\sigma^2} \sim \chi_{n-2}^2$$

$$\mathbb{E}\left[\frac{SS_R}{\sigma^2}\right] = n - 2$$

$$\mathbb{E}\left[\frac{SS_R}{n-2}\right] = \sigma^2$$

MLE :

$$f_{y_1 \dots y_n}(y_1 \dots y_n) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\sum (y_i - (\alpha + \beta x_i))^2 / 2\sigma^2}$$

$$\text{MSE} = \text{MLE}$$

Notazione :

$$\begin{aligned}
 S_{xy} &= \sum_i (x_i - \bar{x})(y_i - \bar{y}) = \dots = \sum_i x_i y_i - n\bar{x}\bar{y} \\
 S_{xx} &= \sum_i (x_i - \bar{x})^2 = \dots = \sum_i x_i^2 - n\bar{x}^2 \\
 S_{yy} &= \sum_i (y_i - \bar{y})^2 = \dots = \sum_i y_i^2 - n\bar{y}^2
 \end{aligned} \tag{7}$$

$$\begin{array}{ll}
 S_{xy} \text{ (Dispersione di x e y)} & S_{xy} \text{ (Dispersione di x)} \\
 & S_{xy} \text{ (Dispersione di y)}
 \end{array}$$

$$A = \bar{y} - B\bar{x} \qquad B = \frac{S_{xy}}{S_{xx}} \qquad SS_R = \frac{S_{xx}S_{yy} - S_{xy}^2}{S_{xx}}$$

$$\text{Inferenza su } \beta \quad \frac{B - \beta}{\sqrt{\frac{\sigma^2}{S_{xx}}}} \sim \mathcal{N}(0, 1) \qquad \frac{SS_R}{\sigma^2} \sim \chi_{n-2}^2$$

$$\frac{\frac{B - \beta}{\sqrt{\frac{\sigma^2}{S_{xx}}}}}{-\sqrt{\frac{SS_R}{j^2(n-2)}}} \sim t_{2-2}$$

$$\sqrt{\frac{(n-2)S_{xx}}{SS_R}} (B - \beta) \sim t_{n-2}$$

$$\beta \in B \pm \sqrt{\frac{SS_R}{(n-2)S_{xx}}} \quad t_{\frac{\alpha}{2}, n-2} \rightarrow \text{Livello di confidenza}$$

Inferenza su α $\frac{A - \alpha}{\sqrt{\frac{\sigma^2 \sum x_i^2}{n S_{xx}}}} \sim \mathcal{N}(0, 1) \quad \frac{SS_R}{\sigma^2} \sim \chi_{n-2}^2$

coefficiente della retta:

$$\alpha \in A \pm \frac{SS_R \sum x_i^2}{\sqrt{n(n-2)S_{xx}}} \sim t_{\frac{\alpha}{2}, n-2} \rightarrow \text{Livello di confidenza}$$

Interferenza su $\alpha + \beta x_0$:

$$\mathbb{E}[A + Bx_0] = \mathbb{E}[A] + x_0 \mathbb{E}[B] = \alpha + \beta x_0$$

$$Var(A + Bx_0) = \dots = \sigma^2 \left[\frac{1}{n} + \frac{(\bar{x} - x_0)^2}{S_{xx}} \right]$$

Distribuzione $A + Bx_0$?

$$A + Bx_0 \sim \mathcal{N}(\alpha + \beta x_0, \sigma^2 \left[\frac{1}{n} + \frac{(\bar{x} - x_0)^2}{S_{xx}} \right])$$

Stima di $\alpha + \beta x_0$

$$\frac{A + Bx_0 - (\alpha + \beta x_0)}{\sqrt{\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right) \left(\frac{SS_R}{n-2} \right)}} \sim t_{n-2}$$

$$\alpha + \beta x_0 \in A + Bx_0 \pm t_{\frac{\alpha}{2}, n-2} \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}} \left(\frac{SS_R}{n-2} \right)$$

Piccolo se i punti sono vicini alla media

8.5.3 Regressione Lineare (e non)

$\{x_i, y_i\}_{i=1}^2 \leftarrow$ punti stocastici

Inferenza $\alpha + \beta x_0 = \mathbb{E}[y] \rightarrow$ non so niente del valore della y in quel punto
 Inferenza $y_0 = y(x_0)\theta$

$$\alpha + \beta x_0 \in A + Bx_0 \pm t_{\frac{\alpha}{2}, n-2} \sqrt{\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right) \frac{SS_R}{n-2}}$$

$\alpha + \beta \mathbf{x}_0 \rightarrow$ Il punto x_0 che sta sulla retta $\alpha + \beta x_0$

Inferenza $y_0 = y(x_0) \rightarrow$ predittivo

$$y \sim \mathcal{N}(\alpha + \beta x_0, \sigma^2)$$

$$A + Bx_0 \sim \mathcal{N}\left(\alpha + \beta x_0, \sigma^2 \left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right)\right)$$

$$y_0 - (A + Bx_0) \sim \mathcal{N}\left(\sigma, \sigma^2 \left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right)\right)$$

$$y_0 = y(x_0) = A + Bx_0 \pm t_{\frac{\alpha}{2}, n-2} \sqrt{\left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right) \frac{SS_R}{n-2}}$$

Coefficiente di determinazione

Definizione: La verifica dei miei valori

Formula generica: $S_{yy} = \sum_{i=1}^n (y_i - \bar{y})^2 \rightarrow$ dispersione di y
 La dispersione è data da due fattori:

- Retta (regressione)

- Rumore

$$SS_R = \sum_{i=1}^n (y_i - (A + Bx_i))^2 \rightarrow \text{Dipende dalla porzione non spiegata della retta}$$

Utilizzo coefficienti di determinazione:

$$R^2 = \frac{S_{yy} - SS_R}{S_{yy}} = 1 - \frac{SS_R}{S_{yy}} \quad 0 \leq R^2 \leq 1$$

Se $R^2 = 1$ la dispersione è data solo dalla retta (*regression*)

Se $R^2 = 0$ la dispersione è data solo dal *rumore*

La retta è migliore più R^2 è vicino a **1**

Coefficiente di correlazione

Formula generica:

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

$$r^2 = \frac{S_{xy}^2}{S_{xx}S_{yy}} = \dots = 1 - \frac{SS_R}{S_{yy}} \rightarrow \text{Dimostrazione matematica di } R^2$$

Analisi dei residui $y - (A + Bx) \rightarrow$ verifico tutti gli errori residui
Per la non linearità

Trasformazione lineare

$$W(t) = ce^{-dt}$$

Dove c e $-dt$ sono *parametri*

$\log(W(t)) = \log(c) - dt \rightarrow$ Prob. soluzione al non lineare $y = \alpha + \beta x$

Rimedio al caso eteroschedastico :

$y_i = \alpha + \beta x_i + e_i \quad e_i \sim \mathcal{N}(0, \sigma_i^2) \rightarrow$ errore in crescita x

$$Var(e_i) = \frac{\sigma^2}{W_i} \sum W_i (y - (A + Bx_0))^2$$

- Regressione lineare multipla

$$- y = \beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 \dots \beta_k \mathbf{x}_k + \mathbf{e}$$

$$- \min \sum_i (y_i - (B_0 + B_1 x_{i1} + B_2 x_{i2} + \dots + B_k x_{ik}))^2$$

- Regressione (lineare) polinomiale

$$- y = \beta_0 = \beta_1 x + \beta_2 x^2 + \dots + \beta_k x^k + e$$

$$- \{ \underline{x_i}, y_i \}_{i=1}^n$$

$$\frac{d}{dB_0} = 0 = \sum_i (y_i - 1 - B_1 x_{i1} + B_2 x_{i2} + \dots + B_k x_{ik})$$

$$\frac{d}{dB_1} = 0 = \sum_i x_{i1}(y_i - B_0 - B_1 x_{i1} + B_2 x_{i2} + \dots + B_k x_{ik})$$

$$x^t x \underline{\beta} = x^t \underline{y} \implies \underline{\beta} = (x^x x)^{-1} x^t \underline{y}$$

AN.O.VA (analysis of variance)

Analisi delle varianze / estensione del test di ipotesi sulle medie

Esempio voti medi degli anni scolastici

Anno.

2020-2021 lockdown

2021-2022 lockdown parziale

2022-2023 presenza

Voti medi.

μ_a

μ_b

μ_c

$$H_0 : \mu_a = \mu_b = \mu_c$$

1) stimatore di σ^2 :

$$\sum_{i=1} \sum_j \frac{(x_{ij} - \mathbb{E}[x_{ij}])^2}{\sigma^2} = \sum_{i=1} \sum_{j=1} \frac{(x_{ij} - \mu_i)^2}{\sigma^2} \sim \chi_{m \cdot n}^2$$

$$SS_W = \sum_{i=1} \sum_{j=1} \frac{(x_{ij} - x_{i*})^2}{\sigma^2} \sim \chi_{m \cdot n - m}^2$$

$$\mathbb{E}\left[\frac{SS_w}{\sigma^2}\right] = n \cdot m - m$$

$$\mathbb{E}\left[\frac{SS_w}{nm-m}\right] = \sigma^2 \quad \text{stimatore 1}$$

2) stimatore di σ^2 supponendo $\mu_1 = \mu_2 = \mu_3 = \dots = \mu_m = \mu$

$$n \sum_{i=1}^m \frac{(x_{i*} - \mu)^2}{\sigma^2} \sim \chi_m^2 \quad x_{**} = \frac{\sum_{i=1}^m \sum_{j=1}^n x_{ij}}{m \cdot n}$$

$$SS_b = n \sum_{i=1}^m (x_{i*} - x_{**})^2 \sim \chi_{m-1}^2$$

$$\mathbb{E}\left[\frac{SS_b}{m-1}\right] = \sigma^2 \rightarrow \text{Stimatore 2}$$

Verifico stimatori :

$$Ts = \frac{SS_b/m - 1}{SS_W/nm - m} \rightarrow \text{intorno a 1 va bene}$$

F Distribution: $F_{m-1, mn-m, \alpha}$

ANOVA :

Se i gruppi sono uguali : $n \text{ camp} = n \cdot m$

Se sono diversi : $n \text{ camp} = \sum_i n_i$

Life testing (Misura di affidabilità)

$$x \geq 0 \mid \text{tempo di vita} \quad \lambda(t) = \frac{f(t)}{1-F(t)}$$

$f(t)$ = Densità di popolazione

$\lambda(t)$ = Intensità di rottura (failure rate)

$$\begin{aligned} P(x \in (t, t + \Delta t) \mid x > t) &= \frac{P(x \in (t, t + \Delta t), x > t)}{P(x > t)} \\ &= \frac{P(x \in (t, t + \Delta t))}{P(x > t)} \\ &\approx \frac{F(t)\Delta t}{1 - F(t)} \end{aligned}$$

Intensità di rottura

Definizione: Densità condizionale di probabilità che un oggetto funzionante almeno fino a t si guasti "subito dopo"

Formula generica:

$$\lambda(t) = \frac{F(t)}{1 - F(t)}$$

Se la distribuzione è esponenziale:

$$\lambda(t) = \frac{\lambda e^{-\lambda t}}{1 - 1 + e^{-\lambda t}} = \lambda \rightarrow \text{dove } \lambda \text{ è una costante}$$

Proprietà $\lambda(t) \Rightarrow F(t)$

$$\lambda(s) = \frac{f(s)}{1 - F(s)} = \frac{F'(s)}{1 - F(s)} = \frac{d}{dS} [-\log(1 - F(s))]$$

$$\int_0^t \lambda(s) ds = -\log(1 - F(s)) + \log(1 - F(s)) = 1 - F(t) = e^{-\int_0^t \lambda(s) ds}$$

Esempio Tasso di mortalità di un fumatore (λ_s) e di un non fumatore (λ_n)

$$\lambda_s(t) = 2\lambda_n(t)$$

$$\begin{aligned}
&= P(\text{Non fumatore di età } \mathbf{A} \text{ vive fino a } \mathbf{B}) \\
&= P(\text{Non fumatore vive fino a } \mathbf{B} \mid \text{è vissuto fino } \mathbf{A}) \\
&= \frac{P(\text{Non fumatore viva fino a } \mathbf{B})}{P(\text{Non fumatore viva fino a } \mathbf{A})} \\
&= \frac{1 - F_N(B)}{1 - F_N(A)} \\
&= \frac{e^{-\int_0^B \lambda(t) dt}}{e^{-\int_0^A \lambda(t) dt}}
\end{aligned}$$

Quindi:

$$P(\text{Non fumatore di età } \mathbf{A} \text{ vive fino a } \mathbf{B}) = e^{-\int_A^B \lambda(t) dt}$$

Per i non fumatori invece:

$$P(\text{Fumatore di età } \mathbf{A} \text{ vive fino a } \mathbf{B}) = e^{-\int_A^B \lambda(t) dt} = P_S$$

Dove $P_S = (P_N)^2 \rightarrow$ quindi la probabilità di sopravvivenza del fumatore è uguale alla probabilità di sopravvivenza del non fumatore *al quadrato*

Probabilità che un non fumatore arrivi ai 60 anni sapendo che è arrivato ai 50:

$$\lambda_N(t) = \frac{1}{20} \quad 50 \leq t \leq 60$$

$$P_N = e^{-\int_{50}^{60} \frac{1}{20} dt} = e^{-\frac{1}{20}(60-50)} = e^{-\frac{1}{2}} \approx 0.607 \approx 61\%$$

$$P_{\leq} = (e^{-\frac{1}{2}})^2 = e^{-1} \approx 0.368 \approx 37\%$$

Stima di affidabilità N oggetti che si possono guastare *indipendenti* tra di loro

Tempi di vita: $\lambda e^{-\lambda t} \quad \lambda = \frac{1}{\theta} \Rightarrow \frac{1}{\theta} e^{-\frac{t}{\theta}}$

Dati a disposizione: $x_1 \leq x_2 \leq x_3 \leq x_4 = r$ $i_1 = 2, i_2 = 3, i_3 = n, i_4 = 1$

Studio la variabile aleatoria X_i , i_j indica quale *oggetto si è guastato* per j-esimo all'istante x_j

(n-r) non si sono guastati \Rightarrow per questi $X_i > x_r$

$$f_{x_1, x_2 \dots x_r}(x_1, x_2, \dots, x_r) = \prod_j^r \frac{1}{\theta} e^{-\frac{x_j}{\theta}} = \frac{1}{\theta^r} e^{-\frac{\sum_j x_j}{\theta}}$$

Ora per i non guasti:

$$P(X_j > x_j \text{ con } j \notin \{i_1, \dots, i_r\}) = \prod_{r+1}^n (1 - F_{X_j}(x_r)) = \left[1 - (1 - e^{-\frac{x_r}{\theta}})\right]$$

$$\log L = -r \log \theta - \frac{1}{\theta} \left[\sum_i^r x_i + (n-r)x_r \right]$$

$$\frac{d \log}{d \theta} = -\frac{r}{\theta} + \frac{1}{\theta^2} \left[\sum_i^r x_i + (n-r)x_r \right] = 0$$

$$-\theta r + \left[\sum_i^r x_i + (n-r)x_r \right] = 0$$

$$\hat{\theta} = \frac{\sum_i^r x_i + (n-r)x_r}{r} = \frac{t}{r} = \frac{TTT}{r} \rightarrow \text{Total Time Test}$$