PROBABILITÀ

EVENTI INCOMPATIBILI $\Rightarrow$ $A \cap B = \emptyset$ $\quad P(A \cap B) = 0$

$$P(A \cup B) = P(A) + P(B)$$

EVENTI COMPATIBILI $\Rightarrow$ $A \cap B \neq \emptyset$

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

EVENTI S-INDIPENDENTI $\Rightarrow$ $P(A|B) = P(A)$

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

EVENTI S-DIPENDENTI $\Rightarrow$ $P(A|B) \neq P(A)$

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

PROBABILITÀ CONDIZIONATA $\Rightarrow$ $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

REGOLA DI BAYES $\Rightarrow$ $P(A) = P(A \cap C) + P(A \cap \bar{C}) =$

$$= P(A|C) \cdot P(C) + P(A|\bar{C}) \cdot P(\bar{C})$$

TEOREMA DI BAYES $\Rightarrow$ $P(E_1|A) = \dfrac{P(A \cap E_1)}{P(A)} =$

$$= \frac{P(A|E_1) \cdot P(E_1)}{P(A|E_1) \cdot P(E_1) + P(A|E_2) \cdot P(E_2) + \dots}$$

MODELLI DI V.A DISCRETE

MODELLO DI V.A. BERNOULLIANA (p=costante)—>v.a. binaria

- pmf $\begin{cases} P(X=0) = q = 1-p \\ P(X=1) = p \end{cases} \Rightarrow \quad p+q=1$

PARAMETRO DA DEFINIRE
p

- MEDIA $E(X) = p$
- VARIANZA $Var(X) = p \cdot q$

MODELLO BINOMIALE (p=costante)—>n volte il modello bernoulliano

$Y = \sum_{i=1}^{m} X_i \longrightarrow X_i$ = V.a. BERNOULLIANA

- pmf $P(y=y) = C_{m,y}\ p^y\ (1-p)^{m-y}$

PARAMETRI DA DEFINIRE
p , m

- MEDIA $E(x) = m \cdot p$
- VARIANZA $Var(x) = m \cdot p \cdot q$

MODELLO GEOMETRICO (p=costante)—>n di successi prima di un insuc.

- pmf $P(X=x) = p^x (1-p)$ edf $F_x(x) = 1 - p^{x+1}$

  → Probabilità di avere un insuccesso dopo una fila ininterrotta di successi

PARAMETRI DA DEFINIRE
p

- pmf $P(X=x) = (1-p)^x \cdot p$

  → Probabilità di avere un successo dopo una fila ininterrotta di insuccessi

Tempo medio di attesa ***prima*** del primo insuccesso → $E(x) = \frac{p}{1-p}$

Tempo medio di attesa ***per*** osservare il primo insuccesso → $E(x) = \frac{1}{1-p}$

MODELLO BINOMIALE NEGATIVO o DI PASCAL (p=costante)

La v.a. ***Y*** indica il numero di successi prima dell'***m-esimo insuccesso***

$$Y = \sum_{m=1}^{m} X_i \longrightarrow \text{v.a. geometriche}$$

- pmf $$P(Y=y) = \binom{y+m-1}{m-1} \cdot p^y \cdot (1-p)^m$$
- MEDIA $$E(Y) = \frac{m \cdot p}{1-p}$$

PARAMETRI DA DEFINIRE
p e m

MODELLO IPERGEOMETRICO (P non costante) —>estrazione con rimessa

- pmf $$P(K=k) = \frac{\binom{D}{k}\binom{N-D}{m-k}}{\binom{N}{m}}$$

PARAMETRI DA DEFINIRE
N < D, N-D
m < k, m-k

MODELLO DI POISSON (P—>0)

Tale v.a. ha a che fare con eventi la cui probabilità tende a 0.

Ipotesi del modello:

$$\begin{cases} \frac{y}{m} \to 0 \\ m \cdot p = \mu = \text{cost.} \\ p \to 0 \\ m \to +\infty \end{cases}$$

1° FORMULAZIONE pmf

$$P(Y=y) = \frac{\mu}{y!} \cdot e^{-\mu}$$

Ipotesi del modello in ambito continuo:

$$\begin{cases} p = \lambda \Delta x \to 0 \\ m = \frac{x}{\Delta x} \to \infty \\ \frac{y}{m} = \frac{y}{x/\Delta x} \to 0 \\ m \cdot p = \lambda x = \mu \end{cases}$$

2° FORMULAZIONE pmf

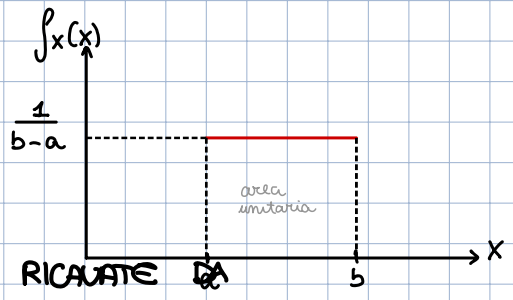
$$P(Y=y) = \frac{(\lambda x)^y}{y!} \cdot e^{-\lambda x}$$

- MEDIA $E(Y) = \mu$
- VARIANZA $E(Y) = \mu$

MODELLI DI V.A. CONTINUE

MODELLO DI V.A. UNIFORME o RETTANGOLARE

pdf $$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{SE } a < x < b \\ 0 & \text{ALTRIMENTI} \end{cases} \qquad a, b \in \mathbb{R}, \; b > a$$



- MEDIA $E(x) = \frac{a+b}{2}$
- VARIANZA $Var(x) = \frac{(b-a)^2}{12}$

Modello v.a. uniforme ridotta (0,1)

Facciamo in modo che $a \equiv 0$ e $b \equiv 1$. La nuova v.a. si ottiene traslando fino all'origine il punto **a** e facendo in modo di avere un lato unitario:

$$y = \frac{x-a}{b-a} \qquad f_y(y) = \begin{cases} 1 & \text{SE } 0 \le y \le 1 \\ 0 & \text{ALTRIMENTI} \end{cases}$$

- MEDIA $E(y) = \frac{1}{2}$
- VARIANZA $Var(y) = \frac{1}{12}$

MODELLO ESPONENZIALE

È un modello che descrive l'intervallo temporale di buon funzionamento in cui non si verificano guasti. Per definire questo modello partiamo dalla v.a. di Poisson e poniamo Y=0 (dove Y è la variabile di conteggio dei guasti, pertanto ponendola uguale a zero affermiamo che non ci sono guasti).

$$P(y=0) = \frac{e^{-\lambda x} \cdot (\lambda x)^0}{0!} = e^{-\lambda x} \quad \Rightarrow \quad P(X \ge x) = e^{-\lambda x}$$

cdf $$F_X(x) = 1 - P(X > x) = 1 - e^{-\lambda x}$$

pdf $$f_X(x) = \lambda e^{-\lambda x}$$

MEDIA $$E(x) = \frac{1}{\lambda}$$

VARIANZA $$Var(x) = \frac{1}{\lambda^2}$$

V.A. GAUSSIANA STANDARD

$$U = \frac{X - \mu}{\sigma}$$

PROPRIETÀ
- SIMMETRICA RISPETTO ALL'ASSE y
- CRESCENTE TRA $-\infty$ e 0
- DECRESCENTE TRA 0 e $+\infty$
- IL MAX DELLA pdf È IN 0

MEDIA $E(U) = 0$

VARIANZA $Var(U) = \sigma^2 = 1$

pdf $f_U(u) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{u^2}{2}}$

GENERICA GAUSSIANA NON STANDARD

pdf $f_X(x) = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sqrt{2\pi} \cdot \sigma}$

MEDIA $E(x) = \mu$

VARIANZA $Var(x) = \sigma^2$

V.A. LOG-NORMALE

pdf $f_y(y) = \frac{1}{\sigma \sqrt{2\pi}\, y} \cdot e^{-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2}$

STATISTICA DESCRITTIVA

MEDIA CAMPIONARIA

$$\bar{X} = \frac{1}{m} \cdot \sum_{i=1}^{m} X_i$$

MEDIA $E(\bar{X}) = \mu$

VARIANZA $Var(\bar{X}) = \frac{\sigma^2}{m}$

VARIANZA CAMPIONARIA

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

$$Var(S^2) = \frac{\sigma^2}{m}(m-1)$$

V.A. INFERENZIALI

CHI-QUADRATO—>inferenza sulla varianza della poolazione

$$K = X^2_r = \sum_{i=1}^{r} U_i^2$$

SE LA POPOLAZIONE È GAUSSIANA $X^2_{r=m-1} = (m-1) \cdot \frac{S^2}{\sigma^2}$

$K \geq 0$

T-STUDENT- >inferenza sulla media della popolazione con varianza incognita

$$T = \frac{U}{\sqrt{X^2/r}}$$

SE LA POPOLAZIONE È GAUSSIANA

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{m}}}$$

$r = m - 1$ MEDIA $E(T) = 0$

Z-FISHER- >inferenza sul rapporto delle varianze di due popolazioni

$$Z = \frac{\chi_1^2/r_1}{\chi_2^2/r_2}$$

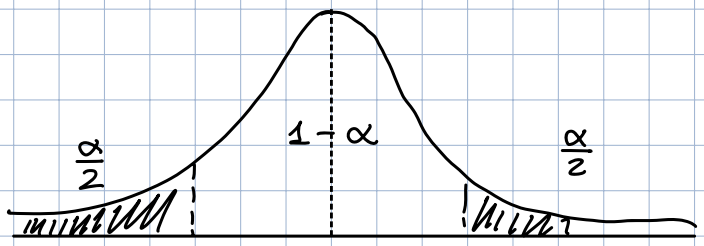
SE LA POPOLAZIONE È GAUSSIANA → $Z_{r_1, r_2} = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2}$

STIMA INTERVALLARE

Popolazione gaussiana

INFERENZA SU μ, CON σ^2 NOTA

$$U = \frac{\bar{X} - \mu}{\sigma/\sqrt{m}} \Rightarrow$$



$$-u < \frac{\bar{x} - \mu}{\sigma/\sqrt{m}} < u \Rightarrow \boxed{\bar{x} - u\cdot\frac{\sigma}{\sqrt{m}} < \mu < \bar{x} + u\frac{\sigma}{\sqrt{m}}}$$

INFERENZA SU μ, CON σ^2 INCOGNITA

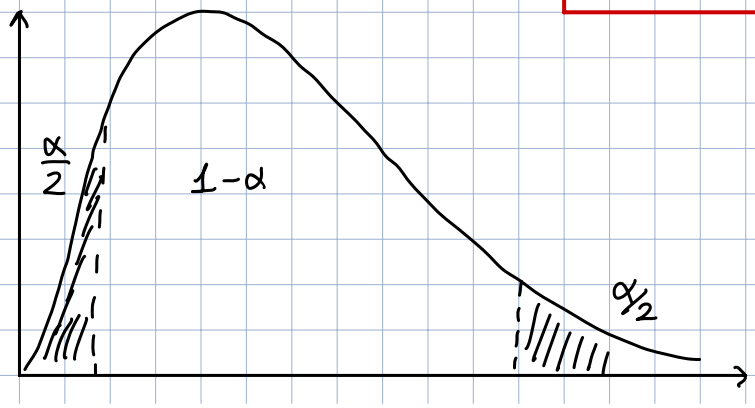
$$T = \frac{\bar{X} - \mu}{S/\sqrt{m}} \qquad \nu = m - 1$$

$$t_{\substack{\alpha/2 \\ \nu=m-1}} < \frac{\bar{x} - \mu}{s/\sqrt{m}} < t_{\substack{\alpha/2 \\ \nu=m-1}} \Rightarrow \boxed{\bar{x} - t_{\substack{\alpha/2 \\ \nu=m-1}} \cdot \frac{s}{\sqrt{m}} < \mu < \bar{x} + t_{\substack{\alpha/2 \\ \nu=m-1}} \cdot \frac{s}{\sqrt{m}}}$$

INFERENZA SU σ^2

$$\chi^2 = (m-1)\frac{s^2}{\sigma^2}$$

$$k_1 < (m-1)\frac{s^2}{\sigma^2} < k_2 \Rightarrow \boxed{\frac{(m-1)s^2}{k_2} < \sigma^2 < \frac{(m-1)s^2}{k_1}}$$



INFERENZA SUL CONFRONTO DI σ_1^2 e σ_2^2

$$Z_{v_1,v_2} = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2}$$

$$Z_{low\ v_1,v_2} < \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} < Z_{upp\ v_1,v_2} \Rightarrow \boxed{\frac{S_2^2}{S_1^2} \cdot Z_L < \frac{\sigma_2^2}{\sigma_1^2} < \frac{S_2^2}{S_1^2} \cdot Z_U}$$

$$Z_{LOWER\ v_1,v_2} = \frac{1}{Z_{UPPER\ v_2,v_1\ \alpha/2}}$$

Popolazione binomiale

Grazie al TLC, il problema binomiale si distribuisce come una normale N di media ***p*** e varianza **p(1-p)/n**. Vogliamo fare inferenza sulla ***p*** della popolazione (probabilità di successo. La funzione ancillare che aiuta a fare ciò è la standardizzazione di $\hat{p}$.

$$-u_{\alpha/2} < \frac{\hat{P} - P}{\sqrt{\frac{\hat{P}(1-\hat{P})}{m}}} < u_{\alpha/2}$$

Popolazione di Poisson

Il parametro su cui fare inferenza è la frequenza media μ.

$$-u_{\alpha/2} < \frac{\hat{\mu} - \mu}{\sqrt{\frac{\hat{\mu}}{m}}} < u_{\alpha/2}$$

Popolazione esponenziale

$$k_1 < 2\lambda m\bar{x} < k_2 \Rightarrow \frac{2m\bar{x}}{k_2} < \frac{1}{\lambda} < \frac{2m\bar{x}}{k_1}$$

TEST D'IPOTESI

ERRORE DI PRIMA SPECIE: Pr(rifiuto Ho | Ho è vera)= α

ERRORE DI SECONDA SPECIE: Pr(accettare Ho | Ho è falsa)= β

LIVELLO DI SIGNIFICATIVITÀ DEL TEST: Pr(accettare Ho | Ho è vera)= $1-\alpha$

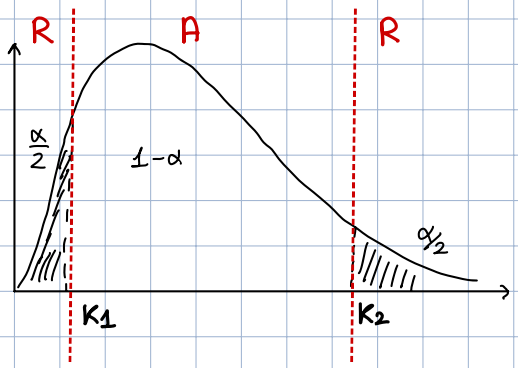
POTENZA DEL TEST: Pr(rifiuto Ho | Ho è falsa)= $1-\beta$, $\alpha \leq 1-\beta \leq 1$

INFERENZA SU σ^2 → Test bilaterale

$H_0 : \sigma^2 = \sigma_0^2$

$H_1 : \sigma^2 \neq \sigma_0^2$

$$\chi^2 = (m-1)\cdot\frac{S^2}{\sigma_2}$$



K_1 : ENTRO CON i g.d.l E $1-\alpha/2$

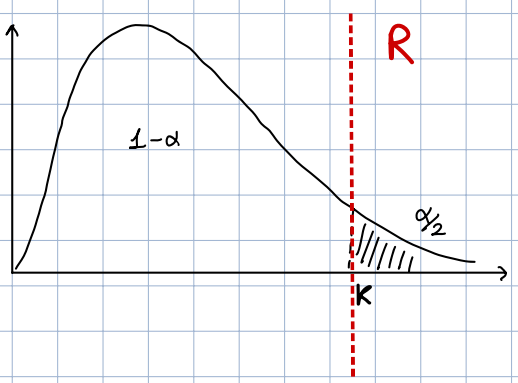
K_2 : ENTRO CON i g.d.l E $\alpha/2$

INFERENZA SU σ^2 → Test unilaterale

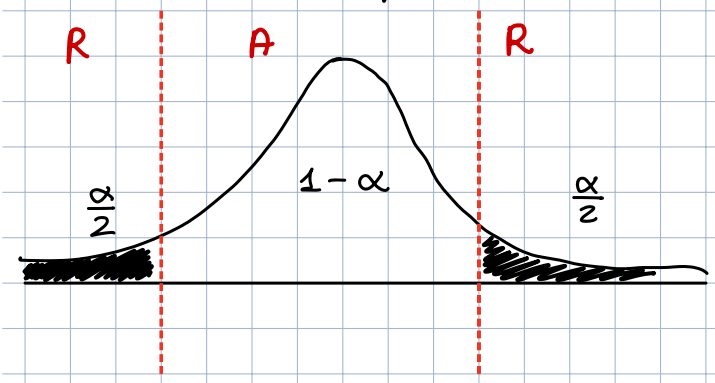
$H_0 : \sigma^2 = \sigma_0^2$

$H_1 : \sigma^2 > \sigma_0^2$

$$\chi^2 = (m-1)\cdot\frac{S^2}{\sigma_2}$$



INFERENZA SU μ CON σ^2 INCOGNITA → Test bilaterale

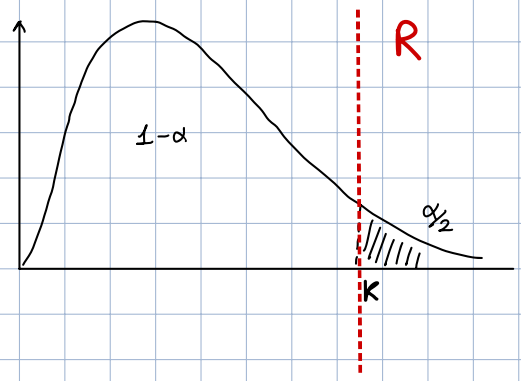


$H_0: \mu = \mu_0$

$H_1: \mu \neq \mu_0$

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

TEST BILATERALE CON DUE POPOLAZIONI (CONFRONTO TRA VARIANZE)



$H_0: \sigma_1^2 = \sigma_2^2$

$H_1: \sigma_1^2 \neq \sigma_2^2$

$$Z = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2}$$

$\nu_1 = n_1 - 1$

$\nu_2 = n_2 - 1$

$$Z_{CALC} = \frac{S_1^2}{S_2^2} \quad \left(\text{SOTTO } H_0 \ \frac{\sigma_1^2}{\sigma_2^2} = 1\right)$$

$$Z_{LOWER} = \frac{1}{Z_{UPP.\,\nu_2,\nu_1}}$$

CONFRONTO TRA MEDIE μ DI DUE POPOLAZIONI

– *se nota* σ^2 e $\sigma_1^2 = \sigma_2^2$

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

$$U = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

– se $\sigma_1^2 \neq \sigma_2^2$

$$U = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

– se σ^2 INCOGNITA

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

$$t_{m_1+m_2-2} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{S_p \cdot \sqrt{\frac{1}{m_1} + \frac{1}{m_2}}}$$

$$S_p^2 = \frac{(m_1 - 1)S_1^2 + (m_2 - 1)S_2^2}{m_1 + m_2 - 2}$$

1. **TROVARE MEDIE** $\bar{x}$ e $\bar{y}$

2. **TROVARE SCARTI** $(x_i - \bar{x})$ e $(y_i - \bar{y})$ → La somma degli scarti deve essere 0

3. **PRODOTTO SCARTI** $(x_i - \bar{x}) \cdot (y_i - \bar{y})$

La somma del prodotto degli scarti é il numeratore della covarianza → $\sum_{i=1}^{m} (x_i - \bar{x})(y_i - \bar{y})$

4. **SCARTI AL QUADRATO** $(x_i - \bar{x})^2$ e $(y_i - \bar{y})^2$

La somma del prodotto degli scarti al quadrato é la devianza →

DEVIANZA

$$Dev(x) = \sum_{i=1}^{m} (x_i - \bar{x})^2$$

$$Dev(y) = \sum_{i=1}^{m} (y_i - \bar{y})^2$$

5. **PRODOTTI** $x_i \cdot y_i$

6. **QUADRATI** x_i^2 e y_i^2

7. **CALCOLO COVARIANZA, VARIANZA E SCARTI QUADRATICI MEDI DI** x **E** y

COVARIANZA

$$Cov(x,y) = \frac{\sum_{i=1}^{m} (x_i - \bar{x})(y_i - \bar{y})}{m}$$

VARIANZA

$$Var(x) = \frac{\sum_{i=1}^{m} (x_i - \bar{x})^2}{m}$$

SCARTO QUADRATICO MEDIO

$$S_x = \sqrt{Var(x)} \qquad S_y = \sqrt{Var(y)}$$

8.STIMA DI $a \longrightarrow \hat{a} = \dfrac{Cov(x,y)}{Var(x)} = \dfrac{\sum_{i=1}^{m}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{m}(x_i-\bar{x})^2}$

9.STIMA DI $b \longrightarrow \hat{b} = \bar{y} - \hat{a}\bar{x}$

10. RETTA DI REGRESSIONE $\longrightarrow y = \hat{a}x + \hat{b}$

11. CALCOLO I RESIDUI $\longrightarrow (y_i - \hat{y}) \longrightarrow$ Somma residui=0

12. RESIDUI AL QUADRATO $\longrightarrow (y_i - \hat{y})^2$

DEVIANZA DELL'ERRORE

$$Dev(E) = \sum_E = \sum_{i=1}^{m}(y_i - \hat{y})^2$$

13. CALCOLO DEI QUADRATI DEGLI SCOSTAMENTI TRA $\hat{y}$ e $\bar{y} \longrightarrow (\hat{y} - \bar{y})$

DEVIANZA DI REGRESSIONE

$$Dev(R) = \sum_{R.F} = \sum_{i=1}^{m}(\hat{y} - \bar{y})^2$$

14. CALCOLO DEVIANZA TOTALE

DEVIANZA TOTALE

$$\sum_T = \sum_E + \sum_R$$

15. CALCOLO IL COEFFICIENTE DI CORRELAZIONE

$$\hat{\rho} = \frac{Cov(x,y)}{\sqrt{Var(x)} \cdot \sqrt{Var(y)}}$$

16. CALCOLO $\longrightarrow$ $\hat{\rho}^2 = \frac{\Sigma_R}{\Sigma_T} = \frac{1-\Sigma_E}{\Sigma_T} = \frac{\hat{a}^2 \cdot Var(x)}{Var(y)}$

17. CALCOLO $\longrightarrow$ $\Sigma_\epsilon = \frac{1-\hat{\rho}^2}{\Sigma_T}$

18. STIMA Var(Z)=Var(Y) $\longrightarrow$ $Var(z) = \frac{\Sigma_\epsilon}{m-2} = \sigma^2$

19. CALCOLO VARIANZA DI $\hat{a}$ **e** $\hat{b}$

$$Var(\hat{a}) = \frac{Var(z)}{Dev(x)} \qquad Var(\hat{b}) = Var(z) \cdot \frac{\frac{1}{m} + \bar{x}^2}{Dev(x)}$$

20. TEST D'IPOTESI E INTERVALLI DI CONFIDENZA DEI PARAMETRI

Su $\hat{a}$: $r = m-2$

$$T = \frac{\hat{a} - a}{\sqrt{Var(\hat{a})}}$$

Su $\hat{b}$ $r = m-2$

$$T = \frac{\hat{b} - b}{\sqrt{Var(\hat{b})}}$$

CONFRONTO T_{CALC} CON T_{TAB} PER DECIDERE SE ACCETTARE H_0

INTERVALLI DI CONFIDENZA

Di $\hat{a}$

$$\hat{a} \pm t_{1-\frac{\alpha}{2}} \cdot \sqrt{Var(\hat{a})}$$

Di $\hat{b}$

$$\hat{b} \pm t_{1-\frac{\alpha}{2}} \cdot \sqrt{Var(\hat{b})}$$

21. TEST DI DIPENDENZA LINEARE

$$Z = \frac{\Sigma_F / 1\,gdl}{\Sigma_E / (m-2)}$$

22. TEST D'IPOTESI E INTERVALLI DI CONFIDENZA DELLA RISPOSTA

$$T_{r=m-2} = \frac{\hat{y}_0 - \mu_{(x_0)}}{\sqrt{Var(\hat{y}_0)}} \qquad Var(\hat{y}_0) = Var(z) \cdot \left[\frac{1}{m} + \frac{(x_0 - \bar{x})^2}{Dev(x)}\right]$$

Intervallo di confidenza della media della variabile aleatoria di risposta y

$$\hat{y}_0 \pm t_{1-\frac{\alpha}{2}} \cdot \sqrt{Var(\hat{y}_0)} \qquad r = m-2$$

Intervallo di confidenza del valore singolo della V.a di risposta y

$$Var(\hat{y}_0 - y_0) = \sigma^2 \cdot \left[1 + \frac{1}{m} + \frac{(x_0 - \bar{x})^2}{Dev(x)}\right]$$

$$T_{r=m-2} = \frac{\hat{y}_0 - y_0}{Var(\hat{y}_0 - y_0)} \qquad \hat{y}_0 \pm t_{1-\frac{\alpha}{2}} \cdot Var(\hat{y}_0 - y_0)$$