Potenza di un test

venerdì 9 maggio 2014 10.10

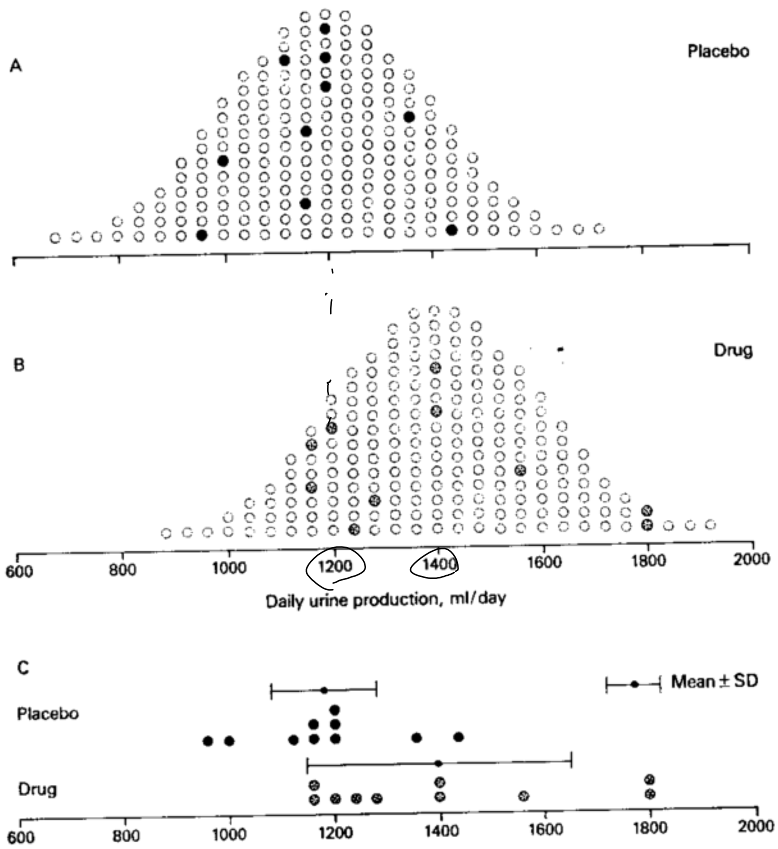
Se $H_0: \mu_X = \mu_Y$ è falsa, siamo esposti al rischio di commettere un errore di 2^A specie; questo succede se $|t| < t_{1-\frac{\alpha}{2}}(\nu)$.

Si chiama potenza di un test la probabilità $1-\beta$ di rifiutare un'ipotesi falsa:

$$1 - \beta = \mathbb{P}_1\{|t| > t_{1-\frac{\alpha}{2}}(\nu)\}$$

$\beta = \mathbb{P}_1\{|t| < t_{1-\frac{\alpha}{2}}(\nu)\}$

prob. di commettere un errore di 2^A specie



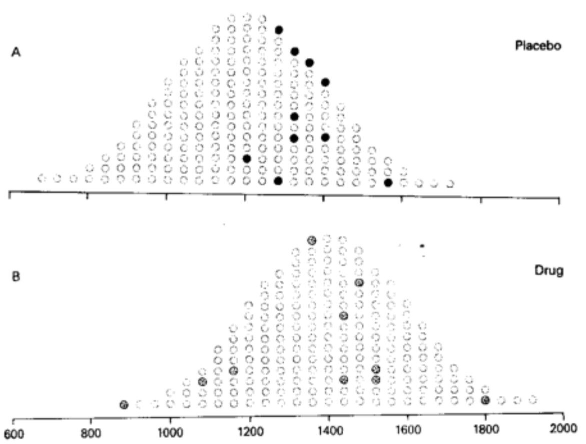
$\sigma = 200$

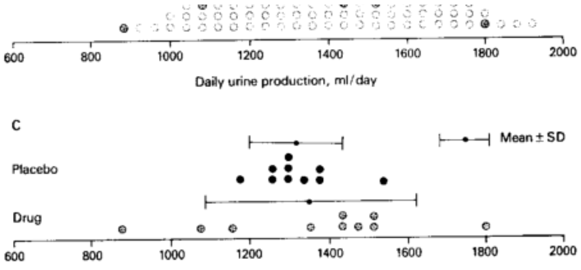
esempio: placebo contro diuretico, $\mu_X = 1200$, $\mu_Y = 1400$. Stavolta è vera $H_1: \mu_X \neq \mu_Y$. Con che probabilità $|t| > t_{1-\frac{\alpha}{2}}(\nu)$

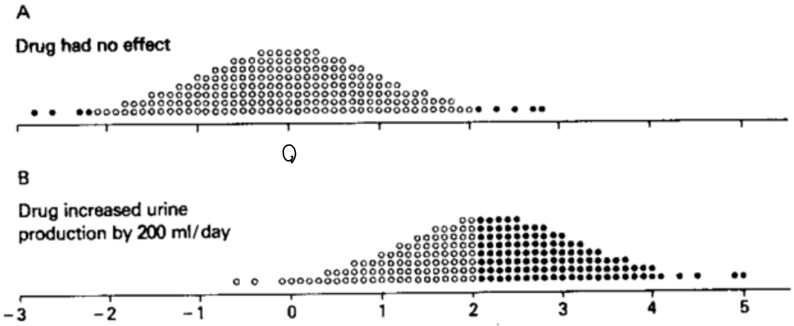
nella 1^A figura, $t = -2{,}201$; se $\alpha = 0{,}05$, il valore critico è $t_{0{,}975}(18) = 2{,}100$; $|t| > t_{1-\frac{\alpha}{2}}(\nu) \Rightarrow H_0$ rif. H_1 acc.

altra figura: $t = -0{,}404$, quindi $|t| < t_{1-\frac{\alpha}{2}}(\nu)$ e commettiamo un errore di 2^A specie.

Supponiamo di fare questo test 200 volte, e otteniamo 200 valori di t:





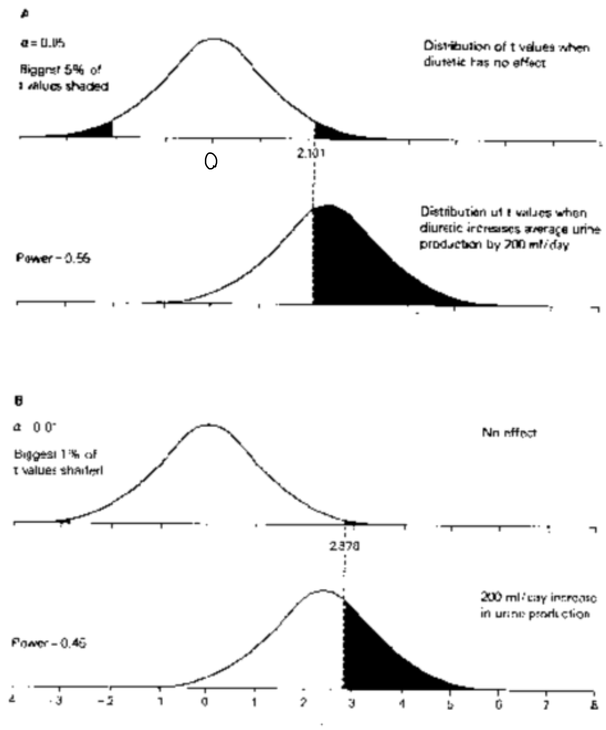


In questo caso, $1-\beta \simeq 0,55$

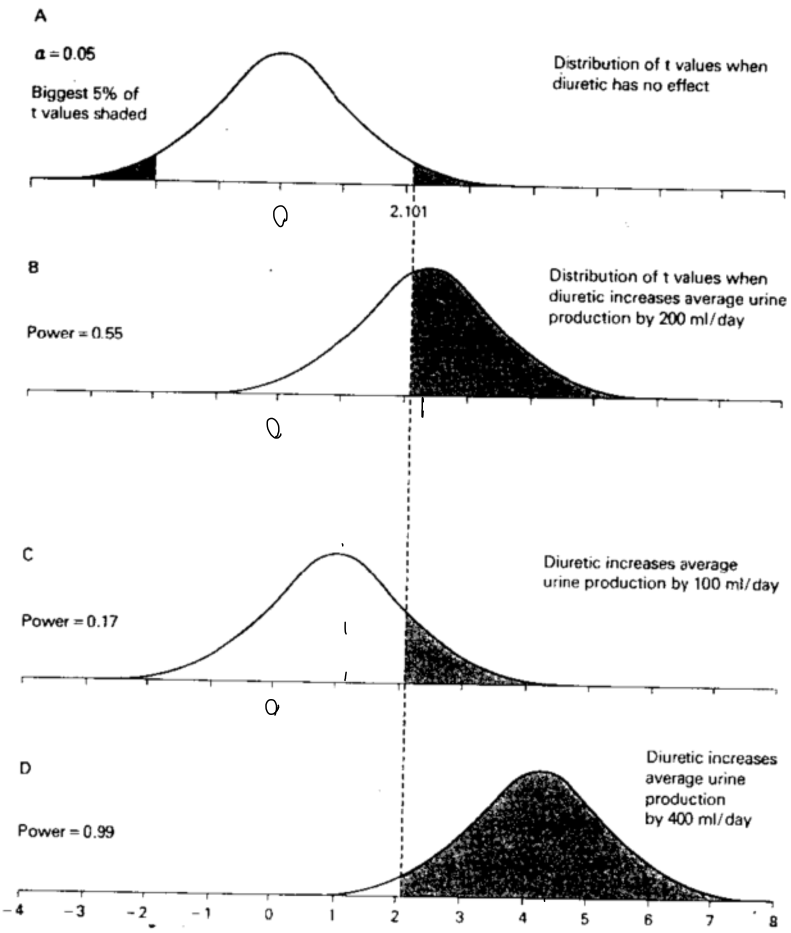
Da cosa dipendono β e $1-\beta$?

β dipende da:

- il livello α: $\searrow$
- la differenza tra i parametri $\delta = |\mu_X - \mu_Y|$: $\searrow$
- la taglia dei campioni n: $\searrow$

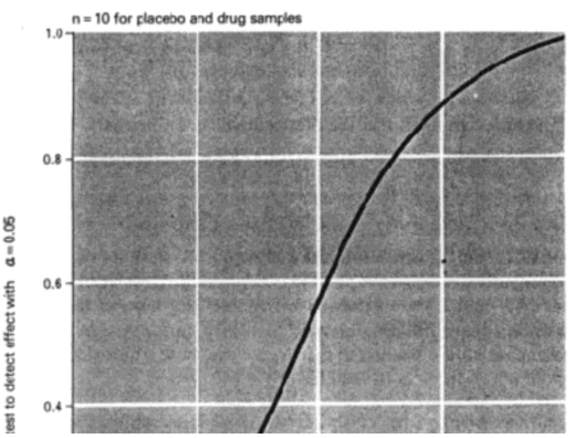


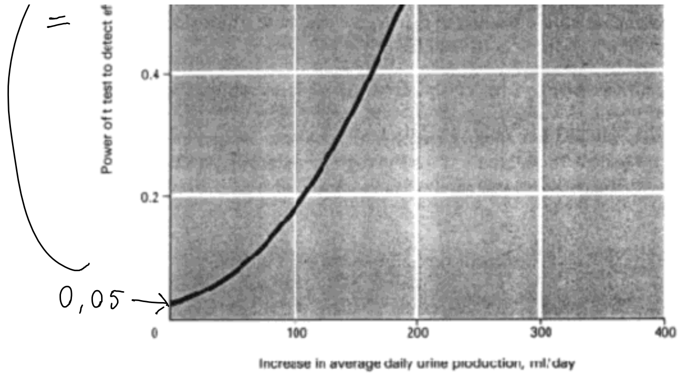
Diminuendo α, aumenta β (e diminuisce la potenza).



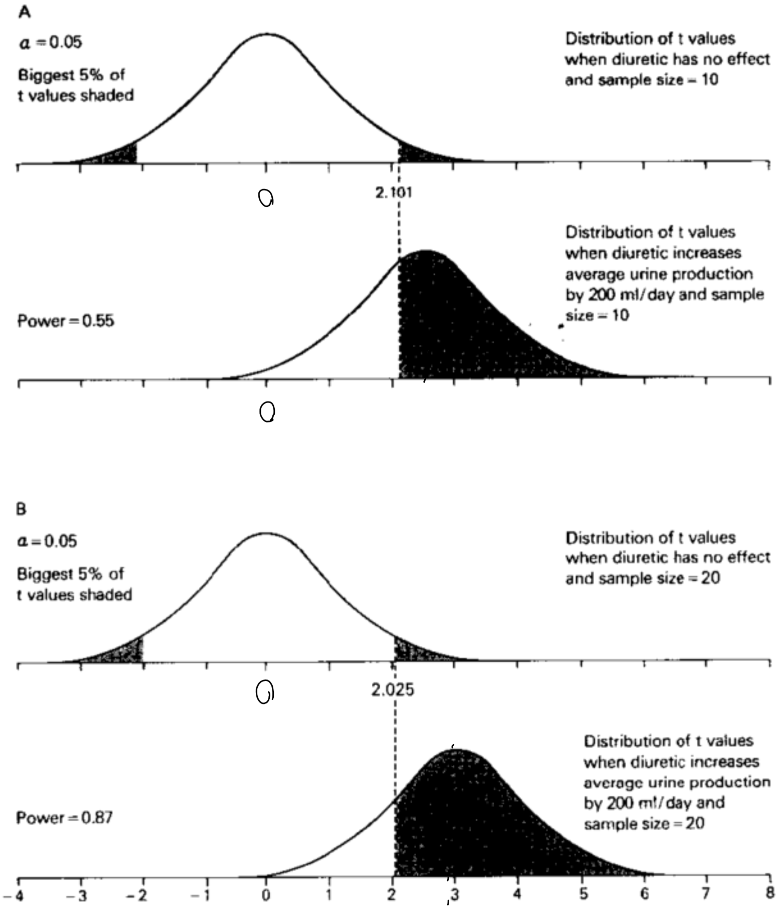
Se aumentiamo $\delta = |\mu_x - \mu_y|$, diminuisce β e aumenta la potenza $1 - \beta$.

Si può ricorrere a <u>grafici della potenza</u>:





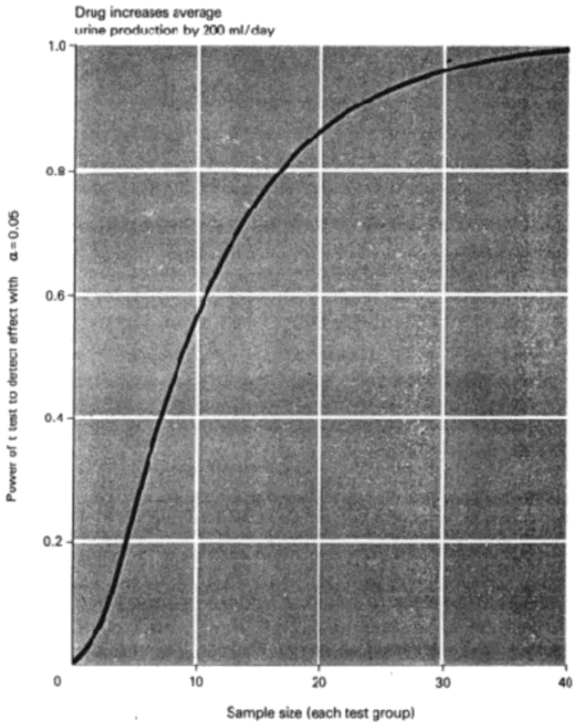
Taglia del campione:

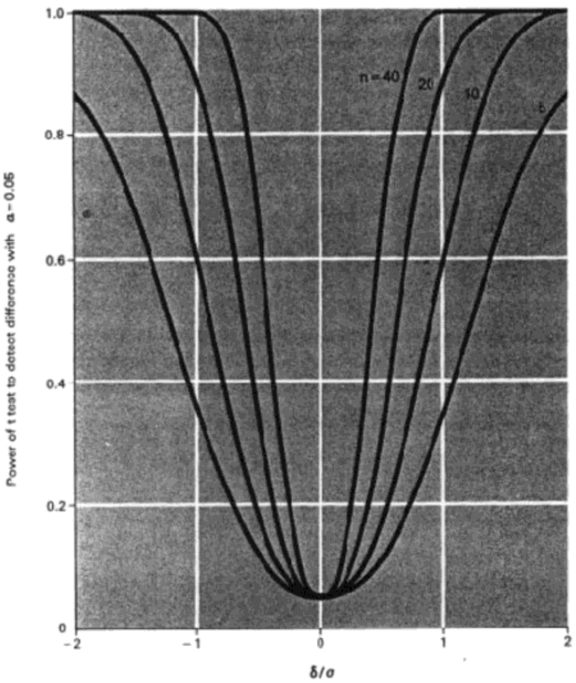


Aumentando la taglia n, la distribuzione di t si sposta verso la regione critica, E il valore critico diminuisce.

Entrambi questi effetti fanno aumentare $1-\beta$.

Grafico della potenza:





Come dipende la potenza da δ e da n?

Supponiamo che $\delta \neq 0$ (qui $\delta = \mu_X - \mu_Y$). Allora

$$t = \frac{\bar{X} - \bar{Y}}{S_{\bar{X}-\bar{Y}}} \nsim t(\nu) \text{ ; che legge ha ?}$$

Sotto H_1, $\bar{X} - \bar{Y} \sim N\left(\mu_X - \mu_Y, \sigma^2\left(\frac{1}{n} + \frac{1}{n}\right)\right)$

(supponiamo che $n_X = n_Y = n$), cioè

$$\bar{X} - \bar{Y} \sim N\left(\delta, \sigma^2 \cdot \frac{2}{n}\right), \text{ quindi } \bar{X} - \bar{Y} - \delta \sim N\left(0, \sigma^2 \cdot \frac{2}{n}\right)$$

$$\Rightarrow t = \frac{\bar{X} - \bar{Y} - \delta + \delta}{S_{\bar{X}-\bar{Y}}} = \underbrace{\frac{\bar{X} - \bar{Y} - \delta}{S_{\bar{X}-\bar{Y}}}}_{\substack{\shortparallel \\ t' \sim t(\nu)}} + \frac{\delta}{S_{\bar{X}-\bar{Y}}}$$

$$t' \sim t(\nu)$$

quanto è lo scostamento $\dfrac{\delta}{S_{\bar{X}-\bar{Y}}}$?

$S_{\bar{X}-\bar{Y}}$ è uno stimatore di $\sigma\sqrt{\frac{2}{n}}$, quindi "per n grande"

$$\frac{\delta}{S_{\bar{X}-\bar{Y}}} \simeq \frac{\delta}{\sigma\sqrt{\frac{2}{n}}} = \sqrt{\frac{n}{2}} \cdot \frac{\delta}{\sigma}$$

Approssimazione normale

Supponiamo che n sia "grande" abbastanza da fare le seguenti approssimazioni:

$$\bar{X} - \bar{Y} \sim N\left(\delta, \sigma^2 \frac{2}{n}\right) \quad (\forall n)$$

$$S_{\bar{X}-\bar{Y}} \approx \sqrt{\frac{2}{n}}\,\sigma$$

Allora $\quad t = \dfrac{\bar{X}-\bar{Y}}{S_{\bar{X}-\bar{Y}}} \approx N\left(\sqrt{\frac{n}{2}}\cdot\frac{\delta}{\sigma}, 1\right)$

Quanto vale $1-\beta = \mathbb{P}_1\left\{|t| > t_{1-\frac{\alpha}{2}}(\nu)\right\}$?

$$\beta = \mathbb{P}_1\left\{|t| < t_{1-\frac{\alpha}{2}}(\nu)\right\} = \mathbb{P}_1\left\{-t_{1-\frac{\alpha}{2}}(\nu) < t < t_{1-\frac{\alpha}{2}}(\nu)\right\} \simeq$$

$$\left(\overset{\text{"}n \simeq \infty\text{"}}{\text{appr. normale}}\right) \simeq \mathbb{P}_1\left\{-q_{1-\frac{\alpha}{2}} < t < q_{1-\frac{\alpha}{2}}\right\} =$$

$$= \mathbb{P}_1\left\{-q_{1-\frac{\alpha}{2}} - \sqrt{\frac{n}{2}}\frac{\delta}{\sigma} < \underbrace{t - \sqrt{\frac{n}{2}}\frac{\delta}{\sigma}}_{\approx N(0,1)} < q_{1-\frac{\alpha}{2}} - \sqrt{\frac{n}{2}}\frac{\delta}{\sigma}\right\} =$$

$$= F_Z\left(q_{1-\frac{\alpha}{2}} - \sqrt{\frac{n}{2}}\frac{\delta}{\sigma}\right) - F_Z\left(-q_{1-\frac{\alpha}{2}} - \sqrt{\frac{n}{2}}\frac{\delta}{\sigma}\right)$$

Se vogliamo invertire questa relazione in funzione di n (o di $\frac{\delta}{\sigma}$), approssimiamo:

- se $\delta > 0$, $-q_{1-\frac{\alpha}{2}} - \sqrt{\frac{n}{2}}\frac{\delta}{\sigma} \ll 0$, e quindi

$$F_2\left(-q_{1-\frac{\alpha}{2}} - \sqrt{\frac{n}{2}}\frac{\delta}{\sigma}\right) \simeq 0 \Rightarrow$$

$$\beta \simeq F_2\left(q_{1-\frac{\alpha}{2}} - \sqrt{\frac{n}{2}}\frac{\delta}{\sigma}\right)$$

$$q_\beta \simeq q_{1-\frac{\alpha}{2}} - \sqrt{\frac{n}{2}}\frac{\delta}{\sigma}$$

$$\sqrt{\frac{n}{2}}\frac{\delta}{\sigma} \simeq q_{1-\frac{\alpha}{2}} - q_\beta = q_{1-\frac{\alpha}{2}} + q_{1-\beta}$$

$$\sqrt{\frac{n}{2}} \simeq \frac{\sigma}{\delta}\left(q_{1-\frac{\alpha}{2}} + q_{1-\beta}\right)$$

$$n \simeq 2\frac{\sigma^2}{\delta^2}\left(q_{1-\frac{\alpha}{2}} + q_{1-\beta}\right)^2 \qquad (*)$$

- se $\delta < 0$, $q_{1-\frac{\alpha}{2}} - \sqrt{\frac{n}{2}}\frac{\delta}{\sigma} \gg 0$, e

$$F_2\left(q_{1-\frac{\alpha}{2}} - \sqrt{\frac{n}{2}}\frac{\delta}{\sigma}\right) \simeq 1 \Longrightarrow$$

$$\beta \simeq 1 - F_2\left(-q_{1-\frac{\alpha}{2}} - \sqrt{\frac{n}{2}}\frac{\delta}{\sigma}\right)$$

$$F_2\left(-q_{1-\frac{\alpha}{2}} - \sqrt{\frac{n}{2}}\frac{\delta}{\sigma}\right) \simeq 1-\beta$$

$$-q_{1-\frac{\alpha}{2}} - \sqrt{\frac{n}{2}}\frac{\delta}{\sigma} \simeq q_{1-\beta}$$

$$-\sqrt{\frac{n}{2}}\frac{\delta}{\sigma} \simeq q_{1-\frac{\alpha}{2}} + q_{1-\beta}$$

$$\frac{n}{2}\frac{\delta^2}{\sigma^2} \simeq \left(q_{1-\frac{\alpha}{2}} + q_{1-\beta}\right)^2$$

e arriviamo ancora a (*).

<u>esercizio 6.3</u> $\mu_x = 45, \mu_y = 50, \sigma_x = \sigma_y = \sigma = 10$

1) $1-\beta = 90\% = 0,9$, $\alpha = 5\% = 0,05$

$$n = 2\left(\frac{\sigma}{\delta}\right)^2\left(q_{1-\frac{\alpha}{2}} + q_{1-\beta}\right)^2 =$$

$$= 2\left(\frac{10}{45-50}\right)^2\left(q_{0,975} + q_{0,9}\right)^2 =$$

$$= 2\cdot(-2)^2(1,96 + 1,28)^2 =$$

$$= 83,98$$

ci vogliono almeno 84 individui per campione

2) $n = 50$

$$n = 2\left(\frac{\sigma}{\delta}\right)^2\left(q_{1-\frac{\alpha}{2}} + q_{1-\beta}\right)^2$$

$$50 = 2\cdot 4\cdot(1,96 + q_{1-\beta})^2$$

$$\frac{25}{4} = (1,96 + q_{1-\beta})^2 =$$

$$\sqrt{\frac{25}{4}} = \frac{5}{2} = 1,96 + q_{1-\beta}$$

$$q_{1-\beta} = \frac{5}{2} - 1,96 = 2,5 - 1,96 = 0,54$$

$$1-\beta = F_z(0,54) = 0,70540 \simeq 70\%$$

3) $$s_{\bar{x}-\bar{y}} = \sqrt{\frac{9,2}{50} + \frac{10,4}{50}} = 0,63$$

$$t = \frac{45,5 - 50.1}{0,63} = \frac{4.6}{0,63} = 7,35$$

$$t_{0,975}(50+50-2) = t_{0,975}(98) \simeq t_{0,975}(100) = 1,984$$

$|t| > 1,984 \Rightarrow H_0$ rif. H_1 acc.

$|t| > 1{,}984 \Rightarrow H_0$ rif. H_1 acc.

4) cerchiamo P:

$t_{0,999}(100) = 3{,}174 < |t| = 2.35$

$\Downarrow$

$\alpha = 0{,}002$

$\Rightarrow \quad P < 0{,}002$