Controllo della Molteplicità nei trial clinici

Biostatistica avanzata per la ricerca clinica

Livio Finos

Outline

Introduzione

Alcune considerazioni preliminari

Microarray study

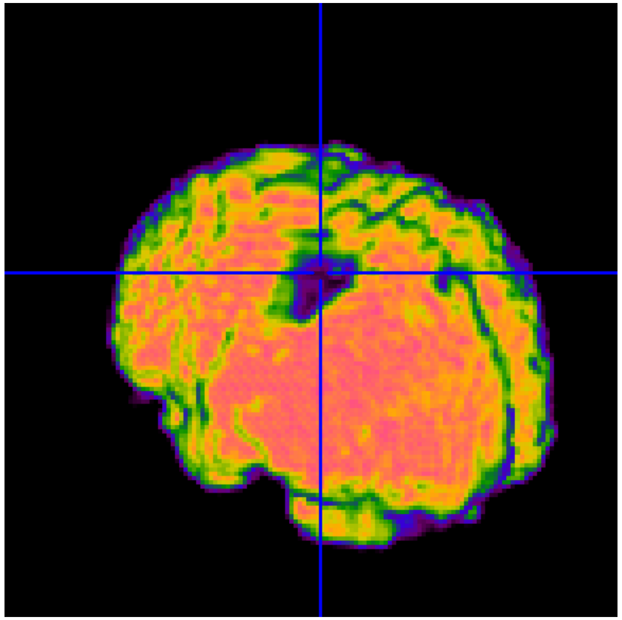
Top 5 genes out of 20000

Gene	p-value
OCIAD2	5.5e-6
NEK3	6.7e-6
TAF5	7.1e-6
FOXD4L6	7.5e-6
ADIG	8.8e-6
⋮	⋮

Small p-value?

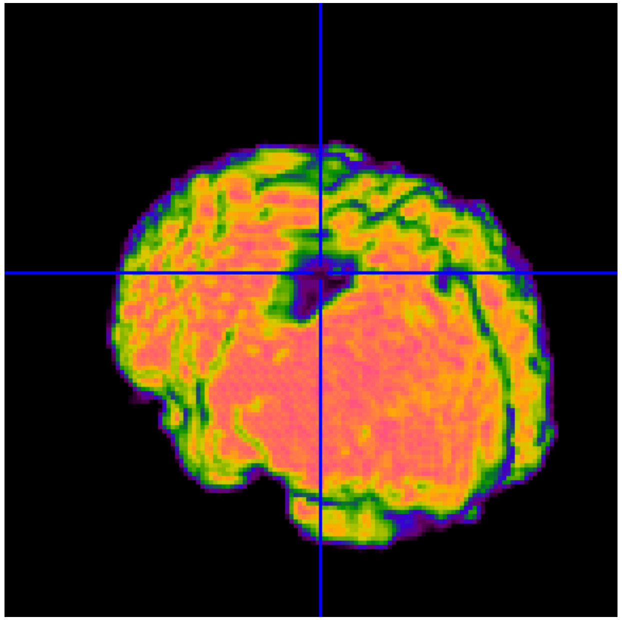
- un p-value di 5.5e-6 è improbabile (evidenza per H_1)
- ma è improbabile anche se consideriamo di averlo cercato tra 20000 test?
- Possiamo veramente dire che OCIAD2 è differentially expressed?
- e a proposito di NEK3?

Ulteriore Esempio: studi fMRI



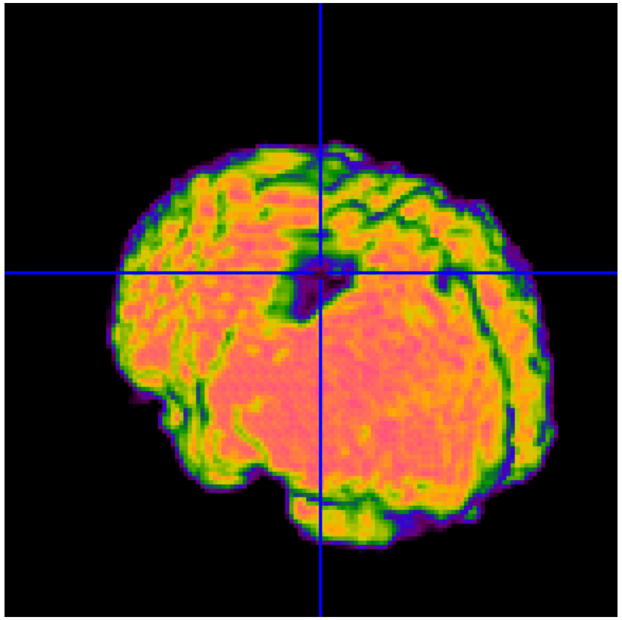
Una mappa di attività celebrale per ogni soggetto

Ulteriore Esempio: studi fMRI



Una mappa di attività celebrale per ogni soggetto
Ogni voxel (punto) produce un p-value

Ulteriore Esempio: studi fMRI



Una mappa di attività celebrale per ogni soggetto
Ogni voxel (punto) produce un p-value
L'output è solitamente una lista dei voxel più attivi
(sui migliaia testati)

Dubbio: necessario controllo della molteplicità?

Altri esempi, più frequenti

Modelli di Regressione (LM e GLM)

Un t-test per ogni Coefficiente di Regressione

Anova

Tutti i Confronti a Coppie (post-hoc)

Ogni volta in cui l'analisi produce più di un p-value

Dubbio: necessario controllo della molteplicità?

. . . e a proposito di trial clinici:

Multiple endpoints

es: più endpoints sono importanti per valutare la guarigione del paziente

Multiple doses

es: Placebo vs Dosi crescenti
vogliamo i Confronti a Coppie (post-hoc)

Multiple sub-groups

es: la terapia funziona solo su sottogruppi della popolazione, ad esempio con un preciso corredo genetico le donne

Nessun Dubbio: necessario controllo della molteplicità!

Outline

Introduzione

Alcune considerazioni preliminari

Verifica di Ipotesi, Un solo test

Due Ipotesi a confronto

- H_0: due gruppi sono Uguali, nessuna relazione tra X e Y,
- H_1: due gruppi sono Diversi, c'è relazione tra X e Y,

Ogni test produce un p-value p,
se $p \leq .05$ ($\alpha = .05$) rifiuto H_0 (e propendo per H_1)

Errori

- **Tipo I** (falso positivo): Rifiuto H_0 quando è Vera
  $P(Errore\ Tipo\ I) = P(p \leq .05|H_0) = .05$
- **Tipo II** (falso negativo): Non Rifiuto H_0 quando è Falsa
  $P(Errore\ Tipo\ II) = P(p > .05|H_1)$
  Potenza:

$$\begin{aligned} P(p \leq .05|H_1) &= 1 - P(p > .05|H_1) \\ &= 1 - P(Errore\ tipo\ II) \end{aligned}$$

Importanza asimmetrica degli errori

Controlliamo la $P(Errore\ tipo\ I)$ (es $\leq .05$)
e cerchiamo il test con massima Potenza (minimo $Errore\ tipo\ II$)

È importante ricordare che

- un p-value significativo ($p \leq \alpha$) ci autorizza a pensare che sia vera H_1, mentre
- un p-value non significativo ($p > \alpha$) NON ci autorizza a pensare che sia vera H_0, semplicemente non abbiamo abbastanza evidenza per rifiutarla.

Errori di Tipo I:

$P(p \leq .05 | H_0 = \text{i 2 gruppi sono Uguali}) =?$

Supponiamo $H_0 : \mu_1 - \mu_2 = 0$ e $H_1 : \mu_1 - \mu_2 < 0$

statistica test $T = \frac{\bar{x}_1 - \bar{x}_2}{\hat{\sigma}}$ ($\hat{\sigma}$ stima della dev std di $\bar{x}_1 - \bar{x}_2$)

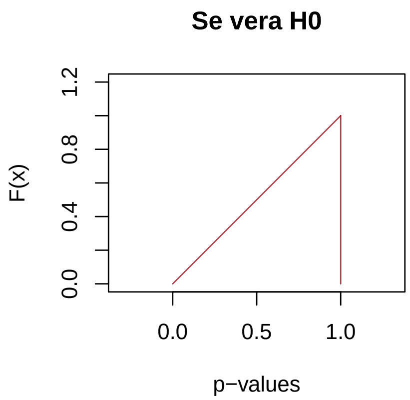
sotto H_0: $T \sim t_{n_1+n_2-2}$, allora

$$P(T \leq t_\alpha | H_0) = \alpha \ \forall \alpha$$

$$P(F(T) \leq F(t_\alpha) | H_0) = \alpha \ \forall \alpha$$

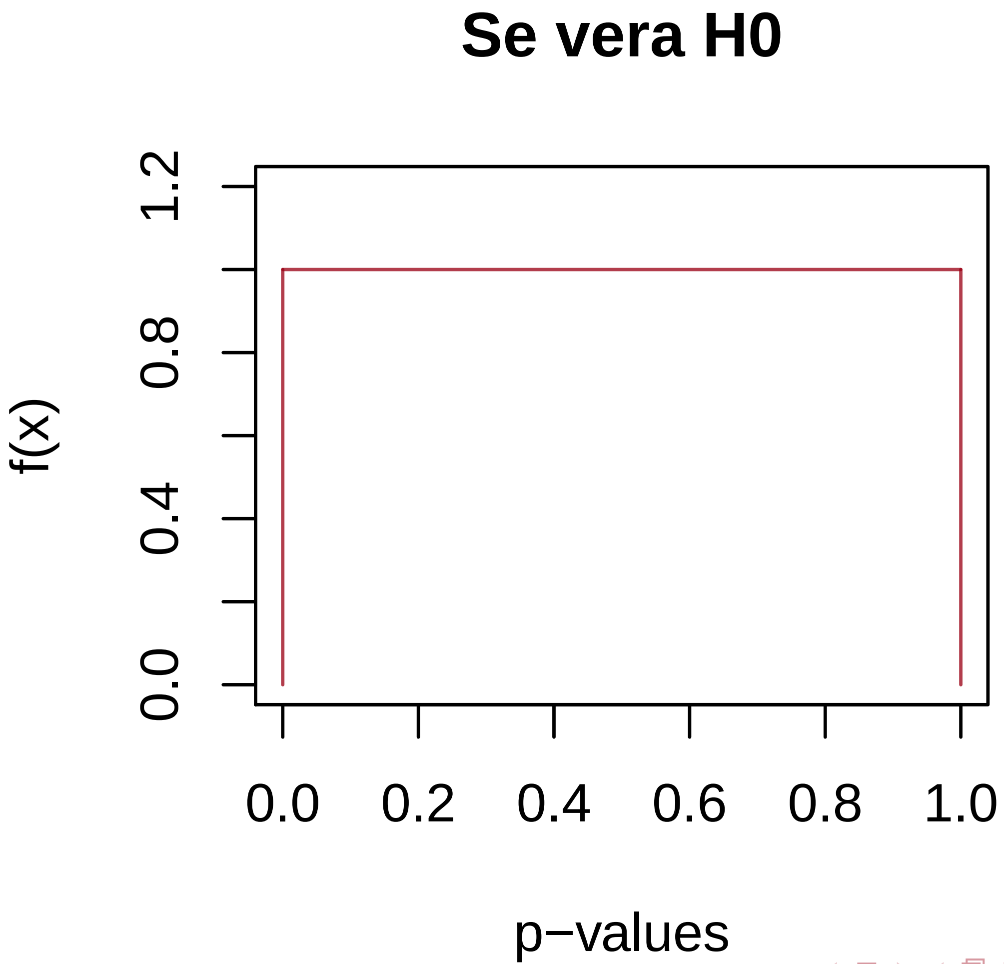
$$P(P \leq \alpha | H_0) = \alpha \ \forall \alpha$$

ne consegue che $P \sim U(0,1)$



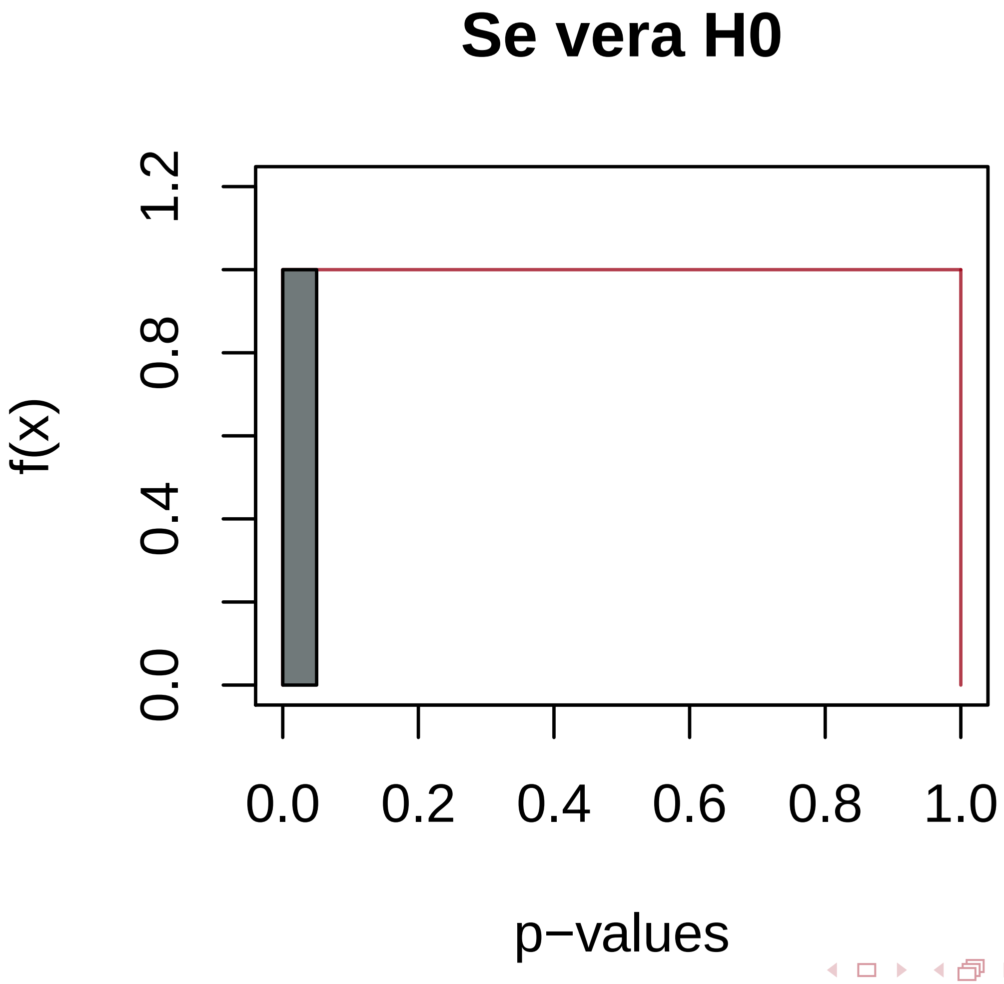
Errori di Tipo I:

Sotto H_0 il p-value è una variabile aleatoria uniforme $U(0,1)$



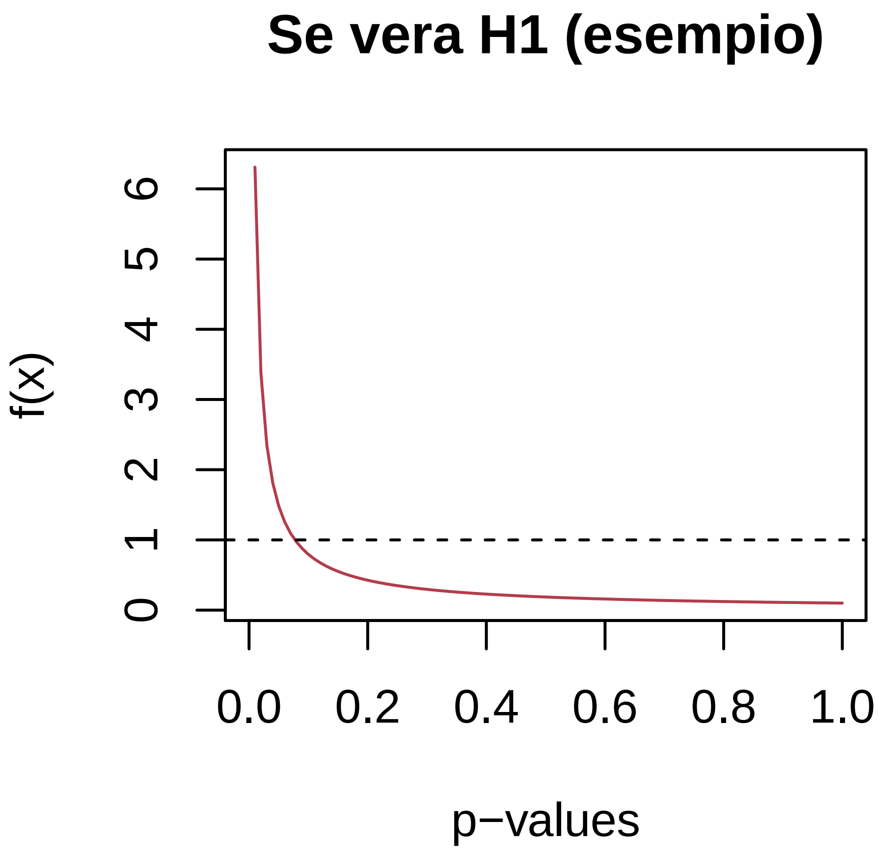
Errori di Tipo I:

Errore di I tipo: $P(p \leq .05|H_0) = .05$



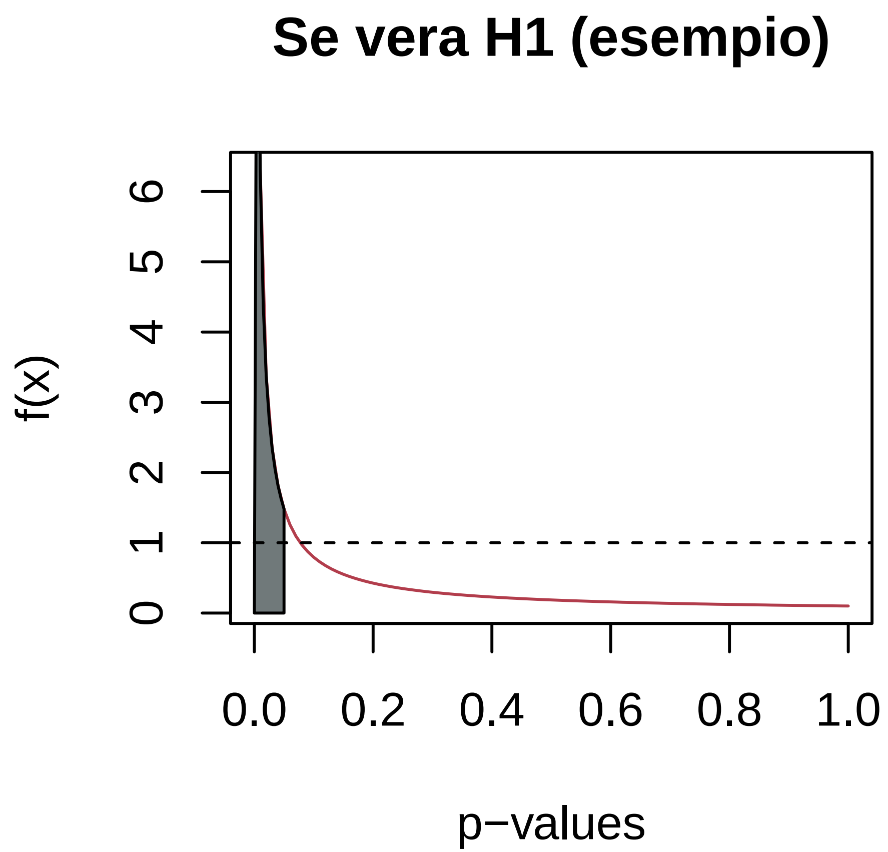
Potenza: $P(p \leq .05|H_1 = 2\ gruppi\ Diversi)$

Sotto H_1 il p-value è stocasticamente inferiore ad una variabile aleatoria uniforme $U(0, 1)$ (Non distorsione del test)



Potenza: $P(p \leq .05|H_1 = 2\ gruppi\ Diversi)$

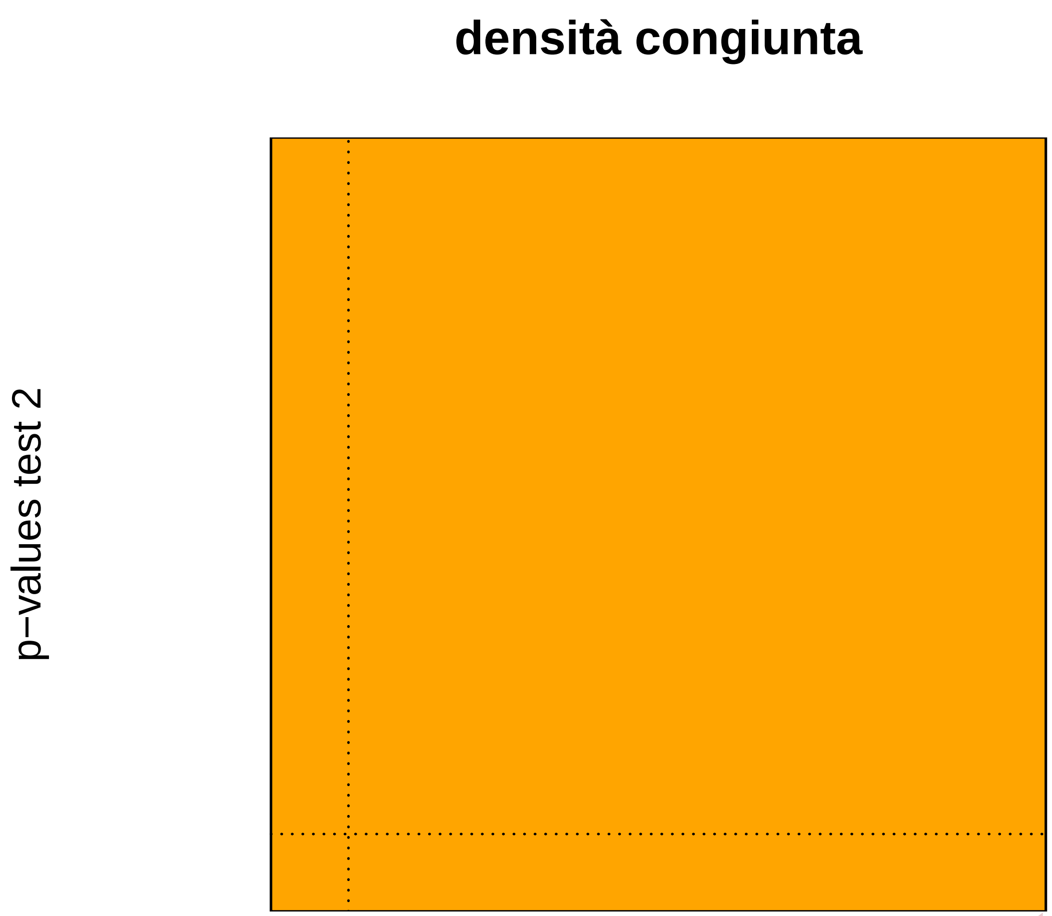
Sotto H_1 $P(p \leq .05|H_1) > .05$, nel nostro caso $= .74$



Errori di Tipo I, Due Test (indipendenti)

Propabilità di ALMENO un (falso) rifiuto?

$$P(p_1 \leq .05 \cup p_2 \leq .05 | H_0) = .05 + .05 - (.05 \cdot .05) =$$
$$1 - (1 - .05)^2 = .0975 = 1 - (1 - \alpha)^2$$



Probabilità di falsi rifiuti

m **p-value indipendenti**

Se rifiuto l'ipotesi quando $p \leq \alpha$

Probabilità ALMENO un falso rifiuto

$P = 1 - (1 - \alpha)^m$

Probabilità di falsi rifiuti

m p-value indipendenti

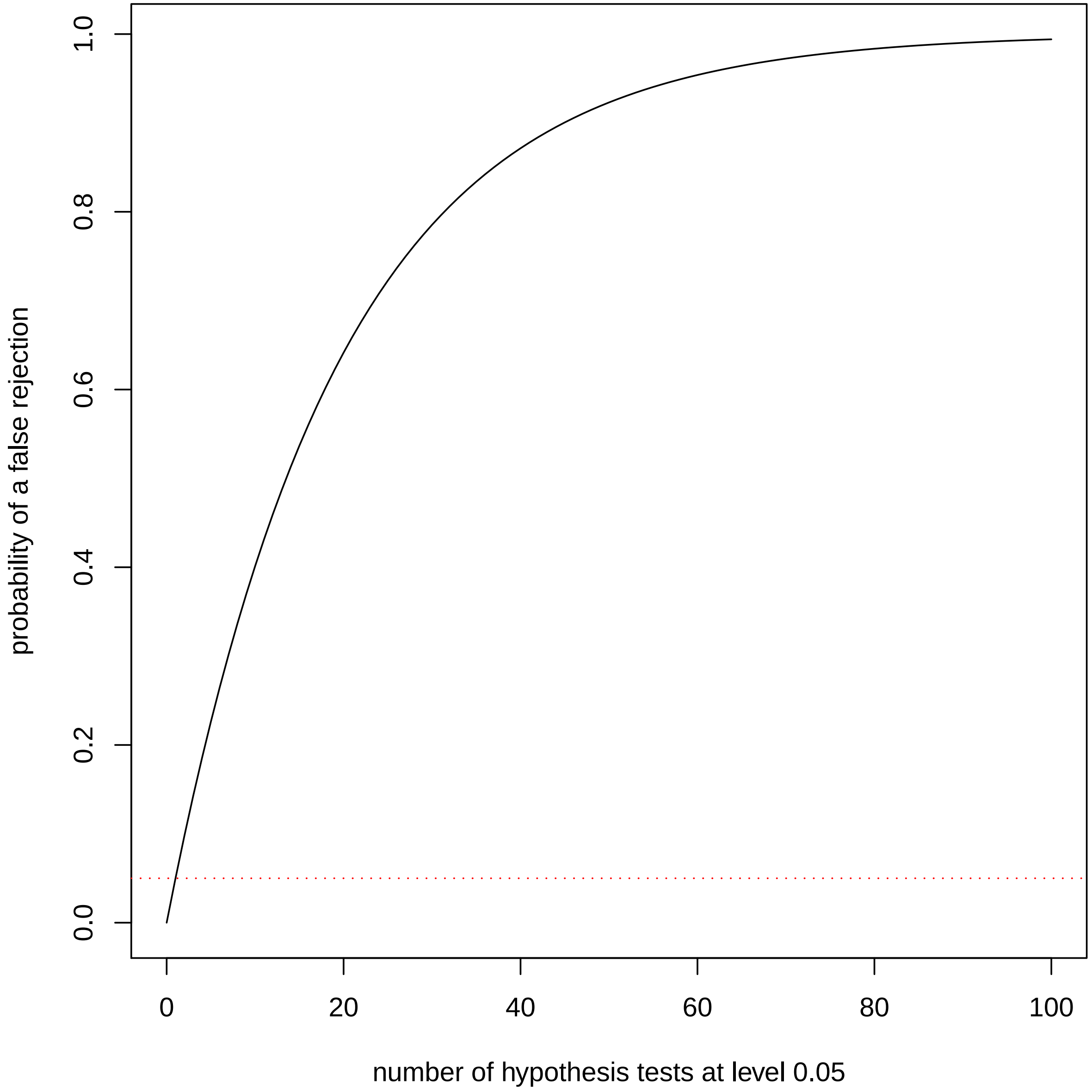
Se rifiuto l'ipotesi quando $p \leq \alpha$

Probabilità ALMENO un falso rifiuto

$P = 1 - (1 - \alpha)^m$

Questo diventa ben presto un problema, se m diventa grande ...

Errori di Tipo I in funzione del numero di test (m)



Type I errors

Come definire l'errore di tipo I quando ci sono molte ipotesi?

Quali procedure controllano questo errore?