

Approccio bayesiano e classico all'analisi ad interim negli studi clinici controllati

Antonella Bodini

`anto@mi.imati.cnr.it`

CNR-IMATI, Milano



- breve introduzione
 - il caso non sequenziale
 - il caso sequenziale
 - approccio classico
 - approccio bayesiano
- } ICON3



- Fase I a carattere esplorativo
- Fase II studi pilota su pochi pazienti
- **Fase III** gran numero di pazienti
- Fase IV tests dopo l'approvazione



- durata pluriennale della sperimentazione
- arrivo sequenziale dei pazienti
- assegnazione randomizzata a 2 (o più) trattamenti (corrente/nuovo)
- risposta (*endpoint*) continua o binaria
- dati censurati
- test di superiorità o di equivalenza



Introduzione III



Ragioni:

- Etiche
- Economiche
- Amministrative
 - fedeltà alla pianificazione
 - fedeltà al protocollo
 - verifica delle ipotesi del disegno
- Nuove informazioni



- Wald (1947, USA) → *SPRT*
Barnard (1946, GB)
- Armitage (1954, 1958, 1975)
Brass (1952, 1958)
- Shaw (1966)
- Pocock (1977)
O'Brien & Fleming (1979)
Lan & DeMets (1983)

Il caso non sequenziale



$$X_{Ai} \sim \mathcal{N}(\mu_A, \sigma^2), \quad X_{Bi} \sim \mathcal{N}(\mu_B, \sigma^2), \quad i = 1, \dots, n$$

osservazioni indipendenti

$$H_0 : \mu_A = \mu_B \quad \text{vs} \quad H_1 : \mu_A \neq \mu_B$$

$$Z = \frac{1}{\sqrt{2n\sigma^2}} \sum_{i=1}^n (X_{Ai} - X_{Bi})$$
$$\sim \mathcal{N} \left((\mu_A - \mu_B) \sqrt{n/2\sigma^2}, 1 \right) \stackrel{H_0}{=} \mathcal{N}(0, 1)$$

$$|Z| \geq z_{1-\alpha/2} \Rightarrow \text{rifiuto } H_0$$

si controlla l'errore di I specie



Il caso sequenziale

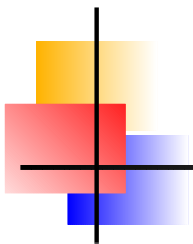


Il monitoraggio continuo è impossibile!

Monitoraggio sequenziale per gruppi:

- K = n.ro di analisi di monitoraggio
- all'analisi k -ma: statistica Z_k
 - se $|Z_k| \geq c_k$, per $k < K$ allora si rifiuta H_0
 - si prosegue fino alla fase finale.

$$c_k : P_{H_0}(|Z_1| \geq c_1 \circ \dots \circ |Z_K| \geq c_K) = \alpha$$



$$c_k \neq z_{1-\alpha/2}$$

$$\alpha = 0.05$$

K	Pocock	O'B & F
1	1.960	1.960
2	2.178	$1.977 \sqrt{K/k}$
3	2.289	$2.004 \sqrt{K/k}$
4	2.361	$2.024 \sqrt{K/k}$
5	2.413	$2.040 \sqrt{K/k}$



N oss., K i.a., una ogni N/K oss.

j – ma analisi: $d(j) = jN/K$ dati

$$X_i \sim \mathcal{N}(\theta, \sigma^2), \quad \sigma^2 \text{ noto}$$

$$Z_j = \frac{1}{\sqrt{d(j)\sigma^2}} \sum_{i=1}^{d(j)} X_i$$

$$\left\{ \begin{array}{l} (Z_1, \dots, Z_K) \sim \mathcal{N}_K \\ Z_i - Z_{i-1} \text{ ind.}; \\ E(Z_j) = \theta \sqrt{d(j)/\sigma^2}, \quad \text{Var}(Z_j) = 1; \\ \text{Corr}(Z_{j_1}, Z_{j_2}) = \sqrt{d(j_1)/d(j_2)} \quad 1 \leq j_1 \leq j_2 \leq K. \end{array} \right.$$



Osservazioni binarie

$$\theta = \log \lambda, \quad \lambda = \text{hazard rate}$$

$$H_0 : \theta = 0 \quad \text{vs} \quad H_1 : \theta \neq 0$$

$$Z_j = S_j / \sqrt{\mathcal{I}_j}$$

$$S_j = d_1(j) - \sum_{i=1}^{d(j)} \frac{n_{0i}(j)}{n_{0i}(j) + n_{1i}(j)}$$

$$\mathcal{I}_j = \sum_{i=1}^{d(j)} \frac{n_{0i}(j) n_{1i}(j)}{[n_{0i}(j) + n_{1i}(j)]^2}$$

$d_1(j)$ decessi nel gruppo di trattamento ($t = 1$)

$n_{ti}(j)$ sopravvissuti nel gruppo $t = 0, 1$



S_j è la logrank score statistics

\mathcal{I}_j informazione per θ

$$\theta \approx 0 \Rightarrow S_j \sim \mathcal{N}(\theta \mathcal{I}_j, \mathcal{I}_j)$$

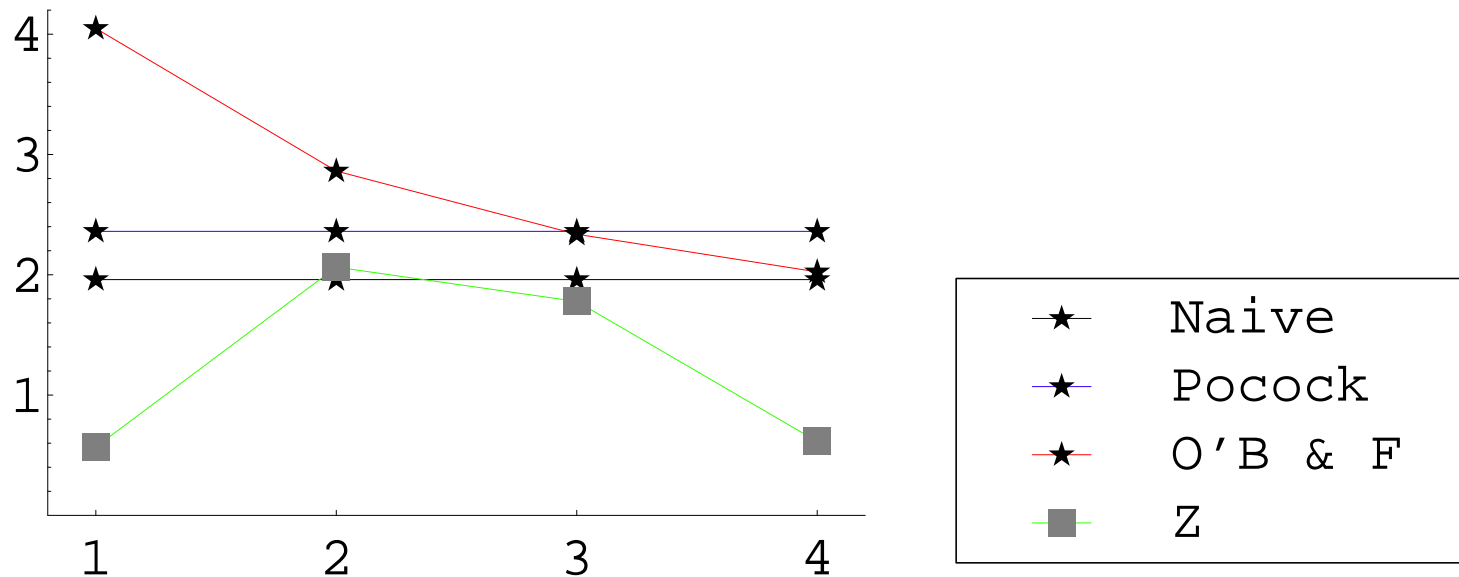
$$\& n_{1i}(j) \approx r n_{0i}(j) \forall i \Rightarrow \mathcal{I}_j \approx d(j) r / (r + 1)^2$$

Harrington *et al.*, 1982; Tsiatis, 1982.

$N = 1286$	1 (tratt.)		0 (contr.)		HR (λ)	\mathcal{I}_j
j	$d_1(j)$	$e_1(j)$	$d_0(j)$	$e_0(j)$	(1)	(2)
1	103	107.818	220	215.182	0.934	71.764
2	196	220.668	447	422.332	0.839	144.829
3	310	336.042	655	628.958	0.886	218.839
4	439	449.388	847	836.612	0.965	292.040

(1) HR: SAS, by Istituto Mario Negri

(2) $r = 0.5; \mathcal{I}_j \approx 2d(j)/9.$



Non si termina mai in anticipo
e si accetta $\theta = 0$, *equivalenza*



Intervalli di confidenza

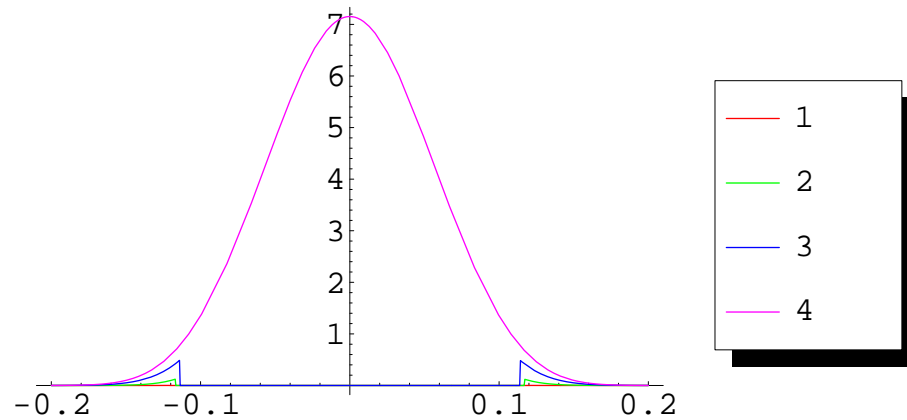
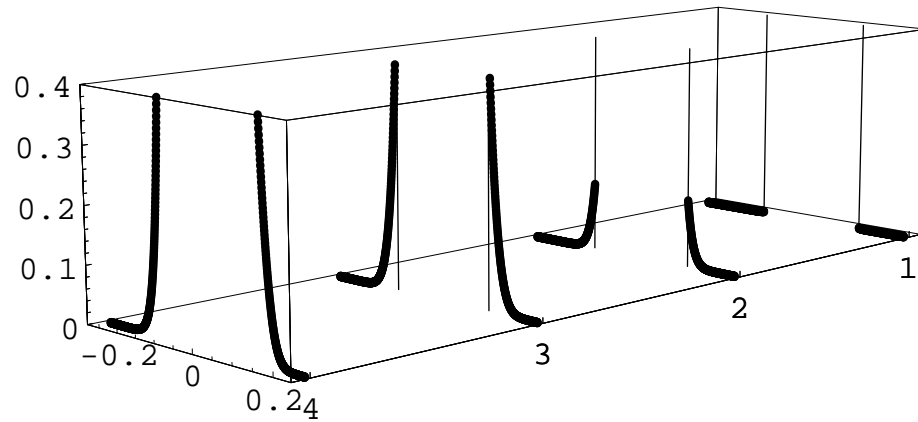


Tsiatis *et al.*, 1984:

H_p. gaussianità, $r = 1$.

Viene determinata la densità della variabile (Z, τ) , con τ tempo d'arresto e Z statistica osservata all'istante di arresto. Quindi, numericamente, si può ottenere l'intervallo di confidenza esatto.

ICON3, cont.



$$IC(95\%) = (0.868, 1.281)$$



Repeated confidence interval approach

Jennison & Turnbull, 1984-1989

Un intervallo di confidenza ad ogni analisi intermedia

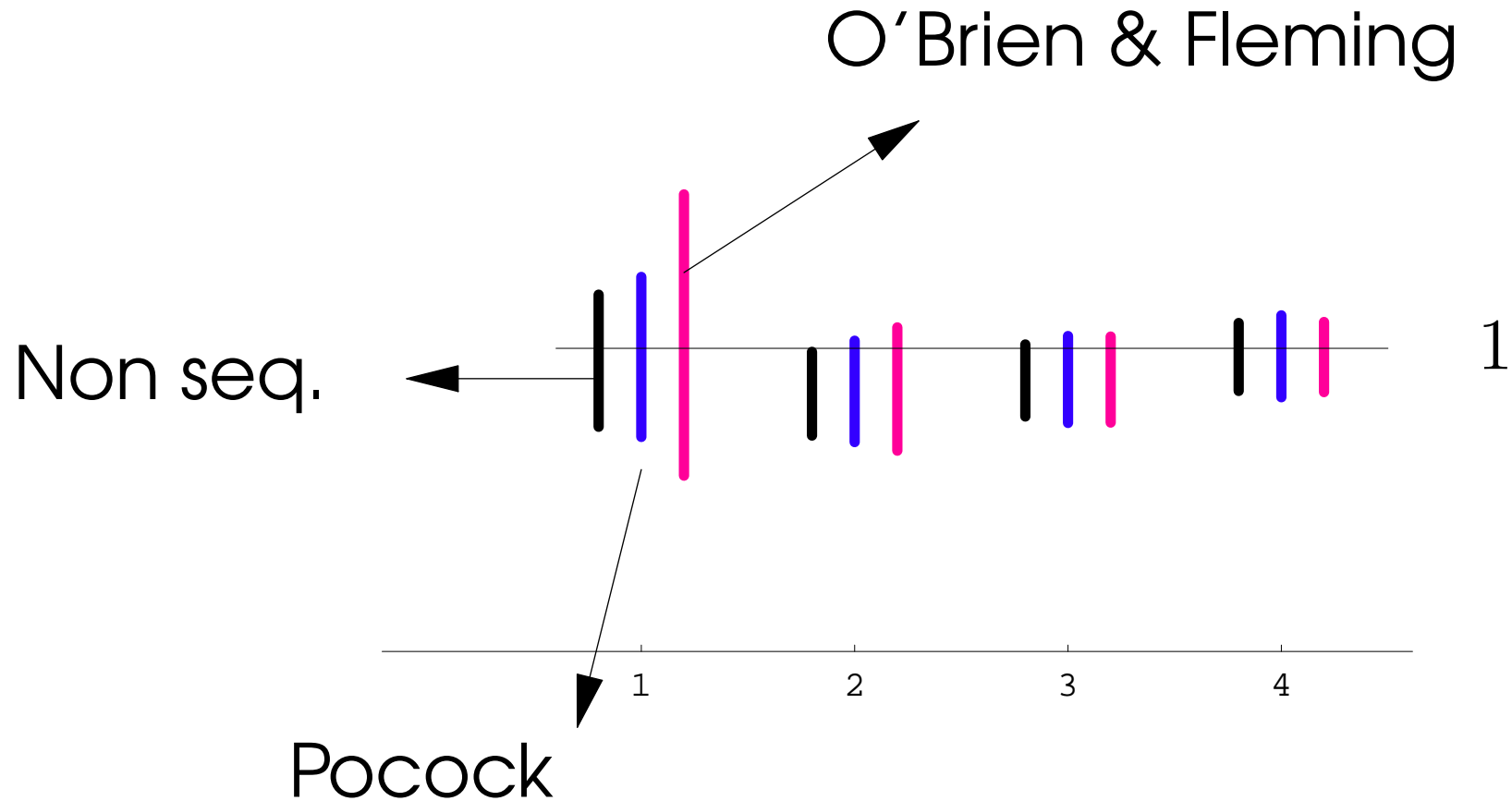
$\{I_1, \dots, I_K\}$ sequenza di RCIs con livello $1 - \alpha$

se

$$P_{\theta}(\theta \in I_j \forall j = 1, \dots, K) = 1 - \alpha$$

ottenuti invertendo un test sequenziale:

$$I_j = \{\theta : |Z_j - \theta\sqrt{I_j})| < c_j\}$$





$\{f_\theta(x) : \theta \in \Theta\}$ modello statistico

estensione del modello:

$p(\theta)$ densità di probabilità (*prior*)

$$\Rightarrow f(x, \theta) = f_\theta(x)p(\theta)$$

Teorema di Bayes:

$$f(\theta|x) \propto f_\theta(x)p(\theta)$$

$$f(\theta|x_1, \dots, x_n) \propto \prod_i f_\theta(x_i)p(\theta)$$

distribuzione a posteriori



- $\hat{\theta} = \int_{\Theta} \theta p(\theta|x_1, \dots, x_n) d\theta$
- mediana, quartili ...
- $I : p(\theta \in I|x_1, \dots, x_n) = 0.95, \quad CrI(95\%)$
- pr. a posteriori di **qualunque** $I \subset \Theta$
- $f(x_{n+1}|x_1, \dots, x_n) =$
$$\int_{\Theta} f_{\theta}(x_{n+1})p(\theta|x_1, \dots, x_n)d\theta$$

distr. predittiva



Interim analysis bayesiana



Spiegelhalter & Freedman, 1988

Ad **ogni** interim analysis:

$$S(j)/\mathcal{I}_j = Z(j)/\sqrt{\mathcal{I}_j} \sim \mathcal{N}(\theta, \mathcal{I}_j^{-1})$$

$$\text{PRIOR: } p_0(\theta) \sim \mathcal{N}(\theta_0, \tau^{-2})$$

$$\mathcal{I}_j = r d(j)/(r + 1)^2 \Rightarrow \tau^2 = r n_0/(r + 1)^2$$

PRIOR \propto AD UNA VEROSIMIGLIANZA DA UN TRIAL
IPOTETICO CON n_0 PAZIENTI E VALORE OSSERVATO
DELLA DIFFERENZA PARI A θ_0



posterior

$$p(\theta | s_j) \sim \mathcal{N} \left(\frac{\tau^2 \theta_0 + s_j}{\tau^2 + \mathcal{I}_j}, \frac{1}{\tau^2 + \mathcal{I}_j} \right)$$

$$\hat{\theta}_j = \frac{(r + 1)^2 s_j + r n_0 \theta_0}{r[n_0 + d(j)]}$$

$$CrI_j = \hat{\theta} \pm \frac{(r + 1) z_{1-\alpha/2}}{\sqrt{r[n_0 + d(j)]}}$$

STOP: $p(\theta < 0 | s_j) < \epsilon$ \circ $p(\theta < 0 | s_j) \geq 1 - \epsilon$

Scelta della prior



- evidenza esterna al trial
- può variare durante il trial

Alcune scelte "naturali": $\mathcal{N}(\theta_0, (r + 1)^2 / rn_0)$

- $\theta_0 = 0, n_0 = 0$ non informativa
- $\theta_0 = 0, n_0 : p_0(\theta > \theta_1) \leq 0.05$ Scettica
- $\theta_0 = \theta_1, n_0$ come sopra. Entusiasta

Altre possibilità.

ICON3, cont.



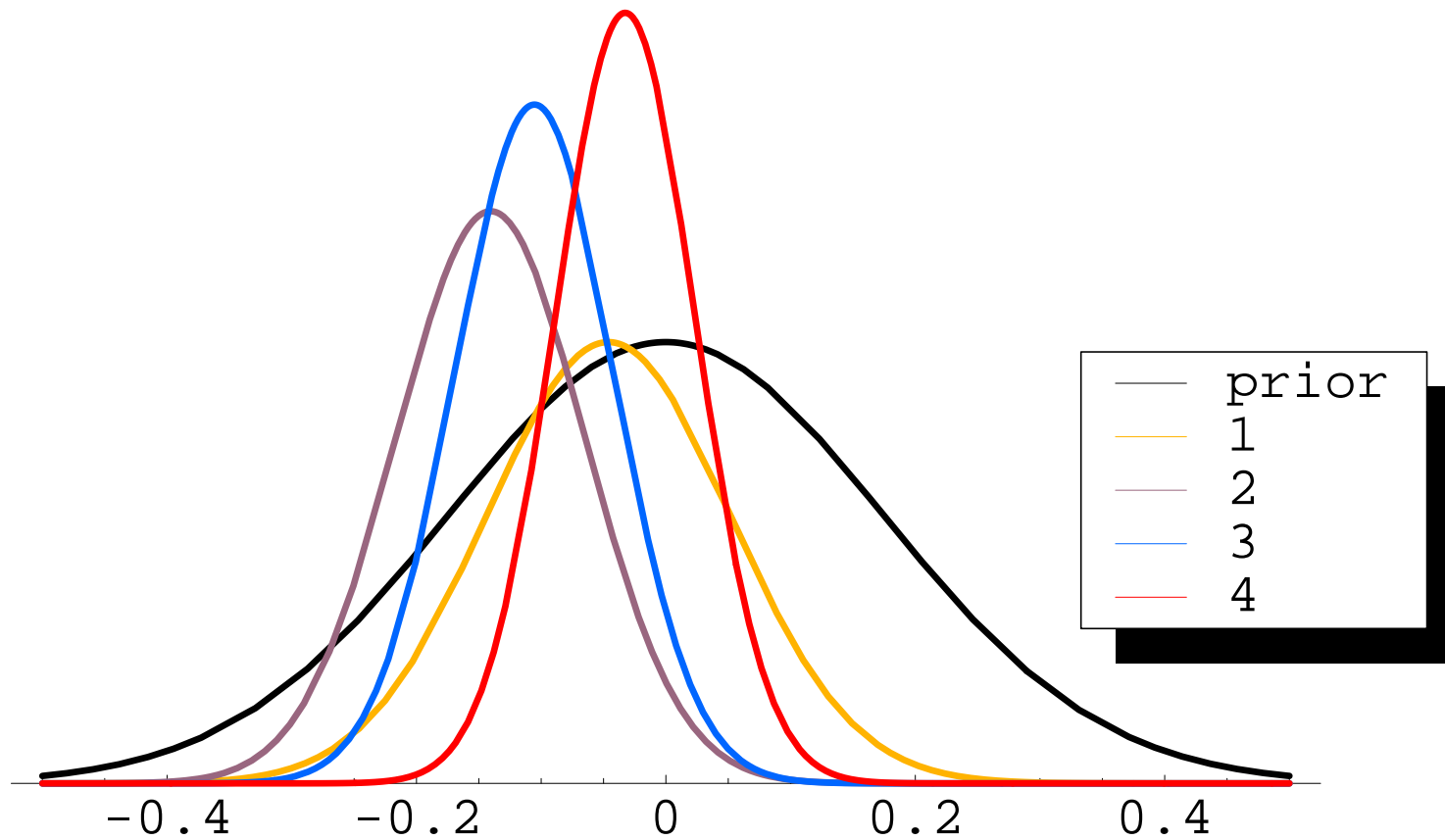
Inferenza per $\lambda = \exp(\theta)$

prior	analisi	$\hat{\lambda}$	CrI(95%)
non informativa $\theta_0 = 0, n_0 = 0$	1	0.935	(0.742, 1.178)
	2	0.841	(0.714, 0.991)
	3	0.886	(0.775, 1.012)
	4	0.964	(0.856, 1.083)
scettica $\theta_0 = 0, n_0 = 147^*$	1	0.955	(0.788, 1.158)
	2	0.869	(0.749, 1.007)
	3	0.900	(0.794, 1.019)
	4	0.968	(0.867, 1.080)
entusiasta $\theta_0 = 0.344^{**}, n_0 = 147^*$	1	1.063	(0.878, 1.288)
	2	0.926	(0.799, 1.074)
	3	0.942	(0.831, 1.067)
	4	1.003	(0.898, 1.119)

* n_0 tale che $p_0(\theta_1, +\infty) \leq 0.025$ per $\theta_1 = \log 1.41$

** $\theta_0 = \theta_1$

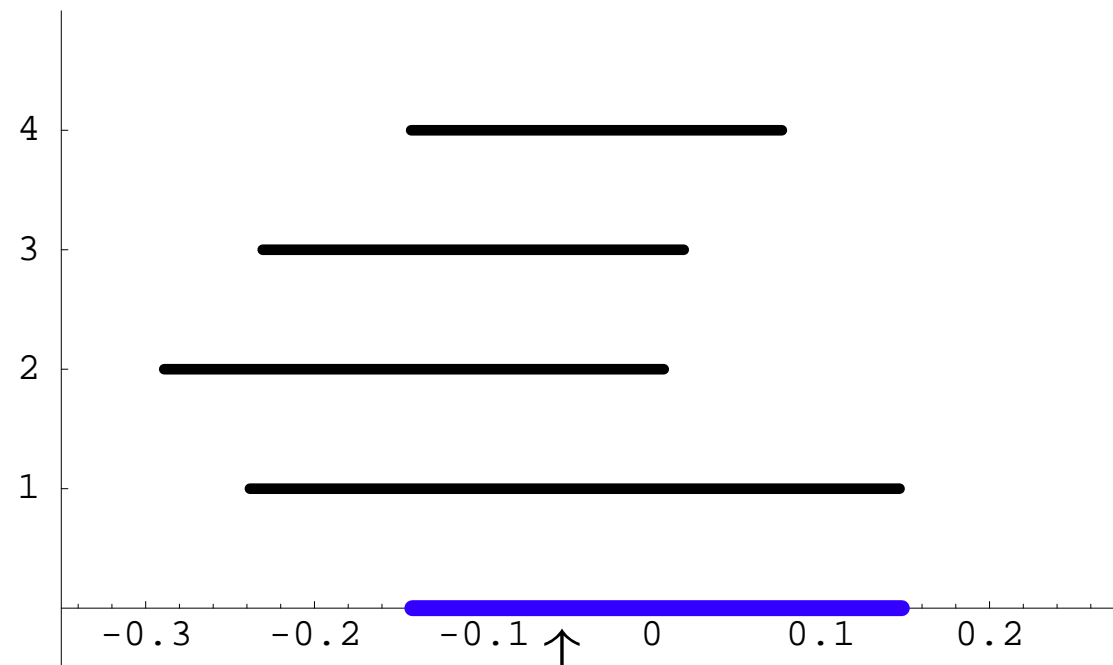
ICON3, cont.



ICON3, cont.



prior scettica:



int. di *equivalenza*

$[-0.142, 0.148]$

ICON3, cont.



Probabilità a posteriori di miglioramento = 0.05, 0.10, 0.20.

prior	j	$\hat{\theta}$	$P(\theta > \theta_{\Delta} s_j)$
non inform. $\theta_0 = 0, n_0 = 0$	1	-0.067	(0.03, 0.001, 0)
	2	-0.173	(0, 0, 0)
	3	-0.121	(0, 0, 0)
	4	-0.037	(0.001, 0, 0)
scettica $\theta_0 = 0, n_0 = 147$	1	-0.046	(0.024, 0.0002, 0)
	2	-0.140	(0, 0, 0)
	3	-0.105	(0, 0, 0)
	4	-0.033	(0.001, 0, 0)
entusiata $\theta_0 = 0.344, n_0 = 147$	1	0.061	(0.19, 0.006, 0)
	2	-0.077	(0.001, 0, 0)
	3	-0.060	(0.0005, 0, 0)
	4	-0.003	(0.005, 0, 0)



- apprendimento dall'esperienza
- incorporare informazioni diverse
- strumenti **aggiuntivi** di valutazione
 - valutazioni di probabilità per θ
 - range di punti di vista da considerare
- elicitazione della prior
- distribuzione predittiva



Bibliografia essenziale



- FAYERS P.M., ASHBY D. AND PARMAR, M.K.B. (1997). Tutorial in Biostatistics. Bayesian Data Monitoring in Clinical Trials. *Statist. Med.* **16**, 1413-1430.
- JENNISON C. AND TURNBULL B.W. (1990). Statistical approaches to Interim monitoring of medical trials: a review and commentary. *Statist. Sc.* **5**, 299-317.
- JENNISON C. AND TURNBULL B.W. (2000). *Group Sequential Methods with Applications to Clinical Trials* Chapman & Hall/CRC. Boca Raton.
- SPIEGELHALTER D.J., FREEDMAN L.S. AND PARMAR, M.K.B (1994). Bayesian approaches to randomized trials (with discussion). *J. Roy. Statist. Soc. A.* **157**, 357-416.



- <http://www.mi.imati.cnr.it/iami/abstracts/04-14.html>
- Approfondire, con l'analisi di dati reali ma anche tramite simulazioni, il contributo dell'approccio bayesiano all'interim analysis.