Estimation de l'effet d'un traitement Statistiques de test et proportionnalité des risques

A. Latouche
Université Versailles St–Quentin, EA 2506

Contexte médical

Essai clinique randomisé

- n patients 2 groupes de traitement
- $n = n_C + n_E$
- Délais de survenue sont censurés
- On observe $min(T, C)$
- Si évènements multiples et exclusifs : évènements concurrents
- On observe aussi l'indicatrice du type d'evt ε

Illustration

Efficacité d'un traitement : réduction de la mortalité cardiovasculaire

n patients

- Infarctus du myocarde ($\varepsilon = 1$)
- Décès de cause cardiovasculaire ($\varepsilon = 2$)
- Décès d'autre causes ($\varepsilon = 3$)
- Censuré ($\varepsilon = 0$)

Effet du traitement

Quantifier, mais sur quelle quantité ?

- Sur un risque ?
- Sur la fonction de survie ?
- Sur l'incidence cumulée ?

Notations

On observe $\tilde{T} = min(T, C)$ et ε

Le risque cause–spécifique pour l'évt 1 dans le groupe Experimental est :

$$\lambda_{1E}(t) = dF_{1E}(t)/S_E(t)$$

où

$$F_{1E}(t) = \Pr(T \leq t, \varepsilon = 1)$$

est l'incidence cumulée.

$S_E(t) = 1 - (F_{1E}(t) + F_{2E}(t))$ *event free survival*

Hypothèse de proportionnalité

Soit h une fonction de risque (*hazard*) , la proportionnalité s'exprime :

$$h_{1E}(t) = h_{1C}(t)\ \mathrm{HR}$$

où HR est le rapport des risques (*hazard ratio*)

$$H_0 : h_{1E}(t) = h_{1C}(t) \iff \mathrm{HR} = 1.$$

Test(s) du log–rank

H_0: les probabilités instantanées de décès sont les mêmes entre les deux groupes

- Analyse de survie : équivalence entre risque et survie
- Test fondé sur les différences pondérées integrées entre
  - Risque cumulé estimé dans le bras de traitement
  - Risque cumulé estimé en réunissant les 2 bras de traitement
- Tarone–Ware (1977), Fleming Harrington (1981)
- Ces tests sont implémentés donc utilisés
- Optimaux pour des risques proportionnels

En pratique

- Risques (cumulés) divergent, convergent
- Log–rank pondéré(s) puissance faible

Supremum log–rank

Soit $V_n(t)$ la statistique de test du log–rank pondéré

- $\sigma_n^2(t)$ la variance de $V_n(t)$
- τ la durée totale de l'essai
- $T_n(t) = V_n(t)/\sigma_n(t)$ log–rank standardisé
- $\sup_{t \in [0,\tau]} |T_n(t)|$ le supremum log–rank

Propriété du supremum log–rank

- Sensibilité accrue à des alternatives stochastiques ordonnées
- Légère augmentation de la taille (5%) mais gain de puissance
- Performance similaire à celle du log–rank, dans une situation où le log–rank est optimale

(Eng et Kosorok, Biometrics. 2005)

Comparaison d'incidence cumulées: test de Gray

$H_0 : F_{1E} = F_{1C}$

De la forme

$$Z_{.} = \int_0^{\tau} W_{.}(u) \left\{ d\hat{\Gamma}_{1.}(u) - d\hat{\Gamma}_1^0(u) \right\},$$

- τ la durée totale de l'essai
- $W_{.}$ pondération
- $\hat{\Gamma}_{1.}(t)$ Risque cumulée estimé dans le bras de traitement "$\cdot$"
- $\hat{\Gamma}_1^0(t)$ Risque cumulée estimé en réunissant les 2 bras de traitement

La fonction de risque est ici : $\alpha_{1.}(t) = dF_{1.}(t)/(1 - F_{1.}(t))$

Test implémenté dans `R` et optimal pour des risques proportionnels

Test de Bajorunaite–Klein

- Modification du test de Gray (*supremum Gray test*)

$$Q_1 = \sup\{|Z_1(t)|,\ t \le \tau\}/\hat{\sigma}(\tau)$$

avec

$$Z_1(t) = \int_0^t W_1(u)\{d\hat{\Gamma}_{11} - d\hat{\Gamma}_1^0\}(u).$$

Sous les hypothèses du Théorème 1 de Gray:

– $Z_1(t)/\sigma(\tau)$ converge vers $B(\sigma(t)/\sigma(\tau))$

– $\sup\limits_{0 \le t \le \tau} \frac{|Z_1(t)|}{\hat{\sigma}_1(\tau)} \xrightarrow{d} \sup\limits_{0 \le t \le 1} |B(t)|$

- Sensibilité accrue à des alternatives ordonnées
- Calcul de la variance simplifiée
- Calcul du nombre de sujets nécessaire (implémenté)

Comportement des tests

Simuler des temps d'évènements sous les hypothèses suivantes pour les Fonctions de risques causes–spécifique et de sous–répartitions

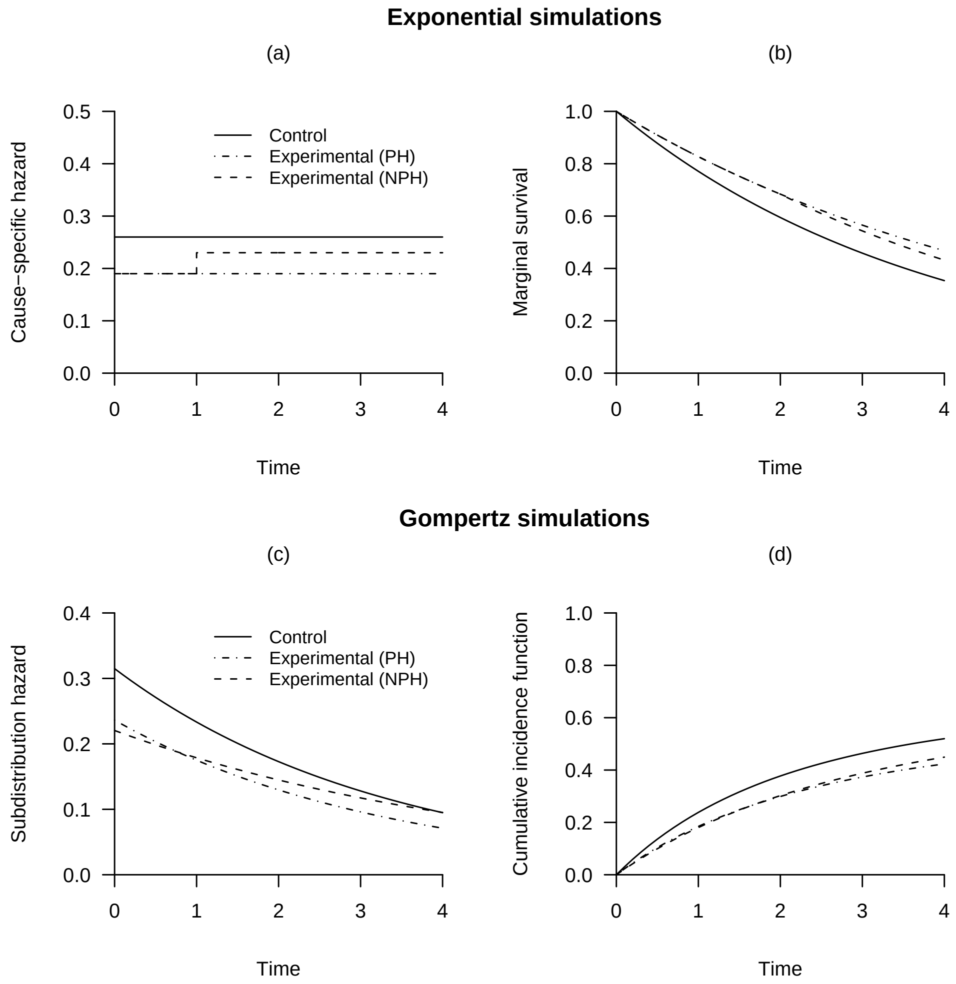
(a) sous H_0

(b) alternative à risque proportionnel

(c) alternative à risques non–proportionnel

Evaluer les performances des 4 tests

Simulation



Exponential simulations
(a)
Cause−specific hazard
0.5
0.4
0.3
0.2
0.1
0.0
0
1
2
3
4
Time
Control
Experimental (PH)
Experimental (NPH)
(b)
Marginal survival
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
4
Time
Gompertz simulations
(c)
Subdistribution hazard
0.4
0.3
0.2
0.1
0.0
0
1
2
3
4
Time
Control
Experimental (PH)
Experimental (NPH)
(d)
Cumulative incidence function
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
4
Time

Simulation : Exponentiel

	$N = 1002$			$N = 1190$		
Test	Size	Power	Power	Size	Power	Power
Log-rank	0.053	0.908	0.796	0.051	0.942	0.847
Supremum log-rank	0.049	0.890	0.808	0.048	0.928	0.861
Gray	0.053	0.880	0.755	0.050	0.923	0.813
Bajorunaite	0.050	0.863	0.783	0.046	0.910	0.835

- Risques proportionnels
- Risques non–proportionnels

Simulation : Gompertz

	$N = 1002$			$N = 1190$		
Test	Size	Power	Power	Size	Power	Power
Log-rank	0.050	0.695	0.674	0.050	0.765	0.742
Supremum log-rank	0.046	0.686	0.697	0.049	0.749	0.763
Gray	0.050	0.860	0.743	0.051	0.912	0.812
Bajorunaite	0.047	0.836	0.767	0.049	0.888	0.826

$$F_{k.}(t) = 1 - \exp\left[\beta_{k.}\left\{1 - \exp(\nu_{k.}t)\right\}/\nu_{k.}\right],$$

où $\nu_{k.} < 0$ et $|\beta_{k.}| < \infty$.

$\Rightarrow \alpha_{1.}(t) = \beta_{1.}\exp(\nu_{1.}t)$

Perspectives

Travail collaboratif avec R. Bajorunaite (Université de Marquette)

- Implementation des tests de Bajorunaite–Klein (p–value)
- Comparaison avec les Tests de type kolmogorov–smirnov (Test de Lin, 1997)
- Version supremum du test de Pepe (1991)

Références

- The versatility of function–indexed Weighted Log–Rank Statistics. Kosorok MR, Lin CY. JASA, 1999
- Klein, JP and Bajorunaite, R. Chapter 16 Inference for Competing Risks. Handbook of Statistics. Vol 25, Advances in Survival Analysis. Elsevier Science, 291-312, 2004
- Two–sample tests of equality of two cumulative incidence functions. R. Bajorunaite, JP Klein. Computational Statistics & Data analysis. 2007
- Sample size calculations in the presence of competing risks. A. Latouche, R. Porcher. Statistics in Medicine, 2007