

TESTS DE COMPARAISON DE COURBES DE SURVIE

Cécile Chauvel ¹ & John O'Quigley ²

*Laboratoire de Statistique Théorique et Appliquée,
Université Pierre et Marie Curie - Paris 6
4 place Jussieu, Tour 15-25, Bureau 228, 75252 Paris Cedex 05*

¹ *cecile.chauvel@upmc.fr*

² *john.oquigley@upmc.fr*

Résumé. Nous décrivons une classe de tests statistiques pour la comparaison de courbes de survie. Cette classe est basée sur la construction du processus du score standardisé d'O'Quigley (2003). La distribution limite de ce processus permet de construire plusieurs tests de la présence d'un effet qui conservent de bonnes puissances même lorsque les hypothèses du modèle à risques proportionnels ne sont pas vérifiées. Un test adaptatif combinant les membres de la classe, dont le log-rank et le log-rank pondéré font partie, peut être construit. Ce test est puissant sous des alternatives à la fois de type risques proportionnels et non-proportionnels. Des simulations pour des échantillons de taille modérée confortent ces conclusions. Une comparaison numérique avec les puissances de tests existants seront présentées, ainsi qu'une illustration sur des données réelles.

Mots-clés. Test Adaptatif, Log-Rank, Risques Non-Proportionnels, Processus du Score, Survie.

Abstract. We describe a class of statistical tests for the comparison of two or more survival curves. The class is based on the construction of the standardized score process of O'Quigley (2003). The study of the limiting distribution of the process implies the construction of several tests for the presence of the effect maintaining good power when the proportional hazards model assumptions are not met. The log-rank test and weighted log-rank tests arise as special members of the class. An adaptive test combining members of the class can be derived, maintaining good power under both proportional hazards and non-proportional hazards alternatives. Simulations under the null as well as different types of alternatives for moderate sample sizes support this conclusion. In addition, numerical power comparisons with existing tests for the presence of an effect are presented. An example is given as an illustration.

Keywords. Adaptive test, Log-Rank, Non-Proportional Hazards, Score Process, Survival.

1 Contexte et résultats

Les essais cliniques randomisés comparant des données de survie ont pour but de tester l'hypothèse nulle d'absence de l'effet du traitement contre l'hypothèse alternative d'une

différence de survie entre le nouveau traitement et le traitement de référence ou le placebo. Le test du log-rank (Mantel 1966) est classiquement utilisé dans cette situation puisqu'il est localement le plus puissant sous le modèle à risques instantanés proportionnels (Peto et Peto 1972). Ce modèle est défini par

$$\lambda(t|Z) = \lambda_0(t) \exp(\beta Z(t)),$$

où $\lambda_0(\cdot)$ est un risque instantané de base, $Z(t)$ un vecteur de covariables pouvant dépendre du temps et β le vecteur de coefficients de régression associé.

Cependant, lorsque les risques ne sont pas proportionnels, les propriétés optimales de ce test ne sont plus vérifiées, pouvant entraîner de grandes pertes de puissance. Afin d'étudier ces situations, nous considérons un modèle plus général appelé modèle à risques non-proportionnels et défini par

$$\lambda(t|Z) = \lambda_0(t) \exp(\beta(t)Z(t)), \tag{1}$$

où le coefficient de régression $\beta(t)$ peut dépendre du temps. Sous ce modèle, on peut définir un processus du score standardisé en étendant les travaux d'O'Quigley (2003), trouvant leur origine dans ceux de Wei (1984). Les propriétés de convergence de ce processus stochastique permettent de construire une classe de tests asymptotiques de l'hypothèse nulle $H_0 : \beta(t) = 0$ contre l'alternative $H_a : \beta(t) \neq 0$.

La statistique d'un des tests de la classe converge presque sûrement vers celle du log-rank. Cette statistique permet d'obtenir un test puissant sous une hypothèse alternative pour laquelle les risques sont proportionnels. Les autres statistiques de la classe sont puissantes sous certaines alternatives raisonnables de type risques non-proportionnels, notamment lorsque l'effet diminue ou change de signe dans le temps. Une combinaison de ces statistiques permet de construire un test adaptatif puissant à la fois sous une hypothèse alternative de type risques proportionnels et risques non-proportionnels. Ce test est adaptatif au sens où la combinaison optimale dépend des données.

2 Présentation de l'exposé

Au cours de l'exposé, nous présenterons le processus stochastique du score standardisé et ses propriétés de convergence. Ceci nous amènera à décrire la classe de tests asymptotiques de la nullité du paramètre du modèle. Une étude de simulation sous différents types d'alternatives sera présentée, dans laquelle nous comparerons ces nouveaux tests à des tests de référence, notamment le log-rank. Nous illustrerons ces tests par une application sur un jeu de données réelles.

Bibliographie

- [1] Mantel, N. (1966), Evaluation of survival data and two new rank order statistics arising in its consideration, *Cancer Chemotherapy Reports*, 50, 163–70.
- [2] O’Quigley, J. (2003), Khmaladze-type graphical evaluation of the proportional hazards assumption, *Biometrika*, 90, 577–584.
- [3] Peto, R. et Peto, J. (1972), Asymptotically efficient rank invariant test procedures, *Journal of the Royal Statistical Society Series A*, 135, 185–207.
- [4] Wei, L. (1984), Testing goodness-of-fit for proportional hazards model with censored observations, *Journal of the American Statistical Association*, 79, 649–652.