
Estimation of survival in age-dependent genetic disease with sporadic cases from pedigree data

Lucas Ducrot*¹ and Grégory Nuel

¹Laboratoire de Probabilités, Statistique et Modélisation – Sorbonne Université, Centre National de la Recherche Scientifique, Université Paris Cité, Sorbonne Université :
UMR_s001, *Centre National de la Recherche Scientifique : UMR_s001, Université Paris Cité :*
UMR_s001 – France

Résumé

Dans le contexte de maladies génétiques avec faible fréquence allélique en population générale et forte pénétrance (maladies mendéliennes par exemple), les approches familiales sont souvent intéressantes. En effet, les patients ont généralement une famille sévèrement touchée par la maladie et sont donc adressés au généticien. Par ailleurs, l'estimation de la survie pour des maladies dépendantes de l'âge est nécessaire pour la mise en place de protocoles médicaux et le suivi des patients.

Le problème principal pour effectuer ces estimations est que les génotypes sont souvent non-observés et doivent être traités comme une variable latente. Dans le cadre spécifique où la maladie ne présente pas de cas sporadique (i.e. seuls les porteurs de mutations délétères peuvent être affectés par la maladie), le problème est plus simple à traiter car un malade est, de fait, automatiquement un porteur de mutation. L'incertitude sur les génotypes repose donc sur les personnes non-affectées. Dans ce scénario, une méthode utilisant des algorithmes d'espérance-maximisation et d'Elston-Stewart a déjà été publiée.

Cependant, la plupart des maladies affectent à la fois les porteurs et non-porteurs de mutations délétères à des fréquences différentes. La méthode proposée a pour but de prendre en compte les cas sporadiques et généraliser l'estimation de la survie pour les maladies génétiques. Pour cela la méthode s'appuie sur deux hypothèses principales: l'incidence de la maladie en population génétique est connue et constante par morceau, le risque instantané relatif entre les porteurs et non-porteurs de mutation est constant par morceau.

Le modèle est un mélange de survies paramétrisé par le risque instantané relatif et par la proportion de porteurs de mutation dans la population. A paramètres fixés, les risques instantanés pour les porteurs et les non-porteurs peuvent être calculés sous la contrainte d'incidence en population générale grâce à une méthode de point fixe. Avec les données de prédigé, les paramètres sont estimés en maximisant la vraisemblance du modèle.

*Intervenant