
NADARAJAH THARSHANNA, Memorial University of NL
Sélection de variables pour l'analyse de données longitudinales

Les données longitudinales de grande dimension présentant un grand nombre de covariables sont de plus en plus fréquentes en biologie. Or il est difficile de calculer la fonction de vraisemblance conjointe pour les données longitudinales, notamment pour les données discrètes corrélées sur les résultats. Dans cette situation, nous proposons une vraisemblance empirique pénalisée fondée sur des équations d'estimation généralisées (GEE). Dans cette méthode, on peut se contenter de spécifier les deux premiers moments marginaux et une structure de corrélation. Des études de simulation montrent que lorsque les hypothèses du modèle sont exactes, la performance de celui-ci est comparable à celle des méthodes existantes et que même en cas d'erreur de spécification du modèle, notre méthode présente des avantages évidents par rapport aux méthodes existantes.