
BEN GOODRICH, Columbia University

Évaluation de l'hypothèse de « données manquantes au hasard » après l'imputation multiple de données manquantes

La plupart des algorithmes d'imputation multiple de valeurs manquantes reposent sur l'hypothèse de « données manquantes au hasard » (MAH), qui affirme grosso modo que la probabilité qu'une donnée soit manquante est conditionnellement indépendante des valeurs de celle-ci. Par conséquent, plusieurs chercheurs croient que les données complétées résultantes ne peuvent pas falsifier l'hypothèse MAH. Cependant, la présence de données observées parmi les données complétées peut constituer un point d'appui. Nous étudions des manières de falsifier l'hypothèse MAH en appliquant des modèles de sélection et une analyse de la structure de covariance aux données complétées. L'efficacité de ces techniques est évaluée à l'aide de simulations.