

Statistique bayésienne distribuée, effet centre et power prior

jeudi 16 novembre 2023 11:25 (25 minutes)

L'exploitation des données de santé suppose le respect de contraintes réglementaires fortes notamment lors de l'analyse des données de plusieurs établissements (ou centres). Porter les analyses statistiques, et les algorithmes d'apprentissage, vers les données, et non plus agréger les données, permet de résoudre ces problèmes réglementaires.

L'inférence bayésienne se prête à ce cas des statistiques, dites distribuées, par la mise en commun des distributions *a posteriori* des différents centres (et non de leurs données). Une difficulté vient quand on cherche un effet centre ou que les données d'un centre sont biaisées. On peut assimiler ce biais à l'hétérogénéité d'un centre par rapport aux autres et utiliser la technique du *power prior*. Cette technique, efficace notamment dans les GLM en conjonction avec le *g-prior* pour les coefficients du prédicteur linéaire, permet de mesurer l'hétérogénéité entre des centres « historiques » et le centre « courant ». Elle pondère la vraisemblance de chaque centre historique par un coefficient α_0 sur lequel un prior Beta est placé.

Si le centre biaisé est parmi les centres historiques le α_0 correspondant sera faible, différent des autres. Le problème est que le biais n'est pas visible si le centre biaisé est le centre courant. On propose alors une « cartographie » de l'hétérogénéité en utilisant tour à tour chaque centre comme courant et les autres comme historiques. Le produit entre α_0 et la taille de chaque centre permet de repérer le centre biaisé, lorsqu'il est en position « historique » par un produit moindre que celui des autres centres. Lorsque le centre biaisé est en position « courante », tous les produits sont du même ordre de grandeur. Le centre biaisé doit ensuite être supprimé ou laissé dans les centres « historiques » pour compenser le biais sur les estimations des paramètres du prédicteur linéaire.

Auteur principal: SAULEAU, Erik-A. (Université de Strasbourg, Laboratoire ICube UMR7357)

Co-auteur: FAYE, Bineta (Hôpitaux Universitaires de Strasbourg)

Orateur: SAULEAU, Erik-A. (Université de Strasbourg, Laboratoire ICube UMR7357)

Classification de Session: Statistique bayésienne