

Projet Statistique Computationnelle (Partie Bayésienne)

Franck Corset

M2 pro SSD

Présentation du sujet

On s'intéresse ici à un modèle de mélange. Le jeu de données que vous allez traiter concerne un indice d'acidité mesuré dans un échantillon de 155 lacs de l'Amérique du Nord. La variable est donnée en log. Ce jeu de données a déjà été traité par Crawford et al. [1].

L'objectif dans ce projet est donc de proposer une modélisation et de faire une analyse bayésienne de ces données.

Vous commenterez également les résultats de vos analyses bayésiennes en indiquant, éventuellement, les hypothèses faites sur les modèles utilisés.

Au final, il faudra rendre un rapport qui détaille

- Une présentation des données,
- votre démarche scientifique ainsi que les hypothèses faites,
- les conclusions des différentes analyses réalisées,
- des perspectives d'améliorations des méthodes proposées,
- une bibliographie donnant vos sources.

Vous pouvez travailler par groupe de 4 maximum.

Idéalement, ce rapport sera fait en Rmarkdown afin de produire un document pdf. Dans ce cas, vous enverrez les deux fichiers (.Rmd et .pdf). Sinon, vous pouvez écrire le document en \LaTeX (les documents word seront quand même acceptés). Il faudra également rendre vos fichiers scripts R. La date

limite pour rendre le projet est le lundi 17 février 2020. Les fichiers seront envoyés à franck.corset@univ-grenoble-alpes.fr.

Références

- [1] Sybil L Crawford, Morris H DeGroot, Joseph B Kadane, and Mitchell J Small. Modeling lake-chemistry distributions : Approximate bayesian methods for estimating a finite-mixture model. *Technometrics*, 34(4) :441–453, 1992.