

Projet Statistique Computationnelle (Partie Bayésienne)

Franck Corset

M2 pro SSD

Présentation du sujet

On s'intéresse ici au modèle de capture recapture (CR). Cette méthode consiste à échantillonner une population de façon répétée. On étudiera ici le cas où l'on échantillonne deux fois. A chaque capture, les individus non marqués sont marqués, les individus déjà marqués sont notés. L'objectif de cette méthode est d'estimer la taille de la population considérée. On considère ici que la population est fermée entre les deux instants de capture et recapture. Généralement, on peut représenter les données sous forme d'un vecteur à deux coordonnées (dans le cas où on a deux échantillonnages). Par exemple, pour un individu i , $x_i = (1, 1)$ si il a été capturé, puis recapturé.

On supposera qu'il y a indépendance entre la capture et la recapture, et que la probabilité d'être capturée est la même à chaque fois (modèle homogène).

Dans ce cadre, vous simulerez un tel modèle et vous mettrez en place une analyse bayésienne de ce modèle.

Pour l'application sur des données réelles, je vous propose de choisir votre jeu de données, que vous présenterez. 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

- votre démarche scientifique ainsi que les hypothèses faites,
- les conclusions des simulations et des applications réelles,
- des perspectives d'améliorations de la méthode proposée,
- une bibliographie donnant vos sources.

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 L^AT_EX (les documents word seront quand même acceptés). Il faudra également rendre vos fichiers scripts R. La date limite pour rendre le projet est le 22 décembre à 16h. Les fichiers seront envoyés à franck.corset@univ-grenoble-alpes.fr.