

Fiche de TD n° 6

TP : les faiseurs de pluie

Le problème

Des expériences¹ ont été menées aux Etats-Unis afin de voir si un traitement des nuages par avion, à partir de nitrate d'argent, était susceptible d'augmenter les précipitations. Les essais consistaient à traiter des nuages alors que la pluie menaçait de tomber et à mesurer la quantité de pluie obtenue. Au total, 52 pluies ont été examinées. Parmi celles-ci, 26 ont été choisies au hasard afin de leur appliquer le traitement au nitrate d'argent.

Les objectifs

Vous devez réaliser une mini-étude de ce problème à partir des données fournies. Comme il s'agit de votre première étude, vous serez guidés pas à pas pour la réalisation de ce travail. On rappelle que ce TD est noté, et compte pour 10% de l'évaluation du cours. Après avoir lancé R, vous ouvrirez un fichier dans l'éditeur fourni qui s'intitulera `TD6_Nom1_Nom2.R`. Ce fichier ainsi que votre compte-rendu sera déposé à la fin du TD sur le portail dans le dossier intitulé `copies TD n°6` ou envoyé par mail à l'adresse `laurent.carraro@telecom-st-etienne.fr` en cas de problème.

Travail à faire Vous êtes autant que possible par groupes de 2, l'un effectuant les commandes en R, l'autre rédigeant en même temps le compte-rendu. Nous vous conseillons d'utiliser le fichier de commandes utilisé pour le cours n°5, intitulé `cours5.R`.

1. Chargez le fichier de données `pluie.txt` et examinez les résumés donnés pour les hauteurs de pluie (commande `summary`). Séparez-les en deux parties (pluies avec ou sans traitement)².
2. Examinez l'histogramme des données séparées, puis les boxplots. L'effet du traitement est-il visible ? les données obtenues sont-elles symétriques ?

¹Voir Simpson, Alsen, Eden : *A bayesian analysis of a multiplicative treatment effect in weather modification*. *Technometrics* 17, p. 161-166 (1975).

²Pour un tableau de données comportant deux colonnes nommées `col1` et `col2`, `col1[col2==1]` fournit les données de `col1` telles que `col2` est égal à 1.

3. En travaillant toujours avec les données séparées en deux groupes, on travaille maintenant avec le logarithme des hauteurs. Obtient-on quelque chose de plus symétrique ?
4. Tracez maintenant les fonctions de répartition empiriques des hauteurs de chacun des deux groupes, sur un même graphique. Qu'observez-vous ? Estimez à l'oeil le décalage entre les deux courbes.
5. En remplaçant les fonctions de répartition empiriques par les densités estimées³, vos observations sont-elles confirmées ? (voir le cours n°5 pour l'estimation de densité).
6. S'il y a un décalage sur les densités du log des hauteurs de pluie, que cela peut-il signifier pour les hauteurs elles-mêmes ?

³commande `density`