

TP2: JACKKNIFE ET BOOTSTRAP

Nous avons vu dans le TP1 une méthode d'estimation du biais de l'estimateur de variance empirique pour la variance. Le but ici est de quantifier le biais (par différentes méthodes) d'un estimateur de corrélation linéaire, lié à la variance. Nous en étudierons aussi la variance, donc la deuxième composante de l'erreur d'un estimateur (biais au carré plus variance de l'estimateur).

Exercice 1 : estimation du biais

Télécharger et installer la librairie `bootstrap` de R. Charger la librairie, puis charger le jeu de données `law`. Sinon aller sur ma page web et télécharger le jeu de données dans *Enseignements*.

- i) Quel est ce jeu de données ? Quelles sont ses dimensions ? A quoi correspondent les colonnes ?
- ii) Calculer le coefficient de corrélation $\hat{\rho}$ entre les colonnes.
- iii) Donner l'estimateur de Quenouille du biais de $\hat{\rho}$, sachant que le biais de Quenouille est défini par
$$Biais = (n - 1) \times (\bar{\rho} - \hat{\rho}),$$
où $\bar{\rho}$ est la moyenne des coefficients de corrélation calculés par une procédure Jackknife, et n est le nombre d'observations.
- iv) Calculer les pseudo-valeurs adaptées au contexte du coefficient de corrélation. Donner l'estimateur Jackknife du biais. Comparer avec la question précédente.

Exercice 2 : estimation de la variance

On s'intéresse maintenant à estimer la variance de l'estimateur du coefficient de corrélation.

- i) Calculer l'estimateur Jackknife de la variance du coefficient de corrélation.
- ii) On s'intéresse dans la suite à un log-ratio défini par

$$\tau = \ln \left(\frac{E[Z]}{E[Y]} \right),$$

où $Y \sim \mathcal{U}(0, 1)$ et $Z \sim \mathcal{Exp}(1)$.

- iii) Calculer un estimateur du log-ratio basé sur la méthode des moments. On prendra $n = 10$ observations pour les simulations.
- iv) Calculer les pseudo-valeurs du ratio par approche Jackknife. En déduire l'estimateur Jackknife de la variance du log-ratio.
- v) Comparer à l'estimateur par moment de la variance obtenu via une procédure Monte Carlo (avec 1000 trajectoires).