

TP Statistiques et Séries Chronologiques
Université de Lorraine

Régression Linéaire Multiple

Clément Dell'Aiera

1 Régression sur variables qualitatives

Le but de cet exercice est d'expliquer la concentration en ozone $O3$ en fonction de la température $T12$ et de la direction du vent $vent$ dans la table *ozone.txt*.

1. Télécharger la table, et effectuer des régressions selon les différents modèles.
2. Tester l'égalité des pentes.
3. Tester l'égalité des ordonnées à l'origine.
4. Analyser les résidus.

1.1 ANOVA à 1 facteur

Nous souhaitons modéliser la concentration en ozone en fonction de la direction du vent.

1. Tracer une boîte à moustaches de la variable $O3$ par rapport aux quatre modalités de la variable $vent$. Le vent semble-t-il avoir une influence sur la concentration en ozone ?
2. On se place dans un modèle d'analyse de la variance à un facteur

$$y_{ij} = \mu + \alpha_j + \epsilon_{ij}$$

- (a) Effectuer la régression linéaire de $O3$ sur $vent$ sous la contrainte $\mu = 0$.
 - (b) Effectuer la régression linéaire de $O3$ sur $vent$ sous la contrainte $\alpha_1 = 0$.
 - (c) Effectuer la régression linéaire de $O3$ sur $vent$ sous la contrainte $\sum n_i \alpha_i = 0$.
 - (d) Effectuer la régression linéaire de $O3$ sur $vent$ sous la contrainte $\sum n_i \alpha_i = 0$.
3. Analyser les résidus afin de constater que l'hypothèse d'homoscédasticité est vérifiée. Pour cela, tracer un boxplot des résidus en fonction de $vent$, les résidus en fonction de $\hat{O}3$, leurs quantiles théoriques ainsi que la distribution des résidus par modalité de $vent$.

1.2 ANOVA à 2 facteurs

Nous voulons maintenant modéliser la concentration en ozone par le vent et la nébulosité, variable à 2 modalités : SOLEIL et NUAGEUX.

1. Procéder à un examen graphique qui puisse déterminer si l'interaction des facteurs influe sur la variable à expliquer. (voir ce qu'est un *profil*)
2. On suppose la gaussianité des résidus.
 - (a) Tester le modèle avec interaction : **mod1**.
 - (b) Tester le modèle sans interaction : **mod2**.
 - (c) Tester le modèle sans effet du facteur *nebulosité* : **mod3**.
3. Grâce à la commande ANOVA de R, effectuer des analyses de la variance entre les modèles **mod1**, **mod2** et **mod3**.
4. Répondez à la problématique.