**Anova à 2 facteurs dans R, quand les hypothèses sont satisfaites**

1. Exploration visuelle des données, boxplots
2. Calcul des moyennes et intervalles de confiances
3. Visualisation préliminaire de l'interaction probable
4. Créer une variable `group` correspondant au croisement des modalités des 2 facteurs
5. On ajuste toujours un modèle ANOVA à 2 facteurs avec un terme d’interaction
6. Vérifier que les effectifs sont égaux ou pas dans chaque groupe (croisement des modalités)
7. Lorsqu’on ajuste le modèle complet (avec interaction), on utilise généralement des carrés de type III => Anona() sur le modèle et non anova().
8. Lorsque les effectifs sont équilibrés, les résultats des carrés de type III et de type II sont strictement identiques. Ici, il ne serait donc pas nécessaire de changer les contrastes. Mais faites-le systématiquement pour en prendre l’habitude.
9. Ajuster l’anova : ml <- lm()
10. Visualiser les résultats, ici Anova(ml)
11. Avant de passer à l’interprétation des résultats, il est nécessaire de vérifier que les hypothèses de validité de l’ANOVA à 2 facteurs sont satisfaites
    1. Les résidus sont indépendants
       1. Les résidus ne doivent pas être corrélés entre eux. Se vérifie à l’aide du plan d’expérience, pas de mesures répétées
       2. Les résidus ne doivent pas être corrélés au facteur étudié
          1. Plot des résidus vs fitted values,
          2. Ou plot des résidus vs les modalités du facteur
    2. Les résidus suivent une loi normale de moyenne 0
       1. Un QQplot et/ou
       2. Un test de normalité comme le test de Shapiro-Wilk
    3. Les résidus relatifs aux différentes modalités sont homogènes (ils ont globalement la même dispersion), autrement dit leur variance est constante.
       1. Graphiquement et/ou
          1. Plot des résidus standardisés vs valeurs prédites
       2. À l’aide d’un test statistique
          1. Le test de Bartlett,
          2. Ou le test de Levene,
          3. Ou encore le test de Fligner-Killeen
             1. Leur Ho : les variances des différents groupes sont globalement identiques
             2. Leur H1 : au moins 2 variances (les variances de 2 modalités – combinaison des 2 facteurs) sont différentes
             3. Pour accepter l’hypothèse d’homogénéité des résidus, il faut une p-value > 0.05.
12. Lorsque toutes les hypothèses sont validées, les résultats peuvent être interprétés.
13. En cas d'interaction significative, il n’est pas possible d’interpréter les effets propres des facteurs individuellement. Dans ce cas de figure, deux solutions sont envisageables :
    1. La première consiste à faire une ANOVA à un facteur sur la variable `group` créée (croisement des modalités des 2 facteurs). Puis, si l’effet est significatif, des comparaisons multiples peuvent être réalisées pour mettre en évidence les moyennes significativement différentes 2 à 2.

Cette méthode peut aboutir à réaliser un grand nombre de comparaisons, dont certaines ne sont pas intéressantes. Comme les p-values sont ajustées pour garder un risque alpha global de 5%, cela peut empêcher la mise en évidence de différences significatives.

* 1. La seconde solution consiste à réaliser les comparaisons des moyennes relatives aux modalités d’un facteur, séparément pour chacune des modalités de l’autre facteur. Méthode plus technique voir plus complexe !
     1. Dans un premier temps, ajuster un modèle ANOVA à un facteur sur la variable `grp` (croisement des modalités), en omettant l’intercept, afin que les paramètres du modèle correspondent aux moyennes des différentes conditions.
     2. Dans un second temps, construire une matrice de contrastes, correspondant aux comparaisons de moyennes souhaitées.
     3. Enfin, le modèle et la matrice sont fournis en argument de la fonction `glht()` du package `multcomp`, pour obtenir ces comparaisons.

Dans cette approche, \*\*on limite donc les comparaisons à celles qui nous intéressent.

1. En cas d'interaction **quantitative** significative, les effets propres des facteurs sont généralement interprétés ( à partir du modèle contenant l’interaction). En fonction de la p-value (inférieure ou supérieure au seuil de significativité choisi) on conclura à la présence d’un effet significatif, ou à l’absence de mise en évidence d’un effet significatif. Si au moins un des effets est significatif, on réalisera les comparaisons multiples correspondantes.
2. En cas d'interaction qualitative non significative :
   1. avant d’interpréter les résultats, on va ajuster à nouveau le modèle de l’ANOVA à 2 facteurs, mais sans le terme d’interaction, puisque celle-ci n’est pas significative.
   2. Par ailleurs, lorsque le modèle ne contient pas de terme d’interaction, on utilise les carrés de type II. Pour cela, il suffit simplement d’utiliser les contrastes par défaut qui sont de type `contr.treatment`.
   3. Comparaisons multiples : Là encore, le package `multcomp` permet de réaliser toutes les comparaisons en une seule fois, et donc d’ajuster les p-values de façon parfaitement adéquate. Pour cela, on réalise deux matrices de contrastes (K1 et K2), une pour chaque facteur, afin de définir les comparaisons souhaitées. Puis on les réunit dans une seule matrice, qui est donnée en argument à la fonction `glht()`.