

TD1 - Données expérimentales

Mathieu Emily - mathieu.emily@agrocampus-ouest.fr

Objectifs : Modèles à effets fixes

Exercice 1 ANALYSE DE VARIANCE À EFFETS FIXES

On souhaite comparer huit eaux minérales gazeuses du point de vue de leur description sensorielle et de leur note hédonique. Les huit eaux gazeuses sont : St Yorre, Badoit, Vichy, Quézac, Arvie, Chateauneuf, Salvetat et Perrier. Ces eaux ont été dégustées par un jury de 16 étudiants. L'appréciation de chaque critère a été évalué entre 0 et 10.

L'évaluation se déroule en deux jours. Chaque jour les 16 juges commencent par décrire les 8 produits (première session de la journée), se reposent pendant un quart d'heure, puis décrivent à nouveau les 8 produits (deuxième session de la journée). Chaque juge effectue donc 16 évaluations par jour.

Les ordres de dégustation des produits ont été établis selon un plan d'expérience qui permet à chaque produit d'être dégusté un même nombre de fois à chaque position (en 1er, 2ème, ..., 8ème) : ceci permet d'éviter d'avoir des effets de rangs confondus avec les effets produits et à chaque produit de ne pas être dégusté toujours après un même produit. Ceci permet d'éviter que l'effet des arrières-effets soit confondus avec les effets produits : par exemple, après avoir goûté un produit très acide, le suivant peut paraître peu acide.

Les données obtenues à l'issue de la séance de dégustation sont disponibles dans le fichier `fich_eau.xls`.

1. Décrire succinctement le jeu de données (nombre d'observations, nature des variables, ...) On va s'intéresser par la suite aux trois variables suivantes : l'intensité de crépitement, la saveur sucrée et l'intensité gustative globale.
2. Représenter graphiquement la liaison entre la variable *produit* et chacune des trois variables précédentes. Que pouvez-vous dire ?
3. On s'intéresse uniquement à l'effet produit. Écrire le modèle d'analyse de la variance étudié. Avec la procédure d'analyse de variance (ANOVA) du logiciel, tester l'hypothèse "il n'y a pas d'effet produit"¹.
 - (a) Pour l'intensité gustative globale, effectuer une comparaison multiple en utilisant les méthodes suivantes :
 - Méthode LSD avec `parwise.t.test` (attention à l'argument `p.adjust`)

```
> pairwise.t.test(Eau$intensite.gustative.globale,Eau$produit,p.adjust="none")
> pairwise.t.test(Eau$intensite.gustative.globale,Eau$produit,p.adjust="bonferroni")
```
 - Méthode HSD de Tukey avec `glht`

```
> require(multcomp)
> modele.1way <- lm(intensite.gustative.globale ~ produit,data=Eau)
> tuk <- glht(modele.1way,linfct=mcp(produit="Tukey"))
> summary(tuk)
> tuk.cld <- cld(tuk)
> plot(tuk.cld)
```
 - D'autres méthodes (dont LSD et HSD) sont également disponible dans le package `agricolae`
 - La méthode `lsmeans` inspirée de celle du logiciel SAS est implémentée dans le package `emmeans`

```
> require(emmeans)
> mod.IGG <- lm(intensite.gustative.globale ~ produit,Eau)
> emmeans(mod.IGG, pairwise ~ produit)
```
4. On ajoute maintenant l'effet de la variable *juge* dans le modèle. Quel peut en être l'intérêt sachant qu'on s'intéresse uniquement à l'effet *produit*? Y a-t-il un effet de la variable produit ?

1. Utiliser la procédure "lm".

5. Avec le modèle à 2 facteurs, comparer les sommes de carrés de type I et les sommes de carrés de type III².
6. Calculer la somme des carrés totale (à partir de la variance de la variable Y) et comparer la à la somme des sommes des carrés (dans le cas des sommes de carrés de type I puis de type III).
On s'intéresse uniquement à la variable saveur sucrée.
7. Donner les coefficients α_i de l'effet produit.³
8. On rajoute l'interaction *juge – produit* dans le modèle. Quel est l'intérêt ? Que pouvez-vous dire ?
9. Utilisation du package `phia` pour des tests post-hoc d'interaction :


```
> library(phia)
> mod.inter <- lm(intensite.gustative.globale ~ produit*juge,Eau)
> interactionMeans(mod.inter) ## Toutes les moyennes d'interaction
> interactionMeans(mod.inter,factors="produit") ## Uniquement les moyennes marginales pour le facteur produit
# test post-hoc - Effets simples
> testInteractions(mod.inter,fixed="juge",across="produit") ## On teste l'effet produit sachant le juge
# test post-hoc - Effets multiples
> testInteractions(mod.inter)
# test post-hoc - Effets simples pairwise
> testInteractions(mod.inter,pairwise="juge",across="produit")
```

Exercice 2 ECRITURE MATRICIELLE ET CONTRASTE (A FAIRE EN R)

1. Importer le jeu de données des poussins : `poussins.txt` (Les variables `Trait` et `Sexe` sont qualitatives)
2. Taper les commandes :


```
> options(contrasts = c("contr.treatment", "contr.treatment"))
> summary(lm(Rdt ~ Trait,data=Poussins))
> options(contrasts = c("contr.sum", "contr.sum"))
> summary(lm(Rdt ~ Trait,data=Poussins))
```

 Qu'en pensez-vous ?
3. Transformer la variable `Trait` en code disjonctif (Vous pourrez utiliser la fonction `tab.disjonctif` du package `FactoMineR`) et retrouver les estimations obtenues à la question précédente à l'aide de la fonction `lm`.
4. En stockant la variable `Trait` en code disjonctif dans une matrice de design X , retrouver les estimations des questions précédentes en appliquant l'estimateur de la régression linéaire : $(X'X)^{-1}X'Y$.
5. On peut également spécifier des contraintes différentes pour chaque facteur par la fonction `contrasts`. Taper :


```
> contrasts(Poussins$Trait) <- contr.sum(3)
> contrasts(Poussins$Sexe) <- contr.treatment(2,base=2)
> summary(lm(Rdt ~ Trait+Sexe,data=Poussins))
```
6. Retrouver les estimations de la question précédente en définissant la bonne matrice de design X et en appliquant $(X'X)^{-1}X'Y$.
7. Taper `by(Poussins$Rdt,Poussins$Trait,mean)`. Comparer les résultats obtenus avec les moyennes de `Rdt` pour chaque modalité de `Trait`. Commentez.

2. Pour avoir les sommes de carré de type III, "charger" la "bibliothèque" `car` et lancer les commandes :
`options(contrasts = c("contr.sum", "contr.sum"))`
`Anova(res.aov,type="III")`

3. Utiliser la fonction "summary.lm".