T. D. n^{o} 5
Analyse d'un plan complet avec facteurs ayant deux modalités

Ces trois exercices sont issus du livre d'exercices de François Husson et de Jérôme Pagès intitulé Statistiques générales pour utilisateurs, *éditions PUR.*

Exercice 1. Introduction à la planification expérimentale

On étudie l'influence de deux facteurs A et B et de leur interaction sur une variable réponse Y. Les deux facteurs ont chacun deux modalités. On réalise alors les quatre essais correspondant à toutes les combinaisons possibles des modalités des deux facteurs. Ce plan est appelé plan d'expériences complet 2^2.

1. Proposer deux modèles pour étudier ces données, l'un avec interaction, l'autre sans. Ecrire ces modèles sous forme indicée. Rappeler les contraintes existant entre les différents paramètres de ces modèles. Combien y-a-t'il de paramètres à estimer ? Quel est le nombre de degrés de liberté ?

On étudie maintenant le modèle avec interaction.

2. Ecrire le modèle sous sa forme matricielle en tenant compte des contraintes précédentes. Détailler chacun des termes de ce modèle et en particulier la matrice X, appelée matrice des effets.
3. Dans la matrice X, quel est le lien entre les colonnes assocées aux facteurs A et B et la colonne de l'interaction AB ?
4. Rappeler l'expression matricielle permettant d'estimer les paramètres du modèle.
5. Calculer la matrice $(X'X)$. En déduire une expression simplifiée de $\hat{\beta}$ ainsi qu'une règle générale permettant d'estimer chacun des effets du modèle.
6. Généraliser la règle précédentepour un plan complet 2^p.

Exercice 2. Construction d'un plan complet 2^3

On étudie l'influence de trois facteurs A, B et C à deux modalités (1 et 2) sur une variable réponse Y. Pour cela on réalise un plan complet. Les résultats sont rassemblés dans le tableau ci-dessous.

A	B	C	Y
1	1	1	38
1	1	2	30
1	2	1	26
1	2	2	19
2	1	1	37
2	1	2	28
2	2	1	24
2	2	2	16

1. Proposer un modèle pour étudier ces données (préciser les degrés de liberté de chacun des termes).
2. Donner la matrice des effets (notée X) associée à votre modèle. Estimer les paramètres de ce modèle.
3. On néglige maintenant toutes les interactions. Ecrire le nouveau modèle et les degrés de liberté associés à chacun des termes. Quelles est la matrice des effets associée à ce modèle ? Quel lien y-a-t'il entre les estimations des paramètres de ce nouveau modèle et celles du modèle précédent ?
4. Estimer la variance résiduelle.
5. Tester la significativité de chacun des coefficients.

Exercice 3. Dépouillement d'un plan d'expériences pour deux facteurs

On étudie l'influence de deux facteurs F_1 et F_2 présentant chacun deux modalités (notées 1 et 2) sur une réponse Y. Les conditions et les résultats sont regroupés dans le tableau ci-dessous.

Essai	F_1	F_2	Y
1	1	1	7
2	1	2	3
3	2	1	1
4	2	2	1

1. A l'aide des matrices présentées ci-dessous, commenter brièvement ce plan.

$$X = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \qquad X' = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

$$X'X = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} \qquad \left(X'X\right)^{-1} = \begin{pmatrix} 1/4 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 \\ 0 & 0 & 1/4 & 0 \\ 0 & 0 & 0 & 1/4 \end{pmatrix}$$

2. Quelle méthode permet de dépouiller ces résultats ?
3. Ecrire le modèle sans interaction.
4. Donner les estimations de α_1, α_2, β_1, β_2.
5. Calculer les valeurs prédites de Y pour les quatre essais. En déduire une estimation de la variance résiduelle. Donner ensuite une estimation de la variance de $\widehat{\alpha_1}$ et $\widehat{\beta_1}$ (usuellement notées $\hat{\sigma}^2_{\widehat{\alpha_1}}$ et $\hat{\sigma}^2_{\widehat{\beta_1}}$)

Exercice 4. Révision de l'analyse de variance à plusieurs facteurs

Une analyse sensorielle a été organisée pour comparer 6 chocolats noirs. Vingt-neuf juges ont évalué ces chocolats deux fois lors de deux séances de dégustations. Ils ont évalué quatorze descripteurs en mettant des notes entre 0 et 10. L'objectif d'une telle analyse est de comparer les produits les uns par rapport aux autres.

Pour chacun des quatorze descripteurs, le même modèle a été construit. Les données étants trop volumineuses, elles n'ont pas été reproduites ici mais sont disponibles dans la bibliothèque de votre groupe libre sous Univ-R.

1. Ecrire le modèle d'analyse de variance comprenant les trois facteurs (juge, produit, séance) ainsi que les interactions d'ordre deux. Indiquer le nombre de degrés de liberté de la résiduelle.
2. Interprétez concrètement ce que signifie chacun des effets et chacune des interactions.
3. Quels sont les descripteurs pour lesquels les juges ont discriminé les produits (au seuil $\alpha = 5\%$) ?
4. Quels sont les descripteurs pour lesquels les juges sont reproductibles, i.e. pour lesquels il y a un consensus entre les juges dans la notation des produits (au seuil $\alpha = 5\%$) ?
5. Quels sont les descripteurs pour lesquels le jury est répétable, i.e. pour lesquels il évalue les produits de la même façon lors des deux séances (au seuil $\alpha = 5\%$) ?