Simulation
Examen de Statistique
Approfondie I
** Corrigé **

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

Exercice 1. Les préférences de biscuits sont-elles les mêmes d'un pays à l'autre ?

1. Dire que les biscuits ne sont pas évalués de la même manière par les deux jurys revient à mettre en évidence une interaction entre les facteurs *biscuit* et *jury*. Si l'on considère l'ensemble des biscuits, le tableau ne fait pas apparaître de répétitions et il est impossible de mettre en évidence une interaction. En revanche, si l'on regroupe les biscuits par origine, alors on peut construire un modèle d'analyse de variance à deux facteurs avec interaction. La variable Y sera l'appréciation et les facteurs seront la nationalité du jury (française ou pakistanaise) et l'origine des biscuits (française ou pakistanaise) ainsi que l'interaction de ces deux facteurs. On étudie donc le modèle :

   $$Y_{i,j,k} = \mu + \alpha_i + \beta_j + \alpha\beta_{i,j} + \epsilon_{i,j,k}, \quad i = 1 \ldots 2,\ j = 1 \ldots 2,\ k = 1 \ldots 4$$

   avec $Y_{i,j,k}$ la moyenne des appréciations du $k$ème biscuit d'origine i évalué par l'ensemble des juges de nationalité j. On postule les hypothèses classiques suivantes pour les résidus :

   $$\forall\ (i,j,k)\ \mathcal{L}(\epsilon_{i,j,k} = \mathcal{N}(0,\sigma)) \text{ et } Cov(\epsilon_{i,j,k}, \epsilon_{r,s,t}) = 0 \text{ si } (i,j,k) \neq (r,s,t)$$

   La variabilité résiduelle est alors celle des biscuits d'une même origine évalués par un même jury (ceci est dû à notre regroupement).

2. Il y a un effet de l'interaction *nationalité du jury* $\times$ *origine des biscuits* mais il n'y a pas d'effet de la nationalité du jury ni d'effet de l'origine des biscuits. Cela signifie qu'aucun des jurys ne met en moyenne des notes plus élevées (pas d'effet jury), et que les biscuits sont en moyenne autant appréciés que les biscuits pakistanais (pas d'effet origine) mais il existe une interaction. Donc les français n'évaluent pas les biscuits comme les pakistanais.

3. Si l'on travaille sur les données brutes, on peut construire le modèle :

   $$Y_{i,j,k,l} = \mu + \alpha_i + \beta_j + \alpha\beta_{i,j} + \epsilon_{i,j,k,l}, \quad i = 1 \ldots 2,\ j = 1 \ldots 2,\ k = 1 \ldots 4, l = 1 \ldots l(j)$$

   avec $Y_{i,j,k,l}$ l'appréciations du $l$ème juge de nationalité j évaluant le $k$ème biscuit d'origine i. On postule les mêmes hypothèses sur les résidus :

   $$\forall\ (i,j,k,l)\ \mathcal{L}(\epsilon_{i,j,k,l} = \mathcal{N}(0,\sigma)) \text{ et } Cov(\epsilon_{i,j,k,l}, \epsilon_{r,s,t,u}) = 0 \text{ si } (i,j,k,l) \neq (r,s,t,u)$$

La variablité résiduelle change donc de signification par rapport à la question 1. En effet deux indices ni'interviennent qu'à travers la résiduelle : à la variabilité des biscuits d'une même origine s'ajoute la variabilité des juges d'un même jury. Mais parallèlement il y a beaucoup plus de données et donc les tests sont plus puissants. Cependant malgré des tests plus puissants on ne met toujours pas en évidence l'effet de l'origine du biscuit ni l'effet de la nationalité du jury. L'interaction entre ces deux facteurs est par contre beaucoup mieux mise en évidence puisue la $p-$valeur associée est maintenant de $5,126 \times 10^{-16}$ (contre 0,025 précedemment).

4. Le graphique demandé représente l'interaction *origine du biscuit* $\times$ *jury*. S'il n'y avait d'interaction, les lignes brisés seraient parallèles. On peut voir ici que les pakistanais apprécient surtout les biscuits pakistanais (biscuits P1, P2, P3 et P4) tandis que les français préfèrent les biscuits français (biscuits F1 à F4). Peut-être y a-t-il une sorte de rejet devant un produit nouveau ou une sur-notation pour des produits nationaux qui sont reconnus et appréciés ?

Exercice 2. Etude de l'appréciation de cocktails de jus de fruits en fonction de leurs saveurs fondamentales et de leur caractère pulpeux

1. Il y a $8 \times 16 \times 4 = 512$ observations.Pour chaque observation, on dispose du nom du juge qui l'a effectuée (variable qualitative), de l'eau concernée (variable qualitative) et de la note mise (variable quantitative).
2. Pour tester la significativité de l'interaction, on construit le test suivant :

   $$H_0 \ : \ \forall(i,j),\ \alpha\beta_{i,j} = 0 \qquad H_1 \ : \ \exists(i_0,j_0) \text{ tels que } \alpha\beta_{i_0,j_0} \neq 0$$

   La statistique de test est :

   $$F = \frac{CM_{\text{interaction}}}{CM_{\text{résiduelle}}}$$

   Ici : $ddl_{\text{interaction}} = (I-1)(J-1) = (8-1)(16-1) = 105$ et $ddl_{\text{résiduelle}} = n-IJ = 512-128 = 384$. Sous l'hypothèse H_0 et une hypothèse de normalité, F suit une loi de Fisher à 105 et 384 degrés de liberté. Décision : la $p-$valeur vaut $1,192 \times 10^{{}^{1}0}$, le test est donc significatif au seuil $\alpha = 0,05$. On considère donc que l'interaction entre le facteur *juge* et le facteur *eau* est significative, c'est-à-dire que tous les juges n'ont pas les mêmes préférences.
3. Les tests sur les effets principaux *eau* et *juge* se construisent de la même façon que le test de l'interaction. Les effets *eau* et *juge* sont donc significatifs : certaines eaux sont globalement plus appréciés que d'autres ; certains juges mettent en moyenne des notes plus élevées que d'autres.
4. L'eau la plus appréciée est l'eau Quézac puisque cette eau a le coefficient $\hat{\alpha}_i$ le plus élevé (0,428). Mais comme l'interaction *juge* $\times$ *eau* est significative, tous les juges ne préfèrent pas forcément l'eau Quézac. A priori, lors d'un rassemblement, il sera bon de prévoir d'autres eaux minérales.