

Instructions :

- Répondez aux questions suivantes à l'aide de SAS et fournissez le code utilisé pour vos analyses dans un fichier (extension .txt, encodage utf8).
- Votre rapport doit être remis en ligne en format PDF et en version papier en classe et ne devrait pas faire plus de 15 pages; toute page excédentaire sera ignorée lors de la correction. Soyez concis mais précis : n'incluez que les sorties pertinentes.
- Les erreurs sont pénalisées même si elles ne sont pas en lien direct avec la question.

3.1 Enseignement de la lecture : les données sont tirées de

J. Baumann, N. Seifert-Kessell, L. Jones (1992), *Effect of Think-Aloud Instruction on Elementary Students' Comprehension Monitoring Abilities*, Journal of Reading Behavior, **24** (2), pp. 143–172.

Ces chercheurs ont fait une étude pour déterminer l'efficacité relative de méthodes d'apprentissage de la lecture. L'échantillon de 66 élèves de quatrième année primaire comporte 32 filles et 34 garçons qui ont été alloués de façon aléatoire aux trois groupes. On s'intéresse à l'amélioration des capacités de lecture par rapport à une méthode d'enseignement traditionnelle (DR). Deux tests ont été administrés avant et pendant l'expérience pour mesurer l'efficacité des méthodes; afin de rendre les résultats comparables, ils ont été repondérés de telle sorte à ce que une note parfaite vaille 1.

Les données contiennent des renseignements sur

- groupe : unité expérimentale, une parmi lecture dirigée - activité de réflexion (DRTA), réflexion à voix haute (TA) et lecture dirigée (DR).
- mpre : moyenne du score de prédiction pré-intervention (standardisé) pour les tests de détection d'erreurs et de compréhension.
- mpost : même que mpre, mais pour les évaluations post-intervention.

Nous sommes intéressés tout d'abord par l'amélioration pour chacune des méthodes à l'aide de deux modèles.

- Dans leur article, Baumann *et al.* font une analyse de variance à un facteur pour les scores pré-intervention (mpre) avec le facteur groupe. Expliquez quel est l'usage d'un tel test dans le contexte de l'étude.
- Soit $dpp = mpost - mpre$ la différence entre résultats standardisés post- et pré-intervention. Ajustez une analyse de variance à un facteur pour dpp avec le facteur groupe (modèle 3.1.1). Écrivez l'équation du modèle ajusté en termes de scores pré- et post-intervention et montrez que le modèle est un cas spécial d'un modèle de régression linéaire pour mpost avec un terme de décalage.
- Comparez le modèle d'analyse de variance à un facteur pour dpp avec groupe (modèle 3.1.1) à un modèle linéaire ayant mpost comme variable réponse et mpre et groupe comme variables explicatives (modèle 3.1.2). Au vu de l'ajustement de ce dernier, est-ce que le modèle d'analyse de variance est adéquat? Justifiez votre réponse.
- Transformez les données de format court à format long; ce dernier est plus convenable pour l'analyse de données longitudinales. En plus de groupe, vos données devraient contenir les colonnes suivantes
 - id : identifiant de l'étudiant.
 - score : moyenne pour l'évaluation.
 - test : variable catégorielle, une de mpost ou mpre, qui renseigne sur le score correspond à la moyenne pré-intervention ou post-intervention.

Le Tableau 1 présente le format final que vous devriez obtenir. Comparez deux modèles pour score en fonction du groupe et de test, en incluant un terme d'interaction entre les deux, mais avec des modèles de covariance intra-individus différents :

- modèle 3.1.3 avec une structure d'équicorrélation pour les erreurs;
- modèle 3.1.4 avec une covariance non structurée.

groupe	test	score	id
DR	mpre	0,23	1
DR	mpost	0,27	1
DR	mpre	0,35	2
DR	mpost	0,42	2
DR	mpre	0,41	3
DR	mpost	0,24	3
DR	mpre	0,57	4
DR	mpost	0,39	4
DR	mpre	0,67	5
DR	mpost	0,56	5

TABLEAU 1 – Premières 10 lignes des données de Baumann en format long

Écrivez la matrice de covariance des erreurs pour le modèle 3.1.3 et rappez les corrélations estimées entre pré-interventions et post-interventions pour un(e) étudiant(e).

- (e) À l'aide des sorties des modèles 3.1.3 et 3.1.4, testez si la variance des scores moyens pré-intervention et post-intervention sont les mêmes. Écrivez le nom du test que vous utilisez, la valeur numérique de la statistique et calculez la valeur- p avant de conclure dans le contexte du problème.
- (f) Puisqu'on a affaire à des données longitudinales, il serait logique de considérer en plus des deux modèles de covariance, un modèle autorégressif d'ordre 1, ou modèle AR(1), pour les erreurs. Est-ce que ça serait utile dans ce cas? Justifiez votre réponse.
- (g) Utilisez le modèle 3.1.4 pour déterminer si les résultats pour les méthodes d'enseignement DRTA et TA sont significativement meilleures que la méthode standard DR.
- (h) Jusqu'à présent, on a supposé que les résultats pré- et post-intervention de tous les élèves avaient la même matrice de covariance. On pourrait cependant supposer que les paramètres de cette matrice de covariance diffèrent d'une méthode d'apprentissage à une autre. Est-ce que les données corroborent cette hypothèse?

3.2 Les données ainsi que la description qui suit ont été adaptées d'un exemple d'OpenBUGS et proviennent de l'étude

H. Goldstein *et al.* (1993). *A Multilevel Analysis of School Examination Results*, Oxford Review of Education, **19** (4), pp. 425–433.

Les auteurs analysent les résultats d'examens d'élèves des écoles dans les districts centraux du Grand Londres, et étudient la variabilité des résultats entre les écoles afin d'en faire un classement. Les scores moyens d'examen de 1978 élèves de 38 écoles différentes ont été standardisés et rendus disponibles pour cette étude. On a pour chaque élève le sexe, son score au test de lecture de Londres (TLL) et un rang pour un test de raisonnement verbal (rv) passé au début de l'année lorsque l'élève était âgé de 11 ans. Les deux tests sont effectués au début de l'année scolaire. Chaque école a été classifiée selon type (école pour filles, école pour garçons ou école mixte, et selon sa dénomination (école anglicane affiliée à Église d'Angleterre, école catholique, école publique, ou autre). Ces deux critères sont utilisés comme variables catégorielles spécifiques à chaque école. Les données goldstein contiennent les variables suivantes :

- `score` : score standardisé à l'examen de fin d'année pour chaque élève,
 - `ecole` : identifiant de l'école,
 - `TLL` : score au test de lecture de Londres
 - `rv` : catégorie du test de raisonnement verbal (1, 2 ou 3, où 1 représente le groupe avec la plus haute aptitude et 3 celui avec la plus faible),
 - `sexe` : sexe de l'élève, soit fille (0) ou garçon (1),
 - `type` : variable catégorielle pour le type de l'école, soit `filles`, `garcons` ou `mixte`.
 - `denom` : variable catégorielle pour la dénomination de l'école, soit Église d'Angleterre (`angli`), catholique (`catho`), école publique (`etat`), ou autre.
- (a) Donnez l'étendue du nombre d'élèves par école et utilisez cette information afin de déterminer s'il est possible d'estimer un effet de groupe fixe pour chaque école.
- (b) Écrivez l'équation du modèle postulé pour la variable `score` incluant les variables `TLL`, `rv`, `sexe`, `type` et `denom` comme effets fixes et un effet aléatoire pour la variable `ecole`. Dans votre modèle, utilisez les catégories de référence `rv=3`, `mixte` pour `type` et `autre` pour `denom`.
- (c) On pourrait considérer un modèle où la variable `rv` a un effet aléatoire plutôt qu'un effet fixe. Lequel des deux modèles vous semble le plus adéquat et qu'elle est la différence conceptuelle entre ces deux modèles?
- (d) En utilisant le modèle ajusté avec un effet aléatoire pour `ecole`, calculez la matrice de covariance estimée pour l'école 37 et expliquez comment les termes de cette matrice ont été obtenus à partir des estimés des paramètres de covariance. Calculez la proportion de la variance totale qui est due à l'effet de groupe pour l'école.
- (e) Produisez un diagramme quantile-quantile normal des effets aléatoires prédits pour l'école. Commentez sur l'hypothèse du modèle quant aux effets aléatoires.
- (f) Le but de l'étude de Goldstein *et al.* (1993) était de classer les écoles londonniennes. Quel est le bienfait de combiner toutes les informations provenant des différentes écoles à l'aide d'effets aléatoires afin d'estimer les scores moyens des élèves?
- (g) Représentez graphiquement les effets de groupe prédits en fonction de l'identifiant de l'école (du groupe) ainsi que les intervalles de prédiction en vous basant sur la formule $\hat{b}_i \pm 1.96se(\hat{b}_i)$ (il se peut que vous ayez besoin de calculer les bornes des intervalles manuellement).
- (h) Quel est le classement des cinq meilleurs écoles?