

## Laboratoire 9 : Régression multiple

Dans le laboratoire précédent, vous avez travaillé sur la régression simple, dans laquelle seulement une variable (c.-à-d., l'intimidation) a été utilisée pour prédire une autre variable (c.-à-d., le rendement en sciences). Dans le présent laboratoire, nous étudierons l'utilisation de deux variables (c.-à-d., l'intimidation et l'environnement d'apprentissage) pour prédire le rendement en sciences; cette façon de faire est appelée « *régression multiple* ». En théorie, l'ajout d'une ou de plusieurs valeurs explicatives devrait améliorer une prédiction. Par exemple, l'âge peut être utilisé pour prédire la santé de quelqu'un, mais il est aussi possible d'utiliser l'âge, le poids, la tension artérielle et le niveau de cholestérol pour prédire la santé.

### 1. Corrélation et consistance interne

Dans le laboratoire 8, vous avez effectué une corrélation et un coefficient alpha de Cronbach pour les items relatifs à l'intimidation; dans le présent laboratoire, vous allez effectuer la même analyse pour les cinq items relatifs aux environnements d'apprentissage (c.-à-d., SQ12A à SQ12 E).

#### Questions :

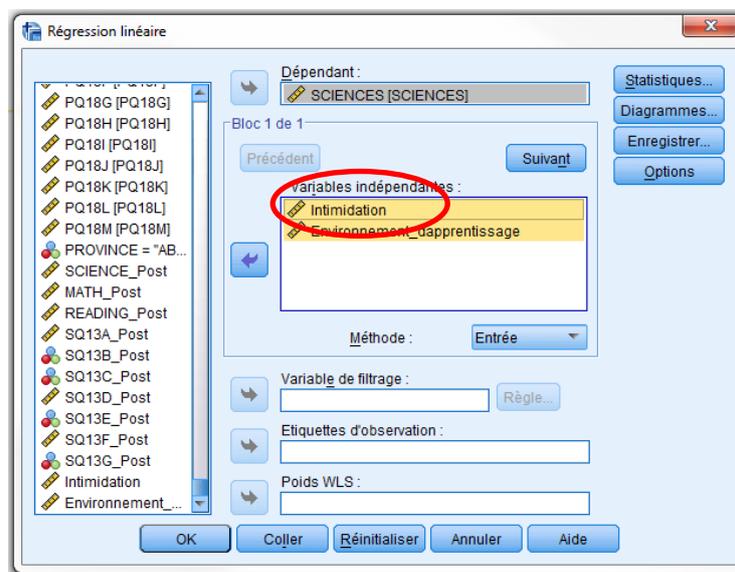
1. Qu'est-ce que le tableau de corrélation du résultat SPSS indique au sujet des items?
2. En regardant le coefficient alpha de Cronbach de la colonne « Alpha de Cronbach en cas de suppression de l'élément » du tableau « Statistiques de total des éléments » pour les items relatifs aux environnements d'apprentissage, quelle est la meilleure valeur alpha possible pouvant être obtenue? Quel(s) item(s) recommanderiez-vous de combiner dans notre moyenne pour obtenir cette valeur?
3. Si vous recommandez d'enlever des items à la question 2, veuillez consulter l'item que vous prévoyez enlever (p. ex., regardez le questionnaire de l'élève ou le manuel de codage) et avancez des hypothèses pour expliquer pourquoi ces items réduiraient la consistance interne de l'ensemble d'items.
4. Veuillez tenir compte des résultats de vos analyses de corrélation et de consistance interne pour recommander les items que vous utiliseriez pour calculer une moyenne pour la variable de l'environnement d'apprentissage.

Au moyen de vos résultats à la question 4, veuillez calculer la moyenne pour la variable de

l'environnement d'apprentissage.

## 2. Régression multiple

La régression multiple utilise le même processus que la régression simple, mais en ajoutant une valeur explicative de plus (l'attractivité) dans la case « Variables indépendantes ». Pour commencer cette analyse, veuillez calculer les moyennes pour l'intimidation et l'environnement d'apprentissage au moyen des recommandations que vous avez formulées aux questions 3 et 4 quant aux items à inclure. Des directives sur la façon de calculer la moyenne d'un ensemble d'items se trouvent dans le laboratoire 8, à la page 3.



### Questions :

5. Au moyen des tableaux « Récapitulatif des modèles » et ANOVA, le modèle global (utilisant l'intimidation et l'environnement d'apprentissage pour prédire le rendement en sciences) est-il significatif pour prédire le rendement? Veuillez répondre en utilisant la phrase suivante : « L'intimidation et l'environnement d'apprentissage [peuvent/ne peuvent pas] prédire le rendement en sciences de façon significative au moyen des analyses statistiques  $r^2$ =[ajouter votre valeur R carré],  $F$ (degré de liberté<sub>régression</sub>, degré de liberté<sub>résiduel</sub>)=[ajouter votre valeur F],  $p$ [</>]0,05. ».
6. Maintenant, nous devons examiner chaque variable explicative utilisée pour déterminer si les deux variables explicatives sont nécessaires (c.-à-d., si la variable de l'intimidation permet à elle seule de prédire le rendement en sciences aussi bien qu'en utilisant les deux variables explicatives, il sera alors préférable d'utiliser seulement l'intimidation; on préférera les modèles simples aux modèles complexes). Au moyen du tableau « Coefficients », veuillez indiquer si chaque variable explicative est significative pour prédire le rendement en sciences. Veuillez répondre en utilisant la

phrase suivante : « La variable relative à l'intimidation [peut/ne peut pas] prédire le rendement en sciences de façon significative en utilisant les analyses statistiques,  $b_1$ =[ajouter votre valeur B],  $t$ =[ajouter votre valeur t],  $p$ [</>]0,05. La variable de l'environnement d'apprentissage [peut/ne peut pas] prédire le rendement en sciences de façon significative en utilisant les analyses statistiques,  $b_1$ =[ ajouter votre valeur B],  $t$ =[ ajouter votre valeur t],  $p$ [</>]0.05. ».

7. Nous devons maintenant interpréter les résultats. Veuillez sélectionner [augmentera/diminuera] et remplir les espaces vides avec les valeurs appropriées (vous pourriez devoir utiliser Google pour cet item) :
- i) En contrôlant la variable de l'environnement d'apprentissage, lorsque l'intimidation est augmentée d'une unité, le rendement en sciences [augmentera/diminuera] de \_\_\_\_\_ unités.
  - ii) En contrôlant la variable de l'intimidation, lorsque l'environnement d'apprentissage déclaré par les élèves est augmenté d'une unité, le rendement [augmentera/diminuera] de \_\_\_\_\_ unités.

### 3. Test du chi carré

Lorsque nous travaillons avec des variables nominales, nous devons utiliser des tests non paramétriques parce qu'elles ne correspondent souvent pas aux hypothèses (c.-à-d., la normalité et l'homogénéité de la variance). Dans cette analyse, nous nous pencherons sur l'enquête menée auprès du personnel enseignant. Plus précisément, nous espérons répondre à la question de recherche suivante : « Y a-t-il une différence entre le nombre de matières scientifiques étudiées par les enseignantes et enseignants (c.-à-d., TQ04) et leur perception d'eux-mêmes en tant que spécialistes en enseignement des sciences (c.-à-d., TQ05A et)? ». Par conséquent, nous avons deux variables nominales : le nombre de cours de sciences suivis au postsecondaire (1-2/3-5/6-9/10+) et la perception des enseignantes et enseignants en tant que spécialistes en enseignement des sciences (éducation et expérience oui/non).

Pour analyser le chi carré de Pearson, cliquez sur Analyse → Statistiques descriptives → Tableaux croisés → transférez chacune des variables dans les cases « Lignes et Colonnes » (*Remarque : n'importe quel ensemble de variables peut aller dans n'importe quelle case.*) → cliquez sur Afficher les diagrammes en bâtons juxtaposés → Statistiques → cliquez sur Chi-deux, Phi et V de Cramer → Poursuivre → Cellules → cochez Observé dans la case Effectifs, puis ligne, position et total dans la case « Pourcentages ») → Poursuivre → cliquez sur OK.

### Questions :

8. Au moyen du tableau de tabulation croisée, veuillez indiquer combien de participantes



et participants font partie de chacune des huit catégories (inclure les pourcentages).  
(Indice, veuillez montrer vos résultats pour TQ05A et TQ05B dans le même tableau.)

9. En utilisant les valeurs chi carré de Pearson, veuillez indiquer s'il existe une association significative entre les variables au moyen de la phrase suivante : « [Il y a/Il n'y a pas] d'association significative entre le nombre de cours liés aux sciences suivis par une enseignante ou un enseignant et sa perception de lui-même en tant que spécialiste en raison de son [éducation/expérience],  $\chi^2$ ([ajouter votre degré de liberté])=[ajouter votre valeur chi carré] p[</>]0,05. ». Vous devrez écrire deux phrases pour cet item puisque vous avez deux items pour TQ05.
10. Qu'est-ce que les valeurs Phi et les valeurs de Cramer indiquent (vous pourriez devoir utiliser Google pour cet item).
11. Qu'est-ce que VOS valeurs Phi et valeurs de Cramer (p. ex., 0,775 et 0,376) indiquent au sujet de la relation entre les variables? Souvent, les graphiques aident à faire comprendre la relation entre les variables aux lectrices et lecteurs qui ne sont pas spécialistes des statistiques.

**Vous avez maintenant terminé le dernier laboratoire pour ce cours.**