

## **Pistes de réflexion pour la mise en œuvre d'un enseignement à distance sur le test du khi carré d'indépendance pour des étudiants en master de sciences de l'éducation**

**Mehdi Khaneboubi**<sup>a</sup>

<sup>a</sup> Laboratoire EMA (EA 4507)  
IUFM de l'université de Cergy-Pontoise  
Site universitaire de Gennevilliers  
Bureau c231,  
ZAC des Barbanniers  
Avenue Marcel Paul  
92230 Gennevilliers  
mehdi.khaneboubi@-cergy.fr

**Mots clefs** : initiation, enseignements à distance, sciences de l'éducation, khi carré.

### **Contexte**

Cette contribution présente la mise en œuvre d'un enseignement d'initiation à l'analyse de données avec  $R$  pour des étudiants en sciences humaines et sociales dans un cours à distance. Ce cours de master professionnel, aussi dupliqué dans un master recherche de sciences de l'éducation, fait partie d'un diplôme conduit conjointement avec l'agence universitaire de la francophonie (AUF) et deux autres institutions d'enseignement supérieur en Belgique et en Suisse.

Les étudiants sont principalement localisés en Afrique et souvent, ne disposent pas de connexions à internet robustes ni d'accès à des bibliothèques universitaires. Tenant compte de ces contraintes, le tchat (clavardage) s'est imposé comme mode d'interaction pédagogique avec les étudiants, pour compléter des supports pédagogiques en ligne. La durée des enseignements est d'environ 6 semaines à raison d'une à deux séances de tchat par semaine. Dans ces diplômes, cet enseignement est le seul à traiter de méthodes quantitatives. C'est aussi le seul enseignement dans lequel les étudiants auront l'occasion de s'initier à la programmation à l'exception d'un autre cours de master professionnel. Les étudiants sont en majorité des enseignants, des formateurs ou des prescripteurs intermédiaires se destinant soit à évoluer dans leurs institutions soit à un doctorat.

Dans ce contexte, quels atouts présente  $R$  par rapport à d'autres logiciels de traitement statistiques ? Quelle progression didactique est pertinente ? Comment articuler instrumentation logicielle et manipulation des savoirs statistiques ? Quelles difficultés peut-on identifier du point de vue des apprenants ? Au travers d'une analyse des tchats, nous verrons quels sont les principaux obstacles et les détails clés à considérer pour mettre œuvre cet enseignement.

### **Mise en œuvre**

Dans notre contexte  $R$  à l'avantage de permettre de communiquer les scripts par tchat et de les examiner collectivement avec les étudiants. De plus, les ressources francophones disponibles en ligne sont doublement adaptées à l'isolement d'apprenants en formation à distance et à des étudiants n'ayant qu'un accès difficile à des bibliothèques universitaires.

D'un point de vue didactique, un certain nombre de préalables doivent être appréhendés pour que les étudiants puissent interpréter les informations produites par le script suivant :

```

# importation des données, dans les deux premières colonnes du fichier
sont des variables qualitatives
read.csv2("variables.csv" , sep=";" , header=TRUE , na.strings="NA")->bdd

# résumé de la base
summary(bdd)
head(bdd)
tail(bdd)

# calcul de l'indicateur de khi deux et de la p-value pour le tableau
constitué par les colonnes 1 et 2 de l'objet « bdd »
chisq.test(table(bdd[,1] , bdd[,2]) , correct = FALSE)->khideux

# tableau des effectifs observés
khideux$observed

# tableau des effectifs théoriques
khideux$expected

# racine carré du tableau de khi deux
khideux$residuals

```

D'une part, il faut les familiariser progressivement avec des manipulations techniques périphériques : installation du logiciel, gestion des formats de fichiers, repérage du répertoire de travail...

Ensuite avec le fonctionnement de *R* : comprendre ce qu'est un objet, distinguer un objet et une fonction, *R* ne répond pas lorsque tout fonctionne, *R* répond des choses incompréhensibles...

Et enfin avec les particularités du test de khi-deux d'indépendance : hypothèse d'indépendance, loi de khi deux, p-value, interprétation des résultats etc.

## Références

- [1] Barnier, J. (2008). *R pour les sociologues (et assimilés)*. Consulté à l'adresse [http://alea.fr.eu.org/j/intro\\_R.html](http://alea.fr.eu.org/j/intro_R.html)
- [2] Cibois, P. (2007). *Les méthodes d'analyse d'enquêtes*. Paris: Presses Universitaires de France.
- [3] Haspekian, M. (2005). An « Instrumental Approach » to Study the Integration of a Computer Tool Into Mathematics Teaching: the Case of Spreadsheets. *International Journal of Computers for Mathematical Learning*, 10(2), 109-141.
- [4] Roditi, E. (2009). Un tableur grapheur pour enseigner les statistiques en sciences humaines et sociales. In G.-L. Baron, É. Bruillard, & L.-O. Pochon (Éd.), *Informatique et progiciels en éducation et en formation* (p. 257–275). Consulté à l'adresse <http://halshs.archives-ouvertes.fr/halshs-00609639>
- [5] Tort, F., Blondel, F. M., & Bruillard, É. (2008). Spreadsheet knowledge and skills of French secondary school students. *Informatics Education-Supporting Computational Thinking*, 305–316. Consulté à l'adresse <http://www.springerlink.com/index/U5457121103G7645.pdf>