

Un enseignement à distance sur le test
de khi deux d'indépendance
en master de sciences humaines
et sociales

Mehdi Khaneboubi
Laboratoire EMA
Université de Cergy-Pontoise

Deuxièmes rencontres R
Lyon - 28/06/2013

Contexte

- Masters à distance en sciences de l' éducation
- Connexions internet des étudiants peu fiables
- Enseignement au format du CNED
 - Support de cours « statiques »
 - Interactions en temps réel par tchats
- 6 semaines, 1H30 par semaine de tchat
- 12 étudiants dans 2 groupes de M2, 6 en M1
- Assiduité relative

Objectif et évaluation du cours

- Être capable de lire des articles ou des ouvrages utilisant des statistiques
- Dépouiller un questionnaire
- Familiariser les étudiants avec les parties les moins nobles du travail sans (trop) les rebuter
- L' évaluation porte notamment sur :
 - la manipulation élémentaire de réponses à un questionnaire (N = 180, 4 variables)
 - l' interprétation des résultats de tests de khi deux

Questionnements

- Quels atouts présente R par rapport à d'autres logiciels de traitements 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 ?

=> Examen des sessions de tchat avec les étudiants

Le cas du tableur

- En collège, Haspekian (2005) estime que l'instrumentation est liée à l'habitude scolaire
- En lycée, Tort & al. (2008) identifient un lien entre savoir mathématique et instrumentation logicielle
- Roditi (2009) présente un enseignement sur le Khi deux en SHS avec un tableur

Intérêts de *R*

- La console, évite des problèmes d' OS et de versions
- Examen des scripts collectivement dans le tchat
- Évite les copies d' écran
- Des ressources en français sur le web
- Ne pas (trop) participer à la promotion des logiciels propriétaires

Combinaison de 3 types de savoirs

- Computer literacy
- Utilisation de R
- Notions statistiques

1 - L'importance des « computer literacy »

PROF : `read.table("nomdevotrefichier.txt" ,
sep=";" , header = TRUE)->tab`

étudiant23 : il ne retrouve pas le fichier qui est
dans Mes Documents

étudiant23 : impossible d'ouvrir le fichier
'nomdevotrefichier.txt' : No such file or
directory

PROF : ok

PROF : forcément

PROF : comment s'appelle votre fichier ?

PROF : comment l'avez-vous nommé ?

1 - L'importance des « computer literacy »

```
read.csv2("variables.csv" , sep=";" ,  
  header=TRUE , na.strings="NA")->bdd
```

Le séparateur de champ : l'obstacle excel

L'emplacement du fichier sur le disque

étudiant5 : Ci-dessous le résultat que ça a donné:

```
utilordiclasse.utilordielev.discipline.age.anais  
plus_d_une_fois_par_sem;oui;autres;[40_50[;1966 : 3  
plus_d_une_fois_par_sem;oui;langues;[20_30[;1980: 3  
plus_d_une_fois_par_sem;oui;langues;[40_50[;1965: 3  
plus_d_une_fois_par_sem;oui;autres;[20_30[;1980 : 2  
plus_d_une_fois_par_sem;oui;autres;[30_40[;1976 : 2  
plus_d_une_fois_par_sem;oui;autres;[30_40[;1978 : 2  
(Other) :158
```

PROF : vous avez oublié sep = ";"

2 - Utilisation de R

La syntaxe

Une fonction appliquée sur un objet ne le modifie pas

R ne répond pas lorsque cela fonctionne

R répond des choses incompréhensibles

PROF : étudiant22 allez y, sur quoi avez-vous buté ?

étudiant22 : avec la fonction: read.csv2
("variables.csv" , sep=";" , header=TRUE ,
na.strings="NA")

étudiant22 : le tableau apparaît avec les variables mais

étudiant22 : table(bdd[,1] , bdd[,2])->tableau

étudiant22 : Erreur dans table(bdd[, 1], bdd[, 2]) :
objet 'bdd' introuvable

PROF : C'est bien mais il manque juste une chose.

===

étudiant20 : j'ai rien, juste ">"

===

étudiant3 : quand je tape getwd

étudiant3 : voilà ce qui est affiché function()internal
(getwd())

3 - Notions statistiques

summary (bdd)

La cohérence des données

Les types de variables

Les non-réponses

Calcul de l'indicateur de khi deux et de la p-value

```
chisq.test(table(bdd[,1] , bdd[,2])) -  
  >khideux ; khideux
```

Ici `table()` est superflu pour R mais signifiante pour un débutant

L'hypothèse d'indépendance

Réalisation d'un test de khi deux dans un tableur

P-value et la table de loi de khi-deux

Interprétation

`khideux$observed`

`khideux$expected`

`khideux$residuals`

Les écarts à l'indépendance

Perspectives

- Réécriture du cours : vers une proto-didactique ?
- Un autre format d'enseignement à distance
- Codeacademy pour les statistiques ?
- To MOOC or not to MOOC ?

Références

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.

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). <http://halshs.archives-ouvertes.fr/halshs-00609639>

Tort, F., Blondel, F. M., & Bruillard, É. (2008). Spreadsheet knowledge and skills of French secondary school students. *Informatics Education-Supporting Computational Thinking*, 305–316. www.springerlink.com/index/U5457121103G7645.pdf