

Motivation

M2R Psychologie, Analyse de données avec R

Ana Karina Fermin

Université Paris Nanterre

<http://fermin.perso.math.cnrs.fr/>

La statistique dans la presse !

Chef économiste de Google : Hal Varian

LE FIGARO · fr

LE FLASH ECO

🏠 > ECONOMIE > CONJONCTURE (ARCHIVES)

Le métier de statisticien sera le plus «sexy» de la décennie

Par [Jean-Pierre Robin](#) | Mis à jour le 18/10/2010 à 11:28 / Publié le 17/10/2010 à 11:27



La Toile et les moteurs de recherche sont d'extraordinaires pourvoyeurs d'informations et de statistiques. Le tout est de savoir s'en servir.

Using R for psychological research

A simple guide to an elegant language

This is one page of a series of tutorials for using R in psychological research. Much of material has also covered in number of [short courses](#) or in a set of [tutorials](#) for specific problems.

site : <http://personality-project.org/r/>

No more excuses: R is better than SPSS for psychology undergrads, and students agree

Posted on [September 10, 2016](#) by [@dalejbarr](#)

Starting this academic year (2016/7), all of our stats teaching in Psychology at the University of Glasgow will use R and RStudio instead of Excel and SPSS.

When we proposed the transition, some staff worried that the scripting nature of R would be too challenging for incoming students, many of whom would be starting the program with little or no programming experience. Staff at other universities who have also been leading similar transitions to R have also told me that they have faced similar pessimism from teaching staff.

I did not share these doubts, because for six years, I have been giving students the choice between R and SPSS in my level 3 course, and found that those who chose to use R and RStudio were mostly able to get up and going on their own, even without very much guidance from me, beyond sending them links to online tutorials. To be sure, students found it challenging at first,

Using R in 12 simple steps for psychological research

(These steps are not meant to limit what can be done with R, but merely to describe how to do the analysis for the most basic of research projects and to give a first experience with R).

1. [Install R](#) on your computer or go to a machine that has it.
2. Download the [psych package](#) as well as other recommended packages from CRAN using the `install.packages` function, or using the package installer in the GUI. To get packages recommended for a particular research field, use the `ctv` package to install a particular task view. Note, these first two steps need to be done only once!
3. Activate the `psych` package or other desired packages using e.g., `library(psych)`. This needs to be done every time you start R. Or, it is possible to modify the startup parameters for R so that certain libraries are loaded automatically.
4. Enter your data using a text editor and save as a text file (perhaps comma delimited if using a spreadsheet program such as Excel or OpenOffice)
5. Read the data file or copy and paste from the clipboard (using, e.g., `read.clipboard`).
6. Find basic [descriptive statistics](#) (e.g., [means](#), [standard deviations](#), [minimum and maximum](#)) using `describe`.
7. Prepare a simple descriptive graph (e.g., a [box plot](#)) of your variables.
8. Find the [correlation matrix](#) to give an overview of relationships (if the number is not too great, a scatter plot matrix or SPLOM plot is very useful, this can be done with `pairs.panels`).
9. If you have an experimental variable, do the [appropriate multiple regression](#) using standardized or at least zero centered scores.
10. If you want to do a [factor analysis or principal components analysis](#) use the `fa` or `principal` functions.
11. To score items and create a scale and find various reliability estimates, use `score.items` and perhaps `omega`.
12. [Graph the results](#)

➤ Formations de l'Ecole Doctorale Connaissance, Langage, Modélisation - ED 139

➤ Formations professionnalisantes

➤ Formations proposées par Yannick Morvan et Laurence Kern

1) Equations structurelles avec R

Partie 1 : Lundi 04 et Mardi 05 janvier 2016

Partie 2 : Lundi 11 et Mardi 12 janvier 2016

Direction Scientifique : Yannick Morvan & Laurence Kern

Formateur : Christophe Lalanne

CLINIQUE PSYCHANALYSE DÉVELOPPEMENT - (CLIPSYD - EA4430)



Les participants apprendront à utiliser ces outils de manière active à l'aide d'exemples pratiques afin d'acquérir une expérience de régressions, des analyses factorielles exploratoires et confirmatoires (pour des données continues et catégorielles - partie 1) ainsi que des modélisations avec des variables latentes plus avancés (Analyse multi-groupes et invariance de la mesure, modèles de médiation et de modulation - partie 2). Cette formation présentera les packages psych et lavaan, qui sont devenus des outils de référence en psychométrie et pour la modélisation par équations structurelles sous R. Le package ggplot2 sera également présent dans la mesure où c'est un module simple et puissant de visualisation des données sous R.

Remarques importantes

- Ce cours n'est pas un cours de statistiques classique. On travaillera avec le logiciel R
- Nous supposons que vous avez déjà une connaissance de certaines méthodes présentées ici.
- Si vous souhaitez des précisions théoriques/méthodologiques à propos d'un certain type d'analyses, nous vous conseillons de voir la doc !

En statistique, comme dans la vraie vie, on se pose des questions, et on essaie d'y répondre. Le statisticien cherche à modéliser...

- 1 Voir et représenter les données.
- 2 Choisir le type de modèle.
- 3 Ajuster le modèle.
- 4 Valider le modèle.
- 5 Selon les besoins, faire de l'inférence (tests, régions de confiance...), de la prédiction etc.

En statistique, comme dans la vraie vie, on se pose des questions, et on essaie d'y répondre. Le statisticien cherche à modéliser...

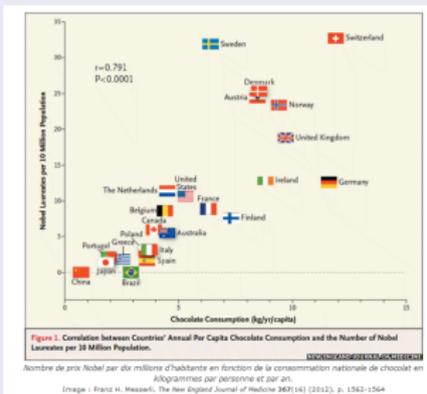
Quel traitement ?



Question : 2 traitements : Traitement A ou Traitement B ?

Question : Plusieurs traitements... Comparer l'effet des traitements (A,B,C,D) sur la dépression. Les traitements ont-ils des effets significativement différents ?

Chocolat et prix Nobel



THE NEW ENGLAND JOURNAL OF MEDICINE

OCCASIONAL NOTES

Chocolate Consumption, Cognitive Function, and Nobel Laureates

Franz H. Messerli, M.D.

Dietary flavonoids, abundant in plant-based foods, have been shown to improve cognitive function. Specifically, a reduction in the risk of dementia, enhanced performance on some cognitive tests, and improved cognitive function in elderly persons with mild impairment have been associated with a regular intake of flavonoids.^{1,2} A subclass of flavonoids called flavanols, which are widely present in cocoa, green tea, red wine, and some fruits, seems to be effective in slowing down or

cause the population of a country is substantially higher than its number of Nobel laureates, the numbers had to be multiplied by 10 million. Thus, the numbers must be read as the number of Nobel laureates for every 10 million persons in a given country.

All Nobel Prizes that were awarded through October 10, 2011, were included. Data on per capita yearly chocolate consumption in 22 countries was obtained from Chocovissime

- Données: 22 pays (The new england journal of medicine)
- Entrée: Consommation de chocolat par habitant
- Sortie: Nombre de lauréats du prix Nobel

Corrélation et causalité? Synonymes? NON!

- Quelle est la cause, quel est l'effet ?
- Un lien statistique n'est pas toujours le signe d'une causalité

Motivation : Un exemple intéressant

Etude aux USA (Baystate Medical Center, 1986)

Quels sont les facteurs pouvant augmenter le risque de donner naissance à des bébés de faible poids inférieur à 2500 g ?

- library MASS de R : 189 individus et 10 variables
 - âge de la mère,
 - race l'ethnicité de la mère,
 - smoke (1 si tabagisme pendant la grossesse, 0 sinon),
 - nombre de visites médicales durant le premier trimestre de grossesse,
 - poids du bébé à la naissance,
 -

Biblio : Venables, W. N. and Ripley, B. D. (2002) Modern Applied Statistics with S. Springer.

Deux questions naturelles se posent avec ces données:

- ① estimer la probabilité qu'un bébé naisse avec un faible poids à partir des variables explicatives
- ② estimer le poids du bébé à la naissance à partir des variables explicatives

En 1966, Sternberg : Mémoire à court terme

- But : Étudier la dynamique de la mémoire à court terme.
- On présente à un individu un ensemble d'items (ex : chiffres à mémoriser) sur un écran.
- Après la présentation de cette liste, on présente un item test pour lequel le sujet doit dire si oui ou non il faisait partie de la liste qui venait d'être présentée.
- Données : (Biblio Howell, fichier Tab2-1.dat)
- Les variables sont les suivantes :
 - RxTime: temps de reconnaissance
 - NStim : nombre d'items à mémoriser
 - YesNo : présence (Yes) ou absence de l'item (No).

Deux questions naturelles se posent avec ces données:

- 1 Peut-on expliquer RxTime en fonction de NStim ?
- 2 Y a-t-il une différence significative entre les temps de réactions des 3 catégories ?

Analyser, interpréter et mettre en forme ses données à l'aide du logiciel R

- La question ici est de comment exploiter l'ensemble des données recueillies au cours de la recherche ?
- Comment faire le lien entre l'ensemble de ces données ?
- Quel est le problème à traiter ?

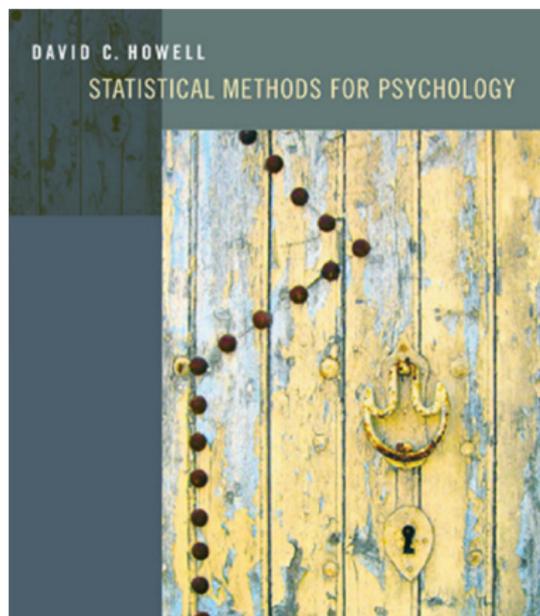
Questions du jour (partie 1)

- Résumer les variables du jeu de données
- Représenter les variables quantitatives et qualitatives.
- Se poser des questions ...
- Modélisation statistique et validation des hypothèses

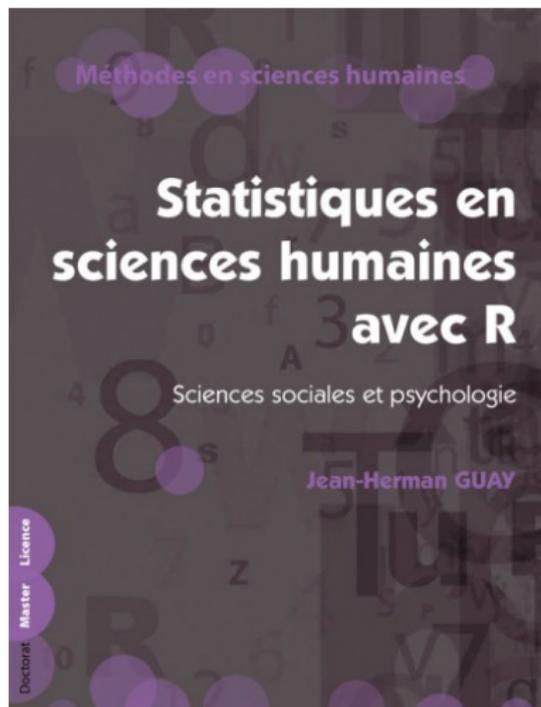
Thèmes abordés dans ce cours-atelier R

- Analyse statistique pour des variables qualitatives et quantitatives sous R
- Test d'hypothèses sous R
- Modèle linéaire : régression, anova sous R
- Sélection de modèles sous
- Régression logistique sous R
-

Statistical Methods for Psychology (Biblio)



Analyse de données avec R (Biblio)



- 1 Learning Statistics with R (Danielle Navarro)
<https://learningstatisticswithr.com/>
- 2 Learning statistics with Jamovi (Danielle Navarro and David Foxcroft) <https://blog.jamovi.org/2018/10/25/learning-statistics-with-jamovi.html>