

## **RAFRAÎCHISSEMENT STATISTIQUE**

**19-20 avril 2016, Université de Genève**

(19 avril : salle M3389 Uni-Mail / 20 avril : salle PS03 Uni-Pignon)

*Gilbert Ritschard, Matthias Studer, Dominique Joye*

Cet atelier propose de faire une synthèse des bonnes pratiques en matière d'analyse de données et de la régression. Pour ce qui est de l'analyse de données, l'atelier traitera des méthodes factorielles (ACP, AFC, ACM, ...), y compris les opérations de rotation, et de la classification automatique (hiérarchique, partitionnement autour de centres, ...). S'agissant de la régression, on examinera la régression linéaire, mais aussi des modèles linéaires plus généraux comme la régression logistique. L'utilisation des méthodes sera illustrée sur des données de LIVES telles que VLV. Pour la partie pratique, on explorera quelques possibilités offertes dans l'environnement R.

Le module accueillera un maximum de 20 participant-e-s.

Les inscriptions se font directement sur intranet à l'adresse suivante :

<https://intranet.lives-nccr.ch/egroupware/spisession>

.....

## **STATISTICAL REVISION**

**April 19-20, 2016, Geneva University**

(April 19 : Room M3389 Uni-Mail / April 20 : Room PS03 Uni-Pignon)

*Gilbert Ritschard, Matthias Studer, Dominique Joye*

This workshop offers a synthesis of best practices in data analysis and regression. In terms of data analysis, the workshop will address the factor analysis (PCA, CA, MCA,...), including the operations of rotation, and automatic classification (hierarchical, partitioning around centres, ... ). With regard to the regression, we will consider the linear regression, but also more general linear models such as logistic regression. The use of the methods will be illustrated on LIVES data such as VLV. In the practical part, a few opportunities in the environment R.

The workshop will accommodate a maximum of 20 participants.

Registrations are directly on the intranet at the following address:

<https://intranet.lives-nccr.ch/egroupware/spisession>

## SCHEDULE

### April 19: Data Analysis (🔄 at Uni-Mail)

#### **Morning: Clustering**

From To Room

9h 12h30 M3389 *Matthias Studer*  
 Some uses of clustering in the social sciences.  
 Main clustering methods.  
 Interpreting the results.  
 Choosing a method and the number of group using cluster quality measures.  
 Some common pitfalls of cluster analysis.

#### **Afternoon: Factorial Methods**

From To Room

14h 17h30 M3389 *Dominique Joye*  
 Principal components.  
 To rotate or not to rotate?  
 Factor models and extensions.

### April 20: Data Analysis (🔄 at Uni-Pignon)

#### **Morning: Regression models**

From To Room

9h 12h30 PS03 *Gilbert Ritschard*  
 Interpreting coefficients of linear and logistic models.  
 Model quality and statistical significance of covariate effects.  
 Presenting regression results.

#### **Afternoon: Modeling strategies**

From To Room

14h 17h30 PS03 *Matthias Studer*  
 Research hypotheses on the relationship between independent and dependent variables.  
 Constructing models.  
 Comparison of model outcomes.  
 Drawing conclusions on the hypotheses