

# Régression et Modélisation

**Professeur : Laurent Lesnard et Milan Bouchet-Valat**

**Année universitaire 2018/2019 : semestre de printemps**

## *Objectif et description de l'enseignement :*

Cet enseignement de 24 heures constitue le deuxième module d'une série de quatre qui, considérés dans leur globalité, visent à donner une formation progressive et méthodique aux aspects fondamentaux de l'analyse statistique univariée, bivariée et multivariée et à sa pratique en sociologie et, plus généralement, en sciences humaines et sociales. L'objectif général poursuivi est de doter les étudiants du Master de sociologie des compétences de base leur permettant, d'une part, d'analyser, à des fins de recherche ou d'étude, les données issues d'une grande enquête statistique ou d'une base de données, d'autre part, de disposer d'un bagage initial suffisamment solide leur permettant ultérieurement d'aborder, seuls ou dans le cadre de formations complémentaires du PAMSS, des méthodes statistiques plus pointues ou plus spécifiques dont ils pourraient avoir besoin dans leur recherche.

Ce deuxième module aborde les aspects fondamentaux de l'analyse de régression ainsi que les principales méthodes d'analyse de régression. Il constitue ainsi une introduction générale à la modélisation statistique. Il se déroule en salle informatique équipée de Stata, en présence de deux enseignants simultanément, les apports de contenu étant ainsi illustrés et mis en pratique par l'utilisation et la maîtrise progressive d'un logiciel statistique généraliste aujourd'hui largement utilisé dans les sciences sociales.

## *Plan de cours (liste des intitulés des séances) :*

- 1 – Utilité et principes fondamentaux des analyses de régression – Rappels d'utilisation de Stata
- 2 – Corrélation et régression linéaire simple
- 3 – Régression linéaire multiple avec des variables explicatives quantitatives et/ou qualitatives
- 4 – Régression linéaire multiple avec interaction : entre deux variables quantitatives, entre deux variables qualitatives, entre une variable quantitative et une variable qualitative
- 5 – Régression linéaire multiple avec interaction (fin)
- 6 – Régression logistique simple et multiple
- 7 – Régression logistique avancée : prédictions et effets marginaux, interaction, problème de l'hétérogénéité non observée
- 8 – Régression logistique avancée (fin)
- 9 – Régressions ordinales et multinomiales
- 10 – Régressions catégorielles avancées
- 11 – Spécification des régressions et limites de la causalité
- 12 – Régressions sur données complexes : l'exemple des panels

***Travail attendu et mode d'évaluation :***

Le module sera validé à partir de la réalisation (par groupe de deux étudiants) d'un mini-mémoire (une vingtaine de pages environ) portant sur une (ou deux) question(s) de recherche empirique qui sera (seront) traitée(s) en utilisant les données d'un ou plusieurs pays, issues de l'Enquête Sociale Européenne (*European Social Survey*, documentation et téléchargement :

<http://www.europeansocialsurvey.org/> et <http://ess.nsd.uib.no/index.jsp>) (vagues 1, 2, 3, 4, 5, 6 et/ou 7). Cette validation est commune au troisième module qui est enseigné simultanément au cours du second semestre du Master 1 de Sociologie. Pour cette raison, le mini-mémoire sera composé de deux parties distinctes, éventuellement, mais non obligatoirement, reliées, l'une utilisant l'analyse de régression, l'autre mettant à profit l'analyse de données et/ou la classification.

*Courte bibliographie sélective*

- Bressoux P., 2010. – *Modélisation statistique appliquée aux sciences sociales*, Bruxelles, De Boeck.
- Cameron A. C., Trivedi P. K., 2010. – *Microeconometrics Using Stata*, College Station (TX), Stata Press.
- Cohen J., Cohen P., West S. G., Aiken L. S., 2003. – *Applied Multiple Regression/ Correlation Analysis for the Behavioral Sciences*, Mahwah (NJ), Erlbaum.
- Howell D. C., 2008. – *Méthodes statistiques en sciences humaines*, Bruxelles, De Boeck.
- Long J. S., 1997. – *Regression Models for Categorical and Limited Dependent Variables*, Thousand Oaks (CA), Sage.
- Long J. S., Freese J., 2014. – *Regression Models for Categorical Dependent Variables Using Stata*, College Station (TX), Stata Press.
- Treiman D. J., 2009. – *Quantitative Data Analysis. Doing Social Research to Test Ideas*, San Francisco, Jossey-Bass.
- Wooldridge J. F., 2012. – *Introductory Econometrics: A Modern Approach*, Mason (OH), South-Western.