Syllabus Module 204 – Advanced Core Module in Biostatistics

Module : 204	**Advanced Core Module in Biostatistics**
Coordinator	Pascal Crépey Professor of epidemiology and biostatistics, EHESP, Métis department pascal.crepey@ehesp.fr
Dates	Wednesday September 2 to September 30, 2020
Credits/ECTS	3 ECTS
Duration	5 days of 6 hours = 30 hours
Location	On-line / EHESP 20 Avenue George Sand 93210 LA PLAINE ST DENIS
Description	If not all MPH students decide to become "biostatisticians", knowledge of biostatistics is required in almost every field of public health and its applications. Therefore, all students have to develop solid knowledge base in biostatistics. This course will present the most fundamental methods used in biostatistics including applied learning exercises by means of computer-based live examples with STATA® or R during all lectures, homework correction of exercises as well as project-based learning within groups. All students will be asked to practice and become familiar with the use of a statistical tool. Various statistical analyses with different sets of data will be conducted, from basic to advanced applications, the latter targeting students who wish to develop an in-depth knowledge of biostatistics and continuing to use biostatistics in further classes or internships. In all cases, public health field examples will highlight that course material is connected to real-life situations.
Course learning objectives	At the end of the module, the students should be able to: • Investigate a public health issue through quantitative data • Make comparisons through basic and multivariate statistical analysis from either R or STATA ® • Interpret and summarize statistical results, with a focus on logistic regression
UE Structure (details of sessions title/spaeker/date/duration)	*Day 1*: Introduction to regression modelling – Computer lab *Day 2*: Sample size and power calculation – Computer lab *Day 3:* Effect modifier – Computer lab *Day 4:* Goodness-of-fit and model building – Computer lab *Day 5:* Sensitivity analysis, Presentation and interpretation of results - Computer lab
Resources	OpenIntro Statistics, third Edition; free download on: https://www.openintro.org/stat/?stat_book=os Self practice requested on STATA or R from various learning files and online resources, notably IDRE at UCLA: http://stats.idre.ucla.edu/r/dae/
Course requirement	Students are required to have followed and validated the M1 core in epidemiology and biostatistics of the MPH, or equivalent if direct entry in M2 program.
Grading and assessment	60% final exam (2-hour table individual assignment) 10% individual mark (weekly homework) 30% group work (individual contribution & group work report) Note also that students will complete a questionnaire that assesses their own and their teammates' contributions to group work. All team members will receive the same grade except if it is clear that a student has not participated effectively (attended and contributed to meetings; made timely, helpful contributions; been constructive, etc.). In that case, the student's grade will be lowered accordingly.
Course policy	**Attendance & punctuality** **Regular and punctual class attendance is a prerequisite for receiving credit in a course.** Students are expected to attend each class. Attendance will be taken at each class. The obligations of attendance and punctuality cover every aspect of the course: - lectures, conferences, group projects, assessments, examinations, as described in EHESP Academic Regulations http://mph.ehesp.fr EHESP Academic Regulation Article. 3).

	If students are not able to make it to class, they are required to send an email to the instructor and to the MPH program coordinating team explaining their absence prior to the scheduled class date. All supporting documents are provided to the end-of-year panel. Students who miss class are responsible for content. Any student who misses a class has the responsibility for obtaining copies of notes, handouts and assignments. If additional assistance is still necessary, an appointment should be scheduled with the instructor. Class time is not to be used to go over material with students who have missed class. **Lateness**: Students who are more than 10 minutes late may be denied access to a class. Repeated late arrivals may be counted as absences (See http://mph.ehesp.fr EHESP Academic Regulation Article. 3 Attendance & Punctuality) **Maximum absences authorized & penalty otherwise** Above 20% of absences will be designated a fail for a given class. The students will be entitled to be reassessed in any failed component(s). If they undertake a reassessment or they retake a module this means that they cannot normally obtain more than the minimum pass mark (i.e. 10 out of 20) **Exceptional circumstances** Absence from any examination or test, or late submission of assignments due to illness, psychological problems, or exceptional personal reasons must be justified; otherwise, students will be penalized, as above mentioned. Students must directly notify their professor or the MPH academic secretariat before the exam or before the assignment deadline. Before accepting the student's justification, the professor or the MPH academic secretariat has the right to request either a certificate from the attending physician or from a psychologist, or from any other relevant person (See http://mph.ehesp.fr EHESP Academic Regulation Article 4 Examinations). **Courtesy**: All cell phones/pages MUST be turned off during class time. Students are required to conduct themselves according to professional standards, eating during class time is not permitted during class time, such as course or group work.
Valuing diversity	Diversity enriches learning. It requires an atmosphere of inclusion and tolerance, which often challenges our own closely-held ideas, as well as our personal comfort zones. The results, however, create a sense of community and promote excellence in the learning environment. This class will follow principles of inclusion, respect, tolerance, and acceptance that support the values of diversity. Diversity includes consideration of: (1) life experiences, including type, variety, uniqueness, duration, personal values, political viewpoints, and intensity; and (2) factors related to "diversity of presence," including, among others, age, economic circumstances, ethnic identification, family educational attainment, disability, gender, geographic origin, maturity, race, religion, sexual orientation and social position.
Course evaluation	EHESP requests that you complete a course evaluation at the end of the school year. Your responses will be anonymous, with feedback provided in the aggregate. Open-ended comments will be shared with instructors, but not identified with individual students. Your participation in course evaluation is an expectation, since providing constructive feedback is a professional obligation. Feedback is critical, moreover, to improving the quality of our courses, as well as for instructor assessment.

Session 1	**Introduction to regression modelling**
Speakers	**Lecturers**: Pascal Crépey, Professor of epidemiology and biostatistics, Métis department Pascal.Crepey@ehesp.fr **Conference speaker**: Simon Combes, Professor of economics, SHS department Simon.Combes@ehesp.fr **Tutors for Group works**: Pascal Crépey, Jean-Baptiste Combes, Mélanie Bertin
Learning Objectives	At the end of the session, the students should be able to: - To explain the rationale for choosing a logistic model accurately - To discuss which question the model could answer - To apply logistic regression with Stata or R
Duration	6 hours
Training methods	Lecture: logistic regression with computer-based live examples with Stata/R Learning exercises: individual, with Stata/R Project based learning: in small groups, with Stata/R

Session 2	**Sample size and power calculation**
Speakers	**Lecturer**: Jonathan Roux, Researcher, Métis department, Jonathan.roux@ehesp.fr **Conference speaker**: Simon Combes, Professor of economics, SHS department Simon.Combes@ehesp.fr **Tutors for Group works**: Pascal Crépey, Jean-Baptiste Combes, Mélanie Bertin
Learning Objectives	At the end of the session, the students should be able to: - To calculate required sample size (a priori) to test a research hypothesis - To estimate power (a posteriori) of a study based on available datasets - To apply these calculations to logistic regression with Stata/R
Duration	6 hours
Training methods	Lecture: sample size and power analysis, with computer-based live examples with Stata Conference: application of logistic regression to study self-perceived health among migrants Learning exercises: individual, with Stata/R Project based learning: in small groups, with Stata/R

Session 3	**Effect modifier**
Speakers	**Lecturer**: Séverine Deguen, Professor of biostatistics, DSET department Severine.Deguen@ehesp.fr **Conference speaker**: Simon Combes, Professor of economics, SHS department Simon.Combes@ehesp.fr **Tutors for Group works**: Pascal Crépey, Jean-Baptiste Combes, Mélanie Bertin
Learning Objectives	At the end of the session, the students should be able to: - To define and asses collinearity and interaction - To detect collinearity and interaction, and deal with them - To apply these concepts and methods to multivariate logistic regression with Stata
Duration	6 hours
Training methods	Lecture: collinearity and interaction, with computer-based live examples with Stata Learning exercises: individual, with Stata/R Project based learning: in small groups, with Stata/R

Session 4	**Goodness of fit and model building**
Speakers	**Lecturer**: Séverine Deguen, Professor of biostatistics, DSET department Severine.Deguen@ehesp.fr **Conference speaker**: Mélanie Bertin, Professor of epidemiology, Métis department, melanie.bertin@ehesp.fr **Tutors for Group works**: Pascal Crépey, Jean-Baptiste Combes, Mélanie Bertin
Learning Objectives	At the end of the session, the students should be able to: - Verify the quality of the model and its ability to make predictions - Apply the Hosmer and Lemeshow test and interpret results - Make a decision on the best model(s) to retain
Duration	6 hours
Training methods	Lecture: GOF, Hosmer & Lemeshow test, choice of model, with computer-based live examples with Stata Learning exercises: individual, with Stata/R Project based learning: in small groups, with Stata/R

Session 5	**Sensitivity analysis - Presentation and interpretation of results**

Speakers	**Lecturer**: Pascal Crépey, Professor of epidemiology and biostatistics, Métis department Pascal.Crepey@ehesp.fr **Tutors for Group works**: Pascal Crépey, Jean-Baptiste Combes, Mélanie Bertin
Learning Objectives	At the end of the session, the students should be able to: - Define and apply sensitivity analysis in the context of multivariate logistic regression - To select and report results of multivariate logistic regression adequately - To discuss main strengths and limitations of the statistical analysis performed
Duration	6 hours
Training methods	Lecture: sensitivity analysis, report and interpretation of results, with computer-based live examples with Stata Learning exercises: individual, with Stata/R Project based learning: in small groups, with Stata/R