

## **SPPH 500 – Analytic Methods in Epidemiological Research Course Syllabus January 2016**

**DAYS & TIMES:** Lectures run every Wednesday 1:00pm-4:00pm in room Woodward 4

**INSTRUCTOR:** Mike Marin

**EMAIL:** [mike.marin@ubc.ca](mailto:mike.marin@ubc.ca), [marin.mike@gmail.com](mailto:marin.mike@gmail.com)

**OFFICE:** The School of Population and Public Health, The University of British Columbia  
Room 263, SPPH Building

**TELEPHONE:** 604-827-1596 (UBC)

**OFFICE HOURS:** TA office hours will be listed on the website and scheduled by the TA once the course has begun. My office hours will be by appointment (email me or talk to me before/after class to arrange a time to meet)

**TEACHING ASSISTANTS:** The teaching assistants (TAs) for this course will be introduced at the beginning of the course. The TAs will hold weekly tutorials/workshops as well as offer regular office hours. Times and locations will be finalized once the course begins. Once this is finalized, it will be shared with you as well as posted on the course website.

**COURSE WEBSITE:** [www.connect.ubc.ca](http://www.connect.ubc.ca) (use your CWL login and access our course page)

### **COURSE OVERVIEW:**

SPPH 500 is a 3 credit graduate course. It is an introduction to regression modeling of epidemiologic data to broaden analytic skills acquired in previous courses. This course covers statistical techniques commonly used in epidemiology and health services research, such as linear, logistic and Poisson regression as well as survival analysis (Cox regression).

Greater emphasis will be placed on conceptual understanding rather than computational ability. In other words, formulating the proper model to fit will be considered more important than the actual fitting of the model using software.

**PREREQUISITES:** SPPH 400, SPPH 502 as well as working knowledge of statistical software.

### **COURSE TEXTBOOK:**

*Vittinghoff E, Glidden D, Shiboski S and McCulloch “Regression Methods in Biostatistics”, Springer 2005*

*Holford T “Multivariate Methods in Epidemiology”, Oxford University Press, 2002.*

*\* I also include a set of notes on course topics*

## **SOME ADDITIONAL REFERENCES:**

1. Rosner B “*Fundamentals of Biostatistics*”, 5<sup>th</sup> Ed, Duxbury, 200
2. Katz MH, “*Multivariable Analysis, A practical Guide for Clinicians*”, Cambridge University Press, 2006

## **LEARNING OBJECTIVES:**

Upon completion of this course, the student should be able to:

- Relate regression methods to appropriate types of epidemiologic data
- Perform linear, logistic, Poisson and Cox regression analysis with multiple variables
- Interpret coefficients of each regression model and summarize results
- Understand the assumptions and limitations of such models
- Communicate effectively with statistical practitioners on regression methods

**\*\* Please see detailed list of Learning Objectives posted on Course Website.**

## **STATISTICAL COMPUTING:**

Statistical computing in this course will be done using R, a freely available, open source ‘clone’ of S-Plus, and can be downloaded for free at [www.r-project.org](http://www.r-project.org). You may also use other softwares (such as SAS, STATA, etc) although the amount of support we can offer for these may be limited.

A set of introductory video tutorials for R can be found on my YouTube Channel:  
<http://www.youtube.com/marinstatlectures>

## **COURSE EVALUATION:**

Assignments (4) 60%

Final Examination: 40%

\* Must pass the final to pass the course

## **ASSIGNMENTS:**

The assignments are designed to help students master the concepts presented in class and gain experience in data analysis, write-up and interpretation. Students are encouraged to work together in teams and discuss their ideas but you will likely find that you will learn the material most thoroughly by first attempting to work out the solutions yourself. Each student is expected to submit his or her own original written solutions, prepared by himself/herself. Academic dishonesty will not be tolerated.

Assignments should be neatly written (or typed). This is not a thesis, but it should still look like something you are proud to have your name on. Some marks will be allocated to organization and presentation of answers.

When preparing solutions related to data analysis, include only those parts of the computer output that are relevant to your answer and highlight or underline the specific items of interest. Alternatively, transcribe those items to another page if you prefer. Computer code used to produce the results is not necessary to include with your assignment.

A marking rubric to the “write-up” portion of the assignments will be provided to you.

Late assignments will be penalized 15% per day late. Assignments will not be accepted beyond 48 hours of the time due (unless a late submission has been discussed and agreed upon with the instructor. Extensions will only be granted under extenuating circumstances (being busy with things does not classify as extenuating circumstances)

### **COURSE NOTES:**

The course consists of a mix of reading through some notes provided online, as well as text book chapters. It will be assumed that you have read through these prior to attending lectures, and to bring a copy with you to class, if necessary. In our meetings, we will work through these and expand on the ideas presented within. It will be useful for you to bring a laptop to lectures, as we will often work through the ideas using software in lectures.

### **COURSE MODULES:**

The course has been divided into three separate modules. The first discussing multiple linear regression, the second discussing logistic regression, and the third and final module discussing time-to-event data (Poisson and Cox regression). A set of notes will be provided for each. You should be reading through each set of module notes prior to attending classes.