

## **SPPH 500 (DL) – Analytic Methods in Epidemiological Research Course Syllabus January 2015**

**DAYS & TIMES:** Days/times of face-to-face meeting will be emailed directly to you as well as appear on the course website.

**LOCATION:** Location for face-to-face meetings will be emailed directly to you as well as appear on the course website.

**INSTRUCTOR:** Mike Marin

**EMAIL:** [mmarin@stat.ubc.ca](mailto:mmarin@stat.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:** Office hours will be on-line and scheduled by the TA once the course begins

**TEACHING ASSISTANTS:** The teaching assistant (TA) for this course will be introduced at the beginning of the course. The TA will hold weekly on-line tutorials as well as be available to offer support over email and Wimba (on-line classroom).

### **COURSE OVERVIEW:**

SPPH 500 DL 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

### **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/user/marinstatlectures>

### **COURSE EVALUATION:**

Assignments (3) 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 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 written solutions (prepared by himself/herself).

- 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.
- Mathematical notation should be used carefully. Make sure your approach to a problem is clearly outlined and explained.
- 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.

### **COURSE WEBSITE:**

This is a distance learning course, so most material will appear on the course website. It is crucial that you visit this website regularly as well as participate in on-line tutorials and discussion forums. On the course website you will find a weekly schedule to help you keep on pace with this course.

The course website can be found at: [www.connect.ubc.ca](http://www.connect.ubc.ca) (use CWL to login)

### **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 the face-to-face meetings. The TA will also hold a few on-line tutorials for each module prior to the face-to-face meetings. In each meeting we will review the module reviewed by yourself and the TA, prior to our meeting.