CAMCCO-L

Intermediate Biostatistics: Strategies for Multivariable Regression Modelling

Part 2

Course Syllabus   **2024-2025**

ONLINE COURSE

MODULE 5
Methodology

DATE/TIME
Thursday, March 27th, 2025
13:30 to 16:30 EDT

LANGUAGE
English

FREE REGISTRATION
www.camccol.com

PRESENTED BY

Dr. Steven Hawken
Associate Professor, University of Ottawa

✉ shawken@ohri.ca

DESCRIPTION

This course will target statisticians, epidemiologists, data scientists and other quantitative researchers/students with a basic familiarity with regression modeling. The course will cover general strategies for fitting prediction models for continuous, categorical and time-to-event outcomes, including: exploratory analysis/data visualization; missing data imputation; covariate selection; model specification; model validation/calibration; handling non-linearity; and choosing between conventional statistical models and machine learning models (and the differences between these types of models). Extensive use of R, RStudio and Frank Harrell's Hmisc and rms r-packages will be used in the course material and case-studies/examples. The course will follow the general philosophy of Regression Modelling Strategies - 2nd Edition textbook by Frank Harrell.

LEARNING OBJECTIVES

- Methods for exploring, describing and understanding your data in preparation for regression modeling
- Fitting multivariable regression models appropriate for continuous, categorical, and time to event outcomes
- Address issues of sample size and overfitting
- Approaches to addressing missing data
- Handling complex non–linear or non–additive relationships
- Testing/quantifying associations between one or more predictors and the response, and interpreting the fitted model
- Model validation and calibration to evaluate predictive accuracy and identify overfitting
- Learn the differences between machine learning and statistical models, and how to choose the best approach for a given problem