EPID 822 APPLIED REGRESSION ANALYSIS

COURSE OUTLINE – WINTER 2023

Course Number: EPID 822

Title: Applied Regression Analysis

Instructors

Name	Office	Phone	Email
Bingshu Chen	Canadian Cancer Trial Group	77703	bechen@ctg.queensu.ca
Wei Tu	Canadian Cancer Trial Group	77399	wtu@ctg.queensu.ca
Zihang Lu	Public Health Sciences	79336	zihang.lu@queensu.ca
SAS Lab Instructors			
Patric Norman	KGH Clinical Research Unit		pan@queensu.ca
Andrew Day	KGH Clinical Research Unit	32442	Andrew.Day@kingstonhsc.ca
Teaching Assistant			
Tristan Derry	Carruther's Hall		15trd1@queensu.ca

Time of Classes
Lecture: Tuesday 9:30 – 11:00am, Thursday 9:30 – 11:00 am.

SAS tutorial:
Thursday 11:30 am to 13:00 pm

TA weekly office hours: To be discussed with students during week 1
TA weekly tutorial: To be discussed with students during week 1

Location of Classes
Lecture/Seminar: Tuesday: Jeffery 234, Thursday: Macintosh-Corry room D216
SAS/Tutorial: Law building 2

Pre-requisites or Co-requisites
EPID-821 and EPID821 SAS lab quiz (pass 50% threshold).

Instructional Objectives
At the end of the course, students should be able to
- Understand the use of some complex statistics in the medical literature.
- Have a basic understanding of the concepts, logic, and numerical steps involved in the development of some commonly used statistical models in epidemiology.

- Apply these statistical models to analyze data for research projects using computer packages and interpret results.

Course Content

Topics include: two-way analysis of variance and covariance, multiple regression, analysis of categorical data, logistic regression, Poisson regression and survival analysis.

Enrolment Limit

No more than 30

Target Group

Graduate students from the Department of Public Health Sciences, and other departments in and outside the faculty of medicine, who needs Biostatistics for their study projects.

Principal Mode of Instruction

Lectures and practicals (3 hours/week) and exercise (take home).

Reference Books

Recommended

Rosner, B. (2011) *Fundamentals of Biostatistics*, 7th Edition. Brooks/Cole, Cengage Learning.

Hassard, T.H. (1991) *Understanding Biostatistics*. St. Louis: Mosby Year Book.

Altman D.G. (1991) *Practical Statistics for Medical Research.* London: Chapman and Hall.

Kleinbaum, D.G., Kupper, L.L. Muller, K.E. and Nizam A. (1998) *Applied Regression Analysis and Other Multivariable Methods* 3rd Edition. Belmont: Duxbury Press.

Selvin, S. (1996). *Statistical Analysis of Epidemiologic Data* 2nd Edition. Oxford: Oxford University Press.

Selvin, S. (1995) *Prictical Biostatistical Methods*. Belmont: Duxbury Press.

Fisher, L.D. and van Belle, G. (1993) *Biostatistics: a Methodology for the Health Sciences*. New York: John Wiley & Sons.

Kleinbaum, D.G. (1994) *Logistic Regression — A Self-Learning Text*. New York: Springer-Verlag.

Frome EL and Checkoway H (1985) Epidemiologic programs for computers and calculators: Use of Poisson regression in estimating incidence rates and ratios. *Am J Epidemiol* 121 (2): 309-23.

Zou G (2004) A modified Poisson regression approach to prospective studies with binary data. *Am J Epidemiol* 149(7): 702-6.

Cantor A. SAS Survival Analysis Techniques for Medical Research, SAS Institute, NC, 2003.

Allison PD Survival Analysis Using the SAS System: A Practical Guide, SAS Institute, NC, 1995.

Method of Assessment

Assessment:	Homework	25%
	SAS Programming	15%
	Mid-term	30%
	Final Exam.	30%

Please note: The minimum passing grade in Graduate School is 70% for this course. The midterm and final exams are closed-book exams. Students are allowed to have one sheet (letter size, double sided) of formulas prepared by themselves and a non-programmable and non-graphical calculator in the exams. Late homework and SAS assignments without valid reasons will only receive 75% of marks if handed in before solutions are posted.

Session Information

Week 1. Two way analysis of variance (Jan 10, 12. Chen)

This session extends the ANOVA technique to analyze continuous data with two classifying variables. Parametric and non- parametric methods will be presented. The concept of interaction will be introduced.

- Two-way ANOVA
- Simple repeated measurement design (randomized blocks design)
- Interactions
- Analysis of matched-pairs
- Propensity score matching
- Assumptions
- Non-parametric two way ANOVA — Friedman's test
- Fixed effects vs random effects
- Mixed effects models

Reading Rosner (2011, Chapter 12.6); Kleinbaum (2013, *et al* Chapter 15); Selvin(1995, Chapter 5); Van Belle et al. (2004, Chapter 10.3).
Assignment 1 due: January 24

Week 2- 6. Multiple regression (Jan 17, 19, 24, 26, 31, Feb 2, 7, 9, 14. Chen)

Multiple regression extends the simple linear regression to allow us to predict the value of a continuous variable from the value of *several* other variables. Variable selection procedure to choose regression model will be introduced. Analysis of residuals and regression diagnostics will also be emphasized.

- Review of simple linear regression and correlation
- Linear regression with two independent variables
- Multiple and partial correlations
- Multiple regression
- Regression, ANOVA and dummy variables

- Testing hypotheses in multiple regression
- Variable selection
- Goodness of fit
- Analysis of residuals and regression diagnostics
- Confounding and interaction in regression
- Dummy variables in regression
- The ANCOVA model
- Test of parallelism
- Using dummy variables
- Interpretation of the model

Reading (Rosner (2011, 11.9 - 11.11; Selvin (1995), Chapter 4; Kleinbaum *et al* (2013, Chapters 8-12, 13-14)
Assignment 2 due: Feb 28.

MID-TERM TEST (Feb 16, Chen)

Feb 16 (9:00 am – 11:30am): Week 1 to week 5. Need book a room from 9am to 12noon.

Reading Week: Feb 20 to 24.

Week 7. Analysis of categorical data using logistic model (Feb 28, Mar 2. Zihang)
- Simple logistic function and logistic regression
- Logistic model and its relation with odds ratio
- Fitting logistic models to contingency tables
- Estimating and interpreting parameters
- Confounding and interaction
- Effect modification

Reading (Salvin(96) Chapter 7; Rosner Chapter 10, 13.8)

Week 8-9. Multiple logistic regression (Mar 7, 9, 14, 16. Zihang)
These sessions emphasize the use of logistic regression to analyze dichotomous response data in relation to multiple independent variables. Selecting and testing appropriate logistic models will be discussed.
- Multiple logistic regression
- Variable selection
- Regression diagnostics and goodness of fit
- Interpretation
- Conditional logistic regression for matched case-control studies

Reading (Kleinbaum *et al* Chapter 23; Rosner 13.8; Salvin (96) Chapter 8-9; Selvin (95), Chapter 10)

Week 10. Poisson regression (Mar 21, 23. Wei)

This session presents the regression approach to analyze the outcome data being frequency counts (such as disease incidence) or rates (incidence rates) in relation to risk factors and covariates (possible confounding variables).

- Person-years
- Analysis of incidence rates
- Poisson distribution
- Adjusted rates and SMR
- Poisson regression
- Poisson regression and logistic regression

Reading Selvin (1995, Chapter 12); Zou (2004), Frome and Checkoway (1985).

Week 11-12. Analysis of time to event data (Mar 28, 30, Apr 4, 6. Wei)

These sessions present methods for the analysis of time to event (censored) data. Life table and Kaplan-Meier estimates and logrank test in comparing survival curves will be introduced. Cox regression model will also be introduced if there is time.

- Censored data
- Kaplan-Meier survival curve
- Life table analysis
- Comparing survival curve using logrank test
- Cox's regression

Reading Rosner (2011, Chapter 14); Allison (1995), Cantor (1997).

FINAL EXAM

Two hours and 45 minutes (*To be confirmed*): Week 6 to week 12.

Course and University Policies

Participation policy
Students are expected to demonstrate professionalism by being present and punctual, and by participating actively in all sessions. Students must inform the Instructor in advance if they will be absent from a session because of illness or other compassionate grounds. If a session is missed, it is up to the student to gather missed information from other students who were present in order to fulfill their assessment requirements.

Academic Integrity
Academic integrity is constituted by five core fundamental values: honesty, trust, fairness, respect and responsibility. Queen's students, faculty, administrators and staff all have responsibilities to support and uphold the fundamental values of academic integrity.
http://www.queensu.ca/calendars/sgsr/AcademicIntegrity Policy.html

Accommodation of Graduate Students with Disabilities

Queen's University is committed to achieving full accessibility for persons with disabilities. Part of this commitment includes arranging academic accommodations for students with disabilities to ensure they have an equitable opportunity to participate in all of their academic activities. If you are a student with a disability and think you may need accommodations, you are strongly encouraged to contact the Student Wellness, Accessibility Services Office and register as early as possible. For more information, including important deadlines, please visit the Accessibility Services website at:
http://www.queensu.ca/studentwellness/accessibility-services

Helpful Resources Available to You

School of Graduate Studies Habitat
This amazing resource brings together resources for living well and staying well in grad school. There is information on: taking care of yourself and managing stress, finding friends and fun, living in Kingston, managing finances, accessibility, building a career and more.
http://www.queensu.ca/sgs/current-students/sgs-habitat

StudentWellness Services
Student Wellness Services supports the personal, academic, and social development of students at Queen's by providing a range of programs and services including accessibility services, counselling services, and health services. Their mission is to provide a welcoming, confidential, and integrated service that is responsive to the needs of students. A full listing of services including how to make an appointment is here:
http://www.queensu.ca/studentwellness/

Student Academic Success Services
Student Academic Success Services (SASS) comprises Learning Strategies and the Writing Centre. They offer academic support to students who wish to develop their skills in critical thinking, reading, learning, studying, writing, and self-management. http://sass.queensu.ca/

Queens Libraries
You can access library materials on-site or via the internet with a Queens net ID. In addition, you can
book study space, borrow materials from other libraries, seek advice on searching for the information you need, and access information about copyright, open access and scholarly communications. For service descriptions visit:
http://library.queensu.ca/

Career Services
Career Services provides career education and employment support services at Queen's for students in
all disciplines. A comprehensive range of services are offered including drop-in career advising, supporting graduate students in making informed decisions about career options, job search strategies, and CVs/resumes. For a full description of services see: http://careers.queensu.ca/