**Intermediate Statistics for Health Service Researchers**

**HAD 5772**

**Winter 2024**

**Instructor**: Alex Hoagland, Ph.D.

Assistant Professor of Health Economics, IHPME

155 College Street, Suite 440

[alexander.hoagland@utoronto.ca](mailto:alexander.hoagland@utoronto.ca)

**Class time and location**: Wednesdays, 9:30am to 11:30am. HSB XXX

**Office hours:** Book appointments at [calendly.com/Hoagland-office-hours/had5744-2022f](https://www.calendly.com/Hoagland-office-hours/had5744-2022f)

* Tuesdays and Thursdays: 1:00pm to 3:00pm. All appointments are by Zoom unless arranged otherwise in advance.
* Please note that I typically respond to emails and other class communication during normal “business hours,” and not on evenings, weekends, or holidays.

**Course Description:** This course is designed to provide introductory background in statistical thinking, and to prepare students in the following areas: correlation; regression; analysis of variance (including factorial, repeated-measures, mixed-design, multivariate); analysis of covariance; logistic regression and factor analysis.

**Course Objectives:**

At the end of this course, students should be able to:

* Demonstrate an understanding (both conceptual and practical) of analytical approaches
* Identify appropriate analytical approaches to address specific research questions;
* Identify appropriate analytical approaches to use with collected data
* Use the computer program R for data management, statistical exploration, and analysis
* Understand, explain, and present analytical results.

**Method of Instruction**: There are lectures but the course is designed to foster a seminar atmosphere. Consequently, in-depth discussion of the theoretical economics literature as it pertains to health economics is expected. Course readings are assigned, and participants are expected to have read these prior to class.

**Evaluation Criteria**

* Assignment 1: 1 group assignment, worth 30% of the final grade.
* Midterm Assessment: done in class, worth 30% of the final grade.
* Individual Final Assessment: done in class, worth 40% of the final grade.

Assignment: There will be one assignment, which can be completed in groups. The purpose of each assignment is for you to demonstrate that you can perform the analyses covered in this course and interpret the results. Late assignments will be **discounted by 10 percentage points per day late.**

Assessments: A final assessment will be administered on the last day of class. Sample exams will be posted prior to the examination. Exams will combine responses to theoretical questions and applied work (in R).

**Grading scale:** Courses taken for graduate credit are assigned a letter grade according to the School of Graduate Studies usage as follows:

|  |  |  |
| --- | --- | --- |
| Letter Grade | Grade Meaning | Numerical Marks (%) |
| A+ | Excellent | 90%-100% |
| A | 85%-89.9% |
| A- | 80%-84.9% |
| B+ | Good | 77%-79.9% |
| B | 73%-76.9% |
| B- | 70%-72.9% |
| FZ | Inadequate | 0-69% |

Note that while course grades may be collectively “curved,” no individual grades will be rounded.

**Course website and textbook:** This course has a GitHub repository that contains all relevant materials; you can access the repo at <https://github.com/alex-hoagland/HAD-5772--Winter-2024>. Our required textbook is *Intuitive Biostatistics: A Nonmathematical Guide to Statistical Thinking* (Harvey Motulsky 2010).

**Alternative Textbook options:**

* Statistical Inference, 2002 (Casella and Berger)
* Regression and Other Stories: <https://avehtari.github.io/ROS-Examples/>, <https://users.aalto.fi/~ave/ROS.pdf>
* Statistical Inference via Data Science: <https://moderndive.com/> -- this looks like the best option for the first half of the course, nothing on mixed methods though. Ask Silver?

**Course Schedule**

|  |  |  |  |
| --- | --- | --- | --- |
| **Session**  **#** | **Date** | **Lecture / Readings** | **Recommended Exercises** |
| 1 | Jan. 10 | Introduction   * Motulsky, Chapters 1-3 |  |
| 2 | Jan. 17 | Visualizing data (+ R IntRoduction)   * Motulsky, Chapters 7-9; 32 * If you are interested in R help, check out Appendix C in Motulsky | Bad/good data viz exercise |
| 3 | Jan. 24 | Measuring uncertainty I: standard errors + confidence intervals   * Motulsky, Chapters 4-6 (emphasis on 4) * Note: working on R code for this lecture | Boostrapping Exercise |
| 4 | Jan. 31 | Measuring uncertainty II: Theory of confidence intervals   * Motulsky, Chapters 12-14 * Additional reading: Chapters 10-11 |  |
| 5 | Feb. 7 | Hypothesis Testing I: *p*-values   * Motulsky, Chapter 15 * Time to catch-up, if needed | One exercise for H testing |
| 6 | Feb. 14 | Hypothesis Testing II: Statistical Significance   * Motulsky, Chapters 16-17 |  |
| N/A | Feb. 20-21 | No class – Winter Reading Week; deadline to drop course without academic penalty |  |
| 7 | Feb. 28 | Hypothesis Testing III: Practical Applications   * Motulsky, Chapters 18-21 (43?) * Time for midterm review if allowed |  |
| 8 | Mar. 6 | **Midterm Exam, In Class** |  |
| 9 | Mar. 13 | Multiple Comparisons + Survival Analysis   * Motulsky, Chapters 22-23, 29 * Examples from Chapters 27, 30, and 31 * Revisit Ch. 5 if still iffy on survival data | Something fun here? |
| 10 | Mar. 20 | Linear Regression I: Bivariate Regression Models   * Motulsky, Chapters 33-35 * If time, brief description of nonlinear regression (36) | One exercise for regression |
| 11 | Mar. 27 | Linear Regression II: Bells and Whistles   * Motulsky, Chapters 37-38 * Include notes on DAGs, Moderation/Mediation, and Bad Controls |  |
| 12 | Apr. 3 | Conclusions + Applications   * Motulsky, Chapters 44-46 * Time for catchup, if needed * Note: is there anything needed here about path/factor analysis? |  |
| 13 | Apr. 10 | **Final Exam, In Class** |  |