Date last changed: 2018-12-12

Department of Political Science
UNIVERSITY OF TORONTO

SPECIAL TOPICS:
Statistical Analysis and Inference for Political Scientists
POL 2800H 1 S
Spring 2019

Instructor: Harald Bathelt
Office: Sidney Smith Hall, Room 3121
Office Hours: Monday, 4:00-5:00 PM and
Tuesday, 1:00-2:00 PM
(or by appointment)
Phone: 416-946-0183
E-Mail: harald.bathelt@utoronto.ca

Lecture Hours: Tuesday, 10:00 AM-1:00 PM
Lecture Room: RW 107 (Ramsay Wright Laboratories)

Course Description:

This course advances important quantitative methods and techniques used in the analysis of empirical data in Political Science and other Social Sciences. It aims to provide a comprehensive understanding of statistical methods for graduate students required to (i) quantify relations and dependencies between variables and (ii) conduct statistical tests in a variety of applications. Students are required to have some background knowledge of research design, basic descriptive statistics, testing and regression analysis at the undergraduate level. The course will help students develop an intuitive, as well as a more formal understanding of these methods. Although formal language will be used, the course does not require in-depth mathematical knowledge. The requirements for the course include a basic understanding of algebra, functions, derivations, calculus, set theory, combinatorics and the sum operator. Prior to the course, it is recommended that students refresh their understanding of regression analysis and hypothesis testing by reviewing Gujarati and Porter (Introduction and Ch. 2), McNabb (Ch. 15), Anderson et al. (Ch. 14), or other similar literature.

The topics of the course include probability distributions, statistical testing and inference, as well as linear and non-linear, simple and multiple regression and correlation techniques. The application of these methods through the use of statistical software (primarily SPSS) will also be part of the course, although this is not the main focus. Canadian census data comprising a large set of socio-economic variables for metropolitan/urban areas for the years 2001, 2006, 2011 and 2016 will be the basis for the analyses conducted in class and for the assignments.

The overall goal of the course is to provide students with an understanding of the theory and methods behind statistical analyses. By the end of the course, students should be able to (i)

conduct their own statistical analyses using different inferential statistics and regression models, (ii) read and understand the key literature in the field, and (iii) extend their statistical and econometric knowledge further through independent reading.

Course Structure:

Although the course is necessarily lecture-based, there will be plenty of room for active student participation in discussions and problem-solving in the classroom. Course evaluation will be based on three assignments, seminar participation and a final take-home exam. To conserve paper, students are encouraged to print all assignments double-sided.

The focus of the course is on understanding the various models, methods and tests, rather than their lab-based application. Assignments are designed to review the important methods discussed in class. In addition to this, students should practice all methods on their own through simple exercises using SPSS and the Canadian census data set or any other database.

All assignments are expected to be handed in as a hardcopy on the due date in class. The due date for the assignments will usually be one week after they are handed out in class. The take-home exam will be handed out during the last class (April 2). The due date for the take-home exam will be April 18.

For the assignments and the take-home exam, a lateness penalty of 3% per day will be deducted from the mark for the first 7 days. After this 7-day period, papers will receive a grade of 0% (special arrangements can only be made based on proper and timely documentation, such as a doctor's note). Students are strongly advised to keep rough and draft work and hard copies of their research papers and assignments until the marked assignments have been returned.

Quercus:

The course uses the management system Quercus which helps to establish efficient communication between instructor and students. One of its advantages is that students can access their marks at any time on an individual basis. To access the Quercus-based course website, students need to go to the UofT portal login page at http://portal.utoronto.ca and log in using their UTORid and password. The Help Desk at the Information Commons and telephone assistance under 416-978-HELP can answer related questions.

Once you have logged into the portal using your UTORid and password, you will find a link on the Quercus "Dashboard" to this course website, along with links to all your other Quercus-based courses. At times, the instructor or teaching assistant may decide to send out important course information by e-mail. To this end, all UofT students are required to have a valid UofT e-mail address. You are responsible for ensuring that your UofT e-mail address is set up AND properly entered in the ACORN system.

Academic Integrity:

Academic integrity is fundamental to learning and scholarship at the University of Toronto. Participating honestly, respectfully, responsibly, and fairly in this academic community ensures that the UofT degree that students earn will be valued as a true indication of their individual academic achievement, and will continue to receive the respect and recognition it deserves. For further details and information on academic integrity, see http://www.artsci.utoronto.ca/osai/students.

Accessibility Needs:

The University of Toronto is committed to accessibility. If students require accommodations for a disability, or have any accessibility concerns about the course, the classroom or course materials, Accessibility Services should be contacted as soon as possible at accessibility.services@utoronto.ca or http://www.accessibility.utoronto.ca/.

Course Evaluation:

The course evaluation will be based on the following aspects:

5%	assignment 1 (handed out in lecture 4: January 29; due lecture 5: February 5)
10%	assignment 2 (handed out in lecture 7: February 26; due in lecture 8: March 5)
15%	assignment 3 (handed out in lecture 10: March 19; due lecture 11: March 26)
60%	take-home exam (handed out in lecture 12: April 2; due Thursday, April 18)
10%	seminar participation

Further details about the course and suggestions from students will be discussed in class.

* * * * *

Required Readings:

It is suggested using Gujarati and Porter (2009) as the primary source book for the course, as it offers a comprehensive overview of econometric methods that extends well beyond the course. It can also be used as a reference book for further research. Alternatively, Wooldridge (2016) offers a very good overview of the topics discussed in class. This book is less formal and easier to read but not as extensive. As the first part of the course focusses on probability distributions and statistical inference and since not all students may have a strong background in these areas, Anderson et al. (2016) offer an easy-to-read introduction into these topics. It may be particularly helpful in the first half of the course and as an introductory text to the more formal readings in the second half.

To become familiar with SPSS, students can use Pallant (2016) as an introduction; Wagner (2017) or Field (2017) are equally possible, or any other SPSS book or online data base. SPSS will not be a core topic in the class; therefore, those who have no experience with this software might need some introductory reading or companion of how to use it. Before buying a book, it is recommended looking into online materials, especially those accessible through the UofT library system. Plenty of instructional materials for SPSS are available online.

*Gujarati, D. N. and Porter, D. (2009): Basic Econometrics. 5th Edition. New York: McGraw-Hill.
A highly recommended book that covers much more than just introductory material. It can be used as a textbook as well as a reference frame for statistical analyses in future research projects. The book goes well beyond the scope of the course.

*Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D. and Cochran, J. J. (2016): Statistics for Business and Economics. 13th Edition. Mason, OH: South-Western (Nelson Education).
A very good, clearly structured, easy to understand, and comprehensive introduction to descriptive statistics, statistical testing and regression analysis at the undergraduate level. The customized edition is less expensive.

Alternative Suggested Readings:

Wooldridge, J. M. (2016): Introductory Econometrics: A Modern Approach. 6th Edition. Mason, OH: South-Western (Nelson Education).
Also, a very good reader on econometrics that covers introductory and some advanced topics. The book is probably a bit easier to read than Gujarati and Porter but does not cover as much advanced material.

Additional Optional Readings:

Agresti, A. and Finlay, B. (2018): Statistical Methods for the Social Sciences. 5th Edition. New York, NY: Pearson.

This text provides a math-based introduction to statistics. It covers topics from elementary probability through logistic regression. Each chapter includes problem sets with examples that are drawn from the social sciences.

Ashenfelter, O., Levine, P. B. and Zimmermann, D. J. (2003): Statistics and Econometrics: Methods and Applications. New York: Wiley.
A systematic, comprehensive and precise introduction to most course topics. The book details the necessary mathematical background and goes substantially beyond what is covered in the course; it is advisable for students with some background in statistics and/or mathematics.

Berman, E. (2018): Essential Statistics for Public Managers and Policy Analysts. 4th Edition. Washington, DC: CQ Press.
Covers a number of course topics well, with many applications and examples. It is well-suited for understanding the rationale behind some of the methods, without the use of formulas. Provides little mathematical reasoning/understanding.

Field, A. (2017): Discovering Statistics using IBM SPSS Statistics. 5th Edition. Thousand Oaks: Sage.
A detailed – sometimes too detailed – introduction to the use of SPSS at a fair price that also includes many procedures not needed in the course.

McNabb, D. (2009): Research Methods for Political Science: Quantitative and Qualitative Methods. 2nd Edition. Armonk, NY: M. E. Sharpe.
A very broad (and thick) overview of both quantitative and qualitative approaches. Part III is the most relevant for this course.

Morgan, G. A., Leech, N. L., Gloeckner, G. W. and Barrett, K.C. (2012): IBM SPSS for Introductory Statistics: Use and Interpretation. 5th Edition. New York: Routledge.
A very good, highly-rated introduction to the basic procedures in SPSS at a fair price.

Pallant, J. (2016): SPSS Survival Manual: A Step by Step Guide to Data Analysis Using IBM SPSS. 6th Edition. Berkshire, England: McGraw Hill.
A well-organized, highly-rated introduction to the basic procedures in SPSS at a very good price.

Salkind, N. J. (2017): Statistics for People Who (Think They) Hate Statistics. 6th Edition. Thousand Oaks: Sage.
A different, sometimes funny, but ultimately serious introduction to statistical methods. It is not just designed for beginners and might be interesting for those students who need some extra motivation to engage in quantitative methods.

Wagner, W. (2017): Using SPSS for Social Statistics and Research Methods. 6th Edition. Thousand Oakes: Pine Forge Press.
A brief introduction to the basic procedures in SPSS; good overview at a fair price.

Examples of Online Instructional Materials for SPSS:

Gaur, A.S. and Gaur, S.S. (2009): Statistical Methods for Practice and Research: A Guide to Data Analysis Using SPSS. 2nd Edition. New Delhi: Response Books.

Hinton, P., McMurray, I. and Brownlow, C. (2014): SPSS Explained. London: Routledge.
Both books are good introductions into using statistical methods using SPSS but like most SPSS books are very light on math. There are many more general SPSS books available online through the UofT library system, almost all of which are sufficient for the course. Among these books, numerous cover specific topics.

IBM (2015): IBM SPSS Statistics 25 Core System User's Guide.
ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/25.0/en/client/Manuals/IBM_SPSS_Statistics_Core_System_User_Guide.pdf
Detailed online instructions that focus on establishing, modifying and administering SPSS databases.

IBM (2015): IBM SPSS Statistics 25 Documentation.
http://www-01.ibm.com/support/docview.wss?uid=swg27049428
Links to various SPSS methods modules with comprehensive overviews each, but often in a condensed form.
Regression Manual:
ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/25.0/en/client/Manuals/IBM_SPSS_Regression.pdf

Other Online Sources:

WolframMathWorld.
http://mathworld.wolfram.com/
A detailed online dictionary which provides answers and explanations regarding mathematical and statistical/probabilistic procedures.

Goodfellow, I., Bengio, Y. and Courville, A. (2016): Deep Learning. Cambridge, MA: MIT Press.
http://www.deeplearningbook.org
This book provides some mathematical foundations for those students interested in more complex mathematical/statistical procedures.

* * * * *

TOPICS: Spring Term 2019

1. Introduction to SPSS: Descriptives, frequencies and crosstabs

2. Probability distributions 1: How to measure and plot chance
(incl. combinatorics, probability concepts, rules for operations, set theory, random and non-random sampling)

3. Probability distributions 2: The binomial, poisson and exponential distributions

4. Statistical inference 1: Normal distribution and test design

5. Statistical inference 2: Parameter tests and related t/F distributions

6. Statistical inference 3: Chi-square test and analysis of variance

7. Linear regression analysis 1: Linear regression/correlation model and statistical tests

8. Linear regression analysis 2: Statistical tests, OLS/ML estimates and non-linear models
(incl. introduction logistics regression)

9. Multiple regression analysis 1: Basic regression/correlation model and tests

10. Multiple regression analysis 2: Model assumptions, dummy variables and multicollinearity
(incl. introduction logit regression)

11. Multiple regression analysis 3: Heteroscedasticity and autocorrelation

12. Multiple regression analysis 4: Specification problems and non-linear models

Readings by Topic: Spring Term 2019

Readings marked by an asterisk (*) represent core readings, which all students are required to read. The other suggested readings provide an alternative perspective on similar material.

1. Introduction to SPSS: Descriptives, frequencies and crosstabs

Task: Familiarization with the database, its organization and variables; conducting simple SPSS analyses

Anderson et al., Chs. 2-3

*SPSS:
Pallant, Chs. 6-8
Wagner, Chs. 1-5
Online SPSS instructions

2. Probability distributions 1: How to measure and plot chance

*Anderson et al., Chs. 4, 5.1-5.4 and 6.1
*Gujarati and Porter, App. A.2-A.5
Wooldridge, App. B.1-B.4

3. Probability distributions 2: The binomial, poisson and exponential distribution

*Anderson et al., Chs. 5.5-5.6 and 6.4
*Gujarati and Porter, App. A.6

4. Statistical inference 1: Normal distribution and test design

*Anderson et al., Chs. 6.2 and 7
*Gujarati and Porter, App. A.6-A.8
McNabb, Ch. 12
Wooldridge, App. C

Assignment 1: handed out in class

5. Statistical inference 2: Parameter tests and related t/F distribution

*Anderson et al., Chs. 9-11
*Gujarati and Porter, Chs. 5.1-5.8 and App. A.6-A.8
Wooldridge, App. B.5

*SPSS:
Pallant, Chs.16-17

Wagner, Ch. 6
Online SPSS instructions

6. Statistical inference 3: Chi-square test and analysis of variance

*Anderson et al., Chs. 12 and 13
*Gujarati and Porter, Chs. 5.7-5.12

*SPSS:
Pallant, Chs. 16 and 18
Wagner, Chs. 6 and 10
Online SPSS instructions

7. Linear regression analysis 1: Linear regression/correlation model and statistical tests

*Anderson et al., Ch. 14
*Gujarati and Porter, Chs. 2, 3 and parts of 4
McNabb, Ch. 14
Wooldridge, Chs. 2 and 3

*SPSS:
Pallant, Chs. 11 and 13
Wagner, Ch. 8
Online SPSS instructions

Assignment 2: handed out in class

8. Linear regression analysis 2: Statistical tests, OLS/ML estimates and non-linear models

Task: Review matrix algebra (*Gujarati and Porter, App. B; Wooldridge, App. D)

*Anderson et al., Ch. 15.9
*Gujarati and Porter, Chs. 4 and 6
Wooldridge, Chs. 4 and 5

*SPSS:
Pallant, Ch. 14
Wagner, Ch. 9
Online SPSS instructions

9. Multiple regression analysis 1: Basic regression/correlation model and tests

Task: Matrix approach to linear regression (*Gujarati and Porter, App. C.1-C.10; Wooldridge, App. E)

*Anderson et al., Ch. 15
*Gujarati and Porter, Chs. 7 and 8

*SPSS:
Pallant, Chs. 12-13
Wagner, Ch. 8
Online SPSS instructions

10. Multiple regression analysis 2: Model assumptions, dummy variables and multicollinearity

*Gujarati and Porter, Chs. 9, 10 and 15
Wooldridge, Ch. 7

Assignment 3: handed out in class

11. Multiple regression analysis 3: Heteroscedasticity and autocorrelation

*Gujarati and Porter, Chs. 11 and 12
Wooldridge, Ch. 8

12. Multiple regression analysis 4: Specification problems and non-linear models

*Gujarati and Porter, Chs. 13 and 14
Wooldridge, Ch. 9

*SPSS:
Pallant, Ch. 14
Wagner, Ch. 8
Online SPSS instructions

Take-Home Exam: handed out in class