Hertie School

Master of International Affairs/Master of Public Policy
Fall Semester 2020
Course Syllabus, Version 29.06.2020

GRAD-C5: Statistics I: Inference and Regression

Prof. Mark Kayser

1. General information

Class time	Group A - Monday 10-12 Group B - Monday 16-18 Group C - Tuesday 12-14 Group D - Tuesday 16-18
Course format	The Statistics I lecture will be taught onsite at the Hertie School and will be live streamed for those unable to be on campus. The lecture will also be recorded for later viewing. Small participatory labs will accompany the lecture and will be taught onsite at the Hertie School, or online via the platform Clickmeeting, depending upon your location. For those attending online labs, Clickmeeting allows for interactive, participatory style teaching. Both the live stream and the labs will cater for different time zones. Online office hours will be available.
Instructor	Prof. Mark Kayser
Instructor's office	2.54
Instructor's e-mail	kayser@hertie-school.org
Instructor's phone number	+49 30 25 92 19 - 307
Assistant	Name: Dayna Sadow Email: sadow@hertie-school.org Phone: +49 30 259 219 307 Room: 2.42.1
Instructor's Office Hours	Thursdays, 13.30-14.30 (to make an appointment please email Dayna Sadow).

Link to Module Handbook **MIA** and **MPP**
Link to **Study, Examination and Admission Rules**

Instructor Information:
Mark Kayser is Professor of Applied Methods and Comparative Politics at the Hertie School. His core research interests address the economic underpinnings of democracy with current projects centering on electoral accountability, economic news, and electoral competitiveness. Before joining the Hertie School faculty, he taught at the University of Rochester and he has held research fellowships at Nuffield College, Oxford and the Center for Advanced Studies in the Behavioral Sciences at Stanford. He has published in the top journals in his field and is the recipient of multiple research awards. He is the former senior editor for comparative political economy of the Oxford Research Encyclopedia and currently serves on several advisory and editorial boards.

2. Course Contents and Learning Objectives

Course contents:
This course offers an introduction to quantitative research methods for public policy. By the end of the semester, students should have a better acquaintance with quantitative methods conducive to empirical policy research and the ability to understand and critique the techniques employed by others. No prior knowledge of statistics is assumed and we will make every effort to address concepts both quantitatively and theoretically. Additionally, every student will attend a weekly lab session run by a teaching assistant in which concepts will be reinforced, where possible, through hands-on application using R. The goal of this course is not only to produce sophisticated consumers of quantitative research but to provide the necessary skills for students to conduct their own research. As such, we eschew the conventional stopping point for an introductory statistics course and spend several weeks on multiple regression. Almost no phenomena in the social world result from singular causality, so a useful applied course must train students in assessing multiple causality. We do so. Attendance in both lecture and lab is mandatory. I urge students to keep up with the readings since the information in this class is highly cumulative.

Labs

Curricular Affairs will initially allocate you to labs and then allow you to switch – by swapping with other students – in order to avoid conflicts with your schedules. If you would like to switch labs, please use the switching partner forum on Moodle to connect with your fellow students. Once you find a switching partner, please fill in the form "switching course request" on MyStudies. It is your responsibility to make sure that by switching courses you do not create time clashes with your other courses.

Teaching style:
This is a lecture course that focuses on instilling an intuitive understanding of fundamental concepts of statistics through the use of examples, visual representations and elementary mathematics.

Prerequisites:
This is an introductory statistics course that should make statistics accessible to all student willing to invest the necessary effort. No prior knowledge of statistics, experience with R or mathematics above the high school level is necessary.

3. Grading and Assignments

Composition of Final Grade:

Homework assignments:	Due:	Submit via Moodle	20%
Mid-term exam	TBA: week of 19.10.20	In Class	20%
Final exam	TBD: week of 14.12.20	In Class	25%
Final data analysis (FDA)	Deadline: 5pm on the Monday following finals week.	Submit by: Moodle & Hard copy	25%
Lab Participation			10%

Assignment Details

Homework assignments
Weekly problem sets are distributed on Moodle. Usually only one question will require human correcting so that the feedback is fast.

Final data analysis
The two exams cover concepts and applications in statistics and research design but do not cover the use or mastery of R. To provide students with an incentive and opportunity to develop their R skills, I will distribute a data set and questions toward the end of the semester for the students to analyze using R. This assignment, similar to a take-home exam, must be completed individually.

Participation Grade
The participation grade is based on the assumption that students take part not as passive consumers of knowledge but as active participants in the exchange, production, and critique of ideas—their own ideas and the ideas of others. Therefore, students should come to class not only having read and viewed the materials assigned for that day but also prepared to discuss the readings of the day and to contribute thoughtfully to the conversation. Participation is marked by its active nature, its consistency, and its quality.

Late submission of assignments:
For each day the final data analysis is turned in late, the grade will be reduced by 10% (e.g. submission two days after the deadline would result in 20% grade deduction). The weekly homework assignments cannot be turned in late because the answers are made public immediately after the submission deadline.

Attendance: Students are expected to be present and prepared for every class session. Active participation during lectures and seminar discussions is essential. If unavoidable circumstances arise which prevent attendance or preparation, the instructor should be advised by email with as much advance notice as possible. Please note that students cannot miss more than two out of 12 course sessions. For further information please consult the Examination Rules §10.

Academic Integrity: The Hertie School is committed to the standards of good academic and ethical conduct. Any violation of these standards shall be subject to disciplinary action. Plagiarism, deceitful actions as well as free-riding in group work are not tolerated. See Examination Rules §16.

Compensation for Disadvantages: If a student furnishes evidence that he or she is not able to take an examination as required in whole or in part due to disability or permanent illness, the Examination Committee may upon written request approve learning accommodation(s). In this respect, the submission of adequate certificates may be required. See Examination Rules §14.

Extenuating circumstances: An extension can be granted due to extenuating circumstances (i.e., for reasons like illness, personal loss or hardship, or caring duties). In such cases, please contact the course instructors and the Examination Office *in advance* of the deadline.

4. General Readings

Some light summer reading (optional):

David Spiegelhalter, The Art of Statistics: Learning from Data (2019 Pelican Books)

Required:

(M) Kenneth J. Meier, Jeffrey L. Brudney and John Bohte, 2014 (9th ed). Applied Statistics for Public and Nonprofit Administration. Thompson Wadsworth. Multiple earlier editions are also acceptable as the book changes very little between editions. Section and page numbers can change a bit between editions.

On library reserve for reference:

(F&W) J. Fox and S. Weisberg, An R Companion to Applied Regression, Third Edition, Sage, 2019.
(AF) Alan Agresti and Barbara Finlay. 2018. Statistical Methods for the Social Sciences. 5th edition. Prentice Hall.
(KW) Paul Kellstedt and Guy Whitten. 2018. The Fundamentals of Political Science Research. Cambridge University Press. 3rd edition.

5. Session Overview

Session	Session Date	Session Title
1	**07-08.09.2020**	Introduction: Motivation and Overview; Measurement: Theory, Validity, Reliability, Levels of Measurement
2	**14-15.09.2020**	Research Design: Experiments & Quasi-Experiments; forming hypotheses; Internal & External Validity; Regression to the mean; Out-of-sample extrapolation; Causality & Spuriousness
3	**21-22.09.2020**	Descriptive Statistics: Levels of Measurement, Measures of Central Tendency; Descriptive Statistics: Measures of Dispersion, Skew; Frequency Distributions; Central Limit Theorem
4	**28-29.09.2020**	Probability and introduction to inference. Z-distribution
5	**05-06.10.2020**	Sampling and Inference. The t-distribution, Confidence Intervals; Sample Proportions & Polling
6	**12-13.10.2020**	Inference Review; Hypothesis testing, testing the difference between two groups
Mid-term Exam Week: 19.10 - 23.10.2020 – no class		
7	**26-27.10.2020**	Categorical Variable Analysis: Cross-tabs, Chi-sq, PRE; Categorical Variable Analysis: Controlling for a third variable
8	**02-03.11.2020**	Regression: Correlation, Bivariate Regression; Goodness of fit & OLS Assumptions
9	**09-10.11.2020**	Multiple Regression; Dummy Regression & Influential Observations
10	**16-17.11.2020**	Multicollinearity, Heteroskedasticity, Correlated Disturbances & Diagnostics
11	**23-24.11.2020**	Model Specification, Interactions, Polynomial Regression

12	**30.11.-01.12.2020**	Overview and Review
Final Exam Week: 14.12 - 18.12.2020 – no class		

6. Course Sessions and Readings

All readings will be accessible on the Moodle course site before semester start. In the case that there is a change in readings, students will be notified by email.

Session 1: Introduction: Motivation and Overview; Measurement: Theory, Validity, Reliability, Levels of Measurement	
Required Readings	Meier 1-2

Session 2: Research Design: Experiments & Quasi-Experiments; forming hypotheses; Internal & External Validity; Regression to the mean; Out-of-sample extrapolation; Causality & Spuriousness	
Required Readings	Meier 3
Session 3: Descriptive Statistics: Levels of Measurement, Measures of Central Tendency; Descriptive Statistics: Measures of Dispersion, Skew; Frequency Distributions; Central Limit Theorem	
Required Readings	Meier 4-6

Session 4: Probability, the Normal Probability Distribution, Z-scores.	
Required Readings	Meier 7; skim 8 & 9

Session 5: Sampling and Inference. The t-distribution, Confidence Intervals; Sample Proportions & Polling	
Required Readings	Meier 10, 12

Session 6: Inference Review; Hypothesis testing, testing the difference between two groups	
Required Readings	Meier 11, 13

Mid-term Exam Week: 19 – 23.10.2020 – no class

Session 7: Categorical Variable Analysis: Cross-tabs, Chi-sq, PRE; Categorical Variable Analysis: Controlling for a third variable	
Required Readings	Meier 14-16

Session 8: Regression: Correlation, Bivariate Regression; Goodness of fit & OLS Assumptions	
Required Readings	Meier 17, 18; FW 4.1, 4.2.1, 4.3

Session 9: Multiple Regression; Dummy Regression & Influential Observations	
Required Readings	Meier 20;FW 4.2.2, 4.5, 4.7, 8.3

Session 10: Multicollinearity, Heteroskedasticity, Correlated Disturbances & Diagnostics	
Required Readings	Meier 20 (again); FW 8.1, 8.2, 8.5, 8.8

Session 11: Model Specification, Interactions, Polynomial Regression	
Required Readings	Meier 20 (again); FW 4.8, 4.6, 4.4

Session 12: Overview and Review

Final Exam Week: 14 - 18.12.2020 – no class

Final Data Analysis: Due at 5pm on the Monday following the exam week