Generalized Linear Models
SURV 699J
2 credits/4 ECTS
Spring 2019

Instructor
Prof. Dr. Thomas Gautschi, gautschi@sowi.uni-mannheim.de

Video lecture by Prof. Dr. Thomas Gautschi

Short Course Description
The main focus of this course lies on the introduction to statistical models and estimators beyond linear regression useful to social and economic scientists. It provides an overview of generalized linear models (GLM) that encompass non-normal response distributions to model functions of the mean. GLMs thus relate the expected mean $E(Y)$ of the dependent variable to the predictor variables via a specific link function. This link function permits the expected mean to be non-linearly related to the predictor variables. Examples for GLMs are the logistic regression, regressions for ordinal data, or regression models for count data. GLMs are generally estimated by use of maximum likelihood estimation. The course thus not only introduces GLMs but starts with an introduction to the principle of maximum likelihood estimation. A good understanding of the classical linear regression model is a prerequisite and required for the course.

Course and Learning Objectives
By the end of the course, students will...
- Understand how to appropriately translate research questions into statistical models
- Be able to apply statistical models appropriate for non-linear problems
- Estimate regression parameters using the maximum likelihood principle
- Perform hypothesis tests for regression models using the maximum likelihood principle
- Be able to identify limitations of non-linear regression models
- Be able to identify violations of the respective regression assumptions of the discussed GLMs
Prerequisites
A sound understanding of linear regression models (OLS) is required. Knowledge in linear algebra and calculus is useful.

Class Structure and Course Concept:
This is an online course using a flipped classroom design. It covers the same material and content as an on-site course but runs differently. In this course, you are responsible for watching video recorded lectures and reading the required literature for each unit and then “attending” mandatory weekly one-hour online meetings where students have the chance to discuss the materials from a unit with the instructor. Just like in an on-site course, homework will be assigned and graded and there will be a final exam at the end of the course.

Although this is an online course where students have more freedom in when they engage with the course materials, students are expected to spend the same amount of time overall on all activities in the course – including preparatory activities (readings, studying), in-class-activities (watching videos, participating in online meetings), and follow-up activities (working on assignments and exams) – as in an on-site course. As a rule of thumb, for each credit offered by a course, students can expect to spend one hour per week on in-class activities and three hours per week on out-of-class activities over the span of a full 12-week term. This is a 2-credit course that runs for 8 weeks. Hence, the total average workload is about 12 hours per week.

Mandatory Weekly Online Meetings:
Thursday, 2:00 PM EST/8:00 PM CET, starting March 7

Meetings will be held online through Zoom. Follow the link to the meeting sessions on the course website on https://www.elms.umd.edu/. If video participation via Internet is not possible, arrangements can be made for students to dial in and join the meetings via telephone.

In preparation for the weekly online meetings, students are expected to watch the lecture videos and read the assigned literature before the start of the meeting. In addition, students are encouraged to post questions about the materials covered in the videos and readings of the week in the forum before the meetings (deadline for posting questions is Thursday, 4:00 AM EST/10:00 AM CET).
Students have the opportunity to use the Zoom meeting room set up for this course to connect with peers outside the scheduled weekly online meetings (e.g., for study groups). Students are encouraged to post the times that they will be using the room to the course website forum to avoid scheduling conflicts. Students are not required to use Zoom and can of course use other online meeting platforms such as Google Hangout or Skype.

**Grading**

Grading will be based on:

- 7 homework assignments (49% of grade total, 7% each)
- Participation in online meetings and submission of questions demonstrating understanding of readings (10% of grade)
- Final Exam (41% of grade)

Students must get a 70% or higher in order to pass the class.

Dates of when assignment will be due are indicated in the syllabus. Extensions will be granted sparingly and are at the instructor's discretion.

**Technical Equipment Needs**

The learning experience in this course will mainly rely on the online interaction between students and the instructor during the weekly online meetings. Therefore we encourage all students in this course to use a web camera and a headset. Decent quality headsets and web cams are available for less than $20 each. We ask students to refrain from using built-in web cams and speakers on their desktops or laptops. We know from our experience in previous online courses that this will reduce the quality of video and audio transmission and therefore will decrease the overall learning experience for all students in the course. In addition, we suggest that students use a wire connection (LAN), if available, when connecting to the online meetings. Wireless connections (WLAN) are usually less stable and might be dropped.

**Long Course Description**

The main focus of this course lies on the introduction to statistical models and estimators beyond linear regression useful to social and economic scientists.
Although very useful, the general linear model (linear regression) is not appropriate if the range of the dependent variable $Y$ is restricted (e.g., binary, ordinal, count) and/or the variance of $Y$ depends on the mean of $Y$. Generalized linear models extend the general linear model to address both of these shortcomings.

The course provides an overview of generalized linear models (GLM) that encompass non-normal response distributions to model functions of the mean of $Y$. GLMs thus relate the expected mean $E(Y)$ of the dependent variable to the predictor variables via a specific link function. This link function is chosen such that it matches the data generating process of the dependent variable $Y$, therefore permitting the expected mean $E(Y)$ to be non-linearly related to the predictor variables. Examples for GLMs are the logistic regression, regressions for ordinal data, or regression models for count data. GLMs are generally estimated by use of maximum likelihood estimation. The course thus not only introduces GLMs but starts with an introduction to the principle of maximum likelihood estimation. A good understanding of the classical linear regression model is a prerequisite and required for the course.

The first two units are dedicated to an introduction to maximum likelihood estimation while the rest of the units will then discuss generalized linear models (GLM) for binary choice decisions (Logit, Probit), ordinal dependent variables, and count data (Poisson, Negative Binomial).

All units will be accompanied by homework assignments to repeat and practice the topics from the units. Any statistic program can be used to solve the homework assignments. Solutions provided by the instructor will use the statistical packages Stata and R.

**Readings**

Required readings will be made available on the course website:

https://www.elms.umd.edu/


**Recommended readings:**


**Academic Conduct**

Clear definitions of the forms of academic misconduct, including cheating and plagiarism, as well as information about disciplinary sanctions for academic misconduct may be found at

[https://www.president.umd.edu/sites/president.umd.edu/files/documents/policies/III-100A.pdf](https://www.president.umd.edu/sites/president.umd.edu/files/documents/policies/III-100A.pdf) (University of Maryland) and


Knowledge of these rules is the responsibility of the student and ignorance of them does not excuse misconduct. The student is expected to be familiar with these guidelines before submitting any written work or taking any exams in this course. Lack of familiarity with these rules in no way constitutes an excuse for acts of misconduct. Charges of plagiarism and other forms of academic misconduct will be dealt with very seriously and may result in oral or written reprimands, a lower or failing grade on the assignment, a lower or failing grade for the course, suspension, and/or, in some cases, expulsion from the university.

**Accommodations for Students with Disabilities**

In order to receive services, students at the University of Maryland must contact the Disability Support Services (DSS) office to register in person for services. Please call
the office to set up an appointment to register with a DSS counselor. Contact the DSS office at 301.314.7682; http://www.counseling.umd.edu/DSS/.

Students at the University of Mannheim should contact the Commissioner and Counsellor for Disabled Students and Students with Chronic Illnesses at http://www.uni-mannheim.de/studienbueros/english/counselling/disabled_persons_and_persons_with_chronic_illnesses/.

Course Evaluation
In an effort to improve the learning experience for students in our online courses, students will be invited to participate in an online course evaluation at the end of the course (in addition to the standard university evaluation survey). Participation is entirely voluntary and highly appreciated.

Class Schedule
Please note that assignments and dates are subject to change. Information (e.g., articles and assignments) posted to the course website supersedes the information noted here.

Daylight saving time begins in the US on March 10, 2019 and clocks are turned forward 1 hour. Daylight saving time begins in Europe on March 31, 2019. Therefore, look carefully at the times of meetings and deadlines!

Week 1: Maximum Likelihood Estimation
Estimator and Variance
Video lecture: available Thursday, February 28, 2019
Online meeting: Thursday, March 7, 2019, 2:00 PM EST/8:00PM CET
Homework assignment 1: due Monday, March 11, 2019, 3:00 PM EDT/8:00 PM CET
Required readings:
Verbeek (2008) Ch. 6.1

Week 2: Maximum Likelihood Estimation
Specification Tests, Model Fit, and Numerical Optimization
Video lecture: available Thursday, March 7
Online meeting: Thursday, March 14, 2019, 3:00 PM EDT/8:00 PM CET
Homework assignment 2: due Monday, March 18, 2019, 3:00 PM EDT/8:00 PM CET

Required readings:
Verbeek (2008) Ch. 6.2
Greene (2008) Ch. 16.6.1-16.6.3

Week 3: Binary Choice Models
LPM, Binary Choices, Logit and Probit Models
Video lecture: available Thursday, March 14
Online meeting: Thursday, March 21, 2019, 3:00 PM EDT/8:00 PM CET
Homework assignment 3: due Monday, March 25, 2019, 3:00 PM EDT/8:00 PM CET

Required readings:
Verbeek (2008) Ch 7.1
Wooldridge (2008) Ch. 17.1

Week 4: Binary Choice Models
Coefficient Interpretation, Statistical Inference, and Goodness-of-Fit
Video lecture: available Thursday, March 21
Online meeting: Thursday, March 28, 2019, 3:00 PM EDT/8:00 PM CET
Homework assignment 4: due Monday, April, 2019, 2:00 PM EDT/8:00 PM CEST

Required readings:
Verbeek (2008) Ch. 7.1
Long (1997) Ch. 3.7-3.8

Week 5: Models for Ordinal Data
Ordered Outcomes, Modelling Strategy, Identification, and Estimation
Video lecture: available Thursday, March 28
Online meeting: Thursday, April 4, 2019, 2:00 PM EDT/8:00 PM CEST
Homework assignment 5: due Monday, April 8, 2019, 2:00 PM EDT/8:00 PM CEST

Required readings:
- Verbeek (2008) Ch. 7.2.1-7.2.3
- Wooldridge (2002) Ch. 15.10
- Long (1997) Ch. 5.1-5.2

Week 6: Models for Ordinal Data
*Inference, Threshold and Coefficient Interpretation, Outcome Probabilities*

Video lecture: available Thursday, April 4

Online meeting: Thursday, April 11, 2019, 2:00 PM EDT/8:00 PM CEST

Homework assignment 6: due Monday, April 15, 2019, 2:00 PM EST/8:00 PM CEST

Required readings:
- Verbeek (2008) Ch. 7.2.1-7.2.3
- Long (1997) Ch. 5

Week 7: Models for Count Data
*Poisson Distribution, Generalization, and Estimator*

Video lecture: available Thursday, April 11

Online meeting: Thursday, April 18, 2019, 2:00 PM EDT/8:00 PM CEST

Homework assignment 7: due Monday, April 22, 2019, 2:00 PM EDT/8:00 PM CEST

Required readings:
- Verbeek (2008) Ch. 7.3;
- Wooldridge (2008) Ch. 17.3
- Cameron and Trivedi (1998) Ch. 3.1, 3.2.1, 3.5.1

Week 8: Models for Count Data
*Variance, Overdispersion, and Negative Binomial*

Video lecture: available Thursday, April 18

Online meeting: Thursday, April 25, 2019, 2:00 PM EDT/8:00 PM CEST
Required readings:
Verbeek (2008) Ch. 7.3

Recommended readings:
Cameron and Trivedi (1998) Ch. 3.2.2, 3.3.1, 3.3.3, 3.4, 4.7

Final Exam
Available: April 19, 2019
Due: May 8, 2019, 2:00 PM EDT/8:00 PM CEST

Note: Student access to the course website will be revoked two weeks after the final exam.