UNIVERSITY OF MANNHEIM
School of Social Sciences

Advanced Quantitative Methods: Introduction

Thomas Gschwend | Lisa-Marie Müller | Domantas Undzėnas

Week 1 - 11 February 2026

Welcome!

Who takes this course?

- M.A. & GESS students interested in learning tools to develop statistical models
- Methods sequence in the political science graduate program:
  1. M.A.: Quantitative Methods (last semester), **required**
  2. M.A.: Advanced Quantitative Methods (this semester), optional
  3. CDSS: additional advanced courses focusing on specific techniques

What's it about?

- Foundation of statistical inference: Using the facts you have to learn about the facts you don't have
- Focus on maximum likelihood theory of inference
- Programming and statistical simulation as practical tools
- Many specific methods & robustness tests
- Learn how to fine-tune existing methods or develop new ones

General Requirements

- Learning in this class is a collective experience. You need to be prepared. Everyone is counting on you!
- Weekly readings: Read slower, take notes. Read by keeping a running list of symbols, equations, and their meaning. Skip no equation! Work in groups to sort out remaining issues.
- *Pre*pare and *post*pare lecture notes
  - Interrupt me as often as necessary!
  - Assume you are the smartest person in the class, and you, eventually, will be.
- You will fail this class if you are not signed up for one of the Labs (no kidding!)
- Six homework assignments: Work in groups! Essential learning experience
- Actively participate in the Lab session.
- Final draft paper (coauthored) + replication material (full paper for more than two co-authors). Paper should be potentially publishable. Consult with me early on about the framing of your contribution, and how to construct a winning argument.

Final draft paper: How to find a topic?

- Hint: start with replicating an existing article.
- Do not replicate the entire article. No replication report. Instead develop your own argument!
- Replicate important aspect of article. Why is it important? Not because of the authors say so but because *you* say so!
- You have to make a case that this is important. How do you know? We are writing for an audience. You have to convince others that this is important.
- Even if authors say that the paper is about X you can convince us that we should instead think about C because it is a more interesting question.
- How to cast an article (big picture) and do all the little details of squaring the terms to come up with the likelihood? Don't loose sight of either side.
- Write down your model!
- Don't trust that the model assumptions are true. Test them!

Course Strategy

- We could teach you the latest and greatest methods, but by the time you graduate ...
  - ... *they* will be old
  - ... or *you* will be old
- We could teach you several years of calculus, linear algebra, mathematical statistics, probability theory, and then start with data analysis. This works great, but not if you wanna be a social scientist.
- Instead, we teach you the *fundamentals*, the underlying *theory of inference*, from which most statistical models are developed. Then we do examples in great detail. Math gets introduced in great depth, but only when needed.
- Read our syllabus carefully.

What is Maximum Likelihood? - Basic Intuition

- Suppose: $Y \sim N(\mu, \sigma^2)$

What is Maximum Likelihood? - Basic Intuition

- Suppose: $Y \sim N(\mu, \sigma^2)$
  - Thus, we *assume* a normal distribution with two parameters:

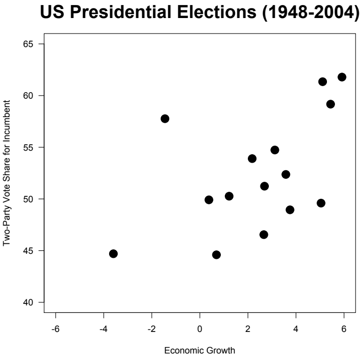
$$\begin{aligned} E[Y] &= \mu \\ Var(Y) &= \sigma^2 \end{aligned}$$

  - We have some observations on Y and we want to estimate μ and σ^2
- Suppose we have made the following observations (say, government approval):

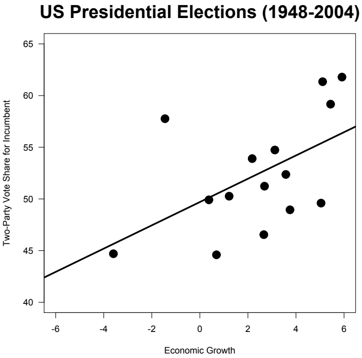
$$Y = \{54, 53, 49, 61, 58\}$$

- Intuitively we wonder about the likelihood of getting these data points if we assume a normal distribution ...
  - ... with $\mu = 100$?
  - ... with $\mu = 55$?
- The basic idea behind *maximum likelihood* is to find the estimate for the parameter values of our chosen (assumed) distribution that *maximizes* the *likelihood* of observing the data we have.

What do we see here?



How to fit a line to a scatterplot?



- We need a rule (ordinary least squares, least absolute values of residuals, ect.)
- We need criteria (unbiasedness, efficiency, ect.)
- We need a theory of inference for a purpose (e.g., causal estimation, prediction)

What is this?

From General Models to Statistical Models: Some Definitions

- Explanatory variables (aka “covariates”, “independent” or “exogenous” variables) are combined into a *design matrix* X
  $X = (1, x_1, \cdots, x_j, \cdots, x_k)$ for $x_j = (x_{1j} \cdots x_{nj})'$. X is $n \times (k+1)$
  - n: Number of observations
  - $(k+1)$: Number of parameters (No. of explanatory variables + 1)
- Dependent (or “outcome”) variable: Y is $n \times 1$
- Y_i is a random variable (before we can observe it)
- y_i is a number (after we can observe it)

Linear Regression Notation

- Standard Version:

$$
\begin{aligned}
Y_i &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \epsilon_i \\
&= \beta_0 + \sum_{j=1}^{k} \beta_j x_{ji} + \epsilon_i \\
&= (1,\ x_{i1}, \cdots, x_{ik}) \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} + \epsilon_i \\
&= X_i \beta + \epsilon_i \qquad\qquad (\mathit{systematic} + \mathit{stochastic}) \\
\epsilon_i &\sim f_N(e_i | 0, \sigma^2)
\end{aligned}
$$

Linear Regression Notation

- Alternative Version:

$$
\begin{aligned}
Y_i &\sim f_N(y_i|\mu_i, \sigma^2) && \text{stochastic} \\
\mu_i &= X_i\beta && \text{systematic}
\end{aligned}
$$

Where is the Uncertainty?

Recall that we can generalize that and write any statistical model as

$$
\begin{aligned}
Y_i &\sim f(y_i|\theta_i, \alpha) \quad \text{stochastic} \\
\theta_i &= g(X_i, \beta) \quad \text{systematic}
\end{aligned}
$$

1. Estimation Uncertainty: Uncertainty about what the true parameters β and α of the model are. Think of it as caused by small samples. Vanishes if N gets larger.
2. Fundamental Uncertainty: Represented by stochastic component of the model. Exists no matter what (even if model is correct and we would have infinite many observations) because of inherent randomness of the world.

OLS Model

Some Definitions and Notation

- Suppose we have the following multiple regression model with $k+1$ parameters (but k independent variables) and $i = 1, \cdots, n$ observations,

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \epsilon_i$$

- For every observation i the relationship between values of the dependent variable y_i and the corresponding values on the covariates $x_{i1}, x_{i2}, \cdots, x_{ik}$ can be written as:

$$\begin{aligned}
y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_k x_{1k} + \epsilon_1 \\
y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \beta_k x_{2k} + \epsilon_2 \\
y_3 &= \beta_0 + \beta_1 x_{31} + \beta_2 x_{32} + \cdots + \beta_k x_{3k} + \epsilon_3 \\
\vdots &= \vdots \\
y_i &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \epsilon_i \\
\vdots &= \vdots \\
y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_k x_{nk} + \epsilon_n
\end{aligned}$$

Some Definitions and Notation

- The system of n equations can be elegantly condensed as follows:

$$\underset{[n\times 1]}{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{pmatrix}} = \underset{[n\times(k+1)]}{\begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & x_{i1} & x_{i2} & \cdots & x_{ik} \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix}} \underset{[(k+1)\times 1]}{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_k \end{pmatrix}} + \underset{[n\times 1]}{\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_i \\ \vdots \\ \epsilon_n \end{pmatrix}}$$

- This can be rewritten as

$$y = X\beta + \epsilon$$

- The model has a systematic component ($X\beta$) and a stochastic component (ϵ)
- We would like to obtain estimates of the population parameters (β), which we denote as $\hat{\beta}$