

Course goal, motivation and course schedule

Course goal (what this course is and is not)

This course is intended to help you develop your statistical literacy—the ability to comprehend and critically evaluate the results of statistical data analyses—especially in the social and behavioral sciences. While you will learn how to apply basic methods for statistical data analysis, this course focuses on preparing you to be a critical and knowledgeable consumer of statistical data analyses, not a skilled producer. Students who intend to go into fields where data-analytic skills, including computational statistics, are necessary are encouraged to consider STAT 311 instead.

Course motivation

As a discipline, statistics focuses on describing and modeling both variability and uncertainty in our world. It includes a wide assortment of theories and methods for summarizing variability in all kinds of data, as well as exploring relationships that exist between variables. In the social and behavioral sciences, these include:

- crime rates by jurisdiction or geopolitical unit (neighborhood, city, state, country);
- incidence and prevalence of particular diseases by community;
- life expectancy and fertility rates by county, state, or country;
- age, sex, gender, ethnic, or SES composition by population;
- unemployment rates over time;
- election victories by incumbency of candidates;
- number of motorists on the roadway at different hours of the day;
- birth rates by date and day of the week;
- educational outcomes based on different levels of access to resources, class size, or attendance at different schools;
- carbon emission levels by country relative to population size or GDP.

Both as voters and as members of various communities, organizations, and institutions, we need to make important decisions based on our understanding of these kinds of information. Consequently, it is important to understand the objectives, built-in assumptions, results, and limitations of the statistical methods used by social scientists to collect and explore such data. It is also important to understand the pitfalls that arise when these methods are applied incorrectly. By understanding these things, you will be better prepared to be a more critical consumer of statistical analyses that you will encounter in popular media and in professional and academic publications.

Course objectives

By the end of this course, you should be able to

- Distinguish between nominal, ordinal, and different kinds of numerical variables;
- Distinguish between sample statistics and population parameters and understand the relationship between these two concepts;
- Identify the strengths and limitations of different strategies researchers use to collect data, as well as how these relate to research questions/goals and statistical inference;
- Evaluate the ability of different statistics to summarize different kinds of data, both numerically and graphically;
- Use the rules of probability theory to model population variability and to explore relationships between variables in different ways;
- Identify parametric probability distributions that are used as models of different data generating processes;
- Understand how statisticians use sampling distributions both to estimate unknown population parameters and to conduct statistical hypothesis tests about such parameters.

Tentative Course Schedule

Note: this schedule is subject to changes based on the class's pace.

Unit 1: Basics of data, variables, research design, and data description (Chapters 1 and 2, OpenIntro Statistics)

Week 1 - Week 3

Unit 2: Understanding data generating processes and uncertainty using probability theory (Chapters 3 and 4, OpenIntro Statistics)

Week 4 - Week 6

Unit 3: Data-based statistical inference for categorical data (Chapters 5-6, OpenIntro Statistics)

Week 7 - Week 9

Unit 4: Data-based statistical inference for numerical data (Chapters 7-8, OpenIntro Statistics)

Week 10 - Week 11

Date			Topic	Chapter covered
27-Sep	Lecture 1	Week 1	Syllabus, class structure, get students to sign up on PollEv	
28-Sep	QS1	Week 1	An overview of big themes in Statistics	
29-Sep	Lecture 2	Week 1	Motivating statistics, data sets, typologies of variables	Unit 1 - Ch. 1
30-Sep				
1-Oct				
2-Oct	Lecture 3	Week 2	Data collection, representative samples, and observational studies	Unit 1 - Ch. 1
3-Oct	QS2	Week 2	WEIRD paper	
4-Oct	Lecture 4	Week 2	Causation: observational studies and more on experiments	Unit 1 - Ch. 1
5-Oct	QS3	Week 2	Banta-Green and Field, 2011	
6-Oct	Lecture 5	Week 2	Descriptive Statistics: Numerical and visual summaries of one variable	Unit 1 - Ch. 2
7-Oct				
8-Oct				
9-Oct	Lecture 6	Week 3	Descriptive Statistics: Numerical and visual summaries of one variable	Unit 1 - Ch. 2
10-Oct	QS4	Week 3	Cleaning, summarizing, and interpreting numerical data	
11-Oct	Lecture 7	Week 3	Descriptive statistics: visual summaries to examine relationships between variables	Unit 1 - Ch. 2
12-Oct	QS5	Week 3	Research ethics in the social sciences	
13-Oct	Lecture 8	Week 3	Study Session - Midterm 1	Unit 1: Ch. 1 & Ch. 2
14-Oct				
15-Oct				
16-Oct	Lecture 9	Week 4	Midterm 1	Unit 1: Ch. 1 & Ch. 2

17-Oct	QS6	Week 4	Midterm 1 solutions + Probability experiment with dice	
18-Oct	Lecture 10	Week 4	Probability theory: introduction, random processes	Unit 2 - Ch. 3
19-Oct	QS7	Week 4	Probability exercises	
20-Oct	Lecture 11	Week 4	Probability theory: rules 1	Unit 2 - Ch. 3
21-Oct				
22-Oct				
23-Oct	Lecture 12	Week 5	Probability theory: rules 2	Unit 2 - Ch. 3
24-Oct	QS8	Week 5	Bayes Theorem	
25-Oct	Lecture 13	Week 5	Probability theory: rules 2	Unit 2 - Ch. 3
26-Oct	QS9	Week 5	Bayes Theorem	
27-Oct	Lecture 14	Week 5	Normal distribution	Unit 2 - Ch. 4
28-Oct				
29-Oct				
30-Oct	Lecture 15	Week 6	Normal distribution	Unit 2 - Ch. 4
31-Oct	QS10	Week 6	Exercises on the Normal distribution	
1-Nov	Lecture 16	Week 6	Normal distribution	Unit 2 - Ch. 4
2-Nov	QS11	Week 6	Exercises on the Normal distribution	
3-Nov	Lecture 17	Week 6	Discrete distributions	Unit 2 - Ch. 4
4-Nov				
5-Nov				
6-Nov	Lecture 18	Week 7	Normal approximation to the binomial distribution	Unit 3 - Ch. 5
7-Nov	QS12	Week 7	Study Session - Midterm 2	
8-Nov	Lecture 19	Week 7	Midterm 2	Unit 2 - Ch. 3 & Ch. 4
9-Nov	QS13	Week 7	Developing intuitions regarding sampling distributions	
10-Nov	UNIVERSITY HOLIDAY: VETERAN'S DAY			

11-Nov				
12-Nov				
13-Nov	Lecture 20	Week 8	Sampling distributions, estimating a population proportion	Unit 3 - Ch. 5
14-Nov	QS14	Week 8	Midterm 2 solutions	
15-Nov	Lecture 21	Week 8	Confidence intervals for a population proportion	Unit 3 - Ch. 5
16-Nov	QS15	Week 8	Practice on Confidence intervals	
17-Nov	Lecture 22	Week 8	Hypothesis testing	Unit 3 - Ch. 6
18-Nov				
19-Nov				
20-Nov	Lecture 23	Week 9	P-values for Inferences on categorical data	Unit 3 - Ch. 6
21-Nov	QS16	Week 9	Practice on hypothesis testing and one-sided hypotheses	
22-Nov	Lecture 24	Week 9	Understanding errors and p-values in hypothesis testing	Unit 4 - Ch. 7
23-Nov	UNIVERSITY HOLIDAY: THANKSGIVING			
24-Nov	UNIVERSITY HOLIDAY: THANKSGIVING			
25-Nov				
26-Nov				
27-Nov	Lecture 25	Week 10	Chi Square test for independence	Unit 4 - Ch. 7
28-Nov	QS17	Week 10	Study Session - Midterm 3	Unit 3 - Ch. 5 & Ch.6
29-Nov	Lecture 26	Week 10	Midterm 3	Unit 3 - Ch. 5 & Ch.6
30-Nov	QS18	Week 10	Midterm 3 solutions	
1-Dec	Lecture 27	Week 10	One-sample means	Unit 4 - Ch. 7
2-Dec				

3-Dec

4-Dec	Lecture 28	Week 11	Difference of two means (paired and independent data)	Unit 4 - Ch. 7
5-Dec	QS19	Week 11	Review of Inference on one or two means	
6-Dec	Lecture 29	Week 11	Comparing means with ANOVA	Unit 4 - Ch. 7
7-Dec	QS20	Week 11	Study session - Final Exam	Unit 1-4, Ch. 1- Ch.7
8-Dec	Lecture 30	Week 11	Intro to linear regression	Unit 4 - Ch. 8