

I306 Statistics for Informatics

Spring 2026 Schedule

2026-03-02

Course Schedule

This schedule outlines the topics, readings, and activities for each class session. Readings and video lectures should be completed *before* class.

Week 1 (Jan 13-15)

13 Jan

Learning Goals

- Define probability in frequentist terms as the proportion of times an event will occur if repeated.
- Define statistics as a mathematical science pertaining to the collection, analysis, interpretation, and presentation of data.
- Describe inference as the fundamental challenge of making justified claims from quantitative data.

Plan

- [20 min] Activity: Guessing ages
- [20 min] Activity: Estimating a big number
- [20 min] Syllabus and schedule overview
- [10 min] Definition of probability and statistics
- [20 min] Q&A and discussion

15 Jan

Learning Goals

- Make sure everyone has Setup R and Rstudio.
- Get started programming in R.
- Define and learn synonyms of *case*, *observational unit*, *variable*, *explanatory* or *response* variables.
- Recognize common *data types* of variables.

Homework Assignment

- Complete the [setup checklist](#).

Reading Assignment

- Read Diez, Çetinkaya-Rundel, and Barr (2019) Preface and Chapter 1, Sections 1.1-1.2 (Case study: stents; Data basics).

Video Lectures

- [200 years of history through health & wealth](#)
- [Using statistics to treat chronic illnesses](#)
- [Case Study: Using Stents to Prevent Strokes](#)
- [Data Basics: Observations, Variable, and Data Matrices](#)
- [Data Collection Principles](#)

Optional Lab

- [Introduction to R and Rstudio](#)

Plan

- [30 min] Troubleshoot R and Rstudio installation
- [45 min] AppliedStatsInteractive tutorials 00-03

If you've got Rstudio working, help somebody who hasn't got it working yet. Once nobody needs help, *find a partner*. You and your partner should work through the first four tutorials in AppliedStatsInteractive.

To load a tutorial run `learnr::run_tutorial("NOTEBOOK_NAME", package = "AppliedStatsInteractive")` in RStudio.

The tutorials for this session have `NOTEBOOK_NAME` equal to:

- `00_StartHere`

- 01_IntroToData
- 02_IntroToR
- 03_DescriptiveNumCat

As you finish each tutorial, click the Submit button and follow the instructions to generate a hash and submit it via Canvas. If you don't finish all the tutorials during class, you can submit them before the assignment in Canvas closes (typically before the next class).

Week 2 (Jan 20-22)

20 Jan

Learning Goals

- Understand the difference between a *population* and a *sample*.
- Recognize strategies for taking useful samples and common sampling pitfalls.
- Know the difference between *causation* and *association*.
- Learn about *experiments* and how randomization lets make causal inferences.

Reading Assignment * Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 1, Sections 1.3-1.4 (Sampling principles and strategies; Experiments).

Video Lectures

- [Sampling principles and strategies](#)
- [Experiments](#)

Plan

- [25 min] Activity 1: An experiment that looks like a survey.
- [25 min] Activity 2: How large is your family?
- [25 min] Activity 3: Spot the Flaw.

22 Jan

Learning Goals

- Understand fundamental statistics for summarizing numerical data: *mean, variance, standard deviation, median, quartiles*.
- Use common visualization techniques for numerical data.
- Critique misleading or poorly designed visualizations.

Reading Assignment * Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 2, Section 2.1 (Examining numerical data)

Video Lecture

- [Examining Numerical Data](#)

Plan

- [30 min] Activity: Critiquing charts from newspapers
 - [45 min] Activity: Social Media Descriptive Statistics
-

Week 3 (Jan 27-29)

Jan 27 class cancelled due to snow day.

29 Jan

Learning Goals

- Understand common visualization techniques for categorical data.
- Define *statistical dependence* and *independence*.
- Use statistical independence to draw conclusions from data.
- Learn how to use *randomization inference* to test a hypothesis.
- Use *bootstrapping* to quantify uncertainty.

Reading Assignment * Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 2, Sections 2.2-2.3 (Considering categorical data; Case study: malaria vaccine).

Video Lectures

- [Considering categorical data](#)
- [Case study: Gender discrimination](#)

Plan

- [25 min] Activity 1:
- [25 min] Activity 2:
- [20 min] Mini-lecture on simulation
- [15 min] Worksheet: Categorical Data & Independence

Homework (due Feb 3)

Complete these AppliedStatsInteractive tutorials and submit your completion hashes via Canvas:

```
learnr::run_tutorial("5_DiscreteDistributions", package = "AppliedStatsInteractive")
learnr::run_tutorial("6_NormalDistribution", package = "AppliedStatsInteractive")
```

Week 4 (Feb 3-5)

3 Feb

Learning Goals

- Understand the definition of probability an outcome *as the proportion of times the outcome would occur if we observed the random process an infinite number of times.*
- Use fundamental *set theory* concepts to define outcomes.
- Learn the *general addition rule* how to calculate the probability of joint and disjoint outcomes.
- Define *probability distributions* over a set of disjoint outcomes.
- Practice *randomization inference* to detect if a dice is fair.

Reading Assignment * Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 3, Section 3.1 (Defining probability).

Video Lecture

- [Probability Intro](#)

Plan

- [30 min] Quiz 1: Data Types & Visualization (Chapters 1-2)
- [40 min] Activity: Possibly unfair dice
- [15 min] Activity: Beano introduction (complete as homework)

Homework (due Feb 5)

Complete this AppliedStatsInteractive tutorial and submit your completion hash via Canvas:

```
learnr::run_tutorial("7_DiscreteDistributionsLab", package = "AppliedStatsInteractive")
```

5 Feb

Learning Goals

- Learn to calculate *conditional probability*, the *probability an outcome given a condition*.
- Define *marginal probability* (probability of a single variable) and *joint probability* (probability of an outcome for more than 1 variable).
- Convert conditional probabilities to and from marginal and joint probabilities using the *general multiplication rule* and *bayes theorem*.
- Use *tree diagrams* to organize the calculation of conditional probabilities.

Reading Assignment * Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 3, Section 3.2 (Conditional probability).

Video Lectures

- [Bayes theorem, the geometry of changing beliefs \(3Blue1Brown\)](#)
- [Probability trees](#)

Plan

- [40 min] Activity: Lie detectors (with tree diagrams)
- [35 min] Activity: Monty Hall (with tree diagrams)

! Milestone 1: Dataset Selection & EDA

Due Friday, Feb 6 at 11:59 PM

Week 5 (Feb 10-12)

10 Feb

Learning Goals

- Understand the structure and properties of *discrete probability distributions*.
- Apply the *binomial distribution* to model repeated trials with two outcomes.
- Calculate probabilities and expected values for binomial random variables.
- Distinguish between scenarios best modeled by *binomial*, *geometric*, or other discrete distributions.
- Recognize how discrete distributions form the foundation before moving to continuous probability distributions.

Reading Assignment

- Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 4, Sections 4.2-4.5 (Geometric, Binomial, Negative Binomial, and Poisson distributions).

Video Lectures

- [OpenIntro: Binomial distribution](#)
- [3Blue1Brown: Binomial distribution](#)
- [JBStatistics: An Introduction to the Poisson Distribution](#)
- [Khan Academy: Geometric distribution](#)

Plan - [30 min] Review quiz - [20 min] Bayes theorem tutorial (Example 3.15 and 3.21) on whiteboard - [40 min] [Discrete distributions coin flip activity](#): hands-on experiments demonstrating binomial, geometric, and Poisson distributions

12 Feb

Learning Goals

- Understand the properties of the *normal distribution* including symmetry, the empirical rule (68-95-99.7), and how mean and standard deviation determine shape.
- Use *Z-scores* to standardize values and compare observations across different normal distributions.
- Calculate probabilities and percentiles for normally distributed variables using `pnorm()` and `qnorm()` in R.
- Assess whether data are approximately normally distributed using graphical tools and diagnostics.
- Understand why the normal distribution is critical for statistical inference.

Reading Assignment

- Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 4, Section 4.1 (Normal distribution).

Video Lectures

- [OpenIntro Statistics: Normal Distribution](#)

Plan

- [15 min] Opening activity: Manual dice rolling in groups (build intuition for sampling)
- [55 min] [AppliedStatsInteractive 6_NormalDistribution \(Parts A, B, and C\)](#) - Interactive tutorial covering normal distribution properties, Z-scores, probability calculations, and applications

- [10 min] Synthesis: Reflection on discoveries and connection to next week's CLT

```
learnr::run_tutorial("6_NormalDistribution", package = "AppliedStatsInteractive")
```

Week 6 (Feb 17-19)

17 Feb

Learning Goals

- Explain the *Law of Large Numbers (LLN)* both conceptually and mathematically: as sample size increases, the sample proportion converges to the true population probability.
- Explain the *Central Limit Theorem (CLT)* and its conditions: the distribution of sample means from any population is approximately normal for large n .
- Demonstrate both concepts empirically through R simulation and pair programming.
- Assess whether a distribution is approximately normal using visualization techniques (histograms with overlaid curves, Q-Q plots).
- Understand why CLT matters for statistical inference and how it underlies hypothesis testing and confidence intervals.

Reading Assignment

- Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 5, Section 5.1 (Foundations for inference and sampling distributions).

Video Lectures

- [OpenIntro Statistics: Variability of the sample proportion](#)
- [3Blue1Brown: The Central Limit Theorem](#)

Plan

- [20 min] Manual dice rolling activity in groups (build concrete intuition for LLN)
- [25 min] Pair coding activity - Phase 2: Simulate Law of Large Numbers with R
- [25 min] Pair coding activity - Phase 3: Simulate CLT with R and assess normality visually
- [15 min] Synthesis: Comparing discoveries to theory

Note: This activity continues next session (Feb 19), starting with the Null Worlds warmup and Phase 5.

Homework (due Feb 19)

Complete these AppliedStatsInteractive tutorials and submit your 2 completion hashes via Canvas:

```
learnr::run_tutorial("9_FoundationsForInference", package = "AppliedStatsInteractive")  
learnr::run_tutorial("10_IntroInferenceLab", package = "AppliedStatsInteractive")
```

19 Feb

Learning Goals

- Connect the CLT to hypothesis testing via simulation
- Compare CLT-based inference to randomization inference
- Build intuition for null distributions and p-values using interactive tools

Plan

- [15 min] Warmup: Explore [Null Worlds: One-Sample Mean](#) with guided prompts (see worksheet)
- [25 min] Finish Phase 3 of CLT worksheet (sampling distributions, normal overlay, SE comparison)
- [10 min] Phase 4: Synthesis discussion
- [25 min] Phase 5: Build a null distribution using simulation; compare CLT-based approach to randomization; connect back to Null Worlds
- [15 min] Discussion: connecting simulation-based and CLT-based inference

No new homework assigned. Tutorials 9 and 10 (assigned Feb 17) are due today.

Week 7 (Feb 24-26)

24 Feb

Learning Goals

- Evaluate student understanding of probability, distributions, and the CLT (Quiz 2)
- Construct confidence intervals using both *bootstrapping* and the *CLT/standard error* approach

- Explore how confidence interval width depends on *sample size* and *variance*
- Compare CIs from bootstrapping to CIs derived from standard errors
- Practice identifying hypotheses, type I/II errors, and interpreting confidence intervals

Reading Assignment

- Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 5, Sections 5.1-5.2 (Point estimates and confidence intervals)

Video Lectures

- [StatQuest: Confidence Intervals](#) - Visual introduction using bootstrapping
- [Khan Academy: Confidence Intervals and Margin of Error](#)

Plan

- [30 min] Quiz 2: Probability, Distributions, and CLT
- [30 min] Live demo: Bootstrap vs. CLT confidence intervals (Parts 1–6 from `ci_exploration_code.qmd`)
- [25 min] Textbook exercises (work in pairs, discuss as a class): 5.17, 5.27, 5.28, 5.30, 5.34, 5.35

Homework (due Mar 3)

Complete these AppliedStatsInteractive tutorials and submit your 2 completion hashes via Canvas:

```
learnr::run_tutorial("11_HTandCIprop", package = "AppliedStatsInteractive")
learnr::run_tutorial("12_InferencePractice", package = "AppliedStatsInteractive")
```

26 Feb

Learning Goals

- Use *Monte Carlo simulation* to empirically estimate Type I error rates and confidence interval coverage
- Compare three approaches to inference for proportions: the *z-test* (normal approximation), *randomization inference*, and *bootstrapping*
- Understand the *success-failure condition* and observe empirically when the normal approximation breaks down
- Extend the comparison to the *difference of two proportions*

Reading Assignment

- Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 6, Sections 6.1-6.2 (Inference for single proportion and difference of proportions)

Video Lectures

- [OpenIntro: Inference for proportions \(covers 1 and 2 proportion scenarios\)](#)

Plan

- [15 min] Mini-lecture: hypothesis testing for proportions, three approaches (whiteboard)
- [25 min] Worksheet Part 1: Monte Carlo simulation — verifying the z-test ([proportions_ht_worksheet.html](#))
- [5 min] Whole-class discussion
- [35 min] Worksheet Part 2: When does the normal approximation break down? + difference of two proportions
- [10 min] Synthesis discussion + wrap-up

Homework (due Mar 3)

Complete these AppliedStatsInteractive tutorials and submit your 2 completion hashes via Canvas:

```
learnr::run_tutorial("11_HTandCIprop", package = "AppliedStatsInteractive")
learnr::run_tutorial("12_InferencePractice", package = "AppliedStatsInteractive")
```

Week 8 (Mar 5, 10)

05 Mar

Learning Goals

- Understand when to use *chi-square tests* vs other hypothesis tests
- Conduct *goodness of fit tests* to compare observed frequencies to expected frequencies
- Construct and analyze *two-way tables* (contingency tables)
- Perform *chi-square tests of independence* for two categorical variables
- Calculate *expected counts* and interpret *chi-square statistics*
- Interpret results and draw conclusions about categorical relationships

Reading Assignment

- Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 6, Sections 6.3-6.4 (Chi-square goodness of fit and two-way tables)

Video Lectures

- [OpenIntro: Chi-square goodness of fit](#)
- [OpenIntro: Chi-square two-way tables](#)

Plan

- [5 min] Introduction: When do we use chi-square tests?
- [35 min] AppliedStatsInteractive 14_ChiSquare - Part A (goodness of fit)
- [35 min] AppliedStatsInteractive 14_ChiSquare - Part B (two-way tables)
- [5 min] Wrap-up and Q&A

```
learnr::run_tutorial("14_ChiSquare", package = "AppliedStatsInteractive")
```

Homework (due Mar 10)

Complete these AppliedStatsInteractive tutorials and submit your 2 completion hashes via Canvas:

```
learnr::run_tutorial("13_InferenceCategoricalLab", package = "AppliedStatsInteractive")  
learnr::run_tutorial("14_ChiSquare", package = "AppliedStatsInteractive")
```

10 Mar

Learning Goals

- Evaluate student understanding of confidence intervals, hypothesis testing for proportions, and chi-square tests (Quiz 3)
- Understand the *t-distribution* and when to use it instead of the normal distribution
- Conduct *one-sample t-tests* for a population mean
- Recognize *paired data* and apply paired t-tests appropriately

Reading Assignment

- Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 7, Sections 7.1-7.3 (Inference for numerical data)

Video Lectures

- [OpenIntro: The t-distribution](#)
- [OpenIntro: Inference for one mean](#)
- [OpenIntro: Paired data](#)

- [OpenIntro: Difference of two means](#)

Plan

- [30 min] Quiz 3: Confidence Intervals, Hypothesis Testing for Proportions, Chi-Square Tests
- [5 min] Introduction: When do we use the t-distribution?
- [45 min] AppliedStatsInteractive 15_HTandCInum

```
learnr::run_tutorial("15_HTandCInum", package = "AppliedStatsInteractive")
```

Homework (due Mar 12)

- Complete AppliedStatsInteractive 15_HTandCInum if not finished in class

! Milestone 2: Data Visualization

Due Friday, Mar 6 at 11:59 PM

Week 9 (Mar 12)

12 Mar

Learning Goals

- Understand *statistical power* and its relationship to Type II error
- Identify factors affecting power: *effect size*, *sample size*, *variance*, and *significance level*
- Apply *Bonferroni* and other corrections for multiple comparisons
- Distinguish between *statistical significance* and *practical significance* (aka theoretical or substantive significance)

Reading Assignment

- Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 7, Section 7.4 (Power calculations for a difference of means)

Video Lectures

- [StatQuest: Statistical Power, Clearly Explained](#)
- [OpenIntro: Multiple comparisons](#)

Plan

- [20 min] Shooting Baskets activity
- [25 min] Multiple Comparisons activity
- [15 min] R demonstration: Bonferroni correction
- [5 min] Wrap-up

Homework (due Mar 24)

Complete these AppliedStatsInteractive tutorials and submit your 2 completion hashes via Canvas:

```
learnr::run_tutorial("15_HTandCInum", package = "AppliedStatsInteractive")
learnr::run_tutorial("16_InferencePractice", package = "AppliedStatsInteractive")
```

Spring Break: March 14-22

Week 10 (Mar 24-26)

24 Mar

Learning Goals

- Understand when *ANOVA* is appropriate (comparing means across 3+ groups)
- Interpret the *F-statistic* and ANOVA table
- Check conditions for ANOVA (independence, approximate normality, equal variance)
- Apply *Tukey's HSD* for pairwise comparisons after ANOVA

Reading Assignment

- Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 7, Section 7.5 (Comparing many means with ANOVA)

Video Lectures

- [OpenIntro: Intro to ANOVA](#)
- [OpenIntro: Conditions for ANOVA](#)

Plan

- [10 min] Warm-up: Why use ANOVA instead of multiple t-tests? (Connect to multiple comparisons activity from last session)

- [45 min] AppliedStatsInteractive 18_ANOVA
- [20 min] R demonstration: Tukey's HSD for ANOVA follow-up (see `tukey_demo.md`)
- [5 min] Wrap-up

```
learnr::run_tutorial("18_ANOVA", package = "AppliedStatsInteractive")
```

Homework (due Mar 26)

- Complete AppliedStatsInteractive 18_ANOVA if not finished in class

26 Mar

Learning Goals

- Demonstrate understanding of ANOVA and Tukey's HSD (Quiz 4)
- Understand the structure of *bivariate data* and when regression is appropriate
- Interpret *scatterplots* and identify linear relationships
- Calculate and interpret *correlation* (r)

Reading Assignment

- Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 8, Sections 8.1-8.2 (Line fitting, residuals, correlation)

Video Lectures

- [OpenIntro: Intro to linear regression](#)
- [OpenIntro: Correlation](#)

Plan

- [30 min] Quiz 4: ANOVA and Tukey's HSD
- [5 min] Introduction: From comparing groups to predicting outcomes
- [40 min] AppliedStatsInteractive 19_LinearRegression (first half)

```
learnr::run_tutorial("19_LinearRegression", package = "AppliedStatsInteractive")
```

Homework (due Mar 31)

- Continue working through AppliedStatsInteractive 19_LinearRegression
-

Week 11 (Mar 31 - Apr 2)

31 Mar

Learning Goals

- Fit a *least squares regression line* to data
- Interpret the *slope* and *intercept* in context
- Calculate and interpret *residuals*
- Understand *R-squared* (R^2) as the proportion of variance explained

Reading Assignment

- Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 8, Sections 8.2-8.3 (Least squares regression, types of outliers)

Video Lectures

- [OpenIntro: Least squares regression](#)
- [OpenIntro: R-squared](#)

Plan

- [10 min] Warm-up: Interpreting correlation from last session
- [35 min] Guessing Ages Regression activity
- [30 min] [OpenIntro Lab: Introduction to Linear Regression](#)

Homework (due Apr 2)

- Complete AppliedStatsInteractive Tutorial 17_InferenceNumericalLab and submit your completion hash via Canvas:

```
learnr::run_tutorial("17_InferenceNumericalLab", package = "AppliedStatsInteractive")
```

- Finish [OpenIntro Lab: Introduction to Linear Regression](#) if not completed in class. Submit your .qmd source file and rendered output (PDF or HTML) to Canvas.

2 Apr

Learning Goals

- Check *conditions for regression inference* using residual plots
- Identify common violations: non-linearity, non-constant variance, non-normality
- Conduct *hypothesis tests* for the regression slope
- Construct and interpret *confidence intervals* for regression parameters

Reading Assignment

- Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 8, Section 8.4 (Inference for linear regression)

Video Lectures

- [OpenIntro: Inference for regression](#)
- [StatQuest: Linear Regression Assumptions](#)

Plan

- [10 min] Warm-up: Why do we need conditions for inference?
- [30 min] Residual Diagnostics activity
- [20 min] Pair exercise: Regression inference concepts
- [15 min] Wrap-up discussion

Homework (due Apr 7)

- [OpenIntro Lab: Introduction to Linear Regression](#) (continued from lesson 20, if not finished). Submit your .qmd source file and rendered output (PDF or HTML) to Canvas.
-

Week 12 (Apr 7-9)

7 Apr

Learning Goals

- Demonstrate understanding of simple linear regression (Quiz 5)
- Understand when and why to use *multiple regression*
- Predict and explain how coefficients change when variables are added
- Interpret coefficients as *marginal effects* (holding other variables constant)

Reading Assignment

- Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 9, Section 9.1 (Introduction to multiple regression)

Video Lectures

- [OpenIntro: Multiple regression intro](#)

Plan

- [30 min] Quiz 5: Simple linear regression
- [5 min] Transition: From one predictor to many—brief introduction
- [45 min] Activity: What Happens When You Add a Variable?

Homework (due Apr 9)

- Complete these AppliedStatsInteractive tutorials and submit your 2 completion hashes via Canvas:

```
learnr::run_tutorial("18_ANOVA", package = "AppliedStatsInteractive")
learnr::run_tutorial("19_LinearRegression", package = "AppliedStatsInteractive")
```

- Read Chapter 9.1-9.2
- [OpenIntro Lab: Multiple Linear Regression](#). Submit your .qmd source file and rendered output (PDF or HTML) to Canvas.

9 Apr

Learning Goals

- Distinguish *confounders* from *mediators* and *colliders*
- Decide when controlling for a variable is appropriate
- Understand *collinearity* and its effect on coefficient estimates
- Explain why R^2 and *adjusted* R^2 don't tell you whether you're estimating the right thing

Reading Assignment

- Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 9, Section 9.2 (Checking model conditions)

Video Lectures

- [Confusing Terms:
- [Causal Diagrams: Causality](#) (10 min) - from *The Effect* by Nick Huntington-Klein
- [Causality: Closing Back Doors](#) (10 min) - controlling for confounders
- [Closing Causal Pathways, and Collider Variables](#) (11 min) - why you shouldn't control for colliders

Plan

- [75 min] Activity: [Control, Confounder, or Something Else?](#)

Homework (due Apr 14)

- Finish the activity (if not completed in class)
- Read Chapter 9.4 (Logistic regression introduction)

! Milestone 3: Statistical Analysis

Due Friday, Apr 10 at 11:59 PM

Week 13 (Apr 14-16)

14 Apr

Learning Goals

- Recognize when the outcome variable is *binary* (yes/no, success/failure)
- Understand why linear regression is inappropriate for binary outcomes
- Fit a logistic regression model using `glm()`
- Convert between probability, odds, and log-odds

Reading Assignment

- Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 9, Section 9.4 (Introduction to logistic regression)

Video Lectures

- [OpenIntro: Logistic regression intro](#)
- [StatQuest: Logistic Regression](#) (optional but excellent)

Plan

- [40 min] Activity: Trashball
- [25 min] In-class worksheet: Probability, Odds, and Log-Odds
- [10 min] Wrap-up: Other binary outcome problems (medical diagnosis, loan default, etc.)

Homework (due Apr 16)

- Finish worksheet if not completed in class

16 Apr

Learning Goals

- Interpret logistic regression output in a real-world context
- Distinguish *inference* (understanding relationships) from *prediction* (forecasting outcomes)

Reading Assignment

- Diez, Çetinkaya-Rundel, and Barr (2019) Chapter 9, Section 9.4 (continued)

Video Lectures

- [StatQuest: Logistic Regression Details](#)

Plan

- [25 min] Activity: Titanic Case Study
- [30 min] Lecture: Inference vs. Prediction

Homework (Quiz 6 on Apr 21)

- Study for Quiz 6
- Prepare project presentations (first group presents Tuesday). See [Presentation Guidelines](#).

Week 14 (Apr 21-23)

21 Apr

Learning Goals

- Demonstrate understanding of multiple and logistic regression (Quiz 6)
- Communicate statistical findings to a general audience

Reading Assignment

- None

Video Lectures

- None

Plan

- [30 min] Quiz 6: Multiple regression and logistic regression

- [45 min] Project presentations (~5 students, 8-9 min each including Q&A). See [Presentation Guidelines](#).

Homework (presentations Apr 23)

- Remaining presenters: finalize presentations for Thursday

23 Apr

Learning Goals

- Communicate statistical findings to a general audience
- Provide constructive feedback on peer presentations

Reading Assignment

- None

Video Lectures

- None

Plan

- [55 min] Project presentations (~6 students, 8-9 min each including Q&A). See [Presentation Guidelines](#).
- [10 min] Course wrap-up and reflection

In-class Presentations during Week 14

! Final Report

Due Friday, May 1 at 11:59 PM

References

Diez, David, Mine Çetinkaya-Rundel, and Christopher D Barr. 2019. *OpenIntro Statistics, Fourth Edition*. self-published. <https://openintro.org/os>.