

Data Analysis and Statistical Inference

Introduction

Sta 101 - Spring 2016

Duke University, Department of Statistical Science

Dr. Çetinkaya-Rundel

Slides posted at http://bit.ly/sta101_s16

- ▶ Professor: Dr. Mine Çetinkaya-Rundel - mine@stat.duke.edu
- ▶ TAs:
 - David Clancy (Head TA)
 - Lindsay Berry
 - Patrick Cardel
 - Anthony Hung
 - Eidan Jacob
 - John Pura
 - Sarah Zimmerman

1

Required materials

- ▶ OpenIntro Statistics, 3rd Edition: <http://openintro.org/os>
- ▶ i>clicker2 - See Google Doc for a list of students selling used clickers (link emailed)
- ▶ (optional) Calculator (just something that can do square roots)

Webpage

http://bit.ly/sta101_s16

- ▶ Pictures and summaries of data
 - *Unit 1 - Intro to data*: Observational studies & non-causal inference, principles of experimental design & causal inference, exploratory data analysis, introduction to simulation-based statistical inference.
- ▶ Mathematics behind statistics
 - *Unit 2 - Probability & distributions*: Basics of probability and chance processes, Bayesian perspective in statistical inference, the normal and binomial distributions.
- ▶ Statistical inference
 - *Unit 3 - Framework for inference*: CLT, sampling distributions, and introduction to theoretical inference.
 - Midterm 1
 - *Unit 4 - Statistical inference for numerical variables*
 - *Unit 5 - Statistical inference for categorical variables*
 - Midterm 2
- ▶ Modeling
 - *Unit 6 - Simple linear regression*: Bivariate correlation and causality, introduction to modeling.
 - *Unit 7 - Multiple linear regression*: More advanced modeling with multiple predictors.
 - Final Exam

- ▶ Set of learning objectives and required and suggested readings, videos, etc. for each unit.
- ▶ Prior to beginning the unit, watch the videos and/or complete the readings and familiarize yourselves with the learning objectives.
- ▶ Begin a new unit with a readiness assessment: individual, then team.
- ▶ Class time: split between lecture, discussion/application, and lab.
- ▶ Complement your learning with problem sets.
- ▶ Wrap up a unit with a performance assessment.

- ▶ Highly functional teams of learners based on survey and pre-test.
- ▶ Team members first point of contact.
- ▶ Application exercises, labs, team readiness assessments, project.
- ▶ Study together, but anything that is not explicitly a team assignment must be your own work.
- ▶ Peer evaluations to ensure that all team members contribute to the success of the group and to address any potential issues early on.
 - If you feel that there are issues within your team, you are encouraged to discuss it with your team members and to bring it to my or your TA's attention ASAP (don't wait till things get worse).

Objective: Two-way communication and instant feedback.

- ▶ Readiness assessments (graded for accuracy)
- ▶ Questions throughout lecture (graded for participation)
 - Get credit for the day you by responding to at least 75% of the questions.
 - Up to three unexcused late arrivals or absences.
- ▶ Register your clicker at <https://www1.iclicker.com/register-clicker> (Student ID = Net ID)
 - If you bought a used clicker, the registration process might ask for a payment. You don't need to pay, you'll be able to register in class.

Objective: Give you independent applied research experience using real data and statistical methods.

- ▶ Proposal: due mid-semester
- ▶ Poster session: last lab of semester
- ▶ Complete in teams, along with peer evaluations to track contribution of each member
- ▶ Must complete the project and score at least 30% of the points on each project in order to pass this class

Midterm 1	Feb 24, Wed
Midterm 2	Mar 30, Wed
Final	May 5, Thu - 7-10pm

- ▶ Exam dates cannot be changed, no make-up exams will be given
- ▶ If you cannot take the exams on these dates you should drop this class
- ▶ Calculator + cheat sheet allowed

- ▶ I will regularly send announcements by email, so make sure to check your email daily.
- ▶ All content related (non-personal) questions should be posted on Piazza.
- ▶ Before posting a new question please make sure to check if your question has already been answered, and answer others' questions.
- ▶ Use informative titles for your posts.
- ▶ It is more efficient to answer most statistical questions "in person" so make use of OH.

- ▶ Prof:
 - Discussion section: Tuesdays, 4 - 5:30pm, Soc Sci 139
 - Office hours: Thursdays, 10:30am - 12pm, Old Chem 213 (my office)

▶ TAs:

TA	Day / time	Location
Sarah Zimmerman	Sun 11am - 1pm	Old Chem 211A
David Clancy	Sun 7:30pm - 9:30pm	Old Chem 025
Drew Jordan	Mon 4:30pm - 5:30pm + 7pm - 8pm	Old Chem 211A
Anthony Hung	Mon 6pm - 8pm	Old Chem 025
Andrew Wong	Mon 7:30pm - 9:30pm	Old Chem 025
Anne Driscoll	Tue 9:30am - 11:30am	Old Chem 211A
Katherine Ahn	Tue 11am - 1pm	Old Chem 211A
Andy Cooper	Tue 2pm - 4pm	Old Chem 211A
Patrick Cardel	Thu 12pm - 2pm	Old Chem 211A
John Pura	Thu 2pm - 3pm + 6pm - 7pm	Old Chem 211A
Lindsay Berry	Thu 3pm - 5pm	Old Chem 211A
Eidan Jacob	Fri 9:30am - 11:30am	Old Chem 211A

Students with disabilities who believe they may need accommodations in this class are encouraged to contact the [Student Disability Access Office](#) at (919) 668-1267 as soon as possible to better ensure that such accommodations can be made.

<http://www.access.duke.edu/students/requesting/index.php>

Any form of academic dishonesty will result in an immediate 0 on the given assignment and will be reported to the Office of Student Conduct. Additional penalties may also be assessed if deemed appropriate. If you have any questions about whether something is or is not allowed, ask me beforehand.

Some examples:

- ▶ Use of disallowed materials (including any form of communication with classmates or accessing the web) during exams and readiness assessments
- ▶ Plagiarism of any kind
- ▶ Use of outside answer keys or solution manuals for the homework

Tips for success

- ▶ Complete the reading before a new unit begins, and then review again after the unit is over.
- ▶ Be an active participant during lectures and labs.
- ▶ Ask questions - during class or office hours, or by email. Ask me, your TAs, and your classmates.
- ▶ Do the problem sets - start early and make sure you attempt and understand all questions.
- ▶ Take each PA and complete practice quizzes (on Coursera) for each unit, and review the feedback for questions you miss.
- ▶ Start your project early and allow adequate time to complete them.
- ▶ Give yourself plenty of time to prepare a good cheat sheet for exams. This requires going through the material and taking the time to review the concepts that you're not comfortable with.
- ▶ Do not procrastinate - don't let a unit go by with unanswered questions as it will just make the following unit's material even more difficult to follow.

To do

- ▶ Download or purchase the textbook
 - Download: <http://openintro.org/os>
 - Purchase: <http://openintro.org/os/amazon>
- ▶ Obtain and register your clicker
 - <https://www1.iclicker.com/register-clicker> (Student ID = Net ID)
 - Note: If you have a used clicker this site might ask you to pay, you don't have to you'll get a chance to register it in class
- ▶ Read the syllabus and let me know if you have any questions
- ▶ Watch/Read/Review the resources for Unit 1
 - RA 1 on Wednesday - not graded (for practice)

Syllabus quiz (10 minutes).

You can refer to the syllabus at http://bit.ly/sta101_s16.