

**SPRING 2019**



# **CE 311S : PROBABILITY AND STATISTICS**

---

Discussion session

M 1:00 – 2:00 PM

**PRIYADARSHAN PATIL**

Teaching Assistant, The University of Texas at Austin

# Introduction

- I am a graduate student working with Dr. Stephen Boyles
- Background: Undergrad in Civil engineering and MS in Transportation engineering
- First time being a TA for CE311S, previous TA experience for CE 301 and CE 321

# Introduce yourself

- Your name
- Year of study / Classification (Freshman, sophomore, etc.)
- One interesting thing about yourself

# Contact information

- Email: [Priyadarshan@utexas.edu](mailto:Priyadarshan@utexas.edu)
- Office: ECJ 6.504
- Office hours: F 11:30-1:30 PM in ECJ 6.406

# Course organization

- Probability – till spring break
- Statistics – post spring break
- Three exams (March 8<sup>th</sup>, April 12<sup>th</sup> and May 17<sup>th</sup>)
- HW assignments (~6-8 homework assignments)
- Canvas quizzes (~6-8 quizzes)
- Reading responses (most weeks will have a reading response due)

# Lab goals

- Enhance learning from Dr. Boyles' lectures
  - Explain confusing concepts
  - Additional problems and examples
- Encourage participation in the course
- Use R for statistical analysis

# General lab structure

- Concept review, practice problems, practice R examples
- For labs requiring R, more time shall be dedicated to R examples
- Technology: Presentation + Writing on board
- You are encouraged to take notes
- Feedback

# Expectations

- Come prepared for the lab (Self-learning and knowing what you don't understand well)
- Stay attentive and ask questions (There are no stupid questions)
- Some don'ts:
  - Don't be late
  - Don't misuse the availability of computers for browsing the internet or social media
  - Don't ask me to solve HW questions exactly (Variants or some hints/ideas are ok)



# A game of prob/stat intuition

- For this lab section only, we shall have a semester long “game” involving a simple yes/no or binary choice bet every week.
- Everyone starts off with 1000 points, you may wager any amount of points you have. The bet shall be double or nothing for the wager amount.
- Goal is to maximize your point total at the end of the semester.

# A game of prob/stat intuition

- At the start of each lab session, the first 5 minutes shall be devoted to this game. The question shall be presented, you may think over it and then submit your answer to me on a piece of paper with your name, the bet amount and the answer by 1:05PM
- We shall then work through the example
- No grades tied to this, maybe a box of cookies to the winner at the end of the semester

# A game of prob/stat intuition

- Any questions?
- Link to your progress: [Google sheets link](#)
- Let's get started

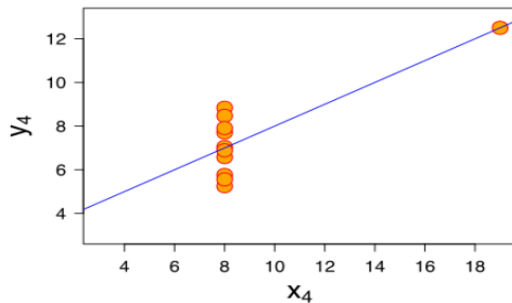
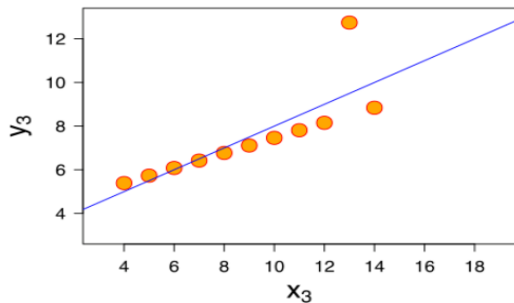
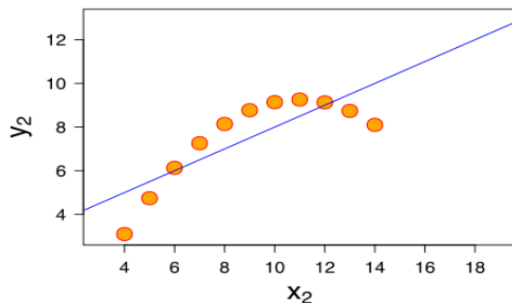
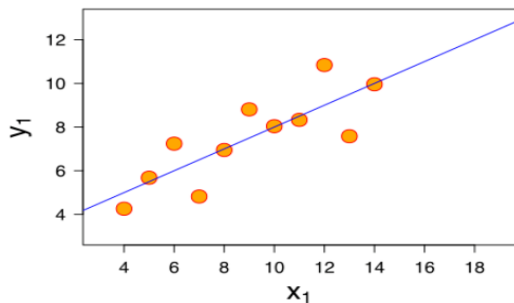
# Week 1: Question

- Can there exist two datasets that have exactly the same mean, variance, correlation, and regression line, but which are comprised of completely different data points?
- A dataset here is defined as a set of  $(x,y)$  points.
- Example dataset:  $\{(1,1), (1,2), (2,2)\}$

# Week 1: Question

Please turn in your Yes/No answer with your wager and name

# Week 1: Answer



- Mean of  $x = 9$
- Mean of  $y = 7.50$
- Variance of  $x = 11$
- Variance of  $y = 4.12$
- Correlation = 0.816
- Linear Regression line:  $y = 3 + 0.5x$

# Raw Data

I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

# Introduction to R and RStudio

- What is R?
- Why do we need to use R?
- Let's switch over to RStudio



# Key Points to remember from R

- Variables
- Scalar math
- Printing and output generation
- PEMDAS / Order of operations
- Vectors and Vector math
- Calling inbuilt functions as well as installing libraries
- Reading a CSV file and working with specific columns