

ISyE 6739 – S18 Statistical Methods (Tuesday and Thursday @ 4:30-5:45 pm)

Instructor: Kamran Paynabar
H. Milton Stewart School of
Industrial and Systems Engineering
Georgia Tech

Kamran.paynabar@isye.gatech.edu
Office: Groseclose 436

About You

- **Name:**
- **Major and Year:**
- **Anything you like to share:**

General Course Information

- **INSTRUCTOR**

- Kamran Paynabar, Office: Groseclose 436,

- Email: kamran.paynabar@isye.gatech.edu

- Office hours: Tue 3:00 to 4:00 pm and Thu 11:00 am to 12:00 pm or by appointment via email

- **TA**: Yuliia Lut, yuliia.lut@gatech.edu.

- Office hours: Wednesday, 11am – 12 pm and Friday 9am – 10am (Main 224)

- **Textbook**

- “Applied Statistics and Probability for Engineers”, 6th Edition, Douglas C. Montgomery, George C. Runger, Wiley, ISBN 978-1118539712, © 2014, 836 pages.

- **COURSE WEBSITE**

- Go to : <http://canvas.gatech.edu/> to find syllabus, lecture notes, and homework assignments/solutions

- **Software**

- Minitab® 17 and R: You can access the software through Vlab at

- <http://vlab.matrix.gatech.edu/>

- **Prerequisite**: Calculus

General Course Information

- **Lecture Videos**

- Available on YouTube. The link of video for each session is posted in advance on Canvas under “calendar”.
 - A short video assignment is given for each class.

- **Group Activities**

- We have group activities in every session.
 - Each group should have 2 students. Send your group information via <https://docs.google.com/spreadsheets/d/1n-bymP0zMCoS6gybM0bGpmr4oBBATe62y9TdwUGuTho/edit?usp=sharing>

Grading Policy

- **Homework (15%) and assignments (5%)** 20%
 - **Quizzes and class participation** 15%
 - **Exam 1 (02/20)** 20%
 - **Exam 2 (04/17)** 25%
 - **Group project (presentation on 04/24, and 05/01, report due on 05/01)** 20%
- There are approximately biweekly homework assignments. The homework should be handed in *in the beginning of the class on the due date*. NO late submission is acceptable. (The lowest score will be dropped)
 - There are quizzes and group activities in each class. Each group should have 2 students.
 - You are encouraged to discuss homework/assignments problems with your fellow students. But your final answers should be based on your own understanding unless it is a group assignment, which will be announced on Canvas. Copying others' work is NOT acceptable and violates the honor code.
 - Requests for re-grading HW/exams/quizzes should be made within a week of returning them.
 - Exams are comprehensive and closed-book. For exam I and II students are allowed to bring one and two (double-sided) sheets for equations, respectively. There are no constraints on the contents of the notes.
 - Detailed information about group project can be found on Canvas.

Course Objectives

- ✓ Summarize and interpret a dataset using descriptive statistics
- ✓ Determine a probability distribution of a population based on a random sample
- ✓ Estimate parameters of a distribution based on a random sample
- ✓ Construct confidence intervals for parameters of a distribution
- ✓ Make a decision about a population based on a random sample
- ✓ Predict a response variable based on one or more predictor variables
- ✓ Identify important factors influencing a response variable

Basic Concepts in Statistics

Statistics

The field of **Statistics** deals with the *collection, presentation, analysis, and use* of data to model systems, make decisions, solve problems, and design products and processes

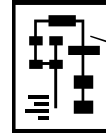
Statistics is the science of data

Examples: Statistics helps us

- Predict the demand of a product / the stock prices
- Select the best supplier with the least lead time (or highest quality)
- Monitor and control a process
- Simulate and model an ER
- Predict election results
- Design new products
- etc.,

Types of Models

- **Physical models:** A smaller or larger physical copy of an object
- **Schematic models:** Represent a system by symbols and signs (e.g., flow diagram, flow process chart)

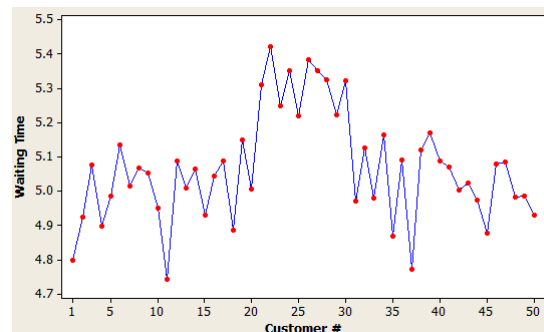


- **Analytical models:** Mathematical and statistical representation of a system
 - **Deterministic:** $F = m \times a$ (based on Physics laws)
 - **Stochastic:** $F = m \times a + e$ (based on both Physics laws and data, more realistic)

Statistics and Statistical Thinking

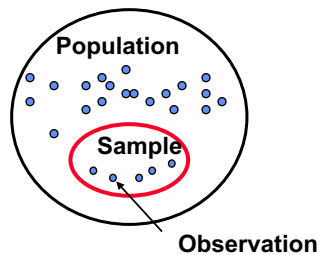
Waiting time of customers in a bank

4.80	4.92	5.08	4.90	4.98	5.14	5.02	5.07	5.05	4.95
4.74	5.09	5.01	5.07	4.93	5.05	5.09	4.89	5.15	5.01
5.31	5.42	5.25	5.35	5.22	5.39	5.35	5.33	5.22	5.32
4.97	5.13	4.98	5.17	4.87	5.09	4.77	5.12	5.17	5.09
5.07	5.00	5.02	4.97	4.88	5.08	5.08	4.98	4.99	4.93

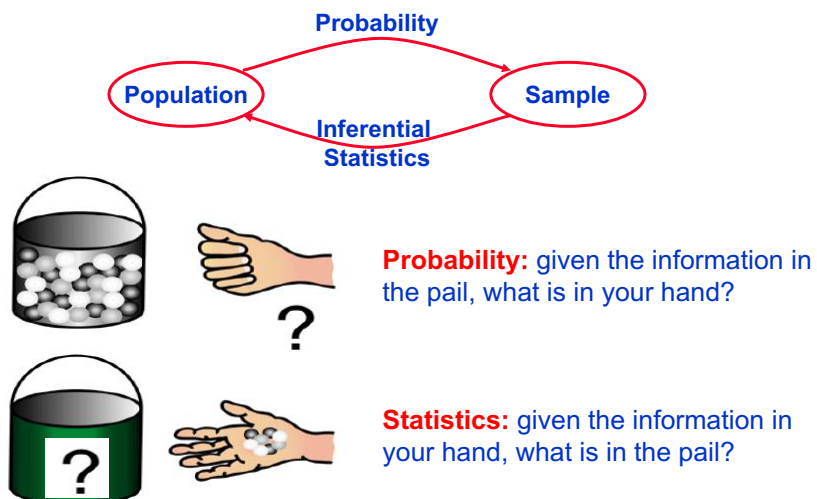


Population Vs. Sample

- **Population:** a finite well-defined group of ALL objects which, although possibly large, can be enumerated in theory (e.g. investigating ALL the bearings manufactured today).
- **Sample:** A sample is a SUBSET of a population (e.g. select 50 out of 1,000 bearings manufactured today).



Probability Vs. Statistics

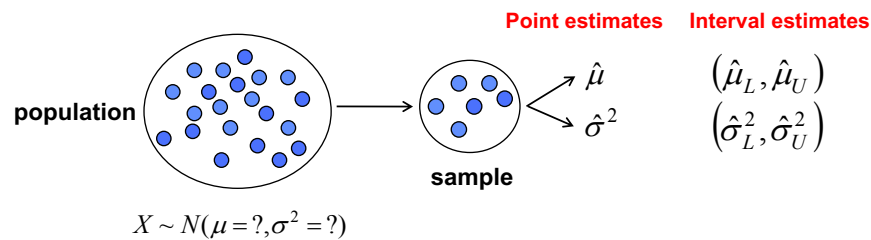


Course Objectives (revisited)

- ✓ Determine a probability distribution of a population based on a random sample
- ✓ Estimate parameters of a distribution based on a random sample
- ✓ Construct confidence intervals for parameters of a distribution

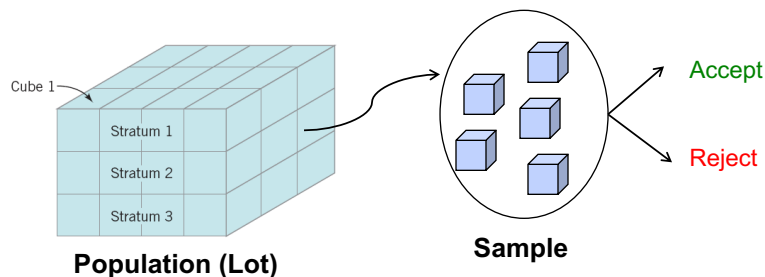
What is the probability that weekly sale of a restaurant is more than 10K?

$$\Pr(X > 10,000)$$



Course Objectives (revisited)

- ✓ Make a decision about a population based on a random sample



Course Objectives (revisited)

- ✓ Predict a response variable based on one or more predictor variables

Given the following data, what is the purity if hydrocarbon level is 1.05?

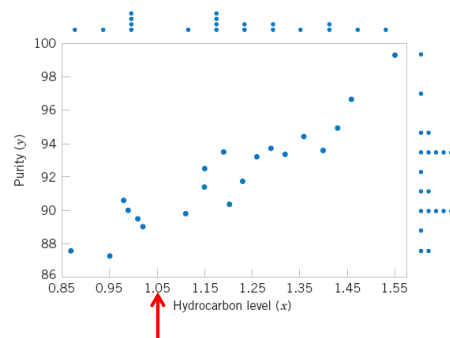


Table 11-1 Oxygen and Hydrocarbon Levels

Observation Number	Hydrocarbon Level x (%)	Purity y (%)
1	0.99	90.01
2	1.02	89.05
3	1.15	91.43
4	1.29	93.74
5	1.46	96.73
6	1.36	94.45
7	0.87	87.59
8	1.23	91.77
9	1.55	99.42
10	1.40	93.65
11	1.19	93.54
12	1.15	92.52
13	0.98	90.56
14	1.01	89.54
15	1.11	89.85
16	1.20	90.39
17	1.26	93.25
18	1.32	93.41
19	1.43	94.98
20	0.95	87.33

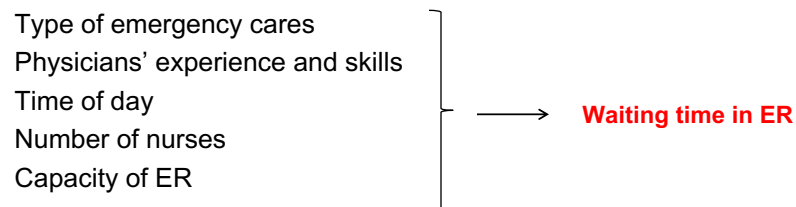
ISyE 6739, Introduction

15

Course Objectives (revisited)

- ✓ Identify important factors influencing a response variable

What are important variables affecting the waiting time in ER?



ISyE 6739, Introduction

16

Data Collection

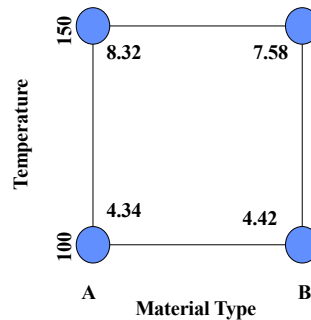
Three basic methods for collecting data:

- A **retrospective** study using historical data
- An **observational** study
- A **designed experiment**

observational study

Person	Height	Weight
1	5'-10"	145
2	6'	185
...

Designed Experiment



Statistics and Statistical Thinking

- Statistical techniques are useful for describing and understanding **variability**.
- By **variability**, we mean successive observations of a system or phenomenon do not produce exactly the same result.
- Statistics gives us a framework for describing this variability and for learning about potential **sources of variability**.

