

ISyE 3770 Statistics and Applications

Instructor: Dan Li
Slides from Prof. Nagi Gebraeel

School of Industrial and Systems Engineering
Georgia Institute of Technology

Spring 2021

1 Syllabus

2 Introduction

- Probability and Statistics

Course Description

- Title: Statistics and Applications
- Book: Applied Statistics and Probability for Engineers, D. C. Montgomery and G. C. Runger
- Instructor: Dan Li (lidan_thu@gatech.edu)
- TA: Xufei Liu (xufeiliu2000@gatech.edu)
- Office Hours: TBA (Check Canvas announcements)

Course Description

- The course objective is to introduce students to basic concepts in probability and statistics, such as counting methods, random variables and their types, how to describe data, how to use data in for estimation descriptive parameters, how to perform statistical inferences (namely hypothesis testing), and how to study relationships among variables.

Course Description

- General topics that will be covered include the following:
 - Basic Probability: Axioms & Rules
 - Random Variables: Discrete & Continuous
 - Joint and Conditional Probability
 - Descriptive Statistics
 - Sampling Distributions
 - Confidence Intervals
 - Hypothesis testing
 - Linear Regression & Analysis of Variance *(depends on course progress)

Grading Policy

Assignments	25%
Midterm Exam I	20%
Midterm Exam II	20%
Final Exam Cumulative	30%
Pop Quizzes & Class Participation	5%

Class Policy

- There are no makeup exams or quizzes.
- Assignments will be turned in as PDF files on Canvas. No late assignments will be accepted.
- There will be two midterm exams held during class. The exact dates will be announced later.
- There will be one final exam and it will be cumulative.
- Honor Code: Students are reminded of the Georgia Tech Honor Code. Please see <http://www.honor.gatech.edu>.

About the Instructor

- Ph.D. Candidate in the School of Industrial and Systems Engineering at Georgia Tech. Advisors: Prof. Nagi Gebraeel and Prof. Kamran Paynabar.
- MS (2020) in Statistics, Georgia Tech.
- BS (2015) in Automotive Engineering, Tsinghua University.
- Research Interest: Cybersecurity, dynamic system modeling and control, anomaly detection, statistical learning, cyber-physical systems, data analysis.
- Please call me Dan – not a professor or a Dr. (yet).

Probability

- Probability: (uncertain world, perfect knowledge of the uncertainty)
 - Counting
 - Random variables, distributions, quantiles, mean variance
 - Conditional probability, Bayes' theorem, base rate fallacy
 - Joint distributions, covariance, correlation, independence
 - Central limit theorem
- Probability help quantify the risks involved in statistical inference, that is, risks involved in decisions made every day.
- Probability provides the framework for the study and application of statistics.

Statistics

- Statistics (data in an uncertain world, perfect/imperfect knowledge of the uncertainty)
 - Point Estimation
 - Statistical Intervals
 - Tests of Hypotheses
- Statistics is a branch of mathematics that concerns the collection, organization, displaying, analysis, interpretation and presentation of data.
- In layman's terms it is the **Science of Data**
- In statistics, we apply probability(probability theory) to draw conclusions from data.

Probability vs Statistics

- **Probability example:** You have a fair coin (equal probability of heads or tails). You will toss it 100 times. What is the probability of 60 or more heads? We can get only a single answer because of the standard computation strategy.
- **Statistics example:** You have a coin of unknown provenance. To investigate whether it is fair you toss it 100 times and count the number of heads. Let's say you count 60 heads. Your job as a statistician is to draw a conclusion (inference) from this data. In this situation, different Statisticians may draw different conclusions because they may use different conclusion forms or may use different methods for predicting the probability (e.g. of landing heads).

What You Will Learn

- Basics of Probability
- Useful Probability Distributions (Discrete & Continuous)
- Descriptive Statistics
- Point Estimation Techniques
- Statistical Inference Techniques (Hypothesis Testing & Confidence Interval)

What You Will Learn

You will be able to solve problems like these by the end of the semester:

- You work for a semiconductor manufacturer that produces controllers used in automobile engine applications. The customer requires that the process fallout or fraction defective at a critical manufacturing step not exceed 0.01 and that the manufacturer demonstrate process capability at this level of quality using Type I error rate $\alpha = 0.05$. You take a random sample of 200 devices and find that three of them are defective. Can your company demonstrate process capability for the customer?

What You Will Learn

- You are analyzing two catalysts to determine how they affect the mean yield of a chemical process. You have run 8 tests and the results of the process yield are given below. Can you decide whether catalysts 1 & 2 have the same yield with 95% confidence?

Observation Number	Catalyst 1	Catalyst 2
1	91.50	89.19
2	94.18	90.95
3	92.18	90.46
4	95.39	93.21
5	91.79	97.19
6	89.07	97.04
7	94.72	91.07
8	89.21	92.75
	$\bar{x}_1 = 92.255$	$\bar{x}_2 = 92.733$
	$s_1 = 2.39$	$s_2 = 2.98$