Lecture 1. Introduction

- What is this course about?
- Logistics
- Questionnaire

Where you have seen "statistics"?

- Application fields of statistics
  - sports (movie "Money Ball")
  - stock market/finance
  - weather forecast
  - machine learning/computer/internet
  - politics
  - biology/medicine

What are statistical methods?

Question 1: what is the true value?

- Digital thermometer takes measurements
- measurements subject to a random error additive to the true value
- If you take 5 measurements, and obtain a sequence of numbers

$$98.2 \quad 98.6 \quad 97.4 \quad 98.2 \quad 97.9 \quad 98.9$$

  then what the true temperature is likely to be?

- voltage of power line, chemical concentration ...

point estimator

Question 2: true or false?

- Apple production line produces a batch of 12 MacBook airs
- a technician claims: a common part - battery, in these MacBooks are entirely defective
- quality test shows that 6 out of 12 MacBooks are defective
- technician said: true of false?

hypothesis test

Question 3: what interval the true value might be in?

- for the same digital thermometer example
- if instead of asking "what is the most likely true temperature"
- we ask "what is a range [a, b], such that the true temperature is most likely to be within"

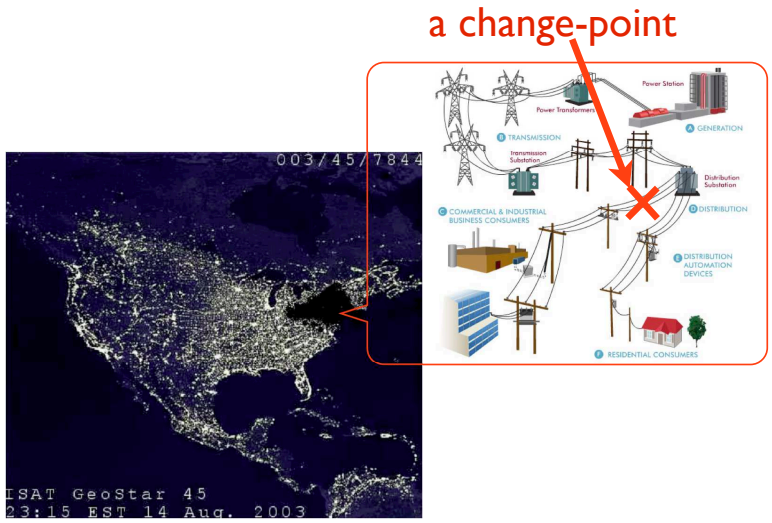
confidence interval

Question 4: design of experiment

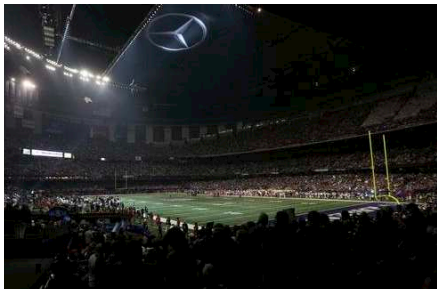
- Color blindness appears in 1% of the people in a certain population.
- How large must a sale be if the population of its containing a color-blind person is to be .95 or more?

how to design experiments

Why engineers care about experiments/measurements



Northeast blackout, 2003

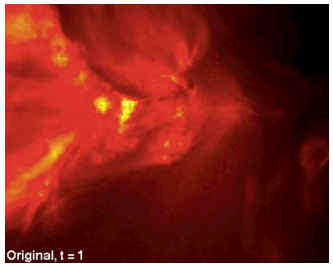
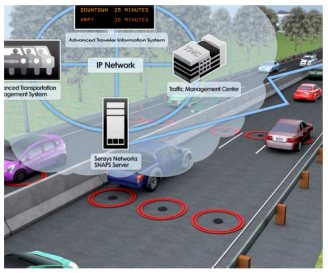
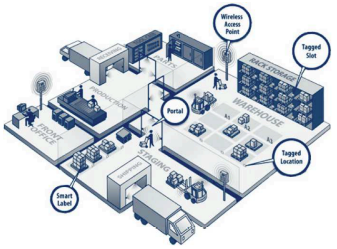


Super Bowl blackout, 2013

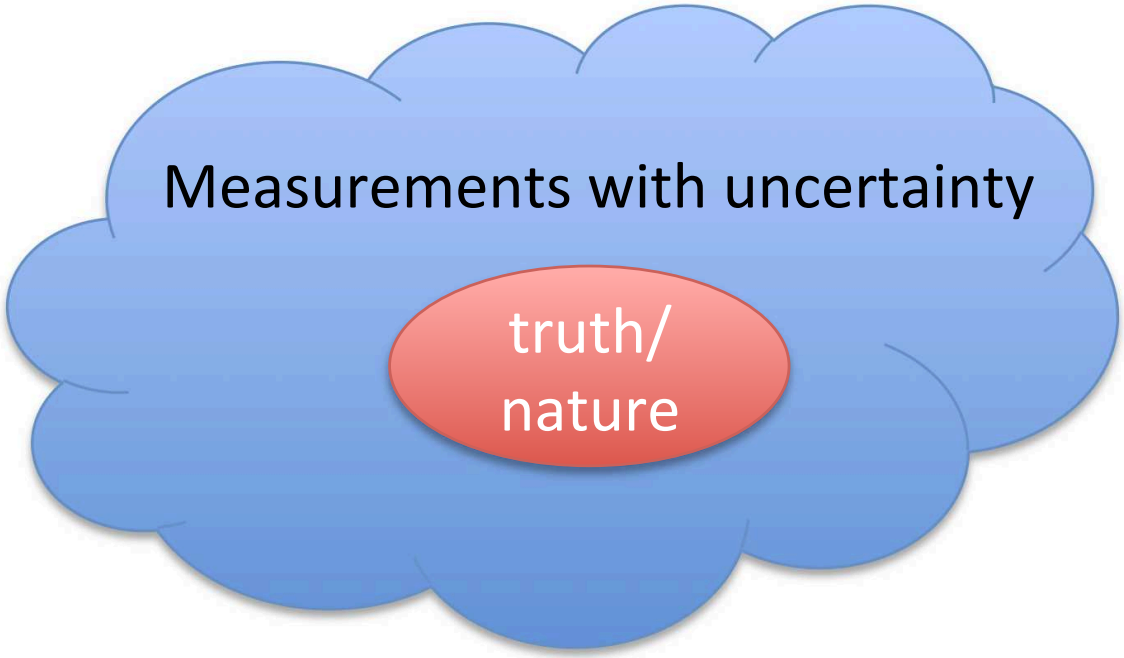
- Northeast blackout of 2003 is a widespread power outage in the northeastern US
- caused by

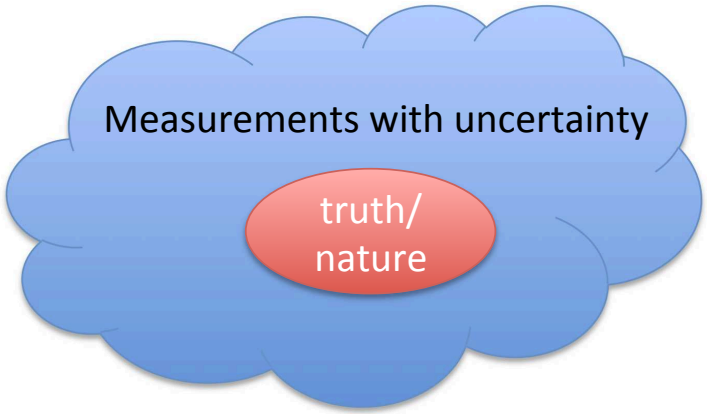
  "...did not recognize the deteriorating condition of the system [in the Ohio part]."
- **change-point**: deteriorating condition of the power system (e.g. break-down of power line)
- detect the change-point in real-time may prevent large-scale blackout

jcwinnie.biz, bravoprojects.co.in

applications	solar flare detection 	power network monitoring 	traffic monitoring 	RFID supply chain management 
sensors	each pixel can be viewed as one sensor.	smart meters 	traffic sensors 	RFID tags 
# sensors	$\sim 60,000$ pixels	$\sim 10^5$ in a subset	$\sim 25,000$ in California	millions

kk.org, wikipedia, ubiu.co.kr, sensysnetworks.com, P. Varaiya


Measurements with uncertainty
truth/
nature



- model the uncertainty: probabilistic model, distribution functions
- make inference about truth: point estimation, hypothesis test, confidence interval

more generally:

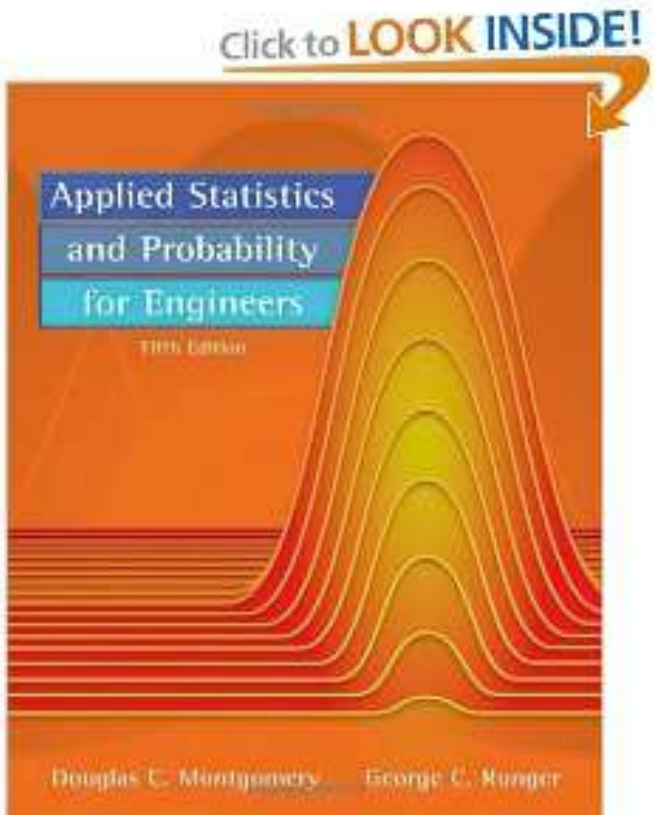
- make inference about two samples, and multiple groups of samples
- interpret input-response: linear regression
  - example: how voltage and current are related

$$V = RI + n$$

Course will cover

- Review of probability theory and random variables (2027)
- Descriptive statistics
- Statistical estimation and sampling distributions
- Statistical confidence intervals of a single population
- Test of hypothesis of a single population
- Comparing two populations
- The analysis of variance
- Simple linear regression
- Multiple linear regression

Textbook



Applied Statistics and Probability for Engineers, 5th Edition, Douglas C. Montgomery, George C. Runger, Wiley

Course Logistics

- Instructor: Prof. Yao Xie, Office: Groseclose 339, Email: yao.xie@isye.gatech.edu
- Instructor office hour: M 4:30-5:30pm, Groseclose 339
- Schedule: MW, 8:05-9:25pm, Instr Center 209
- TA: Caglar Caglayan, Office: Main Building 321, Email: ccaglayan6@mail.gatech.edu
- TA office hour: W 12:00-1:00pm, Office: Main Building 321
- Homework: out M after class, due next W in class
- Grading: Class Attendance 3%, Submitting Teaching Evaluation 2%, Homework - 15%, Computer Example 1 - 10%, Computer Exam 2 - 10%, Midterm 1 - 15%, Midterm 2 - 15%, Final - 30%
- Midterms: 9/23, 11/4 in class
- Computer Exams: 10/21, 11/13 in class
- Final: 12/11 8-10:50am, in classroom