# Probability and Statistics for Engineers
# IndE 315

## Unit 3: Statistical Inference

Nov. 5, 2010

# Announcements

Guest lecturer

- Julie Medero, MSEE (Dec.)
  - researcher in Electrical Engineering
  - studying statistical language processing
  - computer science and linguistics background

- topic:  estimators, bias of estimators, MVUE
  - material covered will be on Exam 3 (7-1,7-2,7-3)

# Today's class topics

- **Random samples**
- Notation summary
- What is a statistic?
- Point estimators
- Bias and variance of point estimators

- Class exercise
- Next assignments

# Reading Q3: Interesting concepts

"I thought random sampling was an interesting and challenging part of this reading. This is because randomization is always a hard task when trying to come up with a study. In a previous statistics class in high school, our teacher emphasized a lot about how important it was to design experiments that truly randomized the sampling so that a critic could not cry 'bias' to the data".– *student response*

# Reading Q2: example of random sample

Would a random sample for height best include both male and female students or only one gender?

- "I would only select students from one gender because the probability distribution of heights is probably going to be different among male and female students."
- "A class of engineering students includes both males and females. In order to successfully represent this population, the sample must include both male and females." – *student responses*

# Random samples

What is a random sample?

- All points in the population equally likely to be chosen
- Individual picks for the sample will have the same distribution and be independent from each other
- Distribution does *not* have to be unimodal

# Today's class topics

- Random samples
- **Notation summary**
- What is a statistic?
- Point estimators
- Bias and variance of point estimators

- Class exercise
- Next assignments

# Confusing concepts from readings

- "A random sample consists of $X_1$, $X_2$, $X_3$... $X_n$ (random variables.) Are these individual observations, samples (each with its own $\bar{x}$), populations, or sample means?"

# Notation summary

- X: A to-be-picked sample of *n* points (random variable)
- $X_i$: A to-be-picked sample point (random variable)
- $\overline{X}$ : The mean of a to-be-picked sample (random variable)

- *x*: A specific sample of *n* points (realization)
- $x_i$: A specific data point that's been picked (realization)
- $\overline{x}$ : The mean of a specific set of data points (realization)

- $\hat{\mu}$ : An estimator's guess of the population mean
- $\mu$ : The actual population mean

# Today's class topics

- Random samples
- Notation explanations
- **What is a statistic?**
- Point estimators
- Bias and variance of point estimators

- Class exercise – estimators
- next assignments

# Reading Q1: what is a statistic?

Is $\bar{x}$, s and normalized z a statistic?

- "a) The first is a statistic because neither the mean, std, or population portion is involved. b) Yes, same as the first reason. c) No, the population mean is involved. "

– *student response*

# Confusing concepts from readings

- "I find the definition of what is a statistic and what constitutes one to still be a little confusing."
- "I am confused about statistics I thought s and σ were the same thing. Why can you use s in statistics but not σ?"
- "Is there any other parameter that can be called a statistic? Is there any other definition of statistic different than that?"

# Statistics

What is a statistic?

- Updated definition: “A statistic is defined as any function of the sample data that does not contain unknown parameters.” (Montgomery)
- Don’t need to know anything about the population
- Example:
  - s is the sample standard deviation (statistic)
  - σ is the population standard deviation (parameter)

# Today's class topics

- Random samples
- Notation summary
- What is a statistic?
- **Point estimators**
- Bias and variance of point estimators

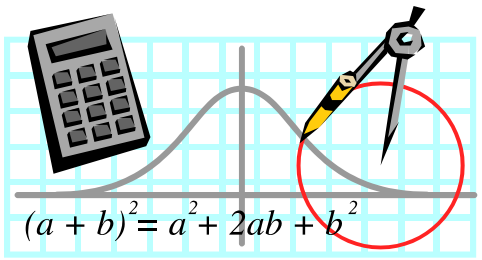
- Class exercise
- Next assignments

# Confusing concepts from readings

- “I found the concept of a point estimate to be confusing. I didn't understand what the purpose of a point estimator is and why they are necessary. Are they just another name for a parameter?”
- “The relationships between a point estimator, point estimate, statistic and random variable was confusing. Section 7-1 didn't do a great job of explaining these definitions. More examples would have been useful to drive home the point.”

# Point estimators

- A single value to estimate a population parameter
- Will generally use a statistic
- Point estimator: method of estimation
- Point estimate: from a specific sample



# Today's class topics

- Random samples
- Notation summary
- What is a statistic?
- Point estimators
- **Bias and variance point estimators**

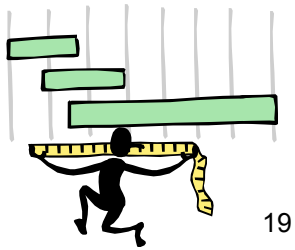
- Class exercise
- Next assignments

# Confusing concepts from readings

- “I understand the difference between the biased and unbiased estimator and what they are equal to. But why? How would we use them in real life?”
- “I’m very confused why a biased estimator would ever be a better estimate of the true value, the book tries to describe a situation where this is true but it was very hard to understand.”
- “Section 7.3 gives several examples of unbiased estimators, but an example of a biased estimators, for contrast, would have been helpful”

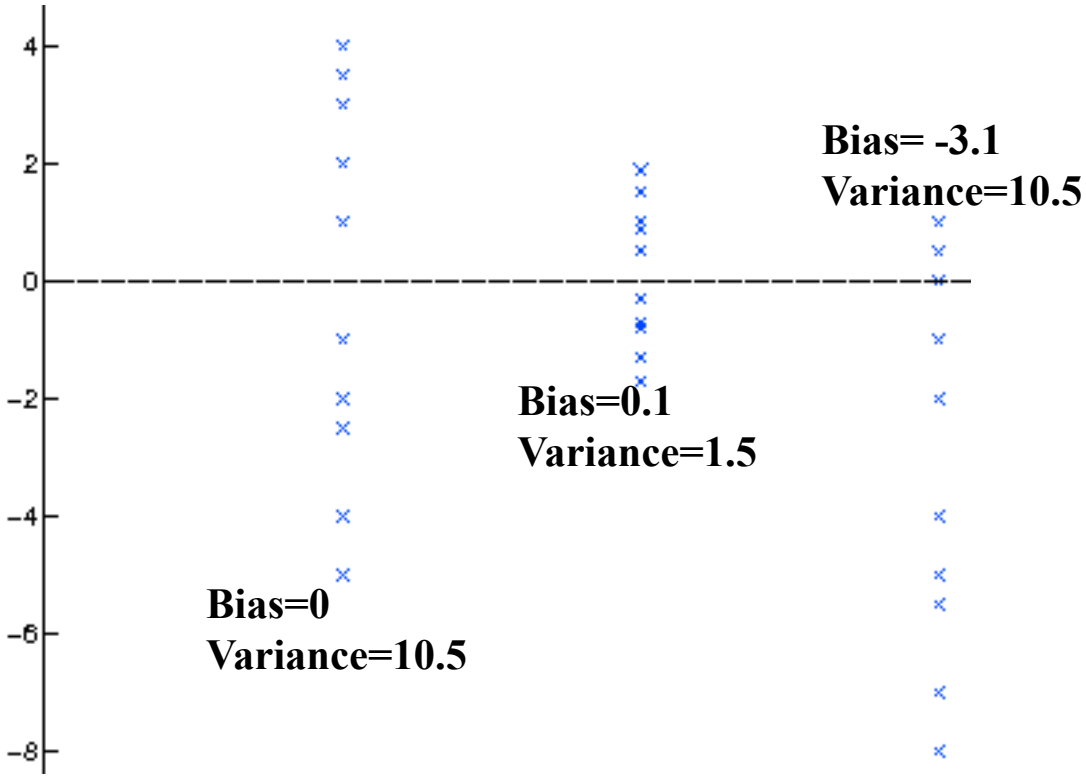
# Bias and variance of point estimators

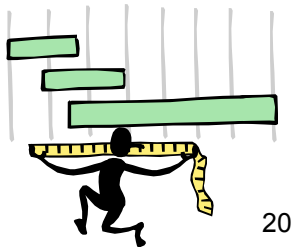
- Bias: difference between expected value of estimator and true parameter value
- Unbiased estimator = zero bias
- Variance: how much estimates will vary from one sample to another
- MVUE: minimum variance unbiased estimator



# Bias and variance of point estimators

Example: Estimators of θ (actual value=0)



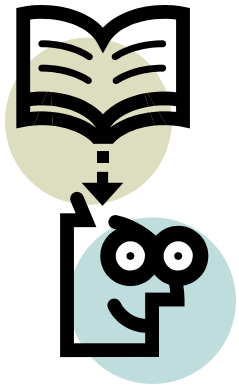
# Today's class topics

- Random samples
- Notation explanations
- What is a statistic?
- Point estimators
- Bias and variance point estimators

- **Class exercise – estimators**
- next assignments

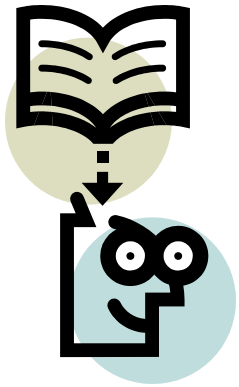
# In-class exercise

- There are *m* balls in a bin, numbered 1 through *m*.
- We get to pick 5 balls from the bin
- Task: Guess the number of balls



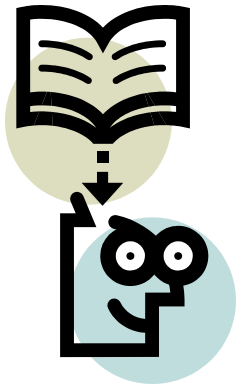
# In-class exercise

- What sort of distribution is this?
- What parameter are we estimating?



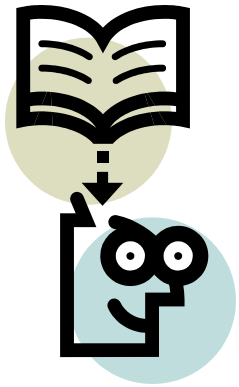
# In-class exercise

- What sort of distribution is this?
  Discrete uniform distribution
- What parameter are we estimating?
  Maximum value of distribution



# In-class exercise

- Come up with 3 estimators
- Calculate each estimator for each data sample
- Plot results



# Today's class topics

- Random samples
- Notation summary
- What is a statistic?
- Point estimators
- Bias and variance point estimators

- Class exercise
- Next assignments

# Next assignments

- HW #5 due on Monday 11/8
- Quiz #5 on Monday 11/8 (a problem like the homework)