

1 Overview and Introduction

Lecture 1:

- Introductions
- Syllabus
- Course
 1. Probability Concepts (This semester)
 - Basic Probability
 - Discrete Probability Distributions
 - Continuous Probability Distributions
 - Multivariate Probability Distributions
 - Transformations of Random Variables
 2. Inferential Statistics (Next Semester)
 - Central Limit Theorem
 - Estimation
 - Hypothesis Testing
 - Evaluating Estimators

2 Statistics Concepts

2.1 Why do we do Statistics?

- When we use statistics, we usually want to know something about a population
 - UCONN students
 - CT residents
 - US population
- But getting this information is not easy
 - Too expensive
 - Too time consuming
- So we take a small sample of the population and use the characteristics of the sample as an approximation of the desired characteristics of the population (This is called inferential statistics)

- In order to understand what our sample tells us about the population of interest, we need to know about the probability *Distribution* of our sample
- Defininton: A probability *Distribution* is is a phrase we use to refer to how likely different possible outcomes of a sample are in relation to other possible outcomes

2.2 What kinds of things do we want to know?

- Usually we want to know about an attribute that every member of the population has
 - Height
 - Gender
 - Weight
 - Residency
- Each of these characteristics is called a *variable*
 - In general, a variable is a value that is unknown
 - For our purposes, a variable is some attribute of interest that every memeber of the population has

2.3 Parameter Vs. Statistic

- When we are interested in a population we often take a sample of the population (a small subset of the population), and use our measurements of the distirbution of the sample to estimate the measurements of the population.
- because we are typically talking about the same kind of measurement of both populations and samples, we use the following definitnons:
 - Defininton: A *Parameter* is a measurement of a population. This is often unknow to us, and always will be. that is why we estimate it.
 - Defininton: A *Statistic* is a measurement of a sample. This is a value we usually know (after collecting a sample) and use it to measure the parameter of interest