Chapter 6 - Lecture 1
Statistics and their distribution

Yuan Huang

January 9th, 2013

Overview of Chapter 6

Chapter 6 : Statistics and Sampling Distributions

1. 6.1 Statistics and Their Distributions
   - Introduce definitions.
2. 6.3 The Distribution of a Linear Combination
   - Tools for proofs.
3. 6.2 The Distribution of the Sample Mean
   - Important properties for sample mean
4. 6.4 Distribution Based on a Normal Random Sample
   - Introduce several important distributions.

- Random variables X_1 and X_2, ...X_n, where n denotes the **sample size**.

- Let say, I have random variables .... So if I get a random sample of size 3, $n = 3$.
- If in observation, X_1 takes value 1, then denote it as $x_1 = 1$. and let's say the values are $x_1 = 2, x_2 = 1, x_3 = 1$.

- What is the difference between X_1 and x_1 ?

Random variable and Observation

- X denotes a **random variable** which is unknown.
- x denotes the **observed value** of the random variable which is known and might be different from sample to sample.

Random variable and Observation

- Random variables have an uncertainty for their values.
- That means two things:
  - You do not know what the value of random variables are until you actually see the observed values in the sample.
  - Any value depending on random variables will be expected to differ from sample to sample.

Random sample = iid

- What a random sample is?
  - All random variables are independent
  - All random variables come from the same distribution **(as from the population)**, that is they are identically distributed

- In short, we write **iid**, which means independent and identically distributed
- Intuitively, random sample is the sample that is representative of the population.

Random sample = iid

- What a random sample is?
  - All random variables are independent
  - All random variables come from the same distribution **(as from the population)**, that is they are identically distributed

- In short, we write **iid**, which means independent and identically distributed
- Intuitively, random sample is the sample that is representative of the population.

Randomness of Sample is always tricky. In this course we just assume it unless otherwise explicitly specified.

Examples

Do you know π ?

Do you know π ? How many digits you can tell?

Examples

Do you know π ? How many digits you can tell?

Required Readings: Are the Digits of π an Independent and Identically Distributed Sequence?

Not example:

1. Convenient Sample: to select sample that is easy to get
    1. Select your family members;
    2. Select your friends and classmates;
    3. Select people you know on Facebook and twitter;
2. Data snooping: Select the part of sample that you prefer, ignore the rest
    1. Learn more about the data snooping: http://data-snooping.martinsewell.com/

Definition of Statistic

- We call **statistic** any quantity whose value can be calculated from sample data. That means a statistic is a function of random variables from our random sample $X_1, \ldots, X_n$.

- Do you think a statistic should be denoted with an upper case letter or a lower case letter?

Examples

A statistic is also a random variable.

1. Sample Mean: $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$
2. Sample Variance: $S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$
3. Other examples: sample quantiles, sample standard deviation, etc.

Sampling Distribution

- The probability distribution of a statistic is called **sampling distribution** to emphasize the fact that it describes how the statistic varies from one random sample to another.

How we find the sampling distribution of a statistic

- Using Probability Rules. (e.g 6.2)
- Simulation Experiments.
- Using known theorems (which is considered an extension of the first case). (section 6.3,6.4)

Example 6.2 page 282

Example 1: Suppose (X_1, X_2) is a random sample of size 2 and each of them has the following probability distribution:

Table: Probability distribution of $X_1(X_2)$

x	40	45	50
$p(x)$	0.2	0.3	0.5

- What is the probability distribution of $\bar{X} = \frac{X_1+X_2}{2}$?
- What is the probability distribution of S^2?

Exercises

- Section 6.1 page 290
  - Hw1(to be continued) 2, 3