

# Lab 05

JIN, ICK HOON

2009. 07. 21.

1. How do we know that the mean and median are both unbiased estimators for the population mean?

The center of both distributions is the center of the parent population.

2. Why do we prefer the mean over the median for this parent population?

The mean has a smaller standard deviation than the median.

3. What is relationship between the standard deviation of the parent population and the standard deviation of the sample mean?

The sample mean standard deviation is  $1/\sqrt{n}$  times the sd of the parent population. The sample mean standard deviation is  $1/\sqrt{n}$  times the sd of the parent population.

4. When do we prefer the median over the mean as an estimate of the center of a distribution?

If a dataset is highly skewed or has outliers, the median is a better estimator for the center.

5. What would the parent population need to look like for the mean not to be the preferred estimator for the population mean?

If the parent population is highly skewed or has numerous outliers, the sample median will be a better estimator for the population mean.

6. If the parent population is normal, what is the distribution of the sample mean if we take a simple random sample?

As long as the parent population is normal, the distribution of the sample mean will be exactly normal - even if we take samples of size 1.

7. If the parent population is uniform, how large of a sample do we need for the distribution of the sample mean to be approximately normal?

Since the uniform distribution is symmetric, a sample of only 10 will make the distribution of the sample mean approximately normal.

8. If the parent population is skewed, how large of a sample do we need for the distribution of the sample mean to be approximately normal?1

The more skewed the parent population, the larger sample is needed. In general, the Central Limit Theorem says if the sample size is sufficiently large, the distribution of the sample mean will be approximately normal no matter what the shape of the parent population as long as the it has a finite mean and variance. 30 is usually considered large enough, but sometimes a larger sample is needed.

9. If we have categorical data, what do we need for the distribution of the sample proportion to be approximately normal?

For categorical data, we must have both  $np \geq 10$  and  $n(1 - p) \geq 10$  for the distribution of the sample proportion to be approximately normal, where  $n$  is the population proportion.