Review of Session 3

- Probability quantifies uncertainty
- Probability in this class based on relative frequency
- Probability theory used to create probability distributions
- Many traits approximate a normal distribution
- When variables is standardized, probabilities can be looked up in table.

Role of Probability in Course

- Draw a sample
- Probability allows us to quantify uncertainty concerning our estimates which is expressed as confidence intervals and tests of hypotheses.
Probability Distribution for Sample Statistics

- Up until now
  - We standardized scores, and
  - Made probability assessment about the scores
  - E.g., What is the probability that a person drawn at random will score above 120 on an I.Q. test.
- In statistical applications
  - The probability distribution is not for a score of a person
  - It is for statistics of samples
- This requires us to imagine a statistic computed on
  - All possible samples of a given size
  - The collection of statistics called sampling distribution.

A Tiny Sampling Example

The following example is unrealistic but hopefully informative.
In this example we look at all samples of size 2 taken from population of size 4. The story line for the example is number of children.

<table>
<thead>
<tr>
<th>Populations:</th>
<th>Means for All Samples of Size 2 (With Replacement)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mother</td>
<td># of Children</td>
</tr>
<tr>
<td>Alice</td>
<td>2</td>
</tr>
<tr>
<td>Betty</td>
<td>4</td>
</tr>
<tr>
<td>Carol</td>
<td>6</td>
</tr>
<tr>
<td>Dee</td>
<td>8</td>
</tr>
</tbody>
</table>

$$\mu = \frac{20}{4} = 5$$  
$$\sigma = \frac{\sqrt{80}}{4} = 2.236$$

Central Limit Theorem

- Means are approximately normally distribute, and the approximation improves as $n$ increases.
Central Limit Theorem

- Means are approximately normally distribute, and the approximation improves as \( n \) increases.
- The mean of the means, \( \mu_x \), is equal to \( \mu \) and is called the Expected Value.
- The standard deviation of the means, \( \sigma_x \), is equal to \( \frac{\sigma}{\sqrt{n}} \) and is called the Standard Error of the Means.

Role of Probability in Course

\[ \bar{X} \rightarrow \mu \]

- \( \bar{X} \) measures the error of the estimate.
- Probability allows us to quantify uncertainty.
- This is used to create confidence intervals and conduct hypothesis tests.

Parametric Approach Uses Theoretical Distribution

- When \( \sigma \) is known, the normal distribution is used to obtain probabilities.
- When \( \sigma \) is unknown, the t distribution is used.
Parametric Approach Uses Theoretical Distribution

- When $\sigma$ is known the normal distribution is used

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

The value of $Z$ can then be compared to a normal distribution during the construction of a confidence interval or when conducting a hypothesis test.

Parametric Approach Uses Theoretical Distribution

- When $\sigma$ is unknown the $t$-distribution is used

$$t = \frac{\bar{X} - \mu}{s / \sqrt{n}}$$

The value of $t$ can then be compared to a $t$ distribution, which has degrees of freedom equal to $n-1$. This can be used during the construction of a confidence interval or when conducting a hypothesis test.

Bootstrapping Uses An Empirical Distribution

- A sample is drawn from a population without replacement.
- The mean of this sample is used to estimate the mean of the population.
- A large number of samples of size $n$
  - Are drawn with replacement
  - From the sample
- An empirical distribution is created from the sample
- Probability estimates are obtained from the empirical distribution.