Central Limit Theorem - Review

Review Component

- **Sample Size Estimation** will be a learning outcome for this course, as part of the "Inference Procedures" Section.
- In preparation, a review of the **Central Limit Theorem** is advisable

Central Limit Theorem - Review

Statistical Estimation

Sample Statistics		Population Parameter
$\overline{x}$	$\longrightarrow$	μ
s	$\longrightarrow$	σ

Use sample statistics to estimate the population parameter.

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}}$$

Figure:

Central Limit Theorem - Review

Statistical Estimation

Sample Statistics		**Population Parameter**
$\overline{x}$	$\longrightarrow$	μ
s	$\longrightarrow$	σ

Use sample statistics to estimate the population parameter.

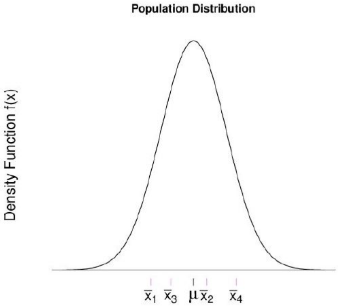
$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}}$$

Central Limit Theorem - Review

Sampling distribution of $\overline{x}$

If we take repeated samples from a population and measure the mean $\overline{x}$ of each sample, we see that most of the sample means obtained are different from each other, but clustered around the population mean μ.



Central Limit Theorem - Review

The sampling distribution of the mean

The sample means $\overline{x}_1, \overline{x}_2, \overline{x}_3 \ldots$ behave like a variable and hence follow a statistical distribution. The probability distribution of the sample means is called the **sampling distribution** of the mean.

The sampling distribution of $\overline{x}$-s itself has a mean and a standard deviation.

Central Limit Theorem - Review

The sampling distribution of the mean

The sampling distribution of the mean is always **Normal**, regardless of the distribution of the population.

The sampling distribution of the mean has itself a mean $\mu_{\overline{X}}$ which equal the population mean μ.

The sampling distribution of the mean has itself a standard deviation, called standard error $\sigma_{\overline{X}}$ (so we don't confuse it with the population standard deviation, σ).

Central Limit Theorem - Review

The Standard Error

The standard error of the sample mean $\overline{x}$ is denoted as $SE(\overline{x})$.

$$SE(\overline{x}) = \frac{\sigma}{\sqrt{n}}$$

This means that if we have a sample of size n=100, the standard error of the sample mean SE($\overline{x}$)=$\frac{\sigma}{\sqrt{100}}$= $\frac{\sigma}{10}$ is 10 times smaller than the standard deviation of the population, σ.

Central Limit Theorem - Review

Conclusions from the sampling experiment

The sampling distribution of the mean $\overline{x}$

- follows a Normal model, i.e. $\overline{x} \sim N$.
- with the same mean, μ, as the population
- but with a reduced standard deviation, $\frac{\sigma}{\sqrt{n}}$, called the standard error of the mean, instead of σ, the population standard deviation

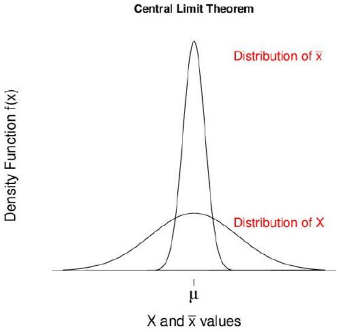
Central Limit Theorem - Review

The sampling distribution of the mean

If $\quad X \sim N(\mu, \sigma^2)$ is a good model for the population
then $\overline{x} \sim N(\mu, \frac{\sigma^2}{n})$ is a good model for the sample means based on n observations

Central Limit Theorem - Review

The Normal model for X and $\overline{x}$



Central Limit Theorem - Review

Central Limit Theorem

Why is the normal density so important?
It describes the behavior of the sample means regardless of the shape of the population from which we sample.

Central Limit Theorem:
When random samples of size **n** are drawn from any distribution, having mean μ and standard deviation σ, then if **n** is large, the sample means tend to form a normal distribution, with mean μ and standard error $\frac{\sigma}{\sqrt{n}}$.