7.1 Introduction

Statistical Inference

Process of drawing conclusions about a population based on information obtained from a sample.

Recall:

Descriptive measures of a population are called

Example:

Descriptive measures calculated from a sample are called

Example:
7.2 **Sampling distribution of a Statistics:**

We would consider the following (inference):

1. Estimation.
2. Hypothesis Testing.

Before talking about inference, the dist'n of the statistics have to be known:

- What is the dist'n of $\bar{X}$?
- What is the dist'n of $\hat{p}$?

The probability dist'n for a statistic is called the.
7.3 Dist’n of Sample Mean and the Central Limit Theorem

Sampling Distribution of $\bar{X}$?

From repeated sampling:

Sample 1 of size $n$ $\rightarrow$ $\bar{x}_1$

Sample 2 of size $n$ $\rightarrow$ $\bar{x}_2$

: $\rightarrow$ :

Sample $m$ of size $n$ $\rightarrow$ $\bar{x}_m$

Questions:

What is the dist'n of $\bar{X}$?
What is the mean of $\bar{X}$?
What is the SD of $\bar{X}$?
Central Limit Theorem (CLT)

If relatively large samples of size $n$ are drawn from **any** population,

- If the popul'n dist'n is normal, the sampling dist'n of $\bar{X}$ will be

- If the population dist'n is non-normal, the sampling dist'n of $\bar{X}$ will be, for large samples ($n \geq 30$),

Mean and SD of $\bar{X}$ ?
The mean of $\bar{X}$ is equal to the mean of the original popul'n. That is:

$$\mu_{\bar{X}} =$$

The variance of $\bar{X}$ is equal to the popul'n variance divided by the sample size. That is:

$$\sigma^2_{\bar{X}} =$$

Hence, the **standard deviation for** $\bar{X}$ **is**:

$$\sigma_{\bar{X}} =$$

The SD of a statistic is called the

$\sigma_{\bar{X}}$ : is called the standard error of the mean.
Example 1

The amount of sulfur in the daily emissions from a power plant has a normal dist’n with a mean of 94 pounds and a SD of 22 pounds.

a. What is the mean and SD of the parent popul'n?

\[
\begin{align*}
\mu &= \text{ } \\
\sigma &= \text{ }
\end{align*}
\]
b. If 5 days are randomly selected and the average sulfur emission is calculated, what is the mean and SD of the sample mean?

Mean: \( \mu_{\bar{X}} = \)

SD: \( \sigma_{\bar{X}} = \)

\( \bar{X} \sim \)
c. Plot the two distributions found in (a) original population and (b) average of 5 observations.
d. Find the probability that on a randomly chosen day, sulfur emissions are more than 100 pounds.

\[ P[X > 100] = ? \]

\[
P[ X > 100] = P \left[ Z > \right]
\]

\[ = P \left[ Z > 0.27 \right] \]

\[ = 0.3936 \]
e. Find the probability that, if five days are randomly selected, their mean emission exceeds 100 pounds.

\[ P(\bar{X} > 100) = ? \]

In this case, the transformation needed to standardized the normal RV is:

\[ Z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \]
\[ P \left[ Z > 0.61 \right] = P \left[ Z > \right] \]

\[ = P \left[ Z > 0.61 \right] \]

\[ = 0.2709 \]
Example 2

The mean and SD of the strength of a packaging material are 55 and 7 pounds, respectively. If 45 specimens of this material are tested,

a. Is it reasonable to assume a normal dist’n for the sample mean $\bar{X}$? Why or why not?

\[
\bar{X} \sim N(\quad )
\]
b. Find the probability that the sample mean strength \( \bar{X} \) will be between 54 and 56 pounds?

\[
P[ ] = ?
\]

\[
= P \left[ \begin{array}{c}
< Z < \\
\end{array} \right]
\]

\[
= 0.8315 - 0.1685
\]

\[
= 0.6630
\]