Normally, we have a population of interest with unknown parameters. This is the underlying distribution or the underlying population.

From this underlying distribution, we draw a sample and calculate a statistic to estimate one of the population parameters.

As an example, \( \bar{x} \) estimates \( \mu \) and \( s^2 \) estimates \( \sigma^2 \). How do we decide which statistic should be used - this is based on something that we call the sampling distribution of a statistic...

Draw an example from Class
6.1 What is a Sampling Distribution and 6.2: Properties of a good estimator:

How do we choose between two estimators? Example; sample mean and the sample median to estimate population mean?

FACT: Neither the mean, nor the median will always fall closer to the population mean; hence we cannot decide on the basis of a single statistic which is the better estimator. (FIGURE 6.1, page 279) We have to base our criteria for choosing an estimator on the sampling distribution of the statistic...

Suppose we draw every possible sample of size n from the underlying population and for each sample, we calculate a statistic. Now note that these statistics are random variables and therefore must have a probability distribution. Collectively, these statistics have a distribution called the **sampling distribution** of the statistic.
Look at a few examples in the book Figure 6.5 page 283 and Figure 6.3, page 280. NOTE how in figure 6.5 the distribution of the sample mean has a smaller spread.

6.2 - Properties of a Sampling Distribution:

Now when we use a single statistic - or a single point - to estimate a parameter, the resulting estimate is called a point estimate. Now any point estimate $T$ has a sampling distribution and will thus have a mean and a variance. Let $E(T) = \mu_T$ and let $\sigma_T$ denote the standard deviation of $T$. Then $\sigma_T$ is often called the standard error of $T$.

6.3 Sampling Distribution of the sample mean

Suppose the statistic of interest is $\bar{x}$. It can be shown that:

1) The mean of the sampling distribution of $\bar{x}$ is $\mu$. That is
\[ \mu = E(\bar{x}) , \]

where \( \mu \) is the mean of the underlying distribution. What does this mean? This property is often called unbiasedness.

In general, let \( T \) be a point estimate for the parameter \( \theta \). Then \( T \) is said to be an unbiased estimate of \( \theta \) if the mean of the sampling distribution of \( T \) is equal to \( \theta \), in other words, \( E(T) = \theta \).
It turns out that both the sample mean and the median are unbiased for the population mean. So now how do we decide on which estimate to use. Okay now if we have to choose between two unbiased estimators, then we pick the one that has a smaller standard error.

Figures 6.6 and 6.7 pp 285

2) The variance of the sampling distribution of $\bar{x}$ is given by

$$ \text{Var}(\bar{x}) = \frac{\sigma^2}{n} $$

where $\sigma^2$ is the variance of the underlying distribution, and $n$ is the sample size. The variance of the median is higher.
Central limit Theorem

For \( n \) sufficiently large (usually greater than 25 or 30), the sampling distribution of \( \bar{x} \) will be approximately normal for any underlying distribution.

\[ \bar{x} \sim N(\mu, \sigma^2/n) \]

If the sample is drawn from a normal distribution, then the distribution of \( \bar{x} \) will be exactly normal regardless of the sample size.
LOOK AT FIGURE 6.9 AND FIGURE 6.10
pp 289 and 291
Example:

a) EG Corp. produces light bulbs with an average lifetime of 450 hours and a standard deviation of 30 hours. If a random sample of 36 bulbs is taken off the assembly line, what is the probability that the average life of the sample is greater than 460 hours?

\[ E(\bar{x}) = 450 \]

standard deviation of \( \bar{x} \) = \( \frac{30}{\sqrt{6}} = 5 \)

Hence \( P(\bar{x} > 460) = P(Z > \ldots) = \ldots \).
b) It is known that the average electricity bill for home owners in the city of Miami is $143.00 with a standard deviation of $50.00. In a random sample of 49 such homeowners, what is the probability that \( \bar{x} \), the sample mean is between $130.00 and $156.00?

Exercises: 6.11, 61.3, 6.21, 6.27, 6.31, 6.35, 6-37, 6.41 (a), (b), 6.43.