INFERENCES ABOUT $\mu$

- One of the major objectives of statistics is to make inferences about the distribution of the elements in a population based on information contained in a sample.

- Numerical summaries that characterize the population distribution are called parameters.

- The population mean $\mu$ and population variance $\sigma^2$ are two important parameters. Others are median, range, mode, etc.
• Methods for **making inferences** are basically designed to answer one of two types of questions:

(a) Approximately what is the value of the parameter?
    or
(b) Is the value of the parameter less than (say) 6?

• Statisticians answer the first question by **estimating** the parameter using the sample.

• The second case might require a **test of a hypothesis**.
Point and Interval Estimation of $\mu$ when $\sigma$ is known and $n$ is large

- Point estimation of $\mu$ does not require that $\sigma$ be known or a large $n$
- The point estimate of $\mu$ is the sample mean $\bar{y}$
- Point estimates by themselves do not tell how much $\bar{y}$ might differ from $\mu$, that is, the accuracy or precision of the estimate
- A measure of accuracy is the difference between sample mean $\bar{y}$ and population mean $\mu$ is called sampling error
A Confidence Interval for $\mu$

- An interval estimate, called a confidence interval, incorporates information about the amount of sampling error in $\bar{y}$
- A confidence interval for $\mu$ takes the form
  $$(\bar{y} - E, \bar{y} + E),$$
  for a number $E$
- An associated number called the confidence coefficient helps assess how likely it is for $\mu$ to be in the interval.
• To derive the specific form of the confidence interval for \( \mu \), for the case when \( \sigma \) is known and \( n \) is large, the CLT result must be used.

• By the CLT, \( Z = (\bar{Y} - \mu)/\sigma_{\bar{Y}} \) is approximately \( \sim N(0, 1) \).

• Let \( Z \) have a \( N(0, 1) \) distribution (exactly).

• Let \( z_{\alpha/2} \) denote the \( 1 - \alpha/2 \) quantile of the standard normal distribution, for a given number \( \alpha \), \( 0 < \alpha < 1 \).

• Then the following probability statement is true:

\[
P \left( -z_{\alpha/2} \leq \frac{\bar{Y} - \mu}{\sigma_{\bar{Y}}} \leq z_{\alpha/2} \right) = 1 - \alpha
\]
Manipulating the inequalities, without changing values, we have

\[ P(\sigma \bar{y} z_{\alpha/2} \geq \mu - \bar{Y} \geq -\sigma \bar{y} z_{\alpha/2}) = 1 - \alpha \]

\[ P(\bar{Y} - \sigma \bar{y} z_{\alpha/2} \leq \mu \leq \bar{Y} + \sigma \bar{y} z_{\alpha/2}) = 1 - \alpha \]

If, for example, \( \alpha = 0.05 \) then

\[ P(\bar{Y} - \frac{\sigma}{\sqrt{n}} z_{0.025} \leq \mu \leq \bar{Y} + \frac{\sigma}{\sqrt{n}} z_{0.025}) = 0.95 \]

and, since \( z_{0.025} = 1.96 \),

\[ P(\bar{Y} - 1.96 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{Y} + 1.96 \frac{\sigma}{\sqrt{n}}) = 0.95 \]
• Using this statement, a \((1-\alpha)100\%\) confidence interval for \(\mu\), with confidence coefficient \(1 - \alpha\) can be calculated using a random sample of data \(y_1, y_2, \ldots, y_n\)

• It is usually written in the form

\[
\left( \bar{y} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2}, \; \bar{y} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \right).
\]

• Example: Suppose \(n = 36\), \(\sigma = 12\), \(\bar{y} = 24.8\)

And for \(\alpha = 0.05\), \(z_{\alpha/2} \equiv z_{0.025} = 1.96\)
• Thus a 95% C.I. for $\mu$ is

$$
\left(24.8 - \frac{12}{\sqrt{36}} \times 1.96, \quad 24.8 + \frac{12}{\sqrt{36}} \times 1.96\right)
$$

• The interval for $\mu$ is: (20.88, 28.72) with confidence coefficient 0.95.

• We might say that we are 95% confident that the population mean $\mu$ is between 20.88 and 28.72.

• But what do we actually mean when we say that we are 95% confident?
Interpretation of a Confidence Interval

- Before the sample is drawn, the probability is \((1 - \alpha)\) that the random interval \((\bar{Y} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2}, \bar{Y} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2})\) will contain \(\mu\) (because \(\bar{Y}\) is a random variable).
- However, once the sample is drawn, and \(\bar{y}\) is calculated, the interval ceases to be random. It is a numerical interval \((\bar{y} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2}, \bar{y} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2})\), calculated specifically for the drawn sample; thus we cannot associate a probability with it.
- We may say that the process which led us to this interval will, on the average, produce an interval containing \(\mu\), \(100(1 - \alpha)\)% of the time.
• It is not correct to say that the interval \((\bar{y} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2}, \bar{y} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2})\) contains \(\mu\) with a specified probability.

• If the process of sampling is repeated many times and 100(1-\(\alpha\))% confidence intervals calculated for each sample, then we are confident that 100(1-\(\alpha\))% of those intervals will contain \(\mu\).

• This idea is illustrated in the picture below:
Notes:

- In the textbook Example 5.1 uses $s$, the sample standard deviation, in place of $\sigma$. Of course, $s$ is a point estimate of $\sigma$. Thus it contains sampling error. However, when $n$ is large, $\sigma$ is sometimes approximated by $s$ because $\sigma$ is not known.

- As the confidence coefficient increases the confidence interval becomes wider. A wider interval estimates $\mu$ less precisely. Thus a 99% confidence interval is less accurate than a 95% confidence interval but one has more confidence that $\mu$ is in the first interval.

- Also note that increasing the sample size $n$ results in a narrower interval (more accurate estimate) for the same confidence coefficient.
Exercise 5.8:
The caffeine content (in mg) was examined for a random sample of 50 cups of black coffee dispensed by a new machine. The mean and standard deviation were 110 mg and 7.1 mg, respectively. Use these data to construct a 98% confidence interval for $\mu$, the mean caffeine content for cups dispensed by the machine.

As the sample size is large we can use the CLT and also approximate $\sigma$, the sample standard deviation of the population by $s = 7.1$

$$n = 50 \quad \bar{y} = 110, \quad \sigma = 7.1, \quad \alpha = .02, \quad z_{0.01} = 2.33$$
\[
\left[ \bar{y} - z_{0.01} \frac{\sigma}{\sqrt{n}}, \ \bar{y} + z_{0.01} \frac{\sigma}{\sqrt{n}} \right]
\]

\[
\left[ 110 - (2.33) \left( \frac{7.1}{\sqrt{50}} \right), \ 110 + (2.33) \left( \frac{7.1}{\sqrt{50}} \right) \right]
\]

(107.66, 112.34)

We are 98% confident that the mean caffeine content for cups dispensed by the machine is between 107.66 and 112.34 mg.
Choosing the Sample Size

- The width of a confidence interval is \((2 \times \sigma \times z_{\alpha/2})/\sqrt{n}\).
- It can be made smaller by changing to a larger \(\alpha\) or increasing sample size \(n\).
- Let us consider selection of \(n\) to achieve a desired width \(2 \cdot E\) for a fixed \(\alpha/\)
- We want \(E\) to at least equal to \(\sigma \times z_{\alpha/2}/\sqrt{n}\)
- Thus we need to select a sample so that

\[
n \geq \frac{(z_{\alpha/2})^2 \sigma^2}{E^2}\]
This will result in an interval no wider than
\((\bar{y} - E, \bar{y} + E)\).

**Example:**
Given \(\sigma = 12\), \(\alpha = 0.05\), \(z_{0.025} = 1.96\)
What sample size \(n\) will give an interval no wider that 5.6?
We set \(E = 2.8\) so

\[
n \geq \frac{(1.96)^2 \cdot 144}{(2.8)^2} = 70.56
\]

Thus the experimenter must choose a sample size \(n = 71\) at least.
Choosing the Sample Size (continued)

Would the sample size $n = 71$ ensure the width to be $\leq 5.6$ if the population has a larger variance? Would $n = 71$ be enough?

The answer is **NO** since the formula

$$\frac{(z_{\alpha/2})^2 \sigma^2}{E^2}$$

involves the population variance $\sigma^2$.

Precision in estimation depends on both $\alpha$ and $\sigma^2$. If the variance of population elements is very small, i.e., the elements are tightly clustered about the population mean $\mu$, then only a small sample is needed for the estimate $\bar{y}$ to be very near $\mu$.  

16
Statistical Tests for $\mu$

- Estimation (either point or interval estimation) was used to help answer a question like “Approximately what is the value of $\mu$?”

- The other kind of question mentioned earlier is “Is it likely that $\mu$ is less than (or greater than) the value $\mu_0$ (a pre-determined value)?”.

- A Test of Hypothesis is used to answer this kind of question. As in estimation, the sample mean $\bar{y}$ of a random sample of $n$ elements from the population is used to answer this question.
Every test of hypothesis features

(a) a **Null Hypothesis** $H_0$ which describes a characteristic of the population (it is believed to be) as it currently exists,

(b) a **Research Hypothesis** (or **Alternative Hypothesis**) $H_a$, which is a proposal about this characteristic, by the person(s) conducting the statistical study.

The idea is that the null hypothesis is presumed to hold unless there is overwhelming evidence in the data in support of the research hypothesis.
Example:

- To determine whether the mean yield per acre (in bushels), $\mu$, of a variety of soybeans increased the current year over the last two years when $\mu$ is believed to be 520 bushels per acre, the following might be tested.

$$H_0 : \mu \leq 520 \quad \text{vs.} \quad H_a : \mu > 520$$

- The fact that $\mu$ may equal a specific value is always included in the null hypothesis. The decision to state whether the data supports the research hypothesis or not is based on a quantity computed from the data called the test statistic.
• If \( \bar{y} \) is in the rejection region then reject \( H_0 \) and say the evidence favors \( H_a \).
• Basically, the test amounts to computing \( \bar{y} \) and looking at its value relative to \( \mu_0 \) and the \( \mu \) values in \( H_a \).
• For example, in the above example if \( \bar{y} < 520 \) we will say there is not sufficient evidence to reject \( H_0 \).
• Even if \( \bar{y} > 520 \), we might still say there is not sufficient evidence to reject \( H_0 \).
• This is ok so long as \( \bar{y} \) is not too much greater than 520.
• “How much is too large?” To decide this, use the probability distribution of \( \bar{Y} \). Begin by picking a small probability \( \alpha \) like \( \alpha = 0.05 \), or \( \alpha = 0.01 \), or \( \alpha = 0.001 \).
• Then reject $H_0$ only when the probability of obtaining a value of $\bar{y}$ larger than the observed value $\bar{y}$ is $\leq \alpha$ when $\mu$ is indeed $\leq 520$, i.e., when $H_0$ is true.

• That is, if $H_0$ is true, the chance of observing a sample that results in a $\bar{y}$ as large as the one calculated should be very small.
How to determine the rejection region

- By the CLT, $\bar{Y}$ is approximately a normal random variable.
- We will ignore the approximation and assume $\bar{Y} \sim N(\mu, \frac{\sigma^2}{n})$.
- Then, for a given $\alpha$,

$$P \left( \frac{\bar{Y} - \mu}{\sigma / \sqrt{n}} > z_\alpha \right) = \alpha$$

where $z_\alpha$ is the $1 - \alpha$ quantile of the standard normal distribution.
- So when $H_0 : \mu \leq \mu_0$ is true ,

$$P \left( \frac{\bar{Y} - \mu_0}{\sigma / \sqrt{n}} > z_\alpha \right) \leq \alpha$$
• That is, \( P(\bar{Y} > \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}) \leq \alpha \). This suggests that we reject \( H_0 : \mu \leq \mu_0 \) in favor of, say \( H_a : \mu > \mu_0 \) when

\[
\bar{y} > \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}
\]

(i.e., \( \bar{y} \) is too much when \( \bar{y} \) exceeds the number \( \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}} \))

• All possible \( \bar{y} \) values satisfying \( \bar{y} > \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}} \) constitute the rejection region.

• Instead of comparing \( \bar{y} \) to \( \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}} \), it is easier to first calculate

\[
z_c = \frac{\bar{y} - \mu_0}{\sigma / \sqrt{n}}
\]
• We see that comparing $\bar{y}$ to $\mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$ is the same as comparing $z_c$ to $z_\alpha$.
• That is, we reject the null hypothesis if $z_c > z_\alpha$ which is the same thing as doing so if $\bar{y} > \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$
• This quantity $z$ is called the test statistic and its value can be calculated using the data and the value $\mu_0$ specified in the null hypothesis $H_0$.
• The notation $z_c$ is used to denote the “computed value” of the test statistic, that is when we plug in the observed value of $\bar{y}$ and obtain a numerical value for $z$. 
Type I Error

- The $\alpha$ corresponds to the probability that the null hypothesis is rejected when actually it is true and is called the probability of committing a Type I error.
- Since the experimenter selects the value of $\alpha$ used in the test procedure, she is able to specify or control the Type I error rate or “how much” of this type of error is permitted in the testing procedure.
Example 5.5

Suppose from a sample of 36 1-acre plots, the yield of corn this year was measured and \( \bar{y} = 573 \) and \( s = 124 \) calculated. Can we conclude that the mean yield of corn for all farms exceeded 520 bushels/acre this year? Here we are going to assume that \( \sigma \) can be approximated by \( s \). Use \( \alpha = .025 \).

Solution:

Set-up five parts of the testing procedure:

1. \( H_0 : \mu \leq 520 \)
2. \( H_a : \mu > 520 \)
3. T.S.: \( z = \frac{\bar{y} - \mu_0}{\sigma / \sqrt{n}} \)
4. R.R: Reject $H_0 : \text{in favor of } H_a : \text{if } z > z_{0.025}$ 
   i.e. the R.R. is $z > 1.96$ since $z_{0.025} = 1.96$

5. Compute the observed value of the test statistic:
   $$z_c = \frac{573-520}{124/\sqrt{36}} = 2.56$$

Decision:
Since $z_c > 1.96$, $z_c$ is in the R.R. Thus $H_0 : \mu \leq 520$ is rejected in favor of the research hypothesis $H_a : \mu > 520$. It is concluded that the mean yield this year exceeds 520 bushels/acre.

Note that the above test procedure is equivalent to determining that the observed value for $\bar{y}$ lies more than 1.96 standard deviations above the mean $\mu_0 = 520$. 
Summary of Test Procedures: (assume $n$ large, and $\sigma$ known.)

Hypotheses:

Case 1: $H_0: \mu \leq \mu_0$ vs. $H_a: \mu > \mu_0$ (right-tailed test)
Case 2: $H_0: \mu \geq \mu_0$ vs. $H_a: \mu < \mu_0$ (left-tailed test)
Case 3: $H_0: \mu = \mu_0$ vs. $H_a: \mu \neq \mu_0$ (two-tailed test)

T.S: $z_c = \frac{\bar{y} - \mu_0}{\sigma/\sqrt{n}}$

R.R: For Type I error probability of $\alpha$:

Case 1: Reject $H_0$ if $z \geq z_\alpha$
Case 2: Reject $H_0$ if $z \leq -z_\alpha$
Case 3: Reject $H_0$ if $|z| \geq z_{\alpha/2}$
Decision: If the computed value of the test statistic, \( z_c \), is in the R.R., then we will reject the null hypothesis \( H_0 \) at the specified \( \alpha \) value. Otherwise, we say we fail to reject \( H_0 \) at the specified \( \alpha \) value.

**Example 5.6** A corporation maintains a large fleet of company cars for its salespeople. To check the average number of miles driven per month per car, a random sample of \( n = 40 \) cars is examined. The mean and standard deviation for the sample are 2,752 miles and 350 miles, respectively. Records for previous years indicate that the average number of miles driven per car per month was 2,600. Use the sample data to test the research hypothesis that the current mean \( \mu \) differs from 2,600. Set \( \alpha = .05 \) and assume that \( \sigma \) can be replaced by \( s \).
Solution  The null hypothesis for this statistical test is $H_0 : \mu = 2,600$ and the research hypothesis is $H_a : \mu \neq 2,600$. Using $\alpha = .05$, the two-tailed rejection region for this test is $|z| > z_{.025}$ or $|z| > 1.96$ and is located as shown below.

To determine how many standard errors our test statistics $\bar{y}$
lies away from $\mu = 2,600$, compute

$$z_c = \frac{\bar{y} - \mu_0}{\sigma/\sqrt{n}} = \frac{2,752 - 2,600}{350/\sqrt{40}} = 2.75.$$ 

Thus $|z_c| = 2.75 > 1.96$ and therefore we reject $H_0 : \mu = 2,600$ at $\alpha = .05$.

It follows that the observed value for $\bar{y}$ lies more than 1.96 standard errors above the mean $\mu = 2,600$, so we reject the null hypothesis in favor of the alternative $H_a : \mu \neq 2,600$. Since $\bar{y} > 2,600$ we conclude that the mean number of miles driven is greater than 2,600.
Level of Significance or the p-value of a Statistical Test

- As an alternative to the formal test where one uses a rejection region based on a specified the Type I error rate $\alpha$, many researchers compute and report the level of significance or the p-value for the test.

- This is the probability, when $H_0$ is true, of observing a statistic as extreme as the one actually observed. Here “extreme” is means “large” or “small” according to the alternative hypothesis $H_a$. 
• Specifically, for a given $\sigma^2$ and $\mu_0$ we find the p-value by:

1. First computing the test statistic, $z_c$,

$$
  z_c = (\bar{y} - \mu_0)/(\sigma/\sqrt{n})
$$

2. a) If $H_a : \mu > \mu_0$, then $p = P(Z > z_c)$.

   b) If $H_a : \mu < \mu_0$, then $p = P(Z < z_c)$.

   c) If $H_a : \mu \neq \mu_0$, then $p = 2P(Z > |z_c|)$.

• A p-value smaller than the pre-specified $\alpha$ value is evidence in favor of rejecting $H_0$. 
Example 5.12:

- In Example 5.7 we tested $H_0 : \mu \leq 380$ vs. $H_a : \mu > 380$
  Calculating the $z$–statistic,

$$z_c = \frac{\bar{y} - 380}{\sigma/\sqrt{n}} = \frac{390 - 380}{35.2/\sqrt{50}} = 2.01$$

- The level of significance, or p-value for this test is

$$p = P(Z > 2.01) = 1 - P(Z < 2.01) = 0.0222$$

- We fail to reject $H_0 : \mu \leq 380$ at $\alpha = .01$ since p-value is not less than .01
Inferences About $\mu$ when $\sigma$ is unknown

- For large sample sizes, it follows from the CLT that $\bar{Y}$ is approximately Normally distributed i.e., $\bar{Y} \sim N(\mu, \frac{\sigma^2}{n})$.
- It follows that the random variable
  \[ T_{n-1} = \left( \frac{\bar{Y} - \mu}{S/\sqrt{n}} \right) \]
  has approximately the Student’s $t$ distribution with $n - 1$ degrees of freedom.
- Here $Y_1, Y_2, \ldots, Y_n$ are sampling random variables and $\bar{Y} = \frac{1}{n} \sum_i Y_i$ is thus a random variable.
The denominator of $T_{n-1}$ is the random variable

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2}.$$ 

If sampling from a Normal distribution, the sample mean $\bar{Y}$ has a Normal distribution (exactly) and therefore $(\bar{Y} - \mu)/(S/\sqrt{n})$ will have a Student’s $t$ distribution (exactly), regardless of sample size.
• In Chapter 5 until now, we have assumed that \( n \) is large and \( \sigma \) is known, or \( n \) is large and \( n \) is large enough also to use \( s \) as an approximation to \( \sigma \).

• Since we used the CLT to derive confidence limits and tests of hypotheses, the exact nature of the sampled population was not required to be specified.

• Now we require that the population distribution to be Normal. We don’t need to know \( \sigma \) or have a large sample, i.e. \( n \) to be large.

• In this situation, for any \( n \), \( (\bar{Y} - \mu)/(S/\sqrt{n}) \) will have a Student’s t-distribution with d.f. = \( n - 1 \).
• Using this fact we can obtain confidence intervals and conduct tests of hypotheses even though \( \sigma \) is unknown.

• Note the difference between \( Z \) and \( T_{n-1} \) is that the parameter \( \sigma \) is in the denominator of \( Z \).

• That is, the point estimator of \( \sigma \), \( S \) is in the denominator of \( T_{n-1} \).
Properties of the Student’s $t$ distribution

- There are many $t$-distributions each specified by a single parameter called degrees of freedom ($df$).
- Like the standard normal population, the distribution is symmetric about 0 and has mean equal to 0.
- The $t$-distribution has variance $df/(df - 1)$, and hence is more variable than the standard normal distribution which has variance equal to 0.
- We say that the $t$-distribution has heavier tails than the standard normal distribution.
- As the degrees of freedom $df$ increases, the $t$-distribution approaches that of the standard normal distribution.
Thus as the sample size $n$ increases the distribution of the $T_{n-1}$ random variable approaches the standard normal distribution.
Table 2, page 1093, gives quantiles 0.90, 0.95, 0.975, 0.99, 0.995, and 0.999 for $t$ distributions with selected $df$.

**Examples:** For $df = 10$

- $P(T_{10} > 1.372) = 0.10$
- $P(T_{10} < 1.372) = 0.90$
- $P(T_{10} > 1.812) = 0.05$

If $\alpha = 0.05$, then $t_\alpha$ satisfies $P(T_{10} > t_{0.05}) = 0.05$ and from Table 2, $t_{0.05} = 1.812$. If $\alpha = 0.05$, then $t_{\alpha/2} \equiv t_{0.025} = 2.228$

**Comparison of Normal and $t$ Quantiles**

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$z_\alpha$</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>60</th>
<th>240</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>1.28</td>
<td>1.37</td>
<td>1.32</td>
<td>1.31</td>
<td>1.30</td>
<td>1.30</td>
<td>1.28</td>
</tr>
<tr>
<td>0.05</td>
<td>1.65</td>
<td>1.81</td>
<td>1.72</td>
<td>1.70</td>
<td>1.68</td>
<td>1.67</td>
<td>1.65</td>
</tr>
<tr>
<td>0.01</td>
<td>2.33</td>
<td>2.76</td>
<td>2.53</td>
<td>2.46</td>
<td>2.42</td>
<td>2.39</td>
<td>2.34</td>
</tr>
</tbody>
</table>
Confidence Interval for $\mu$ based on the t-distribution

- Let the sample size be $n$, and for $\alpha$ be a specified value, say, for e.g. .05.

- A $(1 - \alpha)100\%$ confidence interval for $\mu$ is given by

$$\left( \bar{y} - t_{\alpha/2} \frac{s}{\sqrt{n}}, \bar{y} + t_{\alpha/2} \frac{s}{\sqrt{n}} \right)$$

- This may also be written as \( (\bar{y} - t_{\alpha/2}s_{\bar{y}}, \bar{y} + t_{\alpha/2}s_{\bar{y}}) \)

- Here, $t_{\alpha/2}$ is the $1 - \alpha/2$ percentile of the t-distribution with $n - 1$ degrees of freedom and $s_{\bar{y}}$ is the standard error of the mean.
EXAMPLE 5.17

An airline wants to evaluate the depth perception of its pilots over the age of fifty. A random sample of \( n = 14 \) airline pilots over the age of fifty are asked to judge the distance between two markers placed 20 feet apart at the opposite end of the laboratory. The sample data listed here are the pilots' error (recorded in feet) in judging the distance.

\[
\begin{array}{cccccccc}
2.7 & 2.4 & 1.9 & 2.6 & 2.4 & 1.9 & 2.3 \\
2.2 & 2.5 & 2.3 & 1.8 & 2.5 & 2.0 & 2.2 \\
\end{array}
\]

Use the sample data to place a 95\% confidence interval on \( \mu \), the average error in depth perception for the company's pilots over the age of fifty.

SOLUTION

Before setting up a 95\% confidence interval on \( \mu \), we must first assess the normality assumption by plotting the data in a normal probability plot or a boxplot. Figure 5.20 is a boxplot of the 14 data values. The median line is near the center of the box, the right and left whiskers are approximately the same length, and there are no outliers. The data appear to be a sample from a normal distribution. Thus, it is appropriate to construct the confidence interval based on the \( t \) distribution. You can verify that:

\[
y = 2.26 \quad \text{and} \quad s = .28
\]

**FIGURE 5.20**

Boxplot of distance (with 95\% \( t \) confidence interval for the mean)

![Boxplot](image)

Referring to Table 2 in the Appendix, the \( t \)-value corresponding to \( \alpha = .025 \) and \( df = 13 \) is 2.160. Hence, the 95\% confidence interval for \( \mu \) is

\[
y \pm t_{0.025} \frac{s}{\sqrt{n}} \quad \text{or} \quad 2.26 \pm 2.160 \frac{.28}{\sqrt{14}}
\]

which is the interval 2.26 ± .16, or 2.10 to 2.42. Thus, we are 95\% confident that the average error in the pilots' judgment of the distance is between 2.10 and 2.42 feet.
Test Procedures based on the t-distribution:

Hypotheses:

Case 1: $H_0 : \mu \leq \mu_0$ vs. $H_a : \mu > \mu_0$ (right-tailed test)

Case 2: $H_0 : \mu \geq \mu_0$ vs. $H_a : \mu < \mu_0$ (left-tailed test)

Case 3 $H_0 : \mu = \mu_0$ vs. $H_a : \mu \neq \mu_0$ (two-tailed test)

T.S: $t_c = \frac{\bar{y} - \mu_0}{s/\sqrt{n}}$

R.R: For Type I error probability of $\alpha$:

Case 1: Reject $H_0$ if $t \geq t_{\alpha,(n-1)}$

Case 2: Reject $H_0$ if $t \leq -t_{\alpha,(n-1)}$

Case 3: Reject $H_0$ if $|t| \geq t_{\alpha/2,(n-1)}$
Level of Significance (p-value):

Case 1: \( p = P(T_{n-1} > t_c) \).

Case 2: \( p = P(T_{n-1} < t_c) \).

Case 3: \( p = 2P(T_{n-1} > |t_c|) \).

Exercise 5.15

A massive multistate outbreak of food-borne illness was attributed to *Salmonella enteritidis*. Epidimiologists determined that the source of the illness was ice cream. They sampled nine production runs from the company that produced the ice cream to determine the level of *Salmonella enteritidis* in the ice cream.
These levels (MPN/g) are as follows

\[.593 \quad .142 \quad .329 \quad .691 \quad .231 \quad .793 \quad .519 \quad .392 \quad .418\]

Use the data to determine whether the mean level of *Salmonella enteritidis* in the ice cream is greater than .3 MPN/g with \( \alpha = .01 \)

**Solution:**

Need to test \( H_0 : \mu \leq .3 \) vs. \( H_a : \mu > .3 \)

Because of the small sample size, we need to examine whether the data have been sampled from a normal distribution. To do this a normal probability plot is a good tool.
From the data $\bar{y} = .456$ and $s = .2128$ are computed, giving

$$t_c = \frac{\bar{y} - \mu_0}{s/\sqrt{n}} = \frac{.456 - .3}{.2128/\sqrt{9}} = 2.21$$

Because for the one-tailed test we need $t_{\alpha,(n-1)}$; we look up $t_{.01}$ with $df = 9 - 1 = 8$. It is 2.896

Thus the rejection region is: $t > 2.896$
Since $t_c = 2.214$ does not exceed 2.896 it is not in the R.R. Thus, there is insufficient evidence in the data to reject $H_0$ i.e. to say that the mean level of *Salmonella enteritidis* exceeds the dangerous level of .3 MPN/g.

The p-value to be computed is $P(T_8 > 2.21)$

To calculate this exactly using the t-table is not possible since it is tabulated for only a few values of $\alpha$.

However, we can bound the p-value by noting that, for $(df = 8)$, 2.21 lies between 1.86 and 2.306.

This gives $.025 < p\text{-value} < .05$ showing that the p-value is not less than our $\alpha$ of .01. Thus we fail to reject $H_0$. 
**OC Curve and the Power of a test**

The probabilities of the four possible outcomes of a statistical test are:

<table>
<thead>
<tr>
<th>Decision</th>
<th>Null Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True</td>
</tr>
<tr>
<td>Reject $H_0$</td>
<td>Type I Error $\alpha$</td>
</tr>
<tr>
<td>Accept $H_0$</td>
<td>Correct Decision $1 - \alpha$</td>
</tr>
</tbody>
</table>

- The probabilities of Type I and II error are $\alpha$ and $\beta$, respectively.
• The experimenter can control only the Type I error probability. We do this by specifying an $\alpha$ for an experiment (before the data values are measured).

• When we are testing an hypothesis about $\mu$ using $\alpha$ for the test, the value of Type II error probability $\beta$ depends on the actual value $\mu$ which is not known.

• This is because $\beta$ is the probability of incorrectly accepting $H_0$ when $H_a$ is true.

• When $H_a$ is true, the actual value of $\mu$ may be any value under $H_a$, i.e, a value not specified under $H_0$. Since this value of $\mu$ is unknown, $\beta$ cannot be calculated.
• The implication of this is that when, based on a test, we find that we cannot reject \( H_0 \), we will not say that we accept \( H_0 \).
• Because if we say we accept \( H_0 \) and \( \beta \) turns out to be large, then the probability is large that we will be committing a Type II error.
• Thus the correct way to state the decision is to say, that we fail to reject \( H_0 \).
• In practice, \( \beta \) probabilities for several choices of \( \mu \) (call them \( \beta(\mu) \)) are calculated and plotted in a graph called the OC curve.
• The OC curve can be used to read-off the Type II error probability for a specified set of values sample size \( n \) and \( \alpha \).
Figure 5.12 shows the OC curve for the test of $H_0 : \mu \leq 84$, $H_a : \mu > 84$ for a population with $\sigma = 1.4$. For example, for $\alpha = .05$, $n = 10$, it can be seen that $\beta(84.8) \approx .4$ and that $\beta(84.8)$ decreases as sample size goes from 10 to 25. A conclusion that can be made about this test from the OC curve is that Type II error probability would be $< .1$ for an actual $\mu > 84.8$ for $n = 25$. 
• Figure 5.11 in the textbook (not reproduced here) show how the Type II error probability varies with the value of $\mu$ under the alternative (denoted by $\mu_a$).
• Examples 5.8 and 5.10 show the calculation $\beta$ for a particular value of $\mu_a$ and uses formulas given on page 241 to calculate $\beta$.
• These illustrations use the experiments described in Examples 5.7 and 5.9 (Read pp. 238/243 for full details)
• Another quantity that may be calculated for a test procedure for a specified value of $\mu$ is called the Power of the test and is defined as $1 - \beta(\mu)$.
• The corresponding plot of power against a set of $\mu$ values is called the power curve.
• By definition, power of a test is the probability of rejecting $H_0$ for a specified value of $\mu$ under $H_a$.

• In practice, tests are designed to have large power for some $\mu$ values of interest so that they have small Type II error probabilities.

• We can relate to this idea by thinking of a test as having very good power if it has a very good chance of detecting whether a change in $\mu$ has actually occurred.

• This is usually done by selecting a sample size $n$ to be used for the experiment so that the desired power is achieved for a specified $\mu$ and $\alpha$. 
Using Type II Error Probability $\beta$ Curves

- Consider the Salmonella example again. We have $n = 9$, and $\alpha = .01$. Thus $df = 8$ and we estimate $\sigma \approx .25$.
- We can compute the values of $d$ for several values of $\mu_a$. Then we’ll read $\beta$ for those values of $d$ using the graph in Table 3 in the Appendix (see next slide for the curves for $\alpha = .01$).
- As an example, for $\mu_a = .45$, $d = \frac{|\mu_a - \mu_0|}{\sigma} = \frac{|.45 - .3|}{.25} = .6$.
- Corresponding to $d = .6$ on the horizontal axis, using the curve for $df = 8$ we see that $\beta(.45) = .79$, approx.
- Similarly, for $\mu_a = .45$, $d = 1.0$, and thus $\beta(.55) = .43$, approx. We can construct a table as shown (next slide).
Departures from Normality

- When $\bar{Y}$ is a Normal random variable, i.e., when sampling from a Normal population with mean $\mu_0$, $\frac{\bar{Y} - \mu_0}{S/\sqrt{n}}$ is a $T_{n-1}$ random variable.

- This is a theoretical fact. In practice, however, we never sample exactly from a Normal population, so $\frac{\bar{Y} - \mu_0}{S/\sqrt{n}}$ will be only approximately $T_{n-1}$.

- How much effect can this have on C.I.’s and tests we construct?

- For symmetrically distributed populations and not too small $n$ there is little to worry about.
For highly skewed population distributions the approximation can be terrible especially for small $n$.

<table>
<thead>
<tr>
<th>TABLE 5.6</th>
<th>Level and power values for $t$ test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population Distribution</td>
<td>$n = 10$</td>
</tr>
<tr>
<td></td>
<td>$d$</td>
</tr>
<tr>
<td>Shift</td>
<td>.05</td>
</tr>
<tr>
<td>0</td>
<td>.035</td>
</tr>
<tr>
<td>0.2</td>
<td>.025</td>
</tr>
<tr>
<td>0.6</td>
<td>.007</td>
</tr>
<tr>
<td>0.8</td>
<td></td>
</tr>
</tbody>
</table>

It is recommended that one look at a boxplot, Normal plot, and/or other graphics to see whether severe skewness of the sampling population is indicated. If not, proceed to use the $t$-distribution. If yes, one can use a nonparametric procedure or use a transformation. We will look at some of these later.
Using Confidence Intervals to Test Hypotheses

- We can always look at a \((1-\alpha)100\%\) confidence interval and see what the result of a test would be if we carry out the test.
- For example, consider the \((1-\alpha)100\%\) confidence interval for \(\mu\)
  \[
  \left( \bar{y} - t_{\alpha/2} \frac{s}{\sqrt{n}}, \quad \bar{y} + t_{\alpha/2} \frac{s}{\sqrt{n}} \right)
  \]
- Suppose that \(\mu_0\) is not included in the above confidence interval because
  \[
  \bar{y} + t_{\alpha/2} \frac{s}{\sqrt{n}} < \mu_0
  \]
• By rearranging this we see that this is equivalent to $t_c$ being in the rejection region, i.e.,
\[
\frac{\bar{y} - \mu_0}{s/\sqrt{n}} \geq t_{\alpha/2, (n-1)}
\]

• Observe carefully that this is the same rejection region for the test of $H_0 : \mu \leq \mu_0$ vs. $H_a : \mu > \mu_0$ at level $\alpha/2$.

• That is, we will be rejecting $H_0$ at level $\alpha/2$ if
\[
\frac{\bar{y} - \mu_0}{s/\sqrt{n}} \geq t_{\alpha/2, (n-1)}
\]

• This is equivalent to saying that if $\mu_0$ was not included in a $100(1-\alpha)$% interval, then $H_0$ will be rejected if the test is carried out at $\alpha/2$ level.
• For a two-tailed test, the confidence interval should be based on the same $\alpha$ as the test to make this inference.

**In summary,**

• To test $H_0 : \mu \leq \mu_0$, $H_a : \mu > \mu_0$, at level $\alpha/2$, use a $100(1-\alpha)\%$ confidence interval.

• To test $H_0 : \mu \geq \mu_0$, $H_a : \mu < \mu_0$, at level $\alpha/2$, use a $100(1-\alpha)\%$ confidence interval.

• $H_0 : \mu = \mu_0$, $H_a : \mu \neq \mu_0$, at level $\alpha$, use a $100(1-\alpha)\%$ confidence interval.