W15 QMS 202 Lecture 3

Dr Boža Tasić

January 27, 2015

Hypothesis Testing (One-Sample Test) - Learning Objectives

- In this Chapter we will learn
  - The basic principles of hypothesis testing
  - How to use hypothesis testing to test a mean or proportion
  - The assumptions of each hypothesis-testing procedure, how to evaluate them, and the consequences if they are seriously violated
  - How to avoid the pitfalls involved in hypothesis testing
  - The ethical issues involved in hypothesis testing

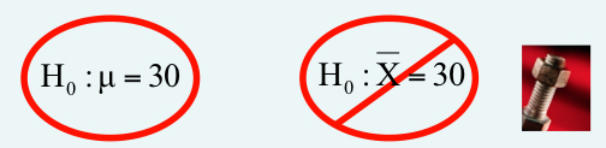
What is a Hypothesis

- A hypothesis is a claim (assertion) about a population parameter:
- Population Mean:
  - The mean GPA of Ryerson 1^{st} year students is $\mu = 3.2$.
- Population Proportion:
  - The proportion of Ryerson continuing education students is $\pi = 0.28$.

The Null Hypothesis H_0

- The **null hypothesis** H_0 states the claim to be tested
- Example:
  - The average diameter of a manufactured bolt is equal to 30mm

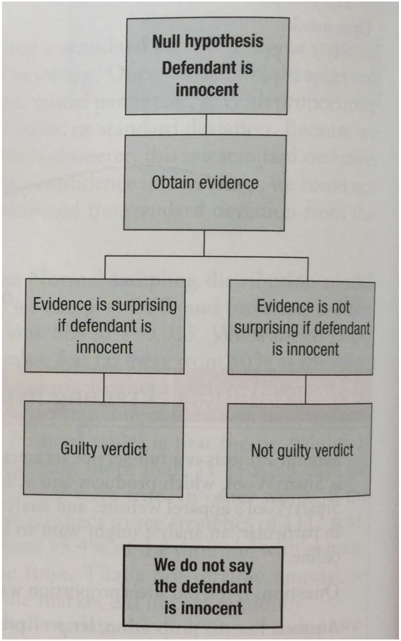
$$H_0 : \mu = 30\text{mm}$$

- The null hypothesis, H_0 always refers to a specified value of the population parameter (such as μ), not a sample statistic (such as $\overline{X}$).
  - 

The Null Hypothesis H_0

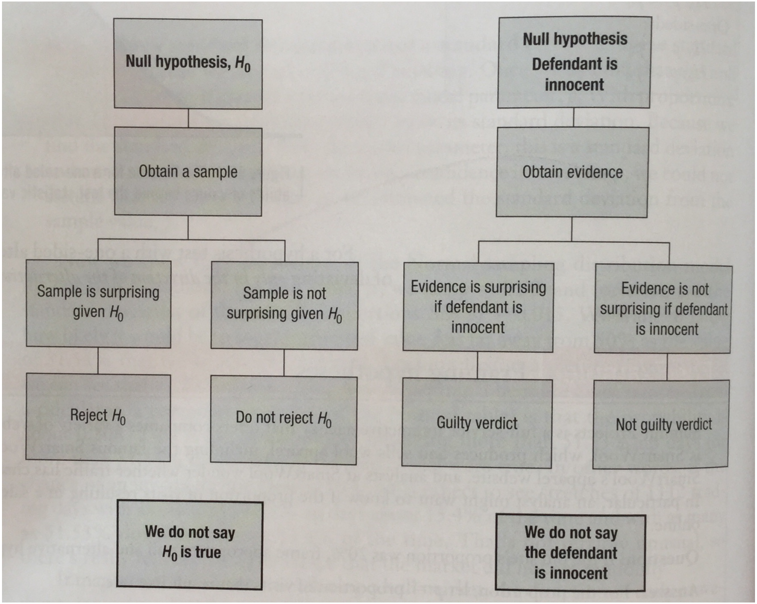
- The null hypothesis H_0 represents the status quo or the current belief in the situation.
- The statement of the null hypothesis H_0 always contains $=$ sign, for example $H_0 : \mu = 30$mm.
- The null hypothesis H_0 may or may not be rejected.
- Begin with the assumption that the null hypothesis H_0 is true
  - Similar to the notion of **innocent until proven guilty**.

A Trial as a Hypothesis Test



-

Hypothesis Testing vs Court Case



The Alternative Hypothesis H_1

- The **alternative hypothesis** H_1 (sometimes also denoted by H_a) is the opposite of the null hypothesis
- Example:
  - The average diameter of a manufactured bolt is not equal to 30mm

$$H_1 : \mu \neq 30\text{mm}$$

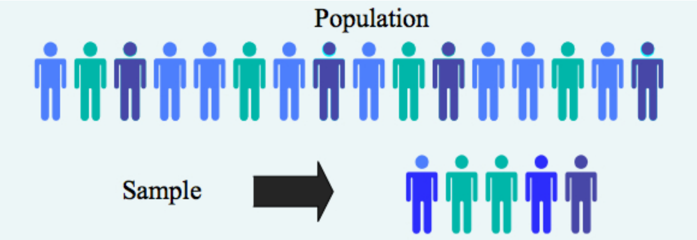
- The alternative hypothesis H_1 challenges the status quo or the current belief in the situation.
- The statement of the alternative hypothesis H_1 never contains $=$ sign, for example $H_1 : \mu \neq 30\text{mm}$
- The alternative hypothesis H_1 may or may not be proven.
- Is generally the hypothesis that the researcher is trying to prove.

The Null and Alternative Hypothesis Testing

- You reject the null hypothesis H_0 when the sample evidence suggests that it is far more likely that the alternative hypothesis H_1 is true. However, failure to reject the null hypothesis is not proof that it is true. We can only conclude that there is insufficient evidence to warrant its rejection.
- If you reject the null hypothesis H_0 you have statistical proof that the alternative hypothesis H_1 is correct.
- If you do not reject the null hypothesis H_0, you have failed to prove the alternative hypothesis H_1. The failure to prove the alternative hypothesis, does not mean that you have proven the the null hypothesis.

The Hypothesis Testing Process

- Claim: The population mean age is 50
- $$H_0 : \mu = 50, \qquad H_1 : \mu \neq 50$$
- We sample the population and find the sample mean

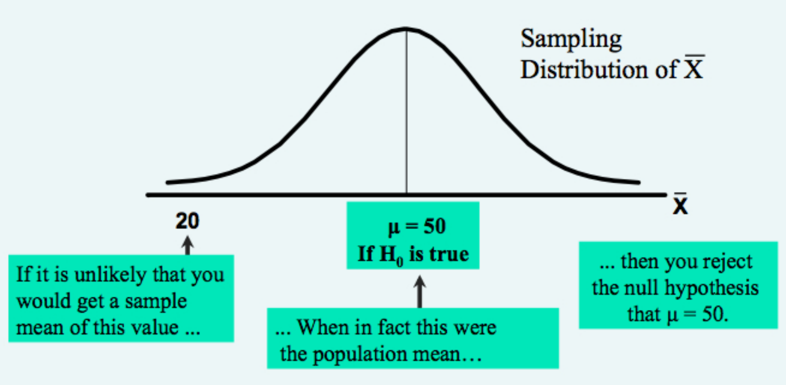


The Hypothesis Testing Process

- Suppose that we found that the sample mean age is $\overline{X} = 20$.
- The sample mean $\overline{X} = 20$ is significantly lower than the claimed mean population age $\mu = 50$.
- If $H_0 : \mu = 50$ were true, than the probability of getting the sample mean $\overline{X} = 20$ is very small! So, we reject the null hypothesis.
- Simply, getting a sample mean $\overline{X} = 20$ is so unlikely if the population mean was $\mu = 50$. So, we conclude that the population mean must not be 50.

The Hypothesis Testing Process

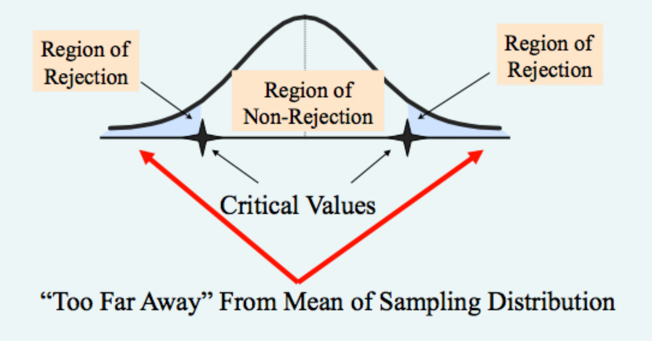
▶



The Hypothesis Testing Process

- If the sample mean is close to the stated population mean, we do not reject the null hypothesis.
- If the sample mean is far from the stated population mean, we do reject the null hypothesis.
- The question is
  - How far is far enough to reject H_0?
- The critical value of a test statistic will help us answer the question of how far is far enough.

The Test Statistic and Critical Values

- The sampling distribution of the test statistic is divided into two regions, a **region of rejection** (AKA the **critical region**) and a **a region of nonrejection**.
- The **critical value** divides the nonrejection region from the rejection region.
- 

Possible Errors in Hypothesis Test Decision Making

- Using a sample statistic to make a decision about the population parameter is not risk free. We can make two types of errors when applying hypothesis testing.
- A **Type I Error** ("False Alarm")
  - Occurs if we reject the null hypothesis H_0, when it is true and should not be rejected. It is considered a serious type of error.
  - The probability of a Type I error occurring is α.
    - It is known as the **level of significance** of the statistical test.
    - It is set by researchers in advance.
    - Most common levels of significance are 0.01, 0.05 and 0.10.
- A **Type II Error** ("Missed Opportunity")
  - Occurs if we do not reject the null hypothesis H_0, when it is false and should be rejected.
  - The probability of a Type II error occurring is β.
  - It is known as the β **risk**.

Possible Errors in Hypothesis Test Decision Making

- The following table summarizes the type of errors

Possible Hypothesis Test Outcomes		
	Actual Situation	
Decision	H_0 True	H_0 False
Do Not Reject H_0	No Error Probability $1 - \alpha$	Type II Error Probability β
Reject H_0	Type I Error Probability α	No Error Probability $1 - \beta$

Possible Errors in Hypothesis Test Decision Making

- The **confidence coefficient** $(1 - \alpha)$ is the probability of not rejecting the null hypothesis H_0 when it is true and should not be rejected.
- The confidence coefficient gives the confidence level that we studied in the previous chapter when constructing confidence intervals.
  - In the example with the mean age, the confidence coefficient measures the probability of concluding that the mean age is 50 when it is actually 50.

Possible Errors in Hypothesis Test Decision Making

- Unlike the Type I error that we control by selecting α, the probability of committing a Type II error β depends on the difference between the hypothesized and actual values of the population parameter. For example if the difference between the hypothesized and actual values of the population parameter is large, β is small.
  - In the example with the mean age $\overline{X} = 20$, there is a small chance β that we will conclude that the mean has not changed from 50.
  - But, if we had $\overline{X} = 49$, there is a large chance β that we will conclude that the mean has not changed from 50.

Possible Errors in Hypothesis Test Decision Making

- The **power of a statistical test** $(1 - \beta)$ is the probability that you will reject the null hypothesis H_0 when it is false and should be rejected.
  - In the example with the mean age $\overline{X} = 20$, the power of a statistical test $(1 - \beta)$ is the probability that you will correctly conclude that the mean age is not 50 when it actually is not 50.

Type I and Type II Errors Relationships

- Type I and Type II errors cannot happen at the same time
- A Type I error can only occur if H_0 is **true**
- A Type II error can only occur if H_0 is **false**
- If Type I error probability α increases, then Type II error probability β decreases

$$\beta \quad \downarrow \quad \text{when} \quad \alpha \quad \uparrow$$

Factors Affecting Type II Error

- The Type II error β increases when the difference between hypothesized parameter and its true value
- $$\beta \quad \uparrow \quad \text{when} \quad \alpha \quad \downarrow$$
- $$\beta \quad \uparrow \quad \text{when} \quad \sigma \quad \uparrow$$
- $$\beta \quad \uparrow \quad \text{when} \quad n \quad \downarrow$$

Quick Question 1- Null Hypothesis

- Which of the following would be an appropriate null hypothesis?

a) The mean of a population is equal to 55.
b) The mean of a sample is equal to 55.
c) The mean of a population is greater than 55.
d) Only A and C are appropriate.

Quick Question 2 - Type I Error

- If a test of hypothesis has a Type I error probability α of 0.01, it means that

a) if the null hypothesis is true, you don't reject it 1% of the time.

b) if the null hypothesis is true, you reject it 1% of the time.

c) if the null hypothesis is false, you don't reject it 1% of the time.

d) if the null hypothesis is false, you reject it 1% of the time.

Quick Question 3 - Type II Error

- If you know that the level of significance α of a test is 5%, you can tell that the probability of committing a Type II error β is

a) 2.5%.

b) 95%.

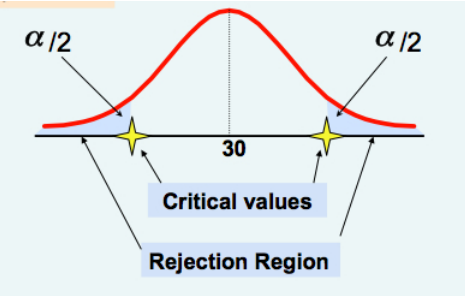
c) 97.5%.

d) unknown.

Level of Significance and the Rejection Region

- Assume that we are testing at the level of significance α, the hypothesis

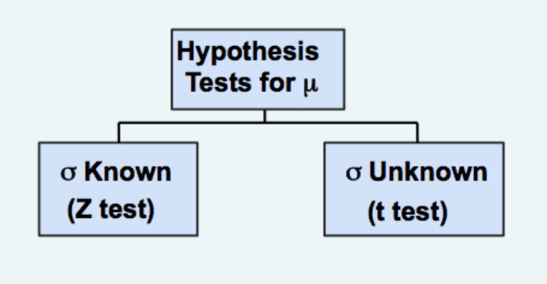
$$H_0 : \mu = 30, \qquad H_1 : \mu \neq 30$$



- This is a two-tail test because there is a rejection region in both tails

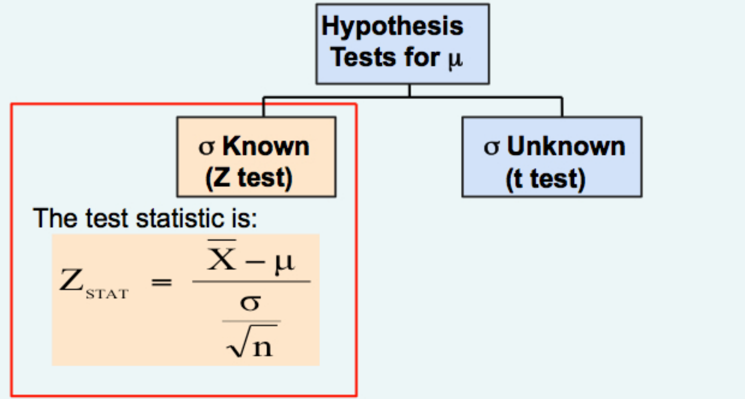
Hypothesis Tests for the Mean

- We will learn



Z Test of Hypothesis for the Mean (σ Known)

- When σ is known we use the **Z test for the mean** if the population is normally distributed. If the population is not normally distributed our sample needs to be large enough!
- We convert sample statistic $\overline{X}$ to a Z_{STAT} test statistic



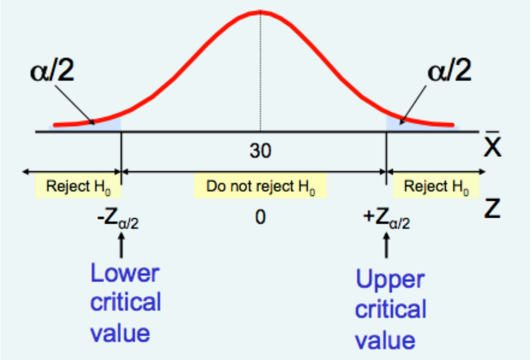
Critical Value Approach to Testing

- For a two-tail test for the mean, σ known:
- Convert sample statistic $\overline{X}$ to a Z_{STAT} test statistic
- Determine the critical Z values for a specified level of significance α
- **Decision Rule**: If the test statistic falls in the rejection region, reject H_0; otherwise do not reject H_0

Critical Value Approach to Testing

- We convert sample statistic $\overline{X}$ to a Z_{STAT} test statistic
- If we are testing at the level of significance α, the hypothesis

$$H_0 : \mu = 30, \qquad H_1 : \mu \neq 30$$



- We have two cutoff values (critical values), defining the regions of rejection

The 6 Step Method of Hypothesis Testing

1. State the null hypothesis, H_0 and the alternative hypothesis, H_1.
2. Choose the level of significance, α, and the sample size, n. The level of significance is based on the relative importance of the risks of committing Type I and Type II errors in the problem.
3. Determine the appropriate test statistic and sampling distribution.
4. Determine the critical values that divide the rejection and nonrejection regions.
5. Collect data and compute the value of the test statistic.
6. Make the statistical decision and state the managerial conclusion. If the test statistic falls into the nonrejection region, do not reject the null hypothesis H_0. If the test statistic falls into the rejection region, reject the null hypothesis. Express the managerial conclusion in the context of the problem.

Critical Value Approach - Example 1

- Test the claim that the true mean diameter of a manufactured bolt is 30 mm. Assume that $\sigma = 0.8$.

1. State the appropriate null and alternative hypotheses.
   - $$H_0 : \mu = 30, \qquad H_1 : \mu \neq 30$$
2. Specify the desired level of significance and the sample size.
   - Suppose that $\alpha = 0.05$ and $n = 100$ are chosen for this test.
3. Determine the appropriate technique
   - σ is assumed known so this is a Z test.
4. Determine the critical values
   - For $\alpha = 0.05$ the critical Z values are ± 1.96.

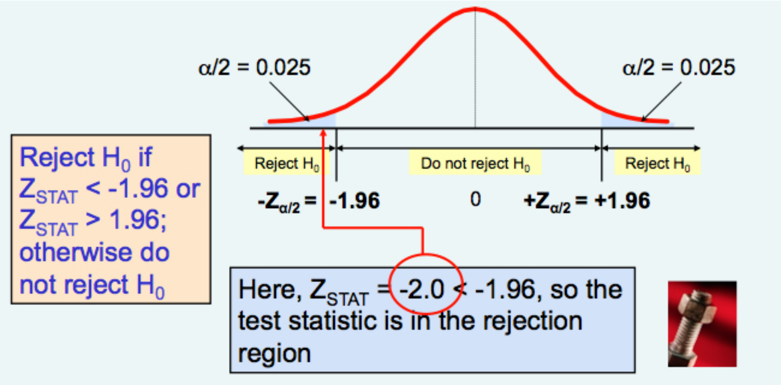
Critical Value Approach - Example 1

5. Collect the data and compute the test statistic
   - Suppose the sample results are $n = 100$, $\overline{X} = 29.84$ ($\sigma = 0.8$ is assumed known)
   - So the test statistic is:

$$Z_{\text{STAT}} = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{29.84 - 30}{\frac{0.8}{\sqrt{100}}} = \frac{-0.16}{0.08} = -2.0$$

6. Is the test statistic in the rejection region?

Critical Value Approach to Testing



6. (continued) Reach a decision and interpret the result
   - Since $Z_{\text{STAT}} = -2.0 < -1.96$, reject the null hypothesis and conclude there is sufficient evidence that the mean diameter of a manufactured bolt is not equal to 30.

The p-Value Approach to Hypothesis Testing

- The fundamental step in our reasoning is the question "**Is the data statistic surprising, given the null hypothesis?**"
- The probability of getting a test statistic equal to or more extreme than the observed sample value given that the null hypothesis H_0 is true is called the ***p*-value**.
- The p-value is also called the **observed level of significance**.

p-Value Approach to Testing: Interpreting the p-Value

- Compare the p-value with the level of significance α
  - If p-value $< \alpha$, reject H_0
  - If p-value $\geq \alpha$, do not reject H_0
- **Remember:**
  - If the p-value is low, then reject H_0

The 5 Step *p*-value Method of Hypothesis Testing

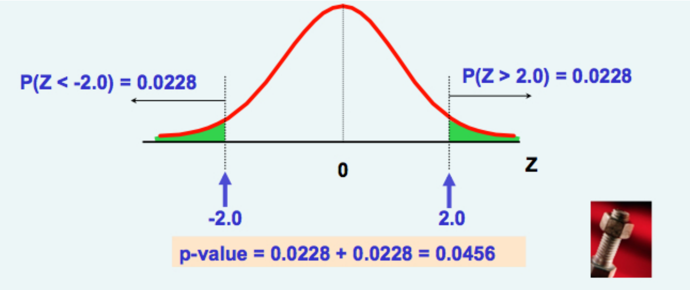
1. State the null hypothesis, H_0 and the alternative hypothesis, H_1.
2. Choose the level of significance, α, and the sample size, n. The level of significance is based on the relative importance of the risks of committing Type I and Type II errors in the problem.
3. Determine the appropriate test statistic and sampling distribution.
4. Collect data and compute the value of the test statistic and the *p*-value
5. Make the statistical decision and state the managerial conclusion. If the *p*-value is $< \alpha$, then reject H_0, otherwise do not reject H_0. State the managerial conclusion in the context of the problem.

p-Value Approach - Example 1

- Test the claim that the true mean diameter of a manufactured bolt is 30 mm. Assume that $\sigma = 0.8$.

1. State the appropriate null and alternative hypotheses.
   - $$H_0 : \mu = 30, \qquad H_1 : \mu \neq 30$$
2. Specify the desired level of significance and the sample size.
   - Suppose that $\alpha = 0.05$ and $n = 100$ are chosen for this test.
3. Determine the appropriate technique
   - σ is assumed known so this is a Z test.
4. Collect the data, compute the test statistic and the p-value
   - Suppose the sample results are $n = 100$, $\overline{X} = 29.84$ ($\sigma = 0.8$ is assumed known) So the test statistic is:
   - $$Z_{\text{STAT}} = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{29.84 - 30}{\frac{0.8}{\sqrt{100}}} = \frac{-0.16}{0.08} = -2.0$$

p-Value Approach - Example 1

4. (continued) Calculate the p-value
   - How likely is it to get a Z_{STAT} of -2 (or something farther from the mean (0), in either direction) if H_0 is true?
   - This means we need to compute $P(Z < -2)$ and $P(Z > 2)$ (it is a two-tail test). Both $P(Z < -2)$ and $P(Z > 2)$ can be computed using your calculator! You will find $P(Z < -2) = 0.0228$ and $P(Z > 2) = 0.0228$.
   - 

   - p-value $= P(Z < -2) + P(Z > 2) = 0.0228 + 0.0228 = 0.0456$

p-Value Approach - Example 1

5. Is the p-value $< \alpha$?
   - Since the p-value $= 0.0456 < \alpha = 0.05$ reject H_0.
   - State the managerial conclusion in the context of the situation.
   - **Conclusion**:There is sufficient evidence to conclude the average diameter of a manufactured bolt is not equal to 30mm.

p-Value Approach - Example 1- Using CASIO Calculator

1. Select **STAT**, **F3** (TEST), **F1**(Z) and then **F1**(1-S). Now select in the **1-Sample ZTest** the following options:
2. Data: **F2** (Variable)
3. μ: **F1 EXE** ($\neq \mu 0$)
4. $\mu 0$: **30 EXE**
5. σ: **0.8 EXE**
6. $\overline{x}$: **29.84 EXE**
7. n: **100 EXE**
8. Save Res: If you do not want to save results in a list press **F1** (None). If you want to save results in a list press **F2** and type the number of list where you want to save results.

8. Execute: key **EXE** or **F1**(Calc).

- You will get the following summary:

p-Value Approach - Example 1- Using CASIO Calculator

- **1-Sample ZTest**
- $\mu \neq 30$
- $z = -2$
- $p = 0.0455$
- $\overline{x} = 29.84$
- $n = 100$
- **Statistical Decision**: Since p-value $= 0.0456 < \alpha = 0.05$ reject H_0.
- **Conclusion**:There is sufficient evidence to conclude the average diameter of a manufactured bolt is not equal to 30mm.

Connection Between Two-Tail Tests and Confidence Intervals

- For $\overline{X} = 29.84$, $\sigma = 0.8$ and $n = 100$, the 95% confidence interval is determine by
- $$\overline{X} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 29.84 \pm 1.96 \frac{0.8}{\sqrt{100}} = 29.84 \pm 0.1568$$
- or in the interval form $29.6832 \leq \mu \leq 29.9968$.
- Since this interval does not contain the hypothesized mean 30, we reject the null hypothesis at $\alpha = 0.05$.

p-Value - Example 2

- A company that makes batteries for computers needs to make sure that the voltage of these batteries is not too low or too high. The ideal voltage for the QMS model battery, which is used in many computers, is 5.60 volts. The process that is used to make these batteries is also used to make many other models of batteries with different voltages. It is known that the standard deviation of the process is 0.18 volts and that the voltage will fit a normal distribution. The process has been just set up to produce 1 million QMS batteries. Before too many are produced the batteries need to be checked to see if the average voltage is close too 5.60 volts. A sample of 25 QMS batteries is tested, and the average voltage is 5.65 volts. Should the process be allowed to continue production of the QMS batteries or should it be adjusted? Use a 5% level of significance.

Quick Question 1 - Example 2

- What type of a parameter is being tested here?

a) σ.

b) μ.

c) π.

d) $\overline{X}$.

Quick Question 2 - Example 2

- Which of the following would be an appropriate alternative hypothesis?

a) $H_0 : \mu = 5.60$ volts.

b) $H_1 : \mu = 5.60$ volts.

c) $H_0 : \mu \neq 5.60$ volts.

d) None of the above.

Quick Question 3 - Example 2

- What is the value of the test statistic?

a) $Z_{\text{STAT}} = 1.96$

b) $Z_{\text{STAT}} = 1.39$

c) $Z_{\text{STAT}} = 2.19$

d) None of the above.

Quick Question 4 - Example 2

- What is the p-value for the test?

a) $p = 0.0165$

b) $p = 0.0649$

c) $p = 0.1898$

d) $p = 0.1649$

Quick Question 5 - Example 2

- What is the statistical decision

a) Reject H_0
b) Accept H_1
c) Reject H_1
d) Fail to reject H_0

Quick Question 6 - Example 2

- What conclusion can you reach?

a) There is not sufficient evidence to conclude that the production process does not need to be adjusted.

b) There is sufficient evidence to conclude that the production process needs to be adjusted.

c) There is sufficient evidence to conclude that the production process does not need to be adjusted.

p-Value - Example 2 - Full Solution

1. State the appropriate null and alternative hypotheses.
   - $$H_0 : \mu = 5.60, \qquad H_1 : \mu \neq 5.60$$
2. Specify the desired level of significance and the sample size.
   - We have $\alpha = 0.05$ and $n = 25$.
3. Determine the appropriate technique
   - $\sigma = 0.18$ is known so this is a Z test.
4. Compute the test statistic and the p-value
   - The test statistic is:
   - $$Z_{\text{STAT}} = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{5.65 - 5.60}{\frac{0.18}{\sqrt{25}}} = \frac{0.05}{0.036} = 1.38888$$

p-Value - Example 2 - Full Solution

4. (continued) Calculate the p-value
   - How likely is it to get a Z_{STAT} of -1.3888 (or something farther from the mean (0), in either direction) if H_0 is true?
   - This means we need to compute $P(Z < -1.3888)$ and $P(Z > 1.3888)$ (it is a two-tail test). You can compute them separately and then add to get the p-value or we can use that

-

$$P(Z < -1.3888) + P(Z > 1.3888) = 1 - P(-1.3888 < Z < 1.3888)$$

- We find $P(-1.3888 < Z < 1.3888) = 0.83513$ and hence

$$\begin{aligned} \text{p-value} &= P(Z < -1.3888) + P(Z > 1.3888) \\ &= 1 - P(-1.3888 < Z < 1.3888) \\ &= 1 - 0.83511 \\ &= 0.16489 \end{aligned}$$

- Since the p-value $= 0.16489 > 0.05$, the conclusion is to not reject the null hypothesis. In other words the production process does not need to be adjusted.

p-Value Approach - Example 2 - Using CASIO Calculator

1. Select **STAT**, **F3** (TEST), **F1**(Z) and then **F1**(1-S). Now select in the **1-Sample ZTest** the following options:
2. Data: **F2** (Variable)
3. μ: **F1 EXE** ($\neq \mu 0$)
4. $\mu 0$: **5.60 EXE**
5. σ: **0.18 EXE**
6. $\overline{x}$: **5.65 EXE**
7. n: **25 EXE**
8. Save Res: If you do not want to save results in a list press **F1** (None). If you want to save results in a list press **F2** and type the number of list where you want to save results.
8. Execute: key **EXE** or **F1**(Calc).

- You will get the following summary:

p-Value Approach - Example 1- Using CASIO Calculator

- **1-Sample ZTest**
- $\mu \neq 5.60$
- $z = 1.3888$
- $p = 0.16487$
- $\overline{x} = 5.65$
- $n = 25$
- **Statistical Decision**: Since p-value $= 0.16487 > \alpha = 0.05$ do not reject H_0.
- **Conclusion**:There is sufficient evidence to conclude that the production process does not need to be adjusted.