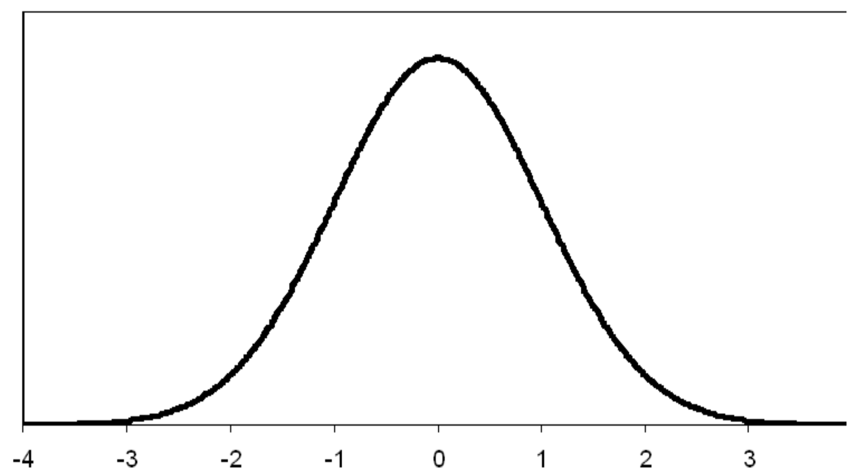
**STAT 206: Chapter 9 (Hypothesis Testing – One-Sample Tests)**

**Ideas in Chapter 8:**

- Basic principles of hypothesis testing
- How to use hypothesis testing to test a mean or proportion
  - Assumptions of each hypothesis-testing procedure
  - How to evaluate them
  - Consequences if they are seriously violated
- Pitfalls & ethical issues involved in hypothesis testing
- How to avoid the pitfalls involved in hypothesis testing
- **Consider:**
  A company claims that it has only a 5% complaint rate for its products. A consumer protection group thinks the percent is higher. A survey of a random sample of 400 product owners shows that 33 had complaints. $\hat{p}$ = ?

  A. $\frac{0.05}{400}$ = 0.0001 B. 0.05 C. $\sqrt{\frac{0.05}{400}}$ D. $\frac{33}{400}$ E. $\sqrt{\frac{0.05(1-0.05)}{400}} = 0.01$

- **Assume that the company's claim is true** and that ***p*** (the population proportion) is really 0.05 (5%) just as the company claims.
  - Remember,
  - The mean of the sampling distribution of $\hat{p}$ is:
  - the standard deviation of the sampling distribution of $\hat{p}$ is:




- **Null Hypothesis, $H_0$ (H-naught)**
  - **Hypothesis test** checks
  - **$H_0$** is ALWAYS about a population parameter, **NOT** a sample statistic
  - Always contains "=", or "≤", or "≥" sign

- Until hypothesis test is completed and the decision is made,

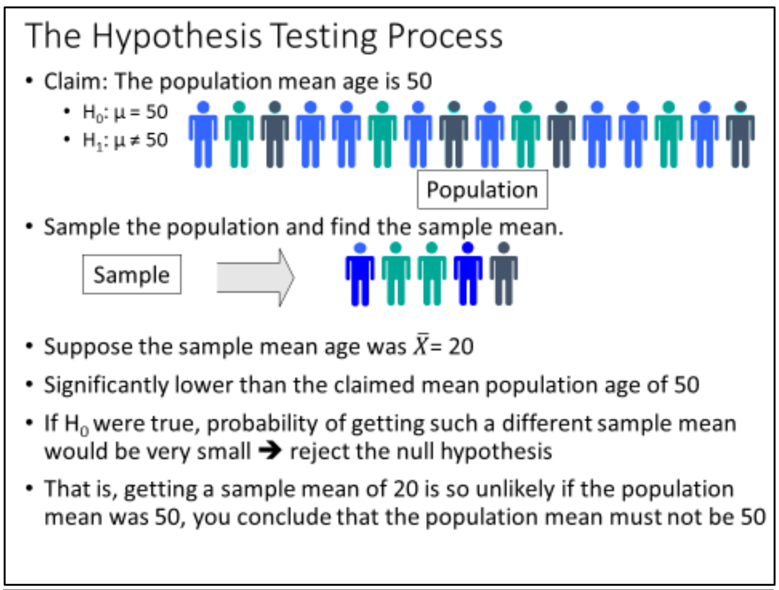
- ASSUMPTION of true $\mathbf{H_0}$ may or may not be REJECTED
  - BUT the ASSUMPTION is NEVER ACCEPTED

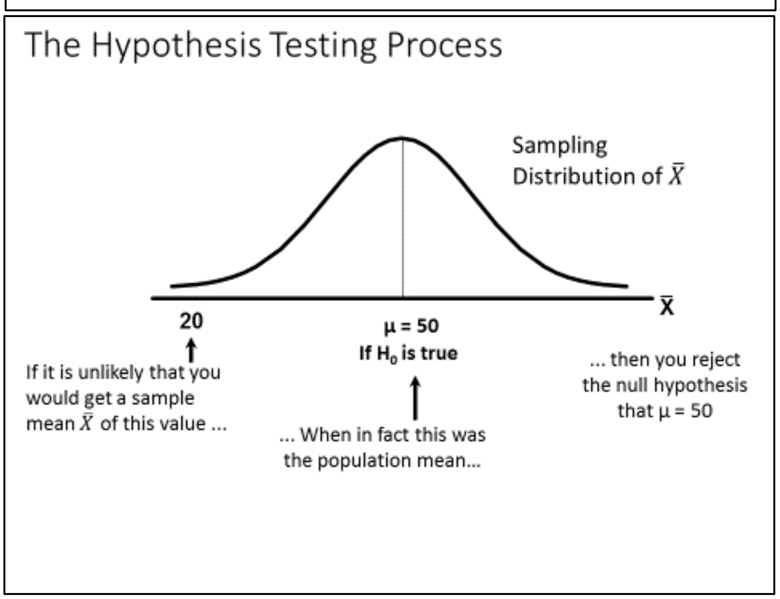
- Examples:
  - Consider a jury trial in which someone is accused of murder
    - $\mathbf{H_0}$:
  - Mean monthly cell phone bill in this city is $\mu$ = $42 (claim by telephone carriers)
    - $\mathbf{H_0}$:
  - Proportion of adults in this city with cell phones is $\boldsymbol{p}$ = 0.68
    - $\mathbf{H_0}$:

- **Alternative Hypothesis, $\mathbf{H_1}$ (or $\mathbf{H_A}$)**
  - Opposite of the null hypothesis
  - Challenges the status quo
  - Never contains the "=", or "≤", or "≥" sign
    - If $H_0$ contains "=" ➔ $H_1$ must contain "≠"
    - If $H_0$ contains "≤" ➔ $H_1$ must contain ">"
    - If $H_0$ contains "≥" ➔ $H_1$ must contain "<"
  - May or may not be proven
  - Is generally the hypothesis that the **researcher is trying to prove**

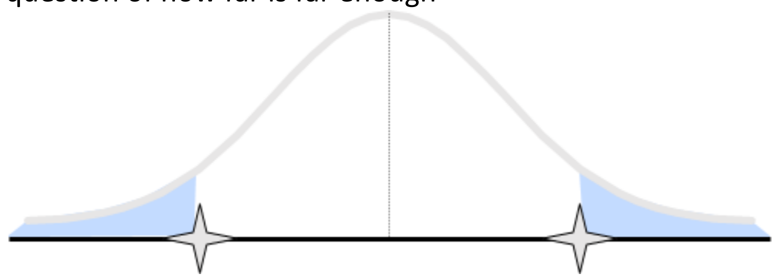
- **Examples:**
  - Consider a jury trial in which someone is accused of murder
    - **$H_0$**: Defendant is innocent
    - **$H_1$**:
  - Mean monthly cell phone bill in this city is $\mu$ = \$42 (claim by telephone carriers)
    - **$H_0$: $\mu$** = \$42
    - **$H_1$**:
    - Proportion of adults in this city with cell phones is ***p*** = 0.68
    - **$H_0$: *p*** = 0.68
    - **$H_1$**:

- **The Hypothesis Testing Process**

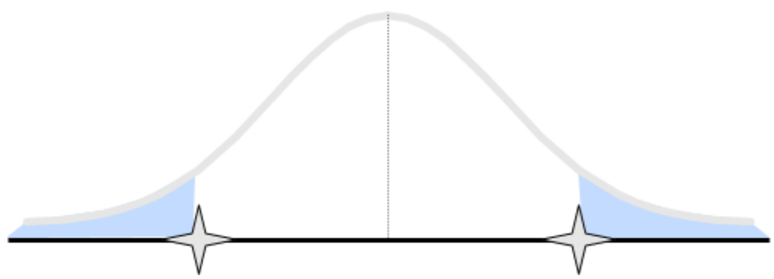


- **Test Statistic and Critical Values**
  - If the sample mean $\bar{X}$ is "close" to the stated (hypothesized) population mean, $H_0$ is NOT rejected
  - If the sample mean $\bar{X}$ is "far" from the stated (hypothesized) population mean, $H_0$ is rejected
  - How far is "far enough" to reject $H_0$?
  - The critical value of a test statistic creates a "line in the sand" for decision making -- it answers the question of how far is far enough



- **Level of Significance (α) and Rejection Region**
  - **α** is the **maximum** value (maximum area in the tails) for which you are willing to reject $H_0$
    - **α** may be spread into both tails ($\frac{\alpha}{2}$) for a 2-tailed test ( = / ≠ )
    - **α** may be concentrated in one tail for a 1-tailed test ( ≤ / > or ≥ / < )
  - Called the **Level of Significance** for the hypothesis test
    - E.g., "At the α=.05 level, we have sufficient evidence to reject…" _or_
      "At the α=.05 level, we do not have sufficient evidence to reject…"

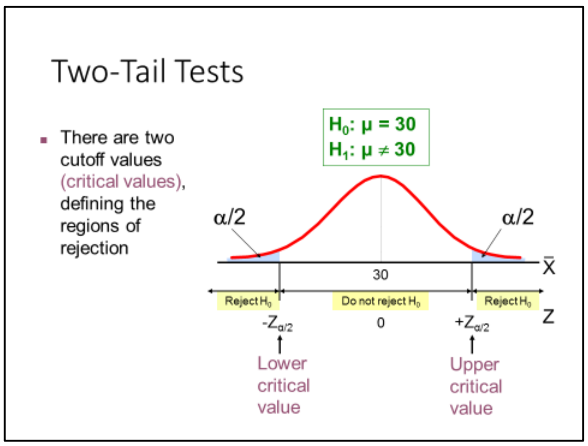


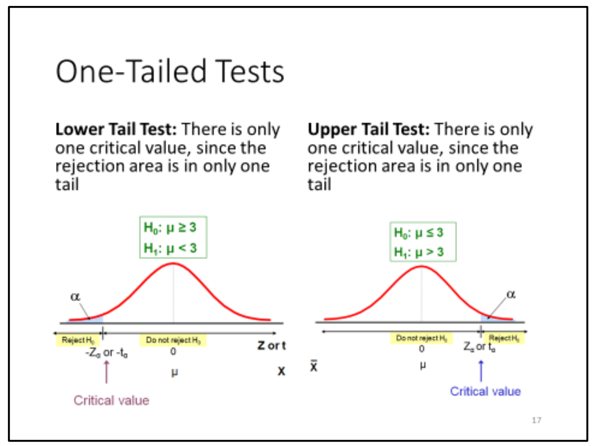
- **5 Steps for <u>Critical Value</u> Approach in Hypothesis Testing**

1. State the null hypothesis, $H_0$ and the alternative hypothesis, $H_1$
2. Choose the level of significance, $\alpha$, and the sample size, n
3. Determine the appropriate test statistic and sampling distribution, calculate test statistic
4. Determine the critical values that divide the rejection and non-rejection regions
5. Make the statistical decision and state the managerial conclusion. If the test statistic falls into the non-rejection 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 Calculations**
  REMEMBER:
  - $z = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$ for sample proportions where n is large enough
  - $z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ for means where σ is known and n≥30 or the underlying population is approximately normal
  - $t = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}$ for means where σ is NOT known and n≥30 or the underlying population is approximately normal
  - **These will be our TEST STATISTICS**

  - Two-Tail Tests One-Tail Tests




- **Example:**
  Gasoline pumped from a supplier's pipeline is supposed to have an octane rating of 87.5. A random sample of 13 days had the following octane readings. Is there evidence, at the .05 level of significance, that the mean octane reading differs from 87.5? ($\bar{X} = 87.08,\ S = 0.649$)
  88.6 86.4 87.2 87.4 87.2 87.6 86.8 86.1 87.4 87.3 86.4 86.6 87.1

- **Example:**
  An engineer wishes to prove that the mean weight of metal components produced by a process is greater than 4.5 oz. A random sample of 10 components produces the data below. Test the engineer's hypothesis using a 0.05 level of significance.
  $(\bar{X} = 4.59,\ \ S = 0.504)$
  4.5 5.6 4.9 3.8 4.1 4.3 4.4 4.7 5.0 4.6

- **Example:**
  A doctor is researching side effects with a new pain medication. A clinical trial including random sample of 340 people who took a new pain relief medication reveals that 23 suffered some side effects. At the α=.05 level of significance, is there evidence that less than 10% of all patients who take the medication will experience side effects?

- **P-Value Approach to Testing**
  - **P-value**:

    - Remember: $\alpha$ is the largest area in the tail(s) for which $H_0$ can be rejected
      - If p-value $<$ $\alpha$ ,

      - If p-value $\geq$ $\alpha$ ,

**P-Value Approach to Testing:**

- **The 5 Step <u>p-value approach</u> to Hypothesis Testing**

1. State the null hypothesis, $H_0$ and the alternative hypothesis, $H_1$

2. Choose the level of significance, $\alpha$, and the sample size, n

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

- **Example:**
  Gasoline pumped from a supplier's pipeline is supposed to have an octane rating of 87.5. A random sample of 13 days had the following octane readings. Is there evidence, at the .05 level of significance, that the mean octane reading differs from 87.5? ($\bar{X} = 87.08$, $S = 0.649$)
  88.6 86.4 87.2 87.4 87.2 87.6 86.8 86.1 87.4 87.3 86.4 86.6 87.1

- **Example:**
  An engineer wishes to prove that the mean weight of metal components produced by a process is greater than 4.5 oz. A random sample of 10 components produces the data below. Test the engineer's hypothesis using a 0.05 level of significance.
  $(\bar{X} = 4.59,\ \ S = 0.504)$
  4.5 5.6 4.9 3.8 4.1 4.3 4.4 4.7 5.0 4.6

- **Example:**
  A doctor is researching side effects with a new pain medication. A clinical trial including random sample of 340 people who took a new pain relief medication reveals that 23 suffered some side effects. At the α=.05 level of significance, is there evidence that less than 10% of all patients who take the medication will experience side effects?

- **What can go wrong?**
  - When dealing with samples and levels of confidence, there is always the possibility of making an error in our hypothesis (significance) test

  - Consider a jury trial in which someone is accused of murder
    - $H_0$: Defendant is innocent
    - $H_A$: Defendant is NOT innocent
    - If the defendant is innocent but is wrongly convicted, a true $H_0$ was rejected in favor of a false $H_A$
    - If the defendant found NOT guilty but truly committed the murder, a false $H_0$ was NOT rejected. That is, we failed to reject a false $H_0$ in favor of a true $H_A$

- **Errors in Hypothesis Testing**
  - **Type I Error**

  - **Type II Error**

  - **POWER** of a statistical test (1-$\boldsymbol{\beta}$) is the probability of rejecting $H_0$ when it is false)
    - POWER =

- **Type I and Type II Errors**

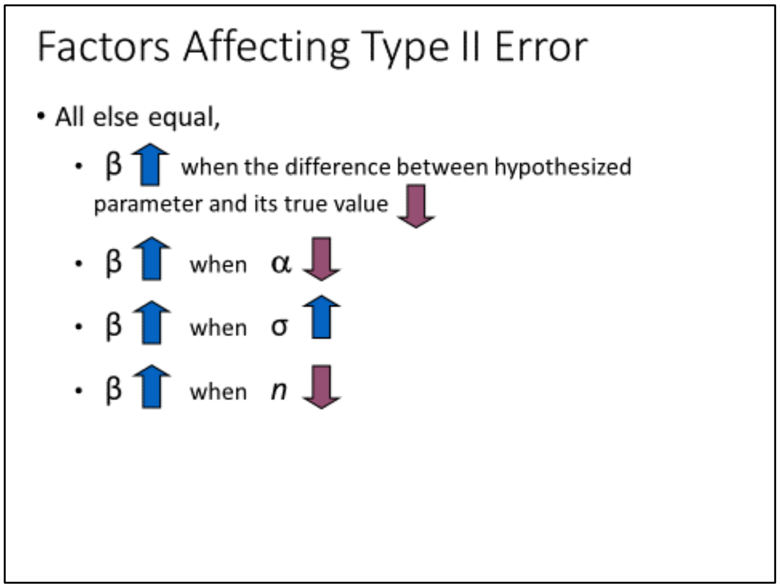
| | **Do NOT Reject $H_0$** | **Reject $H_0$** |
|---|---|---|
| **$H_0$ True** | | |
| **$H_0$ False** | | |

  - We can control the probability of a Type I error by our choice of the significance level

  - The more serious the consequences of a Type I error, the smaller α should be

  - As P(Type I Error) = $\boldsymbol{\alpha}$ goes *Down*, P(Type II Error) goes *Up*

- **Factors Affecting Type II Error**




- **Example 1**
  A planning committee needs to estimate the percentage of students at a large university who will attend an upcoming event so that they can determine an appropriate location for the event. Data is collected to see if there is evidence, at the .05 level of significance, that less than 30% of students will attend this year?
  - Describe a Type I Error for this problem:

  - Potential consequence:

  - Describe a Type II Error for this problem:

  - Potential consequence:

- **Example 2**

An environmentalist takes samples at a nearby river to study the average concentration level of a contaminant. He wants to find out, using a .10 level of significance, if the average concentration level exceeds the acceptable level for safely consuming fish from the river.

- Describe a Type I Error for this problem:

- Potential consequence:

- Describe a Type II Error for this problem:

- Potential consequence:

**Question:** Many consumer groups feel that the U.S. Food and Drug Administration (FDA) drug approval process is too easy and, as a result, too many drugs are approved that are later found to be unsafe. On the other hand, a number of industry lobbyists have pushed for a more lenient approval process so that pharmaceutical companies can get new drugs approved more easily and quickly. Consider the hypotheses

**$H_0$: new, unapproved drug is unsafe**
**$H_A$: new, unapproved drug is safe**

- What is a Type I error in this situation?

    A. Reject $H_0$ that the drug is unsafe when, in fact, $H_0$ is true.
    That is, the drug IS UNSAFE.

    B. Fail to reject $H_0$ that the drug is unsafe when, in fact, $H_0$ is false.
    That is, the drug is SAFE

- What are the consequences of the Type I error?

- What are the consequences of the Type II error

**Question:** Bottles of water have a label stating that the volume is 12 oz. A consumer group suspects the bottles are under-filled and plans to conduct a test using α=0.10. A **Type I** error in this situation would mean:

A. The consumer group concludes the bottles have less than 12 oz. when the true mean amount actually is 12 oz. or more

B. The consumer group does not conclude the bottles have less than 12 oz. when the true mean amount
actually is less than 12 oz.

**Question:** You are the manager of a restaurant that delivers pizza to college dormitory rooms. You have just changed your delivery process in an effort to reduce the mean time between the order and the completion of delivery from the current 25 minutes. A sample of 36 orders using the new delivery process yields a sample mean of 22.4 minutes and a sample standard deviation of 6 minutes. What is the **test statistic** for this test?

A. $z = \frac{\bar{X}-\mu}{\sigma}$

B. $z = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$

C. $z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$

D. $t = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}$

**Question:** You are the manager of a restaurant that delivers pizza to college dormitory rooms. You have just changed your delivery process in an effort to reduce the mean time between the order and the completion of delivery from the current 25 minutes. A sample of 36 orders using the new delivery process yields a sample mean of 22.4 minutes and a sample standard deviation of 6 minutes. What **assumption** do you need to make to perform this test?

A. Sample size n=36>30

B. Underlying data distribution of delivery times is approximately normal

C. Population standard deviation, $\sigma$, is 6 minutes

D. This is a two-side test