Section 7.1 - Introduction to Hypothesis Testing

1. Hypothesis testing: Inferential statistics that test a claim about a parameter.

2. Example:
   We are asked to test a claim that the proportion of U.S. students age 8 to 18 who are aware that software is protected by copyright law is $p = 0.80$. We take a random sample of $n = 1183$ students and find that 1017 of them are aware that software is protected by copyright law. We see that $\hat{p} = 0.86$.

3. 2 most common types of statistical inference:
   - **Confidence interval** interval estimates a population parameter
   - **Hypothesis test** assesses evidence provided by data about population claim

4. Uses of hypothesis testing
   - Marketing division wants to know if a new ad campaign outperforms an old one.
   - Courts inquire about statistical significance in class action discrimination cases.
   - Medical researchers want to know whether a new drug/therapy performs better.

5. Determining the hypotheses.
   - **Null Hypothesis**: an assumption concerning the value of the population parameter being studied (usually represents no effect, no change, no difference, etc.; $\leq, \geq, =$)
     Notation: H_0
   - **Alternative Hypothesis**: the complement of the null hypothesis (must be true if H_0 is false). (states the result for which we hope to find evidence; $<, >, \neq$)
     Notation: H_a

   Note: The null hypothesis may or may not be true. We will carry out a study and then determine if we have strong enough evidence to conclude that the null hypothesis is false (meaning our evidence suggests that H_a is true).

   (a) Example #1:
   A research meteorologist has been studying wind patterns over the Pacific Ocean. Based on these studies, a new route is proposed for commercial airlines going from San Francisco to Honolulu. The new route is intended to reduce flying time by taking advantage of existing wind patterns. It is known that for the old route, the distribution of flying times for a large, four-engine jet has a mean of $\mu = 5.25$ hrs with a standard deviation of 0.6 hours. What null and alternative hypotheses should you use in order to determine if the new route has decreased the average flying time?

   (b) Example #2:
   At a particular state university, the entering students recently have averaged a score of 1000 on their SATs with a standard deviation of 180. Suppose you wish to see if this year's applying students have a different mean SAT score. What null and alternative hypotheses should you use?

(c) Example #3:
Last year, your company's service technicians took an average of 2.6 hours to respond to trouble calls from business customers who had purchased service contracts. What null and alternative hypotheses should you test in order to answer the question, "Does this year's data show a significantly different average response time?"

6. Possible errors in hypothesis testing

Decision	H_0 is true	H_a is true
Reject H_0		
Fail to Reject H_0		

7. **Type I Error**: We reject H_0 when it is true.

8. **Type II Error**: We fail to reject H_0 when it is false.

9. Example #4: What are possible Type I and Type II errors for the situation described in Example #1?

10. **Level of Significance** (α): The maximum allowable probability of making a Type I Error.

11. **Test statistic**: Obtain a simple random sample of n observations from the desired population and calculate the observed sample statistic.

    For example, if we want to test something about a population mean (μ), then we would calculate the sample mean ($\bar{x}$). If we want to test something about a population proportion (p), then we would calculate the sample proportion ($\hat{p}$).

12. Three types of hypothesis tests (p. 338)

    - **left-tailed test**
    - **right-tailed test**
    - **two-tailed test**

13. Example: For Examples #1-3, determine whether each test is left-tailed, right-tailed, or two-tailed.

14. Determining the "strength" of your evidence.

    - P**-value**: the probability of obtaining a sample outcome as extreme or more extreme than the actual observed outcome, *assuming that the null is true.*
    - The evidence is **strong** if the outcome we observe would rarely occur assuming the null hypothesis is true (meaning it is more probable that the alternative hypothesis is true).
    - The evidence is **weak** if the outcome we observe has a high probability of occurring assuming the null hypothesis is true.

    - The P-value measures the strength of the evidence.
        - Smaller P-value $\Rightarrow$ stronger evidence is against H_0.
        - P-value describes the risk of making a mistake if we wrongly reject the null.

15. Draw a conclusion.

    - If the P-value is "small" ($P \leq \alpha$), then we reject H_0 in favor of H_A.
    - If the P-value is "large",($P > \alpha$) then we fail to reject H_0, meaning we cannot conclude H_a.

    Note: You may NEVER conclude that the null is true.