Hypothesis Testing:

1) State the null and alternative hypotheses (one-tailed or two-tailed alternative?)
2) Explain the potential Type I and Type II errors & decide upon an _-level (usually .01, .05, or .10)
3) Calculate an observed test statistic (in this class, either t_{obs} or z_{obs})
4) Sketch the distribution of your test statistic (under H_0) and locate the critical/rejection region
5) If your observed statistic falls in the critical region, reject the null hypothesis

Recommended Steps:

6) Calculate a p-value: P(observing a test statistic that extreme if the null hypothesis is true)
7) Calculate a confidence interval for the population parameter.

Example: The average weight of a watermelon is 18 pounds. A farmer, using a new bio-engineered seed, samples 20 watermelons and calculates an average weight of 19.83 pounds (with a standard deviation of 3.4 pounds). Using α=0.01, determine if the new seed produces significantly heavier watermelons.

1) H_0: $\mu \leq 18$. The new seed does not produce larger watermelons

H_A: $\mu > 18$. The new seed does produce larger watermelons. This is a one-tailed test.

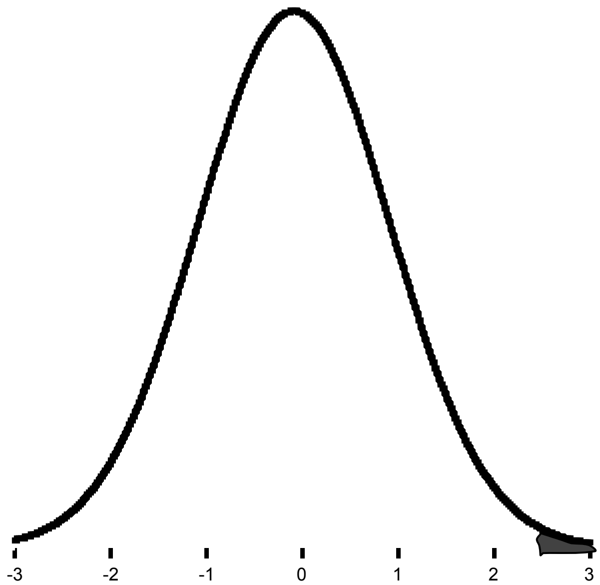
2)

		Our Decision	
		H_0 is true. The new seed is not better.	H_A is true. The new seed is better.
Reality	H_0 is true. Decide the new seed is not better	Correct Decision	_ = Type I Error We incorrectly conclude the new seed is not better We could have had heavier fruit.
	H_A is true. Decide the new seed is better.	_ = Type II Error We incorrectly conclude the new seed is better We buy potentially more expensive seeds for no reason	Correct Decision. Power = 1 – _

3) We only have one sample of observations and σ is unknown, so our test statistic will be a single-sample t-test with n-1 d.f.

$$t_{obs} = \frac{19.83 - 18}{\frac{3.4}{\sqrt{20}}} = \frac{1.83}{0.76} = 2.408$$ is distributed as a t-statistic with 19 degrees of freedom

4) The t-distribution is roughly normal (it looks like the standard normal distribution with fatter tails).



According to my alternate hypothesis, I want to reject H_0 if my observed test statistic is too big. Therefore, I need to label 1% of the distribution as my critical region. Looking in my t-table (page 675 with v = 19 and _ = .01) I see that the critical value of t is 2.539. If it turns out the H_0 is true, I will reject it 1% of the time (the shaded area represents Type I error).

5) My observed test statistic was calculated to be: 2.408. This does not fall in the critical region, so I fail to reject the null hypothesis.

6) P-value = P(observing a t greater than 2.408 given that the null hypothesis is true).

$P(t > 2.408)$ ➔ Look in the t-table and see… $P(t > 2.539) = 0.01$ and $P(t > 2.093) = .025$

So: $.025 < p < .01$

There is a less than 2.5% chance of observing watermelon 19.83 pounds or heavier if the new seeds are worthless

7) It must be true that if I calculate a one-sided 99% confidence interval for the population mean weight of watermelons from my observed sample, the hypothesized population mean will fall within the interval.

$$\overline{X} - t_{\alpha,n-1}\left(\frac{s}{\sqrt{n}}\right) = 19.83 - 2.539\left(\frac{3.4}{\sqrt{20}}\right) = 19.83 - 1.93 = 17.90$$

The one-sided 99% CI is: (17.90 and higher). Since the hypothesized population mean falls within that region, I can't say that the null hypothesis is false.

Complete all 7 steps for the following hypothesis tests:

Group #1: You are the head of the math department at Davenport West high school. You want to see if graphing calculators improve achievement for Algebra II students. Over the past several years, the average score on the Algebra II final exam has been 75.4. This year, you force all teachers to use graphing calculators in their Algebra II classes. At the end of the year, the average final exam score (for 121 students) was 81 points with a standard deviation of 9.2 points. Does it appear as though graphing calculators had an impact on achievement?

Group #2: According to the International Diabetes Research Foundation, an individual has diabetes if their blood glucose concentration is at or above 200 mg/dl. Over the course of 3 months, you sample your blood 25 times and obtain an average glucose concentration of 201.38 with a standard deviation of 7.348 mg/dl. Test the hypothesis that you have diabetes based on your blood glucose concentration.

Optional: The mean monthly wireless phone bill in 1999 was \$45.15. The CEO of Verizon claims the average bill in 2003 is substantially lower. You survey 49 wireless phone users and determine the mean bill to be \$40.24. Assuming σ is \$21.20, test the CEO's claim.

After we discuss hypothesis tests for population proportions, complete the following:

Group #1: Your friend declares himself to be the world's greatest player of "rock-paper-scissors." You decide to challenge him for that title, so you play 510 times. He beats you 183 times in those 510 trials. Is your friend better than average at "rock-paper-scissors?" Use a 5% significance test and then calculate a 95% CI.

Group #2: In order to build a new library, voters in the city of Davenport must pass a levy. City officials conducted a poll of 25 random individuals and found that 64% supported the levy. Can the city be confident that the levy will pass (assuming it only needs a simple majority)? Use a 5% significance test. Next test the same hypothesis, assuming the city polled 200 individuals and found that 64% supported the levy.

Optional: Two competing drugs, Lipitor and Zocor, have been shown to reduce total cholesterol and LDL-cholesterol. It is known that 1.9% of all Zocor users complain of flu-like symptoms when taking the drug. In clinical trials, 11 out of 863 patients taking 10mg of Lipitor daily complained of flu-like symptoms. Is there evidence to support the claim that fewer Lipitor users experience flu-like symptoms?

All of these hypothesis tests were **single-sample hypothesis tests**. We compared a single sample statistic to a hypothesized value of its corresponding parameter.

Our next goal will be to conduct two-sample hypothesis tests (tests in which we compare the means or proportions between two groups).