Lesson 4. Hypothesis Testing – Part 1

1 Hypothesis testing for one population mean

- The goal of **hypothesis testing** is to test competing claims about a parameter
  - If the original claim is true, how likely is it that we would see our data (or something more extreme)?
- Let's focus on when the parameter is a population mean

Example 1. A Keurig machine is supposed to output 6 ounces of coffee when the smallest size is selected. For quality control, one machine is selected to be tested extensively to determine whether its average output is actually 6 ounces. The mean output of 20 cups of coffee is 6.1 ounces, and the standard deviation is 0.3 ounces. Use a significance level of 0.10 to test whether this machine's average output differs from 6 ounces.

General Steps	For Example 1
Step 1: State hypotheses	
State the two competing claims in terms of the parameter of interest. H_0 is the **null hypothesis**, the default. H_A is the **alternative hypothesis**. Equality goes in the null.	
Step 2: Calculate test statistic	
$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$ The test statistic summarizes how far the data is from the null claim, and it is standardized assuming the null is true.	
Step 3: Calculate p-value	
If H_0 is true, our test statistic comes from a t-distribution with $df = n - 1$. The p-value quantifies where our test statistic is in this null distribution. The p-value is a conditional probability: the probability of seeing at least as extreme (in the direction of H_A) a test statistic as ours, **given that H_0 is true**.	

General Steps	For Example 1
Step 4: State conclusion	
A small p-value is evidence against H_0. If the p-value is small enough, we **reject** H_0. Otherwise, we **fail to reject** H_0**.** How small is small enough? The significance level α. We also state the conclusion in context in terms of evidence.	

1.1 Technical conditions to check

- We have the same conditions to check as for a confidence interval for a mean:
    1. Data must be from a simple random sample
    2. Either the population distribution is Normal or the sample size (n) is large

1.2 Statistical significance vs. practical importance

- "Significance" in the statistical sense $\longleftrightarrow$ size of the p-value
- Suppose that we concluded that the sample mean output was (statistically) significantly different from 6 ounces
- This means that the sample mean output is very unlikely to have occured by chance if $\mu = 6$
- It says nothing about the size of the difference

1.3 Type I and Type II errors

	Reject H_0	Fail to reject H_0
H_0 true		
H_0 false		

We control the probability of a Type I error by setting the significance level α