Stat 13, Intro. to Statistical Methods for the Life and Health Sciences.

1. Estimating the mean using the t-test, and ABC song example.
2. Significance level.
3. Type I and Type II errors.

Read chapters 2 and 3.
http://www.stat.ucla.edu/~frederic/13/W23 .
HW2 is due Fri Feb10 at 2pm and is problems 2.3.15, 3.3.18, and 4.1.23.
See day4 notes to make sure you are doing the correct problems.

https://www.youtube.com/watch?v=ho7796-au8U

1. Inference for a Single Quantitative Variable

Section 2.2

Example 2.2:
Estimating Elapsed Time

- Students in a stats class (for their final project) collected data on students' perception of time
- Subjects were told that they'd listen to music and be asked questions when it was over.
- 10 seconds of the Jackson 5's "ABC" and subjects were asked how long they thought it lasted
- Can students accurately estimate the length?

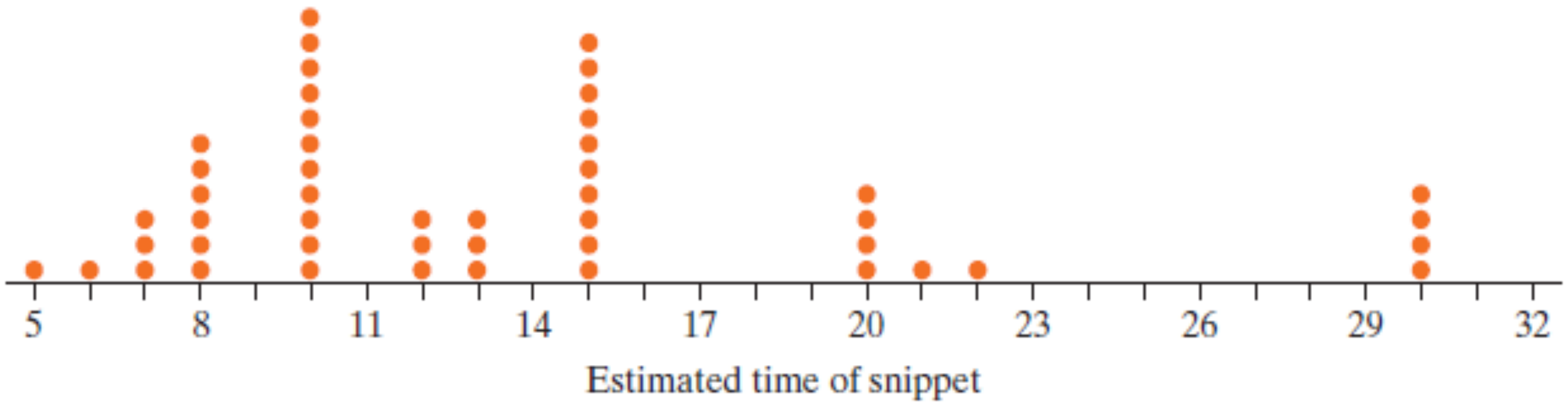
Hypotheses

Null Hypothesis: People will accurately estimate the length of a 10 second-song snippet, on average. ($\mu = 10$ seconds)

Alternative Hypothesis: People will not accurately estimate the length of a 10 second-song snippet, on average. ($\mu \neq 10$ seconds)

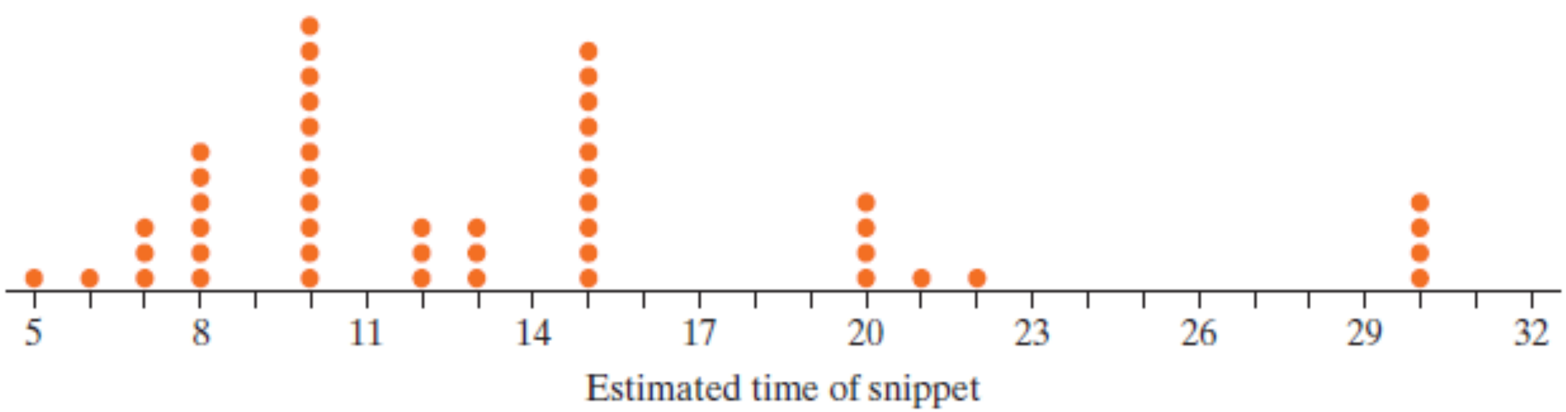
Estimating Time

- A sample of 48 students on campus were subjects and song length estimates were recorded.
- What does a single dot represent?
- What are the observational units? Variable?



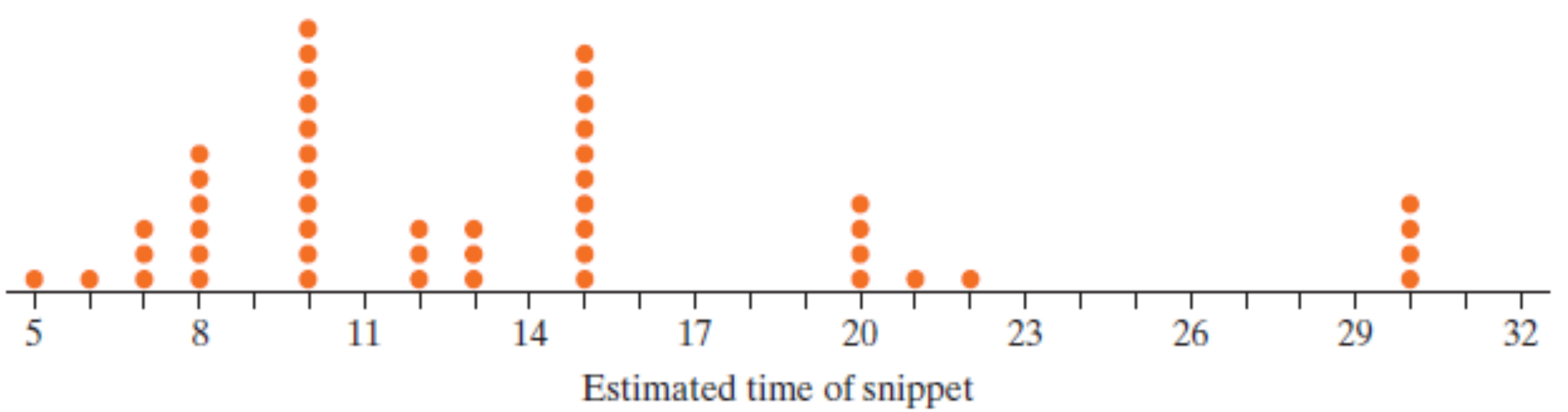
Skewed, mean, median

- The distribution obtained is not symmetric, but is **right skewed.**
- When data are skewed right, the **mean** gets pulled out to the right while the **median** is more resistant to this.



Mean vs Median

- The mean is 13.71 and the median is 12.
- How would these numbers change if one of the people that gave an answer of 30 seconds actually said 300 seconds?
- The standard deviation is 6.5 sec. Also not resistant to outliers.



Inference

- $\mathbf{H_0}$: $\mu = 10$ seconds
- $\mathbf{H_a}$: $\mu \neq 10$ seconds
- Our problem now is, how do we develop a null distribution?
  - Here we don't have population data that reflects our null hypothesis where $\mu = 10$ seconds.
  - All we have is our sample of 48.

Population?

- We need to come up with a large data set that we think our population of time estimates might look like **under a true null.**
- We might assume the population is skewed (like our sample) and has a standard deviation similar to what we found in our sample, but has a mean of 10 seconds.
- The book recommends using an applet for this. We could use *R*, or do a (theory-based) t-test.

Theory-Based Test

- Using simulations to create a population each time we want to run a test of significance is extremely time consuming and cumbersome.
- The null distribution that we developed can be predicted with theory-based methods.
- We know it will be centered on the mean given in the null hypothesis.
- We can also predict its shape and its standard deviation.

t-distribution

- The shape is very much like a normal distribution, but slightly wider in the tails and is called a t-distribution.
- The t-statistic is the standardized statistic we use with a single quantitative variable that looks approximately normal, when the sample size is small, and the statistic can be found using the formula:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

The $s/\sqrt{n}$ (standard deviation of our sample divided by the square root of the sample size) is called the standard error and is an estimate for the standard deviation of the null distribution.

Here $t = \frac{13.71 - 10.0}{6.5/\sqrt{48}} = 3.95.$

p-value = 2*(1-pt(3.95,df=47)) = 0.000261.

Validity Conditions

- The observations must be independent.
- The population must be normally distributed!
- The book says you need the sample size to be at least 20 for the t-test, but this is not technically true. The whole point of the t-test is you can use it even when your sample size is small, provided the two assumptions above hold.

But it is often hard to have any idea if the population is normal without having at least 20 observations.

Estimating Time

Formulate Conclusions.

- Based on our small p-value, we can conclude that our subjects did not accurately estimate the length of a 10-second song snippet and in fact they significantly overestimated it.
- How far can we generalize this?

Summary

- When we test a single quantitative variable, our hypothesis has the following form:
  - **H_0**: μ = some number
  - **H_a**: $\mu \neq$ some number, $\mu <$ something or $\mu >$ something.
- We can get our data (or mean, sample size, and SD for our data) and use the Theory-Based Inference to determine the p-value.
- The p-value we get with this test has the same general meaning as from a test for a single proportion.

2. Significance level

Section 2.3

Significance Level

- We think of a p-value as telling us something about the strength of evidence from a test of significance.
- The lower the p-value the stronger the evidence.
- Some people think of this in more black and white terms. Either we reject the null or not.

Significance Level

- The value that we use to determine how small a p-value needs to be to provide convincing evidence whether or not to reject the null hypothesis is called the **significance level**.
- We reject the null when the p-value is less than or equal to (≤) the significance level.
- The significance level is often represented by the Greek letter alpha, α.

Significance Level

- Typically we use 0.05 for our significance level. There is nothing magical about 0.05. We could set up our test to make it
  - harder to reject the null (smaller significance level say 0.01) or
  - easier (larger significance level say 0.10).

3. Type I and Type II errors

- In medical tests:
  - A type I error is a false positive. (conclude someone has a disease when they don't.)
  - A type II error is a false negative. (conclude someone does not have a disease when they actually do.)
- These types of errors can have very different consequences.

Type I and Type II Errors

TABLE 2.9 A summary of Type I and Type II errors

		What is true (unknown to us)	
		Null hypothesis is true	Null hypothesis is false
What we decide (based on data)	Reject null hypothesis	Type I error (false alarm)	Correct decision
	Do not reject null hypothesis	Correct decision	Type II error (missed opportunity)

•

Type I and Type II errors

TABLE 2.10 Type I and Type II errors summarized in context of jury trial

		What is true (unknown to the jury)	
		Null hypothesis is true (defendant is innocent)	Null hypothesis is false (defendant is guilty)
What jury decides (based on evi-dence)	Reject null hypothesis (Jury finds defendant guilty)	Type I error (false alarm)	Correct decision
	Do not reject null hypothe-sis (Jury finds defendant not guilty)	Correct decision	Type II error (missed opportunity)