

Lesson 49: Confidence Intervals with standard deviation unknown

Opener: insurance article

Robust Procedures: A confidence interval or hypothesis test is called robust if the confidence level or p-value does not change very much when the assumptions of the procedure are violated. This means that the conditions to use the t-test are not met (such as a non-normal population), but the results are still fairly accurate.

- The t procedure is quite robust against non-normality of the population
- The t procedure is NOT robust when there are outliers
- Results improve as sample size, n, increases

General Rules based on sample size:

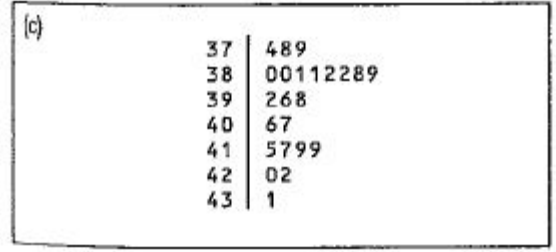
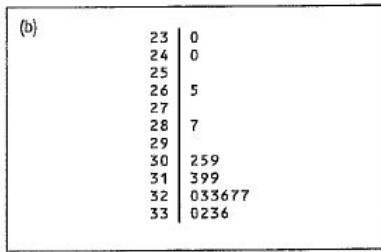
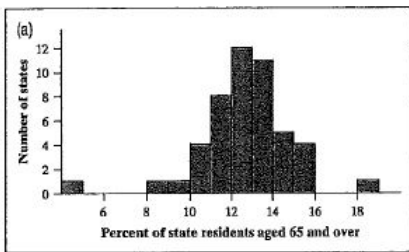
n < 15 only use t-test if population is normal and there are no outliers

15 < n < 30 use t-test unless there are outliers or the population has a strong skew

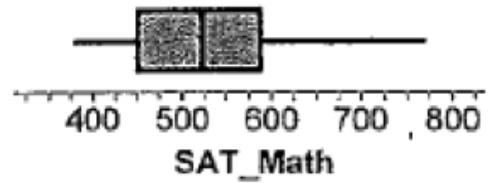
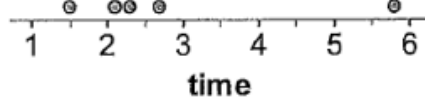
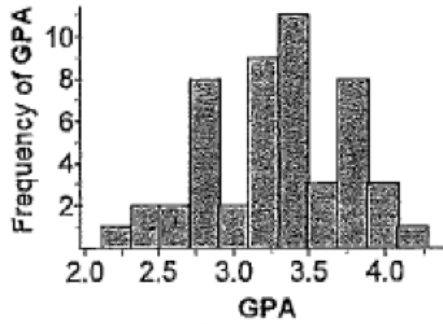
n > 30 use t-test even when skewed, beware of outliers

Use a histogram or boxplot to check for skewness and outliers

Examples – can we use construct a t-interval?



Histogram of the percent of each state's residents who are at least 65 years of age.	Stemplot of the force required to lift 20 boxes at a factory.	Stemplot of the length of a 23 Heliconia flowers.



Histogram of the GPA for 50 randomly selected students.	Dotplot for the amount of time (min) it takes to receive a coffee at a coffee shop.	Boxplot shows the SAT score for a random sample of 20 students.

Example

George is attempting to find out how many text messages the average NWHHS student sends in a month. He takes an SRS of 15 NWHHS students and gets the following numbers

107	116	153	56	44
187	36	145	107	29
109	87	138	121	91

- Since our sample size is less than 30 we need to make sure we do not have outliers and that our sample comes from a normal population. Create a box-plot, a histogram, and a normal probability plot.
- construct a 95% confidence interval for the average number of text messages sent in a month.

Daily Data Collection

Run a matched pair study and find a 95% confidence interval for the true mean difference in "The average hours of sleep on a weekend night and a week night"

Person	Weekend night sleep hours	Week night sleep hours	Difference Weekend - Week	Person	Weekend night sleep hours	Week night sleep hours	Difference Weekend - Week
1				16			
2				17			
3				18			
4				19			
5				20			
6				21			
7				22			
8				23			
9				24			
10				25			
11				26			
12				27			
13				28			
14				29			
15				30			

Parts of the formula:

$$\bar{x} = \underline{\hspace{2cm}} \quad n = \underline{\hspace{2cm}} \quad s_x = \underline{\hspace{2cm}}$$

$$\text{Degrees of freedom} = \underline{\hspace{2cm}} \quad t = \underline{\hspace{2cm}}$$

Interval:

$$\text{Margin of Error} = \underline{\hspace{2cm}}$$

$$\text{Lower} = \underline{\hspace{2cm}}$$

$$\text{Upper} = \underline{\hspace{2cm}}$$

Statement:

Conclusion: