

Preparation Survey

Name: **Key**
 Academic advisor:
 Year in program:
 Research interest:

Computer Information

This course requires the use of a computer please fill in the following information about your computer (and your computer ability) so that I can be better prepared for certain questions. With regard to your primary computer:

What platform is it? (**Bold** one) Mac: **2** PC: **5** Unix/Linux: **0**

Which operating system does it use? _____

Which Internet browser do you use? (Make your primary **RED**, **bold** all you use)

Internet Explorer:	v7 1	v8 4	v9
Firefox:	v3.5	v4 3	v5 1
Opera:	v9	v10	v11
Chrome:	v11	v12	v13 1
Safari:	v4	v5 2	
Other (specify):	_____		

Software Knowledge

As we will be performing real data analysis (and interpretation and presentation), it is necessary that you are able to use certain computer programs. Which of the following programs can you use? Also, for each, please estimate your ability level.

Program Type	Name	Ability Level (Bold one)			
Word processor	MS	None	Low	Medium 2	High 5
Spreadsheet	MS	None	Low	Medium 5	High 2
Presentation	MS	None	Low	Medium 2	High 5
Statistical environments:					
	R	None 4	Low 1	Medium 2	High
	S-Plus	None 7	Low	Medium	High
	SAS	None 4	Low 2	Medium 1	High
	SPSS	None 1	Low 4	Medium 2	High
	STATA	None 5	Low 1	Medium	High

Prerequisite Skillset

The prerequisite for this course is POLS 5013: Quantitative Methods of Political Analysis. In that course, several important statistical topics were covered. To help determine if your statistical preparation is sufficient, please answer the questions in this section.

Definitions

- Type I Error rate: **The probability that you will reject a true null hypothesis**
- The p-value: **The probability that you will obtain results (data) as extreme as the observations — or more extreme — given that the null hypothesis is true**
- The power of a test: **The probability of being able to distinguish between the null and alternative hypotheses**
The probability that you will be able to reject a false null hypothesis
- Correlation: **The level of linear relationship between two variables**
- Heteroskedasticity: **Non-constant variance**

True or False

1. **False:** If the correlation between two variables is zero, then the variables are independent.
2. **False:** An R^2 value of -1.00 indicates that there is no relationship between the two variables.
3. **True:** A p-value less than our selected alpha value indicates a statistically significant relationship.
4. **True:** Three measures of central tendency are the Mean, the Median, and the Mode.
5. **False:** If the data is skewed, the mode is the most appropriate measure of central tendency.

6. **True:** Cross-tabulations are most appropriate for comparing two discrete variables.
7. **False:** Histograms are useful in exploring (and indicating) relationships between two continuous variables.
8. **True:** Two measures of variability are the Inter-Quartile Range (IQR) and the variance.
9. **True:** To determine statistical independence between two discrete variables, chi-squared tests are often used.
10. **False:** To determine whether two means are different, one should use the Analysis of Variance test (ANOVA).

Performance

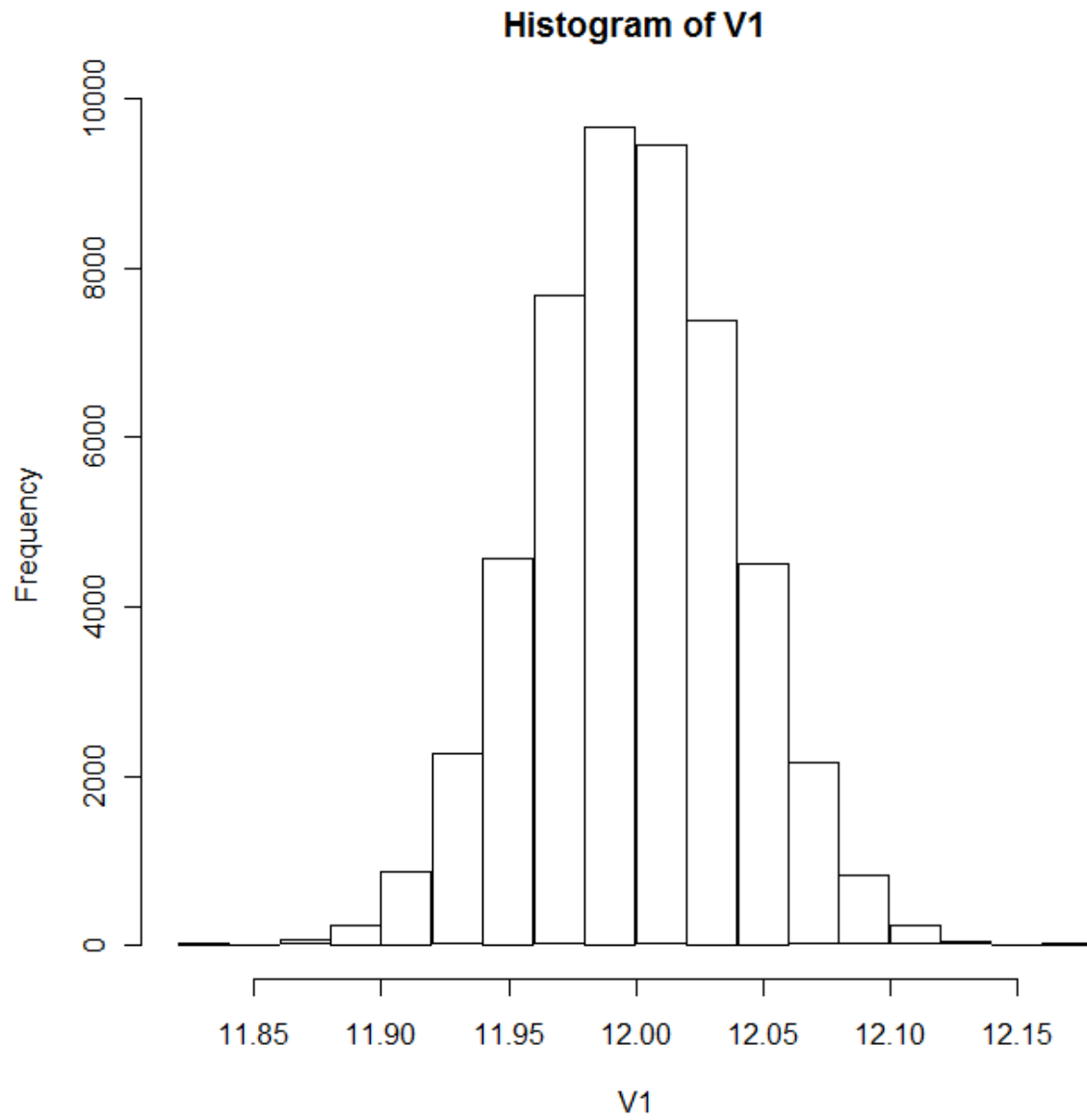
There is a data file located at:

<http://courses.kvasaheim.com/pols6123/data/positioningtubes.csv>

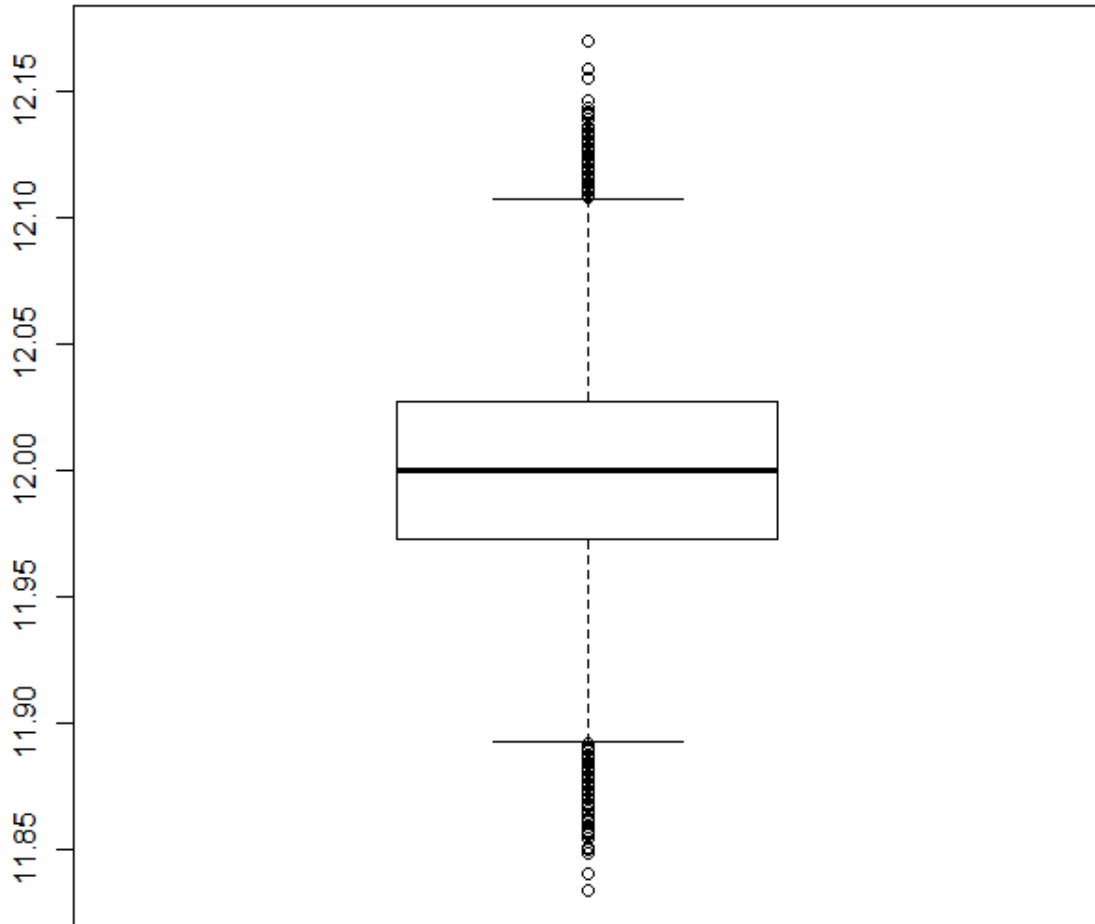
1. Please calculate the following statistics from the data:

Mean	<u>12.000</u>	Variance	<u>0.0016</u>
Median	<u>12</u>	Standard deviation	<u>0.0400</u>
Second Quartile	<u>12</u>	Inter-Quartile range	<u>0.054</u>
Minimum	<u>11.834</u>	Range	<u>0.336</u>
Maximum	<u>12.17</u>	Third Quartile	<u>12.027</u>

2. Please create the following graphs and include them in this document:
a. Histogram:



b. Boxplot:



3. Let us assume our null hypothesis is

H_0 : The population mean is 12.25.

At the usual alpha level ($\alpha=0.05$), do we conclude that the data supports the null hypothesis? Make sure you state the test statistic and the p-value in your conclusion.

The null hypothesis is that the mean of the population from which this data came is 12.25. As we do not know the population variance, we perform a one-sample t-test. The results of this test ($t=-1399$; $df=49999$; $p \ll 0.0001$) suggest that we should reject the null hypothesis at the $\alpha=0.05$ level and conclude that the population mean is not 12.25.

4. Which statistical program did you use to answer these last three questions?

I used R, as did two others. Three used SPSS. Two used Excel.

Finally

Now, please make sure your name and the page number is located at the top-right of each page, except for the first. Save this file as `yourlastname-1022.doc` (use *your* last name in place of `yourlastname`) and email it to your professor as an attachment by August 21, 2011. The subject line for the email needs to be `POLS6123: Prep Survey`. If you fail to do any of these steps, I will assume that you *cannot* do that step.