Midterm 2 Practice Questions

Online S520

1. Wilder and Rypstra (2004) tested the effect of praying mantis excrement on the behavior of wolf spiders. They put 12 wolf spiders in individual containers. Each container had two semicircles of filter paper: one semicircle that had been smeared with praying mantis excrement, and one without excrement. They observed each spider for one hour, and measured its walking speed while it was on each half of the container. They used a t-test at level $\alpha = 0.05$ to see if, on average, there was a difference between walking speed on the paper with excrement and on the paper without excrement.

   (a) What is the experimental unit? What measurements are taken on the experimental units?

   (b) Give null and alternative hypotheses for an appropriate two-tailed t-test. If the null hypothesis is true, what is the distribution of the t-statistic?

   (c) The P-value (significance probability) was calculated to be 0.053, so the null hypothesis was not rejected. From this and the other information given, is it correct to conclude that we are sure that wolf spiders' walking speed is not affected by praying mantis excrement? Explain.

2. In a controversial 2011 paper in the Journal of Personality and Social Psychology, a researcher claimed to have found evidence of precognition. In one experiment, Cornell students sat in front of computer screens with two windows. They were asked to click the window that they thought had a picture behind it; after clicking, a picture would then randomly appear behind one of the two windows. In all, the students correctly clicked the window with the picture 1844 out of 3600 times. However, when the pictures were erotic, the students correctly clicked the window with the picture 828 out of 1560 times.

   Note: Recall that the standard error of a proportion is

   $$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}.$$

   (a) Using the Central Limit Theorem, calculate a 95% confidence interval for the proportion of the time students click correctly for all pictures.

   (b) Using the Central Limit Theorem, calculate a 95% confidence interval for the proportion of the time students click correctly for erotic pictures.

   (c) Another team of researchers wanted to repeat the experiment for erotic pictures, but with a larger sample size, so that the width of a 95% confidence interval for the proportion would be 0.02. How large a sample would they need?

3. The file `ozone.txt` contains 116 measurements of the concentration of ozone (in parts per billion) at Roosevelt Island, New York. After saving the file to your computer, you can load it into R by entering the command:

```
ozone = scan(file.choose())
```

and then selecting the file.

(a) Does the data look like it comes from a normally distributed population? Include a graph to support your answer.

(b) Treating the data as an iid sample, find a 95% confidence interval for the *mean* ozone concentration.

4. The Breakfast Club believes the ages of its members are approximately normally distributed. It sends out a survey to a random sample of 10 of its members. The ages of the sample are:

$$18, 13, 18, 16, 12, 16, 17, 17, 20, 18.$$

The following R output will be useful for this question:

```
> ages = c(18, 13, 18, 16, 12, 16, 17, 17, 20, 18)
> sd(ages)
[1] 2.415229
> pt(0.655, df=9)
[1] 0.7355744
> qt(0.975, df=9)
[1] 2.262157
```

(a) Test the null hypothesis that the mean age of all members of the Breakfast Club is no more than 16, finding a P-value and giving a conclusion.

(b) Find a 95% confidence interval for the mean age of all members of the Breakfast Club.

(c) *The Deadbeat Club sends out a survey to a random sample of 10 of their members. The ages of the sample are:*

$$13, 20, 26, 41, 67, 13, 79, 15, 20, 15.$$

Explain why a t-test would not be the best choice to test the hypothesis that the mean age of all members of the Deadbeat Club is no greater than 30.

5. In a study of a wave power generator, experiments were carried out on scale models in a wave tank to establish how two different choices of mooring method affected the bending stress produced in part of the device. The wave tank could simulate a wide range of sea states and the model system was subjected to the same sample of sea states with each of two mooring methods, one of which was considerably cheaper than the other. The question of interest is whether bending stress differs for the two mooring methods.

   The data frame `waves` contains the following:

```
> waves
method1 method2
1      2.23    1.82
2      2.55    2.42
3      7.99    8.26
4      4.09    3.46
5      9.62    9.77
6      1.59    1.40
7      8.98    8.88
8      0.82    0.87
9     10.83   11.20
10     1.54    1.33
11    10.75   10.32
12     5.79    5.87
13     5.91    6.44
14     5.79    5.87
15     5.50    5.30
16     9.96    9.82
17     1.92    1.69
18     7.38    7.41
```

   Here `method1` and `method2` are the stresses (in Newton meters) for the two different mooring methods. Some summary statistics follow:

```
> summary(waves$method1)
Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.820   2.310   5.790   5.736   8.732  10.830
> summary(waves$method2)
Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.870   1.970   5.870   5.674   8.725  11.200
> summary(waves$method2 - waves$method1)
Min.   1st Qu.   Median     Mean   3rd Qu.     Max.
-0.63000 -0.20750 -0.11500 -0.06167  0.08000  0.53000

> var(waves$method1)
[1] 11.82893
> var(waves$method2)
```

```
[1] 12.54725
> var(waves$method2 - waves$method1)
[1] 0.08414412
```

The data appears close to normal.

(a) (2 points) We want to test for an average difference in stress between the methods. Explain what test we should use and why.

(b) (4 points) Write down the null and alternative hypotheses for a one-sample t-test, and calculate an appropriate t-statistic.

(c) (7 points) Fine the P-value and a 95% confidence interval for the average difference (Method 2 minus Method 1). What do you conclude about the two methods? (Clearly define the parameter you're estimating and the hypotheses you're testing, and give a full and substantive conclusion)

6. In 2014 in the US, there were 4,010,532 recorded births in total. Of these, 327,680 were in June. June has 30 days and 2014 had 365 days. For this question, treat the births in 2014 as an IID sample from a larger population of births (this is not literally true but is a sufficient approximation.)

   (a) Suppose we wish to test the hypothesis that the probability of being born in June is proportional to the number of days in June. Write down mathematical null and alternative hypotheses for this test.

   (b) Using the Central Limit Theorem, calculate an approximate 95% confidence interval for the probability a child is born in June.

   (c) Find the P-value for the test of your hypotheses in (a) using the Binomial probability and normal approximation. Using both this and your confidence interval, explain what you can conclude about the probability of being born in June.

7. The file `IUSalaries.txt` contains the salaries of a random sample of 50 IU Bloomington faculty, along with the salaries of a random sample of 50 IU athletics employees. Download the data from Canvas and read it into R, e.g. by using the following code:

```
salaries = read.table(file.choose(), header=TRUE)
Faculty = salaries$Salary[salaries$Job == "Faculty"]
Athletics = salaries$Salary[salaries$Job == "Athletics"]
```

   (a) Draw side-by-side boxplots of the two sets of salaries, and describe what you see in a sentence.

   (b) Suppose we wish to test for a difference between faculty salaries and athletics salaries. I performed Welch's two sample t-test to test whether the population means for faculty and athletics salaries were equal. I got the following output:

```
> t.test(Faculty, Athletics)

Welch Two Sample t-test

data:  Faculty and Athletics
t = 0.6388, df = 74.075, p-value = 0.5249
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-19826.38  38539.46
sample estimates:
mean of x mean of y
87608.22  78251.68
```

   The P-value produced by this test is 0.5249. Explain why we might *not* trust this P-value. Include graphs to support your answer.

   (c) Suppose we wish to test the hypothesis that the mean of the logged faculty salaries is equal to mean of the logged athletics salaries. Should we use Welch's two-sample t-test, Student's two-sample t-test, or something else? Justify your choice.

   (d) Perform the test you chose in part (c).

   (e) Do IU Bloomington faculty and IU Athletics employees have the same distribution of salaries? Explain.