PH142 2019 Extra Practice Questions

1. Which of the following statements is INCORRECT for the F distribution?

a) The F distribution is skewed to the right
b) It can only take positive values
c) The p-value of F statistics is always the area to the right of the test statistic
d) None of the above

2. Fill in the blank: _____ is a hypothesis test that reshuffles the data to break any relationship between two variables. The null hypothesis of this test is that there is no difference between the two sampling distributions.

Your answer:

3. Circle one term for each blank: For non-parametric methods, _____ [bootstrap/ permutation test] is for hypothesis testing, whereas _____ [bootstrap/permutation test] is for confidence intervals.

Your answer:

4. [1 point] True or false: The chi-squared distribution is not symmetric.

Your answer:

5. True or false: Under the null hypothesis for the unpaired two-sample t-test of the means for populations with unequal variances, the degrees of freedom for the test statistic is equal to n1 + n2 - 2, where n1 and n2 are the sample sizes of the samples.

Your answer:

6. True or false: A test for comparing two proportions can also be conducted using a Chi-square test.

Your answer:

7.True or false: If a 90% confidence interval for a proportion is (0.442, 0.542), then the margin of error is 0.10.

Your answer:

8. True or False: Running a one proportion test means I calculate a one-sided p-value, while running a two proportion test means I calculate a two-sided p-value.

Your answer:

Question 9

Sleep deprivation specialists are interested in determining whether the number of hours a student sleeps per night depends on whether they have an upcoming statistics exam. The researchers posit that the stress of the exam will not influence the students' sleep patterns, but the students disagree. Twelve randomly selected students were asked to record the number of hours they slept on the night two weeks prior to the day of the exam. The same students were also asked to the record the number of hours slept the night before the exam. The data is presented in the table below.

Table 1: Number of hours slept

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11
Two weeks before exam	10.3	10.7	9.9	8.4	8.8	10.1	9.7	9	8	9.2	10.5
Night before exam	5.3	3.9	4.5	4.3	4.4	4.1	2.6	4.6	2.4	2.9	3.6

Assume that all assumptions required for the use of the appropriate procedure to make a confidence interval are met.

a) State the null and alternative hypotheses.

H_0:

H_A:

b) Use the appropriate t-procedure to compute the 95% confidence interval of the difference between the number of hours slept two weeks before the exam and the night of the exam. Interpret the confidence interval.

You'll need to select one of the following critical values to compute the confidence interval:

a) qt(0.95, df = 11) = 1.80
b) qt(0.025, df = 11) = -2.20
c) qt(0.975, df = 12) = 2.18
d) qt(0.975, df = 22) = 2.07

You'll also need the following value: $s = 1.43$

Your answer:

Question 10

Suppose that PH142 has five discussion sections. We want to know whether they have the same exam performance. We select a random sample of students in each of these sections, and run an ANOVA analysis.

a) State the null and alternative hypotheses.

H_0:

H_A:

b) Use information in the provided table cells in addition to other information from the question to complete the ANOVA table.

Term	df	sumsq	meansq	statistic	p.value
Sections	A	1.2	0.3	C	0.01159
Residuals	195	15.6	B	NA	NA

A:

B:

C:

c) How many total students were in the sample?

Your answer:

d) Interpret the p-value in the context of this problem.

Your answer:

Question 11

Suppose you were presented with the following results after running a linear regression using two continuous variables. Which number can be used to quantify a probability having to do with the correlation between freq and temp? State the hypotheses and make a conclusion based on the number you mentioned above.

```
## # A tibble: 2 x 5
##   term        estimate std.error statistic    p.value
##   <chr>          <dbl>     <dbl>     <dbl>      <dbl>
## 1 (Intercept)    -6.19     8.24     -0.751 0.462
## 2 temp            2.33     0.347     6.72  0.00000266
```

Your answer:

Question 12

The approximate frequencies of ABO blood types among all Americans are given in the table below.

O	A	B	AB
44%	42%	10%	4%

While determining how much of each blood type to keep in stock, a particular hospital investigates whether the distribution of blood types among its patients is the same as the distribution of blood types nation-wide. The investigators collect a simple random sample of 250 patients seen at the hospital over the last year. The investigators' results are shown in the table below.

O	A	B	AB	Total
90	110	40	10	250

The investigators perform a χ^2 goodness-of-fit test to determine whether the distribution of blood types among the patients at this hospital differ from the distribution of blood types among Americans in general using a significance level of 0.05.

a) Clearly state the hypotheses and calculate the degrees of freedom and test statistic.

H_0:

H_A:

Your answer:

b) Write the R code you would need to find the p-value for this test.

Your answer:

c) Suppose that the p-value provides strong evidence against the null in favor of the alternative hypothesis. Which of the four blood type categories contributed the most evidence to this difference?

Your answer: