**BIOM 505: Biostatistical Methods I**

**HOMEWORK 3 (Hypothesis Testing) [Due Dec 6, 2018]**

**Dr. Fares Qeadan: fqeadan@salud.unm.edu**

|  |
| --- |
|  |

**Question 1[10 points]:**

A tobacco company says that its best-selling cigarettes contain at most 40 mg of nicotine. Some junior medical students believe that the tobacco company is not saying the truth about its tobacco product. In order to check whether the nicotine content of the cigarettes exceeds 40 mg, the students have obtained a simple random sample of 15 cigarettes which gave a sample average nicotine content of 42.6 mg and standard deviation of 3.7 mg. Assume cigarette nicotine content is normally distributed. Use a 1% level of significance to carry out the appropriate test of significance.

**Answer:**

**Question 2[10 points]:**

A study from the U.S. of 59 professional soccer players found 15 of them to be suffering from depression and anxiety, possibly related to past injuries. Is there evidence that the proportion of all soccer players in the U.S. who suffer from depression and anxiety is over 24%? Use 5% level of significance (α=0.05) to carry out the appropriate test of significance. Also, please explain the conditions for inference in this question.

**Answer:**

**Question 3[2 points]:**

In many situations, the\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ hypothesis is referred to as the research hypothesis, since it is the statement the researcher would really like to validate. Furthermore, in our treatment of hypothesis testing, the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ hypothesis will always be stated as an equality claim.

1. null, alternative
2. null, null
3. alternative, null
4. alternative, alternative

**Answer:**

**Question 4[2 points]:**

A \_\_\_\_\_\_\_\_\_\_\_\_\_ error involves not rejecting the null hypothesis H0 when H0 is false, while a

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ error consists of rejecting the null hypothesis H0 when H0 is true.

1. type I, type II
2. type I, type I
3. type II, type I
4. type II, type II

**Answer:**

**Question 5[2 points]:**

The probabilities of type I and type II errors are traditionally denoted by the Greek letters

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, respectively. Furthermore, a \_\_\_\_\_\_\_\_\_\_\_\_ error is usually more serious than a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ error.

1. α, 1 −β, type I, type II
2. α, β, type II, type I
3. α, β, type I, type II
4. β, α, type I, type II
5. 1 − β, α, type I, type II

**Answer:**

**Question 6[2 points]:**

If the P-value is smaller than or equal to the level of significance α, then the researcher should

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ at level α. Furthermore, if the P-value is larger than the level of significance α, then the researcher should \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ at level α.
.

1. reject H0, fail to reject H0
2. fail to reject H0, reject H0
3. reject H0, fail to reject Ha
4. fail to reject Ha, reject H0
5. reject Ha, fail to reject Ha

**Answer:**

**Question 7[2 points]:**

Which of the following statements is not true?

1. A statistical hypothesis is a claim or assertion either about the value of a single parameter, about the values of several parameters, or about the form of an entire probability distribution.
2. In any hypothesis-testing problem, there are two contradictory hypotheses under consideration.
3. A test of hypothesis is a method for using sample data to decide whether the null hypothesis should be rejected.
4. A type I error consists of not rejecting the null hypothesis H0 when H0 is false.

**Answer:**

**Question 8[2 points]:**

Which of the following statements is true?

1. In general, the null hypothesis H0 is a hypothesis which the researcher tries to disprove, reject or nullify.
2. The alternative hypothesis, denoted by Ha, is the assertion that is contradictory to the null hypothesis H0.
3. The null hypothesis H0 will be rejected in favor of the alternative hypothesis only if sample evidence suggests that H0 is false.
4. If sample evidence does not strongly contradict the null hypothesis H0, we will continue to believe in the truth of H0.
5. All of the above statements are true.

**Answer:**

**Question 9[2 points]:**

Which of the following statements is not correct?

1. It is possible that the null hypothesis may be rejected when it is true.
2. It is impossible that the null hypothesis may be rejected when it is true.
3. It is possible that the null hypothesis may not be rejected when it is false.

**Answer:**

**Question 10[2 points]:**

In hypothesis-testing analysis, a type I error occurs only if

1. the null hypothesis is rejected when it is true
2. the null hypothesis is rejected when it is false
3. the null hypothesis is not rejected when it is false
4. the null hypothesis is not rejected when it is true

**Answer:**

**Question 11[2 points]:**

In hypothesis-testing analysis, a type II error occurs only if

1. the null hypothesis is rejected when it is true
2. the null hypothesis is rejected when it is false
3. the null hypothesis is not rejected when it is false
4. the null hypothesis is not rejected when it is true

**Answer:**

**Question 12[2 points]:**

Suppose that when data from an experiment was analyzed, the P-value for testing H0 :μ ≤50 versus Ha :μ > 50 was calculated as .0244. Which of the following statements is true?

1. H0 is rejected at .10 level.
2. H0 is not rejected at .05 level.
3. H0 is not rejected at .025 level.
4. H0 is rejected at any level α.
5. All of the above statements are true.

**Answer:**

**Question 13[2 points]:**

Which of the following P-values will lead us to reject the null hypothesis at the .05 level?

1. .10
2. .025
3. .075
4. .15
5. Any P-value greater than .05

**Answer:**

**Question 14[2 points]:**

Which of the following statements is not true?

1. The reliability of hypothesis testing procedure in reaching a correct decision can be assessed by studying type I error probability.
2. The process of reaching a decision by using the methodology of classical hypothesis testing involves selecting a level of significance and then rejecting or not rejecting the null hypothesis H0 at that level α.
3. A small P-value would indicate statistical significance in that it would strongly suggest rejection H0 of in favor of Ha.
4. A test with significance level α is one for which the type I error probability is controlled at the specified level.
5. A type I error is usually more serious than a type II error.

**Answer:**

**Question 15[2 points]:**

A physical therapist believes that the average weight of a backpack carried by an UNM sophomore on Monday mornings is at least 19 pounds. In carrying out her hypothesis test at the .05 level, the therapist obtains a p-value of .037. An appropriate formulation of her conclusion is

1. The data support the claim that the average weight of a backpack carried by an UNM sophomore on Monday mornings is at least 19 pounds.
2. The data do not support the claim that the average weight of a backpack carried by an UNM sophomore on Monday mornings is at least 19 pounds.

**Answer:**

**Question 16[10 points]:**

An investigator has conducted a study where it is hypothesized that exposing premature babies to small amounts of caffeine within 24 hours of birth will decrease their length of hospital stay. The sample size for the study was 80. The sample mean length of stay for the infants given caffeine was 40 with s=12 days. The literature reports that premature infants given noncaffeine stimulants within 24 hours had a mean length of stay of 46 days. Does caffeine do a better job at reducing the length of stay? Support your answer with statistical evidence. Assume that length of hospital stay is normally distributed.

**Answer:**

**Question 17[15 points]:**

Consider the following hypertension dataset (<http://www.mathalpha.com/BIOM-505/hypertensionfall17.sas7bdat> ) [Data set 2 on the course webpage]. This dataset is courtesy of Dr Waldon Garris, University of Virginia School of Medicine. Dr Garriss collected the data in a pilot study during his work in the Dominican Republic in 1997. The subjects are persons who came to medical clinics in several villages, for a variety of complaints. Data on gender, age, systolic and diastolic blood pressure were collected.

1. Test whether the true mean systolic blood pressure for males is different than that of females? Assume normality in both groups with unknown but equal variances.
2. What’s the name of the statistical test you have used to answer this question?

**Answer:**

1. Could you reach the same conclusion in (a) by examining the 95% C.I. around the mean difference between males and females? Explain.

**Answer:**

**Question 18[12 points]:**

Using the same hypertension dataset from the previous problem, could you please

1. Test whether the true mean diastolic blood pressure is significantly higher than 80 mm Hg? Assume data is normal with unknown variance.

**Answer:**

1. What’s the name of the statistical test you have used to answer this question?

**Answer:**

**Question 19[12 points]:**

Using the same hypertension dataset from the previous problem, please create a syntax to construct a new variable for hypertension and then answer the following questions:

bp\_cat="Normal" if (sbp<120 & dbp<80)

bp\_cat="Prehypertension" if ((sbp>=120 & sbp<=139) | (dbp>=80 & dbp<=89))

bp\_cat="Stage 1" if ((sbp>=140 & sbp<=159) | (dbp>=90 & dbp<=99))

bp\_cat="Stage 2" if (sbp>=160 | dbp>=100)

hyper=1 if (bp\_cat=="Stage 1" | bp\_cat=="Stage 2")

hyper=0 if (bp\_cat=="Normal" | bp\_cat=="Prehypertension")

1. At the 10% significance level, test whether the true proportion of hypertension is significantly higher than 43%?

**Answer:**

1. What’s the name of the statistical test you have used to answer this question?

**Answer:**

**Question 20[5 points]:**

Using the same hypertension dataset from the previous problem and assuming that you have constructed a new variable for hypertension and a numerical gender variable, please

1. test whether the true rate of hypertension among males is significantly different from that among females at the 1% significance level?

**Answer:**

1. What’s the name of the statistical test you have used to answer this question?

**Answer:**

**Question 21 [EXTRA CREDIT] [up to 4 points]:**

To examine whether a diet and physical activity program, called ZWX, is effective or not for losing weight among adults of the age 40-45 years old, a random sample of subjects of that age group was selected and the weights were recorded at two points, before they participated in the program and then after the finished the program. The data of this study could be created from the following:

input before after

147.33 130.03

184.63 170.43

135.14 123.41

116.23 111.97

129.99 127.42

135.74 128.17

180.50 159.82

136.48 131.50

156.97 147.19

135.32 127.77

150.50 149.91

121.65 119.22

125.13 121.57

148.74 137.03

138.53 138.31

147.62 119.32

146.60 134.86

1. At the 5% significance level, is it safe to say that the diet and physical activity program ZWX is effective for losing weight among adults of the age 40-45 years old? Explain

**Answer:**

1. What’s the name of the statistical test you have used to answer this question?

**Answer:**