**BIOS 6606 MIDTERM ANSWERS**

**Fall 2019**

**Instructions:** You may use only your annotated 3 page study guide and a calculator during this exam. No phones or other electronics are allowed. The exam has 45 questions and is worth 55 points. There are 4 extra credit points from 3 questions at the end of the exam. These extra credit points will count towards your overall grade and will not be included in the total for the exam. The exam is 10 pages (including this one) long. Focus on completing the main part of the exam first before doing the extra credit problems.

***Read and sign below.***

I understand that my participation in this examination and in all academic and professional activities as a UCAMC student is bound by the provisions of the UCAMC Honor Code. I understand that work on this exam is to be done independently. On my honor, I have neither given nor received aid on this examination.

Signature and Date

Print name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Sign name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**EXAM STARTS HERE:**

**\_\_\_\_\_Q1**. A numerical value used as a summary measure for a sample, such as sample [mean](https://brainmass.com/statistics/mean), is known as (Choose one; 1 pt)

A. population parameter   
B. sample parameter  
C. sample statistic  
D. population mean  
E. None of the above answers is correct.

**\_\_\_\_\_Q2**. The \_\_\_\_\_\_\_ is often the preferred measure of central tendency if the data are severely skewed (Choose one; 1 pt)

A. Mean

B. Median

C. Mode

D. Range

E. None of the above answers is correct.

**\_\_\_\_\_Q3**. Which one of these statistics is unaffected by outliers? (Choose one; 1 pt)

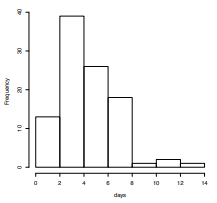
A. Mean

B. Interquartile range

C. Standard deviation

D. Range

E. All are affected



**\_\_\_\_\_Q4**. The histogram above represents the lifespan of a random sample of a particular type of insect. Determine the relationship between the mean and median. (Choose one; 1 pt)

A. mean = median

B. mean ≈ median

C. mean < median

D. mean > median

E. Cannot tell from the information given.

**\_\_\_\_\_Q5**. What is the median and mode of the following distribution of 10 students’ quiz scores (the mean is 14.9). (Short answer; 2 pts)

18, 4, 19, 20, 14, 19, 15, 16, 19, 5

4,5,14,15,16,18,19,19,19,20. The median is 17 and the mode is 19.

**\_\_\_\_\_Q6**. Which of the following statements is a correct null hypothesis? (Choose one; 1 pt)

A. There is no difference between male and female incomes in the sample

B. There is a correlation between males and females incomes

C. There is no difference between male and female incomes in the population

D. Male and female incomes in the population are different

E. None of them

**\_\_\_\_\_Q7**. An experimenter has decided to use a small value of ** like ** = 0*.*01 rather than 0.05 in her experiment. Which of the following is a consequence of her decision? (Choose one, 1 pt)

A. the probability of a Type II error (false negative) is smaller.

B. the probability of a Type I error (false positive) is larger

C. the probability of a Type I error is smaller

D. if the null hypothesis is false, the experimenter will easily be able to reject it

E. even if the null hypothesis is far from being true, it will be hard to reject it.

**\_\_\_\_\_Q8.** Failing to reject the null hypothesis when it is false is: (Choose one; 1 pt)

A. alpha

B. Type I error (false positive)

C. beta

D. Type II error (false negative)

E. none of the above

**\_\_\_\_\_Q9**. Suppose you conduct a significance test for the population mean  and your p-value is 0.07. Given a 0.05 level of significance, which of the following should be your conclusion? (Choose one; 1 pt)

A. Accept HO

B. Reject HA

C. Fail to reject HA

D. Fail to reject H0

E. Reject H0

**\_\_\_\_\_\_Q10**. If the test result for A has a p=0.04 and the result for B has a p=0.0003 then you can conclude that relationship B is stronger than relationship A. (Choose one; 1 pt)

A. True

B. False

**\_\_\_\_\_Q11**. What is the standard deviation of a sampling distribution called? (Choose one; 1 pt)

A. Sampling error

B. Sample error

C. Standard error of the mean

D. Simple error

E. None of the above

**\_\_\_\_\_Q12**. Which of the following is true about a 95% confidence interval of the mean of a given sample: (Choose one, 1 pt)

A. 95 out of 100 sample means will fall within the limits of the confidence interval.

B. There is a 95% chance that the population mean will fall within the limits of the confidence interval.

C. 95 out of 100 population means will fall within the limits of the confidence interval.

D. There is a .05 probability that the population mean falls within the limits of the confidence interval.

E. None of the above

In my textbook, it assumes the mean is a fixed value, which will not be affected by the CI. But the confidence interval is depending on the sampling variance and the mean. so it interprets in another way, that there is 95% of the probability that the (95%) confidence interval will cover the mean.

**\_\_\_\_\_Q13**. What effect would increasing the sample size have on a confidence interval?

A. The confidence interval would increase in size.

B. The confidence interval would decrease in size.

C. The confidence interval is unaffected by sample size.

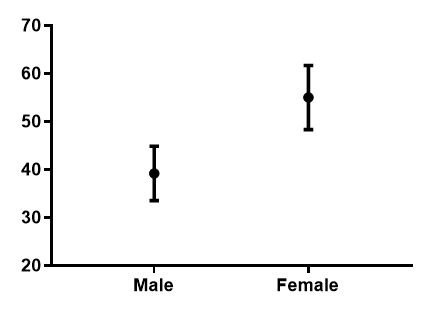
D. The confidence interval could either increase or decrease in size.

E. None of the above

**Q14**. You want to know if the mean values for men and women are different (p<0.05). What do you need to know about the data in the graph below so that you can make a conclusion about statistical significance between men and women? (Short answer; 2 pts)

Type of error bars and sample size

 this file did not show up in my Word file.



**Q15**. A researcher used an unpaired (independent groups) *t*-test to compare hemoglobin levels of males & females. The results showed that *t* = 0.77, *p* = 0.48. Assuming that alpha = 0.05. What decision should the researcher make? (Short answer; 1 pt)

Fail to reject the null hypothesis and accept that the hemoglobin levels are the same by gender

**Q16.** In your answer to Q15, which type of error most possibly they be making? (Short answer; 1 pt)

Type II (false negative)– fail to find a difference when there really is one.

**The following 4 questions Q17-Q20) deal with a weight loss study where triglyceride levels were measured in a group of people at baseline and again at 8 weeks after the subjects started taking a new weight loss drug.**

**Q17**. What type of t-test would you do on these data? What is the non-parametric equivalent test? (Short answer; 2 pts)

Paired t-test; Wilcoxon signed rank

**Q18**. What is the null hypothesis for the t-test? (Short answer; 1 pt)

The difference for the same individual? is in the population is equal to zero

**Q19**. What is the alternative hypothesis for a two-sided test? (Short answer; 1 pt)

The difference in the population is not zero

**Q20**. With p=0.001 and significance set at 0.01, do you reject or fail to reject the null hypothesis? (Short answer; 1 pt)

Reject

**The following 4 questions Q21-Q24) deal with a weight loss study where triglyceride levels were measured in a group of people at baseline and again at 8 weeks after the subjects started taking a new weight loss drug.**

**Q21**. Calculate a one-sample t-test using the summary data below (the data met all test assumptions). Show your work. (Short answer; 1 pt)  
x̄ = [sample mean](http://www.statisticshowto.com/sample-mean/)( Σ xi ) / n = 280  
μ0 = [population mean](http://www.statisticshowto.com/population-mean/) = 300  
s = sample standard deviation = 50

s2 = sample variance =2500  
n = [sample size](http://www.statisticshowto.com/find-sample-size-statistics/) = 15

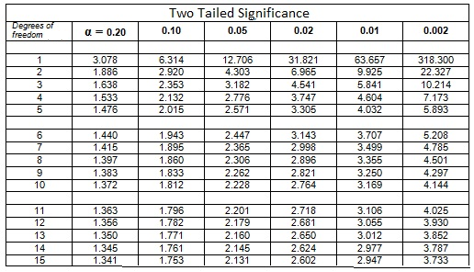
ANSWER: t = (280 – 300)/ (50/√15)

= -20 / 12.909945= -1.549

**Q22**. What are the degrees of freedom for the problem above? (Short answer; 1 pt)

ANSWER: 15-1=14

**Q23**. Find the critical value for the degrees of freedom you calculated in the question Q21. What is the value? Use the standard value for alpha (0.05). (Short answer; 1 pt)



2.145

**Q24**. Do you reject the null hypothesis of no difference in the population means at the 0.05 level for a two-sided test? Why or why not? (Short Answer; 2 pts)

ANSWER: No, fail to reject; the value we calculated does not fall in the rejection region; it is less that the critical value.

**\_\_\_\_\_Q25**. Of the following Z-score values, which one represents the location closest to the mean? (Choose one; 1 pt)

A. Z= 0.5

B. Z= 1,0

C. Z= -1.5

D. Z= -0.3

E. Z= 2.0

**\_\_\_\_\_Q26**. How does the shape of the t distribution compare to the normal distribution? (Choose one; 1 pt)

A. The t distribution is taller and less spread out (with smaller tails),, especially when n is large.

B. The t distribution is taller and less spread out (with smaller tails),, especially when n is small.

C. The t distribution is flatter and more spread out (with bigger tails),, especially when n is large.

D. The t distribution is flatter and more spread out (with bigger tails), especially when n is small.

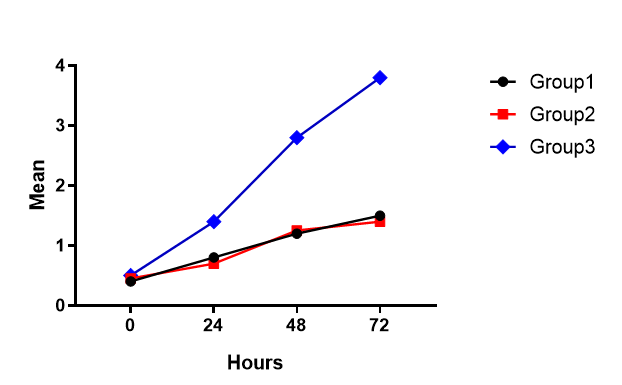
E. None of the above is correct

**Q27**.The distribution of heights of American women aged 18 to 24 is approximately normally distributed with a mean of 65.5 inches and standard deviation of 2.5 inches. Calculate the z-score for a woman six feet tall. Show your work (1 pt) 1foot = 12 inches

=(72-65.5)/2.5 = 2.6

**\_\_\_\_\_Q28**. Looking at the graph below from a repeated measures analysis, which statement is correct? (Choose one; 1 pt)





A. There are no differences by time or by group

B. There is a time main effect only

C. There is a group main effect only

D. There is a time and a group effect but no evidence of interaction between group and time

E. There is a time and a group effect and there is evidence of interaction

**Q29**. You want to do a repeated measures ANOVA. You calculated measures of central tendency and spread for each group in the analysis. The data are in the table below. What evidence do you see that the data may not meet the assumptions of a repeated measures ANOVA? Give all evidence. (Short answer; 2 pts)

Group1 Group2

Mean SD n Mean SD n

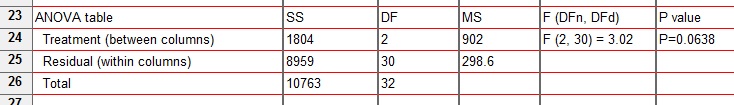
Day0 23.5 17.38 11 20.4 15.91 11

Day5 15.3 23.92 11 27.9 26.65 11

Day12 5.4 4.66 11 34.5 28.02 10

The SD are not equal. Dividing the largest by the smallest (28.02/4.66 = 6.0 – if they say 23.92/4.66 = 5.1, that will be acceptable, just that they are >2) >2. Also, the time intervals are not equal. The design is unbalanced or there is missing data.

**Q30 and Q31: The results from a one-way ANOVA are below.**



**Q30.** How many groups (columns) are in the analysis? (Short answer; 1 pt)

3 (DF+1)

**Q31**. What is the total sample size? (Short answer; 1 pt)

33 (Residual DF + # of groups; or Total DF +1)

**Q32 And Q33**. A researcher measured the IQs of a group of 12 coworkers. She believes that the mean IQ of this group is probably greater than the population average of 100 (we will regard the population IQ is a normal distribution).

**Q32.** What test would you use to test her research hypothesis? (Short1 pt) Z-score test

**Q33.** State the null hypothesis for your test in statistical context? (1 pt) =100.

**\_\_\_\_\_Q34.** In a study, mice are randomly assigned to one of three groups: control, Drug A, or Drub B. After one week, tumor volumes for the three groups are compared to see whether there is difference among the groups. An appropriate statistical test for this comparison is: (Choose one; 1 pt)

A. a repeated measures ANOVA

B. three-way ANOVA

C independent t-test

D. one-way analysis of variance

E. None of the answers above are correct

**The following situation and data apply to Q35-Q40. An experiment in 30 male college students who received training in finger tapping Students were randomized to one of three levels of caffeine. Two hours after caffeine ingestion, the number of taps per minute was recorded. The PI of the study hypothesizes that finger taps will increase with higher levels of caffeine, but he is not certain. He collects the follow data after the experiment.**

caffeine0 caffeine100 caffeine200

242 248 246

245 246 248

244 245 250

248 247 252

247 248 248

248 250 250

242 247 246

244 246 248

246 243 245

242 244 250

**Q35.** What statistical test is appropriate to analyze the data? Give both parametric and non-parametric tests? (Short answer; 2 pts)

One-way ANOVA or Kruskal-Wallis test

**Q36.** What are the null hypotheses for both tests in statistical context? (Short Answer; 2 pts)

The means of the population groups are equal

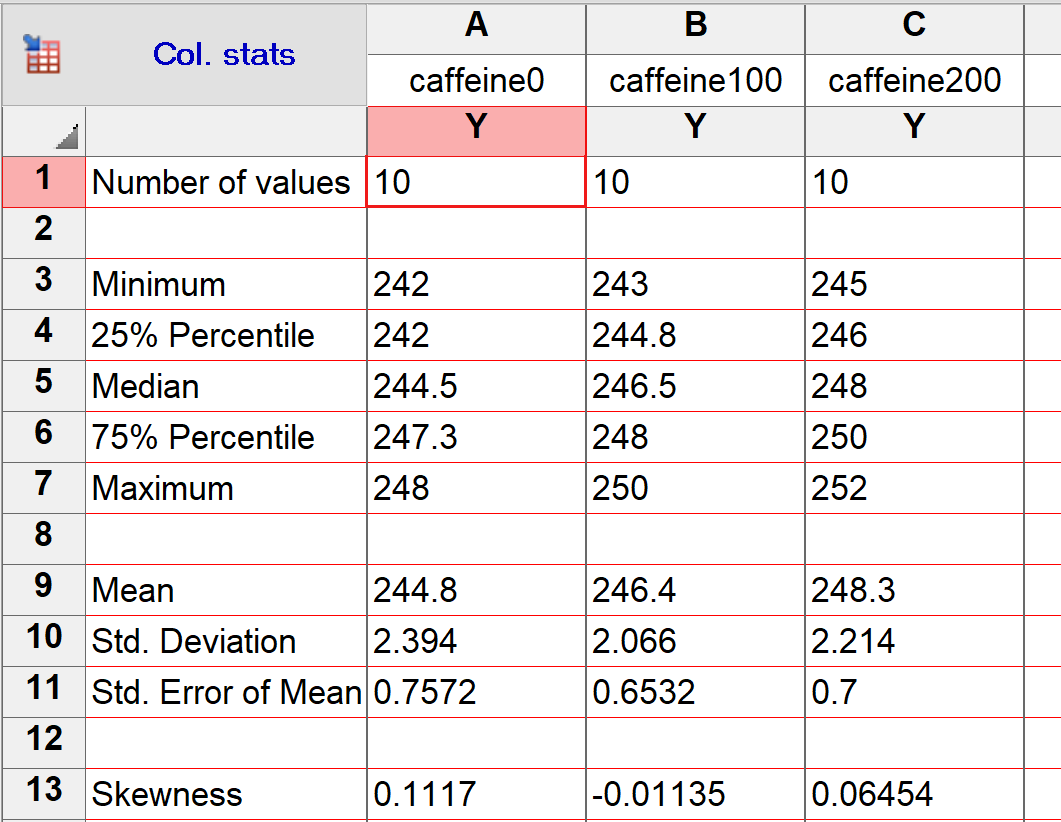
The medians (or distributions) of the population groups are equal

**Q37.** What is the alternate hypothesis for a two-tailed test for both tests? (Short Answer; 2 pts)

At least one mean in the population groups is different from one other

At least one median in the population groups is different from one other

Before you started the analysis, you calculated mean, SD, etc. The results are below.



**Q38.** With the assumptions of normality and homoscedasticity in mind, do you see any evidence from the results above to indicate that these assumptions are met. Why or why not? (Short answer; 2 pts)

Normality – met. The mean and median for each group are close in value and the skewness index is <0.5 for each group.

Homoscedasticity – met. the SD are close if value and the ratio of largest to smallest is <2

**Q39.** Based on your answer above, which is the better test to use on your data – the parametric or non-parametric test? (Short answer; 1 pt)

Parametric one-way ANOVA.

**Q40.** The p-value of the pair-wise tests for the statistical test are all less than 0.005 p=0.006. What do you conclude about the relationship of finger taps to caffeine consumption? (Short answer; 2 pts)

We reject the null hypothesis of no difference between groups. The more caffeine you consume, the more finger taps you can do from a mean of 244.8 taps per minute with no caffeine to 248.3 with the highest dose of caffeine.

Do we need to give them the post hoc test results? Or they just need to answer the at least caffeine0 and caffeine200 will have the most significant difference.

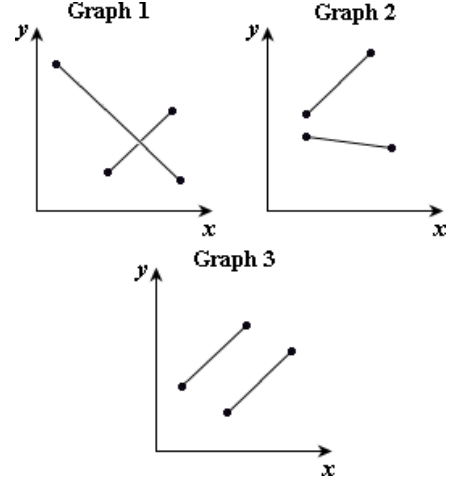
**Q41**. A biologist is interested in studying the effect of different salinity levels and different growth-enhancing nutrients in water on the growth of shrimp. The biologist ordered young tiger shrimp from a supply house for use in the study. Tiger shrimp were randomly placed into one of 12 similar tanks in a controlled environment (n=10 tiger shrimp per tank). The biologist randomly treated each tank with different combinations of two growth-enhancing nutrients (A or B) and two salinity levels (low or high).

What is the experimental unit \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_;

What is the sample size\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ for analysis? (Short answer; 2 pts)

*Experiemental Unit The individual tank; n=12. Extra credit (0.5 pt) if they also add that the observational unit is the individual shrimp.*

**\_\_\_\_\_Q42.** Which graph(s) above show evidence of interaction in a two-way ANOVA? (Choose one, 1 pt)



A. Graph 1 E. Graphs 2 and 3

B. Graph 2 F. Graphs 1 and 3

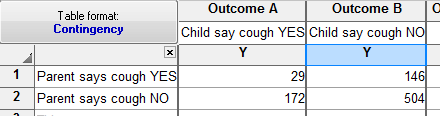
C. Graph 3 G. All

D. Graphs 1 and 2 H. None

**\_\_\_\_\_Q43.** One advantage of the chi-square tests is that they can be used when the data are measured on a nominal scale. (Choose one; 1 pt)

A. True  
B. False

**Q44**.The table below shows data about whether a child coughed first thing in the morning or not in a group of schoolchildren with asthma, for the same individual child as reported both by the child and by the child's parents. A friend ran some statistical tests on the data and gave you the p-values in the box next to the table. Your friend is not sure which test result to use.

****

Chi-square p=0.014

Fisher’s exact test p=0.013

McNemar’s test p=0.16

**\_\_\_ \_\_Q44.** Which is the correct p-value to report (of course you also tell her to report some measure of effect and not to p-hack)? (Choose one; 1 pt)

a. Chi-square only

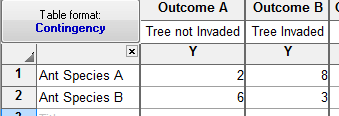
b. Fisher’s exact test only

c. Either Chi-square or Fisher’s exact test

d. McNemar’s test only

e. none of the above

**Q45.** Calculate the expected values for cell A of the table below. Calculate to one decimal. Show your work. (Calculations only; 1 pt)

****

Cell A: = (10x8)/19 = 80/19 = 4.2

**EXTRA CREDIT Q1.** The expected values for Cell B is 5.8, Cell C 3.8, Cell D 5.2. Finish calculating the chi-square statistic below by adding the value for Cell A below. (Note: you should really use the Fisher’s exact test on these data due to the fact that the expected value for at least one of the cells is <5, but the Fisher’s exact is too calculation heavy to do by hand). Show your work. (Calculations only; 2 EC pts)

Cell A = (2-4.2)2/4.2 = (-2.2)2/4.2 = 4.8/4.2 = 1.14

Cell B 0.83

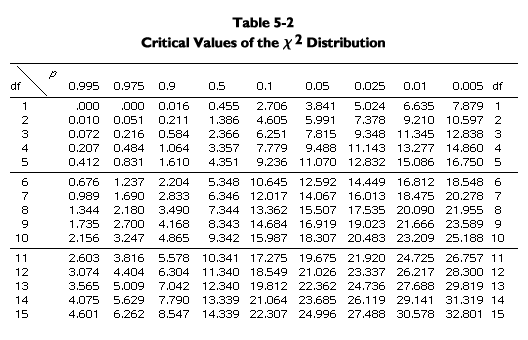
Cell C 1.27

Cell D 0.92

SUM (chi-sq statistic) 4.15

~~+ following question what is the degree of freedom in this case.~~

**EXTRA CREDIT Q2**. What is the critical value for your chi-square test at the 0.05 significance level. (Value only; 1 EC pt)



3.841

**EXTRA CREDIT Q3**. Do you reject or fail to reject the null hypothesis based on the answers to ECQ1 and ECQ2 above? (Short answer; 1 EC pt)

Reject