

NAME: \_\_\_\_\_

PSYC 3101-300

Exam 2  
Spring 2012

You may use scratch paper and a calculator. You may NOT use your book or notes. Show your work. If it is not obvious, CIRCLE YOUR FINAL ANSWERS.

### **PRACTICE QUESTIONS - REVISED**

#### **Short answer/multiple choice questions (~ 8 on exam, 50 points total)**

1) When performing a two-sample t-test, for a given small difference between the two group means, what other two properties of the sample (besides the difference between means of the groups) determine if the effect can be detected?

**Larger sample size  
Smaller within-group variability**

2) When performing an independent-measures ANOVA, which statistic is a measure of effect size? Write an equation for this statistic, and then explain what this statistic is telling you.

**$r^2 = \text{SS}_{\text{bw}} / (\text{SS}_{\text{bw}} + \text{SS}_{\text{wn}})$ . It is interpreted as the amount of variation in the dependent variable that is “explained” (or associated) with the independent variable.**

3) What are two reasons why using an ANOVA is more appropriate than performing multiple t-tests when comparing the means of more than two groups?

**For simplicity of having a single test statistic rather than multiple ones  
To reduce the problem of multiple testing (which inflates the experiment-wise type-I error rate)**

4) What is the main difference between a t-test and a z-test? Describe how this difference relates to the shape of the t and z distributions.

**t-tests are used when the standard deviation (or variance) of the population is unknown and must be estimated from the sample. This increases uncertainty in the test statistic (due to increased uncertainty of what the true SEM should be) and results in fatter tails of the t-distribution (which leads to more extreme t-values than z-values if the null is true).**

5) Briefly explain how increasing sample size influences each of the following. Assume that all other factors are held constant:

A) The size of a z-score in a hypothesis test

**makes the z-score larger**

B) The size of a t-score in a hypothesis test

**makes the t-score larger**

C) The size of Cohen's d.

**has no influence on the Cohen's d (which is a motivation for using it)**

D) The power of a hypothesis test

**increases the power**

6) What does it mean to commit a type-I error?

**It means that you rejected the null hypothesis when the null was true**

7) What circumstance leads you use a dependent t-test instead of a two-sample t-test?

**When the scores from one group are linked to the scores from the other group for some reason (typically, they are linked because the same persons are in both groups)**

8) What circumstance leads you use an ANOVA instead of a two-sample t-test?

**When we want to compare the means of more than two groups**

9) What happens to the shape of the t-distribution as the degrees of freedom get larger?

**The t-distribution becomes more normally distributed**

10) Which of the following is NOT TRUE about an F-statistic:

a. F is always positive

b. Computed as the ratio of Mean Square Between to Mean Square Within

c. F is always greater than 1

d. When degrees of freedom are low, the F distribution is positively skewed

**C – F is always greater than 0, not 1**

11R) Sketch the following distributions, and note where "0" is on the x-axis:

A) normal distribution

B) t-distribution with 5 degrees of freedom

C) F-distribution with 2,20 degrees of freedom

**I won't sketch them here, but A should be known by everyone. The t-dist should have fatter tails. For both of those, "0" is right in the middle. The F should be skewed to the right, and 0 should be at the extreme left of the distribution.**

12) True or False: Increasing the alpha level from  $\alpha = .01$  to  $\alpha = .05$  will produce an increase in the estimated standard error.

**False**

13) True or False: The boundaries for the critical region for a two-tailed test using a t statistic with  $\alpha = .05$  will never be less than  $\pm 1.96$ .

**True**

14R) Which of the following is a fundamental difference between the t statistic and a z-score?

- a. The t statistic uses the sample mean in place of the population mean
- b. The t statistic uses the sample variance in place of the population variance
- c. If the null is true, fewer extreme observations of t are observed than z.
- d. The t statistic is only used for very large samples, whereas the z-score is used for all sample sizes

**B**

15) True or False: If  $H_0: \mu = 0$ , obtaining a 95% confidence interval of (95% CI = -1.2, 0.8) means that you would reject  $H_0$

**False**

16) In ANOVA, assuming that the null hypothesis is true, both types of mean squares ( $MS_{\text{within}}$  and  $MS_{\text{between}}$ ) provide an estimate of what \_\_\_\_\_

- a. variance of individual scores in the population
- b. average differences among means
- c. a measure of central tendency
- d. the overall mean for the set of all scores

**A**

17) True or False: In general, the larger the value for a t statistic (far from zero in either the negative or the positive direction), the less consistent the sample data with the null hypothesis.

**True**

18) What is the sample variance ( $s^2$ ) and the estimated standard error (SE) for a sample with  $SS = 250$  and  $n = 11$ ?

$$s^2 = 250/(11-1) = 25. \text{ SEM} = \sqrt{25/11} = 1.51.$$

19) What if a sample mean falls within the critical region, what would you conclude? (Hint: Your answer should be a statement about the null and/or alternative hypotheses.)

**You would reject the null hypothesis**

20R) True or False: Holding all other factors constant, in a hypothesis test, the smaller the value of the sample variance is, the more likely it is you will reject the null hypothesis.

**True**

21R) True or False: Holding all other factors constant, a large difference between sample means is more likely to produce a large F statistic than a small difference between sample means.

**True**

22) In a two sample t-test, one sample has an  $s_1 = 10$  and the other has an  $s_2 = 12$ . Assuming that both samples are the same size ( $n_1 = n_2 = 10$ ), what is the pooled standard deviation ( $s_p$ ) for this t-test?

$$s_1^2 = 100 \text{ and } s_2^2 = 144. \text{ SS}_1 = 100 \cdot 9 = 900. \text{ SS}_2 = 144 \cdot 9 = 1296. \\ s^2(\text{pooled}) = (900 + 1296)/18 = 122 \\ s(\text{pooled}) = \sqrt{122} = 11.05$$

23) A) If you know that  $t(15) = 2.85$ ,  $m_1 = 20$ , and  $m_2 = 25$ , what is the value of  $S_{(m_1 - m_2)}$ ?

**$2.85 = (25-20)/\text{SEM} = 5/2.85 = 1.75$ . Note that the positive t-value must have meant that numerator was mean2 – mean1. That was not intentional on our part. We'll try not to have questions like this on the test.**

B) Describe in words what this value,  $S_{(m_1 - m_2)}$ , represents.

**It is the standard error of the mean difference. It represents the standard deviation of the sampling distribution of mean differences (hypothetical thought experiment: take a lot of pairs of samples of size  $n_1$  and  $n_2$ . Find the mean of each and subtract them and plot that mean difference. Do this an infinite number of times. The SEMd is the standard deviation of that distribution).**

24R) In an ANOVA, the F-statistic is equal to what?

a. the variance of the sample means

- b. sample means divided by sample variances
- c. mean squared between divided by mean squared within
- d. mean squared between divided by sample means

**C**

25R) Which combination of factors will produce the largest F-ratio?

- i. a large mean differences between groups; ii. small variances within groups
- iii. large variances within groups; iv. large sample size

- a. i, ii, and iii
- b. i, iii, and iv
- c. ii, iii, and iv
- d. i and ii only
- e. i, ii, and iv
- f. i, ii, iii, and iv

**E (i, ii, and iv)**

26) What is the statistical power of a test?

**The probability of rejecting the null given that an alternative hypothesis is true. Note that the power will change depending on which alternative is true (bigger effect sizes mean higher power).**

27) Identify what effects the following scenarios typically have on statistical power:

A) Increase alpha from .05 to .10

**increase power**

B) Decrease the sample size.

**decrease power**

C) Measure your dependent variable with an instrument that is more sensitive, thereby reducing the variation between people in their scores due to measurement error.

**increase power**

D) Larger effect size.

**increase power**

28) When does a repeated-measures t-test improve power over an independent-samples t-test? Why does this occur?

**It increases power when there are stable interpersonal differences in the dependent variable. It occurs because removing this source of variability makes the test more sensitive (by reducing the SEM).**

29) For an independent samples t-test, we find a Cohen's d of +.5. Put into words what this measure of effect size means.

**This means that the mean for one group is one half of a standard deviation above the mean of the other group.**

30) For a single-factor ANOVA, we find an  $r^2$  of .25. Put into words what this measure of effect size means.

**It means that 25% of the variation in the dependent variable is "explained" (or is "associated with") by the levels of the independent variable.**

31) What is the difference between a type-I error and a type-II error? Which one do we typically keep constant in behavioral studies? If we increase the chance of a type-I error, what happens to the chance of a type-II error?

**type-I error is the probability of rejecting the null when the null is true, whereas a type-II error is the probability of failing to reject the null when the alternative is true. We keep the type-I error constant (usually at .05) in behavioral studies. If we increase the chance of a type-I error, the chance of a type-II error necessarily decreases.**

32) For a one-sample t-test that we conduct on height in a sample of 100 individuals, our 95% confidence interval is [66.8,70.2].

A) Put into words what this confidence interval means.

**This says that, with 95% confidence, the true mean in the population lies between 66.8 and 70.2. This means that 95% of similarly constructed CIs would contain the true mean in the population.**

B) What would happen to the above confidence interval if, instead of a 95% confidence interval, we derived a 90% confidence interval?

**The CI would get tighter.**

C) What would typically happen to the above confidence interval if, instead of collecting a sample of 100 individuals, we collected a sample of 10,000 individuals?

**The CI would get tighter**

D) Given the above confidence interval, what is the best estimate of the population mean?

**The best guess for the population mean is the sample mean we got (from the first part of the course), and the sample mean is right in the middle of the CI, 70.**

33) If we know that the variance in the population is equal to 100, the distribution of sample means based on an  $n = 100$  will have a standard error of \_\_\_\_\_ **1** ( $10/\sqrt{100}$ )\_\_\_\_\_

**Interpretation of statistical output questions (1 of these on exam, ~10 points)**

**Note: I will not ask you to re-create R commands on the exam**

34) A researcher examined if students at CU score higher on the math section of the SAT than students at other local universities using a survey. Here are the results of an independent-measures ANOVA comparing the mean SAT math scores. (Note that the “Residuals” row in R is also known as “Within”)

The mean score for CU is 587.5, CSU is 525, and NU is 535

```
      Df Sum Sq Mean Sq F value Pr(>F)
group  2   9017    4508   4.471 0.0448 *
Residuals  9   9075    1008
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Write a 4-sentence summary of the results:

**We are interested in whether there is a difference in average SAT scores between students at CU Boulder, Colorado State, and U Nebraska. We collected data from 12 randomly selected individuals attending the three universities. We found that CU students have the highest SAT scores (mean=587.5) followed by NU (mean=535) and then CSU (mean=525;  $F(2,9)=4.47$ ,  $p=.045$ ;  $\eta^2=.50$ ). We conclude that there is a significant difference between the SAT scores between these universities.**

**Note: Few details were provided for this question, so I filled some in (e.g., “randomly selected”). It is fine to take such liberties if they’re not supplied on the test, although the test questions will be better vetted than this.**

35) Below is an example of a two-sample t-test in R:

```
> t.test(lab_survey$enjoys_news_politics ~ lab_survey$class_status, var.equal=TRUE)
```

Two Sample t-test

```
data: lab_survey$enjoys_news_politics by lab_survey$class_status
t = 0.4637, df = 104, p-value = 0.6439
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.3568480 0.5746405
sample estimates:
mean in group LOW mean in group UPPER
```

3.093023

2.984127

A) Is this a dependent (repeated-measures) samples t-test?

**No**

B) What is the dependent variable?

**lab\_survey\$enjoys\_news\_politics**

C) What is the independent variable?

**lab\_survey\$class\_status**

D) What is the size of the total sample (N)?

**106**

E) What is the mean difference between the two groups?

**.11**

F) Is the null hypothesis contained in the 95 percent confidence interval?

**Yes (null is that the mean difference is 0)**

36) You and your friends are constantly arguing about which cheap beer tastes better – PBR or Rolling Rock. You are a big fan of PBR and decide to survey people walking down the street (as long as they are 21, of course) by asking them to sample one of the beers while blindfolded and rate the taste on a scale of 1 to 7 (7 being the best). Each person you survey only tasted one of the beers. Here are the results:

Person	Beer	Rating
1	RoRo	1
2	RoRo	1
3	RoRo	2
4	RoRo	1
5	PBR	2
6	PBR	3
7	PBR	2
8	PBR	4

A) What would be the best statistical test to perform on these data?

**Independent samples t-test**

B) What are your null and alternative hypotheses?

**$H_0: \mu_1 = \mu_2$**   
 **$H_A: \mu_1 \text{ does not equal } \mu_2$**

You conduct an appropriate statistical test in R and produce these results:

data:

$t = 2.7775$ ,  $df = 6$ ,  $p\text{-value} = 0.0321$

95 percent confidence interval:

0.1785166 2.8214834

sample estimates:

mean of PBR	mean of RoRo
2.75	1.25

C) You tell your friends about the results of your analysis. But one of your friends is curious about how much better PBR tastes than Rolling Rock and you tell him to give you a minute while you compute the effect size. What is Cohen's  $d$  for your test?

**Two ways to get this.**

**WAY 1**

**$2.78 = (2.75 - 1.25)/SEM$**

**$SEM = (2.75 - 1.25)/2.78 = 1.5/2.78 = .54$**

**$.54 = \sqrt{sp^2/4 + sp^2/4} = \sqrt{sp^2/2} = \sqrt{1/2} * sp$**

**$sp = .76$**

**Cohen's  $d = 1.5/.76 = 1.97$**

**WAY 2 – from raw scores:**

**$SS1 = .75$**

**$SS2 = 2.75$**

**$sp^2 = (.75 + 2.75)/6 = .583$**

**$sp = .763$**

**Cohen's  $d = 1.5/.76 = 1.97$**

D) Write a 4-sentence summary for your analysis of PB versus Rolling Rock:

You and your friends are constantly arguing about which cheap beer tastes better – PBR or Rolling Rock. You are a big fan of PBR and decide to survey people walking down the street (as long as they are 21, of course) by asking them to sample one of the beers while blindfolded and rate the taste on a scale of 1 to 7 (7 being the best). Each person you survey only tasted one of the beers.

**My friend and I want to know whether PBR beer or RR beer tastes better. To investigate this, we randomly assigned 4 individuals to a blind taste test of PBR and 4 to a blind taste test of RR and asked them to rate the beer on a 1 (lowest) to 7 (highest) scale. We found that the mean of PBR (2.75) was significantly higher than**

the mean of RR (1.25;  $t(6)=2.78$ ,  $p = .03$ , Cohen's  $d = 1.97$ ). We conclude that PBR tastes better than RR, at least to individuals in this community.

**Performing statistical tests by hand (2 on exam, ~ 40 points total)**

**Note: you will need to choose the correct test to perform on the test.**

37R) A researcher is examining the effects of a new drug in healthy adults. Previous trials have shown greatly enhanced general cognitive function in elderly patients with signs of dementia shortly after having been administered this drug. To evaluate if this drug has any effect in healthy adults, she gives a sample of college students the drug and then gives them IQ tests. IQ is normally distributed in the general population, with a mean of 100, and standard deviation of 15. The average measured IQ after taking the drug is 120.

A) What is the effect size of taking the drug?

**Cohen's  $d = (120-100)/15 = 1.33$**

B) If there are 10 participants in the sample, what is the z-score for this sample? Does the sample mean differ significantly from what is expected by chance?

**$z = (120-100)/(15/\sqrt{10}) = 4.21$ ,  $p < .05$ . Yes, the sample mean is significantly different than what would be expected to arise by chance if the null was true.**

38R) The researcher is concerned that the results of his first sample may not accurately reflect the population because most of her subjects were very healthy (possibly related to above average cognitive function). To try and correct for any bias in her sample, she gives a new set of subjects 2 IQ tests, one before they are administered the drug and one after. Here are their scores:

Subject	Test 1	Test 2	Difference
1	94	93	-1
2	95	97	2
3	97	96	-1
4	103	105	2
5	108	109	1
6	104	106	2
7	109	113	4
8	116	123	7
9	123	124	1
10	140	148	8

A) What is the t-statistic evaluating the effect of the drug?

**mean = 2.5**

**SS = 82.5**

**$s = \sqrt{82.5/9} = 3.03$**

$$t(9) = 2.5/(3.03/\sqrt{10}) = 2.61$$

B) If the critical t-value is 2.283 for an alpha of .05, what is the confidence interval for the mean difference in IQ scores due to the effect of the drug? Can we reject the null hypothesis?

$$CI = 2.5 - 2.283 * .958, 2.5 + 2.283 * .958 = [.312, 4.69]$$

39R) The owner of The Sink (a local bar in Boulder near campus) receives multiple shipments of beer per week. She wants to cut down on unnecessary over-stock, while at the same time being sure enough beer is available. To help inform her decision she examines if people tend to drink more beer on the weekends, compared to the weekday. 5 different customers are surveyed on a random night during the week, and another five on the weekend, and asked to estimate how many beers they plan to drink.

A) Compute an independent samples t-test for the following data to examine if a significant difference exists. What do you conclude? (Note:  $t_{crit}$  for  $\alpha = .05$  and these  $df$  is 2.26).

Weekday night (Sunday-Thursday)	Weekend night (Friday-Saturday)
2	4
3	5
3	8
1	3
3	6
	4

$$sp^2 = (SS1 + SS2)/(df1 + df2) = (3.2 + 16)/(4 + 5) = 2.133$$

$$SEM_d = \sqrt{sp^2/n1 + sp^2/n2} = \sqrt{2.133/5 + 2.133/6} = .884$$

$$t(9) = (2.4 - 5)/.884 = -2.94, p < .05$$

**We reject the null hypothesis**

B) Report the group means and the appropriate measure of effect size.

$$\text{mean1} = 2.4$$

$$\text{mean2} = 5$$

$$\text{Cohen's } d = -2.06/\sqrt{2.133} = 1.78$$

C) What is your 95% confidence interval for your mean difference?

$$CI = [-4.6, -.60]$$

40) An owner of a liquor store wants to examine if there is a significant difference between the amounts of beer purchased depending on whether it is 1) Sunday-Wednesday,

2) Thursday, or 3) Friday-Saturday. Data is collected during each of these time periods from different customers. The number of beers purchased are:

Sun-Wed	Thurs	Fri/Sat
0	2	4
0	3	5
1	2	8
0	1	3
1	2	6
		4

A) Fill in the ANOVA summary table below

Source	Sum Sq	df	Mean Sq	F value
Between groups	60.55	2	30.275	20.5
Within groups	19.2	13	1.48	
Total	79.75	15		

B) Report an appropriate measure of effect size for this finding.

$$\eta^2 = 60.55/79.75 = .76$$

C) Write a four-sentence summary of your findings.

**We are interested in whether the amount of beer purchased depends on the day of a week for a specific liquor store. We collected information on number of beer purchased during a weeklong period. We found that the average amount of beer purchased daily on Fri/Sat (mean=5) is higher than the amount purchased on Thur (mean = 2), which in turn was higher than the amount purchased on the weekday (mean = .4;  $F(2,13)=20.5$ ,  $p<.05$ ). We conclude that there are significant differences in daily beer purchases depending on day of the week at this liquor store.**

41) A psychologist is investigating the hypothesis that children who grow up as the only child in the household develop different personality characteristics than those who grow up in larger families. A sample of  $n=30$  only children is obtained and each child is given a standardized personality test. For the general population, scores on the test vary normally with a mean of 50 and a standard deviation of 15. We find the mean of the sample is 58.

A) Perform the correct statistical test to understand whether there is a significant difference in personality between only children and the rest of the population.

$$z = (58-50)/(15/\sqrt{30}) = 8/2.74 = 2.92$$

B) Find an appropriate measure of effect size for your test above.

$$\text{Cohen's } d = 8/15 = .533$$

C) Write a 4-sentence summary of your findings.

**Children who grow up as only children have been hypothesized to be different in personality than children from larger families. To test this, we collected data on 30 only-children and administered a standardized personality test with known mean of 50 and standard deviation of 15. We found that only children scored significantly higher on this test (58) than the general population ( $z = 2.92$ ,  $p < .05$ ). We conclude that only children do appear to have different personalities than the general population.**

42) Siegel (1990) found that elderly people who owned dogs were less likely to pay visits to their doctors after upsetting events than were those who did not own pets. Similarly, consider the following hypothetical data. A sample of elderly dog owners is compared to a similar group (in terms of age and health) who do not own dogs. The researcher records the number of visits to the doctor during the past year for each person. The data are:

Non-dog owners: 10, 8, 7, 9, 13, 7, 7, 12

Dog owners: 7, 4, 9, 3, 7

Assume that number of doctor visits are normally distributed in the population.

A) Is the number of doctor visits significantly different for dog owners than for non-dog owners? Note that  $t_{crit}$  is 2.20 for an alpha of .05 for these  $df$ .

$$SS1 = 38.9$$

$$SS2 = 24$$

$$sp^2 = (38.9 + 24)/(7 + 4) = 5.71$$

$$t(11) = (9.125 - 6)/\sqrt{5.71/8 + 5.71/5} = 2.29, p < .05$$

B) Provide a proper effect size estimate and provide an interpretation of it in a single sentence.

**Cohen's  $d = 3.125/\sqrt{5.71} = 1.31$ . The mean number of doctor visits for non-dog owners is 1.31 standard deviation units above the mean number of dog owners.**

C) Provide a 95% confidence interval for these findings.

$$CI = 3.125 - 2.2 * 1.36, 3.125 + 2.2 * 1.36 = [.128, 6.12]$$

D) What, in words, does your confidence interval above mean?

**This is the interval in which the true mean of the population lies with 95% confidence.**

43) You want to compare a sample of entrance test scores for a CU Boulder psych class with the entrance test scores at CU Denver. Here are the test scores:

7, 12, 13, 19, 5, 5

The only information you have regarding CU Denver test scores is the mean score, which is 9 points. Assuming both your sample and the CU Denver mean score come from a normal distributed set of scores:

A) What test would you use to test if there is a difference between your sample of CU Boulder scores and the CU Denver mean score?

**single-sample t-test**

B) What is the mean of your sample of CU Boulder scores?

**10.17**

C) How many degrees of freedom does your test have?

**5**

D) What is your alpha level?

**.05**

E) What is the standard error of the mean of your sample of CU Boulder scores?

**SEM = 2.26**

F) What is your test statistic?

**$t(5) = (10.17 - 9)/2.26 = .517$**

G) Do you reject or fail to reject the null hypothesis? (Note: you should be able to answer this without a critical value).

**We fail to reject the null hypothesis. We don't need a critical value here because the t value is so small. Because the tcrit is always  $\geq$  zcrit, if  $\text{abs}(t\text{-value})$  is ever  $< 1.96$ , you know that result is not significant at  $\alpha=.05$ .**

44R) The following independent observations were collected in an experimental study. Use the appropriate test to determine if there is a significant difference between treatments (assume that the scores in the population are normally distributed). Note that the tcrit is 2.31 for these *df*:

Treatment 1                      Treatment 2

2	6
3	10
10	14
8	12
12	18

$$SS1 = 76$$

$$SS2 = 80$$

$$sp^2 = (76+80)/8 = 19.5$$

$$SEMd = \sqrt{(19.5/5 + 19.5/5)} = 2.79$$

$$t(8) = (7 - 12)/2.79 = -1.79, p > .05 \text{ (do not reject the null)}$$

45) For the problem above, assume that each row corresponds to the same person who was measured twice. Perform the appropriate test to determine if there is a significant difference between treatments. Note that the  $t_{crit}$  is 2.78 for these  $df$ .

$$\text{differences} = -4, -7, -4, -4, -6$$

$$SS = 8, df = 4, s^2 = 2, s = 1.4$$

$$SEM = 1.4/\sqrt{5} = .63$$

$$t(4) = -5/.63 = -7.93, p < .05, \text{ reject the null}$$

46R) Suppose I want to run an experiment on the effectiveness of caffeine as performance enhancement in elite runners. One group of subjects consumes 120mg of caffeine dissolved in flavored water prior to running 5 kilometers. The other group of subjects is simply given flavored water (identical, but without the caffeine). Below are the summary data for my study. Calculate a t-statistic and interpret the results of this study using a 4 sentence summary (Note: assume that the scores in the population are normally distributed;  $t_{crit}$  for these  $df$  is 2.01):

<u>Caffeine Group</u>	<u>No Caffeine Group</u>
M = 14.37 minutes	14.81 minutes
SS = 15	SS = 20
n = 25	n = 25

$$sp^2 = (15 + 20)/48 = .729$$

$$SEMd = \sqrt{(.729/25 + .729/25)} = .241$$

$$t(48) = (14.37 - 14.81)/.241 = -1.83, p > .05 \text{ (do not reject the null)}$$

$$\text{Cohen's } d = -.52$$

**It has been hypothesized that caffeine enhances athletic performance. To test this, we gave 25 randomly assigned runners a drink with caffeine in it and 25 runners an identically flavored drink without caffeine and timed their 5 km run. Runners who had caffeine ran faster (mean = 14.37 minutes) than runners who did not have caffeine (mean = 14.81; Cohen's d = -.51), but this was not a significant difference (t(48) = -1.83, ns). We conclude that there is not sufficient evidence to claim that caffeine improves athletic performance as measured by long distance running speed.**

47R) I am running a study on the effects of cell phone usage and driving. I have three independent groups that I put into a driving simulator while I measure the number of errors that each group makes while driving (e.g., exceeding the speed limit, breaking too late, failing to signal, hitting fake pedestrians, etc.). One group is simply instructed to drive the course as accurately as possible (the CONTROL group), a second group is required to maintain a cell phone conversation while driving the course (CELL group), a third is required to maintain a conversation via text messaging while driving the course (TEXT group). The number of errors and summary data for each group are given below.

Control	Cell	Text
1	1	4
0	4	3
1	1	6
2	2	3

A) Perform the appropriate analysis to understand whether there is a difference in errors while driving depending on condition. What is your conclusion? (Note: the  $F_{crit}$  for these degrees of freedom and an alpha of .05 is 4.26).

**SSw = 14**

**SSb = 18.67**

**MSw = 9.33**

**MSb = 1.56**

**F(2,9) = 6, p < .05**

B) Report an appropriate measure of effect size. Put into words what this effect size measure tells you.

**$\eta^2 = 18.67/(18.67+14) = .57$ . 57% of variation in driving errors is associated with the levels of the driving factor.**

48) Your gym is interested in whether or not the addition of a hot tub will make their customers more satisfied. Before announcing the new hot tub, the administration randomly surveys 5 people and rates (on a scale of 1 to 7) their satisfaction with the gym. After they install the new hot tub, the administration finds and resurveys the same people again. The satisfaction ratings are below:

Person	satisfaction before hot tub	Satisfaction after hot tub
1	5	7

2	3	6
3	4	6
4	5	6

A) What is the appropriate type of t-test for these data? Why?

**repeated measures t-test because the same person is in both groups**

B) Calculate the appropriate t-statistic for these data by hand:

**difference scores = -2, -3, -2, -1**

**SD = .816**

**SEM = .816/sqrt(4) = .408**

**t = -2/.408 = 4.91**

C) We could use R (the pt() function) or a table to determine the significance of the t statistic you just found. What piece of information, besides the actual t statistic for found, do you need to determine the p-value?

**We need the degrees of freedom. Here, there are 3 degrees of freedom.**

D) Calculate Cohen's d for the effect you found:

**Cohen's d = -2/.816 = -2.45**

49) A pharmaceutical company has developed a drug that is expected to reduce hunger. To test the drug, three samples of rats are selected with n=10 in each sample. The first sample receives the drug every day. The second sample is given the drug once a week, and the third sample receives no drug at all. The dependent variable is the amount of food eaten by each rat over a 1-month period. These data are analyzed using an ANOVA, and the researchers find an F value of 7.5 and a  $MS_{\text{between}}$  of 15.

A) What are the  $df_{\text{between}}$ ,  $df_{\text{within}}$ , and  $df_{\text{total}}$ ?

**dfb = 2, dfw = 27, dft = 29**

B) What is the  $MS_{\text{within}}$ ?

**$MS_w = 15/7.5 = 2$**

C) What are the  $SS_{\text{between}}$  and  $SS_{\text{within}}$ ?

**$SS_b = MS_b * df_b = 15 * 2 = 30$**

**$SS_w = MS_w * df_w = 2 * 27 = 54$**

D) What is an appropriate measure of effect size for this finding?

**$\eta^2$**