STAT511 – Exam 2
Fall 2019

Honor Pledge: I have not given, received, or used any unauthorized assistance on this exam.

Signature: KEY

Printed Name: ______

Instructions:

- **Open book, open notes, calculator required. No computers or cell phones.**
- **Time limit is 1 hour 50 minutes - strictly enforced!**
- If an answer is in the computer output, use it; don't calculate it by hand.
- Show your work where appropriate. Put your final answer in the box (if provided).
- Make explanations brief and legible.
- All questions are worth 4 points except where noted. Maximum score is 100.
- Computer input/output is provided at the end of the exam.
- The exam contains a total of 10 pages (including computer input/output).
- If you run out of space, you may use page 6 or extra paper.

Questions 1 and 2 (Voters): Suppose it is known that 40% of Colorado adults voted in the 2019 election. Consider a random sample of n = 4 Colorado adults. Let Y be the random variable representing the number who voted out of the sample of size n = 4. Hint: Y follows a binomial distribution.

$n=4, \pi=0.40$

1. Give the mean of Y.

$4 \times 0.4 = 1.6$

$\boxed{1.6}$

2. What is the probability that all 4 of the people in the sample voted. In other words, find P(Y = 4). Note: 0! = 1, 1! = 1, 2! = 2, 3! = 6, 4! = 24. Show your work for full credit.

$$P(Y=4) = \frac{4!}{4!\,0!}(0.4)^4(0.6)^0$$
$$= 0.0256$$

$\boxed{0.0256}$

Question 3 (Prop CC #1):

3. Suppose prior to the election, a poll is being planned to estimate the support for Proposition CC in Colorado. They conjecture that 75% of Colorado voters support Proposition CC. The investigator would like the 95% ME to be 4% (0.04) or less. What is the (minimum) sample size required? Return answer as an integer value.

$$n = \frac{(1.96)^2\, 0.75 * 0.25}{(0.04)^2} = 450.1875$$

OK 450, 469.

$\boxed{451}$

-2 for 230

12

Questions 4 through 8 (Prop CC #2): Now suppose a poll was conducted by contacting a random sample of n = 525 Colorado voters. Of these, 289 support Proposition CC. Let π (pi) be the true proportion of Colorado voters who support the proposition. Use alpha = 0.05. Give your answers to three decimal places.

```
> prop.test(289, 525, correct = FALSE)

1-sample proportions test without continuity correction

data:  289 out of 525, null probability 0.5
X-squared = [redacted], df = 1, p-value = 0.02072
alternative hypothesis: true p is not equal to 0.5
95 percent confidence interval:
 0.5077114 0.5925077
sample estimates:
        p
[redacted]
```

4. Use the information provided to estimate the proportion (π) of Colorado voters who support Proposition CC.

$\hat{\pi} = 289/525 = 0.550$

$\boxed{0.55}$

5. Give the 95% ME corresponding to the estimate from the previous question.

$95\% \text{ ME} = 1.96\sqrt{\frac{0.55(1-0.55)}{525}}$

$= 0.0425$

0.044 OK

$\boxed{0.043}$

6. Briefly explain why the ME from the previous question is larger than 0.04, even though the sample size (n = 525) is greater than what you found in question 3 (Prop CC #1). Note: Two hints for previous questions appear here!

$\hat{\pi} = 0.55$ is closer to 0.5 than our original conjecture of 0.70.

7. Now suppose that we want to test $H_0: \pi \le 0.5$ vs $H_A: \pi > 0.5$.
   A. Calculate the Z test statistic. Show your work for full credit.

$Z = \frac{0.55 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{525}}} = 2.313$

2.291 OK.

$\boxed{2.313}$

   B. Calculate the (one-sided) p-value corresponding to the previous question.

$P(Z > 2.313) = 1 - 0.9896$

$= 0.0104$

0.011 OK

$\boxed{0.010}$

8. The large sample normal approximation is adequate here. Just circle one answer, no need to justify. (2pts)

True   False

22

Questions 9 through 11 (Blood Pressure): A pilot study was done to test a new blood pressure medication. The investigator recruited n = 10 subjects (with systolic blood pressure of 125 or greater) to participate in the study. Each subject has their blood pressure recorded before starting treatment and then 4 weeks after treatment. Use ~~α~~ alpha = 0.05. The R input and output are labeled **Blood Pressure.**

9. A colleague considers the summary statistics and is concerned that the standard deviation is noticeably larger for Post (s = 8.04) than for Pre (s = 5.81). Explain why there is no assumption of equal variances when using the paired t-test.

Paired t-test is based on differences. (Post - Pre for each subject)
So, only the sd of the differences is used.
(Hence a single sd.)

10. Both the paired t-test and Wilcoxon (paired) signed-rank test are shown in the output.

A. Which of these methods has higher power? Briefly justify your response.

Paired t-test has higher power.
- Paired t-test has smaller p-value
- Non-parametric methods known to have lower power.

OK Reduced sample size for Wilcoxon due to diffs = 0.

B. Discuss the difference in assumptions for these two methods. Be specific.

Paired t-test requires assumption that the differences are normally distributed.

C. Based on the output provided, can we evaluate the assumption you identified in the previous question? Briefly discuss.

No! With only means, sds cannot check assumption of normality.
(Some people examined the differences - OK for full credit).

11. Now consider using the two-sample t-test for this analysis. This analysis is not shown in the output.

A. Compared to the paired t-test, would the two-sample t-test have higher or lower power? Briefly justify your response. Hint: Consider the summary statistics.

Two-sample t-test will have lower power.
SD of differences is less than sd of Pre, Post.

B. Would the two-sample t-test be reasonable for this analysis? Briefly discuss. Hint: Think about assumptions.

No, paired obs (Pre, Post) for each subject.
Two-sample t-test requires independent obs.

-2 for unequal variances.

24

Question 12 through 20 (RHI): An experiment was done to compare 4 probiotic treatments (A, B, C, D). Initially, 100 subjects were randomly assigned to one of the 4 treatments. But due to drop out, data was collected for a total of n = 85 subjects. The response of interest was RHI_Diff (comparing post treatment RHI versus pre treatment RHI). RHI stands for reactive hyperemia index which is a measure of endothelial function. Higher RHI generally implies better health. Use $\alpha = 0.05$. The R input and output are labeled **RHI.**

12. A colleague considers the summary statistics and is concerned about unequal sample sizes. He suggests that you "drop some observations to even out the sample sizes." Do you agree or disagree with this suggestion? Just circle one answer, no need to justify. (2pts)

    Agree   Disagree

13. For the majority of analysis, treatments were labeled A, B, C, D. But when creating the boxplots, different numeric labels were "accidentally" used. Resolve the following (numeric) groups to their appropriate (letter) labels. Hint: Consider the summary statistics.

    Group #3 = Trt A

    Group #4 = Trt D

14. Based on the ANOVA output F = 2.30, p = 0.08. State the hypotheses for this test. Be specific.

    H_0: $\mu_A = \mu_B = \mu_C = \mu_D$

    H_A: Not all the means are the same

    -2 for H_A: $\mu_A \neq \mu_B \neq \mu_C \neq \mu_D$

15. Is the assumption of equal variances satisfied? Briefly discuss one piece of evidence provided in the output. As part of the discussion, clearly state the test(s) and/or graph(s) you are considering.

    - levene's test: $p < 0.05$ suggests unequal variances
    - Resids vs Fitted: Slight reverse megaphone, but not bad.

16. Is the assumption of normality satisfied? Briefly discuss one piece of evidence provided in the output. As part of the discussion, clearly state the test(s) and/or graph(s) you are considering.

    - S-W test: $p > 0.05$ supports normality
    - QQ plot of Resids: Very linear, looks good!

0

17. Notice that "emout" object includes 95% confidence intervals (lower.CL, upper.CL). Considering this information, do we have evidence that any of the treatments were effective in increasing RHI? Hint: Consider the fact that our response variable represents a difference (Post – Pre). Discuss.

    No evidence of increase (or difference) for any Trt.
    All of the CIs include zero.

22

18. Based on the "pairs" output, briefly summarize the conclusions of the study. In other words, discuss what treatment pairs show evidence of differences and in what direction. Note: By direction, I just mean which treatment has a higher mean response.

We have evidence that

$\mu_C > \mu_A \quad (p = 0.0285)$

$\mu_D > \mu_A \quad (p = 0.0217)$

19. Before the experiment was started, the investigator planned to compare the average of response for active treatments (A, B, C) versus control treatment (D).

A. Provide appropriate contrast coefficients.

A	B	C	D
1/3	1/3	1/3	-1

B. Use information provided to estimate this contrast. Show your work for full credit.

$$\frac{(-0.279 - 0.010 + 0.177)}{3} - 0.209 = -0.246$$

+0.246 (ok)

$\boxed{-0.246}$

20. The current output shows unadjusted pairwise comparisons. A colleague considers the output and recommends that you use Tukey adjusted pairwise comparisons.

A. Discuss whether this is reasonable and if so why you would want to do it.

Yes, Tukey can be used to control experiment-wise error rate (EER).

B. If you used a Tukey adjustment, would you expect more or fewer "significant" differences as compared to the current (unadjusted) analysis? No need to justify. **(2 pts)**

Fewer differences using Tukey

C. If we are only interested in comparing each of the active treatments (A, B, D) versus control treatment (D), name a multiple testing method that would offer higher power than Tukey. No need to justify. **(2 pts)**

Dunnett

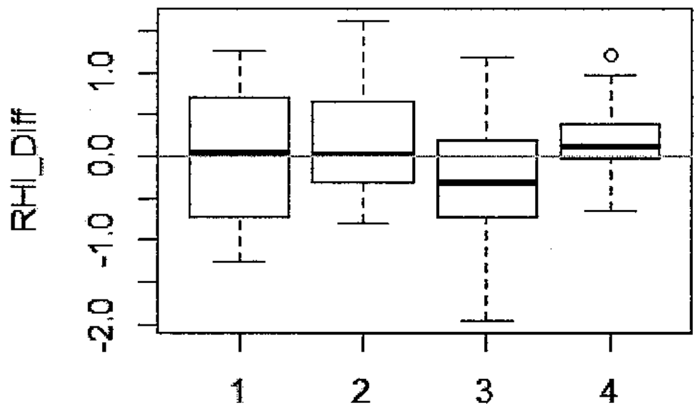
20

Blood Pressure (Questions 9 - 11)

The following variables are included in the data:
Subject: Subject ID (1-10)
Pre: pre-treatment systolic blood pressure (before treatment)
Post: post-treatment systolic blood pressure (after treatment)
Diff: Post – Pre

```
library(tidyverse)
library(coin)
str(BPData)

## Classes 'tbl_df', 'tbl' and 'data.frame':    10 obs. of  4 variables:
##  $ Subject: Factor w/ 10 levels "1","2","3","4",..: 1 2 3 4 5 6 7 8 9 10
##  $ Pre    : num  134 138 138 138 137 140 128 128 126 125
##  $ Post   : num  129 134 139 138 136 141 122 128 120 120
##  $ Diff   : num  -5 -4 1 0 -1 1 -6 0 -6 -5

SumStats <- BPData %>%
        select(-Subject) %>%
        gather(key = "Var", value = "Y") %>%
        group_by(Var) %>%
        summarize(n = n(),
                  mean = mean(Y),
                  sd = sd(Y)) %>%
        arrange(desc(Var))
SumStats

## # A tibble: 3 x 4
##   Var       n  mean    sd
## 1 Pre      10 133.0  5.81
## 2 Post     10 131.0  8.04
## 3 Diff     10  -2.5  2.95

t.test(BPData$Post, BPData$Pre, paired = TRUE)
##  Paired t-test
##
## data:  BPData$Post and BPData$Pre
## t = -2.6769, df = 9, p-value = 0.02534
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -4.6126928 -0.3873072
## sample estimates:
## mean of the differences
##                    -2.5

wilcoxsign_test(Post ~ Pre, data = BPData, distribution = "exact")
##  Exact Wilcoxon-Pratt Signed-Rank Test
##
## data:  y by x (pos, neg)
##   stratified by block
## Z = -1.8541, p-value = 0.07812
## alternative hypothesis: true mu is not equal to 0
```

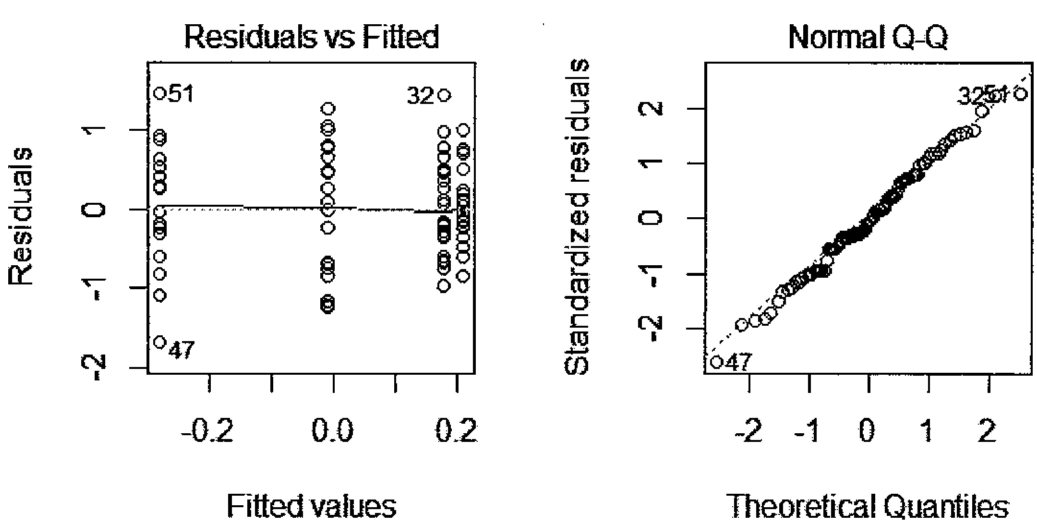
RHI (Questions 12 - 20)




```
library(tidyverse)
library(car)
library(emmeans)
str(RHIData)

## 'data.frame':    85 obs. of  3 variables:
##  $ Subject : int  1 6 14 17 19 24 25 27 32 34 ...
##  $ Trt     : Factor w/ 4 levels "A","B","C","D": 2 2 2 2 2 2 2 2 2 2 ...
##  $ RHI_Diff: num  -0.76 0.25 0.44 -0.23 -1.25 0.79 -0.02 -0.01 1.03 0.46 ...

SumStats <- RHIData %>%
        group_by(Trt) %>%
        summarize(n = n(),
                  mean = mean(RHI_Diff),
                  sd = sd(RHI_Diff),
                  SE = sd/sqrt(n),
                  min = min(RHI_Diff),
                  median = median(RHI_Diff),
                  max = max(RHI_Diff))
SumStats

## # A tibble: 4 x 8
##   Trt       n   mean    sd     SE   min median   max
##   <fct> <int>  <dbl> <dbl>  <dbl> <dbl>  <dbl> <dbl>
## 1 A        19 -0.279  0.764 0.175  -1.95 -0.32   1.18
## 2 B        20 -0.010  0.810 0.181  -1.25  0.04   1.25
## 3 C        24  0.177  0.613 0.125  -0.8   0.015  1.62
## 4 D        22  0.209  0.449 0.0957 -0.65  0.12   1.2
```

RHI continued (Questions 12 - 20)

```
Model <- lm(RHI_Diff ~ Trt, data = RHIData)
anova(Model)

## Analysis of Variance Table
##
## Response: RHI_Diff
##           Df Sum Sq Mean Sq F value  Pr(>F)
## Trt        3  3.057 1.01889  2.3021 0.08327 .
## Residuals 81 35.850 0.44259
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

leveneTest(RHI_Diff ~ Trt, data = RHIData)

## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value  Pr(>F)
## group  3  3.1313 0.03007 *
##       81
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

shapiro.test(Model$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  Model$residuals
## W = 0.99261, p-value = 0.9165

par(mfrow = c(1, 2))
plot(Model, c(1:2))
```



RHI continued (Questions 12 - 20)

```
emout <- emmeans(Model, "Trt")
emout

##  Trt emmean    SE df lower.CL upper.CL
##  A   -0.279 0.153 81  -0.5826   0.0247
##  B   -0.010 0.149 81  -0.3060   0.2860
##  C    0.177 0.136 81  -0.0935   0.4469
##  D    0.209 0.142 81  -0.0736   0.4908
##
## Confidence level used: 0.95

pairs(emout, adjust = "none")

##  contrast estimate    SE df t.ratio p.value
##  A - B      -0.269 0.213 81  -1.262  0.2106
##  A - C      -0.456 0.204 81  -2.230  0.0285
##  A - D      -0.488 0.208 81  -2.340  0.0217
##  B - C      -0.187 0.201 81  -0.927  0.3568
##  B - D      -0.219 0.206 81  -1.064  0.2906
##  C - D      -0.032 0.196 81  -0.163  0.8711
```

Exam2 “Extra” output

This information was not provide in the original exam.

Voters (Question 2)

```
dbinom(4, size = 4, prob = 0.40)

## [1] 0.0256
```

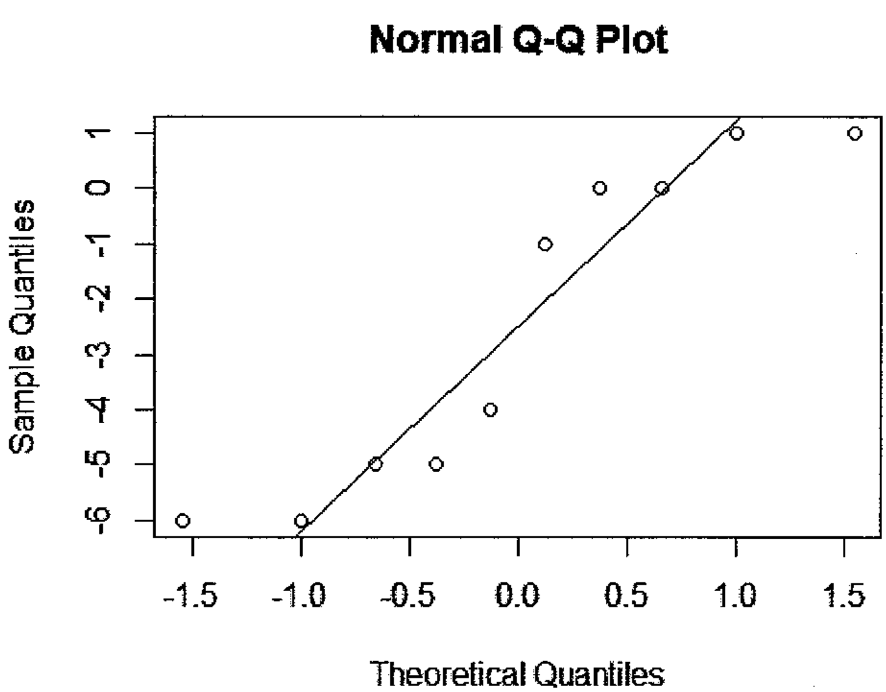
Prop CC #2 (Questions 4-8)

```
prop.test(289, 525, correct = FALSE, alternative = "greater")

##
##  1-sample proportions test without continuity correction
##
## data:  289 out of 525, null probability 0.5
## X-squared = 5.3505, df = 1, p-value = 0.01036
## alternative hypothesis: true p is greater than 0.5
## 95 percent confidence interval:
##  0.5145978 1.0000000
## sample estimates:
##         p
## 0.5504762

sqrt(5.3505)

## [1] 2.313115
```

Blood Pressure (Questions 9 - 11)

```
t.test(BPData$Post, BPData$Pre, paired = TRUE, var.equal = TRUE)

##
##  Paired t-test
##
## data:  BPData$Post and BPData$Pre
## t = -2.6769, df = 9, p-value = 0.02534
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -4.6126928 -0.3873072
## sample estimates:
## mean of the differences
##                    -2.5
```

```
t.test(BPData$Post, BPData$Pre)

##
##  Welch Two Sample t-test
##
## data:  BPData$Post and BPData$Pre
## t = -0.79693, df = 16.38, p-value = 0.4369
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -9.137728  4.137728
## sample estimates:
## mean of x mean of y
##     130.7     133.2

shapiro.test(BPData$Diff)

##
##  Shapiro-Wilk normality test
##
## data:  BPData$Diff
## W = 0.84091, p-value = 0.04525

qqnorm(BPData$Diff);qqline(BPData$Diff)
```



RHI (Questions 12 - 20)

```
library(emmeans)
table(RHIData$Group, RHIData$Trt)

##
##       A  B  C  D
##   1   0 20  0  0
##   2   0  0 24  0
##   3  19  0  0  0
##   4   0  0  0 22

Model <- lm(RHI_Diff ~ Trt, data = RHIData)
emout <- emmeans(Model, "Trt")
contrast(emout,list(
        Q19 = c(1/3, 1/3, 1/3, -1)))

##  contrast estimate    SE df t.ratio p.value
##  Q19        -0.246 0.165 81 -1.492  0.1397

pairs(emout)

##  contrast estimate    SE df t.ratio p.value
##  A - B      -0.269 0.213 81 -1.262  0.5896
##  A - C      -0.456 0.204 81 -2.230  0.1238
##  A - D      -0.488 0.208 81 -2.340  0.0975
##  B - C      -0.187 0.201 81 -0.927  0.7906
##  B - D      -0.219 0.206 81 -1.064  0.7125
##  C - D      -0.032 0.196 81 -0.163  0.9985
##
## P value adjustment: tukey method for comparing a family of 4 estimates
```