Assignment 3 final completed

Shafa

```
summarySE <- function(data=NULL, measurevar, groupvars=NULL, na.rm=FALSE,
                      conf.interval=.95, .drop=TRUE) {
  library(plyr)

  # New version of length which can handle NA's: if na.rm==T, don't count the
m
  length2 <- function (x, na.rm=FALSE) {
    if (na.rm) sum(!is.na(x))
    else       length(x)
  }

  # This does the summary. For each group's data frame, return a vector with
  # N, mean, and sd
  datac <- ddply(data, groupvars, .drop=.drop,
                 .fun = function(xx, col) {
                   c(N    = length2(xx[[col]], na.rm=na.rm),
                     mean = mean   (xx[[col]], na.rm=na.rm),
                     sd   = sd     (xx[[col]], na.rm=na.rm)
                   )
                 },
                 measurevar
  )

  # Rename the "mean" column
  datac <- rename(datac, c("mean" = measurevar))

  datac$se <- datac$sd / sqrt(datac$N)  # Calculate standard error of the mea
n

  # Confidence interval multiplier for standard error
  # Calculate t-statistic for confidence interval:
  # e.g., if conf.interval is .95, use .975 (above/below), and use df=N-1
  ciMult <- qt(conf.interval/2 + .5, datac$N-1)
  datac$ci <- datac$se * ciMult

  return(datac)
}

dat <- read.csv("A3data.csv")

if(!require("psych")) install.packages("psych")
```

```
Loading required package: psych

if(!require("car")) install.packages("car")

Loading required package: car

Loading required package: carData


Attaching package: 'car'

The following object is masked from 'package:psych':

    logit

if(!require("ggplot2")) install.packages("ggplot2")

Loading required package: ggplot2


Attaching package: 'ggplot2'

The following objects are masked from 'package:psych':

    %+%, alpha

if(!require("effsize")) install.packages("effsize")

Loading required package: effsize


Attaching package: 'effsize'

The following object is masked from 'package:psych':

    cohen.d

if(!require("pwr")) install.packages("pwr")

Loading required package: pwr

library(psych)
library(car)
library(ggplot2)
library(effsize)
library(pwr)
```

1. A researcher is testing the efficacy of a new drug that is intended to boost memory performance. First, assume the data come from two groups of 25 participants, randomly assigned to either the control or experimental condition, and the data are ratio-level. Calculate the appropriate t-test to compare the two group means using the appropriate formula and in R. Which formula did you use, and why is it appropriate? Report and discuss the effect size and post hoc power. (8 points)

```
psych::describe(dat)

            vars  n  mean   sd median trimmed  mad   min   max range  skew
memory_ctrl    1 25 78.91 5.87  78.17   78.84 5.78 68.19 91.47 23.28  0.18
memory_trt     2 25 82.59 7.59  83.34   82.72 9.01 65.45 95.73 30.28 -0.23
            kurtosis   se
memory_ctrl    -0.68 1.17
memory_trt     -0.67 1.52

(78.91 - 82.59) / (sqrt(((5.87^2)/25) + ((7.59^2)/25)))

[1] -1.917655

#df = 48
#Tcrit @ df of 40 = 2.021

t.test(dat$memory_ctrl, dat$memory_trt, var.equal = T)


    Two Sample t-test

data:  dat$memory_ctrl and dat$memory_trt
t = -1.915, df = 48, p-value = 0.06146
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -7.5352623  0.1835996
sample estimates:
mean of x mean of y 
 78.90934  82.58517 

effsize::cohen.d(dat$memory_ctrl, dat$memory_trt)


Cohen's d

d estimate: -0.5416397 (medium)
95 percent confidence interval:
      lower       upper 
-1.12066667  0.03738729 

pwr::pwr.t.test(n = 25, d = -.541, sig.level = 0.05, power = ,
                type = c("two.sample"))


     Two-sample t test power calculation 

              n = 25
              d = 0.541
      sig.level = 0.05
          power = 0.4659737
    alternative = two.sided
```

```
NOTE: n is number in *each* group
```

A1. With both the by hand method, and R method, we fail to reject the null hypothesis, and conclude that the groups are not significant different given p = .061

I assumed that all assumptions were met, and used the independent t-test formula which was appropriate because we are comparing two independent groups.

formula: (mean1 - mean2) / sqrt((variance1/n1) + (variance2/n2)), at a df = 48.

After completing effect size and post hoc power analysis, we see that our have a medium effect size, were d = .541, and our post hoc power is .466. This is low powered for our typical target of .80, therefore we would need to increase our sample size if we wished increase our power to .8. Given this information, we shouldn't be very confident in making inferences from this t-test analysis.

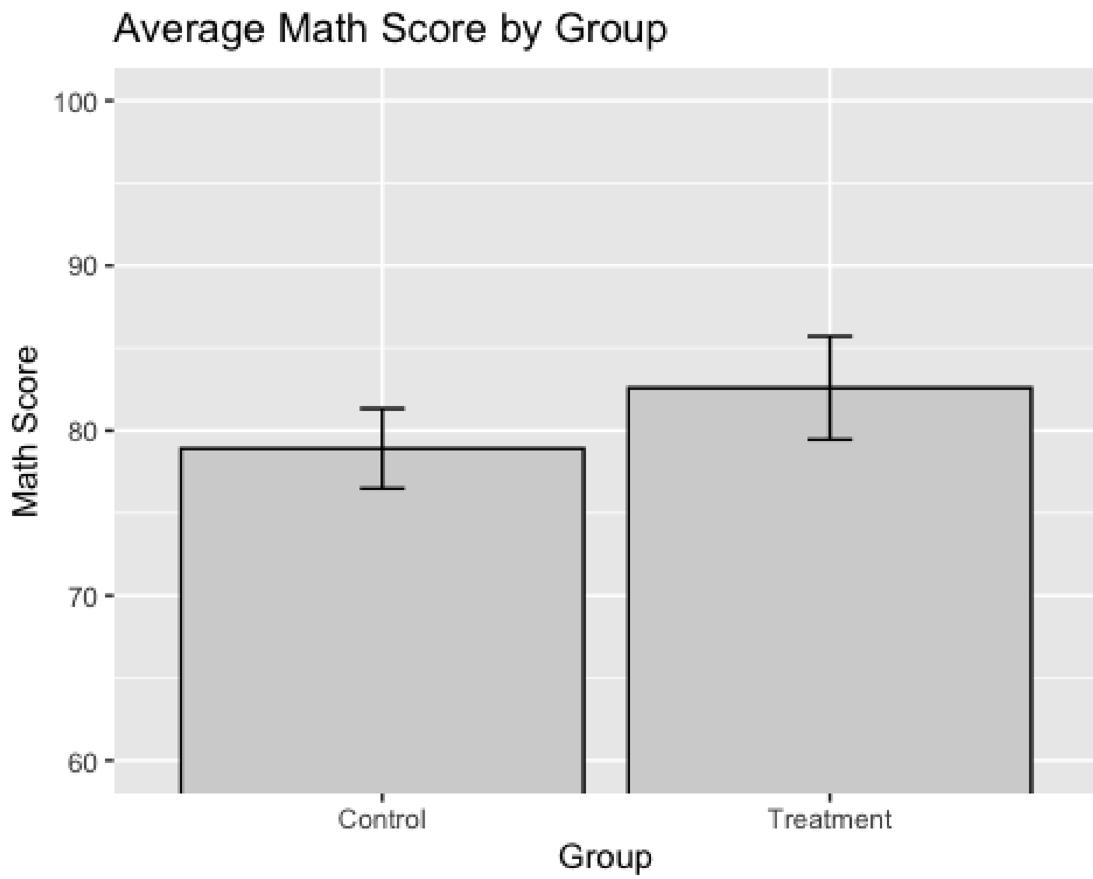
2. Using ggplot, create a graph illustrating the results of your analyses in Question 1. Explain your choices in creating the graph (e.g., type of plot, elements included). (2 points)

```
memdat <- data.frame(Group=(c(rep("Control",25), rep("Treatment",25))),
                       Score=c(dat$memory_ctrl, dat$memory_trt))

mem_sum <- summarySE(memdat, measurevar="Score", groupvars="Group")
mem_sum

      Group  N    Score       sd       se       ci
1   Control 25 78.90934 5.872931 1.174586 2.424227
2 Treatment 25 82.58517 7.590884 1.518177 3.133363

# Bar graph depicting means with SE error bars
barmem <- ggplot(data=mem_sum, aes(x=Group, y=Score)) +
  geom_col(fill="lightgray", col="black") +
  geom_errorbar(aes(ymin=Score-ci, ymax=Score+ci),
              width=.1) +
  coord_cartesian(ylim = c(60, 100)) +
  ggtitle("Average Math Score by Group") +
  labs(x = "Group", y = "Math Score")
barmem
```



A2. Here we see a bar graph comparing the means, with error bars of 95% CI. I am using a bar graph here because it better depicts independent groups better than a line graph, and CI error bars because they can be more conservative (wider spread) than SE error bars. We see that our error bars do overlap, concluding that our means may not be significantly different.

3. Now, analyze the same data assuming it comes from one group of 25 participants tested with and without the drug. Test whether the means differ using the appropriate formula and in R. Report and discuss the effect size and post hoc power. What are the similarities and differences between the present analyses/results and those in Question 1? (8 points)

```
psych::describe(dat$memory_ctrl - dat$memory_trt)

   vars  n  mean   sd median trimmed mad    min   max range skew kurtosis   se
X1    1 25 -3.68 7.8  -2.46   -3.69 7.6 -18.34 11.02 29.36    0    -0.47 1.56

-3.68 / (7.8 / (sqrt(25)))

[1] -2.358974

#df = 48
#Tcrit @ df of 40 = 2.021
```

```
t.test(dat$memory_ctrl, dat$memory_trt, paired=T)


	Paired t-test

data:  dat$memory_ctrl and dat$memory_trt
t = -2.3568, df = 24, p-value = 0.02693
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -6.8948801 -0.4567825
sample estimates:
mean of the differences
              -3.675831

effsize::cohen.d(dat$memory_ctrl, dat$memory_trt, paired = T)


Cohen's d

d estimate: -0.5370066 (medium)
95 percent confidence interval:
      lower       upper
-1.02706202 -0.04695123

pwr::pwr.t.test(n = 25, d = -.537, sig.level = 0.05, power = ,
                type = c("paired"))


     Paired t test power calculation

              n = 25
              d = 0.537
      sig.level = 0.05
          power = 0.7311741
    alternative = two.sided

NOTE: n is number of *pairs*
```

A3. Based on both the formula and R function, we reject the null with p = .026, and conclude that there is evidence for our groups being significantly different. We reach a medium effect size .537 and a post hoc power = .731. This is a medium effect size that we aim for, however the power does fall short of the .80 we aim for. A difference between this analysis and that from question 1 is that this power is greater than was attained in question 1, but a similarity is that it still does not reach the .80 we aim for. However, given that we already archive significance with this number of participants in the within sample t-test, we may not need to increase our sample size.

4. Finally, imagine these data come from two groups of participants but the data are ordinal. How does this change your analysis strategy? Compare the groups' performance using R and discuss statistical significance. (2 points)

```
wilcox.test(dat$memory_ctrl, dat$memory_trt, correct=F, paired=F)


    Wilcoxon rank sum exact test

data:  dat$memory_ctrl and dat$memory_trt
W = 213, p-value = 0.05422
alternative hypothesis: true location shift is not equal to 0
```

A4. If our data was ordinal we would want to complete the Wilcox Rank Sum Test, because to complete a t-test, our data needs to be at least interval. After Wilcox rank sum test we see that our p-value = .0542, which is marginally significant, however we fail to reject the null. But given our marginal significance, a post hoc power test would be useful to probe our analysis.