PSY.308b.DA2

Instructions

You are conducting a slightly different replication of the classic Darley and Batson (1973) social psychology experiment on helping behavior. Your version of this study is going to look at Princeton seminary students' likelihood of helping a person in distress depending on their level of haste (whether they were running early, on time, or late for an appointment) when they encountered the person.

DV = HELPING (ranges from 0 to 6, with higher scores indicating greater helping) IV = HASTE (1 = early; 2 = on time, 3 = late)

NOTE: This is possibly problematic as it cannot be determined if this scale is ordinal or interval in nature. Variability in an ordinal scale may be artificial and not measurable if the quantifier for the interval is not known.

NOTE: R will read the numbers 1, 2, 3 in the HASTE column as numeric values but you want them to be treated as a "factor" (that is, as your categorical IV). Therefore, after you read the data into a dataframe (in this example, called dat), convert the HASTE variable from a numeric variable to a factor variable using the "as.factor" function.

```
#prep

#load data
library(psych)
library(jmv)

##
## Attaching package: 'jmv'

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

## The following object is masked from 'package:stats':
##
##     anova

dat <- read.csv("https://www.dropbox.com/s/zrlmw43dz54v212/PSY.308b.DA2.csv?dl=1")

#View(dat) Row x Col
head(dat)

##   pnum HASTE HELPING
## 1    1     1    3.38
## 2    2     1    3.58
## 3    3     1    3.70
```

```
## 4    4    1    3.65
## 5    5    1    1.84
## 6    6    1    4.66
```

tail(dat)

```
##    pnum HASTE HELPING
## 37  37    3    1.00
## 38  38    3    0.06
## 39  39    3    2.66
## 40  40    3    0.45
## 41  41    3    1.00
## 42  42    3    0.07
```

message ("running 'dim(dat)'")

```
## running 'dim(dat)'
```

dim(dat)

```
## [1] 42  3
```

#encode $HASTE column as a factor because groups are stupidly SPSS numeric and need to be read as factor and not number
message ("before changing $HASTE to factor by running 'class(dat$HASTE)'")

```
## before changing $HASTE to factor by running 'class(dat$HASTE)'
```

class(dat**$**HASTE)

```
## [1] "integer"
```

dat**$**HASTE <- **as.factor**(dat**$**HASTE)
message ("after changing $HASTE to factor by applying 'dat$HASTE <- as.factor(dat$HASTE)' and running 'class(dat$HASTE)'")

```
## after changing $HASTE to factor by applying 'dat$HASTE <- as.factor(dat$HASTE)' and running 'class(dat$HASTE)'
```

class(dat**$**HASTE)

```
## [1] "factor"
```

#Q2
#Descriptives for all data, including histogram - does not check assumptions, but helps for big picture of data
desc <- **descriptives**(dat, vars = **c**('HELPING'), hist = TRUE, sd = TRUE, se = TRUE, skew = TRUE, kurt = TRUE)
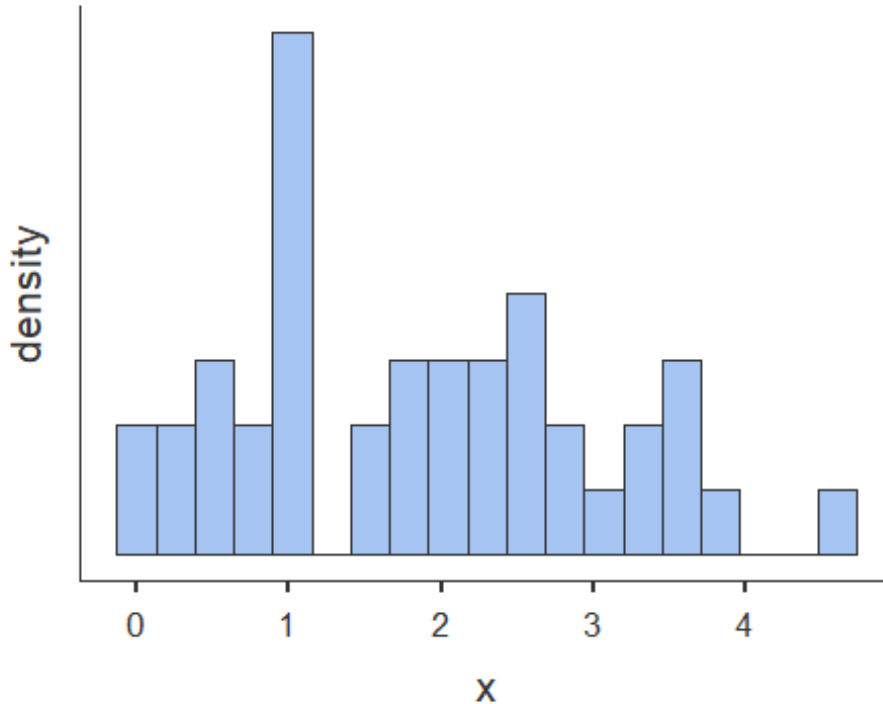desc

```
##
## DESCRIPTIVES
##
## Descriptives
```

```
## ----------------------------------
##                HELPING
## ----------------------------------
##   N                      42
##   Missing                 0
##   Mean                  1.87
##   Std. error mean        0.183
##   Median                1.86
##   Standard deviation     1.19
##   Minimum                0.0600
##   Maximum               4.66
##   Skewness              0.324
##   Std. error skewness    0.365
##   Kurtosis             -0.789
##   Std. error kurtosis    0.717
## ----------------------------------
```



```
#Descriptives by group, including grouped histograms
desc.group <- descriptives(dat, vars = c('HELPING'), splitBy = 'HASTE', hist = TRUE, sd = TRUE, se =
TRUE, skew = TRUE, kurt = TRUE)
desc.group

##
## DESCRIPTIVES
##
## Descriptives
## -------------------------------------------
```

```
##                   HASTE   HELPING
## ------------------------------------------
##   N                1         14
##                    2         14
##                    3         14
##   Missing          1          0
##                    2          0
##                    3          0
##   Mean             1         2.73
##                    2         1.83
##                    3         1.04
##   Std. error mean  1         0.302
##                    2         0.261
##                    3         0.217
##   Median           1         2.80
##                    2         2.04
##                    3         1.00
##   Standard deviation  1       1.13
##                    2         0.978
##                    3         0.813
##   Minimum          1         0.950
##                    2         0.360
##                    3         0.0600
##   Maximum          1         4.66
##                    2         2.94
##                    3         2.66
##   Skewness         1        -0.0860
##                    2        -0.351
##                    3         0.798
##   Std. error skewness  1      0.597
##                    2         0.597
##                    3         0.597
##   Kurtosis         1        -1.06
##                    2        -1.55
##                    3        -0.164
##   Std. error kurtosis  1      1.15
##                    2         1.15
##                    3         1.15
## ------------------------------------------
```
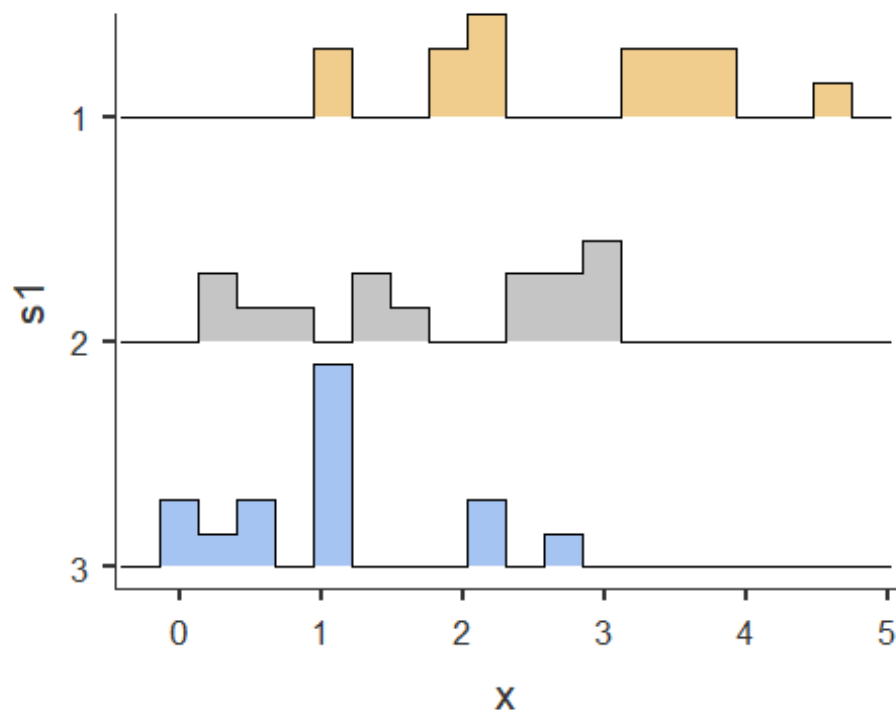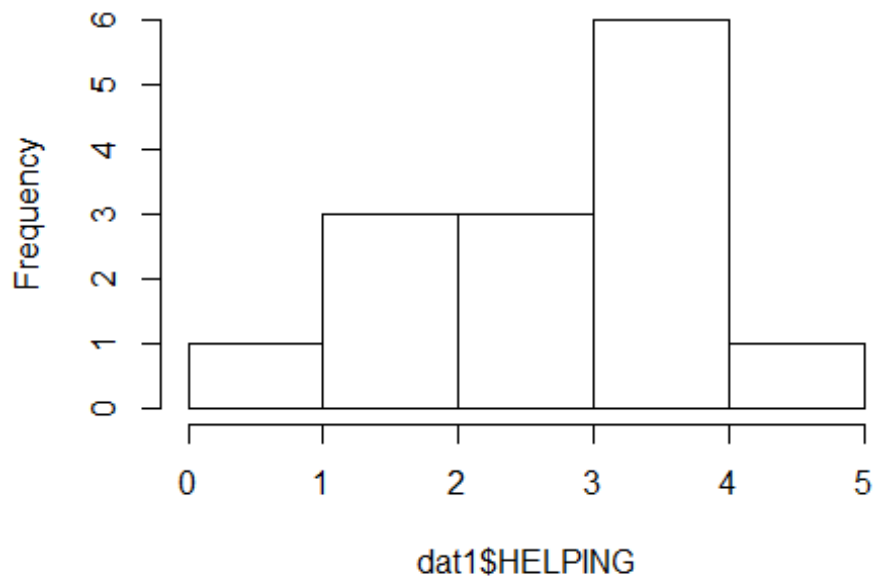
#Subset by group
```
dat1 <- subset(dat, dat$HASTE == "1")
dat2 <- subset(dat, dat$HASTE == "2")
dat3 <- subset(dat, dat$HASTE == "3")

#Individual histograms for each subsetted group
hist(dat1$HELPING)
```
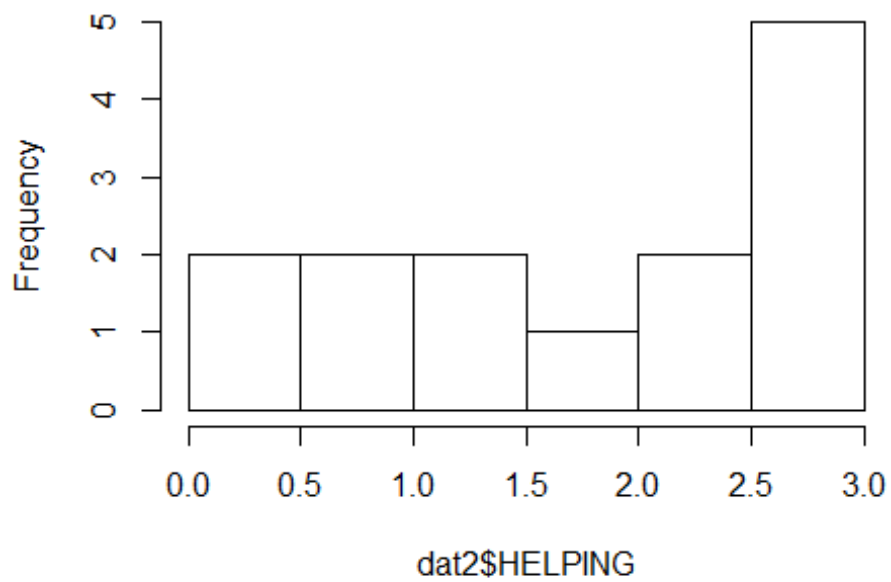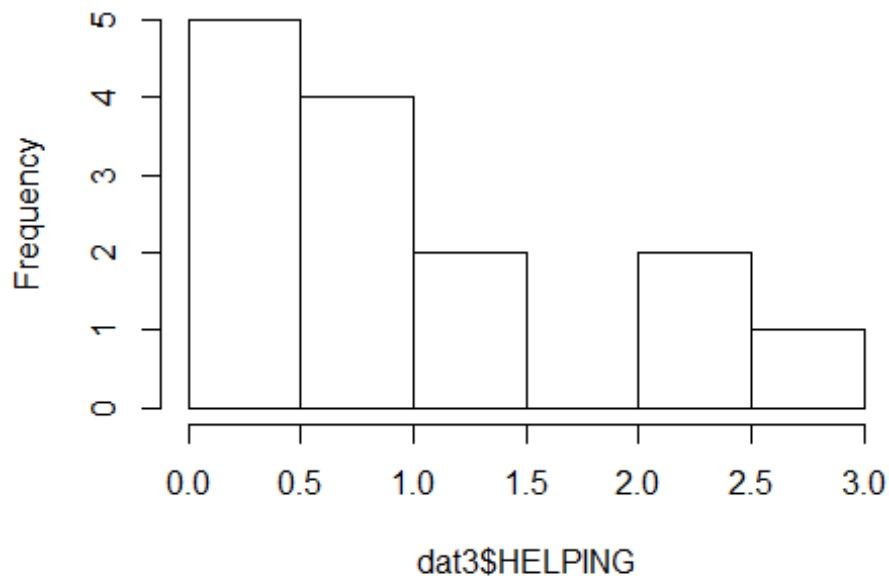
Histogram of dat1$HELPING


```
hist(dat2$HELPING)
```

Histogram of dat2$HELPING


```
hist(dat3$HELPING)
```

Histogram of dat3$HELPING


```
#anova with tukey post-hoc test includes Levene's test for homogeneity
anova(data = dat, dep = 'HELPING', factors = c('HASTE'), effectSize = 'partEta', postHoc = c('HASTE'),
postHocCorr = 'tukey', homo = TRUE)
```

```
##
## ANOVA
##
## ANOVA
## ------------------------------------------------------------------------------
##             Sum of Squares   df   Mean Square   F      p        <U+03B7>²p
## ------------------------------------------------------------------------------
##   HASTE              20.2    2      10.118     10.5   < .001   0.350
##   Residuals          37.6   39       0.965
## ------------------------------------------------------------------------------
##
##
## ASSUMPTION CHECKS
##
## Test for Homogeneity of Variances (Levene's)
## --------------------------------------------
##   F     df1   df2   p
## --------------------------------------------
##   2.71   2     39   0.079
## --------------------------------------------
##
##
## POST HOC TESTS
```

```
##
## Post Hoc Comparisons - HASTE
## ------------------------------------------------------------------------------
##   HASTE      HASTE   Mean Difference   SE     df     t     p-tukey
## ------------------------------------------------------------------------------
##   1      -   2           0.900   0.371   39.0   2.42     0.051
##          -   3           1.699   0.371   39.0   4.58   < .001
##   2      -   3           0.799   0.371   39.0   2.15     0.092
## ------------------------------------------------------------------------------
```

#Q3
#anova with tukey post-hoc test
anova(data = dat, dep = 'HELPING', factors = **c**('HASTE'), effectSize = 'partEta', postHoc = **c**('HASTE'),
postHocCorr = 'tukey', homo = TRUE)

```
##
## ANOVA
##
## ANOVA
## ------------------------------------------------------------------------------
##             Sum of Squares   df   Mean Square   F      p       <U+03B7>²p
## ------------------------------------------------------------------------------
##   HASTE          20.2    2        10.118   10.5   < .001   0.350
##   Residuals      37.6   39         0.965
## ------------------------------------------------------------------------------
##
##
## ASSUMPTION CHECKS
##
## Test for Homogeneity of Variances (Levene's)
## ---------------------------------------------
##   F      df1   df2   p
## ---------------------------------------------
##   2.71    2    39   0.079
## ---------------------------------------------
##
##
## POST HOC TESTS
##
## Post Hoc Comparisons - HASTE
## ------------------------------------------------------------------------------
##   HASTE      HASTE   Mean Difference   SE     df     t     p-tukey
## ------------------------------------------------------------------------------
##   1      -   2           0.900   0.371   39.0   2.42     0.051
##          -   3           1.699   0.371   39.0   4.58   < .001
##   2      -   3           0.799   0.371   39.0   2.15     0.092
## ------------------------------------------------------------------------------
```

#Q6

#Subset for only categories we want for independent samples t-test

```
newdata <- subset(dat, HASTE=="1" | HASTE == "2", select=pnum:HELPING)

#Independent samples t-test
ttestIS(data = newdata, vars = 'HELPING', group = 'HASTE', eqv = TRUE, effectSize = TRUE, ci =
TRUE, desc = TRUE)

##
## INDEPENDENT SAMPLES T-TEST
##
## Independent Samples T-Test
## -------------------------------------------------------------------------------
##                     statistic   df    p      Lower   Upper   Cohen's d
## -------------------------------------------------------------------------------
##   HELPING   Student's t    2.25   26.0   0.033   0.0790   1.72      0.852
## -------------------------------------------------------------------------------
##
##
## ASSUMPTIONS
##
## Test of Equality of Variances (Levene's)
## ----------------------------------------
##           F      df   p
## ----------------------------------------
##   HELPING   0.597    1   0.447
## ----------------------------------------
##   Note. A low p-value suggests a
##   violation of the assumption of
##   equal variances
##
##
## Group Descriptives
## -----------------------------------------------------------
##          Group   N    Mean   Median   SD     SE
## -----------------------------------------------------------
##   HELPING   1      14   2.73    2.80    1.13   0.302
##             2      14   1.83    2.04    0.978  0.261
## -----------------------------------------------------------
```

Questions

1. You are going to be looking for differences between three different groups on one outcome measure of helping. Which analyses would you use for this?

one-way ANOVA with relevant post-hoc analysis if H_0 is rejected.

2. Test the necessary assumptions for this test. Report whether each assumption is violated or not and present the evidence as to why.

Assumptions:

1. Dependent Variable HELPING is continuous - **TRUE**
2. DV is normally distributed for each group on HASTE - **TRUE** 🔻
   - Skew should be between -3 and 3. skew = -.09 for group 1. skew = -.35 for group 2. skew = .80 for group 3
   - kurtosis should be between -10 and 10. kurtosis = -1.06 for group 1. kurtosis = -1.55 for group 2. kurtosis = .16 for group 3.
3. Independence of Observations - participants are not present for both groups - **TRUE**
4. Homogeneity of Variance - **TRUE**
   - Levene's and check the p-value (less than .05 is a violation). $F(2, 39) = .079$, $p = .079$.

3. Is there a significant difference somewhere between these three groups on the outcome measure? Also, test if there is a significant difference between each group (1 & 2, 1 & 3, 2 & 3). Report all the relevant statistics according to APA format (Hint: don't forget to report how big this effect is when applicable).

NOTES: $df_{between}$ = k-1 = **2** df_{within} = k(n-1) = **39**; where k = number of groups, n = number per group; n = 42, k = 3.

NOTES: DV = HELPING (ranges from 0 to 6, with higher scores indicating greater helping); 🔻 IV = HASTE (1 = early; 2 = on time, 3 = late)

A one-way between-subjects ANOVA was used to compare helping behaviors between levels of haste. There was a statistically significant effect for helping behaviors on levels of haste, $F(2, 39) = 10.5$, $p< .001$, $\eta^2 = .35$. There was a medium effect, with levels of haste accounting for 35% of the variance in helping behaviors.

A Tukey HSD post-hoc test was used for pairwise comparisons of levels of haste. Students who were early had significantly higher helping behaviors than students who were late, *M1 - M3* = 1.70, *p<*.001. Students who were early did not have significantly higher helping behavior than students who were on time, *M1 - M2* = .90, *p* = .051. Students who were on time did not have significantly higher helping behavior than students who were late, *M2 - M3* = .80, *p* = .092. Overall, students who were early had the highest level of helping behaviors (*M1* = 2.73, *SD1* = 1.13) and students who were late had the lowest level of helping behaviors (*M3* = 1.04, *SD3* = .81).

4. Please interpret the results of your analyses from #3 for a journal to which you are submitting.

To test our hypotheses regarding the effects of levels of haste on helping behaviors, we conducted a one-way ANOVAs to compare levels of haste across three levels (early, on time, late). There was a significant effect for helping behaviors on levels of haste, $F(2, 39) = 10.5$, $p< .001$, $\eta^2 = .35$. Overall, Tukey HSD Post-hoc pair-wise comparisons indicated that students who were early (*M1* = 2.73, *SD1* = 1.13) had significantly higher helping behaviors than students who were late (*M3* = 1.04, *SD3* = .81), *M1 - M3* = 1.70, *p<*.001. 🔻

5. Now interpret the results of your analyses for an awareness campaign trying to increase the amount strangers help each other (they do not understand statistics that well). What would you conclude?

Based on our tests, running late affects the helping behaviors of students. Students who are on time have the highest amount of helping behaviors whereas students who are late have the lowest number of helping behaviors. Therefore, it seems as though student behaviors that support being early to class should be reinforced in order to maximize helping behaviors.

6. Conduct a t-test between group 1 and group 2 on the outcome measure. Is there a significant difference between group 1 and group 2 according to your t-test? Report all relevant statistics according to APA format (Hint: don't forget to report how big this effect is).

NOTES: df = $n_1 + n_2$ -2

The mean of Group 1 for helping behaviors (M = 2.73) was significantly higher than the mean of Group 2 for helping behaviors (M = 1.83), $t(26)$ = 2.25, p = .033, d = .85.

7. Discuss how your post-hoc analyses in #3 and the corresponding pairwise comparison in #6 differ. Why would you do the analyses in #3 instead of the pairwise comparison in #6 for the overall problem we started with?

The Tukey's post-hoc analysis of Groups 1 and 2 differs quantitatively from the independent samples t-test such that the latter p = .033 which is significant and the former p = .051 which retains the null hypothesis.

The analysis in #3 compares the variance across all levels of the independent variable and then determines if there is a significant difference in variance across all levels of the independent variable. A Tukey's post-hoc analysis determines differences between all combinations of pairwise comparisons while adjusting for α, whereas the independent samples t-test compares the means between samples 1 and 2 independent of and disregarding the variance and subsequent α error corresponsing with comparisons with the third level of the independent variable. Disregarding a level of the independent variable posthoc is bad science and p-hacking.

8. Look at the effect size you used for #6 and the effect size you used for #3. Please briefly describe why you used different effect sizes here and what they each are telling us.

The effect size for question #6 (d = .85) indicated the effect size of only the one pairwise comparison between Groups 1 and 2, whereas the effect sizes for question #3 (η^2 = .35) included the total effect size across all levels of the independent variable. The effect size for a t-test is denoted by d, which indicates the strength of the effect size on the difference

between means of groups, whereas η^2 indicates how levels of the independent variable account for a percentage of the variance in scores of the dependent variable.

9. What is the difference between Tukey's and Bonferroni's procedure? Why would you choose one over the other?

Tukey's procedure adjusts for α error based on pairwise comparisons whereas Bonferroni's procedure asjusts for all possible combinations of comparisons. You would choose Bonferroni's if you want to reduce the chance of having one Type I error to the most extreme conservative level. Otherwise, Tukey's correction may be the best choice.

10. What is "Sum of Squares" and "Mean Squares"? How are they used?

Sum of Squares (SS) is the sum of the squared mean differences, which in conjunction with degrees of freedom is used to calculate Mean Squares (SS/df = MS).

Mean Squares (MS) is variance explained by the sum of squares (SS) divided by degrees of freedom (MS/df). The Mean Squares across groups (MS_A) divided by the Mean Squares of subjects within groups ($MS_{S/A}$) provides us with the F statistic used in ANOVA analysis (F = MS_A / $MS_{S/A}$).