exrcise 2

2023-03-29

Q1

```
  library(haven); library(psych); library(dplyr);
  library(magrittr); library(ggplot2); library(gridExtra)
  library(rstatix); library(multcomp); library(ggeffects)

# Read the data

  extwo = read.csv("Ex2_data.csv")

# Data manipulation
  extwo =
    extwo %>%
    mutate(hasevpsta = ifelse(hasevpst == 99, NA, hasevpst),
           hafrqpsta = ifelse(hafrqpst == 99, NA, hafrqpst),
           mispst = ifelse(hasevpst == 99 | hafrqpst == 99, 1, 0))

 # Convert all 99s to NAs
  extwo =
    extwo %>%
    mutate_all(~replace(., . == 99, NA))


# Frequency tables
  xtabs(~ mispst + male, data = extwo)
```

```
##       male
## mispst  0  1
##      0 60 22
##      1  4  4
```

```
  xtabs(~ mispst + newrx, data = extwo)
```

```
##       newrx
## mispst  0  1
##      0 40 42
##      1  5  3
```

```
  xtabs(~ mispst + topir, data = extwo)
```

```
##       topir
## mispst  0  1
##      0 43 39
##      1  3  5
```

```
  xtabs(~ mispst + raceth, data = extwo)
```

```
##       raceth
## mispst  0  1  2
##      0 40 24 18
##      1  2  4  2
```

```
  xtabs(~ mispst + hasevpre, data = extwo)
```

```
##       hasevpre
## mispst  2  3  4  5  6  7  8  9 10
##      0  1  4 18 16 21  8 10  2  2
##      1  0  1  2  2  0  1  0  1  1
```

```
# Means by group
  extwo %>%
    group_by(mispst) %>%
    summarise(age.mean = mean(age, na.rm = TRUE),
              hafrqpre.mean = mean(hafrqpre, na.rm = TRUE),
              hit6pre.mean = mean(hit6pre, na.rm = TRUE))
```

```
## # A tibble: 2 x 4
##   mispst age.mean hafrqpre.mean hit6pre.mean
##    <dbl>    <dbl>         <dbl>        <dbl>
## 1      0     36.2          21.2         61.7
## 2      1     37.7          22.4         63.8
```

Q2

```
  library(tableone)

  CreateTableOne(vars = c("hafrqpre", "hasevpre", "hit6pre",
                          "raceth", "male", "age", "topir"),
                 strata = c("newrx"), data = extwo,
                           factorVars = c("raceth", "male", "topir"))
```

```
##                         Stratified by newrx
##                          0             1             p      test
##   n                         45            45
##   hafrqpre (mean (SD)) 21.27 (7.96)  21.36 (7.74)   0.957
##   hasevpre (mean (SD))  5.78 (1.81)   5.62 (1.72)   0.677
##   hit6pre (mean (SD))  62.22 (4.01)  61.60 (5.34)   0.534
##   raceth (%)                                       0.149
```

```
##      0                    25 (55.6)     17 (37.8)
##      1                    10 (22.2)     18 (40.0)
##      2                    10 (22.2)     10 (22.2)
##   male = 1 (%)            15 (33.3)     11 (24.4)   0.485
##   age (mean (SD))      36.10 (8.90)  36.54 (9.33)   0.820
##   topir = 1 (%)           24 (53.3)     20 (44.4)   0.527
```

```
  cat("or alternatively...")
```

```
## or alternatively...
```

```
  bothdrugs =
    extwo %>%
    mutate(both =
             case_when(newrx == 1 & topir == 1 ~ 3,
                       newrx == 0 & topir == 1 ~ 2,
                       newrx == 1 & topir == 0 ~ 1, TRUE ~ 0),
           sevpre =
             case_when(hasevpre %in% 8:10 ~ 1,
                       hasevpre < 8 ~ 0))

  CreateTableOne(vars = c("hafrqpre", "hasevpre", "hit6pre",
                          "raceth", "male", "age", "topir"),
                 strata = c("both"), data = bothdrugs,
                           factorVars = c("raceth", "male", "topir"))
```

```
##                        Stratified by both
##                         0             1             2
##   n                        21            25            24
##   hafrqpre (mean (SD)) 22.19 (8.15)  22.16 (8.16)  20.46 (7.87)
##   hasevpre (mean (SD))  5.86 (1.77)   5.76 (2.03)   5.71 (1.88)
##   hit6pre (mean (SD))  61.71 (3.96)  61.60 (5.98)  62.67 (4.07)
##   raceth (%)
##      0                    12 (57.1)     10 (40.0)     13 ( 54.2)
##      1                     3 (14.3)     12 (48.0)      7 ( 29.2)
##      2                     6 (28.6)      3 (12.0)      4 ( 16.7)
##   male = 1 (%)            10 (47.6)      8 (32.0)      5 ( 20.8)
##   age (mean (SD))      33.70 (9.82)  36.74 (9.70)  38.20 (7.61)
##   topir = 1 (%)            0 ( 0.0)      0 ( 0.0)     24 (100.0)
##                        Stratified by both
##                         3              p      test
##   n                        20
##   hafrqpre (mean (SD)) 20.35 (7.26)    0.771
##   hasevpre (mean (SD))  5.45 (1.28)    0.899
##   hit6pre (mean (SD))  61.60 (4.58)    0.842
##   raceth (%)                           0.162
##      0                     7 ( 35.0)
##      1                     6 ( 30.0)
##      2                     7 ( 35.0)
##   male = 1 (%)             3 ( 15.0)   0.096
##   age (mean (SD))      36.28 (9.09)    0.424
##   topir = 1 (%)            20 (100.0)  <0.001
```

Q3

```
 corr.test(extwo %>% dplyr::select("hit6pre", "hit6pst"))
```

```
## Call:corr.test(x = extwo %>% dplyr::select("hit6pre", "hit6pst"))
## Correlation matrix
##         hit6pre hit6pst
## hit6pre    1.00    0.26
## hit6pst    0.26    1.00
## Sample Size
## [1] 90
## Probability values (Entries above the diagonal are adjusted for multiple tests.)
##         hit6pre hit6pst
## hit6pre    0.00    0.01
## hit6pst    0.01    0.00
##
##  To see confidence intervals of the correlations, print with the short=FALSE option
```

Q4

```
# Paired t-test for new drug group
 t.test(Pair(hafrqpre, hafrqpst) ~ 1, data = extwo %>%
          filter(newrx == 1))
```

```
##
##  Paired t-test
##
## data:  Pair(hafrqpre, hafrqpst)
## t = 0.26655, df = 43, p-value = 0.7911
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##  -1.790711  2.336165
## sample estimates:
## mean difference
##       0.2727273
```

```
# Between group t-test
 t.test(hafrqpre ~ newrx, data = extwo)
```

```
##
##  Welch Two Sample t-test
##
## data:  hafrqpre by newrx
## t = -0.05371, df = 87.932, p-value = 0.9573
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -3.377841  3.200063
## sample estimates:
## mean in group 0 mean in group 1
##        21.26667        21.35556
```

```r
# Descriptive stats by group
 CreateContTable(vars = c("hafrqpre", "hafrqpst"),
                 strata = "newrx", data = extwo)
```

```
##                         Stratified by newrx
##                          0            1            p      test
##   n                      45           45
##   hafrqpre (mean (SD)) 21.27 (7.96) 21.36 (7.74)  0.957
##   hafrqpst (mean (SD)) 22.02 (7.36) 21.11 (8.03)  0.586
```

Q5

```r
# Between group t-test
 t.test(hit6pst ~ newrx, data = extwo)
```

```
##
##  Welch Two Sample t-test
##
## data:  hit6pst by newrx
## t = 4.9034, df = 86.419, p-value = 4.374e-06
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  4.003723 9.462944
## sample estimates:
## mean in group 0 mean in group 1
##        59.24444        52.51111
```

```r
# LM
 fit = lm(hit6pst ~ newrx, data = extwo)
 confint(fit)
```

```
##                 2.5 %    97.5 %
## (Intercept) 57.314810 61.174078
## newrx       -9.462248 -4.004419
```

Q6

```r
# Fit lm model with a interaction term
 fit = lm(I(hit6pst - hit6pre) ~ topir * newrx, data = extwo)
 summary(fit)
```

```
##
## Call:
## lm(formula = I(hit6pst - hit6pre) ~ topir * newrx, data = extwo)
##
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -20.9600  -3.8500   0.1667   4.0400  19.0400
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -3.905      1.530  -2.552   0.0125 *
## topir          1.738      2.095   0.830   0.4091
## newrx         -5.135      2.076  -2.474   0.0153 *
## topir:newrx   -1.848      2.969  -0.622   0.5353
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.012 on 86 degrees of freedom
## Multiple R-squared:  0.1713, Adjusted R-squared:  0.1424
## F-statistic: 5.927 on 3 and 86 DF,  p-value: 0.001006
```

```
  confint(fit) # confidence interaval
```

```
##                 2.5 %     97.5 %
## (Intercept) -6.946656 -0.8628674
## topir       -2.427190  5.9033809
## newrx       -9.261465 -1.0090116
## topir:newrx -7.750468  4.0542776
```

```
 # This works too!
  library(apaTables)
  apa.reg.table(fit)
```

```
##
##
## Regression results using I(hit6pst - hit6pre) as the criterion
##
##
##    Predictor      b       b_95%_CI sr2  sr2_95%_CI             Fit
##  (Intercept) -3.90* [-6.95, -0.86]
##        topir   1.74  [-2.43, 5.90] .01 [-.02, .04]
##        newrx -5.14* [-9.26, -1.01] .06 [-.03, .15]
##  topir:newrx  -1.85  [-7.75, 4.05] .00 [-.02, .03]
##                                                      R2 = .171**
##                                                  95% CI[.03,.29]
##
##
## Note. A significant b-weight indicates the semi-partial correlation is also significant.
## b represents unstandardized regression weights.
## sr2 represents the semi-partial correlation squared.
## Square brackets are used to enclose the lower and upper limits of a confidence interval.
## * indicates p < .05. ** indicates p < .01.
##
```

Q7

```
# Fit ANCOVA model
 fit = lm(hit6pst ~ hit6pre + topir + newrx, data = extwo)
 summary(fit)
```

```
##
## Call:
## lm(formula = hit6pst ~ hit6pre + topir + newrx, data = extwo)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -15.0279  -3.5722   0.6953   3.4003  14.6903
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  36.3008     8.9415   4.060 0.000108 ***
## hit6pre       0.3591     0.1430   2.511 0.013918 *
## topir         1.1241     1.3414   0.838 0.404344
## newrx        -6.4100     1.3419  -4.777 7.26e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.328 on 86 degrees of freedom
## Multiple R-squared:  0.2756, Adjusted R-squared:  0.2503
## F-statistic:  10.9 on 3 and 86 DF,  p-value: 3.848e-06
```

```
  confint(fit) # confidence interaval
```

```
##                   2.5 %     97.5 %
## (Intercept) 18.52562283 54.0759921
## hit6pre      0.07477666  0.6434265
## topir       -1.54252601  3.7907954
## newrx       -9.07760274 -3.7423358
```

```
# Calculate adjusted means
 library(emmeans)
 lsmeans =
   emmeans(fit, ~ factor(newrx))
```