Correlation Matrices, Reliability Tests, and t-Tests

Contents

```
library(mosaic)
library(ggplot2)
library(dplyr)

#install.packages("psych")
library(psych)
```

Data inside R packages

To demonstrate the use of the `alpha()` fuction in the `psych` package from computing Cronbach's alpha, we will need a data set with scale items. Many packages have example datasets pre-loaded. The `psych` package has a dataset called `bfi` that has n = 2800 participants responding to 25 personality itmes for the Big Five Inventory.

```
#We can see a list of ALL the datasets current avaliable
data()
```

If we see a dataset we like we can load it into our environment

```
data(bfi)
#?bfi
glimpse(bfi)
```

```
## Observations: 2,800
## Variables: 28
## $ A1        <int> 2, 2, 5, 4, 2, 6, 2, 4, 4, 2, 4, 2, 5, 5, 4, 4, 4, 5...
```

```
## $ A2        <int> 4, 4, 4, 4, 3, 6, 5, 3, 3, 5, 4, 5, 5, 5, 5, 3, 6, 5...
## $ A3        <int> 3, 5, 5, 6, 3, 5, 5, 1, 6, 6, 5, 5, 5, 5, 2, 6, 6, 5...
## $ A4        <int> 4, 2, 4, 5, 4, 6, 3, 5, 3, 6, 6, 5, 6, 6, 2, 6, 2, 4...
## $ A5        <int> 4, 5, 4, 5, 5, 5, 5, 1, 3, 5, 5, 5, 4, 6, 1, 3, 5, 5...
## $ C1        <int> 2, 5, 4, 4, 4, 6, 5, 3, 6, 6, 4, 5, 5, 4, 5, 5, 4, 5...
## $ C2        <int> 3, 4, 5, 4, 4, 6, 4, 2, 6, 5, 3, 4, 4, 4, 5, 5, 4, 5...
## $ C3        <int> 3, 4, 4, 3, 5, 6, 4, 4, 3, 6, 5, 5, 3, 4, 5, 5, 4, 5...
## $ C4        <int> 4, 3, 2, 5, 3, 1, 2, 2, 4, 2, 3, 4, 2, 2, 2, 3, 4, 4...
## $ C5        <int> 4, 4, 5, 5, 2, 3, 3, 4, 5, 1, 2, 5, 2, 1, 2, 5, 4, 3...
## $ E1        <int> 3, 1, 2, 5, 2, 2, 4, 3, 5, 2, 1, 3, 3, 2, 3, 1, 1, 2...
## $ E2        <int> 3, 1, 4, 3, 2, 1, 3, 6, 3, 2, 3, 3, 3, 2, 4, 1, 2, 2...
## $ E3        <int> 3, 6, 4, 4, 5, 6, 4, 4, NA, 4, 2, 4, 3, 4, 3, 6, 5, ...
## $ E4        <int> 4, 4, 4, 4, 4, 5, 5, 2, 4, 5, 5, 5, 2, 6, 6, 6, 5, 6...
## $ E5        <int> 4, 3, 5, 4, 5, 6, 5, 1, 3, 5, 4, 4, 4, 5, 5, 4, 5, 6...
## $ N1        <int> 3, 3, 4, 2, 2, 3, 1, 6, 5, 5, 3, 4, 1, 1, 2, 4, 4, 6...
## $ N2        <int> 4, 3, 5, 5, 3, 5, 2, 3, 5, 5, 3, 5, 2, 1, 4, 5, 4, 5...
## $ N3        <int> 2, 3, 4, 2, 4, 2, 2, 2, 2, 5, 4, 3, 2, 1, 2, 4, 4, 5...
## $ N4        <int> 2, 5, 2, 4, 4, 2, 1, 6, 3, 2, 2, 2, 2, 2, 2, 5, 4, 4...
## $ N5        <int> 3, 5, 3, 1, 3, 3, 1, 4, 3, 4, 3, NA, 2, 1, 3, 5, 5, ...
## $ O1        <int> 3, 4, 4, 3, 3, 4, 5, 3, 6, 5, 5, 4, 4, 5, 5, 6, 5, 5...
## $ O2        <int> 6, 2, 2, 3, 3, 3, 2, 2, 6, 1, 3, 6, 2, 3, 2, 6, 1, 1...
## $ O3        <int> 3, 4, 5, 4, 4, 5, 5, 4, 6, 5, 5, 4, 4, 4, 5, 6, 5, 4...
## $ O4        <int> 4, 3, 5, 3, 3, 6, 6, 5, 6, 5, 6, 5, 5, 4, 5, 3, 6, 5...
## $ O5        <int> 3, 3, 2, 5, 3, 1, 1, 3, 1, 2, 3, 4, 2, 4, 5, 2, 3, 4...
## $ gender    <int> 1, 2, 2, 2, 1, 2, 1, 1, 1, 2, 1, 1, 2, 1, 1, 1, 2, 1...
## $ education <int> NA, NA, NA, NA, NA, 3, NA, 2, 1, NA, 1, NA, NA, NA, ...
## $ age       <int> 16, 18, 17, 17, 17, 21, 18, 19, 19, 17, 21, 16, 16, ...
```

Correlation Matrix

First, we can get descriptive statistics for individual personality scale items. Let's do the agreeableness scale first.

```r
#Get descriptive statistics with mosaic or dplyr here. Get them for Age too.

#Also, get the frequencies for gender and education
```

In addition to descriptive statistics, we would probably also like to get a correlation matrix for items on a scale. I like the function `corr.test()` in the `psych` package. Notice that we are using the `dplyr` `select()` function inside of `corr.test()`.

```r
corr.test(select(bfi, A1, A2, A3, A4, A5))
```

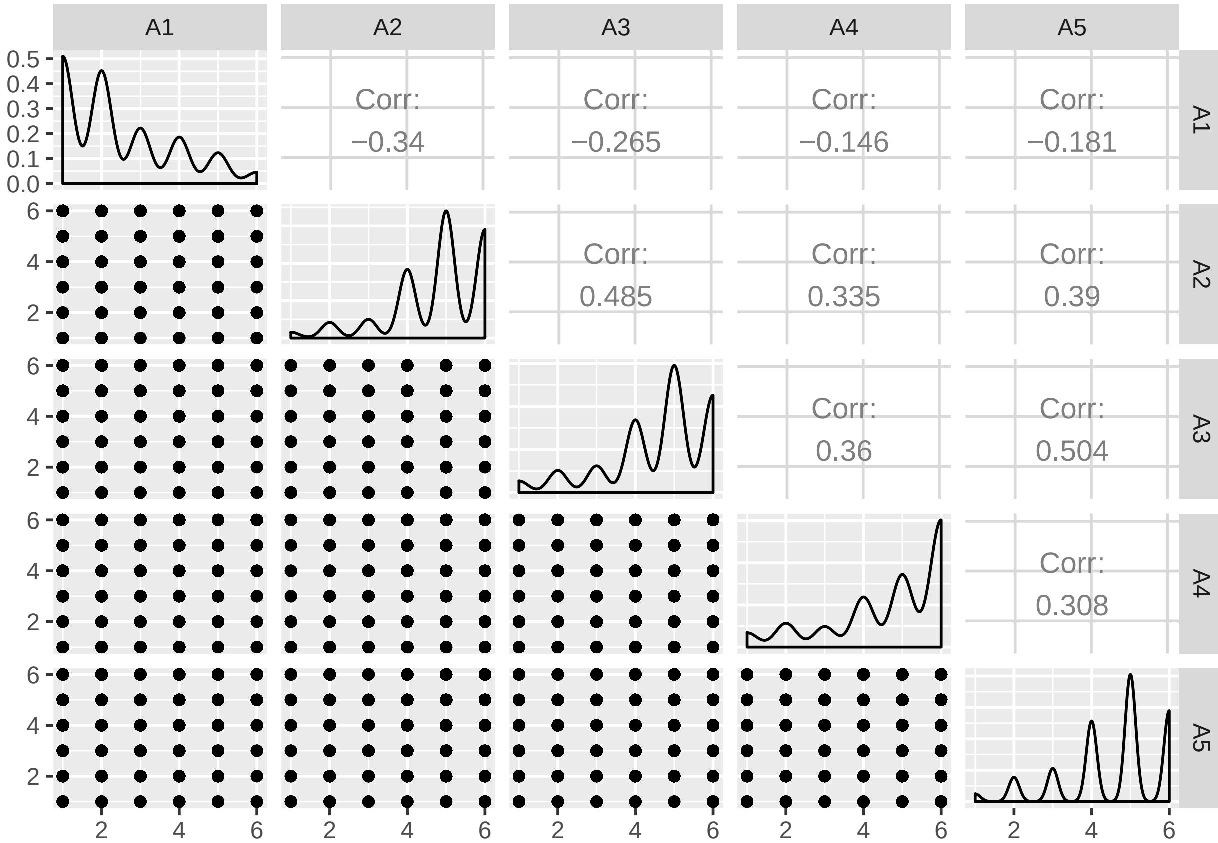
```
## Call:corr.test(x = select(bfi, A1, A2, A3, A4, A5))
## Correlation matrix
```

```
##       A1    A2    A3    A4    A5
## A1  1.00 -0.34 -0.27 -0.15 -0.18
## A2 -0.34  1.00  0.49  0.34  0.39
## A3 -0.27  0.49  1.00  0.36  0.50
## A4 -0.15  0.34  0.36  1.00  0.31
## A5 -0.18  0.39  0.50  0.31  1.00
## Sample Size
##      A1   A2   A3   A4   A5
## A1 2784 2757 2759 2767 2769
## A2 2757 2773 2751 2758 2757
## A3 2759 2751 2774 2759 2758
## A4 2767 2758 2759 2781 2765
## A5 2769 2757 2758 2765 2784
## Probability values (Entries above the diagonal are adjusted for multiple tests.)
##    A1 A2 A3 A4 A5
## A1  0  0  0  0  0
## A2  0  0  0  0  0
## A3  0  0  0  0  0
## A4  0  0  0  0  0
## A5  0  0  0  0  0
##
##  To see confidence intervals of the correlations, print with the short=FALSE option
```

There is also a `pairs()` fucntion in Base R and the `ggpairs()` function in the `GGally` package that is a ggplot correlation matrix/matrix scatterplot provider.

```r
#install.packages("GGally")
library(GGally)

ggpairs(select(bfi, A1, A2, A3, A4, A5))
```



It is clear that we will need to reverse score A1 to assess reliability and to create our agreeableness scale scores. Hint: to reverse score an item use this formula: `A1.r = (min(A1, na.rm = TRUE) + max(A1, na.rm = TRUE)) - A1`.

Now compute correlation matrices for the other 4 personality subscales and create any reverse scored items the need to be created.

```
#correlation matrices and reverse scored items
```

For more practice, find another dataset that sounds interesting and run some correlation matricies.

```
data()
```

Cronbach's alpha

Now that we have our reverse scored items and we have a sense of how the items relate, the next step is to test the scale reliability. We do this with Cronbach's alpha. There is a function called `alpha()` in the `psych` package.

```
alpha(select(bfi, A1.r, A2, A3, A4, A5))
```

```
##
```

```
## Reliability analysis
## Call: alpha(x = select(bfi, A1.r, A2, A3, A4, A5))
##
##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean  sd
##        0.7      0.71    0.68      0.33 2.5 0.009  4.7 0.9
##
##  lower alpha upper     95% confidence boundaries
## 0.69 0.7 0.72
##
##  Reliability if an item is dropped:
##      raw_alpha std.alpha G6(smc) average_r S/N alpha se
## A1.r      0.72      0.73    0.67      0.40 2.6   0.0087
## A2        0.62      0.63    0.58      0.29 1.7   0.0119
## A3        0.60      0.61    0.56      0.28 1.6   0.0124
## A4        0.69      0.69    0.65      0.36 2.3   0.0098
## A5        0.64      0.66    0.61      0.32 1.9   0.0111
##
##  Item statistics
##         n raw.r std.r r.cor r.drop mean  sd
## A1.r 2784  0.58  0.57  0.38   0.31  4.6 1.4
## A2   2773  0.73  0.75  0.67   0.56  4.8 1.2
## A3   2774  0.76  0.77  0.71   0.59  4.6 1.3
## A4   2781  0.65  0.63  0.47   0.39  4.7 1.5
## A5   2784  0.69  0.70  0.60   0.49  4.6 1.3
##
## Non missing response frequency for each item
##         1    2    3    4    5    6 miss
## A1.r 0.03 0.08 0.12 0.14 0.29 0.33 0.01
## A2   0.02 0.05 0.05 0.20 0.37 0.31 0.01
## A3   0.03 0.06 0.07 0.20 0.36 0.27 0.01
## A4   0.05 0.08 0.07 0.16 0.24 0.41 0.01
## A5   0.02 0.07 0.09 0.22 0.35 0.25 0.01
```

With an alphas of .70 we have a reliabile agreeableness scale. The next step would be to create agreeableness scale scores for participants, the average of these 5 items. We can do this with `mutate()`. Note that we cannot use `mean()` here because `mean()` is going to get the mean of a column (a variable). What we want it the mean across columns.

```
bfi <- bfi %>%
  mutate(agreeable = (A1.r + A2 + A3 + A4 + A5)/5)
```

Let's take a look at it with a visualization. Also try facetting by gender.

```
#make some histograms, boxplots, and density plots of agreeableness
```

Now calculate the reliabilities for each of the other 4 scales. Be sure to include the reverse scored items you created where relevant.

```r
#calculate reliabilities
```

Create scale scores and make visualizations of those scale scores.

```r
#scale scores and visualizations
```

t-Tests

Independent Samples t-Test

Next we can ask, are men and women different on agreeableness? To answer this question we need to run an independent samples t-test. We'll use the function `t.test()` in the `mosaic` package. Recall that `mosaic` first needs the formula, `y ~ x`, then the data, `data = dataName`.

```r
t.test(agreeable ~ gender, data = bfi)
```

```
##
##  Welch Two Sample t-test
##
## data:  agreeable by gender
## t = -10.725, df = 1654.5, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.4698058 -0.3245337
## sample estimates:
## mean in group 1 mean in group 2
##        4.377679        4.774848
```

There is a statistically significant difference between women and men on agreeableness, $t(1654.50) = \text{-}10.73$, $p < .001$, with women ($M = 4.77$, $SD = 0.86$) scoring higher than men ($M = 4.38$, $SD = 0.93$). It is possible to code in these numbers such that if the data were updated, the text would update as well.

```r
tmod <- t.test(agreeable ~ gender, data = bfi)
ds <- favstats(agreeable ~ gender, data = bfi)
```

There is a statistically significant difference between women and men on agreeableness, $t(1654.47) = \text{-}10.72$, $p < 0.001$, with women ($M = 4.77$, $SD = 0.86$) scoring higher than men ($M = 4.38$, $SD = 0.93$).

We can also ask for equal variance assumed.

```r
t.test(agreeable ~ gender, data = bfi, var.equal = TRUE)
```

```
##
```

```
##  Two Sample t-test
##
## data:  agreeable by gender
## t = -11.038, df = 2707, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.4677230 -0.3266165
## sample estimates:
## mean in group 1 mean in group 2
##        4.377679        4.774848
```

Test for gender differences for the other 4 scales. Test for gender differences between those who graduated college, and those who did not. Hint: use `mutate()` to create a dichotomous variable as we did for `wise_hus`.

```
#more t-Tests
```

Paired Samples t-Test

To demonstrate paired samples t-test I'd like to use yet another built in dataset. The `sat.act` dataset has information for 700 people on their SAT verbal, SAT quantitative, and ACT scores.

```
data(sat.act)
#?sat.act
glimpse(sat.act)
```

```
## Observations: 700
## Variables: 6
## $ gender    <int> 2, 2, 2, 1, 1, 1, 2, 1, 2, 2, 1, 2, 1, 2, 2, 2, 2, 2...
## $ education <int> 3, 3, 3, 4, 2, 5, 5, 3, 4, 5, 3, 4, 4, 4, 3, 4, 3, 4...
## $ age       <int> 19, 23, 20, 27, 33, 26, 30, 19, 23, 40, 23, 34, 32, ...
## $ ACT       <int> 24, 35, 21, 26, 31, 28, 36, 22, 22, 35, 32, 29, 21, ...
## $ SATV      <int> 500, 600, 480, 550, 600, 640, 610, 520, 400, 730, 76...
## $ SATQ      <int> 500, 500, 470, 520, 550, 640, 500, 560, 600, 800, 71...
```

Let's compare poeple's scores on the SAT verbal to their scores on the SAT quantitative using the `t.test()` function. This time instead of using the formula in the beginning, we add the two variables to be compared separated by a `,`. We also need to add `paired = TRUE`.

```
t.test(SATV, SATQ, data = sat.act, paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  SATV and SATQ
## t = 0.57483, df = 686, p-value = 0.5656
```

```
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -5.116109  9.351917
## sample estimates:
## mean of the differences
##                2.117904
```

There is no staistically significant difference between SAT verbal and SAT quantitative scores, $t(686) = 0.57$, $p = .566$. But perhaps there is a difference for boys only, or for girls only. We can create a dataset of just boys using `filter()` and then another one of just girls. This would be completely fine, but we could also pipe directly into the `t.test()` function. There is a catch: The pipe, `%>%`, puts the resulting dataframe into the first position in the destination function, but we want it in the third position. We can use the `.` symbol to pipe the dataframe into whichever position we'd like.

```
sat.act %>%
  filter(gender == 1) %>%
  t.test(SATV, SATQ, data = ., paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  SATV and SATQ
## t = -3.4934, df = 244, p-value = 0.0005661
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -32.08114  -8.94743
## sample estimates:
## mean of the differences
##               -20.51429
```

```
sat.act %>%
  filter(gender == 2) %>%
  t.test(SATV, SATQ, data = ., paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  SATV and SATQ
## t = 3.1813, df = 441, p-value = 0.00157
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   5.604249 23.721543
## sample estimates:
## mean of the differences
##                 14.6629
```

Explore the differences in `SATV` and `SATQ` for each level of the `education` variable.

```
#differences for each level of education
```