MATH1324 - Class Worksheet

Week 9 - Two-sample and paired samples t-tests

Working in small groups or pairs, complete the following exercises.

Datasets: `Reaction Time Practice.csv` and `Battery.csv` (see Data Repository)

Reaction Time Practice

Activity Description

Does reaction time improve with practice? Export the Reaction Time Practice.csv dataset collected in class (see Data Repository) and import into RStudio. Complete the following exercises.

1. Summarise the First and Second reaction time variables and comment on the trend.

2. Visualise the data using an appropriate plot (see Module 7 course notes for some ideas). Comment on the trend of the data.

3. Compute a difference variable, *d*, and save the variable to the `Reaction.Time.Practice` data object using the following code:

```
> Reaction.Time.Practice$d <-
Reaction.Time.Practice$Average.RT.Second-Reaction.Time.Practice$Average.RT.First
```

4. Perform an appropriate hypothesis test of *d* to determine if RT can improve with practice. Use the 0.05 level of significance.
   a. State the Null and Alternate hypothesis for the appropriate hypothesis test.

   b. Report the test statistic, *df*, *p*-value and 95% CI of the mean difference from the results of the hypothesis test.

   c. Use the *p*-value and 95% *CI* of the mean to make a decision about the Null hypothesis.

   d. Conclude whether or not the results of the hypothesis test were statistically significant and draw a conclusion within the context of the example.

Battery

Activity Description

Reanalyse the `Battery.csv` data used in the demonstration, but, this time, by comparing the mean number of pulses to reach 1.3 V between brands. Assume the data are normally distributed in the population.

5. Summarise the mean pulses to reach 1.30 V between brands.

6. Visualise the data using an appropriate plot. Comment on the difference between brands.

7. Test the assumption of equal variance using the Levene's test.

8. Perform an appropriate hypothesis test compare the mean pulses to reach 1.30 V between brands. Use the 0.05 level of significance.
    a. State the Null and Alternate hypothesis for the appropriate hypothesis test.

    b. Report the test statistic, *df*, *p*-value and 95% CI of the difference between means from the results of the hypothesis test.

    c. Use the *p*-value and 95% *CI* of the mean to make a decision about the Null hypothesis.

    d. Conclude whether or not the results of the hypothesis test were statistically significant and draw a conclusion within the context of the example.

R Hints

Use this example R code if you get stuck:

Reaction Time Practice

```
> favstats(~Average.RT.First, data = Reaction.Time.Practice) #RT first try
> favstats(~Average.RT.Second, data = Reaction.Time.Practice) #RT second try

> Reaction.Time.Practice$d <- Reaction.Time.Practice$Average.RT.Second
-Reaction.Time.Practice$Average.RT.First #Add column of differences

Reaction.Time.Practice_filtered <-subset(Reaction.Time.Practice,
                                         subset = Average.RT.Second < 800)


> favstats(~d, data = Reaction.Time.Practice) #Mean difference

> qqPlot(Reaction.Time.Practice$d, dist="norm") #Check normality if required

> t.test(Reaction.Time.Practice$d, mu = 0)

> install.packages("granova")
> library(granova)

> granova.ds(data.frame(Reaction.Time.Practice$Average.RT.First,
                        Reaction.Time.Practice$Average.RT.Second),
                        xlab = "Reaction Time - First",
                        ylab = "Reaction Time - After")
```

Battery

```
> Battery_sub <- subset(Battery, subset = (Voltage == 1.30)) # Filter 1.3 V

> favstats(~Pulses | Brand, data = Battery_sub) #Summarise data

> boxplot(Pulses ~ Brand, data = Battery_sub) #Visualise

> leveneTest(Pulses ~ Brand, data = Battery_sub) #Levene's test

> t.test(~Pulses | Brand, data = Battery_sub) #Two-sample t-test
```

Solutions and Answers

Reaction Time Practice

1. Summarise the First and Second reaction time variables and comment on the trend.

```
> favstats(~Average.RT.First, data = Reaction.Time.Practice) #RT first try
 min    Q1 median  Q3 max     mean       sd  n missing
 233 294.5    348 467 568 378.9259 103.7018 27       0
> favstats(~Average.RT.Second, data = Reaction.Time.Practice) #RT second try
 min  Q1 median  Q3 max     mean       sd  n missing
 230 292    334 394 823 363.1481 124.2175 27       0
```

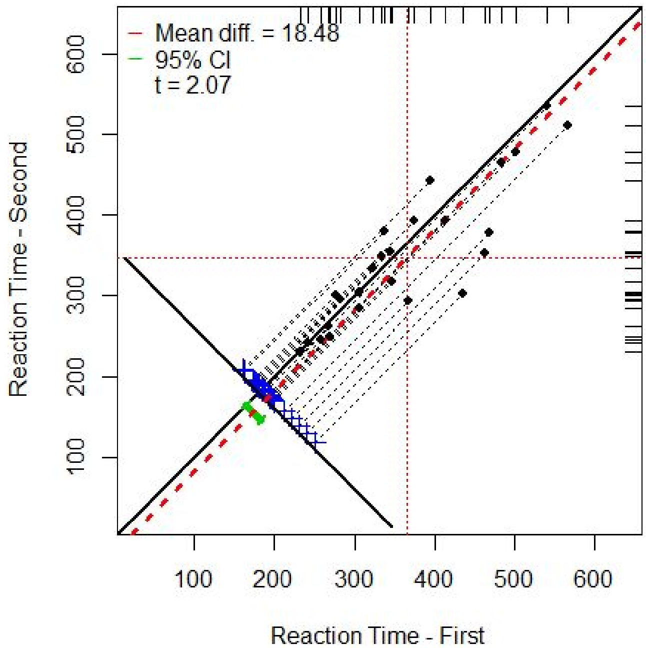
Note the outlier at 823.

2. Visualise the data using an appropriate plot (see Module 7 course notesfor some ideas). Comment on the trend of the data.

With outliers removed...

```
> granova.ds(data.frame(Reaction.Time.Practice_filtered$Average.RT.First,
            Reaction.Time.Practice_filtered$Average.RT.Second),
                   xlab = "Reaction Time - First",
                   ylab = "Reaction Time - Second")
```

Dependent Sample Assessment Plot



3. Compute a difference variable, *d*, and save the variable to the `Reaction.Time.Practice` data object using the following code:

```
> Reaction.Time.Practice$d <-
Reaction.Time.Practice$Average.RT.Second-Reaction.Time.Practice$Average.RT.Fi
rst

> favstats(~d, data = Reaction.Time.Practice) #Mean difference between second
and first
  min  Q1 median Q3 max      mean      sd  n missing
 -277 -28     -7 11 313 -15.77778 92.9327 27       0
```

Note the outlier at 313.

Filter data to remove outliers.

```
> Reaction.Time.Practice_filtered <-subset(Reaction.Time.Practice,
                                          subset = Average.RT.Second < 800 & d
> -200)
```

4. Perform an appropriate hypothesis test of *d* to determine if RT can improve with practice. Use the 0.05 level of significance.
   a. State the Null and Alternate hypothesis for the appropriate hypothesis test.

The statistical hypotheses for the paired samples t-test are as follows:

$$H_0: u_\Delta = 0$$
$$H_A: u_\Delta \neq 0$$

The paired samples t-test was used because the data was paired and we were analysing the average difference between first and second attempts.

b. Report the test statistic, *df*, *p*-value and 95% CI of the mean difference from the results of the hypothesis test.

```
> t.test(Reaction.Time.Practice_filtered$d, mu = 0, alternative =
"two.sided")

      One Sample t-test

data:  Reaction.Time.Practice_filtered$d
t = -2.0654, df = 24, p-value = 0.04985
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -36.94665285  -0.01334715
sample estimates:
mean of x
   -18.48
```

c. Use the *p*-value and 95% *CI* of the mean to make a decision about the Null hypothesis.

Assuming the data are normally distributed in the population (for which there might be doubts), we reject H_0 as the p-value < 0.05 and the 95% CI [-36.95, -0.01] does not capture H_0: $u\Delta = 0$.

d. Conclude whether or not the results of the hypothesis test were statistically significant and draw a conclusion within the context of the example.

The results of the paired samples t-test were statistically significant. If the normality assumption holds, the findings suggest that test takers significantly improve their reaction time ratings on their second attempt.

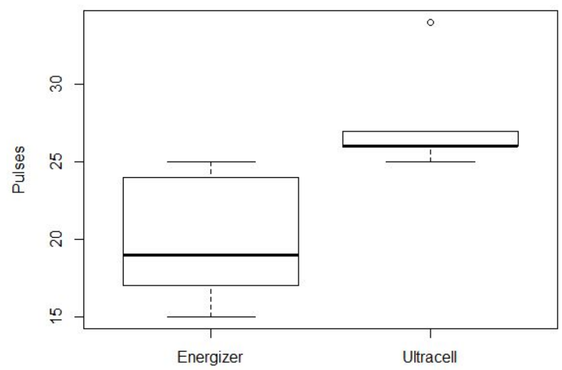
Battery

5. Summarise the mean pulses to reach 1.30 V between brands.

```
> Battery_sub <- subset(Battery, subset = (Voltage == 1.30))
>
> favstats(~Pulses | Brand, data = Battery_sub)
      Brand min Q1 median Q3 max     mean       sd n missing
1 Energizer  15 17     19 24  25 19.66667 3.741657 9       0
2 Ultracell  25 26     26 27  34 27.11111 2.666667 9       0
```

6. Visualise the data using an appropriate plot. Comment on the difference between brands.

```
> boxplot(Pulses ~ Brand, data = Battery_sub,ylab = "Pulses")
```



7. Test the assumption of equal variance using the Levene's test.

```
> leveneTest(Pulses ~ Brand, data = Battery_sub)

Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  1  1.8756 0.1898
      16
```

Equal variance may be assumed because the p-value of the Levene's test was not statistically significant. However, we will still use the Welch two-sample t-test.

8. Perform an appropriate hypothesis test compare the mean pulses to reach 1.30 V between brands. Use the 0.05 level of significance.
   a. State the Null and Alternate hypothesis for the appropriate hypothesis test.

$$H_0: u_{Energizer} - u_{Ultracell} = 0$$

H_A: $u_{Energizer} - u_{Ultracell} \neq 0$

b. Report the test statistic, *df*, *p*-value and 95% CI of the difference between means from the results of the hypothesis test.

```
> t.test(~Pulses | Brand, data = Battery_sub)

      Welch Two Sample t-test

data:  Pulses by Brand
t = -4.8607, df = 14.46, p-value = 0.0002302
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -10.719534  -4.169354
sample estimates:
mean in group Energizer mean in group Ultracell
               19.66667                27.11111
```

c. Use the *p*-value and 95% *CI* of the mean to make a decision about the Null hypothesis.

Reject H_0 because the p-value was statistically significant ($p < .05$) and the 95% CI of the difference between mean [-10.72, -4.17] did not capture H_0: $u_{Energizer} - u_{Ultracell} = 0$.

d. Conclude whether or not the results of the hypothesis test were statistically significant and draw a conclusion within the context of the example.

The results of the study found a statistically significant mean difference between Energizer and Ultracell pulse counts at the 1.3V endpoint. Ultracell batteries performed significantly better on average than the more expensive Energiser batteries.