Part 2

Assessing normality with more than two repeated measures

Friedman’s ANOVA

Assessing the assumption of normality

- Q-Q plots and Shapiro-Wilk test used

When we have more than two condition, complete these steps separately for each condition

Research question

You are a researcher interested in whether a calorie tracking app you have developed works.

You weigh participants (in kg) before exposing them to the app (Timepoint 1), after 3 months (Timepoint 2), after 6 months (Timepoint 3)

Timepoint 1: Start of the study

Timepoint 2: After 3 months

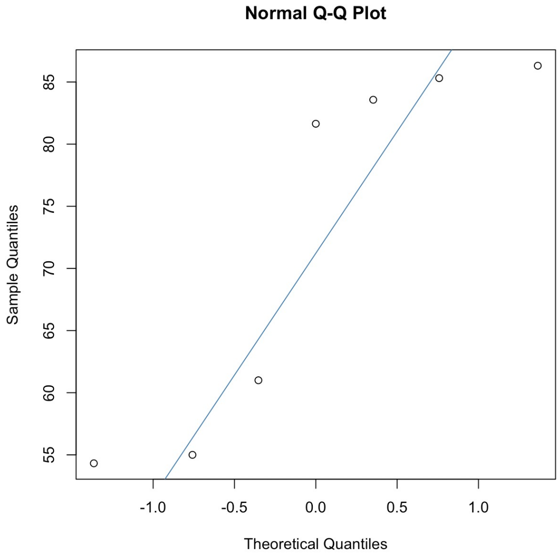
Timepoint 3: After 6 months

Data

Weight		
T1	T2	T3
85.31	59.31	62.41
83.57	60.43	54.34
81.65	80.34	79.65
55.00	53.14	52.12
54.32	51.34	56.32
86.31	75.32	71.34
61.00	60.34	59.34

Assessing the normality assumption
Q-Q plots and Shapiro-Wilk: per timepoint

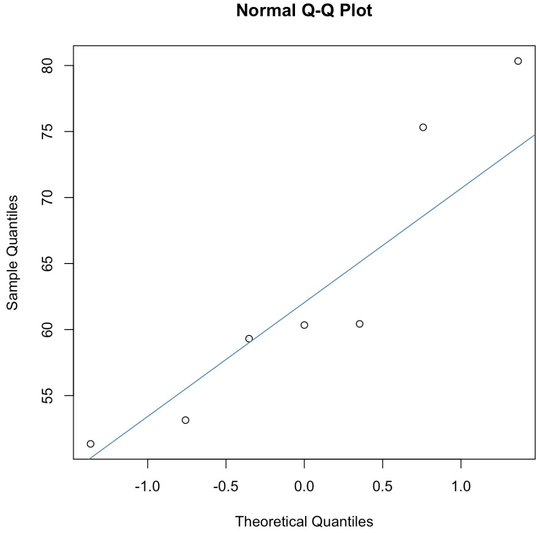
Start



Shapiro-Wilk normality test

```
data:  weight_loss_data$Start
W = 0.78464, p-value = 0.02877
```

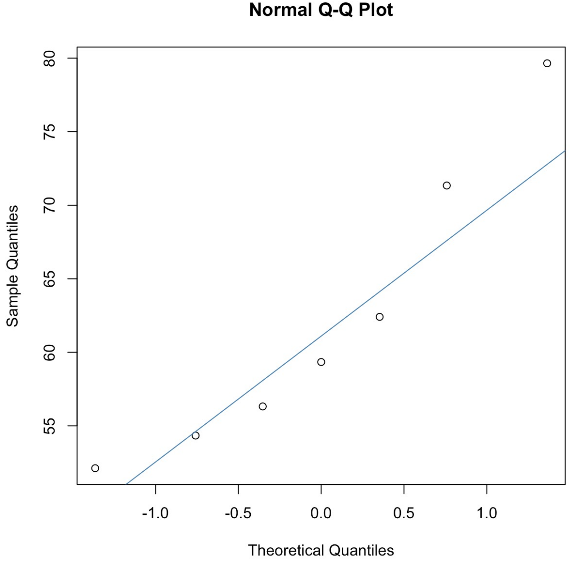
3 months



Shapiro-Wilk normality test

```
data:  weight_loss_data$Three_months
W = 0.87321, p-value = 0.198
```

6 months



Shapiro-Wilk normality test

```
data:  weight_loss_data$Six_months
W = 0.90305, p-value = 0.3499
```

Friedman’s ANOVA

- Alternative to the One-Factor Within-Participants ANOVA
- Appropriate if you have a design with three or more repeated measures

Friedman's ANOVA

Step 1: Rank the data

Weight			Ranks		
T1	T2	T3	T1	T2	T3
85.31	59.31	62.41	3	1	2
83.57	60.43	54.34	3	2	1
81.65	80.34	79.65	3	2	1
55.00	53.14	52.12	3	2	1
54.32	51.34	56.32	2	1	3
86.31	75.32	71.34	3	2	1
61.00	60.34	59.34	3	2	1

Rank data for each participant separately (i.e. each row).

Rank as with the other tests

Friedman's ANOVA

Step 2: Sum the ranks for each timepoint

Weight			Ranks		
T1	T2	T3	T1	T2	T3
85.31	59.31	62.41	3	1	2
83.57	60.43	54.34	3	2	1
81.65	80.34	79.65	3	2	1
55.00	53.14	52.12	3	2	1
54.32	51.34	56.32	2	1	3
86.31	75.32	71.34	3	2	1
61.00	60.34	59.34	3	2	1
SUM OF RANKS			20	12	10

Start = 3+3+3+3+2+3+3 = 20

3 mon = 1+2+2+2+1+2+2 = 12

6 mon = 2+1+1+1+3+1+1 = 10

Friedman's ANOVA

Step 3: Use these values to calculate the test statistic

$$= \left[\frac{12}{Nk(k+1)}\right]\left(R_1^2 + R_2^2 + \ \ldots + R_k^2\right) - 3N(k+1)$$

- N = number of participants
- R_1 = sum of ranks for conditions 1, R_2 = sum of ranks for conditions 2, R_k = simply tells you to repeat this for each conditions
- k = number of conditions

Friedman's ANOVA

Step 3: Use these values to calculate the test statistic

$$= \left[\frac{12}{Nk(k+1)}\right]\left(R_1^2 + R_2^2 + \ \ldots + R_k^2\right) - 3N(k+1)$$

Replace the statistical letters with numbers

$$= \left[\frac{12}{(7*3)(3+1)}\right]\left(20^2 + 12^2 + 10^2\right) - (3*7)(3+1)$$

Friedman's ANOVA

Step 3: Use these values to calculate the test statistic

$$= \left[\frac{12}{(7*3)(3+1)}\right](20^2 + 12^2 + 10^2) - (3*7)(3+1)$$

(7*3)*(3+1) = 84

20*20 = 400

12*12 = 144

10*10 = 100

(3*7)*(3+1) = 84

$$= \frac{12}{84}(400 + 144 + 100) - 84$$

Friedman's ANOVA

Step 3: Use these values to calculate the test statistic

$$= \frac{12}{84}(400 + 144 + 100) - 84$$

400 + 144 + 100 = 644

$$= \frac{12}{84}(644) - 84$$

Friedman's ANOVA

Step 3: Use these values to calculate the test statistic

$$= \frac{12}{84}(644) - 84$$

(12/84)*644 = 92.00

$$= 92.00 - 84 =$$

$$= 8.00$$

Friedman's ANOVA

Step 4: Calculate the degrees of freedom

How are the degrees of freedom calculated?

Degrees of freedom (df) = number of conditions - 1

In our weight loss example, there are 3 conditions:

Df = 3 - 1

Df = 2

Running the analysis in R

Preparing the data

	participant_number	Start	Three_months	Six_months
1	1	85.31	59.31	62.41
2	2	83.57	60.43	54.34
3	3	81.64	80.34	79.65
4	4	55.00	53.14	52.12
5	5	54.32	51.34	56.32
6	6	86.31	75.32	71.34
7	7	61.00	60.34	59.34

We need to get rid of the "participant_number" variable

The function to run the Friedman's ANOVA will only work if the dataframe contains **ONLY** the variables of interest.

Preparing the data

```
weight_loss_data_reduced <- weight_loss_data %>% select(-participant_number)
```

weight_loss_data_reduced:

	Start	Three_months	Six_months
1	85.31	59.31	62.41
2	83.57	60.43	54.34
3	81.64	80.34	79.65
4	55.00	53.14	52.12
5	54.32	51.34	56.32
6	86.31	75.32	71.34
7	61.00	60.34	59.34

Basic code to run Friedman's ANOVA

Keeps output

Function to run test

```
model <- friedman.test(as.matrix(weight_loss_data_reduced))
model
```

Displays output

The Friedman.test function requires the data to be in a matrix. By including "as.matrix" here, it loads the dataframe into this function as a matrix

Dataframe (make sure you have instructed R to use the new, reduced dataset!)

Output from Friedman’s ANOVA

```
	Friedman rank sum test

data:  as.matrix(weight_loss_data_reduced)
Friedman chi-squared = 8, df = 2, p-value = 0.01832
```

Test statistic (X^2F)

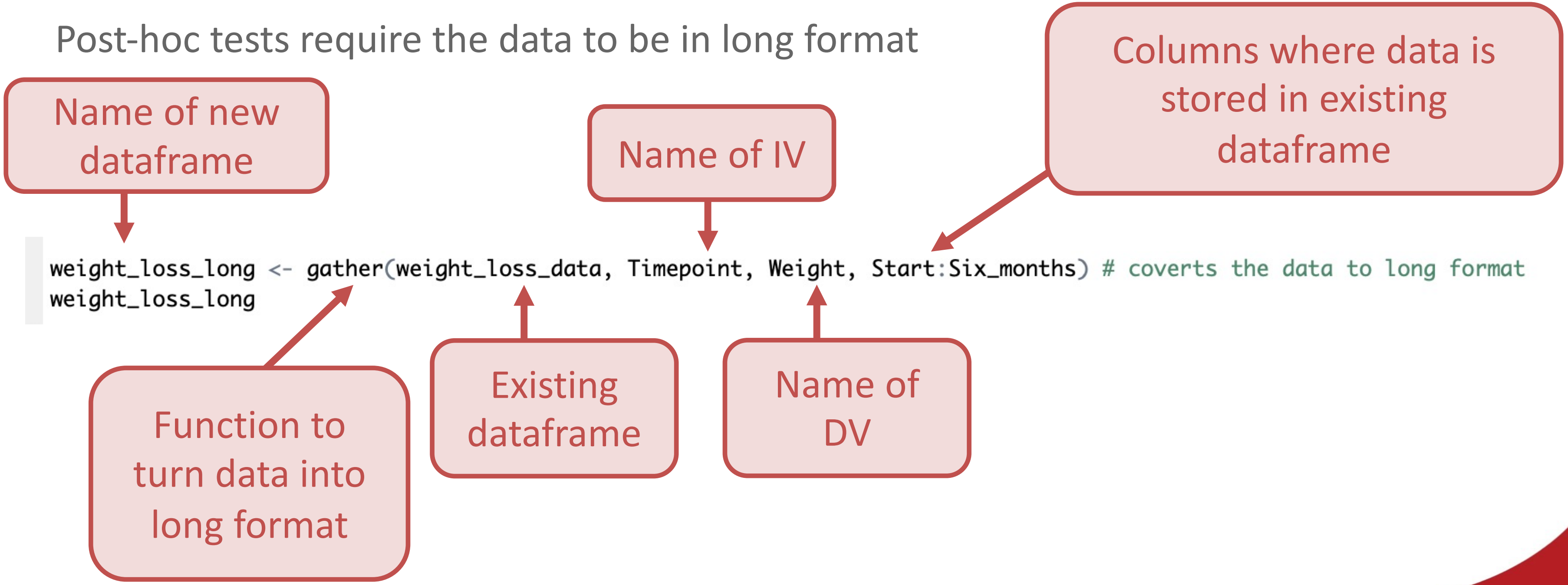
Degrees of freedom

The p-value: A significant effect of timepoint on weight

Where do the differences lie?

We need to perform post-hoc tests

Post-hoc tests require the data to be in long format




```r
weight_loss_long <- gather(weight_loss_data, Timepoint, Weight, Start:Six_months) # coverts the data to long format
weight_loss_long
```

Original vs new dataset

Weight_loss_data:

	participant_number	Start	Three_months	Six_months
1	1	85.31	59.31	62.41
2	2	83.57	60.43	54.34
3	3	81.64	80.34	79.65
4	4	55.00	53.14	52.12
5	5	54.32	51.34	56.32
6	6	86.31	75.32	71.34
7	7	61.00	60.34	59.34

Weight_loss_long:

	participant_number	Timepoint	Weight
1	1	Start	85.31
2	2	Start	83.57
3	3	Start	81.64
4	4	Start	55.00
5	5	Start	54.32
6	6	Start	86.31
7	7	Start	61.00
8	1	Three_months	59.31
9	2	Three_months	60.43
10	3	Three_months	80.34
11	4	Three_months	53.14
12	5	Three_months	51.34
13	6	Three_months	75.32
14	7	Three_months	60.34
15	1	Six_months	62.41
16	2	Six_months	54.34
17	3	Six_months	79.65
18	4	Six_months	52.12
19	5	Six_months	56.32
20	6	Six_months	71.34
21	7	Six_months	59.34

Where do the differences lie?

We can now run the post-hoc tests on the long dataframe (weight_loss_long):

Need to load in the 'PMCMRplus' library

Dependent variable

Column containing the participant number

```
library(PMCMRplus)

frdAllPairsConoverTest(weight_loss_long$Weight, weight_loss_long$Timepoint, weight_loss_long$participant_number, p.adjust = "holm")
```

Function to conduct the post-hoc tests

Independent variable

Which method should R use to correct for multiple comparisons?

Where do the differences lie? Output

Tells you the Conover test has been performed

```
        Pairwise comparisons using Conover's all-pairs test for a two-way balanced complete block design

data: y, groups and blocks

             Six_months Start
Start        0.061      -
Three_months 0.603      0.108

P value adjustment method: holm
```

P-values

P-values adjusted using the Bonferroni-holm correction

P-values are:

Start-6 months = .061
Start to 3 months = .108
3 to 6 months = .603

Effect size

- No easy way to calculate an effect size for the Friedman's ANOVA

Reporting in APA format

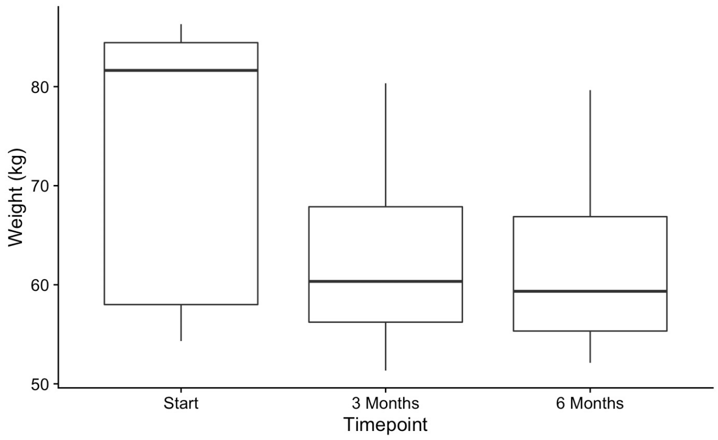
```
        Friedman rank sum test

data:  as.matrix(weight_loss_data_reduced)
Friedman chi-squared = 8, df = 2, p-value = 0.01832
```

```
        Pairwise comparisons using Conc

data: y, groups and blocks

             Six_months Start
Start        0.061      -
Three_months 0.603      0.108

P value adjustment method: holm
```

Friedman's ANOVA revealed that the weight of participants significantly changed over the six months after starting the diet $X^2_F(2) = 8.00$, $p = .018$. Post-hoc comparisons were then conducting using the Conover test. P-values were corrected using Bonferroni-Holm. No significant differences were observed between the start (median = 81.64, range = 54.32-86.31) and the 3 months timepoint (median = 60.34; range = 51.34-80.34; $p = .108$), the start and the 6 months timepoint diet (median = 59.34, range = 52.12-79.65; $p = .061$), or the 3 and the 6 month timepoint ($p = .603$).

Reporting in APA format

Friedman's ANOVA revealed that the weight of participants significantly changed over the six months after starting the diet $X^2_F(2) = 8.00$, p = .018. Post-hoc comparisons were then conducting using the Conover test. P-values were corrected using Bonferroni-Holm. No significant differences were observed between the start (median = 81.64, range = 54.32-86.31) and the 3 months timepoint (median = 60.34; range = 51.34-80.34; p = .108), the start and the 6 months timepoint diet (median = 59.34, range = 52.12-79.65; p = .061), or the 3 and the 6 month timepoint (p = .603).



Lab preparation (~10 minutes)

- Please watch the short lab preparation video prior to your lab
- We will walk through an R script that runs a Friedman's ANOVA

Post-lecture activities

- Now live on Moodle

Thank you for listening!

Please post any questions on the discussion board or on this week's Qualtrics link on Moodle.