Lancaster University

Part 2

Assessing normality when you have two independent groups

How do we assess normality in designs with two independent groups?

1. Quantile-Quantile (Q-Q) plots

2. Shapiro-Wilk test

Data in **each** group should follow a normal distribution

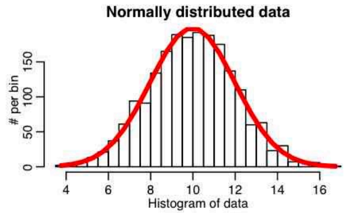
Run these steps separately for each group

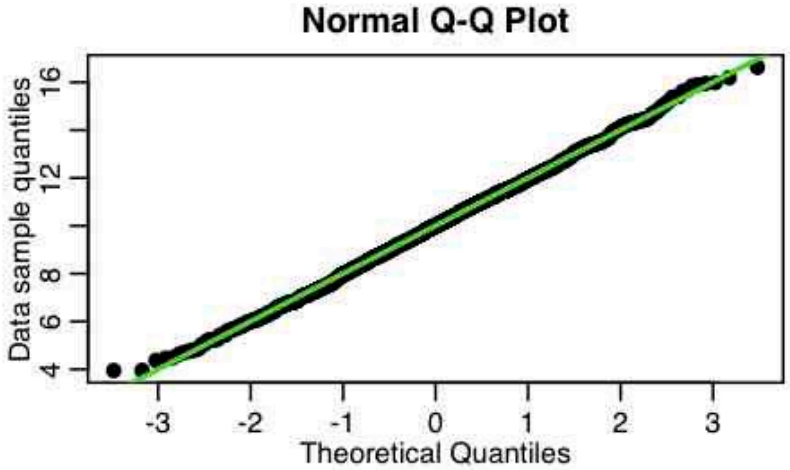
1. Q-Q plots

- Plots values you would expect to get if the distribution was normal against your observed values.

- Expected values are a straight diagonal line – observed values are dots

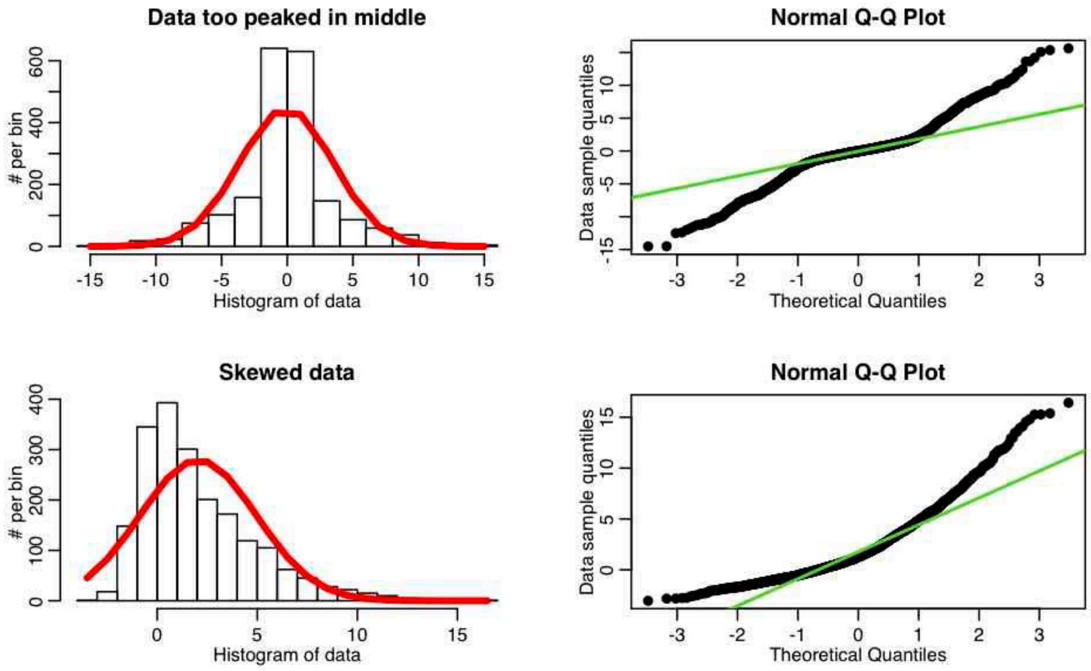
- If normally distributed, dots fall mostly on top of line





1. Quantile-Quantile plots (Q-Q plots)

If not normally distributed, points do not fall on line



2. Shapiro-Wilk test

Test of normality

R code:

Dataframe

```
shapiro.test(data$Outcome)
```

Function to run test

Dependent variable name

R output:

```
> shapiro.test(data$Outcome)

	Shapiro-Wilk normality test

data:  data$Outcome
W = 0.94568, p-value = 0.6178
```

3. Shapiro-Wilk test - Interpretation

```
        Shapiro-Wilk normality test

data:  data$Outcome
W = 0.94568, p-value = 0.6178
```

Null hypothesis: Data came from a normally distributed population

Alternative hypothesis: Data did not come from a normally distributed population

Conduct Shapiro-Wilk test in R and obtain p-value

- $p \leq .05$
  - Data is NOT assumed to come from a normally distributed population
  - Parametric tests not appropriate
  - Consider non-parametric tests
- $p > .05$
  - Data is assumed to come from a normally distributed population
  - Parametric tests may be appropriate – proceed with other assumption checks

Which approach is better?

- Shapiro-Wilk – less subjective, **but** likely to give significant p-values (i.e. deviations from normality) frequently with large sample sizes

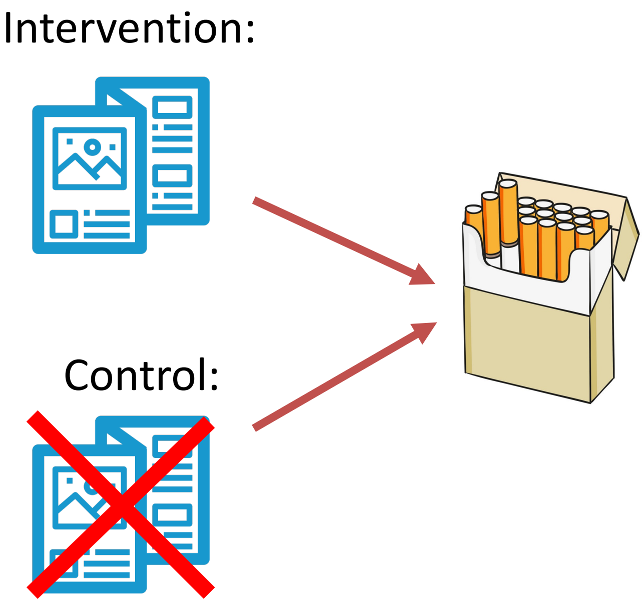
- Consider both outcomes together

- As a general rule of thumb:
  → If the group sample size is <50, rely more on the Shapiro-Wilk test
  → If group sample size >50, rely more on Q-Q plot

An example

You are a researcher interested in smoking behaviours. You develop an intervention to encourage people to stop smoking.

You are interested in whether individuals in the intervention group smoke less cigarettes than individuals in the control group one week later.



Lancaster University

The data

Does a cigarettes intervention (intervention vs control) affect the number of cigarettes consumed?

Group	Frequency
Intervention	0
Intervention	3
Intervention	1
Intervention	0
Intervention	2

Group	Frequency
Control	2
Control	31
Control	29
Control	30
Control	5

Does the data meet the normality assumption? Intervention group

- First limit the data to only the intervention group

```
int <- frequency_data %>% filter(Group == "Intervention")
```

	Group	Frequency
1	Intervention	0
2	Intervention	3
3	Intervention	1
4	Intervention	0
5	Intervention	2

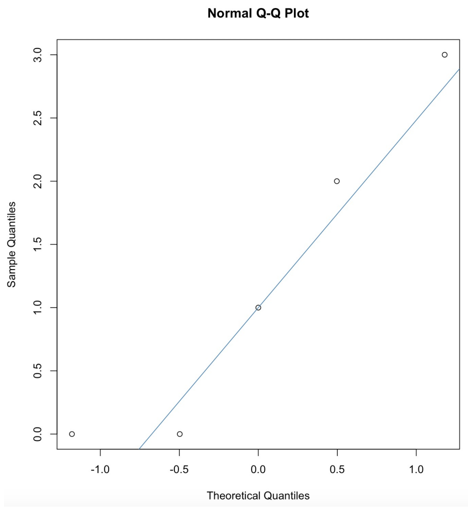
Does the data meet the normality assumption? Intervention group

Produces dots (observed data)

```
qqnorm(int$Frequency)

qqline(int$Frequency, col = "steelblue")
```

Produces line (expected data)



Data appears to meet the normality assumption

```
shapiro.test(int$Frequency)
```

```
Shapiro-Wilk normality test

data:  int$Frequency
W = 0.90202, p-value = 0.4211
```

Repeat for the control group

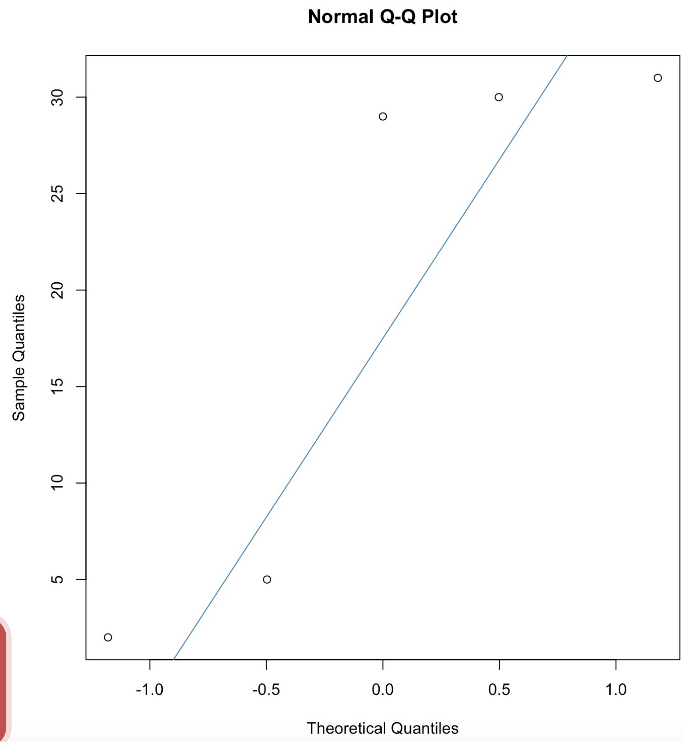
```
control <- frequency_data %>% filter(Group == "Control")

qqnorm(control$Frequency)
qqline(control$Frequency, col = "steelblue")

shapiro.test(control$Frequency)
```

```
        Shapiro-Wilk normality test

data:  control$Frequency
W = 0.75427, p-value = 0.03258
```



Data does not appear to meet the normality assumption

Data for at least one group does not meet the normality assumption. Now what?!

- Conduct the non-parametric alternative of unrelated samples t-test: the Wilcoxon rank-sum test
- This test if you have a design with only two independent groups (and no repeated measures)
- Also known as the Mann-Whitney test