Comparing two means

Parametric and nonparametric

1

Aims and Objectives

Parametric tests

- *t*-tests
  - Independent
  - Dependent
- Rationale for the tests
  - Assumptions
- *t*-tests as a GLM
- Interpretation
- Calculating an effect size
- Reporting results

Nonparametric tests

- Tests based on ranks
  - Wilcoxon rank-sum test
  - Wilcoxon signed-rank test
- Ranking the data

2

Looking at differences

- Instead of looking at relationships, we can also be interested in differences between groups
- The easiest scenario is to compare two groups (*e.g.* control vs. experimental treatment)
  - Independent data
    - Each datapoint constitutes a unique contribution to the sample
  - Dependent data
    - *E.g.* data obtained from repeated measures on same individuals
- This can be done using a *t*-test

3

t-test

- Independent *t*-test
  - Compares two means based on independent data
- Dependent *t*-test
  - Compares two means based on related data
  - Data stem from 'matched' or 'paired' samples, *i.e.* a repeated measures design
- Significance testing
  - Testing the significance of Pearson's correlation coefficient
  - Testing the significance of *b* in regression

4

Rationale for the *t*-test

- Two samples are collected and their means calculated
- If samples come from the same (statistical) population, we expect the means to be approximately equal
  - H_0: There is no difference
- Use the *SE* to gauge variability between sample means
- If difference is larger than expected based on the *SE*:
  1. Sample means in the population fluctuate a lot and we have, by chance, collected two samples atypical of the population
  2. The two samples are representative of their respective populations and come from different populations
- The larger the difference, the more probable the second explanation and if $p_{H_0} < .05$
  - H_0 can be rejected

5

Rationale for the *t*-test

- Recall from Chapter 2
  - Most test-statistics are the ratio of "variance explained" and "variance not explained"

$$\text{test statistic} = \frac{\text{variance explained by the model}}{\text{variance not explained by the model}} = \frac{\text{effect}}{\text{error}}$$

  - For the *t*-test, this looks like

$$t = \frac{\begin{matrix} \textit{observed difference} \\ \textit{between sample means} \end{matrix} - \begin{matrix} \textit{expected difference} \\ \textit{between population means} \end{matrix}}{\begin{matrix} \textit{standard error of difference} \\ \textit{between sample means} \end{matrix}}$$

6

DISCOVERING STATISTICS USING R

Example

- Is arachnophobia specific to real spiders or is a picture enough?
- Predictor variable
  - Exposure to picture of spider
  - Exposure to real spider
- Outcome
  - Anxiety

7

DISCOVERING STATISTICS USING R

The *t*-test as a GLM

Picture group
- The group variable= 0
- b_0= mean of baseline group

Real spider group
- The group variable= 1
- b_1= Difference between means

$$Y_i = b_0 + b_1X_i + \varepsilon_i$$
$$\text{Anxiety}_i = b_0 + b_1\text{Group} + \varepsilon_i$$

$\overline{\text{Anxiety}}_{\text{Picture}} = b_0 + (b_1 \times 0)$

$\overline{\text{Anxiety}}_{\text{Picture}} = b_0$

$b_0 = 40$

$\overline{\text{Anxiety}}_{\text{Real}} = b_0 + (b_1 \times 1)$

$\overline{\text{Anxiety}}_{\text{Real}} = \overline{\text{Anxiety}}_{\text{Picture}} + b_1$

$b_1 = \overline{\text{Anxiety}}_{\text{Real}} - \overline{\text{Anxiety}}_{\text{Picture}}$

$b_1 = 47 – 40 = 7$

8

DISCOVERING STATISTICS USING R

Assumptions of the *t*-test

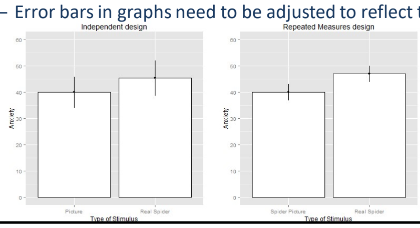
- Both the independent and dependent *t*-test are *parametric tests* that assume:
  - Data are measured at least at the interval level
  - The sampling distribution is normally distributed. In the dependent *t*-test this means that the sampling distribution of the *differences* between scores should be normal, not the scores themselves
- The independent *t*-test also assumes:
  - Variances in the two groups are roughly equal (homogeneity of variance)
  - Scores in the two treatments are independent (because they come from different subjects)

9

DISCOVERING STATISTICS USING R

Plotting data

- A 'complication' with repeated measures design
  - Repeated measures eliminate sampling error due to between-subject differences
  - This reduced *SE* makes statistical tests more sensitive
  - Error bars in graphs need to be adjusted to reflect this



10

DISCOVERING STATISTICS USING R

Example

- Is arachnophobia specific to real spiders or is a picture enough?
- Participants
  - 24 arachnophobic individuals
- Manipulation
  - 12 participants were exposed to a real spider
  - 12 were exposed to a picture of the same spider
- Outcome
  - Anxiety

11

DISCOVERING STATISTICS USING R

The independent *t*-test

- If you have the data for the two groups stored in a single column:

```
> ind.t.test<- t.test(Anxiety~ Group,
                      data= spiderLong)
```

- If you have the data for the two groups stored in two columns:

```
> ind.t.test<- t.test(spiderWide$real,
                      spiderWide$picture)
```

12

DISCOVERING STATISTICS USING R

The independent *t*-test

- Calculating the effect size

$$r = \sqrt{\frac{t^2}{t^2 + df}}$$

- Reporting results
  - On average, participants experienced greater anxiety from real spiders (M= 47.00, SE = 3.18), than from pictures of spiders (M= 40.00, SE = 2.68). This difference was not significant, $t_{(21.4)} = -1.68$, $p > .05$; however, it did represent a medium-sized effect, $r = .34$.

13

DISCOVERING STATISTICS USING R

Example

- Is arachnophobia specific to real spiders or is a picture enough?
- Participants
  - 12 spider phobic individuals
- Manipulation
  - Each participant was exposed to a real spider and a picture of the same spider at two points in time
- Outcome
  - Anxiety

14

DISCOVERING STATISTICS USING R

The dependent *t*-test

- If you have the data for the two treatments stored in a single column:

```
> dep.t.test<- t.test(Anxiety~ Group,
                      data= spiderLong, paired= T)
```

- If you have the data for the two treatments stored in two columns:

```
> dep.t.test<- t.test(spiderWide$real,
                      spiderWide$picture,
                      paired= T)
```

15

The dependent *t*-test

- Calculating the effect size

$$r = \sqrt{\frac{t^2}{t^2 + df}}$$

- Reporting results
  - On average, participants experienced significantly greater anxiety from real spiders (M= 47.00, SE = 3.18) than from pictures of spiders (M= 40.00, SE = 2.68), $t_{(11)} = 2.47$, $p < .05$, $r = .60$.

16

When assumptions are broken

- Independent *t*-test
  - Wilcoxon rank-sum test (= Mann-Whitney test)
- Dependent *t*-test
  - Wilcoxon signed-rank test

Based on ranking the data

- Robust tests
  - Bootstrapping

17

Ranking data

- The tests work on the principle of ranking the data for each group
  - Lowest score= a rank of 1, next lowest score= a rank of 2, and so on
  - Tied ranks are given the same rank: the average of the potential ranks
- For unequal group sizes
  - The test statistic is the sum of ranks in the group that contains the fewest observations
- For equal group sizes
  - The value of the smaller summed rank

18

DISCOVERING STATISTICS USING R

Wilcoxon rank-sum test

- If you have the data for the two treatments stored in a single column

```
> indep.w.test<- wilcox.test(Anxiety~ Group,
                             data= spiderLong)
```

- If you have the data for the two groups stored in two columns

```
> indep.w.test<- wilcox.test(spiderWide$real,
                             spiderWide$picture)
```

19

DISCOVERING STATISTICS USING R

Wilcoxon rank-sum test

- Calculating the effect size

$$r = \frac{z}{\sqrt{N}}$$

- Reporting results
  - Participants did not experience significantly greater anxiety from real spiders (Mdn= 50.0) than from pictures of spiders (Mdn= 40.0), W= 46, p= .14, r= -.30.

20

DISCOVERING STATISTICS USING R

Wilcoxon signed-rank test

- If you have the data for the two treatments stored in a single column:

```
> dep.v.test<- wilcox.test(Anxiety~ Group,
                           data= spiderLong,
                           paired= T)
```

- If you have the data for the two treatments stored in two columns:

```
> dep.v.test<- wilcox.test(spiderWide$real,
                           spiderWide$picture,
                           paired= T)
```

21

Wilcoxon signed-rank test

- Calculating the effect size

$$r = \frac{z}{\sqrt{N}}$$

- Reporting results
  - Participants did not experience significantly greater anxiety from real spiders (Mdn= 50.0) than from pictures of spiders (Mdn= 40.0), W= 8, p= .05, r= -.40.

22

Rest of day…

- Practical Chapter 9 + 15a
  - Read § 9.1, “Cramming Sam’s Tips” and “What Have I discovered about statistics?”
  - Skip sections on R Commander & Wilcox robust methods:
    §9.5.2.3, § 9.5.2.7, § 9.6.3.3, § 9.6.3.7
  - Read § 15.4 + 15.5
  - Solve Smart Alex’s Tasks:
    Chapter 9: 1, 2
    Chapter 15: 1, 2

23