Session 2 – Basic inferential statistics

Thomas J. Faulkenberry, Ph.D.

Tarleton State University

Single sample t-test

Suppose a test prep company claims that their materials will increase your score on a certification test. In the general population, the test is known to have a mean score of $\mu = 200$.

Ten people go through the company's training materials and subsequently take the certification test. Their scores are below:

`216 213 207 194 227 218 193 210 194 199`

Do these data indicate that the company's training materials significantly improve test-takers' scores above the population average of 200?

Hint: look at the `t.test()` function!

Single sample t-test

Solution:

```
x = c(216,213,207,194,227,218,193,210,194,199)
t.test(x, alternative="greater", mu=200)
```

Independent sample t-test

A new drug is supposed to reduce recovery time after surgery. Suppose we measure the recovery time (in days) for patients taking the new drug. Suppose we also measure recovery time for a control group which takes a placebo. The data are as follows:

```
with drug: 15 10 13 7 9 8 21 9 14 8
placebo: 15 14 12 8 14 7 16 10 15 12
```

Do these data indicate that the drug reduces recovery times?

Independent sample t-test

Solution:

```
drug = c(5,10,13,7,9,8,21,9,14,8)
placebo = c(15,14,12,8,14,7,16,10,15,12)
t.test(drug, placebo, alternative="less", var.equal=TRUE)
```

Paired sample t-test

Ten participants rate their enjoyment of a certain statistical software package before and after a day-long training session. Their scores (out of 5 possible points) are listed below:

```
Before: 2 1 4 1 4 3 3 2 3 5
After:  3 0 5 2 5 5 5 4 4 5
```

Was the workshop effective at increasing participant satisfaction with the statistical software package?

Paired sample t-test

Solution:

```
before = c(2,1,4,1,4,3,3,2,3,5)
after = c(3,0,5,2,5,5,5,4,4,5)
t.test(before, after, alternative="less", paired=TRUE)
```

One-way ANOVA

Suppose a committee is trying to score 27 different scholarship applications. As the job is too much work for one grader, suppose 3 are used. The scholarship committee would like to ensure that each grader is using the same grading scale, as otherwise the students aren't being treated equally. One approach to checking if the graders are using the same scale is to randomly assign each grader 9 applications and then compare the scores for the 3 graders, knowing that the differences should be due to chance errors if the graders all grade equally. Suppose the grading scale is on the range 1-5 with 5 being the best and the scores are reported as:

grader A:	4	3	4	5	2	3	4	5
grader B:	4	4	5	5	4	5	4	4
grader C:	3	4	2	4	5	5	4	4

One-way ANOVA

We'll show two methods:

1. use `oneway.test()` function

2. use `aov()` method

One-way ANOVA

Setup:

```
graderA = c(4,3,4,5,2,3,4,5)
graderB = c(4,4,5,5,4,5,4,4)
graderC = c(3,4,2,4,5,5,4,4)
scores = data.frame(graderA, graderB, graderC)
scores = stack(scores)
```

One-way ANOVA

Descriptives:

```
attach(scores)
boxplot(values~ind)
tapply(scores, INDEX=ind, FUN="mean")
```

One-way ANOVA

Method 1:

```
oneway.test(values~ind, data=scores, var.equal = TRUE)
```

One-way ANOVA

Method 2:

```
model = aov(values~ind, data=scores)
summary(model)
```

Factorial ANOVA

140 participants were asked to recall a list of verbally presented words under two experimental manipulations:

- Manipulation 1: half of the participants were given a cue word at study (i.e., an *encoding cue*)
- Manipulation 2: half of the participants were given a cue word at test (i.e., a *retrieval cue*)

Factorial ANOVA

These two manipulations combined to produce 4 experimental conditions:

- cues at both encoding and retrieval
- cue at encoding, but not at retrieval
- no cue at encoding, but cue at retrieval
- no cues at either encoding or retrieval

Factorial ANOVA

The number of words correctly recalled for each participant can be seen in the following dataset

```
data = read.csv("https://git.io/fNbnc")
attach(data)
```

Factorial ANOVA

Descriptives:

```
tapply(correctRecall, INDEX=list(encodingCue,retrievalCue), FUN="mean")
```

Factorial ANOVA

Let's perform a factorial ANOVA on these data.

```
model = aov(correctRecall ~ encodingCue*retrievalCue, data=data)
summary(model)
```

Before next session

We'll need to install the `Tidyverse` package. Type the following in the console:

```
install.packages("tidyverse")
```