

Week 7 Tutorial Questions

PART A - Two-Way ANOVA

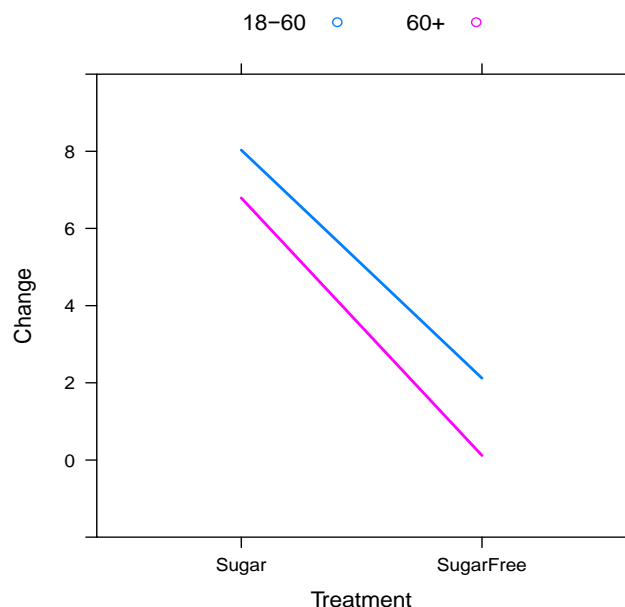
A project within the Islands investigated whether consuming sugar influences memory performance, and if any effects were the same for elderly subjects.

A sample of 40 subjects was randomly chosen from houses in Helluland, with 20 subjects aged 18-60 and 20 aged over 60. Each subject played a pairs memory game with 30 cards as quickly as they could, with the results recorded in seconds. Within each age group, half the subjects were randomly assigned to consume 50 g of lollies, while the other half consumed 50 g of sugar-free lollies. After 15 minutes, they repeated the memory game task.

- a) The table below shows the summary statistics for the decrease in memory game time between the first and second plays:

Age	Treatment	<i>n</i>	Mean (s)	SD (s)
18-60	Sugar	10	8.03	5.054
	Sugar-free	10	2.12	8.140
60+	Sugar	10	6.79	6.289
	Sugar-free	10	0.12	5.547

Draw an interaction effects plot for this data. Use the vertical axis for the mean change in game time and the horizontal axis for the treatment.



- b) Briefly discuss the potential interaction effect from your plot.

- c) To compare the interaction effects of treatment and age group on the decrease in memory game time, a two-way analysis of variance is conducted, giving the following results in R:

```
summary(aov(Change ~ Treatment*Age, data=sugar))
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Treatment	?	395.6	395.6	9.761	0.00351	**
Age	?	26.2	26.2	0.647	0.42630	
Treatment:Age	?	1.4	1.4	0.036	0.85135	
Residuals	?	1459.2	40.5			

What are the degrees of freedom for each of the variance components?

- d) Calculate and interpret the R^2 value of this two-way ANOVA model.
- e) Interpret the three p -values from the ANOVA table. What do you conclude about the effect of consuming sugar on memory performance, and how it relates to age?

Part B: Two-Way ANOVA using RStudio (Based on a past exam paper)

The Maryland Biological Stream Survey conducted a study to investigate the effects of physical and chemical water characteristics on the abundance of Longnose Dace (*Rhinichthys Cataractae*, a freshwater minnow native to North America). The researchers sampled 54 segments of streams, each 50m long. In each 50m segment, they recorded the number of fish, sulfate concentration, and dissolved oxygen level of the stream water. The datafile "DaceFish.csv" contains the following variables.

Sulfate – Sulfate concentrate in three levels (S1 – Less than 20mg/litre, S2 – Less than 40mg/litre, S3 – more than 40 mg/litre).

Oxygen – Dissolved oxygen in three levels (O1 – Less than 8mg/litre, O2 – Less than 10mg/litre, O3 – More than 12g/litre).

Fish – Number of Longnose Dace.

- a) What are the variable types in the data?
- b) How many segments had Sulfate concentration less than 20mg (S1) and dissolved oxygen level less than 10g/litre (O2)
- c) The study uses a Two-Way ANOVA model to investigate the interaction effect of sulfate and oxygen concentrations on the mean number of longnose daces. What are the numerator and denominator degrees of freedom of the F-test statistic to test the statistical significance of the interaction effect?
- d) What do you conclude about the interaction effect?

Part A – Chi-square test

Question 1

In semester 2 2020, students were asked in a Zoom poll to “toss a coin twice and indicate your outcome”. The combined counts from the two streams are shown in the following table:

HH	HT	TH	TT
63	80	64	42

- a) Supposing these were independent tosses of fair coins, how many counts would you expect in each of these four groups?
- a) Calculate the statistic to compare the observed counts to the expected count. Is there any evidence that this process did not work as intended?

Question 2

The following table shows counts of students in the 2018 and 2020 surveys who lived at home or in a shared flat/house:

Year	Home	Shared Flat/House
2018	391	185
2021	550	146

Is there any evidence of a change in living situations between 2018 and 2020?