

community project

encouraging academics to share statistics support resources

All stcp resources are released under a Creative Commons licence

stcp-karadimitriou-interactionsR

The following resources are associated:

Two-way ANOVA in R, two-way ANOVA script and interactions in R script

Interactions in R

Dependent variable: Continuous (scale)

Independent variables: Two categorical (Two-way ANOVA)

An interaction is the combined effect of two independent variables on one dependent variable. Possible interactions can be investigated when carrying out ANOVA with at least two independent grouping variables or multiple regression.

Summarising the effect of two categorical variables on one independent variable

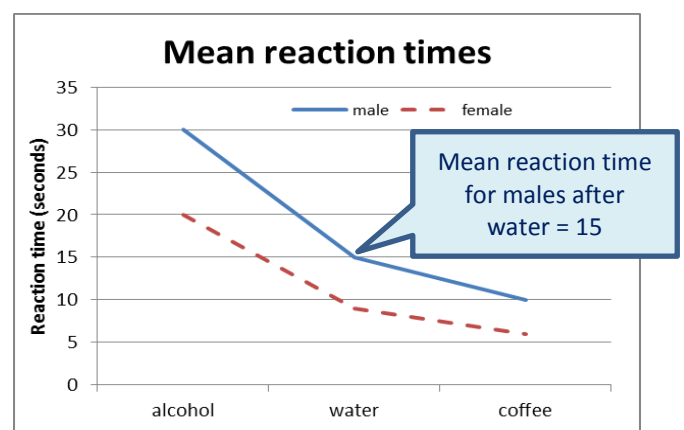
A line chart can be used to compare the means of combinations of two categorical independent variables. It is particularly useful for looking at interaction effects and can also be called an interaction plot or means plot. The lines connect means of each combination.

Example: An experiment was carried out to investigate the effect of drink on reaction times in a driving simulator. Participants were given alcohol, water or coffee. The mean reaction times by group are contained in the table to the right.

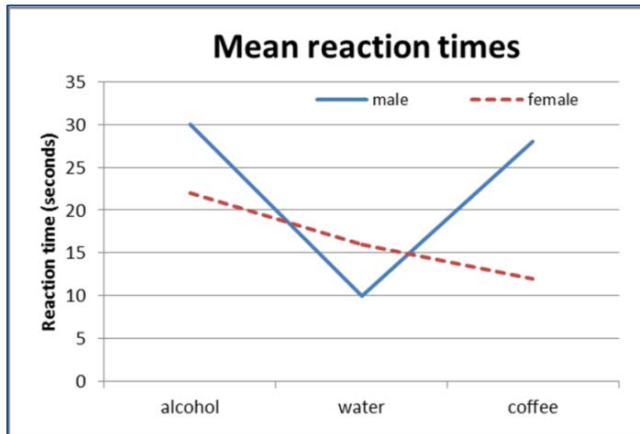
Mean Reaction Times	Male	Female
Alcohol	30	20
Water	15	9
Coffee	10	6

The six means can be displayed in a line/means plot.

For both males and females, the fastest (i.e. lowest) reaction times are after coffee, followed by water then alcohol. Females are faster than males after all three drinks. There is no interaction between gender and drink as the lines are reasonably parallel.



If there was an interaction what would it look like?



An interaction occurs when the lines are not quite so parallel; such that the means of one group do not follow the same pattern as the other group. Here males have their fastest reaction after water, but females have their fastest reaction after coffee. Males are faster than females after water but females are faster after coffee and alcohol.

Producing an interaction plot in R

Data: The data set 'Diet' contains information on 78 people who undertook one of three diets. The dependent variable is weight lost and the independents are the diet and gender. To open the file use the `read.csv()` command and tell R to use the diet dataset until further notice using `attach(dataset)` so that variable names can be used directly.

```
dietR<-read.csv('C:\\...',header=T)
attach(dietR)
```

This example compares weight lost by diet and gender. Tell R that 'Diet' is a factor:

```
Diet<-as.factor(Diet)
```

Tell R that gender is categorical and label the values that gender takes so that the interpretation of the graph can be more straightforward.

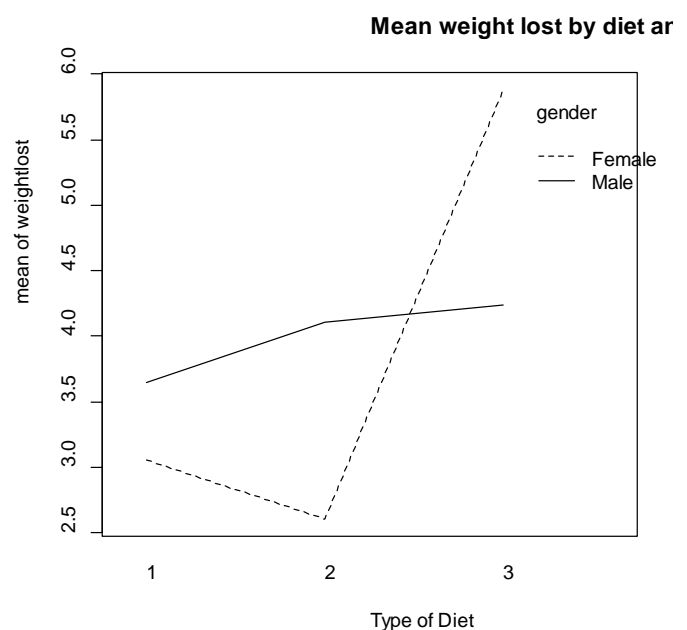
```
gender<-factor(gender,c(0,1),labels=c('Female','Male'))
```

Calculate the weight lost by person (difference in weight before and after the diet) and add the variable to the dataset. Then attach again.

```
dietR$weightlost<-pre.weight-weight6weeks
attach(dietR)
```

To produce the plot in R, use the `interaction.plot(independent1, independent2, dependent)` command. The first term specifies the x axis variable, the second term specifies the variable for the lines and the third one specifies the y axis.

```
interaction.plot(Diet, gender,
weightlost,main="Mean weight lost
by diet and gender",ylab = "mean of
weightlost",xlab = "Type of Diet")
```

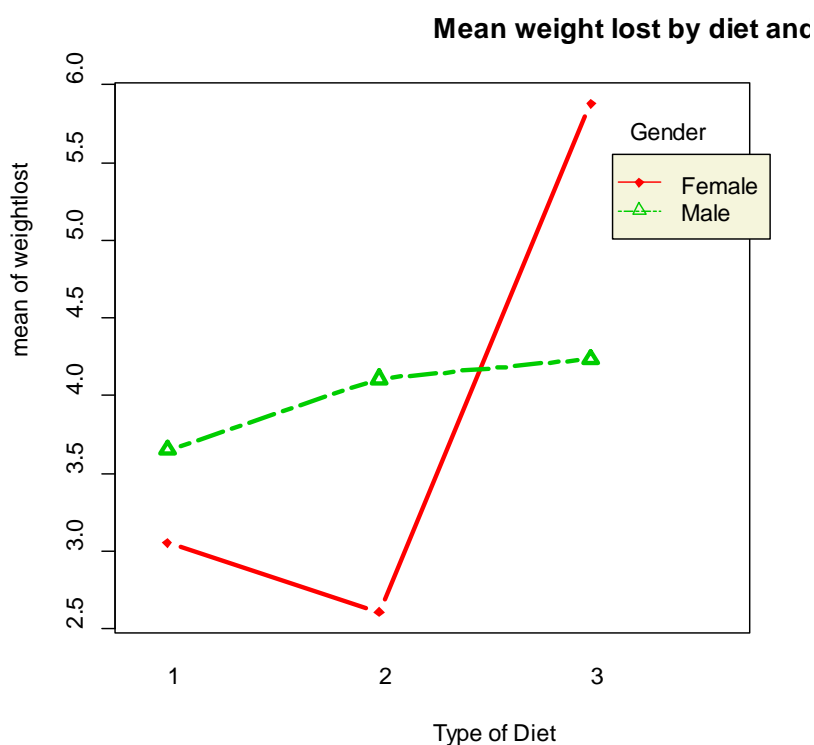


The attributes `lty` and `lwd` specify the two types of lines and the thickness of them respectively. The colour is controlled by `col`. A different type of line can be plotted for each gender using `col=c(a:b)` for colour and `lty=c(a,b)` for line type.

```
interaction.plot(Diet, gender, weightlost, lty = c(1,
12), col=c(2:3), lwd = 3, ylab = "mean of weightlost",
xlab = "Type of Diet", main="Mean weight lost by diet and gender")
```

The `trace.label` attribute changes the legend title from the variable name, `leg.bty = "o"` puts a box around the legend and `leg.bg = " "` controls the background colour of the legend. Use `?interaction.plot` to find out more options.

```
interaction.plot(Diet, gender, weightlost, main="Mean weight lost
by diet and gender", type="b", pch=c(18,24), lty = c(1,
12), col=c(2:3), lwd = 3, trace.label = "Gender", leg.bty="o",
leg.bg="beige", ylab = "mean of weightlost",
xlab = "Type of Diet")
```



There is a clear interaction between diet and gender as the lines are not parallel. For males, there is very little difference between the means for the diets but for females, Diet 3 led to a much higher mean weight loss.

Note: Positive weight lost means the subject lost weight.