Data Analytics Communication Analysing decision.dat experiment data

Gonçalo Hora de Carvalho

2018

1 About the data

Researchers in this experiment wanted to study how participants accumulate evidence for making a decision, and how this mental process manifested in their brains. Participants were shown randomly moving dots, some moving to the right others to the left. Participants are given a button for the 'left' direction and a 'right' arrow for the right direction. The difficulty of the task is dependent on the amount of dots moving synchronously, which the researchers manipulate. Participants are given a difficult and an easy coherence condition, where the coherence is adjusted such that they score on average approximately 70 and 90 percent correct. There was also a control condition were participants are shown the correct answer directly by means of a large arrow indicating the correct direction response.

2 The current task

The task at hand is to check whether participant 1 shows a difference in RT (response time) between the low and high coherence conditions for the dots trials in the random dot motion decision task from which the data are derived. To do this comparison the control condition with arrows is removed using the following (where subj1 refers to the data for subject 1):

```
library(tidyverse)
dat<-read.table("/Users/mvugt/teaching/DataAnalytics&Communication/decision.dat",header=T)
subj1 <- dat %>% filter(subjNo == 1)
onlyDots <- subj1 %>% filter(isDots == 1)
```

Now with the correct data set a t-test will be deployed to find out whether there is a difference in RT between low- and high-coherence trials (these data are not normally distributed, so technically a t-test is not allowed, but for the sake of demonstrating a t-test this fact will be ignored).

1 T-test

T-Test

The test deployed was a paired one sample t-test to check the mean difference between RT and cohFac (if this is zero then there's no significant difference). Degrees of freedom (sample size - df) is 2427, with a resulting P-value of 2.2e-16

The resulting paired sample t-test follows:

```
t.test(subj1$RT,subj1$cohFac, paired=TRUE)

        Paired t-test

data:  subj1$RT and subj1$cohFac
t = 13.629, df = 2427, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.1749615 0.2337693
sample estimates:
mean of the differences
              0.2043654
```

In order to better visualise the data a boxplot was deployed:

The boxplot seems to show that when coherence is factored in there is not much variation in the RT of the subject (less than one second).

The test to be deployed in this scenario must allow for comparison of differences between two samples that differ in a factor. Firstly we deployed the given code:

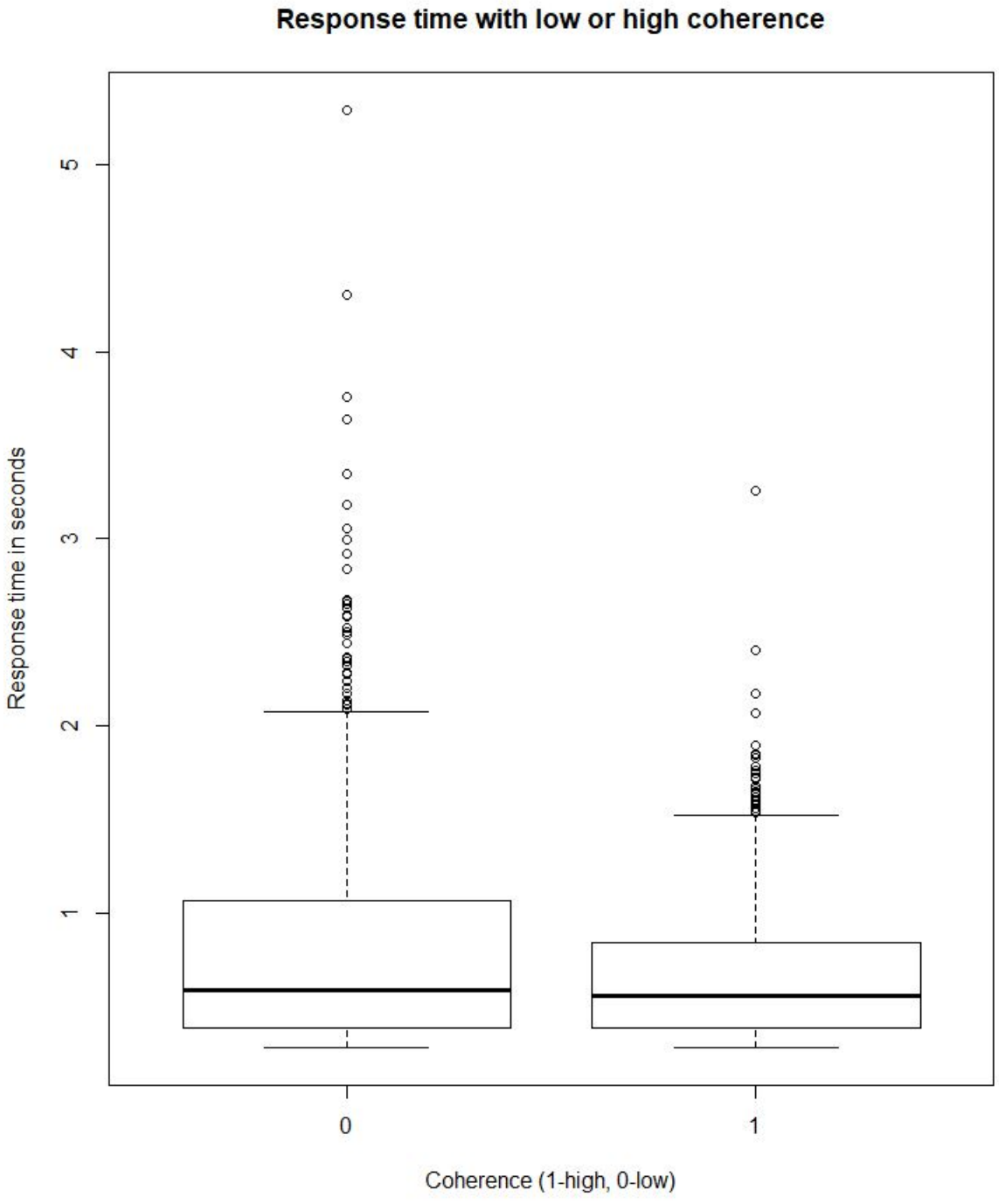
```
onlyDotsAll <- dat[dat$isDots==1,]

avByBlock <- onlyDotsAll  %>% group_by(cohFac, subjNo)  %>% summarise_all(mean)
```

And then in order to observe the data by summarising the outcome of the code given:

```
summary(avByBlock$RT)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.5277  0.7619  0.8589  0.9400  1.1051  2.3619
```

The t.test() function can be used to perform both one and two sample t-tests on vectors of data, but since our data is not normally distributed the wilcox.text() function is used instead since it offers the same basic functionality:


Response time with low or high coherence
Response time in seconds
1
2
3
4
5
0
1
Coherence (1-high, 0-low)

```
t.test(avByBlock$RT ~ avByBlock$cohFac)

        Welch Two Sample t-test

data:  avByBlock$RT by avByBlock$cohFac
t = 3.7045, df = 27.985, p-value = 0.0009232
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.1350647 0.4692105
sample estimates:
mean in group 0 mean in group 1
      1.0911025       0.7889649
```

We can see that the null hypothesis is refuted with a p-value of 0.0009232.

e)

The following code was deployed in order to enable a Cohen's D test:

```
cohensD(avByBlock$RT, avByBlock$cohFac)
[1] 1.04639
```

This function is used to differentiate between two means as a complementary report to the t-test by calculating the Cohen's d measurement of effect size. The result was 1.04639 - which indicates a positive magnitude (different from 0, or rather, there is an effect).

Non-parametric alternatives

The Wilcox test was used since it does not require a normal distribution.

2.1 b)

```
 wilcox.test(avByBlock$RT, avByBlock$cohFac)

        Wilcoxon rank sum test with continuity correction

data:  avByBlock$RT and avByBlock$cohFac
W = 1403, p-value = 0.006273
alternative hypothesis: true location shift is not equal to 0
```

It was found that the means of reaction times are likely to differ when grouping the population by cohFac set to 1 and cohFac set to 0. This is shown from the P-value = 0.006273.

One-way ANOVA

A one-way ANOVA indicated a significant effect of task condition (dots versus control) on RT F(1.05)=13.72 with a p-value: 0.000588. This indicates that there is a significant difference between the groups. The following code was deployed:

```
> anova <- aov(avByBlock$RT ~ factor(avByBlock$cohFac))
> summary(anova)
                          Df Sum Sq Mean Sq F value
factor(avByBlock$cohFac)   1  1.050  1.0498   13.72
Residuals                 44  3.366  0.0765
                            Pr(>F)
factor(avByBlock$cohFac) 0.000588 ***
Residuals
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In order to verify the results, first a residuals vs fits plot is deployed:

It can be seen that there is no relationship between residuals and fitted values which leads to the assumption of homogeneity. The following plot is a mean plot (using plotmeans from the gplots package):

b)

Both p-values were significant, showing that the means differ.

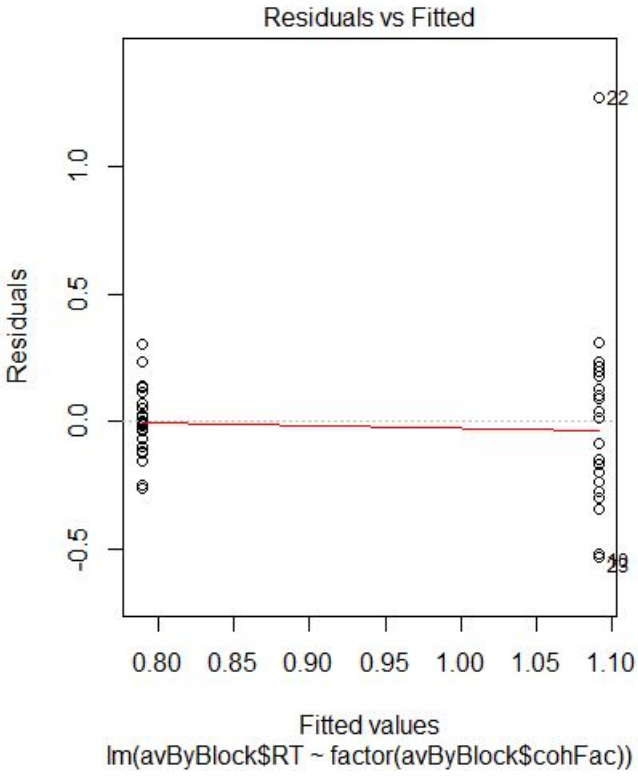
d)

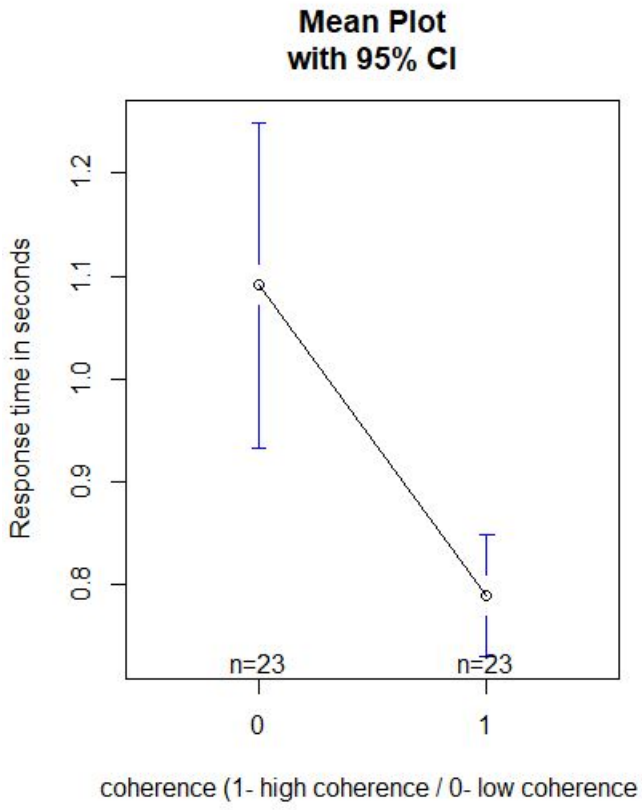
The value of the effect size (approx. 0.24) is close to being considered a medium effect size (=0.30). Which means that the strength of my statistical claim is not null, but at the same time, isn't a strong one.

```
 etaSquared(anova, type = 2, anova = FALSE)
                             eta.sq eta.sq.part
factor(avByBlock$cohFac) 0.2377398   0.2377398
```

Two-way ANOVA

```
res.aov2 <- aov(RT ~ factor(cohFac) + isLeft, data=avByBlock)
> summary(res.aov2)
```


Residuals vs Fitted
Residuals
1.0
0.5
0.0
-0.5
22
0.80
0.85
0.90
0.95
1.00
1.05
1.10
Fitted values
lm(avByBlock$RT ~ factor(avByBlock$cohFac))


Mean Plot
with 95% CI
Response time in seconds
0.8
0.9
1.0
1.1
1.2
n=23
n=23
0
1
coherence (1- high coherence / 0- low coherence

```
               Df Sum Sq Mean Sq F value   Pr(>F)
factor(cohFac)  1  1.050  1.0498  13.433 0.000675 ***
isLeft          1  0.005  0.0054   0.069 0.793970
Residuals      43  3.361  0.0782
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Repeated-measure Anova

a)

In the case of within subjects experiment then a repeated measurements anova is preferable (this case). If there is a between subjects factor then Anova should be performed.

```
summary(aov(RT ~ factor(cohFac) + Error(subjNo/factor(cohFac)), data=avByBlock))

Error: subjNo
          Df    Sum Sq   Mean Sq F value Pr(>F)
Residuals  1 0.0008041 0.0008041

Error: subjNo:factor(cohFac)
               Df Sum Sq Mean Sq
factor(cohFac)  1 0.8769  0.8769

Error: Within
               Df Sum Sq Mean Sq F value Pr(>F)
factor(cohFac)  1  0.185 0.18545   2.323  0.135
Residuals      42  3.353 0.07982
```

There is no significant effect of the coherence on the reaction time according to the test (p-value=0.135 ¿ 0.05).