# *Lab 5*

*June 29, 2017*

## *Lab 5 Data-analysis tips*

To give you an idea about the analyses you will be performing, we will again create simulated data to mimic aspects of the experiment, and then go through the steps of performing the analysis.

We will simulate data for the switching cost. Specifically, we will imagine that women have a smaller switching cost than men. The code below generates sample data for 10 women and 10 men, who each have mean reactions in the repeat and switch conditions. For women, the mean reaction times were 580 ms for repeat and 600 ms for switch sequences, for a total expected switch cost of 20 ms. For men, the mean reaction times were 580 ms for repeat and 650 ms for switch sequences, for a total expected switch cost of 70 ms.

### *Simulating the data*

```
women_switch <-round(rnorm(10,600,20))
women_repeat <-round(rnorm(10,580,20))
men_switch <-round(rnorm(10,650,20))
men_repeat <-round(rnorm(10,580,20))
all_data<-data.frame(Subject=c(rep(seq(1,10,1),2),
                               rep(seq(11,20,1),2)),
                     Gender=rep(c("Female","Male"),each=20),
                     Sequence=rep(rep(c("switch","repeat"),each=10),2),
                     RT=c(women_switch,women_repeat,
                          men_switch,men_repeat))

kable(all_data,format="latex")
```

| Subject | Gender | Sequence | RT |
|---|---|---|---|
| 1 | Female | switch | 572 |
| 2 | Female | switch | 621 |
| 3 | Female | switch | 569 |
| 4 | Female | switch | 609 |
| 5 | Female | switch | 574 |
| 6 | Female | switch | 598 |
| 7 | Female | switch | 590 |
| 8 | Female | switch | 623 |
| 9 | Female | switch | 586 |
| 10 | Female | switch | 590 |
| 1 | Female | repeat | 556 |
| 2 | Female | repeat | 566 |
| 3 | Female | repeat | 588 |
| 4 | Female | repeat | 611 |
| 5 | Female | repeat | 575 |
| 6 | Female | repeat | 582 |
| 7 | Female | repeat | 576 |
| 8 | Female | repeat | 578 |
| 9 | Female | repeat | 558 |
| 10 | Female | repeat | 581 |
| 11 | Male | switch | 660 |
| 12 | Male | switch | 657 |
| 13 | Male | switch | 622 |
| 14 | Male | switch | 651 |
| 15 | Male | switch | 664 |
| 16 | Male | switch | 623 |
| 17 | Male | switch | 642 |
| 18 | Male | switch | 676 |
| 19 | Male | switch | 659 |
| 20 | Male | switch | 618 |
| 11 | Male | repeat | 585 |
| 12 | Male | repeat | 571 |
| 13 | Male | repeat | 575 |
| 14 | Male | repeat | 594 |
| 15 | Male | repeat | 572 |
| 16 | Male | repeat | 605 |
| 17 | Male | repeat | 556 |
| 18 | Male | repeat | 650 |
| 19 | Male | repeat | 594 |
| 20 | Male | repeat | 606 |

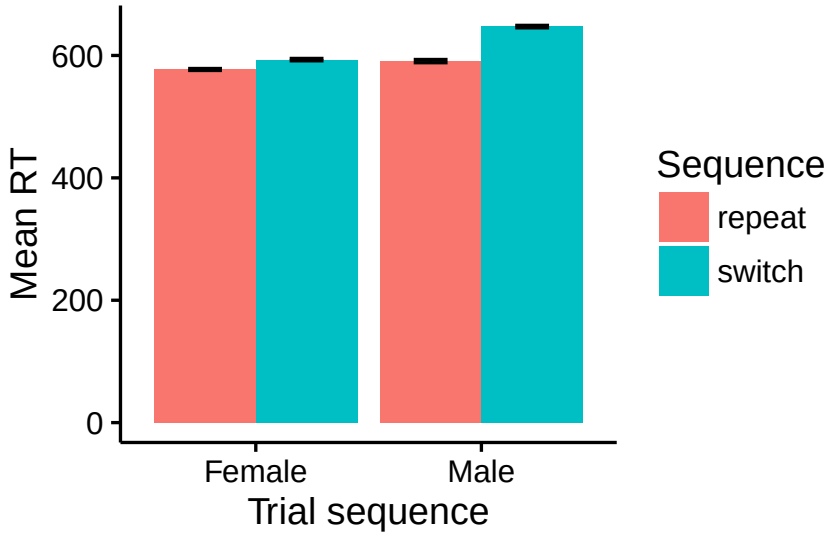
*plotting the data*

We can plot the data at least two ways. See the bar and line graphs below. Note that the x-axis changes between graphs.

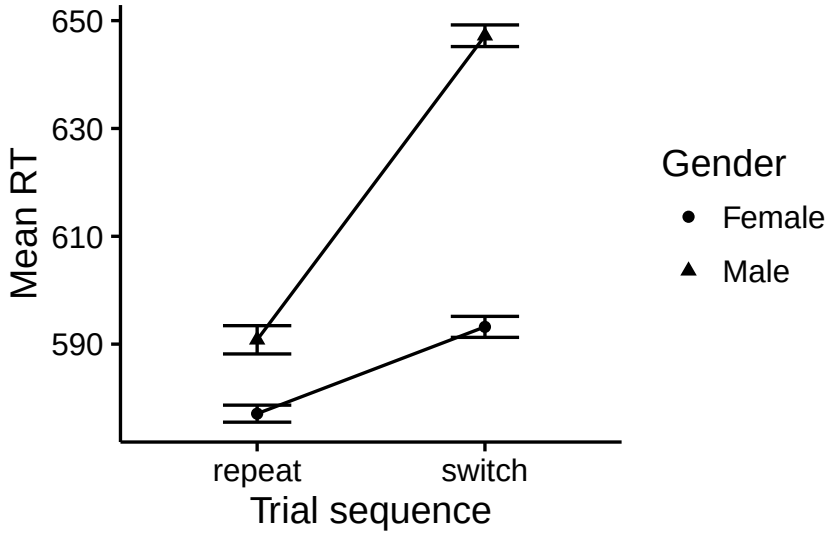
```
library(ggplot2)
library(plyr)
sde<-function(x){sd(x)/length(x)}
plot_means<-ddply(all_data,.(Gender,Sequence),summarise,
                        MeanRT=mean(RT),
                        SE=sde(RT))

limits <- aes(ymax = MeanRT + SE, ymin = MeanRT - SE)

ggplot(plot_means,aes(x=Gender, y=MeanRT, group=Sequence,fill=Sequence))+
  geom_bar(position="dodge",stat="identity")+
  geom_errorbar(limits, width=.3,position=position_dodge(.9))+
  theme_classic(base_size=12)+
  ylab("Mean RT")+
  xlab("Trial sequence")
```




```
ggplot(plot_means,aes(x=Sequence, y=MeanRT, group=Gender,shape=Gender))+
  geom_line()+
  geom_point()+
  geom_errorbar(limits, width=.3)+
  theme_classic(base_size=12)+
  ylab("Mean RT")+
  xlab("Trial sequence")
```




The graphs show that the switch costs (difference between repeat and switch trials) is smaller for women and men. Which is good, because we are simulating the data with this outcome in mind.

*Running the ANOVA*

The next step is to conduct an ANOVA. This design has two factors or independent variables, gender and trial sequence. The gender variable is between-subjects, and the trial sequence variable is within subjects. So, we will run a 2 (Gender: Female vs. Male) x 2 (trial sequence: Repeat vs. Switch) mixed design ANOVA with Gender as the betwen-subjects factor, and trial sequence as the within-subjects factor.

```
library(broom)
all_data$Subject<-as.factor(all_data$Subject)
aov.out<-aov(RT~Gender*Sequence+Error(Subject/Sequence),all_data)
aov_summary<-summary(aov.out)
kable(xtable(aov_summary),format="latex")
```

| | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
|---|---|---|---|---|---|
| Gender | 1 | 11458.225 | 11458.2250 | 22.10104 | 1.78e-04 |
| Residuals | 18 | 9332.050 | 518.4472 | NA | NA |
| Sequence | 1 | 13140.625 | 13140.6250 | 38.47507 | 7.50e-06 |
| Gender:Sequence | 1 | 4060.225 | 4060.2250 | 11.88813 | 2.87e-03 |
| Residuals | 18 | 6147.650 | 341.5361 | NA | NA |

```
mt<-model.tables(aov.out,"means")
mt

## Tables of means
```

```
## Grand mean
##
## 602.075
##
##  Gender
## Gender
## Female   Male
##  585.1  619.0
##
##  Sequence
## Sequence
## repeat switch
##  583.9  620.2
##
##  Gender:Sequence
##         Sequence
## Gender   repeat switch
##   Female 577.1  593.2
##   Male   590.8  647.2
```

The above shows the ANOVA table and the means for the main effects and interaction. We also conduct t.test comparisons to look at the switch costs separately for men and women.

```
FemaleT<-t.test(RT~Sequence,all_data[all_data$Gender=="Female",],paired=TRUE,var.equal=TRUE)
MaleT<-t.test(RT~Sequence,all_data[all_data$Gender=="Male",],paired=TRUE,var.equal=TRUE)
FemaleT
```

```
##
##  Paired t-test
##
## data:  RT by Sequence
## t = -2.3034, df = 9, p-value = 0.04674
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -31.9115634  -0.2884366
## sample estimates:
## mean of the differences
##                   -16.1
```

```
MaleT
```

```
##
##  Paired t-test
##
## data:  RT by Sequence
```

```
## t = -6.0205, df = 9, p-value =
## 0.0001975
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -77.59196 -35.20804
## sample estimates:
## mean of the differences
##                   -56.4
```

*Writing up the results*

The next step is to interpret the results and write them up. Here is an example write-up.

The mean reaction times for each subject in trial sequence condition were submitted to a 2 (Gender: Female vs. Male) x 2 (trial sequence: Repeat vs. Switch) mixed design ANOVA with Gender as the betwen-subjects factor, and trial sequence as the within-subjects factor. Mean reaction times in each condition collapsed across subjects are displayed in Figure 1.

The main effect of gender was significant, F(1, 18) = 22.1, MSE = 518.45, p < 0. Women (585 ms) had faster mean reaction times than men (619 ms).

The main effect of trials sequence was significant, F(1, 18) = 38.48, MSE = 341.54, p < 0. Repeat trials (584) had faster mean reaction times than switch trials (620).

Most important was the significant two-way interaction between gender and trial sequence, F(1, 18) = 11.89, MSE = 341.54, p < 0.003. We interpreted the interaction further by conducting the following comparisons. Women showed a significant switch cost, t(9) = -2.3, p = 0.047, with faster mean reaction times for repeat (577) than switch (593) trials. Men also showed a significant switch cost, t(9) = -6.02, p = 0, with faster mean reaction times for repeat (591) than switch trials (647). The presence of an interaction indicates that the size of the switch cost for women was significantly smaller than the size of the switch cost for men.