6 Initialize R

Initialize R by entering the following commands at the prompt. You must type the commands exactly as shown.

```r
options(contrasts=c("contr.sum","contr.poly") )
load(url("http://psycserv.mcmaster.ca/bennett/psy710/datasets/faces09.Rdata") )
load(url("http://psycserv.mcmaster.ca/bennett/psy710/datasets/police09.Rdata") )
```

7 faces data

An experiment was done to measure the effects of stimulus inversion and contrast reversal on face identification. On each trial, a single face (i.e., the target) was shown for 500 ms, followed by a pair of faces presented simultaneously, side by side. The subject’s task was to indicate which face was the target by pressing one of two buttons as quickly as possible. The orientation of the faces was either upright or inverted; the contrast of the faces was either positive (i.e., normal) or negative (i.e., like a photographic negative). Face orientation and contrast was crossed factorially to create four stimulus conditions: upright-positive, upright-negative, inverted-positive, inverted-negative. The dependent variable was the mean response time (in milliseconds) on all trials in which the subject responded correctly. Forty subjects were assigned randomly to one of the four conditions, with the constraint that there were 10 subjects per condition. The data are stored in the data frame `faces`, with the variables `rt`, `contrast`, `orient`, and `condition`.

Tasks:

1. Use the data and variables in `faces` to recreate Figure 7. Use this figure to evaluate the assumption that the data within each cell are distributed normally.
2. One assumption made by factorial analysis of variance is that variance is constant across cells. Inspection of Figure 7 suggests that this assumption is valid for the `faces` data. Now do a statistical test to evaluate the constant variance assumption.
3. The following code illustrates how to use the `tapply` command to construct a table that lists the number of observations in each cell in the design. Note that the design is balanced:

```r
# note the list of two factors:
with(faces,tapply(rt,list(contrast,orient),length) )
```

```r
## upright inverted
## positive 10 10
## negative 10 10
```

In the next example, we drop `orient` from the list, to calculate the number of observations at each level of `orientation` ignoring the other factor:
Faces Experiment

- Condition: posUp, posInv, negUp, negInv
- RT: 400, 500, 600, 700
(a) Create tables showing the mean and standard deviation of \( rt \) for each cell in the design.

(b) Calculate the marginal means of \( rt \) for each level of contrast and orient.

4. The following code shows how to use \texttt{aov} and \texttt{lm} to conduct an ANOVA that evaluates the main effects of factors A and B and the A × B interaction. In this example, the dependent variable is \( y \) and the data are stored in the data frame \texttt{myData}:

```r
my.aov.01 <- aov(y~A + B + A:B, data=myData)
anova(my.aov.01) # make anova table, or...
summary(my.aov.01) # make anova table
my.lm.01 <- lm(y~A + B + A:B, data=myData)
anova(my.lm.01)
```

(a) Use the data in \texttt{faces} to do an ANOVA that evaluates the main effects of contrast and orientation, and the contrast × orientation interaction, on the dependent variable \( rt \). Explain what each line in the ANOVA table means.

(b) Verify that the results of the ANOVA do not depend on the order of contrast and orientation in the linear model.

5. The following code constructs two nested linear models and then evaluates the difference in the goodness of fit:

```r
faces.aov.05 <- aov(rt~1+orient,data=faces) # reduced model
faces.aov.06 <- aov(rt~1+orient+contrast,data=faces) # full
anova(faces.aov.05,faces.aov.06) # comparison
```

(a) What does the Sum-of-Sq value in the last row of the output represent?

(b) Compare two nested models that estimate the sum-of-squares associated with contrast while ignoring both the main effect of orientation and the contrast × orientation interaction. How does this value compare to the ones calculated previously?

6. The factorial ANOVA assumes that the observations are independent, that the variance is constant across cells, and that the data within each cell are distributed normally. Evaluate the normality assumption for the \texttt{faces} data.
8 analysis of police data

The data.frame police contains data from a hypothetical 3x3 between-subjects, factorial experiment. A police department conducted an experiment to evaluate its humans relations course for new officers. The independent variables were the type of beat to which officers were assigned during the course (factor beat) and the length of the course (factor course). Each subject was assigned randomly to a single combination of beat and course. The factor beat has three levels: innercity, middleclass, and upperclass. The (unordered) factor course also has three levels: short, medium, and long. The dependent variable is attitude toward minority groups following the course. The data frame also contains a variable, id, which is an id number assigned to each subject, and the factor group.1D, which contains a unique name for each cell in the experimental design.

Tasks:

1. Calculate the mean, standard deviation, and n for each cell.
2. Calculate the marginal means for each level of beat and course.
3. Evaluate the analysis of variance’s constant-variance assumption.
4. Use ANOVA to evaluate the effects of beat and course on attitude. Explain your results.
5. Inspect the group means and take a guess about what the beat × course interaction means. You might want to try R’s interaction.plot command, which is described on page 21 in the notes for Chapter 7.