Datasets for Research Methodology

Dr. Marieke van Vugt

January 21, 2015

1 Decision making in a random dot motion task - DECISION

`decision.dat`

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 saw randomly moving dots, of which a subset moved coherently in left-ward or right-ward directions. The researchers manipulated the difficulty of the decision process by changing the coherence of the dots. The more dots were moving coherently, the easier the task. Every participant was given a difficult and an easy coherence condition, where the coherence was adjusted such that they scored on average approximately 70 and 90 percent correct (but while some participants score 70% correct at a coherence of 10%, others score 70% correct at a coherence of 2%). There was also a control condition in which participants did not need to accumulate information about dots movement, but were shown the correct answer directly by means of a large arrow: when this arrow pointed leftwards they had to press the "left" button, while a rightward arrow indicated a "right" response.

Participants were given $0.01 for every correct response (this experiment

Loading dots task..

Figure 1: Sample dots task. When you look at this file in Acrobat you can press "Loading dots task" to see an animation of the task.

was run in the USA, at Princeton University). Participants performed the task in blocks of 4 minutes, during which time they attempted to do as many trials as possible. In other words, responding faster would allow them to do more trials, and therefore potentially earn more money. However, when they went to fast, they started to make errors, thereby reducing their earnings. This scheme therefore forced the participants to adopt an optimal speed-accuracy trade-off.

In the datafile (a text file that can be read in any text editor), every row is a trial that a particular participant did. The meaning of the columns is as follows:

subjNo	subject ID
ER	error (1) or correct (0) response on this trial)
RT	response time in seconds
isLeft	direction of the coherently moving dots; 1= leftward; 0 = rightward
cohVec	coherence of the moving dots (in %)
cohFac	coherence as a factor: low=0, high=1
blocknum	block number (first block = 1)
isDots	1= dots trials; 0 = control condition, in which the correct answer was indicated by means of a clear arrow. In that case the participant barely needs to accumulate evidence to respond.

You can read the datafile into R with the following commands:

```
# set the working directory to the directory in which your data file resides
setwd("/mydatadir")
# read the data, noticing that it has a header
dat<-read.table("decision.dat",header=T)
```

If you only need the data from a single participant, you can use the following code:

```
bySubj <- split(dat,dat$subjNo)
# get the data from the second subject
subj2 = bySubj[[2]]
```

2 Computer system learns to present the user with interesting information - PREFLEARN-ING

`preflearning.csv`

In this experiment, different algorithms were tested that all offered participants articles to read based on their preferences. The participants then had to indicate for every article whether they found that interesting, and the algorithm
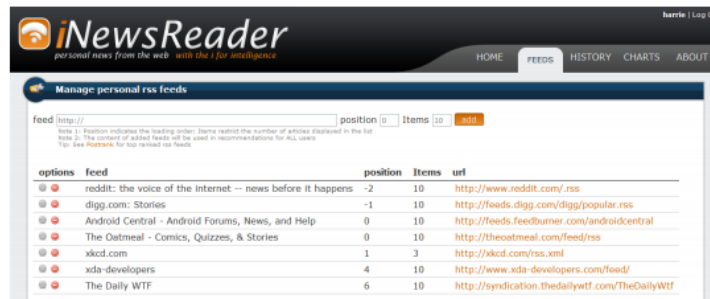
Figure 2: Screenshot of the computer system that learns to offer the user interesting information: an RSS feed of articles.

slowly built up a model of the user's preferences. The dataset consists of the number of "likes" and "dislikes" over time for each of the participants. The dataset also contains the amount of time that a participant spent reading the article. Presumably a participant spends more time reading if s/he likes the article. Participants do not receive feedback on their choices and have no time pressure for making these choices (apart from the fact that probably they don't want to sit there forever...).

Explanation of the different fields in the datafile `preflearning.csv`:

user	participant (who judges the articles)[1]
recommender	not relevant [2]
date	time at which the experiment was done
TotalArticles	number of articles offered
OpenedArticles	number of articles opened by the participant
Likes	number of "likes"
Dislikes	number of "dislikes"
Unvoted	number of articles opened but not judged
ReadingSeconds	amount of time spent reading (in seconds)
ReadingLikeSeconds	amount of time spent reading liked articles
ReadingTime	same as ReadingSeconds but using a different representation of time

You can read in the CSV file with `read.csv`.

Many thanks to Harrie de Groot and Dr. Marco Wiering

3 Scores on reading quizzes - QUIZ

In a previous iteration of Research Methods, students were required to complete a reading quiz about the reading every week. In the datafiles you find the scores on this weekly quiz, for every week a different file. Note that the number of questions differed per week. Students obtained one point per correctly-answered

question, and every row reflects a student. When a student did not do the quiz, this was marked as NA.

    `week1quizGrades.dat`