Bayesian statistics course (Vasishth/Nicenboim)

Date: 25-29 March, 2019

Exercise 1: Bayes Factor hypothesis test for Grodner and Gibson, 2005 data

Consider again the reading time data from the experiment by Grodner and Gibson, 2005. This is the data from their Experiment 1. You can download the paper from [**here**](https://pdfs.semanticscholar.org/98fd/1d9a9191a4e1ae083db538011f333580668b.pdf).

Recall that in this paper, we are interested in the reading time differences between object and subject relatives at the relative clause verb. The expectation from theory is that object relatives (objgap) have longer reading times than subject relatives (subjgap). The explanation for the longer reading times in objgap vs subjgap lies in working memory constraints: it is more difficult to figure out who did what to whom in object relatives than subject relatives.

First, load the data-set provided, and do the preprocessing shown. This gives us the relevant data.

```
library(dplyr)

##
## Attaching package:  'dplyr'
## The following objects are masked from 'package:stats':
##
##     filter, lag
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

gg05e1 <- read.table("data/GrodnerGibson2005E1.csv",sep=",", header=T)
gge1 <- gg05e1 %>% filter(item != 0)

gge1 <- gge1 %>% mutate(word_positionnew = ifelse(item != 15 & word_position > 1
                                                  word_position-1, word_position
#there is a mistake in the coding of word position,
#all items but 15 have regions 10 and higher coded
#as words 11 and higher

## get data from relative clause verb:
gge1crit <- subset(gge1, ( condition == "objgap" & word_position == 6 ) |
```

```
             ( condition == "subjgap" & word_position == 4 ))
gge1crit<-gge1crit[,c(1,2,3,6)]
head(gge1crit)

##    subject item condition rawRT
## 6        1    1    objgap   320
## 19       1    2   subjgap   424
## 34       1    3    objgap   309
## 49       1    4   subjgap   274
## 68       1    5    objgap   333
## 80       1    6   subjgap   266

gge1crit$so<-ifelse(gge1crit$condition=="objgap",1,-1)
```

Use the two brms approaches shown in the slides to do a hypothesis to determine whether there is evidence that object relative clauses take longer to read.

Exercise 2: Chinese Relative clauses (Gisbon and Wu, 2013)

Load the following Chinese RC data we saw in the exercises in 04.01 and subset the relevant data:

```
chineseRC<-read.table("data/gibsonwu2012data.txt")
crit<-subset(chineseRC,region=="headnoun")
crit$region<-factor(crit$region)
head(crit[,c(1,2,3,7)])

##     subj item     type   rt
## 94     1   13  obj-ext 1561
## 221    1    6 subj-ext  959
## 341    1    5  obj-ext  582
## 461    1    9  obj-ext  294
## 621    1   14 subj-ext  438
## 753    1    4 subj-ext  286

crit<-crit[,c(1,2,3,7)]
head(crit)
```

```
##      subj item     type   rt
## 94      1   13  obj-ext 1561
## 221     1    6 subj-ext  959
## 341     1    5  obj-ext  582
## 461     1    9  obj-ext  294
## 621     1   14 subj-ext  438
## 753     1    4 subj-ext  286
```

Use both the brms methods to determine whether *subject* relatives are harder to process than object relatives.

Exercise 3: Chinese Relative clauses replication

Load a replication data-set of the Gibson and Wu data set. We conducted a study that is a direct replication of the original study.

Load the data, and then combine the data from the Gibson and Wu study with our replicaton data:

```
crit_rep<-read.table("data/gibsonwu2012datarepeat.txt")
head(crit_rep)

##     subj item condition pos   rt    region
## 9    1m1   15   obj-ext   8  832 head noun
## 20   1m1    8  subj-ext   8 2131 head noun
## 33   1m1   11   obj-ext   8  553 head noun
## 46   1m1   10  subj-ext   8 1091 head noun
## 62   1m1   16  subj-ext   8  598 head noun
## 75   1m1   14  subj-ext   8  645 head noun

colnames(crit_rep)[3]<-"type"
crit_rep<-crit_rep[,c(1,2,3,5)]
head(crit)

##      subj item     type   rt
## 94      1   13  obj-ext 1561
## 221     1    6 subj-ext  959
## 341     1    5  obj-ext  582
## 461     1    9  obj-ext  294
## 621     1   14 subj-ext  438
## 753     1    4 subj-ext  286
```

```r
head(crit_rep)

##    subj item     type   rt
## 9   1m1   15  obj-ext  832
## 20  1m1    8 subj-ext 2131
## 33  1m1   11  obj-ext  553
## 46  1m1   10 subj-ext 1091
## 62  1m1   16 subj-ext  598
## 75  1m1   14 subj-ext  645

dat<-rbind(crit,crit_rep)
```

Now, using the Bayes Factors approach, test the hypothesis that *subject* relative clauses are harder to process than object relatives.