Skill Improvement 2

Pite

Friday, April 17, 2015

This document will explore how each student would improve in the second day a skill is assigned. The following function is how this analysis is implemented in R.

```
setwd("C:/Users/primavista/Desktop/Flipped Data")
x<-read.csv("output.csv")
library(lubridate)
library(parsedate)
```

```
## Warning: package 'parsedate' was built under R version 3.1.3
```

```
library(arm)
```

```
## Warning: package 'arm' was built under R version 3.1.3
```

```
## Loading required package: MASS
## Loading required package: Matrix
## Loading required package: lme4
```

```
## Warning: package 'lme4' was built under R version 3.1.3
```

```
## Loading required package: Rcpp
##
## arm (Version 1.8-4, built: 2015-04-07)
##
## Working directory is C:/Users/primavista/Desktop/Flipped Data
```

```
x[,5]<-day(as.Date(x[,5],'%d/%m/%y'))
```

We roughly approximate the mean of each day

```
## [1] "Accuracy of day 11 skill 640 is"
```

```
## [1] 0.7824541
```

```
## [1] "Accuracy of day 12 skill 640 is"
```

```
## [1] 0.7781818
```

```
## [1] "Accuracy of day 10 skill 644 is"
```

```
## [1] 0.7693694
```

```
## [1] "Accuracy of day 12 skill 644 is"
```

```
## [1] 0.7711538
```

We still haven't seen much difference. We want to get more information of what is going on with our expected data.

```r
compareDays<-function(day1,day2,skillId,low=0,high=1.0,minAttempt=0,maxAttempt=5){
        y<-x[x$attempt==1,]
        y<-y[y$skill_id==skillId,]
        y<-y[order(y[,5]),]
        index<-((y$created_at)==day1)
        index2<-((y$created_at)==day2)
        firstDay<-y[index,]
        secondDay<-y[index2,]
        meanEach<-vector(length=length(unique(firstDay$student_id)))
        i<-1
        for (student in unique(firstDay$student_id)){
                meanEach[i]=mean(firstDay[firstDay$student_id==student,]$is_correct,na.rm=TRUE)
                i<-i+1
        }
        lowIndex<- (meanEach<=high) & (meanEach>=low)
        meanEach<-meanEach[lowIndex]
        lowScore<-unique(firstDay$student_id)[lowIndex]
        difference<-vector(length=length(unique(lowScore)))
        firstDayAttempt<-vector(length=length(unique(lowScore)))
        i<-1
        for (student in lowScore){
                firstDayAttempt[i]<-sum(firstDay$student_id==student)
                if (sum(secondDay$student_id==student)==0){
                        difference[i]<-NA}
                else if (sum(secondDay$student_id==student)<minAttempt |sum(secondDay$student_id==studen
                        difference[i]<-NA}
                else{
                        difference[i]<-mean(secondDay[secondDay$student_id==student,]$is_correct)-meanEa
                }
                i<-(i+ 1)
        }
        exclude<-is.na(difference)
        difference<-difference[!exclude]
        firstDayAttempt<-firstDayAttempt[!exclude]
        print('mean of the number of each student\'s attempts on the first day')
        print(mean(firstDayAttempt))
        print('Total number of students included in analysis')
        print(length(firstDayAttempt))
        par(mfrow=c(1,2))
        hist(subset(meanEach,!exclude),main="First Day",breaks=(c(0,0.2,0.4,0.6,0.8,1,1.2)-0.01),xlab='
        hist(subset(meanEach,!exclude)+difference,main="Second Day",breaks=(c(0,0.2,0.4,0.6,0.8,1,1.2)-0
        print(t.test(subset(meanEach,!exclude)+difference,subset(meanEach,!exclude),paired=TRUE))
}
```

Basically, what this code does is to choose the data with the following criteria * exclude all attempts that is repeated * choose those attempts on the problem with skill_id as skillId(specify in the argument) * exclude

students whose the number of attempts on the second day is not in $[minAttempt, maxAttempt]$ and those that does not have any attempt on the second day

Then, we plot a histrogram for each day and perform a t.test (with paired observations) to see if the difference between the two groups is statistically significant

Now, we will analyse the mean of each student from each day and compare them(with no restriction on first day's accuracy and the number of attempts on second day.)

```
compareDays(day1=11,day2=12,skillId=640)
```

```
## [1] "mean of the number of each student's attempts on the first day"
## [1] 4.945833
## [1] "Total number of students included in analysis"
## [1] 240
```

First Day   Second Day
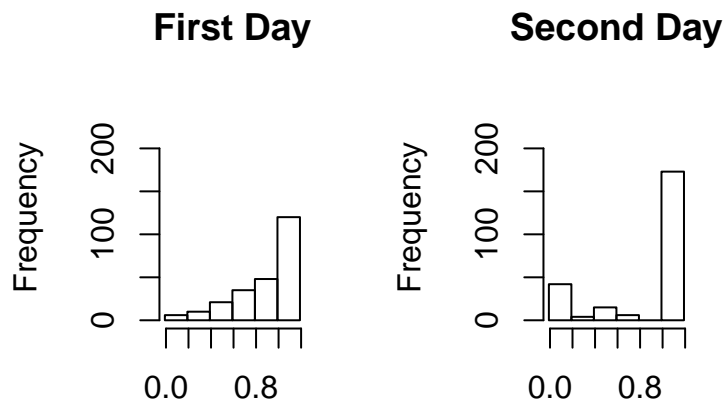

```
##
##   Paired t-test
##
## data:  subset(meanEach, !exclude) + difference and subset(meanEach, !exclude)
## t = -0.6354, df = 239, p-value = 0.5258
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.06776834  0.03471278
## sample estimates:
## mean of the differences
##              -0.01652778
```

```
#for now we just focus on the pattern from day 11,12
```

The difference is not so significant as the confident interval contains 0. This could be due to the fact that there is quite a small(typically 2-3 questions/person) number of questions from skill 640 assigned to each student which makes the rate of accuracy on the second day unstable(for the first day majority of students tend to do all 5). Also, there are only 2 students with at least 4 attempts on the second day. So, we will only

select students that have at least two attempts on the second day.

```
compareDays(day1=11,day2=12,skillId=640,minAttempt=2)
```

```
## [1] "mean of the number of each student's attempts on the first day"
## [1] 4.887931
## [1] "Total number of students included in analysis"
## [1] 116
```
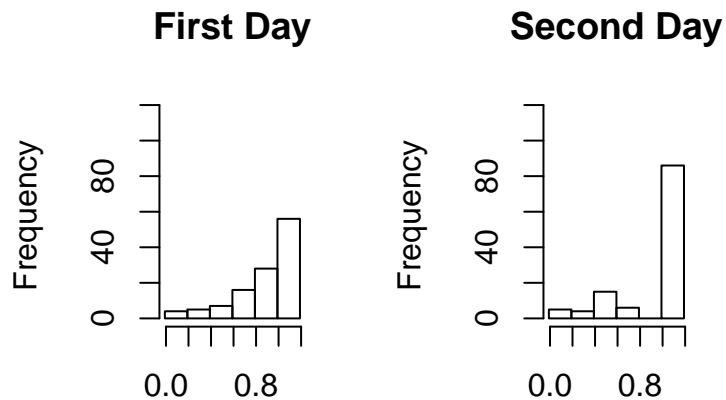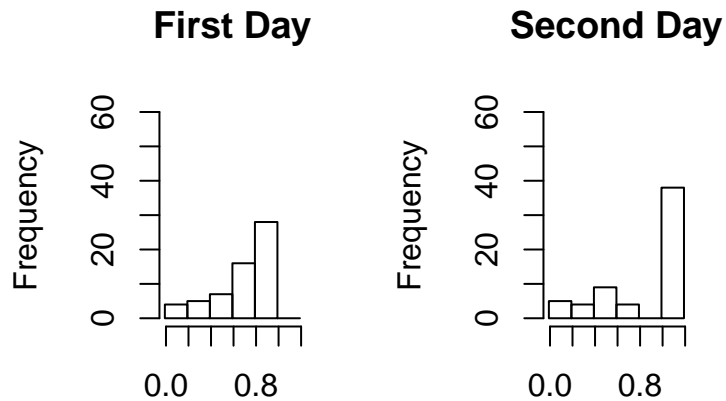


```
##
##  Paired t-test
##
## data:  subset(meanEach, !exclude) + difference and subset(meanEach, !exclude)
## t = 2.1145, df = 115, p-value = 0.03663
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.003834837 0.117429531
## sample estimates:
## mean of the differences
##              0.06063218
```

```
#for now we just focus on the pattern from day 11,12
```

This turns out to have quite a statistical significant with the estimate of 6% increase in accuracy but if we take minAttempt=3 the significant will disappear. We could argue that this is due to the fact that there are only 29 students used in our analysis. Then, we want to see if there is a different effect on students with different score on the first day. We will analyse in two groups, ($<=80$%) and (>80% i.e. 100%).

```
compareDays(day1=11,day2=12,skillId=640,high=0.8,minAttempt=2)
```

```
## [1] "mean of the number of each student's attempts on the first day"
## [1] 4.866667
## [1] "Total number of students included in analysis"
## [1] 60
```

First Day

Second Day


```
##
##  Paired t-test
##
## data:  subset(meanEach, !exclude) + difference and subset(meanEach, !exclude)
## t = 3.8079, df = 59, p-value = 0.000336
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.08462272 0.27204394
## sample estimates:
## mean of the differences
##               0.1783333
```

```
#for now we just focus on the pattern from day 11,12
```

We can see, from the histrograms, that in this case the rate of accuracy tends to increase for people with accuracy $<= 80\%$ in the first day. Now we can see what happens to those with 100% accuracy on the first

day

```r
compareDays(day1=11,day2=12,skillId=640,low=0.81,minAttempt=2)
```

```
## [1] "mean of the number of each student's attempts on the first day"
## [1] 4.910714
## [1] "Total number of students included in analysis"
## [1] 56
```

First Day

Second Day
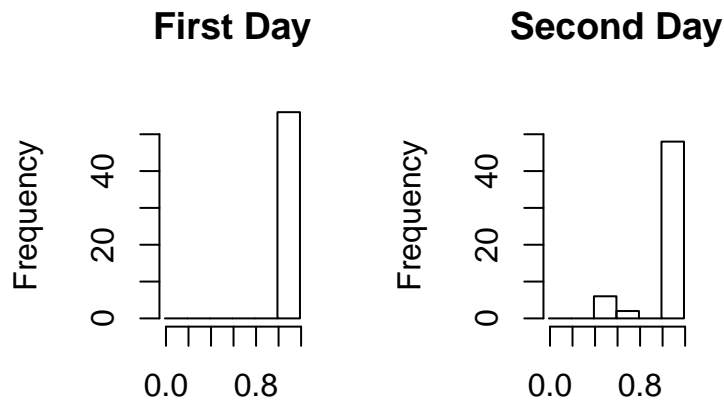

```
##
##  Paired t-test
##
## data:  subset(meanEach, !exclude) + difference and subset(meanEach, !exclude)
## t = -2.9848, df = 55, p-value = 0.004228
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   -0.10943816 -0.02151423
## sample estimates:
## mean of the differences
##              -0.06547619
```

```r
#for now we just focus on the pattern from day 11,12
```

We can see that there is a significant decline. Judging from the histrogram, the decrease in accuracy is caused by those who made one mistake on the second day.

The problem in this analysis is due to the fact that there is quite a small attempts(for each student) on the second day. Alternatively, instead of taking scores for each individuals, we subset only students that is active on both days. The main assumption of this analysis is that two attempts from the same day would be independent and the same in distribution with certain probability.

```
##
##  Exact binomial test
```

```
##
## data:  sum(score2) and length(score2)
## number of successes = 309, number of trials = 387, p-value =
## 0.9498
## alternative hypothesis: true probability of success is not equal to 0.7952822
## 95 percent confidence interval:
##  0.7549703 0.8372945
## sample estimates:
## probability of success
##               0.7984496
```

It turns out that there is not so much significant effect on the matter. This accepts the hypothesis that two days have the same mean with the p-value=0.893(normally reject at p=0.05). It might also be concluded that the rate of accuracy does not decrease even if two skills is assinged on the second day(arguably make the quest harder).

We can also try this with (day1=25,day2=8,skill=630) and (day1=10,day2=12,skill=640)