Homework 4

Ethen Liu, Sai Haran, Sophia Hoffman, Annie Didier, Arindam Bhattacharya

February 1, 2017

Problem 1

```
library(psych)
music<-read.csv("music.csv")
fit1<-factanal(music[,-28],factors=2,score="reg",rotation="varimax")
fit1$loadings # 31, 47 maybe, 49 maybe, 53, 54
```

```
##
## Loadings:
##     Factor1 Factor2
## V28 0.468   0.691
## V29 0.420   0.674
## V30 0.372   0.685
## V31 0.389   0.481
## V32 0.301   0.788
## V33 0.197   0.772
## V34 0.178   0.754
## V35 0.709   0.314
## V36 0.770   0.376
## V37 0.741   0.386
## V38 0.791   0.335
## V39 0.768   0.345
## V40 0.797   0.218
## V41 0.803   0.310
## V42 0.655   0.427
## V43 0.178   0.785
## V44 0.730   0.161
## V45 0.643   0.221
## V46 0.187   0.781
## V47 0.401   0.586
## V48 0.643   0.347
## V49 0.514   0.326
## V50 0.337   0.629
## V51 0.738   0.170
## V52 0.415   0.641
## V53 0.570   0.488
## V54 0.479   0.456
##
##                Factor1 Factor2
## SS loadings       8.65   7.517
## Proportion Var    0.32   0.278
## Cumulative Var    0.32   0.599
```

```
music_alpha <- alpha(music[,-28],check.keys=TRUE) #31, 34, 43, 45, 49
print(music_alpha$alpha.drop)
```

```
##     raw_alpha std.alpha   G6(smc) average_r      S/N    alpha se
```

```
## V28 0.9631798 0.9637899 0.9727955 0.5058592 26.61658 0.001500850
## V29 0.9635396 0.9641297 0.9731183 0.5083041 26.87822 0.001485071
## V30 0.9637488 0.9643465 0.9733998 0.5098754 27.04774 0.001477279
## V31 0.9647507 0.9653253 0.9743759 0.5170829 27.83946 0.001437815
## V32 0.9636538 0.9642789 0.9731113 0.5093849 26.99470 0.001479384
## V33 0.9643837 0.9649624 0.9736212 0.5143886 27.54075 0.001448469
## V34 0.9646682 0.9651532 0.9735325 0.5158020 27.69704 0.001436195
## V35 0.9639520 0.9644518 0.9731699 0.5106419 27.13083 0.001466196
## V36 0.9633445 0.9637926 0.9728749 0.5058785 26.61864 0.001490767
## V37 0.9633670 0.9638845 0.9730121 0.5065380 26.68896 0.001490885
## V38 0.9634592 0.9639247 0.9728420 0.5068265 26.71979 0.001485672
## V39 0.9635338 0.9640027 0.9730166 0.5073878 26.77985 0.001483061
## V40 0.9638987 0.9643865 0.9730473 0.5101665 27.07926 0.001465996
## V41 0.9635283 0.9639648 0.9728960 0.5071148 26.75062 0.001482558
## V42 0.9636802 0.9641834 0.9732045 0.5086922 26.91999 0.001478440
## V43 0.9644278 0.9650300 0.9735629 0.5148887 27.59594 0.001446769
## V44 0.9644730 0.9649841 0.9733450 0.5145489 27.55843 0.001444685
## V45 0.9646251 0.9651449 0.9737383 0.5157402 27.69018 0.001438919
## V46 0.9643636 0.9649308 0.9734128 0.5141551 27.51502 0.001448756
## V47 0.9641167 0.9646677 0.9737198 0.5122199 27.30271 0.001461919
## V48 0.9640553 0.9645557 0.9734032 0.5114002 27.21328 0.001463201
## V49 0.9648153 0.9652879 0.9739868 0.5168044 27.80844 0.001432201
## V50 0.9641935 0.9647842 0.9738411 0.5130752 27.39633 0.001458407
## V51 0.9644495 0.9649533 0.9737558 0.5143212 27.53333 0.001446579
## V52 0.9637542 0.9643265 0.9734354 0.5097300 27.03200 0.001476175
## V53 0.9636827 0.9642422 0.9734386 0.5091181 26.96590 0.001480094
## V54 0.9643143 0.9648644 0.9740283 0.5136655 27.46115 0.001453894
```

Factor analysis and alpha suggests removing 31 and 49, also removing 54 because of promiscuity.

```
# compute alpha for each scale
music1<-music[,c(1,2,3,5,6,7,16,19,20,23,25)]
music2<-music[,c(8,9,10,11,12,13,14,15,17,18,21,24)]
music1_alpha <- alpha(music1)
music2_alpha <- alpha(music2)
print(music1_alpha$alpha.drop)
```

```
##     raw_alpha std.alpha   G6(smc) average_r      S/N    alpha se
## V28 0.9349209 0.9351665 0.9353628 0.5905687 14.42412 0.002715694
## V29 0.9363577 0.9365907 0.9365737 0.5962950 14.77056 0.002656752
## V30 0.9370664 0.9373773 0.9387311 0.5994978 14.96865 0.002631412
## V32 0.9345514 0.9349838 0.9365217 0.5898408 14.38078 0.002740874
## V33 0.9368711 0.9373505 0.9383819 0.5993882 14.96182 0.002641841
## V34 0.9377961 0.9381290 0.9380157 0.6025857 15.16266 0.002600099
## V43 0.9368607 0.9372941 0.9385137 0.5991575 14.94746 0.002642500
## V46 0.9361788 0.9366997 0.9372050 0.5967371 14.79772 0.002671385
## V47 0.9395518 0.9398549 0.9407883 0.6097784 15.62646 0.002527444
## V50 0.9389489 0.9393634 0.9410843 0.6077153 15.49169 0.002557037
## V52 0.9378351 0.9381044 0.9394506 0.6024841 15.15623 0.002602583
```

```
print(music2_alpha$alpha.drop)
```

```
##     raw_alpha std.alpha   G6(smc) average_r      S/N    alpha se
## V35 0.9487241 0.9487833 0.9519246 0.6274328 18.52487 0.002134138
## V36 0.9463672 0.9464578 0.9506805 0.6164155 17.67686 0.002235301
## V37 0.9474630 0.9475803 0.9521443 0.6216915 18.07680 0.002194022
```

```
## V38 0.9463895 0.9465203 0.9505579 0.6167070 17.69867 0.002235036
## V39 0.9470139 0.9471265 0.9512973 0.6195490 17.91305 0.002208672
## V40 0.9477682 0.9478618 0.9518864 0.6230267 18.17978 0.002175424
## V41 0.9463163 0.9463572 0.9503734 0.6159462 17.64182 0.002236369
## V42 0.9494505 0.9495542 0.9532830 0.6311605 18.82327 0.002107415
## V44 0.9499164 0.9500520 0.9520426 0.6335877 19.02083 0.002078028
## V45 0.9521089 0.9520336 0.9543282 0.6434118 19.84791 0.001985027
## V48 0.9502372 0.9503641 0.9536713 0.6351177 19.14671 0.002074305
## V51 0.9497774 0.9498891 0.9540126 0.6327913 18.95572 0.002093699
```

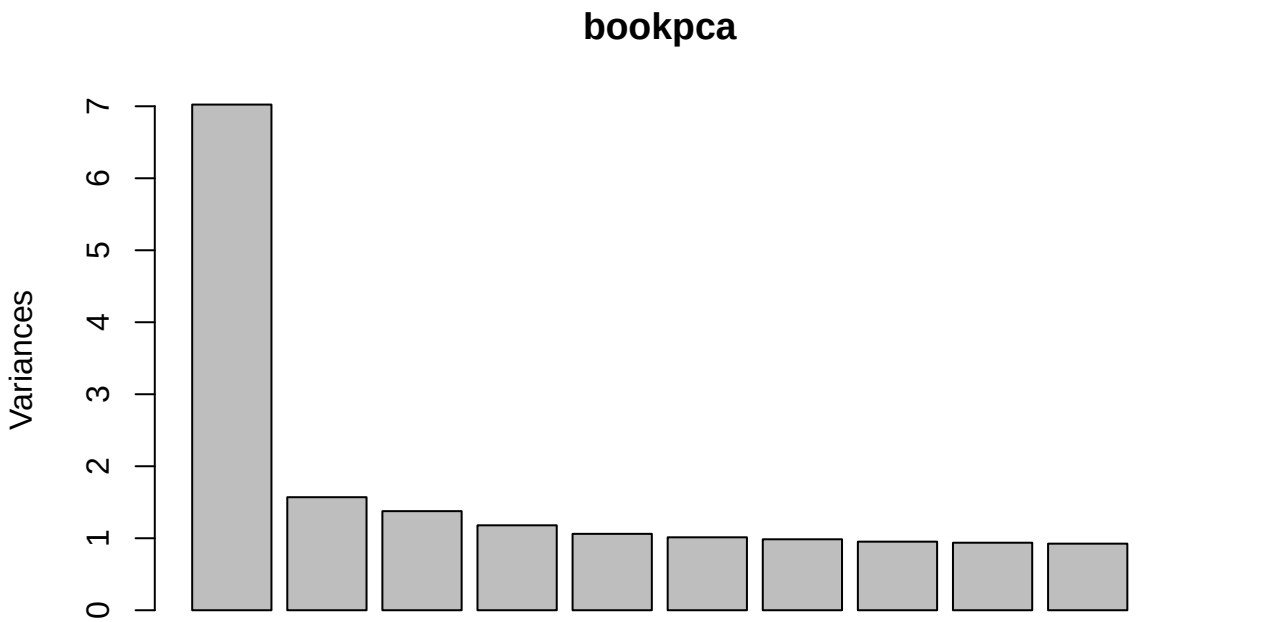
```
fit2<-factanal(music[,-c(4,22,27,28)],factors=2,score="reg",rotation="varimax")
fit2loadings<-fit2$loadings
fit2loadings[fit2loadings<.3]<-0
fit2loadings
```

```
##
## Loadings:
##     Factor1 Factor2
## V28 0.473   0.688
## V29 0.423   0.671
## V30 0.373   0.681
## V32 0.308   0.783
## V33         0.771
## V34         0.756
## V35 0.715   0.309
## V36 0.776   0.371
## V37 0.745   0.382
## V38 0.800   0.328
## V39 0.776   0.340
## V40 0.789
## V41 0.808   0.305
## V42 0.662   0.422
## V43         0.783
## V44 0.722
## V45 0.634
## V46         0.782
## V47 0.404   0.586
## V48 0.646   0.343
## V50 0.336   0.628
## V51 0.734
## V52 0.419   0.636
## V53 0.570   0.483
##
##                Factor1 Factor2
## SS loadings      7.917   6.755
## Proportion Var   0.330   0.281
## Cumulative Var   0.330   0.611
```

- Factor 1 is social aspects of music. Includes questions 28, 29, 30, 32, 33, 34, 43, 46, 47, 50, 52
- Factor 2 is personal aspects. Includes questions 35, 36, 37, 38, 39, 40, 41, 42, 44, 45, 48, 51

Problem 2

```
book<-read.csv("book.csv")
book<-log(book+1)
bookpca<-prcomp(book,scale=TRUE)
plot(bookpca) # let's say 3 factors
```




```
fit3<-factanal(book,factors=3,score="reg",rotation="varimax")
fit3loadings<-fit3$loadings
fit3loadings[fit3loadings<.3]<-0
# fiction, legends, philosophy, religion*, psychology*,
# music*, facsimilie, politics, maps, gamesriddles, sports
# videos, nonbooks
fit3loadings
```

```
##
## Loadings:
##                Factor1 Factor2 Factor3
## fiction1       0.353   0.414   0.328
## classics3              0.429
## cartoons5      0.363
## legends6
## philosophy7    0.428   0.440
## religion8      0.398   0.518
## psychology9    0.312
## linguistics10  0.425   0.308
## art12                  0.443   0.359
## music14                        0.318
## facsimile17
## history19              0.713   0.322
## conthist20             0.509
## economy21      0.466
## politics22
## science23      0.471   0.318
```

```
## compsci26       0.391
## railroads27                     0.347
## maps30
## travelguides31          0.380   0.541
## health35        0.606           0.336
## cooking36                       0.476
## learning37      0.597
## GamesRiddles38
## sports39
## hobby40         0.349           0.578
## nature41                        0.439
## encyclopaedia44 0.436
## videos50
## nonbooks99
##
##                Factor1 Factor2 Factor3
## SS loadings      2.503   2.121   1.720
## Proportion Var   0.083   0.071   0.057
## Cumulative Var   0.083   0.154   0.211
```

```
exclude<-c("fiction1", "legends6", "philosophy7", "facsimile17", "politics22", "maps30",
           "GamesRiddles38", "sports39", "videos50", "nonbooks99","cartoons5", "railroads27")
fit4<-factanal(book[,!names(book)%in%exclude],factors=3,score="reg",rotation="varimax")
fit4$loadings
```

```
##
## Loadings:
##                 Factor1 Factor2 Factor3
## classics3       0.173   0.365   0.195
## religion8       0.354   0.470   0.156
## psychology9     0.287   0.112
## linguistics10   0.453   0.299
## art12           0.104   0.451   0.361
## music14         0.226   0.272   0.287
## history19       0.161   0.808   0.217
## conthist20      0.153   0.568   0.102
## economy21       0.474           0.193
## science23       0.463   0.305   0.192
## compsci26       0.410           0.168
## travelguides31  0.162   0.429   0.490
## health35        0.559   0.109   0.393
## cooking36       0.188   0.117   0.520
## learning37      0.610           0.188
## hobby40         0.302   0.243   0.593
## nature41        0.219   0.179   0.498
## encyclopaedia44 0.453   0.307   0.186
##
##                Factor1 Factor2 Factor3
## SS loadings      2.252   2.223   1.777
## Proportion Var   0.125   0.123   0.099
## Cumulative Var   0.125   0.249   0.347
```

- factor 1 is sciences
- factor 2 is arts/humanities
- factor 3 is hobbies