

MS – E 2112 – Multivariate statistical analysis – Home exercise 7

Srikanth Gadicherla , 546195

March 4th, 2016

Problem 1

Data: The data set consists of 871 individual responses towards science and environment question. We have four variables here with each one having 5 modalities ranging from strongly agree to strongly disagree encoded as 1-5.

a) The complete disjunctive table is formed by having the individual responses as the number of rows and splitting the columns of each variables into their corresponding modalities, so, the columns increase as number of variables times their corresponding modalities.

The complete disjunctive table is for first three rows for the data is given below.

	Believe Science Not Faith/feelings					Modern Science Destruction					Change Damage Nature					Modern Science Side Problems					Sex		Education level						Age					
	SA	A	N	D	SD	SA	A	N	D	SD	SA	A	N	D	SD	SA	A	N	D	SD	M	F	1	2	3	4	5	6	1	2	3	4	5	6
1.	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0
2.	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0
3.	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0

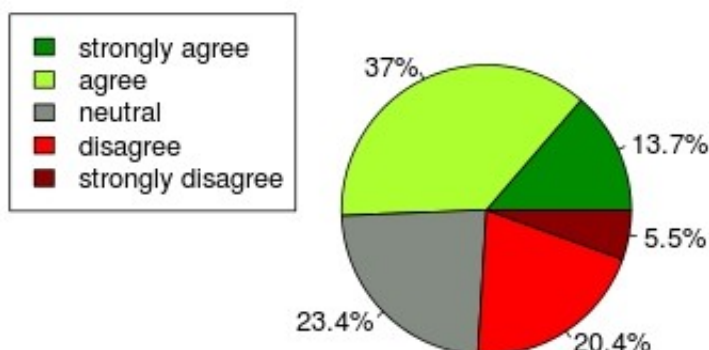
SA – Strongly agree
 A – Agree
 N – Neutral
 D – Disagree
 SD – Strongly Disagree

M – Male
 F – Female

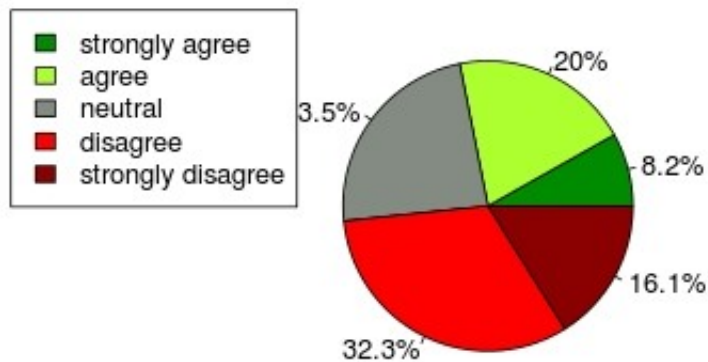
The data also consist of sex, education level and age of the individual. The sex had two modalities i.e., male and female. The education and age had six modalities (details not given in the question paper or mycourses)

b) The multiple correspondence analysis was performed on the data and the plots are given below.

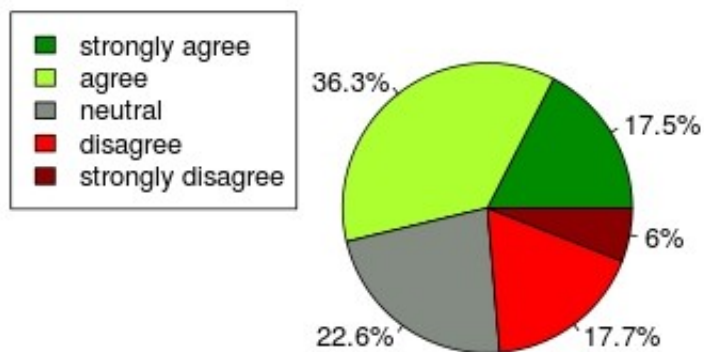
Believe science often but not in faith/feelings



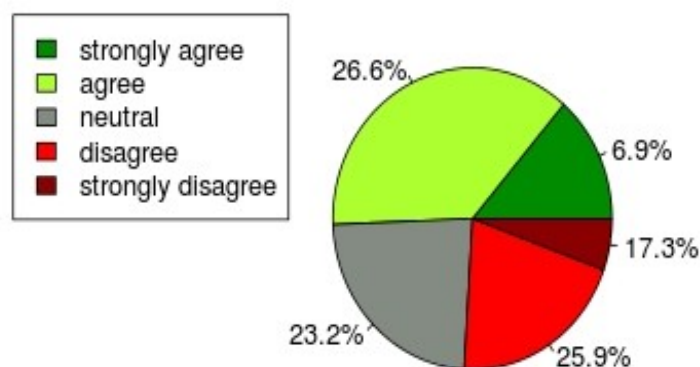
Overall modern science does more harm than good



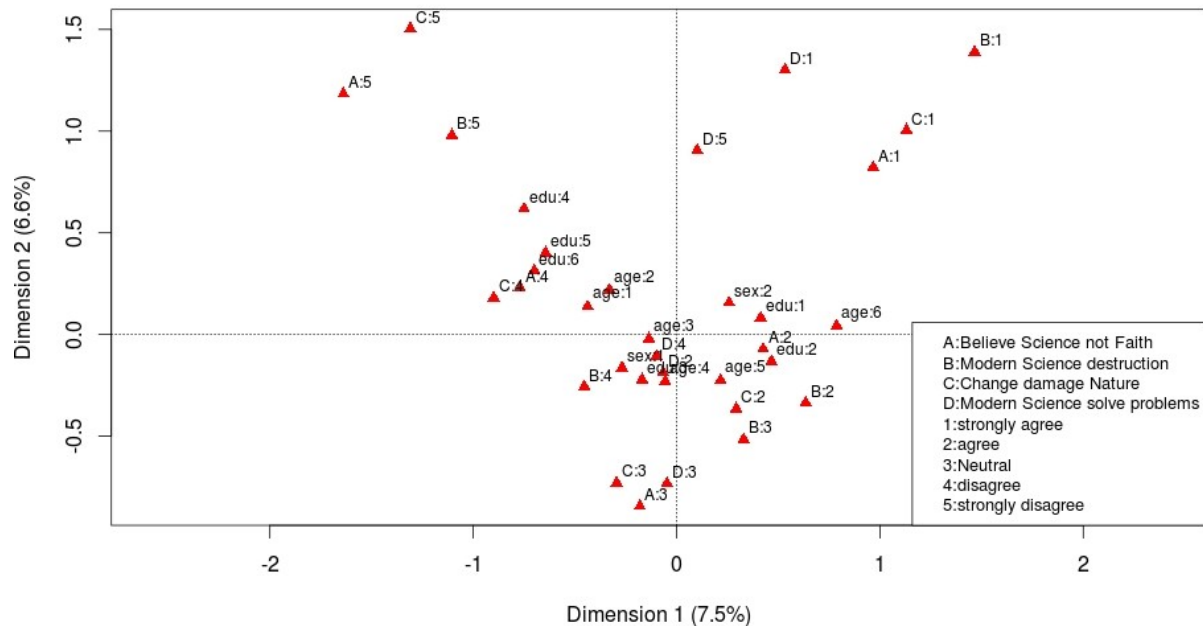
Any change to nature will make things worse



Modern science will solve environment problems with no change



Above plots give the sense of percentage of responses given by individual for each variables. Their modalities with the percentages have been given.



Above is the MCA plot for the data given. The interpretation is as follows.

1. People who strongly believe science but not faith/feelings and also people who strongly disagree human change to nature makes things worse are rare events. You can see from plot "B:1" and "C:5" are farthest away.
2. One interesting observation is (A:1,C:1), (A:2,C:2), (A:3,C:3), (A:4,C:4) and (A:5,C:5) occur together. This means that people who voted for particular option in question A, voted for the same option in question C.
3. The two sex's are diagonally opposite have opposite views. Sex:1 (may be female; not mentioned in question) have views that feeling/faith matter and are in general are in of favour science(See B:4, D:2) While Sex:2 (men) are pessimistic about science(See C:1, B:1) but at the same time are give more preference to science over emotions.
4. edu:4,edu:5, edu:6 (I guess people with high school level education or more), age:1, age:2 (I guess people above 25) are of the opinion that Science in the cause for destruction and there is more than what meets the eye(See A:5, B:5, C:5). The views of these people are completely in perpendicular direction (uncorrelated) to those who believe science is the cause for damage (See A:1, B:1, C:1)
5. age:6, age:3 (may be kids) have no idea of why the comparison is done between science and environment.
6. age:5 (may be in primary school) have mixed views on science as destruction (See B:2, C:2 occurring with B:3)
7. The above analysis also shows that as one grows in education, the opinion on the this matter changes drastically. See edu:1, edu:2, edu:3 and (edu:4,edu:5, edu:6) are all in different directions.

Appendix

The code for the problem solved above.

```
#####
```

```
setwd("//home.org.aalto.fi/gadichs1/data/Desktop/P3_4/MSA/Ex 7")
```

```
data <- read.table('wg93_full1.txt',header=T,sep='\t')
```

```
reponse <- data[,c(1,2,3,4)]
```

```
tableA <- table(data$A)
```

```
tableB <- table(data$B)
```

```
tableC <- table(data$C)
```

```
tableD <- table(data$D)
```

```
labA <- round(100*tableA/sum(tableA),1)
```

```
labB <- round(100*tableB/sum(tableB),1)
```

```
labC <- round(100*tableC/sum(tableC),1)
```

```
labD <- round(100*tableD/sum(tableD),1)
```

```
pielabelsA <- paste(labA, "%", sep="")
```

```
pielabelsB <- paste(labB, "%", sep="")
```

```
pielabelsC <- paste(labC, "%", sep="")
```

```
pielabelsD <- paste(labD, "%", sep="")
```

```
cols <- c("green4", "greenyellow", "honeydew4", "red", "red4")
```

```
names_response <- c("strongly agree", "agree", "Neutral", "disagree", "strongly  
disagree")
```

```
pie(tableA, main="Believe science often but not in  
faith/feelings", col=cols, labels=pielabelsA, cex=0.8)
```

```
legend("topleft", names_response, cex=0.8, fill=cols)
```

```
pie(tableB, main="Overall modern science does more harm than  
good", col=cols, labels=pielabelsB, cex=0.8)
```

```
legend("topleft", names_response, cex=0.8, fill=cols)
```

```
pie(tableC, main="Any change to nature will make things  
worse", col=cols, labels=pielabelsC, cex=0.8)
```

```
legend("topleft", names_response, cex=0.8, fill=cols)
```

```
pie(tableA,main="Modern science will solve environment problems with no  
change",col=cols,labels=pielabelsD,cex=0.8)
```

```
library(ca)
```

```
data.mca <- mjca(data, lambda="indicator")
```

```
((data.mca$sv[1]^2 + data.mca$sv[2]^2) / sum(data.mca$sv^2))
```

```
cats <- apply(data,2,function(x) nlevels(as.factor(x)))
```