

Attitudes towards abortion in the USA

The data are taken from Christensen (1990) and concern a study of attitudes towards legalizing abortion in the USA. The variables are as follows

Age: 18–25, 26–35, 36–45, 46–55, 56–65, > 66.

Race: White, non-white.

Sex: Male, Female.

Opinion: Supports legalized abortion? no, yes, undecided.

After starting MIM we load `mimR` and `Rgraphviz`.

```
> library(mimR)
> library(Rgraphviz)
```

Next we read data from a file and create a `gmData` object.

```
> counts <- as.numeric(unlist(read.table("abortion.txt")))
> abortion <- momentstats(factor = c("Age", "Race", "Sex", "Opinion"),
+   level = c(6, 2, 2, 3), counts = counts)
> gab <- as.gmData(abortion)
> gab
```

```
      name letter factor levels
1     Age      a   TRUE      6
2     Race      b   TRUE      2
3     Sex       c   TRUE      2
4 Opinion      d   TRUE      3
Data origin :      momentstats
```

Before we proceed, we declare the variable `Age` as being ordinal. This may or may not be adequate, but by doing this, the possibility of using ordinal tests is available.

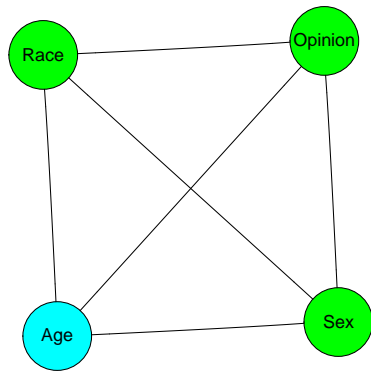
Two of the variables are clearly ordinal. This is declared as follows

```
> ordinal(gab) <- "Age"
> gab
```

```
      name letter factor levels
1     Age      a   TRUE      6
2     Race      b   TRUE      2
3     Sex       c   TRUE      2
4 Opinion      d   TRUE      3
Ordinal      :      Age
Data origin :      momentstats
```

We specify the saturated model as a MIM-model and display it

```
> msat <- mim("../", data = gab)
> display(msat)
```



Next, we try to simplify the model by independence tests for all pairs of variables.

```
> stepwise(msat, arg = "o")
```

Coherent Backward Single-step Selection.

Fixed edges: none.

Critical value: 0.0500

Decomposable mode, Chi-squared tests.

DFs adjusted for sparsity.

Model: abcd

Deviance: 0.0000 DF: 0 P: 1.0000

Edge	Test	Statistic	DF	P
[ab]		22.0885	30	0.8508
[ac]		11.2889	30	0.9992
[ad]		78.0580	40	0.0003 +
[bc]		20.4451	18	0.3083
[bd]		38.2164	24	0.0329 +
[cd]		27.9063	24	0.2640

Formula: Age:Race:Sex:Opinion//

-2logL: 17112.16 DF: 0

Since we may be worried about the validity of the asymptotics, we also use Monte-Carlo tests:

```
> stepwise(msat, arg = "om")
```

Coherent Backward Single-step Selection.

Fixed edges: none.

Critical value: 0.0500

Decomposable mode, Chi-squared tests.

Exact tests, Monte Carlo sampling.

DFs adjusted for sparsity.

Model: abcd

Deviance: 0.0000 DF: 0 P: 1.0000

Edge	Test
------	------

Excluded	Statistic	DF	P
[ab]	22.0885	30	0.9040
[ac]	11.2889	30	0.9990
[ad]	78.0580	40	0.0028 +
[bc]	20.4451	18	0.4450
[bd]	38.2164	24	0.0480 +
[cd]	27.9063	24	0.3770

Formula: Age:Race:Sex:Opinion//

-2logL: 17112.16 DF: 0

In this case, it does seem to matter, so we keep using the Monte-Carlo option and try to make a full stepwise search.

```
> mstep <- stepwise(msat, arg = "m")
```

Coherent Backward Selection.

Fixed edges: none.

Critical value: 0.0500

Decomposable mode, Chi-squared tests.

Exact tests, Monte Carlo sampling.

DFs adjusted for sparsity.

Model: abcd

Deviance: 0.0000 DF: 0 P: 1.0000

Edge	Test		
Excluded	Statistic	DF	P
[ab]	22.0885	30	0.8910
[ac]	11.2889	30	1.0000
[ad]	78.0580	40	0.0000 +
[bc]	20.4451	18	0.4240
[bd]	38.2164	24	0.0444 +
[cd]	27.9063	24	0.4100

Removed edge [ac]

Model: bcd,abd

Deviance: 11.2889 DF: 30 P: 0.9992

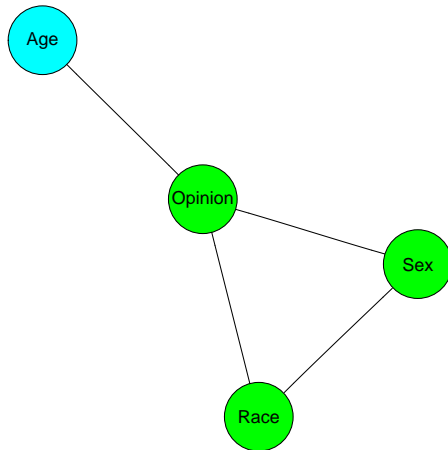
Edge	Test		
Excluded	Statistic	DF	P
[ab]	13.4770	15	0.6260
[bc]	11.8337	3	0.0151 +
[cd]	19.0061	4	0.0020 +

Removed edge [ab]

Selected model: bcd,ad

The selected model has Age independent of Race and Sex for given opinion.

```
> display(mstep)
```



This analysis did not exploit that some of the variables were ordinal. If we repeat the same analysis, but use the appropriate ordinal test, we get a different final model:

```
> mordstep <- stepwise(msat, arg = "mw")
```

Coherent Backward Selection.

Fixed edges: none.

Critical value: 0.0500

Decomposable mode, Chi-squared tests.

Exact tests, Monte Carlo sampling.

DFs adjusted for sparsity.

Model: abcd

Deviance: 0.0000 DF: 0 P: 1.0000

Edge	Test		
Excluded	Statistic	DF	P
[ba]	667365.0000	30	0.0604
[ca]	509529.5000	30	0.7500
[da]	67.0754	8	0.0000 +
[bc]	20.4451	18	0.4428
[bd]	38.2164	24	0.0448 +
[cd]	27.9063	24	0.3942

Removed edge [ca]

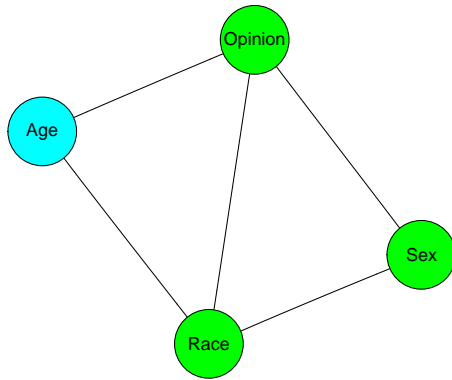
Model: bcd,abd

Deviance: 11.2889 DF: 30 P: 0.9992

Edge	Test		
Excluded	Statistic	DF	P
[ba]	1312637.0000	15	0.0376 +
[bc]	11.8337	3	0.0122 +
[cd]	19.0061	4	0.0000 +

Selected model: bcd,abd

```
> display(mordstep)
```



So, if ordinality is taken into account, a slight dependency between Age and Race seems to be noticeable.

This is not different if we use asymptotic tests.

```
> mordstepas <- stepwise(msat, arg = "w")
```

Coherent Backward Selection.

Fixed edges: none.

Critical value: 0.0500

Decomposable mode, Chi-squared tests.

DFs adjusted for sparsity.

Model: abcd

Deviance: 0.0000 DF: 0 P: 1.0000

Edge	Test		
Excluded	Statistic	DF	P
[ba]	667365.0000	30	0.0638
[ca]	509529.5000	30	0.7507
[da]	67.0754	8	0.0000 +
[bc]	20.4451	18	0.3083
[bd]	38.2164	24	0.0329 +
[cd]	27.9063	24	0.2640

Removed edge [ca]

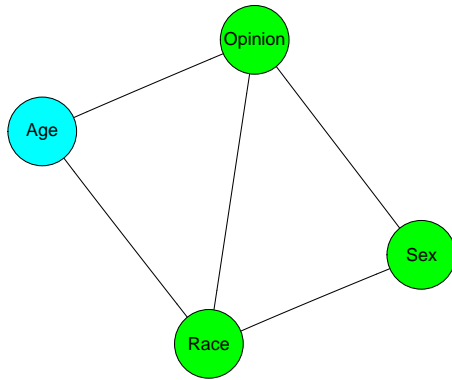
Model: bcd,abd

Deviance: 11.2889 DF: 30 P: 0.9992

Edge	Test		
Excluded	Statistic	DF	P
[ba]	1312637.0000	15	0.0332 +
[bc]	11.8337	3	0.0080 +
[cd]	19.0061	4	0.0008 +

Selected model: bcd,abd

```
> display(mordstepas)
```



To be careful, we proceed to consider the model in mordstep and look at the associated three-way marginal tables and models without second-order interactions.

```

> marg1 <- mim("...", data = gab, marginal = c("Age", "Race", "Opinion"))
> marg12factor <- editmim(marg1, deleteTerm = "Age:Race:Opinion")
> modelTest(marg1, marg12factor)

```

```

Test of H0 : Age:Race:Opinion//
Against    : Race:Opinion + Age:Opinion + Age:Race//

```

```

test: Chi-squared method: asymptotic
stat: 6.506 df: 10 P: 0.771

```

The two-factor model has a fine fit. If we test conditional independence in this model, we get

```

> testdelete("Age:Race", marg12factor)

```

```

test: Chi-squared method: asymptotic
stat: 6.971 df: 5 P: 0.223

```

In this two-factor model we cannot take the ordinality into account and hence do not properly discover the relationship between Age and Race for fixed opinion.

To elaborate further we form the residuals

```

> observed1 <- fitted(marg1)[, 4]
> expected1 <- fitted(marg12factor)[, 4]
> residuals1 <- (observed1 - expected1)/sqrt(expected1)
> residuals1 <- cbind(fitted(marg1)[, 1:3], residuals1)

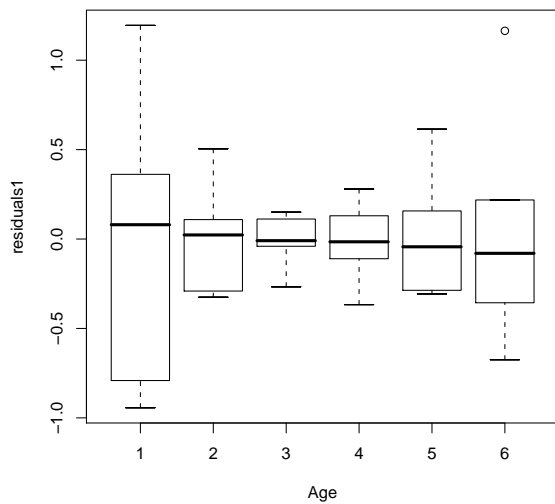
```

and plot them for different age groups

```

> plot(residuals1 ~ Age, data = residuals1)

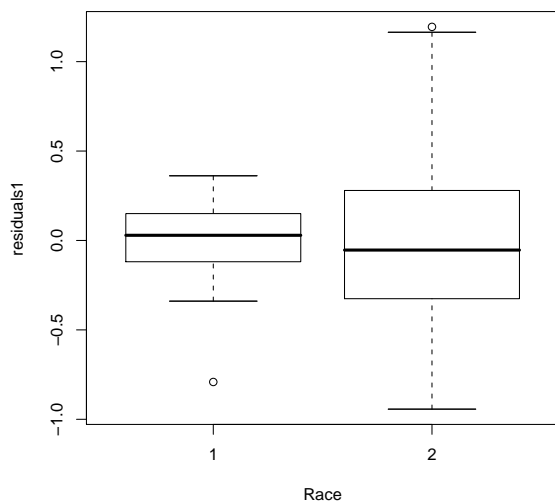
```



This clearly shows the systematic trend as a function of age.

Looking at the residuals by Race, we see a tendency that there is higher variation among the non-white respondents:

```
> plot(residuals1 ~ Race, data = residuals1)
```



For the other marginal we get

```
> marg2 <- mim("../", data = gab, marginal = c("Sex", "Race", "Opinion"))
> marg22factor <- editmim(marg2, deleteTerm = "Sex:Race:Opinion")
> modelTest(marg2, marg22factor)
```

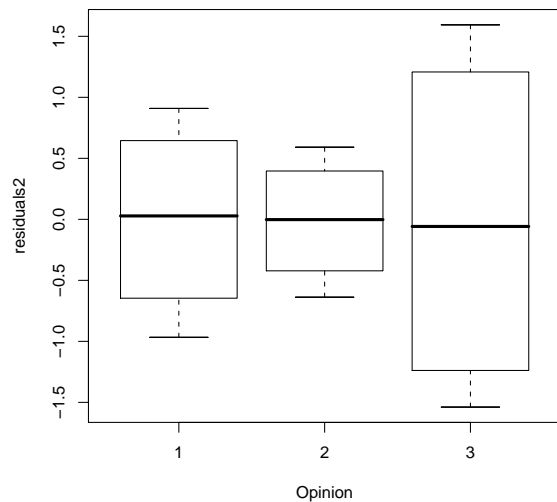
```
Test of H0 : Race:Sex:Opinion//
Against     : Sex:Opinion + Race:Opinion + Race:Sex//
```

```
test: Chi-squared method: asymptotic
stat: 9.482 df: 2 P: 0.009
```

and here even the two-factor model clearly does not fit the data.

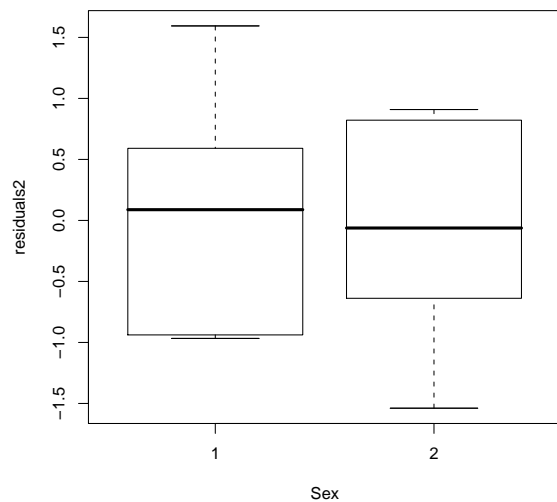
To identify more precisely how the relationship varies, we might also here look closer at the residuals for the two-factor model

```
> observed2 <- fitted(marg2)[, 4]
> expected2 <- fitted(marg22factor)[, 4]
> residuals2 <- (observed2 - expected2)/sqrt(expected2)
> residuals2 <- cbind(fitted(marg2)[, 1:3], residuals2)
> plot(residuals2 ~ Opinion, data = residuals2)
```



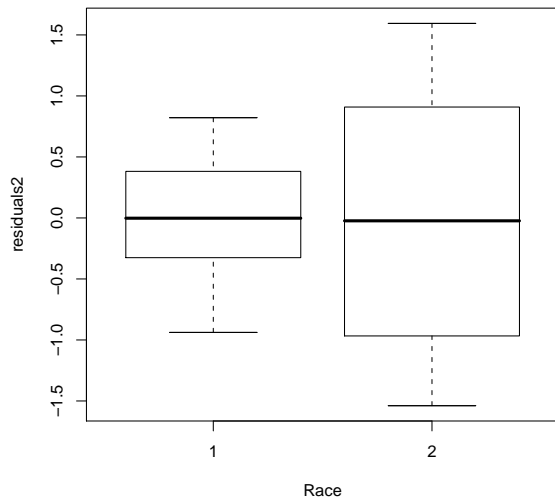
This plot indicates higher variation among the undecided.

```
> plot(residuals2 ~ Sex, data = residuals2)
```



There is also a strong difference between the two sexes.

```
> plot(residuals2 ~ Race, data = residuals2)
```

And again, higher variation among the non-white respondents.

References

Christensen, R. (1990). *Log-Linear Models*. Springer-Verlag, New York.