KU LEUVEN

FACULTY OF SCIENCE

Exam Report

Loglinear models and latent class analysis

Daniel Gerardo GIL SANCHEZ
daniel.gilsanchez@student.kuleuven.be
MSc Statistics

Supervisor: Prof. Femke De Keulenaer

Academic year 2018-2019

Contents

1 Exercise 1

In this exercise, data from the European Social Survey 2008 for **Russia** were used. A five-dimensional table was built with *gender*, *education*, *children at home*, *marital status* and whether the person agrees or disagrees with the following sentence: *"Woman should be prepared to cut down on paid work for the sake of family"*. Table 1 shows the label of each variable, its corresponding codification and the absolute frequency for each category.

Label	Description	Categories	Frequency
Gender (G)	Gender	1: Male	989
		2: Female	1523
Education (E)	Highest level of education	1: Less than or lower secondary	433
		2: Upper secondary	767
		3: Post-secondary	1312
Children (C)	Children at home	1: Yes	891
		2: No	1621
Marital Status (M)	Current marital status	1: Married/living together	1158
		2: Divorced/separated	325
		3: Widowed	501
		4: Never married/never lived together	495
WmnWork (W)	Woman should be to cut down on paid work for sake of family	1: Agree	1648
		2: Neither agree nor disagree	529
		3: Disagree	273

Table 1: Description of variables used in the analysis

In total there is information for 2512 individuals who live in **Russia** but in the combination of these variables to create the frequency table, 97 observations were lost because of the presence of missing values. Thus, this analysis involves 2419 individuals who answered these five variables. Additionally, it is important to mention that the number of categories in the response variable, five, led to a frequency table with 240 cells ($= 5 \times 4 \times 3 \times 2 \times 2$) and an average cell count of 10.08. For this reason and given the high proportion of zeros in the cell counts ($60/240 = 0.25$), the number of categories in the response was reduced to three, by merging the categories "*Agree*" and "*Agree strongly*" into "*Agree*", and "*Disagree*" and "*Disagree strongly*" into "*Disagree*". The final frequency table has 144 cells ($= 3 \times 4 \times 3 \times 2 \times 2$), with an average cell count of 16.8, and which proportion of zeros in cell counts lowered to 14.5% ($= 21/144$). Table 20 shows the frequency table used in this analysis.

The hypothesis in this study is that the perception of whether a woman should sacrifice its work for the sake of family is not the same for people with a different demographic background. For instance, a person who is married may have a different opinion in this regard than a person who has never been married. Also, it is expected that women tend to disagree more than men in this regard. It is also considered the fact that having children may change the meaning of family and therefore the tendency to agree with the sentence presented. Lastly, it is expected that people with higher level of education tend to disagree more than people with lower level of education. Of course, it is part of the hypothesis the possible influence that the combination of these variables may have in the final decision. Thus, it is expected that some interactions between the independent variables appear in the analysis.

Note that the zero cell counts in the final frequency table (Table 20) were considered to be sample zeros, meaning that they are rare combinations that did not occur in this sample but could occur if another sample were taken. For this reason, the value 0.05 was added to all cell counts.

1.1 Aitkin's Simultaneous Testing

The first step was to establish the relationship between all variables in terms of interactions. Since multiple simultaneous tests were performed ($143 = (3 \times 4 \times 3 \times 2 \times 2) - 1$), a correction in the significance level was needed. To do so, an overall γ of 0.45 was considered, which in combination with the number of simultaneous tests led to an overall α of 0.0042 ($= 1 - (1 - 0.45)^{1/143}$). Table 2 shows the results of the Aitkin's Simultaneous Testing procedure.

Null Hypothesis	Calculation of γ	Comparison	Conclusion
H_0: all 1st & higher=0	$\gamma = 1 - (1 - 0.0042)^{143} = 0.450$	$0.00 < 0.450$	Reject
H_0: all 2nd & higher=0	$\gamma = 1 - (1 - 0.0042)^{134} = 0.429$	$0.00 < 0.429$	Reject
H_0: all 3rd & higher=0	$\gamma = 1 - (1 - 0.0042)^{103} = 0.350$	$0.00 < 0.350$	Reject
H_0: all 4th & higher=0	$\gamma = 1 - (1 - 0.0042)^{52} = 0.195$	$0.6569 > 0.195$	No reject
H_0: all 5th	$\gamma = 1 - (1 - 0.0042)^{12} = 0.049$	$0.9927 > 0.049$	No reject

Table 2: Aitkin's Simultaneous Testing

The first step consists in test whether all first order effects and higher order effects are simultaneously equal to zero. Since the p-value of the model, that results from the probability of having a value as large as the L^2 of the model or larger in a χ^2 distribution with the corresponding degrees of freedom, is smaller than the corrected significance level (column 3 in Table 2), it is concluded that a more complex model is needed. Similarly, the second step consists in test whether second order interactions and higher order interactions are simultaneously equal to zero. The p-value is smaller that the corrected significance level, so a complex model is needed. In step 3, third and higher order interactions are simultaneously tested to be equal to zero, resulting a p-value smaller that the corrected significance level, implying that this hypothesis is rejected. Finally in step 4, fourth order and higher order interactions are simultaneously tested to be equal to zero, resulting in a p-value larger than the corrected significance level, not rejecting this hypothesis. Thus, it can be concluded that at least one third order interaction is significant. It is important to mention that this process was also computed using an overall γ of 0.35 and the results were the same.

1.2 Conditional Testing

Since Aitkin's procedure suggested that there was at least one third order interaction, the following step was to find out which interactions were indeed significant. To do so, conditional testing approach was considered on logit models, because a response variable was already defined.

This procedure consists of fitting the most complex model with third order interactions and removing those terms that does not present significant differences in the overall fit. To test whether removing a term produces a significant difference, the difference of the L^2 of each model is computed and contrasted with a χ^2 distribution with k degrees of freedom, where k is the difference in degrees of freedom between models. Note that this procedure can only be used with nested models as in this case. Also, the AIC and BIC were used to compare the complexity of each model. Table 3 presents these results.

Model	L^2	df	p	ΔL^2	Δ df	Δ p	AIC	BIC
1. W|GCEM {GCW, GEW,GMW,CEW, CMW,EMW}	35.68	46	0.8638	–	–	–	−56.32	−322.84
2. W|GCEM {GCW, GEW,GMW,CMW, EMW}	36.01	50	0.9316	0.33	4	0.988	−63.99	−353.69
3. W|GCEM {GEW, GMW,CMW,EMW}	36.22	52	0.9527	0.21	2	0.9009	−67.78	−369.07
4. W|GCEM {GEW, CMW,EMW}	44.46	58	0.9045	8.24	6	0.221	−71.54	−407.6
5. W|GCEM {GW,CMW, EMW}	51.54	62	0.8257	7.08	4	0.1316	−72.46	−431.69
6. W|GCEM {GW,CW, EMW}	62.11	68	0.6782	10.57	6	0.1026	−73.89	−467.89
7. W|GCEM {GW,CW, EW,MW}	83.26	80	0.3796	21.15	12	0.0483	−76.74	−540.27
8. W|GCEM {GW,EW, MW}	87.54	82	0.3174	4.28	2	0.1175	−76.46	−551.58

Table 3: Conditional Testing

This kind of models are hierarchical, meaning that all lower order interactions and main effects are also considered. In each model, the term that has the highest p-value is removed from the model and contrasted with the previous one. The first model has all third order interactions with the response variable. In this case, the interaction with the highest p-value is **CEW** and therefore it is removed. The second model, which does not have **CEW** interaction, does not present significant difference with respect to the first one, so it can be concluded that having children at home in interaction with the level of education does not have an influence in the response variable. This process of removing terms continued until a significant difference was found, see model 7. Up to this point interactions **CEW**, **GCW**, **GMW**, **GEW** and **CMW** did not produce significant differences in the fit of the model. It is not the case of the interaction **EMW** because removing it, led to a significant difference, even though it is borderline 0.0483.

So it can be concluded that the model that has main effects for *gender* and *children* and the interaction between *education* and *marital status* is the final model (model 6 in Table 3). Note that this conclusion is consistent with the conclusion of Aitkin's procedure.

1.3 Interpretation

Having the final model, it was of interest to know what was the influence of each independent variable in the response variable. Given that the response is polytomous with three categories, there are three different logit equations comparing *Agree* with *Neither agree nor disagree*, *Agree* with *Disagree* and *Neither agree nor disagree* with *Disagree*. Since the response variable is ordinal, only adjacent categories are interpreted.

In Table 4 the logit equation for *Agree* vs *Neither agree nor disagree* is presented. The difference between Male and Female is not large because the odds of agreeing that a woman should sacrifice its work in the sake of the family are just 1.094 times higher for males than females, which is practically the same. In a similar way, the odds of agreeing with the sentence are just 1.078 times higher for people that lives with children that people who do not. Note that for gender this difference corresponds to 1.6 percentage points and in living with children this difference is 1.3 percentage points. All these estimates were controlled by education and marital status [1].

The interpretation of the interaction is with respect to the reference category, which in this case is never married people with post-secondary education. So controlling for gender and children at home, the odds of agreeing that a woman should sacrifice its work in the sake of the family are 3.108 times higher for a widowed person with post-secondary education than a never married person with the same level of education. This interpretation holds for every comparison showed in Table 5 with respect to the reference category.

[1] To compute the difference in percentage points, the calculation of the corresponding probabilities of each category was based on marginal odds. To do so, the marginal frequencies for *Agree*, *Neither agree nor disagree* and *Disagree* were 1628, 526 and 265 respectively. Note that these frequencies are not the same as in the Table 2 because of the presence of missing values when the five-dimensional table was built.

	Lambda Parameters			Agree vs Neither		
	Agree	Neither	Disagree	Logit Equation	Logit Diff.	Odds Ratio
WmnWork	1.028	−0.136	−0.892	1.163		
Gender x WmnWork						
Male	0.072	0.027	−0.098	0.045	0.09	1.094
Female	−0.072	−0.027	0.098	(−0.045)	0	1
Children x WmnWork						
Yes	−0.03	−0.067	0.097	0.038	0.075	1.078
No	0.03	0.067	−0.097	(−0.038)	0	1
Education x WmnWork						
Less than lower	0.097	0.072	−0.169	0.025		
Upper secundary	0.009	0.026	−0.035	−0.017		
Post-secundary	−0.105	−0.098	0.204	(−0.007)		
Marital Status x WmnWork						
Married	0.166	0.118	−0.284	0.048		
Divorced	−0.032	0.008	0.024	−0.041		
Widowed	0.116	−0.134	0.018	0.25		
Never married	−0.25	0.007	0.243	(−0.257)		
Education x Marital Status x WmnWork						
Less than lower - Married	0.213	0.213	−0.426	0		
Less than lower - Divorced	−0.184	−0.204	0.387	0.02		
Less than lower - Widowed	0.045	0.074	−0.119	−0.029		
Less than lower - Never married	−0.074	−0.083	0.157	(0.009)		
Upper secundary - Married	−0.046	−0.253	0.299	0.208		
Upper secundary - Divorced	0.18	0.19	−0.37	−0.01		
Upper secundary - Widowed	−0.21	0.185	0.025	−0.394		
Upper secundary - Never married	0.075	−0.121	0.046	(0.196)		
Post-secundary - Married	−0.167	0.04	0.127	(−0.208)		
Post-secundary - Divorced	0.004	0.014	−0.018	(−0.011)		
Post-secundary - Widowed	0.165	−0.258	0.094	(0.423)		
Post-secundary - Never married	−0.001	0.204	−0.203	(−0.205)		

Table 4: Logit equation for *Agree* vs *Neither agree nor disagree*

In addition, it is also important to examine the odds between categories in each variable that composes the interaction. For instance, the odds of agreeing with the sentence in married people are larger as the education level rises (see first row in the last three columns in Table 5). On the contrary, the odds of agreeing with the sentence in widowed people are smaller as the education level rises. This is a clear example of why the interaction between both variables is important (see third row in the last three columns in Table 5).

In a similar way, the odds of agreeing with the sentence are 1.273 times higher for widowed people than divorced people, when the level of education is low, and 2.062 times higher when the level of education is high.

Likewise, in Table 6 the logit equation for *Neither agree nor disagree* vs *Disagree* is presented. Here, the odds of not agreeing nor disagreeing that a woman should sacrifice its work in the sake of the family are 1.284 times higher for males than females, which is almost 28% percent more than in the previous comparison. Additionally, the odds of not agreeing nor disagreeing with the sentence are just 1.388 $(= 1/0.721)$ times higher for people that do not live with children that people who do. So in this case the influence of living with children is the opposite than in the previous case. In percentage points the difference between male and female equals to 5.3 and between living with children or not equals to 7.6.

The interpretation of the interaction is with respect to the reference category, which in this case is again never married people with post-secondary education. So controlling for gender and children at home, the odds of not agreeing nor disagreeing that a woman should sacrifice its work in the sake of the family are 4.106 times higher for a married person with less than secondary education than a never married person with post-secondary education. This interpretation holds for every comparison showed in Table 7 with respect to the reference category.

As before, it is also important to examine the odds between categories in each variable that composes the interaction. For instance, the odds of not agreeing nor disagreeing with the sentence in never married people are smaller as the education level rises (see last row in the last three columns in Table 7). On the contrary, the odds of not agreeing nor disagreeing with the sentence in divorced people are larger as the education level rises. This, again, is a clear example of why the interaction between both variables is important (see second row in the last three columns in Table 7).

In a similar way, the odds of not agreeing nor disagreeing with the sentence are 4.556 times higher for married people than never married people, when the level of education is low, and 1.155 times higher when the level of education is high.

		Lambda parameters			Sum main and inter.			Sum reference cat.			Odds ratio reference cat.			Odds ratio between Education		
Mar. Sta.	**Effect**	**Less**	**Upper**	**Post**	**Less**	**Upper**	**Post**	**Less**	**Upper**	**Post**	**Less**	**Upper**	**Post**	**Less**	**Upper**	**Post**
Married	Edu.	0.025	−0.017	−0.007	0.072	0.238	−0.167	0.541	0.707	0.301	1.717	2.027	1.352	0.847	1.27	1.5
	Mar.	0.048	0.048	0.048												
	Int.	0	0.208	−0.208												
Divorced	Edu.	0.025	−0.017	−0.007	0.004	−0.067	−0.058	0.473	0.401	0.411	1.605	1.494	1.508	1.074	1.064	0.991
	Mar.	−0.041	−0.041	−0.041												
	Int.	0.02	−0.01	−0.011												
Widowed	Edu.	0.025	−0.017	−0.007	0.246	−0.162	0.665	0.714	0.307	1.134	2.043	1.359	3.108	1.503	0.657	0.437
	Mar.	0.25	0.25	0.25												
	Int.	−0.029	−0.394	0.423												
Never married	Edu.	0.025	−0.017	−0.007	−0.223	−0.078	−0.469	0.245	0.391	0	1.278	1.478	1	0.865	1.278	1.478
	Mar.	−0.257	−0.257	−0.257												
	Int.	0.009	0.196	−0.205												
								OR between Widowed and Divorced			1.273	0.91	2.062			

Table 5: Interaction between *Marital status* and *Education* for *Agree* vs *Neither* in the response

	Lambda Parameters			Neither vs Disagree		
	Agree	**Neither**	**Disagree**	**Logit Equation**	**Logit Diff.**	**Odds Ratio**
WmnWork	1.028	−0.136	−0.892	0.757		
Gender x WmnWork						
Male	0.072	0.027	−0.098	0.125	0.250	1.284
Female	−0.072	−0.027	0.098	(−0.125)	0	1
Children x WmnWork						
Yes	−0.03	−0.067	0.097	−0.164	−0.328	0.721
No	0.03	0.067	−0.097	(0.164)	0	1
Education x WmnWork						
Less than lower	0.097	0.072	−0.169	0.241		
Upper secundary	0.009	0.026	−0.035	0.060		
Post-secundary	−0.105	−0.098	0.204	(−0.302)		
Marital Status x WmnWork						
Married	0.166	0.118	−0.284	0.402		
Divorced	−0.032	0.008	0.024	−0.015		
Widowed	0.116	−0.134	0.018	−0.151		
Never married	−0.25	0.007	0.243	(−0.236)		
Education x Marital Status x WmnWork						
Less than lower - Married	0.213	0.213	−0.426	0.639		
Less than lower - Divorced	−0.184	−0.204	0.387	−0.591		
Less than lower - Widowed	0.045	0.074	−0.119	0.192		
Less than lower - Never married	−0.074	−0.083	0.157	(−0.240)		
Upper secundary - Married	−0.046	−0.253	0.299	−0.552		
Upper secundary - Divorced	0.18	0.19	−0.37	0.560		
Upper secundary - Widowed	−0.21	0.185	0.025	0.160		
Upper secundary - Never married	0.075	−0.121	0.046	(−0.168)		
Post-secundary - Married	−0.167	0.04	0.127	(−0.087)		
Post-secundary - Divorced	0.004	0.014	−0.018	(0.032)		
Post-secundary - Widowed	0.165	−0.258	0.094	(−0.352)		
Post-secundary - Never married	−0.001	0.204	−0.203	(0.407)		

Table 6: Logit equation for *Neither agree nor disagree* vs *Disagree*

2 Exercise 2

Having defined the best model to explain the influence of the independent variables on the response, it was of interest to find a new model for the relationship between *gender*, *education*, *children at home* and *marital status*. Therefore, a modified path model is defined in the following sections.

2.1 Theoretical Path Model

Figure 1 shows the suggested path model, where each color represents each submodel considered.

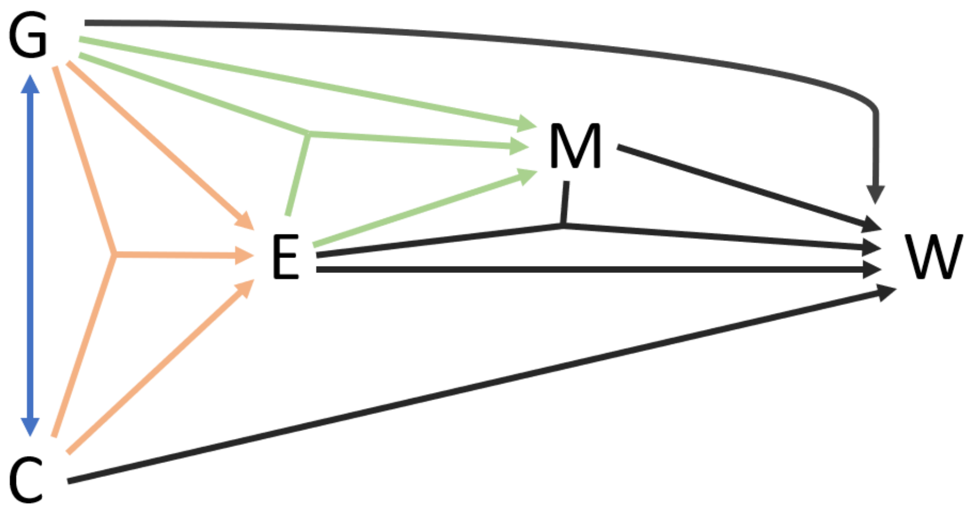

Figure 1: Theoretical path model

Since the global path model cannot be tested in one logit model, a stepwise approach was conducted where each submodel was investigated on a different marginal table. Table 8 presents these results.

Model	L^2	df	p-value
GC {GC}	0.000	0	0.000
E|GC {GCE}	0.000	0	0.000
M|GCE {GEM}	414.591	18	0.000
W|GCEM {GW,CW,EMW}	62.110	68	0.678
TOTAL	476.701	86	0.000

Table 8: Stepwise approach

It is important to mention that different theoretical path models were tested but none of them gave an overall p-value between 0.3 and 0.6, as it is expected. The inclusion of *marital status* turned the overall p-value to zero in all models considered.

		Lambda parameters			Sum main and inter.			Sum reference cat.			Odds ratio reference cat.			Odds ratio between Education		
Mar. Sta.	**Effect**	**Less**	**Upper**	**Post**	**Less**	**Upper**	**Post**	**Less**	**Upper**	**Post**	**Less**	**Upper**	**Post**	**Less**	**Upper**	**Post**
Married	Edu.	0.241	0.06	−0.302	1.282	−0.089	0.014	1.412	0.041	0.144	4.106	1.042	1.155	3.941	3.554	0.902
	Mar.	0.402	0.402	0.402												
	Int.	0.639	−0.552	−0.087												
Divorced	Edu.	0.241	0.06	−0.302	−0.365	0.605	−0.285	−0.235	0.735	−0.155	0.791	2.085	0.856	0.379	0.923	2.435
	Mar.	−0.015	−0.015	−0.015												
	Int.	−0.591	0.56	0.032												
Widowed	Edu.	0.241	0.06	−0.302	0.282	0.069	−0.805	0.413	0.2	−0.675	1.511	1.221	0.509	1.237	2.965	2.397
	Mar.	−0.151	−0.151	−0.151												
	Int.	0.192	0.16	−0.352												
Never married	Edu.	0.241	0.06	−0.302	−0.234	−0.343	−0.13	−0.104	−0.213	0	0.901	0.808	1	1.115	0.901	0.808
	Mar.	−0.236	−0.236	−0.236												
	Int.	−0.24	−0.168	0.407												
					OR between Married and Never Married						4.556	1.289	1.155			

Table 7: Interaction between *Marital status* and *Education* for *Neither* vs *Disagree* in the response

2.2 Best path model

The next step was to find a better path model by removing those terms that does not produce significant differences in the overall fit, as it was conducted on exercise 1. This is performed by conditional testing and Table 9 present these results.

Model	L^2	df	p	ΔL^2	Δ df	Δ p	AIC	BIC
Theoretical	476.7	86	0.000	–	–	–	304.7	−193.59
Removing GCE	477.69	88	0.000	0.99	2	0.609	301.69	−208.19
Removing GEM	483.98	94	0.000	6.29	6	0.392	295.98	−248.66
Removing GC	485.54	95	0.000	1.56	1	0.211	295.54	−254.9

Table 9: Conditional Testing

It can be seen that removing the interaction between *gender*, *children* and *education* did not produce significant differences. The same happened with the interaction between *gender*, *education* and *marital status*, removing this term did not change the fit of the initial model. Finally, removing the interaction between *gender* and living with *children* did not change significantly the overall fit either. The final model can be represented as:

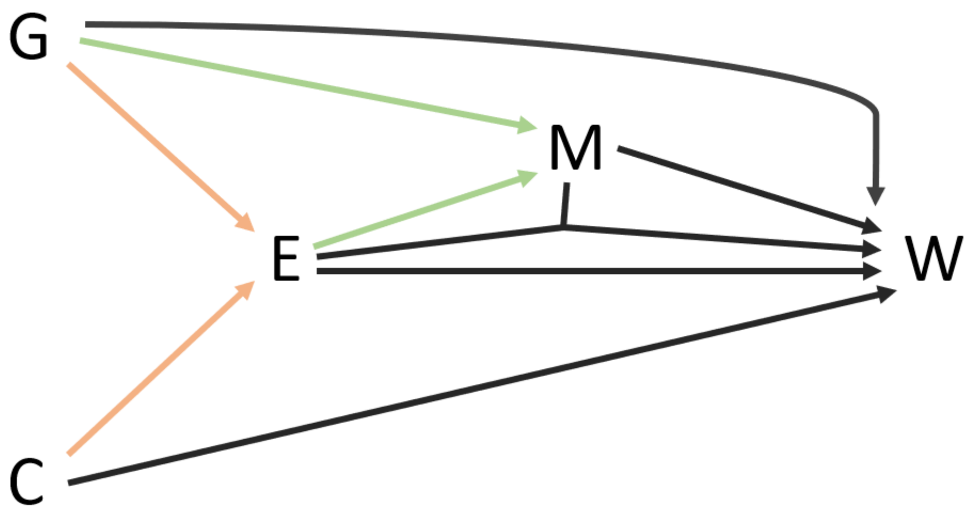

Figure 2: Modified path model

2.3 Interpretation

Note that this final model has the same terms that the model in exercise 1, the main difference is how some variables cause others. In the following, a description of each submodel is presented, beginning with the simpler model to the complex one. The last submodel is not interpreted because it is the same as in the exercise 1.

Table 10 shows the estimates for *gender* and *children* at home. On average, the odds of finding a female in the sample are 1.2395 times higher than expected on the basis of the overall effect. Likewise, the odds of finding people that do not live with children in the same are 1.3308 times higher than expected on the basis of the overall effect.

	Lambda	**Tau**
Gender		
Male	-0.2147	0.8068
Female	0.2147	1.2395
Children		
Yes	-0.2857	0.7515
No	0.2857	1.3308

Table 10: Parameter estimation. Submodel 1

Table 11 shows the logit equation for the second submodel, which involves the influence of *gender* in *education* and living with *children* in *education*. Since the response variable in this submodel is ordinal, only adjacent categories are interpreted.

It can be seen that the odds of having less than secondary education are $1.504(= 1/0.665)$ times higher for females than males, controlling by living with a *children*. In addition, the odds of having upper secondary education are 1.354 times higher for males than for females. So it can be concluded, that even though women are more likely to have lower level of education than upper secondary, it is also likely to find more women in post secondary than in upper secondary. Or in other words, it is more likely to find men with upper secondary than in any other level of education.

In regard to the second variable considered in the model, it can be concluded that the odds of having less than secondary education are $3.083(= 1/0.324)$ times higher for people that do not live with children that people who do, controlling by *gender*. Also, the odds of having upper secondary are $1.319(= 1/0.758)$ times higher for people that do not live with children that people who do. Thus, in general, people that do not live with children are more likely to have lower levels of education.

	Lambda Parameters			Less vs Upper		
	Less	**Upper**	**Post**	**Logit Equation**	**Logit Diff.**	**Odds Ratio**
Education	-0.8	0.131	0.668	-0.931		
Gender x Education						
Male	-0.086	0.119	-0.033	-0.204	-0.408	0.665
Female	0.086	-0.119	0.033	(0.204)	0	1
Children x Education						
Yes	-0.421	0.142	0.28	-0.563	-1.126	0.324
No	0.421	-0.142	-0.28	(0.563)	0	1

(a) Logit equation for *less than secondary* vs *upper secondary*

	Lambda Parameters			Upper vs Post		
	Less	**Upper**	**Post**	**Logit Equation**	**Logit Diff.**	**Odds Ratio**
Education	-0.8	0.131	0.668	-0.537		
Gender x Education						
Male	-0.086	0.119	-0.033	0.151	0.303	1.354
Female	0.086	-0.119	0.033	(-0.151)	0	1
Children x Education						
Yes	-0.421	0.142	0.28	-0.138	-0.277	0.758
No	0.421	-0.142	-0.28	(0.138)	0	1

(b) Logit equation for *upper secondary* vs *Post secondary*

Table 11: Logit equation for Submodel 2

Table 12 shows the logit equation for the third submodel, which involves the influence of *gender* in *marital status* and *education* in *marital status*. Since the response variable in this submodel is nominal with four categories, there are six possible logit equations. In this case, only two comparisons are going to be shown: Married vs Never Married and Divorced vs Widowed.

In regard to the first comparison, the odds of being married are 1.223 times higher for males than females, controlling by education. The odds of being married are $2.444 (= 1/0.409)$ times higher for people with post-secondary education than people with less than secondary education, and $2.062 (= 1/0.485)$ times higher for people with post-secondary education than people with upper secondary. Hence, it can be concluded that males with post secondary education are more likely to be married.

With respect to the second comparison, the odds of being divorced are 2.768 times higher for males than females, controlling by education. And the odds of being divorced are $8.208 (= 1/0.122)$ times higher for people with post-secondary education than people with less than secondary education, and $1.306 (= 1/0.766)$ times higher for people with post-secondary education than people with upper secondary. Thus, it can also be concluded that males with post secondary education are more likely to be divorced. It is actually an intuitive conclusion because if males with high level of education are more likely to be married, they are by consequence more likely to be divorced.

	Married	**Lambda Parameters** **Divorced**	**Widowed**	**Never married**	**Married vs Never married** **Logit Equation**	**Logit Diff.**	**Odds Ratio**
Marital Status	0.792	−0.654	−0.203	0.0647	0.728		
Gender x Marital Status							
Male	0.341	−0.036	−0.545	0.2399	0.101	0.201	1.223
Female	−0.341	0.036	0.545	−0.2399	(−0.101)	0	1
Education x Marital Status							
Less	−0.342	−0.493	0.822	0.013	−0.355	−0.894	0.409
Upper	0.1	0.07	−0.454	0.2846	−0.184	−0.724	0.485
Post	0.242	0.423	−0.367	−0.2976	(0.539)	0	1

(a) Logit equation for *Married* vs *Never married*

	Married	**Lambda Parameters** **Divorced**	**Widowed**	**Never married**	**Divorced vs Widowed** **Logit Equation**	**Logit Diff.**	**Odds Ratio**
Marital Status	0.792	−0.654	−0.203	0.0647	−0.452		
Gender x Marital Status							
Male	0.341	−0.036	−0.545	0.2399	0.509	1.018	2.768
Female	−0.341	0.036	0.545	−0.2399	(−0.509)	0	1
Education x Marital Status							
Less	−0.342	−0.493	0.822	0.013	−1.314	−2.105	0.122
Upper	0.1	0.07	−0.454	0.2846	0.524	−0.267	0.766
Post	0.242	0.423	−0.367	−0.2976	(0.791)	0	1

(b) Logit equation for *Divorced* vs *Widowed*

Table 12: Logit equation for Submodel 3

To conclude, different hypotheses were tested by doing this kind of analysis. The first hypothesis was that a person who is married may have a different opinion in regard to woman that should sacrifice its work in the sake of the family than a person who has never been married. This hypothesis was confirmed and complemented: there is indeed a difference in opinion between married and never married people, and more generally in marital status, which is also influenced by the level of education, because the interaction is significant.

The second hypothesis was that women tend to disagree more than men in this regard, but it was not actually proved because the odds of agreeing the sentence were similar to neither agreeing nor disagreeing. And the odds of neither agree nor disagree vs disagree were slightly higher for males than females. This term was borderline not significant in the model, so there is still doubt about the difference in opinion.

The third hypothesis consisted in the possible change of opinion with respect to the sentence, which could be influenced by the fact of living with children. Even though the variable was kept in the model, it seemed that it does not really influence the position of the people in regard to woman sacrificing its work for the sake of the family.

The last hypothesis was that people with higher level of education tend to disagree more than people with lower level of education and the answer is: it depends. Mainly because of the interaction with marital status, which means that the position of people in regard to the sentence is influenced by a combination of both variables.

3 Exercise 3

In this exercise, data from the General Social Survey 2008 collected in the USA were used. A five-dimensional frequency table was built with different variables with the purpose of measuring through a latent variable how respectful a person can be with respect to someone else's rights. Table 13 shows the label of each variable, its corresponding codification and the absolute frequency of each category.

Label	Description	Categories	Frequency
Women (W)	Abortion if woman wants	1: Yes	550
	for any reason	2: No	748
Communist (C)	Should communist teacher	1: Fired	500
	be fired	2: Not fired	783
Anti-religious (A)	Allow anti-religious book	1: Removed	353
	in library	2: Not Removed	962
Race-diff (R)	Differences due to lack of	1: Yes	600
	education	2: No	693
Homosexual (H)	Allow homosexual to speak	1: Allowed	1097
		2: Not Allowed	225

Table 13: Description of variables used in the analysis

In total there is information for 2023 individuals but in the frequency table those registers that present missing values in any of the variables considered in the analysis were removed. As a result, this analysis involves only 567 individuals who answered these five variables. Table 21 shows the frequency table used in this analysis.

The motivation to choose these questions was to represent how people feel about specific rights for some populations that are constantly discriminated, such as: women, anti-religious people, blacks/African-Americans, Homosexuals and people with different political/economical interests, in this case communists. The purpose of this section is to test whether these items actually measure one concept, which in this case is respect toward other people.

3.1 Exploratory analysis

The first step consisted of fitting different models with the objective of select the best one. First a model without a latent variable is fitted, called complete independence model, with the purpose of being the benchmark for other models with latent variables. The second model considered a latent variable with two classes, the third model considered a latent variable with three classes and so on. Table 14 shows the comparison of these models.

Model	L^2	X^2	df	p-value	AIC	BIC
Complete independence	235.29	327.24	26	0	183.29	70.44
2-class model	26.72	24.4	20	0.1434	−13.28	−100.09
3-class model	20.38	17.87	14	0.1186	−7.62	−68.38
4-class model	11.24	9.3	8	0.1887	−4.7645	−39.4874

Table 14: Comparison of latent class models

This comparison of models is based on the overall p-value, AIC and BIC criteria. The p-value of the complete independence model is zero, so it is discarded. The model with a latent variable with two classes seems to be the best choice, because is the one that presents lower AIC and BIC and an acceptable p-value.

Having chosen this model, different random starting values were considered to estimate the corresponding parameters with the purpose of avoiding local optima solutions. As a result the L^2, AIC and BIC were stable implying that the solution is good. Table 15 presents the estimated latent class probabilities and the conditional probabilities.

	X1	X2
	0.5664	0.4336
C 1	0.1852	0.6773
C 2	0.8148	0.3227
A 1	0.0488	0.5869
A 2	0.9512	0.4131
R 1	0.5489	0.369
R 2	0.4511	0.631
H 1	0.9889	0.6241
H 2	0.0111	0.3759
W 1	0.6028	0.14
W 2	0.3972	0.86

Table 15: Conditional probabilities with 2 latent classes

This model shows that 56.64% of the population belongs to latent class 1 and 43.36% to latent class 2. From the conditional probabilities it can be concluded that the first latent class represents those people who respect other people's rights because of the higher probabilities in each of the indicators. In this case, a person that belongs to this latent class has a 95% probability of not removing a book of an anti-religious person from the library and a 95% probability of allowing a homosexual to make a speech in his community.

On the other hand the second latent class represents those people who do not respect other people's rights. In this case, a person that belongs to this class has a 86% probability of not admitting a legal abortion if the woman wants it for any reason and a 67% probability of thinking that a communist should not be teaching in a college.

With this model, if one used this classification rule to assign cases to the two latent classes, it could be expected to classify 11.28% of the total sample incorrectly. This is actually 74% better than assigning all respondents to the largest latent class.

3.2 Testing significance of effects

It was of interest to know whether a small number of indicators measure the same underlying concept. To do so, the conditional testing approach was used by comparing the significance of the difference in L^2 of each model. Each variable was removed one at a time and compared with the model that has all five variables. This comparison could be conducted with this approach because all models were nested.

Table 16 presents the corresponding results as well as the AIC and BIC for each model. Since all differences are significant with respect to the model with all manifest variables, it can be concluded that each of the five indicators included in the model is significantly related to the latent variable.

Model	$\mathbf{L^2}$	**df**	**p**	$\mathbf{\Delta L^2}$	**Δ df**	**Δ p**	**AIC**	**BIC**
Unrestricted	26.72	20	0.1434	–	–	–	−13.28	−100.09
Removing C	106.1	21	0.0000	79.38	1	0.0000	64.1	−27.04
Removing A	127.33	21	0.0000	100.61	1	0.0000	85.33	−5.82
Removing R	38.32	21	0.0118	11.6	1	0.0007	−3.68	−94.83
Removing H	104.14	21	0.0000	77.42	1	0.0000	62.14	−29.01
Removing W	101.63	21	0.0000	74.92	1	0.0000	59.63	−31.51

Table 16: Conditional Testing

3.3 Confirmatory analysis

Some restrictions were considered in the model with the goal of improving the fit of the model. To test whether the restriction improved the fit, conditional testing procedure was conducted under the same conditions mentioned before. Now, taking Table 15 as a reference, the following restrictions were imposed:

1. Set to 1 the conditional probability of the second category in the anti-religious variable in latent class 1.

2. Set to 1 the conditional probability of the first category in the homosexual variable in latent class 1.

3. Restrict the conditional probability of the first category of the variable communist to be equal to the first category of the variable homosexual in latent class 2.

4. Restrict the conditional probability of the first category of the variable anti-religious in latent class 2 to be equal to the first category of the variable women in latent class 1.

These restrictions are shown in Table 17 and the corresponding results are in Table 18.

	X1	**X2**
	0.5664	0.4336
C 1	0.1852	0.6773
C 2	0.8148	0.3227
A 1	0.0488	0.5869
A 2	0.9512	0.4131
R 1	0.5489	0.369
R 2	0.4511	0.631
H 1	0.9889	0.6241
H 2	0.0111	0.3759
W 1	0.6028	0.14
W 2	0.3972	0.86

Table 17: Restrictions imposed in the model

In Table 18 can be seen that the first restriction produced significant changes in the fit of the model, so it is not considered. The remaining restrictions improved the model, thus the estimated latent class probabilities and the conditional probabilities of the final model are presented in Table 19.

Model	L^2	**df**	**p**	ΔL^2	**Δ df**	**Δ p**	**AIC**	**BIC**
Unrestricted	26.72	20	0.1434	–	–	–	−13.28	−100.09
Restriction 1	31.07	21	0.0725	4.35	1	0.0369	−10.93	−102.08
Restriction 2	27.46	21	0.1563	0.74	1	0.3905	−14.54	−105.69
Restriction 3	27.76	22	0.1838	0.3	1	0.5823	−16.24	−111.73
Restriction 4	28.74	23	0.1891	0.98	1	0.3221	−17.26	−117.09

Table 18: Testing restrictions

These results suggest that 54.52% of the population belongs to latent class 1 and 45.48% to latent class 2. From the conditional probabilities it can be concluded that the interpretation did not change with respect to the model showed in the first numeral (see Table 15), where the first latent class represents those who respect other people's rights and the second latent class represents those who do not respect other people's rights.

With this model, if one used this classification rule to assign cases to the two latent classes, it could be expected to classify 11.15% of the total sample incorrectly. This is actually 75.48% better than assigning all respondents to the largest latent class.

	X1	**X2**
	0.5452	0.4548
C 1	0.1837	0.642
C 2	0.8163	0.358
A 1	0.0354	0.5936
A 2	0.9646	0.4064
R 1	0.5498	0.3764
R 2	0.4502	0.6236
H 1	1	0.642
H 2	0	0.358
W 1	0.5936	0.1571
W 2	0.4064	0.8429

Table 19: Conditional probabilities with 2 latent classes and imposed restrictions

To conclude, a five-dimensional frequency table was built with indicators on how people feel about specific rights for some populations that are constantly discriminated. This table was used to conduct a latent class analysis that results in two classes that represents two different kinds of people: those who respect other people's rights and those who do not it. An attempt to remove some of the indicators was performed without success because removing each of the manifest variables led to significant changes in the model. Four restrictions were finally imposed, and only three of them were kept in the model.

4 Appendix

Gender	Children	Education	Marital Status	WmnWork Agree	Neither	Disagree
Male	Yes	Less than secondary	Married	11	2	0
			Divorced	0	0	0
			Widowed	6	2	0
			Never Married	1	1	0
		Upper secondary	Married	74	18	14
			Divorced	4	0	0
			Widowed	1	1	1
			Never Married	2	2	0
		Post secondary	Married	120	35	21
			Divorced	7	1	0
			Widowed	5	0	0
			Never Married	0	0	0
	No	Less than secondary	Married	33	10	1
			Divorced	6	3	2
			Widowed	17	6	0
			Never Married	30	11	2
		Upper secondary	Married	59	20	5
			Divorced	21	7	1
			Widowed	7	4	1
			Never Married	63	19	10
		Post secondary	Married	94	35	12
			Divorced	27	11	8
			Widowed	17	3	2
			Never Married	53	18	6
Female	Yes	Less than secondary	Married	13	2	0
			Divorced	2	1	0
			Widowed	14	3	4
			Never Married	0	0	0
		Upper secondary	Married	67	16	6
			Divorced	19	4	3
			Widowed	14	5	3
			Never Married	7	5	1
		Post secondary	Married	127	42	28
			Divorced	52	23	15
			Widowed	34	8	9
			Never Married	6	7	3
	No	Less than secondary	Married	31	12	2
			Divorced	8	1	1
			Widowed	99	24	8
			Never Married	19	8	8
		Upper secondary	Married	58	9	8
			Divorced	15	9	1
			Widowed	43	15	6
			Never Married	60	20	16
		Post secondary	Married	93	48	16
			Divorced	47	10	7
			Widowed	93	14	16
			Never Married	49	31	18

Table 20: Frequency table ESS 2008

Communist	Anti-religious	Race-diff	Homosexual	Women Yes	Women No
Fired	Removed	Yes	Allowed	8	16
			Not Allowed	0	12
		No	Allowed	6	32
			Not Allowed	5	25
	Not removed	Yes	Allowed	16	29
			Not Allowed	3	9
		No	Allowed	23	29
			Not Allowed	1	12
Not Fired	Removed	Yes	Allowed	6	14
			Not Allowed	0	4
		No	Allowed	3	21
			Not Allowed	1	7
	Not removed	Yes	Allowed	92	51
			Not Allowed	1	6
		No	Allowed	61	64
			Not Allowed	2	8

Table 21: Frequency table GSS 2008

- **LEM Script Exercise 1**

```
* G = Gender (1 = Male, 2 = Female);
* C = Children living at home or not (1 = Respondent lives with children, 2 = Does not);
* E = Education (1 = Less than lower secondary education or lower secondary education
*                 2 = Upper secondary education completed
*                 3 = Post-secondary non-tertiary education completed
*          or tertiary eduaction completed);
* M = Marital Status (1 = Maried/living together
*                     2 = Divorced/separated
*                     3 = Widowed
*                     4 = Never maried/never lived together);
* W = Woman should be prepared to cut down on paid work for sake of family
*                (1 = Agree
*                 2 = Neither agree nor disagree
*                 3 = Disagree);

*************************************************
*******      AITKINS TESTING PROCEDURE      *******
*************************************************
*man 5
*dim 2 2 3 4 3
*lab G C E M W
* 1st order and higher = 0
* mod {}
* 2nd order and higher = 0
* mod {G,C,E,M,W}
* 3rd order and higher = 0
* mod {GC,GE,GM,GW,CE,CM,CW,EM,EW,MW}
* 4th order and higher = 0
* mod {GCE,GCM,GCW,GEM,GEW,GMW,CEM,CEW,CMW,EMW}
* 5th order and higher = 0
* mod {GCEM,GCEW,GCMW,GEMW,CEMW}
*rec 144
*rco
***nse  * in case I need it


*************************************************
```

```
*******      CONDITIONAL TESTING              *******
****************************************************
* At least one third-order interaction is significant
* Most complex model
*mod {GCE,GCM,GCW,GEM,GEW,GMW,CEM,CEW,CMW,EMW}

* Logit notation
*mod W|GCEM {GCW,GEW,GMW,CEW,CMW,EMW} * (1)
*mod W|GCEM {GCW,GEW,GMW,CMW,EMW} * (2) Remove CEW
*mod W|GCEM {GEW,GMW,CMW,EMW} * (3) Remove GCW
*mod W|GCEM {GEW,CMW,EMW} * (4) Remove GMW
*mod W|GCEM {GW,CMW,EMW} * (5) Remove GEW
mod W|GCEM {GW,CW,EMW} * (6) Remove CMW  * Chosen
*mod W|GCEM {GW,CW,EW,MW} * (7) Remove EMW
*mod W|GCEM {GW,EW,MW} * (8) Remove CW
dat [ 1 1       1       1       1       11.05
1       1       1       1       2       2.05
1       1       1       1       3       0.05
1       1       1       2       1       0.05
1       1       1       2       2       0.05
1       1       1       2       3       0.05
1       1       1       3       1       6.05
1       1       1       3       2       2.05
1       1       1       3       3       0.05
1       1       1       4       1       1.05
1       1       1       4       2       1.05
1       1       1       4       3       0.05
1       1       2       1       1       74.05
1       1       2       1       2       18.05
1       1       2       1       3       14.05
1       1       2       2       1       4.05
1       1       2       2       2       0.05
1       1       2       2       3       0.05
1       1       2       3       1       1.05
1       1       2       3       2       1.05
1       1       2       3       3       1.05
1       1       2       4       1       2.05
1       1       2       4       2       2.05
1       1       2       4       3       0.05
1       1       3       1       1       120.05
1       1       3       1       2       35.05
1       1       3       1       3       21.05
1       1       3       2       1       7.05
1       1       3       2       2       1.05
1       1       3       2       3       0.05
1       1       3       3       1       5.05
1       1       3       3       2       0.05
1       1       3       3       3       0.05
1       1       3       4       1       0.05
1       1       3       4       2       0.05
1       1       3       4       3       0.05
1       2       1       1       1       33.05
1       2       1       1       2       10.05
1       2       1       1       3       1.05
1       2       1       2       1       6.05
1       2       1       2       2       3.05
1       2       1       2       3       2.05
1       2       1       3       1       17.05
1       2       1       3       2       6.05
1       2       1       3       3       0.05
1       2       1       4       1       30.05
1       2       1       4       2       11.05
1       2       1       4       3       2.05
1       2       2       1       1       59.05
1       2       2       1       2       20.05
1       2       2       1       3       5.05
1       2       2       2       1       21.05
1       2       2       2       2       7.05
1       2       2       2       3       1.05
1       2       2       3       1       7.05
1       2       2       3       2       4.05
```

1	2	2	3	3	1.05
1	2	2	4	1	63.05
1	2	2	4	2	19.05
1	2	2	4	3	10.05
1	2	3	1	1	94.05
1	2	3	1	2	35.05
1	2	3	1	3	12.05
1	2	3	2	1	27.05
1	2	3	2	2	11.05
1	2	3	2	3	8.05
1	2	3	3	1	17.05
1	2	3	3	2	3.05
1	2	3	3	3	2.05
1	2	3	4	1	53.05
1	2	3	4	2	18.05
1	2	3	4	3	6.05
2	1	1	1	1	13.05
2	1	1	1	2	2.05
2	1	1	1	3	0.05
2	1	1	2	1	2.05
2	1	1	2	2	1.05
2	1	1	2	3	0.05
2	1	1	3	1	14.05
2	1	1	3	2	3.05
2	1	1	3	3	4.05
2	1	1	4	1	0.05
2	1	1	4	2	0.05
2	1	1	4	3	0.05
2	1	2	1	1	67.05
2	1	2	1	2	16.05
2	1	2	1	3	6.05
2	1	2	2	1	19.05
2	1	2	2	2	4.05
2	1	2	2	3	3.05
2	1	2	3	1	14.05
2	1	2	3	2	5.05
2	1	2	3	3	3.05
2	1	2	4	1	7.05
2	1	2	4	2	5.05
2	1	2	4	3	1.05
2	1	3	1	1	127.05
2	1	3	1	2	42.05
2	1	3	1	3	28.05
2	1	3	2	1	52.05
2	1	3	2	2	23.05
2	1	3	2	3	15.05
2	1	3	3	1	34.05
2	1	3	3	2	8.05
2	1	3	3	3	9.05
2	1	3	4	1	6.05
2	1	3	4	2	7.05
2	1	3	4	3	3.05
2	2	1	1	1	31.05
2	2	1	1	2	12.05
2	2	1	1	3	2.05
2	2	1	2	1	8.05
2	2	1	2	2	1.05
2	2	1	2	3	1.05
2	2	1	3	1	99.05
2	2	1	3	2	24.05
2	2	1	3	3	8.05
2	2	1	4	1	19.05
2	2	1	4	2	8.05
2	2	1	4	3	8.05
2	2	2	1	1	58.05
2	2	2	1	2	9.05
2	2	2	1	3	8.05
2	2	2	2	1	15.05
2	2	2	2	2	9.05
2	2	2	2	3	1.05
2	2	2	3	1	43.05

```
2         2         2         3         2         15.05
2         2         2         3         3         6.05
2         2         2         4         1         60.05
2         2         2         4         2         20.05
2         2         2         4         3         16.05
2         2         3         1         1         93.05
2         2         3         1         2         48.05
2         2         3         1         3         16.05
2         2         3         2         1         47.05
2         2         3         2         2         10.05
2         2         3         2         3         7.05
2         2         3         3         1         93.05
2         2         3         3         2         14.05
2         2         3         3         3         16.05
2         2         3         4         1         49.05
2         2         3         4         2         31.05
2         2         3         4         3         18.05
]
```

- **LEM Script Exercise 2**

```
* G = Gender (1 = Male, 2 = Female);
* C = Children living at home or not (1 = Respondent lives with children, 2 = Does not);
* E = Education (1 = Less than lower secondary education or lower secondary education
*                  2 = Upper secondary education completed
*                  3 = Post-secondary non-tertiary education completed
*           or tertiary eduaction completed);
* M = Marital Status (1 = Maried/living together
*                       2 = Divorced/separated
*                       3 = Widowed
*                       4 = Never maried/never lived together);
* W = Woman should be prepared to cut down on paid work for sake of family
*                 (1 = Agree
*                  2 = Neither agree nor disagree
*                  3 = Disagree);
*man 5
*dim 2 2 3 4 3
*lab G C E M W
*mod W|GCEM {GW,CW,EMW} * (10) remove CMW  * Chosen

* PATH MODEL
* Model 1
* 1.1
*man 2
*dim 2 2
*lab G C
*mod GC
*rec 144
*ski [3 4 5]
*rco

* 1.2
*man 3
*dim 2 2 3
*lab G C E
*mod E|GC {GCE}
*rec 144
*ski [4 5]
*rco

* 1.3
*man 4
*dim 2 2 3 4
*lab G C E M
*mod M|GCE {GEM}
*rec 144
*ski [5]
*rco

* 1.4
*man 5
```

```
*dim 2 2 3 4 3
*lab G C E M W
*mod W|GCEM {GW,CW,EMW}
*rec 144
*rco

* Final Model 1
*man 5
*dim 2 2 3 4 3
*lab G C E M W
*mod     GC {GC}
*        E|GC {GCE}
*        M|GCE {GEM}
*        W|GCEM {GW,CW,EMW}
*rec 144
*rco

*****
* Model 2
* 2.1 is the same as 1.1

* 2.2
*man 3
*dim 2 2 3
*lab G C E
*mod E|GC {GE,CE}
*rec 144
*ski [4 5]
*rco

* 2.3 is the same as 1.3
* 2.4 is the same as 1.4

* Final Model 2
*man 5
*dim 2 2 3 4 3
*lab G C E M W
*mod     GC {GC}
*        E|GC {GE,CE}
*        M|GCE {GEM}
*        W|GCEM {GW,CW,EMW}
*rec 144
*rco

*****
* Model 3
* 3.1 is the same as 2.1
* 3.2 is the same as 2.2

* 3.3
*man 4
*dim 2 2 3 4
*lab G C E M
*mod M|GCE {GM,EM}
*rec 144
*ski [5]
*rco

* 3.4 is the same as 2.4

* Final Model 3
*man 5
*dim 2 2 3 4 3
*lab G C E M W
*mod     GC {GC}
*        E|GC {GE,CE}
*        M|GCE {GM,EM}
*        W|GCEM {GW,CW,EMW}
*rec 144
*rco
```

```
*****
* Model 4
* 4.1
*man 2
*dim 2 2
*lab G C
*mod GC {G,C}
*rec 144
*ski [3 4 5]
*rco

* 4.2 is the same as 3.2
* 4.3 is the same as 3.3
* 4.4 is the same as 3.4

* Final Model 4
*man 5
*dim 2 2 3 4 3
*lab G C E M W
*mod      GC {G,C}
*         E|GC {GE,CE}
*         M|GCE {GM,EM}
*         W|GCEM {GW,CW,EMW}
*rec 144
*rco

* THE DATA IS THE SAME THAT IN THE FIRST EXERCISE
```

- **LEM Script Exercise 3**

```
* Latent class analysis
* C: Comunist (1: Fired, 2:Not fired)
* A: Anti-Religious (1: Removed 2: Not removed)
* R: Black people have chance of education (1: Yes, 2: No)
* H: Homosexual allow to speak (1: Allowed, 2: Not allowed)
* W: Abortion for any reason (1: Yes, 2: No)

**************************************************
*******      TESTING LATENT CLASSES        *******
**************************************************
* Independent model
*man 5
*dim 2 2 2 2 2
*lab C A R H W
*mod {C,A,R,H,W}

* Latent with 2 classes  * Chosen
*lat 1
*man 5
*dim 2 2 2 2 2 2
*lab X C A R H W
*mod X C|X A|X R|X H|X W|X

* Latent with 3 classes
*lat 1
*man 5
*dim 3 2 2 2 2 2
*lab X C A R H W
*mod X C|X A|X R|X H|X W|X

* Latent with 4 classes
*lat 1
*man 5
*dim 4 2 2 2 2 2
*lab X C A R H W
*mod X C|X A|X R|X H|X W|X

**************************************************
***   CONDITIONAL TESTING FOR EACH VARIABLE   ****
**************************************************
* Conditional testing for each variable
```

```
* Latent with 2 classes  * Unrestricted model
*lat 1
*man 5
*dim 2 2 2 2 2 2
*lab X C A R H W
*mod X C|X A|X R|X H|X W|X

* Removing C
*lat 1
*man 5
*dim 2 2 2 2 2 2
*lab X C A R H W
*mod X C A|X R|X H|X W|X

* Removing A
*lat 1
*man 5
*dim 2 2 2 2 2 2
*lab X C A R H W
*mod X C|X A R|X H|X W|X

* Removing R
*lat 1
*man 5
*dim 2 2 2 2 2 2
*lab X C A R H W
*mod X C|X A|X R H|X W|X

* Removing H
*lat 1
*man 5
*dim 2 2 2 2 2 2
*lab X C A R H W
*mod X C|X A|X R|X H W|X

* Removing W
*lat 1
*man 5
*dim 2 2 2 2 2 2
*lab X C A R H W
*mod X C|X A|X R|X H|X W


**************************************************
*******         CONFIRMATORY ANALYSIS      *******
**************************************************
* Confirmatory analysis: Restrictions
* Latent with 2 classes  * Unrestricted model
*lat 1
*man 5
*dim 2 2 2 2 2 2
*lab X C A R H W
*mod     X
*        C|X
*        A|X
*        R|X
*        H|X
*        W|X
*sta X [.4 .3]
*sta CX [.8 .2 .7 .3]
*sta AX [.8 .2 .7 .3]
*sta RX [.8 .2 .7 .3]
*sta HX [.8 .2 .7 .3]
*sta WX [.8 .2 .7 .3]

* Restriction A, Latent class 1 * Significant change in L2
*lat 1
*man 5
*dim 2 2 2 2 2 2
*lab X C A R H W
*mod     X
```

```
*        C|X
*        A|X eq2
*        R|X
*        H|X
*        W|X
*des [0 -1 0 0]
*sta X [.4 .3]
*sta CX [.8 .2 .7 .3]
*sta AX [0 1 .7 .3]
*sta RX [.8 .2 .7 .3]
*sta HX [.8 .2 .7 .3]
*sta WX [.8 .2 .7 .3]

* Restriction H, Latent class 1 * Not significant change in L2
*lat 1
*man 5
*dim 2 2 2 2 2 2
*lab X C A R H W
*mod     X
*        C|X
*        A|X
*        R|X
*        H|X eq2
*        W|X
*des [-1 0 0 0]
*sta X [.4 .3]
*sta CX [.8 .2 .7 .3]
*sta AX [.8 .2 .7 .3]
*sta RX [.8 .2 .7 .3]
*sta HX [1 0 .7 .3]
*sta WX [.8 .2 .7 .3]

* Restriction C1 = H1, Latent class 2. Same prob.
*lat 1
*man 5
*dim 2 2 2 2 2 2
*lab X C A R H W
*mod     X
*        C|X eq2
*        A|X
*        R|X
*        H|X eq2
*        W|X
*des [0 0 2 0
*-1 0 2 0]
*sta X [.4 .3]
*sta CX [.8 .2 .7 .3]
*sta AX [.8 .2 .7 .3]
*sta RX [.8 .2 .7 .3]
*sta HX [1 0 .7 .3]
*sta WX [.8 .2 .7 .3]

* Restriction A1 in class 2 = W1 in class 1
lat 1
man 5
dim 2 2 2 2 2 2
lab X C A R H W
mod      X
         C|X eq2
         A|X eq2
         R|X
         H|X eq2
         W|X eq2
des [0 0 2 0
0 0 3 0
-1 0 2 0
3 0 0 0]
sta X [.4 .3]
sta CX [.8 .2 .7 .3]
sta AX [.8 .2 .7 .3]
sta RX [.8 .2 .7 .3]
```

```
sta HX [1 0 .7 .3]
sta WX [.8 .2 .7 .3]


dat [8  16
0       12
6       32
5       25
16      29
3       9
23      29
1       12
6       14
0       4
3       21
1       7
92      51
1       6
61      64
2       8
]
```