Latent Class Analysis: Model Selection (part 2)

Jeroen K. Vermunt

Department of Methodology and Statistics, Tilburg University

www.jeroenvermunt.nl

Introduction

- In the previous video I demonstrated how to use the different types of statistics

- In this video I will give more details on the computation of these statistics.

- I will also pay attention to (parametric) bootstrapping of p values

Four types of statistics

- Information criteria (BIC, AIC, AIC3)

- Goodness-of-fit tests (L-squared, X-squared)
  - Including bootstrap p values

- Bivariate residuals (BVRs)

- Likelihood-ratio (-2LLdiff) tests
  - Including bootstrap p values

Information criteria

- Balancing model fit (-2LL value) and model complexity (Npar)

BIC = -2*LL* + ln(*N*) * *Npar*
AIC = -2*LL* + 2 * *Npar*
AIC3 = -2*LL* + 3 * *Npar*

$Npar = (C-1) + C \cdot \sum_{j=1}^{J}(M_j - 1)$; where $M_j = \text{number of categories item } j$

- BIC, AIC and AIC3 differ in the "penalty" for the number of parameters
- Note that ln(*N*) will (almost) always be larger than 3 (thus …?)

- Sometimes information criteria are computed as L^2 - w * df

Depression.sav example (2 class model)

- N=1710; ln(N)=7.4442

- Npar = 1 + 2 * (1+1+1+1+1) = 11

- BIC = -2*-4370.4561 + 7.4442 * 11 = 8822.7990
- AIC = -2*-4370.4561 + 2 * 11 = 8762.9122
- AIC3 = -2*-4370.4561 + 3 * 11 = 8773.9122

Goodness-of-fit tests

- H0: the model with C classes
- H1: the "saturated" model

- Observed frequency for pattern *p*: n_p
- Estimated frequency under model with C classes: $\mu_p = N * P(\mathbf{y}_p)$

- Likelihood-ratio chi-squared: $L^2 = 2 \sum_p n_p \ln \frac{n_p}{\mu_p}$
- Pearson chi-squared: $X^2 = \sum_p \frac{(n_p - \mu_p)^2}{\mu_p}$

Goodness-of-fit tests (DF)

- *df* = degrees of freedom
- *Npar* = number of parameters
- M_j = number of categories item j

- *df* = number of patterns - 1 – *Npar* = $(\prod_{j=1}^{J} M_j) - 1 - Npar$

  with $Npar = (C - 1) + C \cdot \sum_{j=1}^{J} (M_j - 1)$

- Watch out with sparseness! L^2 and X^2 will give very different p values. Use bootstrap p values in that case.

Depression.sav example (2 class model)

- $L^2 =$ 2*{64*ln(64/28.7075) + 60*ln(60/76.6320) + … = 103.9374
- $X^2 =$ (64−28.7075)^2/28.7075 + (60−76.6320)^2/76.6320 + … = 112.0766
- *df* = 2*2*2*2*2-1 - 11 = 31 – 11 = 20

Bivariate Residuals (BVRs)

- Pearson chi-squared divided by “df” is computed in all two-way tables

- Estimated frequencies in a two table can be obtained by applying the LC model equation to the pair concerned:

$$N \cdot P(y_j, y_{j'}) = N \cdot \sum_{c=1}^{C} P(X = c)\, P(y_j \mid X = c)\, P(y_{j'} \mid X = c)$$

Depression.sav example (2 class model)

appetite	hopeless	n	P(y)	mu	BVR
poor appetite	feeling hopeless	103	0.04432	75.7895	9.769289
poor appetite	hopeful	225	0.14749	252.212	2.935967
good appetite	feeling hopeless	112	0.08138	139.152	5.298028
good appetite	hopeful	1270	0.72681	1242.85	0.59324
		1710	1		18.59652

Likelihood-ratio (LR) tests

- H0: the model with C classes
- H1: the model with C+1 classes

- LR = -2LLdiff = -2LL(C classes) - -2LL(C+1 classes)

- Not allowed to use a “standard” p-value
- Alternative: Vuong-Lo-Mendell-Rubin (VLMR) p-value

- Better: bootstrap p-value

Depression.sav example (1-4 class model)

	LL	VLMR	p-value
1-Cluster	-4813.17		
2-Cluster	-4370.46	885.421	0
3-Cluster	-4331.2	78.5168	0
4-Cluster	-4323.92	14.5627	0.0109

Note: Latent GOLD provides VLMR test when running models for range of classes

Bootstrap p-value for L^2 (X^2) and/or -2LLdiff

- *Non-parametric* or naïve bootstrap: sample *N* observations with replacement from the data set. We do not use this one here!

- *Parametric* bootstrap: generate samples of *N* observations from the assumed population model (the H0 model).

- In fact, we mimic the assumed sampling mechanism and construct the distribution of the test statistic via MC simulation.

Bootstrapping L^2 (X^2) and -2LLdiff

- Generate say 500 samples of size *N* from the C-class model

- For L^2 and/or X^2
  - For each sample, estimate the C-class model and compute the L^2 (X^2) value
  - The p value is the proportion of bootstrap samples with a L^2 (X^2) value larger than the one in your sample

- For -2LLdiff
  - For each sample, estimate the C-class and C+1-class models and compute -2LLdiff value
  - The p value is the proportion of bootstrap samples with a -2LLdiff value larger than the one in your sample

Depression.sav example (1-4 class model)

- In Latent GOLD, you can request bootstrap p-values either for an estimated model or for a range of models with the bootstrap option in Output

	L^2	df	p-value	Bootstrap p	VLMR	p-value	-2LL Diff	Bootstrap p
1-Cluster	989.3584	26	6.70E-192	0				
2-Cluster	103.9374	20	2.50E-13	0	885.421	0	885.421	0
3-Cluster	25.4206	14	0.031	0.032	78.5168	0	78.5168	0
4-Cluster	10.8579	8	0.21	0.362	14.5627	0.0109	14.5627	0.02

- Note that in this application sparseness is not a problem.