

Latent Class Analysis

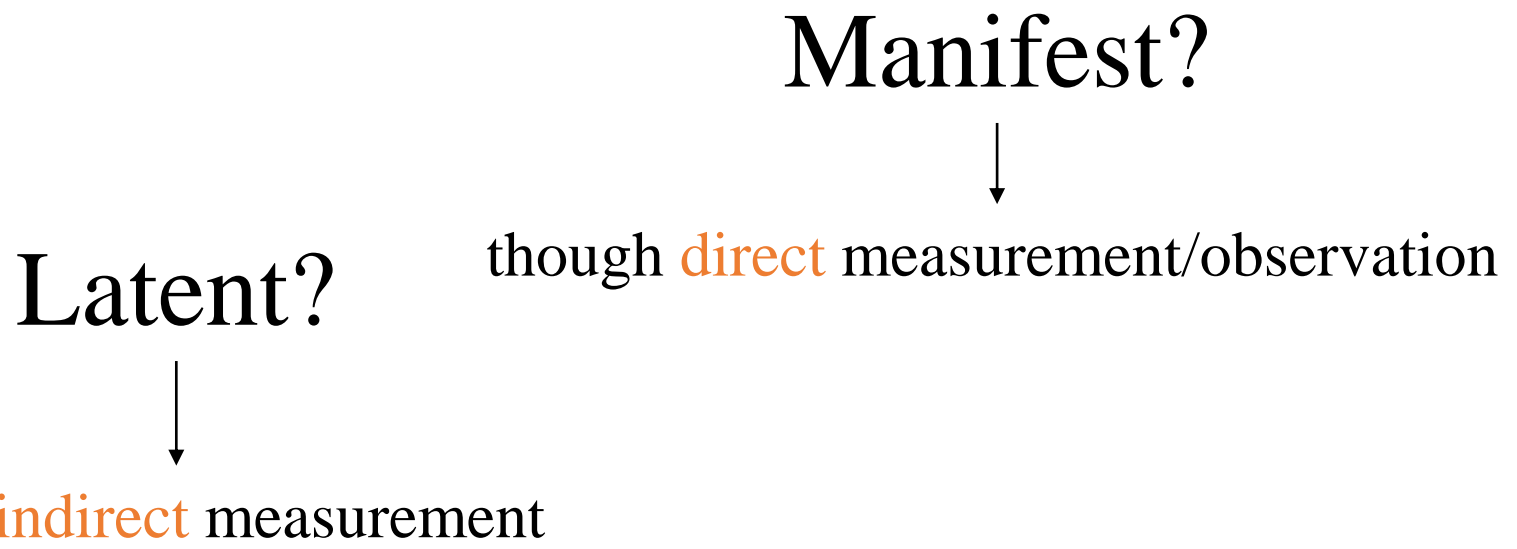
——from manifest variables to latent variables

Reporter: 黄颖诗

- **Related concepts**
- **What we have already knew**
- **What we want to know**
- **LatentGOLD**

Related concepts

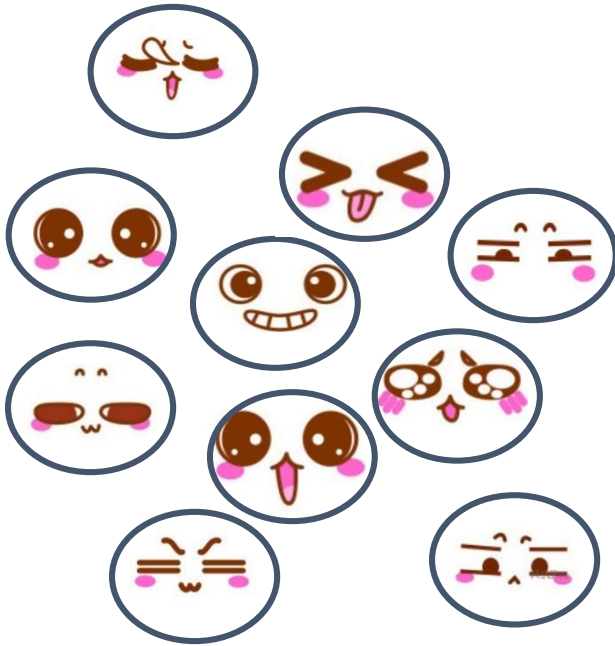
:-D



Already knew

As for continuous variable











172 TOT COLLEGE 012 A 8E5SDIC



(10 items for math ability)



✓ Describe from 10 aspects

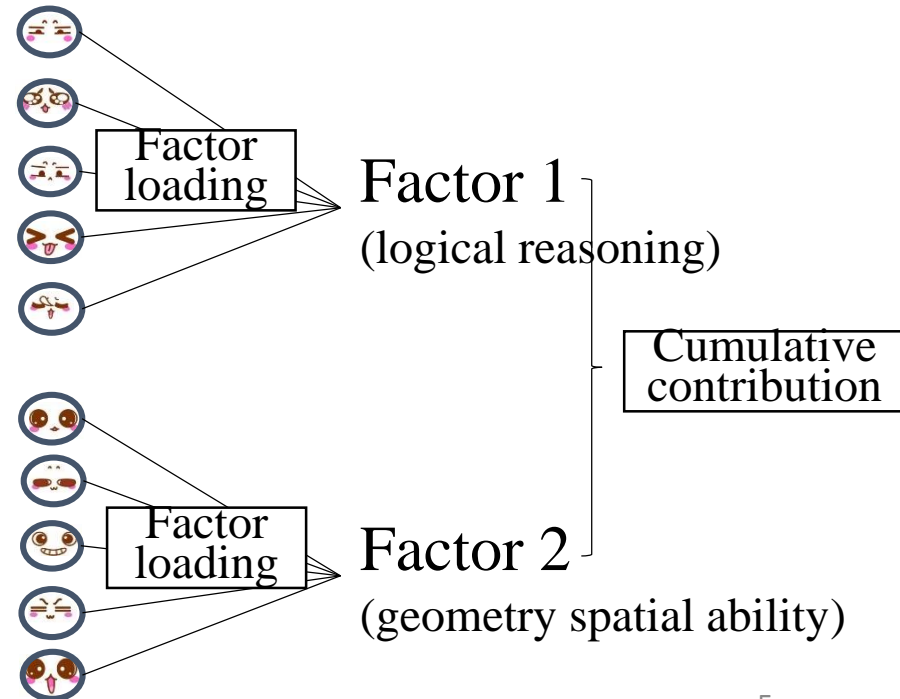
- 
- 
- 
- 
- 
- 
- 
- 
- 
- 

Already knew

As for continuous variable



✓ Common compound/linear transform
(simple structure)



Want to know

As for categorical variable

How about the **categorical variable**?

&

How to have a cluster of the **case**?



Latent class analysis(Lazarsfeld, 1950)

Want to know

Latent class analysis(LCA)

EMILIE GILBERT (UNIVERSITY OF YORK)

Latent class analysis

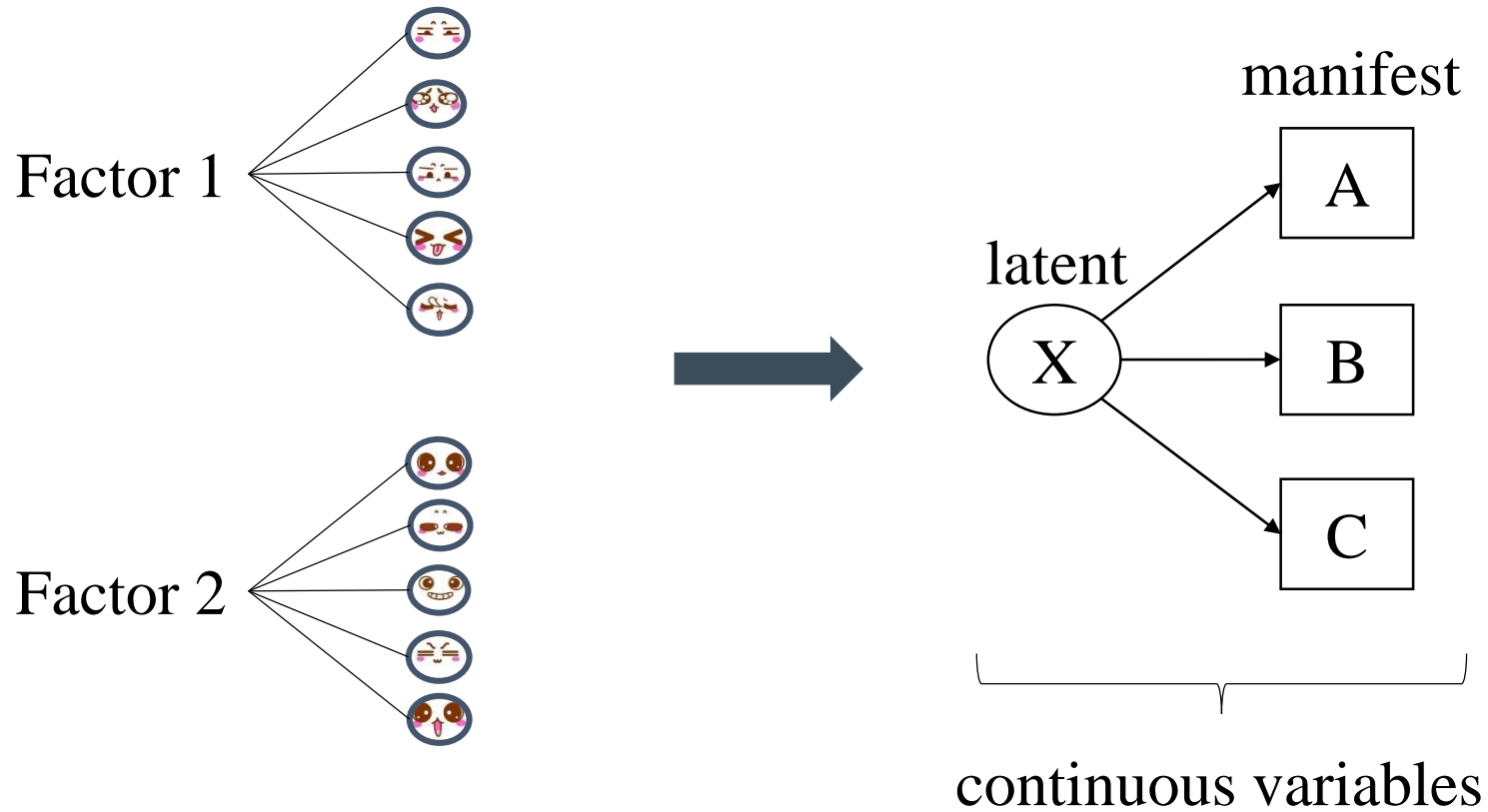
assumption

parameters

analysis process

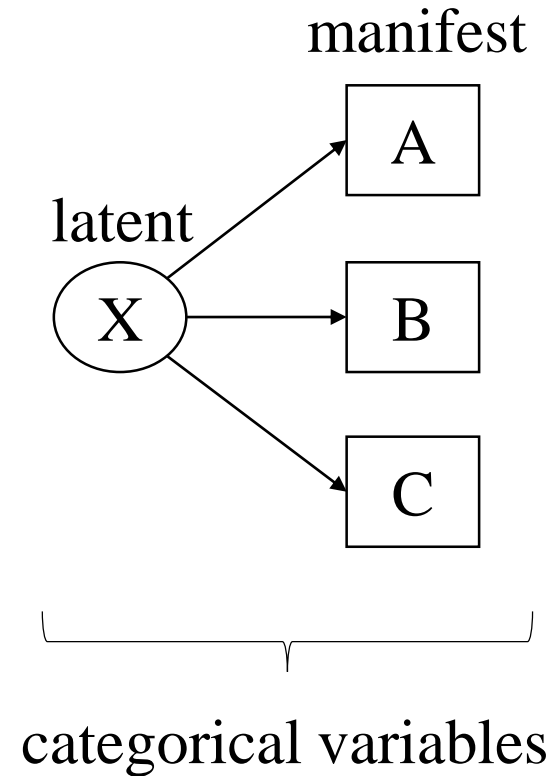
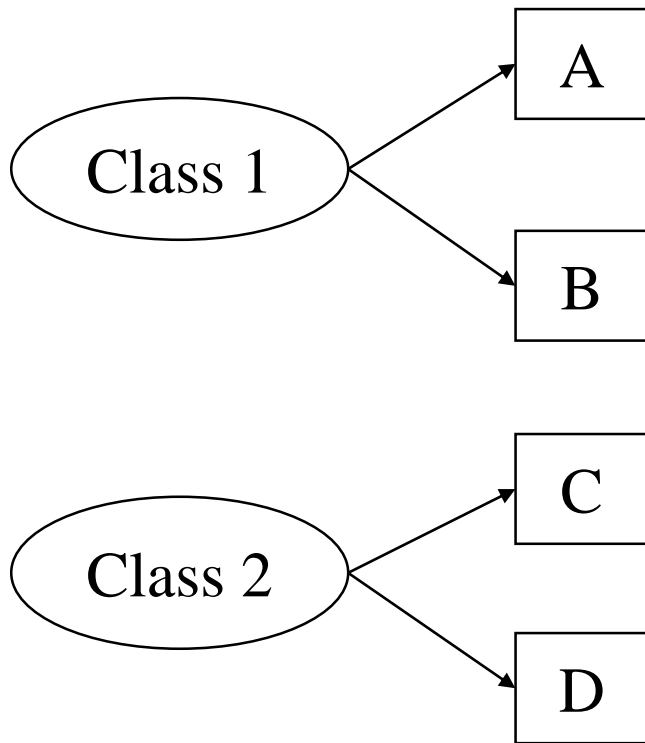
Want to know

Conceptualization & assumption



Want to know

Conceptualization & assumption



Want to know

Data table

Factor analysis

Person	Item				
	1	2	3	...	i
1	X_{11}	X_{12}	X_{13}	...	X_{1i}
2	X_{21}	X_{22}	X_{23}	...	X_{2i}
3	X_{31}	X_{32}	X_{33}	...	X_{3i}
...
p	X_{p1}	X_{p2}	X_{p3}	...	X_{pi}

Observed scores \rightarrow individual difference

Want to know

Data table

Latent class analysis

Response	Gender	
	male	female
yes	89 (58.17%)	64 (41.83%)
no	56 (37.09%)	95 (62.91%)

Frequency/probability → individual difference

Want to know

Assumption (Rolf Langeheine, 1988)

1 Class size

The population consists of m latent classes of unknown size W_j ($j = 1, m$).

2 Class-specific item probabilities

Within each class j , each item i has a **specific probability** of occurrence for each of its possible outcomes.

3 Local independence

Within **each class** j the manifest variables are postulated to be **independent**.

Want to know

Parameters (Rolf Langeheine, 1988)

Latent class probabilities

→ Cumulative contribution

The w 's, that is, unconditional probabilities of being in latent class j ;

Conditional probabilities

→ Factor loading

The p 's, that is, conditional probabilities of making a particular response on an item i .

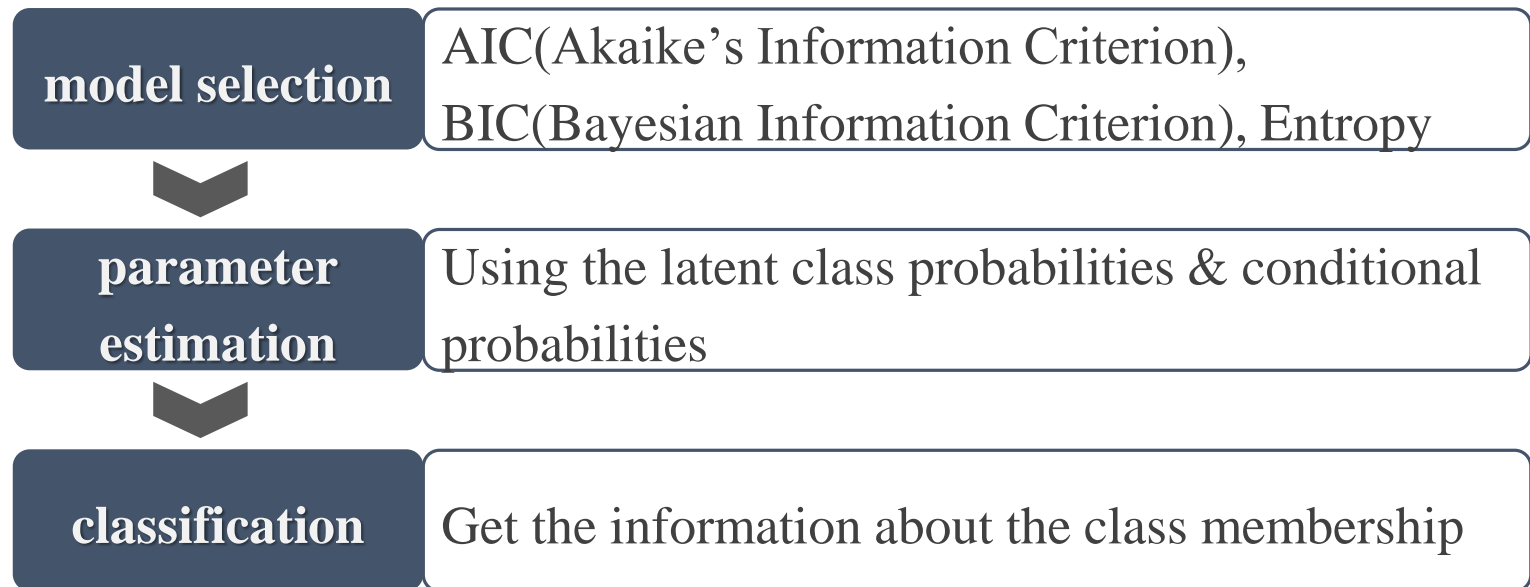


Maximum Likelihood (ML)

Want to know

Analysis process

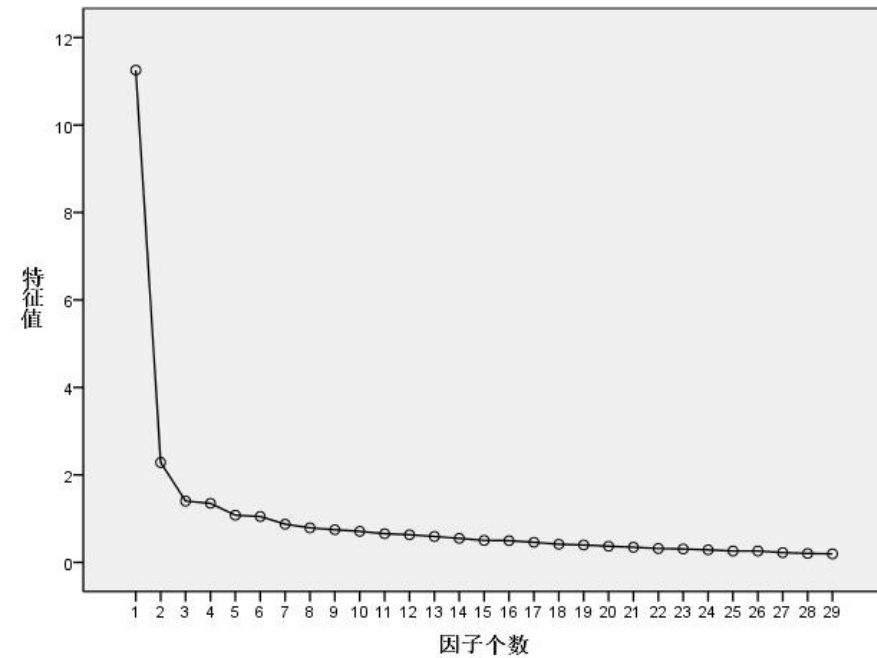
17/11/2022 10:00:22



Want to know

Model selection

FACTORS TO CONSIDER



FA

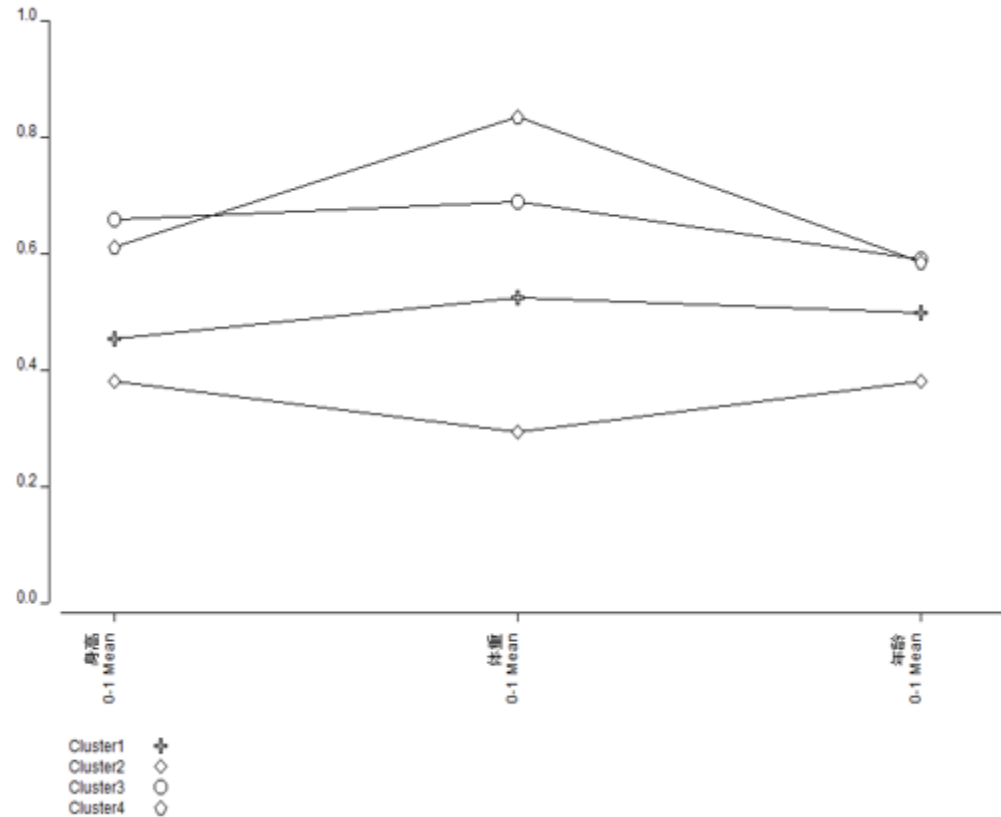
Model	AIC	BIC	Entropy
1	7063.259	7105.274	
2	6508.669	6596.201	0.928
3	6157.773	6290.821	0.942
4	6014.392	6192.956	0.934
5	5973.560	6197.641	0.888

LCA

→ Search for the inflection point & highest entropy

Want to know

Parameter estimation



- The **higher** value of the conditional probabilities indicates a **greater tendency**.

Want to know

Class membership

Using the posterior probabilities of the participant in the latent class

	Class 1	Class 2	Class 3
p_1	0.80	0.10	0.10
p_2	0.02	0.91	0.07



p_1 belongs to the class 1;

p_2 belongs to the class 2.

LatentGOLD

Create an SPSS data file

sexend.sav [数据集1] - SPSS Statistics 数据编辑器

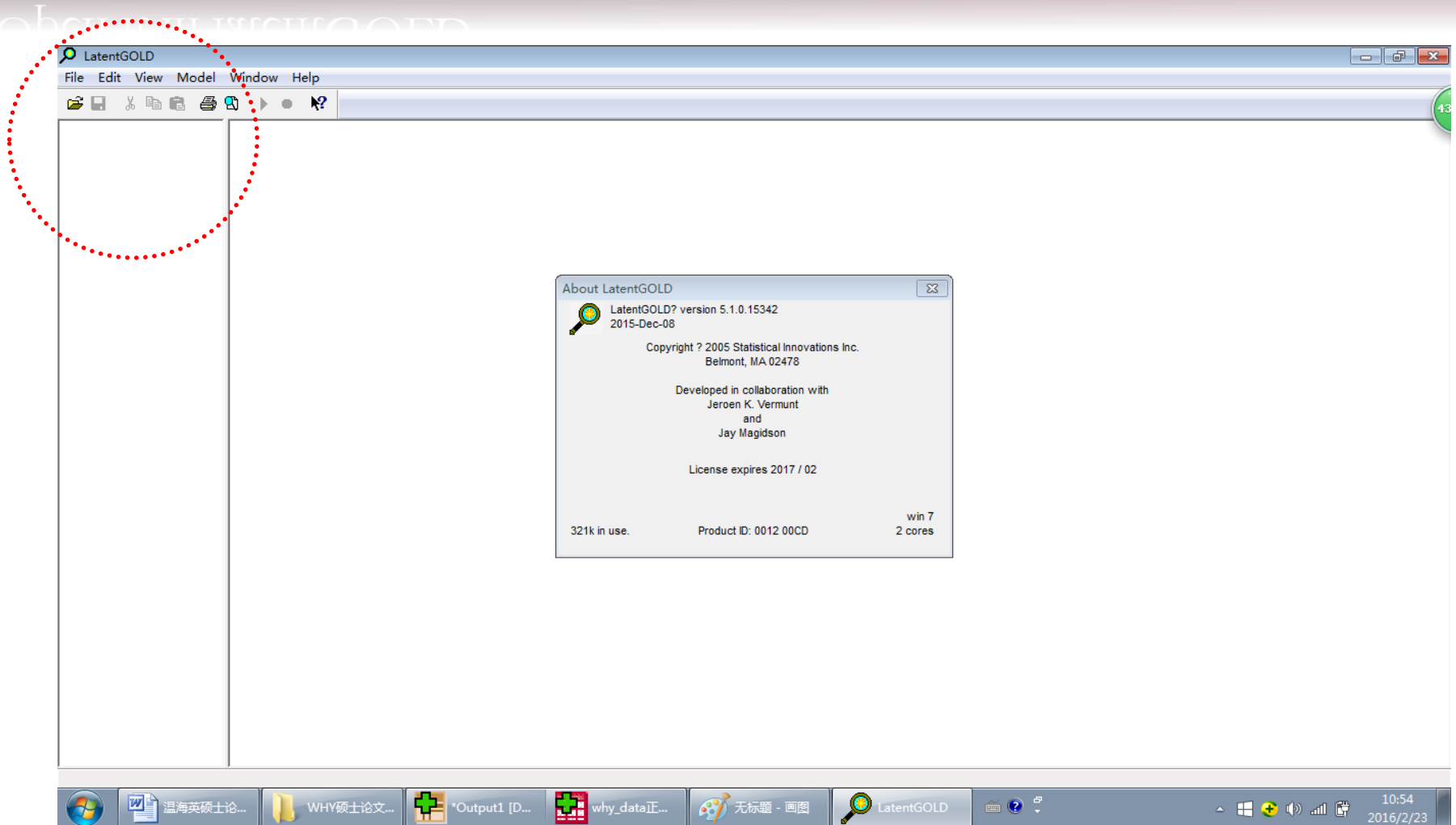
文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 图形(G) 实用程序(U) 附加内容(O) 窗口(W) 帮助

45 :

	item1	item2	item3	region	race	age	变量
31	3.0000	1.0000	2.0000	1	2	1	
32	1.0000	1.0000	4.0000	2	2	2	
33	1.0000	1.0000	1.0000	1	2	1	
34	2.0000	1.0000	4.0000	2	2	1	
35	3.0000	1.0000	1.0000	1	2	1	
36	2.0000	3.0000	1.0000	2	2	2	
37	1.0000	3.0000	1.0000	2	2	1	
38	2.0000	2.0000	4.0000	1	2	2	
39	3.0000	2.0000	2.0000	2	2	2	
40	2.0000	3.0000	1.0000	2	1	2	
41	3.0000	1.0000	2.0000	1	1	2	
42	1.0000	2.0000	4.0000	1	2	2	
43	2.0000	2.0000	4.0000	2	1	2	
44	3.0000	1.0000	4.0000	2	1	2	
45	2.0000	1.0000	2.0000	2	1	1	
46	2.0000	1.0000	2.0000	1	1	2	

LatentGOLD

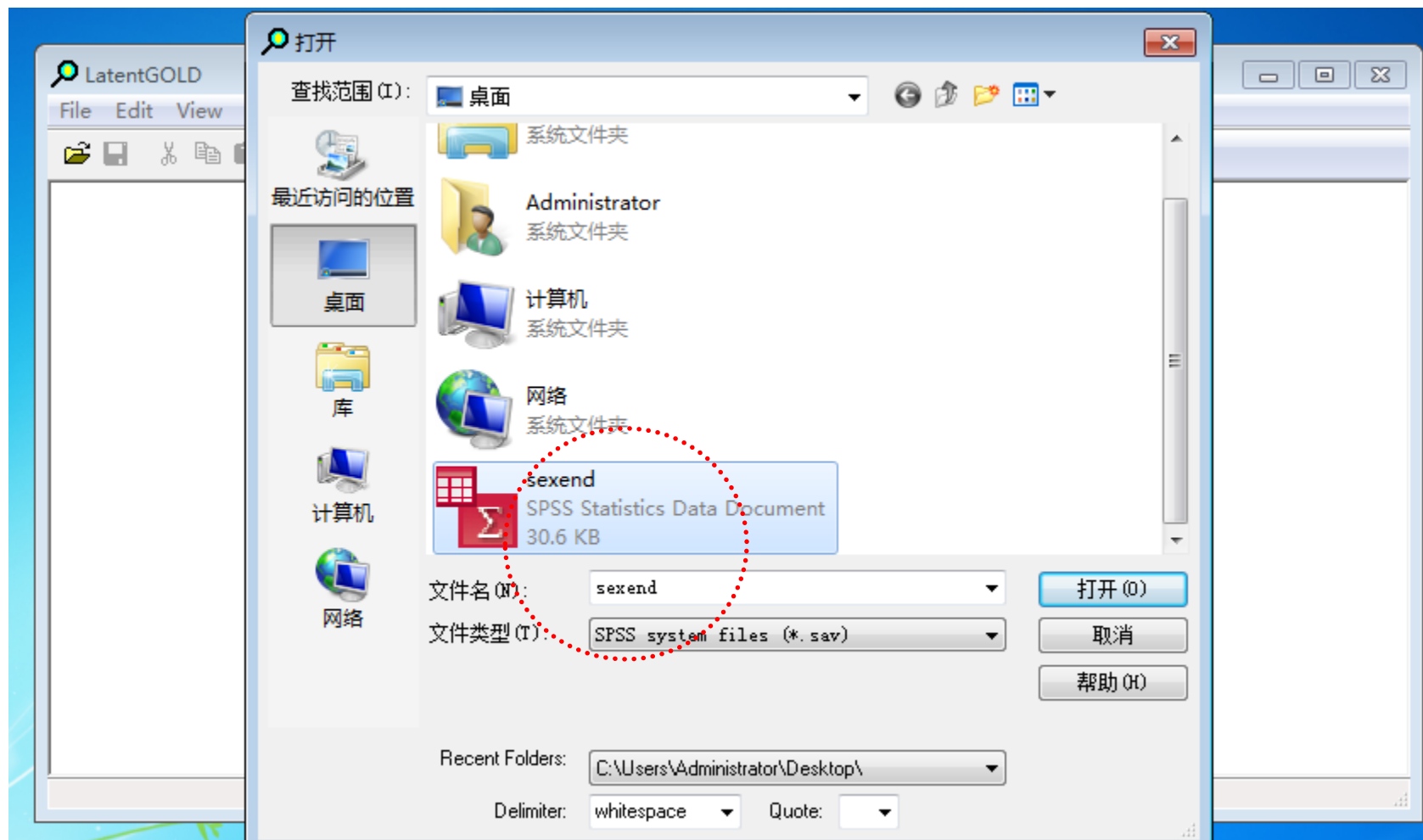
Open it in latentGOLD



LatentGOLD

Open it in latentGOLD

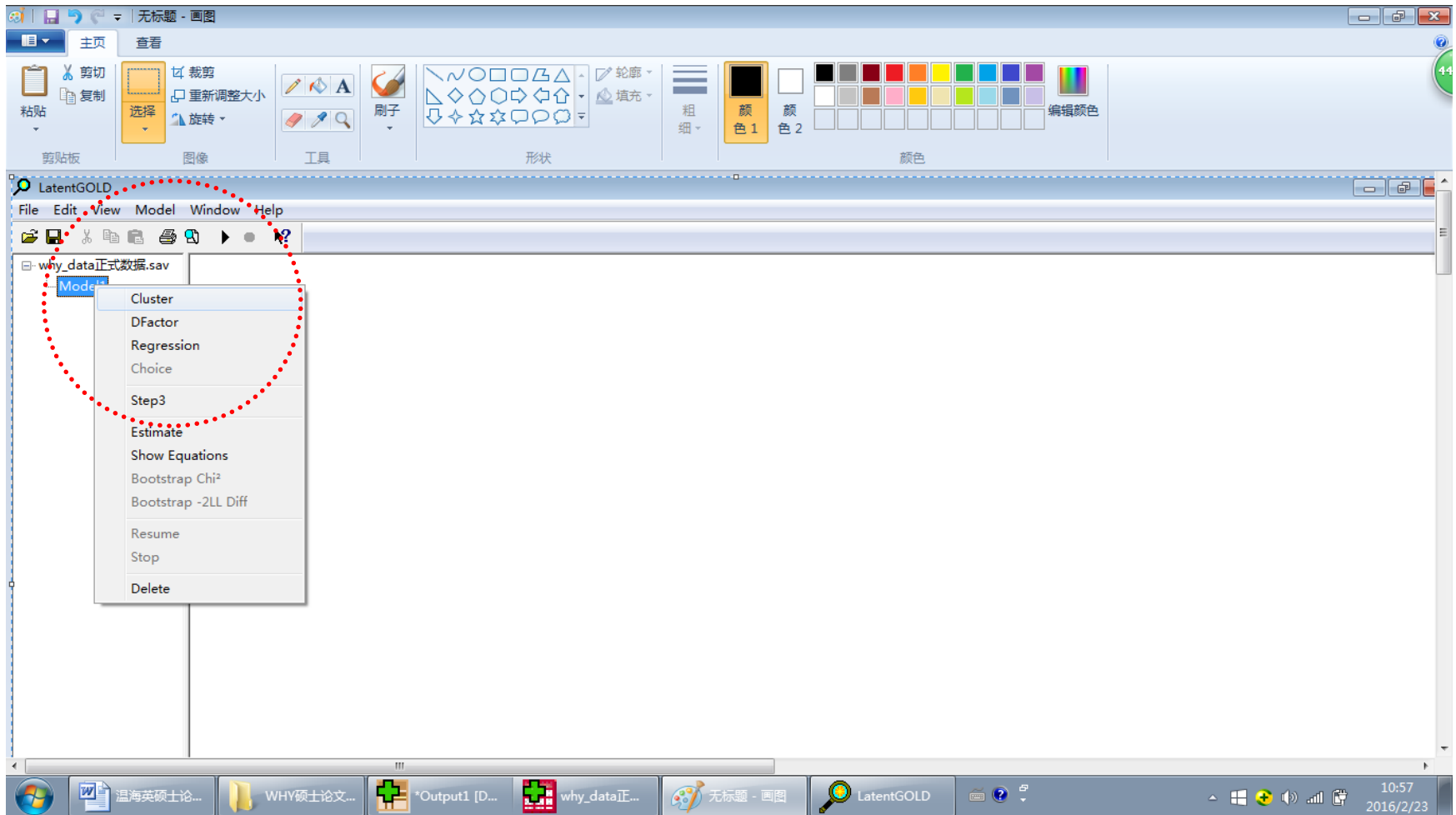
Open it in latentGOLD



LatentGOLD

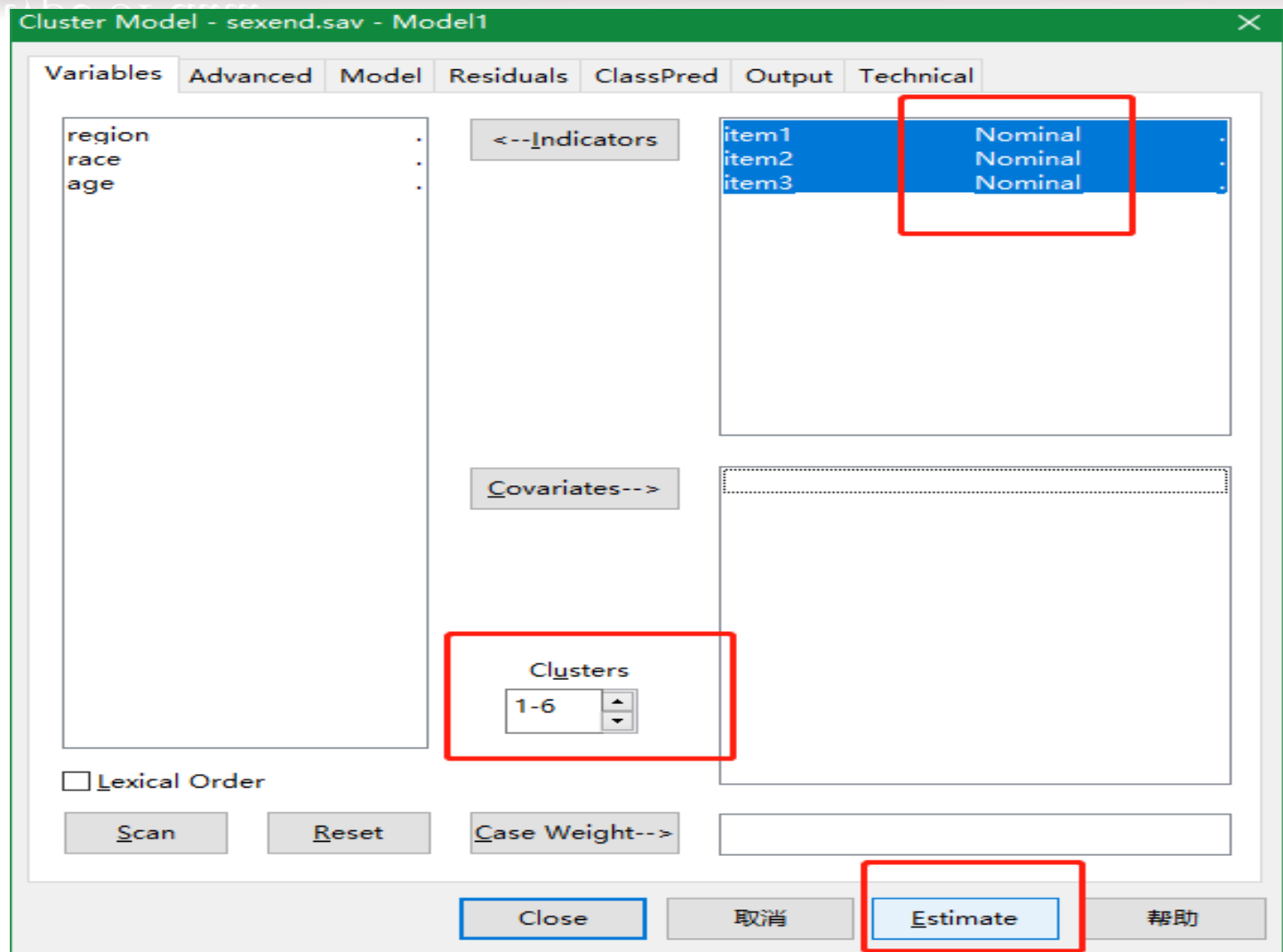
Choose cluster

CLUSTER CLUSTER



LatentGOLD

Choose type of data



LatentGOLD

Sort out the goodness of fit index

LatentGOLD

File Edit View Model Window Help

ny_data[正式数据].sav

- Model1 - LL = -5400.3289
- Model2 - LL = -4881.2537
- Model3 - LL = -4623.8585
- Model4 - LL = -4520.0271
- Model5 - LL = -4422.8116
- Model6 - LL = -4355.4191
- Model7 - LL = -4276.4303**
- Model8

7-Cluster Model

Number of cases	119
Number of parameters (Npar)	230
Activated Constraints	0
Random Seed	506062
Best Start Seed	506062

Log-likelihood Statistics

Log-likelihood (LL)	-4276.4303
Log-prior	-60.8953
Log-posterior	-4337.3256
BIC (based on LL)	9652.0591
AIC (based on LL)	9012.8607
AIC3 (based on LL)	9242.8607
CAIC (based on LL)	9882.0591
SABIC (based on LL)	8924.9401

Classification Statistics

Classification errors	0.0130
Reduction of errors (Lambda)	0.9838
Entropy R-squared	0.9824
Standard R-squared	0.9783
Classification log-likelihood	-4280.4301
Entropy	3.9997
CLC	8560.8601
AWE	11449.2569
ICL-BIC	9660.0585

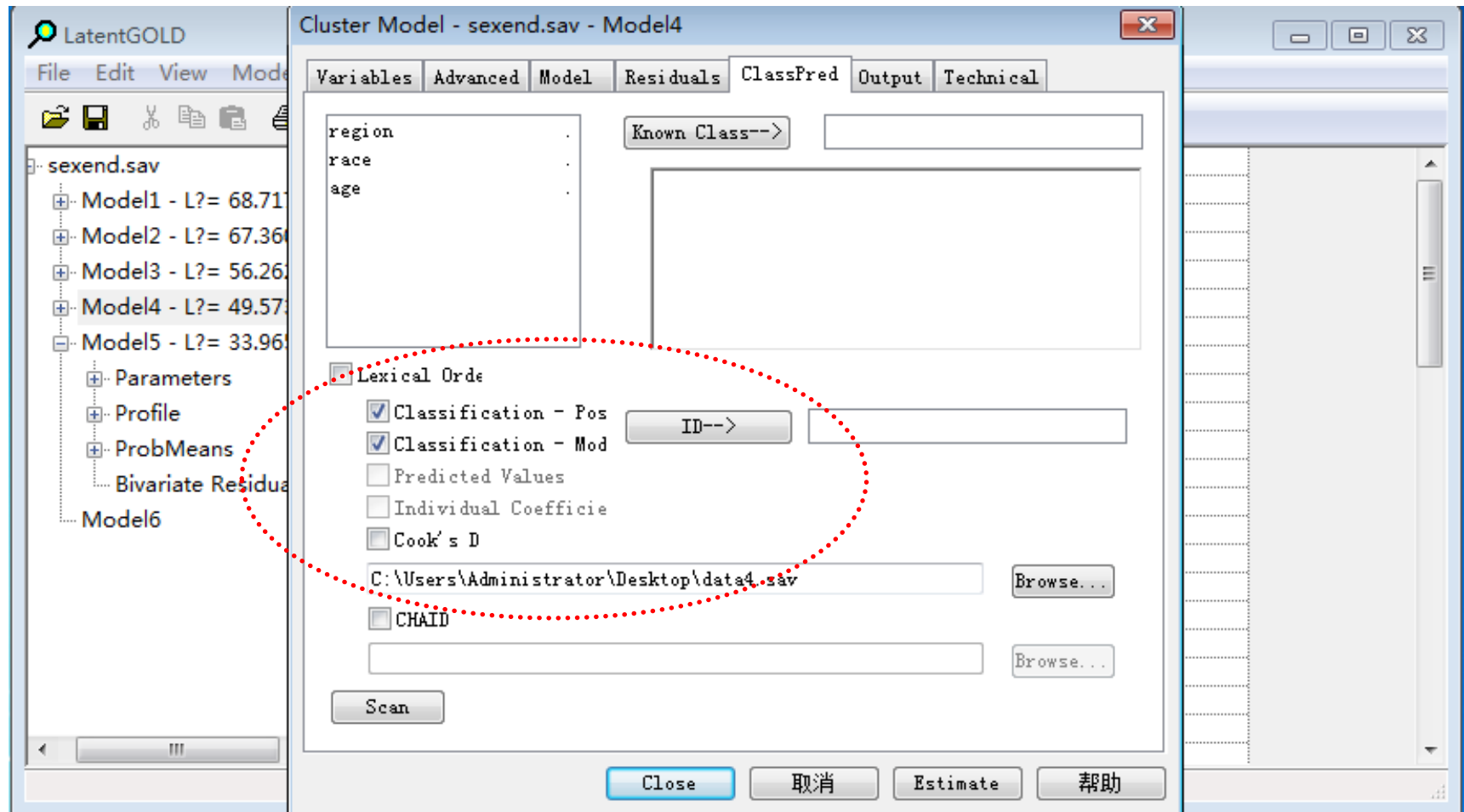
Classification Table

Latent	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7	Total
Cluster1	23.3559	0.0756	0.1911	0.0415	0.0000	0.0000	0.0000	23.6641
Cluster2	0.5092	22.8915	0.0000	0.0000	0.0000	0.0000	0.0074	23.4082
Cluster3	0.0689	0.0000	18.3910	0.0474	0.0754	0.0000	0.0000	18.5828
Cluster4	0.0660	0.0000	0.0035	14.9111	0.0000	0.0000	0.0000	14.9806
Cluster5	0.0000	0.0000	0.4143	0.0000	13.9238	0.0070	0.0000	14.3451
Cluster6	0.0000	0.0000	0.0000	0.0000	0.0007	13.9930	0.0000	13.9937
Cluster7	0.0000	0.0329	0.0000	0.0000	0.0000	0.0000	9.9926	10.0255
Total	24.0000	23.0000	19.0000	15.0000	14.0000	14.0000	10.0000	119.0000

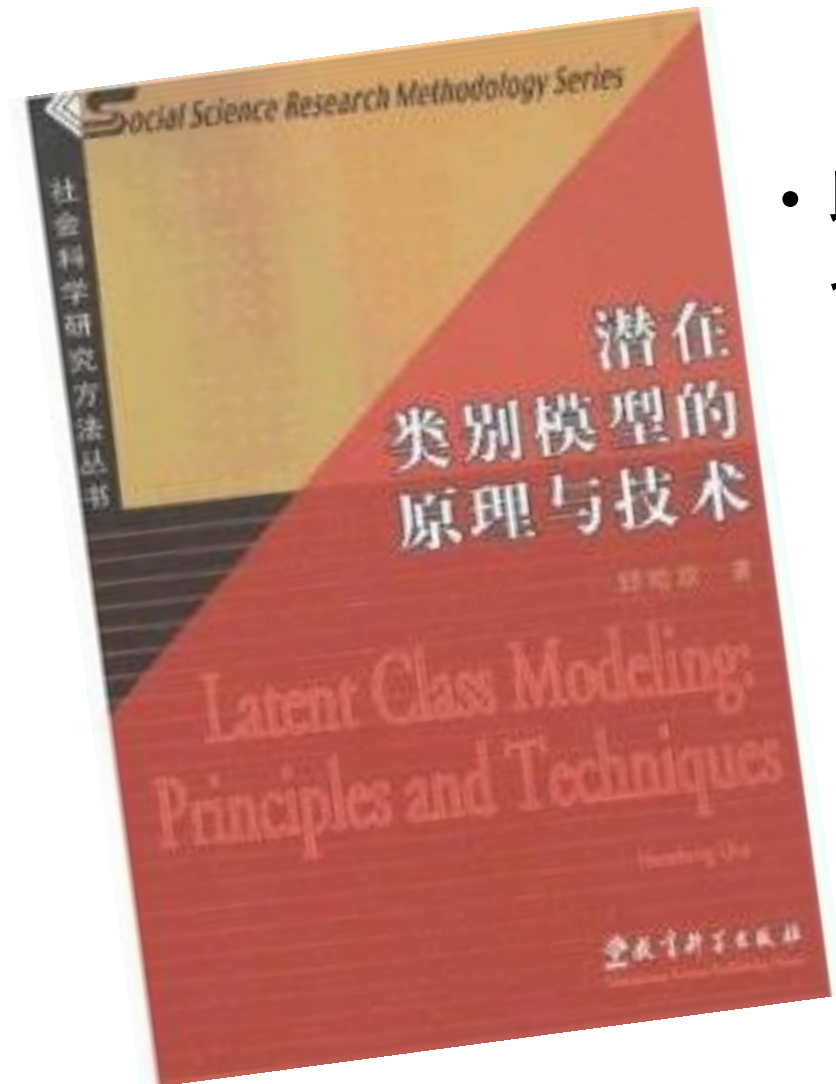
Taskbar: 温海英硕士论... WHY硕士论文... *Output1 [D... why_data正... 无标题 - 画图 LatentGOLD 11:01 2016/2/23

LatentGOLD

Produce an “SAV” file



Recommend



- 邱皓政. (2008). *潜在类别模型的原理与技术*. 北京: 教育科学出版社.

Thanks!