Latent Class Analysis

——from manifest variables to latent variables

Reporter: 黄颖诗

Contents

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

Related concepts

:-D

Manifest? though direct measurement/observation Latent?

though indirect measurement

Already knew

As for continuous variable

TIS TOT COMMINGUES VARIADIC



(10 items for math ability)

✓ Describe from 10 aspects

Already knew

As for continuous variable



(10 items for math ability)

Common compound/linear transform (simple structure)



As for categorical variable

The for calegorical variable

How about the categorical variable? & How to have a cluster of the case? Latent class analysis(Lazarsfeld, 1950)

Latent class analysis(LCA)

Datetit class aftaryors (DCT)

Latent class analysis assumption parameters analysis process

Conceptualization & assumption

Conceptualization & assumption



Conceptualization & assumption

Conceptualization & assumption



categorical variables

Data table

Factor analysis

D			Item		
Person	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}
•••	•••	•••	•••	•••	•••
р	X_{p1}	X_{p2}	X_{p3}		X_{pi}

Observed scores \rightarrow individual difference

Data table

Latent class analysis

Dermone	Gender						
Response	male	female					
Ves	89	64					
<i>y</i> c <i>s</i>	(58.17%)	(41.83%)					
no	56	95					
no	(37.09%)	(62.91%)					

$Frequency/probability \rightarrow individual \ difference$

1

2

3

Assumption(Rolf Langeheine, 1988)

7155umpuon(non Langeneme, 1700)

Class size

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



Within each class *j*, each item *i* has a specific probability

of occurrence for each of its possible outcomes.

Local independence

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

Parameters(Rolf Langeheine, 1988)

r arameuers(non rangeneme, 1700)



Analysis process

manysis process

model coloction	AIC(Akaike's Information Criterion),						
	BIC(Bayesian Information Criterion), Entropy						
parameter	Using the latent class probabilities & conditional						
estimation	probabilities						
classification	Get the information about the class membership						

Model selection

MODOL SCICCUOIL



 \rightarrow Search for the inflection point & highest entropy

Parameter estimation

1.0 0.8 0.6 0.4 0.2 0.0 e and a main and a main a m Main a m Main a m 体重 0-1 Mean 年間 0-1 Mean Clusterd

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

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.

Create an SPSS data file

CICALC ALLOL OD UALA THE

sexend.s	av [数据集1] - S	PSS Statistics	数据编辑器				
文件(E) 编辑	■(E) 祝图(⊻)		[]) 分析(A)	图形(<u>G</u>) 实用和	夏序(U) 附加内	春(0) 窗口(\/)	帮助
📂 🔜 🚉	📴 🕈 🔿	浩 📭 📑 🖌	Ma 🔸 📩	📰 🦺 📰 🤇 🤋	🐳 🕗 👟 👒	/	
45 :							
	item1	item2	item3	region	race	age	重变
31	3.0000	1.0000	2.0000		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	
	2 0000	2,0000	4 0000	4	4	2	

Open it in latentGOLD

Open it in latentGOLD

Open it in faterito OED

Choose cluster

CHOOSE CIUSICI

øi 🖬 🄊 (° 🕫	- 一无标题 - 画图							
■▼ 主页	查看							Q
	近保 ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓	 ▲ ▲ ▲ ▲	へ○□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□	→ → 粗 細→ 色1 色				44
剪贴板	图像	工具	形状		颜色			
LatentGOLD				0				
File Edit View	/ Model Window He	lp						
🖨 🖶 🕺 🗈	i 🖻 🎒 🅄 🕨 🔹	N?						E
- why_dataIFst	数据.sav Cluster DFactor Regression Choice Step3 Estimate Show Equations Bootstrap Chi ² Bootstrap -2LL Diff Resume Stop Delete							
								-
•)(III)	1 -			4
👩 📑	温海英硕士论 📘 V	VHY硕士论文	*Output1 [D <mark>计</mark> why_data正	🔊 无标题 - 画图	LatentGOLD	ã 🕐 🛱	- 🕂 🕀 📶 🕻	10:57

Choose type of data

Variables Advanced I	odel Residuals ClassPred Output Technical
region race age	 <u>I</u>ndicators item1 item2 Nominal Nominal
	<u>C</u> ovariates>
	Cl <u>u</u> sters
Lexical Order	

Sort out the goodness of fit index

Prior Cost Very data Product P	LatentGOLD	We have the									
Image: Second	ie Edit view wodei	window Heip									
Jobal Cluster Model Cluster Model Model 1. Ll = -500.329 Member of cases 119 Model 2. Ll = -482.8587 Random Seed 230 Model 1. Ll = -422.0021 Random Seed 96062 Model 2. Ll = -422.0126 Rendom Seed 96062 Model 1. Ll = -422.0121 Bert Start Seed 0 Model 2. Ll = -4325.4131 Log-interime (Hor) -278.433 Model 2. Ll = -432.614 Log-interime (Hor) -4278.433 Porobles Log-interime (Hor) -4278.433 Borsiat Residuals Extinated Solution	j 🛛 🕺 🗗 🗍 🗍	⊠ ▶ ● №?									
Model1 · LI = 5400.3287 Immber of cases 19 Immber of cases 19 Model2 · LI = 4831.2537 Activated Constraints 0 Immber of cases 19 Model3 · LI = 4420.0217 Best Start Seed 506062 Immber of cases 10 Immber of cases Model5 · LI = 4422.8115 Best Start Seed 506062 Immber of cases 10 Immber of cases Model5 · LI = 4422.8115 Log-sikelihood (LI) -4276.4503 Immber of cases Immber of cases Immber of cases Portine Portine Log-sikelihood (LI) 9276.4503 Immber of cases Immber of cases Portine Portine Seed Cont LI 992.0507 Immber of cases Immber of cases Bicrinate Residuals Estimate Residuals CAC (based on LL) 992.0507 Immber of cases Immber of cases Bicrinate Residuals CAC (based on LL) 992.0507 Immber of cases Immber of cases Immber of cases Model8 UL = 4975.310 Immber of cases Bicrinet Ovalues Not (based on LL) 992.0507 <	data正式数据.sav	7-Cluster Model									
Model2 Li Li Li Li Li Li Model2 Li -4223853 Member of parameters (lipar) 20	4odel1 - 11 = -5400 3289										
Models Lise Log Log <thlog<< td=""><td>A-J-12 11 - 4991 2527</td><td>Number of cases</td><td>119</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></thlog<<>	A-J-12 11 - 4991 2527	Number of cases	119								
Models Lackadde Activated Constraints 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	/iodel2 - LL = -4881.2557	Number of parameters (Npar)	230								
Macha Li	/lodel3 - LL = -4623,8585	Activated Constraints	0								
Bit of all 1 Dest start see Double Dest start see Double Model 6 L1 -4325.4131 Log-linelhood Statistics	Aodel4 - LL = -4520.0271	Random Seed	506062								
Model6 LL -4335.4191 Log-likelihood Statistics - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -	Nodel5 - LL = -4422.8116	Best Start Seed	506062								
Model7 L1 = 4276.4303 L0.9-init 4.08 865 Image: constraint of the second	Nodel6 - LL = -4355.4191	Log likelihood Statistics									
Parameters Log-prite 40.8853 Profile Log-posterior 4337 3256	Aodel7 - LL = -4276.4303	Log-likelihood (LL)	-4276 4303								
Log-posterior 437/3256 Image: Constraint of the second of	- Parameters	Log Internition (EE)	-60.8953								
Prolue BIC (based on LL) 99520591 Image: Construction of the second	Deefle	Log-posterior	-4337.3256								
Bry Problems AC (based on LL) 9012 8007 Image: Control of the second	Proille	BIC (based on LL)	9652.0591								
Bivariate Residuals Estimated/Jalues-Mode Ballic (based on LL) 9920 2807 Image: Construction of the construct	■ ProbMeans	AIC (based on LL)	9012.8607								
LestimatedValues-Mode CAIC (based on LL) 9892.0591 Medel8 SABIC (based on LL) 8924.9401 Classification Statistics Clusters	Bivariate Residuals	AIC3 (based on LL)	9242.8607								
ABBIC (based on LL) 8924 9401 Image: Construction of the second of the	EstimatedValues-Mode	CAIC (based on LL)	9882.0591								
Classification Statistics Clusters Clusters Clusters Clusters Clusters Clusters Classification errors (Lambda) 0.9338 C C C C C Reduction of errors (Lambda) 0.9338 C C C C C Entropy R-squared 0.9624 C C C C C Standard R-squared 0.9783 C C C C C C Classification log-likelihood -4280.4301 C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C	Nedel8	SABIC (based on LL)	8924.9401								
Classification Statistics Clusters O I I I I I Classification errors 0.0130 I I I I I Reduction of errors (Lambda) 0.9838 I I I I I Entropy R-squared 0.9824 I I I I I I Standard R-squared 0.9783 I I I I I I I Classification log-likelihood -4280.4301 I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I											
Classification errors 0.0130 0 0 0 0 0 0 Reduction of errors (Lambda) 0.9824 0.9824 0 0 0 0 Entropy R-squared 0.9824 0.9824 0 0 0 0 0 Standard R-squared 0.9733 0 0 0 0 0 0 Classification log-likelihood 4280.4301 0 0 0 0 0 0 0 Classification log-likelihood 4280.4301 0 0 0 0 0 0 0 0 Classification log-likelihood 4280.4301 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		Classification Statistics	Clusters								
Herduction of errors (Lamitoda) 0.9823		Classification errors	0.0130								
Littery Suparal 0.9783 Standard R-squared 0.9783		Entropy Requered	0.9030								
Classification log-likelihood 4220.4301 Entropy 3.9997		Standard P-squared	0.9024								
International Entropy 3.9997 Image: marked		Classification log-likelihood	-4280,4301								
CLC 8560.8601 Image: Constraint of the second of the seco		Entropy	3.9997								
AWE 11449.2569 Image: CL_BC 9660.055 Image: CL_BC 9660.055 Image: CL_BC 9660.055 Image: CL_BC Image:		CLC	8560.8601								
ICL-BIC 9660.0585 ICL ICL ICL 9660.0585 ICL ICL ICL ICL Classification Table Modal ICL ICL ICL ICL ICL ICL Latent Cluster1 Cluster2 Cluster3 Cluster4 Cluster6 Cluster6 Cluster7 Total Cluster2 0.0595 0.0765 0.1911 0.0415 0.0000 0.0000 23.6841 Cluster3 0.0689 0.0000 10.0001 0.0000 0.0000 18.5828 Cluster3 0.0689 0.0000 14.3911 0.0000 0.0000 14.3966 Cluster4 0.0600 0.0000 0.0000 13.2928 0.0000 14.3951 Cluster6 0.0000 0.0000 0.0000 0.0000 13.9930 13.9937 Cluster7 0.0000 0.0000 0.0000 0.0000 10.9937		AWE	11449.2569								
Classification Table Model Icon Clusters Clusters Clusters Clusters Clusters Classification Table Clusters Clusters Clusters Clusters Clusters Clusters Total Clusters Clusters Clusters Clusters Clusters Clusters Clusters Total Clusters 0.0000 22.8915 0.0000 0.0000 0.0000 23.6841 Clusters 0.0689 0.0000 13.3910 0.0474 0.0000 0.0000 18.5828 Clusters 0.0600 0.0000 0.0141 0.0000 13.923 0.0000 14.3810 Clusters 0.0000 0.0000 0.0000 13.923 0.0000 14.3810 Clusters 0.0000 0.0000 0.0000 13.923 0.0000 14.3937 Clusters 0.0000 0.0000 0.0000 0.0000 13.923 0.0000 13.9337		ICL-BIC	9660.0585								
Classification Table Model Model Cluster3 Cluster4 Cluster5 Cluster6 Cluster6 Cluster6 Latent Cluster1 Cluster3 Cluster3 Cluster4 Cluster5 Cluster6 Cluster6 Cluster6 Cluster6 Total Cluster1 23.3559 0.0756 0.1911 0.0415 0.0000 0.0000 23.6641 Cluster2 0.5092 22.8915 0.0000 0.0000 0.0000 18.528 Cluster3 0.0689 0.0000 13.3910 0.0474 0.0000 0.0000 14.3806 Cluster4 0.0600 0.0000 13.9238 0.0070 14.3816 Cluster5 0.0000 0.0000 0.0000 13.9238 0.0000 14.3816 Cluster6 0.0000 0.0000 0.0000 13.9238 0.0000 13.9337 Cluster6 0.0000 0.0000 0.0000 0.0000 13.9237 0.0255		01	N 1-1								
Cluster1			Modal	Cluster?	Cluster?	Cluster4	Cluster	Cluster	Cluster7	Total	
Cluster1 2.0.300 0.0010 0.0017 0.0000 0.0004 0.23.4082 Cluster2 0.0689 0.0000 18.3910 0.0017 0.0000 0.0004 23.4082 Cluster3 0.0689 0.0000 18.3910 0.0474 0.0000 0.0000 18.5828 Cluster4 0.0660 0.0000 0.0000 13.9238 0.0000 14.3906 Cluster5 0.0000 0.0000 0.0000 13.9238 0.0000 14.3931 Cluster6 0.0000 0.0000 0.0000 0.0000 13.9238 10.0001 13.9333 Cluster7 0.0000 0.0000 0.0000 0.0000 13.9333 10.0010 13.9337		Cluster1	23 3440	0.0756	0 1911	0.0415	0.0000	0 0000	0.0000	23.6641	
Cluster3 0.0689 0.0000 18.3910 0.0754 0.0000 18.5828 Cluster4 0.0660 0.0000 0.0351 14.9111 0.0000 0.0000 14.5828 Cluster5 0.0000 0.0000 0.4143 0.0000 13.9238 0.0000 14.3451 Cluster6 0.0000 0.0000 0.0000 0.0000 13.9393 0.0000 14.3451 Cluster6 0.0000 0.0000 0.0000 0.0000 13.9393 0.0000 13.9393 Cluster6 0.0000 0.0200 0.0000 0.0000 13.9393 10.0001 13.9374		Cluster?	0.5092	22 8915	0.0000	0.0000	0.0000	0.0000	0.0074	23.4082	
Cluster4 0.0660 0.0000 0.0005 14.9111 0.0000 0.0000 14.9806 Cluster5 0.0000 0.0000 0.4143 0.0000 13.9238 0.0070 14.3451 Cluster6 0.0000 0.0000 0.0000 0.0000 13.9330 0.0000 13.9337 Cluster7 0.0000 0.0329 0.0000 0.0000 0.0000 9.9926 10.0255		Cluster3	0.0689	0.0000	18,3910	0.0474	0.0754	0.0000	0.0000	18.5828	
Cluster5 0.0000 0.0141 0.0000 13.9238 0.0070 0.0000 14.3451 Cluster6 0.0000 0.0000 0.0000 0.0000 13.9238 0.0070 13.9937 Cluster7 0.0000 0.0329 0.0000 0.0000 0.0000 9.9926 10.0255		Cluster4	0.0660	0.0000	0.0035	14.9111	0.0000	0.0000	0.0000	14.9806	
Cluster6 0.0000 0.0000 0.0000 0.0000 13.9930 0.0000 13.9937 Cluster7 0.0000 0.0329 0.0000 0.0000 0.0000 9.9926 10.0255		Cluster5	0.0000	0.0000	0.4143	0.0000	13.9238	0.0070	0.0000	14.3451	
Cluster7 0.0000 0.0329 0.0000 0.0000 0.0000 0.0000 9.9926 10.0255		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 14.0000 14.0000 19.0000 19.0000		Total	24.0000	23.0000	19.0000	15.0000	14.0000	14.0000	10.0000	119.0000	

Produce an "SAV" file

rioduce all OTV IIIC

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!