LCA Examples

Lecture 13
10/16/06

Outline
Assumptions

\[
L(\pi, p \mid Y) = \prod_{i=1}^{N} \left[ \sum_{m=1}^{M} \pi_m \prod_{k=1}^{K} p_{km}^{y_{ik}} (1 - p_{km})^{(1-y_{ik})} \right]
\]

- Independent individuals (product over i)
- Internal Homogeneity (p_{km})
- Conditional Independence (product over k)

Examples of latent class analysis

- “Explore whether or not there are underlying clinical constructs that distinguish OCD-related subgroups”
Subjects and Methods

- OCD cases (n=80), control subjects (n=73) and their respective first degree relatives
- 450 subjects had complete data on all 10 relevant diagnoses
  - OCD, OC-personality disorder, tic disorder (TD), panic disorder or agoraphobia (PD/AG), generalized anxiety disorder (GAD), separation anxiety disorder (SAD), recurrent major depressive disorder (RMDD), hypochondriasis or body dysmorphic disorder (SOM), pathologic skin picking or nail biting (PSP/NB), anorexia nervosa or bulimia nervosa (ED)
  - Selected because of known association to OCD.
  - Some combined based on high redundancy
- Mplus used for model estimation
- Tried 2, 3, 4, 5 class models
- Based model choice on AIC and ‘scientific’ interpretation
Conclusion

• Best fitting model is four class structure. (1) Minimal disorder, (2) predominant RMDD and GAD, (3) “highly comorbid”, (4) PD/AG and TD. First three classes are ordinal, and the 4th class is qualitatively distinct.

Comments

• Some classes SMALL
• Did investigate construct validity by assigning individuals to classes and comparing to other variables.
• Some items seem perhaps not so related to classes (i.e. ‘bad items’) – TD, OCPD, ED
• Correlated individuals (family study)
Confirm Existing Depression Diagnoses


Purpose: To examine the DSM-III diagnostic criteria for major depression in two samples.

~90% of suicides in developed countries had a mental disorder at the time (Healthy People 2010 - SAMHSA)
ECA – The Sample

• 5 community mental health catchment areas
• Household epidemiological surveys
• Sample: Baltimore (3198)
• Symptoms 30 days prior to interview

Prevalence of Indicators in Baltimore

• Dysphoria – sad for 2 wks (4%)
• Appetite change – (6%)
• Sleep problems – (11%)
• Moving slowly or too much – (6%)
• Interest in sex – (2%)
• Tired out – (7%)
• Worthless – (3%)
• Concentration/Thinking (5%)
• Suicide – (9%)
# Model Fit

<table>
<thead>
<tr>
<th>Model</th>
<th>Chi-Square</th>
<th>DF</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independence</td>
<td>2416</td>
<td>511</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two class unrestricted</td>
<td>559</td>
<td>492</td>
<td>88</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two class restricted (C)</td>
<td>1045</td>
<td>493</td>
<td>97</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three class unrestricted</td>
<td>400</td>
<td>482</td>
<td>83</td>
<td>15</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Three class unrestricted (C)</td>
<td>436</td>
<td>483</td>
<td>86</td>
<td>12</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Three class restricted (BC)</td>
<td>995</td>
<td>485</td>
<td>97</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Four class unrestricted</td>
<td>376</td>
<td>474</td>
<td>82</td>
<td>14</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Restricted (C)</td>
<td>385</td>
<td>473</td>
<td>82</td>
<td>15</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Restrictions: (C) If a subject is in the depressed class, he must have dysphoria. (BC) Dysphoria is required for either of the non-normal classes.

## Model Selection

- Chi-Square Statistics used for comparing n+1 class model versus n class model
- For nested models, (I.e., when one model is a constrained version of another, you can take the difference in chi-square between the two models, and the df = difference in df between the two models.
A Look at Assignments

• Which symptom patterns get you into Class C (Major Depression)?
  – All patterns with dysphoria and 5 other symptoms were in class C and also would be considered Major Depression by DSM-III.
  – Certain patterns with dysphoria and only 4 other symptoms would be considered Depression by DSM-III but would not be in class C. All but one of these patterns includes sleep: sleep may not indicate serious depression
  – Certain patterns are in Class C but not DSM-III—all of these include concentration, suggesting it might be a more serious indicator of depression

Conclusions

• A three class model fits the observed data best, and the third class strongly resembles DSM-III Major Depression
• Certain symptoms may indicate either more or less serious depression based on the results of the latent class analysis: this might guide treatment
• Results were replicated in Raleigh-Durham, see paper for methods and details.
Latent Class Analysis of Lifetime Depressive Symptoms in the National Comorbidity Survey


- National Comorbidity Study
- N=2836 (excluded asymptomatic people)
- Used worst lifetime episode
  - At least two weeks
  - Impairment or help-seeking
  - One or more depression sx
- 6 class model with 14 indicators
Conditional Dependence

- Can motivate fitting of more classes than necessary
- Fix:
  - “joint” items (as in Eaton, 1989)
  - Polytomous indicators
  - Drop items (as in Netstadt, 2003)
  - Other methods being developed
Examine Subtypes in Different Groups

Genetics Section, Institute of Psychiatry, London

Purpose: To examine the types and prevalences of schizophrenia in men and women

Camberwell Register

- Area of London south of the Thames
- 447 first contact patients with psychotic illness from 1965 to 1984
Simultaneous Latent Class Models

Forces some classes to be all male, or all female, by setting the starting conditional probabilities of being male to either 0 or 1. Can also force “homogeneity” – I.e., force male and female classes to have the same conditional probabilities, and then see how the fit compares to unconstrained models.

Non-Homogenous Models

How many parameters? \((4-1)+(4*3)\)
Homogenous Models

How many parameters? (4-1)+(2*3)

Prevalence of Indicators

- Males (51%)
- Family History (8%)
- Restricted Affect (11%)
- Persecutory Delusion (77%)
- Poor Social Adjustment (37%)
- Dysphoria (48%)
- Early Onset (37%)
- Winter Birth (44%)
## Model Fit

<table>
<thead>
<tr>
<th>Model</th>
<th>Chi-Square</th>
<th>L-Square</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1. 1 Class per Gender, Total Homogeneity</td>
<td>465.18</td>
<td>325.85</td>
<td>247</td>
</tr>
<tr>
<td>M2. 1 Class per Gender Unconstrained</td>
<td>419.39</td>
<td>265.42</td>
<td>240</td>
</tr>
<tr>
<td>M3. 2 Classes per Gender, Total Homogeneity</td>
<td>264.18</td>
<td>222.99</td>
<td>238</td>
</tr>
<tr>
<td>M4. 2 Classes per Gender Unconstrained</td>
<td>250.41</td>
<td>176.96</td>
<td>225</td>
</tr>
<tr>
<td>M5. 3 Classes per Gender, Total Homogeneity</td>
<td>245.34</td>
<td>204.51</td>
<td>230</td>
</tr>
<tr>
<td>M6. 3 Classes per Gender Unconstrained</td>
<td>260.68</td>
<td>157.67</td>
<td>214</td>
</tr>
</tbody>
</table>

## Parameter Estimates M4

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Class 1</td>
<td>Class 2</td>
</tr>
<tr>
<td><strong>Prevalence</strong></td>
<td>0.22</td>
<td>0.29</td>
</tr>
<tr>
<td>Family History</td>
<td>0.11</td>
<td>0.03</td>
</tr>
<tr>
<td>Restricted Affect</td>
<td>0.30</td>
<td>0.06</td>
</tr>
<tr>
<td>Persecutory Delusions</td>
<td>0.65</td>
<td>0.83</td>
</tr>
<tr>
<td>Social Adjustment</td>
<td>0.72</td>
<td>0.29</td>
</tr>
<tr>
<td>Dysphoria</td>
<td>0.50</td>
<td>0.34</td>
</tr>
<tr>
<td>Early Onset</td>
<td>0.78</td>
<td>0.20</td>
</tr>
<tr>
<td>Winter Birth</td>
<td>0.36</td>
<td>0.51</td>
</tr>
</tbody>
</table>
Parameter Estimates M5

<table>
<thead>
<tr>
<th></th>
<th>Class 1</th>
<th>Class 2</th>
<th>Class 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prevalence-Men</strong></td>
<td>0.29</td>
<td>0.22</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Prevalence-Women</strong></td>
<td>0.13</td>
<td>0.23</td>
<td>0.13</td>
</tr>
<tr>
<td>Family History</td>
<td>0.10</td>
<td>0.07</td>
<td>0.01</td>
</tr>
<tr>
<td>Restricted Affect</td>
<td>0.22</td>
<td>0.03</td>
<td>0.00</td>
</tr>
<tr>
<td>Persecutory Delusions</td>
<td>0.61</td>
<td>0.88</td>
<td>0.93</td>
</tr>
<tr>
<td>Social Adjustment</td>
<td>0.60</td>
<td>0.20</td>
<td>0.23</td>
</tr>
<tr>
<td>Dysphoria</td>
<td>0.50</td>
<td>0.31</td>
<td>0.98</td>
</tr>
<tr>
<td>Early Onset</td>
<td>0.74</td>
<td>0.09</td>
<td>0.16</td>
</tr>
<tr>
<td>Winter Birth</td>
<td>0.41</td>
<td>0.51</td>
<td>0.28</td>
</tr>
</tbody>
</table>

Conclusions

- M5 is easier to interpret.
- Class 1 is characterized by + family history, early onset, restricted affect, poor social adjustment, and male:female ratio is 2:1
- Class 2 is characterized by persecutory delusions, winter birth, and similar prevalence in men and women.
- Class 3 is found predominantly in women, and is characterized by dysphoria, persecutory delusions.
4056 subjects aged 16-20 from 212 communities

Goal: identify subtypes of underage drinkers, in order to determine what characteristics indicate problem drinking

Indicators – individual, peer, and family factors.
Choosing number of classes

• Comparison of median absolute % difference between observed and expected odds ratios between indicators. (15.7%, 6.6%, 3.6%)
• Examination of classification error (6%, 16%, 21%)
• AIC
Comments

- Good sample size
- Good selection of number of classes
- Calculation of standard errors that take into account clustering.
- Cross-sectional study – what would make this analysis even better?

Reliability and Validity in LCA

- Large samples sizes = multi-site, many interviewers
- Does replication of results speak to reliability or validity?
- How to establish validity of classes?
  - Internal construct validity
  - External construct validity
- How might latent classes be used to establish validity of a new scale?