Mixture Modeling of Heterogeneity in Schizophrenia Data

Shane T. Jensen

Department of Statistics
The Wharton School
University of Pennsylvania

stjensen@wharton.upenn.edu

Collaborators: Donald B. Rubin and Mark F. Lenzenweger
Outline

1. Introduction to Schizophrenia and Schizotypes
2. Mixture Models for Discrete Measures
3. Application to Cornell Study
4. Mixture Models for Continuous Measures
5. Application to Westchester Study
6. Summary and Future Work
Schizophrenia and Schizotypes

- Schizophrenia is a mental illness characterized by:
  - early onset (< 30 years old)
  - thought disorder
  - hallucinations
  - social withdrawal
  - delusions
  - lack of emotional response

- Schizotypes have the tendency to become schizophrenic but have only shown “watered-down” symptoms.

- Psychologists are interested in the risk that schizotypic individuals will develop schizophrenia as well as prediction based on several measures of interest.

- Base rate for schizotypy in general population is believed to be somewhere between 10 and 15%.
First Dataset: Cornell Experiment

- Schizotypes detected using *Perceptual Aberration Survey (PAS)*
- Discrete performance measures for all classified schizotypes ($n_s = 24$) and random sample of classified normal ($n_n = 21$) were collected on several tasks:
  1. FMS: Failure to Maintain Set on Wisconsin card-sorting task
  2. ETD: Eye Tracking Dysfunction for following a moving target
  3. CPT: Reaction times requiring sustained attention to a stimulus
  4. TDI: Thought disordered responses in a verbal description
  5. Delayed Response: delay in response to a visual stimulus
PAS Classification Survey

• Individuals must answer 35 true or false questions related to possible past occurrences of abnormal visual, tactile, or auditory perceptions of their body or their environment.

• Often used to classify individuals since abnormal perceptions have been observed in individuals at risk for schizophrenia.

• All individuals scoring more than 2 SDs above the mean score were classified as schizotypic.

• We hypothesize that not all classified schizotypes are true schizotypes ie. classified schizotype group is a mixture of true schizotypes and true normals.
Theoretical Basis for a Mixture Model

1. “There is a strong theoretical basis for the overall distribution to contain a mixture of two components/classes.”

2. “This would correspond to persons carrying a genuine liability for schizophrenia versus those who do not carry it – it is essentially a strong claim regarding latent structure.”

3. “a single, unimodal latent distribution (no contain evidence of a mixture) would correspond to a simple severity continuum model that likely is subserved by a polygenic system”

- We will parse the mixture by using a subset of the collected performance measures while keeping other measures out of model for later validation
Mixture Model Assumptions I

- Classification is correct for all individuals classified as normal.
- Normal individuals misclassified as schizotypes have the same model on their performance measures as correctly classified normals.

[Diagram]

All Subjects

\[ \phi \]

Classified Normals

\[ 1 - \phi \] known

Classified Schizotypes

\[ 1 - \lambda \]

Normals

\[ \lambda \] unknown

True Schizotypes
• Performance on different tasks is assumed to be independent for normal individuals, while no independence is assumed for the schizotypic group

• Several other measures could have been used in the model, but the two measures FMS and ETD were considered to best fit this independence assumption in normal individuals

• Both FMS AND ETD are discrete measures: $FMS = 1, \ldots, 5$ and $ETD = 1, \ldots, 8$
Underlying Probability Model

- Each individual can be represented by a single count in either a true normal or schizotype table of FMS and ETD performance.

\[ N^*_{i,j} \sim \text{Multinomial}(\pi^N_{i,j}) \quad \text{Independence model on } \pi^N_{i,j} \]

\[ S^*_{i,j} \sim \text{Multinomial}(\pi^S_{i,j}) \quad \text{Saturated model on } \pi^S_{i,j} \]
Parameters of Interest

- Unknown parameters are $\lambda$, the proportion of true schizotypes among classified schizotypes, and $\pi^N, \pi^S$, the cell probabilities for the true normal and true schizotype tables.

- These parameters can not be directly estimated from our observed counts $N_{ij}, S_{ij}$, since we do not know which of the classified schizotypes are truly schizotypic.

- Estimating parameters would be easy if we had complete data ie. knowing which classified schizotypes are truly schizotypic.
Missing Data Framework

• We augment our observed data with $I$, a vector of missing indicator variables for true schizotypal status.

$$I_k = \begin{cases} 
1 & \text{if } k\text{-th subject true schizotypal} \\
0 & \text{if } k\text{-th subject true normal}
\end{cases}$$

• Filling in missing data $I$ gives us the complete data $N_{ij}^*$ and $S_{ij}^*$

• The Expectation-Maximization (EM) algorithm is very useful for computing maximum likelihood estimates of parameters in the presence of missing data.
Maximum Likelihood Estimation (EM)

- E-step takes expectation of missing $I_k$ for each individual $k$, given observed values and current parameter values:

$$E[I_k|\lambda, \pi, N, S] = \frac{P(FMS = i, ETD = j|I_k = 1) \cdot P(I_k = 1)}{P(FMS = i, ETD = j)} = \frac{\pi_{ij}^S \cdot (1 - \phi) \lambda}{\pi_{ij}^S \cdot (1 - \phi) \lambda + \pi_{ij}^N \cdot (\phi + (1 - \phi)(1 - \lambda))}$$

- These expectations are used to convert observed $N_{ij}$ and $S_{ij}$ into complete $N_{ij}^\star$ and $S_{ij}^\star$ by splitting classified schizotypes.

- M-step calculates MLE of parameters from complete data:

  1. $\hat{\pi}_{ij}^S = S_{ij}^\star / S^\star$ due to fully saturated model
  2. $\hat{\pi}_{ij}^N = (N_i^\star / N^\star_\cdot) \times (N^\star_j / N^\star_\cdot)$ due to independence model
  3. $\hat{\lambda} = \sum_k E[I_k] / n_s$ \textit{ie.} fraction of true schizotypes amongst the classified schizotypes.
### Maximum Likelihood Inference

- **Estimated proportion of true schizotypes** $\hat{\lambda} = 0.38$

#### Classified Normals

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.38</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.10</td>
</tr>
<tr>
<td>2</td>
<td>0.14</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>3</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
</tbody>
</table>

#### Classified Schizotypes

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.08</td>
<td>0.08</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.08</td>
</tr>
<tr>
<td>2</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>3</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>4</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>5</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>6</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
</tbody>
</table>

#### True Normals

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.21</td>
<td>0.07</td>
<td>0.09</td>
<td>0.01</td>
<td>0.13</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>2</td>
<td>0.16</td>
<td>0.05</td>
<td>0.07</td>
<td>0.01</td>
<td>0.09</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>3</td>
<td>0.04</td>
<td>0.01</td>
<td>0.02</td>
<td>0.00</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
</tbody>
</table>

#### True Schizotypes

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.10</td>
<td>0.10</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>2</td>
<td>0.21</td>
<td>0.10</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>3</td>
<td>0.04</td>
<td>0.02</td>
<td>0.02</td>
<td>0.00</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>4</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>5</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>6</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
</tbody>
</table>

- Algorithm converged to the same final values for several starting values, suggesting the likelihood is unimodal.
- No estimates of variability in results!
Bayesian Approach

- Parameters are random variables that follow prior distributions:
  \[ \lambda \sim \text{Unif}(0, 1) \quad \pi^S_{ij} \sim \text{Dirichlet}(\theta_{11}, \ldots, \theta_{IJ}) \quad \pi^N_{ij} \sim \text{Dirichlet}(\psi_{11}, \ldots, \psi_{IJ}) \]

- Prior is combined with the observed likelihood to form the posterior distribution:
  \[ \text{posterior} \propto \text{likelihood} \cdot \text{prior} \]
  \[ p(\lambda, \pi^N, \pi^S|N, S) \propto p(N, S|\lambda, \pi^N, \pi^S) \cdot p(\lambda, \pi^N, \pi^S) \]

- The posterior distribution \( p(\lambda, \pi^N, \pi^S|N, S) \) does not have a simple analytical form

- Our strategy is to estimate \( p(\lambda, \pi^N, \pi^S, I|N, S) \), where \( I \) is the missing vector of indicator variables for true schizotype status
Gibbs Sampler

- A Gibbs sampler can be used to obtain draws that converge to the joint posterior distribution $p(\lambda, \pi^N, \pi^S, I|N, S)$

- Iteratively sample from the conditional distribution of each parameter given all other parameters:
  
  1. $I_k|\pi^S_{ij}, \pi^N_{ij}, \lambda \sim \text{Bernoulli}(p)$ where $p = \frac{\pi^S_{ij} \cdot (1-\phi) \lambda}{\pi^S_{ij} \cdot (1-\phi) \lambda + \pi^N_{ij} \cdot (\phi + (1-\phi) (1-\lambda))}$
  2. Imputing the missing data I gives complete data $N^*_{ij}$ and $S^*_{ij}$
  3. $\lambda|I, N, S \sim \text{Beta}(S^*_{..} + 1, S_{..} - S^*_{..} + 1)$
  4. $\pi^S_{ij}|I, S \sim \text{Dirichlet}(S^*_{11} + \theta_{11}, \ldots, S^*_{IJ} + \theta_{IJ})$
  5. $\pi^N_{ij}|I, N = \alpha_i \times \beta_j$ where $\alpha_i \sim \text{Dirichlet}(N^*_1. + \psi_1., \ldots, N^*_{I.} + \psi_{I.})$ and $\beta_j \sim \text{Dirichlet}(N^*_{.1} + \theta_{.1}, \ldots, N^*_{.J} + \theta_{.J})$
Bayesian Inference

• Posterior estimates very similar to MLE estimates, but we now also have a measure of estimation variability

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>SD</th>
<th>95% P.I.</th>
<th>Parameter</th>
<th>Mean</th>
<th>SD</th>
<th>95% P.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal Cell (1,1)</td>
<td>0.2238</td>
<td>0.0567</td>
<td>( 0.1297 , 0.3443 )</td>
<td>Schizy Cell (1,1)</td>
<td>0.0000</td>
<td>0.0000</td>
<td>( 0.0000 , 0.0000 )</td>
</tr>
<tr>
<td>Normal Cell (2,1)</td>
<td>0.1512</td>
<td>0.0435</td>
<td>( 0.0749 , 0.2415 )</td>
<td>Schizy Cell (1,2)</td>
<td>0.0000</td>
<td>0.0000</td>
<td>( 0.0000 , 0.0000 )</td>
</tr>
<tr>
<td>Normal Cell (3,1)</td>
<td>0.0445</td>
<td>0.0225</td>
<td>( 0.0122 , 0.0981 )</td>
<td>Schizy Cell (2,3)</td>
<td>0.2282</td>
<td>0.1279</td>
<td>( 0.0301 , 0.5217 )</td>
</tr>
<tr>
<td>Normal Cell (1,2)</td>
<td>0.0733</td>
<td>0.0342</td>
<td>( 0.0240 , 0.1618 )</td>
<td>Schizy Cell (6,3)</td>
<td>0.1115</td>
<td>0.0966</td>
<td>( 0.0044 , 0.3567 )</td>
</tr>
<tr>
<td>Normal Cell (1,5)</td>
<td>0.1037</td>
<td>0.0374</td>
<td>( 0.0471 , 0.1827 )</td>
<td>Schizy Cell (2,5)</td>
<td>0.0000</td>
<td>0.0000</td>
<td>( 0.0000 , 0.0000 )</td>
</tr>
<tr>
<td>Normal Cell (2,2)</td>
<td>0.0498</td>
<td>0.0245</td>
<td>( 0.0148 , 0.1091 )</td>
<td>Schizy Cell (3,5)</td>
<td>0.0001</td>
<td>0.0024</td>
<td>( 0.0000 , 0.0000 )</td>
</tr>
<tr>
<td>Normal Cell (3,2)</td>
<td>0.0147</td>
<td>0.0101</td>
<td>( 0.0027 , 0.0394 )</td>
<td>Schizy Cell (4,6)</td>
<td>0.1182</td>
<td>0.1030</td>
<td>( 0.0034 , 0.3849 )</td>
</tr>
<tr>
<td>Normal Cell (2,8)</td>
<td>0.0803</td>
<td>0.0308</td>
<td>( 0.0304 , 0.1516 )</td>
<td>Schizy Cell (5,7)</td>
<td>0.1061</td>
<td>0.0947</td>
<td>( 0.0037 , 0.3480 )</td>
</tr>
<tr>
<td>Normal Cell (3,7)</td>
<td>0.0027</td>
<td>0.0030</td>
<td>( 0.0001 , 0.0111 )</td>
<td>Schizy Cell (1,8)</td>
<td>0.0000</td>
<td>0.0000</td>
<td>( 0.0000 , 0.0000 )</td>
</tr>
<tr>
<td>Normal Cell (3,8)</td>
<td>0.0236</td>
<td>0.0134</td>
<td>( 0.0055 , 0.0561 )</td>
<td>Schizy Cell (5,8)</td>
<td>0.1054</td>
<td>0.0917</td>
<td>( 0.0033 , 0.3370 )</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>0.3847</td>
<td>0.0949</td>
<td>( 0.2108 , 0.5730 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

• Confirming EM results, posterior cell probabilities have shrunk towards zero in many cells of the true schizotype table
Validation with other measures

• Posterior probabilities of true schizotypy were calculated for each individual that was used to re-classify schizotype group

![Graph showing posterior probabilities of true schizotypy](image)

• Examine new groups on other measures not in mixture model

<table>
<thead>
<tr>
<th>Measure</th>
<th>True Normals</th>
<th>False Positives</th>
<th>True Schizotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean  SD</td>
<td>Mean  SD</td>
<td>Mean  SD</td>
</tr>
<tr>
<td>CPT-IP Reaction Time</td>
<td>543.53 53.69</td>
<td>572.48 60.18</td>
<td>595.44 59.19</td>
</tr>
<tr>
<td>Delayed Response Task</td>
<td>93.26 5.40</td>
<td>90.85 7.95</td>
<td>86.65 8.70</td>
</tr>
<tr>
<td>Total Thought Disorder</td>
<td>2.57 3.17</td>
<td>5.79 8.27</td>
<td>9.50 15.01</td>
</tr>
</tbody>
</table>

• New classification has reduced heterogeneity in the other measures that were not even included in model
Second Dataset: Population Survey

- OMNI Inventory - self-report survey containing questions pertaining to pathological variables of schizotypy
  - cheaper than sending clinicians into general population
  - respondents can still be followed-up by clinicians
  - self-selection bias and problems with self-assessment

- Ads for survey were placed in weekly newspaper in Westchester/Putnam counties of New York as well as several other locations in the US (eg. Californian Naval Base).

- Dataset consists of 3146 adult respondents, 80% from Westchester/Putnam county in New York
Variables of interest are continuous positive scores on 9 diagnostic criteria for Schizotypal Personality Disorder.

Each test item scored in a directional manner such that low scores mean the absence of pathology and higher scores meant the presence of pathology.

Each variable was associated with several items in the OMNI inventory survey, and an individual's score for a particular variable is an average across these several questions.

Also have measures on several other variables: ethnicity, marital status, education, sex.
Demographic Variables

**SEX**

**ETHNIC**

**MARITAL**

**EDUCATE**

Shane T. Jensen

January 26, 2006
1. No Perspective: neutral events have special meaning
2. Odd Beliefs: e.g. you can change traffic lights
3. Unusual Perceptions: including bodily illusions
4. Thought Disorder: odd thinking and speech
5. Paranoia: Suspiciousness or paranoid ideation
6. Constricted affect: inappropriate or no emotional expression
7. Odd Behaviour: e.g. wearing a parka in July
8. Isolation: lack of close friends other than immediate family
9. Social Anxiety: fears that don’t diminish with familiarity
Pathology Variable Distributions

1. Perspective
2. Odd Beliefs 0.228
3. Perceptions 0.399 0.367
4. T.Disorder 0.449 0.149 0.384
5. Paranoia 0.549 0.243 0.327 0.340
6. C.Affect 0.081 -0.040 0.053 0.145 0.104
7. Odd Behave 0.379 0.309 0.363 0.345 0.348 0.348
8. Isolation 0.046 -0.032 0.029 0.083 0.135 0.247 -0.053
9. Anxiety 0.216 0.062 0.116 0.161 0.307 0.138 0.187 0.203

Shane T. Jensen

January 26, 2006
Multivariate Normal/T.Normal Mixture

• \( Y_i \) from \( i \)-th individual is a vector with \( k \) variables

• Two component multivariate mixture:

  1. Normals: \( Y_i \sim N_k(1, \Sigma_N) \cdot I(Y_i>1) \)

  2. Schizotypes: \( Y_i \sim N_k(\mu, \Sigma_S) \cdot I(Y_i>1) \)

• \( P(\text{schizotype}) = \lambda \) and \( P(\text{normal}) = 1 - \lambda \)
Major Model Issues

• Variable Selection - Dimension of mixture model
  – each variable modelled individually as univariate mixture?
  – all nine variables modelled together?
  – some subset of variables?

• Modelling Variance Parameters $\Sigma_S$ and $\Sigma_N$
  – unequal or equal marginal variances (diagonal entries) between normals and schizotype components?
  – Off-diagonal elements of $\Sigma_N$ should be zero since pathology variables shouldn’t be correlated in normal individuals
  – correlation or no correlation in schizotyes?
Implementation of Mixture Model

- Parameters of interest are $\lambda$, $\mu$, $\Sigma_S$ and $\Sigma_N$

- Missing Data: $I_i$ from individual $i$ is 1 if they are a schizotype

$$I_i = \begin{cases} 
1 & \text{wp } \lambda \\
0 & \text{wp } 1 - \lambda 
\end{cases}$$

- Maximum Likelihood estimation via EM algorithm

- Bayesian estimation of parameters via Gibbs sampling, with prior distributions

$$p(\lambda) = 1 \quad p(\mu, \Sigma_S) = \prod_{j=1}^{k} p(\mu_j, \sigma^2_{Sj}) = \prod_{j=1}^{k} \sigma^{-2}_{Sj} \quad p(\Sigma_N) = \prod_{j=1}^{k} p(\sigma^2_{Nj}) = \prod_{j=1}^{k} \sigma^{-2}_{Nj}$$
**EM Algorithm**

- **E-step:** calculate expected value of complete data loglikelihood conditional on observed data and current MLE estimates:

  \[
  E(I_i|Y_i, \hat{\theta}) = P(I_i = 1|Y_i, \hat{\theta}) = \frac{P(Y_i|I_i = 1, \hat{\theta}) \cdot P(I_i = 1|\hat{\theta})}{P(Y_i|I_i = 1, \hat{\theta}) \cdot P(I_i = 1|\hat{\theta}) + P(Y_i|I_i = 0, \hat{\theta}) \cdot P(I_i = 0|\hat{\theta})} = \frac{f_S(Y_i|\hat{\theta}) \cdot \hat{\lambda}}{f_S(Y_i|\hat{\theta}) \cdot \hat{\lambda} + f_N(Y_i|\hat{\theta}) \cdot (1 - \hat{\lambda})}
  \]

- **M-step:** fill in missing \( I_i \) with \( E(I_i|Y_i, \hat{\theta}) \) and calculate MLEs:

  \[
  \hat{\lambda} = \frac{\sum I_i}{n} \quad \hat{\mu}_j = \frac{\sum Y_{ij} \cdot I_i}{n_S} \quad \hat{\sigma}_{Sj}^2 = \frac{\sum (Y_{ij} - \hat{\mu}_j)^2 \cdot I_i}{n_S} \quad \hat{\sigma}_{Nj}^2 = \frac{\sum (Y_{ij} - 1)^2 \cdot (1 - I_i)}{n_N}
  \]

- Iterate between E-step and M-step until convergence

- **If we assume** \( \sigma_{Sj}^2 = \sigma_{Nj}^2 = \sigma_j^2 \), **then** \( \sigma_j^2 = (n_S \cdot \hat{\sigma}_{Sj}^2 + n_N \cdot \hat{\sigma}_{Nj}^2) / n \)

- **Does not account for truncation of schizotype component**
Gibbs Sampling Strategy

1. $I_i|\lambda, \mu, \sigma_S, \sigma_N, Y_i \sim \text{Bernoulli}(p_i)$

2. $\lambda|\mu, \sigma_S, \sigma_N, I, Y \sim \text{Beta}(n_S + 1, n_N + 1)$

3. $\mu_j|\sigma_{j_S}^2, I, Y \sim \text{Normal}(\sum_i Y_{ij} \cdot I_i/n_S, \sigma_{j_S}^2/n_S)$

4. $\sigma_{j_S}^2|\mu_j, I, Y \sim \text{Inv-}\chi^2(n_S, \sum_i (Y_{ij} - \hat{\mu}_j)^2 \cdot I_i)$

5. $\sigma_{N_j}^2|I, Y \sim \text{Inv-}\chi^2(n_N, \sum_i (Y_{ij} - 1)^2 \cdot (1 - I_i))$

- If we assume $\sigma_{j_S}^2 = \sigma_{N_j}^2 = \sigma_j^2$, then

$$\sigma_j^2 \sim \text{Inv-}\chi^2 \left( n, \sum_i (Y_{ij} - \hat{\mu}_j)^2 \cdot I_i + \sum_i (Y_{ij} - 1)^2 \cdot (1 - I_i) \right)$$
Univariate Mixture Models

• Constrain $\sigma_{Sj} = \sigma_{Nj}$ so that Schizotype tail probability is greater than tail probability for Normal individuals

• Very different values of $\lambda$ for different variables
Weakness of Univariate Model

- The proportion of schizotypes varies quite a bit across the different measures
- Individuals could be classified as schizotype based on some of their measures and as normal based on the rest
- This doesn’t make much sense since an individual should either be schizotype or normal
- Need to use information across variables for each individual
- Next step is to fit a multivariate mixture model using all nine pathology variables
Multivariate Mixture Model

- Several variables do not really seem to fulfill the mixture model assumptions
Variable Selection

• Find variables that are most valid and best fit assumptions

• 1st Tier: Perspective, Odd Beliefs, Perceptions, Isolation
  – pile up at lowest values, consistent with assumptions
  – easiest variables for people to assess about themselves

• 2nd Tier: Thought Disorder, Odd Behaviour, Social Anxiety
  – difficult to self-assess, usually measured by clinicians
  – many normals have high social anxiety

• 3rd Tier (Worst): Paranoia, Constricted Affect
  – very difficult for individual to assess their own paranoia
  – constricted affect not well assessed: only one OMNI question
Looking at different models

- focus on the 1st tier variables, but also fit models for the 2nd tier variables as well as both together

- fit both equal and unequal variance versions of each model

<table>
<thead>
<tr>
<th>Model</th>
<th>Mode</th>
<th>SD</th>
<th>Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Tier</td>
<td>0.16</td>
<td>0.01</td>
<td>(0.13, 0.18)</td>
</tr>
<tr>
<td>1st Tier Equal Variances</td>
<td>0.10</td>
<td>0.01</td>
<td>(0.07, 0.12)</td>
</tr>
<tr>
<td>2nd Tier</td>
<td>0.45</td>
<td>0.02</td>
<td>(0.42, 0.49)</td>
</tr>
<tr>
<td>2nd Tier Equal Variances</td>
<td>0.76</td>
<td>0.01</td>
<td>(0.74, 0.79)</td>
</tr>
<tr>
<td>Both Tiers</td>
<td>0.32</td>
<td>0.01</td>
<td>(0.29, 0.34)</td>
</tr>
<tr>
<td>Both Tiers Equal Variances</td>
<td>0.24</td>
<td>0.01</td>
<td>(0.20, 0.26)</td>
</tr>
</tbody>
</table>

- generally less difference between models based of same tier
Estimate $P(\text{schizotype})$ for each individual under each model.

Again, less difference between models within same tier.
• Take 1st tier, equal variance ($\lambda = 0.10$) as best model and partition individuals based on posterior probabilities
All Variables After Partitioning

- Look at the difference in the other pathology variables between putative normals and schizotypes.

- Normals and schizotypes well separated, even those not included in mixture model (interesting exception: Isolation).
Future Work

1. Model with schizotype variance forced to be larger than normal variance:
   \[ \sigma^2_{Sj} = \sigma^2_{Nj} + \tau_j \text{ where } \tau_j > 0 \]
   
   - constrains tail probabilities without forcing equal variances
   
   - psychologically valid?

2. Implement correlation within schizotype group
   
   - Metropolis-Hastings algorithm or WinBugs

3. Incorporate demographic covariates into mixture model
   
   - different vectors of regression coefficients between components
   
   - adjust estimates to fit overall US population

4. Account for truncation of schizotype component


http://stat.wharton.upenn.edu/~stjensen/