IDENTIFYING LOCAL DEPENDENCE WITH A SCORE TEST STATISTIC BASED ON THE BIFACTOR 2-PARAMETER LOGISTIC MODEL

Yang Liu    David Thissen

L. L. Thurstone Psychometric Laboratory
Department of Psychology
The University of North Carolina, Chapel Hill

July 21, 2011
1. Introduction

2. Simulation design

3. Results

4. Discussion
1 Introduction

2 Simulation design

3 Results

4 Discussion
Local dependence (LD) is the violation of local independence

(Strong) local independence (SLI, McDonald, 1981, 1982)
- Responses to different items are independent conditional on the latent variable(s) of interest

Practically, only test the violation of pairwise independence in some certain parametric forms
- Bifactor model (Gibbons & Hedecker, 1992)
- Threshold shift model (Glas & Suárez Falcón, 2003)
- LD $X^2$ statistic (Chen & Thissen, 1997)
Local dependence (LD) is the violation of local independence.

(Strong) local independence (SLI, McDonald, 1981, 1982)
- Responses to different items are independent conditional on the latent variable(s) of interest.

Practically, only test the violation of pairwise independence in some certain parametric forms:
- Bifactor model (Gibbons & Hedecker, 1992)
- Threshold shift model (Glas & Suárez Falcón, 2003)
- LD $X^2$ statistic (Chen & Thissen, 1997)
Local dependence (LD) is the violation of local independence

(Strong) local independence (SLI, McDonald, 1981, 1982)

- Responses to different items are independent conditional on the latent variable(s) of interest

Practically, only test the violation of pairwise independence in some certain parametric forms

- Bifactor model (Gibbons & Hedecker, 1992)
- Threshold shift model (Glas & Suárez Falcón, 2003)
- LD $X^2$ statistic (Chen & Thissen, 1997)
Local dependence (LD) is the violation of local independence.

(Strong) local independence (SLI, McDonald, 1981, 1982)
- Responses to different items are independent conditional on the latent variable(s) of interest.

Practically, only test the violation of pairwise independence in some certain parametric forms:
- Bifactor model (Gibbons & Hedecker, 1992)
- Threshold shift model (Glas & Suárez Falcón, 2003)
- LD $X^2$ statistic (Chen & Thissen, 1997)
For item pair $p$ and $q$, a secondary factor is added to the original 2-parameter logistic (2PL) model

$$T_p(x_p | \theta_1, \theta_2) = \frac{1}{1 + \exp[(-1)^{x_p}(a'_p \theta_1 + a_{pq} \theta_2 + c_p)]}$$

$$T_q(x_q | \theta_1, \theta_2) = \frac{1}{1 + \exp[(-1)^{x_q}(a'_q \theta_1 \pm a_{pq} \theta_2 + c_q)]}$$

- Identification constraint: equal absolute value for the secondary slopes
  - Same sign: positive LD
  - Opposite sign: negative LD
- Model for underlying local dependence (ULD, Thissen et al., 1992)
- Testing hypotheses $H_0 : a_{pq} = 0$ vs. $H_{1,B} : a_{pq} \neq 0$
For item pair $p$ and $q$, a secondary factor is added to the original 2-parameter logistic (2PL) model

$$T_p(x_p|\theta_1, \theta_2) = \frac{1}{1 + \exp[(-1)^x_p(a_p' \theta_1 + a_{pq} \theta_2 + c_p)]}$$

$$T_q(x_q|\theta_1, \theta_2) = \frac{1}{1 + \exp[(-1)^x_q(a_q' \theta_1 \pm a_{pq} \theta_2 + c_q)]}$$

Identification constraint: equal absolute value for the secondary slopes
- Same sign: positive LD
- Opposite sign: negative LD

Model for underlying local dependence (ULD, Thissen et al., 1992)

Testing hypotheses $H_0 : a_{pq} = 0$ vs. $H_{1,B} : a_{pq} \neq 0$
For item pair \( p \) and \( q \), a secondary factor is added to the original 2-parameter logistic (2PL) model

\[
T_p(x_p|\theta_1, \theta_2) = \frac{1}{1 + \exp[(-1)^{x_p}(a'_p \theta_1 + a_{pq} \theta_2 + c_p)]}
\]

\[
T_q(x_q|\theta_1, \theta_2) = \frac{1}{1 + \exp[(-1)^{x_q}(a'_q \theta_1 \pm a_{pq} \theta_2 + c_q)]}
\]

Identification constraint: equal absolute value for the secondary slopes
- Same sign: positive LD
- Opposite sign: negative LD

Model for underlying local dependence (ULD, Thissen et al., 1992)

Testing hypotheses \( H_0 : a_{pq} = 0 \) vs. \( H_{1,B} : a_{pq} \neq 0 \)
For item pair $p$ and $q$, a secondary factor is added to the original 2-parameter logistic (2PL) model

$$T_p(x_p | \theta_1, \theta_2) = \frac{1}{1 + \exp[(-1)^{x_p}(a'_p \theta_1 + a_{pq} \theta_2 + c_p)]]}$$

$$T_q(x_q | \theta_1, \theta_2) = \frac{1}{1 + \exp[(-1)^{x_q}(a'_q \theta_1 \pm a_{pq} \theta_2 + c_q)]}$$

Identification constraint: equal absolute value for the secondary slopes
- Same sign: positive LD
- Opposite sign: negative LD

Model for underlying local dependence (ULD, Thissen et al., 1992)

Testing hypotheses $H_0 : a_{pq} = 0$ vs. $H_{1,B} : a_{pq} \neq 0$
Answering the first item (i.e. item $p$) correctly facilitates the response to the second item (i.e. item $q$)

$$T_q(x_q|\theta_1, x_p) = \frac{1}{1 + \exp[(-1)^x_q(a'_q \theta_1 + c_q + \delta_{pq} x_p)]]}$$

Model for surface local dependence (SLD, Thissen et al., 1992)

Testing hypotheses $H_0 : \delta_{pq} = 0$ vs. $H_{1,T} : \delta_{pq} \neq 0$
Answering the first item (i.e. item $p$) correctly facilitates the response to the second item (i.e. item $q$)

$$T_q(x_q|\theta_1,x_p) = \frac{1}{1 + \exp\left((1)^x_q(a'_q \theta_1 + c_q + \delta_{pq}x_p)\right)}$$

Model for surface local dependence (SLD, Thissen et al., 1992)

Testing hypotheses $H_0 : \delta_{pq} = 0$ vs. $H_{1,T} : \delta_{pq} \neq 0$
Answering the first item (i.e. item $p$) correctly facilitates the response to the second item (i.e. item $q$)

$$T_q(x_q|\theta_1, x_p) = \frac{1}{1 + \exp[(-1)^x_q(a'_q \theta_1 + c_q + \delta_{pq} x_p)]}$$

Model for surface local dependence (SLD, Thissen et al., 1992)

Testing hypotheses $H_0: \delta_{pq} = 0$ vs. $H_{1,T}: \delta_{pq} \neq 0$
Consider the 2-way marginal table for item $p$ and $q$ as a multinomial model: $\text{Multinom}(4; \pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})$

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\pi_{00}$</td>
<td>$\pi_{01}$</td>
</tr>
<tr>
<td>1</td>
<td>$\pi_{10}$</td>
<td>$\pi_{11}$</td>
</tr>
</tbody>
</table>

Testing hypotheses $H_0: \pi_{xp|xq} = \pi_{x^*p|xq}^*$ vs. $H_{1,M}: \pi_{xp|xq} \neq \pi_{x^*p|xq}^*, \forall x_p, x_q$

Expected cell probability: $\pi_{x^*p|xq}^* = \int_{\theta_1} T_p(x_p|\theta_1)T_q(x_q|\theta_1)\phi(\theta_1)d\theta_1$

Pearson’s $X^2 = N \sum_{x_p=0}^1 \sum_{x_q=0}^1 \frac{(\hat{\pi}_{xp|xq} - \pi_{x^*p|xq}^*)^2}{\pi_{x^*p|xq}^*}$

Observed cell probability: $\hat{\pi}_{xp|xq}$

Chen & Thissen (1997) suggested to use $\chi_1^2$ as an approximation of the null distribution.
Consider the 2-way marginal table for item $p$ and $q$ as a multinomial model: $\text{Multinom}(4; \pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})$

<table>
<thead>
<tr>
<th>Item $q$</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\pi_{00}$</td>
<td>$\pi_{01}$</td>
</tr>
<tr>
<td>1</td>
<td>$\pi_{10}$</td>
<td>$\pi_{11}$</td>
</tr>
</tbody>
</table>

- Testing hypotheses $H_0 : \pi_{x_px_q} = \pi^*_{x_px_q}$ vs. $H_{1,M} : \pi_{x_px_q} \neq \pi^*_{x_px_q}, \forall x_p, x_q$

- Expected cell probability: $\pi^*_{x_px_q} = \int_{\theta_1} T_p(x_p|\theta_1)T_q(x_q|\theta_1)\phi(\theta_1)d\theta_1$

- Pearson’s $X^2 = N \sum_{x_p=0}^{1} \sum_{x_q=0}^{1} \frac{(\hat{\pi}_{x_px_q} - \pi^*_{x_px_q})^2}{\pi^*_{x_px_q}}$

- Observed cell probability: $\hat{\pi}_{x_px_q}$

- Chen & Thissen (1997) suggested to use $\chi^2_1$ as an approximation of the null distribution.
Consider the 2-way marginal table for item $p$ and $q$ as a multinomial model: $\text{Multinom}(4; \pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})$

<table>
<thead>
<tr>
<th>Item $p$</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\pi_{00}$</td>
<td>$\pi_{01}$</td>
</tr>
<tr>
<td>1</td>
<td>$\pi_{10}$</td>
<td>$\pi_{11}$</td>
</tr>
</tbody>
</table>

Testing hypotheses $H_0: \pi_{x_p x_q} = \pi_{x_p x_q}^*$ vs. $H_{1,M}: \pi_{x_p x_q} \neq \pi_{x_p x_q}^*, \forall x_p, x_q$

Expected cell probability: $\pi_{x_p x_q}^* = \int_{\theta_1} T_p(x_p | \theta_1) T_q(x_q | \theta_1) d\theta_1$

Pearson's $X^2 = N \sum_{x_p=0}^{1} \sum_{x_q=0}^{1} \frac{(\hat{\pi}_{x_p x_q} - \pi_{x_p x_q}^*)^2}{\pi_{x_p x_q}^*}$

Chen & Thissen (1997) suggested to use $\chi^2_1$ as an approximation of the null distribution.
Consider the 2-way marginal table for item $p$ and $q$ as a multinomial model: $\text{Multinom}(4; \pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})$

<table>
<thead>
<tr>
<th>Item $p$</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\pi_{00}$</td>
<td>$\pi_{01}$</td>
</tr>
<tr>
<td>1</td>
<td>$\pi_{10}$</td>
<td>$\pi_{11}$</td>
</tr>
</tbody>
</table>

- Testing hypotheses $H_0: \pi_{x_p x_q} = \pi^*_{x_p x_q}$ vs. $H_{1,M}: \pi_{x_p x_q} \neq \pi^*_{x_p x_q}$, $\forall x_p, x_q$

- Expected cell probability: $\pi^*_{x_p x_q} = \int_{\theta_1} T_p(x_p|\theta_1)T_q(x_q|\theta_1)\phi(\theta_1)d\theta_1$

- Pearson’s $X^2 = N \sum_{x_p=0}^{1} \sum_{x_q=0}^{1} \left( \hat{\pi}_{x_p x_q} - \pi^*_{x_p x_q} \right)^2 / \pi^*_{x_p x_q}$

- Observed cell probability: $\hat{\pi}_{x_p x_q}$

- Chen & Thissen (1997) suggested to use $\chi_1^2$ as an approximation of the null distribution
Score test (Rao, 1948; a.k.a. Lagrange multiplier test)

1. Get the restricted (i.e. in the constrained parameter space implied by $H_0$) maximum likelihood estimates of all parameters $\hat{\eta}_0$—equivalent to fitting a locally independent 2PL model.

2. Evaluate the gradient $\nabla(\eta)$ and Hessian $H(\eta)$ using the estimates from step 1—gradient vector and Hessian matrix based on the LD models need to be derived.

3. Compute score test statistic $S = \nabla'(\hat{\eta}_0)H^{-1}(\hat{\eta}_0)\nabla(\hat{\eta}_0) \xrightarrow{\mathcal{D}} \chi_1^2$

Related issues in the current project

- For the bifactor model, the first derivative w.r.t. $a_{pq}$ is 0 at $a_{pq} = 0$; as a result, the score test statistics is evaluated at 0.0001 instead of 0.
- Cross-product approximation of Hessian is used (see Bock & Lieberman, 1970).
Score test (Rao, 1948; a.k.a. Lagrange multiplier test)

1. Get the restricted (i.e. in the constrained parameter space implied by $H_0$) maximum likelihood estimates of all parameters $\hat{\eta}_0$—equivalent to fitting a locally independent 2PL model
2. Evaluate the gradient $\nabla(\eta)$ and Hessian $H(\eta)$ using the estimates from step 1—gradient vector and Hessian matrix based on the LD models need to be derived
3. Compute score test statistic $S = \nabla'(\hat{\eta}_0)H^{-1}(\hat{\eta}_0)\nabla(\hat{\eta}_0) \overset{D}{\rightarrow} \chi^2_1$

Related issues in the current project

- For the bifactor model, the first derivative w.r.t. $a_{pq}$ is 0 at $a_{pq} = 0$; as a result, the score test statistics is evaluated at 0.0001 instead of 0
- Cross-product approximation of Hessian is used (see Bock & Lieberman, 1970)
Score test (Rao, 1948; a.k.a. Lagrange multiplier test)

1. Get the restricted (i.e. in the constrained parameter space implied by $H_0$) maximum likelihood estimates of all parameters $\hat{\eta}_0$—equivalent to fitting a locally independent 2PL model

2. Evaluate the gradient $\nabla(\eta)$ and Hessian $H(\eta)$ using the estimates from step 1—gradient vector and Hessian matrix based on the LD models need to be derived

3. Compute score test statistic $S = \nabla' (\hat{\eta}_0) H^{-1} (\hat{\eta}_0) \nabla (\hat{\eta}_0) \sim \chi^2_1$

Related issues in the current project

- For the bifactor model, the first derivative w.r.t. $a_{pq}$ is 0 at $a_{pq} = 0$; as a result, the score test statistics is evaluated at 0.0001 instead of 0
- Cross-product approximation of Hessian is used (see Bock & Lieberman, 1970)
Score test (Rao, 1948; a.k.a. Lagrange multiplier test)

1. Get the restricted (i.e. in the constrained parameter space implied by $H_0$) maximum likelihood estimates of all parameters $\hat{\eta}_0$—equivalent to fitting a locally independent 2PL model

2. Evaluate the gradient $\nabla(\eta)$ and Hessian $H(\eta)$ using the estimates from step 1—gradient vector and Hessian matrix based on the LD models need to be derived

3. Compute score test statistic $S = \nabla'(\hat{\eta}_0)H^{-1}(\hat{\eta}_0)\nabla(\hat{\eta}_0) \overset{D}{\rightarrow} \chi^2_1$

Related issues in the current project

- For the bifactor model, the first derivative w/r.t. $a_{pq}$ is 0 at $a_{pq} = 0$; as a result, the score test statistics is evaluated at 0.0001 instead of 0

- Cross-product approximation of Hessian is used (see Bock & Lieberman, 1970)
Score test (Rao, 1948; a.k.a. Lagrange multiplier test)

1. Get the restricted (i.e. in the constrained parameter space implied by $H_0$) maximum likelihood estimates of all parameters $\hat{\eta}_0$—equivalent to fitting a locally independent 2PL model

2. Evaluate the gradient $\nabla(\eta)$ and Hessian $H(\eta)$ using the estimates from step 1—gradient vector and Hessian matrix based on the LD models need to be derived

3. Compute score test statistic $S = \nabla'(\hat{\eta}_0)H^{-1}(\hat{\eta}_0)\nabla(\hat{\eta}_0) \xrightarrow{D} \chi^2_1$

Related issues in the current project

- For the bifactor model, the first derivative w/r.t. $a_{pq}$ is 0 at $a_{pq} = 0$; as a result, the score test statistics is evaluated at 0.0001 instead of 0
- Cross-product approximation of Hessian is used (see Bock & Lieberman, 1970)
Score test (Rao, 1948; a.k.a. Lagrange multiplier test)

1. Get the restricted (i.e. in the constrained parameter space implied by $H_0$) maximum likelihood estimates of all parameters $\hat{\eta}_0$—equivalent to fitting a locally independent 2PL model
2. Evaluate the gradient $\nabla(\eta)$ and Hessian $H(\eta)$ using the estimates from step 1—gradient vector and Hessian matrix based on the LD models need to be derived
3. Compute score test statistic $S = \nabla'(\hat{\eta}_0)H^{-1}(\hat{\eta}_0)\nabla(\hat{\eta}_0) \xrightarrow{D} \chi^2_1$

Related issues in the current project

- For the bifactor model, the first derivative w/r.t. $a_{pq}$ is 0 at $a_{pq} = 0$; as a result, the score test statistics is evaluated at 0.0001 instead of 0
- Cross-product approximation of Hessian is used (see Bock & Lieberman, 1970)
Score test (Rao, 1948; a.k.a. Lagrange multiplier test)

1. Get the restricted (i.e. in the constrained parameter space implied by $H_0$) maximum likelihood estimates of all parameters $\hat{\eta}_0$—equivalent to fitting a locally independent 2PL model.

2. Evaluate the gradient $\nabla(\eta)$ and Hessian $H(\eta)$ using the estimates from step 1—gradient vector and Hessian matrix based on the LD models need to be derived.

3. Compute score test statistic $S = \nabla'(\hat{\eta}_0)H^{-1}(\hat{\eta}_0)\nabla(\hat{\eta}_0) \overset{D}{\rightarrow} \chi^2_1$

Related issues in the current project

- For the bifactor model, the first derivative w.r.t. $a_{pq}$ is 0 at $a_{pq} = 0$; as a result, the score test statistics is evaluated at 0.0001 instead of 0.
- Cross-product approximation of Hessian is used (see Bock & Lieberman, 1970)
Score test (Rao, 1948; a.k.a. Lagrange multiplier test)

1. Get the restricted (i.e. in the constrained parameter space implied by $H_0$) maximum likelihood estimates of all parameters $\hat{\eta}_0$—equivalent to fitting a locally independent 2PL model

2. Evaluate the gradient $\nabla(\eta)$ and Hessian $H(\eta)$ using the estimates from step 1—gradient vector and Hessian matrix based on the LD models need to be derived

3. Compute score test statistic $S = \nabla'(\hat{\eta}_0)H^{-1}(\hat{\eta}_0)\nabla(\hat{\eta}_0) \xrightarrow{D} \chi^2_1$

Related issues in the current project

- For the bifactor model, the first derivative w.r.t. $a_{pq}$ is 0 at $a_{pq} = 0$; as a result, the score test statistic is evaluated at 0.0001 instead of 0

- Cross-product approximation of Hessian is used (see Bock & Lieberman, 1970)
Score test (Rao, 1948; a.k.a. Lagrange multiplier test)

1. Get the restricted (i.e., in the constrained parameter space implied by $H_0$) maximum likelihood estimates of all parameters $\hat{\eta}_0$—equivalent to fitting a locally independent 2PL model

2. Evaluate the gradient $\nabla(\eta)$ and Hessian $H(\eta)$ using the estimates from step 1—gradient vector and Hessian matrix based on the LD models need to be derived

3. Compute score test statistic $S = \nabla' (\hat{\eta}_0) H^{-1} (\hat{\eta}_0) \nabla (\hat{\eta}_0) \xrightarrow{D} \chi^2_1$

Related issues in the current project

- For the bifactor model, the first derivative w.r.t. $a_{pq}$ is 0 at $a_{pq} = 0$; as a result, the score test statistics is evaluated at 0.0001 instead of 0
- Cross-product approximation of Hessian is used (see Bock & Lieberman, 1970)
OVERVIEW

- Three statistics
  - $S_b$: score statistic based on bifactor model
  - $S_t$: score statistic based on threshold shift model
  - $X^2$: LD $X^2$ statistic

- Null distribution and type I error rate
  - Null case: Locally independent data

- Power
  - ULD case: data generated with bifactor model
  - SLD case: data generated with threshold shift model
Generating distribution (see Chen & Thissen, 1997)

- Slopes: $a \sim \log \mathcal{N}(0, 0.5)$, truncated to $[0.2, 5.18]$
- Thresholds: $b \sim \mathcal{N}(0, 1.5)$, truncated to $[-2, 2]$
- Intercepts: $c = -ab$

Conditions

<table>
<thead>
<tr>
<th>No. of items</th>
<th>Sample sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>200, 500, 1000</td>
</tr>
<tr>
<td>25</td>
<td>200, 500, 1000</td>
</tr>
<tr>
<td>50</td>
<td>200, 500, 1000</td>
</tr>
</tbody>
</table>

Design: No. of items × Sample sizes

No. of replications: 1000
NULL CASE

- Generating distribution (see Chen & Thissen, 1997)
  - Slopes: $a \sim \log N(0,0.5)$, truncated to $[0.2, 5.18]$
  - Thresholds: $b \sim N(0,1.5)$, truncated to $[-2, 2]$
  - Intercepts: $c = -ab$

- Conditions

<table>
<thead>
<tr>
<th>No. of items</th>
<th>Sample sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>200, 500, 1000</td>
</tr>
<tr>
<td>25</td>
<td>200, 500, 1000</td>
</tr>
<tr>
<td>50</td>
<td>200, 500, 1000</td>
</tr>
</tbody>
</table>

Design: No. of items $\times$ Sample sizes

- No. of replications: 1000
Generating distribution (see Chen & Thissen, 1997)

- Slopes: \( a \sim \log \mathcal{N}(0,0.5) \), truncated to \([0.2, 5.18]\)
- Thresholds: \( b \sim \mathcal{N}(0,1.5) \), truncated to \([-2, 2]\)
- Intercepts: \( c = -ab \)

Conditions

<table>
<thead>
<tr>
<th>No. of items</th>
<th>Sample sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>200, 500, 1000</td>
</tr>
<tr>
<td>25</td>
<td>200, 500, 1000</td>
</tr>
<tr>
<td>50</td>
<td>200, 500, 1000</td>
</tr>
</tbody>
</table>

Design: No. of items \( \times \) Sample sizes

No. of replications: 1000
ULD AND SLD CASES

- Generating distribution
  - Secondary loadings: $\lambda_{pq} \sim N(\mu_\lambda, 0.01)$, truncated to $(\mu_\lambda - 0.2, \mu_\lambda + 0.2)$
  - Secondary slope: $a_{pq} = \frac{1.702 \lambda_{pq}}{\sqrt{(1 - \lambda_{pq})^2}}$
  - Threshold shift: $|\delta_{pq}| \approx \frac{\lambda_{pq}^2}{(1 - \lambda_p^2)}$

<table>
<thead>
<tr>
<th>No. of items</th>
<th>Sample sizes</th>
<th>Strength of LD</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>200, 500, 1000</td>
<td>$\mu_\lambda = 0.3, 0.5, 0.7$</td>
</tr>
<tr>
<td>25</td>
<td>200, 500, 1000</td>
<td>$\mu_\lambda = 0.3, 0.5, 0.7$</td>
</tr>
<tr>
<td>50</td>
<td>200, 500, 1000</td>
<td>$\mu_\lambda = 0.3, 0.5, 0.7$</td>
</tr>
</tbody>
</table>

Design: No. of items × Sample sizes × Strength

- No. of replications: 1000
ULD AND SLD CASES

- Generating distribution
  - Secondary loadings: $\lambda_{pq} \sim \mathcal{N}(\mu_\lambda, 0.01)$, truncated to $(\mu_\lambda - 0.2, \mu_\lambda + 0.2)$
  - Secondary slope: $a_{pq} = \frac{1.702 \lambda_{pq}}{\sqrt{(1 - \lambda_{pq})^2}}$
  - Threshold shift: $|\delta_{pq}| \approx \frac{\lambda^2_{pq}}{(1 - \lambda^2_p)}$

<table>
<thead>
<tr>
<th>No. of items</th>
<th>Sample sizes</th>
<th>Strength of LD</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>200, 500, 1000</td>
<td>$\mu_\lambda = 0.3, 0.5, 0.7$</td>
</tr>
<tr>
<td>25</td>
<td>200, 500, 1000</td>
<td>$\mu_\lambda = 0.3, 0.5, 0.7$</td>
</tr>
<tr>
<td>50</td>
<td>200, 500, 1000</td>
<td>$\mu_\lambda = 0.3, 0.5, 0.7$</td>
</tr>
</tbody>
</table>

Design: No. of items × Sample sizes × Strength

No. of replications: 1000
ULD AND SLD CASES

- Generating distribution
  - Secondary loadings: $\lambda_{pq} \sim N(\mu_\lambda, 0.01)$, truncated to $(\mu_\lambda - 0.2, \mu_\lambda + 0.2)$
  - Secondary slope: $a_{pq} = \frac{1.702 \lambda_{pq}}{\sqrt{(1 - \lambda_{pq})^2}}$
  - Threshold shift: $|\delta_{pq}| \approx \frac{\lambda_{pq}^2}{(1 - \lambda_p^2)}$

<table>
<thead>
<tr>
<th>No. of items</th>
<th>Sample sizes</th>
<th>Strength of LD</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>200, 500, 1000</td>
<td>$\mu_\lambda = 0.3, 0.5, 0.7$</td>
</tr>
<tr>
<td>25</td>
<td>200, 500, 1000</td>
<td>$\mu_\lambda = 0.3, 0.5, 0.7$</td>
</tr>
<tr>
<td>50</td>
<td>200, 500, 1000</td>
<td>$\mu_\lambda = 0.3, 0.5, 0.7$</td>
</tr>
</tbody>
</table>

Design: No. of items $\times$ Sample sizes $\times$ Strength

- No. of replications: 1000
OUTLINE

1. Introduction

2. Simulation design

3. Results

4. Discussion
### Empirical Quantiles: 10 Items

<table>
<thead>
<tr>
<th></th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N=200</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_b$</td>
<td>0.13</td>
<td>0.49</td>
<td>1.48</td>
<td>3.28</td>
<td>4.95</td>
<td>10.63</td>
</tr>
<tr>
<td>$S_t$</td>
<td>0.13</td>
<td>0.50</td>
<td>1.45</td>
<td>3.38</td>
<td>5.16</td>
<td>29.86</td>
</tr>
<tr>
<td>$X^2$</td>
<td>0.06</td>
<td>0.25</td>
<td>0.71</td>
<td>1.46</td>
<td>2.21</td>
<td>4.10</td>
</tr>
<tr>
<td><strong>N=500</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_b$</td>
<td>0.11</td>
<td>0.46</td>
<td>1.29</td>
<td>2.56</td>
<td>3.73</td>
<td>8.96</td>
</tr>
<tr>
<td>$S_t$</td>
<td>0.10</td>
<td>0.45</td>
<td>1.31</td>
<td>2.67</td>
<td>3.88</td>
<td>8.28</td>
</tr>
<tr>
<td>$X^2$</td>
<td>0.05</td>
<td>0.26</td>
<td>0.71</td>
<td>1.52</td>
<td>2.26</td>
<td>4.69</td>
</tr>
<tr>
<td><strong>N=1000</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_b$</td>
<td>0.11</td>
<td>0.46</td>
<td>1.34</td>
<td>2.85</td>
<td>4.03</td>
<td>7.46</td>
</tr>
<tr>
<td>$S_t$</td>
<td>0.11</td>
<td>0.48</td>
<td>1.33</td>
<td>2.94</td>
<td>4.16</td>
<td>7.35</td>
</tr>
<tr>
<td>$X^2$</td>
<td>0.06</td>
<td>0.27</td>
<td>0.75</td>
<td>1.77</td>
<td>2.39</td>
<td>4.03</td>
</tr>
</tbody>
</table>

<p>| $\chi^2_1$ | 0.10  | 0.45  | 1.32  | 2.71  | 3.84  | 6.63  |</p>
<table>
<thead>
<tr>
<th></th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_b$</td>
<td>0.14</td>
<td>0.65</td>
<td>2.16</td>
<td>4.32</td>
<td>6.12</td>
<td>19.84</td>
</tr>
<tr>
<td>$S_t$</td>
<td>0.15</td>
<td>0.65</td>
<td>2.10</td>
<td>4.28</td>
<td>6.52</td>
<td>45.26</td>
</tr>
<tr>
<td>$X^2$</td>
<td>0.08</td>
<td>0.32</td>
<td>1.01</td>
<td>2.34</td>
<td>2.89</td>
<td>4.08</td>
</tr>
<tr>
<td>N=500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_b$</td>
<td>0.10</td>
<td>0.49</td>
<td>1.61</td>
<td>3.46</td>
<td>5.28</td>
<td>11.27</td>
</tr>
<tr>
<td>$S_t$</td>
<td>0.11</td>
<td>0.50</td>
<td>1.68</td>
<td>3.43</td>
<td>5.16</td>
<td>11.44</td>
</tr>
<tr>
<td>$X^2$</td>
<td>0.07</td>
<td>0.31</td>
<td>0.96</td>
<td>2.08</td>
<td>2.83</td>
<td>4.78</td>
</tr>
<tr>
<td>N=1000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_b$</td>
<td>0.09</td>
<td>0.48</td>
<td>1.44</td>
<td>2.82</td>
<td>4.37</td>
<td>6.79</td>
</tr>
<tr>
<td>$S_t$</td>
<td>0.09</td>
<td>0.48</td>
<td>1.37</td>
<td>2.87</td>
<td>4.46</td>
<td>7.11</td>
</tr>
<tr>
<td>$X^2$</td>
<td>0.07</td>
<td>0.32</td>
<td>0.93</td>
<td>1.83</td>
<td>2.67</td>
<td>4.35</td>
</tr>
<tr>
<td>$\chi^2_1$</td>
<td>0.10</td>
<td>0.45</td>
<td>1.32</td>
<td>2.71</td>
<td>3.84</td>
<td>6.63</td>
</tr>
</tbody>
</table>
**EMPIRICAL QUANTILES: 50 ITEMS**

<table>
<thead>
<tr>
<th></th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_b$</td>
<td>0.19</td>
<td>0.84</td>
<td>2.88</td>
<td>6.53</td>
<td>10.47</td>
<td>325.45</td>
</tr>
<tr>
<td>$S_t$</td>
<td>0.19</td>
<td>0.83</td>
<td>2.83</td>
<td>7.12</td>
<td>10.84</td>
<td>240.26</td>
</tr>
<tr>
<td>$X^2$</td>
<td>0.11</td>
<td>0.39</td>
<td>1.07</td>
<td>2.29</td>
<td>3.35</td>
<td>42.08</td>
</tr>
<tr>
<td>N=500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_b$</td>
<td>0.11</td>
<td>0.54</td>
<td>1.70</td>
<td>3.46</td>
<td>5.26</td>
<td>10.56</td>
</tr>
<tr>
<td>$S_t$</td>
<td>0.12</td>
<td>0.58</td>
<td>1.70</td>
<td>3.48</td>
<td>5.38</td>
<td>11.85</td>
</tr>
<tr>
<td>$X^2$</td>
<td>0.11</td>
<td>0.39</td>
<td>1.06</td>
<td>1.99</td>
<td>2.89</td>
<td>5.58</td>
</tr>
<tr>
<td>N=1000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_b$</td>
<td>0.11</td>
<td>0.53</td>
<td>1.58</td>
<td>3.34</td>
<td>4.80</td>
<td>8.79</td>
</tr>
<tr>
<td>$S_t$</td>
<td>0.13</td>
<td>0.52</td>
<td>1.58</td>
<td>3.34</td>
<td>4.87</td>
<td>10.57</td>
</tr>
<tr>
<td>$X^2$</td>
<td>0.14</td>
<td>0.40</td>
<td>1.05</td>
<td>2.03</td>
<td>3.12</td>
<td>5.61</td>
</tr>
<tr>
<td>$\chi_1^2$</td>
<td>0.10</td>
<td>0.45</td>
<td>1.32</td>
<td>2.71</td>
<td>3.84</td>
<td>6.63</td>
</tr>
</tbody>
</table>
ROC CURVE: WEAK SLD
ROC CURVE: MODERATE SLD

SLD: N = 200, J = 10, \( \mu_x = 0.5 \) (\( \delta_{pq} = 0.34 \))
- Bifactor S_b
- Threshold shift S_t
- LD X^2

SLD: N = 500, J = 10, \( \mu_x = 0.5 \) (\( \delta_{pq} = 0.34 \))
- Bifactor S_b
- Threshold shift S_t
- LD X^2

SLD: N = 1000, J = 10, \( \mu_x = 0.5 \) (\( \delta_{pq} = 0.34 \))
- Bifactor S_b
- Threshold shift S_t
- LD X^2

SLD: N = 200, J = 25, \( \mu_x = 0.5 \) (\( \delta_{pq} = 0.34 \))
- Bifactor S_b
- Threshold shift S_t
- LD X^2

SLD: N = 500, J = 25, \( \mu_x = 0.5 \) (\( \delta_{pq} = 0.34 \))
- Bifactor S_b
- Threshold shift S_t
- LD X^2

SLD: N = 1000, J = 25, \( \mu_x = 0.5 \) (\( \delta_{pq} = 0.34 \))
- Bifactor S_b
- Threshold shift S_t
- LD X^2

SLD: N = 200, J = 50, \( \mu_x = 0.5 \) (\( \delta_{pq} = 0.34 \))
- Bifactor S_b
- Threshold shift S_t
- LD X^2

SLD: N = 500, J = 50, \( \mu_x = 0.5 \) (\( \delta_{pq} = 0.34 \))
- Bifactor S_b
- Threshold shift S_t
- LD X^2

SLD: N = 1000, J = 50, \( \mu_x = 0.5 \) (\( \delta_{pq} = 0.34 \))
- Bifactor S_b
- Threshold shift S_t
- LD X^2
ROC CURVE: WEAK ULD

ULD: N = 200, J = 10, \( \mu_\lambda = 0.3 \) (\( a_{pq} = 0.54 \))
- Bifactor \( S_B \)
- Threshold shift \( S_t \)
- LD \( X^2 \)

ULD: N = 500, J = 10, \( \mu_\lambda = 0.3 \) (\( a_{pq} = 0.54 \))
- Bifactor \( S_B \)
- Threshold shift \( S_t \)
- LD \( X^2 \)

ULD: N = 1000, J = 10, \( \mu_\lambda = 0.3 \) (\( a_{pq} = 0.54 \))
- Bifactor \( S_B \)
- Threshold shift \( S_t \)
- LD \( X^2 \)

ULD: N = 200, J = 25, \( \mu_\lambda = 0.3 \) (\( a_{pq} = 0.54 \))
- Bifactor \( S_B \)
- Threshold shift \( S_t \)
- LD \( X^2 \)

ULD: N = 500, J = 25, \( \mu_\lambda = 0.3 \) (\( a_{pq} = 0.54 \))
- Bifactor \( S_B \)
- Threshold shift \( S_t \)
- LD \( X^2 \)

ULD: N = 1000, J = 25, \( \mu_\lambda = 0.3 \) (\( a_{pq} = 0.54 \))
- Bifactor \( S_B \)
- Threshold shift \( S_t \)
- LD \( X^2 \)

ULD: N = 200, J = 50, \( \mu_\lambda = 0.3 \) (\( a_{pq} = 0.54 \))
- Bifactor \( S_B \)
- Threshold shift \( S_t \)
- LD \( X^2 \)

ULD: N = 500, J = 50, \( \mu_\lambda = 0.3 \) (\( a_{pq} = 0.54 \))
- Bifactor \( S_B \)
- Threshold shift \( S_t \)
- LD \( X^2 \)

ULD: N = 1000, J = 50, \( \mu_\lambda = 0.3 \) (\( a_{pq} = 0.54 \))
- Bifactor \( S_B \)
- Threshold shift \( S_t \)
- LD \( X^2 \)
ROC CURVE: MODERATE ULD

ULD: N = 200, J = 10, \( \mu \approx 0.5 \) (\( a_{pq} = 0.98 \))
- Bifactor \( S_b \)
- Threshold shift \( S_t \)
- LD \( X^2 \)

ULD: N = 500, J = 10, \( \mu \approx 0.5 \) (\( a_{pq} = 0.98 \))
- Bifactor \( S_b \)
- Threshold shift \( S_t \)
- LD \( X^2 \)

ULD: N = 1000, J = 10, \( \mu \approx 0.5 \) (\( a_{pq} = 0.98 \))
- Bifactor \( S_b \)
- Threshold shift \( S_t \)
- LD \( X^2 \)

ULD: N = 200, J = 25, \( \mu \approx 0.5 \) (\( a_{pq} = 0.98 \))
- Bifactor \( S_b \)
- Threshold shift \( S_t \)
- LD \( X^2 \)

ULD: N = 500, J = 25, \( \mu \approx 0.5 \) (\( a_{pq} = 0.98 \))
- Bifactor \( S_b \)
- Threshold shift \( S_t \)
- LD \( X^2 \)

ULD: N = 1000, J = 25, \( \mu \approx 0.5 \) (\( a_{pq} = 0.98 \))
- Bifactor \( S_b \)
- Threshold shift \( S_t \)
- LD \( X^2 \)

ULD: N = 200, J = 50, \( \mu \approx 0.5 \) (\( a_{pq} = 0.98 \))
- Bifactor \( S_b \)
- Threshold shift \( S_t \)
- LD \( X^2 \)

ULD: N = 500, J = 50, \( \mu \approx 0.5 \) (\( a_{pq} = 0.98 \))
- Bifactor \( S_b \)
- Threshold shift \( S_t \)
- LD \( X^2 \)

ULD: N = 1000, J = 50, \( \mu \approx 0.5 \) (\( a_{pq} = 0.98 \))
- Bifactor \( S_b \)
- Threshold shift \( S_t \)
- LD \( X^2 \)
ROC CURVE: STRONG ULD

ULD: N = 200, J = 10, $\mu_\lambda = 0.7$ ($a_{pq} = 1.67$)
- Bifactor $S_b$
- Threshold shift $S_t$
- LD $X^2$

ULD: N = 500, J = 10, $\mu_\lambda = 0.7$ ($a_{pq} = 1.67$)
- Bifactor $S_b$
- Threshold shift $S_t$
- LD $X^2$

ULD: N = 1000, J = 10, $\mu_\lambda = 0.7$ ($a_{pq} = 1.67$)
- Bifactor $S_b$
- Threshold shift $S_t$
- LD $X^2$

ULD: N = 200, J = 25, $\mu_\lambda = 0.7$ ($a_{pq} = 1.67$)
- Bifactor $S_b$
- Threshold shift $S_t$
- LD $X^2$

ULD: N = 500, J = 25, $\mu_\lambda = 0.7$ ($a_{pq} = 1.67$)
- Bifactor $S_b$
- Threshold shift $S_t$
- LD $X^2$

ULD: N = 1000, J = 25, $\mu_\lambda = 0.7$ ($a_{pq} = 1.67$)
- Bifactor $S_b$
- Threshold shift $S_t$
- LD $X^2$

ULD: N = 200, J = 50, $\mu_\lambda = 0.7$ ($a_{pq} = 1.67$)
- Bifactor $S_b$
- Threshold shift $S_t$
- LD $X^2$

ULD: N = 500, J = 50, $\mu_\lambda = 0.7$ ($a_{pq} = 1.67$)
- Bifactor $S_b$
- Threshold shift $S_t$
- LD $X^2$

ULD: N = 1000, J = 50, $\mu_\lambda = 0.7$ ($a_{pq} = 1.67$)
- Bifactor $S_b$
- Threshold shift $S_t$
- LD $X^2$
1. Introduction
2. Simulation design
3. Results
4. Discussion
Conclusions

- LD $X^2$ is the easiest to compute, and it works fine.
- Threshold shift score test statistic $S_t$ is the second easiest to compute (i.e. unidimensional), and works fine for both SLD and ULD cases.
- Bifactor score test statistic $S_b$ is the hardest to compute (i.e. with one more secondary dimension), and provides no advantage over the other two.

Future directions

- Generalize to polytomous IRT models—results may differ.
DISCUSSION

- **Conclusions**
  - LD $X^2$ is the easiest to compute, and it works fine
  - Threshold shift score test statistic $S_t$ is the second easiest to compute (i.e. unidimensional), and works fine for both SLD and ULD cases
  - Bifactor score test statistic $S_b$ is the hardest to compute (i.e. with one more secondary dimension), and provides no advantage over the other two

- **Future directions**
  - Generalize to polytomous IRT models—results may differ
Conclusions

- LD $X^2$ is the easiest to compute, and it works fine
- Threshold shift score test statistic $S_t$ is the second easiest to compute (i.e. unidimensional), and works fine for both SLD and ULD cases
- Bifactor score test statistic $S_b$ is the hardest to compute (i.e. with one more secondary dimension), and provides no advantage over the other two

Future directions

- Generalize to polytomous IRT models—results may differ
Conclusions

- LD $X^2$ is the easiest to compute, and it works fine.
- Threshold shift score test statistic $S_t$ is the second easiest to compute (i.e. unidimensional), and works fine for both SLD and ULD cases.
- Bifactor score test statistic $S_b$ is the hardest to compute (i.e. with one more secondary dimension), and provides no advantage over the other two.

Future directions

- Generalize to polytomous IRT models—results may differ.
DISCUSSION

- Conclusions
  - LD $X^2$ is the easiest to compute, and it works fine
  - Threshold shift score test statistic $S_t$ is the second easiest to compute (i.e. unidimensional), and works fine for both SLD and ULD cases
  - Bifactor score test statistic $S_b$ is the hardest to compute (i.e. with one more secondary dimension), and provides no advantage over the other two

- Future directions
  - Generalize to polytomous IRT models—results may differ
Questions & Comments?
Thanks!