Further Readings On Path Analysis With Categorical Outcomes

MacKinnon, D.P., Lockwood, C.M., Brown, C.H., Wang, W., & Hoffman, J.M. (2007). The intermediate endpoint effect in logistic and probit regression. Clinical Trials, 4, 499-513.

Xie, Y. (1989). Structural equation models for ordinal variables. Sociological Methods & Research, 17, 325-352.

85

Categorical Observed And Continuous Latent Variables

86

Continuous Latent Variable Analysis With Categorical Outcomes

Model Identification

- EFA, CFA, and SEM the same as for continuous outcomes
- Multiple group and models for longitudinal data require invariance of measurement thresholds and loadings, requiring threshold structure (and scale factor parameters)

Interpretation

- Estimated coefficients – sign, significance most important
- Estimated coefficients can be converted to probabilities

87

Continuous Latent Variable Analysis With Categorical Outcomes (Continued)

Estimation

- Maximum likelihood computational burden increases significantly with number of factors
- Weighted least squares computation burden increases significantly with the number of variables

Model Fit

- Only chi-square studied
- Simulation studies needed for TLI, CFI, RMSEA, SRMR, and WRMR (see, however, Yu, 2002)

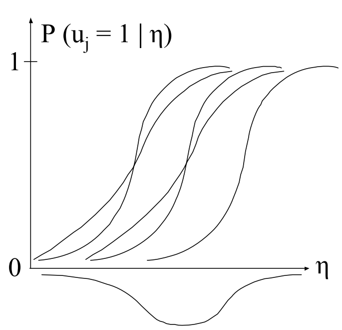
88

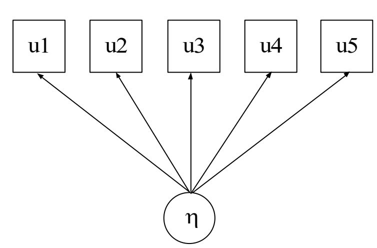
Item Response Theory

89

Item Response Theory

Latent trait modeling
Factor analysis with categorical outcomes





90

Item Response Theory (Continued)

IRT typically does not use the full SEM model

$$u_i^* = \boldsymbol{\nu} + \boldsymbol{\Lambda}\, \boldsymbol{\eta}_i\, (+ \boldsymbol{K}\, \mathbf{x}_i\,) + \varepsilon_i\,, \qquad (127)$$

$$\boldsymbol{\eta}_i = \boldsymbol{\alpha} + (\, \boldsymbol{B}\boldsymbol{\eta}_i + \boldsymbol{\Gamma}\, \mathbf{x}_i) + \zeta_i\,, \qquad (128)$$

and typically considers a single η (see, however, Bock, Gibbons, & Muraki, 1988). Aims:

- Item parameter estimation (ML): Calibration
- Estimation of η values: Scoring
- Assessment of information function
- Test equating
- DIF analysis

91

IRT Models And Estimators In Mplus

- ML (full information estimation): Logit and probit links
- WLS (limited information estimation): Probit link

92

Translating Factor Analysis Parameters In Mplus To IRT Parameters

- IRT calls the continuous latent variable θ
- 2-parameter logistic IRT model uses

$$P(u = 1 \mid \theta) = \frac{1}{1 + e^{-D\,a(\theta - b)}}$$

with D = 1.7 to make a, b close to those of probit

a discrimination

b difficulty

- 2-parameter normal ogive IRT model uses

$$P(u = 1 \mid \theta) = \Phi\left[a(\theta - b)\right]$$

- Typically $\theta \sim N(0,1)$

93

Translating Factor Analysis Parameters To IRT Parameters (Continued)

- The Mplus factor analysis model uses

$$P(u = 1 \mid \eta) = \frac{1}{1 + e^{-(-\tau + \lambda\eta)}} \quad \text{for logit}$$

$$P(u = 1 \mid \eta) = \Phi\left[(-\tau + \lambda\eta)\,\theta^{-1/2}\right] \text{ for probit}$$

where θ is the residual variance

The logit conversion is:

$$a = \lambda\sqrt{\psi} / D$$
$$b = (\tau - \lambda\alpha) / \lambda\sqrt{\psi}$$

The probit conversion is:

$$a = \lambda\sqrt{\psi}\,\theta^{-1/2}$$
$$b = (\tau - \lambda\alpha) / \lambda\sqrt{\psi}$$

- Conversion automatically done in Mplus

94

Testing The Model Against Data

- Model fit to frequency tables. Overall test against data
  - When the model contains only **u**, summing over the cells,

$$\chi^2_P = \sum_i \frac{(o_i - e_i)^2}{e_i}, \qquad (82)$$

$$\chi^2_{LR} = 2 \sum_i o_i \, log \, o_i / e_i . \qquad (83)$$

A cell that has non-zero observed frequency and expected frequency less than .01 is not included in the χ^2 computation as the default. With missing data on **u**, the EM algorithm described in Little and Rubin (1987; chapter 9.3, pp. 181-185) is used to compute the estimated frequencies in the unrestricted multinomial model. In this case, a test of MCAR for the unrestricted model is also provided (Little & Rubin, 1987, pp. 192-193).

- Model fit to univariate and bivariate frequency tables. Mplus TECH10

95

Antisocial Behavior (ASB) Data

The Antisocial Behavior (ASB) data were taken from the National Longitudinal Survey of Youth (NLSY) that is sponsored by the Bureau of Labor Statistics. These data are made available to the public by Ohio State University. The data were obtained as a multistage probability sample with oversampling of blacks, Hispanics, and economically disadvantaged non-blacks and non-Hispanics.

Data for the analysis include 15 of the 17 antisocial behavior items that were collected in 1980 when respondents were between the ages of 16 and 23 and the background variables of age, gender and ethnicity. The ASB items assessed the frequency of various behaviors during the past year. A sample of 7,326 respondents has complete data on the antisocial behavior items and the background variables of age, gender, and ethnicity. Following is a list of the 15 items:

96

Antisocial Behavior (ASB) Data (Continued)

Damaged property
Fighting
Shoplifting
Stole < $50
Stole > $50
Seriously threaten
Intent to injure
Use marijuana
Use other drugs
Sold marijuana
Sold hard drugs
"Con" someone
Take auto
Broken into building
Held stolen goods

These items were dichotomized 0/1 with 0 representing never in the last year. An EFA suggested three factors: property offense, person offense, and drug offense.

97

Input For IRT Analysis Of Eight ASB Property Offense Items

```
TITLE:      2-parameter logistic IRT
            for 8 property offense items
DATA:       FILE = asb.dat;
            FORMAT = 34X 54F2.0;
VARIABLE:   NAMES = property fight shoplift lt50 gt50 force threat
            injure pot drug soldpot solddrug con auto bldg goods
            gambling
            dsm1-dsm22 sex black hisp single divorce dropout
            college onset f1 f2 f3
            age94 cohort dep abuse;
            USEVAR = property shoplift lt50 gt50 con auto bldg
            goods;
            CATEGORICAL = property-goods;
ANALYSIS:   ESTIMATOR = MLR;
MODEL:      f BY property-goods*;
            f@1;
OUTPUT:     TECH1 TECH8 TECH10;
PLOT:       TYPE = PLOT3;
```

98

Output Excerpts IRT Analysis Of Eight ASB Property Offense Items

```
TESTS OF MODEL FIT
Loglikelihood
   H0 Value                                        -19758.361
   H0 Scaling Correction Factor for MLR                 0.996
Information Criteria
   Number of Free Parameters                               16
   Akaike (AIC)                                     39548.722
   Bayesian (BIC)                                   39659.109
   Sample-Size Adjusted BIC                         39608.265
      (n* = (n + 2) / 24)
Chi-Square Test of Model Fit for the Binary and Ordered
   Categorical (Ordinal) Outcomes
   Pearson Chi-Square
   Value                                              324.381
   Degrees of Freedom                                     239
   P-Value                                             0.0002
```

99

Output Excerpts IRT Analysis Of Eight ASB Property Offense Items (Continued)

```
    Likelihood Ratio Chi-Square
    Value                                        327.053
    Degrees of Freedom                               239
    P-Value                                       0.0001
MODEL RESULTS                                              Two-Tailed
                    Estimate       S.E.   Est./S.E.        P-Value
F          BY
     PROPERTY          2.032      0.084      24.060          0.000
     SHOPLIFT          1.712      0.068      25.115          0.000
     LT50              1.850      0.076      24.411          0.000
     GT50              2.472      0.139      17.773          0.000
     CON               1.180      0.051      23.148          0.000
     AUTO              1.383      0.070      19.702          0.000
     BLDG              2.741      0.151      18.119          0.000
     GOODS             2.472      0.116      21.339          0.000
```

100

Output Excerpts IRT Analysis Of Eight ASB Property Offense Items (Continued)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Thresholds				
PROPERTY$1	2.398	0.073	32.803	0.000
SHOPLIFT$1	1.529	0.049	31.125	0.000
LT50$1	2.252	0.065	34.509	0.000
GT50$1	5.054	0.195	25.912	0.000
CON$1	1.560	0.041	37.894	0.000
AUTO$1	3.144	0.079	39.948	0.000
BLDG$1	5.185	0.208	24.983	0.000
GOODS$1	3.691	0.126	29.316	0.000
Variances				
F	1.000	0.000	999.000	999.000

101

Output Excerpts IRT Analysis Of Eight ASB Property Offense Items (Continued)

IRT PARAMETERIZATION IN TWO-PARAMETER LOGISTIC METRIC WHERE THE LOGIT IS 1.7*DISCRIMINATION*(THETA - DIFFICULTY)

Item Discriminations	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
F BY				
PROPERTY	1.195	0.050	24.060	0.000
SHOPLIFT	1.007	0.040	25.115	0.000
LT50	1.088	0.045	24.411	0.000
GT50	1.454	0.082	17.773	0.000
CON	0.694	0.030	23.148	0.000
AUTO	0.813	0.041	19.702	0.000
BLDG	1.612	0.089	18.119	0.000
GOODS	1.454	0.068	21.339	0.000

102

Output Excerpts IRT Analysis Of Eight ASB Property Offense Items (Continued)

Item Difficulties	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
PROPERTY$1	1.180	0.031	38.268	0.000
SHOPLIFT$1	0.893	0.029	31.309	0.000
LT50$1	1.217	0.033	36.604	0.000
GT50$1	2.044	0.053	38.588	0.000
CON$1	1.322	0.048	27.809	0.000
AUTO$1	2.274	0.081	28.232	0.000
BLDG$1	1.891	0.045	42.204	0.000
GOODS$1	1.493	0.035	43.045	0.000
Variances				
F	1.000	0.000	0.000	1.000

103

Output Excerpts IRT Analysis Of Eight ASB Property Offense Items (Continued)

TECHNICAL 10 OUTPUT
MODEL FIT INFORMATION FOR THE LATENT CLASS INDICATOR MODEL PART
RESPONSE PATTERNS

No.	Pattern	No.	Pattern	No.	Pattern	No.	Pattern
1	00000000	2	10100000	3	00001101	4	00000010
5	01100000	6	00001000	7	10001010	8	00010001
9	10100010	10	11000000	11	10101110	12	11100010
13	11010111	14	10000000	15	11110001	16	10000001

104

Output Excerpts IRT Analysis Of Eight ASB Property Offense Items (Continued)

RESPONSE PATTERN FREQUENCIES AND CHI-SQURE CONTRIBUTIONS

Response Pattern	Frequency Observed	Frequency Estimated	Standardized Residual (z-score)	Chi-square Pearson	Contribution Loglikelihood
1	3581.00	3565.17	0.37	0.07	31.73
2	60.00	57.05	0.39	0.15	6.05
3	2.00	3.12	-0.77	0.59	-2.14
4	18.00	17.65	0.08	0.01	0.71
5	137.00	110.30	2.56	6.46	59.39
6	476.00	495.86	-0.92	0.80	-38.92

Output Excerpts IRT Analysis Of Eight ASB Property Offense Items (Continued)

BIVARIATE MODEL FIT INFORMATION

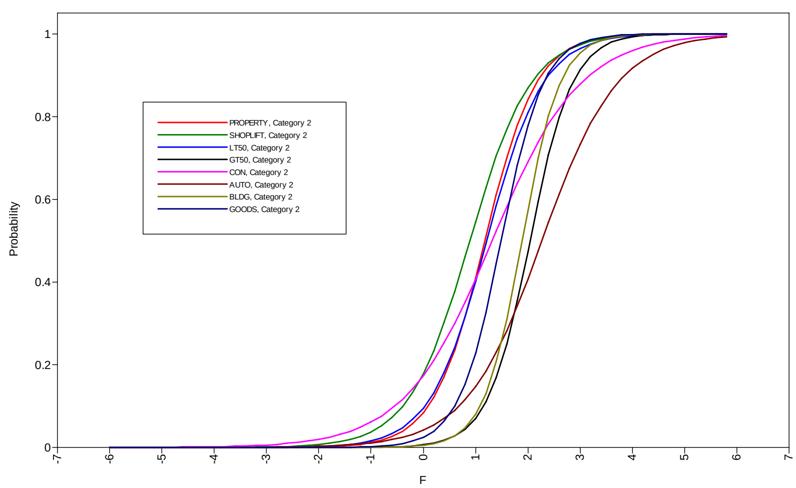
VARIABLE PROPERTY	VARIABLE SHOPLIFT	Estimated Probabilities H1	Estimated Probabilities H0	Standardized Residual (z-score)
Category 1	Category 1	0.656	0.655	0.157
Category 1	Category 2	0.159	0.160	-0.176
Category 2	Category 1	0.080	0.081	-0.285
Category 2	Category 2	0.105	0.104	0.222
Bivariate Pearson Chi-Square				0.153
Bivariate Log-Likelihood Chi-Square				0.077

Output Excerpts IRT Analysis Of Eight ASB Property Offense Items (Continued)

			Estimated Probabilities		
					Standardized
VARIABLE	VARIABLE		H1	H0	Residual
PROPERTY	SHOPLIFT				(z-score)
LT50	GT50				
Category 1	Category 1		0.799	0.795	0.873
Category 1	Category 2		0.014	0.018	-2.615
Category 2	Category 1		0.152	0.156	-0.945
Category 2	Category 2		0.035	0.032	1.912
Bivariate Pearson Chi-Square					11.167
Bivariate Log-Likelihood Chi-Square					5.806

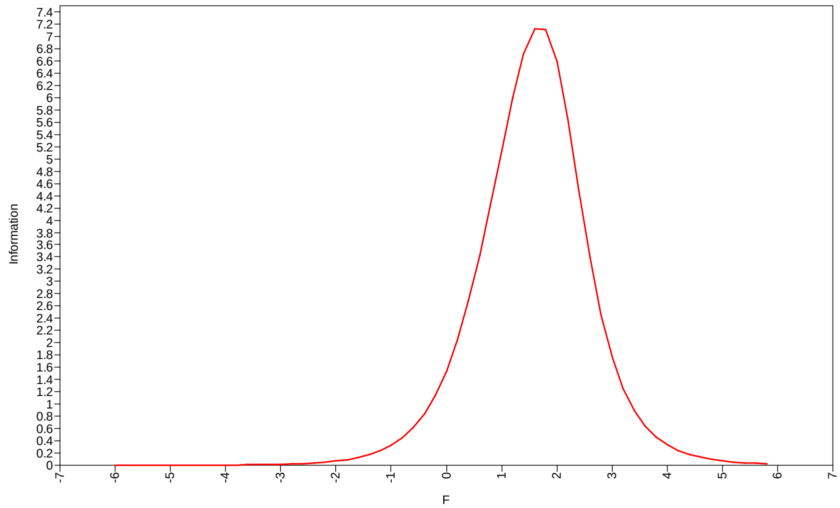
107

Item Characteristic Curves



108

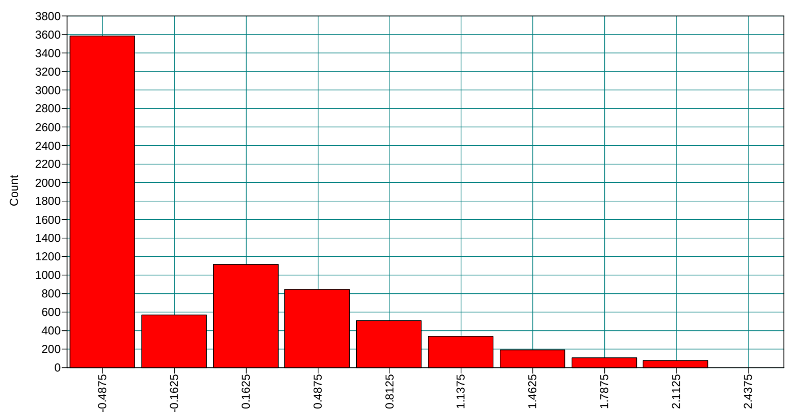
Test Information Curve



109

Histogram For Estimated Factor Scores Using The Expected A Posteriori Method

Prior (normal) + Data = Posterior



110

Further Readings On IRT

Baker, F.B. & Kim, S.H. (2004). Item response theory. Parameter estimation techniques. Second edition. New York: Marcel Dekker.

Bock, R.D. (1997). A brief history of item response theory. Educational Measurement: Issues and Practice, 16, 21-33.

du Toit, M. (2003). IRT from SSI. Lincolnwood, IL: Scientific Software International, Inc. (BILOG, MULTILOG, PARSCALE, TESTFACT)

Embretson, S. E., & Reise, S. P. (2000). Item response theory for psychologists. Mahwah, NJ: Erlbaum.

Hambleton, R.K. & Swaminathan, H. (1985). Item response theory. Boston: Kluwer-Nijhoff.

MacIntosh, R. & Hashim, S. (2003). Variance estimation for converting MIMIC model parameters to IRT parameters in DIF analysis. Applied Psychological Measurement, 27, 372-379.

Muthén, B., Kao, Chih-Fen, & Burstein, L. (1991). Instructional sensitivity in mathematics achievement test items: Applications of a new IRT-based detection technique. Journal of Educational Measurement, 28, 1-22. (#35)

Further Readings On IRT (Continued)

Muthén, B. & Asparouhov, T. (2002). Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus. Mplus Web Note #4 (www.statmodel.com).

Takane, Y. & DeLeeuw, J. (1987). On the relationship between item response theory and factor analysis of discretized variables. Psychometrika, 52, 393-408.

Exploratory Factor Analysis

Exploratory Factor Analysis For Outcomes That Are Categorical, Censored, Counts

Rotation of the factor loading matrix as with continuous outcomes

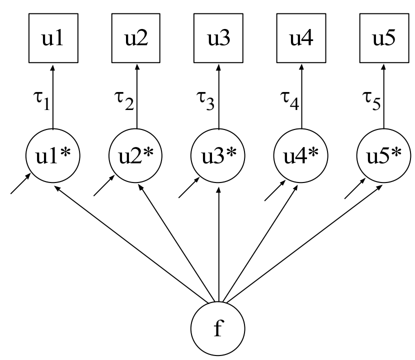
- Maximum-likelihood estimation
  - Computationally feasible for only a few factors, but can handle many items
  - Frequency table testing typically not useful
- Limited-information weighted least square estimation
  - Computationally feasible for many factors, but not huge number of items
  - Testing against bivariate tables
  - Modification indices for residual correlations

Assumptions Behind ML And WLS

Note that when assuming normal factors and using probit links, ML uses the same model as WLS. This is because normal factors and probit links result in multivariate normal u* variables. For model estimation, WLS uses the limited information of first- and second-order moments, thresholds and sample correlations of the multivariate normal u* variables (tetrachoric, polychoric, and polyserial correlations), whereas ML uses full information from all moments of the data.

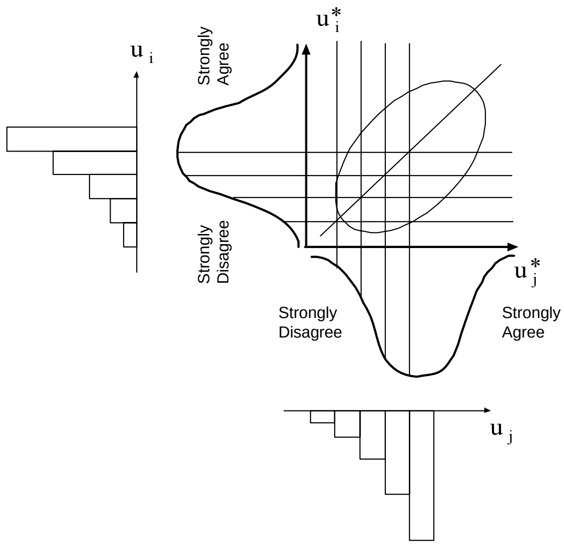
115

Latent Response Variable Formulation Of A Factor Model



116

Latent Response Variable Correlations



117

Sample Statistics With Categorical Outcomes And Weighted Least Squares Estimation

- Types of u^* correlations (normality assumed)
  - Both dichotomous – tetrachoric
  - Both polytomous – polychoric
  - One dichotomous, one continuous – biserial
  - One polytomous, one continuous – polyserial

- Analysis choices
  - Case A – no x variables – use u^* correlations
  - Case B – x variables present
    - – Use u^* correlations (full normality of u^* and x assumed)
    - – Use regression-based statistics (conditional normality of u^* given x assumed)

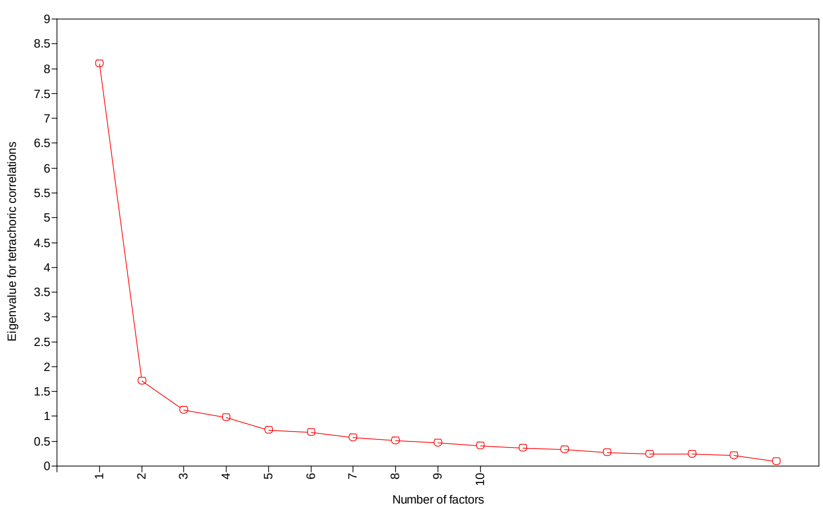
118

Exploratory Factor Analysis Of 17 ASB Items Using WLSM

```
TITLE:      EFA using WLSM
DATA:       FILE = asb.dat;
            FORMAT = 34X 54F2.0;
VARIABLE:   NAMES = property fight shoplift lt50 gt50 force threat
            injure pot drug
            soldpot solddrug con auto bldg goods gambling
            dsm1-dsm22 sex black hisp single divorce dropout
            college onset f1 f2 f3
            age94 cohort dep abuse;
            USEVAR = property-gambling;
            CATEGORICAL = property-gambling;
ANALYSIS:   TYPE = EFA 1 5;
OUTPUT:     MODINDICES;
PLOT:       TYPE = PLOT3;
```

119

Eigenvalue Plot For Tetrachoric Correlations Among 17 ASB Items



120

Output Excerpts 3- And 4-Factor WLSM EFA Of 17 ASB Items

```
EXPLORATORY FACTOR ANALYSIS WITH 3 FACTOR(S):

TESTS OF MODEL FIT

Chi-Square Test of Model Fit

          Value                              584.356*
          Degrees of Freedom                      88
          P-Value                             0.0000

*  The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV
   cannot be used for chi-square difference tests.  MLM, MLR and
   WLSM chi-square difference testing is described in the Mplus
   Technical Appendices at www.statmodel.com.  See chi-square
   difference testing in the index of the Mplus User's Guide.

Chi-Square Test of Model Fit for the Baseline Model

          Value                            53652.583
          Degrees of Freedom                     136
          P-Value                             0.0000
```

121

Output Excerpts 3- And 4-Factor WLSM EFA Of 17 ASB Items (Continued)

```
CFI/TLI

          CFI                                  0.991
          TLI                                  0.986

Number of Free Parameters                         48

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                             0.028

SRMR (Standardized Root Mean Square Residual)

          Value                                0.045

MINIMUM ROTATION FUNCTION VALUE       0.08510
```

122

Output Excerpts 3- And 4-Factor WLSM EFA Of 17 ASB Items (Continued)

QUARTIMIN ROTATED LOADINGS

	1	2	3
PROPERTY	**0.669**	0.179	-0.036
FIGHT	0.266	**0.548**	-0.121
SHOPLIFT	**0.600**	-0.028	0.185
LT50	**0.818**	-0.185	0.046
GT50	**0.807**	0.003	0.016
FORCE	0.379	0.344	0.000
THREAT	-0.008	**0.821**	0.049
INJURE	-0.022	**0.761**	0.101
POT	-0.051	0.001	**0.903**
DRUG	-0.021	-0.020	**0.897**
SOLDPOT	0.126	0.058	**0.759**
SOLDDRUG	0.175	0.083	**0.606**
CON	0.460	0.228	-0.065

123

Output Excerpts 3- And 4-Factor WLSM EFA Of 17 ASB Items (Continued)

	1	2	3
AUTO	**0.460**	0.139	0.073
BLDG	**0.797**	0.033	0.017
GOODS	0.700	0.109	0.066
GAMBLING	0.314	0.327	0.092

QUARTIMIN FACTOR CORRELATIONS

1	1.000		
2	0.598	1.000	
3	0.614	0.371	1.000

124

Output Excerpts 3- And 4-Factor WLSM EFA Of 17 ASB Items (Continued)

```
EXPLORATORY FACTOR ANALYSIS WITH 4 FACTOR(S):
TESTS OF MODEL FIT
Chi-Square Test of Model Fit
              Value                             303.340*
              Degrees of Freedom                      74
              P-Value                             0.0000

*  The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV
cannot be used for chi-square difference tests.  MLM, MLR and
WLSM chi-square difference testing is described in the Mplus
Technical Appendices at www.statmodel.com.  See chi-square
difference testing in the index of the Mplus User's Guide.
```


Output Excerpts 3- And 4-Factor WLSM EFA Of 17 ASB Items (Continued)

```
Chi-Square Test of Model Fit for the Baseline Model
          Value                         53652.583
          Degrees of Freedom                  136
          P-Value                          0.0000
CFI/TLI
          CFI                               0.996
          TLI                               0.992
Number of Free Parameters                      62
RMSEA (Root Mean Square Error Of Approximation)
          Estimate                          0.021
SRMR (Standardized Root Mean Square Residual)
          Value                             0.026
MINIMUM ROTATION FUNCTION VALUE       0.19546
```

Output Excerpts 3- And 4-Factor WLSM EFA Of 17 ASB Items (Continued)

QUARTIMIN ROTATED LOADINGS

	1	2	3	4
PROPERTY	**0.670**	0.191	-0.006	-0.043
FIGHT	0.290	**0.537**	-0.060	-0.098
SHOPLIFT	**0.679**	-0.001	0.225	-0.159
LT50	**0.817**	-0.152	0.066	-0.049
GT50	**0.762**	-0.008	-0.036	0.154
FORCE	0.257	0.288	-0.195	0.491
THREAT	0.003	**0.858**	0.101	-0.078
INJURE	-0.036	**0.728**	0.056	0.162
POT	0.041	0.074	**0.923**	-0.069
DRUG	0.051	0.007	**0.717**	0.227
SOLDPOT	0.149	0.070	**0.598**	0.281
SOLDDRUG	0.065	-0.037	0.269	**0.791**
CON	0.420	0.223	-0.072	0.081

127

Output Excerpts 3- And 4-Factor WLSM EFA Of 17 ASB Items (Continued)

	1	2	3	4
AUTO	**0.446**	0.138	0.051	0.074
BLDG	**0.770**	0.042	0.010	0.055
GOODS	**0.662**	0.109	0.030	0.126
GAMBLING	0.208	0.270	-0.083	0.449

QUARTIMIN FACTOR CORRELATIONS

1	1.000			
2	0.571	1.000		
3	0.485	0.230	1.000	
4	0.481	0.312	0.376	1.000

128

Practical Issues In The Analysis Of Categorical Outcomes

129

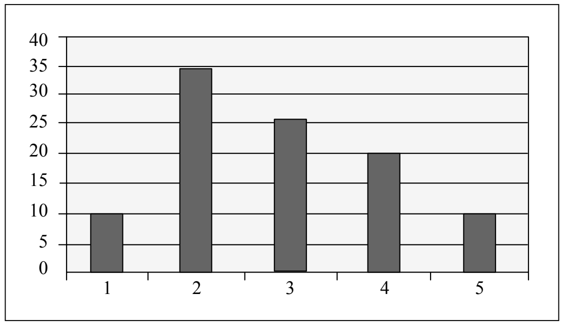
Overview Of Practical Issues In The Analysis Of Categorical Outcomes

- When Is A Variable Best Treated As Categorical?
  - Less dependent on number of categories than the presence of floor and ceiling effects
  - When the aim is to estimate probabilities or odds
- What's Wrong With Treating Categorical Variables As Continuous Variables?
  - Correlations will be attenuated particularly when there are floor and ceiling effects
  - Can lead to factors that reflect item difficulty extremeness
  - Predicted probabilities can be outside the 0/1 range

130

Approaches To Use With Categorical Data

- Data that lead to incorrect standard errors and chi-square under normality assumption

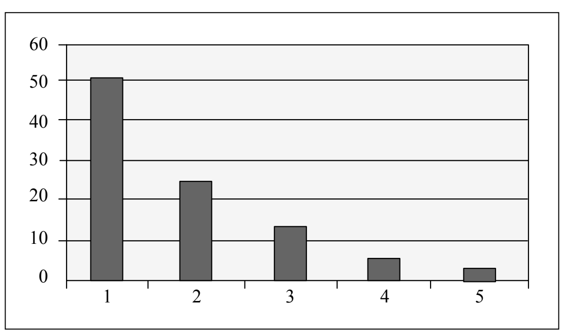


- Transform variable and treat as a continuous variable
- Treat as a continuous variable and use non-normality robust maximum likelihood estimation

131

Approaches To Use With Categorical Data (Continued)

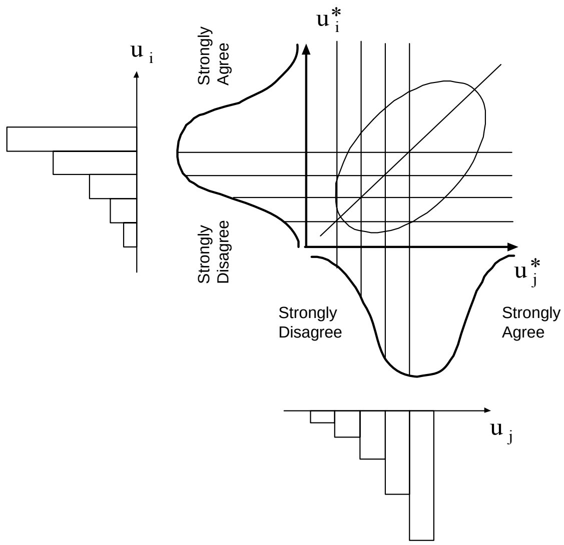
- Data that lead to incorrect standard errors, chi-square, and parameter estimates under normality assumption



- Treat as a categorical variable

132

Latent Response Variable Correlations



Distortions Of Underlying Correlation Structure

Pearson product-moment correlations unsuited to categorical variables due to limitation in range.

Example: $P\,(u_1) = 0.5, \quad P\,(u_2 = 1) = 0.2$

Gives max Pearson correlation = 0.5

		Variable 1		
		0	1	
Variable 2	0	50	30	
	1	0	20	20
			50	100

Distortions Of Underlying Correlation Structure (Continued)

Phi coefficient (Pearson correlation):

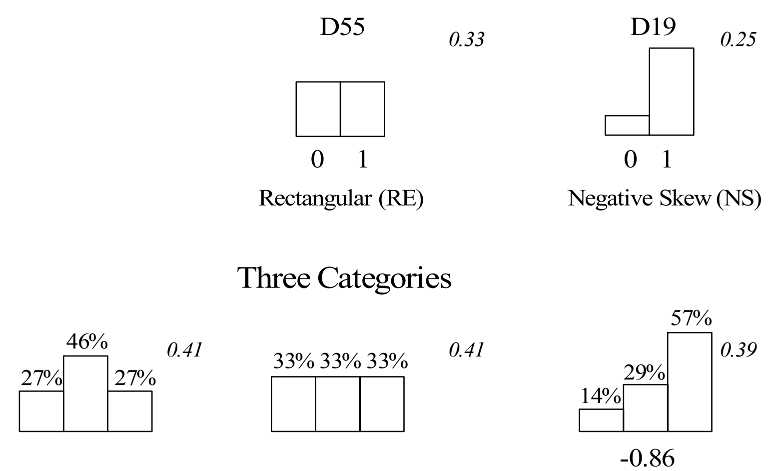
$$R = \frac{\text{Cov}(u_1, u_2)}{\text{SD}(u_1)\,\text{SD}(u_2)} =$$

$$\frac{P(u_1 = 1 \text{ and } u_2 = 1) - P(u_1 = 1)\,P(u_2 = 1)}{\sqrt{P(u_1 = 1)\,[1 - P(u_1 = 1)]}\;\sqrt{P(u_2 = 1)\,[1 - P(u_2 = 1)]}}$$

$$R_{max.} = \frac{0.2 - 0.5 \text{ x } 0.2}{\sqrt{.5 \text{ x } .5}\;\sqrt{.2 \text{ x } .8}} = \frac{0.1}{0.2} = 0.5$$

Correlational Attenuation

Correlation between underlying continuous u^* variables = 0.5



Correlational Attenuation (Continued)

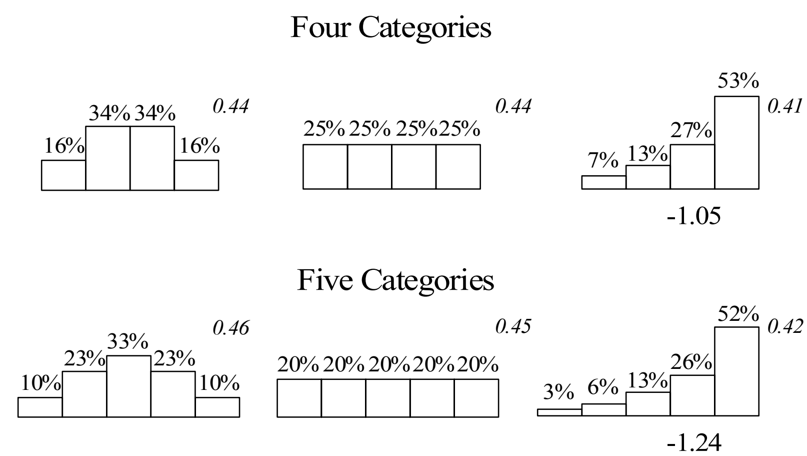


Table 1 (Part 2)
Pearson Correlations for True Correlations = 0.50

	D19	D28	D37	D46	D55	D64	D73	D82	D91	3SY	3RE	3NS	3PS
D19	25												
D28	26	30											
D37	26	30	32										
D46	24	30	32	33									
D55	23	28	31	33	33								
D64	20	26	30	23	33	33							
D73	18	23	27	30	31	32	32						
D82	15	20	23	26	28	30	30	30					
D91	10	15	18	20	22	24	26	26	25				
3SY	26	32	35	36	37	36	35	32	26	41			
3RE	25	31	35	36	37	36	35	31	25	41	41		
3NS	29	33	35	36	35	33	30	26	20	39	39	39	
3PS	20	26	30	33	35	36	35	33	29	39	39	34	39
4SY	27	33	36	38	38	38	36	33	27	43	43	40	40
4RE	26	33	36	38	38	38	36	33	26	42	42	40	40
4NS	30	35	36	36	35	33	30	27	20	40	39	40	34
3PS	20	27	31	34	35	36	36	35	30	40	39	34	40
5SY	28	34	37	38	39	38	37	34	28	44	43	41	41
4RE	27	33	37	38	39	38	37	33	27	43	43	41	41
5NS	31	35	36	36	35	33	30	26	20	40	39	40	34
5PS	20	26	30	33	35	36	36	35	31	40	39	34	40
CON	29	35	38	39	40	39	38	35	29	45	45	42	42
	D19	D28	D37	D46	D55	D64	D73	D82	D91	3SY	3RE	3NS	3PS

Pearson Correlations for True Correlations = 0.50

	4SY	4RE	4NS	4PS	5SY	5RE	5NS	5PS	CON
4SY	44								
4RE	44	44							
4NS	41	41	41						
4PS	41	41	35	41					
5SY	45	45	42	42	46				
5RE	45	45	41	41	46	45			
5NS	41	40	42	34	42	41	42		
5PS	41	41	34	42	42	41	34	42	
CON	47	46	43	43	48	47	44	44	50
	4SY	4RE	4NS	4PS	5SY	5RE	5NS	5PS	CON

139

Approaches To Use With Categorical Outcomes

- Items, Testlets, Sums, Or Factor Scores?
  - A sum of at least 15 unidimensional items is reliable
  - Testlets can be used as continuous indicators
  - Factor scores can be estimated as in IRT
- Sample Size
  - Larger than for continuous variables
  - Univariate and bivariate distributions should contain several observations per cell

140

Further Readings On Factor Analysis Of Categorical Outcomes

Bock, R.D., Gibbons, R., & Muraki, E.J. (1998). Full information item factor analysis. Applied Psychological Measurement, 12, 261-280.

Flora, D.B. & Curran, P.J., (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. Psychological Methods, 9, 466-491.

Muthén, B. (1989). Dichotomous factor analysis of symptom data. In Eaton & Bohrnstedt (Eds.), Latent variable models for dichotomous outcomes: Analysis of data from the epidemiological Catchment Area program (pp.19-65), a special issue of Sociological Methods & Research, 18, 19-65.

Muthen, B. & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. British Journal of Mathematical and Statistical Psychology, 38, 171-189.

Muthen, B. & Kaplan, D. (1992). A comparison of some methodologies for the factor analysis of non-normal Likert variables: A note on the size of the model. British Journal of Mathematical and Statistical Psychology, 45, 19-30.

CFA With Covariates (MIMIC)