**Incorporação de indicadores
categóricos ordinais em
modelos de equações estruturais**

Bruno Cesar Bistaffa

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
MESTRE EM CIÊNCIAS

Programa: Estatística

Orientador: Profa. Dra. Lúcia Pereira Barroso

São Paulo, Dezembro de 2010

Incorporação de indicadores categóricos ordinais em modelos de equações estruturais

Este exemplar corresponde à redação
final da dissertação devidamente corrigida
e defendida por Bruno Cesar Bistaffa
e aprovada pela Comissão Julgadora.

Comissão Julgadora:

- Profa. Dra. Lúcia Pereira Barroso (orientadora) - IME-USP
- Profa. Dra. Julia Maria Pavan Soler - IME-USP
- Prof. Dr. Marcelo Ângelo Cirillo - UFLA

Agradecimentos

Em primeiro lugar gostaria de agradecer a Deus, senhor maior, e a toda a minha família: meu pai Cesar, minha mãe Eliana e minha irmã Maria Julia. Obrigado por tudo o que fizeram e continuam fazendo por mim; jamais teria chegado até aqui sem o amor e carinho de vocês.

Obrigado à Teledata Informações e Tecnologia (Telecheque) e ao Itaú-Unibanco por me permitirem conciliar minha vida acadêmica e minha vida profissional, deixando eu realizar os dois maiores sonhos da minha vida: fazer o mestrado e trabalhar com estatística. Gostaria de agradecer a todos os meus colegas de trabalho, Walter Alfieri, José Luiz Praxedes, Carlos Valdesuso, Genaro Lins Dueire e em especial aos meus gestores imediatos Wilson Fukushima e Rodolfo Arashiro Rodriguez, aprendi coisas com vocês que levarei de lição para o resto da vida.

Obrigado a todos os professores do IME-USP com quem tive o prazer de estudar: Vanderlei da Costa Bueno, Chang Chiann, Nikolai Valtchev Kolev, Carlos Alberto de Bragança Pereira, Sergio Wechsler e em especial à professora Lúcia Pereira Barroso por ser esta pessoa maravilhosa, que sempre esteve de portas abertas a me ajudar na elaboração desta dissertação. Gostaria de agradecer também a todos os meus colegas de mestrado, em especial ao amigo Marcos Alves dos Santos, pela inestimável ajuda durante o mestrado.

Por fim gostaria de agradecer a minha namorada Luana Isabel Naves. Obrigado por entender minhas ausências, minhas crises, meus sonhos esquisitos e meu amor pela estatística; este trabalho jamais teria terminado se você não estivesse do meu lado me apoiando nas minhas decisões, mesmo sem entender a maioria delas.

CRANEGAM, obrigado pela proteção e oração. Vocês sabem que estão em todos os lugares em que vou.

Resumo

A modelagem de equações estruturais é uma técnica estatística multivariada que permite analisar variáveis que não podem ser medidas diretamente, mas que podem ser estimadas através de indicadores. Dado o poder que esta técnica tem em acomodar diversas situações em um único modelo, sua aplicação vem crescendo nas diversas áreas do conhecimento. Diante disto, este trabalho teve por objetivo avaliar a incorporação de indicadores categóricos ordinais em modelos de equações estruturais, fazendo um resumo dos principais procedimentos teóricos e subjetivos presentes no processo de estimação de um modelo, avaliando as suposições violadas quando indicadores ordinais são utilizados para estimar variáveis latentes e criando diretrizes que devem ser seguidas para a correta estimação dos parâmetros do modelo. Mostramos que as correlações especiais (correlação tetracórica, correlação policórica, correlação biserial e correlação poliserial) são as melhores escolhas como medida de associação entre indicadores, que estimam com maior precisão a correlação entre duas variáveis, em comparação à correlação de Pearson, e que são robustas a desvios de simetria e curtose. Por fim aplicamos os conceitos apresentados ao longo deste estudo a dois modelos hipotéticos com o objetivo de avaliar as diferenças entre os parâmetros estimados quando um modelo é ajustado utilizando a matriz de correlações especiais em substituição à matriz de correlação de Pearson.

Palavras-chave: Modelagem de Equações Estruturais, Correlação Tetracórica, Correlação Policórica, Correlação Biserial, Correlação Poliserial.

Abstract

The structural equation modeling is a multivariate statistical technique that allows us to analyze variables that can't be measured directly but can be estimated through indicators. Given the power that this technique has to accommodate several situations in a single model, its application has increased in several areas of the knowledge. At first, this study aimed to evaluate the incorporation of ordinal categorical indicators in structural equation models, making a summary of the major theoretical and subjective procedures of estimating the present model, assessing the assumptions that are violated when ordinal indicators are used to estimate latent variables and creating guidelines to be followed to correct estimation of model parameters. We show that the special correlations (tetrachoric correlation, polychoric correlation, biserial correlation and poliserrial correlation) are the best choices as a measure of association between indicators, that estimate more accurately the correlation between two variables, compared to Pearson's correlation, and that they are robust to deviations from symmetry and kurtosis. Finally, we apply the concepts presented in this study to two hypothetical models to evaluate the differences between the estimated parameters when a model is adjusted using the special correlation matrix substituting the Pearson's correlation matrix.

Keywords: Structural Equation Modeling, Tetrachoric Correlation, Polychoric Correlation, Biserial Correlation, Poliserrial Correlation.

Sumário

1	Introdução	1
2	Modelagem de Equações Estruturais	3
2.1	Histórico	4
2.2	Exemplo Intuitivo	5
2.3	Modelo Matemático	8
2.4	Diagrama de Caminhos	10
2.4.1	Elementos de um Diagrama de Caminhos	10
2.4.2	Equações na Forma Matricial	11
2.4.3	Efeitos Direto, Indireto e Total	13
2.5	Matriz de Covariância Implícita	14
2.6	Validade e Confiabilidade	15
2.7	Relações entre as Variáveis	16
2.7.1	Isolamento	17
2.7.2	Associação	17
2.7.3	Direção	17
2.7.4	Diferenças entre Indicadores Reflexivos e Formativos	18
2.8	Identificação do Modelo	19
2.8.1	Regra t	19
2.8.2	Regra 2 Passos	19
2.8.3	Recomendações para Identificação	21
2.9	Estimação do Modelo	22
2.9.1	Máxima Verossimilhança	22
2.9.2	Mínimos Quadrados Generalizados	23
2.9.3	Mínimos Quadrados Não-Ponderados	24
2.9.4	Suposição sobre a Distribuição dos Dados	24
2.9.5	Mínimos Quadrados Ponderados	25
2.9.6	Mínimos Quadrados Ponderados Diagonalmente	26
2.9.7	Estimadores Elípticos	27
2.9.8	Covariâncias <i>versus</i> Correlações	28
2.10	Avaliação do Ajuste do Modelo	28

2.10.1	Avaliação dos Parâmetros Estimados	29
2.10.2	Ajuste Geral do Modelo	31
2.10.3	Medidas de Ajuste Alternativas	32
2.11	Reespecificação do Modelo	38
2.11.1	Teste da Razão de Verossimilhanças	38
2.11.2	Teste do Multiplicador de Lagrange	39
2.11.3	Teste de Wald	40
2.11.4	Limitações dos Testes Estatísticos	40
2.12	Resumo do Capítulo	41
3	Incorporação de Variáveis Categóricas	43
3.1	Suposições Violadas e suas Consequências	43
3.2	Procedimentos Corretivos	44
3.3	Resumo do Capítulo	46
4	Correlações Especiais	47
4.1	Exemplo Intuitivo	47
4.2	Correlação Policórica	49
4.3	Correlação Tetracórica	56
4.4	Correlação Poliserial	60
4.5	Correlação Biserial	65
4.6	Desvios de Normalidade	69
4.6.1	Desvios de Curtose	69
4.6.2	Desvios de Simetria	75
4.7	Resumo do Capítulo	81
5	Simulações para o Modelo de Equações Estruturais	83
5.1	Metodologia Utilizada	83
5.2	Modelo 1	84
5.3	Modelo 2	93
5.4	Resumo do Capítulo	105
6	Conclusões	107
A	Matriz de Correlações Especiais	109
B	Função para Ajustar Modelos SEM	115
C	Matriz de Pesos	125
	Referências Bibliográficas	127

Capítulo 1

Introdução

A Modelagem de Equações Estruturais (SEM, do Inglês *Structural Equation Modeling*) é uma técnica estatística multivariada que permite acomodar muitas situações em um único conjunto de equações. Graças ao seu poder de análise, vem ganhando popularidade nos últimos anos ao ponto de encontrarmos muitas aplicações da técnica a problemas de diversas áreas do conhecimento.

Ela se diferencia das demais técnicas multivariadas pelo fato que é possível analisar relações entre variáveis que não conseguimos medir diretamente, mas que podemos estimar através de indicadores.

Dado o poder que esta técnica tem, é imprescindível que o pesquisador conheça e compreenda o impacto que cada decisão, tomada no processo de estimação do modelo, tem sobre os resultados finais.

Este trabalho tem por objetivo avaliar a incorporação de indicadores categóricos ordinais em modelos de equações estruturais, criando diretrizes que deveriam ser adotadas para a correta estimação dos modelos estudados.

Inicialmente, no Capítulo 2, fazemos uma ampla revisão dos principais procedimentos envolvidos no processo de estimação de um modelo de equações estruturais. Apresentamos o diagrama de caminhos, as suposições inerentes ao modelo, as funções de discrepância utilizadas no processo de estimação dos parâmetros, as medidas de qualidade do ajuste obtido e técnicas para reespecificação do modelo quando o ajuste não é razoável.

No Capítulo 3, avaliamos as suposições violadas quando indicadores categóricos ordinais são utilizados para estimar variáveis latentes e mostramos que nestes casos o procedimento mais indicado é calcular correlações especiais (correlação tetracórica, correlação policórica, correlação biserial e correlação poliserial) como medida de associação entre os indicadores do modelo.

No Capítulo 4, apresentamos as principais formas de estimação das correlações especiais e realizamos estudos de simulação para cada uma destas medidas a fim de determinarmos as vantagens em utilizá-las em substituição à correlação de Pearson, considerando os indicadores ordinais como variáveis contínuas; também avaliamos se as correlações especiais são robustas

a desvios de simetria e curtose.

No Capítulo 5, aplicamos os conceitos apresentados ao longo deste estudo a dois modelos hipotéticos, avaliando as diferenças nos parâmetros estimados quando utilizamos a matriz de correlações especiais, em substituição à correlação de Pearson, e estimamos o modelo utilizando as diferentes funções de discrepância existentes.

Por fim, apresentamos nossas conclusões e sugestões para estudos futuros.

Capítulo 2

Modelagem de Equações Estruturais

A Modelagem de Equações Estruturais é uma técnica estatística multivariada que possibilita analisar várias relações de dependência entre si (ela permite que uma variável dependente em uma relação seja a variável independente em outra) e aceita a incorporação de variáveis que não podem ser mensuradas diretamente.

Ao contrário das demais metodologias multivariadas, ela é uma técnica baseada em conceitos teóricos e é indicada para análises confirmatórias onde o conhecimento substantivo do pesquisador tem papel decisivo em, praticamente, todas as etapas; desta forma é importante que o pesquisador conheça e compreenda todos os procedimentos que envolvem a modelagem de equações estruturais.

A SEM possui duas partes: (1) o modelo de mensuração e (2) o modelo estrutural.

Enquanto o modelo de mensuração especifica como as variáveis latentes são mensuradas pelas variáveis observadas, descrevendo sua validade e confiabilidade, o modelo estrutural especifica as relações de causa e efeito entre as variáveis latentes, apresentando os efeitos causais e o total da variância não explicada.

Goldberger (1972) apresenta três situações em que SEM é indicada:

- Quando existem variáveis explicativas não-observáveis;
- Quando as variáveis observáveis possuem erros de medida;
- Quando existe interdependência entre as variáveis observadas.

Bollen (1989) afirma que é possível verificar que SEM engloba algumas das principais técnicas estatísticas multivariadas:

- Análise de Regressão;
- Sistema de Equações Simultâneas;
- Análise Fatorial Confirmatória;
- Análise de Correlação Canônica;

- Análise de Painel de Dados;
- Análise de Variância;
- Análise de Covariância;
- Modelo de Indicadores Múltiplos.

O objetivo deste capítulo é apresentar um resumo dos procedimentos e conceitos envolvidos na aplicação da modelagem de equações estruturais.

2.1 Histórico

Não existe um consenso sobre quem é o precursor da modelagem de equações estruturais; esta pergunta fica mais complicada de ser respondida à medida que esta técnica multivariada continua sendo desenvolvida até hoje.

De forma sucinta, pode-se afirmar que SEM é consequência da evolução de três componentes que estão presentes no modelo geral. Estes elementos são:

- A análise de caminhos;
- A síntese conceitual de variável latente e modelos de mensuração;
- Os procedimentos de estimação.

Traçando o caminho de cada um dos componentes de SEM, é possível ter uma ideia sobre a origem da técnica.

O inventor da análise de caminhos é o biométrico Sewall Wright (1918, 1921, 1934, 1960) e ela é composta de 3 partes: o diagrama de caminhos, as equações relacionando correlações (ou covariâncias) aos parâmetros e a decomposição dos efeitos.

O diagrama de caminhos é uma simples representação do sistema de equações simultâneas; usando esta representação pictórica, Wright propôs um conjunto de regras que relacionam as correlações aos parâmetros do modelo. Com os parâmetros estimados é possível distinguir os efeitos direto, indireto e total de uma variável em outra; o efeito direto é a influência de uma variável em outra; o efeito indireto é a influência de uma variável em outra com o intermédio de pelo menos uma outra variável; o efeito total é a soma dos efeitos direto e indireto.

A primeira aplicação da análise de caminhos foi feita por Wright em 1918, entretanto a maioria dos artigos, utilizando a técnica, são datados da década de 1960.

Um novo impulso para a técnica foi dado quando Jöreskog (1973), Keesing (1972) e Wiley (1973) desenvolveram modelos de equações estruturais mais gerais, incorporando diagrama de caminhos em suas representações.

Além da análise de caminhos, a síntese de variáveis latentes e modelos de mensuração são essenciais para SEM. A análise fatorial tradicional, desenvolvida por Spearman (1904),

ênfatizava a relação entre fatores latentes e variáveis observadas, todavia não existe nenhuma referência à relação estrutural entre as variáveis latentes.

Em econometria o foco era a relação estrutural entre as variáveis observadas, com eventual referência a situações com erros nas variáveis. Os exemplos de aplicação de Wright evidenciaram que os modelos de econometria podiam ser identificados e estimados. Desta forma, foram desenvolvidos diversos modelos que acomodavam muitas situações entre as décadas de 1960 e 1970. Entretanto, estes trabalhos não especificavam um modelo geral que poderia ser aplicado a qualquer problema.

Foi somente nos trabalhos de Jöreskog (1973), Keesing (1972) e Wiley (1973) que o modelo geral foi proposto. Este modelo era dividido em duas partes. A primeira era composta de um modelo de variáveis latentes, similar aos modelos de equações simultâneas da econometria, e a segunda era composta de um modelo de mensuração que apontava indicadores como efeitos das variáveis latentes, como na análise fatorial. Eles desenvolveram expressões matriciais que permitiram a aplicabilidade deste modelo a inúmeros problemas individuais.

Bentler e Bonett (1980), McArdle e McDonald (1984) e outros autores propuseram representações alternativas de equações estruturais gerais.

A primeira aplicação de Wright utilizou de procedimentos *ad hoc* para encontrar estimativas dos parâmetros e, neste ponto, os trabalhos em econometria e psicometria se tornaram indispensáveis. Em econometria a propriedade dos estimadores das equações estruturais com variáveis observadas foi estabelecida; por exemplo, por Goldberger (1964). Em psicometria os trabalhos de Lawley (1940), Anderson e Rubin (1956) e Jöreskog (1969) ajudaram a formular testes estatísticos para a análise fatorial; Bock e Borgman (1966) desenvolveram uma análise da estrutura de covariância com o objetivo de estimar a variância de variáveis latentes; Jöreskog (1973) propôs estimadores de máxima verossimilhança baseados em distribuições multinormais das variáveis observadas; Jöreskog e Goldberger (1972) e Browne (1974, 1982, 1984) propuseram estimadores de mínimos quadrados generalizados para uma maior flexibilidade nas aplicações; Browne (1982, 1984) sugeriu estimadores que assumiam distribuições elípticas ou arbitrárias; Bentler (1983) sugeriu estimadores que tratam produtos de momentos de variáveis observadas; Muthén (1984, 1987), entre outros, generalizou estes modelos para variáveis observadas limitadas ou ordinais.

Além de todos os autores que ajudaram a desenvolver a modelagem de equações estruturais, os softwares estatísticos ajudaram a popularizar a técnica. Os principais são o LISREL (Jöreskog e Sorbom, 1986), o Proc Calis e o Proc Tcalis do SAS, o AMOS da SPSS e o pacote *sem* do R.

2.2 Exemplo Intuitivo

Este exemplo foi retirado de Hair *et al.* (1998) e mostra o poder que a modelagem de equações estruturais possui em acomodar múltiplas relações de dependência inter-relacionadas

em um único modelo.

Suponha que administradores de uma empresa estejam interessados em aumentar a retenção e a produtividade de seus empregados através de um melhor entendimento das atitudes e motivações com relação à empresa.

De uma pesquisa inicial feita junto aos funcionários dos Recursos Humanos (RH) foram identificadas três atitudes que deveriam ser indicativos da satisfação dos funcionários. Estas atitudes foram: satisfação com o emprego, compromisso organizacional e probabilidade de mudança de emprego. Munido destas informações, eles desenvolveram relações de dependência conectando cada atitude com seus antecedentes. A Figura 2.1, denominada diagrama de caminhos, é uma representação pictórica das relações de inter-dependência sendo que as setas retilíneas da figura representam o impacto das variáveis independentes nas variáveis dependentes e as retas curvilíneas representam correlações entre as variáveis.

Figura 2.1: Diagrama de caminhos da relação de atitudes dos empregados.

O diagrama de caminhos é a base para a análise de caminhos, um procedimento que estima a intensidade das relações representadas através de uma figura. A análise de caminhos utiliza a matriz de correlações entre as variáveis para calcular estas intensidades pelo fato de que a correlação entre duas variáveis quaisquer podem ser representadas como a soma dos caminhos compostos que conectam esses pontos.

Um caminho composto pode ser definido pelas seguintes regras:

- Depois de ir adiante através de uma seta, o caminho não pode retroceder, entretanto o caminho pode ir para trás quantas vezes for necessário antes de ir para frente;
- Não pode ir através da mesma variável mais do que uma vez;
- Pode incluir somente uma seta curvada (correlação entre duas variáveis).

Para simplificar nosso exemplo utilizaremos apenas uma relação de dependência que faz parte do modelo completo, ela pode ser visualizada na Figura 2.2.

Figura 2.2: Diagrama de caminhos da relação causal entre X_1 e X_2 com Y_1 .

As variáveis X_1 e X_2 são correlacionadas entre si e prevêem a variável dependente Y_1 ; esta relação pode ser escrita como:

$$Y_1 = b_1X_1 + b_2X_2.$$

Pelas regras da análise de caminhos podemos utilizar as correlações entre os constructos para estimar as relações causais representadas pelos coeficientes b_1 e b_2 ; para facilitar a compreensão, os caminhos causais são representados por A , B e C .

Suponha que os constructos sejam correlacionados entre si pela matriz de correlação dada pela Tabela 2.1.

Tabela 2.1: Correlações entre as variáveis X_1 , X_2 e Y_1 .

	X_1	X_2	Y_1
X_1	1		
X_2	0,6	1	
Y_1	0,5	0,7	1

Desta forma podemos formar três equações:

$$\begin{cases} Cor_{X_1, X_2} = A \\ Cor_{X_1, Y_1} = B + AC \\ Cor_{X_2, Y_1} = C + AB \end{cases}$$

e resolvendo este sistema encontramos os seguintes valores:

$$\begin{cases} A = 0,60 \\ B = 0,33 \\ C = 0,53 \end{cases}.$$

A abordagem apresentada neste exemplo permite que a análise de caminhos resolva qualquer relação causal baseando-se apenas nas correlações e não importando quão grande seja o diagrama de caminhos.

2.3 Modelo Matemático

A notação utilizada ao longo de todo este capítulo é baseada em [Bollen \(1989\)](#).

O modelo completo de SEM é composto por um sistema de equações simultâneas que possuem variáveis aleatórias e parâmetros estruturais.

Os parâmetros estruturais medem a associação entre as variáveis do modelo. Eles são constantes invariantes que fornecem uma medida da relação “causal” entre variáveis.

As variáveis aleatórias representam conceitos unidimensionais e podem ser de duas formas: observadas ou latentes. As **variáveis observadas**, ou indicadores, possuem erros de mensuração e são utilizadas para prever as variáveis latentes. É suposto que elas são perfeitamente correlacionadas (ou pelo menos altamente correlacionadas) com a variável latente que medem. As **variáveis latentes**, ou constructos, não são medidas diretamente e podem ser classificadas em endógenas ou exógenas.

As **variáveis latentes exógenas** são apenas preditas ou a “causa” de outras variáveis ou constructos do modelo; dizemos que os fatores que as determinam estão fora do modelo. As **variáveis latentes endógenas** são determinadas por variáveis que estão presentes no modelo e são explicadas parcialmente pelo modelo; desta forma elas possuem erro de mensuração.

É importante ressaltar que as equações representadas em SEM são lineares nas variáveis e nos parâmetros, entretanto existem modelos que permitem a incorporação de relações não-lineares e de constantes nas equações.

Desta forma, o modelo completo de SEM é representado pelo seguinte sistema de equações lineares:

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}, \quad (2.1)$$

$$\mathbf{y} = \boldsymbol{\Lambda}_y\boldsymbol{\eta} + \boldsymbol{\epsilon}, \quad (2.2)$$

$$\mathbf{x} = \boldsymbol{\Lambda}_x\boldsymbol{\xi} + \boldsymbol{\delta}, \quad (2.3)$$

Sendo que:

- $\boldsymbol{\eta}' = (\eta_1, \dots, \eta_m)$ são as variáveis latentes endógenas;
- $\boldsymbol{\xi}' = (\xi_1, \dots, \xi_n)$ são as variáveis latentes exógenas;
- $\mathbf{B}_{(m \times m)}$ é a matriz de coeficientes que representam os efeitos diretos das variáveis $\boldsymbol{\eta}$ em outras variáveis $\boldsymbol{\eta}$;
- $\boldsymbol{\Gamma}_{(m \times n)}$ é a matriz de coeficientes que representam os efeitos diretos das variáveis $\boldsymbol{\xi}$ em variáveis $\boldsymbol{\eta}$;

- $\zeta' = (\zeta_1, \dots, \zeta_m)$ é o vetor aleatório de erros nas equações estruturais;
- $y' = (y_1, \dots, y_p)$ e $x' = (x_1, \dots, x_q)$ são os vetores de variáveis observadas;
- Λ_y é a matriz de coeficientes que representam o impacto das variáveis η nas variáveis y ;
- Λ_x é a matriz de coeficientes que representam o impacto das variáveis ξ nas variáveis x ;
- $\epsilon' = (\epsilon_1, \dots, \epsilon_p)$ e $\delta' = (\delta_1, \dots, \delta_q)$ são os vetores de erro de mensuração.

Com as seguintes pressuposições:

- ζ e ξ são não-correlacionados;
- $I - B$ é não singular;
- $E[\zeta] = 0$;
- $E[\eta] = 0$;
- $E[\xi] = 0$;
- ϵ é não-correlacionado com η ;
- δ é não-correlacionado com ξ ;
- ϵ e δ são não-correlacionados entre si, mas podem ter suas respectivas variáveis correlacionadas entre si;
- ζ , ϵ e δ são mutuamente não-correlacionados.

E com as seguintes matrizes de covariância:

- $\Phi_{(n \times n)} = E[\xi\xi']$;
- $\Psi_{(m \times m)} = E[\zeta\zeta']$;
- $\Theta_{\epsilon(p \times p)} = E[\epsilon\epsilon']$;
- $\Theta_{\delta(q \times q)} = E[\delta\delta']$.

A equação (2.1) refere-se ao modelo estrutural e as equações (2.2) e (2.3) referem-se ao modelo de mensuração.

Os coeficientes de Λ_y e de Λ_x são coeficientes de regressão, eles indicam o efeito de uma variável latente em um indicador. Para podermos interpretar completamente estes coeficientes é necessário associar uma escala à variável latente. Geralmente associa-se a escala da variável latente a um dos indicadores ou a variância da variável latente.

2.4 Diagrama de Caminhos

O diagrama de caminhos é uma representação do sistema de equações simultâneas sob estudo. Talvez sua principal vantagem é a de representar, em uma figura, as relações “causais” assumidas como verdadeiras no modelo.

É importante ressaltar que em um diagrama de caminhos é assumido que todas as relações causais estão indicadas e que estas relações são lineares.

2.4.1 Elementos de um Diagrama de Caminhos

Para compreender um diagrama de caminhos e seu poder em sumarizar todas as relações em uma única figura é necessário compreender todos os símbolos que podem ser utilizados nesta representação.

Um diagrama de caminhos pode possuir os seguintes símbolos:

- Variáveis latentes são representadas por elipses;
- Variáveis observadas são representadas por retângulos;
- Variáveis de erro são representadas dentro da figura;
- Flexas retilíneas únicas significam que a variável da base da flexa “causa” a variável do outro extremo;
- Flexas retilíneas duplas significam que uma variável “causa” e é “causada” pela outra;
- Flexas curvilíneas de duplo sentido significam correlação entre as variáveis.

A Figura 2.3 representa um diagrama com todos os elementos descritos acima.

Pela Figura 2.3, temos que:

- ξ_1, ξ_2, η_1 e η_2 são variáveis latentes;
- $x_1, x_2, x_3, x_4, y_1, y_2, y_3$ e y_4 são indicadores;
- $\delta_1, \delta_2, \delta_3, \delta_4, \epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4, \zeta_1$ e ζ_2 são variáveis de erro;
- ξ_1 e ξ_2 têm influência direta sobre η_1 ;
- ξ_1 tem influência direta sobre η_2, x_1 e x_2 ;
- ξ_2 tem influência direta sobre x_3 e x_4 ;
- η_1 e η_2 têm influência recíproca;
- η_1 tem influência direta sobre y_1 e y_2 ;
- η_2 tem influência direta sobre y_3 e y_4 ;

Figura 2.3: Diagrama de caminhos do modelo completo.

- ξ_1 e ξ_2 podem ser correlacionados;

Observe que as variáveis ξ_1 e ξ_2 não possuem nenhuma flexa retilínea apontando para elas, desta forma são denominadas variáveis latentes exógenas; já as variáveis η_1 e η_2 sofrem influência de outras variáveis latentes, desta forma são denominadas variáveis latentes endógenas.

2.4.2 Equações na Forma Matricial

Uma vez que todas as relações foram representadas no diagrama de caminhos, os esforços devem se concentrar em transformar estas relações nas equações do modelo estrutural e de mensuração.

Da Figura 2.3 podemos escrever as equações (2.1), (2.2) e (2.3).

O modelo estrutural, que representa as relações entre as variáveis latentes, possui as seguintes equações:

$$\begin{aligned} \eta_1 &= \beta_{12}\eta_2 + \gamma_{11}\xi_1 + \gamma_{12}\xi_2 + \zeta_1, \\ \eta_2 &= \beta_{21}\eta_1 + \gamma_{21}\xi_1 + \gamma_{22}\xi_2 + \zeta_2, \end{aligned} \quad (2.4)$$

ou em forma matricial:

$$\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} 0 & \beta_{12} \\ \beta_{21} & 0 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} + \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & 0 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix}. \quad (2.5)$$

Com as seguintes matrizes de covariância:

$$\mathbf{\Phi} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \quad (2.6)$$

e

$$\mathbf{\Psi} = \begin{pmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{pmatrix}. \quad (2.7)$$

As equações do modelo de mensuração, que especificam como as variáveis latentes são medidas pelas variáveis observadas, são:

$$\begin{aligned} y_1 &= \lambda_{11}^{(y)} \eta_1 + \epsilon_1 & x_1 &= \lambda_{11}^{(x)} \xi_1 + \delta_1 \\ y_2 &= \lambda_{21}^{(y)} \eta_1 + \epsilon_2 & x_2 &= \lambda_{21}^{(x)} \xi_1 + \delta_2 \\ y_3 &= \lambda_{32}^{(y)} \eta_2 + \epsilon_3 & x_3 &= \lambda_{32}^{(x)} \xi_2 + \delta_3 \\ y_4 &= \lambda_{42}^{(y)} \eta_2 + \epsilon_4 & x_4 &= \lambda_{42}^{(x)} \xi_2 + \delta_4 \end{aligned} \quad (2.8)$$

ou em forma matricial:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} \lambda_{11}^{(y)} & 0 \\ \lambda_{21}^{(y)} & 0 \\ 0 & \lambda_{32}^{(y)} \\ 0 & \lambda_{42}^{(y)} \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{pmatrix} \quad (2.9)$$

e

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \lambda_{11}^{(x)} & 0 \\ \lambda_{21}^{(x)} & 0 \\ 0 & \lambda_{32}^{(x)} \\ 0 & \lambda_{42}^{(x)} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{pmatrix}. \quad (2.10)$$

Com as seguintes matrizes de covariância:

$$\mathbf{\Theta}_{\epsilon} = \begin{pmatrix} \theta_{11}^{(\epsilon)} & 0 & 0 & 0 \\ 0 & \theta_{22}^{(\epsilon)} & 0 & 0 \\ 0 & 0 & \theta_{33}^{(\epsilon)} & 0 \\ 0 & 0 & 0 & \theta_{44}^{(\epsilon)} \end{pmatrix} \quad (2.11)$$

e

$$\mathbf{\Theta}_{\delta} = \begin{pmatrix} \theta_{11}^{(\delta)} & 0 & 0 & 0 \\ 0 & \theta_{22}^{(\delta)} & 0 & 0 \\ 0 & 0 & \theta_{33}^{(\delta)} & 0 \\ 0 & 0 & 0 & \theta_{44}^{(\delta)} \end{pmatrix}. \quad (2.12)$$

Os parâmetros do modelo que são fixados em 0, ou em qualquer outro valor, são denominados **parâmetros fixos**. Quando algum parâmetro é restringido ao valor de outro parâmetro (por exemplo, $\gamma_{11} = \gamma_{21}$), dizemos que estes são **parâmetros restritos**. Os

demais parâmetros, que não possuem restrições, são denominados **parâmetros livres**.

Quando o modelo de equações estruturais possui uma relação recíproca, tal como a de η_1 e η_2 , dizemos que o modelo é **não-recursivo** (uma relação indireta, ou composta, recíproca também faz com que o modelo seja considerado não-recursivo); estes modelos são mais complicados de serem analisados.

Para finalizar, se as relações de um modelo são bem definidas, ele é dito ser **estável**, caso contrário é denominado **instável**.

2.4.3 Efeitos Direto, Indireto e Total

Como já dito anteriormente, em um diagrama de caminhos existem três tipos de efeito: (1) direto, (2) indireto e (3) total.

O efeito direto é a influência de uma variável em outra sem o intermédio de nenhuma outra. O efeito indireto é a influência de uma variável em outra pelo intermédio de pelo menos uma outra variável. O efeito total é a soma dos efeitos direto e indireto.

Podemos resumir estes efeitos em uma simples equação:

$$\text{Efeito Total} = \text{Efeito Direto} + \text{Efeito Indireto}. \quad (2.13)$$

A fim de ilustrar estes efeitos, considere o diagrama de caminhos da Figura 2.3. O efeito direto de ξ_1 em η_1 é γ_{11} ; o efeito indireto de ξ_1 em η_1 é $\phi_{21}\gamma_{12}$; e o efeito total é $\gamma_{11} + \phi_{21}\gamma_{12}$.

É importante entender e compreender cada tipo de efeito a fim de ter uma melhor compreensão das relações entre as variáveis presentes no modelo.

Bollen (1989) apresenta a Tabela 2.2, que sumariza todos os efeitos em um modelo completo.

Tabela 2.2: Fórmulas para o cálculo dos efeitos direto, indireto e total.

Efeitos de ξ	Efeitos em		
	η	y	x
Direto	Γ	$\mathbf{0}$	Λ_x
Indireto	$(\mathbf{I} - \mathbf{B})^{-1}\Gamma - \Gamma$	$\Lambda_y(\mathbf{I} - \mathbf{B})^{-1}\Gamma$	$\mathbf{0}$
Total	$(\mathbf{I} - \mathbf{B})^{-1}\Gamma$	$\Lambda_y(\mathbf{I} - \mathbf{B})^{-1}\Gamma$	Λ_x
Efeitos de η	η	y	x
	$\mathbf{B}$	Λ_y	$\mathbf{0}$
Direto	$\mathbf{B}$	Λ_y	$\mathbf{0}$
Indireto	$(\mathbf{I} - \mathbf{B})^{-1} - \mathbf{I} - \mathbf{B}$	$\Lambda_y(\mathbf{I} - \mathbf{B})^{-1} - \Lambda_y$	$\mathbf{0}$
Total	$(\mathbf{I} - \mathbf{B})^{-1} - \mathbf{I}$	$\Lambda_y(\mathbf{I} - \mathbf{B})^{-1}$	$\mathbf{0}$

Uma condição suficiente, mas não necessária, para que todos estes efeitos possam ser calculados é que o valor absoluto do maior autovalor da matriz $\mathbf{B}$ seja menor que 1 (Bentler e Freeman, 1983).

2.5 Matriz de Covariância Implícita

A hipótese básica da modelagem de equações estruturais é:

$$\Sigma = \Sigma(\Theta), \quad (2.14)$$

sendo que Σ é a matriz de covariância populacional de $\mathbf{x}$ e $\mathbf{y}$ e $\Sigma(\Theta)$ é a matriz de covariância escrita como função dos parâmetros do modelo. A relação entre Σ e $\Sigma(\Theta)$ é a base para o entendimento da identificação, estimação e avaliação do modelo ajustado.

Como Σ é a matriz de covariâncias das variáveis observadas, podemos particioná-la tal que:

$$\Sigma = \begin{pmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{pmatrix}$$

e, através da álgebra de matrizes podemos escrever Σ como função dos parâmetros do modelo.

Considere primeiramente $\Sigma_{yy}(\Theta)$, a matriz de covariância implícita de $\mathbf{y}$:

$$\begin{aligned} \Sigma_{yy}(\Theta) &= E[\mathbf{y}\mathbf{y}'] \\ &= E[(\Lambda_y \boldsymbol{\eta} + \boldsymbol{\epsilon})(\Lambda_y \boldsymbol{\eta} + \boldsymbol{\epsilon})'] \\ &= E[(\Lambda_y \boldsymbol{\eta} + \boldsymbol{\epsilon})(\boldsymbol{\eta}' \Lambda_y' + \boldsymbol{\epsilon}')] \\ &= E[\Lambda_y \boldsymbol{\eta} \boldsymbol{\eta}' \Lambda_y' + \Lambda_y \boldsymbol{\eta} \boldsymbol{\epsilon}' + \boldsymbol{\epsilon} \boldsymbol{\eta}' \Lambda_y' + \boldsymbol{\epsilon} \boldsymbol{\epsilon}'] \\ &= \Lambda_y E[\boldsymbol{\eta} \boldsymbol{\eta}'] \Lambda_y' + \Lambda_y E[\boldsymbol{\eta} \boldsymbol{\epsilon}'] + E[\boldsymbol{\epsilon} \boldsymbol{\eta}'] \Lambda_y' + E[\boldsymbol{\epsilon} \boldsymbol{\epsilon}'] \\ &= \Lambda_y E[\boldsymbol{\eta} \boldsymbol{\eta}'] \Lambda_y' + \Theta_\epsilon. \end{aligned} \quad (2.15)$$

De (2.1), temos que:

$$\boldsymbol{\eta} = (\mathbf{I} - \mathbf{B})^{-1}(\Gamma \boldsymbol{\xi} + \boldsymbol{\zeta}). \quad (2.16)$$

Desta forma, substituindo (2.16) em (2.15), temos:

$$\begin{aligned} \Sigma_{yy}(\Theta) &= \Lambda_y E[\boldsymbol{\eta} \boldsymbol{\eta}'] \Lambda_y' + \Theta_\epsilon \\ &= \Lambda_y E \left[(\mathbf{I} - \mathbf{B})^{-1} (\Gamma \boldsymbol{\xi} + \boldsymbol{\zeta}) ((\mathbf{I} - \mathbf{B})^{-1} (\Gamma \boldsymbol{\xi} + \boldsymbol{\zeta}))' \right] \Lambda_y' + \Theta_\epsilon \\ &= \Lambda_y (\mathbf{I} - \mathbf{B})^{-1} E[(\Gamma \boldsymbol{\xi} + \boldsymbol{\zeta})(\Gamma \boldsymbol{\xi} + \boldsymbol{\zeta})'] (\mathbf{I} - \mathbf{B})^{-1'} \Lambda_y' + \Theta_\epsilon \\ &= \Lambda_y (\mathbf{I} - \mathbf{B})^{-1} \{ \Gamma E[\boldsymbol{\xi} \boldsymbol{\xi}'] \Gamma' + \Gamma E[\boldsymbol{\xi} \boldsymbol{\zeta}'] + E[\boldsymbol{\zeta} \boldsymbol{\xi}'] \Gamma' + E[\boldsymbol{\zeta} \boldsymbol{\zeta}'] \} (\mathbf{I} - \mathbf{B})^{-1'} \Lambda_y' + \Theta_\epsilon \\ &= \Lambda_y (\mathbf{I} - \mathbf{B})^{-1} [\Gamma \Phi \Gamma' + \Psi] (\mathbf{I} - \mathbf{B})^{-1'} \Lambda_y' + \Theta_\epsilon. \end{aligned} \quad (2.17)$$

A matriz de covariância implícita de $\mathbf{x}$, $\Sigma_{xx}(\Theta)$ é dada por:

$$\begin{aligned} \Sigma_{xx}(\Theta) &= E(\mathbf{x} \mathbf{x}') \\ &= E[(\Lambda_x \boldsymbol{\xi} + \boldsymbol{\zeta})(\Lambda_x \boldsymbol{\xi} + \boldsymbol{\zeta})'] \end{aligned}$$

$$\begin{aligned}
&= E[(\Lambda_x \xi + \zeta)(\xi' \Lambda_x' + \zeta')] \\
&= \Lambda_x E[\xi \xi'] \Lambda_x' + \Lambda_x E[\xi \zeta'] + E[\zeta \xi'] \Lambda_x' + E[\zeta \zeta'] \\
&= \Lambda_x \Phi \Lambda_x' + \Theta_\delta.
\end{aligned} \tag{2.18}$$

A parte final é a matriz de covariância implícita é $\Sigma_{xy}(\Theta)$, as covariâncias implícitas de $\mathbf{x}$ com $\mathbf{y}$:

$$\begin{aligned}
\Sigma_{xy}(\Theta) &= E(\mathbf{x}\mathbf{y}') \\
&= E[(\Lambda_y \eta + \epsilon)(\Lambda_x \xi + \zeta)'] \\
&= E[(\Lambda_y \eta + \epsilon)(\xi' \Lambda_x' + \zeta')] \\
&= E[\Lambda_y \eta \xi' \Lambda_x + \Lambda_y \eta \zeta' + \epsilon \xi' \Lambda_x' + \epsilon \zeta'] \\
&= \Lambda_y E[\eta \xi'] \Lambda_x' \\
&= \Lambda_y E[(I - B)^{-1}(\Gamma \xi + \zeta) \xi'] \Lambda_x' \\
&= \Lambda_y (I - B)^{-1} E[\Gamma \xi \xi' + \zeta \xi'] \Lambda_x' \\
&= \Lambda_y (I - B)^{-1} \Gamma \Phi \Lambda_x'.
\end{aligned} \tag{2.19}$$

Com os resultados de (2.17), (2.18) e (2.19), a matriz de covariância implícita de $\mathbf{x}$ e $\mathbf{y}$ pode ser escrita como:

$$\Sigma(\Theta) = \begin{pmatrix} \Lambda_y (I - B)^{-1} [\Gamma \Phi \Gamma' + \Psi] (I - B)^{-1'} \Lambda_y' \Theta_\epsilon & \Lambda_y (I - B)^{-1} \Gamma \Phi \Lambda_x' \\ \Lambda_x \Phi \Gamma' (I - B)^{-1'} \Lambda_y' & \Lambda_x \Phi \Lambda_x' + \Theta_\delta \end{pmatrix}. \tag{2.20}$$

Na prática, como não conhecemos as verdadeiras covariâncias populacionais, resolvemos o sistema $\mathbf{S} = \Sigma(\Theta)$ para encontrarmos as estimativas dos parâmetros do modelo, em que $\mathbf{S}$ é a matriz de covariância amostral usual.

2.6 Validade e Confiabilidade

Um modelo de equações estruturais é dividido na parte de mensuração e na parte estrutural. À medida que o modelo de mensuração indica como é que as variáveis latentes são medidas, é importante garantir a validade e a confiabilidade dos indicadores a fim de determinar o quão adequados eles são como medidas de uma variável latente.

A validade de um indicador diz respeito ao fato de que ele está medindo o que se propôs a medir.

Bollen (1989) e Sanches (2005) afirmam que, mesmo existindo algumas medidas desenvolvidas para tentar mensurar a validade de uma medida, não é possível promover a validade. Entretanto é possível desenvolver um forte suporte para ela.

A confiabilidade pode ser vista como a consistência da mensuração de um modelo. Isto não é garantia da validade de um modelo, mas ao contrário, a confiabilidade só pode ser

aplicada uma vez que garantimos a validade do indicadores.

Existem várias formas de se estimar a confiabilidade de medidas, entretanto a mais utilizada e aceita é o **alpha de Cronback**, que é calculado por:

$$\rho_{H,H}^2 = \left(\frac{q}{q-1} \right) \left(1 - \frac{\sum_{i=1}^q Var(X_i)}{Var(\sum_{i=1}^q X_i)} \right), \quad (2.21)$$

sendo que q é o número de indicadores do constructo.

2.7 Relações entre as Variáveis

A modelagem de equações estruturais possui muitas suposições que necessitam ser verificadas a fim de que as conclusões obtidas através do modelo sejam válidas.

Uma das suposições conceitualmente mais importantes é a de que existe uma estrutura de causalidade nos modelos de equações estruturais.

Como já dito anteriormente, quando existe uma relação de causa entre uma variável latente e uma variável observada, dizemos que a variável observada é um indicador da variável latente. A questão é que podem existir dois tipos de indicadores: os (1) indicadores reflexivos e os (2) indicadores formativos.

Os **indicadores reflexivos** são causados pela variável latente e são denominados reflexivos porque representam uma manifestação da variável latente; são os mais comuns e formam a base da análise fatorial. Os **indicadores formativos** são a causa da variável latente.

A Figura 2.4 apresenta dois diagramas de caminhos, uma com indicadores formativos e outra com indicadores reflexivos.

Figura 2.4: Diagramas de caminho com indicadores reflexivos e formativos.

Dentre as várias definições que existem sobre causa, utilizamos a de [Sanches \(2005\)](#). “Se tivermos duas variáveis, X e Y , isoladas de qualquer influência externa, e se a cada mudança de X , Y também sobre influência, então dizemos que X causa Y ”.

Desta simples definição, obtemos os elementos básicos para a verificação da existência de uma relação causal entre duas variáveis: (1) isolamento, (2) associação e (3) direção.

2.7.1 Isolamento

O isolamento se refere ao fato de que só conseguimos determinar a relação causal entre duas variáveis se elas estiverem isoladas de qualquer outro tipo de influência; entretanto é praticamente impossível isolar todas as variáveis de um modelo de equações estruturais das influências externas.

Desta forma, o isolamento total pode ser substituído por um pseudo-isolamento e incorporamos um distúrbio ao modelo que supomos representar o efeito desconhecido, e pequeno, das causas externas não incluídas no modelo. Para estes erros assume-se que são não correlacionados entre si e não correlacionados com as variáveis exógenas.

Embora a suposição de isolamento possa ser relaxada pela suposição de pseudo-isolamento, uma suposição ainda existe e deve ser respeitada. Bollen (1989) apresenta um exemplo em que ele demonstra que quando a suposição de pseudo-isolamento é violada, a estimação dos parâmetros do modelo fica comprometida.

2.7.2 Associação

À medida que a possível causa e efeito estão isolados de qualquer forma de influência, necessitamos que as duas variáveis sejam associadas para que possamos determinar a relação de causa entre as variáveis.

Bollen (1989) afirma que podem existir relações de causa sem que exista associação; esta situação se deve ao fato de que um indicador pode ser causado por mais de uma variável latente; quando estas relações possuem magnitude parecidas, mas com sinais contrários, o efeito de uma acaba anulando o efeito de outra.

Uma relação de causa pode ser justificada por uma correlação parcial entre as variáveis uma vez que apenas a correlação bivariada não é uma condição necessária, nem suficiente, para determinar a relação causal. A verificação prática desta associação pode ser dificultada pelos erros amostrais ou pela multicolinearidade presente no modelo.

2.7.3 Direção

A possibilidade de uma associação ser causal repousa na necessidade de haver uma direção para esta causa, ou seja, a causa do modelo deve preceder o efeito avaliado.

Quando temos um modelo hipotético e desejamos avaliá-lo em SEM, necessitamos explicitar suas relações através de um diagrama de caminhos e na figura ficará evidenciada a direção da causa. É importante que a direção representada no diagrama também seja corretamente especificada na forma algébrica do problema com o objetivo de que o modelo ajustado reflita a teoria analisada. Desta forma verificamos que a definição da direção da causa é um ponto crítico e deve ser feita com a devida atenção.

A direção da causa é o que diferencia os indicadores reflexivos dos indicadores formativos.

2.7.4 Diferenças entre Indicadores Reflexivos e Formativos

Ao longo desta seção fizemos uma pequena introdução do que são indicadores reflexivos, indicadores formativos e como verificar se existe uma relação causal entre duas variáveis; os esforços se concentram agora em explicar, de forma sucinta, as principais diferenças entre estes dois tipos de indicadores. Leitores interessados no assunto são encorajados a ler [Bollen \(1989\)](#) e [Sanches \(2005\)](#).

A utilização de indicadores reflexivos na maioria das aplicações de SEM se deve ao fato de que a técnica incorporou variáveis latentes da análise fatorial e nessa técnica não existem indicadores formativos. A questão crucial é que a utilização de um indicador formativo como um indicador reflexivo (ou seja, indicando erroneamente a direção causal entre duas variáveis) acaba por produzir um ajuste muito ruim em um modelo que teria um bom ajuste caso a relação causal fosse corretamente especificada ([Sanches, 2005](#)).

A primeira diferença entre indicadores reflexivos e formativos aparece já na definição de variável latente. Quando utilizamos indicadores reflexivos, a variável latente é definida como uma variável que causa os indicadores observados e suas causadoras estão fora do modelo; já quando temos indicadores formativos, a variável latente é não observável e pode ser encarada como uma combinação linear de seus indicadores. De forma sucinta podemos afirmar que nos casos em que a variável latente possui somente indicadores reflexivos, seu conceito é muito mais abstrato do que quando a variável latente se relaciona com indicadores formativos.

Na escolha de indicadores para as variáveis latentes, [Bollen \(1989\)](#) afirma que é preferível escolher variáveis observadas que sejam correlacionadas com o constructo, entretanto esta regra só é válida para a escolha de indicadores reflexivos. [Bollen e Lennox \(1991\)](#) afirmam que quando temos indicadores formativos no modelo sob estudo, a procura por indicadores correlacionados com a variável latente pode ocasionar a omissão de alguma informação importante para o modelo.

Se de um lado não necessitamos que os indicadores formativos sejam correlacionados com a variável latente, por outro a multicolinearidade pode dificultar a distinção do impacto de cada indicador na variável latente, uma vez que existe variância comum entre eles. Nos casos em que a multicolinearidade entre indicadores reflexivos existe, as consequências são mínimas pelo fato de que esta estrutura de correlação é justificada pela própria natureza da relação entre os indicadores e a variável latente.

Nos casos em que o objetivo final da análise é obter valores estimados para a variável latente, [Bollen e Lennox \(1991\)](#) indicam obter estas novas variáveis através de somas não ponderadas de seus indicadores. Quando os indicadores são reflexivos, estes indicadores são denominados **escalas**; já quando os indicadores são formativos, eles são denominados **índices**.

2.8 Identificação do Modelo

Antes de estimarmos os parâmetros do modelo de equações estruturais desenvolvido anteriormente, necessitamos determinar se ele é, ou não, identificável.

Teoricamente a identificação tem por objetivo demonstrar que os parâmetros desconhecidos do modelo são funções apenas de parâmetros identificáveis e que estas funções levam somente a soluções únicas, ou seja, que não são impróprias ou indefinidas; na prática as regras que apresentaremos nesta seção servirão muito mais para afirmar que o modelo não é identificável.

Quando um parâmetro é função apenas de parâmetros que podem ser estimados e esta função leva a um valor único, dizemos que o parâmetro é **identificado**. Se todos os parâmetros do modelo forem identificados, dizemos que o modelo é **identificado**.

Se um parâmetro é uma função que não leva a uma solução única, dizemos que o parâmetro é **sub-identificado**. Se ao menos um parâmetro do modelo é sub-identificado, dizemos que o modelo é **sub-identificado**.

Já quando temos excesso de informação para estimarmos um parâmetro, dizemos que ele é **superidentificado** e se isto ocorre com pelo menos um dos parâmetros, dizemos que o modelo é **superidentificado**.

Para todas as regras que apresentaremos a seguir, assumimos que cada variável latente está associada a uma escala, pois sem esta suposição, identificação não é possível.

2.8.1 Regra t

A regra t é uma regra necessária, mas não suficiente. Ela diz que o número de parâmetros desconhecidos tem de ser menor ou igual ao número de elementos não redundantes da matriz de covariância, ou seja:

$$t \leq \frac{(p+q)(p+q+1)}{2}, \quad (2.22)$$

sendo que:

- t é o número de parâmetros livres do modelo;
- p é o número de variáveis observáveis dos η 's;
- q é o número de variáveis observáveis dos ξ 's.

2.8.2 Regra 2 Passos

A regra 2 passos é composta de duas partes; na primeira parte é verificado se os parâmetros de mensuração são identificados e na segunda parte é verificado se os parâmetros do modelo estrutural também são identificados.

No primeiro passo tratamos o modelo SEM como uma análise fatorial confirmatória. Isto se faz considerando as variáveis observadas $\mathbf{x}$ e $\mathbf{y}$ como variáveis presentes somente no vetor

$\mathbf{x}$ e as variáveis ξ e η somente como variáveis ξ . Com o modelo reformulado, testamos a identificabilidade do modelo com as seguintes regras:

Regra t

Análoga à apresentada anteriormente.

Regra 3 Indicadores

Uma condição suficiente para identificar um modelo com apenas uma variável latente é o modelo ter pelo menos três indicadores com cargas diferentes de zero e Θ_δ ser diagonal.

Para um modelo com duas ou mais variáveis latentes, as condições suficientes são:

- três ou mais indicadores por variável latente;
- cada linha de $\Lambda_{\mathbf{x}}$ tem um, e apenas um, elemento diferente de zero;
- Θ_δ é uma matriz diagonal.

Nesta regra não existem restrições à matriz Φ .

Regra 2 Indicadores

Esta regra é uma condição suficiente, mas não necessária para modelos com mais de uma variável latente. As condições são:

- Cada linha de $\Lambda_{\mathbf{x}}$ possui um, e apenas um, valor diferente de zero;
- Existem pelo menos dois indicadores por variável latente;
- Cada linha de Φ tem pelo menos um elemento diferente de zero fora da diagonal;
- Θ_δ é uma matriz diagonal.

Se for possível estabelecer identificação no primeiro passo, vamos para o segundo passo, senão a regra de identificação não é aplicável.

No segundo passo, as equações do modelo de mensuração são tratadas como se fossem equações estruturais de variáveis observadas, ou seja, assumimos que cada variável latente é uma variável observada que é perfeitamente medida. Com as reformulações do modelo, testamos a identificabilidade com as seguintes regras:

Regra t

Análoga à apresentada anteriormente.

Regra B nula

A regra B nula é uma condição suficiente, mas não necessária, para a identificação do modelo. Ela diz que se a matriz B for nula, os parâmetros desconhecidos são identificáveis.

Regra Recursiva

Da mesma forma que a regra anterior, esta é uma regra suficiente.

Se a matriz B for triangular inferior e Ψ for diagonal, então o modelo é identificável.

Condição de Ordem e de Posto

Exceto pela regra t, as outras regras apresentadas fazem restrições a B ou a Ψ . Modelos não-recursivos falham em satisfazer estas restrições, mas deve haver uma identificação estabelecida em algum outro caminho do modelo. Estas regras são aplicáveis em modelos que assumem ser sem erro de mensuração e que as variáveis exógenas ($\mathbf{x}$) são não-correlacionadas com os erros (ζ).

A condição de ordem diz que uma condição necessária para uma equação ser identificada é que o número de variáveis excluídas da equação seja pelo menos $(p - 1)$.

Para verificar a condição de ordem, utilizamos a matriz $C = [(I - B), -\Gamma]$. Para cada linha da matriz resultante, conta-se o número de zeros. Se cada linha desta matriz possuir $(p - 1)$ ou mais zeros, a condição de ordem foi satisfeita.

Para verificarmos a condição de posto para a i -ésima equação, eliminamos todas as colunas da matriz C que não possuem zeros na i -ésima linha de C e usando as colunas restantes criamos a matriz C_i . Uma condição necessária e suficiente para a identificação da i -ésima equação é que o posto de C_i seja igual a $(p - 1)$.

Se todas as equações são identificadas, então o modelo é identificado.

Se o primeiro passo mostra que os parâmetros de mensuração são identificados e o segundo passo mostra que os parâmetros do modelo de variáveis latentes também é identificável, isto é suficiente para identificar o modelo proposto.

É importante salientar que existem modelos SEM que não atendem a regra 2 passos e ainda assim podem ser estimados.

2.8.3 Recomendações para Identificação

Enquanto nenhuma regra necessária e suficiente for criada para a identificação de modelos gerais de SEM, a identificação continuará sendo um desafio.

Como diretriz para a identificação dos modelos, recomendamos que pelo menos uma das regras que apresentamos nesta seção seja utilizada.

Caso alguma das regras dê indícios de que o modelo não é identificado, recomendamos utilizar as regras propostas por Jöreskog e Sörbom (1986). Eles sugerem que se estime o

modelo SEM com a matriz $\mathbf{S}$ usual e que seja armazenada a matriz de covariância estimada $\hat{\Sigma}$ (substituindo as estimativas de Θ em $\Sigma(\Theta)$, obtemos a matriz de covariância implícita $\hat{\Sigma}$); depois substitua a matriz $\mathbf{S}$ pela matriz $\hat{\Sigma}$, estimada no passo anterior, e re-estime o modelo; se o modelo é identificado, as novas estimativas devem ser, aproximadamente, iguais às estimadas da primeira vez.

Também é possível estimar modelos com diferentes soluções iniciais e verificar se as soluções estão convergindo para um mesmo valor ou estimar o modelo com sub-amostras e verificar se as soluções são consistentes.

Os testes empíricos apresentados são aplicáveis a todos os tipos de modelos de equações estruturais.

2.9 Estimação do Modelo

Como já dito anteriormente, a estimação do modelo completo de SEM repousa nas relações entre a matriz de covariância Σ e os parâmetros estruturais Θ , como ficou evidenciado em (2.20).

Se o modelo sob estudo é correto e os parâmetros populacionais são conhecidos, temos que $\Sigma = \Sigma(\Theta)$, entretanto como não conhecemos Σ , utilizamos a matriz de covariância amostral $\mathbf{S}$ para estimarmos os parâmetros do modelo.

As funções de ajuste, F , que apresentaremos são baseadas em $\mathbf{S}$ e em $\Sigma(\Theta)$ e são denominadas **funções de discrepância**.

As funções de discrepância possuem as seguintes propriedades:

- F é um escalar, para cada Θ fixado;
- $F \geq 0$;
- $F = 0$ se e somente se $\Sigma(\Theta) = \mathbf{S}$;
- F é contínua em $\Sigma(\Theta)$ e em $\mathbf{S}$.

2.9.1 Máxima Verossimilhança

A função de ajuste mais utilizada em modelos SEM é a função baseada em máxima verossimilhança (ML, do inglês *Maximum Likelihood*), e que depende da suposição de que os indicadores seguem uma distribuição normal multivariada.

Neste caso a função a ser minimizada é:

$$F_{ML} = \log |\Sigma(\Theta)| + \text{tr}(\mathbf{S}\Sigma^{-1}(\Theta)) - \log |\mathbf{S}| - (p + q). \quad (2.23)$$

As propriedades dos estimadores de máxima verossimilhança são:

- Mesmo podendo serem viciados para pequenas amostras, os estimadores são assintoticamente não-viciados;
- São consistentes;
- São assintoticamente eficientes;
- São assintoticamente normais;
- São invariantes.

Como os estimadores são assintoticamente normais, quando o tamanho amostral aumenta é possível realizar testes de significância estatística para cada um dos parâmetros estimados.

Além dos testes individuais, é possível realizar um teste para o ajuste global do modelo. A hipótese nula, H_0 , do teste é de que $\Sigma = \Sigma(\Theta)$ e sob esta hipótese temos que:

$$(n-1)F_{ML} \sim \chi^2 \left(\frac{1}{2}(p+q)(p+q+1) - t \right). \quad (2.24)$$

Neste caso, se a hipótese nula não for rejeitada, temos evidências de que o modelo teve um bom ajuste.

É importante ressaltar que este teste deve ser utilizado com parcimônia; é necessário ter um tamanho amostral suficientemente grande e que as variáveis sejam normais¹.

2.9.2 Mínimos Quadrados Generalizados

A estimação por mínimos quadrados generalizados (GLS, do inglês *Generalized Least Square*) tem por objetivo minimizar a diferença entre os valores observados em $\mathbf{S}$ e os correspondentes elementos preditos em $\Sigma(\Theta)$.

A função a ser minimizada é:

$$F_{GLS} = \frac{1}{2} \text{tr} \left(\{[\mathbf{S} - \Sigma(\Theta)]\mathbf{W}^{-1}\}^2 \right), \quad (2.25)$$

sendo que $\mathbf{W}^{-1}$ é uma matriz de pesos para a matriz residual. Diversas matrizes de pesos podem ser utilizadas, entretanto a mais usual é $\mathbf{W}^{-1} = \mathbf{S}^{-1}$.

As estimativas obtidas pelo método GLS possuem as mesmas propriedades das estimativas obtidas pelo método ML e os testes individuais e de ajuste global, apresentados anteriormente, continuam válidos.

Sob a suposição de normalidade multivariada, tanto os estimadores de ML, quando os estimadores de GLS, são ótimos quando o tamanho amostral é grande (Bollen, 1989).

¹Relaxamento destas suposições serão discutidas posteriormente.

2.9.3 Mínimos Quadrados Não-Ponderados

A função de ajuste para o método de mínimos quadrados não-ponderados (ULS, do inglês *Unweighted Least Square*) é:

$$F_{ULS} = \frac{1}{2} \text{tr} [(\mathbf{S} - \mathbf{\Sigma}(\mathbf{\Theta}))^2]. \quad (2.26)$$

Esta função minimiza a soma de quadrados de cada elemento da matriz residual $(\mathbf{S} - \mathbf{\Sigma}(\mathbf{\Theta}))$.

Mesmo sendo a mais simples de todas as funções de discrepância, ela determina estimadores consistentes de $\mathbf{\Theta}$ e não exige a suposição de que as variáveis observadas seguem alguma distribuição em particular. Em contrapartida, ela não possui os estimadores assintoticamente mais eficientes, não possui escala invariante e o teste χ^2 não pode ser avaliado.

2.9.4 Suposição sobre a Distribuição dos Dados

Até agora nos limitamos a dizer que as variáveis observadas deveriam ser normalmente distribuídas para que os resultados das estimativas obtidas pelo método de ML e GLS fossem válidas; a partir de agora explicitaremos as consequências de se violar estas suposições.

No desenvolvimento da função F_{ML} , a distribuição normal multivariada das variáveis observadas é considerada como sendo verdadeira; quando essa suposição não é válida, mas os dados não possuem curtose excessiva, as funções F_{ML} e F_{GLS} ainda podem ser avaliadas. Sob estas condições, os estimadores obtidos são consistentes e assintoticamente eficientes, além de ser possível realizar testes estatísticos para os parâmetros e para o ajuste global do modelo sob estudo.

Browne (1982, 1984) apresenta a Tabela 2.3, que sumariza todas as propriedades dos estimadores ML e GLS.

Tabela 2.3: *Propriedades dos estimadores ML e GLS.*

Distribuição das Var. observadas	Propriedade dos Estimadores ML e GLS			
	Consistência	Ef. Assintótica	aCov($\hat{\mathbf{\Theta}}$)	Teste χ^2
Multinormal	Sim	Sim	Correto	Correto
Sem Curtose	Sim	Sim	Correto	Correto
Elíptica	Sim	Sim	Incorreto	Incorreto
“Arbitrária”	Sim	Não	Incorreto	Incorreto

Bollen (1989) afirma que quando o tamanho amostral cresce, as estimativas obtidas pelos métodos de ML, GLS e ULS convergem para o verdadeiro parâmetro populacional, mesmo que a distribuição dos dados não seja normal.

Desta forma podemos afirmar que os resultados numéricos das estimativas obtidas por qualquer um dos métodos de estimação possíveis para SEM serão aproximadamente iguais, o que acaba por produzir interpretações gerais similares. O que muda é que é necessário que

algumas suposições sejam feitas a fim de que a medida de ajuste global e os erros-padrão das estimativas dos parâmetros sejam assintoticamente corretas.

Em suma, temos que a não-normalidade dos dados não afeta a consistência dos estimadores F_{ML} , F_{GLS} e F_{ULS} , mas a curtose excessiva acaba por eliminar a eficiência assintótica e inviabiliza os testes estatísticos para os parâmetros e para o ajuste global do modelo.

O leitor é encorajado a testar a normalidade e o nível de curtose dos dados a fim de garantir que as propriedades dos estimadores obtidos pelos métodos ML e GLS sejam válidas.

Nos casos em que estas suposições não possam ser verificadas podemos testar algumas opções de correção. Uma opção é a transformação das variáveis não-normais. Outra alternativa é empregar reamostragem *Bootstrap* ou *Jackknife* e produzir testes estatísticos não-paramétricos. Entretanto é necessário avaliar o tempo gasto para que estes tipos de procedimentos sejam realizados.

Uma terceira forma é utilizar estimadores alternativos que permitam que os dados sejam não-normais, mas que mesmo assim sejam assintoticamente eficientes. Os métodos de mínimos quadrados ponderados (WLS, do inglês *Weighted Least Square*), mínimos quadrados ponderados diagonalmente (DWLS, do inglês *Diagonal Weighted Least Square*) e estimadores elípticos (E) são algumas das opções possíveis (Bollen, 1989).

2.9.5 Mínimos Quadrados Ponderados

As funções de discrepância em modelos de equações estruturais são casos particulares de uma função de discrepância global. Esta função pode ser escrita como:

$$\begin{aligned} F &= (\mathbf{s} - \boldsymbol{\sigma}(\boldsymbol{\theta}))' \mathbf{W}^{-1} (\mathbf{s} - \boldsymbol{\sigma}(\boldsymbol{\theta})) \\ &= \sum_{g=1}^k \sum_{h=1}^g \sum_{i=1}^k \sum_{j=1}^i w^{gh,ij} (s_{gh} - \sigma_{gh})(s_{ij} - \sigma_{ij}), \end{aligned} \quad (2.27)$$

sendo que:

- $\mathbf{s}' = (s_{11}, \dots, s_{kk})$ é um vetor com os elementos da matriz triangular inferior de $\mathbf{S}$;
- $\boldsymbol{\sigma}(\boldsymbol{\theta})' = (\sigma_{11}, \dots, \sigma_{kk})$ é um vetor de elementos correspondentes de $\boldsymbol{\Sigma}(\boldsymbol{\theta})$;
- $w^{gh,ij}$ são elementos de uma matriz de pesos positiva-definida $\mathbf{W}^{-1}$;
- $u = \frac{1}{2}k(k+1)$.

Desta forma, a matriz $\mathbf{W}^{-1}$ deve ser tal que seus elementos sejam estimativas consistentes da covariância assintótica entre s_{gh} e s_{ij} .

A covariância assintótica entre s_{gh} e s_{ij} é:

$$\text{acov}(s_{ij}, s_{gh}) = n^{-1}(\sigma_{ijgh} - \sigma_{ij}\sigma_{gh}), \quad (2.28)$$

sendo que σ_{ijgh} é o momento de quarta ordem em torno da média.

Quando cada um dos elementos de $\mathbf{W}$ é uma estimativa de (2.28), temos que estamos estimando os parâmetros do modelo pelo método WLS. Este método também é comumente chamado de **função de distribuição arbitrária** ou **estimador livre de distribuição**.

Os elementos de (2.28) podem ser estimados por:

$$\hat{s}_{ijgh} = \frac{1}{n} \sum_{t=1}^n (\mathbf{Z}_{it} - \bar{\mathbf{Z}}_i)(\mathbf{Z}_{jt} - \bar{\mathbf{Z}}_j)(\mathbf{Z}_{gt} - \bar{\mathbf{Z}}_g)(\mathbf{Z}_{ht} - \bar{\mathbf{Z}}_h) \quad (2.29)$$

e

$$\hat{s}_{ij} = \frac{1}{n} \sum_{t=1}^n (\mathbf{Z}_{it} - \bar{\mathbf{Z}}_i)(\mathbf{Z}_{jt} - \bar{\mathbf{Z}}_j), \quad (2.30)$$

sendo que:

- $\mathbf{Z}$ é a matriz de dados observados;
- $\bar{\mathbf{Z}} = \frac{1}{n} \mathbf{Z}' \mathbf{1}$ é o vetor de médias;
- n é o tamanho amostral.

A vantagem dos estimadores de WLS é que são feitas suposições mínimas sobre a distribuição dos dados. Fornecendo uma matriz de pesos como a dada anteriormente, ele se transforma em um estimador eficiente que fornece uma matriz de covariâncias para as estimativas de Θ e um teste χ^2 de ajuste global do modelo. Já a desvantagem de se utilizar este método está no tempo computacional necessário para inverter uma matriz quadrada e estimar o modelo.

Existem indícios de que o tamanho amostral necessário para a convergência das estimativas, obtidas neste método, é maior do que para F_{ML} , F_{GLS} e F_{ULS} ; não se tem muitas ideias sobre as propriedades dos estimadores de F_{WLS} quando o tamanho amostral é pequeno e não é claro qual é o ganho de performance, utilizando F_{WLS} em comparação às demais formas de estimação, quando existem desvios moderados de normalidade (Bollen, 1989).

2.9.6 Mínimos Quadrados Ponderados Diagonalmente

Como dito anteriormente, o método WLS tem a desvantagem de se ter de inverter uma matriz que pode ser muito grande dependendo do número de variáveis observadas.

Uma forma alternativa é calcular somente as variâncias assintóticas dos coeficientes estimados e utilizar a função (2.31), que é uma simplificação do modelo geral (2.27).

$$\sum_{g=1}^k \sum_{h=1}^k \frac{1}{w_{gh}} (s_{gh} - \sigma_{gh})^2, \quad (2.31)$$

sendo que w_{gh} é um estimativa da variância assintótica de s_{gh} .

A função (2.31) corresponde a utilizar uma matriz de pesos diagonal $\mathbf{W}^{-1}$.

Este método é considerado um meio termo entre ULS e WLS e não é correto avaliar o teste χ^2 .

Tanto as estimativas obtidas através do método WLS, quanto as estimativas obtidas pelo método DWLS, são não-eficientes.

2.9.7 Estimadores Elípticos

A fórmula geral de uma densidade elíptica é:

$$c|\mathbf{V}|^{-\frac{1}{2}}h\{(\mathbf{z}-\boldsymbol{\mu})'\mathbf{V}^{-1}(\mathbf{z}-\boldsymbol{\mu})\}, \quad (2.32)$$

sendo que:

- c é uma constante;
- h é uma função não-negativa;
- $\mathbf{z}$ é um vetor de dados com média $\boldsymbol{\mu}$ e matriz positiva-definida $\mathbf{V}$.

As distribuições elípticas incluem uma vasta gama de distribuições, possuem coeficiente de assimetria 0, mas podem ter curtose diferente da curtose da distribuição normal multivariada.

Neste caso a função a ser ajustada se resume a:

$$F_E = \frac{1}{2}(c+1)^{-1}\text{tr}[(\mathbf{S}-\boldsymbol{\Sigma}(\boldsymbol{\Theta}))\mathbf{V}^{-1}]^2 - c_1 [\text{tr}(\mathbf{S}-\boldsymbol{\Sigma}(\boldsymbol{\Theta}))\mathbf{V}^{-1}]^2, \quad (2.33)$$

sendo que:

- c é um parâmetro de curtose comum;
- $c_1 = \frac{c}{[4(c+1)^2+2(p+q)k(k+1)]}$;
- $\mathbf{V}$ é uma matriz de pesos.

O parâmetro de curtose comum pode ser estimado por:

$$\hat{c} = \frac{b_{2,(p+q)} - (p+q)(p+q+2)}{(p+q)(p+q+2)}, \quad (2.34)$$

sendo que $b_{2,(p+q)} = \left(\frac{1}{n}\right) \sum_{i=1}^n \{(\mathbf{Z}_i - \bar{\mathbf{Z}})'\mathbf{S}^{-1}(\mathbf{Z}_i - \bar{\mathbf{Z}})\}^2$.

O parâmetro de curtose comum também pode ser estimado por:

$$\hat{c} = \frac{1}{3(p+q)} \sum_{i=1}^{p+q} b_{2i}, \quad (2.35)$$

sendo que b_2 são as estimativas de curtose univariadas.

Escolhendo $\mathbf{V}$ como um estimador consistente de Σ , obtemos um estimador assintoticamente eficiente de Θ e $(n-1)F_E$ possui distribuição assintoticamente χ^2 . A escolha mais comuns de $\mathbf{V}$ é $\mathbf{S}$ (Bollen, 1989).

2.9.8 Covariâncias *versus* Correlações

Além de escolher qual método será utilizado para estimar os parâmetros do modelo de equações estruturais, é necessário determinar qual matriz de entrada será utilizada.

A matriz de covariâncias tem como vantagem fornecer comparações válidas entre diferentes populações, ou amostras, mas a interpretação dos coeficientes fica comprometida.

A matriz de correlações é mais indicada quando o objetivo do estudo é compreender as relações entre os constructos do modelo, mas não explicar a variância total de uma variável latente.

Todas as formas de estimação apresentadas nas seções anteriores permitem que qualquer matriz de entrada seja utilizada, entretanto deve-se ter cuidado na interpretação dos testes estatísticos quando se utiliza a matriz de correlação e quando as estimativas são obtidas pelos métodos ML ou GLS (Bollen, 1989).

As estimativas obtidas pelo método WLS também são válidas para a matriz de correlações, desde que a matriz de pesos reflita a covariância assintótica entre as correlações r_{gh} e r_{ij} .

A falta de dados pode ser um problema potencial na estimação das correlações ou covariâncias na medida em que acabam por diminuir o tamanho amostral. Hair *et al.* (1998) afirma que os algoritmos EM introduzem o menor viés nos modelos estimados, mas que as opções *pairwise* ou *listwise* também têm bom resultado quando a proporção de dados faltantes não é grande.

Como referência de tamanho amostral, Hair *et al.* (1998) recomenda 200 ou mais observações, com um aumento caso o modelo seja muito complexo, se existem suspeitas de má especificação do modelo, se os dados exigem características não-normais ou se algum procedimento alternativo, tal como WLS, for empregado.

2.10 Avaliação do Ajuste do Modelo

O próximo passo na modelagem de equações estruturais é a avaliação do ajuste do modelo estimado por alguma das técnicas apresentadas anteriormente.

É importante que antes de se fazer qualquer interpretação do modelo estimado, seja avaliado se os resultados obtidos são válidos.

Nesta etapa os coeficientes estimados e a força das relações inseridas no modelo são examinadas. Esta etapa pode ser dividida em 3 partes: (1) avaliação dos parâmetros estimados, (2) ajuste geral do modelo e (3) medidas alternativas de ajuste.

2.10.1 Avaliação dos Parâmetros Estimados

Embora muitas vezes possa passar despercebido pelo pesquisador, avaliar o sinal e a magnitude dos parâmetros estimados pelo modelo pode evitar maiores trabalhos caso o modelo ajustado não seja válido.

Dado o caráter confirmatório que a modelagem de equações estruturais possui, muitas vezes se tem de antemão uma idéia, mesmo que subjetiva, dos sinais e da magnitude dos parâmetros do modelo. Caso alguma estimativa se desvie muito do valor que se espera obter, pode ser que exista alguma hipótese defeituosa no modelo testado ou um modelo mal-especificado.

Existem várias formas de má especificação em um modelo de equações estruturais, dentre as quais podemos citar:

- Quando as estimativas são soluções impróprias, no sentido de que os valores observados são impossíveis de serem verificados na população sob estudo;
- Quando as variâncias estimadas são menores que zero ou não significantes, para qualquer um dos constructos do modelo;
- Quando a correlação entre duas variáveis está fora do intervalo $[-1, 1]$;
- Quando existem erros-padrão muito grandes, associados a qualquer um dos coeficientes estimados.

As soluções impróprias podem ser causadas por vários motivos. Primeiro, pode acontecer que a amostra analisada não reflita consistentemente a população sob estudo. Segundo, os dados possuem *outliers*, ou observações influentes, que distorcem a matriz analisada e, conseqüentemente, os parâmetros estimados. Terceiro, realmente existe uma má especificação do modelo; nestes casos é necessária uma reespecificação do modelo, que será discutida posteriormente, entretanto podemos adiantar que qualquer reespecificação depende muito mais do conhecimento específico do problema do que de qualquer informação obtida através de técnicas estatísticas.

Outra forma de avaliar os parâmetros estimados é através de testes estatísticos. Nos casos em que o modelo é estimado por ML, GLS ou WLS é possível realizar um teste t para os parâmetros individualmente.

A matriz de covariância das estimativas é derivada da inversa da matriz de informação e, desta forma, a diagonal principal da matriz contém as variâncias assintóticas das estimativas. A estatística do teste é:

$$t = \frac{\hat{\theta}_i - \theta_i}{s_{\hat{\theta}_i}}, \quad (2.36)$$

sendo que:

- $\hat{\theta}_i$ é o parâmetro estimado;

- θ_i é o valor hipotetizado;
- $s_{\hat{\theta}_i}$ é o erro-padrão assintótico do estimador.

Já a matriz de covariância assintótica dos parâmetros estimados é dada por:

$$\left(\frac{2}{n-1} \right) \left\{ E \left[\frac{\partial^2 F_{ML}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right] \right\}^{-1}. \quad (2.37)$$

Substituindo $\hat{\boldsymbol{\theta}}$ por $\boldsymbol{\theta}$ na equação (2.37), temos uma matriz de covariâncias assintóticas que estima as variâncias assintóticas dos parâmetros $\hat{\boldsymbol{\theta}}$ na diagonal principal e as covariâncias assintóticas fora da diagonal principal (Bollen, 1989).

Boomsma (1983) afirma que quando analisamos o modelo utilizando a matriz de correlação, obtemos estimativas imprecisas das variâncias e covariâncias dos parâmetros estimados. Lawley e Maxwell (1971), Jennrich e Thayer (1973) e Browne (1982) apresentam correções nos erros-padrão quando correlações são analisadas.

Caso a significância estatística não seja alcançada para algum dos parâmetros estimados, pode-se avaliar a exclusão desta relação e reestimar o modelo ou, nos casos de indicadores, aplicar algum tipo de transformação objetivando um melhor ajuste do modelo.

À medida que temos a matriz de covariância dos parâmetros estimados, é possível reproduzir a matriz de correlações assintóticas destas estimativas.

A fórmula para o cálculo das correlações assintóticas é:

$$\text{aCor}(\hat{\theta}_i, \hat{\theta}_j) = \frac{\text{aCov}(\hat{\theta}_i, \hat{\theta}_j)}{\sqrt{\text{aVar}(\hat{\theta}_i) \text{aVar}(\hat{\theta}_j)}}. \quad (2.38)$$

A interpretação desta matriz é análoga a da matriz de correlações dos coeficientes estimados em regressão. Correlações que são extremamente altas indicam que duas estimativas são altamente associadas, indicando que existe colinearidade no modelo.

Diretrizes indicam que correlações, em módulo, acima de 0,70 já são indícios de alta colinearidade.

Também é possível avaliar o ajuste de cada um dos indicadores do modelo através do **coeficiente de correlação múltipla** de cada variável. Ele é calculado por:

$$R_{x_i}^2 = 1 - \frac{\text{Var}(\zeta_i)}{\hat{\sigma}_{ii}}, \quad (2.39)$$

sendo que

- $\text{Var}(\zeta_i)$ é a variância do erro associado à variável indicadora x_i ;
- $\hat{\sigma}_{ii}$ é a variância da variável indicadora x_i predita pelo modelo.

De maneira geral, o objetivo é encontrar e utilizar medidas com alto coeficiente de correlação múltipla, não existindo valores de referência para esta medida.

2.10.2 Ajuste Geral do Modelo

O ajuste geral do modelo tem por objetivo verificar se a hipótese de estrutura de variância está respeitada, ou seja, as medidas de ajuste avaliam se a matriz de covariância ajustada pelo modelo, $\hat{\Sigma}$, é condizente com a matriz de covariância amostral, S .

As medidas de ajuste têm a vantagem de serem medidas que avaliam o ajuste do modelo como um todo, permitindo que sejam detectadas falhas no ajuste do modelo que não podem ser detectadas em testes de parâmetros estimados, entretanto elas têm as desvantagens de não poderem ser aplicadas em modelos exatamente identificados e de poderem diferir dos ajustes dos componentes do modelo.

Resíduos

A principal hipótese de um modelo de equações estruturais é $\Sigma = \Sigma(\Theta)$, neste caso, quando a hipótese é verdadeira, esperamos que $\Sigma - \hat{\Sigma} = \mathbf{0}$, ou seja, que a matriz de resíduos populacionais seja nula.

Quando encontramos algum resíduo populacional diferente de zero, isto significa que ele foi mal estimado pelo modelo.

O objetivo é que os resíduos estejam próximos de zero, indicando um bom ajuste do modelo, entretanto eles podem variar muito por causa de flutuações amostrais e da escala das variáveis.

Para contornar este tipo de problema, Jöreskog e Sorbom (1986) desenvolveram resíduos padronizados, que levam em consideração o tamanho da amostra e corrige as variáveis a fim de que todas fiquem na mesma escala. A fórmula para calcular os resíduos populacionais padronizados é:

$$e_{i,j} = \frac{(s_{ij} - \hat{\sigma}_{ij})}{\left[\frac{1}{n}(\hat{\sigma}_{ii}\hat{\sigma}_{jj} + \hat{\sigma}_{ij})\right]^{\frac{1}{2}}}. \quad (2.40)$$

Jöreskog e Sorbom (1986) indicam que se existirem muitos resíduos padronizados maiores, em módulo, que 2 ou 3, há indícios de que o modelo foi mal especificado.

A principal diferença entre os resíduos populacionais padronizados e a matriz de resíduos obtida quando o modelo é ajustado utilizando a matriz de correlações, é que a primeira pondera a medida pelo tamanho amostral e a segunda somente padroniza a influência da escala das variáveis.

Raiz do Quadrado Médio Residual

A raiz do quadrado médio residual (RMR, do inglês *Root Mean-square Residual*) é uma medida utilizada para quantificar as diferenças entre $\hat{\Sigma}$ e S . Ela é calculada por:

$$RMR = \left[2 \sum_{i=1}^k \sum_{j=1}^i \frac{(s_{ij} - \hat{\sigma}_{ij})^2}{(p+q)(p+q+1)} \right]^{\frac{1}{2}}. \quad (2.41)$$

Valores próximos de zero indicam bom ajuste.

Da mesma forma que os resíduos padronizados, esta medida pode ser afetada pela escala das variáveis. Desta forma é aconselhável, nos casos em que existe este tipo de problema, utilizar os resíduos da matriz de correlação.

Teste χ^2

O teste χ^2 é uma forma de avaliarmos, estatisticamente, a hipótese $\Sigma = \Sigma(\Theta)$. Como já dito anteriormente, $(n - 1)F \sim \chi^2 \left(\frac{1}{2}(p + q)(p + q + 1) - t \right)$ quando F é estimada por ML, GLS ou WLS.

Neste teste, o objetivo é não rejeitar a hipótese nula. Geralmente são utilizados níveis de significância mais altos, como 0,05 ou 0,10, para se avaliar o teste χ^2 .

É importante interpretar com cautela os resultados deste teste, pois a estatística χ^2 pode ser afetada por vários fatores. O teste χ^2 é sensível a desvios severos de normalidade, mais especificamente quando existem grandes assimetrias no conjunto de dados.

Um outro fator que pode distorcer a estatística do teste é o tamanho amostral; uma amostra muito grande pode aumentar a confiança de que a matriz residual é zero, entretanto a significância substantiva da diferença pode levar a rejeições de diferenças que são insignificantes. De outro lado, um tamanho amostral pequeno pode reduzir o poder do teste em detectar diferenças significativas.

2.10.3 Medidas de Ajuste Alternativas

Em modelagem de equações estruturais não existe uma única medida que permita dizer que o ajuste é bom, ou não. Durante as últimas décadas muitos pesquisadores vêm desenvolvendo medidas que visam avaliar o ajuste do modelo estimado.

Estas medidas são vistas sob 3 perspectivas:

- Ajuste global;
- Ajuste comparativo a um modelo *baseline*²;
- Parcimônia do modelo.

Dado o grande número de medidas existentes (e ao consenso de que nenhuma delas, individualmente, tem o poder de avaliar o ajuste de um modelo), é recomendado avaliar todas elas, compreendendo o que cada uma está medindo e tendo em mente que elas são medidas descritivas.

²Um modelo *baseline* é um modelo que possui um ajuste ruim, utilizado para comparação com o modelo sob estudo. Ele pode variar de acordo com os objetivos do estudo. Geralmente ele supõe que existe uma variável latente por indicador, erros e variáveis latentes não correlacionados e parâmetros ligando os indicadores às variáveis latentes todos iguais a 1.

Leitores interessados neste assunto são encorajados a ler [Melhado \(2004\)](#) e [Bollen \(1989\)](#) para uma melhor compreensão de todas as medidas de ajuste alternativas que serão apresentadas.

Discrepância Mínima da Amostra

A discrepância mínima da amostra (CMIN, do inglês *Minimum Discrepancy*) é calculada por:

$$\hat{c} = n\hat{F}, \quad (2.42)$$

sendo que $\hat{F}$ é o valor da função de discrepância calculada para o modelo estudado.

Não existem valores de referência para esta medida. Ela é calculada para o modelo sob estudo, para o modelo *baseline* e para o modelo saturado³. Os valores resultantes são comparados entre si. Também pode ser utilizada para comparar modelos concorrentes.

Índice de Ajuste Normalizado

O índice de ajuste normalizado (NFI, do inglês *Normalized Fit Indices*) foi proposto por [Bentler e Bonett \(1980\)](#) e é calculado por:

$$\Delta_1 = \frac{F_b - F_m}{F_b} = \frac{\chi_b^2 - \chi_m^2}{\chi_b^2}, \quad (2.43)$$

sendo que:

- F_b é o valor da função de discrepância do modelo *baseline*;
- F_m é o valor da função de discrepância do modelo sob estudo.

Ela tem por objetivo medir o quanto a função de discrepância reduz quando passamos do modelo *baseline* para o modelo estudado.

Quanto mais próximo de 1, melhor é o ajuste do modelo. [Bentler e Bonett \(1980\)](#) afirmam que valores acima de 0,90 já indicam bom ajuste do modelo.

Índice de Ajuste Corrigido

O índice de ajuste corrigido (IFI, do inglês *Incremental Fit Indices*) foi proposto por [Bollen \(1989\)](#) e tem por objetivo diminuir a influência do tamanho amostral e da adição de parâmetros no modelo que podem distorcer o índice Δ_1 . Ele é calculado por:

$$\Delta_2 = \frac{F_b - F_m}{F_b - [gl_m/(n-1)]} = \frac{\chi_b^2 - \chi_m^2}{\chi_b^2 - gl_m}, \quad (2.44)$$

³Um modelo saturado é tal que não existem restrições sob os momentos populacionais, ou seja, é um modelo em que o número de equações é igual ao número de parâmetros distintos e, desta forma, sempre se ajusta perfeitamente ao conjunto de dados.

sendo que gl_m são os graus de liberdade do modelo estudado.

Ao contrário de Δ_1 , este índice não varia necessariamente entre 0 e 1. Valores muito acima de 1 indicam superajuste do modelo.

Podem ser encontradas grandes diferenças entre Δ_1 e Δ_2 , principalmente com tamanhos amostrais pequenos ou moderados; a diferença diminui à medida que o tamanho da amostra aumenta.

Quando o modelo é válido, Δ_2 tende a 1.

Índice de Ajuste Relativo

O índice de ajuste relativo (RFI, do inglês *Relative Fit Indices*) foi proposto por [Bollen \(1989\)](#) e é calculado por:

$$\rho_1 = \frac{(F_b/gl_b) - (F_m/gl_m)}{(F_b/gl_b)} = \frac{(\chi_b^2/gl_b) - (\chi_m^2/gl_m)}{(\chi_b^2/gl_b)}. \quad (2.45)$$

Este índice tem por objetivo medir a redução da função de discrepância quando passamos do modelo *baseline* para o modelo sob estudo, levando em consideração os graus de liberdade.

Ele varia de 0 a 1, sendo que valores próximos de 1 indicam bom ajuste do modelo.

Índice de Tucker-Lewis

O índice de Tucker-Lewis (TLI, do inglês *Tucker-Lewis Indices*) foi proposto por [Tucker e Lewis \(1973\)](#) e por [Bentler e Bonett \(1980\)](#). Ele é calculado por:

$$\rho_2 = \frac{(F_b/gl_b) - (F_m/gl_m)}{(F_b/gl_b) - [1/(n-1)]} = \frac{(\chi_b^2/gl_b) - (\chi_m^2/gl_m)}{(\chi_b^2/gl_b) - 1}. \quad (2.46)$$

Este índice tem o mesmo objetivo de que ρ_1 , só que ele é menos dependente do tamanho amostral.

Valores muito maiores que 1 indicam superajustamento, ao passo que valores muito menores que 1 indicam má especificação do modelo.

Podem ser encontradas diferenças grandes entre ρ_1 e ρ_2 . Entretanto é esperado encontrar pequenas diferenças à medida que o tamanho amostral aumenta.

O valor de ρ_2 converge para 1 quando o modelo é válido, independentemente do tamanho amostral. Bentler e Bonett (1980) afirmam que valores acima de 0,90 já indicam bom ajuste.

Índice de Qualidade do Ajuste e Índice de Qualidade do Ajuste Corrigido

O índice de qualidade do ajuste (GFI, do inglês *Goodness of Fit Index*) e o índice de qualidade do ajuste corrigido (AGFI, do inglês *Adjusted Goodness of Fit Index*) são calculados para cada uma das funções de discrepância utilizadas na estimação do modelo.

Jöreskog e Sorbom (1986) propuseram medidas GFI e AGFI para modelos ajustados por ML e ULS, que são calculadas por:

$$GFI_{ML} = 1 - \frac{\text{tr} \left[\left(\hat{\Sigma}^{-1} \mathbf{S} - \mathbf{I} \right)^2 \right]}{\text{tr} \left[\left(\hat{\Sigma}^{-1} \mathbf{S} \right)^2 \right]}; \quad (2.47)$$

$$AGFI_{ML} = 1 - \left[\frac{(p+q)(p+q+1)}{2gl_m} \right] [1 - GFI_{ML}]; \quad (2.48)$$

$$GFI_{ULS} = 1 - \frac{\text{tr} \left[\left(\mathbf{S} - \hat{\Sigma} \right)^2 \right]}{\text{tr} [\mathbf{S}^2]}; \quad (2.49)$$

$$AGFI_{ULS} = 1 - \left[\frac{(p+q)(p+q+1)}{2gl_m} \right] [1 - GFI_{ULS}]. \quad (2.50)$$

Tanaka e Huba (1985) propuseram estas medidas para modelos ajustados por GLS, que são calculadas por:

$$GFI_{GLS} = 1 - \frac{\text{tr} \left[\left(\mathbf{I} - \hat{\Sigma} \mathbf{S}^{-1} \right)^2 \right]}{(p+q)}; \quad (2.51)$$

$$AGFI_{GLS} = 1 - \left[\frac{(p+q)(p+q+1)}{2gl_m} \right] [1 - GFI_{GLS}]. \quad (2.52)$$

O índice GFI mede a quantidade relativa de covariâncias em $\mathbf{S}$ que são preditas por $\hat{\Sigma}$ e a medida AGFI ajusta-se aos graus de liberdade do modelo e a quantidade de variáveis observadas.

Estas medidas podem ser negativas e são iguais a 1 quando $\mathbf{S} = \hat{\Sigma}$.

Índice de Qualidade do Ajuste de Parcimônia

O índice de qualidade do ajuste de parcimônia (PGFI, do inglês *Parsimony Goodness of Fit Index*) é calculado por:

$$PGFI = GFI \frac{2gl_m}{(p+q)(p+q+1)}. \quad (2.53)$$

As medidas de parcimônia em geral servem para comparar modelos com diferentes números de parâmetros ou para determinar o impacto da adição de parâmetros ao modelo.

Ele varia de 0 a 1, dificilmente atinge valores muito altos e não existem valores de referência.

Índice de Ajuste Normalizado de Parcimônia

O índice de ajuste normalizado de parcimônia (PNFI, do inglês *Parsimony Normalized Fit Indices*) é calculado por:

$$PNFI = NFI \frac{gl_m}{gl_b}. \quad (2.54)$$

Esta medida é utilizada principalmente para comparar modelos com diferentes graus de liberdade. Diferenças entre 0,06 a 0,09 entre as medidas PNFI podem significar diferenças significativas entre os modelos comparados.

Quanto maior o valor do índice, melhor é o ajuste, valores mínimos aceitáveis estão entre 0,50 e 0,60.

Índice de Ajuste Comparativo

O índice de ajuste comparativo (CFI, do inglês *comparative Fit Indices*) é calculado por:

$$CFI = 1 - \frac{\max(\hat{c}_m - gl_m, 0)}{\max(\hat{c}_b - gl_b, 0)}. \quad (2.55)$$

Este índice varia entre 0 e 1. Valores aceitáveis estão acima de 0,90.

Raiz do Erro Quadrático Médio de Aproximação

A raiz do erro quadrático médio de aproximação (RMSEA, do inglês *Root Mean Square Error of Approximation*) é calculada por:

$$RMSEA = \sqrt{\frac{\hat{F}_\theta}{gl_m}}, \quad (2.56)$$

sendo que $\hat{F}_\theta = \max\left(\frac{\hat{c} - gl_m}{n}, 0\right)$.

Melhado (2004) afirma que esta medida tem sido reconhecida como um dos melhores critérios informativos para ser utilizado em modelagem de equações estruturais.

Valores abaixo de 0,05 indicam bom ajuste, valores entre 0,05 e 0,10 indicam ajuste razoável e valores acima de 0,10 indicam ajuste pobre.

χ^2 Relativo

O χ^2 relativo, χ_r^2 , é calculado por:

$$\chi_r^2 = \frac{\chi^2}{gl_m}. \quad (2.57)$$

Não existe consenso sobre quais valores indicam um bom ajuste, entretanto as recomendações são razões de 3,2 ou menores (Bollen, 1989).

χ^2 Padronizado

O χ^2 padronizado, χ_p^2 , é calculado por:

$$\chi_r^2 = \frac{\chi^2 - gl_m}{\sqrt{2gl_m}}. \quad (2.58)$$

Esta medida possui as mesmas características do χ_r^2 . Não existe consenso sobre quais valores indicam bom ajuste.

Critério de Informação de Akaike

O critério de informação de Akaike (AIC, do inglês *Akaike Information Criterion*) é calculado por:

$$AIC = \hat{c} + 2t. \quad (2.59)$$

Esta medida é utilizada para comparação de modelos. Recomenda-se escolher o modelo que apresentar o menor valor AIC.

Critério de Informação de Bayes

O critério de informação de Bayes (BIC, do inglês *Bayes Information Criterion*) é calculado por:

$$BIC = \hat{c} + t \ln(n(p+q)). \quad (2.60)$$

Ela tende a escolher modelos mais parcimoniosos e recomenda-se escolher o modelo que apresentar o menor valor BIC.

Critério de Browni-Cudeck

O critério de Browni-Cudeck (BCC, do inglês *Browni-Cudeck Criterion*) é calculado por:

$$BCC = \hat{c} + 2t \frac{b \left(\frac{(p+q)(p+q+3)}{n-p-q-2} \right)}{(p+q)(p+q+3)}. \quad (2.61)$$

Esta medida é utilizada para comparar modelos. Recomenda-se escolher o modelo que apresentar o menor valor BCC.

Critério da Informação de Akaike Consistente

O critério da informação de Akaike consistente (CAIC, do inglês *Consistent Akaike Information Criterion*) é calculado por:

$$CAIC = \hat{c} + t(\ln(n) + 1). \quad (2.62)$$

Da mesma forma que AIC, BIC e BCC, esta medida é utilizada para comparar modelos. Recomenda-se escolher o modelo que apresentar o menor valor CAIC.

2.11 Reespecificação do Modelo

Pode acontecer que na avaliação do modelo encontremos evidências de que ele não foi bem ajustado. Nestes casos é necessária uma reespecificação do modelo com o intuito de aumentar o ajuste ou a sua correspondência com a teoria.

Sanches (2005) afirma que o procedimento mais correto neste tipo de situação é coletar os dados novamente e reestimar o modelo, entretanto nem sempre este tipo de procedimento é possível.

Uma forma potencial e rica de determinar a reespecificação de um modelo é utilizar o conhecimento teórico do pesquisador (Bollen, 1989).

As medidas apresentadas anteriormente servem somente para comparar modelos, mas elas são descritivas e não permitem a realização de testes estatísticos. Para contornar este problema, várias técnicas foram desenvolvidas para ajudar o pesquisador no processo de reespecificação do modelo, dentre elas podemos citar: (1) o teste da razão de verossimilhanças, (2) o teste do multiplicador de Lagrange e (3) o teste de Wald.

Estes testes são equivalentes entre si e levam em consideração as diferenças entre as estimativas χ^2 do modelo proposto e do modelo reespecificado. Estes testes têm a pressuposição de que são aplicados para modelos aninhados⁴ e têm caráter muito mais exploratório do que inferencial.

De uma forma geral, é necessário tomar cuidado com qualquer procedimento descrito nesta seção, uma vez que reespecificações aumentam a chance de se obter um modelo que seja significativo somente pelo acaso (Sanches, 2005). É recomendado que qualquer reespecificação só seja realizada caso exista embasamento teórico.

2.11.1 Teste da Razão de Verossimilhanças

O teste da razão de verossimilhanças pode ser utilizado para determinar qual dos modelos reespecificados tem o maior decréscimo na estatística χ^2 .

Uma forma é estimar todos os possíveis modelos identificáveis que eliminam uma restrição, computar a função de discrepância destes modelos e calcular o teste da razão de verossimilhanças. Caso nenhum modelo apresente diminuição significativa, pode-se refazer o mesmo procedimento eliminando duas restrições e assim sucessivamente.

⁴Um modelo que necessita que uma função de seus parâmetros livres seja igual a outros parâmetros livres (ou iguais a uma constante) são aninhados com um modelo idêntico, só que sem estas restrições. O caso mais simples ocorre quando os parâmetros livres de um modelo são um subconjunto dos parâmetros livres de um segundo modelo.

A estatística do teste pode ser escrita como:

$$\chi^2 = (n - 1)(F_m - F_r), \quad (2.63)$$

sendo que:

- F_m é o valor da função de discrepância avaliada para o modelo proposto;
- F_r é o valor da função de discrepância avaliada para o modelo reespecificado.

Esta estatística possui distribuição χ^2 com os graus de liberdade igual à diferença entre os graus de liberdade do modelo proposto e do modelo reespecificado.

Tal como o teste χ^2 usual, este teste tem o poder de detectar modelos restritivos falsos em grandes amostras e tem a desvantagem de ter de ser computado para uma grande quantidade de modelos.

2.11.2 Teste do Multiplicador de Lagrange

O teste do multiplicador de Lagrange compara o modelo proposto com um modelo menos restritivo, sem a necessidade de ter de determinar as estimativas do modelo menos restritivo.

A estatística do teste é baseada no log irrestrito da função de verossimilhança e pode ser escrita como:

$$LM = \frac{n - 1}{2} \left(\frac{\partial F_{ML}}{\partial \boldsymbol{\theta}} \right)' \left[E \left(\frac{\partial^2 F_{ML}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right) \right]^{-1} \left(\frac{\partial F_{ML}}{\partial \boldsymbol{\theta}} \right), \quad (2.64)$$

em que $\boldsymbol{\theta}$ é avaliado por $\hat{\boldsymbol{\theta}}_m$.

A distribuição limite da estatística LM é χ^2 com graus de liberdade igual à diferença entre os graus de liberdade do modelo com e sem restrições.

Para facilitar a aplicação deste teste, foi proposto o teste do multiplicador de Lagrange univariado. Desta forma, a estatística do teste se reduz a:

$$LM = \left[\frac{\partial \log L(\boldsymbol{\theta})}{\partial \theta_i} \right] \left[\mathbf{I}^{-1} \left(\hat{\boldsymbol{\theta}}_m \right) \right]_{ii}, \quad (2.65)$$

sendo que $\mathbf{I}^{-1}$ é a inversa da matriz de informação.

No teste univariado, a única diferença entre o modelo restrito e irrestrito está na limitação em θ_i . Desta forma a estatística acima possui distribuição $\chi^2(1)$ e é utilizada para testar qual a melhora no ajuste quando o parâmetro θ_i é livre.

O teste do multiplicador de Lagrange é equivalente ao índice de modificação.

Ele é utilizado para determinar quais das restrições que se forem consideradas livres, no modelo proposto, levam à maior redução na estimativa χ^2 do ajuste global do modelo. Este procedimento é realizado até que o modelo tenha um ajuste adequado e ele pode ser considerado como um processo *stepwise* de regressão.

2.11.3 Teste de Wald

O teste de Wald avalia quais restrições podem ser incorporadas ao modelo proposto que levam ao menor aumento da estatística χ^2 do modelo ajustado.

Suponha que as limitações colocadas no modelo irrestrito podem ser sumarizadas em um vetor $r(\boldsymbol{\theta})$, em que a dimensão deste vetor é menor que a dimensão de $\boldsymbol{\theta}$. O vetor $r(\boldsymbol{\theta})$ pode ser encarado como uma função destes parâmetros.

O teste de Wald determina a extensão do afastamento de $\hat{\boldsymbol{\theta}}_m$ às restrições impostas pelo modelo aninhado.

A estatística do teste é:

$$W = \left[r(\hat{\boldsymbol{\theta}}_m) \right]' \left\{ \left[\frac{\partial r(\hat{\boldsymbol{\theta}}_m)}{\partial \hat{\boldsymbol{\theta}}_m} \right]' \text{aCov}(\hat{\boldsymbol{\theta}}_m) \left[\frac{\partial r(\hat{\boldsymbol{\theta}}_m)}{\partial \hat{\boldsymbol{\theta}}_m} \right] \right\}^{-1} \left[r(\hat{\boldsymbol{\theta}}_m) \right], \quad (2.66)$$

sendo que:

- $r(\hat{\boldsymbol{\theta}}_m)$ é $r(\boldsymbol{\theta})$ avaliada em $\hat{\boldsymbol{\theta}}_m$;
- $\text{aCov}(\hat{\boldsymbol{\theta}}_m)$ é uma estimativa da matriz de covariância assintótica de $\hat{\boldsymbol{\theta}}_m$.

O termo entre chaves é uma estimativa da matriz de covariância assintótica de $r(\hat{\boldsymbol{\theta}}_m)$.

A estatística W tem distribuição χ^2 com graus de liberdade igual ao número de limitações em $r(\boldsymbol{\theta})$ quando o modelo restritivo é válido.

2.11.4 Limitações dos Testes Estatísticos

Embora os testes apresentados anteriormente possam ser convidativos à primeira vista, é necessário apontar suas limitações a fim de que seus resultados exploratórios sejam avaliados antes de se determinar uma reespecificação no modelo proposto.

Primeiro, a ordem em que os parâmetros são considerados livres ou restritos pode afetar a significância dos parâmetros restantes pelo fato de que as estimativas dos parâmetros são correlacionadas entre si (Saris *et al.*, 1979).

Segundo, a probabilidade associada aos testes do multiplicador de Lagrange e Wald são provavelmente não acuradas quando se utiliza procedimentos do tipo *stepwise*.

Terceiro, estes testes avaliam apenas mudanças na estatística χ^2 do ajuste global do modelo, mas não o tamanho das mudanças nos parâmetros estimados. Desta forma, uma estratégia prudente é examinar os parâmetros estimados em conjunto com a estatística χ^2 do ajuste do modelo (Bollen, 1989).

Quarto, nos casos em que um modelo não é válido, o teste do multiplicador de Lagrange possui versões viciadas que podem ser obtidas pelas mesmas restrições aplicadas no modelo correto (Bollen, 1989).

Quinto, os testes do multiplicador de Lagrange e de Wald não têm poder de detectar mudanças complexas que podem ser necessárias em modelos de equações estruturais.

A última limitação é que estes testes podem introduzir, ou remover, parâmetros de um modelo sem nenhuma interpretação clara.

2.12 Resumo do Capítulo

Ao longo deste capítulo fizemos um resumo de todos os procedimentos envolvidos na modelagem de equações estruturais.

Inicialmente foram apresentadas as equações do modelo de mensuração e do modelo estrutural, o diagrama de caminhos e a matriz de covariância implícita.

Logo após comentamos sobre como verificar validade e confiabilidade dos indicadores do modelo, além das relações entre variáveis observadas e constructos.

Foram explicitadas as regras para verificar se o modelo era identificável, os principais procedimentos para estimação dos parâmetros e regras para avaliar o ajuste do modelo.

Por fim apresentamos algumas técnicas de reespecificação de modelos para os casos em que não se consegue um bom ajuste para o modelo proposto inicialmente.

Buscou-se fazer um resumo das principais recomendações feitas sobre cada um dos tópicos discutidos neste capítulo. Entretanto o leitor é encorajado a avaliar cada uma das decisões envolvidas na análise de equações estruturais levando sempre em consideração a teoria que está sendo estudada através do modelo.

Capítulo 3

Incorporação de Variáveis Categóricas

Uma das suposições do modelo de equações estruturais é a de que as variáveis observadas são contínuas. Entretanto, podem existir situações em que se faz necessária a incorporação de indicadores categóricos que sejam relevantes para se estimar os constructos do modelo.

O pesquisador pode se sentir tentado a dar números para as categorias da variável ordinal, estimar a matriz de covariância, ou correlação, considerando a variável categórica como contínua e estimar o modelo de equações estruturais.

O problema é que fazendo este tipo de procedimento perde-se a confiabilidade nos resultados e invalida-se qualquer interpretação das estimativas obtidas.

Diante desta situação, os objetivos deste capítulo são: (1) examinar quais suposições são violadas com a incorporação de indicadores categóricos, avaliando suas consequências, e (2) apresentar alguns dos procedimentos corretivos.

Este texto se restringe à discussão de variáveis latentes contínuas com indicadores dicotômicos ou ordinais e sua extensão a modelos de equações estruturais; para leitores interessados em relações mais complexas, recomenda-se a leitura de Lazarsfeld e Henry (1968), Goodman (1972), Bishop *et al.* (1975), Amemiya (1981) e Maddala (1983).

3.1 Suposições Violadas e suas Consequências

Seja $\mathbf{y}^*$ um vetor de indicadores contínuos de $\boldsymbol{\eta}$ que se ajusta ao modelo de mensuração usual, ou seja:

$$\mathbf{y}^* = \mathbf{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\epsilon}, \quad (3.1)$$

sendo que:

- $E(\boldsymbol{\epsilon}) = \mathbf{0}$;
- $\boldsymbol{\epsilon}$ é não correlacionado com $\boldsymbol{\eta}$;
- $\boldsymbol{\eta}$ é da forma ajustada.

Suponha agora que não observamos $\mathbf{y}^*$, mas um vetor $\mathbf{y}$ com algumas, ou todas, as versões categóricas das variáveis presentes em $\mathbf{y}^*$. A variável $\mathbf{y}_1^*$, por exemplo, pode ser normalmente distribuída, enquanto que $\mathbf{y}_1$ é uma variável ordinal de quatro categorias.

Desta forma $\mathbf{y}^* \neq \mathbf{y}$ e, conseqüentemente:

$$\mathbf{y} \neq \Lambda_{\mathbf{y}}\boldsymbol{\eta} + \boldsymbol{\epsilon}. \quad (3.2)$$

A primeira consequência da desigualdade (3.2) é que o modelo de mensuração não é válido para $\mathbf{y}$. A segunda consequência é que a distribuição das variáveis ordinais geralmente difere da distribuição da variável latente contínua. Para este tipo de situações é possível estimar os parâmetros do modelo utilizando estimadores assintoticamente livres de distribuição (ADF, do inglês *Asymptotic Distribution Free*), entretanto as variáveis ordinais podem criar erros heterocedásticos e os estimadores ADF supõem homocedasticidade (Bollen, 1989).

A terceira consequência é que a $\text{aCov}(s_{ij}, s_{gh})$ difere da $\text{aCov}(s_{ij}^*, s_{gh}^*)$.

A última, e mais séria consequência é que a hipótese de estrutura de covariância é violada. Em um modelo de equações estruturais assume-se que para $\boldsymbol{\Sigma}^*$, temos $\boldsymbol{\Sigma}^* = \boldsymbol{\Sigma}(\boldsymbol{\theta})$. Entretanto, para $\boldsymbol{\Sigma}$, $\boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}^*$ e, desta forma, $\boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}(\boldsymbol{\theta})$ - o que significa que a hipótese de estrutura de covariância é válida para os indicadores latentes contínuos, mas não necessariamente para os indicadores ordinais.

Olsson (1979), Boomsma (1983), Johnson e Creech (1983), Muthén e Kaplan (1985), Babakus *et al.* (1987) e outros vêm, ao longo dos anos, realizando estudos simulados com o objetivo de testar a robustez dos ajustes quando se considera uma variável ordinal como contínua.

De uma maneira geral estes estudos indicam que quando existem curtose e assimetria excessiva nas versões ordinais dos indicadores latentes contínuos, os testes χ^2 , do ajuste global do modelo, e t , de significância estatística dos parâmetros, obtidas por F_{ML} ou F_{GLS} são adversamente afetadas; a estatística χ^2 parece ser mais afetada pelos desvios de curtose e assimetria do que, necessariamente, pelo número de categorias da variável ordinal; quando a estrutura de covariância é válida para os indicadores contínuos latentes, mas não para os indicadores ordinais observados, os coeficientes padronizados estimados parecem ser atenuados e esta atenuação é inversamente proporcional ao número de categorias; e quando existem poucas categorias com alta carga fatorial inicial, as funções de discrepância produzem erros de mensuração correlacionados com maior frequência do que quando se utiliza os indicadores contínuos. Estes resultados devem ser encarados como preliminares pelo fato de ainda não terem sido generalizados a um conjunto mais geral de modelos (Bollen, 1989).

3.2 Procedimentos Corretivos

Na seção anterior, descrevemos as consequências por se utilizar indicadores ordinais em substituição a indicadores contínuos.

Resumidamente temos que:

$$\mathbf{y} \neq \Lambda_{\mathbf{y}}\boldsymbol{\eta} + \boldsymbol{\epsilon}; \quad (3.3)$$

$$\mathbf{x} \neq \Lambda_{\mathbf{x}}\boldsymbol{\xi} + \boldsymbol{\delta}; \quad (3.4)$$

$$\Sigma \neq \Sigma(\boldsymbol{\theta}); \quad (3.5)$$

$$\text{aCov}(s_{ij}, s_{gh}) \neq \text{aCov}(s_{ij}^*, s_{gh}^*). \quad (3.6)$$

Os procedimentos corretivos que vamos descrever a partir de agora têm como objetivo contornar estes problemas.

Para começar, nós discutiremos (3.3) e (3.4). Neste caso uma relação linear associando $\mathbf{y}(\mathbf{x})$ com $\boldsymbol{\eta}(\boldsymbol{\xi})$ não é a escolha mais indicada. É necessária uma função não-linear relacionando os indicadores ordinais observados aos seus respectivos indicadores contínuos latentes.

Para indicadores ordinais podemos determinar os **limiares discretizantes**¹. Eles podem ser estimados se conhecemos as distribuições dos indicadores latentes contínuos e a proporção amostral de casos em cada categoria. A suposição mais comum é que $\mathbf{y}^*$ e $\mathbf{x}^*$ são normalmente distribuídos com a média 0 e variância 1. Desta forma as estimativas dos limiares são obtidas por:

$$a_i = \Phi^{-1} \left(\sum_{k=1}^i \frac{n_k}{n} \right), \quad i = 1, \dots, c-1; \quad (3.7)$$

sendo que:

- $\Phi^{-1}(z)$ é a inversa da distribuição acumulada normal padrão;
- n_k é número de observações na k -ésima categoria;
- n é o número total de observações;
- c é o número de categorias da variável ordinal.

Em suma, para corrigir o problema de que $\mathbf{y} \neq \Lambda_{\mathbf{y}}\boldsymbol{\eta} + \boldsymbol{\epsilon}$ e $\mathbf{x} \neq \Lambda_{\mathbf{x}}\boldsymbol{\xi} + \boldsymbol{\delta}$, quando alguns indicadores são ordinais, nós adicionamos um modelo limiar para o modelo de mensuração usual e este modelo limiar relaciona as variáveis ordinais às suas contrapartes latentes.

A consequência de que a hipótese de estrutura de covariância usualmente não é verdadeira para a matriz de covariância das variáveis ordinais observadas ($\Sigma \neq \Sigma(\boldsymbol{\theta})$), mas é verdadeira para os indicadores latentes contínuos ($\Sigma^* = \Sigma(\boldsymbol{\theta})$), pode ser contornada encontrando um estimador consistente para Σ^* .

Para estimar Σ^* , que possui indicadores ordinais, nós podemos assumir uma distribuição para os indicadores contínuos latentes, geralmente multinormal, e calcular as relações entre as variáveis através de correlações especiais.

¹Valores para os quais uma variável sai de uma categoria e vai para a outra.

Se ambas as variáveis são contínuas, a correlação de Pearson é aplicável; se ambas as variáveis são ordinais, utiliza-se a correlação policórica (se as variáveis ordinais são dicotômicas utiliza-se a correlação tetracórica); se uma variável é contínua e a outra é ordinal utiliza-se a correlação poliserial (se a variável ordinal for dicotômica utiliza-se a correlação biserial).

A matriz de correlações especiais é um estimador consistente de Σ^* , o que permite testar a hipótese de que $(\Sigma^* = \Sigma(\theta))$.

A utilização de $\hat{\Sigma}^*$ leva a estimadores consistentes de θ com qualquer uma das funções de ajuste existentes. Entretanto, os erros-padrão, o teste t , o teste χ^2 e outros testes de significância estatística não são corretos para F_{ML} , F_{GLS} ou F_{ULS} .

3.3 Resumo do Capítulo

Este capítulo teve por objetivo avaliar as suposições violadas quando consideramos indicadores ordinais ao invés de indicadores contínuos e suas consequências na estimação do modelo e na validade das estimativas obtidas.

Após o levantamento das suposições violadas, apresentamos os procedimentos necessários para que fosse possível incorporar indicadores ordinais na modelagem de equações estruturais.

Resumidamente, a estimação da correlação entre os indicadores utilizando as correlações especiais (correlação tetracórica, correlação policórica, correlação biserial e correlação poliserial) e a suposição adicional de um modelo limiar relacionando os indicadores ordinais às suas contrapartes latentes, permitem que o modelo de equações estruturais seja corretamente estimado.

Capítulo 4

Correlações Especiais

As correlações biserial e poliserial são medidas de associação bivariada utilizadas quando uma das variáveis é ordinal e a outra é intervalar; já as correlações tetracórica e policórica são medidas utilizadas quando ambas as variáveis são ordinais.

A origem destes coeficientes é atribuído a [Pearson \(1900\)](#). Ele observou que em alguns casos era muito complicado medir precisamente o valor de uma variável em uma escala quantitativa, mas extremamente fácil de associar estas variáveis a categorias ordenadas. Para estes tipos de variáveis, Pearson assumiu que elas eram funções de distribuições normais latentes e, em seguida, estimou a correlação produto-momento entre as variáveis normais.

A correlação tetracórica foi desenvolvida como uma estimativa do parâmetro de correlação de uma distribuição normal bivariada oriunda de dados em forma de uma tabela 2×2 e a correlação policórica como uma generalização para uma tabela $r \times s$.

No desenvolvimento da correlação biserial, o parâmetro da distribuição normal bivariada é estimado com dados em que uma das variáveis normais é da forma dicotômica e a outra está em sua forma contínua original. A correlação poliserial é uma generalização com uma variável dividida em r categorias ordenadas, em vez de apenas 2.

O objetivo deste capítulo é apresentar as formas de cálculo das estimativas das correlações especiais, apresentando alguns exemplos, avaliando a vantagem de utilizá-las em substituição à correlação de Pearson e encontrando indícios de que elas são robustas quando existem alguns desvios de normalidade.

4.1 Exemplo Intuitivo

Suponha que dois psiquiatras estão fazendo um diagnóstico de presença-ausência de depressão em um grupo de pacientes. Embora o resultado deste teste seja dicotômico, temos que o nível de depressão é continuamente distribuído na população.

A Figura 4.1 explica como é feita a discretização de uma variável latente, ou seja, para o diagnóstico de um caso em particular, um psiquiatra considera o nível de depressão (Y), relativo a um limiar (t); se o nível julgado é maior do que o limiar, um diagnóstico positivo

Figura 4.1: Variável Contínua Latente (Y : Severidade da Depressão) e limiar discretizante.

é feito, caso contrário é dado um diagnóstico negativo.

A Figura 4.2 retrata a situação do julgamento de 2 psiquiatras.

Figura 4.2: Distribuição conjunta da Severidade da Depressão para 2 psiquiatras (Y_1 e Y_2) e dois limiares (t_1 e t_2).

Os valores a , b , c e d , na Tabela 4.1, representam a proporção de casos contidos em cada uma das regiões definidas pelos limiares. Estas proporções podem ser resumidas em uma tabela de contingência.

Uma vez que conhecemos as proporções a , b , c e d , a Tabela 4.1 é uma simples forma de estimarmos o modelo representado pela Figura 4.2. Especificamente, necessitamos estimar os limiares discretizantes t_1 e t_2 e um terceiro parâmetro, ρ^* , que determina a inclinação da elipse, sendo que ρ^* é o coeficiente de correlação tetracórica.

A correlação tetracórica pode ser vista como a correlação entre o julgamento da severidade da depressão entre os dois psiquiatras.

O princípio de estimação é bastante simples; basicamente o programa de computador testa várias combinações de ρ^* , t_1 e t_2 até que sejam encontrados valores para os quais

Tabela 4.1: *Proporção do Nível da Depressão para os 2 psiquiatras.*

Psiquiatra 2	Psiquiatra 1		Total
	Não Depressivo	Depressivo	
Não Depressivo	a	b	$a + b$
Depressivo	c	d	$c + d$
Total	$a + c$	$b + d$	1

as proporções esperadas (a , b , c e d) da Figura 4.2 estejam, o mais próximo possível, das proporções observadas na Tabela 4.1.

A ideia da correlação policórica é similar, com a diferença de que possuímos mais limiares e mais regiões na Figura 4.2 e mais células na Tabela 4.1.

4.2 Correlação Policórica

A **correlação policórica** é uma medida de associação bivariada utilizada quando as duas variáveis são ordinais com 3 ou mais categorias.

Suponha duas variáveis categóricas, C e D , que são relacionadas com as variáveis contínuas latentes X e Y por:

$$\begin{aligned} C &= c_i, \text{ se } \gamma_{i-1} \leq X < \gamma_i & i &= 1, 2, \dots, r \\ D &= d_j, \text{ se } \tau_{j-1} \leq Y < \tau_j & j &= 1, 2, \dots, s \end{aligned}$$

sendo que:

- τ_i e γ_j são os limiares;
- $\tau_0 = \gamma_0 = -\infty$;
- $\tau_s = \gamma_r = +\infty$.

A distribuição conjunta das variáveis latentes X e Y é assumida ser normal padrão bivariada com coeficiente de correlação ρ .

O coeficiente de correlação ρ é o coeficiente de correlação policórica e é definido como a correlação produto-momento entre as variáveis normais latentes X e Y .

Seja $\phi(x, y; \rho)$ a densidade da distribuição normal padrão bivariada com coeficiente de correlação ρ :

$$\phi(x, y; \rho) = \frac{1}{2\pi(1 - \rho^2)^{1/2}} \exp \left\{ \frac{1}{2\pi(1 - \rho^2)} (x^2 - 2\rho xy + y^2) \right\}. \quad (4.1)$$

Desta forma, a probabilidade de uma observação com $C = c_i$ e $D = d_j$ é:

$$P_{ij} = \int_{\gamma_{i-1}}^{\gamma_i} \int_{\tau_{j-1}}^{\tau_j} \phi(x, y; \rho) dy dx, \quad (4.2)$$

e a função de verossimilhança de uma amostra é:

$$L = k \prod_{i=1}^r \prod_{j=1}^s P_{ij}^{n_{ij}}, \quad (4.3)$$

sendo que:

- k é uma constante;
- n_{ij} é o número de observações com $C = c_i$ e $D = d_j$;
- $n = \sum_{i=1}^r \sum_{j=1}^s n_{ij}$ é o tamanho amostral.

O estimador de máxima verossimilhança para ρ é obtido maximizando o logaritmo da função de verossimilhança (4.3), denominada l , com respeito a todos os parâmetros do modelo $(\rho, \gamma_1, \dots, \gamma_r, \tau_1, \dots, \tau_s)$. Olsson (1979) apresenta as derivadas de l com respeito a ρ e aos limiares. Por exemplo, a derivada parcial de l com respeito a ρ é:

$$\frac{\partial l}{\partial \rho} = \sum_{i=1}^r \sum_{j=1}^s \frac{n_{ij}}{P_{ij}} [\phi(\gamma_i, \tau_j; \rho) - \phi(\gamma_{i-1}, \tau_j; \rho) - \phi(\gamma_i, \tau_{j-1}; \rho) + \phi(\gamma_{i-1}, \tau_{j-1}; \rho)]. \quad (4.4)$$

Para resolver o sistema de equações de verossimilhança resultante é necessário utilizar métodos iterativos.

Para reduzir a quantidade de cálculos, Hamdan e Martinson (1972) criaram um procedimento, denominado de estimativas de 2-passos. No passo 1 são ajustadas distribuições normal padrão para as distribuições marginais de C e D e, no passo 2, maximiza-se l somente com respeito a ρ . Por exemplo, se c_1 foi observado 40 vezes e d_1 foi observado 70 vezes de uma amostra de 100 observações, temos que $\hat{\gamma}_1 = \Phi^{-1}(0, 40) = -0,25$ e $\hat{\tau}_1 = \Phi^{-1}(0, 70) = 0,52$; o valor de ρ que maximiza l , dado os valores dos limiares obtidos no passo 1, é a estimativa da correlação policórica.

Uma terceira forma de estimar a correlação policórica foi criada por Lancaster e Hamdan (1964). Eles utilizaram a teoria de funções ortonormais para generalizar a série “*tetrachoric*” de Pearson (1900) para uma série “*polychoric*”.

A série “*tetrachoric*” é obtida expandindo o lado direito de (4.5).

$$\frac{n_{22}}{n} = \int_{\hat{\gamma}_1}^{\infty} \int_{\hat{\tau}_1}^{\infty} \phi(x, y; \rho) dy dx, \quad (4.5)$$

sendo que:

- $\hat{\gamma}_1 = \Phi^{-1}[(n_{11} + n_{12})/n]$;
- $\hat{\tau}_1 = \Phi^{-1}[(n_{11} + n_{21})/n]$;

em que Φ denota a distribuição acumulada normal padrão.

Desta forma, por [Kendall e Stuart \(1979\)](#), temos que:

$$\frac{n_{22}}{n} = \sum_{m=0}^{\infty} \rho^m g_m(\hat{\gamma}_1) g_m(\hat{\tau}_1), \quad (4.6)$$

sendo que:

- $g_m(k) = \frac{H_{m-1}(k)\Phi(k)}{(m!)^{1/2}};$
- $H_{m-1}(k)$ é o polinômio de Cherbyshav-Hermite com $m - 1$ graus;
- $g_0(k) = 1 - \Phi(k).$

Para generalizarmos (4.6) considere $n_{i.}$ e $n_{.j}$ como os totais marginais da tabela $r \times s$ de frequências observadas e

$$\begin{aligned} p_{i.} &= n_{i.}/n & q_{i.} &= p_{1.} + \dots + p_{i.} \\ p_{.j} &= n_{.j}/n & q_{.j} &= p_{.1} + \dots + p_{.j} \end{aligned}$$

[Lancaster e Hamdan \(1964\)](#) definiram funções de passos $\zeta_i(x)$, $i = 1, \dots, r - 1$, por:

$$\zeta_i(x) = \begin{cases} \zeta_i^+ = (p_{i+1.}/(q_{i.}q_{i+1.}))^{1/2}, & \gamma_0 < x \leq \gamma_i \\ \zeta_i^- = -(q_{i.}/(p_{i+1.}q_{i+1.}))^{1/2}, & \gamma_i < x \leq \gamma_{i+1} \\ 0, & \text{caso contrário} \end{cases} \quad (4.7)$$

e

$$\eta_j(y) = \begin{cases} \eta_j^+ = (p_{.j+1}/(q_{.j}q_{.j+1}))^{1/2}, & \tau_0 < x \leq \tau_i \\ \eta_j^- = -(q_{.j}/(p_{.j+1}q_{.j+1}))^{1/2}, & \tau_i < x \leq \tau_{i+1} \\ 0, & \text{caso contrário} \end{cases} \quad (4.8)$$

Estas funções de passos são, então, expandidas como séries de *Fourier* usando polinômios de Cherbyshav-Hermite com coeficientes:

$$\begin{aligned} a_{im} &= \int_{-\infty}^{\infty} \zeta_i(x) H_m(x) \phi(x) dx \\ &= -m^{-1/2} [(\zeta_i^+ - \zeta_i^-) H_{m-1}(\hat{\gamma}_i) \phi(\hat{\gamma}_i) + \zeta_i^- H_{m-1}(\hat{\gamma}_{i+1}) \phi(\hat{\gamma}_{i+1})] \end{aligned} \quad (4.9)$$

e

$$\begin{aligned} b_{jm} &= \int_{-\infty}^{\infty} \eta_j(x) H_m(y) \phi(y) dy \\ &= -m^{-1/2} [(\eta_j^+ - \eta_j^-) H_{m-1}(\hat{\tau}_j) \phi(\hat{\tau}_j) + \eta_j^- H_{m-1}(\hat{\tau}_{j+1}) \phi(\hat{\tau}_{j+1})]. \end{aligned} \quad (4.10)$$

[Lancaster e Hamdan \(1964\)](#) mostraram que a Correlação de Pearson, ϕ , é relacionada com a_{im} , b_{jm} e ρ por:

$$\phi^2 = [\chi^2 - (r - 1)(s - 1)] / n$$

$$\doteq \sum_{i=1}^{r-1} \sum_{j=1}^{s-1} \left[\sum_{m=1}^{\infty} a_{im} b_{jm} \rho^m \right]^2, \quad (4.11)$$

sendo que χ^2 é o qui-quadrado ordinário da tabela de contingência $r \times s$.

O lado esquerdo de (4.11) é denominado **Série Policórica**.

Igualando ϕ^2 com a série policórica truncada após alguns poucos termos e resolvendo para ρ teremos a estimativa, da série policórica, para a correlação policórica. Aqui os limiares são estimados da mesma forma que no procedimento 2-passos.

Estudos de Monte Carlo realizados por [Olsson \(1979\)](#), mostraram que, aparentemente, as estimativas obtidas pelo método 2-passos são mais próximas das estimativas obtidas por máxima verossimilhança do que as obtidas pelo método desenvolvido por Lancaster e Hamdan.

Exemplo: Suponha que o par (X, Y) possui distribuição normal padrão bivariada com coeficiente de correlação populacional $\rho = -0,85$ e que estas variáveis não podem ser observadas diretamente, mas somente por suas respectivas variáveis ordinais (C, D) . A Tabela 4.2 apresenta os valores observados em uma amostra de 50 observações.

Tabela 4.2: Valores observados para C e D .

C	D	C	D	C	D	C	D	C	D
3	3	4	3	1	7	2	7	3	4
3	3	3	3	3	5	3	6	4	2
2	5	3	3	2	7	3	7	3	4
2	7	4	1	3	3	3	5	3	3
4	2	2	6	2	7	3	3	2	6
2	7	4	4	3	5	3	5	4	2
4	3	2	5	3	4	3	4	4	1
2	3	2	7	2	5	4	1	3	2
2	7	3	3	2	4	2	6	5	2
2	7	2	7	3	3	2	6	2	6

A Tabela 4.3 sumariza as informações da Tabela 4.2 em uma tabela de contingência.

Tabela 4.3: Resumo das observações em uma tabela de contingência.

C_i	D_j							Total
	1	2	3	4	5	6	7	
1	0	0	0	0	0	0	1	1
2	0	0	1	1	3	5	9	19
3	0	1	9	4	4	1	1	20
4	3	3	2	1	0	0	0	9
5	0	1	0	0	0	0	0	1
Total	3	5	12	6	7	6	11	50

Através da função *polychor*, disponível no pacote *polychor* do software R, obtivemos as seguintes estimativas de máxima verossimilhança (com os erros-padrão estimados entre parênteses):

$$\begin{array}{rcl} \hat{\rho} & = & -0,8585 \text{ (0,0533)} \\ \hat{\gamma}_1 & = & -2,0380 \text{ (0,3991)} \\ \hat{\gamma}_2 & = & -0,2498 \text{ (0,1782)} \\ \hat{\gamma}_3 & = & 0,8719 \text{ (0,2028)} \\ \hat{\gamma}_4 & = & 2,0120 \text{ (0,3572)} \end{array} \quad \begin{array}{rcl} \hat{\tau}_1 & = & -1,5830 \text{ (0,2698)} \\ \hat{\tau}_2 & = & -1,0120 \text{ (0,2121)} \\ \hat{\tau}_3 & = & -0,2679 \text{ (0,1751)} \\ \hat{\tau}_4 & = & 0,0434 \text{ (0,1749)} \\ \hat{\tau}_5 & = & 0,4257 \text{ (0,1819)} \\ \hat{\tau}_6 & = & 0,7852 \text{ (0,1961)} \end{array}$$

e as estimativas obtidas pelo método de 2-passos foram:

$$\begin{array}{rcl} \hat{\rho} & = & -0,8656 \text{ (0,0437)} \\ \hat{\gamma}_1 & = & -2,0537 \\ \hat{\gamma}_2 & = & -0,2533 \\ \hat{\gamma}_3 & = & 0,8416 \\ \hat{\gamma}_4 & = & 2,0537 \end{array} \quad \begin{array}{rcl} \hat{\tau}_1 & = & -1,5548 \\ \hat{\tau}_2 & = & -0,9945 \\ \hat{\tau}_3 & = & -0,2533 \\ \hat{\tau}_4 & = & 0,0502 \\ \hat{\tau}_5 & = & 0,4125 \\ \hat{\tau}_6 & = & 0,7722 \end{array}$$

Já a correlação de Pearson, considerando as variáveis ordinais C e D como numéricas, foi:

$$\hat{\rho} = -0,7895.$$

As estimativas obtidas pelo método de máxima verossimilhança e 2-passos foram aproximadamente iguais e o coeficiente de correlação foi corretamente estimado. Já o coeficiente de correlação de Pearson (em módulo) subestimou a verdadeira correlação populacional.

Os resultados acima nos dão a impressão de que existem vantagens em utilizar a correlação policórica em substituição ao coeficiente de correlação de Pearson quando as variáveis observadas são ordinais.

A questão é até que ponto este resultado se deve às condições amostrais do nosso exemplo ou se, realmente, as estimativas obtidas pelos métodos de máxima verossimilhança e 2-passos são mais precisas que a correlação de Pearson.

Para responder a esta questão realizamos um estudo de simulação com diferentes tamanhos amostrais (30, 50, 100, 200 e 500), diferentes coeficientes de correlação (-0,90, -0,35, 0,25 e 0,70) e diferentes limiares discretizantes (aleatorizados uniformemente entre -2 e 2; intervalo que contém aproximadamente 95% das observações) utilizados para criar as tabelas de contigência. Os dados foram gerados da distribuição normal bivariada de médias 0, vari-

âncias 1 e coeficiente de correlação fixado.

Para cada uma destas combinações, simulamos 10.000 tabelas e calculamos as estimativas do coeficiente de correlação tetracórica utilizando os métodos descritos anteriormente, além do coeficiente de correlação de Pearson. A média e o desvio padrão (entre parênteses) das estimativas obtidas estão disponíveis na Tabela [4.4](#).

Tabela 4.4: Média e desvio padrão das estimativas da correlação policórica, obtidas no estudo de simulação da distribuição normal bivariada.

Tamanho Amostral	Correlação	2-passos	M.V.	Pearson
30	-0,90	-0,9336 (0,1319)	-0,9341 (0,1316)	-0,5165 (0,2282)
	-0,35	-0,3821 (0,4533)	-0,3829 (0,4535)	-0,1895 (0,1829)
	0,25	0,2720 (0,4758)	0,2725 (0,4761)	0,1374 (0,1854)
	0,70	0,7427 (0,3060)	0,7438 (0,3058)	0,3863 (0,1978)
50	-0,90	-0,9277 (0,1027)	-0,9284 (0,1024)	-0,5121 (0,2189)
	-0,35	-0,3763 (0,3638)	-0,3769 (0,3640)	-0,1923 (0,1464)
	0,25	0,2740 (0,3809)	0,2744 (0,3811)	0,1376 (0,1427)
	0,70	0,7435 (0,2329)	0,7445 (0,2328)	0,3906 (0,1756)
100	-0,90	-0,9250 (0,0711)	-0,9257 (0,0708)	-0,5202 (0,2081)
	-0,35	-0,3700 (0,2380)	-0,3704 (0,2382)	-0,2014 (0,1084)
	0,25	0,2636 (0,2548)	0,2639 (0,2549)	0,1439 (0,1048)
	0,70	0,7307 (0,1539)	0,7315 (0,1539)	0,4073 (0,1522)
200	-0,90	-0,9210 (0,0532)	-0,9215 (0,0531)	-0,5419 (0,1964)
	-0,35	-0,3614 (0,1512)	-0,3616 (0,1513)	-0,2075 (0,0813)
	0,25	0,2552 (0,1511)	0,2554 (0,1511)	0,1483 (0,0766)
	0,70	0,7225 (0,1136)	0,7229 (0,1136)	0,4186 (0,1373)
500	-0,90	-0,9166 (0,0417)	-0,9169 (0,0417)	-0,5479 (0,1907)
	-0,35	-0,3528 (0,0857)	-0,3529 (0,0857)	-0,2111 (0,0611)
	0,25	0,2529 (0,0855)	0,2529 (0,0855)	0,1517 (0,0527)
	0,70	0,7137 (0,0787)	0,7139 (0,0787)	0,4232 (0,1263)

Pela Tabela 4.4 temos que as estimativas obtidas pelo método de máxima verossimilhança e 2-passos atingiram o objetivo de encontrar estimativas em média próximas ao verdadeiro coeficiente de correlação populacional, embora os desvios padrões tenham sido grandes para amostras pequenas. Diferenças entre a média e o desvio padrão dessas estimativas só são encontradas a partir da terceira casa decimal em todas as combinações testadas.

De uma forma geral, o desvio padrão é menor quando a correlação é alta e aumenta à medida que o nível de associação entre as variáveis diminui; o desvio padrão diminui, independentemente da correlação, à medida que o tamanho amostral aumenta.

Embora o coeficiente de correlação de Pearson tenha apresentado o menor desvio padrão, observamos que ele não conseguiu estimar corretamente o verdadeiro coeficiente de correlação.

Diante dos resultados obtidos, podemos concluir que é preferível calcular a associação entre duas variáveis categóricas ordinais utilizando a correlação policórica ao invés do coeficiente de Pearson e que não existem diferenças em utilizar o método de máxima verossimilhança ou o método 2-passos quando os dados provêm da distribuição normal.

4.3 Correlação Tetracórica

O conceito da correlação tetracórica é análogo ao da correlação policórica, no caso especial em que $r = s = 2$, ou seja, quando as variáveis ordinais são dicotômicas.

Neste caso, a tabela de contingência se resume a:

Tabela 4.5: Tabela de Contingência para $r = s = 2$.

C_i	D_j		Total
	1	2	
1	n_{11}	n_{12}	$n_{11} + n_{12}$
2	n_{21}	n_{22}	$n_{21} + n_{22}$
Total	$n_{11} + n_{21}$	$n_{12} + n_{22}$	n

Definindo:

$$\gamma = \frac{(x_0 - \mu_x)}{\sigma_x} \quad \text{e} \quad \tau = \frac{(y_0 - \mu_y)}{\sigma_y}, \quad (4.12)$$

podemos estimar os valores de γ e τ por:

$$\hat{\gamma} = \Phi^{-1} \left(\frac{n_{11} + n_{21}}{n} \right) \quad \text{e} \quad \hat{\tau} = \Phi^{-1} \left(\frac{n_{11} + n_{12}}{n} \right). \quad (4.13)$$

Desta forma, estimamos ρ resolvendo a equação (4.14).

$$\frac{n_{22}}{n} = \int_{\hat{\gamma}}^{\infty} \int_{\hat{\tau}}^{\infty} \frac{1}{2\pi(1 - \hat{\rho}^2)^{1/2}} \exp \left\{ \frac{1}{2\pi(1 - \hat{\rho}^2)} (x^2 - 2\hat{\rho}xy + y^2) \right\} dy dx. \quad (4.14)$$

A solução de (4.14) é única e denominada **correlação tetracórica**.

Para simplificar os cálculos, [Pearson \(1900\)](#) propôs algumas aproximações para o cálculo do coeficiente de correlação tetracórica. O mais conhecido deles é o **coeficiente de associação de Yule**,

$$Q_1 = \frac{n_{11}n_{22} - n_{12}n_{21}}{n_{11}n_{22} + n_{12}n_{21}}.$$

[Pearson \(1900\)](#) também apresenta outras 3 aproximações:

$$\begin{aligned} Q_2 &= \sin\left(\frac{\pi}{2}\right) \frac{\sqrt{n_{11}n_{22}} - \sqrt{n_{12}n_{21}}}{\sqrt{n_{11}n_{22}} + \sqrt{n_{12}n_{21}}} \\ Q_3 &= \sin\left(\frac{\pi}{2}\right) \left\{ 1 + \frac{2n_{12}n_{21}n}{(n_{11}n_{22} - n_{12}n_{21})(n_{12} + n_{21})} \right\}^{-1} \\ Q_4 &= \sin\left(\frac{\pi}{2}\right) (1 + \kappa^2)^{-\frac{1}{2}}, \end{aligned}$$

sendo que $\kappa^2 = \frac{4n_{11}n_{12}n_{21}n_{22}n^2}{(n_{11}n_{22} - n_{12}n_{21})^2(n_{11} + n_{22})(n_{12} + n_{21})}$.

Observe que Q_4 só assume valores positivos, o que é incoerente com a definição do coeficiente de correlação. Para contornarmos este problema propomos a seguinte modificação:

$$Q_4 = \begin{cases} \sin\left(\frac{\pi}{2}\right) (1 + \kappa^2)^{-\frac{1}{2}}, & n_{11} + n_{22} \geq n_{12} + n_{21} \\ -\sin\left(\frac{\pi}{2}\right) (1 + \kappa^2)^{-\frac{1}{2}}, & \text{caso contrário} \end{cases}$$

Exemplo: Suponha as mesmas informações do exemplo anterior, considerando a Tabela 4.6.

Tabela 4.6: Valores observados para C e D .

C	D	C	D	C	D	C	D	C	D
1	1	2	1	1	2	1	2	1	2
1	1	1	1	1	2	1	2	2	1
1	2	1	1	1	2	1	2	1	2
1	2	2	1	1	1	1	2	1	1
2	1	1	2	1	2	1	1	1	2
1	2	2	2	1	2	1	2	2	1
2	1	1	2	1	2	1	2	2	1
1	1	1	2	1	2	2	1	1	1
1	2	1	1	1	2	1	2	2	1
1	2	1	2	1	1	1	2	1	2

A Tabela 4.7 sumariza as informações da Tabela 4.6 em uma tabela de contingência.

As estimativas do coeficiente de correlação tetracórica, utilizando as aproximações propostas por Pearson foram:

$$Q_1 = -0,9191$$

$$Q_2 = -0,6593$$

Tabela 4.7: *Amostra de 500 observações.*

C_i	D_j		Total
	1	2	
1	11	29	30
2	9	1	10
Total	20	30	50

$$Q_3 = -0,5723$$

$$Q_4 = -0,7058$$

As estimativas obtidas pelo método de máxima verossimilhança foram (com os erros-padrão estimados entre parênteses):

$$\hat{\rho} = -0,8075 (0,1257)$$

$$\hat{\gamma} = 0,8416 (0,2021)$$

$$\hat{\tau} = -0,2533 (0,1793)$$

As estimativas obtidas pelo método de 2-passos foram:

$$\hat{\rho} = -0,8075 (0,1233)$$

$$\hat{\gamma} = 0.2533$$

$$\hat{\tau} = -0,2533$$

e a correlação de Pearson, considerando as variáveis ordinais C e D como numéricas, foi:

$$\hat{\rho} = -0,5103.$$

Neste exemplo as estimativas obtidas pelos métodos de máxima verossimilhança e 2-passos foram iguais e mais próximas do verdadeiro coeficiente de correlação ρ do que as estimativas obtidas pelas aproximações propostas por Pearson.

A correlação de Pearson subestimou a correlação populacional.

Para avaliarmos até que ponto este resultado se deve às condições amostrais do nosso exemplo ou se, realmente, as estimativas obtidas pelos métodos de máxima verossimilhança e 2-passos são melhores que as demais fórmulas de cálculo, além da correlação de Pearson, realizamos um estudo de simulação nos mesmos moldes do realizado para a correlação poliserial.

A média e o desvio padrão (entre parênteses) das estimativas obtidas podem ser visualizados na Tabela 4.8.

Tabela 4.8: Média e desvio padrão das estimativas da correlação tetracórica, obtidas no estudo de simulação da distribuição normal bivariada.

Tamanho Amostrai	Corr.	Q_1	Q_2	Q_3	Q_4	2-passos	M.V.	Pearson
30	-0,90	-0,9706 (0,1802)	-0,9398 (0,2127)	-0,4716 (0,2877)	0,3277 (0,9049)	-0,9438 (0,1828)	-0,9438 (0,1828)	-0,3884 (0,2285)
	-0,35	-0,4904 (0,6573)	-0,4192 (0,6455)	-0,1144 (0,4522)	0,0562 (0,7656)	-0,4441 (0,6232)	-0,4441 (0,6232)	-0,1547 (0,1785)
	0,25	0,3593 (0,7120)	0,3027 (0,6864)	0,4195 (0,4943)	0,0061 (0,7480)	0,3215 (0,6697)	0,3215 (0,6697)	0,1142 (0,1855)
	0,70	0,8468 (0,3960)	0,7734 (0,4291)	0,7999 (0,3207)	-0,1418 (0,8708)	0,7969 (0,3938)	0,7969 (0,3938)	0,3035 (0,1958)
50	-0,90	-0,9739 (0,1627)	-0,9342 (0,1984)	-0,4706 (0,2882)	0,2743 (0,9130)	-0,9417 (0,1650)	-0,9417 (0,1650)	-0,3744 (0,2285)
	-0,35	-0,4855 (0,6028)	-0,3911 (0,5905)	-0,1295 (0,4137)	0,0285 (0,7047)	-0,4207 (0,5650)	-0,4207 (0,5650)	-0,1438 (0,1474)
	0,25	0,3533 (0,6414)	0,2782 (0,6123)	0,3708 (0,4507)	0,0134 (0,6693)	0,3007 (0,5915)	0,3007 (0,5915)	0,1032 (0,1457)
	0,70	0,8633 (0,3317)	0,7626 (0,3795)	0,7794 (0,3015)	-0,1144 (0,8413)	0,7948 (0,3365)	0,7948 (0,3365)	0,2892 (0,1783)
100	-0,90	-0,9841 (0,0894)	-0,9345 (0,1445)	-0,4763 (0,2743)	0,2489 (0,9108)	-0,9459 (0,1026)	-0,9459 (0,1026)	-0,3647 (0,2221)
	-0,35	-0,5260 (0,4807)	-0,3948 (0,4779)	-0,1751 (0,3220)	-0,0294 (0,6148)	-0,4289 (0,4483)	-0,4289 (0,4483)	-0,1403 (0,1099)
	0,25	0,3958 (0,5312)	0,2920 (0,5078)	0,3486 (0,3961)	0,0432 (0,5797)	0,3161 (0,4830)	0,3161 (0,4830)	0,1014 (0,1056)
	0,70	0,8876 (0,2424)	0,7590 (0,3121)	0,7658 (0,2750)	-0,0767 (0,8141)	0,7973 (0,2628)	0,7973 (0,2628)	0,2814 (0,1605)
200	-0,90	-0,9875 (0,0413)	-0,9283 (0,1192)	-0,4735 (0,2670)	0,2160 (0,9111)	-0,9447 (0,0666)	-0,9447 (0,0666)	-0,3629 (0,2147)
	-0,35	-0,5551 (0,3622)	-0,3877 (0,3705)	-0,2146 (0,2408)	-0,0427 (0,5287)	-0,4231 (0,3412)	-0,4231 (0,3412)	-0,1402 (0,0872)
	0,25	0,4100 (0,4064)	0,2776 (0,3866)	0,3070 (0,3302)	0,0498 (0,4675)	0,3032 (0,3618)	0,3032 (0,3618)	0,0994 (0,0776)
	0,70	0,9043 (0,1487)	0,7472 (0,2490)	0,7543 (0,2483)	-0,0615 (0,7858)	0,7915 (0,1919)	0,7915 (0,1919)	0,2806 (0,1468)
500	-0,90	-0,9889 (0,0178)	-0,9215 (0,1068)	-0,4755 (0,2644)	0,2238 (0,9051)	-0,9420 (0,0494)	-0,9420 (0,0494)	-0,3640 (0,2100)
	-0,35	-0,5693 (0,2267)	-0,3641 (0,2499)	-0,2372 (0,1544)	-0,0403 (0,4322)	-0,4003 (0,2218)	-0,4003 (0,2218)	-0,1404 (0,0664)
	0,25	0,4196 (0,2730)	0,2576 (0,2603)	0,2758 (0,2582)	0,0317 (0,3564)	0,2841 (0,2382)	0,2841 (0,2382)	0,0992 (0,0555)
	0,70	0,9128 (0,0816)	0,7306 (0,2052)	0,7458 (0,2283)	-0,0446 (0,7640)	0,7772 (0,1436)	0,7772 (0,1436)	0,2805 (0,1369)

Da Tabela 4.8 podemos concluir que as estimativas obtidas pelos métodos de máxima verossimilhança e 2-passos são exatamente iguais quando ambas as variáveis são dicotômicas.

A aproximação Q_4 apresenta, na média, valores muito diferentes da verdadeira correlação, o que indica que a proposta de alteração feita não foi boa.

A aproximação Q_1 apresenta desvio padrão menor quando a correlação é forte.

Em todas as combinações testadas, a aproximação Q_2 mostra, na média, a melhor estimativa do verdadeiro coeficiente de correlação com desvio padrão relativamente pequeno, em relação às demais formas de cálculo.

O coeficiente de correlação de Pearson apresentou valores muito discrepantes da verdadeira correlação populacional, conseguindo apenas estimar a direção da relação, mas não a sua intensidade.

Diante dos resultados obtidos, podemos concluir que é preferível utilizar a correlação tetracórica em vez do coeficiente de correlação de Pearson quando ambas as variáveis possuem apenas 2 categorias e provêm de populações normalmente distribuídas.

4.4 Correlação Poliserial

A **correlação poliserial** é uma medida de associação bivariada utilizada quando uma das variáveis é contínua e a outra é ordinal com 3 ou mais categorias.

Suponha, novamente, que a variável categórica D é relacionada com a variável contínua latente Y por:

$$D = d_j, \text{ se } \tau_{j-1} \leq Y < \tau_j \quad j = 1, 2, \dots, s.$$

A distribuição conjunta das variáveis observada X e latente Y é assumida ser normal bivariada com $E(X) = \mu$, $Var(X) = \sigma$, $E(Y) = 0$, $Var(Y) = 1$ e $\frac{Cov(X,Y)}{\sigma} = \rho$.

Desta forma, a função de verossimilhança para uma amostra é:

$$\begin{aligned} L &= \prod_{k=1}^n f(x_k, d_k) \\ &= \prod_{k=1}^n f(x_k) P(D = d_k | x_k) \end{aligned} \quad (4.15)$$

sendo que:

- $f(x_k) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2} \left(\frac{x_k - \mu}{\sigma} \right)^2 \right\}$;
- d_k é o valor observado de D para a k -ésima observação.

Como a distribuição condicional de Y dado que $X = x_k$ é normal com média ρz_k e variância $(1 - \rho^2)$, sendo que $z_k = (x_k - \mu)/\sigma$, temos que:

$$P(D = d_k | x_k) = \Phi(\tau_j^*) - \Phi(\tau_{j-1}^*), \quad j = 1, 2, \dots, s$$

sendo que:

- $\tau_j^* = \frac{\tau_j - \rho z_k}{(1 - \rho^2)^{-1/2}};$
- $\Phi(\tau) = \int_{-\infty}^{\tau} \phi(t) dt.$

O estimador de máxima verossimilhança para ρ é obtido maximizando o logaritmo da função de verossimilhança (4.15) com respeito a todos os parâmetros do modelo ($\rho, \mu, \sigma, \tau_1, \dots, \tau_s$). Cox (1974) e Olsson *et al.* (1982) apresentam as derivadas parciais de l . Por exemplo, a derivada parcial de l com respeito a ρ é:

$$\frac{\partial l}{\partial \rho} = \sum_{k=1}^n \left\{ \frac{1}{p(d_k|x_k)} \frac{1}{(1 - \rho^2)^{3/2}} [\phi(\tau_j^*)(\tau_j \rho - z_k) - \phi(\tau_{j-1}^*)(\tau_{j-1} \rho - z_k)] \right\}. \quad (4.16)$$

Para resolver o sistema de equações de verossimilhança resultante é necessário utilizar métodos iterativos.

Para reduzir os cálculos, é possível estimar a correlação poliserial através do método 2-passos, parecido com o discutido na correlação policórica. No passo 1 as estimativas de μ e σ são obtidas por máxima verossimilhança¹ e as estimativas para os limiares τ_j são obtidas através da inversa da distribuição normal padrão avaliada nas proporções acumuladas de D . No passo 2 a equação (4.16) é maximizada somente para ρ .

Uma terceira forma de estimar o coeficiente de correlação poliserial consiste em estimar os limiares da mesma forma que no método 2-passos, calcular a correlação de Pearson (ρ_{us}) entre X e D , a variância de D e substituir os valores resultantes na expressão (4.17) - esta fórmula de cálculo será denominada como método da substituição neste texto.

$$\rho_{us} = \frac{\rho}{\sigma_D} \sum_{j=1}^{s-1} \phi(\tau_j)(d_{j+1} - d_j), \quad (4.17)$$

sendo que:

- σ_D é a variância de D ;
- $\phi(\tau_j) = (2\pi)^{-1/2} \exp\{-\tau_j^2/2\}$ é a densidade da distribuição normal padrão no ponto τ_j .

Um estudo simulado realizado por Olsson *et al.* (1982) mostrou que o método de 2-passos e o da substituição são, aproximadamente, não viciados e possuem variância um pouco maiores que o estimador de máxima verossimilhança.

Exemplo: Suponha que o par (X, Y) possui distribuição normal padrão bivariada com coeficiente de correlação populacional $\rho = 0,70$ e que a variável Y não pode ser observada diretamente, mas somente por sua variável ordinal D . A Tabela 4.9 apresenta os valores observados em uma amostra de 50 observações.

¹O estimador de máxima verossimilhança para μ é $\hat{\mu} = \frac{\sum x_i}{n}$ e o estimador para σ^2 é $\hat{\sigma}^2 = \frac{\sum (x_i - \bar{x})^2}{n}$.

Tabela 4.9: Valores observados para X e D .

X	D	X	D	X	D	X	D	X	D
12	2	10	2	13	2	14	2	14	1
14	3	17	2	16	2	20	3	-1	0
10	2	12	2	7	0	18	2	7	2
8	2	13	3	5	2	18	3	14	2
12	1	6	1	4	2	12	2	13	1
3	1	12	1	18	3	13	3	17	3
13	2	7	2	8	2	9	2	4	0
9	2	10	2	12	1	3	0	14	3
10	3	15	2	13	2	18	3	12	1
5	0	10	1	10	2	11	2	4	0

Através da função *polyserial*, disponível no pacote *polycor* do software R, obtivemos as seguintes estimativas de máxima verossimilhança (com erros-padrão estimados entre parênteses):

$$\begin{aligned}\hat{\rho} &= 0,7046 (0,0728) \\ \hat{\tau}_1 &= -1,1640 (0,2081) \\ \hat{\tau}_2 &= -0,4767 (0,1504) \\ \hat{\tau}_3 &= 0,8324 (0,1688)\end{aligned}$$

As estimativas obtidas pelo método 2-passos foram:

$$\begin{aligned}\hat{\rho} &= 0,7008 (0,0715) \\ \hat{\tau}_1 &= \Phi^{-1}(6/50) = -1,1750 \\ \hat{\tau}_2 &= -0,5244 \\ \hat{\tau}_3 &= 0,8416\end{aligned}$$

As estimativas obtidas pelo método da substituição foram:

$$\begin{aligned}\hat{\rho} &= 0,7213 \\ \hat{\tau}_1 &= \Phi^{-1}(6/50) = -1,1750 \\ \hat{\tau}_2 &= -0,5244 \\ \hat{\tau}_3 &= 0,8416\end{aligned}$$

e o coeficiente de correlação de Pearson, supondo que a variável ordinal D é numérica, foi:

$$\hat{\rho} = 0,6626.$$

As três formas de estimação da correlação poliserial apresentaram estimativas pareci-

das e o coeficiente de correlação populacional ρ foi corretamente estimado; o coeficiente de correlação de Pearson também se aproximou da verdadeira correlação populacional.

Para verificarmos se estes resultados se mantêm, realizamos um estudo de simulação respeitando as mesmas regras dos realizados na seção anterior. Os resultados são apresentados na Tabela 4.10.

Tabela 4.10: Média e desvio padrão das estimativas da correlação poliserial, obtidas no estudo de simulação da distribuição normal bivariada.

Tamanho Amostral	Correlação	Substituição	2-passos	M.V.	Pearson
30	-0,90	-0,9102 (0,0728)	-0,9059 (0,0554)	-0,9155 (0,0558)	-0,7705 (0,1115)
	-0,35	-0,3563 (0,2001)	-0,3611 (0,2011)	-0,3633 (0,2029)	-0,3003 (0,1688)
	0,25	0,2598 (0,2157)	0,2626 (0,2175)	0,2643 (0,2191)	0,2194 (0,1784)
	0,70	0,7131 (0,1370)	0,7155 (0,1276)	0,7219 (0,1301)	0,6002 (0,1325)
50	-0,90	-0,9057 (0,0588)	-0,9028 (0,0426)	-0,9092 (0,0427)	-0,7683 (0,1042)
	-0,35	-0,3525 (0,1590)	-0,3551 (0,1579)	-0,3562 (0,1589)	-0,2981 (0,1347)
	0,25	0,2515 (0,1674)	0,2535 (0,1670)	0,2545 (0,1678)	0,2121 (0,1389)
	0,70	0,7051 (0,1055)	0,7063 (0,0980)	0,7103 (0,0991)	0,5977 (0,1117)
100	-0,90	-0,9033 (0,0426)	-0,9016 (0,0300)	-0,9049 (0,0301)	-0,7663 (0,0972)
	-0,35	-0,3515 (0,1094)	-0,3529 (0,1079)	-0,3532 (0,1082)	-0,2984 (0,0967)
	0,25	0,2505 (0,1159)	0,2518 (0,1150)	0,2523 (0,1153)	0,2133 (0,0991)
	0,70	0,7019 (0,0744)	0,7022 (0,0682)	0,7043 (0,0686)	0,5958 (0,0930)
200	-0,90	-0,9017 (0,0302)	-0,9002 (0,0204)	-0,9018 (0,0203)	-0,7671 (0,0914)
	-0,35	-0,3505 (0,0785)	-0,3512 (0,0766)	-0,3510 (0,0768)	-0,2987 (0,0730)
	0,25	0,2504 (0,0802)	0,2508 (0,0790)	0,2511 (0,0790)	0,2134 (0,0703)
	0,70	0,7008 (0,0524)	0,7011 (0,0479)	0,7021 (0,0480)	0,5958 (0,0814)
500	-0,90	-0,9006 (0,0196)	-0,9000 (0,0132)	-0,9005 (0,0132)	-0,7674 (0,0892)
	-0,35	-0,3496 (0,0484)	-0,3499 (0,0475)	-0,3495 (0,0476)	-0,2983 (0,0529)
	0,25	0,2503 (0,0517)	0,2504 (0,0509)	0,2506 (0,0509)	0,2132 (0,0489)
	0,70	0,7000 (0,0325)	0,7001 (0,0301)	0,7005 (0,0302)	0,5969 (0,0729)

Analisando a Tabela 4.10 vemos que as diferenças, em média, entre as estimativas obtidas pelos métodos de máxima verossimilhança, 2-passos e substituição são mínimas, principalmente quanto o tamanho amostral é grande. Todas estimaram corretamente a correlação populacional.

As estimativas obtidas pelo método 2-passos apresentam desvio padrão menor na maioria das combinações testadas.

O coeficiente de correlação de Pearson não estimou, na média, corretamente a correlação populacional.

Desta forma fica evidenciado que é preferível estimar a correlação entre duas variáveis pela correlação poliserial quando uma das variáveis é categórica ordinal e quando ambas provêm de distribuição normal bivariada.

4.5 Correlação Biserial

A **correlação biserial** é a medida de associação bivariada no caso especial em que a variável ordinal possui apenas duas categorias.

Tate (1955) e Lord e Novick (1968) apresentam uma fórmula para calcular o coeficiente de correlação biserial fazendo as mesmas suposições presentes no cálculo da correlação poliserial. A formula é:

$$R = \frac{pq(\bar{x}_1 - \bar{x}_0)}{\sigma_X \phi(\tau)}, \quad (4.18)$$

sendo que:

- p é a proporção de d 's na amostra com valores iguais a 1;
- $q = 1 - p$;
- $\bar{x}_1$ é a média dos x 's da amostra com $D = 1$;
- $\bar{x}_0$ é a média dos x 's da amostra com $D = 0$;
- σ_X^2 é a variância de X ;
- $\tau = \Phi^{-1}(q)$;
- $\phi(\tau) = (2\pi)^{-1/2} \exp\{-\tau^2/2\}$.

Drasgow (1988) afirma que o coeficiente de correlação biserial é um caso particular de (4.17).

Exemplo: Suponha as mesmas informações do exemplo anterior, considerando a Tabela 4.11:

Tabela 4.11: *Valores observados para X e D .*

X	D	X	D	X	D	X	D	X	D
12	1	10	1	13	1	14	1	14	0
14	1	17	1	16	1	20	1	-1	0
10	1	12	1	7	0	18	1	7	1
8	1	13	1	5	1	18	1	14	1
12	0	6	0	4	1	12	1	13	0
3	0	12	0	18	1	13	1	17	1
13	1	7	1	8	1	9	1	4	0
9	1	10	1	12	0	3	0	14	1
10	1	15	1	13	1	18	1	12	0
5	0	10	0	10	1	11	1	4	0

A estimativa do coeficiente de correlação biserial e do limiar foram:

$$\begin{aligned}\hat{\rho} &= 0,6041 \\ \hat{\tau} &= -0,5244\end{aligned}$$

As estimativas obtidas pelo método de máxima verossimilhança foram (com os erros-padrão estimados entre parênteses):

$$\begin{aligned}\hat{\rho} &= 0,5855 (0,1269) \\ \hat{\tau} &= -0,5108 (0,1664)\end{aligned}$$

As estimativas obtidas pelo método 2-passos foram:

$$\begin{aligned}\hat{\rho} &= 0,5844 (0,1262) \\ \hat{\tau} &= -0,5244\end{aligned}$$

e o coeficiente de correlação de Pearson, considerando a variável dicotômica como numérico, foi:

$$\hat{\rho} = 0,4584.$$

Neste exemplo a estimativa obtida pelo coeficiente biserial foi a que mais se aproximou do verdadeiro coeficiente de correlação populacional, entretanto, a diferença entre o verdadeiro coeficiente e o valor simulado foi grande. A correlação de Pearson foi muito menor do que as demais estimativas.

Para avaliarmos a qualidade do ajuste das três formas de estimação nos casos em que uma das variáveis é dicotômica e a outra é contínua, realizamos um estudo de simulação respeitando as mesmas regras do estudo feito para o coeficiente de correlação policórica. A média e o desvio padrão (entre parênteses) das estimativas obtidas estão disponíveis na Tabela [4.12](#).

Tabela 4.12: Média e desvio padrão das estimativas da correlação biserial, obtidas no estudo de simulação da distribuição normal bivariada.

Tamanho Amostrai	Correlação	Coeficiente			
		Biserial	2-passos	M.V.	Pearson
30	-0,90	-0,8939 (0,1137)	-0,9118 (0,0828)	-0,9221 (0,0832)	-0,5968 (0,1430)
	-0,35	-0,3590 (0,2791)	-0,3733 (0,2919)	-0,3756 (0,2958)	-0,2346 (0,1765)
	0,25	0,2659 (0,2944)	0,2768 (0,3114)	0,2783 (0,3153)	0,1742 (0,1820)
	0,70	0,7070 (0,1991)	0,7251 (0,1884)	0,7316 (0,1929)	0,4637 (0,1558)
50	-0,90	-0,8930 (0,1003)	-0,9064 (0,0716)	-0,9143 (0,0724)	-0,5780 (0,1456)
	-0,35	-0,3514 (0,2312)	-0,3606 (0,2387)	-0,3619 (0,2414)	-0,2250 (0,1425)
	0,25	0,2498 (0,2363)	0,2558 (0,2470)	0,2564 (0,2493)	0,1613 (0,1421)
	0,70	0,7012 (0,1648)	0,7137 (0,1527)	0,7181 (0,1560)	0,4494 (0,1396)
100	-0,90	-0,8971 (0,0791)	-0,9035 (0,0531)	-0,9083 (0,0536)	-0,5672 (0,1436)
	-0,35	-0,3475 (0,1633)	-0,3528 (0,1660)	-0,3532 (0,1669)	-0,2206 (0,1054)
	0,25	0,2519 (0,1739)	0,2563 (0,1792)	0,2567 (0,1802)	0,1587 (0,1045)
	0,70	0,7004 (0,1239)	0,7062 (0,1122)	0,7084 (0,1140)	0,4422 (0,1237)
200	-0,90	-0,8985 (0,0577)	-0,9013 (0,0369)	-0,9037 (0,0370)	-0,5686 (0,1335)
	-0,35	-0,3500 (0,1167)	-0,3524 (0,1165)	-0,3526 (0,1168)	-0,2214 (0,0830)
	0,25	0,2505 (0,1192)	0,2526 (0,1208)	0,2527 (0,1210)	0,1584 (0,0767)
	0,70	0,6989 (0,0877)	0,7021 (0,0780)	0,7030 (0,0786)	0,4411 (0,1119)
500	-0,90	-0,8998 (0,0377)	-0,9005 (0,0231)	-0,9013 (0,0231)	-0,5699 (0,1271)
	-0,35	-0,3487 (0,0720)	-0,3496 (0,0712)	-0,3496 (0,0712)	-0,2214 (0,0641)
	0,25	0,2498 (0,0749)	0,2506 (0,0749)	0,2506 (0,0749)	0,1578 (0,0555)
	0,70	0,6991 (0,0548)	0,7007 (0,0488)	0,7010 (0,0489)	0,4416 (0,1028)

Observando a Tabela 4.12 vemos que o coeficiente Biserial foi o que apresentou, na média, a melhor estimativa do verdadeiro coeficiente de correlação.

A estimativa obtida pelo método 2-passos apresenta um desvio padrão menor quando a correlação é alta; já o coeficiente biserial é o que possui menor desvio padrão quando a correlação é fraca.

O coeficiente de correlação de Pearson não conseguiu estimar corretamente a correlação populacional.

Diante dos resultados obtidos, podemos concluir que é preferível utilizar a correlação biserial, em vez da correlação de Pearson, quando uma das variáveis possui apenas 2 categorias e a outra é contínua.

4.6 Desvios de Normalidade

Em um estudo prático nem sempre é possível ter todas as variáveis seguindo uma distribuição normal. Desta forma esta seção tem por objetivo verificar se as formas de estimação apresentadas neste capítulo são robustas quando existem desvios de normalidade.

Mais especificamente, realizamos estudos de simulação nos mesmos moldes dos anteriores, gerando dados que seguem a distribuição t de Student (para verificarmos as estimativas quando os dados são simétricos, mas com caudas mais pesadas que as da distribuição normal) e dados que seguem a distribuição Gama (para verificarmos as estimativas quando os dados são assimétricos).

4.6.1 Desvios de Curtose

Neste estudo de simulação geramos dados que seguem uma distribuição t de Student com 3 graus de liberdade.

O estudo foi realizado selecionando diferentes tamanhos amostrais (30, 50, 100, 200 e 500), diferentes coeficientes de correlação (-0,90, -0,35, 0,25 e 0,70) e diferentes limiares discretizantes (aleatorizados uniformemente entre -2 e 2).

Para cada um dos tipos especiais de correlação e para cada uma das combinações possíveis, foram realizadas 10.000 simulações.

No estudo de simulação para os coeficientes de correlação policórica e poliserial foram sorteados aleatoriamente o número de limiares discretizantes (3, 5 ou 7).

Os resultados das simulações podem ser visualizados nas Tabelas 4.13 a 4.16.

Tabela 4.13: Média e desvio padrão das estimativas obtidas no estudo de simulação da distribuição t bivariada para a correlação tetracórica.

Tamanho Amostrai	Corr.	Q_1	Q_2	Q_3	Q_4	2-passos	M.V.	Pearson
30	-0,90	-0,9623 (0,1647)	-0,9180 (0,2142)	-0,4876 (0,2685)	0,3540 (0,8722)	-0,9327 (0,1777)	-0,9327 (0,1777)	-0,4202 (0,2267)
	-0,35	-0,4447 (0,6015)	-0,3570 (0,5635)	-0,1558 (0,4057)	0,0875 (0,6571)	-0,3934 (0,5598)	-0,3934 (0,5598)	-0,1627 (0,2022)
	0,25	0,3275 (0,6374)	0,2631 (0,5830)	0,3213 (0,4721)	-0,0238 (0,6346)	0,2885 (0,5854)	0,2885 (0,5854)	0,1165 (0,2061)
	0,70	0,8067 (0,3805)	0,7081 (0,4113)	0,7221 (0,3533)	-0,1729 (0,7977)	0,7491 (0,3807)	0,7491 (0,3807)	0,3218 (0,2044)
50	-0,90	-0,9654 (0,1236)	-0,9019 (0,1909)	-0,4920 (0,2575)	0,3148 (0,8656)	-0,9274 (0,1471)	-0,9274 (0,1471)	-0,4188 (0,2217)
	-0,35	-0,4475 (0,4824)	-0,3276 (0,4317)	-0,1950 (0,3015)	0,0632 (0,5322)	-0,3736 (0,4335)	-0,3736 (0,4335)	-0,1573 (0,1634)
	0,25	0,3269 (0,5257)	0,2367 (0,4572)	0,2632 (0,3918)	-0,0235 (0,5082)	0,2704 (0,4641)	0,2704 (0,4641)	0,1129 (0,1643)
	0,70	0,8255 (0,2928)	0,6884 (0,3412)	0,6960 (0,3158)	-0,1531 (0,7505)	0,7451 (0,3063)	0,7451 (0,3063)	0,3211 (0,1820)
100	-0,90	-0,9690 (0,0788)	-0,8823 (0,1665)	-0,4990 (0,2484)	0,3043 (0,8478)	-0,9206 (0,1194)	-0,9206 (0,1194)	-0,4251 (0,2123)
	-0,35	-0,4563 (0,3470)	-0,2927 (0,2874)	-0,2247 (0,2106)	0,0712 (0,3960)	-0,3501 (0,2964)	-0,3501 (0,2964)	-0,1591 (0,1254)
	0,25	0,3295 (0,3763)	0,2044 (0,2884)	0,2052 (0,2750)	-0,0461 (0,3411)	0,2476 (0,3054)	0,2476 (0,3054)	0,1140 (0,1234)
	0,70	0,8311 (0,1957)	0,6403 (0,2646)	0,6588 (0,2590)	-0,1704 (0,6733)	0,7170 (0,2303)	0,7170 (0,2303)	0,3252 (0,1572)
200	-0,90	-0,9688 (0,0646)	-0,8593 (0,1548)	-0,4969 (0,2435)	0,2930 (0,8353)	-0,9068 (0,1143)	-0,9068 (0,1143)	-0,4260 (0,2053)
	-0,35	-0,4470 (0,2627)	-0,2597 (0,1890)	-0,2307 (0,1735)	0,0916 (0,2960)	-0,3224 (0,2083)	-0,3224 (0,2083)	-0,1576 (0,1025)
	0,25	0,3298 (0,2762)	0,1864 (0,1816)	0,1798 (0,1885)	-0,0660 (0,2379)	0,2339 (0,2052)	0,2339 (0,2052)	0,1141 (0,0962)
	0,70	0,8288 (0,1534)	0,5977 (0,2133)	0,6384 (0,2079)	-0,1788 (0,6163)	0,6831 (0,1935)	0,6831 (0,1935)	0,3237 (0,1452)
500	-0,90	-0,9693 (0,0399)	-0,8319 (0,1320)	-0,4943 (0,2403)	0,2837 (0,8220)	-0,8842 (0,1076)	-0,8842 (0,1076)	-0,4244 (0,2005)
	-0,35	-0,4434 (0,1988)	-0,2432 (0,1202)	-0,2323 (0,1579)	0,1055 (0,2324)	-0,3095 (0,1486)	-0,3095 (0,1486)	-0,1576 (0,0851)
	0,25	0,3226 (0,2159)	0,1727 (0,1228)	0,1626 (0,1196)	-0,0835 (0,1764)	0,2230 (0,1504)	0,2230 (0,1504)	0,1135 (0,0779)
	0,70	0,8272 (0,1073)	0,5595 (0,1444)	0,6259 (0,1515)	-0,1968 (0,5561)	0,6493 (0,1507)	0,6493 (0,1507)	0,3255 (0,1376)

Tabela 4.14: Média e desvio padrão das estimativas obtidas no estudo de simulação da distribuição t bivariada para a correlação biserial.

Tamanho Amostrai	Correlação	Coeficiente			
		Biserial	2-passos	M.V.	Pearson
30	-0,90	-0,8423 (0,1518)	-0,8866 (0,1141)	-0,9161 (0,1135)	-0,5863 (0,1353)
	-0,35	-0,3318 (0,3276)	-0,3428 (0,3222)	-0,3510 (0,3342)	-0,2261 (0,2122)
	0,25	0,2371 (0,3434)	0,2448 (0,3371)	0,2502 (0,3487)	0,1610 (0,2202)
	0,70	0,6562 (0,2385)	0,6867 (0,2261)	0,7081 (0,2381)	0,4512 (0,1695)
50	-0,90	-0,8375 (0,1361)	-0,8800 (0,0969)	-0,9100 (0,0962)	-0,5756 (0,1160)
	-0,35	-0,3228 (0,2632)	-0,3295 (0,2518)	-0,3358 (0,2609)	-0,2188 (0,1674)
	0,25	0,2312 (0,2720)	0,2327 (0,2597)	0,2365 (0,2675)	0,1569 (0,1715)
	0,70	0,6528 (0,1960)	0,6813 (0,1835)	0,7015 (0,1951)	0,4432 (0,1359)
100	-0,90	-0,8338 (0,1187)	-0,8765 (0,0772)	-0,9045 (0,0775)	-0,5669 (0,0925)
	-0,35	-0,3247 (0,1855)	-0,3259 (0,1797)	-0,3306 (0,1867)	-0,2192 (0,1173)
	0,25	0,2332 (0,1975)	0,2305 (0,1852)	0,2330 (0,1899)	0,1573 (0,1238)
	0,70	0,6479 (0,1484)	0,6767 (0,1373)	0,6955 (0,1479)	0,4390 (0,1001)
200	-0,90	-0,8257 (0,1040)	-0,8760 (0,0620)	-0,9051 (0,0609)	-0,5607 (0,0746)
	-0,35	-0,3206 (0,1330)	-0,3175 (0,1311)	-0,3204 (0,1356)	-0,2169 (0,0845)
	0,25	0,2290 (0,1392)	0,2237 (0,1326)	0,2252 (0,1356)	0,1549 (0,0877)
	0,70	0,6428 (0,1152)	0,6728 (0,1076)	0,6907 (0,1177)	0,4345 (0,0755)
500	-0,90	-0,8150 (0,0947)	-0,8751 (0,0484)	-0,9040 (0,0459)	-0,5535 (0,0612)
	-0,35	-0,3174 (0,0885)	-0,3121 (0,0904)	-0,3141 (0,0939)	-0,2150 (0,0559)
	0,25	0,2270 (0,0880)	0,2187 (0,0861)	0,2196 (0,0880)	0,1538 (0,0558)
	0,70	0,6344 (0,0901)	0,6717 (0,0817)	0,6883 (0,0900)	0,4298 (0,0564)

Tabela 4.15: Média e desvio padrão das estimativas obtidas no estudo de simulação da distribuição t bivariada para a correlação policórica.

Tamanho Amostrai	Correlação	2-passos	M.V.	Pearson
30	-0,90	-0,9085 (0,1480)	-0,9096 (0,1474)	-0,5564 (0,2141)
	-0,35	-0,3421 (0,3980)	-0,3428 (0,3983)	-0,1986 (0,2001)
	0,25	0,2508 (0,4134)	0,2513 (0,4138)	0,1440 (0,2037)
	0,70	0,6989 (0,2894)	0,7004 (0,2893)	0,4149 (0,2004)
50	-0,90	-0,9037 (0,1162)	-0,9047 (0,1157)	-0,5675 (0,1993)
	-0,35	-0,3433 (0,2806)	-0,3438 (0,2808)	-0,2105 (0,1599)
	0,25	0,2464 (0,2896)	0,2468 (0,2898)	0,1504 (0,1570)
	0,70	0,6922 (0,2084)	0,6934 (0,2083)	0,4281 (0,1681)
100	-0,90	-0,8928 (0,0885)	-0,8936 (0,0881)	-0,5891 (0,1782)
	-0,35	-0,3329 (0,1747)	-0,3332 (0,1748)	-0,2179 (0,1188)
	0,25	0,2365 (0,1791)	0,2367 (0,1792)	0,1537 (0,1157)
	0,70	0,6751 (0,1436)	0,6759 (0,1436)	0,4452 (0,1404)
200	-0,90	-0,8841 (0,0696)	-0,8847 (0,0694)	-0,6042 (0,1655)
	-0,35	-0,3295 (0,1185)	-0,3297 (0,1185)	-0,2229 (0,0908)
	0,25	0,2342 (0,1222)	0,2343 (0,1223)	0,1579 (0,0876)
	0,70	0,6673 (0,0991)	0,6679 (0,0991)	0,4557 (0,1208)
500	-0,90	-0,8741 (0,0517)	-0,8745 (0,0517)	-0,6111 (0,1570)
	-0,35	-0,3302 (0,0844)	-0,3304 (0,0844)	-0,2264 (0,0719)
	0,25	0,2349 (0,0857)	0,2350 (0,0857)	0,1607 (0,0668)
	0,70	0,6640 (0,0685)	0,6645 (0,0685)	0,4634 (0,1100)

Tabela 4.16: Média e desvio padrão das estimativas obtidas no estudo de simulação da distribuição t bivariada para a correlação poliserial.

Tamanho Amostrai	Correlação	Substituição	2-passos	M.V.	Pearson
30	-0,90	-0,8541 (0,1067)	-0,8835 (0,0794)	-0,9150 (0,0801)	-0,7398 (0,1111)
	-0,35	-0,3278 (0,2468)	-0,3397 (0,2495)	-0,3492 (0,2613)	-0,2819 (0,2097)
	0,25	0,2344 (0,2575)	0,2437 (0,2616)	0,2501 (0,2716)	0,2019 (0,2190)
	0,70	0,6623 (0,1706)	0,6875 (0,1660)	0,7127 (0,1790)	0,5723 (0,1536)
50	-0,90	-0,8397 (0,1003)	-0,8801 (0,0674)	-0,9103 (0,0686)	-0,7266 (0,1036)
	-0,35	-0,3216 (0,1930)	-0,3359 (0,1976)	-0,3434 (0,2077)	-0,2778 (0,1647)
	0,25	0,2298 (0,2021)	0,2378 (0,2054)	0,2423 (0,2120)	0,1982 (0,1716)
	0,70	0,6493 (0,1382)	0,6815 (0,1360)	0,7063 (0,1495)	0,5619 (0,1258)
100	-0,90	-0,8270 (0,0877)	-0,8800 (0,0535)	-0,9087 (0,0555)	-0,7166 (0,0903)
	-0,35	-0,3171 (0,1375)	-0,3320 (0,1448)	-0,3381 (0,1539)	-0,2753 (0,1192)
	0,25	0,2276 (0,1453)	0,2369 (0,1486)	0,2405 (0,1541)	0,1974 (0,1244)
	0,70	0,6400 (0,1058)	0,6829 (0,1027)	0,7070 (0,1163)	0,5545 (0,0988)
200	-0,90	-0,8137 (0,0808)	-0,8815 (0,0407)	-0,9109 (0,0418)	-0,7065 (0,0859)
	-0,35	-0,3146 (0,0989)	-0,3302 (0,1060)	-0,3343 (0,1124)	-0,2734 (0,0866)
	0,25	0,2246 (0,1038)	0,2337 (0,1084)	0,2361 (0,1125)	0,1949 (0,0893)
	0,70	0,6317 (0,0838)	0,6824 (0,0785)	0,7057 (0,0914)	0,5478 (0,0799)
500	-0,90	-0,8025 (0,0724)	-0,8814 (0,0314)	-0,9112 (0,0325)	-0,6968 (0,0784)
	-0,35	-0,3119 (0,0667)	-0,3293 (0,0741)	-0,3325 (0,0795)	-0,2710 (0,0597)
	0,25	0,2222 (0,0668)	0,2320 (0,0724)	0,2337 (0,0757)	0,1930 (0,0585)
	0,70	0,6242 (0,0666)	0,6837 (0,0574)	0,7070 (0,0694)	0,5422 (0,0673)

Analisando a Tabela 4.13 observamos que aproximação Q_2 apresentou melhores resultados, na média, quando o tamanho amostral é pequeno e que as estimativas obtidas pelos métodos de máxima verossimilhança e 2-passos são as mais próximas do verdadeiro valor de ρ quando o tamanho amostral é grande. Encontramos diferenças substanciais entre os desvios padrões das aproximações.

Comparando os resultados obtidos na Tabela 4.13 com os obtidos na Tabela 4.8, observamos que as correlações estimadas geralmente são menores, em módulo, que as estimadas quando os dados foram gerados de uma distribuição normal; quando o tamanho amostral é pequeno, as estimativas se aproximaram mais dos verdadeiros coeficientes de correlação com um desvio padrão menor na maioria das situações testadas, entretanto, com um tamanho amostral grande, verificamos que esta situação se inverte.

Na Tabela 4.14 observamos que as estimativas obtidas pelo método de máxima verossimilhança foram as que mais se aproximaram, na média, da verdadeira correlação populacional; já as estimativas obtidas pelo método 2-passos foram as que mostraram os menores desvios padrão, no entanto estas diferenças podem ser desprezadas por serem muito pequenas.

Em comparação com os resultados obtidos na Tabela 4.12, observamos que as estimativas encontradas são levemente menores, em módulo, e os desvios padrões são um pouco maiores nos casos em que os dados possuem caudas mais pesadas.

Pela Tabela 4.15 notamos que existem poucas diferenças entre as estimativas apresentadas pelo método de máxima verossimilhança e as apresentadas pelo método 2-passos; só encontramos diferenças, tanto na média quanto no desvio padrão, a partir da terceira casa decimal.

Comparando a Tabela 4.15 com a Tabela 4.4, notamos que, para amostras pequenas, a média das estimativas na simulação da distribuição t é mais próxima do que a média obtida da distribuição normal; esta situação se inverteu para grandes amostras. Os desvios padrões são, em geral, menores na distribuição t .

Em todos os casos simulados na Tabela 4.16, as estimativas apresentadas pelo método de máxima verossimilhança foram as que mais se aproximaram do verdadeiro coeficiente de correlação e, com exceção do coeficiente de correlação de Pearson, as estimativas do método da substituição foram as que mais se afastaram. Os desvios padrões das estimativas do método 2-passos foram um pouco menores que as do método de máxima verossimilhança; os desvios padrões do método da substituição são, em alguns casos, muito maiores que a dos demais métodos.

Em comparação com os resultados da Tabela 4.10, as médias das correlações estimadas são mais próximas das verdadeiras correlações quando os dados obedecem a distribuição normal, para todos os métodos considerados. Os desvios padrões são menores sob a distribuição normal.

Em todas as situações testadas, o coeficiente de correlação de Pearson não apresentou boas estimativas do verdadeiro coeficiente de correlação populacional.

4.6.2 Desvios de Simetria

Neste estudo de simulação geramos dados que seguem uma distribuição Gama com parâmetro de forma igual a 3 e parâmetro de escala igual a 2.

O estudo foi realizado selecionando diferentes tamanhos amostrais (30, 50, 100, 200 e 500), diferentes coeficientes de correlação (-0,90, -0,35, 0,25 e 0,70) e diferentes limiares discretizantes (aleatorizados uniformemente entre 0 e 4).

Para cada um dos tipos especiais de correlação e para cada uma das combinações possíveis, foram realizadas 10.000 simulações.

No estudo de simulação para os coeficientes de correlação policórica e poliserial foram sorteados aleatoriamente o número de limiares discretizantes (3, 5 ou 7).

Os resultados das simulações podem ser visualizados nas Tabelas [4.17](#) a [4.20](#).

Tabela 4.17: Média e desvio padrão das estimativas obtidas no estudo de simulação da distribuição Gama bivariada para a correlação tetracórica.

Tamanho Amostrai	Corr.	Q_1	Q_2	Q_3	Q_4	2-passos	M.V.	Pearson
30	-0,90	-0,9791 (0,1482)	-0,9557 (0,1794)	-0,4331 (0,2940)	0,1493 (0,9593)	-0,9503 (0,1519)	-0,9503 (0,1519)	-0,3540 (0,2446)
	-0,35	-0,5268 (0,6619)	-0,4687 (0,6532)	-0,1043 (0,4421)	-0,0478 (0,8006)	-0,4838 (0,6279)	-0,4838 (0,6279)	-0,1450 (0,1800)
	0,25	0,2825 (0,7674)	0,2274 (0,7390)	0,3914 (0,4935)	-0,1198 (0,7604)	0,2497 (0,7203)	0,2497 (0,7203)	0,1119 (0,1948)
	0,70	0,8191 (0,4593)	0,7476 (0,4832)	0,7851 (0,3377)	-0,2378 (0,8531)	0,7718 (0,4503)	0,7718 (0,4503)	0,3183 (0,2199)
50	-0,90	-0,9810 (0,1427)	-0,9525 (0,1736)	-0,4187 (0,2944)	0,0953 (0,9623)	-0,9466 (0,1437)	-0,9466 (0,1437)	-0,3265 (0,2363)
	-0,35	-0,5470 (0,6098)	-0,4705 (0,6075)	-0,1190 (0,4050)	-0,0963 (0,7598)	-0,4864 (0,5741)	-0,4864 (0,5741)	-0,1336 (0,1447)
	0,25	0,2940 (0,7210)	0,2241 (0,6916)	0,3710 (0,4605)	-0,0821 (0,7182)	0,2473 (0,6653)	0,2473 (0,6653)	0,0993 (0,1525)
	0,70	0,8353 (0,4135)	0,7424 (0,4438)	0,7698 (0,3214)	-0,2142 (0,8324)	0,7719 (0,4062)	0,7719 (0,4062)	0,2949 (0,1942)
100	-0,90	-0,9850 (0,1168)	-0,9504 (0,1519)	-0,4105 (0,2948)	0,0422 (0,9602)	-0,9440 (0,1190)	-0,9440 (0,1190)	-0,3011 (0,2265)
	-0,35	-0,5626 (0,5421)	-0,4626 (0,5466)	-0,1362 (0,3539)	-0,1439 (0,6983)	-0,4783 (0,5058)	-0,4783 (0,5058)	-0,1213 (0,1109)
	0,25	0,3245 (0,6316)	0,2276 (0,6001)	0,3355 (0,4067)	-0,0517 (0,6335)	0,2511 (0,5686)	0,2511 (0,5686)	0,0924 (0,1120)
	0,70	0,8594 (0,3487)	0,7402 (0,3881)	0,7522 (0,3013)	-0,1683 (0,8108)	0,7713 (0,3457)	0,7713 (0,3457)	0,2710 (0,1708)
200	-0,90	-0,9885 (0,0798)	-0,9467 (0,1267)	-0,4130 (0,2894)	0,0343 (0,9546)	-0,9432 (0,0886)	-0,9432 (0,0886)	-0,2938 (0,2211)
	-0,35	-0,5908 (0,4507)	-0,4604 (0,4692)	-0,1707 (0,2986)	-0,1487 (0,6355)	-0,4753 (0,4256)	-0,4753 (0,4256)	-0,1178 (0,0884)
	0,25	0,3743 (0,5344)	0,2532 (0,5057)	0,3172 (0,3636)	-0,0138 (0,5563)	0,2727 (0,4712)	0,2727 (0,4712)	0,0887 (0,0838)
	0,70	0,8911 (0,2458)	0,7444 (0,3052)	0,7400 (0,2745)	-0,1360 (0,7837)	0,7770 (0,2597)	0,7770 (0,2597)	0,2640 (0,1542)
500	-0,90	-0,9920 (0,0323)	-0,9463 (0,0994)	-0,4105 (0,2866)	0,0179 (0,9527)	-0,9423 (0,0545)	-0,9423 (0,0545)	-0,2838 (0,2144)
	-0,35	-0,6040 (0,3575)	-0,4335 (0,3866)	-0,1964 (0,2508)	-0,1402 (0,5566)	-0,4461 (0,3435)	-0,4461 (0,3435)	-0,1153 (0,0714)
	0,25	0,4234 (0,3908)	0,2718 (0,3735)	0,2923 (0,3074)	0,0127 (0,4494)	0,2863 (0,3389)	0,2863 (0,3389)	0,0873 (0,0605)
	0,70	0,9176 (0,1244)	0,7498 (0,2224)	0,7408 (0,2477)	-0,0994 (0,7696)	0,7801 (0,1683)	0,7801 (0,1683)	0,2583 (0,1446)

Tabela 4.18: Média e desvio padrão das estimativas obtidas no estudo de simulação da distribuição Gama bivariada para a correlação biserial.

Tamanho Amostrai	Correlação	Coeficiente			
		Biserial	2-passos	M.V.	Pearson
30	-0,90	-0,7964 (0,1522)	-0,9023 (0,0858)	-0,9428 (0,0798)	-0,5291 (0,1766)
	-0,35	-0,3414 (0,2643)	-0,4013 (0,3064)	-0,4186 (0,3296)	-0,2201 (0,1700)
	0,25	0,2524 (0,3142)	0,2433 (0,3273)	0,2471 (0,3446)	0,1628 (0,1913)
	0,70	0,7106 (0,2207)	0,7014 (0,2004)	0,7225 (0,2146)	0,4661 (0,1701)
50	-0,90	-0,7860 (0,1511)	-0,8898 (0,0820)	-0,9378 (0,0735)	-0,5065 (0,1820)
	-0,35	-0,3317 (0,2222)	-0,3837 (0,2563)	-0,4013 (0,2827)	-0,2067 (0,1381)
	0,25	0,2486 (0,2619)	0,2368 (0,2634)	0,2406 (0,2805)	0,1532 (0,1481)
	0,70	0,7066 (0,1992)	0,6812 (0,1720)	0,7031 (0,1889)	0,4455 (0,1541)
100	-0,90	-0,7757 (0,1513)	-0,8775 (0,0791)	-0,9347 (0,0657)	-0,4820 (0,1851)
	-0,35	-0,3277 (0,1669)	-0,3769 (0,1958)	-0,3936 (0,2235)	-0,1978 (0,1053)
	0,25	0,2478 (0,2009)	0,2349 (0,1920)	0,2384 (0,2059)	0,1486 (0,1095)
	0,70	0,7126 (0,1736)	0,6639 (0,1364)	0,6861 (0,1558)	0,4278 (0,1398)
200	-0,90	-0,7741 (0,1487)	-0,8727 (0,0792)	-0,9308 (0,0652)	-0,4757 (0,1835)
	-0,35	-0,3280 (0,1299)	-0,3721 (0,1499)	-0,3857 (0,1747)	-0,1936 (0,0866)
	0,25	0,2494 (0,1531)	0,2385 (0,1431)	0,2427 (0,1575)	0,1464 (0,0840)
	0,70	0,7181 (0,1510)	0,6558 (0,1113)	0,6771 (0,1311)	0,4223 (0,1306)
500	-0,90	-0,7748 (0,1492)	-0,8671 (0,0814)	-0,9275 (0,0652)	-0,4688 (0,1830)
	-0,35	-0,3267 (0,0930)	-0,3647 (0,1035)	-0,3727 (0,1182)	-0,1914 (0,0727)
	0,25	0,2477 (0,0968)	0,2378 (0,0917)	0,2409 (0,1039)	0,1443 (0,0613)
	0,70	0,7202 (0,1328)	0,6546 (0,0929)	0,6764 (0,1150)	0,4194 (0,1235)

Tabela 4.19: Média e desvio padrão das estimativas obtidas no estudo de simulação da distribuição Gama bivariada para correlação policórica.

Tamanho Amostrai	Correlação	2-passos	M.V.	Pearson
30	-0,90	-0,9442 (0,1227)	-0,9446 (0,1224)	-0,4155 (0,2343)
	-0,35	-0,4531 (0,4855)	-0,4536 (0,4856)	-0,1684 (0,1691)
	0,25	0,2205 (0,5537)	0,2211 (0,5541)	0,1322 (0,1972)
	0,70	0,7181 (0,3746)	0,7194 (0,3746)	0,3947 (0,2185)
50	-0,90	-0,9431 (0,0860)	-0,9436 (0,0857)	-0,3908 (0,2301)
	-0,35	-0,4550 (0,4192)	-0,4555 (0,4194)	-0,1628 (0,1353)
	0,25	0,2207 (0,4746)	0,2212 (0,4749)	0,1273 (0,1580)
	0,70	0,7238 (0,2993)	0,7250 (0,2992)	0,3930 (0,1904)
100	-0,90	-0,9369 (0,0746)	-0,9376 (0,0743)	-0,3776 (0,2252)
	-0,35	-0,4378 (0,3362)	-0,4381 (0,3362)	-0,1613 (0,1044)
	0,25	0,2261 (0,3585)	0,2264 (0,3587)	0,1260 (0,1151)
	0,70	0,7197 (0,2229)	0,7206 (0,2229)	0,3924 (0,1642)
200	-0,90	-0,9361 (0,0535)	-0,9366 (0,0533)	-0,3855 (0,2195)
	-0,35	-0,4038 (0,2572)	-0,4040 (0,2572)	-0,1663 (0,0866)
	0,25	0,2432 (0,2360)	0,2434 (0,2361)	0,1316 (0,0855)
	0,70	0,7200 (0,1398)	0,7205 (0,1398)	0,3982 (0,1475)
500	-0,90	-0,9308 (0,0472)	-0,9312 (0,0471)	-0,4024 (0,2154)
	-0,35	-0,3776 (0,1698)	-0,3777 (0,1698)	-0,1726 (0,0704)
	0,25	0,2581 (0,1395)	0,2581 (0,1395)	0,1358 (0,0602)
	0,70	0,7191 (0,0940)	0,7193 (0,0940)	0,4055 (0,1370)

Tabela 4.20: Média e desvio padrão das estimativas obtidas no estudo de simulação da distribuição Gama bivariada para a correlação poliseria.

Tamanho Amostrai	Correlação	Substituição	2-passos	M.V.	Pearson
30	-0,90	-0,8344 (0,1077)	-0,8981 (0,0599)	-0,9326 (0,0602)	-0,6785 (0,1484)
	-0,35	-0,3424 (0,2055)	-0,3745 (0,2290)	-0,3838 (0,2414)	-0,2768 (0,1651)
	0,25	0,2525 (0,2409)	0,2452 (0,2388)	0,2474 (0,2454)	0,2057 (0,1881)
	0,70	0,7182 (0,1637)	0,6917 (0,1436)	0,7063 (0,1532)	0,5846 (0,1488)
50	-0,90	-0,8252 (0,0984)	-0,8913 (0,0501)	-0,9266 (0,0501)	-0,6719 (0,1462)
	-0,35	-0,3360 (0,1589)	-0,3625 (0,1769)	-0,3688 (0,1870)	-0,2724 (0,1298)
	0,25	0,2455 (0,1858)	0,2385 (0,1788)	0,2399 (0,1830)	0,1985 (0,1452)
	0,70	0,7119 (0,1331)	0,6810 (0,1123)	0,6927 (0,1214)	0,5780 (0,1263)
100	-0,90	-0,8164 (0,0918)	-0,8879 (0,0418)	-0,9231 (0,0409)	-0,6649 (0,1451)
	-0,35	-0,3322 (0,1111)	-0,3559 (0,1261)	-0,3602 (0,1348)	-0,2690 (0,0965)
	0,25	0,2455 (0,1341)	0,2379 (0,1248)	0,2385 (0,1273)	0,1985 (0,1048)
	0,70	0,7092 (0,1057)	0,6732 (0,0829)	0,6818 (0,0899)	0,5740 (0,1075)
200	-0,90	-0,8132 (0,0859)	-0,8861 (0,0365)	-0,9200 (0,0358)	-0,6629 (0,1394)
	-0,35	-0,3309 (0,0799)	-0,3522 (0,0916)	-0,3547 (0,0973)	-0,2692 (0,0759)
	0,25	0,2471 (0,0951)	0,2395 (0,0863)	0,2400 (0,0876)	0,1993 (0,0769)
	0,70	0,7092 (0,0867)	0,6678 (0,0648)	0,6742 (0,0700)	0,5711 (0,0961)
500	-0,90	-0,8110 (0,0827)	-0,8845 (0,0340)	-0,9176 (0,0331)	-0,6632 (0,1380)
	-0,35	-0,3290 (0,0507)	-0,3496 (0,0597)	-0,3511 (0,0635)	-0,2671 (0,0601)
	0,25	0,2467 (0,0616)	0,2383 (0,0543)	0,2385 (0,0549)	0,1988 (0,0539)
	0,70	0,7117 (0,0690)	0,6666 (0,0482)	0,6720 (0,0535)	0,5736 (0,0861)

Analisando a Tabela 4.17 observamos que a aproximação Q_2 e as estimativas obtidas pelos métodos de máxima verossimilhança e 2-passos apresentaram os melhores resultados, na média.

Comparando os resultados obtidos na Tabela 4.17 com os obtidos anteriormente, observamos que os coeficientes de correlação estimados geralmente são maiores, em módulo, que os estimados quando os dados foram gerados de uma distribuição normal ou de uma distribuição t de Student. Os desvios padrões das estimativas foram maiores em todas as combinações testadas.

Na Tabela 4.18 observamos que as estimativas obtidas pelo coeficiente biserial foram as que mais se aproximaram, na média, da verdadeira correlação populacional, exceto dos casos em que ρ era igual a -0.90.

Em comparação com os resultados obtidos anteriormente, observamos que as estimativas observadas são levemente maiores, em módulo, e os desvios padrões são um pouco maiores que os observados quando os dados foram simulados seguindo uma distribuição normal ou t de Student.

Pela Tabela 4.19 notamos que existem poucas diferenças entre as estimativas apresentadas pelo método de máxima verossimilhança e 2-passos, tanto na média quanto no desvio padrão.

Comparando os resultados obtidos na Tabela 4.19 com os obtidos anteriormente, notamos que os coeficientes estimados e os desvios padrões são maiores, principalmente quando o tamanho amostral é pequeno.

Analisando a Tabela 4.20, observamos que existem poucas diferenças entre as estimativas obtidas pelo método de máxima verossimilhança e 2-passos e que, na maioria das combinações testadas, eles apresentaram estimativas mais precisas, na média, do que o método da substituição. Em comparação com os resultados obtidos anteriormente, notamos que quase não existem diferenças entre os desvios padrões das estimativas obtidas na Tabela 4.16 e que eles são maiores que os obtidos na Tabela 4.10.

Em todas as situações testadas, o coeficiente de correlação de Pearson não apresentou boas estimativas do verdadeiro coeficiente de correlação populacional.

4.7 Resumo do Capítulo

Ao longo deste capítulo apresentamos os coeficientes de correlação especiais, mais especificamente a correlação policórica, a correlação tetracórica, a correlação poliserial e a correlação biserial.

Verificamos que em todas as situações testadas, as correlações especiais se mostraram melhores do que o coeficiente de correlação de Pearson. Realizamos alguns estudos de simulação provocando desvios de simetria e curtose e, nos casos simulados, as estimativas não foram muito afetadas. Para se ter melhores conclusões, um estudo de robustez mais detalhado seria indicado, entretanto este não era o objetivo deste trabalho.

Capítulo 5

Simulações para o Modelo de Equações Estruturais

Nos capítulos anteriores discutimos a incorporação de indicadores categóricos ordinais em modelagem de equações estruturais e verificamos que as correlações especiais são preferíveis à correlação de Pearson, uma vez que as correlações tetracórica, policórica, biserial e poli-serial medem com maior acurácia a associação entre variáveis quando existem informações sumarizadas em categorias no conjunto de dados.

Muitos pesquisadores utilizam escalas ordinais (Escala de Likert, por exemplo) no desenvolvimento de questionários com o objetivo de padronizar as respostas dos respondentes e, desta forma, facilitar a análise dos dados. O problema reside em considerar estas escalas como variáveis contínuas, ainda mais quando uma pequena quantidade de categorias é empregada. Existem pesquisadores que afirmam que os respondentes podem não compreender os níveis pareados como sendo equidistantes, além de que em alguns tipos de pesquisa os sujeitos pesquisados evitam categorias mais extremas ou preferem dar respostas mais favoráveis a si ou ao assunto de que a pesquisa está tratando.

Este capítulo tem por objetivo comparar os parâmetros estimados utilizando a correlação de Pearson e as correlações especiais quando o modelo é estimado pelas diferentes funções de discrepância apresentadas no Capítulo 1.

5.1 Metodologia Utilizada

A fim de responder a pergunta de qual função de discrepância e qual medida de associação deve ser utilizada em SEM quando existem variáveis categóricas ordinais, foi desenvolvida uma metodologia de simulação parecida com a de [Sanches \(2005\)](#).

Para estimarmos o modelo sob estudo utilizando as diferentes funções de discrepância, foram feitas modificações na função *sem*, disponível no pacote *sem* do software R. Foram incluídos parâmetros na função que indicam qual função de discrepância deve ser utilizada, qual matriz de pesos deve ser considerada (nos casos em que o modelo é ajustado por

WLS ou por DWLS) e qual é o coeficiente de curtose comum (no caso em que o modelo é ajustado por E). Utilizando a função *summary* é possível calcular as principais medidas de ajuste do modelo, entretanto é necessário analisar estes resultados com parcimônia. A função modificada está disponível no Apêndice B.

A ideia geral foi simular um conjunto de dados com uma determinada estrutura e estimar o modelo de equações estruturais com a matriz de correlações dos dados simulados originalmente para as diferentes funções de discrepância; este procedimento foi realizado com o intuito de determinarmos qual valor teríamos estimado para os parâmetros do modelo se indicadores ordinais não tivessem sido utilizados na análise. Feito isto, categorizamos algumas das variáveis e reestimamos o modelo com a matriz de correlação de Pearson (considerando os indicadores ordinais como contínuos) e com a matriz de correlações especiais. A função *cor_especiais*, disponibilizada no Apêndice A, calcula a matriz de correlações especiais a partir de um conjunto de dados.

A matriz de correlações especiais foi calculada utilizando as correlações especiais discutidas no capítulo anterior. Quando duas variáveis são dicotômicas, foi calculada a correlação tetracórica; quando duas variáveis são categóricas ordinais com mais de 2 categorias, foi calculada a correlação policórica; quando uma das variáveis é dicotômica e a outra é contínua, foi calculada a correlação biserial; e quando uma variável é categórica ordinal e a outra é contínua, foi calculada a correlação poliserial. Para este estudo de simulação a correlação policórica foi estimada pela aproximação Q_2 de Pearson e as demais correlações especiais foram estimadas pelo método de 2-passos.

Como o tempo computacional é grande, foram realizadas 1.000 simulações do modelo estudado e os parâmetros estimados foram sumarizados em uma tabela com a média e os desvios padrões.

Para verificarmos se o modelo é robusto a desvios de simetria e curtose, o estudo acima foi refeito assumindo que as variáveis (observadas e latentes) seguem uma distribuição t de Student e Gama, tal como foi feito no estudo de simulação das correlações especiais. Mesmo supondo que a distribuição das variáveis possa mudar, a estrutura de erros permaneceu inalterada, ou seja, supusemos que os erros eram provenientes da distribuição normal padrão.

5.2 Modelo 1

O primeiro modelo que avaliamos possui 3 indicadores formativos, 3 indicadores reflexivos e apenas uma variável latente; o diagrama de caminhos do modelo pode ser visualizado na Figura 5.1.

As variáveis X_1 , X_2 e X_3 e os erros ζ_1 , ϵ_1 , ϵ_2 e ϵ_3 da Figura 5.1 foram gerados independentemente supondo que os dados seguem uma distribuição normal padrão.

Os valores populacionais dos parâmetros foram: $\gamma_{11} = 0,5$, $\gamma_{12} = 1,0$, $\gamma_{13} = 1,5$, $\lambda_{21}^{(Y)} = 0,8$, $\lambda_{31}^{(Y)} = 1,3$ e $\lambda_{11}^{(Y)}$ foi fixado em 1 a fim de que o modelo fosse identificável.

Figura 5.1: Diagrama de caminhos do Modelo 1.

Os indicadores reflexivos Y_1 , Y_2 e Y_3 foram calculados utilizando os parâmetros populacionais, os indicadores formativos e os erros simulados.

Munido de todas as variáveis observadas, foram simuladas 1.000 amostras, cada uma com 500 observações, do diagrama de caminhos da Figura 5.1.

Inicialmente ajustamos o modelo considerando a matriz de correlação de Pearson calculada utilizando os dados originais (Coluna Original da Tabela 5.1). Estes valores não são iguais aos valores fixados para os parâmetros, e nem poderiam ser, pois o ajuste foi feito sobre a matriz de correlação e não sobre a matriz de covariância. São os valores de referência com os quais as outras estimativas devem ser comparadas.

Após o modelo ser ajustado com os dados originais, as variáveis X_1 , X_2 , Y_1 e Y_2 foram categorizadas. As variáveis X_1 e Y_1 foram consideradas como sendo dicotômicas e as variáveis X_2 e Y_2 foram divididas em 5 categorias; os limiares dicretizantes foram sorteados aleatoriamente entre -3 e 3. Este intervalo contém, em média, 87% das observações, pois as distribuições possuem variância diferente de 1.

Com as novas variáveis, o modelo foi ajustado novamente utilizando a matriz de correlações especiais (coluna Corr. Especial da Tabela 5.1) e a matriz de correlação de Pearson (coluna Corr. Pearson da Tabela 5.1), considerando as variáveis ordinais como contínuas.

A média e o desvio padrão (entre parênteses) dos parâmetros estimados, para cada uma das matrizes de correlação e função de discrepância utilizada, podem ser visualizados nas Tabelas. 5.1 e 5.2.

Tabela 5.1: Média e desvio padrão dos parâmetros estimados para o Modelo 1, obtidas no estudo de simulação da distribuição normal.

Parâmetros	Função de Discrepância					
	ML			GLS		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
γ_{11}	0,2129 (0,0214)	0,1062 (0,0297)	0,2058 (0,0485)	0,2132 (0,0217)	0,1067 (0,0298)	0,2153 (0,0547)
γ_{12}	0,4267 (0,0232)	0,2897 (0,0334)	0,4283 (0,0362)	0,4274 (0,0233)	0,2914 (0,0336)	0,4319 (0,0386)
γ_{13}	0,6390 (0,0228)	0,4959 (0,0327)	0,6392 (0,0352)	0,6399 (0,0228)	0,4985 (0,0333)	0,6408 (0,0368)
λ_{21}	0,9529 (0,0148)	1,1088 (0,0728)	0,9535 (0,0290)	0,9529 (0,0149)	1,1089 (0,0728)	0,9535 (0,0323)
λ_{31}	1,0554 (0,0136)	1,3600 (0,0848)	1,0586 (0,0361)	1,0552 (0,0136)	1,3568 (0,0847)	1,0574 (0,0389)
ψ_{11}	0,1812 (0,0161)	0,1475 (0,0221)	0,1819 (0,0283)	0,1806 (0,0161)	0,1468 (0,0219)	0,1784 (0,0283)
θ_{11}	0,1822 (0,0180)	0,5085 (0,0513)	0,1833 (0,0405)	0,1797 (0,0178)	0,4921 (0,0529)	0,1628 (0,0418)
θ_{22}	0,2575 (0,0227)	0,3988 (0,0570)	0,2583 (0,0347)	0,2537 (0,0225)	0,3855 (0,0549)	0,2424 (0,0353)
θ_{33}	0,0894 (0,0122)	0,0992 (0,0254)	0,0865 (0,0278)	0,0886 (0,0121)	0,0992 (0,0248)	0,0817 (0,0251)
Parâmetros	ULS			WLS		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
γ_{11}	0,2128 (0,0219)	0,1061 (0,0311)	0,1995 (0,0478)	0,2128 (0,0216)	0,1056 (0,0424)	0,1701 (0,0776)
γ_{12}	0,4270 (0,0234)	0,2899 (0,0348)	0,4276 (0,0369)	0,4268 (0,0233)	0,2876 (0,0455)	0,4284 (0,0550)
γ_{13}	0,6389 (0,0230)	0,4954 (0,0331)	0,6393 (0,0352)	0,6387 (0,0229)	0,4913 (0,0364)	0,6409 (0,0436)
λ_{21}	0,9529 (0,0155)	1,1098 (0,0742)	0,9584 (0,0313)	0,9531 (0,0150)	1,1117 (0,1185)	0,9936 (0,1671)
λ_{31}	1,0553 (0,0137)	1,3478 (0,0960)	1,0605 (0,0359)	1,0558 (0,0158)	1,3164 (0,1568)	1,0948 (0,1729)
ψ_{11}	0,1811 (0,0167)	0,1549 (0,0267)	0,1821 (0,0293)	0,1799 (0,0163)	0,1754 (0,0551)	0,1879 (0,2293)
θ_{11}	0,1822 (0,0184)	0,5013 (0,0572)	0,1862 (0,0420)	0,1853 (0,0193)	0,4855 (0,0738)	0,1830 (0,2414)
θ_{22}	0,2575 (0,0231)	0,3895 (0,0611)	0,2535 (0,0349)	0,2605 (0,0237)	0,3741 (0,0785)	0,2246 (0,0851)
θ_{33}	0,0896 (0,0123)	0,1042 (0,0290)	0,0867 (0,0245)	0,0912 (0,0151)	-0,0194 (0,1839)	0,0884 (0,2686)

Tabela 5.2: *Continuação da Tabela 5.1.*

Parâmetros	Função de Discrepância					
	DWLS			E		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
γ_{11}	0,2130 (0,0215)	0,1057 (0,0401)	0,1748 (0,0735)	0,2132 (0,0217)	0,1067 (0,0298)	0,2152 (0,0546)
γ_{12}	0,4268 (0,0234)	0,2868 (0,0430)	0,4268 (0,0514)	0,4274 (0,0233)	0,2913 (0,0336)	0,4318 (0,0386)
γ_{13}	0,6389 (0,0229)	0,4903 (0,0352)	0,6385 (0,0420)	0,6399 (0,0228)	0,4985 (0,0332)	0,6406 (0,0368)
λ_{21}	0,9530 (0,0153)	1,0994 (0,1153)	0,9933 (0,1703)	0,9529 (0,0149)	1,1089 (0,0728)	0,9536 (0,0326)
λ_{31}	1,0557 (0,0158)	1,3074 (0,1561)	1,0933 (0,1688)	1,0552 (0,0136)	1,3569 (0,0846)	1,0576 (0,0391)
ψ_{11}	0,1810 (0,0161)	0,1795 (0,0615)	0,1928 (0,2266)	0,1806 (0,0161)	0,1466 (0,0218)	0,1778 (0,0286)
θ_{11}	0,1824 (0,0189)	0,4822 (0,0808)	0,1820 (0,2469)	0,1797 (0,0178)	0,4914 (0,0530)	0,1623 (0,0419)
θ_{22}	0,2574 (0,0233)	0,3807 (0,0941)	0,2248 (0,1359)	0,2537 (0,0225)	0,3849 (0,0548)	0,2416 (0,0356)
θ_{33}	0,0892 (0,0143)	0,1340 (0,0786)	0,0621 (0,1332)	0,0885 (0,0121)	0,0990 (0,0247)	0,0813 (0,0250)

Analisando as Tabelas 5.1 e 5.2 notamos que as estimativas obtidas utilizando a matriz de correlação dos dados originais e as correlações especiais foram aproximadamente iguais, para todas as funções de discrepância analisadas. Quando o modelo foi ajustado utilizando a matriz de correlação de Pearson dos dados categorizados, os parâmetros variaram em relação ao que era esperado, independentemente da função de discrepância.

O desvio padrão das estimativas foi aproximadamente igual para todas as funções de discrepância, variando apenas entre as diferentes matrizes de correlação utilizadas para estimar o Modelo 1. Os desvios padrões apresentados pela matriz de correlações especiais foi um pouco maior do que a dos dados originais, exceto nos casos em que as funções WLS e DWLS foram utilizadas; nestes casos o desvio padrão foi muito maior que nas demais funções de discrepância.

Com o intuito de termos indicações sobre a robustez dos resultados quando existem desvios de simetria e curtose no conjunto de dados, o estudo de simulação anterior foi realizado mais duas vezes; o primeiro considerou que as variáveis X_1 , X_2 e X_3 seguiam uma distribuição t de Student com 3 graus de liberdade e o segundo considerou que elas seguiam uma distribuição Gama com parâmetro de forma igual a 3 e de escala igual a 2. Em ambos os casos não foi suposta nenhuma estrutura de covariância entre as variáveis e os erros do modelo (δ_1 , ϵ_1 , ϵ_2 e ϵ_3) foram considerados como sendo normalmente distribuídos.

As variáveis foram categorizadas da mesma forma que anteriormente. Quando os indicadores formativos seguiam uma distribuição t de Student, os limiares discretizantes foram sorteados aleatoriamente no intervalo -3 e 3 e quando as variáveis seguiam uma distribuição Gama, os limiares foram sorteados aleatoriamente no intervalo 0 e 4.

A média e o desvio padrão (entre parênteses) dos parâmetros estimados podem ser visualizados nas Tabelas 5.3 a 5.6.

Tabela 5.3: Média e desvio padrão dos parâmetros estimados para o Modelo 1, obtidas no estudo de simulação da distribuição t .

Parâmetros	Função de Discrepância					
	ML			GLS		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
γ_{11}	0,2490 (0,0397)	0,1160 (0,0288)	0,2203 (0,0464)	0,2492 (0,0398)	0,1144 (0,0238)	0,2207 (0,0682)
γ_{12}	0,4926 (0,0502)	0,2950 (0,0441)	0,4632 (0,0401)	0,4927 (0,0502)	0,2919 (0,0397)	0,4810 (0,0549)
γ_{13}	0,7193 (0,0406)	0,5374 (0,0817)	0,7314 (0,0433)	0,7198 (0,0406)	0,5327 (0,0286)	0,7424 (0,0489)
λ_{21}	0,9761 (0,0072)	1,0746 (0,1744)	0,9637 (0,0273)	0,9761 (0,0072)	1,1065 (0,0517)	0,9569 (0,0446)
λ_{31}	1,0249 (0,0065)	1,4085 (0,1812)	1,0190 (0,0308)	1,0248 (0,0065)	1,3821 (0,1009)	1,0044 (0,0462)
ψ_{11}	0,0864 (0,0125)	0,0364 (0,4221)	0,1060 (0,0447)	0,0861 (0,0125)	0,1050 (0,0225)	0,0893 (0,0418)
θ_{11}	0,0865 (0,0127)	0,5196 (0,0805)	0,0917 (0,0418)	0,0853 (0,0126)	0,4456 (0,0578)	0,0606 (0,0356)
θ_{22}	0,1296 (0,0177)	0,3991 (0,0842)	0,1601 (0,0787)	0,1277 (0,0175)	0,3487 (0,0611)	0,1075 (0,0447)
θ_{33}	0,0406 (0,0075)	-0,0050 (0,4581)	0,0586 (0,0409)	0,0401 (0,0075)	0,0772 (0,0422)	0,0406 (0,0204)
Parâmetros	ULS			WLS		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
γ_{11}	0,2492 (0,0398)	0,1276 (0,0248)	0,2206 (0,0451)	0,2502 (0,0399)	0,1349 (0,0390)	0,1866 (0,0648)
γ_{12}	0,4924 (0,0504)	0,3180 (0,0333)	0,4632 (0,0375)	0,4944 (0,0506)	0,3316 (0,0444)	0,4739 (0,0522)
γ_{13}	0,7193 (0,0407)	0,5156 (0,0266)	0,7298 (0,0394)	0,7221 (0,0412)	0,4807 (0,0341)	0,7403 (0,0451)
λ_{21}	0,9762 (0,0074)	1,1043 (0,0517)	0,9729 (0,0220)	0,9752 (0,0086)	1,1327 (0,1077)	1,0293 (0,0933)
λ_{31}	1,0249 (0,0066)	1,3073 (0,0830)	1,0279 (0,0261)	1,0170 (0,0120)	1,1738 (0,1456)	1,1147 (0,0979)
ψ_{11}	0,0864 (0,0127)	0,1389 (0,0240)	0,0962 (0,0367)	0,0849 (0,0126)	0,2118 (0,0613)	0,0230 (0,0703)
θ_{11}	0,0865 (0,0129)	0,4756 (0,0490)	0,1060 (0,0315)	0,0807 (0,0209)	0,4259 (0,0723)	0,1635 (0,0721)
θ_{22}	0,1294 (0,0178)	0,3616 (0,0622)	0,1543 (0,0298)	0,1259 (0,0283)	0,2878 (0,0874)	0,1247 (0,0723)
θ_{33}	0,0406 (0,0076)	0,1102 (0,0449)	0,0564 (0,0218)	0,0524 (0,0134)	-0,5139 (0,5799)	0,2605 (0,6167)

Tabela 5.4: *Continuação da Tabela 5.3.*

Parâmetros	Função de Discrepância					
	DWLS			E		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
γ_{11}	0,2494 (0,0398)	0,1345 (0,0371)	0,1917 (0,0612)	0,2222 (6,7420)	0,1200 (0,1919)	0,2485 (0,7995)
γ_{12}	0,4928 (0,0501)	0,3287 (0,0426)	0,4694 (0,0493)	0,6059 (8,0295)	0,2896 (0,0629)	0,4632 (0,4496)
γ_{13}	0,7194 (0,0407)	0,4755 (0,0345)	0,7348 (0,0424)	1,6442 (7,6921)	0,5287 (0,0763)	0,7555 (0,4877)
λ_{21}	0,9761 (0,0076)	1,1001 (0,1365)	1,0728 (0,1314)	21,6656 (158,5875)	1,1083 (0,0748)	0,9643 (0,0455)
λ_{31}	1,0248 (0,0079)	1,1426 (0,1612)	1,1521 (0,1297)	24,0501 (167,6531)	1,3877 (0,1084)	1,0073 (0,0559)
ψ_{11}	0,0859 (0,0125)	0,2319 (0,0769)	0,0042 (0,0889)	-348,3748 (1347,9213)	-0,1754 (8,7854)	-3,5937 (84,0760)
θ_{11}	0,0864 (0,0135)	0,4118 (0,0909)	0,1950 (0,1032)	-1077,7478 (3180,3987)	-8,1114 (270,2354)	-16,6121 (372,5773)
θ_{22}	0,1295 (0,0179)	0,2978 (0,1077)	0,0872 (0,1142)	-1312,1188 (3727,9241)	-5,7963 (194,0352)	-21,3025 (480,1412)
θ_{33}	0,0407 (0,0086)	0,2224 (0,0865)	-0,0207 (0,1051)	-456,2781 (1604,8036)	-1,6076 (53,2065)	-6,3245 (203,2715)

Tabela 5.5: Média e desvio padrão dos parâmetros estimados para o Modelo 1, obtidas no estudo de simulação da distribuição Gama.

Parâmetros	Função de Discrepância					
	ML			GLS		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
γ_{11}	0,2011 (0,0251)	0,0782 (0,0709)	0,2088 (0,1847)	0,2014 (0,0252)	0,0774 (0,0309)	0,2114 (0,0958)
γ_{12}	0,4033 (0,0278)	0,2077 (0,0935)	0,4036 (0,0554)	0,4041 (0,0280)	0,2065 (0,0563)	0,4103 (0,0644)
γ_{13}	0,6027 (0,0261)	0,3637 (0,1103)	0,6014 (0,0773)	0,6037 (0,0262)	0,3576 (0,0868)	0,6091 (0,0596)
λ_{21}	0,9440 (0,0191)	1,2699 (0,5249)	0,9330 (0,0684)	0,9441 (0,0192)	1,3192 (0,4479)	0,9304 (0,0730)
λ_{31}	1,0665 (0,0178)	1,9640 (0,9240)	1,0773 (0,0875)	1,0663 (0,0178)	1,9435 (0,6690)	1,0692 (0,0869)
ψ_{11}	0,2163 (0,0193)	0,0789 (1,1155)	0,1830 (0,2627)	0,2158 (0,0193)	0,1094 (0,0469)	0,1859 (0,0532)
θ_{11}	0,2166 (0,0223)	0,7171 (0,6631)	0,2254 (0,3325)	0,2138 (0,0221)	0,6537 (0,1193)	0,1928 (0,0683)
θ_{22}	0,3020 (0,0272)	0,5262 (0,3048)	0,3380 (0,0610)	0,2976 (0,0270)	0,5328 (0,0807)	0,2937 (0,0589)
θ_{33}	0,1095 (0,0164)	-0,0268 (1,1806)	0,1165 (0,1037)	0,1084 (0,0162)	0,0952 (0,0363)	0,1042 (0,0415)
Parâmetros	ULS			WLS		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
γ_{11}	0,2011 (0,0255)	0,0816 (0,0328)	0,1989 (0,0757)	0,2011 (0,0252)	0,0847 (0,0466)	0,1855 (0,1217)
γ_{12}	0,4034 (0,0280)	0,2115 (0,0563)	0,4009 (0,0565)	0,4035 (0,0280)	0,2103 (0,0665)	0,3937 (0,0885)
γ_{13}	0,6027 (0,0261)	0,3515 (0,0820)	0,6033 (0,0493)	0,6027 (0,0263)	0,3214 (0,0894)	0,6075 (0,0679)
λ_{21}	0,9442 (0,0198)	1,3076 (0,4267)	0,9427 (0,0551)	0,9438 (0,0193)	1,4486 (0,5982)	0,9355 (0,2692)
λ_{31}	1,0664 (0,0180)	1,8431 (0,6128)	1,0909 (0,0695)	1,0667 (0,0209)	1,6777 (0,6383)	1,0795 (0,3180)
ψ_{11}	0,2161 (0,0201)	0,1331 (0,0549)	0,1817 (0,0528)	0,2149 (0,0197)	0,2016 (0,0891)	0,1844 (0,3042)
θ_{11}	0,2167 (0,0228)	0,6813 (0,1276)	0,2494 (0,0688)	0,2191 (0,0241)	0,6378 (0,1491)	0,2316 (0,3080)
θ_{22}	0,3018 (0,0275)	0,5428 (0,0907)	0,3363 (0,0494)	0,3050 (0,0276)	0,4462 (0,1280)	0,3096 (0,1469)
θ_{33}	0,1098 (0,0165)	0,1114 (0,0415)	0,1132 (0,0417)	0,1112 (0,0193)	-0,4265 (0,3823)	0,1272 (0,5384)

Tabela 5.6: *Continuação da Tabela 5.5.*

Parâmetros	Função de Discrepância					
	DWLS			E		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
γ_{11}	0,2011 (0,0251)	0,0834 (0,0439)	0,1891 (0,1171)	0,2013 (0,0251)	0,0773 (0,0310)	0,2114 (0,0921)
γ_{12}	0,4036 (0,0280)	0,2037 (0,0643)	0,3979 (0,0846)	0,4040 (0,0280)	0,2059 (0,0564)	0,4103 (0,0618)
γ_{131}	0,6029 (0,0264)	0,3148 (0,0891)	0,6119 (0,0649)	0,6037 (0,0263)	0,3567 (0,0868)	0,6090 (0,0583)
λ_{21}	0,9438 (0,0196)	1,3729 (0,5277)	0,9419 (0,2853)	0,9440 (0,0192)	1,3204 (0,4493)	0,9323 (0,0705)
λ_{31}	1,0665 (0,0204)	1,5646 (0,5393)	1,0937 (0,2306)	1,0662 (0,0178)	1,9495 (0,6745)	1,0700 (0,0820)
ψ_{11}	0,2160 (0,0193)	0,2084 (0,0930)	0,2516 (1,2775)	0,2158 (0,0193)	0,1079 (0,0466)	0,1829 (0,0525)
θ_{11}	0,2165 (0,0234)	0,6304 (0,1515)	0,1628 (1,2776)	0,2136 (0,0220)	0,6458 (0,1180)	0,1893 (0,0682)
θ_{22}	0,3022 (0,0277)	0,4310 (0,1326)	0,3492 (0,2072)	0,2975 (0,0270)	0,5263 (0,0799)	0,2883 (0,0619)
θ_{33}	0,1094 (0,0186)	0,2501 (0,1264)	0,1216 (0,1686)	0,1083 (0,0162)	0,0935 (0,0360)	0,1023 (0,0401)

Analisando as Tabelas 5.3 e 5.4, notamos que as estimativas obtidas utilizando os dados originais e as correlações especiais foram aproximadamente iguais, com uma diferença maior quando as funções de discrepância WLS e DWLS foram utilizadas.

Da mesma forma que anteriormente, a matriz de correlação de Pearson das variáveis categóricas não estimou corretamente os parâmetros do Modelo 1.

A função de discrepância E apresentou estimativas muito diferentes entre si e em comparação com as demais funções de discrepância. Analisando os resultados observamos que alguns erros foram simulados com valores negativos (o que é incoerente com a definição de SEM) e com desvios padrões muito grandes; esta situação ocorreu com as três matrizes de correlações utilizadas.

Em comparação com os parâmetros estimados na Tabela 5.1, observamos poucas diferenças entre os parâmetros (com exceção dos obtidos quando o modelo foi estimado por E) e um leve aumento dos desvios padrões.

Quando os indicadores formativos foram gerados seguindo uma distribuição Gama, as estimativas obtidas utilizando os dados originais e as correlações especiais foram aproximadamente iguais; os desvios padrões foram bem maiores quando foi utilizada a matriz de correlações especiais; e a correlação de Pearson dos dados categorizados subestimou os parâmetros do modelo estudado.

De uma forma geral os parâmetros estimados foram um pouco menores do que quando os dados foram gerados seguindo uma distribuição normal padrão ou t de Student.

5.3 Modelo 2

O segundo modelo estudado é uma análise fatorial confirmatória com dois constructos e oito indicadores; o diagrama de caminhos do modelo pode ser visualizado na Figura 5.2.

Figura 5.2: Diagrama de caminhos do Modelo 2.

Neste modelo as variáveis e os erros dos indicadores foram simulados independentemente, supondo que os dados seguem uma distribuição normal padrão.

Os valores populacionais dos parâmetros foram: $\lambda_{11}^{(X)} = 0,3$, $\lambda_{21}^{(X)} = 0,6$, $\lambda_{31}^{(X)} = 0,9$, $\lambda_{41}^{(X)} = 1,2$, $\lambda_{52}^{(X)} = 0,7$, $\lambda_{62}^{(X)} = -1,8$, $\lambda_{72}^{(X)} = -0,5$ e $\lambda_{82}^{(X)} = 0,2$.

Os indicadores da análise fatorial confirmatória foram calculados utilizando os parâmetros populacionais, os erros e os valores simulados para os constructos.

O procedimento foi análogo ao do Modelo 1, ou seja, foram geradas 1.000 amostras com 500 observações e o modelo foi estimado utilizando as mesmas funções de discrepância anteriores.

Após o modelo ser estimado com as variáveis originais, as variáveis X_1 , X_2 , X_4 , X_5 , X_6 e X_8 foram categorizadas. As variáveis X_1 e X_5 foram consideradas como sendo dicotômicas e as demais foram divididas em 5 categorias; os limiares discretizantes foram sorteados aleatoriamente entre -3 e 3.

Com o novo conjunto de dados, o modelo foi estimado utilizando a matriz de correlações especiais e a matriz de correlação de Pearson, considerando as variáveis categóricas ordinais como contínuas. A média e o desvio padrão (entre parênteses) dos parâmetros estimados podem ser visualizados nas Tabelas 5.7 a 5.9.

Tabela 5.7: Média e desvio padrão dos parâmetros estimados para o Modelo 2, obtidas no estudo de simulação da distribuição normal.

Parâmetros	Função de Discrepância					
	ML			GLS		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
$\lambda_{11}^{(X)}$	0,2871 (0,0514)	0,1761 (0,0739)	0,2989 (0,1265)	0,2850 (0,0515)	0,1755 (0,0744)	0,3001 (0,3620)
$\lambda_{21}^{(X)}$	0,5140 (0,0445)	0,4326 (0,0703)	0,5067 (0,1210)	0,5111 (0,0447)	0,4299 (0,0710)	0,5124 (0,2705)
$\lambda_{31}^{(X)}$	0,6682 (0,0442)	0,6686 (0,0592)	0,6581 (0,1562)	0,6638 (0,0444)	0,6641 (0,0607)	0,6814 (0,4325)
$\lambda_{41}^{(X)}$	0,7667 (0,0476)	0,6731 (0,0784)	0,7757 (0,2714)	0,7613 (0,0476)	0,6683 (0,0788)	0,7558 (0,2037)
$\lambda_{52}^{(X)}$	-0,5700 (0,0543)	-0,3525 (0,1146)	-0,5574 (0,4729)	-0,5651 (0,0508)	-0,3502 (0,1140)	-0,5521 (0,5703)
$\lambda_{62}^{(X)}$	0,8830 (0,0640)	0,8039 (0,1901)	0,9294 (0,5892)	0,8750 (0,0627)	0,7987 (0,2194)	0,9098 (0,6931)
$\lambda_{72}^{(X)}$	0,4451 (0,0475)	0,4452 (0,0729)	0,4199 (0,2559)	0,4420 (0,0484)	0,4431 (0,0741)	0,4322 (0,3424)
$\lambda_{82}^{(X)}$	-0,1969 (0,0568)	-0,1571 (0,0600)	-0,1991 (0,2246)	-0,1948 (0,0483)	-0,1564 (0,0608)	-0,1526 (0,4663)
δ_1	0,9150 (0,0295)	0,9635 (0,0254)	0,8951 (0,0926)	0,8938 (0,0304)	0,9404 (0,0285)	0,6476 (1,8214)
δ_2	0,7330 (0,0496)	0,8079 (0,0580)	0,7260 (0,1320)	0,7181 (0,0454)	0,7907 (0,0582)	0,5952 (1,9475)
δ_3	0,5506 (0,0625)	0,5494 (0,0798)	0,5326 (0,2642)	0,5402 (0,0583)	0,5383 (0,0807)	0,3027 (3,3023)
δ_4	0,4090 (0,0739)	0,5408 (0,1033)	0,3135 (1,5315)	0,4017 (0,0716)	0,5302 (0,1020)	0,3344 (0,9575)
δ_5	0,6731 (0,0579)	0,8626 (0,0767)	0,5541 (1,9609)	0,6596 (0,0574)	0,8447 (0,0758)	0,2662 (4,0961)
δ_6	0,2175 (0,1129)	0,3176 (0,8309)	-0,0838 (3,0705)	0,2143 (0,1111)	0,2975 (1,2251)	-0,3543 (5,4626)
δ_7	0,7994 (0,0431)	0,7965 (0,0647)	0,7970 (0,0982)	0,7831 (0,0427)	0,7799 (0,0645)	0,6515 (2,8848)
δ_8	0,9593 (0,0189)	0,9717 (0,0193)	0,9551 (0,0634)	0,9370 (0,0222)	0,9495 (0,0219)	0,6831 (3,3290)

Tabela 5.8: *Continuação da Tabela 5.7.*

Parâmetros	Função de Discrepância					
	ULS			WLS		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
$\lambda_{11}^{(X)}$	0,2873 (0,0508)	0,1761 (0,0751)	0,2952 (0,1123)	0,2871 (0,0507)	0,1804 (0,2706)	0,3333 (0,6572)
$\lambda_{21}^{(X)}$	0,5147 (0,0440)	0,4322 (0,0767)	0,5171 (0,0683)	0,5145 (0,0437)	0,4453 (0,7026)	0,5476 (0,6883)
$\lambda_{31}^{(X)}$	0,6685 (0,0440)	0,6660 (0,0804)	0,6676 (0,0699)	0,6685 (0,0435)	0,5873 (0,5328)	0,6219 (0,6255)
$\lambda_{41}^{(X)}$	0,7663 (0,0465)	0,6711 (0,0904)	0,7675 (0,0801)	0,7669 (0,0466)	0,6067 (0,5015)	0,7340 (0,5866)
$\lambda_{52}^{(X)}$	-0,5707 (0,0507)	-0,3539 (0,1143)	-0,5574 (0,1843)	-0,5697 (0,0504)	-0,3637 (0,2837)	-0,4415 (0,8204)
$\lambda_{62}^{(X)}$	0,8804 (0,0618)	0,7998 (0,2213)	0,8919 (0,4646)	0,8821 (0,0616)	0,7133 (0,2895)	0,5783 (0,6241)
$\lambda_{72}^{(X)}$	0,4456 (0,0473)	0,4470 (0,0717)	0,4346 (0,1111)	0,4454 (0,0472)	0,4017 (0,1676)	0,2855 (0,3163)
$\lambda_{82}^{(X)}$	-0,1965 (0,0487)	-0,1580 (0,0603)	-0,1854 (0,1001)	-0,1959 (0,0479)	-0,1443 (0,2297)	-0,0804 (0,2053)
δ_1	0,9149 (0,0292)	0,9633 (0,0254)	0,9003 (0,0719)	0,9150 (0,0292)	0,8943 (0,6354)	0,4574 (2,6686)
δ_2	0,7331 (0,0453)	0,8073 (0,0580)	0,7279 (0,0719)	0,7334 (0,0449)	0,3085 (8,2085)	0,2268 (7,3427)
δ_3	0,5511 (0,0586)	0,5500 (0,0790)	0,5494 (0,1027)	0,5512 (0,0579)	0,3713 (6,5516)	0,2223 (11,5366)
δ_4	0,4106 (0,0716)	0,5414 (0,1024)	0,4045 (0,1271)	0,4098 (0,0718)	0,3808 (5,1350)	0,1174 (7,1610)
δ_5	0,6717 (0,0579)	0,8617 (0,0767)	0,6553 (0,1409)	0,6729 (0,0574)	0,7873 (0,7010)	0,1326 (3,5097)
δ_6	0,2211 (0,1104)	0,3115 (1,3605)	-0,0111 (4,5622)	0,2182 (0,1104)	0,4079 (0,2772)	0,2768 (0,4455)
δ_7	0,7992 (0,0421)	0,7951 (0,0639)	0,7988 (0,0652)	0,7994 (0,0421)	0,8104 (0,0799)	0,8186 (0,0841)
δ_8	0,9590 (0,0195)	0,9714 (0,0195)	0,9556 (0,0351)	0,9593 (0,0191)	0,9264 (1,3178)	0,9514 (0,0452)

Tabela 5.9: *Continuação da Tabela 5.7.*

Parâmetros	Função de Discrepância					
	DWLS			E		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
$\lambda_{11}^{(X)}$	0,2864 (0,0539)	0,1702 (0,3620)	0,2635 (1,4285)	0,2847 (0,0514)	0,4898 (18,2631)	-0,0740 (19,6596)
$\lambda_{21}^{(X)}$	0,5127 (0,0479)	0,3842 (0,4906)	0,4959 (1,0216)	0,5104 (0,0446)	-0,3012 (13,8729)	0,0161 (18,0464)
$\lambda_{31}^{(X)}$	0,6695 (0,0519)	0,5492 (0,3040)	0,5156 (0,5744)	0,6630 (0,0443)	-1,2100 (16,3354)	0,8645 (19,1351)
$\lambda_{41}^{(X)}$	0,7648 (0,0573)	0,6026 (0,5660)	0,6156 (0,8902)	0,7604 (0,0475)	0,0569 (15,8024)	3,6227 (29,6793)
$\lambda_{52}^{(X)}$	-0,5696 (0,0595)	-0,3139 (0,5514)	-0,3838 (1,4055)	-0,5645 (0,0508)	-0,3126 (17,6211)	-0,1339 (24,3046)
$\lambda_{62}^{(X)}$	0,8851 (0,0851)	0,6282 (0,5688)	0,4666 (0,9365)	0,8739 (0,0626)	1,3033 (17,3014)	3,5468 (27,0315)
$\lambda_{72}^{(X)}$	0,4438 (0,0510)	0,4077 (0,7099)	0,2420 (0,4379)	0,4415 (0,0483)	0,0146 (14,7248)	-0,9963 (19,0875)
$\lambda_{82}^{(X)}$	-0,1948 (0,0500)	-0,1212 (0,4839)	0,0123 (1,7625)	-0,1946 (0,0483)	0,5244 (16,5627)	-0,8619 (18,5512)
δ_1	0,9282 (0,0366)	0,8147 (0,7374)	-1,1905 (12,0108)	0,8916 (0,0305)	-1445,5244 (3159,2527)	-993,5816 (2614,1322)
δ_2	0,7462 (0,0534)	0,5274 (3,8036)	-0,4937 (9,6259)	0,7163 (0,0454)	-1050,4135 (2371,7366)	-963,8715 (2685,3272)
δ_3	0,5578 (0,0721)	0,4916 (0,7355)	0,2711 (3,9021)	0,5388 (0,0582)	-948,8331 (2822,3889)	-883,6974 (2839,7235)
δ_4	0,4198 (0,0910)	0,2361 (3,5673)	-0,2479 (8,8146)	0,4007 (0,0714)	-880,8772 (2854,5518)	-1146,6831 (4996,1600)
δ_5	0,6811 (0,0702)	0,5807 (2,1102)	-1,1593 (9,2884)	0,6580 (0,0573)	-1581,4520 (3470,3026)	-1046,6333 (3692,9491)
δ_6	0,2142 (0,1610)	0,2425 (4,0344)	-0,0224 (5,4849)	0,2138 (0,1109)	-1197,9852 (3256,5527)	-1247,7602 (4780,9123)
δ_7	0,8119 (0,0488)	0,2694 (6,1004)	0,6223 (1,9590)	0,7812 (0,0426)	-1084,0077 (2774,9132)	-1084,1319 (2876,8878)
δ_8	0,9731 (0,0274)	0,6884 (3,2771)	-2,5563 (71,9221)	0,9347 (0,0224)	-1448,9191 (3160,3074)	-1137,5318 (3009,8481)

Analisando as Tabelas 5.7 a 5.9 notamos que quando o modelo foi ajustado com os dados originais, os parâmetros estimados foram aproximadamente iguais, independentemente da função de discrepância utilizada. Os desvios padrões das estimativas também variaram pouco.

Quando o modelo foi ajustado utilizando a matriz de correlação de Pearson dos dados categorizados, observamos que os métodos ML, GLS e ULS apresentam estimativas dos parâmetros parecidas, WLS e DWLS apresentam estimativas que se desviaram um pouco dos parâmetros e E não conseguiu estimar, na média, os parâmetros do modelo. De maneira geral os parâmetros foram menores do que os obtidos quando o modelo foi ajustado utilizando a matriz de correlação dos dados originais.

Já quando o modelo foi ajustado utilizando a matriz de correlações especiais notamos que os métodos ML, GLS e ULS apresentaram estimativas muito parecidas entre si e em comparação aos parâmetros estimados com os dados originais; as estimativas obtidas por WLS e DWLS variaram um pouco em relação aos valores anteriores; e E não conseguiu estimar os parâmetros do modelo. O desvio padrão das estimativas obtidas por ULS foi muito menor do que as estimativas obtidas pelas demais funções de discrepância.

Para ter indicação se o modelo estudado é robusto a desvios de assimetria e curtose, o estudo de simulação foi realizado novamente supondo que os constructos seguissem uma distribuição t de Student com 3 graus de liberdade e depois uma distribuição Gama com parâmetro de forma igual a 3 e parâmetro de escala igual a 2. As demais condições não foram alteradas.

A média e o desvio padrão das estimativas obtidas podem ser visualizados nas Tabelas 5.10 a 5.15.

Tabela 5.10: Média e desvio padrão dos parâmetros estimados para o Modelo 2, obtidas no estudo de simulação da distribuição t de Student.

Parâmetros	Função de Discrepância					
	ML			GLS		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
$\lambda_{11}^{(X)}$	0,5617 (0,5855)	0,2677 (0,0703)	0,4762 (0,1885)	0,4479 (0,0624)	0,2663 (0,0703)	0,4407 (0,2061)
$\lambda_{21}^{(X)}$	0,7215 (0,7503)	0,5821 (0,0558)	0,6473 (0,3490)	0,7030 (0,0497)	0,5783 (0,0560)	0,6509 (0,0821)
$\lambda_{31}^{(X)}$	0,7012 (0,6827)	0,7909 (0,0582)	0,8030 (0,2669)	0,8260 (0,0361)	0,7857 (0,0578)	0,8124 (0,1267)
$\lambda_{41}^{(X)}$	0,8089 (0,4900)	0,7210 (0,0581)	0,8062 (0,3334)	0,8870 (0,0287)	0,7168 (0,0576)	0,8148 (0,0861)
$\lambda_{52}^{(X)}$	-0,8053 (0,9390)	-0,4741 (0,0798)	-0,7350 (0,4453)	-0,6640 (0,3548)	-0,4708 (0,0796)	-0,7129 (0,4408)
$\lambda_{62}^{(X)}$	1,0338 (2,0911)	0,7633 (0,0866)	0,9074 (0,3451)	0,8391 (0,4262)	0,7579 (0,0864)	0,8373 (0,2726)
$\lambda_{72}^{(X)}$	0,7342 (1,4168)	0,6066 (0,0754)	0,6303 (0,2868)	0,5567 (0,3088)	0,6029 (0,0755)	0,5889 (0,1614)
$\lambda_{82}^{(X)}$	-0,3952 (0,6930)	-0,2468 (0,0622)	-0,3530 (0,3977)	-0,2690 (0,1714)	-0,2457 (0,0626)	-0,2788 (0,2017)
δ_1	0,7813 (0,6978)	0,9234 (0,0356)	0,7604 (0,2505)	0,7739 (0,0611)	0,8972 (0,0398)	0,6308 (0,5756)
δ_2	0,4598 (1,9851)	0,6575 (0,0658)	0,4993 (0,4396)	0,4851 (0,0706)	0,6425 (0,0624)	0,5054 (0,0832)
δ_3	0,2046 (1,0466)	0,3704 (0,0906)	0,2252 (0,4872)	0,3000 (0,0584)	0,3627 (0,0867)	0,2802 (0,4709)
δ_4	0,1127 (0,9543)	0,4764 (0,0829)	0,2206 (0,4247)	0,1967 (0,0493)	0,4649 (0,0787)	0,2791 (0,0962)
δ_5	0,2520 (4,6695)	0,7688 (0,0724)	0,3510 (0,5839)	0,4159 (0,0682)	0,7523 (0,0709)	0,2089 (2,9058)
δ_6	-0,2862 (1,8542)	0,4099 (0,1286)	0,1510 (0,5488)	0,0983 (0,0476)	0,4010 (0,1262)	0,1724 (1,0001)
δ_7	0,4164 (1,6625)	0,6263 (0,0960)	0,5786 (0,3342)	0,5769 (0,0705)	0,6118 (0,0943)	0,5760 (0,4492)
δ_8	0,8718 (0,5602)	0,9352 (0,0311)	0,9034 (0,2379)	0,8763 (0,0408)	0,9110 (0,0337)	0,7979 (0,7791)

Tabela 5.11: *Continuação da Tabela 5.10.*

Parâmetros	Função de Discrepância					
	ULS			WLS		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
$\lambda_{11}^{(X)}$	0,4515 (0,0627)	0,2660 (0,0698)	0,4456 (0,1144)	0,4517 (0,0696)	0,2550 (0,4290)	0,4401 (1,2300)
$\lambda_{21}^{(X)}$	0,7085 (0,0494)	0,5822 (0,0559)	0,6682 (0,0554)	0,7092 (0,0610)	0,5124 (0,3965)	0,5702 (0,5582)
$\lambda_{31}^{(X)}$	0,8321 (0,0358)	0,7923 (0,0571)	0,8280 (0,0503)	0,8311 (0,0536)	0,6102 (0,3711)	0,5974 (0,7920)
$\lambda_{41}^{(X)}$	0,8936 (0,0282)	0,7205 (0,0577)	0,8257 (0,0476)	0,8937 (0,0512)	0,6459 (0,5886)	0,7118 (0,6778)
$\lambda_{52}^{(X)}$	-0,7572 (0,0456)	-0,4725 (0,0781)	-0,7188 (0,1443)	-0,7563 (0,0607)	-0,3351 (0,5240)	-0,1415 (1,3849)
$\lambda_{62}^{(X)}$	0,9476 (0,0254)	0,7610 (0,0856)	0,8637 (0,1196)	0,9508 (0,0589)	0,5428 (0,4898)	0,3529 (0,9676)
$\lambda_{72}^{(X)}$	0,6395 (0,0545)	0,6086 (0,0732)	0,6111 (0,0943)	0,6384 (0,0636)	0,4065 (0,3857)	0,2290 (0,6005)
$\lambda_{82}^{(X)}$	-0,3163 (0,0580)	-0,2502 (0,0621)	-0,2963 (0,0902)	-0,3160 (0,0642)	-0,1059 (0,5560)	0,0686 (1,0584)
δ_1	0,7922 (0,0615)	0,9244 (0,0353)	0,7883 (0,1154)	0,8038 (0,0780)	0,7224 (1,2553)	-0,8058 (7,4535)
δ_2	0,4956 (0,0721)	0,6580 (0,0637)	0,5504 (0,0765)	0,5038 (0,0967)	0,5082 (2,0484)	0,2967 (3,8168)
δ_3	0,3063 (0,0601)	0,3690 (0,0872)	0,3120 (0,0834)	0,3122 (0,0986)	0,0095 (1,1803)	-0,6330 (7,5370)
δ_4	0,2006 (0,0504)	0,4776 (0,0813)	0,3160 (0,0790)	0,2059 (0,1017)	0,1493 (5,3125)	-0,0655 (6,7363)
δ_5	0,4246 (0,0699)	0,7706 (0,0706)	0,4625 (0,1542)	0,4328 (0,1026)	0,5970 (1,3206)	-0,9775 (9,7028)
δ_6	0,1013 (0,0480)	0,4135 (0,1268)	0,2397 (0,1192)	0,0966 (0,1209)	0,4222 (0,3246)	-0,0811 (5,1366)
δ_7	0,5881 (0,0718)	0,6242 (0,0930)	0,6177 (0,0733)	0,5964 (0,0904)	0,5221 (0,3620)	0,3716 (2,9499)
δ_8	0,8966 (0,0404)	0,9335 (0,0312)	0,9041 (0,0476)	0,9085 (0,0584)	0,6431 (3,5532)	-0,1592 (11,7001)

Tabela 5.12: *Continuação da Tabela 5.10.*

Parâmetros	Função de Discrepância					
	DWLS			E		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
$\lambda_{11}^{(X)}$	0,4514 (0,0625)	0,2540 (0,0933)	0,4411 (0,2110)	10,3440 (93,4139)	0,4912 (17,1907)	0,7649 (33,7169)
$\lambda_{21}^{(X)}$	0,7084 (0,0492)	0,5737 (0,2554)	0,6709 (0,3550)	16,5541 (100,2822)	-0,0050 (17,6113)	3,1823 (35,0556)
$\lambda_{31}^{(X)}$	0,8322 (0,0355)	0,7912 (0,2219)	0,8211 (0,1554)	16,5721 (93,0957)	-0,4129 (18,4153)	5,3199 (40,2090)
$\lambda_{41}^{(X)}$	0,8937 (0,0282)	0,6949 (0,2035)	0,8163 (0,1671)	18,0336 (97,3976)	0,2781 (20,1321)	5,1471 (37,0106)
$\lambda_{52}^{(X)}$	-0,7571 (0,0455)	-0,3537 (0,3299)	-0,2371 (0,7361)	-13,5877 (91,1008)	-0,8044 (21,1793)	-3,4643 (36,3845)
$\lambda_{62}^{(X)}$	0,9483 (0,0256)	0,5996 (0,4823)	0,3814 (0,7882)	9,8799 (77,6911)	1,4313 (20,9351)	8,1309 (42,5993)
$\lambda_{72}^{(X)}$	0,6392 (0,0543)	0,4640 (0,3907)	0,2502 (0,5573)	10,1766 (76,8471)	0,7182 (17,1747)	4,4688 (34,2821)
$\lambda_{82}^{(X)}$	-0,3159 (0,0575)	-0,1689 (0,1929)	-0,0716 (0,2990)	2,9495 (85,0924)	0,6589 (17,9932)	-0,3854 (25,2031)
δ_1	0,7923 (0,0612)	0,9268 (0,0371)	0,7610 (0,7006)	-275,9230 (4221,9313)	-1408,3798 (3084,4420)	-640,6320 (3101,9455)
δ_2	0,4957 (0,0718)	0,6056 (1,4484)	0,4240 (4,1330)	62,0367 (3342,1743)	-1153,4256 (2535,0677)	-700,7701 (2888,5585)
δ_3	0,3062 (0,0596)	0,3243 (0,2009)	0,3027 (0,1941)	-14,2476 (2627,5962)	-932,9037 (2470,6983)	-660,4108 (3572,3468)
δ_4	0,2005 (0,0503)	0,4758 (0,2365)	0,3058 (0,2654)	-401,9328 (5242,4663)	-936,7757 (2180,8735)	-521,0407 (2519,6808)
δ_5	0,4248 (0,0697)	0,7662 (0,0786)	0,4023 (1,4488)	-212,1254 (2752,8287)	-1658,1469 (3627,1039)	-665,9690 (3744,0742)
δ_6	0,1001 (0,0485)	0,4088 (0,1750)	0,2345 (0,1933)	-460,5798 (5939,7843)	-1160,7186 (2909,3052)	-869,9195 (4388,4193)
δ_7	0,5885 (0,0716)	0,6318 (0,1094)	0,6269 (0,0963)	-157,0704 (3616,6672)	-1143,2286 (2602,3667)	-813,7873 (3070,1027)
δ_8	0,8969 (0,0400)	0,9342 (0,0411)	0,9056 (0,0572)	-371,0539 (4307,0247)	-1493,2321 (3257,8586)	-870,5017 (3483,9997)

Tabela 5.13: Média e desvio padrão dos parâmetros estimados para o Modelo 2, obtidas no estudo de simulação da distribuição Gama.

Parâmetros	Função de Discrepância					
	ML			GLS		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
$\lambda_{11}^{(X)}$	0,2502 (0,0542)	0,1521 (0,0774)	0,2534 (0,1297)	0,2490 (0,0548)	0,1514 (0,0762)	0,2604 (0,3860)
$\lambda_{21}^{(X)}$	0,4610 (0,0495)	0,3730 (0,0870)	0,4480 (0,2178)	0,4581 (0,0493)	0,3703 (0,0855)	0,4553 (0,4113)
$\lambda_{31}^{(X)}$	0,6159 (0,0519)	0,6399 (0,2866)	0,6336 (0,3281)	0,6120 (0,0523)	0,6383 (0,3328)	0,6514 (0,6072)
$\lambda_{41}^{(X)}$	0,7218 (0,0559)	0,5255 (0,1450)	0,6652 (0,3285)	0,7163 (0,0558)	0,5208 (0,1420)	0,6776 (0,6334)
$\lambda_{52}^{(X)}$	-0,5145 (0,0580)	-0,3073 (0,2600)	-0,4185 (0,6650)	-0,5109 (0,0587)	-0,3058 (0,2241)	-0,4467 (0,8103)
$\lambda_{62}^{(X)}$	0,8501 (0,0812)	0,5979 (0,4598)	0,7716 (0,8032)	0,8426 (0,0815)	0,5949 (0,5977)	0,7667 (1,1381)
$\lambda_{72}^{(X)}$	0,3964 (0,0549)	0,3849 (0,2138)	0,3670 (0,3801)	0,3940 (0,0551)	0,3845 (0,4175)	0,4014 (0,7359)
$\lambda_{82}^{(X)}$	-0,1690 (0,0529)	-0,1298 (0,1497)	-0,1404 (0,2834)	-0,1682 (0,0538)	-0,1261 (0,0937)	-0,1129 (0,4280)
δ_1	0,9345 (0,0271)	0,9702 (0,0441)	0,9165 (0,1117)	0,9120 (0,0283)	0,9484 (0,0271)	0,6621 (2,4080)
δ_2	0,7851 (0,0456)	0,8530 (0,0659)	0,7488 (1,3569)	0,7683 (0,0453)	0,8359 (0,0606)	0,5567 (3,2180)
δ_3	0,6180 (0,0636)	0,5094 (1,3656)	0,5076 (1,7958)	0,6050 (0,0635)	0,4651 (2,1305)	0,1637 (5,1623)
δ_4	0,4759 (0,0817)	0,7039 (0,1393)	0,4503 (2,0134)	0,4669 (0,0806)	0,6904 (0,1358)	0,0781 (5,4468)
δ_5	0,7319 (0,0591)	0,8471 (0,7883)	0,4148 (3,1777)	0,7167 (0,0589)	0,8372 (0,7984)	0,0513 (5,8915)
δ_6	0,2707 (0,1429)	0,4320 (1,9034)	-0,2232 (4,6260)	0,2668 (0,1417)	0,2721 (3,2607)	-0,9477 (9,0156)
δ_7	0,8398 (0,0431)	0,8117 (0,5622)	0,7289 (2,7418)	0,8219 (0,0437)	0,6591 (2,3572)	0,2523 (6,9517)
δ_8	0,9686 (0,0184)	0,9759 (0,0257)	0,9575 (0,1388)	0,9459 (0,0221)	0,9533 (0,0259)	0,7335 (3,7526)

Tabela 5.14: *Continuação da Tabela 5.13.*

Parâmetros	Função de Discrepância					
	ULS			WLS		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
$\lambda_{11}^{(X)}$	0,2503 (0,0542)	0,0372 (0,1658)	0,2481 (0,1375)	0,2493 (0,0571)	0,1312 (0,4563)	0,2954 (1,2469)
$\lambda_{21}^{(X)}$	0,4610 (0,0496)	0,0955 (0,3713)	0,4568 (0,4402)	0,4596 (0,0539)	0,3745 (0,6729)	0,4496 (1,5774)
$\lambda_{31}^{(X)}$	0,6164 (0,0517)	0,1035 (0,7036)	0,6173 (0,3527)	0,6164 (0,0579)	0,5128 (0,4886)	0,5119 (0,9159)
$\lambda_{41}^{(X)}$	0,7212 (0,0554)	0,1634 (0,5185)	0,6433 (0,3464)	0,7208 (0,0637)	0,5023 (0,6922)	0,6997 (1,4151)
$\lambda_{52}^{(X)}$	-0,5157 (0,0575)	-0,3075 (0,2685)	-0,4313 (0,6298)	-0,5137 (0,0702)	-0,3145 (0,6988)	-0,4351 (1,5301)
$\lambda_{62}^{(X)}$	0,8471 (0,0796)	0,5838 (0,5155)	0,7731 (0,9876)	0,8493 (0,1105)	0,4437 (0,7752)	0,4955 (1,2874)
$\lambda_{72}^{(X)}$	0,3972 (0,0550)	0,3826 (0,2218)	0,3656 (0,3261)	0,3947 (0,0615)	0,2856 (0,4299)	0,2348 (0,3408)
$\lambda_{82}^{(X)}$	-0,1703 (0,0542)	-0,1276 (0,0923)	-0,1349 (0,1513)	-0,1685 (0,0573)	-0,0941 (0,4113)	-0,0236 (0,9115)
δ_1	0,9344 (0,0272)	0,9711 (0,0242)	0,9196 (0,0846)	0,9477 (0,0336)	0,7475 (1,2860)	-0,7281 (7,2162)
δ_2	0,7850 (0,0457)	0,8531 (0,0604)	0,5978 (4,4697)	0,7964 (0,0529)	0,3569 (4,5976)	-1,7622 (30,6221)
δ_3	0,6174 (0,0635)	0,4947 (1,9854)	0,4946 (2,6147)	0,6260 (0,0743)	0,4029 (2,8604)	-0,2651 (10,2659)
δ_4	0,4768 (0,0808)	0,7047 (0,1360)	0,4663 (3,3658)	0,4846 (0,0945)	0,2175 (4,6943)	-1,5956 (12,9116)
δ_5	0,7307 (0,0588)	0,8334 (1,2327)	0,4177 (5,1716)	0,7412 (0,0665)	0,3970 (2,7391)	-1,5672 (9,1770)
δ_6	0,2761 (0,1388)	0,3937 (2,8426)	-0,5722 (8,7513)	0,2723 (0,1858)	0,1369 (4,2335)	-1,3379 (14,7506)
δ_7	0,8392 (0,0434)	0,8044 (0,8706)	0,7601 (2,6069)	0,8517 (0,0481)	0,6854 (3,9355)	0,7988 (1,0444)
δ_8	0,9681 (0,0189)	0,9752 (0,0227)	0,9589 (0,0573)	0,9806 (0,0247)	0,7407 (2,0826)	-0,1270 (11,9113)

Tabela 5.15: *Continuação da Tabela 5.13.*

Parâmetros	Função de Discrepância					
	DWLS			E		
	Original	Corr. Pearson	Corr. Especial	Original	Corr. Pearson	Corr. Especial
$\lambda_{11}^{(X)}$	0,2502 (0,0539)	0,1036 (0,3266)	0,2654 (0,7515)	0,2486 (0,0547)	0,3906 (16,0811)	-0,4194 (15,7176)
$\lambda_{21}^{(X)}$	0,4613 (0,0492)	0,3586 (0,8655)	0,5335 (1,1456)	0,4574 (0,0492)	0,4666 (14,3143)	0,1121 (15,1237)
$\lambda_{31}^{(X)}$	0,6158 (0,0513)	0,4719 (0,8290)	0,5313 (0,6683)	0,6111 (0,0523)	0,8694 (17,0163)	0,3592 (14,1630)
$\lambda_{41}^{(X)}$	0,7216 (0,0553)	0,4669 (0,8173)	0,6941 (1,3998)	0,7153 (0,0558)	0,2658 (15,0636)	0,6418 (17,4444)
$\lambda_{52}^{(X)}$	-0,5148 (0,0575)	-0,3359 (0,5359)	-0,4314 (1,0504)	-0,5094 (0,0653)	0,4360 (18,7116)	0,6641 (19,2945)
$\lambda_{62}^{(X)}$	0,8496 (0,0800)	0,5218 (0,6685)	0,5247 (0,9615)	0,8399 (0,0957)	0,0523 (14,9662)	1,5371 (15,0757)
$\lambda_{72}^{(X)}$	0,3967 (0,0547)	0,3071 (0,2287)	0,2481 (0,2840)	0,3930 (0,0580)	0,1050 (15,0657)	-0,0818 (14,0495)
$\lambda_{82}^{(X)}$	-0,1690 (0,0528)	-0,1071 (0,2494)	-0,0737 (0,5906)	-0,1676 (0,0550)	-0,0423 (17,7247)	-0,1317 (14,6621)
δ_1	0,9345 (0,0270)	0,8827 (1,0964)	0,3654 (3,6306)	0,9093 (0,0284)	-1313,0750 (2980,6175)	-1040,1734 (2656,6829)
δ_2	0,7847 (0,0453)	0,1230 (7,9227)	-0,5957 (12,3673)	0,7660 (0,0452)	-992,0743 (2293,2279)	-1055,5325 (2798,3339)
δ_3	0,6181 (0,0628)	0,0907 (8,5357)	0,2715 (7,7999)	0,6033 (0,0633)	-942,6085 (2517,6122)	-976,7982 (2493,8739)
δ_4	0,4762 (0,0808)	0,1147 (6,8711)	-1,4393 (17,7767)	0,4655 (0,0804)	-930,5906 (2300,9912)	-938,6759 (2431,8946)
δ_5	0,7317 (0,0586)	0,6003 (1,9678)	-0,2885 (6,3427)	0,7146 (0,0588)	-1540,3197 (3471,6644)	-1204,6693 (3174,9657)
δ_6	0,2719 (0,1404)	0,2812 (3,8746)	-0,1989 (7,3790)	0,2660 (0,1413)	-1178,8425 (2853,8545)	-1023,1927 (2575,3463)
δ_7	0,8397 (0,0430)	0,8535 (0,1035)	0,8579 (0,1072)	0,8195 (0,0437)	-1151,4269 (2676,7489)	-1128,8582 (2672,5279)
δ_8	0,9686 (0,0184)	0,9264 (1,0819)	0,6461 (5,1891)	0,9432 (0,0223)	-1409,0637 (3216,1706)	-1320,3845 (3104,2922)

Analisando as Tabelas 5.10 a 5.12 notamos que os parâmetros estimados por ML, quando o modelo foi ajustado utilizando os dados originais, foram um pouco maiores do que os obtidos por GLS, ULS, WLS e DWLS. Em contrapartida, o desvio padrão destas foram muito menores do que os obtidos por ML.

A função de discrepância E não conseguiu estimar, na média, o modelo corretamente.

Em comparação com as Tabelas 5.1 e 5.2, observamos que os coeficientes estimados foram maiores que os obtidos quando os dados eram normalmente distribuídos.

Quando o modelo foi estimado utilizando as correlações de Pearson para as variáveis categóricas, os parâmetros obtidos foram, na média, menores que os obtidos com os dados originais. No caso em que o modelo foi ajustado utilizando a matriz de correlações especiais, notamos que ML, GLS e ULS apresentaram valores parecidos e WLS e DWLS se desviaram um pouco dos parâmetros estimados com os dados originais. O desvio padrão das estimativas obtidas por ULS foi bem menor que nos demais métodos.

Pelas Tabelas 5.13 a 5.15 observamos que quando o modelo foi ajustado com a matriz de correlações dos dados originais, os parâmetros estimados foram aproximadamente iguais (na média e no desvio padrão) para todas as funções de discrepância. De maneira geral as estimativas foram um pouco menores do que quando as variáveis do modelo eram normalmente distribuídas.

Já quando o modelo foi estimado utilizando a matriz de correlações especiais, notamos estimativas parecidas quando as funções de discrepância ML, GLS e ULS foram utilizadas, nesta situação com desvios padrões parecidos. As estimativas obtidas por WLS e DWLS se desviaram um pouco em relação às obtidas anteriormente, principalmente sobre as variáveis relacionadas ao constructo ξ_2 . A função de discrepância E não conseguiu estimar, na média, os parâmetros do modelo.

5.4 Resumo do Capítulo

Neste capítulo realizamos estudos de simulação que utilizaram todos os conceitos trabalhados nos capítulos anteriores.

Foi verificado, em dois modelos de equações estruturais hipotéticos, como se comportam as estimativas dos parâmetros quando o modelo é estimado utilizando a matriz de correlações especiais, ou a matriz de correlação de Pearson, e as diferentes funções de discrepância citadas neste texto. Também foi avaliado como é que estas estimativas se comportam quando existem desvios de normalidade no conjunto de dados.

De maneira geral, notamos que quando os dados são normalmente distribuídos e o modelo é ajustado com a matriz de correlação dos dados originais, não há diferença entre qual função de discrepância utilizar. Quando os modelos foram gerados com desvios de simetria e curtose observamos que as estimativas variaram um pouco, mas os resultados obtidos foram parecidos para todos os métodos de estimação.

Quando algumas das variáveis foram categorizadas e o modelo reestimado utilizando a matriz de correlações especiais, notamos que as estimativas continuaram próximas dos valores obtidos com os dados originais, mas com um aumento do desvio padrão. Quando desvios de normalidade foram incluídos no modelo, observamos que as estimativas continuaram próximas dos parâmetros dos dados originais, mas com alguns desvios em relação aos dados simulados seguindo uma distribuição normal.

Em todas as situações testadas a matriz de correlação de Pearson, considerando os indicadores categóricos como contínuos, subestimou os parâmetros dos modelos estudados.

Nos casos em que o modelo foi ajustado por E, notamos que algumas poucas estimativas com valores muito atípicos acabaram por onerar a média dos valores encontrados. Esta situação também foi observada quando os modelos de equações estruturais foram estimados por WLS e DWLS, embora em escala muito menor.

Capítulo 6

Conclusões

Este trabalho teve como objetivo principal avaliar a inclusão de indicadores categóricos ordinais em modelos de equações estruturais, criando diretrizes que deveriam ser seguidas em aplicações práticas.

Inicialmente fizemos uma revisão dos principais procedimentos envolvidos em SEM. Apresentamos as equações matriciais, o diagrama de caminhos, as suposições inerentes ao modelo, as principais funções de discrepância utilizadas no processo de estimação, algumas medidas de qualidade do ajuste e algumas técnicas de reespecificação do modelo quando um ajuste razoável não é obtido. De uma forma geral, buscou-se fazer um resumo das principais recomendações encontradas na literatura sobre o assunto.

Logo após, avaliamos as suposições violadas, em modelos de equações estruturais, quando variáveis categóricas ordinais são utilizadas como indicadores contínuos no modelo. Nestes casos o procedimento corretivo mais indicado é supor um modelo limiar relacionando estes indicadores às suas contrapartes latentes e estimar a matriz de correlações populacional através de correlações especiais, mais especificamente as correlações tetracórica, policórica, biserial e poliserial.

Fizemos um resumo completo das correlações especiais apresentando as principais formas de estimação destes coeficientes. Realizamos estudos de simulação para cada uma destas correlações e observamos que quando possuímos variáveis categóricas ordinais, elas estimam a relação entre as variáveis com maior precisão do que a correlação de Pearson, considerando os indicadores ordinais como contínuos; fizemos também alguns estudos de simulação considerando desvios de simetria e curtose nos dados.

Por fim aplicamos os conceitos apresentados neste texto a dois modelos de equações estruturais hipotéticos. Notamos que, realmente, não existem grandes diferenças, entre as estimativas obtidas, quando o modelo é ajustado pelas diferentes funções de discrepância. Quando transformamos alguns dos indicadores em variáveis ordinais, notamos que é preferível utilizar a matriz de correlações especiais à matriz de correlação de Pearson para estimar o modelo; neste tipo de situação observamos que quando existem desvios de normalidade, os parâmetros são estimados com alguns desvios.

Quando o modelo foi ajustado utilizando a matriz de correlações especiais, as funções de discrepância WLS, DWLS e, principalmente E, apresentaram poucas estimativas com valores inconsistentes, mas que acabaram por prejudicar as estimativas médias e seus respectivos desvios padrões. De maneira geral a função ULS foi a que apresentou as estimativas mais coerentes, com um desvio padrão relativamente pequeno.

Como sugestão para estudos futuros, recomendamos que sejam realizados mais estudos de simulação, parecidos com o proposto, aplicando a uma gama maior de modelos de equações estruturais, variando o tamanho amostral e o número de indicadores categóricos ordinais a fim de validarmos os resultados encontrados no presente estudo.

Sugerimos também um estudo de robustez mais amplo sobre o comportamento dos estimadores sob desvios de simetria e curtose.

Apêndice A

Matriz de Correlações Especiais

A função *cor_especiais*, implementada no software R, calcula a matriz de correlações especiais a partir de um conjunto de dados em que as variáveis categóricas são do tipo *factor* (para maiores detalhes ver o *help* do software). Ela depende das funções *polychor* e *polyserial*, disponíveis no pacote *Polycor*.

Para duas variáveis dicotômicas a rotina calcula a correlação tetracórica; para duas variáveis categóricas ordinais calcula-se a correlação policórica; quando uma variável é dicotômica e a outra é contínua, calcula-se a correlação biserial; e quando uma variável é categórica ordinal e a outra é contínua, estima-se a correlação poliserial.

Os parâmetros de entrada da função *cor_especiais* são:

- *dados* = Conjunto de dados com as variáveis categóricas sendo do tipo *factor*;
- *m_tetra* = Método de estimação da correlação tetracórica;
- *m_poli* = Método de estimação da correlação policórica;
- *m_biserial* = Método de estimação da correlação biserial;
- *m_poliserial* = Método de estimação da correlação poliserial.

Os parâmetros *m_tetra*, *m_poli*, *m_biserial* e *m_poliserial* devem ser declarados como texto e os possíveis valores que podem receber são:

- “Q1” = Estimação da correlação policórica pela aproximação Q_1 de Pearson;
- “Q2” = Estimação da correlação policórica pela aproximação Q_2 de Pearson;
- “Q3” = Estimação da correlação policórica pela aproximação Q_3 de Pearson;
- “Q4” = Estimação da correlação policórica pela aproximação Q_4 de Pearson;
- “SUBS” = Estimação da correlação biserial pelo coeficiente biserial e da correlação poliserial pelo método da substituição;

- “STEPS” = Estimação das correlações especiais pelo método 2-passos;
- “ML” = Estimação das Correlações especiais pelo método de máxima verossimilhança.

Os códigos da função são:

```
cor_especiais = function(dados, m_tetra, m_poli, m_biserial, m_poliserial)
{
  # Função interna para calcular a correlação policórica
  policorica = function(n, method)
  {
    if (method == 'ML' | method == 'MV')
    { f = polychor(n, ML = T) } else
    { f = polychor(n, ML = F) }
    f
  }
  # Função interna para estimar a correlação tetracórica
  tetracorica = function(n, method)
  {
    # variáveis auxiliares utilizadas para o cálculo da correlação tetracórica
    a1 = n[1,1] * n[2,2]
    a2 = n[1,2] * n[2,1]
    a3 = n[1,1] + n[2,2]
    a4 = n[1,2] + n[2,1]
    sum = sum(n)
    prod = prod(n)
    aux = 1 + ( 2*a2*sum) / ((a1 - a2) * a4)
    k = (4*prod*sum^2) / ( (a1 - a2)^2 * a3 * a4 )
    if (method == 'Q1')
    { f = (a1 - a2) / (a1 + a2) }
    else if (method == 'Q2')
    { f = sin(pi / 2) * ((sqrt(a1) - sqrt(a2)) / (sqrt(a1) + sqrt(a2))) }
    else if (method == 'Q3')
    { if (a4 != 0)
      { f = sin(pi / 2) * (1 / aux) } else
      { f = NA } }
    else if (method == 'Q4')
    { if (a3 != 0 & a4 != 0)
      { if (k >= -1)
        { f = sin(pi / 2) * (1 / sqrt(1 + k)) } else
        { f = NA } }
      } else
      { f = NA }
    }
  }
}
```

```

if (is.na(f)) f = NA else if (a3 > a4) f = f * -1}
else if (method == 'STEPS')
{ f = polychor(n) }
else if (method == 'ML' | method == 'MV')
{ f = polychor(n, ML = T)}
f
}

# Função interna para calcular as correlações biserial e poliserial
serial = function(x, y, method, control = list(), std.err = FALSE,
maxcor = 0.9999, bins = 4)
{
if (method == 'STEPS')
{ f = function(pars) {
rho = pars[1]
if (abs(rho) > maxcor) rho = sign(rho) * maxcor
cts = if (length(pars) == 1)
c(-Inf, cuts, Inf)
else c(-Inf, pars[-1], Inf)
tau = (matrix(cts, n, s + 1, byrow = TRUE) - matrix(rho * z, n, s + 1))/sqrt(1 - rho^2)
- sum(log(dnorm(z) * (pnorm(tau[cbind(indices, y + 1)]) - pnorm(tau[cbind(indices, y)]))))
}
if (!is.numeric(x)) stop("x must be numeric")
valid = complete.cases(x, y)
x = x[valid]
y = y[valid]
z = scale(x)
tab = table(y)
n = sum(tab)
s = length(tab)
if (s < 2)
{ warning("y has fewer than 2 levels")
return(NA) }
if (sum(0 != tab) < 2)
{ warning("y has fewer than 2 levels with data")
return(NA)}
indices = 1:n
cuts = qnorm(cumsum(tab)/n)[-s]
y = as.numeric(as.factor(y))
rho = sqrt((n - 1)/n) * sd(y) * cor(x, y)/sum(dnorm(cuts))
v = optimise(f, interval = c(-1, 1))$minimum }
else if (method == 'ML' | method == 'MV')
{ v = polyserial(x, y, ML = T) }

```

```

else if (method == 'Substituição' method == 'SUBS' method == 'SUBSTITUIÇÃO' )
{v = polyserial(x, y, ML = F)}
v
}
# Rotina para calcular a matriz de correlações especiais
dados = dados[complete.cases(dados),]
m = dim(dados)[2]
cor_especial = matrix(ncol=m,nrow=m)
method = matrix(ncol=m,nrow=m)
for (i in 1:m)
for (j in 1:i)
{if (i == j)
{cor_especial[i,j] = 1
method[i,j] = 'NA'} else
{
auxi = is.factor(dados[,i])
auxj = is.factor(dados[,j])
if (auxi) cati = length(table(dados[,i]))
if (auxj) catj = length(table(dados[,j]))
if (auxi & auxj)
{if (cati == 2 & catj == 2)
{ f = tetracorica(n = table(dados[,i], dados[,j]), method = m_tetra)
auxm = 'Tetracórica'
} else
{ f = policorica(n = table(dados[,i], dados[,j]), method = m_poli)
auxm = 'Policórica'}
} else if (auxi | auxj)
{ if (auxi)
{if (cati == 2)
{ f = serial(dados[,j], dados[,i], method = m_biserial)
auxm = 'Biserial'} else
{ f = serial(dados[,j], dados[,i], method = m_poliserial)
auxm = 'Poliserial'}}} else
{ if (catj == 2)
{ f = serial(dados[,i], dados[,j], method = m_biserial)
auxm = 'Biserial'} else
{ f = serial(dados[,i], dados[,j], method = m_poliserial)
auxm = 'Poliserial'}}} else
{ f = cor(dados[,i],dados[,j])
auxm = 'Pearson'}
cor_especial[i,j] = cor_especial[j,i] = f
method[i,j] = method[j,i] = auxm

```

```
}}  
resultado = list()  
dimnames(cor_especial) = list(names(dados),names(dados))  
dimnames(method) = list(names(dados),names(dados))  
resultado$cor = cor_especial  
resultado$tipo = method  
resultado  
}
```


Apêndice B

Função para Ajustar Modelos SEM

A função *sem*, disponível no pacote *sem* do software R, estima modelos de equações estruturais somente por máxima verossimilhança. A partir do código original da função foram feitas modificações com o objetivo de que a função passasse a estimar os modelos pelas demais funções de discrepância apresentadas neste texto.

A nova função foi dividida em duas partes. A primeira função, denominada *sem_modificada*, arruma os dados de entrada e executa a segunda função, denominada *aux*. A função *aux* é a que estima o modelo de equações estruturais.

Utilizando a função *summary* em conjunto com a função *sem_modificada*, é possível calcular as principais medidas de qualidade do ajuste e testes estatísticos para os parâmetros e para o modelo geral.

Além dos parâmetros originais da função *sem*¹, foram incorporados os seguintes parâmetros:

- *method* = Método de Estimação do Modelo de Equações Estruturais;
- *acov* = Matriz de Pesos, utilizada quando o modelo é estimado por WLS ou DWLS;
- *curtose* = Coeficiente de Curtose Comum, utilizado quando o modelo é estimado por E.

O parâmetro *method* deve ser declarado como texto e os possíveis valores que pode receber são:

- “ML” = Estimação do Modelo por Máxima Verossimilhança;
- “GLS” = Estimação do Modelo por Mínimos Quadrados Genaralizados;
- “ULS” = Estimação do Modelo por Mínimos Quadrados Não-Ponderados;
- “WLS” = Estimação do Modelo por Mínimos Quadrados Ponderados;
- “DWLS” = Estimação do Modelo por Mínimos Quadrados Ponderados Diagonalmente;

¹Para uma descrição completa dos parâmetros da função *sem*, veja o *help* da função.

- “E” = Estimação do Modelo por Estimadores Elípticos.

O coeficiente de correlação comum pode ser calculado por qualquer uma das formas apresentadas anteriormente.

Os códigos da função *sem_modificada* e *aux* são:

```
sem_modificada = function (ram, S, N, obs.variables = rownames(S), fixed.x = NULL,
  debug = FALSE, method, acov = NA, curtose = NA, ...)
{
  parse.path = function(path) {
    path.1 = gsub("-", , gsub(", , path))
    direction = if (regexpr("<>", path.1) > 0) 2
    else if (regexpr("<", path.1) > 0) -1
    else if (regexpr(">", path.1) > 0) 1
    else stop(paste("ill-formed path:", path))
    path.1 = strsplit(path.1, "[<>]")[1]
    list(first = path.1[1], second = path.1[length(path.1)], direction = direction)}
  if (!(is.matrix(ram)) | ncol(ram) != 3)
    stop("ram argument must be a 3-column matrix")
  startvalues = as.numeric(ram[, 3])
  par.names = ram[, 2]
  n.paths = length(par.names)
  heads = from = to = rep(0, n.paths)
  for (p in 1:n.paths) {
    path = parse.path(ram[p, 1])
    heads[p] = abs(path$direction)
    to[p] = path$second
    from[p] = path$first
    if (path$direction == -1) {
      to[p] = path$first
      from[p] = path$second } }
  ram = matrix(0, p, 5)
  all.vars = unique(c(to, from))
  latent.vars = setdiff(all.vars, obs.variables)
  not.used = setdiff(obs.variables, all.vars)
  if (length(not.used) > 0) {
    rownames(S) = colnames(S) = obs.variables
    obs.variables = setdiff(obs.variables, not.used)
    S = S[obs.variables, obs.variables]
    warning("The following observed variables are
    in the input covariance or raw-moment matrix"
    , "but do not appear in the model: n", paste(not.used, collapse = ", "), " n")
  }
  vars = c(obs.variables, latent.vars)
```



```

pars = na.omit(unique(par.names))
ram[, 1] = heads
ram[, 2] = apply(outer(vars, to, "=="), 2, which)
ram[, 3] = apply(outer(vars, from, "=="), 2, which)
par.nos = apply(outer(pars, par.names, "=="), 2, which)
if (length(par.nos) > 0)
ram[, 4] = unlist(lapply(par.nos, function(x) if (length(x) == 0) 0
else x))
ram[, 5] = startvalues
colnames(ram) = c("heads", "to", "from", "parameter", "start")
if (!is.null(fixed.x))
fixed.x = apply(outer(vars, fixed.x, "=="), 2, which)
n = length(obs.variables)
m = length(all.vars)
t = length(pars)
if (debug) {
cat(" n observed variables: n")
print(paste(paste(1:n, ":", sep = ), obs.variables, sep = ))
cat(" n")
if (m > n) {
cat(" n latent variables: n")
print(paste(paste((n + 1):m, ":", sep = ), latent.vars, sep = ))
cat(" n") }
cat(" n parameters: n")
print(paste(paste(1:t, ":", sep = ), pars, sep = ))
cat(" n n RAM: n")
print(ram) }
aux = (ram = ram, S = S, N = N, param.names = pars, var.names = vars,
fixed.x = fixed.x, debug = debug, acov = acov, method = method, curtose = curtose, ...)
}

```

```

aux = function (ram, S, N, param.names = paste("Param", 1:t, sep = ),
var.names = paste("V", 1:m, sep = ), fixed.x = NULL, raw = FALSE,
debug = FALSE, analytic.gradient = TRUE, warn = FALSE, maxiter = 500,
par.size = c("ones", "startvalues"), refit = TRUE, start.tol = 1e-06,
method, acov, curtose, ...)
{
ord = function(x) 1 + apply(outer(unique(x), x, "<"), 2, sum)
is.triangular = function(X)
{is.matrix(X) && (nrow(X) == ncol(X)) && (all(0 == X[upper.tri(X)])) ||
(all(0 == X[lower.tri(X)]))}
S = unclass(S)

```

```

if (nrow(S) > 1 && is.triangular(S))
S = S + t(S) - diag(diag(S))
if (!isSymmetric(S))
stop("S must be a square triangular or symmetric matrix")
if (qr(S)$rank < ncol(S))
warning("S is numerically singular: expect problems")
if (any(eigen(S, symmetric = TRUE, only.values = TRUE)$values <= 0))
warning("S is not positive-definite: expect problems")
if ((!is.matrix(ram)) | ncol(ram) != 5 | (!is.numeric(ram)))
stop("ram argument must be a 5-column numeric matrix")
par.size = if (missing(par.size)) {
range = range(diag(S))
if (range[2]/range[1] > 100) "startvalues"
else "ones"}
else match.arg(par.size)
n = nrow(S)
observed = 1:n
n.fix = length(fixed.x)
if (!is.null(fixed.x)) {
for (i in 1:n.fix) {
for (j in 1:i) {
ram = rbind(ram, c(2, fixed.x[i], fixed.x[j], 0, S[fixed.x[i], fixed.x[j]]))}}
m = max(ram[, c(2, 3)])
missing.variances = setdiff(1:m, ram[, 2][ram[, 2] == ram[, 3]])
if (length(missing.variances) > 0)
warning(paste("The following variables have no variance or error-variance parameter
(double-headed arrow): n", paste(var.names[missing.variances], collapse = ", "),
" nThe model is almost surely misspecified; check also for missing covariances. n"))
t = max(ram[, 4])
df = n * (n + 1)/2 - t - n.fix * (n.fix + 1)/2
if (df < 0)
stop(paste("The model has negative degrees of freedom =", df))
J = matrix(0, n, m)
correct = matrix(2, m, m)
diag(correct) = 1
J[cbind(1:n, observed)] = 1
par.posn = sapply(1:t, function(i) which(ram[, 4] == i)[1])
colnames(ram) = c("heads", "to", "from", "parameter", "start value")
rownames(ram) = rep(, nrow(ram))
if (length(param.names) > 0) rownames(ram)[par.posn] = param.names
fixed = ram[, 4] == 0
sel.free = ram[, 4]

```

```

sel.free[fixed] = 1
one.head = ram[, 1] == 1
one.free = which(!(fixed) & one.head)
two.free = which(!(fixed) & !one.head)
arrows.1 = ram[one.head, c(2, 3), drop = FALSE]
arrows.2 = ram[!one.head, c(2, 3), drop = FALSE]
arrows.2t = ram[!one.head, c(3, 2), drop = FALSE]
arrows.1.free = ram[one.free, c(2, 3), drop = FALSE]
arrows.2.free = ram[two.free, c(2, 3), drop = FALSE]
sel.free.1 = sel.free[one.free]
sel.free.2 = sel.free[two.free]
unique.free.1 = unique(sel.free.1)
unique.free.2 = unique(sel.free.2)

# Calculos auxiliares utilizados nas funções de maximização
matrix = matrix( rep(0 ,(n*(n+1)/2)^2), nrow = (n*(n+1)/2),ncol = (n*(n+1)/2))
dimnames(matrix) = dimnames(acov)
curtose1 = curtose / ((4 * (curtose + 1)^2) + (2 * n * (n+1)))

# Função a ser minimizada
objective.1 = function(par) {
  A = P = matrix(0, m, m)
  val = ifelse(fixed, ram[, 5], par[sel.free])
  A[arrows.1] = val[one.head]
  P[arrows.2t] = P[arrows.2] = val[!one.head]
  I.Ainv = solve(diag(m) - A)
  C = J %*% I.Ainv %*% P %*% t(I.Ainv) %*% t(J)
  Cinv = solve(C)
  if (method == 'ML')
  { f = sum(diag(S %*% Cinv)) + log(det(C)) - log(det(S)) - n }
  else if (method == 'GLS')
  { f = 0.5 * sum(diag(((S - C)%*%solve(S))%*%((S - C)%*%solve(S)))) }
  else if (method == 'ULS')
  { f = 0.5 * sum(diag(((S - C))%*%((S - C)))) }
  else if (method == 'WLS')
  { aux = matrix
    for (gg in 1:n)
    for (hh in 1:gg)
    for (ii in 1:n)
    for (jj in 1:ii)
    aux[paste('s',gg,',',hh,sep=''),paste('s',ii,',',jj,sep='')] =
    acov[paste('s',gg,',',hh,sep=''),paste('s',ii,',',jj,sep='')]*(S[gg,hh]-C[gg,hh])*(S[ii,jj]-C[ii,jj])
  }
}

```

```

f = sum(aux) }
else if (method == 'DWLS')
{ aux = matrix
for (gg in 1:n)
for (hh in 1:gg)
aux[paste('s',gg,',',hh,sep=''),paste('s',gg,',',hh,sep='')] =
acov[paste('s',gg,',',hh,sep=''),paste('s',gg,',',hh,sep='')] * ((S[gg,hh]-C[gg,hh])^2)
f = sum(aux) }
else if (method == 'E')
{ f = (0.5*(curtose+1)^-1*sum(diag(((S-C)%*% solve(S))%*%((S-C)%*% solve(S))))
-curtose1*(sum(diag((S-C)%*% solve(S))))^2 }
else if (method == 'ERLS') { f =
(0.5*(curtose+1)^-1*sum(diag(((S-C)%*% solve(C))%*%((S-C)%*% solve(C))))
-curtose1*(sum(diag((S-C)%*% solve(C))))^2 }
else if (method == 'ELS')
{f = (0.5*(curtose+1)^-1 * sum(diag(((S-C))%*%((S-C))))-curtose1*(sum(diag((S-C))))^2 }
attributes(f) = list(C = C, A = A, P = P)
f }

```

```

result = list()
result$var.names = var.names
result$ram = ram
rownames(S) = colnames(S) = var.names[observed]
result$S = S
result$J = J
result$n.fix = n.fix
result$n = n
result$N = N
result$m = m
result$t = t
result$raw = raw
if (length(param.names) == 0) {
warning("there are no free parameters in the model")
obj = objective.1(NULL)} else
{ start = if (any(is.na(ram[, 5][par.posn])))
startvalues(S, ram, debug = debug, tol = start.tol)
else ram[, 5][par.posn]
if (!warn) {
save.warn = options(warn = -1)
on.exit(options(save.warn))}
typsize = if (par.size == "startvalues") abs(start)
else rep(1, t)

```

```

res = nlm(if (analytic.gradient)
objective.2
else objective.1, start, hessian = TRUE, iterlim = maxiter,
print.level = if (debug) 2 else 0, typsize = typsize, ...)
convergence = res$code
if (!warn)
options(save.warn)
par = res$estimate
names(par) = param.names
obj = objective.1(par)
ram[par.posn, 5] = start
par.code = paste(var.names[ram[, 2]], c("<—", "<—>")[ram[, 1]], var.names[ram[, 3]])
result$coeff = par
result$par.posn = par.posn
result$convergence = convergence
result$iterations = res$iterations
result$raw = raw
if (convergence > 2)
warning(paste("Optimization may not have converged; nlm return code = ",
res$code, ". Consult ?nlm. n", sep = ))
qr.hess = try(qr(res$hessian), silent = TRUE)
if (class(qr.hess) == "try-error") {
warning("Could not compute QR decomposition of Hessian.
nOptimization probably did not converge. n")
cov = matrix(NA, t, t)
colnames(cov) = rownames(cov) = param.names
result$cov = cov}
else if (qr.hess$rank < t) {
warning("singular Hessian: model is probably underidentified. n")
cov = matrix(NA, t, t)
colnames(cov) = rownames(cov) = param.names
result$cov = cov
which.aliased = qr.hess$pivot[-(1:qr.hess$rank)]
aliased = param.names[which.aliased]
position.aliased = is.element(ram[, 4], which.aliased)
if (refit) {
warning("refitting without aliased parameters. n")
ram.refit = ram[!position.aliased, ]
par.refit = ram.refit[, 4]
par.refit[par.refit != 0] = ord(par.refit[par.refit != 0])
ram.refit[, 4] = par.refit
iterations = result$iterations

```

```

result = Recall(ram.refit, S, N, param.names = param.names[-which.aliased],
var.names = var.names, debug = debug, analytic.gradient = analytic.gradient,
warn = warn, maxiter = maxiter, par.size = par.size, refit = TRUE, ...)
result$iterations = iterations + result$iterations
aliased = c(aliased, result$aliased) }
result$aliased = aliased }
else {
cov = (2/(N - (!raw))) * solve(res$hessian)
colnames(cov) = rownames(cov) = param.names
result$cov = cov
if (any(diag(cov) < 0)) {
result$aliased = c(param.names[diag(cov) < 0],
result$aliased)
warning("Negative parameter variances. nModel is probably underidentified. n")}}}}

# Calculando o Valor da Função de Discrepância - Utiliado no Teste Qui-Quadrado
if (method == 'WLS' | method == 'DWLS')
result$scriterion = c(obj)/(N-1) else
result$scriterion = c(obj)

C = attr(obj, "C")
rownames(C) = colnames(C) = var.names[observed]
result$C = C
A = attr(obj, "A")
rownames(A) = colnames(A) = var.names
result$A = A
P = attr(obj, "P")
rownames(P) = colnames(P) = var.names
result$P = P

# Salvando no output a matriz de pesos utilizada
result$Pesos = acov

if (!raw) {
CC = diag(diag(S))

# Calculando o Qui-Quadrado do Modelo Nulo
if (method == 'ML' | method == 'MV')
result$schisqNull = (N - 1) * (sum(diag(S %%% solve(CC)))
+ log(det(CC)) - log(det(S)) - n)
else if (method == 'GLS')
result$schisqNull = (N - 1) * 0.5 *
sum(diag( ((S%% CC)%%% solve(S)) %%% ((S%% CC)%%% solve(S)) ))

```

```

else if (method == 'ULS')
result$chisqNull = (N - 1)*0.5*sum(diag( (S%% CC) %% (S%% CC) ))
else if (method == 'WLS')
{ aux = matrix
for (gg in 1:n)
for (hh in 1:gg)
for (ii in 1:n)
for (jj in 1:ii)
aux[paste('s',gg,',',hh,sep=''),paste('s',ii,',',jj,sep='')] =
acov[paste('s',gg,',',hh,sep=''),paste('s',ii,',',jj,sep='')]*( S[gg,hh]-CC[gg,hh] )*( S[ii,jj]-CC[ii,jj])
result$chisqNull = sum(aux) }
else if (method == 'DWLS')
{ aux = matrix
for (gg in 1:n)
for (hh in 1:gg)
aux[paste('s',gg,',',hh,sep=''),paste('s',gg,',',hh,sep='')] =
acov[paste('s',gg,',',hh,sep=''),paste('s',gg,',',hh,sep='')] * (( S[gg,hh] - CC[gg,hh] )^2)
result$chisqNull = sum(aux) }
else if (method == 'E')
result$chisqNull =
(0.5*(curtose-1)^-1*sum(diag(((S-CC)%% solve(S))%%((S-CC)%% solve(S))))
-curtose1*(sum(diag((S-CC)%% solve(S))))^2
else if (method == 'ERLS')
result$chisqNull =
(0.5*(curtose-1)^-1*sum(diag(((S-CC)%% solve(CC))%%((S-CC)%% solve(CC))))
-curtose1*(sum(diag((S-CC)%% solve(CC))))^2
else if (method == 'ELS')
result$chisqNull = (0.5*(curtose-1)^-1*sum(diag(((S-CC))%%((S-CC))))
-curtose1*(sum(diag((S-CC))))^2 }

class(result) = "sem"
result
}

```


Apêndice C

Matriz de Pesos

A função *pesos*, implementada no software R, calcula a matriz de pesos que é utilizada nos métodos WLS e DWLS para estimação de modelos de equações estruturais.

Os parâmetros de entrada da função são:

- *dados* = Conjunto de dados;
- *method* = Método que será utilizado para estimar o modelo SEM posteriormente.

O parâmetro *method* deve ser declarado como texto e os possíveis valores que pode receber são:

- “WLS” = Calcula a matriz de pesos completa, utilizada quando um modelo SEM é estimado por WLS;
- “DWLS” = Calcula a matriz diagonal de pesos, utilizada quando um modelo SEM é estimado por DWLS;

Os códigos da função são:

```
pesos = function(dados,method)
{
  #informações iniciais
  n = dim(dados)[1]
  p = dim(dados)[2]
  tam = p*(p+1)/2
  # criando e dando nome das colunas para a matriz de pesos
  W = matrix(ncol = tam ,nrow = tam)
  k=1
  nomes = vector()
  for (i in 1:p)
  for (j in 1:p)
  if (i >= j)
```

```

{ nomes[k] = paste('s',i,',',j,sep="")
k = k+1 }
dimnames(W) = list(nomes,nomes)
# Criando uma matriz com as diferenças entre os valores observados e a média
auxiliar = dados
media = mean(dados)
for (i in 1:p) auxiliar[,i] = dados[,i] - rep(media[i],n)
# Calculando a matriz de pesos
for (a in 1:p)
for (b in a:p)
for (c in 1:p)
for (d in c:p)
{if (method=='WLS' | (method=='DWLS' & a==c & b==d))
{ s_abcd = sum(auxiliar[,a] * auxiliar[,b] * auxiliar[,c] * auxiliar[,d])/n
s_ab = sum(auxiliar[,a] * auxiliar[,b])/n
s_cd = sum(auxiliar[,c] * auxiliar[,d])/n
valor = (1/n) * (s_abcd - s_ab*s_cd)
} else if (metodo=='DWLS')
valor = 0 else
valor = NA
W[paste('s',b,',',a,sep=""),paste('s',d,',',c,sep="")] = valor
}
W_inv = solve(W)
W_inv
}

```

Referências Bibliográficas

- Amemiya(1981)** T. Amemiya. Qualitative response models: A survey. *Journal of Economic Literature*, 19:483–536. Citado na pág. [43](#)
- Anderson e Rubin(1956)** T. W. Anderson e H. Rubin. Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium for Mathematical Statistics Problems*, 5:111–150. Citado na pág. [5](#)
- Babakus et al.(1987)** E. C. Babakus, E. Ferguson Jr., e K. G. Jöreskog. The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *Journal of Marketing Research*, 24:222–228. Citado na pág. [44](#)
- Bentler(1983)** P. M. Bentler. Simultaneous equation systems as momnt structure models. *Journal of Econometrics*, 22:13–42. Citado na pág. [5](#)
- Bentler e Bonett(1980)** P. M. Bentler e D. G. Bonett. Significance tests and Goodness-of-Fit in the analysis of covariance structures. *Psychological Bulletin*, 88:588–600. Citado na pág. [5](#), [33](#), [34](#)
- Bentler e Freeman(1983)** P. M. Bentler e E. H. Freeman. Tests for stability in linear structural equation system. *Psychometrika*, 48:143–145. Citado na pág. [13](#)
- Bishop et al.(1975)** T. Bishop, S. Fienberg, e P. Holland. *Discrete Multivariate Analysis*. Cambridge, MA: MIT Press. Citado na pág. [43](#)
- Bock e Borgman(1966)** R. D. Bock e R. E. Borgman. Analysis of covariance structures. *Psychometrika*, 31:507–534. Citado na pág. [5](#)
- Bollen(1989)** K. A. Bollen. *Structural Equations with Latent Variables*. Wiley, New York, USA. Citado na pág. [3](#), [8](#), [13](#), [15](#), [17](#), [18](#), [23](#), [24](#), [25](#), [26](#), [28](#), [30](#), [33](#), [34](#), [36](#), [38](#), [40](#), [44](#)
- Bollen e Lennox(1991)** K. A. Bollen e R. Lennox. Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, 110:2:305–314. Citado na pág. [18](#)
- Boomsma(1983)** A. Boomsma. *On the Rubustness of LISREL (Maximum Likelihood Estimation) Against Small Sample Size and Nonnormlity*. Sociometric Research Foundation, Amsterdam. Citado na pág. [30](#), [44](#)
- Browne(1974)** M. W. Browne. Generalizes least square estimators in the analysis of covariance structures. *South Africa Statistical Journal*, 8:1–24. Citado na pág. [5](#)
- Browne(1982)** M. W. Browne. Covariance structures. Em *Topics in Multivariate Analysis*, páginas 72–141, Cambridge. Cambridge University Press. Citado na pág. [5](#), [24](#), [30](#)

- Browne(1984)** M. W. Browne. Asymptotic distribution free methods in analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37:62–83. Citado na pág. [5](#), [24](#)
- Cox(1974)** N. R. Cox. Estimation of the correlation between a continuous and a discrete variable. *Biometrics*, 30:171–178. Citado na pág. [61](#)
- Dragow(1988)** F. Dragow. Polychoric and polyserial correlations. Em *Encyclopedia of Statistical Sciences*, volume 7, páginas 69–74, New York. John Wiley. Citado na pág. [65](#)
- Goldberger(1964)** A. S. Goldberger. *Econometric Theory*. Wiley, New York, USA. Citado na pág. [5](#)
- Goldberger(1972)** A. S. Goldberger. Structural equation methods in the social sciences. *Econometrica*, 40:979–1001. Citado na pág. [3](#)
- Goodman(1972)** L. A. Goodman. A general model for the analysis of surveys. *American Journal of Sociology*, 77:1035–1086. Citado na pág. [43](#)
- Hair et al.(1998)** J. F. Hair, B. Black, B. Babin, R. E. Anderson, e R. L. Tatham. *Multivariate Data Analysis*. Prentice Hall, 5 edição. Citado na pág. [5](#), [28](#)
- Hamdan e Martinson(1972)** M. A. Hamdan e E. O. Martinson. Maximum likelihood and some other asymptotically efficient estimators of correlation in two way contingency tables. *Journal of Statistical Computation and Simulation*, 1:45–54. Citado na pág. [50](#)
- Jennrich e Thayer(1973)** R. I. Jennrich e D. T. Thayer. A note on lawley’s formulas for standard errors in maximum likelihood factor analysis. *Psychometrika*, 38:571–580. Citado na pág. [30](#)
- Johnson e Creech(1983)** D. R. Johnson e J. C. Creech. Ordinal measures in multiple indicators models: A simulation study of categorization error. *American Sociological Review*, 48:398–407. Citado na pág. [44](#)
- Jöreskog(1969)** K. G. Jöreskog. A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34:183–202. Citado na pág. [5](#)
- Jöreskog(1973)** K. G. Jöreskog. A general method for estimating a linear structural equation system. Em *Structural Equation Models in the Social Sciences*, páginas 85–112, New York, USA. Academic Press. Citado na pág. [4](#), [5](#)
- Jöreskog e Goldberger(1972)** K. G. Jöreskog e A. S. Goldberger. Factor analysis by generalized least square. *Psychometrika*, 37:243–260. Citado na pág. [5](#)
- Jöreskog e Sorbom(1986)** K. G. Jöreskog e D. Sorbom. *Lisrel VI: Analysis of Linear Structural Relationships by Maximum Likelihood and Least Square Methods*, 1986. Citado na pág. [5](#), [21](#), [31](#), [34](#)
- Keesing(1972)** J. W. Keesing. *Maximum Likelihood Approaches to Causal Analysis*. Tese de Doutorado, Department of Education, University of Chicago. Citado na pág. [4](#), [5](#)
- Kendall e Stuart(1979)** M. G. Kendall e A. Stuart. *The Advanced Theory of Statistics. Vol 2: Inference and Relationship*. Macmillan, New York, USA, 4 edição. Citado na pág. [51](#)

- Lancaster e Hamdan(1964)** H. O. Lancaster e M. A. Hamdan. Estimation of the correlation coefficient in contingency tables with possibly nonmetrical characters. *Psychometrika*, 29:283–391. Citado na pág. 50, 51
- Lawley(1940)** D. N. Lawley. The estimation of factor loadings by the method of maximum likelihood. *Proceedings of the Royal Society of Edinburgh*, 60:64–82. Citado na pág. 5
- Lawley e Maxwell(1971)** D. N. Lawley e E. A. Maxwell. *Factor Analysis as a Statistical Method*. Butterworth, London. Citado na pág. 30
- Lazarsfeld e Henry(1968)** P. F. Lazarsfeld e N. W. Henry. *Latent Structure Analysis*. Houghton Mifflin, Boston. Citado na pág. 43
- Lord e Novick(1968)** F. M. Lord e M. R. Novick. *Statistical Theories of Mental Test Scores*. Addison-Wesley, Mass. Citado na pág. 65
- Maddala(1983)** G. S. Maddala. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press. Citado na pág. 43
- McArdle e McDonald(1984)** J. J. McArdle e R. P. McDonald. Some algebraic properties of the reticular action model for moment structures. *British Journal of Mathematical and Statistical Psychology*, 37:234–251. Citado na pág. 5
- Melhado(2004)** T. T. Melhado. Medidas de ajuste de modelos de equações estruturais. Dissertação de Mestrado, Instituto de Matemática e Estatística, Universidade de São Paulo. Citado na pág. 33, 36
- Muthén(1984)** B. Muthén. A general structure equation model with dichotomous, ordered categorical and continuous latent variable indicators. *Psychometrika*, 49:115–132. Citado na pág. 5
- Muthén(1987)** B. Muthén. *LISCOMP: Analysis of Linear Structural Equations with a Comprehensive Measurement Model*, 1987. Mooresville, IN: Scientific Software, Inc. Citado na pág. 5
- Muthén e Kaplan(1985)** B. Muthén e D. Kaplan. A comparison of some methodologies for the factor analysis of non-normal likert variables. *British Journal of Mathematical and Statistical Psychology*, 38:171–189. Citado na pág. 44
- Olsson(1979)** U. Olsson. Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, 44:443–460. Citado na pág. 44, 50, 52
- Olsson et al.(1982)** U. Olsson, F. Drasgow, e N. J. Dorans. The polyserial correlation coefficient. *Psychometrika*, 47:337–347. Citado na pág. 61
- Pearson(1900)** K. Pearson. Philos. *Trans. R. Soc. Lond. A*, 195:79–150. Citado na pág. 47, 50, 57
- Sanches(2005)** M. R. Sanches. Indicadores formativos em modelos de equações estruturais. Dissertação de Mestrado, Instituto de Matemática e Estatística, Universidade de São Paulo. Citado na pág. 15, 16, 18, 38, 83
- Saris et al.(1979)** W. E. Saris, W. M. Pijper, e P. Zegwaard. Detection of specification errors in linear structural equation models. Em *Sociological Methodology 1979*, páginas 151–171, San Francisco. Jossey-Bass. Citado na pág. 40

- Spearman(1904)** C. Spearman. General intelligence, objectively determined and mensured. *American Journal of Psychology*, 15:201–293. Citado na pág. [4](#)
- Tanaka e Huba(1985)** J. S. Tanaka e G. J. Huba. A fit index for covariance structure models under arbitrary gls estimation. *British Journal of Mathematical and Statistical Psychology*, 38:197–201. Citado na pág. [35](#)
- Tate(1955)** R. F. Tate. Devices the maximum likelihood estimator of the biserial correlation and its asymptotic properties. *Biometrika*, 42:205–216. Citado na pág. [65](#)
- Tuker e Lewis(1973)** L. R. Tuker e C. Lewis. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38:1–10. Citado na pág. [34](#)
- Wiley(1973)** D. E. Wiley. The identification problem for structural equation models with unmeasured variables. Em *Structural Equation Models in the Social Sciences*, páginas 69–83, New York, USA. Academic Press. Citado na pág. [4](#), [5](#)
- Wright(1918)** S. Wright. On the nature of size factors. *Genetics*, 3:367–374. Citado na pág. [4](#)
- Wright(1921)** S. Wright. Correlation and causation. *Journal of Agricultural Research*, 20: 557–585. Citado na pág. [4](#)
- Wright(1934)** S. Wright. The method of path coefficients. *Annals of Mathematical Statistics*, 5:161–215. Citado na pág. [4](#)
- Wright(1960)** S. Wright. Path coefficients and path regressions: Alternative or complementary concepts? *Biometrics*, 16:189–202. Citado na pág. [4](#)