

Equações Estruturais: Uma aplicação a perfis de consumidores de café

Gislene Araujo Pereira ¹

Letícia Lima Milani ²

Marcelo Ângelo Cirillo ^{2 4}

1 Introdução

O café sempre teve um papel importante para a economia brasileira, porém, sua participação relativa diminuiu gradativamente, dada a contínua diversificação da pauta de exportações brasileira [4]. Mesmo assim, o café continua sendo importante gerador de divisas para o país, visto que em 2013 contribuiu com 6,7% do valor total das exportações do agronegócio brasileiro. Já a produção de café no Brasil correspondeu a 33,60% do total mundial, tornando o país o maior produtor e exportador. O Brasil destacou-se também como o segundo maior consumidor de café, com um consumo anual de 21 milhões de sacas de 60 kg [1]. Desta forma, este trabalho tem por objetivo aplicar a modelagem de equações estruturais ao estudo de perfis de consumidores de café. Com este propósito foi utilizada uma pesquisa realizada em 2011, com uma amostra de 322 consumidores de café com idades a partir de 18 anos, na cidade de Varginha/MG.

A modelagem de equações estruturais (SEM-*Structural Equation Modeling*) é uma metodologia que permite a incorporação de variáveis latentes ou constructos (variável que não podem ser medidas diretamente, mas pode ser representada ou medida pelas variáveis observáveis), e também possibilita analisar várias relações de dependência entre si. De uma forma simplista [3] descrever a SEM como uma combinação das metodologias clássicas de análise fatorial, que define um modelo de medida que operacionaliza variáveis latentes ou constructos, e de regressão linear, que estabelece, no modelo estrutural, a relação entre as diferentes variáveis sob estudo. Como a SEM permite acomodar múltiplas relações de dependência inter-relacionadas em um só modelo, se aproxima muito da regressão múltipla. Entretanto uma única relação (equação) é estimada na regressão múltipla, já com a SEM pode estimar várias equações ao mesmo tempo, e elas podem ser inter-relacionadas, ou seja, a variável dependente em uma equação pode ser independente em outra. Segundo [2], a SEM é composta de um sistema de equações lineares formado em dois sub-sistemas, representados respectivamente pelas variáveis latentes, dadas no modelo estrutural, que especificam as relações de causa e efeito entre as variáveis latentes, apresentando os efeitos causais e o total da variância não explicada; e pelas variáveis observáveis descritas

¹DEX - UFLA. e-mail: gislene.araujo.p@gmail.com

²DEX - UFLA.

⁴Agradecimento ao CNPq, Fapemig e CAPES pelo apoio financeiro.

pelo modelo de mensuração, que especifica como as variáveis latentes são mensuradas pelas variáveis observadas, decrevendo sua validade e confiabilidade. Em se tratando dos métodos de estimação, ainda de de acordo [2] os mais citados na literatura, são métodos de estimação de Máxima Verossimilhança e Mínimos Quadrados .

2 Material e Métodos

Com o intuito de analisar perfis de consumidores de café, foi utilizada uma pesquisa realizada no ano de 2011, com uma amostra de 322 consumidores de café com idades a partir de 18 anos, na cidade de Varginha/MG . Esta pesquisa foi realizada no desenvolvimento de um projeto de iniciação científica aprovado pela Universidade Federal de Alfenas-MG (Unifal-MG). Vale ressaltar que o questionário utilizado nesta pesquisa teve aprovação do Comitê de Ética em Pesquisa (CEP) da Unifal-MG.

Com base nas respostas por meio de categorias (conforme Tabela 1 e Tabela 2), um modelo de equações estruturais apenas com variáveis observadas, foi proposto. Mantendo as especificações referentes a consideração das variáveis endógenas e exógenas, respectivamente conforme descrito nas Tabelas 1 e 2, para uma melhor compreensão do modelo estrutural proposto, na Figura 1 segue uma representação gráfica, na qual, as linhas que unem as variáveis definem os relacionamentos entre elas e o sentido da causa e efeito.

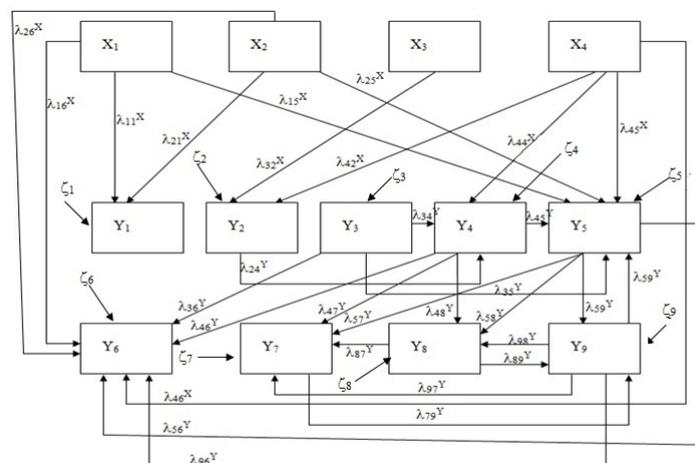


Figura 1- Diagrama de caminhos do modelo estrutural proposto.

Tabela 1- Categorias das questões considerando as variáveis exógenas observadas.

Variáveis: Questões	Categorias				
X1: Sexo	Feminino:1	Masculino:2			
X2:Faixa-Etária	18-27 anos:1	28-37 anos:2	38-47 anos:3	48-57 anos:4	acima 58 anos:5
X3:Estado-Civil	Solteiro:1	Casado:2	Viúvo:3	Divorciado:4	Outros:5
X4: Fumante	Sim:1	Não:2			

Tabela 2- Categorias das questões considerando as variáveis endógenas observadas.

Variáveis:Questões	Categorias
Y1: Onde inicio o consumo de café?	Casa:1; Trabalho:2; Outro:3
Y2: Com que frequência você toma café?	Alguns dias:1; Todos os dias:2; Várias vezes ao dia:3
Y3: Qual horário você costuma tomar café?	Manhã:1; Tarde:2; Noite:3; Vários Horários:4
Y4: Qual a média de xícaras de café consumidas por dia?	Uma:1; Duas:2; Três:3; Quatro:4; Mais de Quatro:5
Y5:O que prefere no café?	Gosto:1; Aroma:2; Gosto e Aroma:3
Y6: Como prefere o café?	Com Leite:1; Cappucino:2; Puro:3; Outros:4
Y7: Quais beneficios atribui ao café?	Estimulante:1; Melhora do raciocínio:2; Nenhum:3
Y8: Quais maleficios atribui ao café?	Insônia:1; Escurecimento dos Dentes:2; Gastrite:3; Nenhum:4
Y9: Você classifica um café como de qualidade por meio da:	Cor:1; Marca:2; Aroma:3; Gosto:4; Rendimento:5

As equações estruturais para o modelo proposto são descritas a seguir, considerando X_i ($i = 1, \dots, 4$) como variáveis exógenas observadas e Y_i ($i = 1, \dots, 9$) como variáveis endógenas observadas.

$$\begin{aligned}
 Y_1 &= \lambda_{11}^X X_1 + \lambda_{21}^X X_2 + \zeta_1 & Y_7 &= \lambda_{47}^Y Y_4 + \lambda_{57}^Y Y_5 + \lambda_{87}^Y Y_8 + \lambda_{97}^Y Y_9 + \zeta_7 \\
 Y_2 &= \lambda_{32}^X X_3 + \lambda_{42}^X X_4 + \zeta_2 & Y_8 &= \lambda_{48}^Y Y_4 + \lambda_{58}^Y Y_5 + \lambda_{98}^Y Y_9 + \zeta_8 \\
 Y_3 &= \lambda_{33}^X X_3 + \lambda_{43}^X X_4 + \zeta_3 & Y_9 &= \lambda_{59}^Y Y_5 + \lambda_{79}^Y Y_7 + \lambda_{89}^Y Y_8 + \zeta_9 \\
 Y_4 &= \lambda_{44}^X X_4 + \lambda_{24}^Y Y_2 + \lambda_{34}^Y Y_3 + \zeta_4 \\
 Y_5 &= \lambda_{15}^X X_1 + \lambda_{25}^X X_2 + \lambda_{45}^X X_4 + \lambda_{35}^Y Y_3 + \lambda_{45}^Y Y_4 + \lambda_{95}^Y Y_9 + \zeta_5 \\
 Y_6 &= \lambda_{16}^X X_1 + \lambda_{26}^X X_2 + \lambda_{46}^X X_4 + \lambda_{36}^Y Y_3 + \lambda_{46}^Y Y_4 + \lambda_{56}^Y Y_5 + \beta_{96} Y_9 + \zeta_6
 \end{aligned} \tag{1}$$

Assumindo que as variáveis exógenas X_i ($i = 1, \dots, 4$) são não correlacionadas com os erros estruturais ζ_i ($i = 1, \dots, 9$), utilizou-se o método de mínimos quadrados para estimação dos parâmetros do modelo proposto .

3 Resultados e Discussões

Por meio dos resultados apresentados na Tabela 3, observa-se que com exceção das equações para Y_5 e Y_6 todas as variáveis consideradas para as outras equações propostas, ao nível de 5% são estatisticamente significativas.

Para a equação Y_5 , observa-se que apenas a variável exógena X_4 ("Fumante") e a variável endógena Y_4 ("Qual a média de xícaras de café consumidas por dia?") influenciam as respostas relacionadas a questão "O que você prefere no café?". Já para a equação Y_6 , observa-se que

apenas a variável exógena X_2 ("Faixa-etária") e as variáveis endógenas Y_3 ("Com que frequência você toma café?") e Y_5 ("O que você prefere no café?") não influenciam as respostas relacionadas a questão "Como você prefere seu café?".

Tabela 3- Estimativas dos parâmetros para o modelo proposto segundo as equações em (1).

Y1 e seus indicadores- $R^2 = 0,7996$				
Parâmetro	Estimativas	Erro Padrão	Valor t	Pr(> t)
λ_{11}^X	0,6713	0,0444	15,1390	$< 2E - 16$
λ_{21}^X	0,0802	0,0239	3,3540	0,0009
Y2 e seus indicadores- $R^2 = 0,6616$				
Parâmetro	Estimativas	Erro Padrão	Valor t	Pr(> t)
λ_{32}^X	0,7433	0,0751	9,8930	$< 2E - 16$
λ_{42}^X	0,1926	0,0836	2,3030	0,0219
Y3 e seus indicadores- $R^2 = 0,8235$				
Parâmetro	Estimativas	Erro Padrão	Valor t	Pr(> t)
λ_{33}^X	0,4270	0,0454	9,4140	$< 2E - 16$
λ_{43}^X	0,4876	0,0505	9,6590	$< 2E - 16$
Y4 e seus indicadores- $R^2 = 0,9245$				
Parâmetro	Estimativas	Erro Padrão	Valor t	Pr(> t)
λ_{44}^X	-0,2929	0,0456	-6,4250	4,79E-10
λ_{24}^Y	0,3391	0,0357	9,5060	$< 2E - 16$
λ_{34}^Y	1,1850	0,0598	19,8280	$< 2E - 16$
Y5 e seus indicadores- $R^2 = 0,8166$				
Parâmetro	Estimativas	Erro Padrão	Valor t	Pr(> t)
λ_{15}^X	-0,0771	0,0930	-0,8290	0,4077
λ_{25}^X	0,0293	0,0349	0,8400	0,4014
λ_{45}^X	0,6873	0,0906	7,5840	3,76E-13
λ_{35}^Y	0,0922	0,1050	0,8780	0,3804
λ_{45}^Y	0,2043	0,0611	3,3430	0,0009
λ_{95}^Y	0,0509	0,0428	1,1890	0,2352
Y6 e seus indicadores- $R^2 = 0,8938$				
Parâmetro	Estimativas	Erro Padrão	Valor t	Pr(> t)
λ_{16}^X	0,4822	0,1028	4,6920	4,04E-06
λ_{26}^X	-0,0383	0,0386	-0,9920	0,3218
λ_{46}^X	0,4173	0,1087	3,8370	0,0002
λ_{36}^Y	-0,0910	0,1160	-0,7840	0,4336
λ_{46}^Y	0,2871	0,0686	4,1850	3,71E-05
λ_{56}^Y	0,1084	0,0621	1,7460	0,0819
λ_{96}^Y	0,1455	0,0473	3,0750	0,0023

Tabela 3:Continuação.

Y7 e seus indicadores- $R^2 = 0,8191$				
Parâmetro	Estimativas	Erro Padrão	Valor t	Pr(> t)
λ_{47}^Y	-0,2243	0,0445	-5,0470	7,56E-07
λ_{57}^Y	0,1315	0,0638	2,0590	0,0403
λ_{87}^Y	0,3622	0,0507	7,1500	5,98E-12
λ_{97}^Y	0,3198	0,0497	6,4310	4,65E-10
Y8 e seus indicadores- $R^2 = 0,9035$				
Parâmetro	Estimativas	Erro Padrão	Valor t	Pr(> t)
λ_{48}^Y	0,2125	0,0477	4,4590	1,14E-05
λ_{58}^Y	0,4199	0,0665	6,3110	9,27E-10
λ_{98}^Y	0,6129	0,0429	14,2730	< 2E - 16
Y9 e seus indicadores- $R^2 = 0,8977$				
Parâmetro	Estimativas	Erro Padrão	Valor t	Pr(> t)
λ_{59}^Y	0,3296	0,0654	5,0360	7,97E-07
λ_{79}^Y	0,3086	0,0562	5,4860	8,37E-08
λ_{89}^Y	0,5411	0,0476	11,3660	< 2E - 16

4 Conclusões

O modelo proposto de uma forma geral se comportou relativamente bem, e além disso, as relações entre as variáveis exógenas e endógenas avaliadas, podem contribuir para que novos modelos de equações estruturais sejam desenvolvidos.

Referências

- [1] ABIC - ASSOCIAÇÃO BRASILEIRA DA INDÚSTRIA DE CAFÉ. Disponível em: <<http://www.abic.com.br>>. Acesso em fev.2014.
- [2] BOLLEN, K. **Structural Equations with Latent Variables**. New York,USA: Wiley, 1989.
- [3] MARÔCO, J. **Análise de Equações Estruturais: Fundamentos teóricos, software aplicações..** Pêro Pinheiro: ReportNumber,Ltda, 2010.
- [4] SANTOS,V.D.;GOMES.M.F.M.;BRAGA,M.J.;SILVEIRA,S.F.R. Análise do setor de produção e processamento de café em Minas Gerais: *uma abordagem matriz insumo-produto*. **RESR**, Piracicaba, v.47, n.2, p. 363-388, 2009.