**FUNDAÇÃO GETÚLIO VARGAS**

**Business Analytics – Big Data**

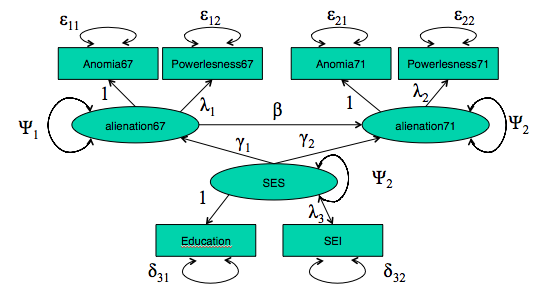
**Prof João Fernando Serrajordia Rocha de Mello**

**Disciplina: Modelagem Estatística Avançada – Trabalho em grupo**

**Grupo:**

**Para as questões que seguem, trabalharemos com os dados sociológicos da pesquisa da Wheaton.**

Considere o modelo de equações estruturais representado pelo diagrama abaixo. Considere que trabalharemos com o pacote SEM no R. Considere que nesta modelagem as variáveis são consideradas em mesma escala, portanto trabalharemos com a matriz de covariâncias (não com a matriz de correlações).



Siglas:

SES – Status Sócio Econômico

SEI - Índice sócio econômico

Questão 1

1. Quantos elementos possui a matriz de covariâncias das variáveis manifestas? (Lembre-se que não estamos trabalhando com a matriz de correlações) (0,5 pt)
   1. 6
   2. 13
   3. 15
   4. 21
   5. 27
2. Quantos parâmetros livres (i.e. a serem estimados e sem restrições) possui o modelo apresentado no diagrama? (0,5 pt)
   1. 6
   2. 13
   3. 15
   4. 17
   5. 21
3. Quantos graus de liberdade o modelo possui? (elementos da matriz de covariâncias menos parâmetros livres do modelo) (0,5 pt)
   1. 4
   2. 6
   3. 8
   4. 10
   5. 12

Questão 2:

1. Considere a saída do pacote “sem” do R. Considerando o Qui-quadrado de 76,1, qual alternativa traz uma afirmativa correta? (1,0 pt)
   1. O *p-value* é maior que 5%, indicando que o ajuste do modelo é bom. Além disso, pode-se também observar indicadores favoráveis na estatística de qualidade de ajuste (GOF>99%).
   2. Embora o *p-value* seja maior que 5%, é necessário observar outras estatísticas de qualidade de ajuste para assegurar a adequação do modelo.
   3. Embora o *p-value* tenha sido menor que 0,0001, o modelo tenha ajuste aceitável pois a estatística de qualidade de ajuste (GOF) é maior que 95%.
   4. O ajuste é ruim pois o *p-value* associado ao Qui-quadrado é menor que 0,0001 e a estatística de qualidade de ajuste *ajustada* é menor que 95%.
   5. O ajuste é considerado bom pois o *p-value* associado ao Qui-quadrado é menor que 0,001 e a estatística de qualidade de ajuste *ajustada* é menor que 95%.
2. Avaliando as cargas fatoriais padronizadas (utilize a função standardizedCoefficients) , pode-se afirmar: (1,0 pt)
   1. São altas para as variáveis ANOMIA67, POWLES67, ANOMIA71, POWLES71 e EDUC (maiores que 70%) e somente para a variável SEI está abaixo de 70%, mas próximo desta referência.
   2. Todas as cargas são menores que 70%, indicando um modelo de mensuração muito pobre.
   3. Todas as variáveis apresentam valores considerados muito altos para cargas fatoriais: todos maiores que 50%.
   4. Embora todas as cargas fatoriais sejam altas (maiores que 70%), o modelo de mensuração é ruim devido ao baixo ajuste geral observado pela estatística Qui-quadrado.
   5. As variáveis que medem o construto SES possuem as maiores cargas fatoriais.
3. Avaliando os índices de modificação, marque a alternativa correta: (1,0 pt)
   1. Não há parâmetros relevantes a serem adicionados.
   2. Há parâmetros pouco relevantes a serem adicionados, mas como não são explicáveis em termos de negócio, não devem ser incorporados no modelo.
   3. Há parâmetros de regressão com alta relevância mas que não fazem sentido no modelo, pois relacionam variáveis manifestas. No entanto, as correlação correspondentes entre as duas variáveis observadas em momentos distintos no tempo são relevantes e fazem sentido por ser a mesma variável em momentos diferentes.
   4. Os índices mostram parâmetros de variância faltantes para as variáveis não observáveis. Estas variáveis devem melhorar o ajuste geral se consideradas.
   5. A análise evidencia a falta de parâmetros de correlação entre os construtos, o que não pode ser considerado uma vez que já estão presentes parâmetros de regressão equivalentes.

Questão 3:

Considere o modelo alterado considerando-se as correlações entre os parâmetros de correlação entre Powerlesness67 com Powerlesness71 e Anomia67 com Animia71.

1. Após a alteração, Quantos elementos possui a matriz de covariâncias das variáveis manifestas? (Lembre-se que não estamos trabalhando com a matriz de correlações) (0,5 pt)
   1. 6
   2. 13
   3. 15
   4. 21
   5. 27
2. Após a alteração, quantos parâmetros livres (i.e. a serem estimados e sem restrições) possui o modelo apresentado no diagrama? (0,5 pt)
   1. 6
   2. 13
   3. 15
   4. 17
   5. 21
3. Após a alteração, quantos graus de liberdade o modelo possui? (i.e. elementos da matriz de covariâncias menos parâmetros livres do modelo) (0,5 pt)
   1. 4
   2. 6
   3. 8
   4. 10
   5. 12

Questão 4:

1. Sobre o modelo alterado, estimado pelo pacote SEM do R, com a matriz de covariâncias do data frame Wheaton, o que se pode afirmar? (1,0 pt)
   1. O p-*value* é maior que 5%, indicando que o ajuste do modelo é bom. Além disso, pode-se também observar indicadores favoráveis na estatística de qualidade de ajuste (GOF>99%).
   2. Embora o p-*value* seja maior que 5%, é necessário observar outras estatísticas de qualidade de ajuste para assegurar a adequação do modelo.
   3. Embora o p-*value* tenha sido menor que 0,0001, o modelo tenha ajuste aceitável pois a estatística de qualidade de ajuste (GOF) é maior que 95%.
   4. O ajuste é ruim pois o p-*value* associado ao qui-quadrado é menor que 0,0001 e as estatísticas de qualidade de ajuste são menores que 90%.
   5. O ajuste é considerado bom pois o p-*value* associado ao qui-quadrado é menor que 0,0001 e as estatísticas de qualidade de ajuste são menores que 90%.
2. Avaliando as cargas fatoriais padronizadas (utilize a função standardizedCoefficients) , pode-se afirmar: (1,0 pt)
   1. São altas para as variáveis ANOMIA67, POWLES67, ANOMIA71, POWLES71 e EDUC (maiores que 70%) e somente para a variável SEI está abaixo de 70%, mas pode-se considerar um bom modelo de mensuração.
   2. Todas as cargas são menores que 70%, indicando um modelo de mensuração muito pobre.
   3. Todas as variáveis apresentam valores considerados muito altos para cargas fatoriais: todos maiores que 50%.
   4. Embora todas as cargas fatoriais sejam altas, o modelo de mensuração é ruim devido ao baixo ajuste geral observado pela estatística qui-quadrado.
   5. As variáveis que medem o construto SES possuem as maiores cargas fatoriais.
3. Avaliando os índices de modificação, marque a alternativa correta: (1,0 pt)
   1. Não há parâmetros relevantes a serem adicionados.
   2. Há parâmetros pouco relevantes a serem adicionados, mas como não são explicáveis em termos de negócio, não devem ser incorporados no modelo.
   3. Há parâmetros de regressão com alta relevância mas que não fazem sentido no modelo. Há também parâmetros de alta relevância de correlação entre as variáveis observadas em momentos distintos no tempo que podem ser inseridos no modelo.
   4. Os índices mostram parâmetros de variância faltantes para as variáveis não observáveis. Estas variáveis devem melhorar o ajuste geral se consideradas.
   5. A análise evidencia a falta de parâmetros de correlação entre os construtos, o que não pode ser considerado uma vez que já estejam presentes parâmetros de regressão.

Questão 5: Considerando este modelo como final, o que se pode afirmar sobre a relação entre os **construtos**? (1.0 pt)

1. São estatisticamente significantes as relações Anomia67 e Powerlesness67 causando Alienation67, Anomia71 e Powerlesness71 causando Alienation 71 e SEI e Education causando SES .
2. As relações SES causa Alienation67 e Alienation71, Alienation 67 causa Alienation71 são estatisticamente significantes.
3. São significantes apenas SES causando Alienation71 e Alienation67 causando Alienation71.
4. Nenhuma relação foi encontrada como estatisticamente significante entre os construtos.
5. São significantes Alienation71 causando Alienation67, Alienation71 causando SES e Alienation71 causando SES.