

9^o CONGRESSO BRASILEIRO
de **AVALIAÇÃO PSICOLÓGICA**

25-28, JUNHO/2019



UCSAL, **SALVADOR**/BA
CAMPUS PITUAÇU

DESAFIOS NA PESQUISA E NA PRÁTICA EM AVALIAÇÃO PSICOLÓGICA

Pacote lavaan

Dr. Felipe Valentini &
Dr. Nelson Hauck

Pacote lavaan

- Latent Variable Analysis
- Desenvolvido por Yves Rosseel
- Autor possui uma página virtual sobre o pacote lavaan, com vários tutoriais: <http://lavaan.ugent.be/>
- Pacote lavaan faz muitos tipos de análises confirmatórias
- Fórum com perguntas e respostas:
<https://groups.google.com/forum/#!forum/lavaan>

Amostras de comportamentos é o fim que realmente queremos avaliar?

Variáveis Observadas – comportamentos tratados como diretamente observados

Ex. Saio bastante com os amigos

Variáveis Latentes – Variável estimada e não observada diretamente

Ex. Extroversão

Latente

Observado



Estimar o que não
é observado
diretamente

Parâmetros dos
itens



Latente

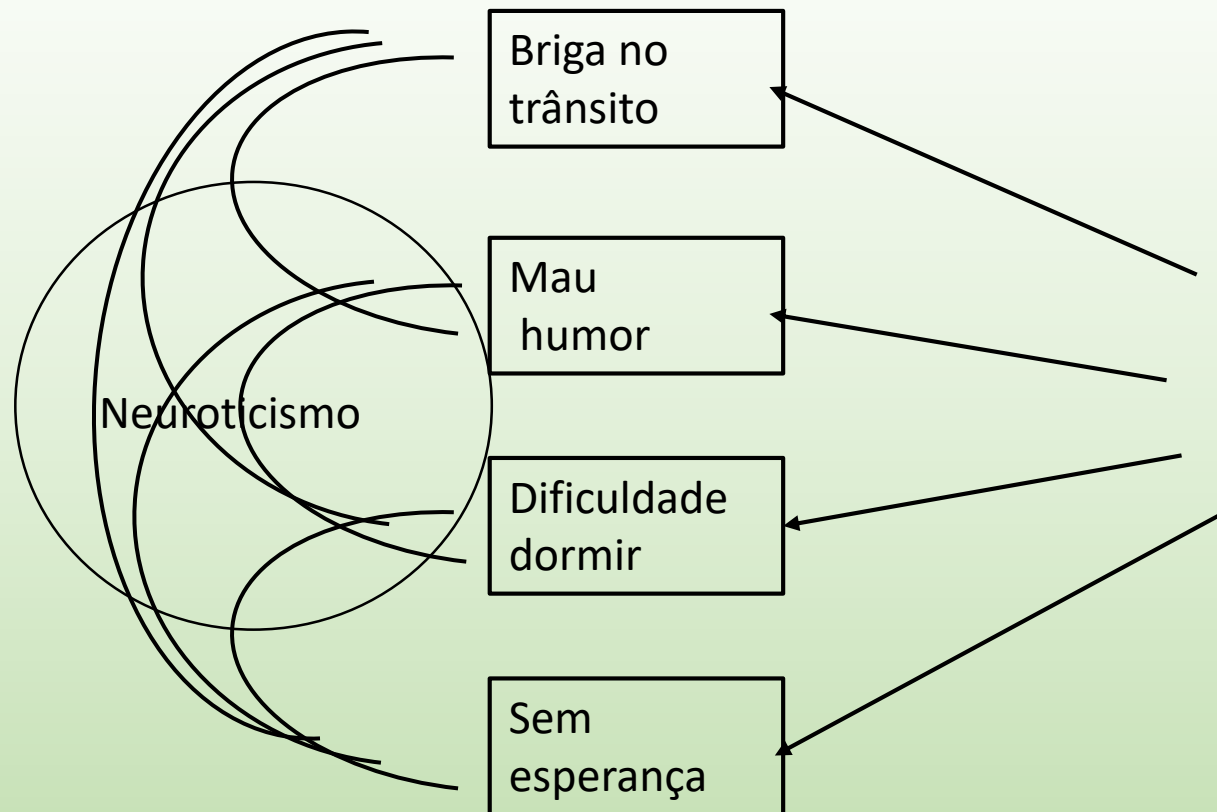


Observado

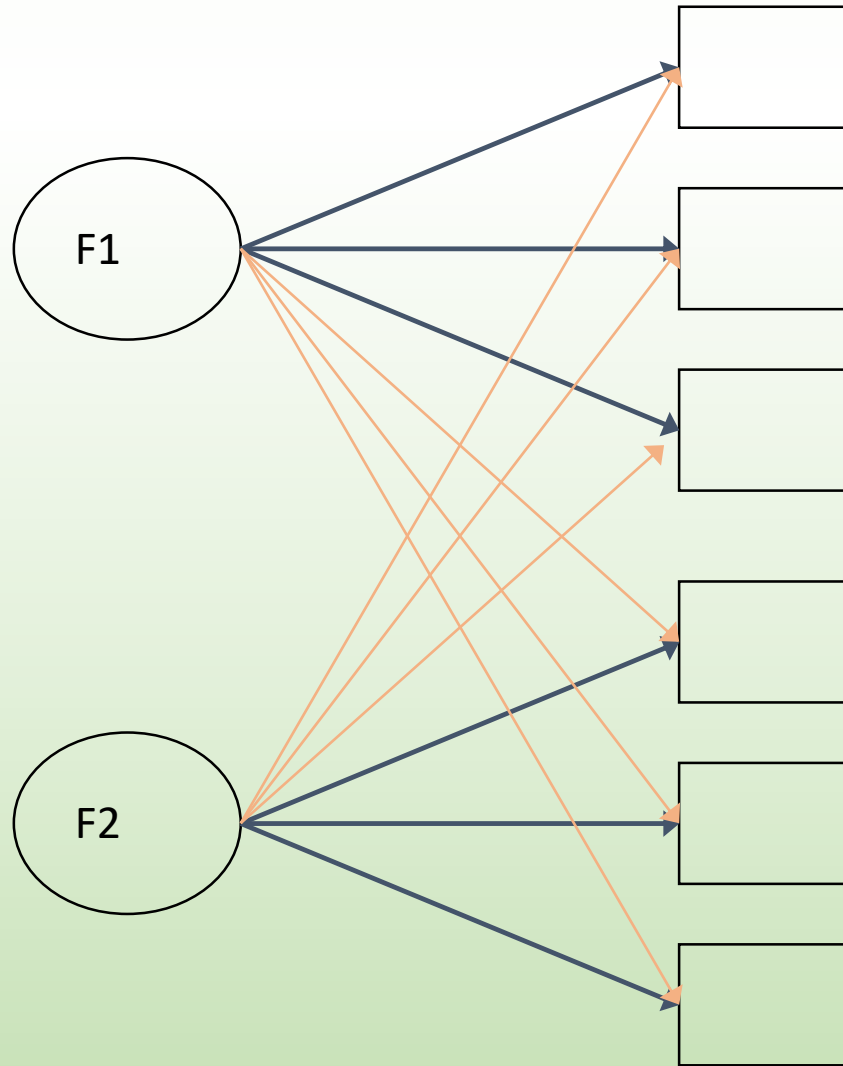
Modelagem de Variáveis Latentes

... Construindo Pontes

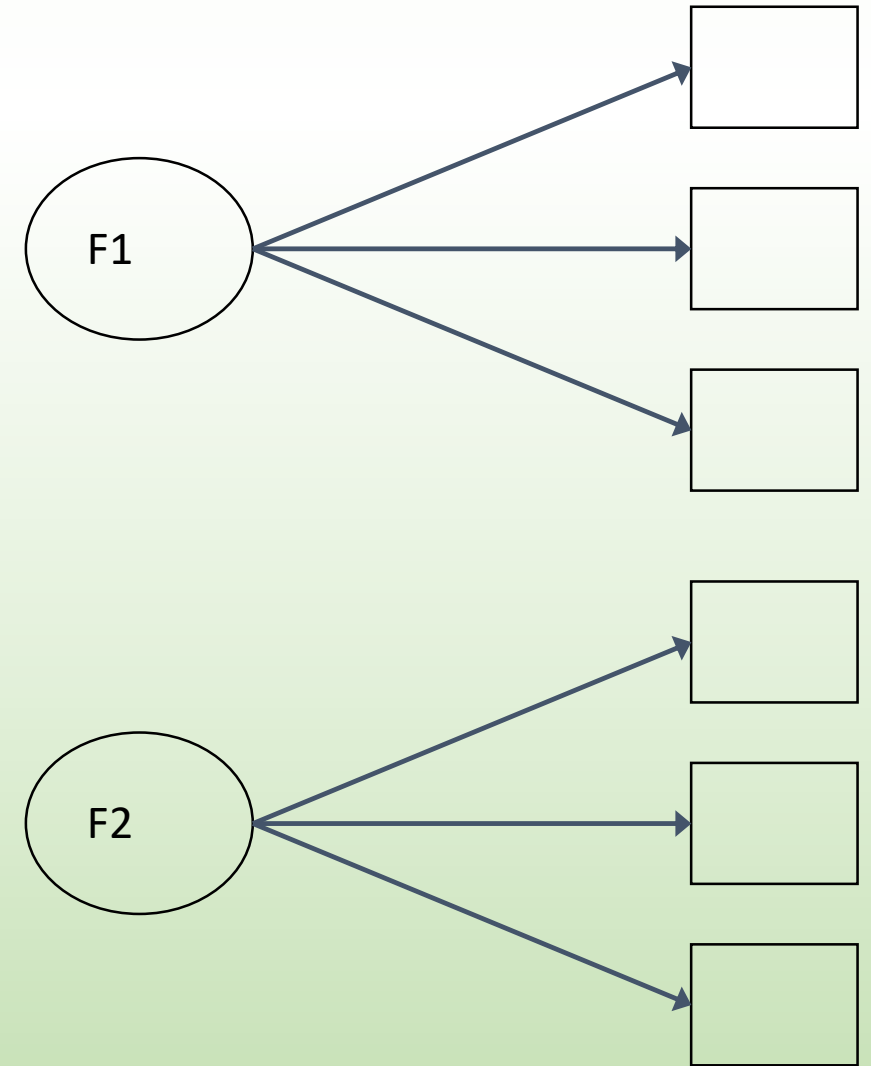
- *Novamente...*
 - Variável latente é uma variável estimada, sem observação direta disponível

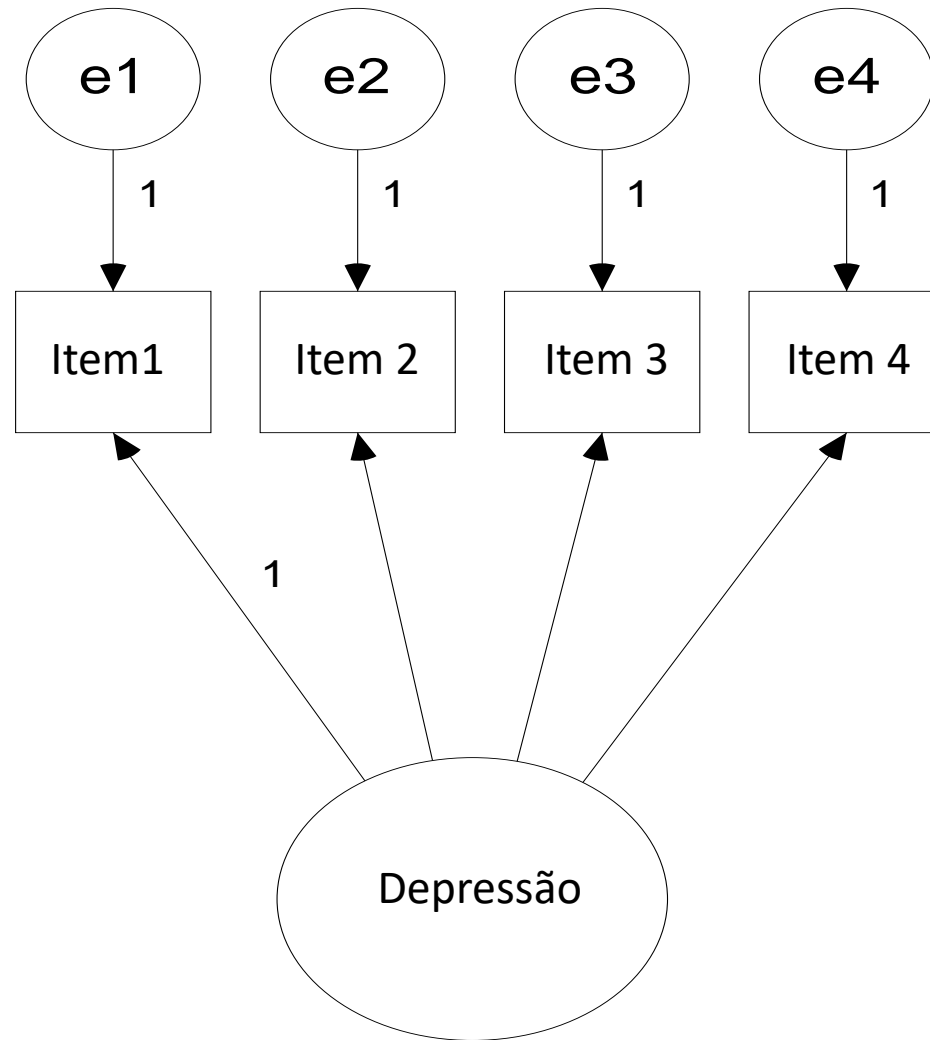


AFE



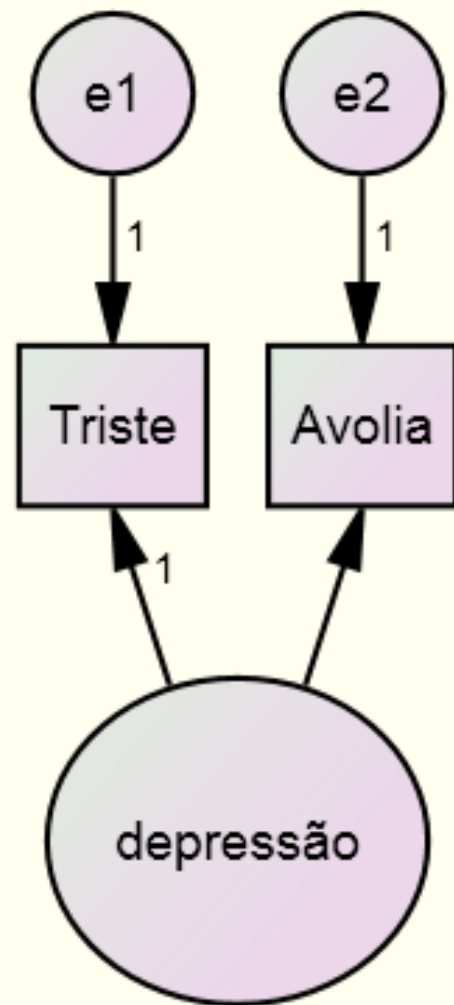
AFC





Exemplo 1 no lavaan

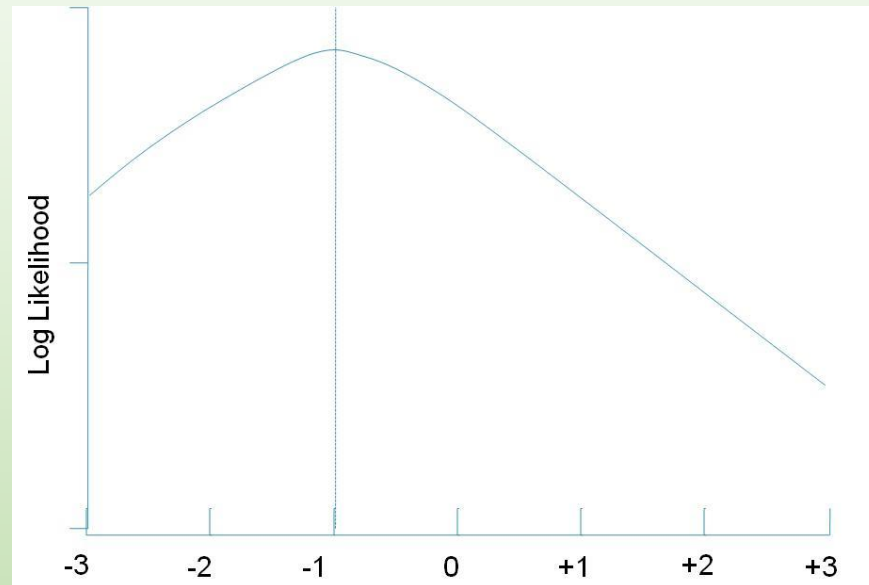
Identificação do Modelo



Exemplo 2

O modelo é identificado se os parâmetros podem ser computados

Se infinitas possibilidades de fixar pesos e parâmetros são plausíveis, os parâmetros são matematicamente indeterminados – modelo não identificado (*underidentified*)



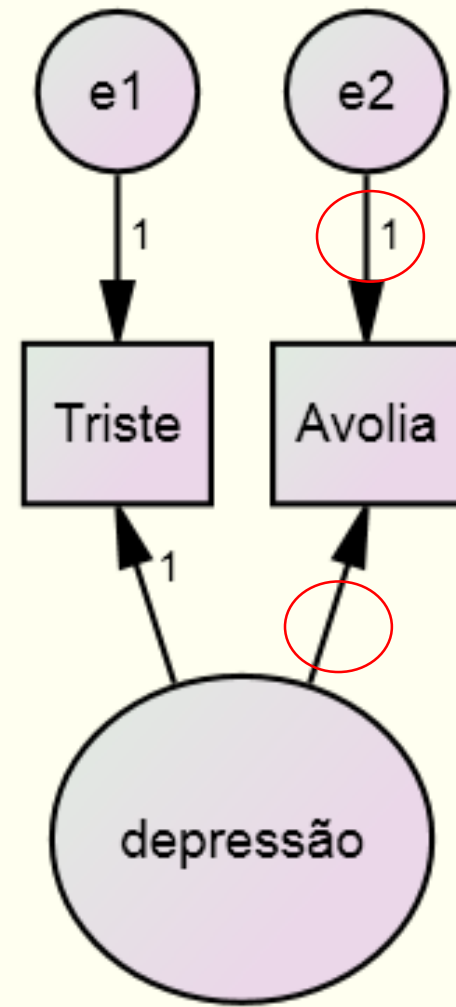
$$\frac{[p \cdot (p + 1)]}{2}$$

Pontos de Informação
Graus de liberdade Total
Sample moments

Quanto dinheiro eu tenho na minha poupança para
gastar com a estimação do meu modelo ?

- Quanto custa para estimar o meu modelo ????

Cada parâmetro (ex: pesos, coeficiente, variância ou covariância) estimado utiliza 1 grau de liberdade (*sample moment*).

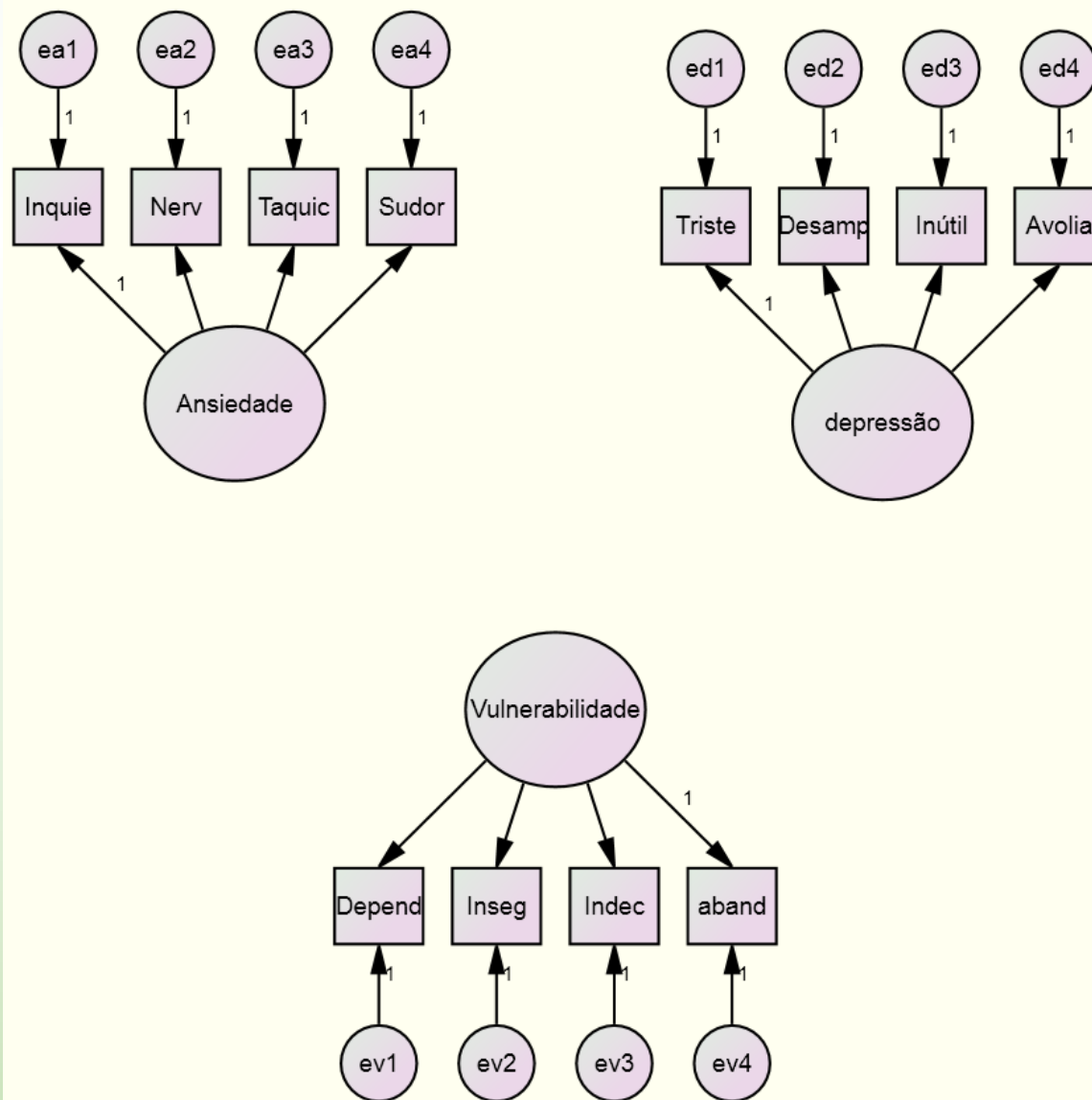


Fixado (Constrained)

Livre para ser
estimado

Exemplo 2

- Overidentified



Fit

✓ Os dados ajustam-se ao modelo?

$$\Sigma \approx S$$

Matriz sigma-hat \approx Matriz S

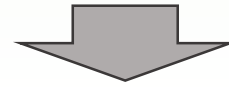
O quanto a matriz das correlações (ou covariâncias) imposta pelo modelo está próxima da matriz dos dados originais.

\neq entre Σ e S = Matriz residual

	Item1	Item2	Item3
Item1			
Item2	Cov		
Item3	Cov	Cov	

	Item1	Item2	Item3
Item1			
Item2	Cov		
Item3	Cov	Cov	

Chi-quadrado pode ser utilizado para acessar o tamanho da discrepância entre Σ e S (confirmar a hipótese nula)



Entretanto o Chi-quadrado é sensível ao tamanho da amostra e a não-normalidade multivariada.



Ao menos 24 indicadores de ajuste foram propostos nos últimos 20 anos para complementar o χ^2 .



Alguns deles consideram o número de parâmetros estimados



Enquanto outros compara o fit dos modelos estimados a outros modelos relacionados

Sugestões:

$\chi^2/df < 2$ (aceitável < 5 , desejável $< 1,96$)

Comparative-Fit-Index (CFI) $> 0,95; 0,90; 0,8$

Tucker Lewis Index (TLI) $> 0,95; 0,90; 0,8$

Root Mean Square Error Approximation (RMSEA) $< 0,05; 0,10$

Quando os indicadores de fit não são muito bons...

SEM também fornece informações de onde potenciais erros de especificação foram cometidos.

Dois erros comuns:

1- Estimar parâmetros inúteis.

Solução: verificar o valor de t . Valores maiores do que 2, normalmente, são significativos. Se *razão crítica* < 2 , rever o parâmetro.

2- Deixar de estimar parâmetros importantes.

Solução: verificar os indicadores de modificação (*modification indices*). Quanto diminuiria o valor de χ^2 caso o parâmetro fosse estimado.

Beba com moderação!!! - Pode-se gerar modelos não replicáveis. Mudanças devem ser feitas quando se encontra uma justificativa teórica

Parâmetros estimados

Coeficientes não padronizados

Os parâmetros são estatisticamente significativos?

Em outras palavras, devo mantê-los no modelo ?

Coeficientes padronizados

Mantêm uma métrica única – permite comparações (no mesmo modelo).

Teoria de Estimação de Parâmetro

Maximun Likelihood (máxima verossimilhança - ML)

Generalized least squares (mínimos quadrados - GLS)

Asympotically distributio-free (ADF)

Robust Weighted Least Squares (WLSMV)

Diferenças nas suposições e propriedades teóricas

ML e GLS pressupõe, por exemplo, normalidade multivariada,
ADF e WLSMV não.

- Maximun Likelihood (máxima verossimilhança - ML)
- Maximiza a probabilidade que os dados (observados) tenham sido retirados da população
- Estimação iterativa – estima parâmetros simultaneamente
 - Começa o processo com ‘parâmetros iniciais’ (‘priors’)
 - A cada iteração, calcula-se novos parâmetros que melhoram a estimação
 - Portanto, a cada iteração é maximizada a probabilidade que os dados tenham sido retirados da população (ajuste do modelo).
 - Convergência do modelo: ponto no qual as novas estimativas de parâmetros não aumentam consideravelmente o ajuste do modelo.

- Métodos de estimação
- WLSMV – Robustustus Weighted Least Squares (ou Weighted least Squares Mean and Variance Adjusted)
 - Usa apenas a diagonal da matriz empírica
 - Estudo de simulação – Muthen- 150 – 200 participantes (para 10 a 15 indicadores)
- Indicadores podem ser ordinais ou binários

- Exemplo 4, itens:

Item 1	Alegre
Item 2	Comunicativo
Item 3	Falante
Item 4	Envolvente
Item 5	Bondoso
Item 6	Belo
Item 7	Amável
Item 8	Gentil
Item 9	Convencional
Item 10	Criativo
Item 11	Doidão
Item 12	Inteligente

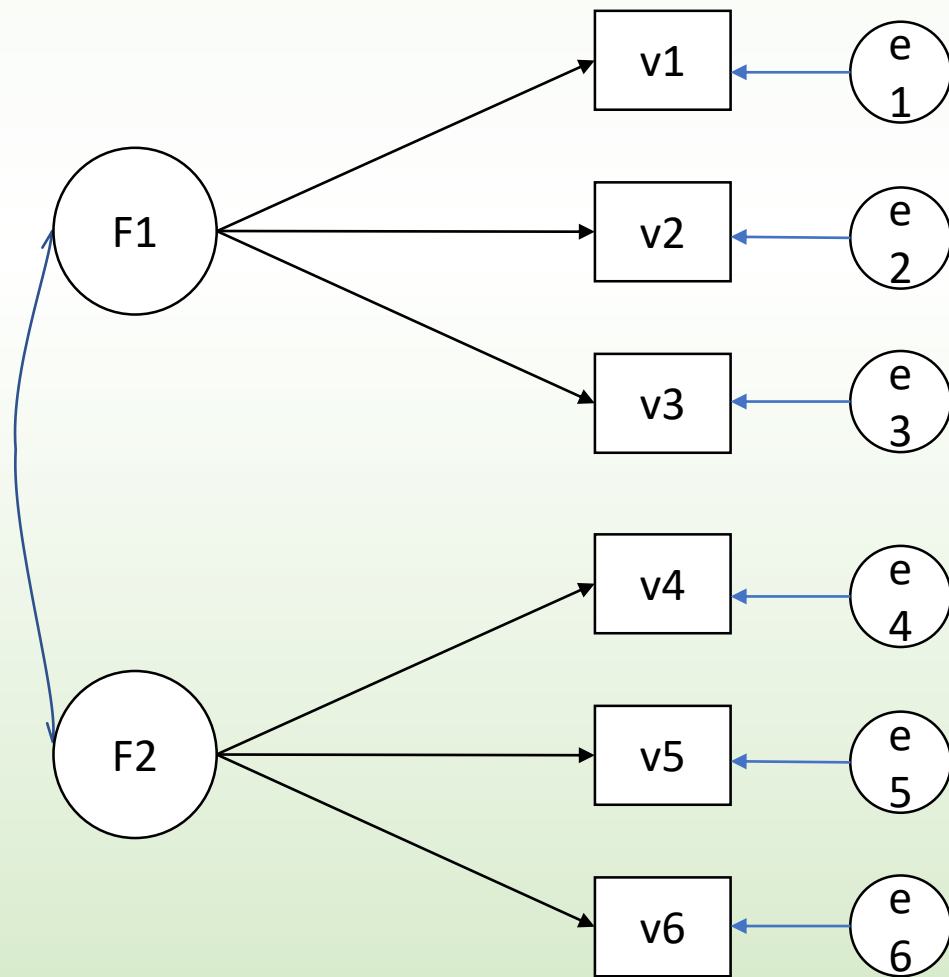
Sintaxe

- Duas partes
 - Modelo dentro de um objeto
 - Rodar a função

```
cfa(banco_de_dados, objeto_do_modelo)
```

Sintaxe para o modelo

- Correlação
 - $X \sim Y$
- Regressão
 - $Y = X$
- Fator e indicadores
 - $F1 \sim v1 + v2 + v3$
- Correlações residuais entre itens
 - $v1 \sim v2$
- Variância
 - $F1 \sim F1$



$$F1 = \sim v1 + v2 + v3$$

$$F2 = \sim v4 + v5 + v6$$

$$F1 \sim\sim F2$$

Se os fatores forem ortogonais:

$$F1 \sim\sim 0 * F2$$

Se quisermos correlacionar os resíduos dos itens 5 e 6

$$F1 = \sim v1 + v2 + v3$$

$$F2 = \sim v4 + v5 + v6$$

$$F1 \sim\sim F2$$

$$v5 \sim\sim v6$$

Se quisermos liberar um parâmetro

$$F1 = \sim NA * v1 + v2 + v3$$

Exercícios

Baixar do github

Para além do básico...

Bootstrapping

- ✓ *Computer-Intensive Techniques*
- ✓ O computador cria diversas sub-amostras do banco original e repete as análises em cada uma.
- ✓ É utilizado em outros testes e técnicas estatísticas.

Bootstrapping

- ✓ Por que utilizar?

1. Quanto mais não normal forem os dados maior o χ^2 da ML e da GLS
2. Amostras pequenas tendem a inflar o χ^2
3. Dados não normais subestimam o TLI e CFI.
4. Dados não normais podem causar baixos EP.

- ✓ Quando a distribuição é não normal as estimativas de *bootstrapping* são menos enviesadas do que as ML.

Entretanto... Cuidado!!!

- ✓ Lembre-se que a sub-amostra é apenas uma parte da amostra original. Esta amostra original deve:
 - ✓ Ser representativa da população
 - ✓ Independência das observações
 - ✓ Quando os dados são normais o procedimento produz mais vieses do que outros métodos de estimação.
 - ✓ O procedimento não 'tapa furos' de amostra pequenas

Em suma... Nada de milagres!!!

lavaan apresenta duas maneiras:

- usar os argumentos `se = "bootstrap"` OU `test = "bootstrap"` dentro da função `cfa`

```
cfa(model = exemplo5, data=data_exemplo,  
     se="bootstrap", bootstrap=100)
```

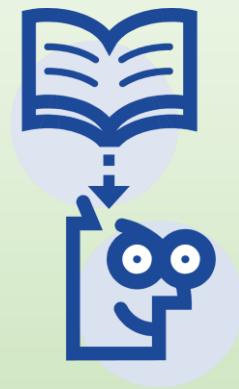
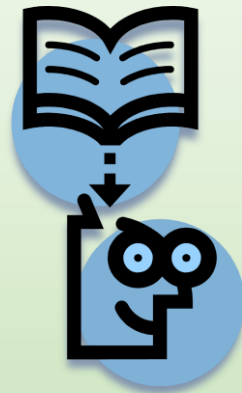
- usar a função `bootstrapLavaan()` para um modelo já estimado

```
bootstrapLavaan(cfa_ex5b, R=100)
```


Multigrupos

A estrutura do instrumento e os parâmetros variam de população para população ?

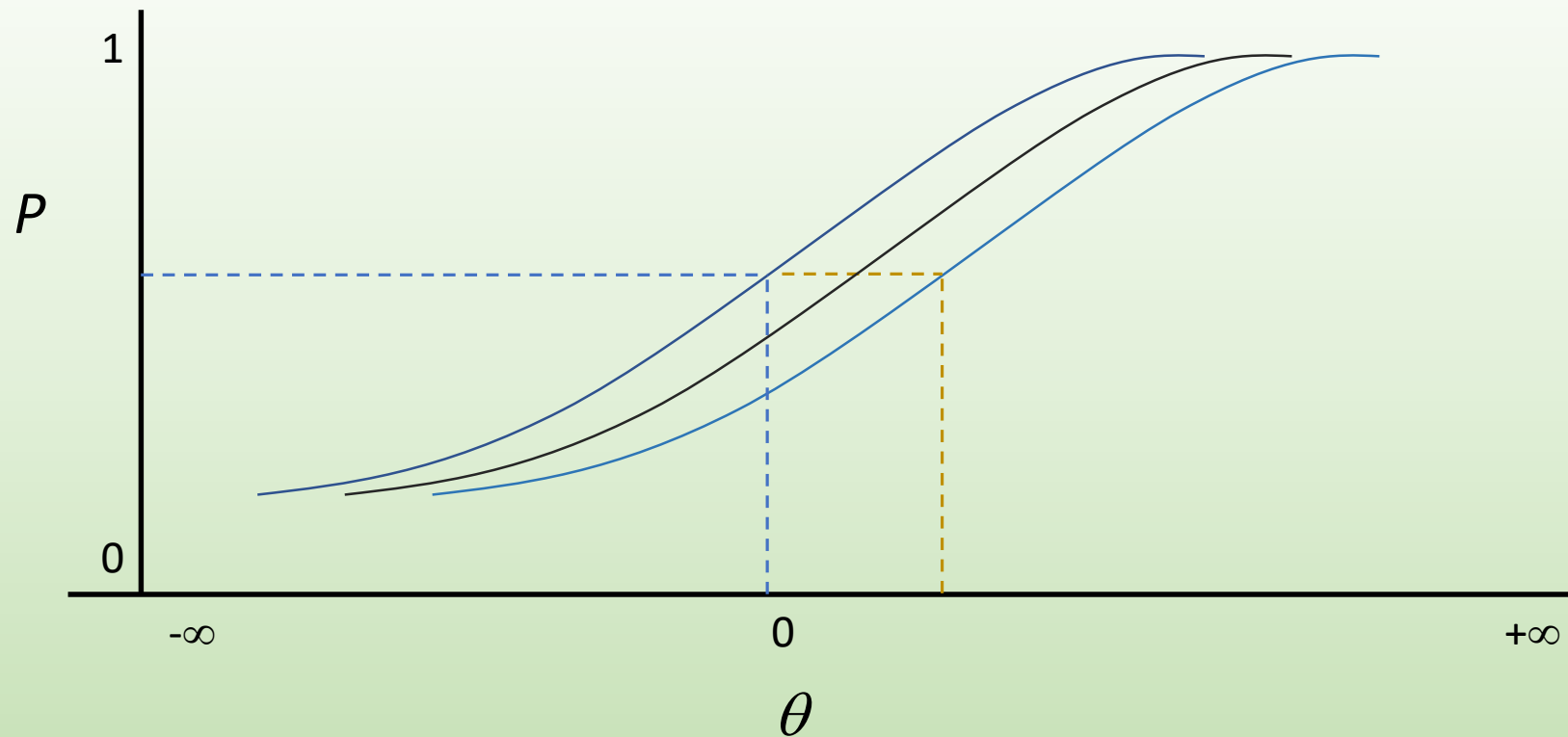
Se: θ_m = θ_t



Deve... P (item 1) = P (item 1)

Multigrupos

A estrutura do instrumento e os parâmetros variam de população para população ?



Tipo	Definição (Fixar)
1 - Invariância Configural	Número de itens e fatores
2 – Invariância Métrica (Weak Invariance)	Número de itens e fatores + cargas fatoriais
3 – Invariância Escalar (Strong Invariance)	Número de itens e fatores + cargas fatoriais + Interceptos
4 – Invariância Residual (Strict Invariance)	Número de itens e fatores + cargas fatoriais + Interceptos + variância residual

lavaan:

Usar o argumento 'group' na função CFA

```
cfa(model = exemplo5, data=data_exemplo,  
    group="variável_do_grupo",  
    group.equal=c("loadings" , "intercepts", "residuals"))
```

OU usar a função (do pacote semTools)

```
measurementInvariance(model = exemplo5,  
    data=data_exemplo, group="variável_do_grupo")
```

lavaan:

Para testar invariância parcial, acrescentar o argumento **group.partial** na função.

```
measurementInvariance(model = exemplo6,  
  data=data_exemplo, group= "gênero",  
  group.partial=c("f1 =~ A2", "D2~1"))
```



Obrigado!