

结构方程模型

- 1 结构方程模型 (SEM) 基础介绍
 - 1.1 基本介绍
 - 1.2 不同方法的对比
 - 1.3 三种变量
 - 1.4 两种模型
- 2 论文中结构方程模型的基本步骤
- 3 Amos 中绘制结构方程模型操作
- 4 结构方程模型样本数据准备与导入
- 5 样本数据的信度分析
- 6 样本数据的效度检验
 - 6.1 探索性因子分析 (EFA)
 - 6.1.1 操作步骤
 - 6.1.2 观察指标
 - 6.2 验证性因子分析 (CFA)
 - 6.2.1 结构效度 (模型适配度)
 - 1 基本假设
 - 2 参数指标
 - 3 操作步骤
 - 6.2.2 组合效度 / 收敛效度
 - 6.2.3 区别效度 / 区分效度
- 7 结构方程模型适配度指标
- 8 结构方程模型的修正与迭代
- 9 路径分析结构解读
 - 9.1 解读说明
 - 9.2 非标准化系数
 - 9.3 非标准化系数
- 10 模型与统计指标输出与解读

1 结构方程模型 (SEM) 基础介绍

1.1 基本介绍

结构方程模型 (Structural Equation Modeling) 是基于变量的协方差矩阵来分析变量之间关系的一种统计方法. 这种方法适用于研究

- 1. 多个因变量受多个自变量的影响.
- 2. 一个自变量通过什么途径影响到因变量.
- 3. 自变量对因变量的影响有何不同等多种复杂关系.

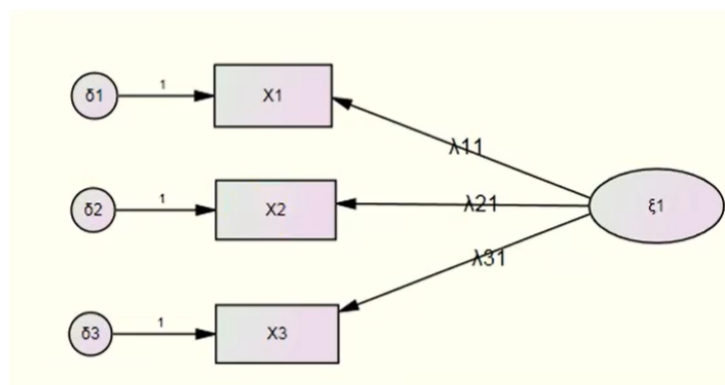
1.2 不同方法的对比

方法	描述
回归分析	多个自变量对一个因变量

方法	描述
路径分析	多个自变量对多个因变量
结构方程模型	路径分析 + 测量关系
探索性因子分析 (EFA)	测量关系 (非经典量表)
验证性因子分析 (CFA)	测量关系 (经典量表)

1.3 三种变量

1. **潜在变量** (latent variable), 即无法观测的变量. 该变量用圆形 (椭圆形) 表示.
 1. 外因潜在变量: 作为因, 用 ζ 表示.
 2. 内因潜在变量: 作为果, 用 η 表示.
2. **观测变量** (manifest variables), 又称为指标变量 (indicator variables) 或可测量变量 (measured variables). 该变量用矩形表示.
3. **残差变异项**: 用圆形表示, 记为 ζ .



1.4 两种模型

1. **测量模型**: 由潜在变量和观测变量则称.
2. **结构模型**: 潜在变量间因果关系模型的说明.

2 论文中结构方程模型的基本步骤

1. 构建理论模型
2. 提出研究假设
3. 给出变量操作定义
4. 样本数据收集
5. 信度与效度检验
6. 模型适配度检验
7. 模型调整与修正
8. 路径系数分析

3 Amos 中绘制结构方程模型操作

如果遇到没有出现乱码, 但是在 Amos 输入中文就会报错的情况, 可以参考如下做法:

系统设置 - 语言 - 管理语言设置 - 更改系统区域设置, 确保当前系统区域设置为 "中文(简体, 中国)" 后, 取消勾选 "Beta 版" 后重启即可.

注: 可以选中结构方程中的部分内容, 复制后直接粘贴图片到其它软件.

4 结构方程模型样本数据准备与导入

将 Excel 数据导入 SPSS, 保存后再导入 Amos. 注意删除空行.

缺失数据如果较少的话, 可以剔除.

5 样本数据的信度分析

信度 (reliability) 指检测结果的一致性、稳定性和可靠性.

在 SPSS 中: 分析 - 刻度 - 可靠性分析, 其中系数 Cronbach $\alpha \in [0, 1]$, 低于 0.6 一般认为一致信度不足, 0.7 - 0.8 表示具有相当的信度, 0.8 - 0.9 表示信度非常好.

6 样本数据的效度检验

效度 (validity) 即有效性.

6.1 探索性因子分析 (EFA)

6.1.1 操作步骤

SPSS: 分析 - 降维 - 因子分析

- 描述: 勾选 KMO 和 Bartlett 球形度检验.
- 抽取: 勾选碎石图. 调整因子的固定数量.
- 旋转: 基于最大方差; 输出选中载荷图.
- 得分: 选中 "显示因子得分系数矩阵"
- 选项: 缺失值使用均值替换; 按大小顺序; 取消小系数.

6.1.2 观察指标

1. KMO (Kaiser-Meyer-Olkin) 的值大于 0.5 / 0.6 即合格.
Bartlett 球形度检验的值 (Sig.) 小于 0.05 即合格.
2. 解释的总方差: 初始特征值的合计大于 1, 对应旋转平方和载入累积大于 60% / 70%.
3. 碎石图: 寻找 "拐点".
4. 旋转成分矩阵: 大于 0.5 / 0.4 即合格.

6.2 验证性因子分析 (CFA)

检验方面:

1. 结构效度
2. 组合效度 / 收敛效度
3. 区别效度 / 区分效度

6.2.1 结构效度 (模型适配度)

1 基本假设

基本假设: 样本共变异数矩阵 \approx 模型共变异数矩阵

1. 反映模型相似度的指标: GFI, CFI, TLI... > 0.9 (> 0.5 最低, > 0.8 可接受)
2. 反映模型差异性的指标: RMSEA, SRMR < 0.05 (< 0.5 最低, < 0.08 可接受)

2 参数指标

指标名称	临界值	全称
绝对指标		
χ^2 (卡方)	越小越好	
GFI (拟合优度指数)	> 0.9	Goodness of Fit Index
AGFI (修正拟合优度指数)	> 0.9	Adjusted Goodness of Fit Index
RMR (残差均方根)	< 0.05	Root Mean Square Residual
RMSEA (近似误差均方根)	< 0.05	Root Mean Square Error of Approximation
相对指标		
NFI (规范拟合指数)	> 0.9	Normed Fit Index
NNFI (非规范拟合指数)	> 0.9	Non-Normed Fit Index 或 TLI (Tucker-Lewis Index)
IFI (增值拟合指数)	> 0.9	Incremental Fit Index
CFI (比较拟合指数)	> 0.9	Comparative Fit Index
信息指标		
AIC (赤池信息准则)	越小越好	Akaike Information Criterion
ECVI (交叉验证指数)	越小越好	... Cross Validation Index

3 操作步骤

在 AMOS 中, 打开 Analysis Properties 窗口,

- Estimation 窗口
 - 如果勾选 Estimate means and intercepts, 则无法计算 GFI 和 AGFI 的值.
 - 如果取消勾选后报错, 可能是因为 SPSS 导入数据时有缺失值, 可以剔除后重新运行.
 - 如果没有缺失值, 可能是因为从 Excel 导至 SPSS 时有空行, 去 SPSS 删了重新导入到 AMOS 即可.
- 在 Output 窗口
 - 建议所有选项全部勾选.
 - 如果报错要求取消勾选 Modification indices, 则可能是有缺失值或空行.

注: 中间一栏中, Iteration 次数如果是 1 或 49 则说明有问题.

6.2.2 组合效度 / 收敛效度

组合效度 (Composite Reliability, CR)

- 构面内部一致性, 相当于 Cronbach's α .
- CR 越高表示构面内部一致性越高, 越收敛.
- 大于 0.7 即可.

平均编译数萃取量 AVE (Average of Variance Extracted)

- 潜在变量对观察变量解释能力的平均.
- AVE 越高, 收敛效度越高.
- 大于 0.5 即可 (0.36 - 0.5 为接受门槛).

$$AVE = \frac{\sum \lambda_i^2}{\sum \lambda_i^2 + \sum \delta(\varepsilon_i)} = \frac{\sum L_i^2}{n}.$$

6.2.3 区别效度 / 区分效度

7 结构方程模型适配度指标

8 结构方程模型的修正与迭代

整体模型适配度 \approx 每个测量模型适配度 (CFA) 的均值

因子载荷量 > 0.5 则合格.

SMC > 0.36 则合格.

残差不独立: 勾选 Modification indices, 在 view text 窗口中查看相关性.

9 路径分析结构解读

9.1 解读说明

非标准化系数一般用于查看有无显著性.

标准化系数则是用于查看重要性.

9.2 非标准化系数

- 违犯估计
- 误差项的估计系数不能为负数
- 因子负荷量和结构路径系数显著
- 自变量增加一个单位, 因变量增加几个单位

9.3 非标准化系数

- 因子载荷量 > 0.6 (0.7 以上理想; 探索性问卷 0.6), 不能过于接近 1 (原则上不超过 0.95).
- 测量模型的 SMC (题目信度) > 0.36 , 不要超过 0.9.
- 结构模型的 SMC: 0.19 (small), 0.33 (medium), 0.67 (large) (主要看显著性)
- 自变量每增加一个标准差, 因变量增加几个单位.

10 模型与统计指标输出与解读

SEM 使用注意事项

- 样本规模: 中 / 大规模样本, 150+
- 每个潜在变量至少对应 3 个观察变量, 5-7 个为佳.
- 每个观察变量的测量尺度颗粒越细越好, 李克特 7 级量表.
- 潜在变量尽量控制在 7 个以内, 测量题目尽量在 20 个左右.
- 每个观察变量一般只能对应一个潜在变量 (Cross-loading); 如果对应两个的话, 运行之后如果系数小于 0.4, 则建议删去这个.
- 自变量一般都具有相关性.
- 每个潜在变量都需要设置一个参考指标, 因素负载量为 1 (剩余为估计参数)
- 尽量采用一阶模型, 或者引用经典二阶模型.