



结构方程模型

城市分析方法系列课程

苏州大学 王灿

如何掌握一个定量分析方法？

基本思想

了解big picture: 它是什么，能解决什么问题

技术细节

清楚它具体是怎么实现的、相关的公式和数据论证

工具软件

知道输入输出，会操作，必要时写代码

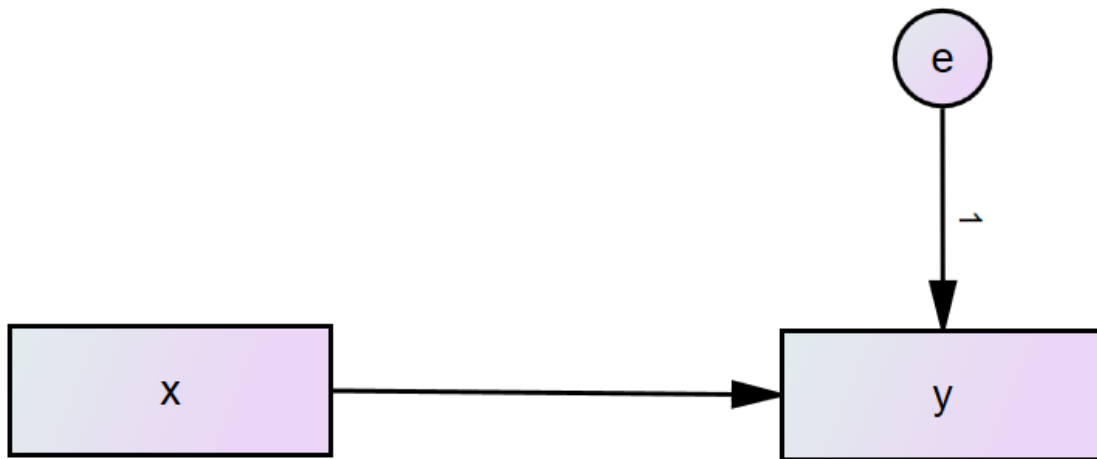
实践应用

能够在实践中合理地应用

课程政策

- 所有课件均会提供，请按需做笔记
- 如有问题，请务必提问
- 小作业：
 - 会做：一个简单的指定数据练习
 - 看得懂：一篇自选文献阅读

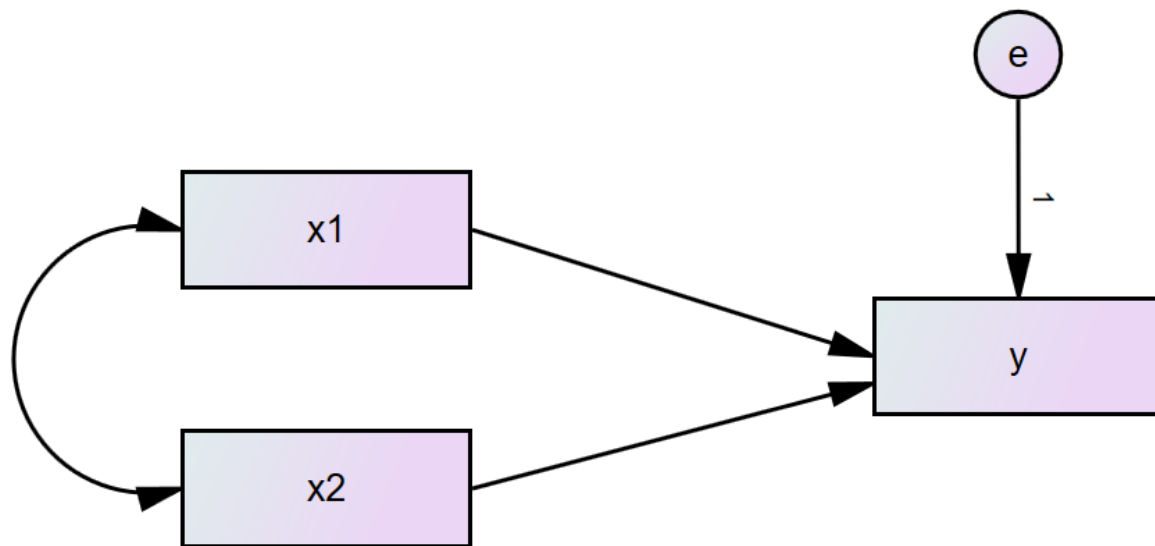
回顾：线性回归和路径分析



一元线性回归

$$y = a + bx + e$$

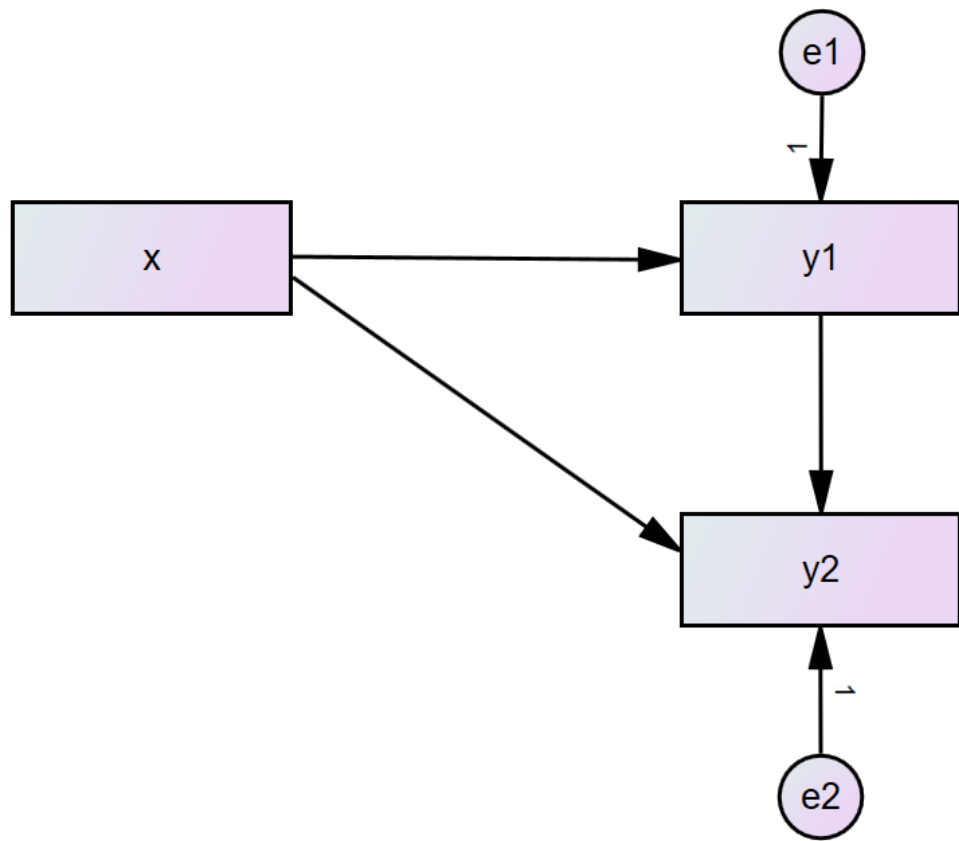
回顾：线性回归和路径分析



多元线性回归

$$y = a + b_1x_1 + b_2x_2 + e$$

回顾：线性回归和路径分析

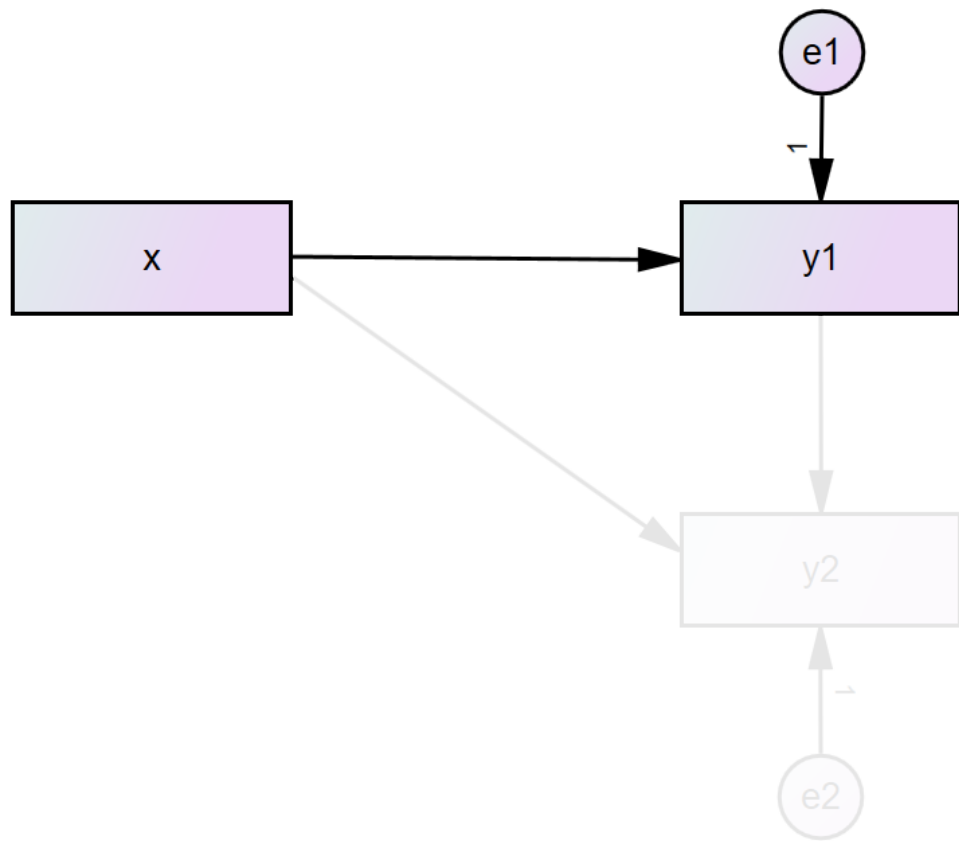


复杂的线性回归

$$y_1 = a_1 + bx + e_1$$

$$y_2 = a_2 + b_1y_1 + b_2x + e_2$$

回顾：线性回归和路径分析

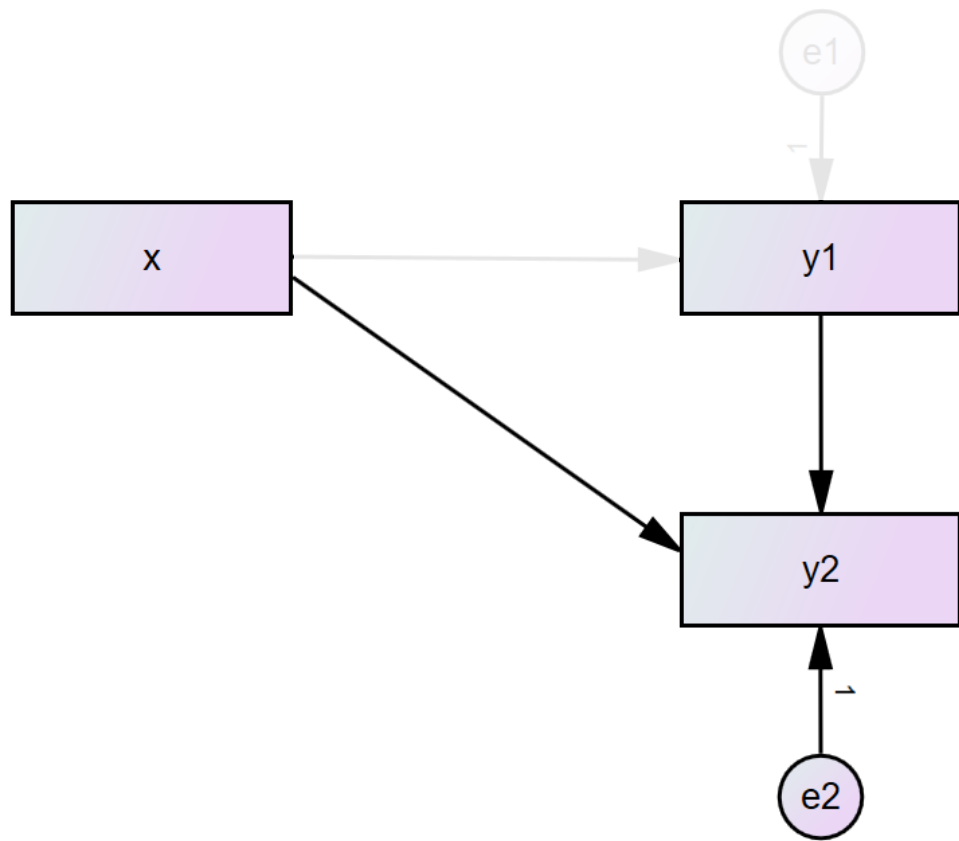


复杂的线性回归

$$y_1 = a_1 + bx + e_1$$

$$y_2 = a_2 + b_1y_1 + b_2x + e_2$$

回顾：线性回归和路径分析



复杂的线性回归

$$y_1 = a_1 + bx + e_1$$

$$y_2 = a_2 + b_1y_1 + b_2x + e_2$$

某种意义上，

“线性回归” 是 “路径分析” 的特例，

“路径分析” 是 “结构方程” 的特例。

path diagram

“路径分析” 是用**路径图**表示的复杂 “线性回归” ,

“结构方程” 是可以使用**潜变量**的 “路径分析” 。

latent variable / construct

大纲

- 什么是潜变量？
- 结构方程模型能做什么？
- 结构方程模型如何工作？
 - 测量模型
 - 结构模型
- 如何在AMOS中估计结构方程模型？

什么是潜变量?

苏州市2022年**GDP**为23958.3亿元

苏大的**留学生人数**为1273人

某小区去年发生的**盗窃案数量**为35起

张三的**城市设计课分数**为90分



能够被直接测度的**可观察变量**

什么是潜变量?

苏州市的**城市竞争力**?

苏大的**国际化水平**?

某小区居民的**安全感**?

张三的**设计能力**?



不能够被直接测度的**潜变量**

什么是潜变量?

可观察变量

可以被直接测度

observable variable 可观察变量

manifest variable 显变量

indicator 参照

可以直接测度吗?

潜变量

无法直接测度, 需借助可观察变量反映

unobservable variable 不可观察变量

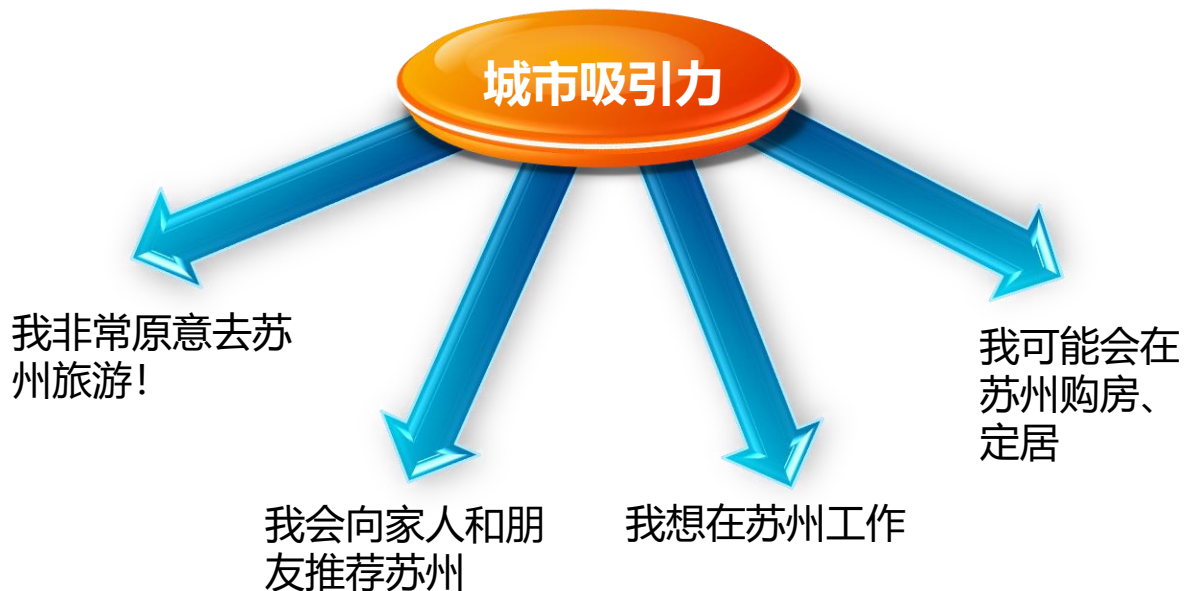
latent variable 潜变量

construct 概念/构面

什么是潜变量？

如何测度潜变量？

如果苏州的城市吸引力很高，那么会有哪些表现呢？



利用问卷/量表

你同意以下说法吗？

得分	分类表述
1	完全不同意
2	较不同意
3	略不同意
4	不清楚
5	略同意
6	较同意
7	完全同意

结构方程模型能做什么？

结构方程模型

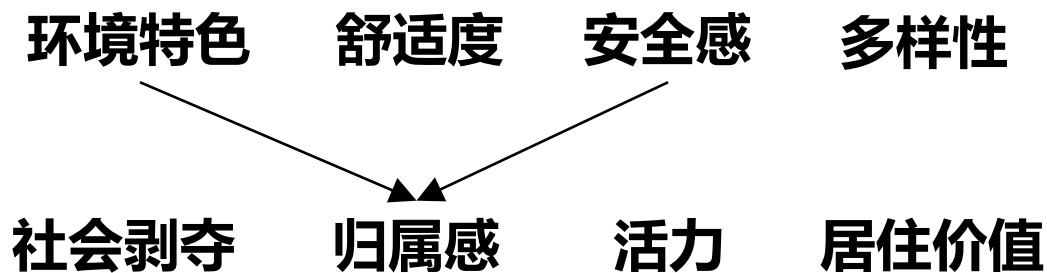
(structural equation model, SEM)

可以被认为使用**潜变量**的**路径分析**。

结构方程模型能做什么？

为什么要用SEM？

- 社会学、城市研究中有很多潜变量，它们是复杂多面的。
- 我们关心这些潜变量之间的关系。



结构方程模型能做什么？

为什么要用SEM？

- 传统统计方法存在问题。

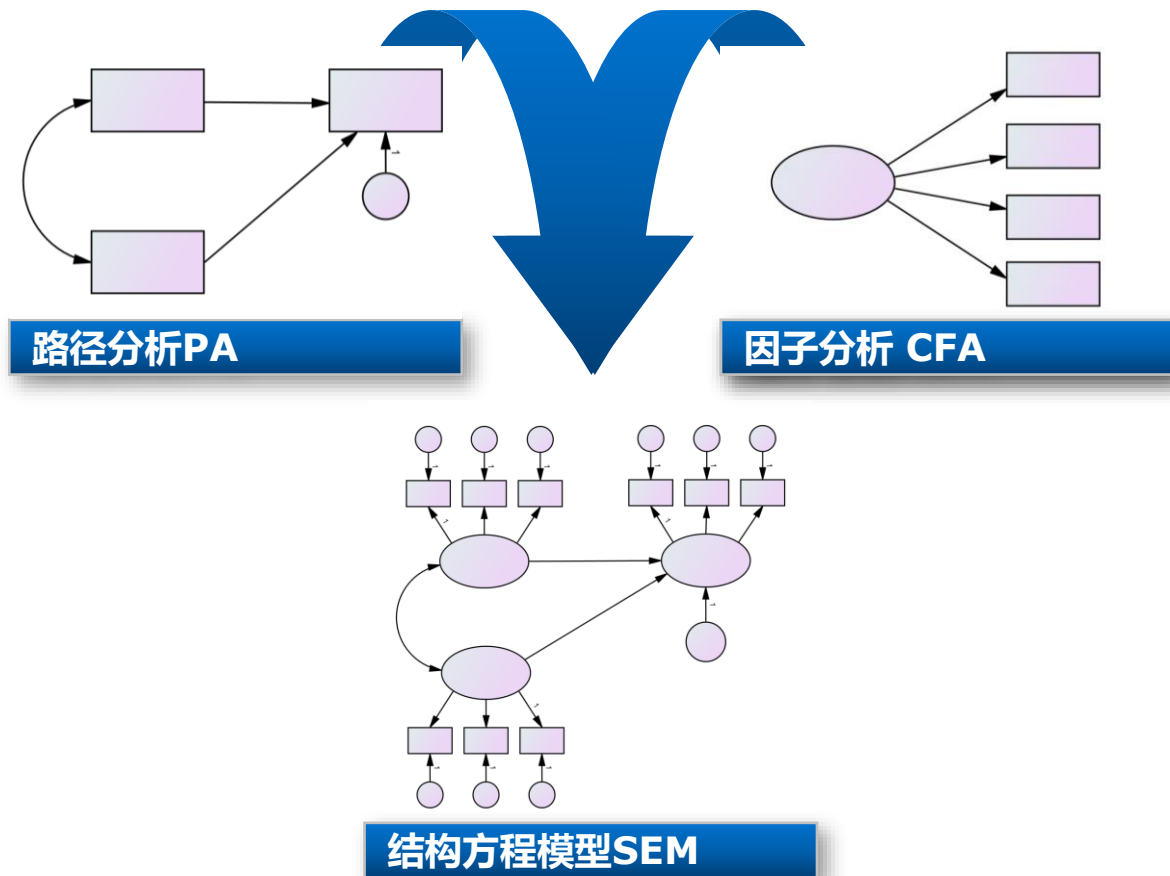


结构方程模型能做什么？

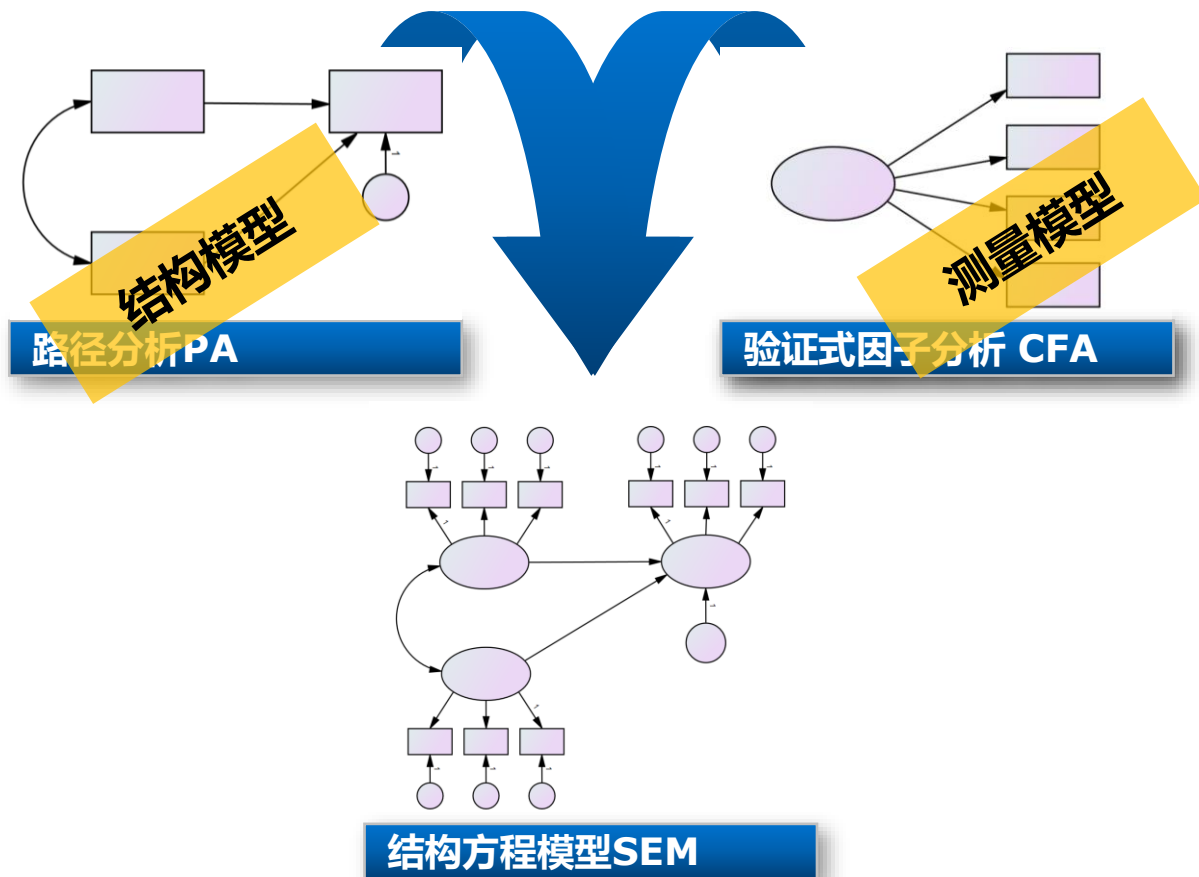
SEM特别适用于：

- 需要对复杂、多面的概念建模时；
- 需要同时考虑一系列的关系时，而不仅仅是一个因变量和几个自变量关系时。

结构方程模型是如何工作的？



结构方程模型是如何工作的？



结构方程模型是如何工作的？

- X1 我们小区是经过精心设计的。
- X2 我们小区的住宅楼别具一格。
- X3 我们小区的户外景化与众不同。
- X4 大家从不会混淆我们小区和其他小区。
- X5 我们小区很少发生犯罪事件。
- X6 我不担心家中财物被盗。
- X7 我们小区的安保服务令人放心。
- X8 即使深夜回家，我也不觉得害怕。
- X9 我对我们社区有很深的依恋。
- X10 我很了解我们社区的情况。
- X11 我希望付出时间，为社区做贡献。
- X12 我和社区里的邻居相处得很愉快。

您同意X1-X12的说法吗？

- 1. 完全不同意
- 2. 较不同意
- 3. 略不同意
- 4. 不清楚
- 5. 略同意
- 6. 较同意
- 7. 完全同意

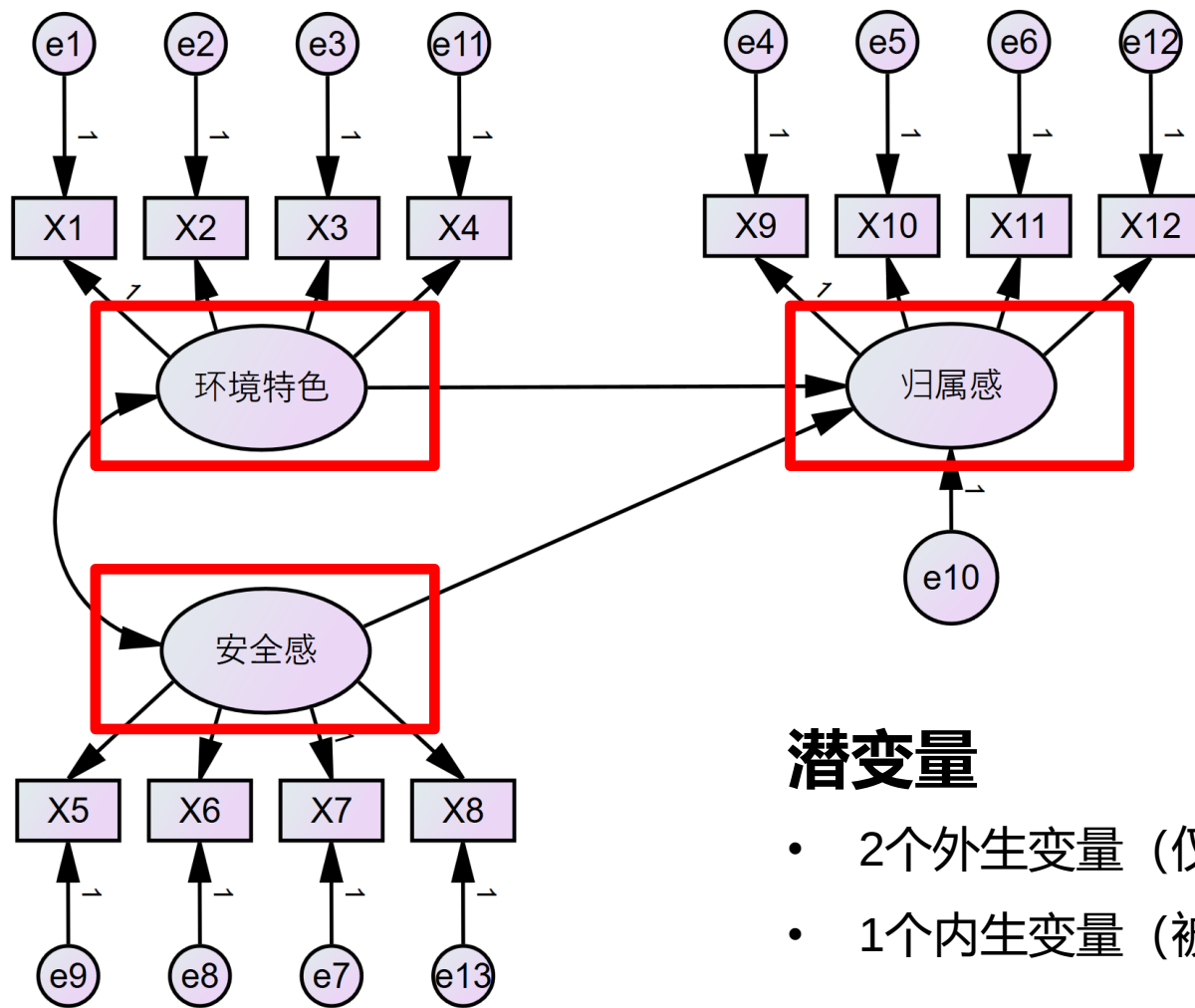
结构方程模型是如何工作的？

X1	我们小区是经过精心设计的。	5
X2	我们小区的住宅材料别具一格。	4
X3	我们小区的绿化与众不同。	7
X4	大家从不会混淆我们小区和其他小区。	6
X5	我们小区很少发生犯罪事件。	6
X6	我不担心家中财物被盗。	4
X7	我们小区的安保服务令人放心。	3
X8	即使深夜回家，我也不觉得害怕。	3
X9	我对我们社区有很深的依恋。	5
X10	我很了解我们社区的情况。	4
X11	我希望付出时间，为社区做贡献。	5
X12	我和社区里的邻居相处得很愉快。	3

您同意X1-X12的说法吗？

- 1. 完全不同意
- 2. 较不同意
- 3. 略不同意
- 4. 不清楚
- 5. 略同意
- 6. 较同意
- 7. 完全同意

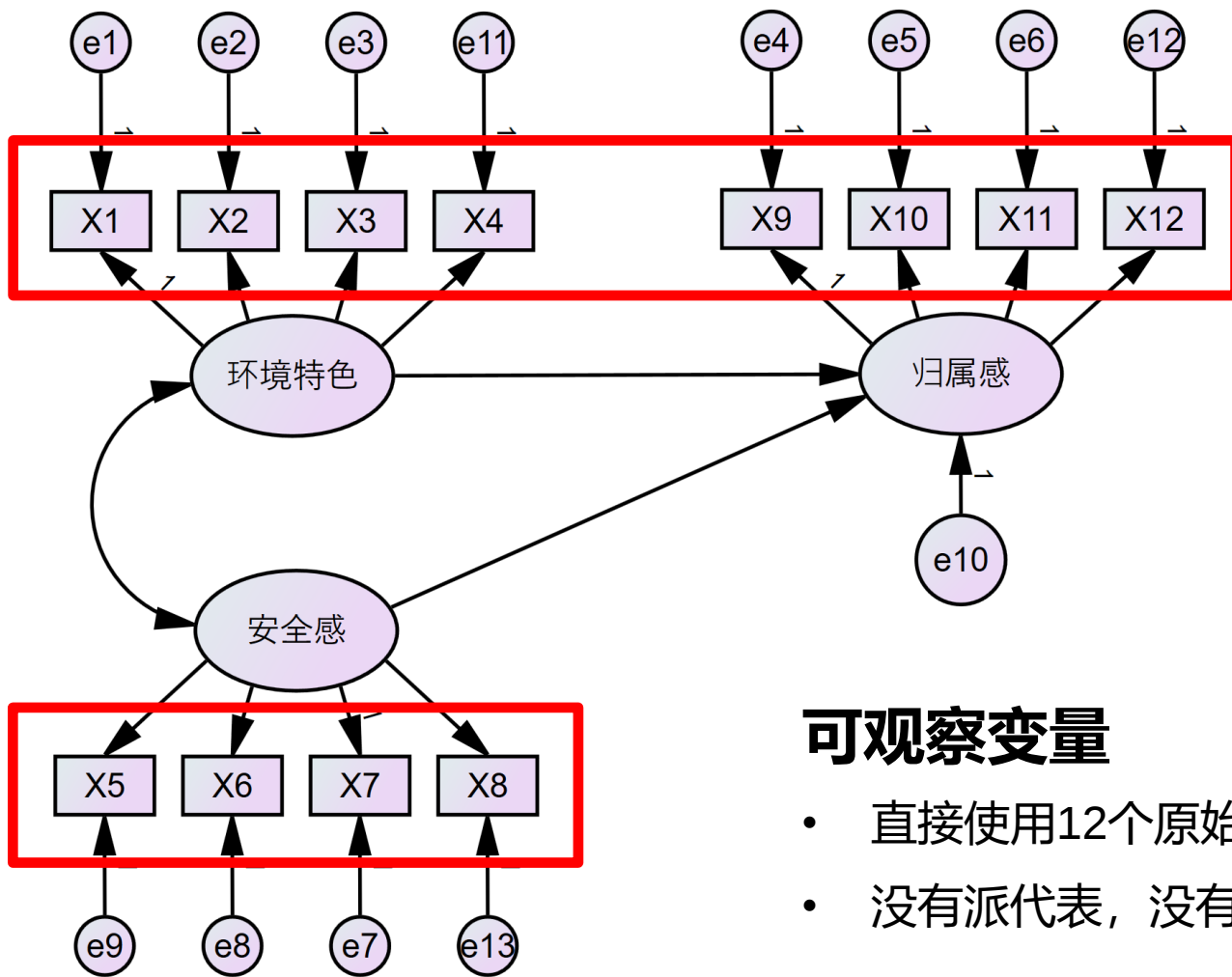
结构方程模型是如何工作的？



潜变量

- 2个外生变量 (仅解释其他变量)
- 1个内生变量 (被其他变量解释)

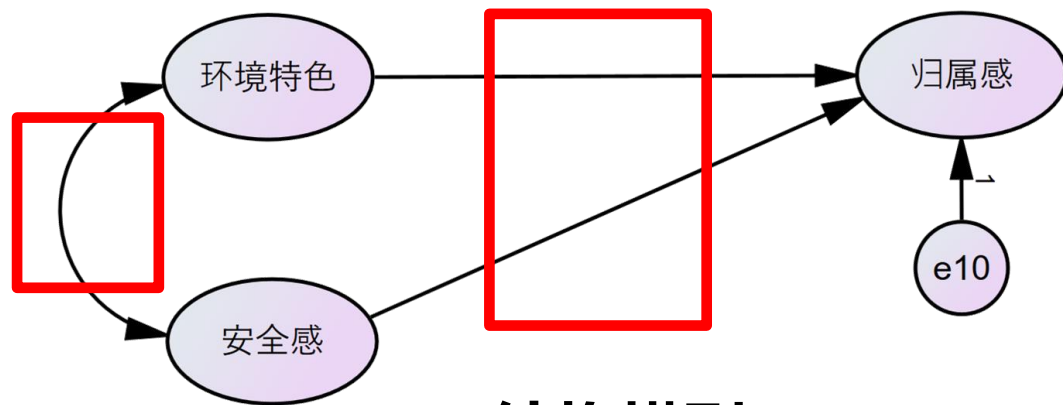
结构方程模型是如何工作的？



可观察变量

- 直接使用12个原始变量
- 没有派代表，没有求平均

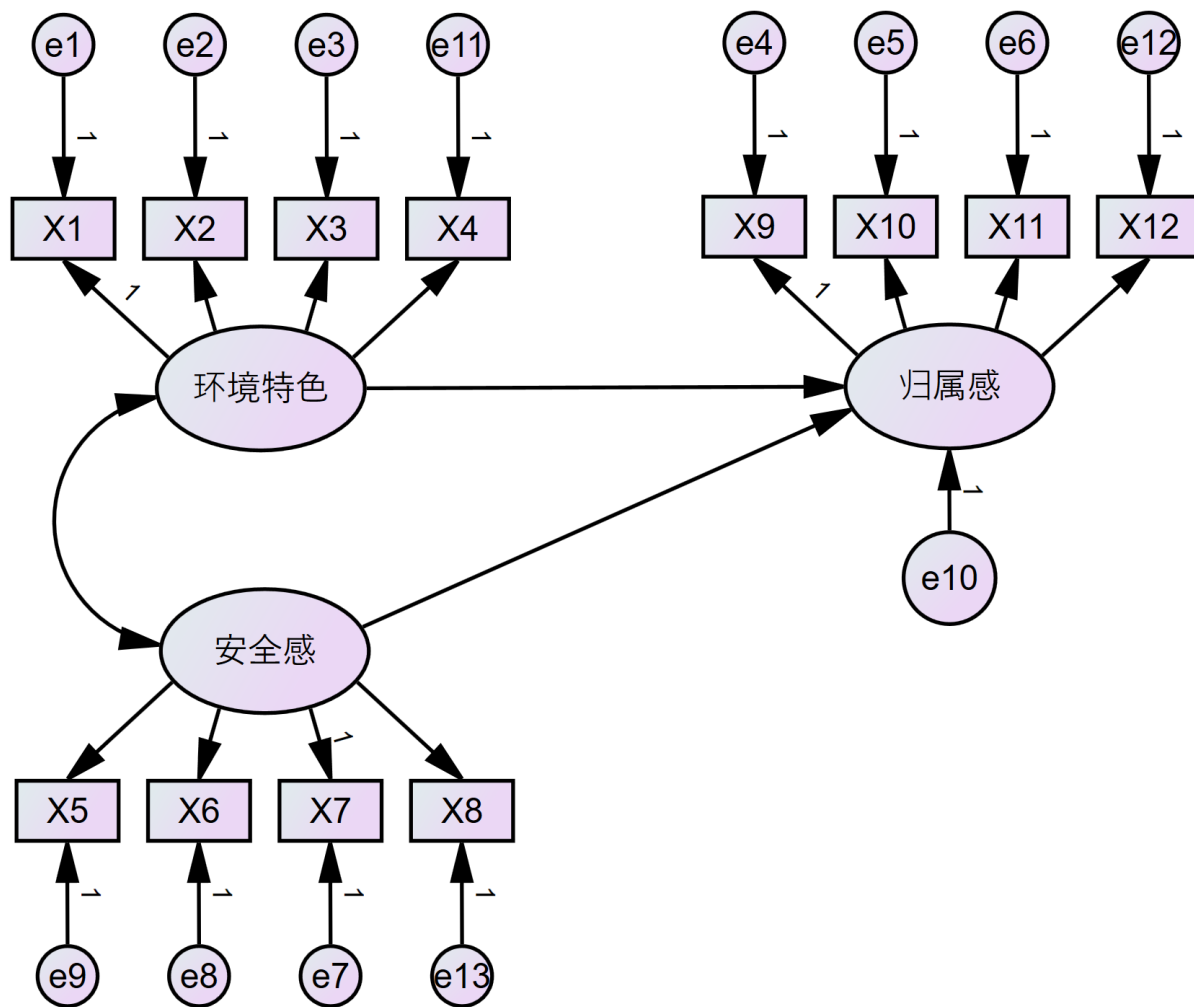
结构方程模型是如何工作的？



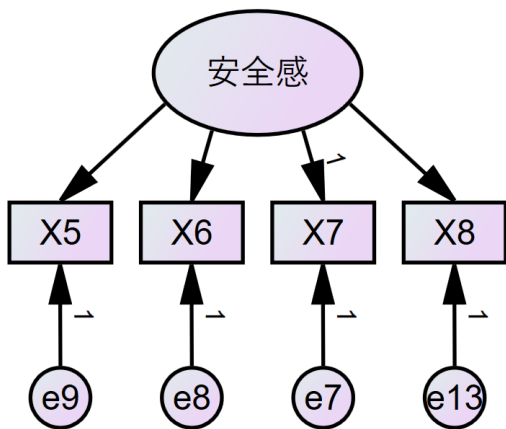
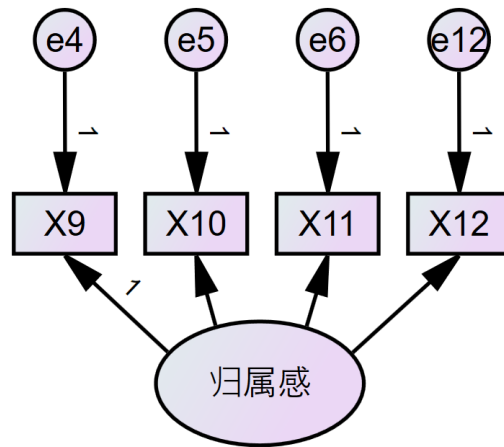
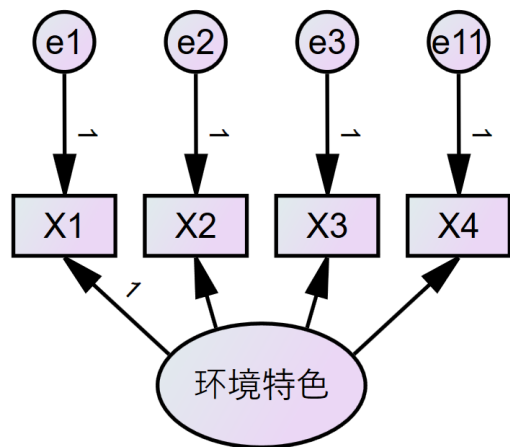
结构模型 (归宿)

- 分析潜变量之间的影响机制；
- 单箭头表示回归关系（分析重点）；
- 双箭头表相关关系。

结构方程模型是如何工作的？



结构方程模型是如何工作的？

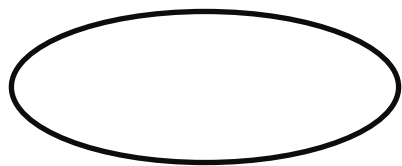


测量模型（基础）

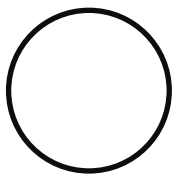
- 用可观察变量对概念进行测量；
- 可观察变量用矩阵表示；
- 潜变量/概念用椭圆表示；
- 误差项用圆形表示。

结构方程模型是如何工作的？

SEM中的路径图符号：



潜变量/概念
latent variable



误差
error



可观察变量
observed variable



相关/协变
covariance



回归
regression

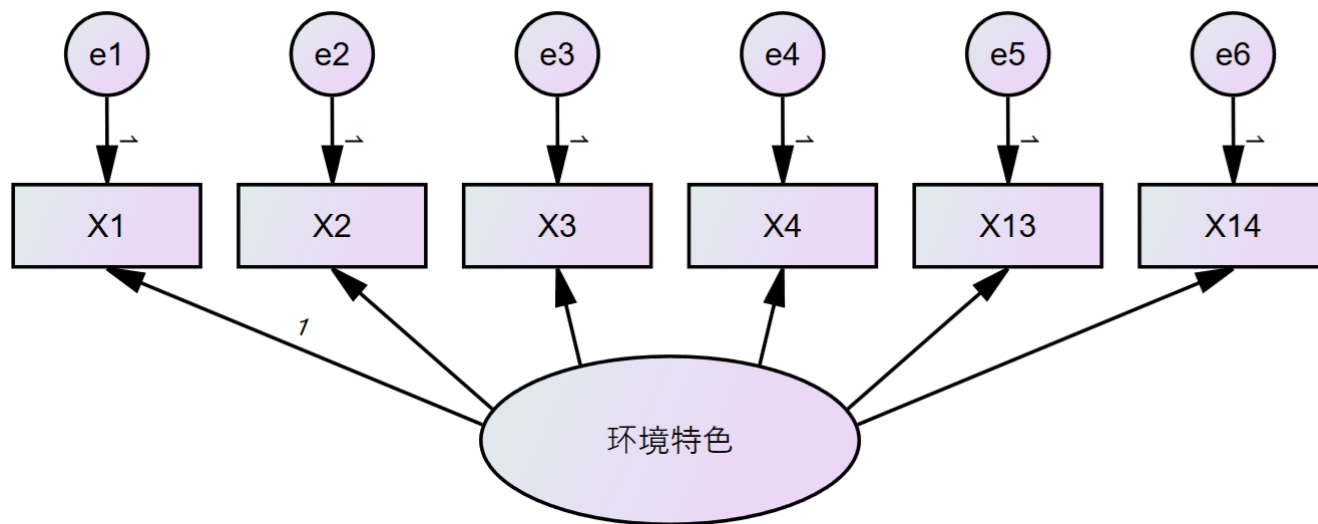
结构方程模型是如何工作的？

SEM的建模步骤：两步走

1. 建立满意的测量模型
 - 对每个潜变量分别进行验证式因子分析
 - 评估收敛效度和区别效度
2. 建立结构模型
 - 评估拟合优度
 - 分析回归系数
 - 考虑更复杂的后续分析

测量模型 Measurement Model

验证式因子分析 Confirmatory Factor Analysis (CFA)

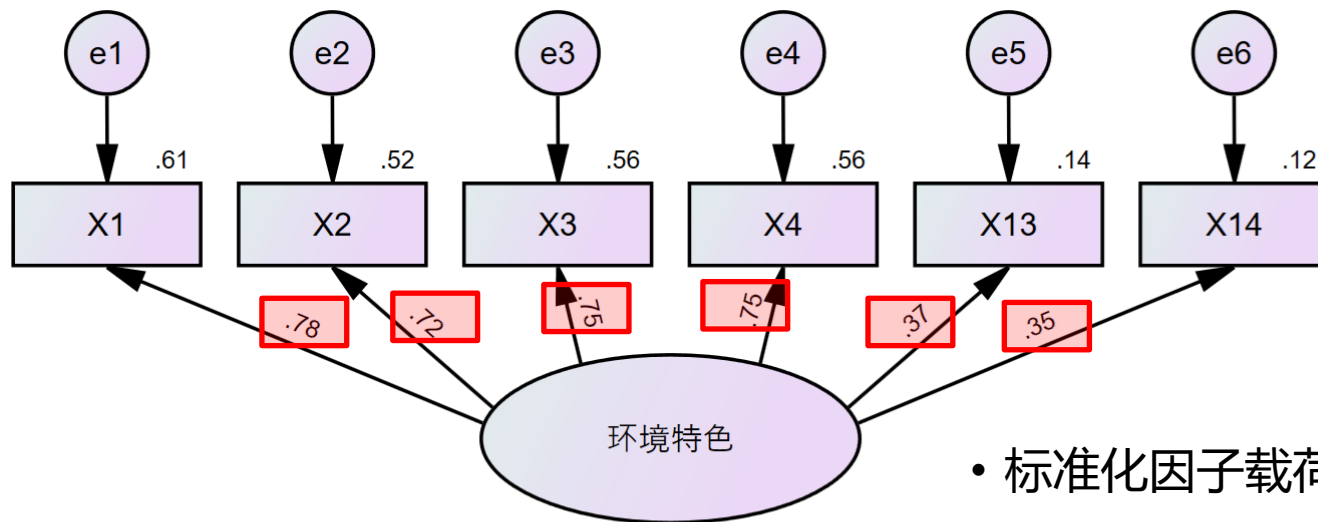


CFI=\cfi

TLI=\tli

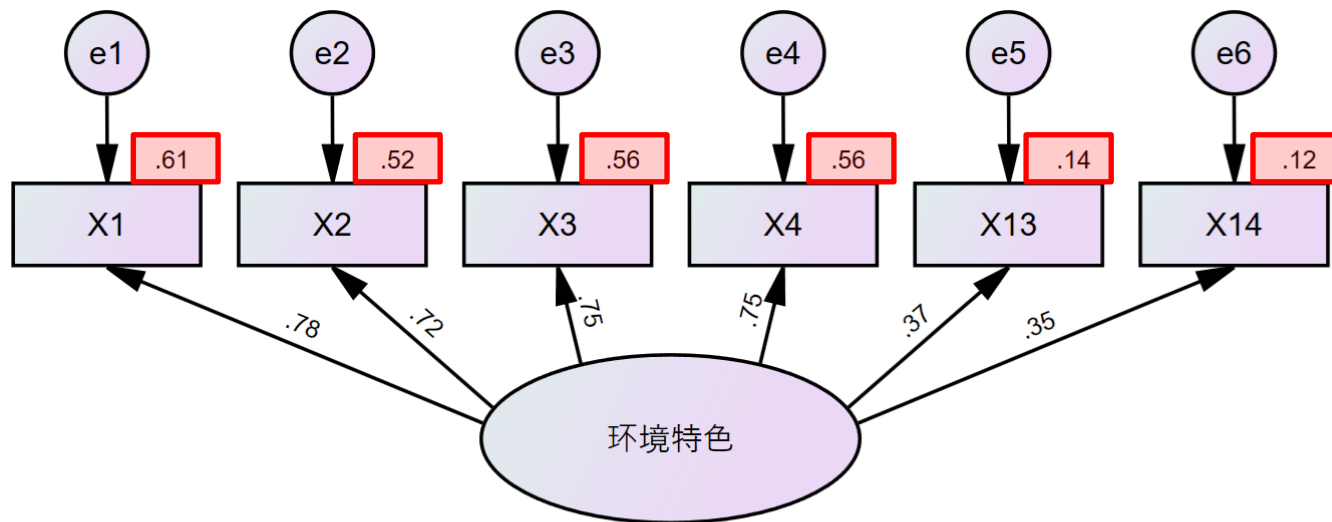
RMSEA=\rmsea

验证式因子分析 Confirmatory Factor Analysis (CFA)



CFI=.462
TLI=.103
RMSEA=.471

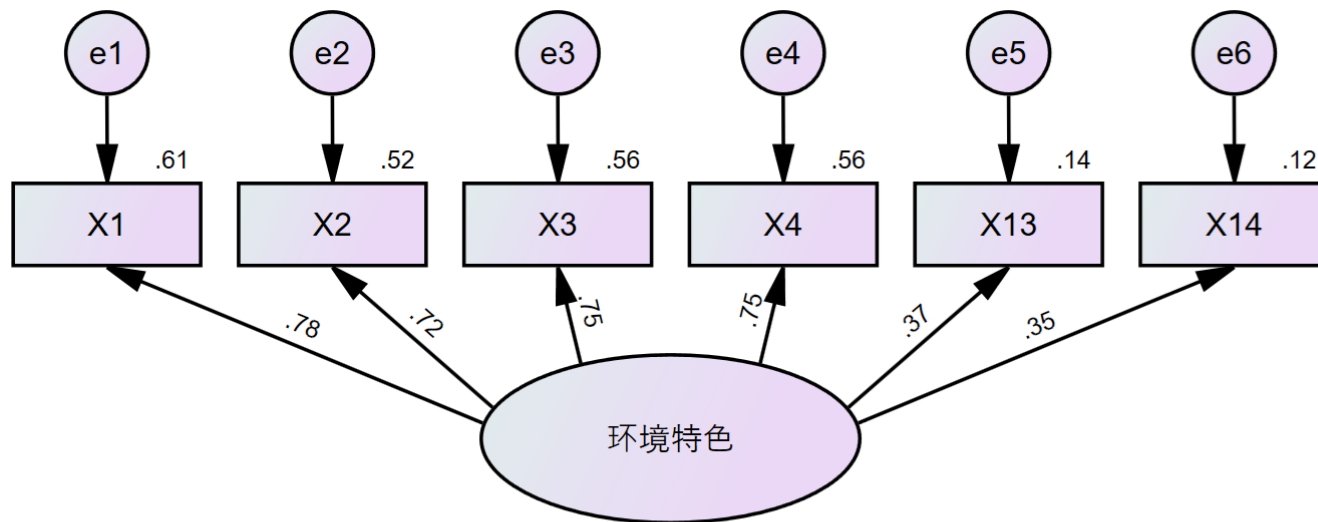
验证式因子分析 Confirmatory Factor Analysis (CFA)



CFI=.462
TLI=.103
RMSEA=.471

- SMC (squared multiple correlation)
 - 相当于 R^2 , >0.5为佳

验证式因子分析 Confirmatory Factor Analysis (CFA)

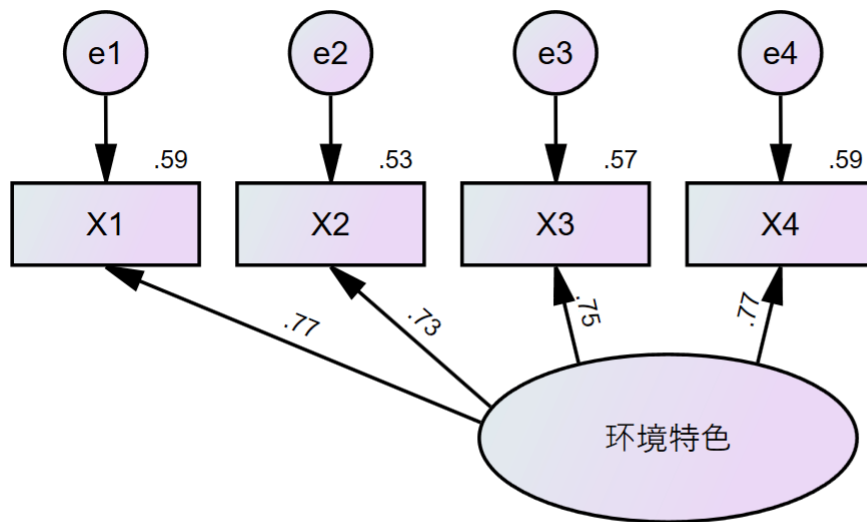


- 拟合优度 (goodness-of-fit)

- CFI, TLI > 0.9
- RMSEA < 0.08

CFI=.462
TLI=.103
RMSEA=.471

验证式因子分析 Confirmatory Factor Analysis (CFA)



CFI=.992
TLI=.976
RMSEA=.078

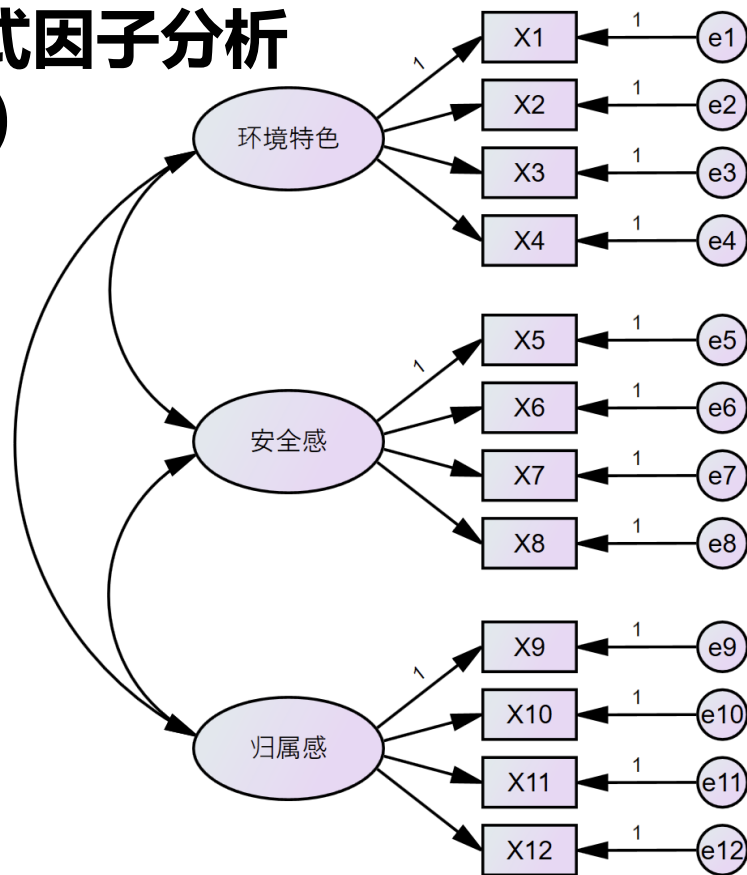
- 标准化因子载荷 (std. factor loading)
 - 类似于相关系数
 - >0.7为佳, <0.45剔除
- SMC (squared multiple correlation)
 - 相当于 R^2 , >0.5为佳
- 拟合优度 (goodness-of-fit)
 - CFI, TLI > 0.9
 - RMSEA < 0.08

问卷设计的建议

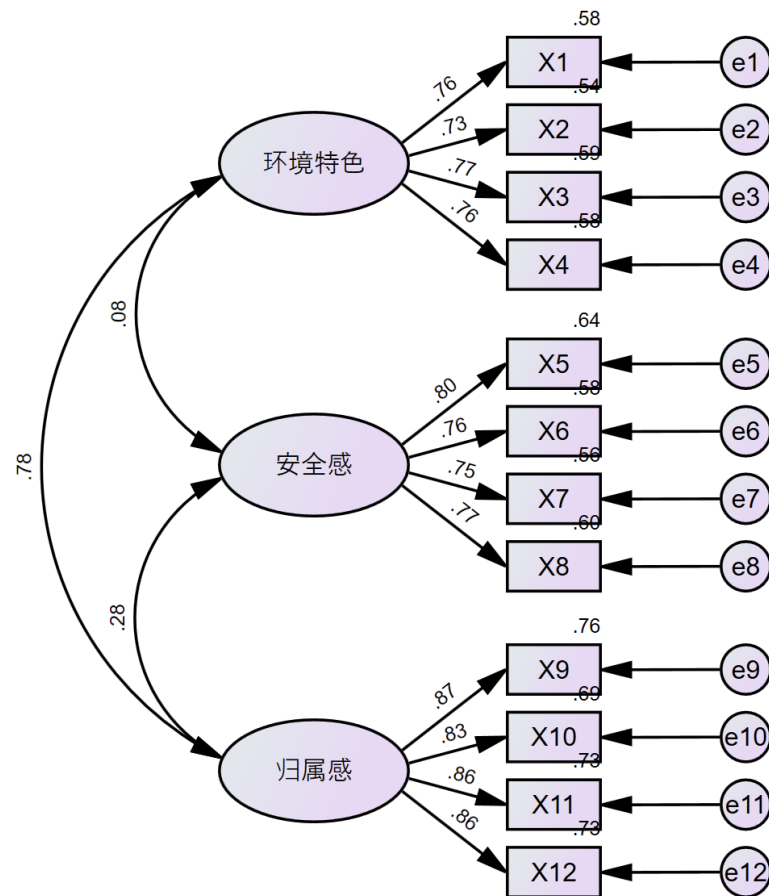
- 每个潜变量至少需要3题，5-7题为佳
- 量表最好为7点尺度
- 最好有一定的理论支持

测量模型 Measurement Model

验证式因子分析 (CFA)



CFI=\cfi
TLI=\tli
RMSEA=\rmsea



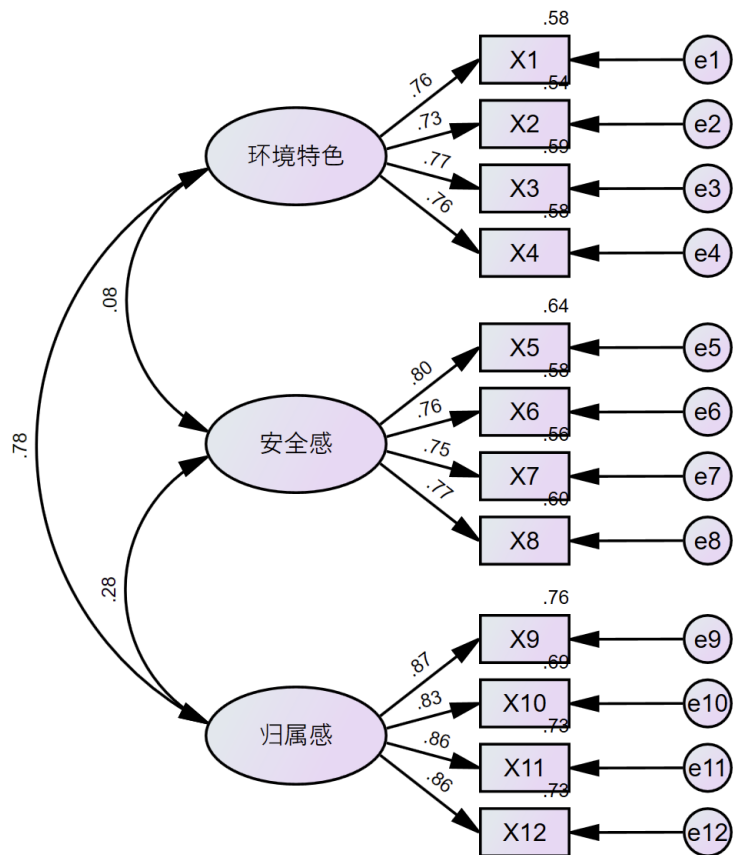
CFI=.980
TLI=.973
RMSEA=.052

收敛效度 (Convergence Validity)

- 强调同一概念下各指标的**内部一致性**；
- 若收敛效度较好，说明可观察变量确实是概念的有效反映指标，测量效果较好；
- 检验方法：
 - 标准化因素负荷
 - SMC (R^2)
 - 组成信度 (CR)
 - 平均变异萃取量 (AVE)

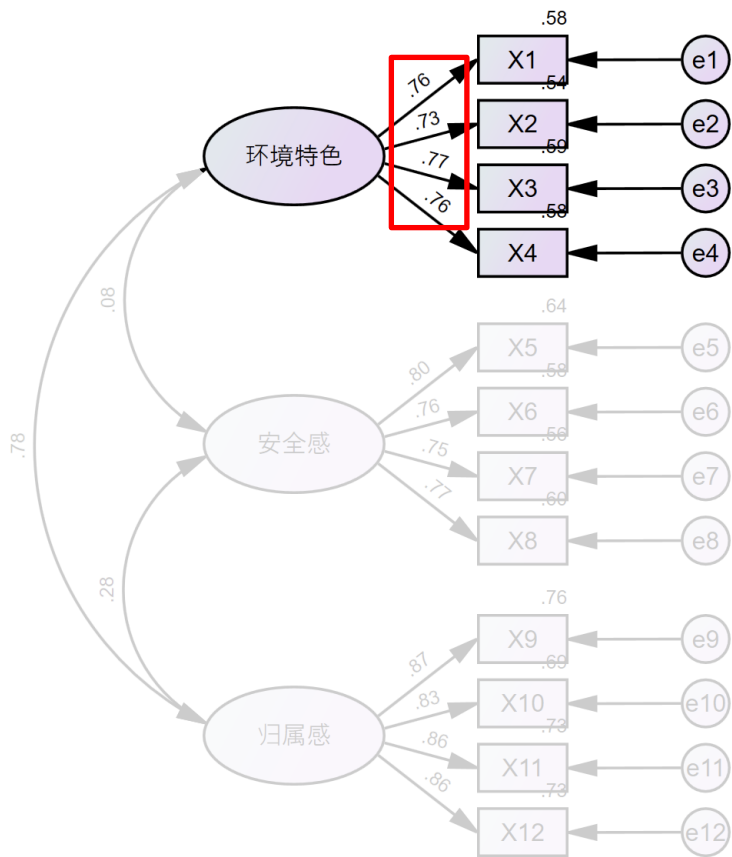


收敛效度 (Convergence Validity)



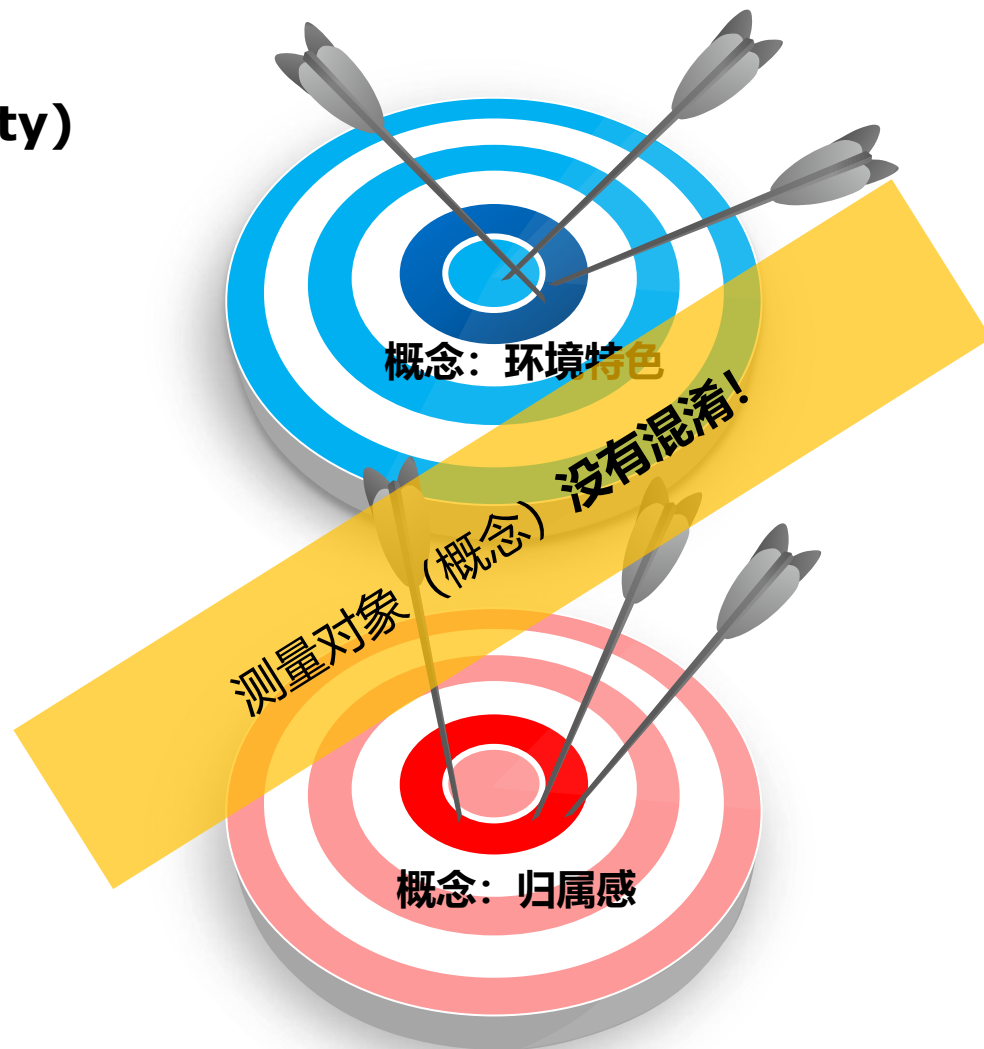
测量模型 Measurement Model

收斂效度 (Convergence Validity)

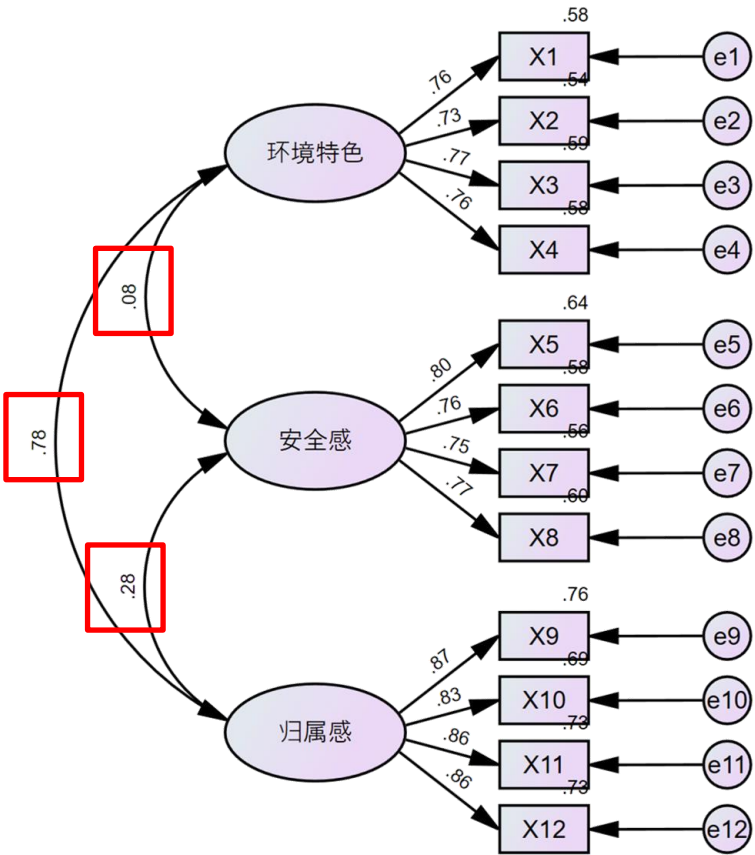
[illegible]

区别效度 (Discriminate Validity)

- 强调不同概念之间的**差异性 (弱相关性)**，防止主观上创造出过于同质的概念；
- 若区别效度较好，说明两个概念没有混淆，可以同时成立（前提是各自的收敛效度达标），否则应删掉一个，或进行合并；
- 检验方法：
 - 1) 直接观察相关系数，是否大于0.85(较不严格)，如果大于0.85，则没有足够的区别效度；
 - 2) 检查相关系数的95%置信区间是否越过1，如果越过1，则没有足够的区别效度。



区别效度 (Discriminate Validity)



Correlations

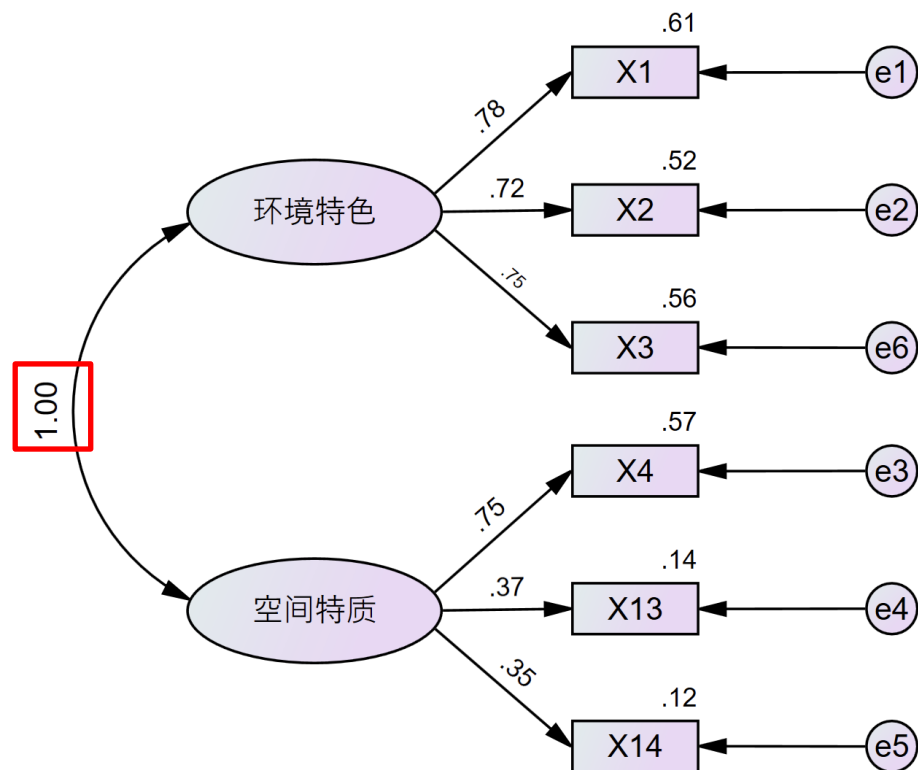
Parameter		Estimate	Lower	Upper	P
Characteristic <--> Safety		.083	-.039	.187	.129
Characteristic <--> Belongness		.784	.733	.823	.020
Safety <--> Belongness		.284	.182	.382	.020

没有跨1，具有区别效度

测量模型 Measurement Model

区别效度 (Discriminate Validity)

没有区别效度的例子：概念混淆



Correlations

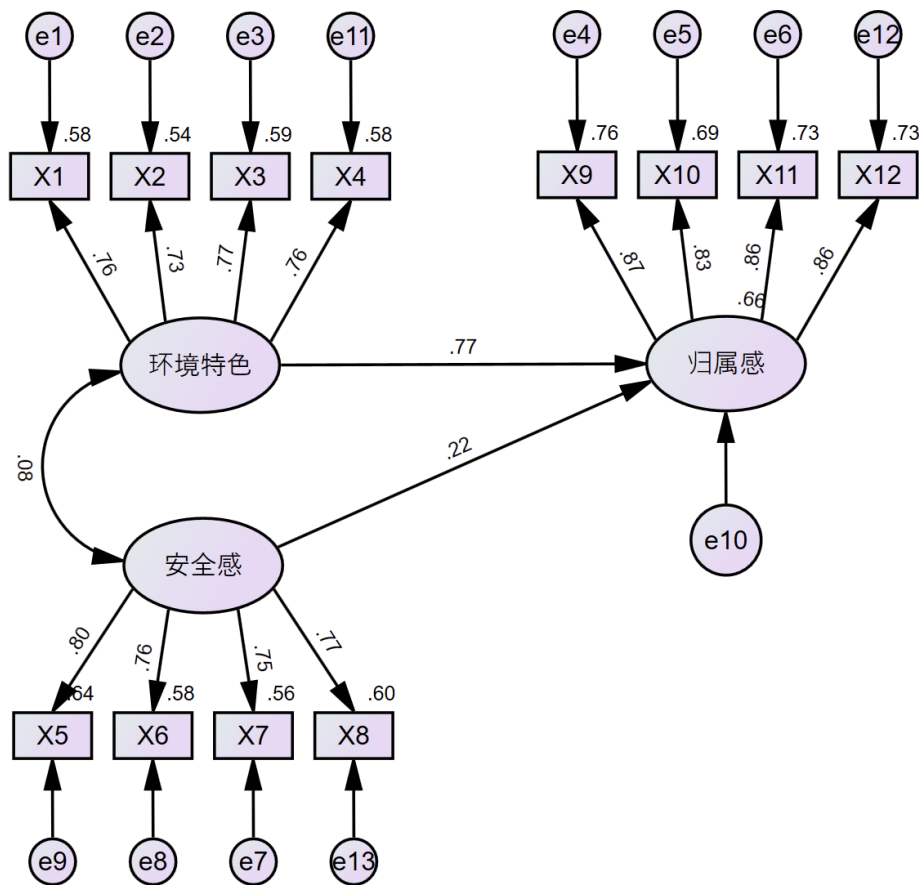
Parameter	Estimate	Lower	Upper	P
F1 <--> F2	.996	.371	1.092	.010

每个概念都独一无二地成立
并且被准确测度



结构模型

结构模型 Structural Model

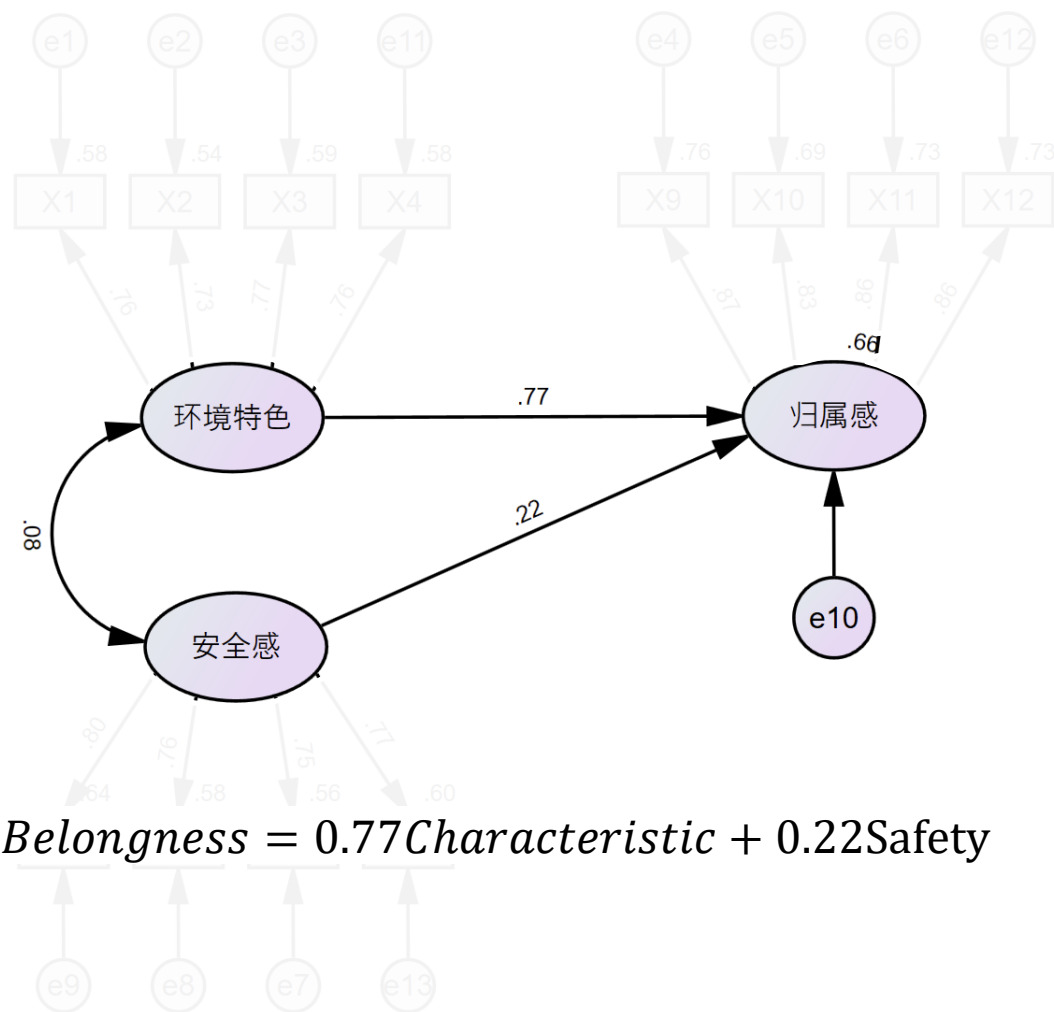


- 拟合优度 (goodness-of-fit)

- CFI, TLI > 0.9
- RMSEA < 0.08

CFI=.980
TLI=.973
RMSEA=.052

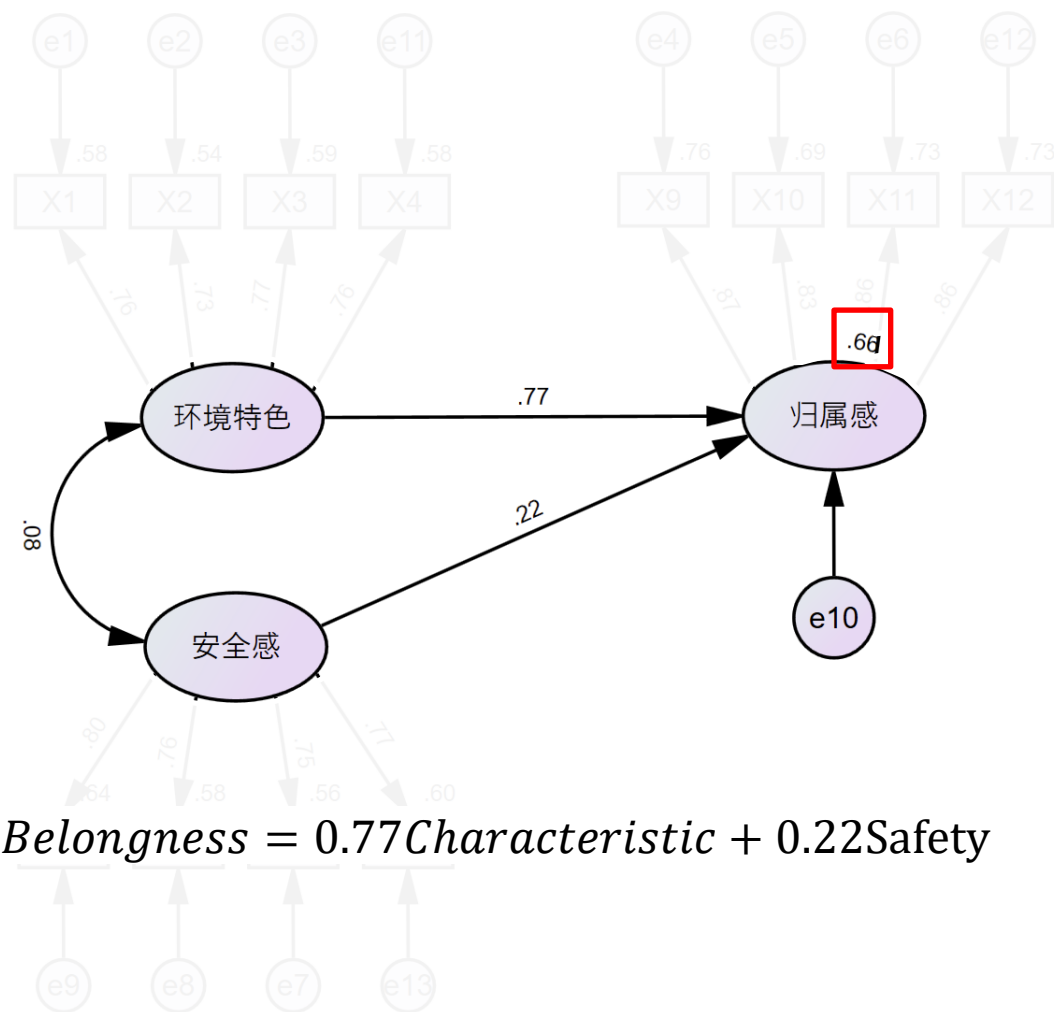
结构模型 Structural Model



标准化路径系数

- 效应的方向（符号）
 - 环境特色水平越高，归属感水平越高。
 - 安全感水平越高，归属感水平越高。
- 效应的大小（绝对值）
 - 环境特色的影响力大于安全感。

结构模型 Structural Model



标准化路径系数

- 效应的方向（符号）
 - 环境特色水平越高，归属感水平越高。
 - 安全感水平越高，归属感水平越高。
- 效应的大小（绝对值）
 - 环境特色的影响力大于安全感。

R² (SMC)

- 66%的归属感水平变异可以被解释。

结构模型 Structural Model

Regression Weights

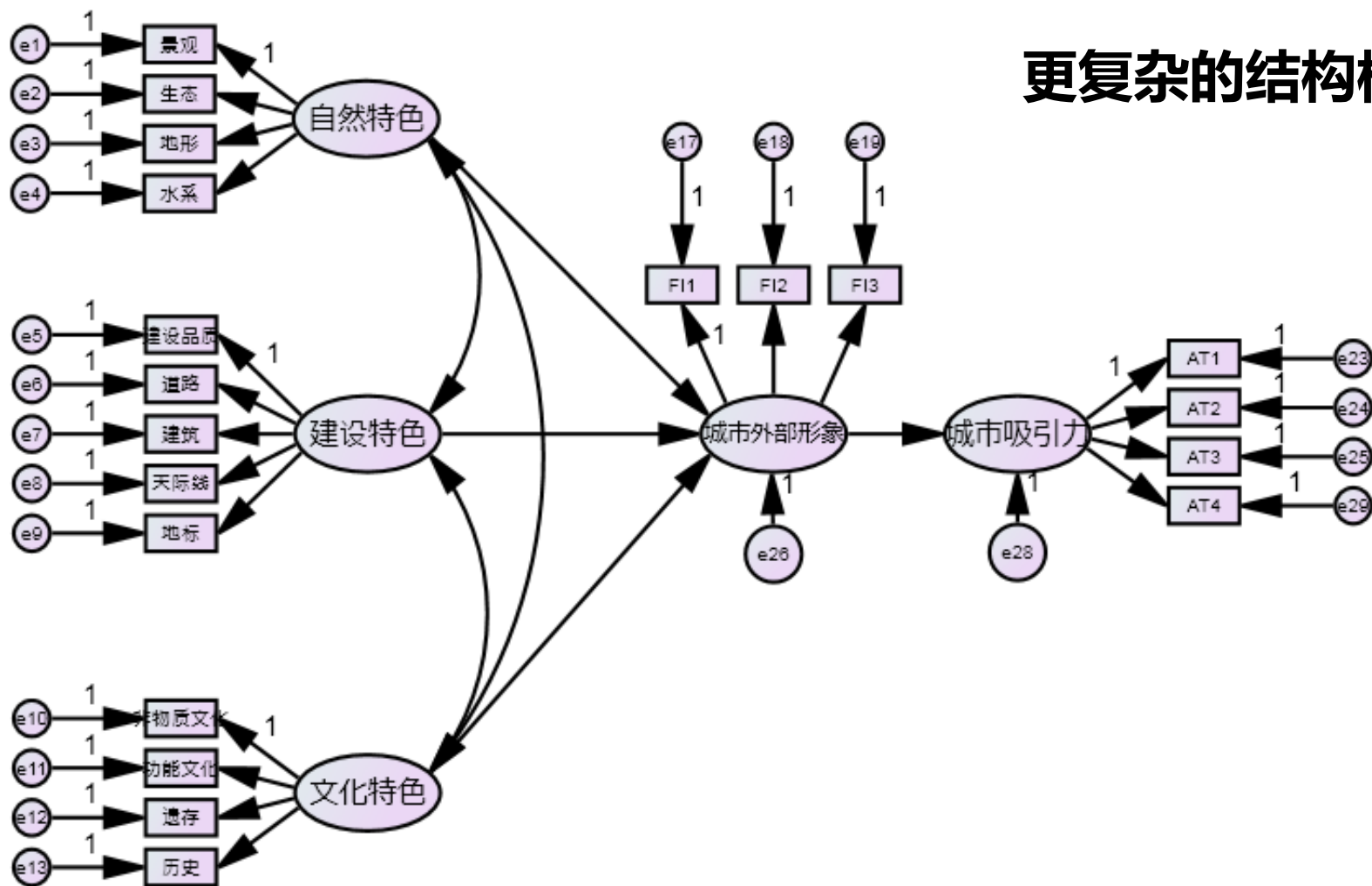
		Estimate	S.E.	C.R.	PLabel
B_elongness	<--- C_characteristic	.788	.051	15.380	***
B_elongness	<--- Safety	.205	.035	5.939	***
X1	<--- C_characteristic	1.000			
X2	<--- C_characteristic	1.077	.067	15.964	***
X3	<--- C_characteristic	.983	.059	16.714	***
X9	<--- B_elongness	1.000			
X10	<--- B_elongness	1.070	.045	23.782	***
X11	<--- B_elongness	1.140	.046	25.042	***
X7	<--- Safety	.950	.057	16.613	***
X6	<--- Safety	.903	.053	17.002	***
X5	<--- Safety	1.000			
X4	<--- C_characteristic	1.060	.064	16.556	***
X12	<--- B_elongness	1.093	.044	25.042	***
X8	<--- Safety	.990	.057	17.264	***

非标准化路径系数(AMOS)

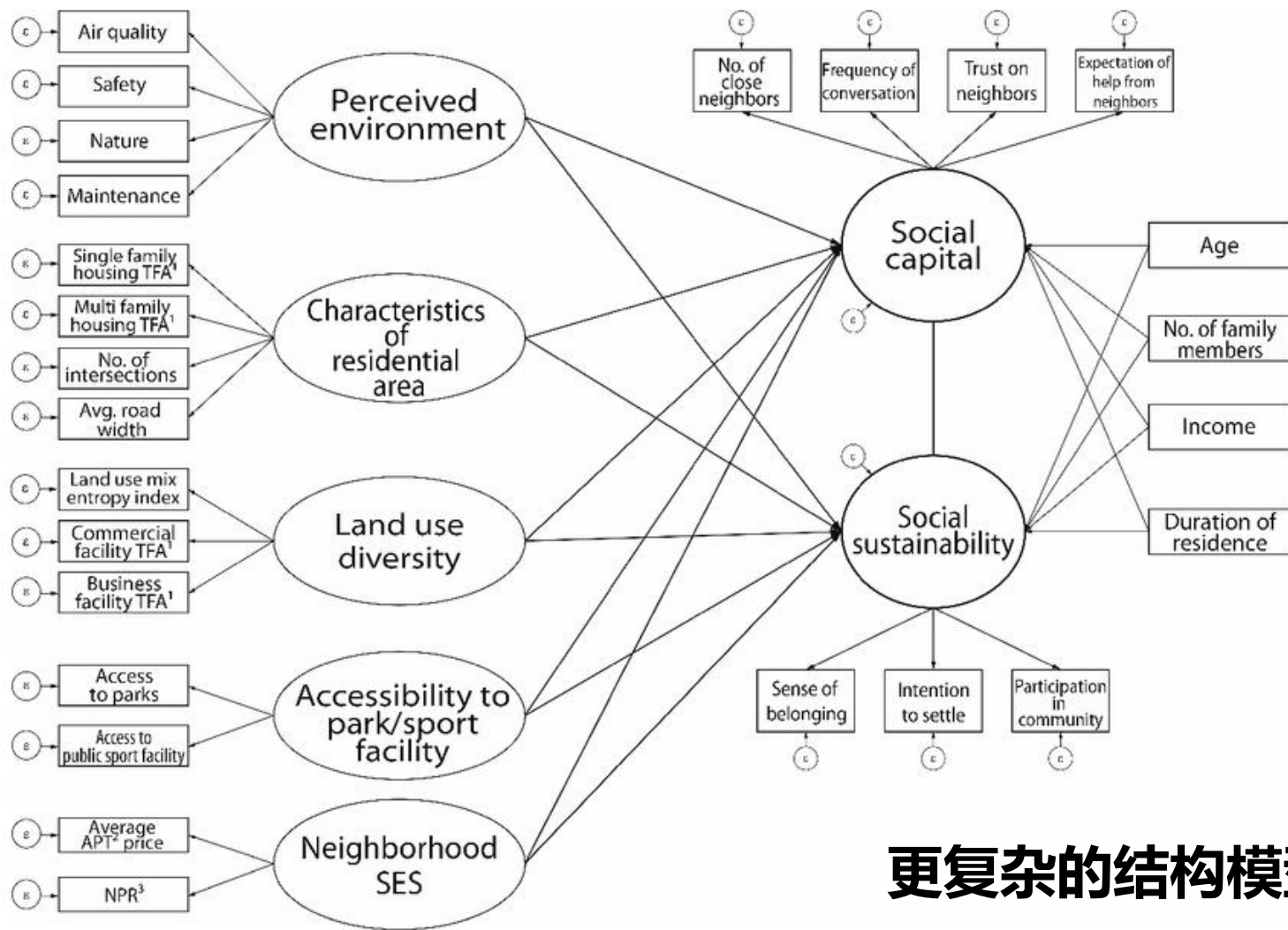
- 效应的显著性 (P值)
 - 环境特色和安全感对归属感的影响均在1%的水平下显著。

结构模型 Structural Model

更复杂的结构模型

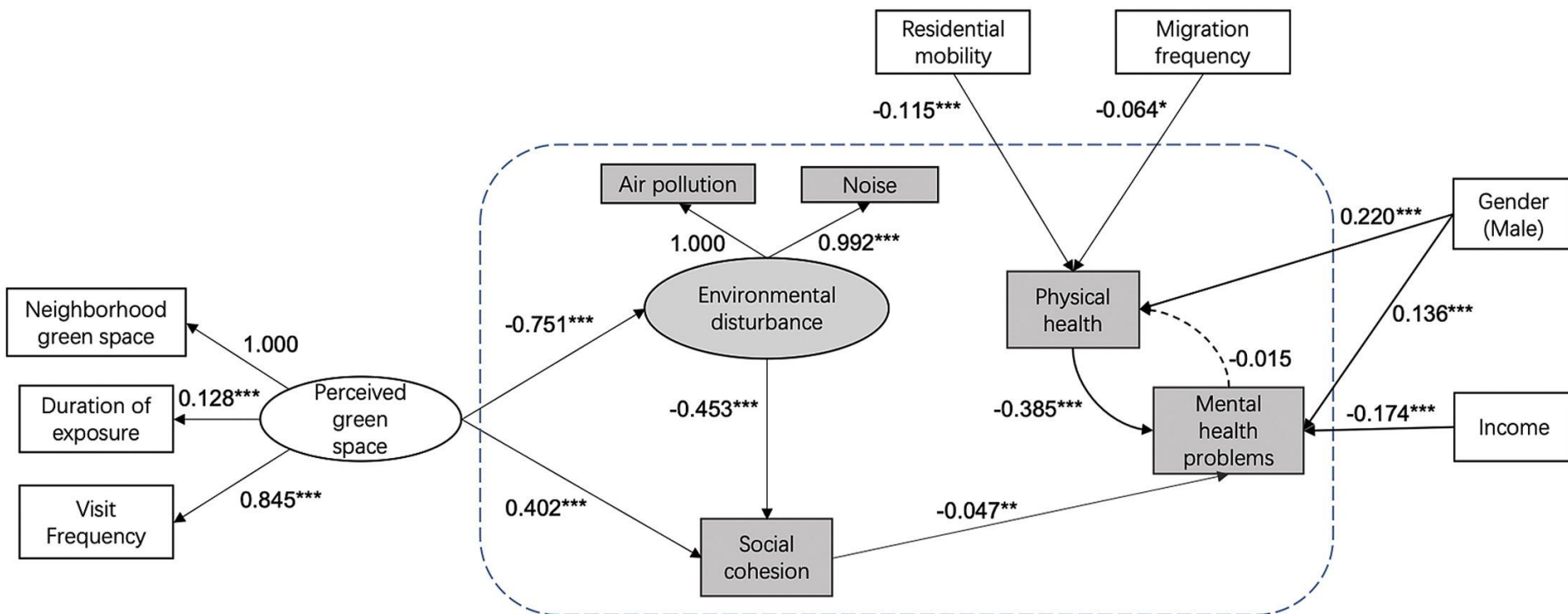


结构模型 Structural Model



更复杂的结构模型

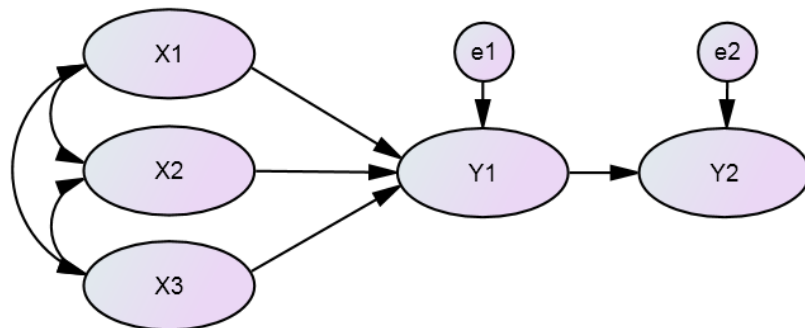
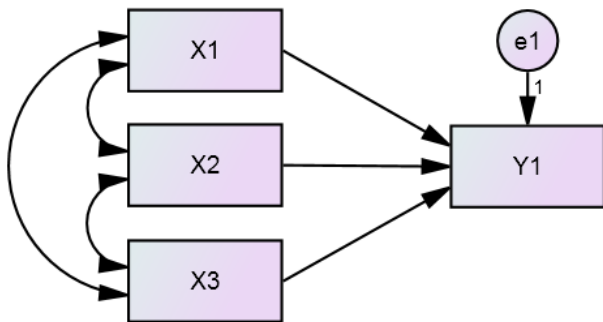
结构模型 Structural Model



更复杂的结构模型

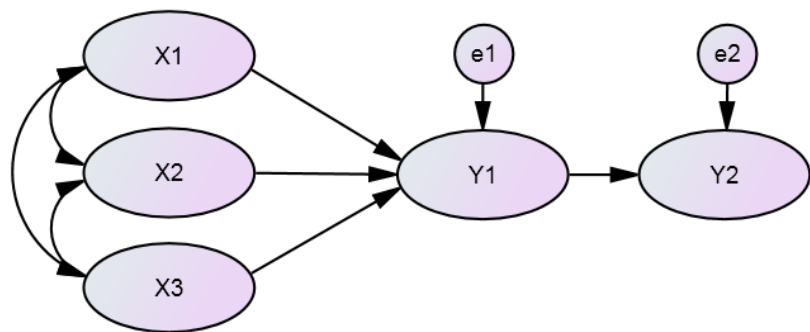
总结

回归分析 → 路径分析 → 结构方程模型

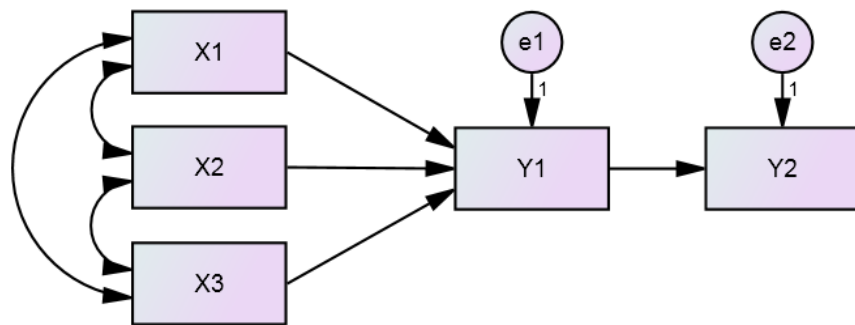


总结

回归分析 → 路径分析 → 结构方程模型



观察变量路径分析



结构方程模型
(潜变量路径分析)

结构方程模型擅于处理**概念**之间的**复杂关系**。

总结

SEM建模过程

1. 建立理论模型
2. 问卷设计和数据收集
3. 建立满意的测量模型（验证式因子分析）
4. 估计结构模型

总结

SEM技术要点

