

分层分析

分层分析回答的问题是：有哪些因素影响所研究的危险因素（X）与结果变量（Y）的关系？即：有没有混杂因素？有没有效应修饰因子？参看[流行病学分析思路](#)。

指定了一系列分层（分类型）变量，系统在每个分层变量的每个层级水平上对所研究的危险因素与结果变量进行回归分析。目的是：

- 1) 发现效应修饰因子：如果所研究的危险因素（X）对结果变量（Y）的作用在分层变量的不同的层级水平上有显著不同，表示该分层变量是效应修饰因子。具体做法是首先看不同层级水平上 X 对 Y 的回归系数或比值比的差异，如果差异较大，做交互作用检验，看差异是否有显著性。
- 2) 发现混杂因素：如果单因素分析危险因素（X）对结果变量（Y）有作用，但按分层变量分层分析，每个层级水平上所研究的危险因素（X）对结果变量（Y）都没有作用，意味着该分层变量是混杂因素。

系统将自动检测结局变量的类型（如两分类变量、连续变量等），再自动默认选择合适的回归模型（如 Logistic 回归或线性回归模型）。用户可以对分布类型和联系函数自行定义。还可以调用 Cox 模型做生存分析或条件 logistic 回归分析。

用户可以定义表格输出格式，包括要报告的结果、行列编排、小数点等。

例（输入界面）：

逐个分层因素分层回归分析 ?

I. 标题:

II. 选择分析对象:

III. 选择结果变量 (Y): Cox 模型生存分析(事件=1)

变量	分布	联系函数
Systolic BP, mmhg	Gaussian	Identity
Diastolic BP, mmhg	Gaussian	Identity

IV. 选择危险因素 (X):

变量
Body mass index, kg/m2

选择时间变量:

或, 开始时间:

结束时间:

VI. 选择分层变量 (行):

变量
SEX
SMOKE
Occupation
Education

V. 选择调整变量:

变量
Age, years

VI. 选择列分层变量:

VII. 选择列设计:

VIII. 选择输出内容与格式:

IX. 精确到小数点:

如用GEE

研究对象编号:

内部相关类型:

输出结果:

分层分析

	样本量 (%)	Systolic BP, mmhg	Diastolic BP, mmhg
SEX			
Male	394 (49.7%)	-0.2 (-1.2, 0.8) 0.704	-0.2 (-0.8, 0.3) 0.388
Female	399 (50.3%)	1.3 (0.5, 2.1) <0.001	0.5 (0.1, 0.9) 0.012
SMOKE			
no	478 (60.7%)	0.8 (0.1, 1.4) 0.032	0.4 (0.0, 0.8) 0.038
yes	309 (39.3%)	0.2 (-0.9, 1.4) 0.688	-0.1 (-0.7, 0.5) 0.760
Occupation			
farmer	412 (52%)	1.5 (0.7, 2.3) <0.001	0.6 (0.2, 1.0) 0.003
others	381 (48%)	-0.6 (-1.6, 0.3) 0.210	-0.4 (-0.9, 0.2) 0.189
Education			
elementary of lower	326 (41.4%)	1.5 (0.5, 2.6) 0.004	0.6 (0.1, 1.2) 0.018
middle	248 (31.5%)	0.6 (-0.4, 1.6) 0.237	0.0 (-0.6, 0.6) 0.956

high or above	214 (27.2%)	-0.5 (-1.4, 0.5) 0.354	0.1 (-0.5, 0.7) 0.712
---------------	-------------	------------------------	-----------------------

For Systolic BP, mmhg Diastolic BP, mmhg: 回归系数 (95%可信区间) P 值
 结局变量: Systolic BP, mmhg 和 Diastolic BP, mmhg
 危险因素: Body mass index, kg/m²
 调整变量: Age, years
 于 2015-07-24 使用《易俾统计》软件(www.empowerstats.com) 和 R 软件生成。

使用的样本量

	Systolic BP, mmhg	Diastolic BP, mmhg
SEX		
Male	394	394
Female	399	399
SMOKE		
no	478	478
yes	309	309
Occupation		
farmer	412	412
others	381	381
Education		
elementary of lower	326	326
middle	248	248
high or above	214	214

结果解释:

以上结果均是通过回归分析得出的。结果变量 (Y) 有 Systolic BP 与 Diastolic BP, 危险因素 (X) 为 Body mass index。表中结果为: 回归系数 (95%可信区间) P 值。

以 Systolic BP 为例, 按 SEX 分层分析, 在男性 (Male) 中, Body mass index 每增加 1kg/m², Systolic BP 降低 0.2mmHg, 95%可信区间 -1.2 至 0.8mmHg, P=0.704; 女性 (Female) 中, Body mass index 每增加 1kg/m², Systolic BP 增加 1.3 mmHg, 95%可信区间 0.5 至 2.1, P<0.001。