

扫描数据关联关系

使用该模块，用户指定结果变量 Y，逐个检查数据中其它变量与 Y 的关联关系，寻找与之有关联的变量。该模块有如下特点：

- 1) Y 变量类型可以是多种，如正态分布、二项分布、泊松分布等，软件自动识别变量类型，给出默认的分布与联系函数，用户也可以重新定义分布类型与联系函数。
- 2) Y 可以是时间依赖的生存状态，此时需要给定时间变量，软件将调用 cox 模型分析数据，寻找与生存状态有关的变量。
- 3) 可以使用 Cox 模型分析配对的病例对照研究资料，此时需要给定配对组编号。
- 4) 如果研究对象之间存在内部相关性，调用广义估计方程（GEE）分析。
- 5) 可以给出要调整的变量，逐个分析其它变量，在调整了这些变量后，与 Y 的关系。
- 6) 结果以图表形式展示，并按关联关系的 p 值排序，方便用户快速查询。

例 2. 打开易侬软件自带的练习项目“demo”，调用数据分析菜单下的“扫描数据关联关系”模块。选择变量(Y)为 SBP（收缩压）、DBP（舒张压），选择调整变量为 AGE（年龄）、SEX（性别）、EDU（文化程度），如下图所示，点击“查看结果”。

扫描关联关系

标题: 扫描关联关系

选择分析对象: 所有数据记录

应变量(Y)

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

要调整的变量

Cox 模型生存分析(事件=1)

选择时间变量: []

或开始时间: []

结束时间: []

如用GEE

研究对象编号: []

内部相关类型: []

刷新 保存 查看结果

下图是对数据库中每个变量与 Y 关系的检查结果，最显著的放在最上面，给出回归系数（如为二分类的 Y，给出 OR）及 95%可信区间与 P 值。如果被检查的变量是连续性的，同时给出其与 Y 的曲线拟合图。

Scan for associations with Systolic BP, mmHg					Graphics	
P.value	Variable	Result				
*** 0	HBP High BP	factor(HBP) coeff. 95%low 95%upp p 38.00 35.00 40.00 0.0000				
*** 0	SNP1	factor(SNP1)1 coeff. 95%low 95%upp p 8.00 5.00 11.00 0e+00 factor(SNP1)2 12.00 5.00 19.00 4e-04				
*** 0	A11	factor(A11) coeff. 95%low 95%upp p 9.00 5.00 12.00 0.0000				
*** 0	SNP2	factor(SNP2)1 coeff. 95%low 95%upp p -6.00 -9.00 -4.00 0.0000 factor(SNP2)2 -11.00 -17.00 -6.00 0.0000				
*** 0.5215 (3e-04)	SUBJ	SUBJ coeff. 95%low 95%upp p 0.00 0.00 0.00 0.5235 edf Ref. df F p-value s(SUBJ) 2.7 3.36 5.84 3e-04				
*** 0.5018 (3e-04)	FMYID	FMYID coeff. 95%low 95%upp p 0.00 0.00 0.00 0.5038 edf Ref. df F p-value s(FMYID) 2.7 3.36 5.84 3e-04				

与“诊断数据关联关系”模块相比，“扫描数据关联关系”模块虽然不能直接选点分组进行可视化数据分析，但是可以做生存分析和广义估计方程（GEE），并且可以得到曲线拟合的结果，这对于快速挖掘数据信息，很有帮助。