
利用SPSS软件实现随机化过程

作者：陶立元 赵一鸣 来源：临床流行病学和循证医学

本文原地址：<https://www.iikx.com/news/statistics/6221.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

利用SPSS软件实现随机化过程

。随机过程是临床研究过程中至关重要的一部分。在前面的内容中我们也讲过了很多关于随机和随机数的内容，请大家参阅旧的目录。本期将跟大家一起讨论一下如何通过SPSS来实现随机化过程。

在说之前，先跟大家聊一聊随机数的类型，随机数有真随机数(True Random Number)和伪随机数(Pseudo Random Number)之分。我们平时所使用的无论什么程序产生的随机数都是伪随机数，它是有“种子”的，种子决定了随机数的固定序列。而真随机数很多情况下只能看老天的眼色，比如掷骰子、布朗运动、量子效应、放射性衰变等等，所以说我们在临床研究中常用的软件产生随机数的方法都是伪随机数，伪随机数虽然不具有完全的随机性，但是已经足够用了。

SPSS随机化

直接进行抽样或分组

1、利用“选择个案”窗口实现

2、利用“复杂抽样”窗口实现

利用syntax语句实现

按要求生成随机数字

3、利用“计算变量”窗口实现

利用syntax语句实现

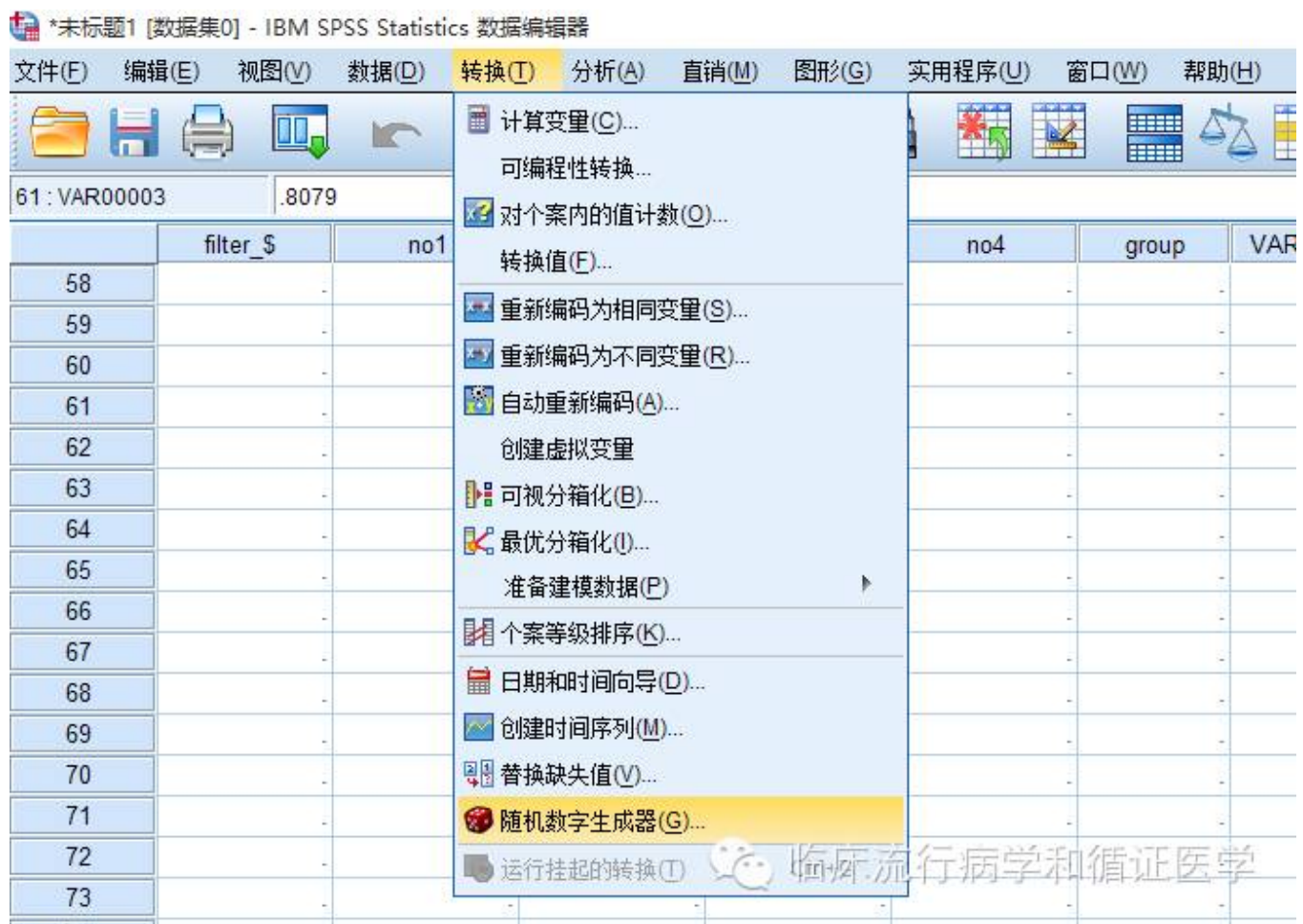
临床流行病学和循证医学

总体上SPSS为随机化提供了两类过程，第一类过程直接进行随机抽样或分组，第二类过程只生成随机数字，需要研究者依据随机数字自己去进行随机抽样或分组。

在上述过程中，我们都可以选择采用“随机种子数”或“固定种子数”两种方式来产生随机数。采用“随机种子数”时，随机数字不可重现;采用“固定种子数”时，随机数字可以重现出来。由于SPSS的syntax不是所有人都常用，在此就不说了，其实大家可以通过点击paste选项保存出任意计算过程的syntax语句。下面分别介绍其它三种随机数生成方法:

1、利用“选择个案”窗口实现

a. 设定种子数：顺序点击转换、随机数字生成器，勾选Mersenne Twister(相比兼容SPSS12更可靠)，勾选固定值(设定一个自己的种子数)，点击确定。这样就设定好了种子数。



b. 依次点击数据、选择个案，点选随机个案样本(D)，点击样本，点选其中一个并填入数值，点击继续，点击确定即可。如下图：

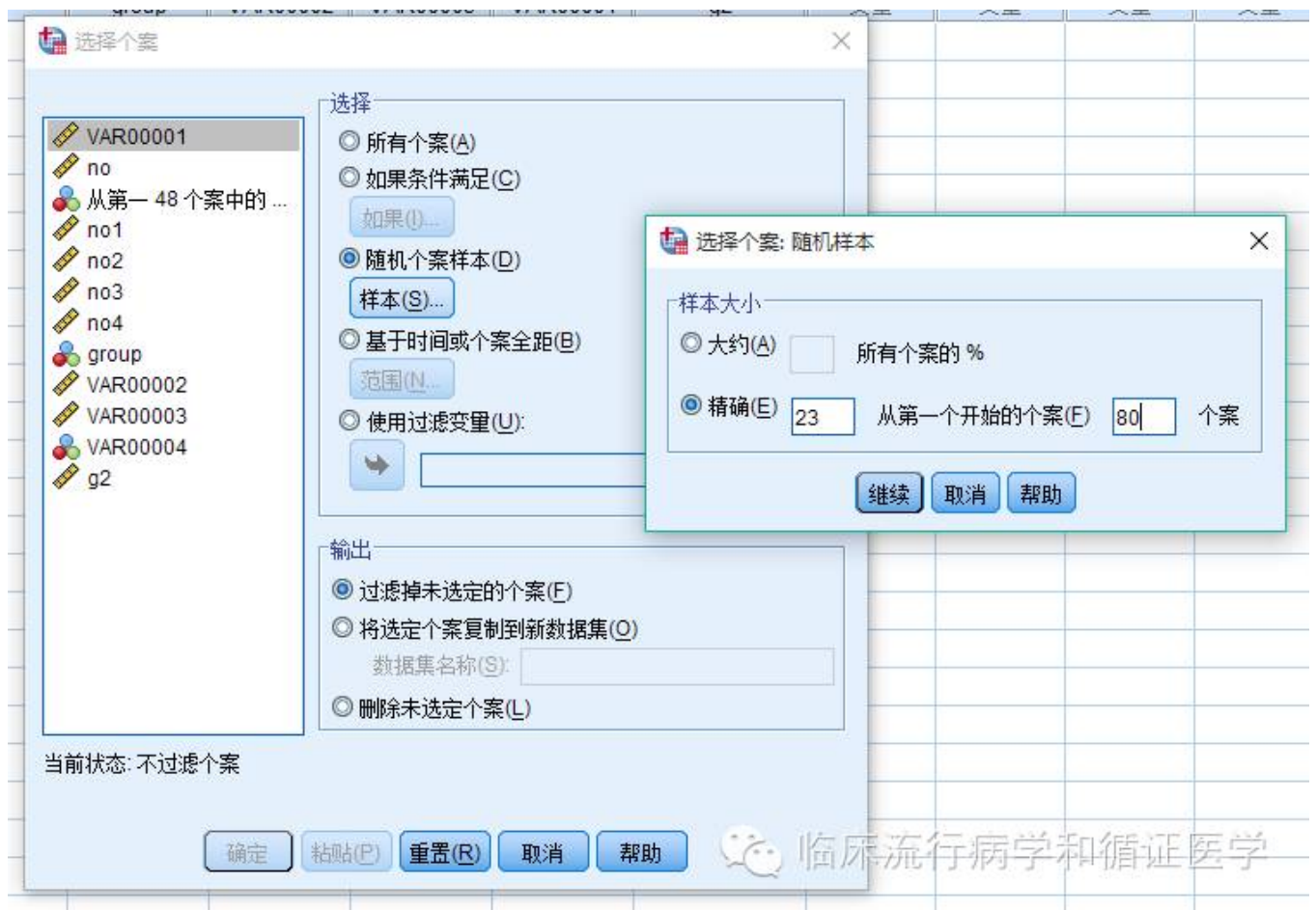
文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 直销(M) 图形(G) 实用程序(U)

61: VAR00003 .8079

	filter_\$	no3	no4
58			
59			
60			
61			
62			
63			
64			
65			
66			
67			
68			
69			
70			
71			
72			
73			
74			
75			
76			
77			
78			
79			
80			
81			
82			
83			

- 定义变量属性(V)...
- 设置未知测量级别(L)...
- 复制数据属性(C)...
- 新建定制属性(B)...
- 定义日期(E)...
- 定义多响应集(M)...
- 验证(L) ▶
- 标识重复个案(U)...
- 标识异常个案(I)...
- 比较数据集(P)...
- 排序个案(O)...
- 排序变量(B)...
- 变换(N)...
- 合并文件(G) ▶
- 重组(R)...
- 搜索权重...
- 倾向得分匹配...
- 个案控制匹配...
- 汇总(A)...
- 拆分为文件
- 正交设计(H) ▶
- 复制数据集(D)
- 拆分文件(F)...
- 选择个案(S)...
- 加权个案(W)...

临床流行病学和循证医学



2、利用“复杂抽样”窗口来实现;

复杂抽样窗口可实现分层抽样，并保存出抽样计划和抽样结果文件。按照其提示一步步点击即可。为了简化，后面过程不再截图。

61: VAR00003	filter_\$	no1	no4	group
58	-	-	-	-
59	-	-	-	-
60	-	-	-	-
61	-	-	-	-
62	-	-	-	-
63	-	-	-	-
64	-	-	-	-
65	-	-	-	-
66	-	-	-	-
67	-	-	-	-
68	-	-	-	-
69	-	-	-	-
70	-	-	-	-
71	-	-	-	-
72	-	-	-	-
73	-	-	-	-
74	-	-	-	-
75	-	-	-	-
76	-	-	-	-
77	-	-	-	-
78	-	-	-	-
79	-	-	-	-
80	-	-	-	-
81	-	-	-	-
82	-	-	-	-
83	-	-	-	-
84	-	-	-	-
85	-	-	-	-
86	-	-	-	-

报告(P)

描述统计(E)

表(B)

比较平均值(M)

一般线性模型(G)

广义线性模型(Z)

混合模型(X)

相关(C)

回归(R)

对数线性模型(O)

神经网络(W)

分类(F)

降维(D)

度量(A)

非参数检验(N)

预测(I)

生存函数(S)

多重响应(U)

缺失值分析(Y)...

多重插补(T)

复杂抽样(L)

模拟(I)...

质量控制(Q)

ROC 曲线图(V)...

选择样本(S)...

准备分析(P)...

频率(F)...

描述(D)...

交叉表格(C)...

比率(R)...

一般线性模型(G)...

Logistic 回归...

序数回归(O)...

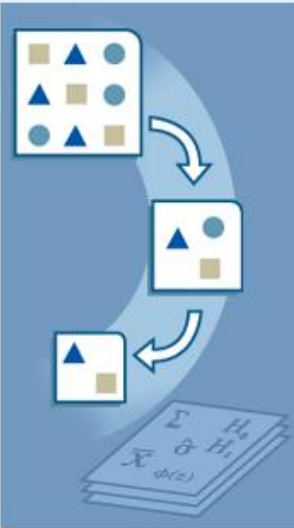
Cox 回归...

 抽样向导

×

欢迎使用抽样向导

抽样向导帮助您设计和选择复杂样本。您的选择将保存到计划文件中，可以在分析时用于表明数据如何抽样。您也可以使用该向导根据现有计划来修改抽样计划或抽取一个样本。



您希望做什么？

☒ 设计样本(D)

如果没有创建计划文件，选择此选项。您将可以选择抽取该样本。

文件(I):

☐ 编辑样本设计(E)

如果您想添加、删除或修改现有规划的阶段，选择此选项。您将可以选择抽取该样本。

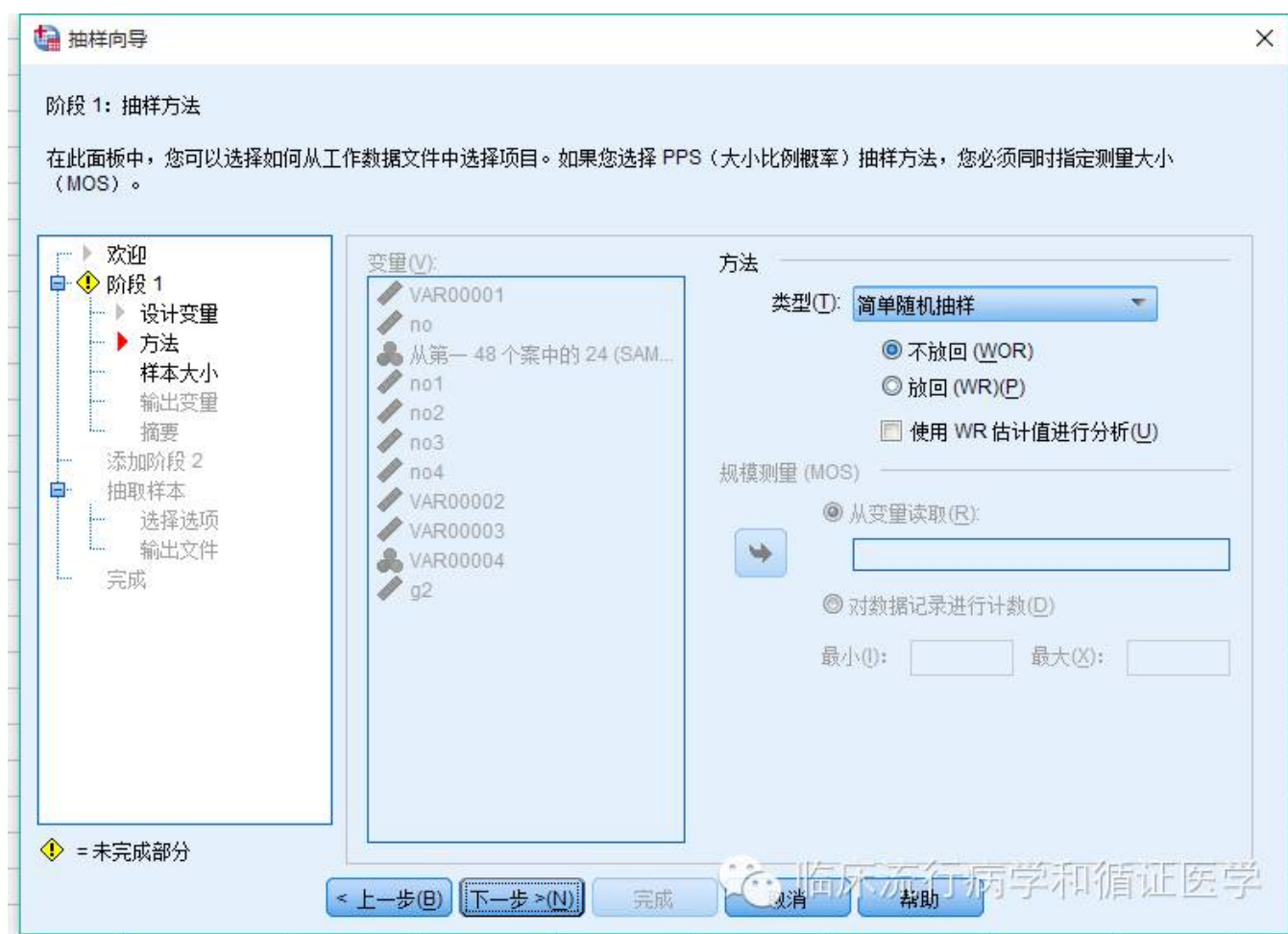
文件(L):

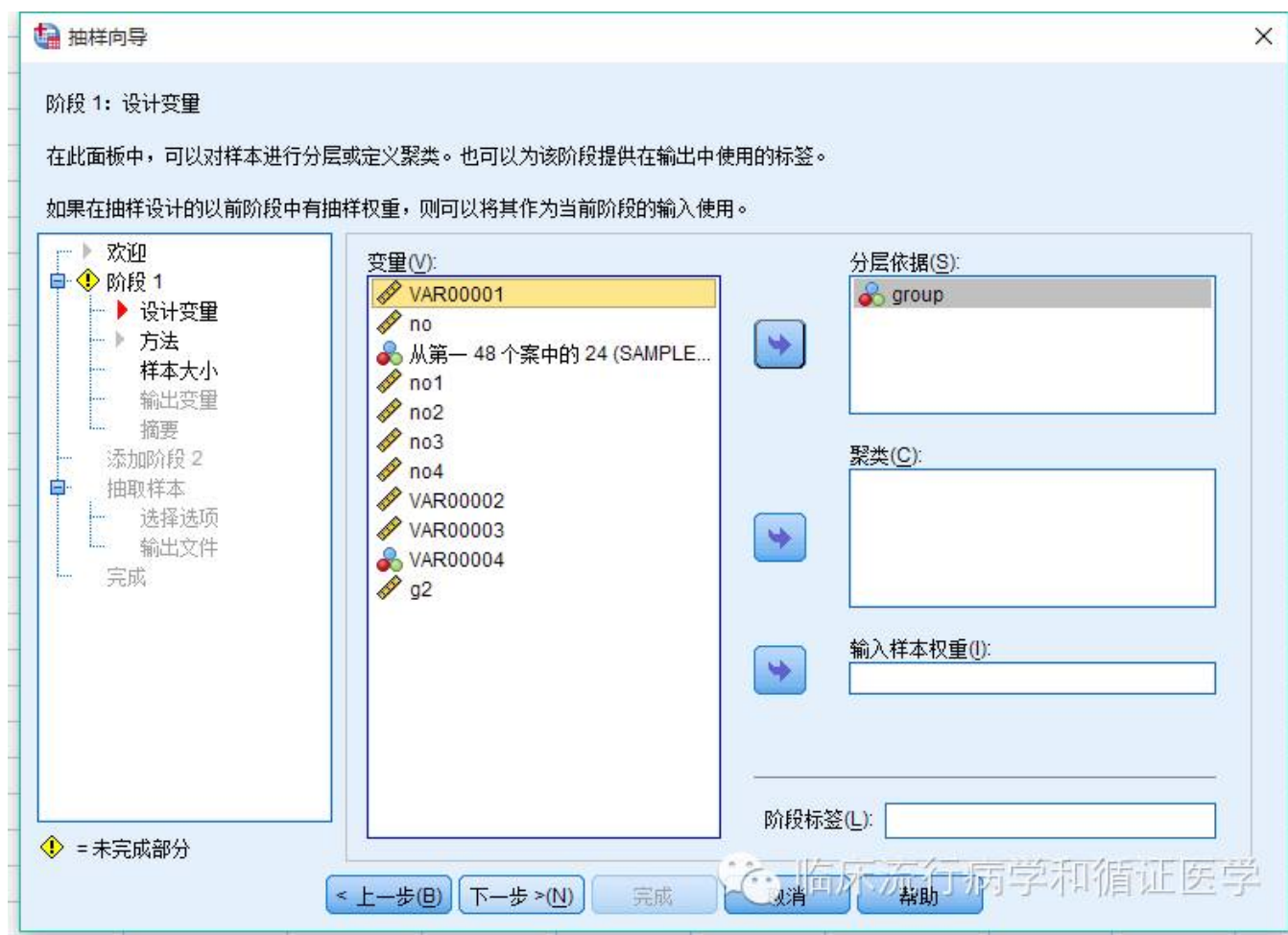
☐ 抽取样本

如果您已经拥有计划文件并且想绘制样本，选择此选项。

文件(E):

临床流行病学和循证医学





3、利用“计算变量”窗口来实现

a. 设定种子数，方法同上。

b. 点击转换、计算变量，在目标变量中给变量起一个名字，在函数组中点选随机数字，点选任意一个随机函数，填入适当的条件，点击确定即可。我个人常用Rv.Normal或Rv.Uniform函数。

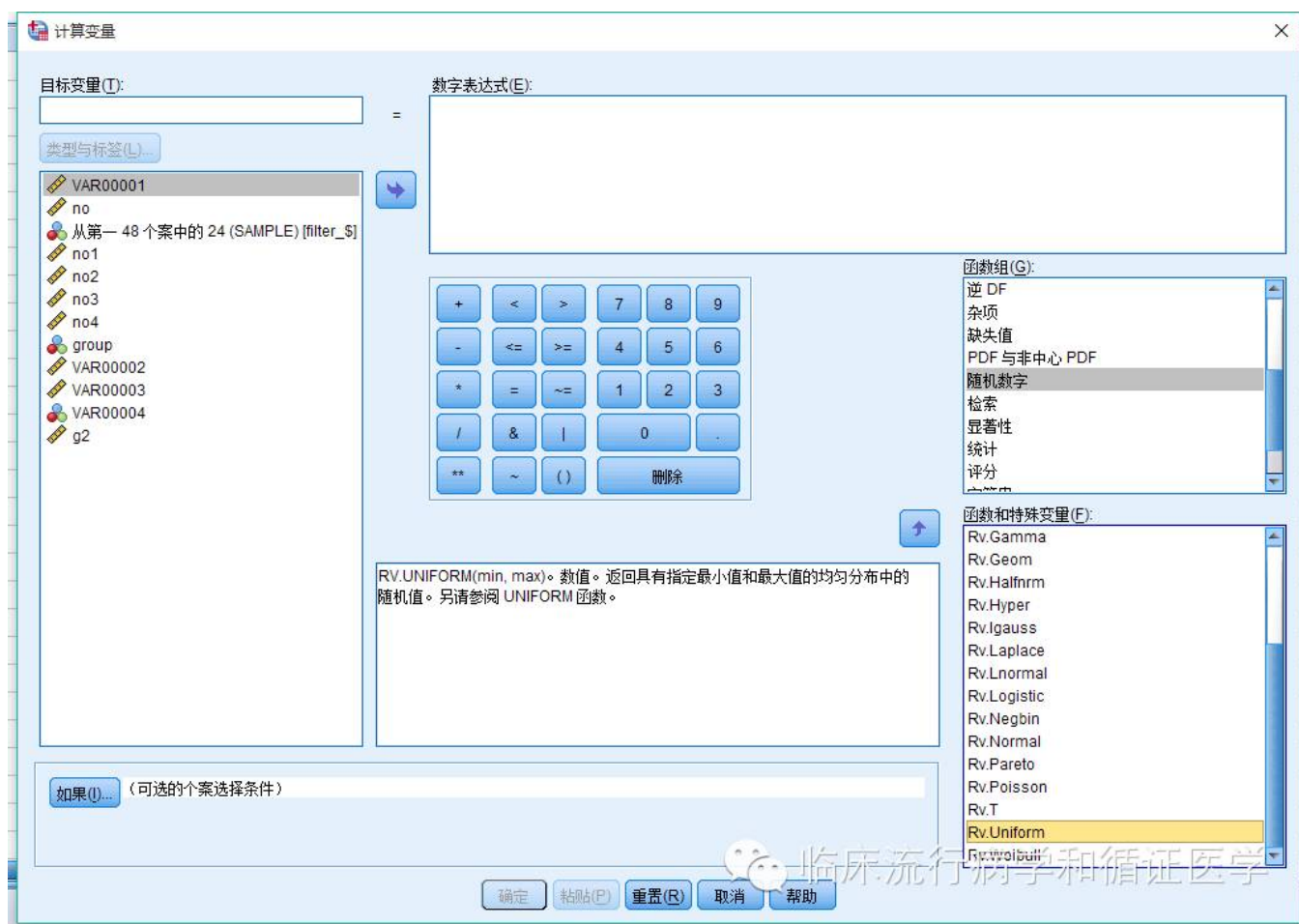
文件(E) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 直销(M) 图形(G) 实用程序(U) 窗口(W) 帮助(H)

61: VAR00003 .8079

	filter_\$	no1	no4	group	VA
58	-	-	-	-	-
59	-	-	-	-	-
60	-	-	-	-	-
61	-	-	-	-	-
62	-	-	-	-	-
63	-	-	-	-	-
64	-	-	-	-	-
65	-	-	-	-	-
66	-	-	-	-	-
67	-	-	-	-	-
68	-	-	-	-	-
69	-	-	-	-	-
70	-	-	-	-	-
71	-	-	-	-	-
72	-	-	-	-	-
73	-	-	-	-	-
74	-	-	-	-	-

计算变量(C)...
 可编程性转换...
 对个案内的值计数(O)...
 转换值(F)...
 重新编码为相同变量(S)...
 重新编码为不同变量(R)...
 自动重新编码(A)...
 创建虚拟变量
 可视分箱化(B)...
 最优分箱化(I)...
 准备建模数据(P)
 个案等级排序(K)...
 日期和时间向导(D)...
 创建时间序列(M)...
 替换缺失值(V)...
 随机数字生成器(G)...
 运行挂起的转换(T) Ctrl+G

临床流行病学和循证医学



c. 获得随机数字之后，你可以采用任意规则去进行随机抽样和分组，当然这个规则是你在生成随机数字之前就定好的，如你打算选取排序前50%的个案入A组等。

更多统计方法 请访问 <https://www.iikx.com/news/statistics/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](https://www.iikx.com)转发