
随机分组编制方法哪家强

作者：张华 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/1823.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！

随机分组

堪称临床研究史上最伟大的发明，随机分组中最重要的一个工作就是随机序列的产生，合适的随机序列产生方法是关键。你知道的随机序列编制方法有哪些？哪个方法你认为最好？在这里我个人认为“合适”具备4个维度：科学

(使用科学的方法产生，而不能是“随意”)，可重复

(可在同样的条件下产生相同的序列)，均衡性好，不可预测

(不能通过已经公开的序列号预测下一个研究对象的组别)。科学和不可预测两个维度比较好理解。均衡性好指总体分组例数的比例符合我们的预期，不能相差太大，同时在选取随机序列某一段，也具备均衡性，如AAAAAABBBBBB序列是不可取的。可重复主要用于质量控制的需要，当我们说我们的序列是随机产生的，对于其它研究者或者杂志编辑怎么能知道是否是“科学”产生的？还是研究者“随意编制”的？下面我们看几种随机方法的优缺点。

投硬币或者查询随机数字表法

这种方法做法是：事先确定一个规则，然后通过投硬币或者查询随机数字表确定每个研究对象的分组。这种方法是最原始的随机方法，也是最科学的随机方法，但却不是不合适的方法。原因在于：1.易受主管因素影响，即可能违反规则，通过重复投掷硬币，得到想要分组结果；2.不具备可重复性，这种序列结果重复出现的可能性非常小；3.组间例数不均衡，两组例数可能相同，也可能相差较大，特别对于小样本的序列。

通过软件产生的简单随机序列

现在SAS、R、Stata等软件都可以按我们设置的条件产生随机序列。(简单随机方法见公众号第7期文章。)这种方法不再受主观因素的影响，同时设定了“种子数”后可以重复产生相同的序列，便分组序列号可以被其它人验证，方便进行质量控制。但科学性却稍微下降，它不是真正意义上的随机，而是伪随机方法，是计算机从后台随机序列中取出部分随机数字通过一定的规则产生随机序列，而从哪个地方取随机数字不是随机的，而且可以设置“种子数”固定取数的位置。其次简单随机序列可能存在局部均衡性不好的问题，如AAAAAABBBBBB序列也可能出现。

通过软件产生的区组随机序列

这种方法是通过将随机序列分成N段，每段称作一个区组，如12个研究对象的序列可以分为三个区组，每个区组4个研究对象，在每个区组内保持分组的比例的条件下进行随机，从而达到总体和局部都能保持均衡性。(区组随机方法见公众号第8期文章。)如ABBA BABA BBAA序列，每四

个序列里包括两个A和B。但这种方法也存在问题，除了科学性下降外，每个区组是固定的比例，如知道了前3个序列号后，第4个可“预测”。

通过软件产生的区组随机序列是现在应用最多的临床研究随机分组方法，对于它的缺陷，现在也有人提出一些解决方法，比如混合区组方法。这种方法是设置区组大小不同，如区组4和6都同时存在，而且区组大小也是随机产生的，这样我们不知道区组大小时，无法预测下一个序号的分组。另外一种方法是在生成的随机序列中随机插入几个不同长度的序列，打乱原来固定区组序列，从而使研究参与者无法“预测”。

现在我中心根据研究研究的需要，已经编制了相关的程序，可实现简单随机、区组随机、混合区组随机和插补区组随机等随机序列产生方法，服务于临床研究人员，如果有需要可以联系我们。如果您有更好的随机序列产生方法，也可以我们联系并一起探讨更合适的随机序列产生方法。

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