

第 1 篇 统计分析基础

第 1 章 应确保数据是值得分析的

一提起统计分析，很多人脑海中闪现出来的第一个问题是：应该选择什么统计分析方法分析面前的资料？这是对统计学一知半解的一个明显标志！统计分析应该被使用在“值得分析的资料”上，而不是被使用在任何资料上，更不是被使用在不知任何专业背景的所谓“大数据”上！

1.1 什么是数据和/或统计资料

1.1.1 数据不等于统计资料

【问题】 下面列出了 300 个数据，请问：利用这些数据可以做哪些统计分析？

0.6 0.8 0.9 1.1 1.2 1.4 1.5 1.6 1.8 1.9 2.1 2.2 2.4 2.5
2.6 2.8 2.9 3.1 3.2 3.4 3.5 3.6 3.8 3.9 10.1 10.2 10.4 10.1
4.2 4.2 4.3 4.4 4.5 4.5 4.6 4.7 4.8 4.8 4.9 5.0 5.1 5.2
5.2 5.3 5.4 5.5 5.5 5.6 5.7 5.8 5.8 5.9 6.0 6.1 6.2 6.2
6.3 6.4 6.5 6.5 6.6 6.7 6.8 6.8 6.9 7.0 7.1 7.2 7.2 7.3
7.4 7.5 7.5 7.6 7.7 7.8 8.1 8.1 8.2 8.3 8.3 8.4 8.5 8.5
8.6 8.7 8.7 8.8 8.9 8.9 9.0 9.1 9.1 9.2 9.3 9.3 9.4 9.5
9.5 9.6 9.7 9.7 9.8 9.9 9.9 10.0 10.1 10.1 10.2 10.3 10.3 10.4
10.5 10.5 10.6 10.7 10.7 10.8 10.9 10.9 11.0 11.1 11.1 11.2 11.3 11.3
11.4 11.5 11.5 11.6 11.7 11.7 11.8 11.9 12.1 12.2 12.2 12.3 12.4 12.5
12.5 12.6 12.7 12.8 12.8 12.9 13.0 13.1 13.2 13.2 13.3 13.4 13.5 13.5
13.6 13.7 13.8 13.8 13.9 14.0 14.1 14.2 14.2 14.3 14.4 14.5 14.5 14.6
14.7 14.8 14.8 14.9 15.0 15.1 15.2 15.2 15.3 15.4 15.5 15.5 15.6 15.7
15.8 15.8 16.1 16.2 16.3 16.3 16.4 16.5 16.6 16.7 16.8 16.8 16.9 17.0
17.1 17.2 17.3 17.3 17.4 17.5 17.6 17.7 17.8 17.8 17.9 18.0 18.1 18.2
18.3 18.3 18.4 18.5 18.6 18.7 18.8 18.8 18.9 19.0 19.1 19.2 19.3 19.3

19.4 19.5 19.6 19.7 19.8 20.2 20.3 20.5 20.7 20.8 21.0 21.2 21.3 21.5
 21.7 21.8 22.0 22.2 22.3 22.5 22.7 22.8 23.0 23.2 23.3 23.5 23.7 24.2
 24.4 24.7 24.9 25.1 25.3 25.6 25.8 26.0 26.2 26.4 26.7 26.9 27.1 27.3
 28.3 28.6 28.9 29.1 29.4 29.7 30.0 30.3 30.6 30.9 31.1 31.4 32.4 32.8
 33.2 37.6 34.0 34.4 34.8 35.2 35.6 27.8 27.7 37.7 38.2 38.8 39.3 9.9
 40.6 9.9 41.9 42.5 45.3 46.5

【解答】 研究奥林匹克数学的人们会试图找出任何几个相邻数据之间的内在关系，最理想的结果是“推导”出一个“递推公式”，给定数据所在的序号 i ，就能推算出 x_i 等于多少。而纯数理统计学家们会不假思索地说，可以绘制频数直方图或箱式图，形象化地呈现这些定量数据的频数分布规律；还可计算三种区间（即置信区间、容许区间和预测区间），进而做出某些统计推断；若给定了标准值，还可对样本所抽自的总体均值与给定标准值之间差别是否具有统计学意义进行差异性分析，简称为单样本均值的假设检验。

其实，上面两种人所做的“统计工作”都有一个重要的前提条件，那就是此数据是值得分析的。然而，并非所有数据都是值得分析的！换一句话说，对数据进行分析之前，必须证明前提条件是成立的，否则，就是数字游戏！

事实上，值得分析的数据最起码要明确如下几条前提条件：收集这些数据的目的是什么？数据取自的总体是什么？这些数据分别在哪些条件下收集到的？数据的专业含义是什么？数据的名称是什么？数据的性质（定量或定性）是什么？若属于定量数据其单位是什么？测量数据的精确度有多高？由此可知，不明确这些前提条件的“一串数”可以称为“数据”；而明确这些前提条件的“一串数”才可以称为“统计资料”。

现在，全球有很多人在炒作“大数据”。大多数人或研究课题所呼喊的“大数据”，其实对应的正是前述提及的“数据”，而不是值得分析的“统计资料”！

1.1.2 统计资料的要素

1. 统计资料的基本要素

统计资料的基本要素有三条，即变量名称、专业含义和度量单位。例如，前面提及的那 300 个数据。若用变量 x 代表每个数据，且已知其专业含义为“尿汞值”，又知其单位是“ $\mu\text{g/L}$ ”，此时就可以说，这 300 个数据提供的信息已基本清晰了，它们基本上可以被称为一组统计资料了。

2. 统计资料的全部要素

严格地说，仅仅知道一组统计资料的基本要素还是不能确定它是否值得分析。例如，前面提及的那 300 个数据，即使知道了它们是 x （尿汞值， $\mu\text{g/L}$ ）的具体取值，但不知道收集这些数据的目的是什么、这些数据所测自的受试对象是什么、这些受试对象是否来自同一个总体。如果研究目的是估计某地区正常成年人尿汞值的 80% 参考值范围，这个地区全部正常成年人构成了一个特定的总体，而这 300 个数据所测自的受试对象正是从这个总体中完全随机或按重要非试验因素分层随机抽取的（目的是为了提提高样本对于总体的代表性），那么便有理由认为，对这 300 个数据组成的统计资料进行各种统计表达与描述和各种统计分析，是值得的。

由此可知，统计资料的全部要素除了它应具备基本要素外，还应具备：有明确定义的与特

定研究目的对应的总体；采取随机的方法从总体中获得足够大样本含量的样本；确保样本对于总体具有极好的代表性。总而言之，统计资料的全部要素应包括以下八条：

①研究目的；②特定总体；③合适的样本含量；④随机获取样本；⑤样本对于总体有极好的代表性；⑥变量名称；⑦专业名称；⑧单位（一般仅在定量变量时要求提供变量的单位，在一定程度上反映了测定数值的准确程度）。

1.2 确保数据值得分析的第一道关——制订科学完善的课题设计方案

1.2.1 什么叫科学研究

科学研究就是人类探索未知事物和现象内在规律和变化与发展规律的一个漫长而又艰辛的过程。在这个过程中，人类前仆后继，不断求索，提出了一个又一个科学问题，通过不懈努力，攻克了一个又一个科学堡垒，使人类社会不断进步，不断迈向新的征程。

1.2.2 科学研究与课题之间是什么关系

古人常出自个人兴趣和爱好，自觉地开展某些研究，一般都是小规模；而当今人类，绝大多数都出自某个集团（如某个国家、军队）的利益需求，开展大规模的科学研究，一般都需要一个团队来共同实施，有时甚至需要全球共同齐心协力，方可实现，如人类基因组计划项目。显然，大规模的科学研究计划，参与的人员多、耗时长、需要的经费多，必须在物质条件和精神条件都处于极好的配置状态下，才有可能实现预期的研究目标，获得预期的研究成果，而体现这种“配置”的具体研究计划或任务被称为“课题”。从而，“科学研究”这一笼统的概念，就被一系列具体的科研课题所取代了。所以，一般来说，科学研究都需要以课题或项目的方式来体现或落实。

1.2.3 做课题之前为什么要制订课题设计方案

一个大中型科研课题，虽然只有一个最终的研究目标，但通常该目标都可被分解成若干甚至几十或几百个具体的小目标，需要具有多学科知识结构的众多人才进行一年甚至数年艰苦卓绝的探索和研究，才有可能实现（当然，也有可能失败）。事实上，任何科研课题都是一个“系统工程”。如果边做边看，干不下去再重新开始，这种思维方式和做事风格，会使失败的概率必然大于成功的机会。所以，要想尽可能有把握做好一个科研课题，事先必须制订出科学完善的课题设计方案。

1.2.4 课题设计方案有哪些种类

课题设计方案包括课题框架设计方案和课题技术设计方案两大类。对于任何一个科研课题而言，要想圆满地完成事先确定的研究目标，要考虑的问题或方面往往都是很多的，稍有不慎，一旦拉下了某些重要的方面或在某些方面考虑不周，就可能导致整个课题研究前功尽弃，因此，事先制订出科学完善的科研设计方案，显得尤为重要。

科研设计的涵盖面非常广，其基本内容概括起来如图1-1和图1-2所示，分别是从结构上和功能上来划分的。图1-2实际上就是图1-1中“课题技术设计方案”的一种变形。

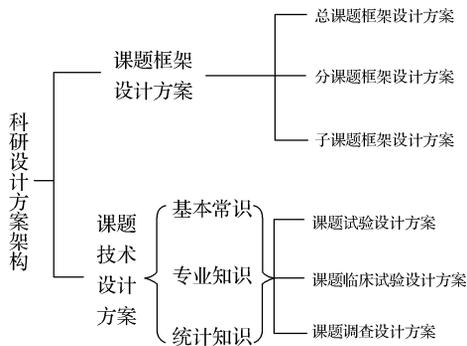


图 1-1 科研设计方案架构图

在图 1-2 中，专业设计包括基本常识与专业知识；统计设计包括试验性和观察性研究设计，而试验性研究设计又根据受试对象的不同，可进一步划分为试验设计(以动物或样品为受试对象)和临床试验设计(以健康人或某病患者为受试对象)。

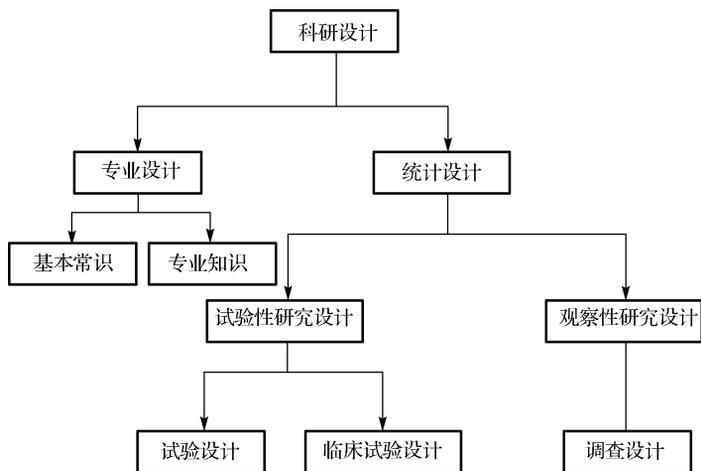


图 1-2 科研设计框图

在图 1-2 中，似乎仅涉及统计学知识，其实不然，任何一个科研课题，只具备统计学知识而不具备与拟研究问题有关的专业知识的制订者，制订不出“有用”的“统计设计方案”。事实上，制订统计设计方案的人，不仅需要具备丰富的专业知识，更需要具备正常人的思维和心态，要用够用足“基本常识”，并且要会将基本常识、专业知识和统计知识相互融合，才有可能制订出具有极高使用价值的统计设计方案，如图 1-3 所示。



图 1-3 课题统计设计方案的内核

1.2.5 科学完善的科研设计方案的标志

在科研课题获得批准之后且课题正式实施之前，应制订出科学完善的课题设计方案。

一个课题设计方案制订得是否科学完善，应有一些重要的标志。其具体内容如下：

(1)人力、物力、财力和时间满足设计要求。也就是说，申请书上计划的科研条件，在实际科研环境中均能一一兑现。例如，检测某种物质需要一种昂贵的仪器设备，而本课题组已经拥有、可以借到或可以请他人代为检测，若国内没有一个单位拥有，则属于“空头支票”，不具备完成科研课题的基本条件。又比如，完成此科研课题需要10个专业领域的专家参与，而现有研究人员仅涉及6个专业领域，另外4个专业领域的研究人员目前几乎不可能得到。笔者在检查重大和重点科研课题时经常发现，一些涉及多因素多指标的复杂科研课题，按照所制订的研究目标，必须有高水平的统计学专家和计算机专家花较多时间自始至终参与，并花大气力研究和攻关，方可达到事先确定的研究目的。然而，在那些课题组的全部参与人员中，却找不到这两类专家，这属于人力或技术水平上达不到设计要求。再比如，要想完成某项难度较大的科研课题，至少需要5年时间，而申请时考虑不周，计划3年完成，这属于时间上不符合设计要求，事先夸下海口，而无力实现。

(2)试验设计的“三要素(受试对象、影响因素、试验效应)”、“四原则(随机原则、对照原则、重复原则、均衡原则)”、“设计类型(安排因素及其水平的一种架构)”和“质量控制(确保所获得的科研数据准确可靠的措施)”均符合专业和统计学要求。这方面的内容很多，请参见《医学统计学》相关专著。这里仅举一例，拟使用某种动物制造一种用于试验的动物模型，但这种动物很特殊，根据目前的条件无法获得，这就属于三要素中的“受试对象”不符合设计要求。

必须强调指出的是，三要素中的“受试对象”，应选定合适的种类并把握质量标准 and 数量要求。质量标准中包括五类标准的制订(纳入标准、排除标准、中止标准、终止标准、剔除标准)；数量要求指有根据地估算出每个小组中需要的合适样本含量。“影响因素”中应特别重视找全找准重要非试验因素，并尽可能按重要非试验因素对受试对象进行分层随机化。“试验效应”方面，注意三类指标(诊断性指标、疗效评价指标和安全性评价指标)都应有主次之分，还应注意指标的测定时间和测定方法。

在四原则方面，应巧妙地使用分层随机化，确定合理的对照组，有根据地估计各小组所需要的样本含量和确保组间具有高度的均衡性。

(3)重要的影响因素(包括试验因素和非试验因素)和观测指标没有遗漏，并做了合理安排。影响因素和试验效应属于三要素中的两个重要要素，再一次特别提出来，是为了强调其重要性。因为在很多科研课题的设计中，研究者常将某些重要的试验因素和体现试验效应的观测指标遗漏掉了，导致科研结果回答不了需要回答的问题，甚至很容易得出错误的结论。

(4)必须找准找全重要的非试验因素并使之都得到很有效的控制。很多科研课题的结论经不起时间和实践的检验，问题的症结就在于重要的非试验因素未得到有效的控制，它们“污染”了试验结果，从而得出了不够正确甚至错误的结论。

(5)研究过程中可能出现的各种情况都已考虑在内，并有相应的对策。这点也是至关重要的。比如，在新药临床试验研究中，若中途退出的受试者人数较多，可能会导致整个临床试验前功尽弃，一旦发现可能出现较高比例的受试者退出临床试验，应预先想好对策，尽可能降低退出比例。再比如，某些受试者的某些测定指标的数值明显偏高或偏低，应事先对此种可能的

异常现象有所警惕，必要时，应对所有受试者进行多次重复观测，用每位受试者该指标多次测定值的平均值取代其测定指标的数值，以降低个别受试者在个别时间点上出现难以处理的异常值。有时，极少数异常值可能会颠覆整个研究结论。

(6)对操作方法，试验数据的收集、整理、分析等均有一套明确的规定和方法。尤其是有多个单位、多位研究者合作的研究课题，统一规定试验操作规程，统一培训，仪器和试剂统一质量指标和校准，有规范化且符合试验设计要求的收集和整理资料的表格，对获得准确可靠试验资料是极其重要的。有了数据后，由谁来进行数据处理、采用哪些统计分析方法分析，事先都应做到胸中有数，不要等到试验都做完了，再来考虑请谁做数据分析。因为到那时，很可能这个科研课题的数据已无分析的价值了。为什么会这样呢？若科研设计有严重错误或质量控制有问题或收集资料的方法不正确，都可能得到包含很多错误信息的资料，无论用何种高级的统计分析方法去分析，其结论都是荒谬的。也就是说，做统计分析的人应参与开始的科研设计，应在科研过程中参与质量控制，应协助研究人员确保科研资料的原始性和准确、可靠、完整。

1.3 确保数据值得分析的第二道关——实时进行严格的过程质量控制

1.3.1 必须严格控制课题实施过程中的质量

至于科研课题实施过程中的质量控制，事先考虑的内容和拟采取的对策应该包含在科研设计方案之中。若在实施过程中新出现一些影响结果正确性的问题但的确未包括在事先制订的科研设计方案之中，应临时补充和修改与之有关的科研设计方案的内容，尽可能把“突发事件”所造成的坏影响降到最低程度。

在科研工作中，较多的属于试验性研究。首先，通过图 1-4 粗略地了解质量控制的大致内容；然后，针对试验研究过程中如何做好质量控制进行详细论述。



图 1-4 课题实施过程中质量控制的内容

1.3.2 进行质量控制的必要性

由图 1-4 可知，无论试验设计方案制订得多么科学严谨，它毕竟是纸上的东西，要把纸上写的东西落到实处，需要通过实施者所持有的态度和采取的具体行动，需要受试对象（特别是人）的全力配合，需要环境和条件处于设计者所希望的状态。然而，在较长时间的试验过程中，有诸多方面都或多或少会发生一些变化，一旦这些变化超出了“允许”的范围，试验结果就会偏离其真实值较远，结论的可信度就必然会大打折扣。因此，对试验过程中出现的可能会影响

结果准确度的任何问题，都应能及时发现，并能在尽可能短的时间内消除其影响。因此，在试验过程中，进行质量控制是非常必要的。

1.3.3 进行质量控制的环节与措施

需要进行质量控制的环节包括：控制研究者的责任心、心理状态和技术水平，控制受试者的心理状态和不依从性，控制环境的变化(如温度、湿度、光线、清洁度等)，控制试验条件的改变(如仪器的状态发生改变、试剂过期、观测指标的单位不同、测定的方法不同等)。这些环节都对试验结果可能造成不利影响。

对试验过程的质量进行控制的措施有：对不同单位的多名研究者应进行统一的培训，内容包括增强责任心教育、调适好心理状态(在言行上不应有个人的倾向性或偏好)、接受正规的技术培训，并要求每位研究者应自觉遵守统一的操作规程。对于动物或样品类受试对象而言，应当选用与研究目的和试验因素相对应(同质)的动物或样品；对于某病患者类受试对象而言，应当制订严格的纳入与排除标准；应尽可能采取双盲法(受试对象和研究者都不知道具体的分组结果)或三盲法(受试对象、研究者和数据统计分析者都不知道具体的分组结果)分配受试对象，尽可能减少或消除来自研究者，特别是受试者心理的影响。对于环境因素，应尽可能创造条件，使试验研究在特定的温度、湿度、光线和清洁度条件下进行；对于仪器设备、试剂应尽可能校准，并经常对其进行检查和调试；对于指标的测定单位、测定方法应完全统一，实在不能统一的，应采取合适的方法进行换算；对于主要评价指标，若测定方法无法统一，应确定一个有资质的测定部门来完成测试，称为“中心试验室”，这是消除多个试验研究单位测定结果参差不齐的重要措施之一。

1.4 确保数据值得分析的第三道关——确保数据的原始性没有被破坏

1.4.1 应有切实可行的措施确保收集的数据具有原始性

对于试验研究课题，应有与试验设计方案匹配的记录原始数据的规范化表格；对于临床试验研究课题，应有与临床试验设计方案匹配的临床病例报告表(简称CRF)；对于调查研究课题，应有与调查设计方案匹配的调查表。

1.4.2 与常见试验设计类型对应的规范化统计表

1. 单组设计

【例1-1】 为探讨结核病人的营养状况，某研究者对2000年收治的45例病人的营养状况进行了调查，调查的指标包括白蛋白(ALB)、白蛋白/球蛋白(A/G)比值。数据见表1-1，已知这些指标的下限分别为：白蛋白35g/L、白蛋白/球蛋白1.5。试评价结核病人的营养状况。

2. 配对设计

【例1-2】 对血小板活化模型大鼠以ASA进行试验性治疗，以血浆TXB₂(ng/L)为指标，其结果见表1-2。

表 1-1 45 例结核患者的营养状况

| ID | ALB | A/G | ID | ALB | A/G | ID | ALB | A/G |
|----|---------|--------|----|---------|--------|----|---------|--------|
| 1 | 30.2527 | 0.9332 | 16 | 29.4046 | 0.4953 | 31 | 37.1407 | 0.4976 |
| 2 | 37.2461 | 1.1546 | 17 | 31.7185 | 1.3290 | 32 | 24.2056 | 1.0685 |
| 3 | 31.9332 | 1.1154 | 18 | 26.3070 | 0.9545 | 33 | 34.1015 | 0.6992 |
| 4 | 38.2970 | 0.8270 | 19 | 42.2041 | 1.0580 | 34 | 36.4848 | 0.8745 |
| 5 | 29.7258 | 1.4642 | 20 | 28.2290 | 0.8142 | 35 | 41.3322 | 1.0341 |
| 6 | 40.2492 | 1.1738 | 21 | 32.0035 | 0.9595 | 36 | 28.3449 | 1.1210 |
| 7 | 29.6518 | 1.7670 | 22 | 33.4976 | 0.8141 | 37 | 41.3502 | 1.1723 |
| 8 | 33.9290 | 0.7915 | 23 | 28.7575 | 0.9331 | 38 | 47.1740 | 0.5704 |
| 9 | 35.8982 | 1.4726 | 24 | 23.6329 | 1.3659 | 39 | 25.1091 | 0.9331 |
| 10 | 33.2792 | 1.2376 | 25 | 32.3452 | 0.7535 | 40 | 30.5990 | 0.9196 |
| 11 | 33.0777 | 0.3949 | 26 | 35.2587 | 0.7430 | 41 | 33.3743 | 0.6501 |
| 12 | 23.7641 | 0.8388 | 27 | 33.1799 | 1.3128 | 42 | 28.5072 | 0.6796 |
| 13 | 41.5968 | 1.1160 | 28 | 34.1699 | 0.8658 | 43 | 34.8188 | 0.9514 |
| 14 | 39.0836 | 1.2769 | 29 | 31.8015 | 1.2527 | 44 | 29.3374 | 0.6549 |
| 15 | 33.1353 | 1.2509 | 30 | 31.5403 | 0.9285 | 45 | 33.9863 | 1.0733 |

3. 单因素 K 水平设计 ($K \geq 2$)

【例 1-3】 试就表 1-3 资料说明大白鼠感染脊髓灰质炎病毒后, 再进行伤寒或百日咳预防接种是否影响生存日数。

表 1-2 大鼠血小板活化模型 ASA 治疗前后血浆 TXB₂ 的变化 (ng/L)

| 大鼠号 | 血浆 TXB ₂ (ng/L) | |
|-----|----------------------------|-----|
| | 给药前 | 给药后 |
| 1 | 250 | 184 |
| 2 | 226 | 205 |
| ... | ... | ... |
| 10 | 176 | 176 |

表 1-3 感染脊髓灰质炎病毒后的 3 组大鼠接受不同处理后的生存情况

| 编号 | 生存日数 | | |
|----|------------|---------|------|
| | 处理: 接种伤寒疫苗 | 接种百日咳疫苗 | 空白对照 |
| 1 | 5 | 6 | 8 |
| 2 | 7 | 6 | 9 |
| 3 | 8 | 7 | 10 |
| 4 | 11 | 10 | 12 |
| 5 | 11 | 16 | 14 |
| 6 | 12 | 11 | 16 |

4. 无重复试验随机区组设计

【例 1-4】 某研究者欲比较 3 种抗癌药物对小白鼠肉瘤的抑瘤疗效, 先将 15 只染有肉瘤的小白鼠按体重大小配成 5 个区组, 使每个区组内的 3 只小白鼠体重最接近, 然后随机决定每个区组中的 3 只小白鼠接受 3 种药物的治疗, 以肉瘤的质量为指标, 试验结果见表 1-4。试比较不同抗癌药物对小白鼠肉瘤的抑瘤效果之间的差别是否有统计学意义。

5. 有重复试验随机区组设计

【例 1-5】 现有 3 种降低转氨酶的药物 A、B、C, 为了考察它们对甲型肝炎和乙型肝炎患者转氨酶降低的程度之间的差别是否有统计学意义, 收集到的试验数据见表 1-5 (即从两型患者的总体中各随机抽取 9 例, 然后分别随机均分入 3 个药物组中去)。此为具有 3 次独立重复试验的随机区组设计 (以前称为 3×2 析因设计, 但不够严格, 因为“型别”是一个属性因素)。其中, 药物是试验因素, 型别是属性因素。

表 1-4 不同药物作用后小白鼠肉瘤质量

| 区 组 | 肉瘤质量(g) | | | |
|-----|---------|------|------|------|
| | 药物种类: | A | B | C |
| 1 | | 0.82 | 0.65 | 0.51 |
| 2 | | 0.73 | 0.54 | 0.23 |
| 3 | | 0.43 | 0.34 | 0.28 |
| 4 | | 0.41 | 0.21 | 0.31 |
| 5 | | 0.68 | 0.43 | 0.24 |

表 1-5 3 种药物对两型肝炎患者转氨酶降低效果的试验结果

| 肝炎类别 | 转氨酶降低值(kPa) | | | | | | | | | |
|------|-------------|-----|----|-----|-----|----|-----|-----|----|----|
| | 药物种类: | A 药 | | | B 药 | | | C 药 | | |
| 甲型 | | 100 | 85 | 90 | 120 | 90 | 110 | 50 | 60 | 40 |
| 乙型 | | 65 | 75 | 100 | 45 | 30 | 50 | 50 | 60 | 45 |

6. 具有一个重复测量因素的单因素设计

【例 1-6】 某外科医生在每个病人手术过程中的 3 个不同阶段测得病人的门静脉压力资料见表 1-6。研究者希望考察 3 个阶段门静脉压力之间差别是否具有统计学意义。

表 1-6 10 名病人在手术过程中的 3 个不同阶段门静脉压力的测定结果

| 病 例 号 | 门静脉压力(cmH ₂ O) | | | |
|-------|---------------------------|----------------|----------------|----------------|
| | A(阶段): | A ₁ | A ₂ | A ₃ |
| 1 | | 40 | 36 | 36 |
| 2 | | 19 | 17 | 18 |
| ... | | ... | ... | ... |
| 10 | | 36 | 37 | 38 |

7. 平衡不完全随机区组设计

【例 1-7】 用 4 种方法治疗脚气, 受试者为两脚都患脚气的患者, 每只脚接受一种处理, 观察指标为治疗效果评分。设计格式和资料见表 1-7。

表 1-7 四种方法治疗脚气的评分

| 受试者编号 | 疗 效 评 分 | | | | |
|-------|---------|---|---|---|---|
| | 治疗方法: | 甲 | 乙 | 丙 | 丁 |
| 1 | | 5 | 3 | — | — |
| 2 | | — | — | 4 | 7 |
| 3 | | 2 | — | 6 | — |
| 4 | | — | 4 | — | 8 |
| 5 | | 3 | — | — | 7 |
| 6 | | — | 2 | 5 | — |

8. 无重复试验的双因素设计

【例 1-8】 将落叶松苗木栽在 4 块不同的苗床(B)上, 每块苗床的苗木又分别使用 3 种不同的肥料(A)以观察肥效差异, 2 年后, 于每一苗床的各施肥小区内用随机方法抽取一株, 测量其高度, 资料见表 1-8。

表 1-8 不同肥料 A 和不同苗床 B 对落叶松苗高影响的测定结果

| 肥料 A | 苗高 (cm) | | | | | |
|----------------|---------|----------------|----------------|----------------|----------------|-----|
| | 苗床 B: | B ₁ | B ₂ | B ₃ | B ₄ | 合计 |
| A ₁ | | 55 | 47 | 47 | 53 | 202 |
| A ₂ | | 63 | 54 | 57 | 58 | 232 |
| A ₃ | | 52 | 42 | 41 | 48 | 183 |
| 合计 | | 170 | 143 | 145 | 159 | 617 |

【特别提示】 当试验中同时涉及两个试验因素时，通常都需要先考察它们之间的交互作用是否具有统计学意义；而当未做独立重复试验时，对其定量指标的试验结果进行统计分析是无法考察交互作用项的效应大小的。因此，基于此设计所得出的结论是否正确取决于以下两点：其一，两试验因素之间的交互作用是否可忽略不计；其二，各试验条件下的试验结果是否十分稳定，其误差在专业上允许的范围之内。

9. 不可拆分的单个体型无重复取样的拉丁方设计

【例 1-9】 试评价 4 台秤的测定结果是否一致。样品质量为观测指标，设计和资料见表 1-9。

表 1-9 4 台秤称样品质量的一致性试验

| 样品代号 | 秤的代号 (样品质量, mg) | | | | |
|------|-----------------|------|------|------|------|
| | 测定顺序: | 1 | 2 | 3 | 4 |
| 1 | | C() | B() | A() | D() |
| 2 | | B() | A() | D() | C() |
| 3 | | A() | D() | C() | B() |
| 4 | | D() | C() | B() | A() |

【特别提示】 表中括号内填写样品质量；所选取的 4 个样品的质量最好不等，且包含实际使用时质量的最小值和最大值，这样有利于评价结果在一定质量范围内的准确性。

10. 可以拆分的单个体型有重复取样的拉丁方设计

【例 1-10】 观察 A、B、C 3 种中药的促凝作用，以生理盐水作为对照(D)。每次试验时，从动物身上抽取血样，往血样中加入不同药物(包括生理盐水)，观测指标为“血凝时间(s)”。现取品种相同、体重接近的雄性白兔 4 只，每只白兔的血样被均分成 4 份，以不同的顺序接受不同的药物处理，以血浆复钙凝固时间(s)为指标，设计和资料见表 1-10。

表 1-10 不同中药组测得的血凝时间

| 白兔编号 | 中药代号 (血凝时间(s)) | | | | |
|------|----------------|--------|--------|--------|--------|
| | 试验顺序号: | 1 | 2 | 3 | 4 |
| 1 | | C(85) | D(102) | A(90) | B(100) |
| 2 | | D(108) | A(88) | B(110) | C(80) |
| 3 | | A(95) | B(98) | C(82) | D(110) |
| 4 | | B(105) | C(78) | D(104) | A(92) |

11. 多个体型无重复试验的拉丁方设计

【例 1-11】 为了研究 5 个不同剂量的甲状腺提取液对豚鼠甲状腺重的影响，考虑到鼠的种系和体重对观测指标可能有一定的影响，设计试验时，最好将这两个重要的非处理因素一并

安排。根据专业知识得知,这三个因素之间的交互作用可忽略不计,设计格式和收集的资料见表 1-11。注意:每种种系的鼠随机抽取了 5 只。

表 1-11 5 个不同剂量的甲状腺提取液对豚鼠甲状腺重的影响

| 种 系 | 甲状腺提取液的剂量(字母)(甲状腺重(g/200g 体重)) | | | | | |
|-----|--------------------------------|-------|-------|-------|-------|-------|
| | 体重: | 1 | 2 | 3 | 4 | 5 |
| 1 | | C(65) | E(85) | A(57) | B(49) | D(79) |
| 2 | | E(82) | B(63) | D(77) | C(70) | A(46) |
| 3 | | A(73) | D(68) | C(51) | E(76) | B(52) |
| 4 | | D(92) | C(67) | B(63) | A(41) | E(68) |
| 5 | | B(81) | A(56) | E(99) | D(75) | C(66) |

注: A、B、C、D、E 所代表的具体剂量从略。

12. 单个体型有重复试验的成组与配对 2×2 交叉设计

【例 1-12】 某麻醉科医生研究催醒宁对氟呱啉的作用,选用生理盐水(记为 C)作为催醒宁(记为 T)的对照药,拟用 6 只大鼠作为受试对象。先将 6 只大鼠随机均分成两组,每组 3 只;再用随机方法决定第 1 组 3 只大鼠接受两种处理的顺序,第 2 组中的 3 只大鼠接受两种处理的顺序相反。设计格式与资料见表 1-12。

【例 1-13】 为研究 A(90402 中药复方)、B(安慰剂)两种处理对提高高原劳动能力的影 响,以条件近似的 10 名健康人作为受试对象,把受试对象和测定顺序(冬季、春季)作为两个重要的非处理因素,每人都服用两种药物各一次。试验设计方案如下:将条件最接近的每两人配成一对,用随机的方法确定每对中一人使用 A、B 药物的顺序,另一个人服药的顺序相反。设计格式和试验结果见表 1-13,其观测指标为 PWC170(即把心跳校正到 170 次/min 时能做的功)。

表 1-12 考察催醒宁对氟呱啉的作用的试验结果

| 大鼠编号 | 是否用催醒宁与出现蹲和走动的时间(单位) | | |
|------|----------------------|-------|-------|
| | 试验顺序: | 第 1 次 | 第 2 次 |
| 1 | | T(15) | C(27) |
| 3 | | T(17) | C(28) |
| 6 | | T(5) | C(26) |
| 2 | | C(31) | T(25) |
| 4 | | C(30) | T(14) |
| 5 | | C(25) | T(18) |

注:(1, 2)、…、(5, 6)号分别为配对组;T 为用催醒宁, C 为不用催醒宁。

表 1-13 两种药物对提高高原劳动能力的试验结果

| 受试者编号 | 药物与 PWC170 | | | |
|-------|------------|-------|---|-------|
| | 试验时间: | 冬 | 季 | 春 |
| 1 | A | 159.4 | B | 153.8 |
| 2 | B | 129.4 | A | 159.8 |
| 3 | B | 122.1 | A | 137.6 |
| 4 | A | 130.4 | B | 137.8 |
| 5 | A | 150.7 | B | 140.2 |
| 6 | B | 129.9 | A | 146.1 |
| 7 | B | 166.3 | A | 210.9 |
| 8 | A | 208.7 | B | 169.9 |
| 9 | A | 180.1 | B | 150.7 |
| 10 | B | 143.4 | A | 150.8 |

注:(1, 2)、…、(9, 10)号分别为配对组;A 为 90402 中药复方, B 为安慰剂。

13. 多个体型有重复试验的 2×2 交叉设计

【例 1-14】 假定某公司原生产药物 B, 现对其剂型进行改造, 生产药物 A。用 AUC(血药浓度 - 时间曲线下面积)作为评价指标。试验目的是为了评价 A 药与 B 药是否具有相等的生物利用度。试验设计方案如下:选取 24 个双胞胎家庭(每个家庭有两个受试者属于同一组双

胞胎), 将他们随机均分成两组, 用药顺序为: 第 1 组每个家庭中的两个受试者在第 1 与第 2 试验周期分别接受 A 药与 B 药, 用 AB 表示; 第 2 组每个家庭中的两个受试者在第 1 与第 2 试验周期分别接受 B 药与 A 药, 用 BA 表示。在两个试验周期期间设立“洗脱期(无实际意义)”。设计格式与资料见表 1-14。

表 1-14 A、B 两种剂型药物影响下的血药浓度 - 时间曲线下面积 (AUC, $\mu\text{g}/\text{mL}$)

| 组 别 | 家庭编号 | AUC, $\mu\text{g}/\text{mL}$ | | 组 别 | 家庭编号 | AUC, $\mu\text{g}/\text{mL}$ | |
|-----|------|------------------------------|-------|-----|------|------------------------------|-------|
| | | 周 期 1 | 周 期 2 | | | 周 期 1 | 周 期 2 |
| AB | 1 | 84.4 | 70.9 | BA | 3 | 66.6 | 63.6 |
| | 2 | 101.0 | 92.7 | | 5 | 117.0 | 107.3 |
| | ... | ... | ... | | ... | ... | ... |
| | 21 | 65.4 | 63.0 | | 22 | 56.8 | 61.5 |
| | 23 | 61.9 | 51.9 | | 24 | 74.7 | 76.0 |

14. 嵌套(系统分组)设计

(1) 试验因素存在自然属性上的嵌套关系。

【例 1-15】 嵌套设计的第一个例子。为研究萝卜叶子中 M 物质的含量, 随机采集了 3 棵萝卜(设为 A_1 、 A_2 、 A_3), 各棵萝卜中随机取几片叶子(设为 B_1 、 B_2 、 B_3) 做测定, 结果见表 1-15。显然, M 物质含量的多少, 不仅取决于所取自的叶子, 更主要地取决于所取自的萝卜, 即不同萝卜之间的变异大于同一棵萝卜不同叶子之间的变异。

表 1-15 萝卜叶子中 M 物质的测定结果

| 编 号 | M 物质含量(单位) | | | | | | | |
|-----|-----------------------|------|-----------------------------|------|------|-----------------------|------|--|
| | A_1 (B_1 B_2) | | A_2 (B_1 B_2 B_3) | | | A_3 (B_1 B_2) | | |
| 1 | 3.28 | 3.52 | 2.46 | 1.87 | 2.19 | 2.77 | 3.74 | |
| 2 | 3.09 | 3.48 | 2.44 | 1.92 | 2.19 | 2.66 | 3.44 | |
| 3 | 3.31 | 4.07 | . | 2.10 | . | . | . | |

(2) 试验因素对定量结果指标的影响存在主次关系。

【例 1-16】 为了研究两种药物[人参二醇皂苷(PSD)、人参三醇皂苷(PTSD)], 分别取 6 个不同剂量对生物体骨髓红系祖细胞增殖的影响情况, 每个生物体仅接受一种药物的一个特定剂量的处理, 设计格式见表 1-16(药物种类之间的差别大于药物浓度之间的差别)。

表 1-16 PDS 与 PTS 对骨髓红系祖细胞的增殖作用 ($\bar{x} \pm s$, $n=10$)

| 药 物 种 类 | 骨髓红系祖细胞的含量(个/ 0.5×10^5) | | | | | | |
|---------|-----------------------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| | 药物浓度 (ng/ml): | 0.0 | 2.5 | 12.5 | 25.0 | 50.0 | 100.0 |
| PDS | | 115.2 \pm 20.6 | 143.8 \pm 20.8 | 144.3 \pm 29.6 | 154.3 \pm 28.2 | 178.5 \pm 30.2 | 165.9 \pm 22.4 |
| PTS | | 115.2 \pm 20.6 | 95.9 \pm 12.9 | 89.1 \pm 14.0 | 85.8 \pm 14.3 | 49.5 \pm 12.9 | 6.5 \pm 2.7 |

15. 裂区(分割)设计

(1) 试验因素施加于受试对象时存在先后顺序之分。

【例 1-17】 嵌套设计的第一个例子。为研究不同瘤株的生瘤效果和不同浓度蛇毒的抑瘤作用, 先将 48 只小鼠按原始体重分成 3 个随机区组(每个区组内的 16 只小鼠体重最接近), 将

每个区组内的16只小鼠随机地均分成4组，分别接种4种不同的瘤株，观察肿瘤生长情况。1d后再对接种同一种瘤株的4只小鼠分别腹腔注射4种不同浓度的蛇毒，连续用蛇毒抑瘤10d，停药1d后解剖测瘤重。设计格式和资料见表1-17。

表1-17 不同瘤株与不同浓度的蛇毒共同作用后对小鼠抑瘤效果的影响

| 瘤株 A | 蛇毒浓度 B | 瘤重(g) | | | |
|-----------------------|----------------|-------|------|------|------|
| | | 随机区组: | 1 | 2 | 3 |
| A ₁ (S180) | B ₁ | | 0.80 | 0.76 | 0.36 |
| | B ₂ | | 0.36 | 0.26 | 0.31 |
| | B ₃ | | 0.17 | 0.28 | 0.16 |
| | B ₄ | | 0.28 | 0.13 | 0.11 |
| A ₂ (HS) | B ₁ | | 0.74 | 0.43 | 0.57 |
| | B ₂ | | 0.50 | 0.46 | 0.32 |
| | B ₃ | | 0.42 | 0.20 | 0.20 |
| | B ₄ | | 0.36 | 0.26 | 0.32 |
| A ₃ (EC) | B ₁ | | 0.31 | 0.55 | 0.32 |
| | B ₂ | | 0.20 | 0.15 | 0.20 |
| | B ₃ | | 0.38 | 0.18 | 0.26 |
| | B ₄ | | 0.25 | 0.21 | 0.14 |
| A ₄ (ARS) | B ₁ | | 0.48 | 0.57 | 0.33 |
| | B ₂ | | 0.18 | 0.30 | 0.29 |
| | B ₃ | | 0.44 | 0.27 | 0.27 |
| | B ₄ | | 0.22 | 0.30 | 0.37 |

注：在本例中，瘤株先作用于受试对象，蛇毒后作用于受试对象。

(2) 一个完整试验被整批重复。

【例1-18】 嵌套设计的第二个例子。在某项试验中，涉及药物种类和药物剂量，但试验需要在不同时间重复进行。显然，时间改变是较慢的，而在每个确定的时间点上，药物种类与剂量的水平改变都是很方便的。若用大鼠作为受试对象，可选取来自大鼠的重要非试验因素作为子区组因素，还可以不同天做试验作为大区组因素，将药物种类与剂量进行组合构成析因结构，在以天为大区组因素的各水平下，再以时间段和时刻为两层子区组因素，按含区组因素析因设计分配大鼠。设计格式和资料见表1-18。

表1-18 不同药物、剂量和作用时间下大鼠胸膜腔内积液量

| 日期 | 时段 | 时刻 | 剂量 | 胸膜腔内积液量(\bar{x} , ml, $n=5$) | | | | | | | |
|-----|-----|-----|-----|----------------------------------|-----|------|-------|-----|-----|------|-----|
| | | | | 药物: | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 第1天 | 上午 | 1 | 1 | 5.7 | 8.6 | 6.9 | 6.6 | 6.7 | 7.4 | 5.7 | 6.7 |
| | | 1 | 2 | 5.1 | 7.2 | 6.8 | 6.4 | 6.6 | 8.7 | 6.7 | 7.0 |
| | 2 | 1 | 8.4 | 9.6 | 9.3 | 11.1 | 121.5 | 8.7 | 9.3 | 9.5 | |
| | 2 | 2 | 7.3 | 8.7 | 7.9 | 6.9 | 8.9 | 9.5 | 8.3 | 11.3 | |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |

续表

| 日期 | 时段 | 时刻 | 剂量 | 胸膜腔内积液量 (\bar{x} , ml, $n=5$) | | | | | | | | |
|-------|-----|-----|-----|-----------------------------------|------|-----|------|------|------|------|------|-----|
| | | | | 药物: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 第 2 天 | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| | 下午 | 1 | 1 | 5.8 | 6.8 | 7.0 | 8.5 | 7.8 | 7.3 | 6.4 | 8.5 | |
| | | 1 | 2 | 5.4 | 7.9 | 8.0 | 6.4 | 8.4 | 7.1 | 6.4 | 7.2 | |
| | | 2 | 1 | 9.1 | 10.8 | 6.9 | 12.2 | 9.9 | 10.4 | 10.6 | 10.5 | |
| | | 2 | 2 | 5.3 | 10.4 | 8.2 | 8.1 | 10.9 | 9.8 | 8.4 | 14.6 | |

说明: 第 1 行的省略号代表省略了第 1 天下午的 4 行结果, 第 2 行的省略号代表省略了第 2 天上午的 4 行结果。

16. 重复测量设计

(1) “时间”为重复测量因素。

【例 1-19】 具有一个重复测量因素的单因素设计实例。某外科医生在每个病人的手术过程中的 3 个不同阶段测得病人门静脉压力资料见表 1-16, 该资料所代表的试验属于具有一个重复测量的单因素设计。

【例 1-20】 具有一个重复测量因素的两因素设计实例。资料见表 1-19, 每行中的 3 个数据测自同一只大鼠。该资料属于具有一个重复测量的两因素设计。值得注意的是, 若第 1 个时间点是用药前, 则将该行数据作为“协变量”的值更为妥当。

表 1-19 大鼠分别接受某药物 3 种不同剂量后在 3 个不同时间点上心率的测定结果

| 药物剂量 | 鼠号 | 心率(次/分) | | | |
|------|-----|-------------|-----|-----|-----|
| | | 给药后时间(min): | 0 | 2 | 4 |
| 低剂量组 | 1 | | 205 | 246 | 215 |
| | 2 | | 313 | 338 | 340 |
| | ... | | ... | ... | ... |
| 中剂量组 | 7 | | 277 | 268 | 275 |
| | 8 | | 339 | 296 | 284 |
| | ... | | ... | ... | ... |
| 高剂量组 | 13 | | 221 | 193 | 205 |
| | 14 | | 278 | 245 | 253 |
| | ... | | ... | ... | ... |

【例 1-21】 具有一个重复测量因素的三因素设计实例。某研究者为了探讨两种不同处理(因素 A)和不同剂量(因素 B)作用下不同时间点(因素 T)上 NFS-60 细胞凋亡情况, 将全部样品随机均分为 6 个组, 每组含 3 个样品。对各组中的每个样品在处理后的 5 个不同时间点上观测 200 个细胞, 数出凋亡细胞个数。因素 A 分为 A_1 (单纯照射)、 A_2 (照射 + 过钒酸钠(Perv)), 各处理组的照射剂量(因素 B, Gy)又分为 0、3、10。设计格式和资料见表 1-20。

(2) “处理”与“部位”组合起来构成复合型重复测量因素。

【例 1-22】 具有两个重复测量因素的两因素设计实例。用贲门癌患者的标本制成液体, 在 3 种不同处理条件下观测鸡胚背根神经节与鸡胚交感神经节中长出突起神经节的比例。现有贲门癌患者 10 例, 将每人的标本均分成 3 份, 分别给予 3 种不同的处理(因素 A), 即 A_1 (加入 100ng/ml 神经生长因子)、 A_2 (加入 200ng/ml 神经生长因子)和 A_3 (单用贲门癌培养液); 并对每种处理后的标本中的两种类型神经节(因素 B), 即 B_1 (背根神经节)与 B_2 (交感神经节), 观测长出突起神经节的比例(Y)。设计格式和资料见表 1-21。

表 1-20 不同处理和不同剂量作用下不同时间点上 NFS-60 细胞凋亡情况

| 处理组(因素 A) | 照射剂量(Gy) | 样品编号 | 200 个细胞中凋亡细胞数 | | | | |
|----------------|----------|------|---------------|-----------|-----------|-----------|-----------|
| | | | 时间 T(h): | $T_1(12)$ | $T_2(24)$ | $T_3(36)$ | $T_4(48)$ |
| A_1 (照射) | 0 | 1 | 3 | 7 | 7 | 10 | 17 |
| | | 2 | 5 | 3 | 5 | 7 | 18 |
| | | 3 | 5 | 6 | 8 | 14 | 14 |
| | 3 | 1 | 4 | 7 | 10 | 20 | 33 |
| | | 2 | 3 | 4 | 14 | 22 | 37 |
| | | 3 | 3 | 5 | 9 | 19 | 31 |
| | 10 | 1 | 6 | 18 | 38 | 83 | 101 |
| | | 2 | 4 | 16 | 42 | 91 | 120 |
| | | 3 | 5 | 18 | 45 | 92 | 117 |
| A_2 (加 Perv) | 0 | 1 | 4 | 6 | 4 | 12 | 14 |
| | | 2 | 2 | 4 | 6 | 14 | 17 |
| | | 3 | 3 | 7 | 3 | 11 | 19 |
| | 3 | 1 | 5 | 7 | 7 | 14 | 17 |
| | | 2 | 3 | 5 | 9 | 12 | 19 |
| | | 3 | 7 | 6 | 6 | 15 | 16 |
| | 10 | 1 | 5 | 10 | 16 | 33 | 49 |
| | | 2 | 4 | 12 | 12 | 37 | 44 |
| | | 3 | 3 | 9 | 14 | 36 | 52 |

表 1-21 贲门癌患者的标本经 3 种处理后两种神经节中长出突起的神神经节的比例

| 病 例 号 | Y(长出突起的神神经节的比例) | | | | | |
|-------|-----------------|-------|------------|-------|------------|-------|
| | $A_1(B_1)$ | B_2 | $A_2(B_1)$ | B_2 | $A_3(B_1)$ | B_2 |
| 1 | 0.50 | 0.43 | 0.50 | 0.43 | 0.80 | 0.50 |
| 2 | 0.63 | 0.38 | 0.55 | 0.50 | 0.71 | 0.63 |
| 3 | 0.50 | 0.50 | 0.54 | 0.50 | 0.83 | 0.67 |
| 4 | 0.43 | 0.43 | 0.50 | 0.38 | 0.70 | 0.57 |
| 5 | 0.50 | 0.40 | 0.50 | 0.57 | 0.70 | 0.71 |
| 6 | 0.44 | 0.38 | 0.63 | 0.50 | 0.77 | 0.63 |
| 7 | 0.43 | 0.50 | 0.50 | 0.38 | 0.69 | 0.50 |
| 8 | 0.63 | 0.44 | 0.40 | 0.38 | 0.70 | 0.50 |
| 9 | 0.44 | 0.50 | 0.50 | 0.44 | 0.67 | 0.50 |
| 10 | 0.44 | 0.43 | 0.45 | 0.40 | 0.60 | 0.75 |

(3) “处理”与“时间”组合起来构成复合型重复测量因素。

【例 1-23】 具有两个重复测量的三因素设计实例。某研究者为了研究阿托品累计用量对有机磷药物中毒后的治疗效果，将 10 只大鼠随机均分成两组，第 1 组 5 只接受有机磷农药使之中毒，第 2 组 5 只未接受有机磷农药。两组大鼠均给予 1 次乙酰胆碱，同时，4 次给予阿托品治疗，每次均为 2mg/kg，在每次给予阿托品治疗后的 2min、5min、10min 观测膈肌对电刺激产生的反应，用“面积”表示，面积越大，治疗效果越好。设计格式和资料见表 1-22。

表 1-22 连续 4 次给阿托品治疗不同时间点上膈肌对电刺激产生反应的观测结果

| 中毒与否(A) | 大鼠编号 | 面积(mm ²) | | | | | | |
|---------|------|----------------------|-------------------|------|------|-------------------|------|------|
| | | B与C: | B ₁ (2 | 5 | 10) | B ₂ (2 | 5 | 10) |
| 是 | 1 | | 48.0 | 48.6 | 51.0 | 52.8 | 50.6 | 49.7 |
| | 2 | | 45.1 | 47.7 | 48.7 | 47.1 | 45.7 | 42.3 |
| | 3 | | 45.1 | 47.7 | 48.7 | 46.0 | 46.0 | 43.7 |
| | 4 | | 41.4 | 41.1 | 42.0 | 39.1 | 38.0 | 35.4 |
| | 5 | | 57.7 | 60.6 | 61.4 | 58.6 | 56.8 | 52.0 |
| 否 | 1 | | 64.8 | 66.3 | 67.1 | 66.0 | 65.1 | 62.0 |
| | 2 | | 59.7 | 59.7 | 59.1 | 59.4 | 59.1 | 59.4 |
| | 3 | | 53.1 | 54.6 | 54.6 | 52.3 | 51.7 | 49.1 |
| | 4 | | 61.1 | 62.9 | 60.6 | 56.8 | 52.8 | 47.1 |
| | 5 | | 84.6 | 85.1 | 84.3 | 80.3 | 77.1 | 72.3 |

注:表中右边还有 6 列数据未列出,其表头为“B₃(2 5 10)min”和“B₄(2 5 10)min”。

【分析与解答】在前面介绍的 5 个重复测量的实例中,例 1-6 和例 1-21 中没有试验分组因素。这 5 个实例的共同特点是:至少有 1 个重复测量因素,还有 1 个单个体型区组因素(测自每 1 个个体的多个数据形成区组因素的 1 个水平,它们必须按先后顺序写在同一行上)。

17. 析因设计

【例 1-24】 2×2 析因设计实例。将 40 只雌性小鼠完全随机地均分为 4 组,每组 10 只。第 1 组为空白对照组,即不给予任何药物治疗;第 2 组为仅给予一定剂量的激素治疗;第 3 组为仅给予一定剂量的某种中药治疗;第 4 组为既给予激素又给予中药(剂量同前面两组)治疗。观测指标为“排卵数”。设计格式见表 1-23。

表 1-23 中药用否与激素用否对雌性小鼠排卵数的影响情况

| 中药用否 | 排卵数(个, $\bar{x} \pm s$) | | |
|------|--------------------------|---------------|--------------|
| | 激素用否: | 用 | 否 |
| 用 | | 38.40 ± 26.91 | 10.10 ± 5.32 |
| 否 | | 37.90 ± 27.18 | 8.90 ± 5.20 |

【例 1-25】 2×3×4 析因设计实例。某试验同时涉及 A、B、C 3 个地位平等的试验因素, A 分为 2 个水平, B 分为 3 个水平, C 分为 4 个水平,观测指标为 OD 值,受试对象为样品,不同试验条件下均独立地重复做了 2 个样品,设计格式和资料见表 1-24,此为三因素析因设计。

表 1-24 3 个试验因素作用下 OD 值的测定结果

| 因素 A 与 B | | OD 值 | | | | | | | |
|----------------|----------------|-------|------|----------------|------|----------------|------|----------------|------|
| | | 因素 C: | | C ₁ | | C ₂ | | C ₃ | |
| A ₁ | B ₁ | 0.39 | 0.41 | 0.37 | 0.39 | 0.42 | 0.38 | 0.44 | 0.41 |
| | B ₂ | 0.37 | 0.36 | 0.43 | 0.45 | 0.41 | 0.37 | 0.42 | 0.39 |
| | B ₃ | 0.45 | 0.43 | 0.46 | 0.39 | 0.38 | 0.35 | 0.39 | 0.37 |
| A ₂ | B ₁ | 0.36 | 0.41 | 0.45 | 0.36 | 0.41 | 0.45 | 0.41 | 0.46 |
| | B ₂ | 0.42 | 0.37 | 0.38 | 0.41 | 0.38 | 0.36 | 0.43 | 0.38 |
| | B ₃ | 0.37 | 0.43 | 0.36 | 0.39 | 0.43 | 0.42 | 0.35 | 0.37 |

特别说明,标准的析因设计具有如下7个特点:

第一,试验因素的个数 ≥ 2 。

第二,每个试验因素的水平数 ≥ 2 且可以不等。

第三,不同的试验条件数(或组合数)等于全部试验因素的水平数之乘积。

第四,各试验条件下至少要做两次独立重复试验。

第五,全部受试对象被完全随机地分配进入任何一个试验条件组中去,各小组中的受试对象个数可以不等,但最好相等。

第六,试验时,任何一次试验都将涉及任何一个试验因素的某个水平,即全部试验因素同时施加。

第七,进行统计分析时,假定全部试验因素对观测结果的影响是地位平等的,即没有主次之分。

上述特点决定了析因设计分别具有一个明显的优点和缺点。优点:可以分析每个试验因素的主效应和试验因素之间各级交互作用的效应大小;缺点:总试验次数太多,试验费用增大且费时费力。

18. 含区组因素的析因设计

【例 1-26】 有人研究猪食用不同饲料对体重增加量的影响。A(大豆粉+不同含量的蛋白质): A_1 (加14%蛋白质)、 A_2 (加12%蛋白质);B(玉米中己氨酸的含量): B_1 (含0.6%己氨酸)、 B_2 (缺乏己氨酸)。共有4种不同的饲料配方。用24头猪作为受试对象,按猪的初始体重由轻到重排序,并形成6个随机区组,每组中的4只猪体重接近,随机地分入4个饲料组中去,在喂养期间,尽可能使每只猪每餐都能吃饱。设计格式和资料见表1-25,该资料所取自的试验设计类型为含区组因素的析因设计。

表 1-25 两种饲料的不同配方对猪平均日增重量的影响结果

| 随机区组编号 | 平均日增重量(×500g) | | | |
|--------|---------------|--------|------------|--------|
| | $A_1(B_1)$ | $B_2)$ | $A_2(B_1)$ | $B_2)$ |
| 1 | 1.38 | 1.52 | 1.22 | 1.11 |
| ... | ... | ... | ... | ... |
| 6 | 1.47 | 1.53 | 1.16 | 0.99 |

说明:因篇幅所限,其他设计类型(如分式析因设计、正交设计和均匀设计等)从略。

1.5 常见不值得分析的数据种类

1.5.1 人为编造的数据是不值得分析的

当我们试图运用统计分析方法处理数据时,首先要问的问题不是选择什么统计分析方法处理此数据合适,而是此资料是否值得分析。

【例 1-27】 对于表1-26所示资料,有人曾采用3种分析策略(参见下面张飞给出的处理方法)得出了两个相互矛盾的结论,这一命题被称为“Symposon 悖论”。事实上,对于该命题,具有聪明脑瓜的人(简称张飞)和具有正常心态、思维的人(简称百姓)会给出截然不同的解读方法和答案。

表 1-26 同时按年龄和性别分层后吸烟与否与是否患肺癌的调查结果

| 年龄(岁) | 吸烟与否 | 例 数 | | | | | |
|-------|------|-----|----|-----|----|---|----|
| | | 男性: | | 女性: | | | |
| | | 患 | 未患 | 患 | 未患 | 患 | 未患 |
| ≤40 | 吸烟 | 5 | 5 | 40 | 50 | | |
| | 不吸烟 | 60 | 55 | 5 | 5 | | |
| >40 | 吸烟 | 30 | 10 | 5 | 55 | | |
| | 不吸烟 | 30 | 5 | 5 | 35 | | |

【分析与解答】 张飞这样处理: 分别按 3 种策略分析此资料, 即策略 1: 只考察吸烟与不吸烟患肺癌的概率之间的差别; 策略 2: 在按性别分层的基础上, 再采用策略 1 分析; 策略 3: 先按性别分层, 后按年龄分层, 再采用策略 1 分析。得到两种自相矛盾的结论, 结论 1: 吸烟有利于健康(策略 1 与策略 3); 结论 2: 吸烟有害于健康(策略 2)。

百姓这样处理: 首先提出一个问题, 此资料是否值得分析? 经过审读, 发现它不值得分析, 因为它是造假的产物! 何以见得? 此资料中有两个疑点, 由此可推测其纯属人为编造的“调查资料”。疑点之一, 调查了 ≤40 岁男性吸烟者 10 人, 居然有一半人患了肺癌; 调查了 ≤40 岁女性不吸烟者 10 人, 居然也有一半人患了肺癌。这样的调查结果只有在“肺癌病房”中才有可能获得, 因为肺癌病房中有 5 位患者, 另 5 位是照顾病人的家属或护工。疑点之二, 随机调查 400 位对象, 再按“年龄”、“性别”、“吸烟与否”和“是否患肺癌”将其划分成 16 个小组, 每个小组中的人数的个位数不是 0 就是 5(由基本常识可知, 一个正整数的个位数有 0、1、2、…、9 共 10 种可能性), 出现这种“奇特现象”的概率是极低的, 大约为 $P = \left(\frac{2}{10}\right)^{16} = (0.2)^{16} = 6.5536 \times 10^{-12}$ 。以如此低的概率出现的“奇特现象”, 在实际调查资料中几乎是不可能出现的!

另外, 即便表 1-26 中未出现上述两个疑点, 而且调查资料是在科学完善的调查设计方案指导下获得的, 张飞的 3 个分析策略都是不正确的, 它们均属于“用单因素分析方法取代多因素分析方法”的错误。正确的分析方法是直接分析 3 个原因(“年龄”、“性别”、“吸烟与否”)对“是否患肺癌”这个“二值结果变量”的影响是否具有统计学意义。具体地说, 可以选用对数线性模型分析, 也可以选用多重 logistic 回归模型分析。若分析目的仅仅是希望考察吸烟与不吸烟得肺癌的概率之间的差别是否具有统计学意义, 还可以选用 CMH 校正的 χ^2 检验。此法是把性别和年龄视为两个分层因素, 对每一层(研究层内吸烟与不吸烟患肺癌概率之间的差别)做一些基本计算, 然后再进行合并计算, 最后得出汇总分析的结果和结论(即在消除分层因素影响的前提下, 分析吸烟与不吸烟患肺癌概率之间的差别是否具有统计学意义)。

事实上, 基于表 1-26 的调查设计, 即使在样本含量、样本的代表性和资料的准确性等方面都做得很好, 统计分析方法选用得也正确, 其结论也不一定就很科学。因为导致一个人是否患肺癌, 还有很多可能的影响因素未加以考察, 如是否有肺癌的家族史、生活环境(如空气污染情况、饮水的质量、饮食构成、饮食习惯、生活习惯)、锻炼身体情况、职业、劳动强度、家庭经济状况、家庭生活是否幸福美满、人际关系是否良好、心态是否平稳、睡眠质量和每晚平均深睡眠时间长短等。这一切也都与一个人“是否患肺癌”有很大关系, 它们被统称为“重要的非试验因素”。在任何一项调查或试验研究中, 很关键的一个问题是: 是否“找准找全”了影响研究目的实现的主要评价指标的“重要非试验因素”; 不仅如此, 还必须在尽可能大的样本含量且对总体有极好代表性的受试对象身上准确地获得试验因素、全部重要非试验因素和主要评价指标的数值。所有这一切, 都与“数据是否值得分析”有密切关联。