



Practical Nonparametric Statistics

# 实用非参数统计

(第3版)

[美] W.J.Conover 著

崔恒建 译

**TURING**

图灵数学·统计学丛书

Practical Nonparametric Statistics

# 实用非参数统计

(第3版)

[美] W. J. Conover 著

崔恒建 译



人民邮电出版社

POSTS & TELECOM PRESS

## 图书在版编目 (CIP) 数据

实用非参数统计: 第3版 / (美) 康诺华著; 崔恒建译.

—北京: 人民邮电出版社, 2006.4

(图灵数学·统计学丛书)

ISBN 7-115-14616-0

I. 实... II. ①康...②崔... III. 非参数统计 IV. 0212.7

中国版本图书馆 CIP 数据核字 (2006) 第 022972 号

### 内 容 提 要

非参数统计为有效地分析试验设计及其实际问题中所获得的数据提供了丰富而有说服力的统计工具, 而本书则从问题背景与动机、方法引进、理论基础、计算机实现、应用实例、文献综述等诸多方面介绍了非参数统计方法, 其内容包括: 基于二项分布的检验、列联表、秩检验、Kolmogorov-Smirnov 型统计量等. 本书内容丰富、思路清晰、层次分明. 在强调实用性的同时, 突出了应用方法与理论相结合. 书中的正文和习题中都提供了大量的实际案例, 书中最后有许多统计用表以及奇数号习题解答和术语索引.

本书作为非参数统计的基础教材, 适用于统计学专业的高年级本科生和研究生课程的教学, 也可提供给从事统计学应用与研究、数据的分析处理及其相关领域的专业人员阅读与参考.

图灵数学·统计学丛书

### 实用非参数统计 (第3版)

- 
- ◆ 著 [美] W. J. Conover
  - 译 崔恒建
  - 责任编辑 王丽萍 蔡红艳
  - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号  
邮编 100061 电子函件 315@ptpress.com.cn  
网址 <http://www.ptpress.com.cn>  
北京铭成印刷有限公司印刷  
新华书店总店北京发行所经销
  - ◆ 开本: 700×1000 1/16  
印张: 26.75  
字数: 540 千字 2006 年 4 月第 1 版  
印数: 1-4 000 册 2006 年 4 月北京第 1 次印刷  
著作权合同登记号 图字: 01-2005-5222 号

ISBN 7-115-14616-0/TP · 5307

定价: 59.00 元

读者服务热线: (010) 88593802 印装质量热线: (010) 67129223

# 版 权 声 明

Original edition, entitled *Practical Nonparametric Statistics* by W. J. Conover, ISBN 0-471-16068-7, published by Wiley Publishing, Inc.

Copyright © 1999 John Wiley & Sons, Inc.

All rights reserved. This translation published under license.

Translation edition published by POSTS & TELECOM PRESS Copyright 2006.

本书简体中文版由 Wiley Publishing, Inc. 授权人民邮电出版社独家出版。

版权所有，侵权必究。

## 某些非参数检验的位置快速查找图

含有度量的类型 (见 2.1 节) 查找图中页码为英文原书页码, 与书中页边标注的页码一致。

所获样本类型	所含的假设检验	名义度量 (根据范畴分离观测)	有序度量* (观测从小到大顺序排列)	区间度量* (观测的数值含有意义)
一样本 $X_1, \dots, X_n$	均值 (中位数) (均值的置信区间) 非随机性 拟合优度 $[F(x)]$ 的置信界	二项检验 (p. 124) $P$ 的置信区间 (p. 130) $\chi^2$ 检验 (p. 242) $\chi^2$ 检验 (p. 240)	分位数检验 (p. 136) $x_p$ 的置信区间 (p. 143) Cox 和 Stuart 检验 (p. 170) Sperman $\rho$ (p. 323) Kolmogorov 检验 (pp. 430, 435) $F(x)$ 的置信界 (p. 438)	Wilcoxon 检验 (p. 352) 均值的置信区间 (p. 306)  Lilliefors 检验 (pp. 443, 448) Shapiro-Wilk (p. 450)
成对观测或两配对样本 $(X_1, Y_1), \dots, (X_n, Y_n)$	均值 (中位数) (两均值差的置信区间)  回归斜率 独立性  相依性度量	McNemar 检验 (p. 166)  $P$ 的置信区间 (p. 130)  $\chi^2$ 检验 (p. 204) Fisher 精确检验 (pp. 188, 209) Mantel-Haenszel 检验 (p. 192) 见 4.4 节	符号检验 (p. 157)  $x_p$ 的置信区间 (p. 143)  单调回归 (p. 344) 符号检验 (p. 172) Spearman $\rho$ (p. 316) Kendall $\tau$ (p. 321) 见 5.4 节	Wilcoxon 检验 (p. 352) Van der Waerden 检验 (p. 400) 随机化检验 (p. 412) 差的置信区间 (p. 360)  斜率检验 (p. 335) 见 5.11 节

多维观测, 或随机完全 区组设计	均值 (中位数)	Cochran 检验 (p. 251)	Friedman (p. 369) 有序备择假设的 Page 检验 (p. 380)	Quade 检验 (p. 373)
两随机样本 $X_1, \dots, X_n$ 和 $Y_1, \dots, Y_m$ (也可见几个随机 样本的检验)	相依性度量 and 独立性检验 均值 (中位数) (均值的置信区间) 方差 同分布	见 4.4 节 $\chi^2$ 检验 (pp. 180, 199) $P_1 - P_2$ 的置信区间 (p. 135)	见 5.4 节 (特别是 p. 327) Mann-Whitney 检验 (p. 272) Mann-Whitney 置信区间 (p. 281) 平方秩检验 (p. 300) Klotz 检验 (p. 401) Smirnov 检验 (p. 456) Cramer's - von Mises 检验 (p. 463)	随机化检验 (p. 409)
几个随机样本	均值 (中位数) 方差 同分布	$\chi^2$ 检验 (pp. 180, 199)	中位数检验 (p. 218) Kruskal-Wallis 检验 (p. 288) Van der Waerden 检验 (p. 397) 有序备择假设的 Jonckheere- Terpstra 检验 (p. 325) 平方秩检验 (p. 303)	
其他类型	均值 (中位数)	多方式列联表 (p. 2.14) 对数线性模型 (p. 259)	扩展的中位数检验 (p. 224) BIBD 的 Durbin 检验 (p. 388) 试验设计的秩变换 (p. 417)	

\*列在“名义度量”之下的方法这里也可以使用。  
+列在“名义度量”和“有序度量”之下的方法这里也可以使用。

## 译者介绍

**崔恒建** 北京师范大学数学科学学院教授，博士生导师，统计与金融数学系主任。1986年获北京师范大学数理统计硕士学位，1993年获中国科学院系统科学研究所数理统计博士学位，1995年北京师范大学博士后出站。现任中国概率统计学会常务理事，生存分析学会理事，美国数学会评论员，世界泛华统计学会会员。

主要从事统计学、数理统计学和数学方面的研究、教学和管理工作。在国内外主要学术刊物上发表论文70余篇。2004年7月作为当地组委会主席首次在中国成功举办“国际稳健统计大会（北京2004）”。曾先后访问过日本东京大学和创价大学、中国香港城市大学和香港大学、中国台湾的“中央研究院”和台湾中山大学、意大利ICTP中心、澳大利亚Latrobe大学、美国伊利诺伊大学（UIUC）和亚利桑那州立大学、新加坡国立大学等，并与上述学校进行过多次合作研究。其业绩和成果入选多部国内外名人录和检索系统。

# 译者序

随着时代的发展,许多参数统计方法已不能满足实际应用的需要,甚至很多时候人们根本就不知道应用何种参数统计方法,这时非参数统计方法就显得尤为重要.由于它对模型的要求甚少,并且具有稳健性和效率高的特点,因此它将越来越成为实用工作者们所喜欢的统计方法.

原书著者威廉姆 J. 康诺华从 1981 年至今是美国得克萨斯理工大学资深 (Horn 讲席) 统计学教授,曾任爱荷华州立大学 (Iowa State University) 和加利福尼亚大学戴维斯分校 (University of California at Davis) 客座教授. 曾任《美国统计学杂志》(JASA),《美国统计学家》(*The American Statistician*),《技术计量学》(*Technometrics*),《生物统计学》(*Biometrics*),《统计年鉴》(*The Annals of Statistics*),《统计通讯》(*Communications in Statistics*),《统计计算和模拟杂志》(*Journal of Statistical Computation and Simulation*),《加拿大统计杂志》(*The Canadian Journal of Statistics*),《决策科学》(*Decision Sciences*) 等杂志的编委. 主持过十几次重大的统计学术年会,并多次在这种会议上担任主席. 作为特邀代表出席过近 20 次重大统计学术年会及其他主要学术会议并在会议上作特邀报告. 仅在世界主要顶级统计杂志上就发表了 40 多篇学术论文,另外在其他学术会议论文集及其他文献上还有多篇论文发表. 作者不仅是非参数统计领域的权威、著名统计学家,同时也是一名优秀的统计教育学家,在不少国家的多所大学作过讲学. 他用此书对研究生进行过多次教学,并十分讲究教学方法. 我对作者严谨求实的工作作风、优秀的教学方法及取得的成就深表敬慕.

原著可以说是作者多年来从事非参数统计研究和教学的经验总结 (本译本译自原著的第三版). 其特点是内容丰富、思路清晰、层次分明,侧重内在的统计思想的阐述. 尤其是涉及到的假设检验部分,采用了问题背景与动机、方法引进、假定条件、假设、检验统计量、零分布、应用实例、计算机实现、理论基础、文献综述这样的顺序编写,显得条理清晰,实用易懂. 本书在强调实用性的同时,又突出了应用方法与理论的结合,同时也介绍了如何进行统计计算. 书中的正文和习题中都提供了大量的实际案例,书中的最后有许多统计用表以及奇数号习题解答和术语索引. 这是一本非参数统计方面难得的好书.

中文译稿是译者 2006 年初完成的. 在整个翻译过程中,得到我的学生杨艳霞、王川、冯峥晖、陈霞、王婷、钟平寿、郭艳梅和张娟的大力帮助,他们分别帮助我进行录入和部分初译,在此我对他们深表谢意. 对人民邮电出版社图灵公司和王丽萍编辑为本书中译本的出版所付出的辛勤努力表示衷心感谢. 同时也十分感谢我的



## 2 译者序

---

妻子陈秋华对我的支持和鼓励.

由于译者水平有限,在翻译过程中一定会存在不少缺点和不当之处,恳请同行专家、老师及读者们给予批评指正.

译者 2006年2月8日

于北京师范大学数学科学学院统计与金融数学系

# 序

30 多年前当我决定写一本关于非参数统计的书时，我希望它能作为一学期的非参数（或者不涉及分布的）统计的教科书；同时我也希望它能够成为查询非参数统计中最有用方法的快捷参考书，这样，科研工作者们能够很容易地找到他们所需要的内容。由此，该书的第一版于 1971 年由 John Wiley & Sons 出版，当时的编辑是我以前大学的室友 Joe Frank。

我和 Joe 都没有预料到我们这本书会取得如此巨大的成功。我曾向另外 6 家出版社投递了我的手稿，可他们对这本书没有任何兴趣。但是 Joe Frank 和 John Wiley & Sons 给了我这个机会，为此，我从内心里感谢他们。

这本书经历了前两版和几十次重印，深获好评，人们称它理论严密、讲解清晰，是一本值得信赖的教科书。与此同时，世界各地的科研人员把它作为一本便捷的参考书，用以了解如何以及何时使用最常用的非参数方法，并从中找到清晰简明的说明。

现在该是第三版问世的时候了，第二版已经出版了将近 20 年，需要进行一些更新。在第三版中我加入了一些经受了时间的检验并且现在在实践中常被使用的方法，如  $2 \times 2$  列联表的 Fisher 精确检验，联合几种列联表的 Mantel-Haenszel 检验，生存曲线的 Kaplan-Meier 估计，Jonckheere-Terpstra 检验和有序备择的 Page 检验，还有一些有关自助法（bootstrap）的讨论。第三版还加进了一些新的习题和思考题，补充了新的例题并更新了参考书目。在第二版的基础上去掉了一些被认为不再实用的方法，主要是因为这些方法只能用于容量相同的独立样本，而大多数数据集并不具有这个性质。应许多读者的要求，以前出现在第 1 章和第 2 章中的证明过程也缩减了许多。

经过深思熟虑，我在第三版中加入了一些计算机软件包的内容。一方面，读者将发现计算机说明非常有用；另一方面，计算机软件包的变化比书要快得多，计算机的提示也很容易过时。由此使用了一个折中的办法：只加了一些应保持几年不变的有用的计算机提示，而有关个人计算机软件包的更全面深入的信息还应在期刊和杂志的文章中查阅。

书中提及了几个含有大量非参数程序的计算机软件包，还有很多类似的软件包这里没有列出。有兴趣的读者可以上因特网获得这些专为非参数统计分析的软件包设计的更多信息，很多情况下，可以从网站下载一些软件包的演示版。为了避免很快过时，下面列出了一些写这本书时较常用的统计软件包以及相关的网址。

*Minitab*——<http://www.minitab.com>

*PASS* 或 *NCSS*——<http://www.ncss.com>

*Resampling Stats*——<http://www.statistics.com>

*SAS*——<http://www.sas.com/rnd>

*SPSS*——<http://www.spss.com>

*STATA*——<http://www.stata.com>

*STATISTICA*——<http://www.statsoft.com>

*StatMost*——<http://www.dataxiom.com>

*StatXact*——<http://www.cytel.com>

*SYSTAT*——<http://www.spss.com/software/science/systat>

*S-Plus*——<http://www.mathsoft.com/splus>

*STATA* 的网站含有与上述及其他软件提供者的链接。

和前几版一样，阅读和理解这本书需要有较好的大学代数的知识，以及要求不高的数学能力。编写第 1 章和第 2 章的目的就是使学生的知识水平提高，以便能够理解书中其余章节的理论和方法。

该书已经成功地用作本科生和研究生的教材。在本科生的教学中，大多数的教师发现“思考题”与“理论”部分对学生来说非常具有挑战性，所以在课程中他们往往略去了这些材料。我自己已多次给研究生讲过这本书，仅有一次略去了第 1 章和第 2 章，结果十分糟糕。另外有些教师告诉我，他们略去了第 1 章和第 2 章，认为不会产生问题，但是他们的学生却时常告诉我说，他们需要自学前面的章节才能理解后面的内容。

当我讲授这门课时，讲完前两章安排了一次测试，讲完第 3、4 章安排了第二次测试，第三次测试是放在第 5 章后，最后一次则是针对整本书的，包括第 6 章在内。在很多情况下，我顺利地将本书中的复习题作为学生课外测评题的基础。讲授这门课时，我没有使用电脑，但是与我交流过的教师中有一半在不同程度地使用电脑授课。

在这本书的编写过程中，有很多人的工作值得我感谢，在这里我不能一一列举。首先，我想到的是给我支持和鼓励的妻子 Susan，以及这些年与我一起工作的许多 Wiley 的编辑，特别是 Joe Frank, Andy Ford 和 Brad Wiley II，还有很多在撰写这版和前几版时给我建议的人，有 Barney Bissinger, Ben Duran, Ron Iman, Mark Johnson, Hossien Mansouri，还有很多读者、老师和从业人员。在此对他们表示我诚挚的谢意。还要感谢上苍指引着我的人生，正是它常常反对我的意愿，让我经历种种磨练，我才走到现在。

W. J. Conover

# 前言

在词典中“科学”(science)的一个定义是“由观测、试验、归纳得到的事实”。今天在追求科学的过程中，全社会投入了大量的时间、资金和资源。这种追求很容易使人沮丧，因为正如许多人所知道的，观测过程、实验和结论不总是只依赖于“事实”，两个人做同一个试验，用同一套观测方法，可能得到两个不同的结论。

例如，把一只老鼠放入一个有两扇门的围栏中，并且把两个门都关上。一扇门涂成红色，一扇门涂成蓝色。然后给这只老鼠播放青年人喜欢的流行歌曲20分钟，在这之后，打开两扇门，老鼠逃出了围栏。试验者记下了这只老鼠选择逃出的门的颜色。这个试验被重复了10次，每次试验使用的是不同的老鼠。

做完10次试验后，试验者发现老鼠从红色的门逃出的次数有7次，然后就得出结论：老鼠在这种状况下更喜欢红色。然而，一个同事无意中听到了这个结论，就打趣道：“如果我掷一枚硬币10次，有7次正面朝上，并且在抛掷之前，我哼唱‘Yankee Doodle’，那么你的结论是我的口哨声使得硬币正面朝上吗？”从老鼠选门和掷硬币试验的类比中，这位试验者发现了自己的错误，并且体会到试验的结论只不过是偶然性的结果。

后来，他又做了第二个试验。他向10只老鼠的血液中注射了某种药物，五分钟以后，发现有7只老鼠死亡了，而另外3只老鼠依然很健康。他想起了自己先前做的那个试验，尽管有7只老鼠死掉了，他还是断定这样的—个结果具有很大的偶然性，认为这也并不代表这种药物具有危险性。

他的同事忍不住了，说道：“在你的第一次试验中，如果不用音乐，每只老鼠选择红门的概率都是0.5，所以可以与掷硬币试验类比。在这个试验中，如果这种药物没有作用的话，老鼠在5分钟之内死亡的概率确实很小。既然你的10次试验中，这种概率很小的事情发生了7次，那么我想这种药物注射能够导致死亡的结论是比较可靠的。”

科学研究即是如此。不久，大多数科学家意识到，表达一个试验结果最理想的方式（像先前的试验一样）类似如下的说法：“在没有经过处理的情况下，与我得到的结论一样的试验次数在1000次中至多有3次，所以我断言：我的处理是有效的。”在这种情况下，每一个了解这个试验的科学家就知道这个试验中的主观结论有多少。

科学中有一个被称作“统计”的领域，就是对科学家所得出的结论中的主观性提供一种度量方法，从而可以从“观点”中分离出“科学”的成分。这可以通过对试验建立理论“模型”来实现，例如被称为“掷硬币”的模型，它就是为了上面讨论过的第一个试验而建立的。对于各种可能发生的结果，在假设只有随机因素影响

试验结果的情况下，既没有音乐也没有药物注射，为了定义什么是“偶然”（可能性），概率被运用到了这个模型中。于是，试验者们在判断一个试验结果是否是由所进行的处理所导致的，或者在不经过处理的情况下，仅由随机因素是否能够得到相同的结果时，就有了一个客观的标准。

尽管有时对某些试验很难找到一个合适的理论模型，但是真正的难点在于定义了模型后，怎样寻找与之相关的概率。很多合理的模型，建立后找不到概率上的解决办法。于是，统计学家们为了找到所期望的概率，经常对模型稍作调整，并且希望调整后的模型仍然能够很好地描述实际情况。对于这些“近似问题”，能够找到确切的解决办法。因此，在统计学中的这一部分有时被称为“参数统计”，其中包含一些我们熟知的检验，例如  $t$  检验， $F$  检验等等。

在 20 世纪 30 年代后期，寻找概率这一问题的另一个方法开始吸引人们的注意，这个方法包括对模型做很小的改动（如果必要的话），用一些并不复杂的方法去找出所期望的概率，或者至少找到一个很好的逼近。这样对于确切的问题我们就找到了近似解答，而在参数统计中是利用那些近似问题得到确切解答。这就是统计方法中新的一部分——非参数统计。

非参数方法已经成为实验室中需要进行统计分析的应用科学家们的基本工具。当做出错误结论的代价很高时，应用科学家们十分关注他们所用的统计方法，且希望那些方法并不是基于那些看起来无效或者无法证实的假设。

非参数统计方法除了涉及的模型较简单这个优点外，计算量也较少，因此比其他的统计方法用起来更简单和快捷。非参数统计方法的第三个优点是不需要用到高等代数的知识就可以演绎得很严密。理解这个统计方法所蕴含的理论的人，一般很少在不适当的时候应用它，而且如果这个模型还没有其他统计学者研究过，也更容易研究出自己的统计方法。

还有一些需要用到较多高等数学的非参数统计可以不加推导地讲述，但是我们在合适的地方都标注了参考文献，可以从中找到证明。

第四个优点，也是最重要的一个，就是在参数模型的假设不成立的情况下，它比参数方法更有效。而有效之处在于试验者有可能证明他在试验中所需要证明的结论。当参数方法不太适用的时候，非参数方法比参数方法能更有效地利用数据。

这就涉及为什么本书没有包含许多基于游程（run）的非参数检验。游程是一个相似观测的序列。例如，如果每天的 DJIA（道·琼斯工业指数）都被记录下来，上升游程（run up）是指关于 DJIA 一串连续的测量，它的每一个观测都比前一天的要高。可以类似地定义下降游程（run down）。总的游程数指这两种游程的总数目，可以作为当天至下一天观测的独立性的度量。

游程检验可以用来检验观测的独立性，检验两种样本是否来自同一个总体（通过将两种样本从小到大排序，然后数出来自同一个样本观测的游程数），或可用于其他的情形。基于游程数目的检验比较容易实现（游程数比较容易计算），并且游程数

的概率分布常常是精确已知的。然而，游程检验不是很有效，每种情况都可以被更有效的非参数方法所代替。因为基于游程数的检验不是很实用，所以这本书中没有包含这种方法。

上述参数统计与非参数统计的对比仅仅是尝试着给出一个本书主题的大致思想。关于统计学中这两个分支不同之处的详细讨论将在第2章中展开，同时也将对科学试验体系进行更细致的讨论。为了在第2章中给出一些例子和说明，需要一些概率基本部分的预备知识，第1章讲述了这些内容。

从第3章往后，会经常用到第1、2章中引入的概念。后面几章讲述的是不同的非参数方法，是按照所分析的模型的类型来编排的，而不是按照所进行的试验的类型编排。很多人提到非参数方法时，想到的仅仅是第5、6章中介绍的基于秩的检验方法，这是因为第3、4章中给出的用来分析定性（名义）数据的方法必然是非参数的。这些章节后的一些习题揭示了用来分析定性和定量数据的参数和非参数统计方法间一些鲜为人知但很有用的联系。

为方便那些想知道分析中可能用到了哪些技术的试验者，在书的前面给出了一张交叉参考表，表中根据所要解决的问题的类型列出了书中介绍的方法。

对于想了解更多有关非参数方法的读者，可参照本书的参考书目。Savage(1962)所写的参考书目对于想获得每个论题的早期文献的人是很有帮助的。

# 目 录

第 1 章 概率论 .....	1	4.6 相关观测的 Cochran 检验 .....	180
导言 .....	1	4.7 其他分析方法讨论 .....	186
1.1 计数 .....	1	4.8 第 3、4 章复习题 .....	187
1.2 概率 .....	7	第 5 章 秩检验 .....	194
1.3 随机变量 .....	14	导言 .....	194
1.4 随机变量的性质 .....	21	5.1 两个独立样本 .....	195
1.5 连续型随机变量 .....	34	5.2 多个独立样本 .....	207
1.6 第 1 章复习题 .....	43	5.3 等方差检验 .....	216
第 2 章 统计推断 .....	46	5.4 秩相关性度量 .....	225
导言 .....	46	5.5 非参数线性回归方法 .....	240
2.1 总体、样本与统计量 .....	46	5.6 单调回归方法 .....	250
2.2 估计 .....	54	5.7 一样本或配对情形 .....	255
2.3 假设检验 .....	66	5.8 多个相关的样本 .....	268
2.4 假设检验的性质 .....	74	5.9 平衡的不完全区组设计 .....	283
2.5 非参数统计评述 .....	81	5.10 A. R. E. 不低于 1 的检验 .....	290
2.6 第 2 章复习题 .....	85	5.11 Fisher 随机化方法 .....	299
第 3 章 基于二项分布的检验 .....	88	5.12 秩变换的讨论 .....	307
导言 .....	88	5.13 第 5 章复习题 .....	310
3.1 二项检验与 $p$ 的估计 .....	88	第 6 章 Kolmogorov-Smirnov 型	
3.2 分位数检验和 $\chi_p$ 的估计 .....	97	统计量 .....	317
3.3 容忍限 .....	107	导言 .....	317
3.4 符号检验 .....	112	6.1 Kolmogorov 拟合优度检验 .....	317
3.5 符号检验的一些变形 .....	118	6.2 分布族的拟合优度检验 .....	327
第 4 章 列联表 .....	128	6.3 两组独立样本的检验 .....	338
导言 .....	128	6.4 第 1 章至第 6 章复习题 .....	346
4.1 $2 \times 2$ 列联表 .....	128	附表 .....	353
4.2 $r \times c$ 列联表 .....	142	奇数号习题解答 .....	399
4.3 中位数检验 .....	156	索引 .....	402
4.4 相依性度量 .....	163	参考文献 <sup>1</sup>	
4.5 $\chi^2$ 拟合优度检验 .....	172		

1. 限于篇幅, 本书中参考文献请读者登录图灵网站 [www.turingbook.com](http://www.turingbook.com) 查阅下载.

# 第1章 概 率 论

## 导 言

非参数统计方法中一个诱人的特性是：你并不需要成为一个概率论方面的专家就能理解非参数方法所蕴含的理论。只要掌握了一些易学的基本概念，非参数方法的基本理论很容易被理解。本章介绍这些基本概念，所需要的只是耐心、信心和较好的高中代数知识。

这本书是这样安排的，读者可以直接找到所需要的统计方法，然后从头到尾一步一步按着说明操作。然而如果这样的话，不一定能知其然更不能知其所以然，从而通常导致错误处理了数据和得出不合理的结论。在第1、2章，读者应该全面理解所使用的非参数方法，甚至能对其稍作修改，使它能更好地应用于所分析的特殊数据集。

对学习每一节的建议：通读教材，然后演算例子，最后做每一节后的习题和思考题。这样可以为下一节的学习做准备，也可以增强前面所提到的耐心和信心。

## 1.1 计 数

计算概率的过程通常依赖于计数，如通常的计数，“1、2、3”等等。而通常的计数方法在一些复杂的情形下将变得十分冗长乏味，所以本节介绍一些有技巧性的计数方法，用于处理这种复杂的情形。

5

掷一枚硬币时，我们仅考虑2种可能的结果：或者出现正面（H），或者出现反面（T）。如果掷一次，可能出现2种结果：H或T；如果掷两次，会出现4种可能：HH, HT, TH, TT，其中HT表示掷第1次时出现H，掷第2次时出现T。每多掷一次，则可能出现的结果数将是原来的2倍，因为最后一次投掷总有2种可能的结果。所以，如果投掷 $n$ 次硬币，结果会出现 $2^n$ 种可能。

### 试验

为讨论一般化，我们把掷硬币作为试验的一个例子。无论是掷一次、两次还是 $n$ 次，这个过程都被看作是一个试验，这样投掷3次硬币也是一次试验，它是3个仅掷一次硬币的独立试验的组合。我们称较短的试验为基本试验，“试验”由基本试验组合而成。一般来说，一个试验是遵循一套已设计好的规则的过程，试验前并不知道



遵循这些规则的最终结果.

## 模型

很少人会认真地考虑掷硬币作为一个试验其自身所具有的价值. 实际上, 掷硬币的价值在于它为许多不同情况下的不同模型提供了原型. 如果我们考虑的是一个均匀硬币, 即每一面出现的可能性是一样的, 这个试验就与下面的试验类似: 老鼠对于两个门的选择; 顾客对两种商品的选择; 教师判断两种教学法方法哪种更有效; 市场分析师预测周一行情将是看涨还是走跌等等.

如果硬币是不均匀的, 即一面比另一面更容易朝上, 这种模型将适用于更广泛的一类试验. 例如: 给老鼠的血液注射药物来检验这种药物是否是致命的; 找病人来检验一种新的治疗法; 消费者从几种产品里选购一种, 这些产品中只有一种是由某公司制造的等等. 每种情况我们关心的结果只有两个, 如“生”或“死”, “痊愈”或“未痊愈”, “某公司的品牌”或“其他品牌”, 两种结果发生的可能性不必相等.

6

在本章和下一章, 我们将讨论掷硬币、掷骰子、从罐中取筹码、将球放入箱中等几种有价值的模型, 其价值在于它们是大量复杂模型的简单而实用的原型, 这些复杂模型来自电子物理、心理学、社会学、教育学、生物学、经济学、化学等许多不同领域的实验. Feller(1968)很好地研究了这些模型的差异性. 这些内容本章先做部分介绍, 因为需要其他非参数方法的知识, 所以大部分说明要放到后面的几个章节中.

## 事件

这样, 我们把掷硬币作为一个试验, 并把每一次投掷作为一个基本试验, 把一个或多个基本试验或整个试验的结果称为事件(event). 刚描述的掷硬币试验包含  $n$  个基本试验, 其中每个基本试验的结果是事件“H”或“T”. 事件的组合本身也是一个事件, 所以试验中  $2^n$  个可能结果的每一个都可以看作一个事件. 其他事件的例子有“至少一次正面朝上”、“第4次反面朝上”、“正面朝上的次数至少是反面朝上的两倍”等等.

推而广之, 可得如下准则:

**准则 1.1.1** 如果一个试验包含  $n$  个基本试验, 且每一个基本试验有  $k$  种可能的结果, 则整个试验有  $k^n$  个可能的结果.

### 例 1.1.1

如果一个试验由 7 个基本试验组成, 每个基本试验是向 3 个盒子之一投一个球. 第一次投会产生 3 个不同的结果, 前两次投共有  $3^2 = 9$  种可能的结果. 这样推广到由 7 次投掷组成的试验, 它的试验结果就有  $3^7 = 2187$  种不同的可能. ■

现在, 考虑一个盒子装有  $n$  个筹码, 筹码编号从 1 到  $n$ . 从盒子中先取出一个筹码放到桌子上, 可以看到它的编号. 这个筹码是从  $n$  个筹码中任取的一个, 所以有  $n$