

JIYU ZHI DE TONGJI TUIDUAN

〔美〕T. P. 海特曼斯波格/著

杨永信/译

史宁中/审校

# 基于秩的统计推断

于秩的统计推断

□ 东北师范大学出版社



# 基于秩的统计推断

[美] T. P. 海特曼斯波格 著

杨永信 译

史宁中 审校

东北师范大学出版社

1995 · 长春

(吉)新登字 12 号

基于秩的统计推断

JIYU ZHI DE TONGJI TUIDUAN

[美] T. P. 海特曼斯波格著 杨永信译

---

责任编辑: 杨述春      封面设计: 魏国强      责任校对: 张中敏

东北师范大学出版社出版  
(长春市斯大林大街 110 号)

吉林省新华书店发行  
吉林工学院印刷厂制版

(邮政编码: 130024)

吉新月历公司印刷分公司印刷厂印刷

---

开本: 850×1168 毫米 1/32

1995 年 5 月第 1 版

印张: 12

1995 年 5 月第 1 次印刷

字数: 312 千

印数: 0 001—1 000 册

---

ISBN 7-5602-1585-8/O · 82

定价: 10.00 元

“悲哀啊悲哀，欠考虑的推断只能带来痛苦的后果。”

选自《现代抒情歌剧：小鸡的寓言》

作词：阿利西娅·卡平特

作曲：格里格·史密斯

## 中文版序

非参数统计为试验设计中产生的数据分析提供了丰富的统计工具。其中包括单样本位置模型、双样本位置模型、以及像双因子分类 (two way layout) 和多元回归这样的多样本试验设计。还包括点估计、置信区间、假设检验、多重比较、及有关离群值 (outlier) 识别和模型评论的回归诊断。近几年来, 书中的许多方法已有新的发展, 而且可以应用于数据分析。

本书提供了上述统计工具的理论基础, 而且包括检验及置信区间所需要的分布理论。为估价统计方法的优劣, 我们引入了有效性、势、及稳健性理论中的影响曲线及崩溃 (breakdown)。可以证明, 秩方法具有很高的有效性和很高的势。同时, 在具有有界影响函数和正崩溃点这个意义上看, 秩方法也是稳健的。通过阅读本书所提供的材料, 读者将了解有关现代稳健性理论的思想、及更多的有效性及势的传统思想。

希望阅读本书, 使学生和研究人员逐步建立对非参数方法的广泛性和重要性的深刻理解和认识。更希望有人继续从事这方面的研究, 并把非参数方法扩展到新的领域。

托马斯 P. 海特曼斯伯格

1994 年 8 月

于宾夕法尼亚州立大学统计系

## 译者的话

本书是根据美国威利出版公司(John Wiley & Sons, Inc.)1984年出版的托马斯 P. 海特曼斯伯格(Thomas P. Hettmansperger)所著《基于秩的统计推断》(Statistical Inference Based on Ranks) (英文版)一书译成的。

原著作者, 海特曼斯伯格博士, 1939年8月出生于美国印第安纳州; 1961年于印第安纳大学(Indiana U.)获数学学士学位, 1965年于爱荷华大学(U. of Iowa)获数学硕士学位, 1967年于爱荷华大学(U. of Iowa)获统计学博士学位; 现为宾夕法尼亚州立大学(The Pennsylvania State U.)统计系教授(曾任该系系主任); 曾任普林斯顿大学(Princeton U.)、加州大学伯克利分校(U. C. Berkeley)、瑞士波恩大学(U. of Bern)、华盛顿大学(U. of Washington, Seattle)、澳大利亚国立大学(Australian National U., Canberra)等大学的客座教授; 曾任美国统计学会理事; 曾任《美国统计学会杂志》(JASA)、《美国统计学家》(The American Statistician)、《教育统计杂志》(J. of Educational Statistics)等主要统计学刊物的编委; 现为《统计通讯》(Communications in Statistics)和《非参数统计杂志》(J. of Nonparametric Statistics)的编委; 曾主持过十几次重大的统计学学术年会, 并在多次这种会议上担任主席; 曾作为特邀代表出席过近四十次重大的统计学学术年会及其他主要学术会议并在会议上做特邀学术报告; 仅在世界主要权威性统计学杂志上就发表了六十多篇学术论文; 另外在

其他学术会议论文集及其他文献上还有二十多篇论文发表；曾在五个国家的三十多所大学讲学。作者不仅是当今非参数理论的权威，世界级的统计学家，也是一名优秀的教育家。作者曾多次获教学奖。至今为止，作者培养出十七名博士，其中，一些博士分别在美国国家标准局、贝尔实验室、通用软件公司、康州大学、加州大学戴维斯分校（U. C. Davis）等大学、公司和部门任职。作为作者的学生，译者对作者严谨的科学态度、勤奋的工作作风、优秀的教学方法、和蔼谦逊及对学生的热情关怀留下了深刻印象。

原著是作者多年来从事非参数理论研究及教学的经验总结。其特点之一是侧重内在的统计思想的阐述，清楚易懂。原著曾被 D. S. Shmerling 译成俄文在前苏联出版。欧美的许多大学至今仍把原著当作统计专业研究生的教材使用。原著曾受到国际统计学会的《书评》（Short Book Reviews）、瑞典的 *Elementa*、美国的《数学评论》（Math. Reviews）、《新书评论》（New Book in Review）、《技术统计》（Technometrics）、《生物统计》（Biometrics）、《美国科学家》（American Scientist）、《美国统计学会杂志》（JASA）、《心理统计》（Psychometrika）等杂志及书评的评论。卡耐基梅隆大学（Carnegie-Mellon U.）的 Dorota M. Dabrowska 博士在为原著写的评论中说：“这是将秩方法广泛应用于统计实践的一本极好的入门教材。……本书包含大量例题及练习，作为研究生一二年级的教材极为合适。”

中文译稿是译者在 1992 年 4 月完成的。在整个翻译过程中，深得史宁中博士的帮助和指导。史宁中博士审阅了全部译稿，同时也得到原著作者的合作，改正了原著中的一些印刷错误。译者在此，对史宁中博士和帮助译者抄写并整理了部分译稿的刘薇女士表示深深的感谢和敬意，对东北师范大学出版社的朋友们为中译本的出版所做出的努力表示衷心感谢。

由于译者水平所限，在翻译过程中，在书中许多专业术语目前尚无现成的中文译法的情况下，很可能会有许多不当之处。译者欢迎同行的前辈、专家及老师们给予指正。

译者 1994 年 8 月  
于宾夕法尼亚州立大学统计系

## 序 言

以秩为根据的非参数方法的历史相当短, 大约有 40 年左右。这里, 我想提一提几个重大的贡献。这里提到的只是主要的, 不是全部。这一领域仍将是一个活跃的研究领域, 而且在很大程度上归功于几个关键性的成就。

本书所介绍和评价的非参数方法开始于 1945 年 F. Wilcoxon 的工作及 1947 年 H. B. Mann 和 D. R. Whitney 的工作。50 年代, 有人研究了关于位置 (location) 的非参数检验, 其方法是利用 Pitman 渐近有效性来估价检验的局部势性质。在一系列的文章中, J. L. Hodges 和 E. L. Lehmann 发现了一个令人惊奇的结果, 即与正态模型中的  $t$  检验相比较, 秩检验能经受住有效性的较小损失, 而且在粗尾模型 (heavy-tailed model) 中, 秩检验可能更有效。大约在这个时候, 非参数检验开始得到数据分析家的一些承认, 并且找到了进入初等教材最后一章的途径。第一本论述应用非参数统计的书是由 S. Siegel 在 1956 年出版的。这是一个很大的成功, 尤其是在行为科学家们中间。事实上, 从 1961 年到 1972 年, 该书被引用了 1824 次, 在被引用的次数最多的数学和统计学书目上名列第二。从 1970 年开始, 论述非参数方法的新书以大约每年一本的速度出现了。

60 年代, Hodges 和 Lehmann 从秩检验统计量出发, 导出了位置参数的点估计和置信区间。他们还证明, 这种估计方法从原来的检验统计量那里继承了有效性性质。他们还发现, 根据 J. W. Tukey、P. J. Huber 及 F. Hampel 为衡量估计量的稳定性而提出的新准则, 这些估计量是稳健的。研究一般秩计分检验统计量的

构造,需要渐近分布理论。这期间, J. Hajek 为这个渐近分布理论提出了一个新的有效方法。

为了分析已设计的试验, Hodges 和 Lehmann 在 60 年代初引入了排列的秩检验 (aligned rank tests)。M. L. Puri 和 P. K. Sen 又广泛研究了 this 检验。J. N. Adichie 对简单回归模型, 提出并讨论了秩检验及其相对应的估计。

在 70 年代, 先前的研究工作加强了, 而且还扩展到了线性模型中的秩检验和秩估计。在线性模型所需要的渐近分布理论中, 很多内容都起源于 J. Jureckova 于 1970 年以前发表的基本结果。以她的研究为基础, 有可能产生出统一的以秩为根据的分析复杂数据集的方法。有希望的话, 在 80 年代, 我们将看到计算机的使用及这些有效的、稳健的统计方法的更加广泛的应用。

本书的主要目的是系统介绍统一连贯的 (以秩为根据的) 统计方法, 以便在各种试验场合中进行统计推断。本书首先介绍的是单样本位置模型, 然后是双样本位置模型、单因素和双因素分类及一般线性模型。最后一章系统介绍多元位置模型的方法。在所有的场合里, 对于每个模型, 检验和估计都是作为相互关联的一组方法而一起阐述的。

由于需要, 我们引入了数理统计中的一些基本工具和结果。这些工具分为两种: 一是衡量统计方法的统计性质所需要的, 二是衡量稳定性所需要的。前一种工具主要是渐近相对有效性和渐近局部势。后一种主要是影响曲线和容忍限 (influence curve and the tolerance) (即崩溃 breakdown)。稳定性准则是稳健统计方法现代理论的核心。本书对所介绍的所有方法的有效性性质都作了描述。我们广泛系统地介绍了单样本位置模型中的稳健性性质, 而对简单回归模型中的稳健性性质的讨论却比较简短。这样做的目的是帮助学生逐渐形成有效性及稳定性方面的使用知识。

因为这是一本研究统计推断及数据分析的教材, 所以书的内

容是围绕统计方法编写的。深入理解了这些方法及其在简单模型中的性质，读者将会为讨论一般线性模型中的统计方法而做好准备。我们严格阐述了以秩之和为根据的统计方法。这些方法包括 Wilcoxon 符号秩统计量、Mann-Whitney-Wilcoxon 统计量、Kruskal-Wallis 统计量、Friedman 统计量及线性模型中以残差为根据的秩检验。我们讨论了更一般的秩计分和及其来源，并把前者综合在后者的讨论当中。有两个理由使我们的讨论集中在秩数和上：一是研究人员应用最普遍的是秩数和；二是具有很少量的数学经验就可探索秩数和的性质。

一般来说，在论述非参数统计的教材当中，无论是在理论上还是在应用上，都没有系统处理包括多元回归及方差设计分析在内的线性模型。因为多数数据分析是在线性模型的情况下进行的，所以不系统处理线性模型是一个严重的遗漏。认真的研究人员还没有更加广泛地应用非参数方法，部分原因是他们对线性模型缺乏系统了解。本书详述了这些方法。此外，在不远的将来，有可能找到统计软件来实现这些方法。Minitab 统计计算系统已经包含了简单设计所需要用的主要非参数方法。将来，它还会包含能做秩检验及秩估计的秩回归指令。因此，本书所介绍的统计方法将全部可以使用。而且，研究人员可以用这些方法去分析复杂的数据集。

本书包含很多练习和问题。我们清楚地叙述了练习中的主要结果。因此，对不愿意花时间推导那些等式的读者来说，仍可找到那些等式。书后是一个附录。附录中是数理统计主要部分的重要结果(不带证明)。所有重要的统计方法都有带数据的应用例子。

前三章包括单样本位置模型和双样本位置模型，详述了有限样本理论及渐近分布理论，并且导出了检验、点估计及置信区间。通过渐近有效性、影响曲线和容忍限，我们还探讨了这些检验及估计的性质。这部分内容可以在研究生一、二年级的一个学期中

完成讲授。本书所需要的预备知识是数理统计方面的一个初级课程和现代微积分。

多数稳健性内容被安排在每节的最后。如果希望得到更多的统计推断,那么可跃过这些稳健性材料,这样可以论及单因素及双因素分类设计方面的论题(第4章),或者可以论及单样本及双样本单变量检验的多元形式(第6章)。线性模型中的秩方法(第5章)可以在下一学期讲授。第5章的内容需要统计学的较深的背景知识。读者应具备用矩阵表示线性模型的预备知识。

本书是从宾州州立大学最近这15年多的讲稿中产生的。我对这个题目的兴趣开始于本人在爱荷华(Iowa)大学期间,与 Bob Hogg 和 Tim Robertson 一起做的讨论和研究。我感谢他们长时间的鼓励及和我的交谈。这些年和我一起工作的研究生们为我提供了澄清许多思想所需要的连续不断的挑战。我特别感谢 Joe McKean 和 Jay Aubuchon。

我还要感谢宾州州立大学统计系主任 Bill Harkness 的不断支持。海军研究署(Office of Naval Research)资助了第5章中出现的部分研究成果,在此向他们表示诚挚的谢意。John Wiley 出版社的 Bea Shube 在与本书有关的工作中给予了很大帮助。最后,感谢打字员 Jane Uhrine、Peggy Lynch、Bonnie Cain 和 Barbara Itinger。特别是 Bonnie Henninger, 努力誊写了我的笔记。

托马斯 P. 海特曼斯伯格

1984年1月于宾夕法尼亚,

斯泰特科利奇

# 目 录

## 1 具有任意、连续分布的

### 单样本位置模型 ..... 1

- 1.1 引言 ..... 1
- 1.2 符号检验及其分布 ..... 3
- 1.3 统计检验的相合性 ..... 6
- 1.4 一个最优势检验 ..... 9
- 1.5 估计 ..... 13
- 1.6 稳定性 ..... 19
- 1.7 小结及数据内相关性的效应 ..... 26
- 1.8 练习 ..... 29

## 2 具有对称、连续分布的

### 单样本位置模型 ..... 33

- 2.1 引言 ..... 33
- 2.2 Wilcoxon 符号秩检验 ..... 35
- 2.3 由 Wilcoxon 符号秩统计量导出的  
点估计和区间估计 ..... 43
- 2.4 秩检验和秩估计的稳定性性质 ..... 47
- 2.5 Wilcoxon 符号秩统计量的一般渐近理论 ..... 53
- 2.6 渐近相对有效性 ..... 70
- 2.7 Wilcoxon 符号秩统计量的渐近线性性,

## 2 目 录

---

|                                      |            |
|--------------------------------------|------------|
| 小结及相关性效应·····                        | 85         |
| 2.8 一般计分统计量·····                     | 95         |
| 2.9 一般计分统计量的有效性 ·····                | 113        |
| 2.10 练习·····                         | 136        |
| <b>3 双样本位置模型 ·····</b>               | <b>145</b> |
| 3.1 引言 ·····                         | 145        |
| 3.2 Mann-Whitney-Wilcoxon 秩统计量 ····· | 147        |
| 3.3 秩在备选假设下的分布 ·····                 | 156        |
| 3.4 一般计分 ·····                       | 163        |
| 3.5 备选假设下的渐近分布理论 ·····               | 174        |
| 3.6 设计的比较 ·····                      | 186        |
| 3.7 练习 ·····                         | 189        |
| <b>4 单因素分类、双因素分类及秩相关 ·····</b>       | <b>196</b> |
| 4.1 引言 ·····                         | 196        |
| 4.2 单因素分类——Kruskal-Wallis 检验 ·····   | 197        |
| 4.3 双因素分类——Friedman 检验 ·····         | 213        |
| 4.4 秩相关与秩关联 ·····                    | 221        |
| 4.5 练习 ·····                         | 238        |
| <b>5 线性模型 ·····</b>                  | <b>243</b> |
| 5.1 引言和简单回归 ·····                    | 243        |
| 5.2 线性模型中的秩估计 ·····                  | 253        |
| 5.3 线性模型中的秩检验 ·····                  | 274        |
| 5.4 补充及例子 ·····                      | 289        |
| 5.5 练习 ·····                         | 302        |

---

|                                   |            |
|-----------------------------------|------------|
| <b>6 多元位置模型 .....</b>             | <b>306</b> |
| 6.1 引言 .....                      | 306        |
| 6.2 单样本多元位置模型 .....               | 307        |
| 6.3 双样本多元位置模型 .....               | 315        |
| 6.4 练习 .....                      | 320        |
| <br><b>附录：数学和数理统计学中的若干结果.....</b> | <b>324</b> |
| <br><b>参考文献.....</b>              | <b>339</b> |
| <br><b>名词索引.....</b>              | <b>352</b> |

## 第1章

## 具有任意、连续分布的单样本位置模型

## 1.1 引言

先考虑单个总体的情况,并且希望做出最少的假设.我们只假定总体分布是连续的,对总体分布的位置做出统计推断.

因此需要对任意一连续分布的位置定义一个度量.因为没有对总体分布的形状做出假设,所以均值和中位数这两个最普通的度量不一定重合.在分布对称的情形中,它们是重合的,而且自然地确定了总体分布的中心.

在一般背景中,中位数有两点优于均值.第一,作为大约将总体分布平分的那个点,中位数永远存在.但均值不一定存在,Cauchy 分布就是一个例子.第二,对于来自总体分布的轻微干扰,中位数是非常稳定的.进而,如果总体中存在异常值,或存在过失误差(gross errors),这些异常值或过失误差对中位数所产生的影响将会很小,却可能给总体均值带来极大变化.本章将详细讨论均值和中位数的稳定性性质.目前,我们用总体中位数  $\theta$  来做位置的度量.

这一章将介绍  $\theta$  的一个检验、点估计及置信区间,并研究它们的性质.本章出现的一些想法将贯穿全书,而简单位置模型则为这些想法的讨论提供了一个非常好的素材.本章还讨论检验的显著水平和势的渐近性,还包括对检验的相合性(consistency)以及由检

验统计量导出估计方法的讨论.

假设  $X$  是具有任意连续分布函数  $F(x) = P(X \leq x)$  的随机变量, 定义中位数为一点  $\theta$ , 满足

$$P(X \leq \theta) = P(X \geq \theta) \geq 1/2. \quad (1.1.1)$$

我们称  $\theta$  为  $X$  的中位数, 也称  $\theta$  为  $F$  的中位数.

一般来说, 可能存在许多这样的  $\theta$  满足中位数的条件 (练习 1.8.1). 当这种情况发生时, 应指定一个法则来选取一个特殊的中位数.

在本书中, 我们假定中位数是唯一的, 而且  $F$  绝对连续,  $f(x) = F'(x)$  是它的密度函数. 将这一类分布记为

$$\Omega_0 = \left\{ F : F \text{ 绝对连续且 } F(0) = \frac{1}{2} \text{ 唯一} \right\}. \quad (1.1.2)$$

设抽样模型是一个由独立同分布 (i. i. d.) 随机变量组成的随机样本  $X_1, \dots, X_n$ , 每个  $X_i$  的分布都是  $F(x - \theta)$ ,  $F \in \Omega_0$ . 我们考虑的统计推断问题是

$$\begin{aligned} H_0 : \theta = 0 & \quad \text{对} \\ H_A : \theta > 0 & \end{aligned} \quad (1.1.3)$$

的检验.

这是需要考虑的最一般的单边假设. 检验问题  $H_0 : \theta = \theta_0$  对  $H_A : \theta > \theta_0$ , 其中  $\theta_0$  为一特定值, 可归结为 (1.1.3). 事实上, 在  $H_0$  之下,  $Y_1 = X_1 - \theta_0, \dots, Y_n = X_n - \theta_0$  是来自  $F \in \Omega_0$  的一个样本. 我们主要讨论单边检验, 是因为如何将其结果推广到相应的双边检验问题一般来说是清楚的. 最后注意: 零假设和备选假设都是复杂的.  $H_0$  只指出, 样本来自某一任意绝对连续的具有唯一的中位数  $\theta$  的分布;  $H_A$  只指明任意一个具有正中位数的分布.

**【例 1.1.1】** 在 20 世纪 50 年代初, G. V. T. Matthews 和一些人研究了鸟的飞行试验. 关于这项研究的解释, 见 Matthews