

《Robust Statistics》一书评介*

陈希孺

(中国科学技术大学)

Peter J. Huber: Robust Statistics, John Wiley, 1981

Robust Statistics (姑译为“稳健统计”。这个名词在国内统计界尚无确定的译法)的概念并无公认的正式定义,原因在相当程度上是因为,这个词的含义可以从种种不同的侧重点对它进行解释,不同的解释引出不同的定义,但总的说来,它是一种以可能在效率上作一定的牺牲为代价,而求得更大的可靠性为特征的统计方法。这种思想的萌芽可以追溯得很早,某些重要的稳健统计方法,如修剪平均(Trimmed mean),很早以前就有人使用,“Robust”一词本身是G. E. P. Box在1953年引进的, Tukey在1960年写的文章[1]对推动这个分支的建立起了重大的作用。然而,由于缺乏数学上的严密理论,直到1964年,统计学界,特别是其主要兴趣在于数学理论的那部分人,总觉得这个东西不大“正规”,好象有点登不得大雅之堂似的。1964年Huber发表了著名的工作[2],对位置参数的稳健估计的Minimax理论作了严格的处理,引起了理论统计学者的重视,从那时起这方面的研究工作进展很快,积累了大量文献。因此P. J. Huber,即我们下面要评介的专著的作者,确然可认为是稳健统计最重要的奠基者和推动者之一。

近十余年来,在稳健统计方面虽然也出现了少量专著,例如[3],[4]。不过[3]只是讨论了位置参数的情况且偏重模拟结果,而[4]只是一个简单的介绍,想要对这方面的现代面貌有较全面的理解,必须翻检大量文献,Huber新著《Robust Statistics》的出版,填补了这方面的空缺,相信会对这个分支的定型化、“普及”,以及进一步的研究工作,起相当的促进作用。

全书共11章,正文293页,各章标题如下:1. 一般概念。2. 弱拓扑及其距离化。3. 估计的基本类型。4. 位置参数估计的渐近Minimax理论。5. 刻度参数。6. 多参数问题,特别是位置一刻度参数同时估计。7. 回归。8. 稳健的相关和协差阵。9. 设计的稳健性。10. 精确的小样本结果。11. 杂题。附有索引及参考文献134条。

关于本书的性质与特点,最好是引作者自己的一段话来说明:“本书对理论和应用统计学者提供关于稳健性的一坚实基础,其写法是理论性的,但着重点在于概念,而不在数

* 1982年4月6日收到。

学上的完整性”。本书总结了稳健统计研究方面到目前为止的主要成果，但依本文作者的观点看，本书的主要意义，首先还不在于材料方面，而在于材料的组织，概念和观点的阐述上，以使读者对这个分支有一种正确而实在的观点，这一点很重要。因为，特别像“稳健性”这样发端于实际应用且在统计这座大厦中属于“上层建筑”性质的东西，如果只注意其数学方面，就会丧失其现实性且变得索然无味了。

Huber 从开始写作本书到最后出版，前后经历了近十年的时间，他边教边写，这十年正是稳健统计迅速发展的时期，而他又处在这个发展的中心和带头人的地位，因此这本书是精心之作。本书行文简练清楚，这样他能以相对说来较小篇幅包含了这么多的内容。例如第一章，总共只 18 页的篇幅，开头以 Tukey[1]中的一个著名例子说明稳健统计的必要性，接着说明稳健统计方法要达到的目的，然后从属性和数量的角度对稳健性的概念作了分析，引进了像“影响曲线”，“崩溃点”这样的基本概念，讨论了稳健估计的优良性及计算方法等。读者通过这样很短几页，就对整个稳健性问题的全貌有了一个初步的概念，而可用于指导阅读以后的章节。第二章也值得一提，如作者自己所说，本书的重点在于“分布稳健性”，即：当模型中关于分布的假设有少许背离时，统计方法的性能不应有显著变化。因此，怎样去衡量分布之间的差异就成为一个根本问题。在第二章中，作者对这个问题的有关材料，如弱拓扑、Lévy 与 Prohorov 度量、有界 Lipshitz 度量等，作了系统的阐述，这些材料在目前出版的其他著作中只是散见，这个内容，即使对于其主要兴趣不在稳健统计方面的同志，也是很有用的。其它章节的内容，读者可在前列标题中窥见一斑，由篇幅关系不便在此作较仔细的介绍。

对于具有一定数学知识（相当于大学二年级的数学知识及复旦大学《概率论》第一册的、主要兴趣在于应用的人来说，本书可使他对稳健统计方法有一个系统了解，明白它可以做到什么，不可以做到什么，那些方法目前看来在应用上较有前途等。本书适合于作为统计专业“稳健性统计”课程的教材和主要参考书，作为青年教师扩大视野提高水平进修的材料，至于专门研究者（指一个未甚入门但对这分支感兴趣的研究者），本文作者的看法是：本书可以带领他走很长一段路，指引他进一步研究从何处入手，但不能代替专门文献。如果他想研究稳健统计中的前沿问题，他需要在阅读本书的基础上进一步钻研有关的杂志文献。

参 考 文 献

- [1] Tukey, J. W. A Survey of Sampling from contaminated distributions, in «Contributions to Probability and Statistics», Stanford Univ. Press.
- [2] Huber, P. J. Robust estimation of a location parameter, *Ann. Math. Statist.* 35(1964), p.73.
- [3] Andrews, D. F. et al, Robust Estimation of Location Survey and Advances, Princeton Univ. Press, 1972.
- [4] Huber, P. J., Robust Statistical Procedures. Reginal Conference Series in Applied Mathematics No. 27. Soc. Industr. Appl. Math., Philadelphia, Penn.