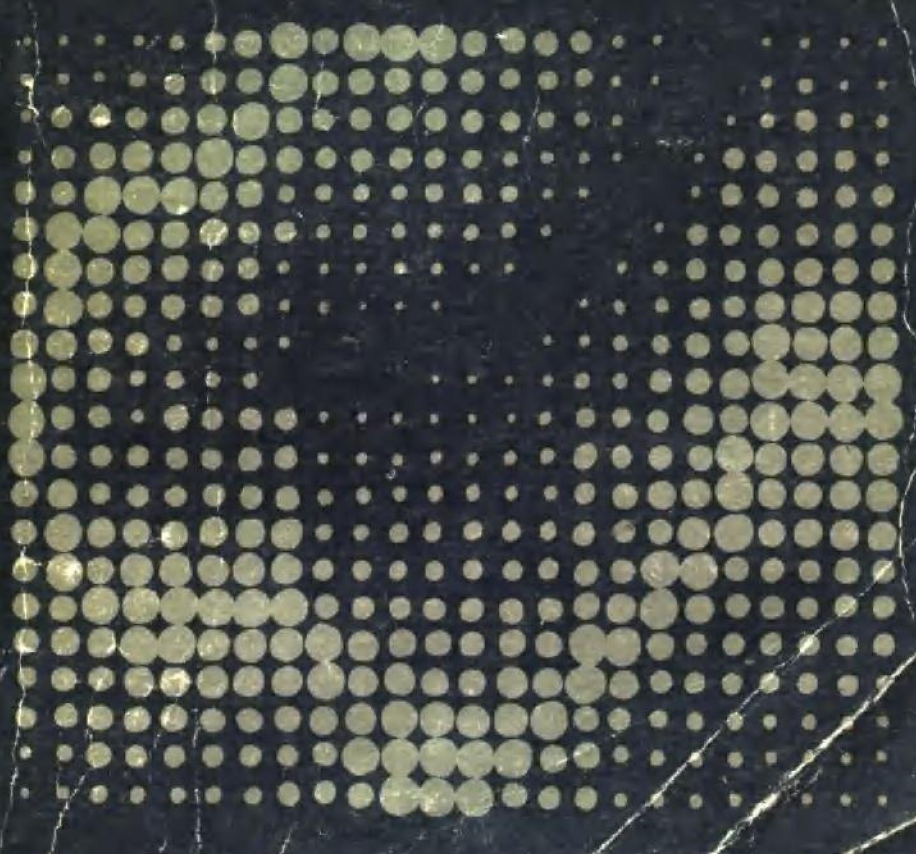


美 J·G·H·维恩堡 J·A·休麦克 D·奥尔特曼 著

常学将 胡文明 王明生 詹森 刘兰亭 译

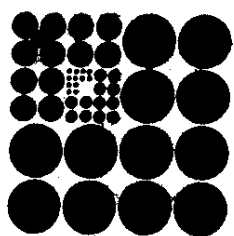
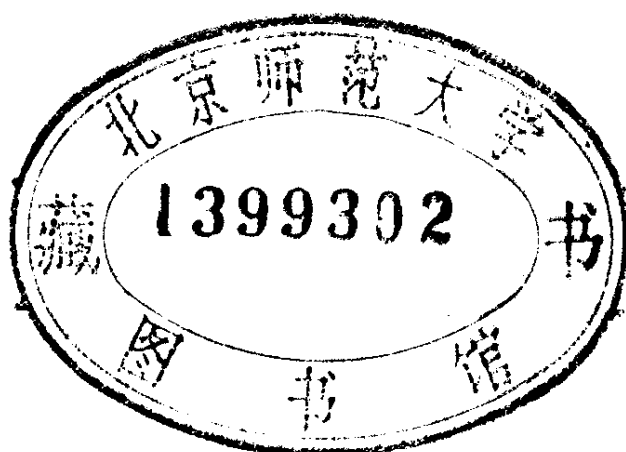
数理统计 初级教程



数理统计 初级教程

[美]G·H·维恩堡 J·A·休麦克 D·奥尔特曼 著

常学将 胡文明 王明生 詹森 刘兰亭 译



山西人民出版社

数理统计初级教程

G·H·维恩堡

〔美〕J·A·休麦克 著

D·奥尔特曼

常学将 胡文明 王明生 译
詹 森 刘兰亭

*

山西人民出版社出版 (太原并州北路十一号)

山西省新华书店发行 山西新华印刷厂印刷

*

开本: 850×1168 1/32 印张: 20.375 字数: 481千字

1986年8月第1版 1986年8月太原第1次印刷

印数: 1—17,000册

书号: 4088.136 定价: 3.90元

中译本序

统计是认识社会的一种重要手段，是正确制定社会经济发展决策的重要依据。经济越发展，越需要加强统计。经济越搞活，越需要发挥统计的监督作用。

统计工作要做到数字准确，资料丰富，信息灵通，须从多方面着手。其中重要的一着是统计干部要掌握渊博的知识和多方面的业务技能，包括掌握数理统计的理论、方法和应用技能。

数理统计与社会经济统计在理论与方法上有很多共同之处。通晓数理统计的人对社会经济统计的理论和方法就会理解得更深刻，运用得更自如，技巧更娴熟，就会如鱼得水，如虎生翼，在统计工作这个领域里，得到更多的自由。

近几年来，统计干部中学习数理统计的人渐渐多起来了，这是十分可喜的现象。数理统计知识的普及与应用，对提高我们的统计理论水平及业务工作能力，无疑将是一个极大的推动。基于这样的认识，山西大学常学将教授和山西省统计局胡文明同志等翻译了这本书，目的在于给广大统计人员提供一个初级的、系统的教材。

本书浅显易懂，不仅适合广大理论工作者、教学工作者和实际工作者在研究、教学及工作中参考，而且适合初学者自学使用。本书在着重解释概念的同时，还写出了必要的算法，并以实际生活中的有趣事例反复讲解，辅之以知识范围十分广泛的习题，生动活泼，便于学习领会。

本书的出版，将对加快我国统计工作现代化的步伐，开创统计工作的新局面，起到一定的作用。

张 塞

1985.6.12

作者原序

本书系统阐述了初级统计教程应包括的各种常见课题。它既适合低年级使用，也适合高年级使用。由于着眼点在于深入解释概念，因此对读者的数学知识要求甚少。我们衷心希望给学生提供一个清晰的来龙去脉，而又不使他们象通常学习此课程那样感到苦恼。采用的办法是等到把每个新概念彻底搞清楚之后，才提出各种困难的问题，以期尽量减少学生常发生的“数学愁”。通过这种办法，再针对各种实际情况，细心安排一些例子并使用幽默的语言，使学生在学此书时感到是一种享受。

虽然本书的前三版是成功的，但在修订本中，我们还是做了重大的修改。许多修改是为了保证跟上时代的潮流。例如着重增加了非参数方法。此外，还反映了使用本书的教师和学生的很好的意见。前几版的优点，特别是写得有意思的部分，均予保留。

这一版明显的改变是题目。前六章讲授描述统计方法。第七、第八章专门讲授正态分布与中心极限定理。在第九章中，可以学到概率论的广泛知识。紧跟着在第十章讲授判定与风险的初步知识；这一章是本书的一个独到之处。第十一章到第十四章有最常用的参数统计检验与估计。回归、相关、线性预测则放在第十五章到第十七章。最后两章是非参数检验。第四版的一个新特点是包括四个复习，放在全书的适当位置。

本版编入一些新内容，例如从分组数据求百分等级的方法，拟合优度的 χ^2 检验，曼-惠特尼检验，威尔科克斯符号检验。

其他章节也做了扩充或改进。有关概率与二项分布的正态逼近的段落都增加了内容。讲判定与风险的一章，现在增加了涉及样本均值分布做判定时，第Ⅰ类与第Ⅱ类错误概率的计算。在常规的假设检验方法之外，增加了附加步骤，从而在结果的解释方面有新的着重点。整个的六步检验用许多例子加以说明。

全部练习都经过重新整理。由第二章开始，每章末尾有两组习题。每组习题都是用该章内容可以解决的。两组习题是类似的，在每组中也有相类似的练习。单数题号的习题有答案。双数题号的习题对教师来说答案是容易得到的。有足够的习题既可供学生练习，也可供教师在考核时使用。习题涉及很多领域的知识，这恰好反映了统计方法的宽广的实用性。对于严肃认真的学生，这样会增强他们学习此门课程的信心和兴趣。（下略）

G·H·维恩堡

J·A·休麦克

D·奥尔特曼

目 录

中译本序.....(1)

作者原序.....(1)

第一章 绪论.....(1)

1.1 你与统计(1)

1.2 统计史概述(2)

1.3 描述统计与抽样统计(5)

1.4 随机样本和分层随机样本(7)

1.5 统计方法的滥用(8)

1.6 对统计的怀疑主义与对这种观点的反驳(9)

习题(11)

第二章 平均数, 中位数, 众数.....(14)

2.1 定义(14)

2.2 平均数的两个性质(17)

2.3 项的改变对平均数的影响(21)

2.4 平均数、中位数及众数的应用(26)

2.5 记号(27)

习题(28)

第三章 可变性及两种可变性测度.....(36)

3.1 可变性(36)

3.2	方差	(38)
3.3	标准差	(41)
3.4	改变分布中的项如何影响标准差及方差	(43)
3.5	方差的应用	(45)
3.6	记号	(45)
3.7	方差的计算公式	(49)
	习题	(50)
第四章 说明分布中项的位置的方法		(58)
4.1	百分等级与百分位	(58)
4.2	z 变量	(62)
4.3	z 变量平均数	(66)
4.4	z 变量与百分等级之间无固定关系	(68)
4.5	z 变量集合的平均数与标准差	(70)
4.6	标准变量 (或 T 变量)	(72)
4.7	应用	(72)
4.8	记号	(73)
	习题	(74)
第五章 数据分组与作图		(81)
5.1	离散变量与连续变量	(81)
5.2	离散变量值的列表与作图	(83)
5.3	舍入	(85)
5.4	数据分组——离散情况	(87)
5.5	数据分组——连续情况	(91)
5.6	作为渐近图形的曲线	(95)
5.7	面积的意义	(97)

5.8	以数据确定分布图形的形状	(99)
	习题	(102)
 第六章 用分组数据计算各种指标(113)		
6.1	引言	(113)
6.2	用直方图求中位数与百分位数	(113)
6.3	中位数及其它百分位数的计算	(116)
6.4	四分位数与十分位数	(119)
6.5	由分组数据求百分等级	(120)
6.6	累积曲线	(121)
6.7	分组数据的众数	(123)
6.8	分组数据的平均数——直接方法	(123)
6.9	分组数据的平均数——编码方法	(126)
6.10	关于分组数据的方差与标准差——直接方法	(130)
6.11	分组数据的方差与标准差——编码方式	(133)
	习题	(135)
 总复习 I		
 第七章 正态分布(153)		
7.1	理论分布	(153)
7.2	正态分布与正态曲线	(153)
7.3	正态分布的四个性质	(156)
7.4	使用正态分布表作关于项的推断	(158)
7.5	射击的分布	(167)
7.6	由百分等级确定 z 变量的值	(170)
	习题	(172)

第八章	样本和的分布及样本平均数的分布	(182)
8.1	中心极限定理	(182)
8.2	如何实际使用定理 8.1	(184)
8.3	样本平均数的中心极限定理	(186)
8.4	样本平均数分布的性质	(188)
8.5	确定一个样本平均数在其样本分布中的位置	(192)
	习题	(196)
第九章	概率	(204)
9.1	概率	(204)
9.2	日常说法中的“概率”一词	(207)
9.3	样本空间和事件	(208)
9.4	概率与抽样试验	(211)
9.5	概率和正态分布	(213)
9.6	关于样本平均的概率陈述	(215)
9.7	二项分布的正态逼近	(218)
	习题	(223)
第十章	判定和风险	(228)
10.1	引言	(228)
10.2	假设检验	(231)
10.3	关于均匀硬币和有偏硬币检验的最终结果	(235)
10.4	利用样本平均值的分布计算各类错误	(242)
10.5	关于风险的最后的话	(246)
	习题	(247)

总复习 II	(253)
--------------	-------

第十一章 假设检验方法	(262)
--------------------------	-------

11.1 引言	(262)
11.2 零假设	(263)
11.3 什么时候拒绝零假设, 显著性水平	(264)
11.4 三个假设检验问题	(273)
11.5 某些新说法	(279)
11.6 单边检验	(279)
习题	(283)

第十二章 估计	(290)
----------------------	-------

12.1 引言	(290)
12.2 估计的无偏性	(290)
12.3 自由度概念	(296)
12.4 方差的估计	(299)
12.5 s^2 的计算公式	(303)
12.6 区间估计	(304)
12.7 总体均值的区间估计	(306)
习题	(311)

第十三章 t 分布与t检验	(317)
----------------------------	-------

13.1 引言	(317)
13.2 t 分布	(320)
13.3 两平均数之差的t检验	(325)
13.4 平均数之差的t检验的应用	(331)
13.5 匹配分组的t检验	(336)

习题.....	(340)
第十四章 F 检验和方差分析.....	(349)
14.1 引言.....	(349)
14.2 F 分布的理论模型.....	(350)
14.3 两个以上的样本均值的比较.....	(358)
14.4 方差分析引论.....	(360)
14.5 方差分析技术.....	(367)
习题.....	(373)
 总复习 III	(383)
 第十五章 回归与预测.....	(395)
15.1 引言.....	(395)
15.2 盲目预测.....	(396)
15.3 预定变量和预测变量.....	(400)
15.4 回归.....	(411)
习题.....	(415)
 第十六章 相关.....	(422)
16.1 相关的概念.....	(422)
16.2 两 z 变量取值集合之间的相关.....	(433)
16.3 相关系数.....	(436)
16.4 相关系数的计算.....	(440)
16.5 相关系数的显著性检验.....	(445)
习题.....	(447)

第十七章 相关与线性回归.....(455)

17.1 引言.....(455)

17.2 相关系数的几何意义.....(459)

17.3 原始数据散布图的最优拟合直线.....(462)

17.4 最优拟合直线与预测.....(467)

17.5 相关与线性预测的精度.....(473)

习题.....(477)

第十八章 独立性的 χ^2 检验和拟合优度.....(484)

18.1 引言.....(484)

18.2 独立性.....(485)

18.3 相关的实例.....(486)

18.4 以列联表表示调查结果.....(490)

18.5 期望频数.....(491)

18.6 χ^2 检验.....(493)

18.7 更复杂的 χ^2 检验.....(497)

18.8 χ^2 检验与“拟合优度”.....(502)

习题.....(506)

第十九章 非参数统计检验.....(513)

19.1 引言.....(513)

19.2 曼-惠特尼检验.....(514)

19.3 符号检验——“事先”和“事后”资料.....(518)

19.4 威尔柯克斯符号秩检验——自然对样本.....(523)

19.5 秩相关检验——皮尔逊相关系数 r
的显著性检验的另一种形式.....(527)

19.6 随机性的游程检验.....(532)

习题	(537)
总复习 IV	(546)
附录	(564)
公式表	(564)
平方根表	(570)
正态曲线面积	(572)
正态分布表	(574)
t 分布	(575)
F 值	(576)
χ^2 值	(580)
u 的临界值 (总游程)	(581)
数学复习	(582)
部分习题的答案	(586)
索引	(629)
译后记	

第一章 绪 论

1.1 你与统计

读者在接触本书以前，对于统计这个课题的知识，可能甚少，也可能全然不懂，甚至会把“统计”一词混同于一堆数据、表格或图表。你可能听到过新闻分析家、政治家以至体育节目广播员的播音，他们都要提出统计数字，好象他们是很可信赖的。如生活费用、人口增长和棒球选手的打击率等等都有统计。事实上，考查你所关心的最近几天的每个实际问题，似乎都有统计。

没有经过预先处理的数据，看起来会使人不得要领。但经过处理的数据就会更有意义，更有用处。比如生活费的“指数”，棒球的“平均”打击率等，这些数字称为“统计”才更合适。

“指数”做为生活水平的测度，“平均数”做为某棒球选手本领的测度。这样的测度，通常称为“统计量”。这是关于统计的一种理解。这一点在以后的课文中将予以充分发挥。

“统计”一词也常常当作统计方法的缩写来用。统计方法是一种用于收集、表示（图解）、分析和解释数据的方法。比如，一个棒球手处于攻方，由各次攻击的数据，得到该选手最后打击率的方法，就是统计方法。

我们主要关心对统计概念与统计方法的了解。这些概念与方法大都可以借助常识的推理去理解。我们主要的目的，是指出所以做出这种推理的理由。事实上，统计方法可以说成是常识推理

对数据分析的运用。

从本质上讲，统计方法就是从一大堆资料中提取有用信息。你可能正从事某个方面的工作，想从你或别人搜集到的资料中导出有用的结论。在这种情况下，你就会感觉到统计知识的价值。为什么你需要熟悉统计问题呢？我们请你注意两个主要原因。

大多数专业领域的文献，都包括该领域研究者的实验结果，所公布的报告往往使用统计语言。这就要求任何试图使用这些统计资料的读者，具有起码的统计知识。如果需要决定应否采取某种新的试验方法时，统计知识更是不可缺少的。

不仅专业文献中广泛使用统计，报纸和刊物也常常使用统计术语，在提出“平均数”和说明图表时，尤其如此。从本章稍后一节里可以看到，在许多例子中，统计资料不仅为人们所利用，甚至常被滥用。读者在阅读这类材料时，自然应予注意。

对统计知识的这种需要，我们写作本书时牢记在心。我们的目的是能使读者对统计知识有一个清晰的理解。

当然，学习基础统计需要用到一定的数学。附录Ⅷ是一个复习。对初学统计学的读者来说，先读读基础算术及代数公式的简述，可能会感到舒服一些。本书把符号表述控制在最少限度内，并且在导入时作详尽解释。本书并非数学专著。我们只想帮助你变成一个具有统计学方面知识的人。

1.2 统计史概述

统计的发展有点象语言本身的发展。统计起源于古代。例如，平均数的应用在毕达哥拉斯（Pythagoras）时代已有相当了解。早在圣经时代就提到过统计调查。象语言一样，由于需要，统计逐渐发展起来。随着社会的日益复杂，人们对准确总结和以