



五南商学Top



统计学

吴冬友 杨玉坤 著



中国统计出版社



统计学

[台湾] 吴冬友、杨玉坤 著



中国税务出版社

图书在版编目(CIP)数据

统计学/吴冬友著. —北京: 中国税务出版社, 2005. 3

ISBN 7-80117-728-2

I. 统... II. 吴... III. 统计学 IV. C8

中国版本图书馆 CIP 数据核字(2005)第 018666 号

本书为(台湾)五南图书公司授权中国税务出版社在大陆地区出版发行简体字版本。

著作权合同登记号图字:01-2005-1883 号

版权所有·侵权必究

书 名: 统计学

作 者: 吴冬友 杨玉坤

责任编辑: 刘 莉

特约编审: 李 德

责任校对: 孙荣杰 刘 坤

出版发行: 中国税务出版社

经 销: 北京中税五南文化发展有限责任公司

地 址: 北京市宣武区宣武门西大街 28 号大成广场 9 号门 1904 室

邮 编: 100053

电 话: 010-63601825 转 604/606

传 真: 010-63601825 转 601

e-mail: wunan@vip.sina.com

印 刷: 北京神剑印刷厂

规 格: 720 毫米×1000 毫米

印 张: 43.75

字 数: 647 千字

版 次: 2005 年 4 月第 1 版 2005 年 4 月北京第 1 次印刷

书 号: ISBN 7-80117-728-2/F. 703

定 价: 68.00 元

如发现有印装错误 可随时寄经销处调换

运筹帷幄 决胜千里
商场如战 快意驰骋

科学的数据 理性的判断
正确的决策 科学的工具



以 科学的工具 量化的数据
做 理性的判断 正确的决策

三版序

由多年统计学的教学经验得知同学们在学习上的困扰有：(1)基本概念的模糊不清，(2)统计理论理解的障碍，(3)计算流程掌握的失控，(4)理论与实际的应用困境。所以，如何以深入浅出的方式使同学们能在概念、理论、方法及应用上有一个整体的认识，是我们写这本《统计学》的目的。例如，同学们对统计假设检定的过程有许多困惑及误解，诸如假设的拟定、检定统计量的选择、拒绝域及接受域的意义、显著水准及显著概率（ p 值）的差别、型一及型二决策失误的计算等等。在本书第 11 章第 1 节中，通过一步一步清楚的说明、定义及 20 个搭配的相关例题，同学们研读之后，对于上述困惑将会有柳暗花明的清醒。此外，我们也尽可能使用流程图协助同学们充分掌握概念、理论及演算流程的连贯性。例如，在参数估计、假设检定、变异数分析、回归分析、卡方检定、无母数检定等章节中，本书通过 40 多个流程图来引导概念思考的主轴，并在步骤程序说明中陈述相关的理论，使同学们能将概念、理论、方法及流程紧密结合，融会贯通。

第 3 版新增内容有：(1)第 1 章除了统计资料更新外，还改写了大部分的内容，例如统计学、概率论及统计方法间关系的说明；(2)第 4 章增加了平方和及变异数速算公式、变异系数、动差、偏态及峰态；(3)第 6 章则增加了动差母函数。在不影响全书结构及保留教学弹性的前提下，这些增加的内容均列在各章的附录中。

为了提升教学效率及命题的方便，我们也为老师们提供了教学附件。(1)辅助投影片：定义、定理、图、表及流程图，(2)题库：各章题目均以难易程度（简单、中等、难）分类，(3)习题解答。

感谢辅仁大学大学部及研究所同学在使用本书的过程中提出的许多宝贵的意见，尤其是罗梓豪同学的细心校对，是同学们的认真投入使得本书第3版在内容上更趋完整。最后，感谢五南图书张毓芬副总编辑、陈舒榆编辑及校对们在编辑、校稿、排版及打字上的细心投入，使得本书在编排上更具可读性。

吴冬友

杨玉坤

2004年5月

再版序

本书首版承蒙诸位统计学授课教师及同学的厚爱，一年内即再版。无可讳言，首版虽几经校对仍有许多漏网的错误，对于诸位先生及同学们热心帮忙找出这些错误，致上我们最诚挚的谢意。为了感谢各方所给予的肯定及建议，我们除了更审慎仔细校对外，并于再版中增加了以下附录：(1)增加索引(index)以方便读者查阅及前后对照，(2)增加各分配的重要特性及它们间的相互关系，(3)分别将四大连续分配(常态、 t 、 χ^2 及 F)及6个间断分配(超几何、伯努力、二项、卜瓦松、几何及负二项)彼此间的关系用简单明了的关联图来说明，以达化繁为简、加深印象的效果。

吴冬友

杨玉坤

2003年6月

序

统计学是一门数量学科，也是一门应用学科。作为一本统计学教科书，如果偏重数量方法的解说而缺乏实例应用，就好像只提供数量技巧但未告知如何将它应用在解决实际问题上；反之，如果偏重实际问题但对数量方法的说明不够完整，则又好像只告知它的用途但欠缺清晰的操作概念，所以，如何在两者之间取得平衡是本书的最高指导原则。

本书在内容与结构上的特色有三：(1)详细完整的数量方法操作流程。从收集资料、计算公式到陈述结论，尽量以图表或流程图的方式来说明，并在每一步骤中穿插相关的统计概念及定理，使你在有一个完整的计算流程之余同时也知道它的统计理论根据。例如，第6章及第7章中随机变数与样本空间关系以结构转换图加以说明；第9章、第10章的估计问题，第11章、第12章的检定问题，第16章的卡方检定问题及第17章的无母数统计方法，皆以步骤程序配合流程图逐步说明；第13章变异数分析更进一步将资料处理用图表方式与流程图加以整合；第14章及第15章回归分析则是以一个简洁的回归分析结构图来贯穿所有回归分析要探讨的主题。(2)广泛的应用实例。本书中的例题与习题除了纯粹为了熟习计算的题型外，大多以模拟实际问题之形态使读者在运用统计方法之余也充分体会如何将统计方法应用于日常决策。例如，第5章有计算电子信箱地址的概率问题及罹患艾滋病概率的贝氏定理应用问题，第6章有概率分配处理物流仓储的管理问题，第9章至第17章的例题及习题广泛涵盖市场行销、财务分析、生产管理、社会治安、环境品质、教学实验等主题。(3)提供深入研究所应具备的基础训练。本书某些章节的附录有相当多统计理论的入门知识，提供有兴趣的读者具备更扎实的专业基础。例如，第5章附录了相当完整的集合运算公式，使你在面对艰深的

概率问题时能以集合的操作来处理；第六章附录了统计推理中非常重要的期望值运算（E-operator）基本功及其在统计推理上的应用；第7章附录了一些深入分析必备的概率分配及其统计性质；第8章附录了本书后续各章统计推论中所根据的基础定理；第13章附录了3种实验设计平方和的推导，从这些推导中可以使你除完全掌握运算（ Σ -operator）的操作要领之外还可以充分体会运算在统计中的重要性；第14章及第15章分别附录了简单回归及多元回归分析的基础理论。此外，本书末的附录也有一些基本的数学公式作为参考。

多做习题是深入了解及自我检验学习成果的最佳途径。鉴于此，我们另编有统计学习题解答作为辅助教材。这本习题解答的特色有二：(1)重点摘要提示。各章题解之前皆有定义、定理、重点提示及相关的架构流程方便查阅。(2)详细完整的解题过程。所有的题解尽可能配合概念、步骤及流程，使读者在反复练习中逐步强化学习效果。所以这本统计学习题解答，不但是在学同学们的辅助材料，对于有志于自我进修或准备考试的读者，也是一本非常适合的参考书。

感谢辅仁大学管理学研究所刘美文在研究所学业繁重之际，细心地为本书进行校对，更感谢五南图书副总编辑张毓芬，编辑、打字及校对为本书所投入的幕后工作。

吴冬友

杨玉坤

2002年5月

目 录

第 1 章 绪 论 ————— 1

1.1 统计学与概率论 / 1

1.2 统计学的内容 / 2

第 2 章 收集资料 ————— 9

2.1 母群体与样本 / 9

2.2 资料形态 / 11

2.3 资料来源及抽样方式 / 12

第 3 章 叙述统计之一 ————— 17

3.1 单变数——质的资料 / 18

3.2 单变数——量的资料 / 23

3.3 双变数资料 / 34

第 4 章 叙述统计之二 ————— 45

4.1 资料的中心值 / 45

4.2 资料的分散度 / 50

4.3 资料的百分位数 / 56

4.4 标准差的应用 / 60

附录 4-1 平方和及变异数的速算公式 / 65

附录 4-2 变异系数 / 66

附录 4-3 动差 / 68

附录 4-4 偏态与峰态 / 74

第 5 章 概率导论

83

- 5.1 概率专有名词 / 83
- 5.2 事件的概率值 / 86
- 5.3 复合事件——事件的运算 / 89
- 5.4 复合事件的概率运算法则——加法法则 / 94
- 5.5 联合概率, 条件概率与独立事件 / 100
- 5.6 乘法法则与树型图 / 110
- 5.7 贝氏定理 / 117
- 附录 5-1 集合基本定义 / 127
- 附录 5-2 集合运算的基本定律 / 128
- 附录 5-3 集合运算公式 / 130

第 6 章 随机变数及概率分配

133

- 6.1 随机变数 / 133
- 6.2 间断随机变数的概率分配 / 135
- 6.3 间断随机变数的期望值 / 142
- 6.4 间断随机变数的变异数 / 148
- 6.5 连续随机变数的概率分配 / 151
- 6.6 连续随机变数的期望值 / 154
- 6.7 连续随机变数的变异数 / 155
- 6.8 期望值与变异数的计算公式 / 156
- 附录 6-1 单一变项期望运算性质 / 165
- 附录 6-2 双变项期望运算性质 / 167
- 附录 6-3 相关系数 / 171
- 附录 6-4 契比雪夫定理 / 172
- 附录 6-5 动差母函数 / 173

第 7 章 几个常用的随机变数及其概率分配

181

- 7.1 伯努力试验 / 181
- 7.2 二项分配 / 185
- 7.3 卜瓦松分配 / 194

- 7.4 常态分配 / 203
- 7.5 近似分配 / 212
- 附录 7-1 多项分配 / 219
- 附录 7-2 几何分配 / 220
- 附录 7-3 超几何分配 / 222
- 附录 7-4 负二项分配 / 224
- 附录 7-5 均匀分配 / 227
- 附录 7-6 伽偶分配 / 229
- 附录 7-7 指数分配 / 231

第 8 章 抽样分配 ————— 235

- 8.1 抽样分配 / 237
- 8.2 $\bar{X}$ 的统计性质 / 246
- 8.3 中央极限定理 / 249
- 附录 8-1 大数法则 / 256
- 附录 8-2 常态分配的抽样分配 / 257
- 附录 8-3 抽样分配间的关系 / 258

第 9 章 估计(一) ————— 261

- 9.1 母群体平均数 μ 的估计之一 / 265
- 9.2 t 分配 / 276
- 9.3 母群体平均数 μ 的估计之二 / 281
- 9.4 母群体平均数 μ 的估计之三 / 285
- 9.5 母群体比例值 p 的估计 / 288
- 9.6 χ^2 分配 / 292
- 9.7 母群体变异数 σ^2 的估计 / 296
- 9.8 样本数的决定 / 299

第 10 章 估计(二) ————— 309

- 10.1 两个母群体平均数差的估计之一 / 309
- 10.2 两个母群体平均数差的估计之二 / 312

- 10.3 两个母群体平均数差的估计之三 / 315
- 10.4 两个母群体平均数差的估计之四 / 319
- 10.5 两母群体比例差的估计 / 321
- 10.6 F 分配 / 325
- 10.7 两个独立常态母群体变异数比值 $\frac{\sigma_1^2}{\sigma_2^2}$ 之估计 / 330

第 11 章 统计假设检定(一) 339

- 11.1 检定的基本观念 / 339
- 11.2 母群体平均数 μ 的检定之一 / 361
- 11.3 母群体平均数 μ 的检定之二 / 367
- 11.4 母群体平均数 μ 的检定之三 / 371
- 11.5 母群体比例值 p 的检定 / 375
- 11.6 母群体变异数 σ^2 的检定 / 380

第 12 章 统计假设检定(二) 389

- 12.1 两个母群体平均数差的检定之一 / 389
- 12.2 两个母群体平均数差的检定之二 / 395
- 12.3 两个母群体平均数差的检定之三 / 399
- 12.4 两个母群体平均数差的检定之四 / 404
- 12.5 两个母群体平均数差的检定之五 / 409
- 12.6 两个母群体比例差的检定 / 416
- 12.7 两个母群体变异数比值的检定 / 422

第 13 章 变异数分析 433

- 13.1 变异数分析的基本观念 / 435
- 13.2 单因子实验设计(一维变异数分析) / 438
- 13.3 区集设计(二维变异数分析) / 452
- 13.4 双因子设计 / 464
- 附录 13-1 单因子设计平方和恒等式 / 487
- 附录 13-2 区集设计平方和恒等式 / 488
- 附录 13-3 双因子设计平方和恒等式 / 490

第 14 章 简单回归与相关 ————— 495

- 14.1 基本假设 / 498
- 14.2 参数估计 / 501
- 14.3 统计性质 / 507
- 14.4 回归线的解释能力 / 514
- 14.5 检定与区间估计 / 517
- 14.6 预测 / 521
- 14.7 相关分析 / 525
- 附录 14-1 最小平方回归线 / 535
- 附录 14-2 最小残差平方和 (SSE) / 535
- 附录 14-3 简单回归线斜率 β_1 的一般检定程序 / 536
- 附录 14-4 简单回归线斜率 β_1 的区间估计 / 538
- 附录 14-5 简单回归线 $\beta_1=0$ 之 t 检定与 F 检定的相通性 / 538
- 附录 14-6 $X=x_0$ 时平均反应 $E(Y|X=x_0)$ 的 $100(1-\alpha)\%$ 信赖区间 / 539

第 15 章 多元线性回归 ————— 541

- 15.1 基本假设 / 542
- 15.2 参数估计 / 544
- 15.3 统计性质 / 548
- 15.4 多元回归判定系数 / 553
- 15.5 检定与区间估计 / 556
- 附录 15-1 多元回归模式的矩阵表示法 / 561
- 附录 15-2 例题 15-2-1 的矩阵计算法 / 562
- 附录 15-3 最小平方回归式系数的抽样分配 / 563
- 附录 15-4 多元回归 SST, SSR 及 SSE / 564

第 16 章 卡方检定 ————— 567

- 16.1 多项分配参数结构检定 / 567
- 16.2 独立性检定 / 572
- 16.3 适合度检定 / 577

第 17 章 无母数检定方法 593

- 17.1 单一母群体中心值的无母数检定 / 594
- 17.2 两个母群体是否相同的无母数检定方法 / 599
- 17.3 多个母群体是否相同的无母数检定方法 / 614
- 17.4 两变项是否相关的无母数检定方法 / 618

附 录 一 627

- 一、 Σ 运算法则 / 628
- 二、阶乘与组合公式 / 630
- 三、二项式定理及多项式定理 / 631
- 四、泰勒展式 (Taylor's series) / 633
- 五、概率分配的特性及相互关系 / 635

附 录 二 645

- 表 A-1 标准常态表 / 646
- 表 A-2 t 分配概率表 / 647
- 表 A-3 χ^2 分配概率表 / 648
- 表 A-4 F 分配概率表 / 650
- 表 A-5 二项分配概率表 / 658
- 表 A-6 卜瓦松分配概率表 / 668
- 表 A-7 乱数表 / 673

参考文献 675

第1章 绪论

1.1 统计学与概率论

概率论的起源，不像苹果落在牛顿头上发现地心引力那么传奇，也不像阿基米德在浴缸中发现浮力原理那么浪漫。它的起源只是一个偶然。时间回溯到 17 世纪中叶，狂热赌徒狄梅内（C. De Mere）被赌注分配问题所困惑，遂求教于巴斯卡（B. Pascal），巴斯卡便与另一位数学家费马（P. Fermat）以排列组合（permutation and combination）来分析这个问题，他们对这个问题的探讨便是近代概率理论的开端。到了 19 世纪，古典的概率理论已不敷使用，促使概率论越过了机会游戏的框架而更上一层楼，柯莫葛洛夫（Kolmogorov）将概率论与现代集合论（set theory）、测度论（measure theory）等相结合，为概率论建立了公设化的基础。时至今日，在许多实际问题中，本身就包含了机会因素，也就是所谓的随机现象，而概率论便是讨论各种随机现象，并为它们建立理论架构。更重要的是，概率论是统计分析的基础。

统计学（statistics）一词来自于拉丁文 status，它的原始意义是“状态（state）”。自古以来，统计学一直被视为是以图、表及数据来描述一个国家的经济、人口、社会等资料的。时至今日，各国政府、研究机构及工商企业都有专责机构从事统计分析工作，依循统计方法的步骤（图 1-1-1），即搜集—汇整—解释—推论，来深入了解问题或现象的本质，进而有助于提升决策品质，其范围则涵盖了经济、管理、财务、金融、教育、社会福