

统 计 学 经 典 译 丛

律师统计学

(第2版)

Statistics for Lawyers
Second Edition

迈克尔·O·芬克尔斯坦 (Michael O. Finkelstein) 布鲁斯·莱文 (Bruce Levin) 著
钟卫 译 袁卫 校



D915/15

2008

统 计 学 经 典 译 丛

律师统计学

(第2版)

Statistics for Lawyers
Second Edition

迈克尔·O·芬克尔斯坦 (Michael O. Finkelstein) 布鲁斯·莱文 (Bruce Levin) 著
钟卫 译 袁卫 校



中国人民大学出版社

· 北京 ·

图书在版编目 (CIP) 数据

律师统计学/芬克尔斯坦, 莱文著; 钟卫译, 袁卫校. 2 版

北京: 中国人民大学出版社, 2008

(统计学经典译丛)

ISBN 978-7-300-09446-5

I. 律…

II. ①芬… ②莱… ③钟… ④袁…

III. 统计学-应用-诉讼法

IV. D915

中国版本图书馆 CIP 数据核字 (2008) 第 094934 号

统计学经典译丛

律师统计学 (第 2 版)

迈克尔·O·芬克尔斯坦 著

布鲁斯·莱文

钟卫 译

袁卫 校

出版发行 中国人民大学出版社

社 址 北京中关村大街 31 号

邮政编码 100080

电 话 010-62511242 (总编室)

010-62511398 (质管部)

010-82501766 (邮购部)

010-62514148 (门市部)

010-62515195 (发行公司)

010-62515275 (盗版举报)

网 址 <http://www.crup.com.cn>

<http://www.ttrnet.com>(人大教研网)

经 销 新华书店

印 刷 河北三河市新世纪印务有限公司

规 格 155 mm×235 mm 16 开本

版 次 2008 年 7 月第 1 版

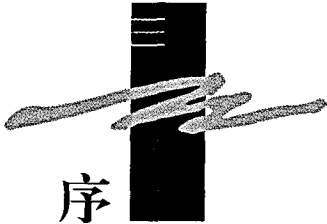
印 张 39.25 插页 1

印 次 2008 年 7 月第 1 次印刷

字 数 619 000

定 价 75.00 元

版权所有 侵权必究 印装差错 负责调换



总 序

统计学作为数据的科学，不仅在理、工、农、医和社会科学的经济、管理、教育、心理等各个领域都有着广泛而成功的应用，而且在纯文科的历史、语言和法律等领域也日益成为重要的工具。这一方面是由于人文与社会科学的多数学科越来越重视对数据的分析，强调定性与定量的对立统一，这体现为对统计学的需求不断增长；另一方面是由于统计方法的体系也在不断完善、创新，特别是与计算机技术的结合，使得统计方法的应用更加简捷方便，应用领域不断扩展，这体现为统计方法的应用和服务能力不断加强。



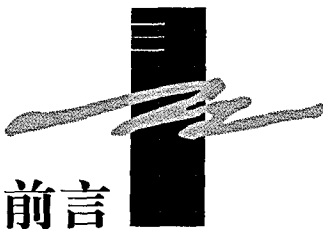
最近几年,一批国外优秀的统计教材和著作被翻译介绍给国内的读者,受到读者的喜爱和欢迎,同时也促进了我国高校统计教材的建设。纵观翻译出版的统计教材和图书,以经济管理类应用统计的教材居多。这既是我国现阶段经济管理教育,特别是 MBA、EMBA 教育火爆的体现,也反映出其他领域统计应用的相对落后。

中国人民大学出版社推出的这套统计译丛选择了国际著名出版公司 Springer 和 John Wiley & Sons 近年出版的统计教材和应用图书,特别是在工程、行为科学、法律和智能科学等领域的优秀统计教材与图书,其中有些领域在我国基本上还是空白。例如,统计方法在法律诉讼中的应用对我们就极有启发。丛书选择了两本法律应用的经典著作进行翻译,一本是《律师统计学》(第2版),另一本是《用数字证明——法律与诉讼中的实证方法》。前一本按照统计方法的体系,精选了100多个法律诉讼中的实际案例对统计方法的应用进行了深入浅出的介绍;后一本则侧重于从描述统计方法的角度介绍其在法律中的应用。其他几本书既有经过不断再版、不断教学实践修改完善的优秀教材,如《行为统计学基础》(第9版)、《工程统计学》(第3版)、《商务统计学》(第4版),也有对某个专题的介绍和讨论,如《智慧、基因与成功:钟形曲线与统计》等。

我们衷心地希望通过这套译丛的翻译出版能够更全面地反映整个统计学科,特别是应用统计学的现状,同时也衷心地祝愿这套译丛能够促进统计方法在我国更多领域的应用。

袁卫

2005年3月
中国人民大学静园



译者前言

几年前在查阅 Springer 的统计学出版物时，我们发现并买了一本《律师统计学》(*Statistics for Lawyers*)，看后觉得其中的案例很生动，就将这本书的部分案例应用于中国人民大学等学校的 MBA 和 EMBA 的教学中。学生们对于这样的案例很感兴趣，对统计方法也有了更深刻的理解。随后，应中国人民大学法学院的邀请，我们又专门给法学院的学生作了“统计方法在法律和法庭审判中的应用”的报告，受到了老师和学生的普遍欢迎。同时他们也提出了当下适合法律应用的统计学参考资料很少的问题。

在国外，将统计学应用于法律诉讼已有



几十年的历史，统计方法在法律领域日益成为重要的工具。这一方面是由于法律越来越重视对数据的分析，强调定性与定量的统一；另一方面是由于统计方法的体系也在不断完善、创新，使统计方法的应用和服务能力不断加强，在法律诉讼中起着越来越重要的作用。在国内，虽然统计方法在经济学、管理学、社会学、遗传学、流行病学等领域发挥着越来越重要的作用，但是把统计应用于法庭审判还是比较新的课题。

近年来，中国的司法改革和发展步伐不断加快，在司法实践中对证据不仅要进行认真的定性分析，也要进行数量的分析。将定量分析技术应用于法律领域已成为国内外法律工作者的共识。但是目前国内相关的书籍和参考文献还比较少，于是我们萌生了将这本书译成中文介绍给国内读者的想法。中国人民大学出版社贺耀敏社长非常支持这项工作，迅速和 Springer 联系版权等事宜，使得这本书最终能够顺利地出版并和读者见面。

《律师统计学》按照统计方法的体系，精选了大量法律诉讼中的实际案例，对统计方法的应用进行了深入浅出的介绍。本书的最大特点就是应用性强。它通过大量实际审判案例介绍统计方法，更有利于读者对方法的理解。本书的第二个特点是通俗易懂，案例生动。作者回避了烦琐的数学推导，采用深入浅出、循序渐进的案例介绍统计方法，即使没有概率和统计学基础的读者也能够看懂。

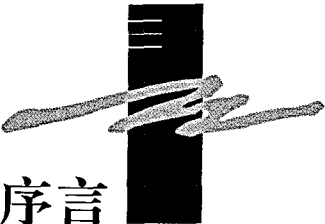
本书不仅可以作为高等院校应用统计学和法学相关专业的教材和参考文献，也可供统计和司法实际工作者参考。需要特别指出的是，用统计的方法、数量分析的方法在法庭作证时，必须十分严谨、慎重，统计与数据分析仅提供证据，无法代替法官或陪审团的裁决。

原书附录 2 中含有部分统计学常用的表，读者可以登录中国人民大学出版社工商管理分社网站 www.rdjg.com.cn 免费下载。

感谢中国人民大学出版社的编辑，他们的工作效率和认真的工作态度值得我们学习。在本书的翻译过程中，赵建斌、张俊岭、麻丽珍、王孟欣、吕忠伟、陈晓亮、张加刘等同学都付出了辛勤的劳动，在此一并表示感谢。

本书由钟卫翻译，袁卫负责组织并做全书的校稿。由于我们的专业水平和能力有限，书中的不妥及错误之处，敬请读者批评指正。

袁卫 钟卫
于中国人民大学



第二版序言

本书第一版出版后的 10 年中，法律领域对统计学证据的关注程度已经超过了法律的起源——民事权利法。第二版融入了该领域中许多新的方法，如 DNA 证据（2.1.1, 3.1.2 和 3.2.2），有毒物诉讼案件的流行病学研究（第 10 章），校正人口普查的统计模型（9.2.1），以及样本投票案例（13.2.3）。在一宗涉及有毒物质索赔的诉讼案——Daubert 案中，最高法院开创性地运用统计证据，其中流行病学的研究发挥了关键的作用，这标志着统计证据在科学证据的神圣殿堂中已占有十分重要的位置。本案中，最高法院不再以原来学术界普遍认可的 Frye 检

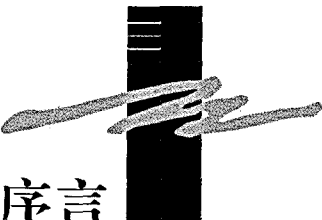


验作为通过资格审查的基础，而是明确地给陪审团一个全新的任务：他们必须评估提供的证据是否可靠相关。新的体系使法官处于一个十分尴尬的境地，不仅要懂得科学知识的前沿，还要站在证人的立场上去体会和判断一个专家所说的话的科学依据。幸运的是，大概在 Daubert 案一年后，即 1994 年，联邦司法中心出版并赠予所有法官一本《科学证据参考手册》，这本手册大概是统计方法运用的启蒙读物。2000 年，该书的新版本出人意料地成为畅销书。觉得本书很难懂的读者可以翻阅该书以获得有用的指导，至少在部分章节上如此。

但是新的、以案例为中心的统计学应用只是其发展中的一部分。从长远来看，更重要的在于司法体系自身对于统计研究的不断深入。这种研究能提供人们以洞察力，使他们有时甚至敢于怀疑平时看来非常神圣的司法机构。5.6.3 节给出了这样一个例子——对预期陪审员的绝对挑战，而作者通过分析数据发现其中绝大部分“绝对挑战”实际上仅为“猜测”而已。另外一个例子就是在本书写作之时，新闻媒介正热衷于报道由哥伦比亚大学法学院承担的一项涉及面甚广的课题——死刑的统计研究。该项研究为人们展示了一个令人震惊的现状——在一些能导致死刑判决的犯罪审判中，普遍存在一些严重失误。该研究几乎可以影响即将进行的立法工作，并有可能为时下关于罚金的辩论提供重要资料。读者应当注意到，在这两个研究中，正是在各案例细节分析中使用了统计分析方法才得出引人注目的结论。

和第一版一样，本版新增部分的很多素材来源于这些案例涉及的一些律师、统计学家以及经济学家，我们对他们给予我们工作的慷慨支持表示感谢。在这里，我们需要特别感谢 Orley Ashenfelter, David Baldus, William Fairley, David Freedman 以及 Sol Schreiber，他们提供了相关材料并对相关工作给予了指导。

迈克·O·芬克尔斯坦
布鲁斯·莱文



第一版序言

就法律的理性研究而言，刻板的教条主义者可能是对现在的人而言，而未来的人们将是精通统计学和经济学的大师。

——O. W. 福尔摩斯
法律之道（1897）

本书的目的是给律师和未来的律师们介绍一些在法律纠纷中的统计分析方法。全书在各章节中穿插着法律诉讼案例研究及数学说明。所研究内容包括概述一些真实案例（但不全是），这些概述以问题的形式提出，以期引起讨论。它们被用来阐明法律领域，而在这其中统计已经发挥了一定作用（至少

已经有此发展前景)；同时，它们也用来说明一系列进行量化推理的方法。另外，本书涉及的内容还包括对一些法律系统以及法律法规的影响的统计研究。我们尽可能都给出了数据摘录，以使读者能够清楚地以第一手资料进行统计计算和统计推断。本文没有给出司法观点，因为它们与本书主要关注的统计问题无关。另一方面，我们给出了一些司法上的失策，以使读者能够监督司法工作者，同时作为新生专家通过修正法律中的不足，来提高自身水平。

案例研究中的大部分数据的计算不需要读者有概率或统计知识。对于那些没有详细介绍的案例，在统计注释中提供了技术工具。统计注释部分用于说明基本文献中的问题（以相当精练的方式）；而其余部分超越了这个目标。对于更为随意的、详尽的，或者更为宽泛的讨论，读者可能想参看一些统计教程，本书部分章节以及参考文献给出了一些参考材料。关于案例研究中的一些数学计算问题我们在附录 1 中给出。其他不涉及计算的法律问题和统计问题大部分留给读者来讨论。

除去其中一些需用脑力研究的话题外，律师及未来的律师们可能会问，他们是否必须掌握这本书所教的这么多的统计学知识。当然，没必要全部知道。但对越来越多的法律学者、律师、法官甚至立法人员而言，熟悉统计思想能解释福尔摩斯的判断艺术，这不是一种义务（责任）而是一种需要。在多种研究领域，我们的知识需要用统计角度评价出来的数据来表示。那些在一般社会（大众）上普遍适用的准则已被法庭采用。各个研究领域的经济学家、社会学家、遗传学家、流行病学家和其他学者，都在各自精通的领域不断进行验证，并利用统计手段进行描述和推断。尤其是在经济学以及涉及经济的问题中，大量的数据、计算机的存在和当前的流行趋势已经推动了复杂经济计量模型的产生，而这些模型都发表在一些著名的期刊杂志上。但即使那些最完美、最经典的模型，也要依赖一些不可靠的假设，一旦进行检验，势必影响整个模型的有效性。

通常，诉讼中的统计问题不是由统计学家来阐述，而是由其他学科的专家、仅知道一些统计学知识的律师或法庭来阐述。这种实现方法的随意性使统计学与从法庭上所获得的其他专门知识产生了差别，同时也必然增加模型不当假设或其他简单错误的可能性。相对于对统计知之甚少、全靠咨询他人来做决策的人来说，那些对统计了解较多的律师们能

更有效地提出有价值的研究，更好地避免一些复杂模型中的严重错误，而且能把结论浅显易懂地告知决策者。尽管这些律师通常无须进行数学计算，但他们对于统计基本原则的领悟（这是必需的）的最佳来源就是深刻分析研究数据。

统计确实很重要吗？这个问题有时会惹恼统计学家。在法律背景下，问题在于在研究过程中是否给予统计模型客观的评价，以及在决策时是否给予统计结论足够的重视。遗憾的是，从现有记录来看，情况很不好，甚至相反。部分案件中法庭比较认同统计证据；但在一些很重要的公众诉讼案件中，一些很好的统计模型却往往不被法官接受（或非常差的模型反而被接受了）。90年前福尔摩斯预计的未来律师的情况，如今还未实现。

尽管在一些颇受关注的诉讼案中发生了一些无视统计证据的事件，但不可避免的是，基于数据的研究分析将继续被大多数学术机构所推崇，被越来越多的公众诉讼案所采纳。法庭对数据分析、统计推断的准则理解越透彻，对证据的研究就越精确、越具有分辨性。而且，人们希望统计工作的现状适时地影响人们对公正的理解，即使在法庭上也是如此。尽管一个有深厚统计学识的律师不会因为统计而舍弃法律基本准则，但对于研究对象的更多了解势必有助于理性、有学识等因素的积累，继而能认识事物的本质、不断获取知识。无论如何，这就是本书的最大目的。

本书的很多素材来源于一些相应的案例涉及的律师和统计学家。我们感谢他们慷慨大方地将自己的论文提供给我们。其中尤其要感谢 David Baldus, Jack Boger, Will Fairley, David Freedman, Elizabeth HOLTzman, Jay Kadane 和 Jack Weinstein。同时也要感谢 Joseph Fleiss, Mervyn Susser, Zena Stein, 以及他们各自所在研究所——生物统计处、哥伦比亚大学医药系 Sergievsk 中心、纽约州心理研究所。

我们还要特别感谢 Margaret Murphy，她快速录入手稿并不断修改；感谢 Marcia Schoen，她对计算部分进行了录入；还要感谢 Stephen Sullivan, Lynn Cushman 和 Matthew Herenstein，他们核对了引文部分；以及 Ann Kinney，她修改了案例研究部分并核对了计算部分。

衷心感激我们的家人——Claire, Katie, Matthew, Betty, Joby, Laura 和 Julie, 感谢他们在本书的长期写作期间的耐心、鼓励与支持。

迈克尔·O·芬克尔斯坦
布鲁斯·莱文

目 录



第 1 章 描述统计	(1)
1.1 描述统计导论	(1)
1.2 中心位置的测度	(3)
1.3 离散度的测度	(20)
1.4 相关性的测度	(31)
1.5 两个比例差异程度的测度	(40)
第 2 章 计数问题	(48)
2.1 排列与组合	(48)
2.2 波动理论	(60)
第 3 章 概率论基础	(64)
3.1 概率计算的基本原理	(64)
3.2 选择效应	(81)
3.3 贝叶斯定理	(83)



3.4	甄别机制和诊断检验	(89)
3.5	蒙特卡洛方法	(96)
3.6	概率基础	(102)
第4章	一些概率分布	(106)
4.1	概率分布简介	(106)
4.2	二项分布	(110)
4.3	正态分布和中心极限定理	(121)
4.4	统计假设检验	(128)
4.5	超几何分布	(132)
4.6	正态性检验	(144)
4.7	泊松分布	(152)
4.8	几何分布和指数分布	(159)
第5章	两个比例的统计推断	(165)
5.1	两个比例相等的 Fisher 精确检验	(165)
5.2	两个比例相等的卡方检验和 z 分数检验	(169)
5.3	比例的置信区间	(178)
5.4	假设检验中的统计功效	(196)
5.5	法律上的显著性和统计显著性	(203)
5.6	极大似然估计	(207)
第6章	多重比例的比较	(216)
6.1	用 χ^2 统计量检验拟合优度	(216)
6.2	邦弗朗尼不等式和多重比较	(229)
6.3	关联度的测度方法: ϕ^2 统计量和 τ_B 统计量	(234)
第7章	均值比较	(241)
7.1	学生氏 t 检验: 假设检验和置信区间	(241)
7.2	方差分析的多重比较	(251)
第8章	对独立层的数据合并	(258)
8.1	合并数据的 M-H 检验与 Fisher 检验	(258)
8.2	整合分析	(271)
第9章	抽样问题	(278)
9.1	随机抽样理论	(278)

9.2 捕获再捕获	(300)
第 10 章 流行病统计学	(306)
10.1 导论	(306)
10.2 特异危险度	(308)
10.3 流行病统计学的主要起因	(316)
第 11 章 生存分析	(345)
11.1 死亡密度函数、生存函数和危险率函数	(345)
11.2 比例危险模型	(353)
11.3 量化风险评估	(361)
第 12 章 非参数方法	(370)
12.1 符号检验	(370)
12.2 威尔科克森符号秩检验	(373)
12.3 威尔科克森秩和检验	(375)
12.4 斯皮尔曼等级相关系数	(381)
第 13 章 回归模型	(383)
13.1 多元回归模型简介	(383)
13.2 回归方程系数的估计和解释	(391)
13.3 回归方程的不确定性测度	(401)
13.4 回归系数的统计显著性	(407)
13.5 回归方程中的解释变量的选取	(412)
13.6 阅读多元回归的计算机输出内容	(419)
13.7 置信区间和预测区间	(432)
13.8 回归模型的假设	(438)
13.9 变量变换	(444)
第 14 章 复杂回归模型	(453)
14.1 时间序列	(453)
14.2 交互模型	(471)
14.3 就业歧视案中的备择模型	(478)
14.4 局部加权回归	(482)
14.5 雇佣回归模型中的无法识别偏差	(485)
14.6 方程组	(492)



14.7	Logit 和 probit 回归	(496)
14.8	泊松回归	(512)
14.9	刀切法、交叉验证法、自助法	(515)
附录 1 案例的计算和注释		(522)
参考文献		(601)