

统计学系列

统计学的世界

让你学会用手中的少量数据，
对重大问题做出明智的决策。

第五版
Fifth Edition

STATISTICS
Concepts and Controversies

[美] 戴维·S·穆尔 著
(David S. Moore)



中信出版社
CITIC PUBLISHING HOUSE

统计学系列

统计学的世界

让你学会用手中的少量数据，
对重大问题做出明智的决策。

第五版
Fifth Edition

STATISTICS
Concepts and Controversies

[美] 戴维·S·穆尔 著
郑惟厚 译

中信出版社
CITIC PUBLISHING HOUSE

图书在版编目 (CIP) 数据

统计学的世界/[美]穆尔著;郑惟厚译. —北京:中信出版社, 2003. 9

书名原文: Statistics: concepts and controversies

ISBN 7-80073-954-6

I. 统… II. ①穆… ②郑… III. 统计学 - 通俗读物 IV. C8-49

中国版本图书馆 CIP 数据核字(2003)第 084824 号

Statistics: concepts and controversies

© 2001, by W. H. Freeman and Company.

Simplified Chinese Translation copyright © 2003 by CITIC Publishing House

Published by arrangement with W. H. FREEMAN AND COMPANY, a division of BEDFORD.

FREEMAN AND W. H. FREEMAN AND COMPANY LLC.

ALL RIGHTS RESERVED.

本书译文由天下文化出版公司授权使用。

统计学的世界(第五版)

TONGJIXUE DE SHIJIE

著 者: [美]戴维·S·穆尔

译 者: 郑惟厚

责任编辑: 谢楠

出版发行: 中信出版社(北京市朝阳区东外大街亮马河南路14号塔园外交办公大楼 邮编 100600)

经 销 者: 中信联合发行有限公司

承 印 者: 中国电影出版社印刷厂

开 本: 787mm × 1092mm 1/16 印 张: 41.5 字 数: 543千字

版 次: 2003年11月第1版 印 次: 2003年11月第1次印刷

京权图字: 01-2003-5930

书 号: ISBN 7-80073-954-6/F·600

定 价: 79.00元

版权所有·侵权必究

凡购本社图书,如有缺页、倒页、脱页,由发行公司负责退换。服务热线:010-85322521

E-mail: sales@citicpub.com

010-85322522

写给教师

统计可以当作公共课程来教

《统计学的世界》是一本公共课程的统计(statistics)书，也就是说，是非数学系学生公共教育的一部分。这本书的内容，是从我在给普度大学(Purdue University)文学院一、二年级学生开课的经验中，累积发展出来的。很高兴的是其他许多教师也觉得，这本书对各领域的学生都有用，甚至对哲学系和药学系的学生也是如此。这本第五版经过了大幅度的修订，差不多可以算是本新书。不过目标仍和前几版相同，就是要把统计以不同的面貌呈现给大家，即不把统计当作专业的工具，而是当作受过教育的人应有的文化素养。

统计和文科的关系

统计在一般人的心目中，是最不“文科”的学科了。要是有人说统计的好话，多半是在说它的用处。医疗领域的专家必须有统计知识，才能了解医学研究报告的内容；经理人需要统计，因为大量的数字需要经过有效浓缩，才能看得出所以然；老百姓需要了解统计，才能知道民意调查和消费者物价指数(CPI, Consumer Price Index)到底是怎么回事。因为数据(data)和机遇(chance)无所不在，所以我们的广告词可以这样写：每个人都用得到统计，甚至因统计而获利多多。

以上说的是事实。我还可以进一步地说，对于大多数的学生来讲，这本书所持的强调“观念”，以及用言语来说明的方式，帮助学生为将来学习统计所打下的基础，要优于一般以“方法”为主的介绍方式。美国统计协会(American Statistical Association)和美国数学学会(Mathematical Association of American)的联合课程



委员会就曾建议，任何统计的入门课程，都应该“强调如何做统计思考”而且内容应该“多一些数据和观念，少一点公式和推导过程”。这本书就像这项建议所说的，它的内涵符合素质教育的精神，也就是有许多观念、许多思考，而只有简单的数据、少量的公式，且没有任何正式推导。

应该学习统计观念的另一个理由是：统计其实属于文科领域。人文教育强调基本的思维能力，也就是可以应用在各种情境的一般探索方法。传统的人文学科中呈现了这些方法，比如文学和历史研究、人类社会的政治分析及社会分析、探讨大自然的实验科学，以及数学领域中抽象概念及演绎的力量。统计可归类于文科领域，因为根据不确定的经验数据(empirical data)做推论，用的也是类似的一般思维方法。这本书的两个主题——数据及机遇，普遍存在于我们的日常经验当中。虽然我们数学工具来处理数据和机遇问题，但数学所提供的观念，却不是纯然数学的。事实上，心理学家已经提出具说服力的论点，认为若想要加强我们对日常生活中所遇到的数据或机遇现象做出合理推论的能力，精通数学并没有多大帮助。

这本书在作者的能力范围和读者可接受的限度下，提出的观点是：统计是一种独立且基本的思考方法。书的重点是在统计思考，也有人称为认识数(quantitative literacy)*。

* 译注：literacy 指认识字，quantitative literacy 就指认识数。

本书的类型

统计学的书有的专讲理论，有的专讲方法，但这本书两者都不是。《统计学的世界》讲的是统计概念、统计推理，还有统计与公共政策和包括从药学到社会学等人文科学的关系。我加入了许多基本的绘图及数值技巧，来落实概念，支持推理。学生通过处理数据，可以学到怎么样来看数据。不过我可没有让技巧超越观念变成主角。我打算用言语来教而不是用代数来教，鼓励讨论甚至争辩，而不是只教计算，当然有一些计算是必须的。而从目录就可以看得出，本书的涵盖范围比传统的统计教科书广得多。以素质教育的精神来看，我宁广不宁精。

虽然本书的内容不那么正经八百，但它仍是一本教科书。书的编排是针对有系统的学习所规划，书中有许多习题，而且其中



有不少是要求学生讨论或做判断。即使对于那些主动阅读本书且乐在其中的同学来说，也应该做做习题而不是只读正文。而教师们请小心，虽然本书所用到的数学层次很低，但它并不是一本轻松的教科书。本书重点放在观念和推理上，教起来反而会比满篇都是公式的教科书还辛苦。

前言

什么是统计

统计是用来处理数据的。数据由数字组成，但它不仅是单纯的数字而已。数据是有内容的数字。比如说，光是 10.5 这个数字本身并没什么含意，但是假如我们得知，一个朋友的新生儿出生时重 10.5 磅，我们会恭喜她生了个健康宝宝。我们根据数字，配合上下文并和常识衔接，就可以做出判断。我们知道 10.5 磅重的婴儿相当大，而且也知道婴儿不太可能重 10.5 英两或 10.5 公斤。数字加上上下文才能提供有效的信息。

统计是从数据中找出信息，并且做出结论。我们用的工具是图表和计算，加上常识判断。我们先来简单地快速看一下，媒体和热门政治或社会议题如何处理数据和统计研究，借此展开讨论统计的序幕。稍后我们会对这里提到的例子做更详尽的审视。

数据胜过轶闻

信仰不能取代数字

斯潘塞 (Henry Spencer)

轶闻引人注目，是因为它很突出，所以会深入人心。轶闻能使议题人性化，所以新闻报道常常以轶闻当开场或结尾。但它并不足以当成做决定的根据，而且正因为它们很突出，所以反而常常会产生误导。应该注意的是，声明的背后有没有数据支持，而不是有没有动人的故事。

住在高压电线附近会导致儿童得白血病吗？美国国家癌症研究所 (National Cancer Institute) 花费 5 年的时间和 500 万美元为这个问题搜集资料。结论是：在白血病和暴露在高压电线所产生的磁



场之间，找不到相关关系。在《新英格兰医学期刊》(*New England Journal of Medicine*)上和这篇研究报告同时登出的评论中严厉提出，“应该马上停止浪费我们的研究资源”在这个问题上。

现在试着比较以下两者的影响：一是电视上对于一项耗时5年，花费500万美元的调查结果的新闻报道，另一是电视访问一位能说会道的母亲，她的孩子得了白血病，而且他们恰巧住在高压电线附近。在大众心目中，每次都是轶闻得胜。但我们应该要心存疑问。数据比轶闻可靠，因为数据可以有系统地描绘出整体的情况，而轶闻只聚焦于少数特例。

我忍不住要加上一句“数据胜过自封的专家”。大部分媒体心目中的平衡报道，就是指在不同立场的两方面，各找一位“专家”来发表简短的评论。我们永远也没法知道，是否某位专家表达的是真正具有代表性的意见，而另一位只是照顾私人利益的骗子。而且由于媒体对于制造冲突的喜好，社会大众现在以为，对应每一位专家，必有立场相反的另一位专家。如果你真正关心某个议题，应该试着判断数据透露何种信息，以及数据的品质如何。当然！的确有许多议题悬而未决，但也有许多议题只在不重视证据的人心中才悬而未决。你至少可以了解一下这些“专家”的背景，还有他们所引用的研究结果，是否曾刊载在审编制度严谨的期刊上。

数据从何而来非常重要

数字不会说谎，但说谎的人会想出办法。

格罗夫纳(Charles Grosvenor)

数据是数字，而数字总显得牢不可破。其实有的如此，有的却不是。任何统计研究中最重要的事，就是数据的来源。当专栏作者蓝德丝(Ann Landers)问她的读者：如果重新来过的话是否仍要生孩子？而回答的人当中有70%坚决说“不要”的时候，你对于蓝德丝引述自泪迹斑斑的信中，投书人泣诉他们的孩子如何像野兽一样时，不妨一笑置之。蓝德丝从事的是娱乐业，她邀请读者回答这个问题时，回应最热烈的应该是后悔有了孩子的父母。大部分的父母并不后悔有小孩。我们知道这个事实，是因为曾经



有人对许多父母做过意见调查，而且为了不偏向任何一个答案，访问的父母是随机(at random)抽取的。民意调查当然不是没有瑕疵，这点我们以后会谈到的，但是比起邀请满肚子气的人来回答问题，民意调查显然高明多了。

即使是信誉卓著的期刊，也不一定能对坏数据免疫。《美国医学会期刊》(*Journal of the American Medical Association*)曾登过一篇文章，声称将冷却的液体经过管子打进胃里，可以缓解溃疡症状。病人的确对这种治疗有反应，但只是因为病人通常对自己信任的医师的权威信服，因此对他给的任何治疗都有反应。也就是说安慰剂*(placebo)发生了作用。后来终于有人起疑，做了有适当对照组(control)的研究，有些病人接受这项治疗，有些接受安慰剂，结果使用安慰剂这组的“表现”还稍好些呢。“没有比较，就没有结论”，是判断医学研究的一个很好的出发点。比如说，我对于最近突然盛行的“自然疗法”就有点存疑。这些疗法当中，极少会经过比较试验(comparative trial)，来证明它们不仅仅是装在有漂亮植物图片的瓶子里卖的安慰剂而已。

*译注：没有实际治疗作用的假治疗。

小心潜在变量

我的钱够我这辈子用了，只要我不买东西。

梅森(Jackie Mason)

有报道说，在美国，设有赌场的县犯罪率较高。有位大学老师说，在网上修课的学生，比在教室里修课的学生表现好。政府的报告强调，受教育多的民众，比起受教育少的民众，赚钱要多得多。最好不要太快做结论。要先问：“是不是有什么他们没有告诉我的，可以解释这些事？”

有赌场的县的确犯罪率较高，但是在城市或是较穷的县，犯罪率照样比较高。怎样的县会有赌场呢？这些县是否在设立赌场之前，就有高犯罪率了呢？网上修课的学生学得比较好，但是和在教室修课的学生比起来，这些学生年龄比较大，底子也比较好，因此表现好很正常。教育程度高的人的确钱赚得多。但受教育多的人和受教育少的人比起来，平均来说他们的父母受的教育较多，也比较有钱。他们在较好的环境长大，上较好的学校。这些有利



条件让他们得到更多的教育，而即使他们不受这么多教育，这些条件也可能帮他们赚很多钱。

以上提到的这些研究，都报告了两个变量之间的关联，并且引导我们推断出，其中一个变量影响了另一个变量的结论。“赌场会增加犯罪率”以及“想要有钱就多读点书”是他们传达的信息。这些信息有可能是正确的。但也说不定我们看到的关联，大部分可以由隐藏在背景中的一些变量解释，比如会接受赌场的县的本质，以及高教育程度者生来就具备的有利条件。好的统计研究会考虑许多的背景变量。这当然要有技巧，不过你至少可以找一找，看他们有没有做。

变异无所不在

当事实改变时，我就改变主意。您呢？

凯恩斯(John Maynard Keynes)

如果你的舌下温超过 37°C ，是不是就代表你在发烧呢？也许不是。每个人的“正常”体温都会有一些差异。你自己的体温也有变化，早上6点时稍高些，下午6点就较低。美国政府宣布上个月失业率上升0.1个百分点，而新开发的建筑项目下降了3个百分点。股市因此遽升或遽降。股市波动常常是不理性的。政府的数据是根据好的样本得到的，是好的估计，但并不是百分之百的事实。同样的调查再做一遍，结果可能会有一些差别。而和经济有关的事件原本就会上上下下，而影响它的因素包括气候、罢工事件、假日以及各式各样其他原因。

很多人都像股市一样，会对数据的一些微小的变动做过度反应，而事实上这些改变并不是实质的改变，根本微不足道。以下是全美国最大的市场研究公司其领导人尼尔森(Arthur Nielsen)的经验之谈：

太多商界人士都对所有白纸黑字印出来的数字都同样的信任。他们认为数字就代表“事实”，要他们用“概率”(probability)的观点来看事情有点困难。他们未能看清，数字只是一种简化的表示法，它描述一个范



围，这范围说明我们对潜在情况的实际了解有多少。

变异是无所不在的。个体之间有差别，而对同一个个体多度量几次，结果也会不一样，并且几乎每件事都会随着时间而不同。至于一些自封的专家对每天股市变化的深入分析，或者对一场其实是最后一秒才定输赢的球赛，却硬要归罪于输球的球队的能力或特征，你尽可以对这些评论嗤之以鼻。

结论并不是百分之百的

数学定律不能百分之百确实地用在现实生活里；能百分之百确实地用数学定律描述的，就不是现实生活。

爱因斯坦 (Albert Einstein)

因为变异无所不在，所以统计结论并不是绝对的。大部分中年妇女会定期做乳房 X 光摄影，以期能早期发现乳癌。但乳房摄影是否真的可以减低死于乳癌的风险？高品质的统计研究发现，对于 50—64 岁的妇女来说，乳房摄影可以减少 26% 的死亡率。但这个数字是该年龄层妇女的一个平均数字。因为个别差异到处都有，所以对于不同的妇女来说，结果可能大不相同。有些每年做乳房摄影的妇女死于乳癌，而有一辈子没做过乳房摄影的妇女，却活到 100 岁，最后是因为骑摩托车出车祸而死亡。

总结报告事实上说的是“乳房摄影可以将乳癌死亡率减少 26% [95% 置信区间 (confidence interval) 为 17%—34%]”。而 26% 这个数字，根据尼尔森的说明，是“某范围的简化说法，这范围描述我们对隐藏情况的实际了解有多少”。在此例当中，这个范围是 17%—34%，而我们有 95% 的信心，真正的比例会落在这个范围内。也就是说，我们相当有把握，但不完全确定。一旦你超越新闻报道的层次，可以找找看有没有诸如“95% 信心”及“有统计上的显著意义 (statistically significant)”等字眼，这代表这个研究的结果虽不能说百分之百确定，但已相当有把握。



数据可反映社会价值

要用统计骗人很容易。但是不用统计，骗人更容易。

莫斯提勒(Frederick Mosteller)

好的数据确实胜过轶闻。比起轶闻和光大声嚷嚷预测未来，数据要客观得多。统计和其他的公开论述比起来，它根据事实且较科学又较理性。对于争论性的议题，统计研究应该比其他大部分证据受到更多的重视。不过世界上没有百分之百客观这回事。社会环境会影响我们对于要度量什么及如何度量的决定，因此也就影响到统计。

以自杀率来说，在不同国家间就有很大的差别。各国公布的自杀率有差别，似乎大部分要归因于社会观念，而不是因为自杀率真的有差别。自杀人数的计算，是根据死亡证明书。填写死亡证明书(证明书上的细节，依不同的国家或不同的州而有别)的官员，对于比如说像没有目击证人的溺死或摔死事件，可以决定要追究的程度。在视自杀为耻辱的地方，就有较多的自杀被报道为意外死亡。比方说，在大部分人信奉天主教的国家，所公布的自杀率就比其他国家低。日本文化中，有遇到羞辱时就光荣自杀的传统，这种传统使人们较不视自杀为耻辱，就使得日本的自杀事件，被报道的比例较高。而在其他国家，可能因为社会价值观的改变，使得自杀事件的计数增加。一个愈来愈普遍的看法是，抑郁症应该被视为生理上的疾病，而不是个性上的弱点，而自杀是这种病的悲剧性结束，不是道德瑕疵。因此使得家属和医生更愿意把自杀列为死因。

在不像自杀这么敏感的问题上，社会价值观依然可能影响到数据。美国的失业率，是每个月由劳工统计局(Bureau of Labor Statistics)用一个很大且很专业的全国性样本算出来的。但是“失业”的定义是什么呢?它是说虽然你想工作，却没有工作，而且过去两周曾经积极找工作。如果你已有两周没在找工作，你就不算是失业，而只是“不属于劳动人口”。这样定义失业率，反映了我们加诸于工作的价值观。如果换个方式定义，可能会得到很不



一样的失业率。

我的意思，不是叫你不要相信失业率的数字。失业率的定义已经稳定了好一段时间，所以我们可以从它看出趋势。在各个国家之间，这个定义也大致吻合，所以我們也可以在国际间做比较。数据是由不受政治干扰的专业人士算出来的。失业率是重要且有用的信息。我要强调的是，并不是每件重要的事都可以只用数字表示，还有，把事情简化为数字的人，会受到各种压力的影响，不管有意识还是无意识。

统计学的世界

简明目录

| | |
|---------------------------|------------|
| 第一部分 产生数据 | 1 |
| 第 1 章 数据从何而来? | 2 |
| 第 2 章 好样本和坏样本 | 18 |
| 第 3 章 样本告诉我们什么? | 34 |
| 第 4 章 真实世界中的抽样调查 | 58 |
| 第 5 章 实验面面观 | 84 |
| 第 6 章 真实世界中的实验 | 106 |
| 第 7 章 数据伦理 | 128 |
| 第 8 章 度量 | 150 |
| 第 9 章 数字合不合理? | 174 |
| 第一部分 复习 | 192 |
| | |
| 第二部分 整合数据 | 205 |
| 第 10 章 好的图及坏的图 | 206 |
| 第 11 章 用图形呈现分布 | 232 |
| 第 12 章 用数字描述分布 | 254 |
| 第 13 章 正态分布 | 282 |
| 第 14 章 描述相关关系的方法:散布图和相关系数 | 306 |
| 第 15 章 描述相关关系:回归、预测及因果关系 | 330 |
| 第 16 章 消费者物价指数和政府统计 | 358 |
| 第二部分 复习 | 382 |
| | |
| 第三部分 机遇 | 399 |
| 第 17 章 考虑可能性 | 400 |
| 第 18 章 概率模型 | 420 |



| | | |
|-------------|-------------|------------|
| 第 19 章 | 模拟 | 436 |
| 第 20 章 | 赌场的优势：期望值 | 454 |
| 第三部分 | 复习 | 472 |
| 第四部分 | 推论 | 483 |
| 第 21 章 | 什么是置信区间 | 484 |
| 第 22 章 | 什么是显著性检验 | 504 |
| 第 23 章 | 统计推论的使用与滥用 | 522 |
| 第 24 章 | 双向表及卡方检验* | 540 |
| 第 25 章 | 有关总体平均数的推论* | 562 |
| 第四部分 | 复习 | 578 |

*本章内容较深，可列为选读。

统计学的世界

详细目录

| | | |
|------|--------------|----|
| 写给教师 | 统计可以当作公共课程来教 | I |
| 前言 | 什么是统计 | IV |
| 统计与你 | 这本书里谈些什么? | XI |

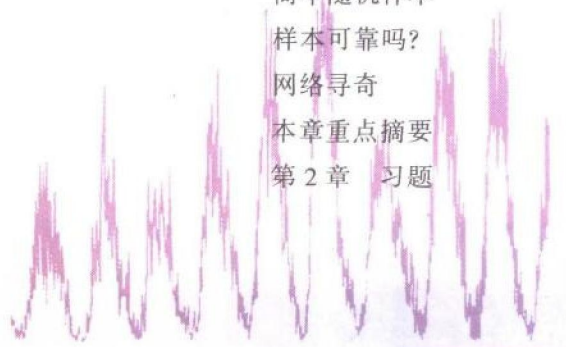
第一部分 产生数据 1

第1章 数据从何而来? 2

| | |
|--------------|----|
| 来说说数据: 个体和变量 | 3 |
| 观测研究 | 4 |
| 抽样调查 | 6 |
| 普查 | 9 |
| 实验 | 10 |
| 网络寻奇 | 12 |
| 本章重点摘要 | 13 |
| 第1章 习题 | 14 |

第2章 好样本和坏样本 18

| | |
|----------|----|
| 怎样可取得坏样本 | 19 |
| 简单随机样本 | 21 |
| 样本可靠吗? | 25 |
| 网络寻奇 | 26 |
| 本章重点摘要 | 27 |
| 第2章 习题 | 28 |





| | |
|-----------------------|-----------|
| 第3章 样本告诉我们什么? | 34 |
| 从样本到总体 | 35 |
| 抽样变异 | 36 |
| 误差界限 | 40 |
| 置信叙述 | 43 |
| 从大总体抽样 | 45 |
| 统计学上的争议：选举民意调查该禁止吗？ | 46 |
| 网络寻奇 | 47 |
| 本章重点摘要 | 48 |
| 第3章 习题 | 49 |
| | |
| 第4章 真实世界中的抽样调查 | 58 |
| 抽样调查怎样出错 | 59 |
| 抽样误差 | 60 |
| 非抽样误差 | 62 |
| 问题的措辞 | 65 |
| 如何应对非抽样误差 | 66 |
| 真实世界中的抽样设计 | 67 |
| 相信调查结果之前该问的问题 | 71 |
| 网络寻奇 | 72 |
| 本章重点摘要 | 73 |
| 第4章 习题 | 74 |
| | |
| 第5章 实验面面观 | 84 |
| 谈谈实验 | 85 |
| 怎么样做坏实验 | 86 |
| 随机化比较实验 | 89 |
| 实验设计的逻辑 | 91 |
| 统计显著性 | 93 |
| 只能观测的时候怎么办 | 93 |
| 网络寻奇 | 96 |
| 本章重点摘要 | 97 |
| 第5章 习题 | 98 |