

<<统计数字会撒谎>>

图书基本信息

书名：<<统计数字会撒谎>>

13位ISBN编号：9787507420852

10位ISBN编号：750742085X

出版时间：2009-3

出版时间：中国城市出版社

作者：[美] 达莱尔·哈夫

页数：154

译者：廖颖林

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

## &lt;&lt;统计数字会撒谎&gt;&gt;

## 前言

我的岳父从爱荷华州到加利福尼亚州不久便对我说：“你们这儿治安不好。”在他所阅读的关于加州的报道的确如此。

但是，这些报道通常来自一份爱荷华州的报纸。

这份报纸不会轻易忽略掉加州发生的任何犯罪行为，虽然它也报道本州的谋杀案，但看起来它更乐意大肆渲染加州出现的同类情况，而且还因此而闻名。

我岳父得出的这个结论是建立在明显有偏样本基础之上的，是一个随意的统计结论。

类似于其他更为精致的统计结论，也存在着证据与结果不匹配的问题，因为这些结论都在假定：报纸专栏中对犯罪行为的报道是测量犯罪率的工具。

几年前，十来个调查人员独立地发表了关于抗组织胺药的试验数据。

所有的数据都证明，在经过抗组织胺药物治疗后，相当高比例的感冒能够治愈。

这一结论引起了传媒的大肆宣传报道，抗组织胺药的广告铺天盖地，医药界也掀起了此类药物的生产热潮。

人们对健康永恒不变的追求造就了这种热潮，但奇怪的是，人们拒绝越过统计资料去注意一下早就了解的事实。

正如一位幽默的非医学权威人士，亨利·G·菲尔森(Henry G. Felsen)不久前所指出的——正确的治疗的确能在7天内治愈感冒，但是即使不进行任何治疗，感冒也能在一个星期内痊愈。

统计这种神秘的语言，在一个靠事实说话的社会里是如此地吸引眼球，但有时它却被人利用，并成为恶意夸大或简化事实、迷惑他人的工具。

在报告社会经济趋势、商业状况、民意调查和普查的大量数据时，统计方法或者统计术语是必不可少的。

但如果作者不能正确理解并恰当地使用这些统计语言，而读者又并不能真正了解这些术语的含义，那么，统计结果只能是废话一堆。

在科技主题的文章中，经常充斥着滥用统计资料的现象，每一个数据都很有力地打消你的疑虑，都张着嘴告诉你这是对的，但是人们很难联想到这样一种场景：灯光如豆的实验室中，“白衣英雄们”仍在不计报酬、不辞辛劳地包装这些数据。

统计，就像最神奇的化妆术，只要略施粉黛，东施尤胜西施。

巧妙伪装的统计资料胜过希特勒的弥天大谎，虽然它也会引起人们的误解，但制造它的人却能巧妙脱身。

这本书是一本如何利用统计瞒天过海的入门读物。

看上去，它很像骗子的行骗宝典。

但或许我可以模仿一个已退休的窃贼——如何神不知、鬼不觉地撬开一把锁，他的回忆录达到了研究生课程的水平——替这本书说句公道话：毕竟，骗子对于行骗的技巧早已胸有成竹，而诚实的人出于自卫也应该掌握它。

## <<统计数字会撒谎>>

### 内容概要

“自从使用了某某牌牙膏，我们的蛀牙减少了23%”；“1924级的耶鲁毕业生平均年收入是25111美元”；真相果真如此？

数字的魔力决不仅止于此！

“根据相同的收入取样，计算方法也完全正确，同一个小区居民的年平均收入可以是2000英镑，也能是10000英镑”；只需施展一点小技巧，就能让数据面目全非！

你是否执着于打听智力测验的结果？

你是否对电视上那些专家所言的各种标准坚信不疑？

你是否认为抽烟与成绩的好坏直接相关？

抑或受高等教育与晚婚有必然联系？

你是否确信眼见为实？

如果图形能让事实的1：2瞬间变为眼前的1：8呢？

你还确信自己的感觉吗？

这本“故事书”般的经典之作将为您轻松揭开谜底……作为揭露数据造假经典工具书，达莱尔·哈夫这本《统计数字会撒谎》在50年代一出版，就成为有些人的“心头大患”，他们在全世界对其进行围追堵截，因为他们相当害怕读者看到书中的真相。

它诞生之后，不仅被奉为“世界统计学普及读物第一畅销书”、“25本投资经典之一”、“经典社会学读本”，而且被誉为“数据造假曝光宝典”，在世界畅销50年而不衰。

达莱尔·哈夫用风趣的插图和通俗的语言把高深的统计写得像“故事书”一样精彩，这种独具一格的行文和编排使得此书更具魅力。

此外，书中还给你揭露了大量至今仍被销售员、专家、记者或者广告撰稿人频频使用的“行骗工具”：有偏的样本，可以挑选的平均数，遗漏某些重要的数据，混淆因果关系，滥用一维图形等。

更为重要的是，达莱尔·哈夫在书中破解了如何识破虚假统计数据。

当你面对虚假但貌似真实的统计数据时，谁说的？

他是如何知道的？

遗漏了什么？

是否有人偷换了概念？

这个资料有意义么？

达莱尔·哈夫提出的这个“五步法”，堪称识破数据花言巧语的“武功秘籍”，它可以让您：知悉惯用的“行骗技巧”洞悉统计数据的真实面目 轻松理财、管理属下、彻底读懂经济报道 不再被形形色色的数据所忽悠，不再让自己的决策南辕北辙

## <<统计数字会撒谎>>

### 作者简介

达莱尔·哈夫 (Darrell Huff)

美国统计专家，于1913年出生在美国爱荷华州，毕业于爱荷华州立大学 (the State University of Iowa)，获得学士学位和硕士学位，在此期间他由于成绩优异加入了美国大学优等生的荣誉学会 (Phi Beta Kappa)，同时还参加了社会心理学、统计学以及智力测验等研究项目。

达莱尔·哈夫的文章多见于《哈泼斯》、《星期六邮报》、《时尚先生》以及《纽约时报》等美国顶尖媒体。

1963年，由于他的贡献被授予国家学院钟奖 (National School Bell Award)

## &lt;&lt;统计数字会撒谎&gt;&gt;

## 书籍目录

Chapter 1内在有偏的样本我们来看一则以前的新闻报道，“1924级的耶鲁毕业生平均年收入为25111美元。

”好家伙，这些人干得真不赖！

大家都把孩子送去耶鲁和牛津吧，那就是高薪的代名词，年纪轻轻就能享受幸福的生活。

可是，等一等，在充满怀疑的惊鸿一瞥后，关于该数字的两个疑点凸现出来：它居然惊人的精确，它也大得令人难以置信。

Chapter 2精心挑选的平均数房地产商费尽心思大力宣传，你所居住的小区是高档小区，你邻居的年均收入是10000英镑；而该区域的纳税人委员会却反复向政府强调，这里居民的平均年收入只有2000英镑，是该减减税赋了。

到底谁在说谎？

事实上，他们都是诚实的。

两个数字都是正规的平均数，来自相同的居民，根据相同的收入，计算方法也完全正确，奥妙何在？

Chapter 3没有披露的数据“自从使用了多克斯牌牙膏，我们的蛀牙减少了23%。

”让我再引用一篇文章的标题——《现在就来预测孩子将来长多高》。

“只需要利用现有的身高，再查表中的比例即可。

”标题如是写道。

没有比这个“23%”和“身高表”更无聊的广告和报道了！

可他们看起来是如此的煞有介事，问题出在哪？

Chapter 4毫无意义的工作彼得和琳达做了公认最好的智力测验——斯坦福-比内测验，结果是琳达的智商为101，彼得的只有98。

专家告诉我们智商的平均数是100，即100意味着“正常”。

于是进一步推断，琳达是比较聪明的孩子，彼得是个笨孩子。

相信我，任何类似的结论纯粹都是胡说。

Chapter 5令人惊奇的图形一张图告诉你朴实的10%的增长，而另一张却看起来是让人振奋的100%的增长，别怀疑你的眼睛，截然不同的两幅图说的可是同一回事！

Chapter 6一维图形的滥用数字是2<sup>1</sup>，但视觉效果却是8<sup>1</sup>。

嘴上说的是125倍，看起来却是3倍……或许你正在被这些图形所振奋着。

Chapter 7不完全匹配的资料一篇来自著名实验室的报告：在11秒钟内仅仅半盎司该药的剂量就杀死了试管中31108个病菌。

随处可见某种榨汁机的广告：“经过实验室的证明”该榨汁机的“榨汁功能增强了26%。

”……听起来真不错，这是货真价实的“挂羊头卖狗肉”。

Chapter 8相关关系的误解抽烟与大学成绩；独身与上大学；身上的跳蚤与健康；房屋顶上白鹤鸟巢的个数与荷兰某个家庭中已出生孩子的人数；马萨诸塞州长老教会会长的收入与哈瓦那朗姆酒的价格…

…是风牛马不相及？

还是亦步亦趋、息息相关？

真实的结果一定让你大跌眼镜。

Chapter 9如何进行统计操纵怎样在一年内获得22500美元的总收入？

你只需1个妻子（或丈夫）和13个孩子。

“现在就购买圣诞礼物，你将节省100%。

”精明的商家居然成了圣诞老人，开始免费馈赠了。

有些人很害怕你知晓其中的“技巧”。

Chapter 10如何反驳统计资料谁说的？

他是如何知道的？

遗漏了什么？

是否有人偷换了概念？

## <<统计数字会撒谎>>

这个资料有意义吗？

提这5个问题，凭双眼就识破并揭穿虚假的统计资料；更为重要的是在具有欺骗性的数据海洋中找出可靠有用的资料，不再让你的关键决策南辕北辙。

## &lt;&lt;统计数字会撒谎&gt;&gt;

## 章节摘录

第一章 内在有偏的样本 有一个装着红、白两色豆子的桶，如果你想要准确知道这个桶中两种豆子的数量，你惟一能文摘一：这个平均年收入精确得如此吓人。让我们来研究一则以前的新闻报道：“1924级的耶鲁毕业生平均年收入为25111美元”，要知道几年前的钱是更值钱的。

好家伙，他们干得真不赖！

可是，等一等，这个令人印象深刻的数字到底意味着什么？

是否像表面看到的那样，足以证明如果你把你的男孩送进耶鲁大学，或者牛津大学，那么在年老时你就不需要辛苦地上班，甚至他将来年老时也不用上班？

在充满怀疑的惊鸿一瞥后，关于该数字的两个疑点凸现出来：它惊人的精确；它大得令人难以置信。

对一群相隔千山万水的人，了解他们的平均收入，而且竟然精确到以元为单位，这几乎不太可能。就算是自己去年的收入，除非全部来自薪水，否则也很难知道得如此准确。

但是对于年收入25000美元的阶层来说，投资渠道更广，因此他们的收入不可能完全来自于薪水。

而且毫无疑问，这个可爱的平均数出自耶鲁人之口。

即使1924年他们在纽海文（纽海文是美国东北部康涅狄格州的一个城市，耶鲁大学就坐落在这个城市）。

接受过优良的教育，也很难保证四分之一世纪后，他们还能坚持说真话。

当问及收入时，有些人出于虚荣或者天生乐观而夸大数据；有些人却故意缩小数字，特别当涉及征所得税问题时，往往会犹豫不决，生怕与其他文件填报的数据不符，谁知道税务员又看到了什么？

这两种趋势——夸大与缩小，也许将相互抵消，但这种可能性很小。

一般而言，一种趋势总会强于另一种，但我们很难猜测哪种趋势将胜出。

我们试着来解释这个数字，单凭常识就知道这个数字与现实出入很大。

现在，让我们找找最大误差的可能来源。

是什么使那些实际上收入也许只有25111美元一半的人们最终会拥有如此丰厚的平均收入？

让我们来揭开这神秘的面纱。

可以肯定的是：耶鲁毕业生的报道基于对某个样本的分析，因为常识告诉我们，没有人能够掌握所有仍在世的1924级学生的情况，25年后，他们中的许多人已经消失在茫茫人海中。

并且，在那些能够取得联系的人中，许多人根本不会回答问卷，特别是一个涉及隐私的问卷。

一般情况下，邮寄问卷的回收率达到5%~10%就已经相当可观了。

也许这个调查的回收率会高些，但也不可能达到100%。

因此，这个收入数据建立在一个样本之上：由能够取得联系并愿意回答问卷的耶鲁学生组成。

那么，这个样本的代表性强吗？

也就是说，能否假设样本与未被样本包括的那些人——无法联系的人或者不愿意回答的人——具有同等的收入水平？

那些在耶鲁大学毕业生通讯录上被注明“地址不祥”的迷路小羔羊是谁呢？

他们是高收入阶层吗？

华尔街的金融家、公司领导层，亦或是制造企业或公用事业的总裁？

不，要找到富人的地址根本不难。

这个班级最显赫的人，即使忽略了与校友办公室联系，他们的地址也可以通过查《美国名人录》

（Who's Who in America）或其他参考资料找到。

因此，我们可以较合理地推测，那些被遗漏的人是获取耶鲁文学学士以后25年间没能实现自己光辉梦想的人，他们是小职员、技工、流浪汉、失业的酒鬼、仅仅得以糊口的作家或艺术家……将六七个甚至更多这种人的收入相加才可能达到25111美元。

他们不会在班级的联谊会上注册，仅仅是因为他们支付不起路费。

又是谁会将调查问卷丢进最近的废纸篓？

## &lt;&lt;统计数字会撒谎&gt;&gt;

我们不太肯定，但是猜想他们中大部分人并没有赚到足以炫耀的数目。

很明显，样本遗漏了可能降低平均收入的两类人。

让我们见识一下25111美元的庐山真面目：如果它是一个真实的数据，它也仅仅代表了1924级耶鲁学生中能够联系上的，并愿意站出来说出收入的一个特殊群体。

当然，它的真实性还需要满足这个假定：这些绅士们说的都是真话。

我们能否过于轻率地做出这样的假定呢？

来自抽样理论的一个分支，即市场研究的经验告诉我们，人们会说真话的假定往往是不可靠的。

文摘二：我有两个平均年收入，你想听哪个？

我相信你不是一个势利小人，而我也并不做房地产生意。

但请让我们作这样的假定，并且假设，此刻你正在一条我熟知的街上看房子。

对你的情况进行了初步判断后，我巧舌如簧、费尽心思地让你相信附近居民的平均年收入大约有10000英镑。

也许这坚定了你要在此居住的信心，不管怎样，买卖最终成交了，那美妙的数字也被牢记在你的脑海

而且，既然你已经买下了房子——你有那么一点势利，当与朋友聊天时，你就会不经意地流露出你居住的地点：我住在一个相当棒的高收入小区。

一年左右过后，我们又见面了。

作为某纳税者委员会的成员，我正在四处奔走，为降低税率、降低财产估价，或降低公共交通费用而呼吁。

我的理由很简单，我们支付不起各种上涨的费用，毕竟，附近居民的平均年收入只有2000英镑。

也许你会加入到我们委员会的工作中来——你不仅势利，而且还挺吝啬。

但是，当听到那可恨的2000英镑时，你也禁不住大吃一惊。

到底是我现在撒谎了呢？

还是一年前撒了谎？

其实这两次你都无法怪罪于我，利用统计撒谎的妙处被展现得淋漓尽致。

无论是10000英镑，还是2000英镑，它们都是正规的平均数，计算方法也完全正确。

两个数字都基于相同的数据，来自相同的居民，根据相同的收入。

所有都是相同的，但显然其中有一个数据令人误解，足以与弥天大谎相媲美。

我的花招就是两次分别使用了不同的平均数，“平均数”这个词宽泛的涵义帮了大忙。

当一个家伙希望用数据影响公众观点，或者向其他人推销广告版面，平均数便是一个经常被使用的伎俩，虽然偶尔是出于无心，但更多的时候是明知故犯。

所以，当你被告知某个数是平均数时，除非能说出它的具体种类——均值，中位数，还是众数，否则你对它的具体涵义仍知之甚少。

在这个例子中，不合适的“平均数”实际上是毫无意义的，只要碰到关于收入的数据，这种情况就经常出现。

在我卖给你房子附近的居民区里，这两个平均数相差甚远，问题出在哪？

因为收入是显著偏斜的。

你的邻居中大多数都是小农、在附近村庄上班的工薪阶层或是靠养老金为生的退休老人，但却有3户邻居是来度周末的百万富翁，就是这3户人家的收入提高了总收入，相应抬高了算术平均数。

这样一来，均值达到了绝大多数家庭遥不可及的水平。

几乎每个人都低于平均数，虽然这听起来像是笑话或者文学修辞，但却是事实。

文摘三：现在广告：本产品能有效杀死31108个细菌。

胡扯。

如果你想证明某事，却发现没有能力办到，那么就试着解释其他相关事情，并假装它们是一回事

在统计资料与人类思维冲撞所引起的耀眼光芒中，几乎没有人会发现它们的区别。

不完全匹配的资料是一种保证你处在有利位置上的武器，而且屡试不爽。

## &lt;&lt;统计数字会撒谎&gt;&gt;

你无法证明你的秘方能够治疗感冒，但你却可以在报纸上（用大字标题）刊登一篇得到保证的实验室报告，“在11秒内仅仅半盎司的该药剂量就杀死了试管中31108个细菌”。

当你这样做时，首先要确定这是一家著名的实验室，或者至少有一个让人印象深刻的名字，然后将这个报告全文复印下来，并在旁边放上一个穿着白大褂医生的照片。

但记住千万别提这里面的小把戏，而且你也没有责任指出：在试管中很有效的抗菌剂，在人的咽喉里就不会发挥作用，特别是为了防止该药灼烧喉咙而按照说明书进行稀释后。

也不要为报道杀死了哪些细菌而迷惑，谁会知道哪种细菌引起了感冒？

说不定感冒根本与细菌无关。

实际上，不管到底是什么引起了感冒，试管内各种各样的细菌与感冒的制造者之间的关系仍是未知数。

但大家并不会深究，特别是正在流鼻涕时。

也许这个例子太明显了，大家很容易识破它。

然而不完全匹配的资料通常不会以这种面目出现在广告中，下面，让我们欣赏一个更高明的手法。

假设在美国种族歧视情绪抬头的某个时期，你却受雇“证明”相反的结论——这并不是件难事。

你可以自己进行一个民意调查，或者更好的做法是聘请一家有名望的调查机构替你调查。

对总体中普通的横截面样本进行提问，看他们“是否认为黑人与白人有相同的就业机会”，你可以不断的重新调查，以确保得到一个你满意的“人们关于种族问题的报告”。

普林斯顿民意调查研究所（Princeton's Office of Public Opinion Research）曾经测试过同样的问题。

调查结果能够成为以下观点的有趣证据：事情总是表里不一，在民意调查中尤甚。

每个被访者在被问及关于就业机会的问题时，还回答了其他的问题，这些问题旨在了解被调查者真实的种族歧视态度。

调查结果表明：那些种族歧视情绪最强烈的人中，大部分的人对就业机会问题的回答为“是”。

（实际上黑人的同情者中三分之二的人并不认为黑人与白人有同等的机会，而种族歧视者中三分之二的人认为就业机会相同。

）很明显，虽然你根据这个调查得到了一些关于人们种族歧视态度的有趣结论，但是对于黑人的就业情况你仍然知之甚少。

于是你将发现：在调查期间如果种族歧视情绪高昂，将有更多的人认为黑人与白人就业机会相同。

于是你宣布结论：调查显示，黑人一直受到公平的对待。

在巧妙运用不完全匹配的资料后，你得到了出色的结论。

如同上例，实际情况越糟，你的调查却越能让它看上去令人欣慰。

## <<统计数字会撒谎>>

### 媒体关注与评论

哈夫先生用如此生动的、充满人情味的方式来论述统计这个干巴巴的课题，真是一剂灵丹妙药...我们太需要这本书了，它虽然娱乐性强浅易读，却十分具有说服力。

——美国著名管理杂志 《管理评论》作者和制图者倾注了全力，给大家提供了一本十分轻松活泼的读物和卡通画。

它们能给你带来娱乐，又能引发思考，而且还揭穿了许多统计方法的谎言。

——美国著名图书杂志 《图书期刊》这是一本善意破坏性的书，读完它后，你对于‘万能统计’的信任将大大降低。

——美国权威政论杂志 《大西洋》有3种谎言：谎言、糟糕透顶的谎言和统计资料。

——英国前首相本杰明·迪斯雷利（Disraeli） 对于追求效率的公民而言，统计思维总有一天会和读写能力一样必要。

——美国著名科幻作家H·G·威尔斯（H.G.Wells） 使我们陷入麻烦的通常并非我们不知道的事情，而是那些我们知道得不确切的事情。

——美国作家阿蒂默斯·沃德（Artemus Ward） 我需要完成一个很大的课题——统计学，但却感到我的写作功底十分有限，如果不牺牲准确性和完整性，就很难让人理解。

——英国著名科学家相关系数“r”创立者弗朗西斯·高尔顿（Sir Francis Galton）

## &lt;&lt;统计数字会撒谎&gt;&gt;

## 编辑推荐

揭秘最隐蔽的数据造假方法，畅销美国50年的伟大时代经典，世界统计学普及读物第一畅销书  
像“故事书”一样生动精彩，充满了“马克吐温”式的幽默讽刺案例 “25本投资经典之一”、“经典社会学读本” 商业人士必知的数据造假权威读本，公开统计专家才明白的专业统计知识  
莞尔一笑中让你知晓深奥的统计学基本原理，练就揭露“虚假数据”的火眼金睛 假数据也可以披上羊的外衣！

房价，工资，增长率，利润率……你都有可能被这些统计数字狠狠地忽悠一把！

然而，假的真不了！

翻翻报纸，《房价数据满天飞 市民抱怨看不懂》《“平均工资水平”有多少可信性？

》、《长春白领月薪1500？

标准太低数据存疑》、《平均工资越高越令人恐惧》……类似于这样的质疑越来越多，我们要如何才能看清这些漫天飞舞的数据背后的谎言？

《统计数字会撒谎》正是这样一本能让你辨清真假数据的首选必读书。

《统计数字会撒谎》一书是由美国统计专家达莱尔·哈夫写就的传世之作，该书引发的“编造虚假信息”话题受到美国社会持续普遍的关注和美国权威媒体的激烈争论。

它自50年代出版以来就不断一版再版，并被翻译成多种文字，在世界的影响力持久不衰，被誉为美国商业人士、研修人员的重要入门必修书之一。

书里面大胆地揭露了至今仍然被销售员、广告撰稿人、记者甚至专家频频使用的大量的统计操纵技巧，同时还配有别具一格的风趣插图以及众多幽默的案例。

神秘的统计学在这里被哈夫像讲故事一样一一道来，莞尔一笑中让你知晓深奥的统计学基本原理，掌握揭露“虚假数据”的最有力武器…… 桑普拉斯曾说过，统计分析让一切假象原形毕露。

但遗憾的是，统计未必能够揭示真实，有时候还可能成为假象的帮凶。

当我们面对生活中形形色色的统计数据的时候，还要多保持一些理智和清醒，并要有所保留地看待问题。

因为“如果一个人以种种肯定的立论开始，他必将终止于各种怀疑；但如果他愿意抱着怀疑的态度开始，那么他必将获得肯定的结论。

”（培根语）

<<统计数字会撒谎>>

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>