

统计之美



[统计之美_下载链接1](#)

著者:李舰

出版者:电子工业出版社

出版时间:2019-3

装帧:平装

isbn:9787121354045

《统计之美：人工智能时代的科学思维》每一章节都通过某个历史经典故事、网络热点事件、行业真实案例等背景知识引出对知识点的介绍，涵盖了经典的统计学模型和前沿的数据科学技术，并结合大数据实践领域的经验，构造了一套以实际应用为导向的知识体系，帮助读者轻松地掌握各分析方法的思想精髓，并能快速地应用到实际的工作中去。

作者简介:

李舰，统计学本科、软件工程硕士、商学博士，数据科学领域的多年从业者，见证并参与了统计学从冷门专业到显学的过程。是“统计之都”社区的核心成员之一、开源社区的活跃贡献者，致力于数据科学在实体行业中的应用。

海恩，在硅谷著名互联网企业从事数据分析工作，业余时间亦喜欢自己抓数据做点好玩的分析。

目录: 第1章 统计与科学1

1.1 随机的世界 2

1.1.1 打台球的物理学家，确定、不确定与随机2

1.1.2 上帝掷骰子吗，决定论与随机性4

1.1.3 连环杀手的归案，随机与均匀 6

1.1.4 扔硬币的数学家们，大数定律9

1.2 认识概率 11

1.2.1 游戏奖金的分配，概率论的发展 11

1.2.2 6连号和14连号，概率的计算 13

1.2.3 主持人背后的山羊，条件概率 16

1.2.4 寻找失踪的核潜艇，贝叶斯概率18

1.3 统计思想和模型 20

1.3.1 女士品茶，假设检验 20

1.3.2 “渣男”去死，最大似然 23

1.3.3 六西格玛的奇迹，小概率 25

1.3.4 牛顿的苹果，模型拟合27

1.4 统计与科学 29

1.4.1 智多星与神机军师，统计预测29

1.4.2 深蓝与阿尔法狗，归纳和演绎 31

1.4.3 中药与西药，临床试验33

1.4.4 所有模型都是错的，科学和否定 36

第2章 数据与数学38

2.1 数据与空间 39

2.1.1 多维世界的虫子，坐标和向量 39

2.1.2 黑客帝国和变形金刚，矩阵简介 42

2.1.3 九章算术与线性方程，线性方程组 45

2.1.4 二十八宿与黄道十二宫，线性变换 48

2.2 随机变量和分布 51

2.2.1 伯努利的硬币，随机分布 51

2.2.2 相亲多少次与神奇的37，离散型分布 54

2.2.3 棣莫弗的正态，连续型分布 56

2.2.4 醉鬼的步伐，随机过程 58

2.3 认识数据 59

2.3.1 忒修斯之船，数据、测量与变量 59

2.3.2 从性别到体重，数据的尺度 61

2.3.3 周岁与虚岁，连续变量与离散变量 63

2.3.4 一份体检记录，数据分析的基本数据结构 65

2.4 数理统计基础 66

2.4.1 管中窥豹与一叶知秋，总体和样本 66

2.4.2 恶赌鬼的诡计，数字特征 68

2.4.3 被平均的工资，统计量 70

2.4.4 小李飞刀与孔雀翎，参数估计	72
第3章 数据可视化	76
3.1 历史上的统计图形	77
3.1.1 河图与洛书，可视化简介	77
3.1.2 伦敦霍乱的防治，空间可视化	78
3.1.3 南丁格尔的玫瑰，玫瑰花瓣图	79
3.1.4 拿破仑远征，自定义统计图形	81
3.2 数据与可视化	83
3.2.1 女王的裙子，数据可视化	83
3.2.2 画布与宣纸，图形设备	84
3.2.3 深水王子与针眼画师，绘图语言的变迁	86
3.2.4 “挑战者”号航天飞机，直观的可视化	88
3.3 基础统计图形	90
3.3.1 老忠实喷泉的秘密，分布图	90
3.3.2 统计图形的奠基人，条形图和饼图	91
3.3.3 古老国度的诗云，坐标变换	93
3.3.4 飞翔的动态气泡图，动态可视化	95
3.4 数据之间的关系	97
3.4.1 东上相的轨道，散点图	97
3.4.2 五十州的最高峰，箱线图	99
3.4.3 泰坦尼克号的幸存者，马赛克图	100
3.4.4 切尔诺夫的笑脸，样本关系的可视化	102
第4章 模型与方法	104
4.1 常用统计模型	105
4.1.1 穿杨与射雕，回归模型	105
4.1.2 降维攻击，主成分分析	108
4.1.3 顾客就是上帝，路径模型	110
4.1.4 股票的走势，时间序列	114
4.2 机器学习	116
4.2.1 啤酒和尿布的传说，关联规则	116
4.2.2 寻找“白富美”，聚类分析	118
4.2.3 宁可错杀与绝不放过，分类效果评估	121
4.2.4 树木与森林，常用分类算法	124
4.3 人工智能	128
4.3.1 人工智能的三起两落，AI 发展史	128
4.3.2 深度学习的前生今世，深度学习简史	130
4.3.3 神秘的神经，神经网络简介	132
4.3.4 美丽的滤镜，卷积神经网络与深度学习	135
4.4 其他分析方法	139
4.4.1 茶、酒与百事可乐，随机试验方法	139
4.4.2 蒙特卡罗和原子弹，蒙特卡罗方法	142
4.4.3 医生的笔迹，文本分析	143
4.4.4 沙漠里的飞碟，最优化方法	146
第5章 大数据时代	149
5.1 技术的变迁	150
5.1.1 统计学的滥觞统计学简介	150
5.1.2 信息时代的来临计算机科学简介	151
5.1.3 数据挖掘和商业智能，商业智能简介	153
5.1.4 大数据时代新纪元，大数据简介	154
5.2 分析工具	156
5.2.1 谁说菜鸟不会数据分析Excel 简介	156
5.2.2 群雄逐鹿的分析软件统计软件和BI 系统	158
5.2.3 全栈工程师的最爱，Python 简介	160
5.2.4 本书作者最爱的R，R语言简介	162

- 5.3 计算框架 164
 - 5.3.1 冰箱里的大象，可扩容的数据分析164
 - 5.3.2 将兵与将将，并行计算 165
 - 5.3.3 电老虎和电蚂蚁，大型机和云计算 167
 - 5.3.4 摩尔定律的未来GPU，计算框架170
- 5.4 大数据行业应用 172
 - 5.4.1 互联网的兴起，互联网概览 172
 - 5.4.2 流量的起点，搜索引擎173
 - 5.4.3 收入的来源，精准广告 175
 - 5.4.4 猜你喜欢和投其所好，推荐系统 177
- 第6章 数据的陷阱180
 - 6.1 一叶障目 181
 - 6.1.1 神奇的天蝎座，规律的背后 181
 - 6.1.2 赢家的诅咒，悖论与分布182
 - 6.1.3 打飞机的油价，选择性关注184
 - 6.1.4 和女神的缘分，频率与巧合 185
 - 6.2 相关与因果 187
 - 6.2.1 芳华与热饮，遗漏的关键变量187
 - 6.2.2 热帖的秘密，不存在的相关 188
 - 6.2.3 雪与火的城市，地理决定的因果 189
 - 6.2.4 名字很重要吗，背后的关键因素 190
 - 6.3 样本和调查192
 - 6.3.1 测不准的美国大选，选择性抽样 192
 - 6.3.2 不对称的杜蕾斯，数据无反应偏差194
 - 6.3.3 幸运儿的传奇，幸存者偏差195
 - 6.3.4 哈佛校长的辞职，样本方差的影响197
 - 6.4 图形的误导 198
 - 6.4.1 收入的变化，被掩盖的数据 198
 - 6.4.2 收费站与汽车站，视觉的误区200
 - 6.4.3 东莞的逃亡，隐含信息的误导201
 - 6.4.4 有毒的拟合，图形与模型203
 - 参考文献207
 - 索引211
 - • • • • (收起)

[统计之美_下载链接1](#)

标签

人工智能

统计

统计学

科普

数学

数据挖掘

思维

2019

评论

文笔差，故事、背景知识不会讲。没学过统计的看不懂，学过统计的又觉得讲得不到位。引用几十个中文维基百科的定义，绝对无法忍受！第74页说区间估计是1934年提出来的，第150页说区间估计是1928年提出的。逗谁？

前言很赞，大部分案例很好，完全避开公式不可能，扣一分。

作者谙熟统计和科学史，观察全面，思考深入，文笔深厚。著作本身是难得的统计科普佳作，填补了市场空白，应载入中文统计类书籍史册。只是全书戛然而止，总有意犹未尽之感，也令人格外期待作者的未来之作。

不够深入

讲故事，没什么新意。

冲前言给作者打五星，文章干货少了点。
倒是激发了我几个灵感，1.演绎和归纳两种思维范式在人类历史上交替领先，各领风骚。随着人们对随机性的认识，现在是归纳法的代表量子力学，大数据和人工智能占上风的时代。2.中国传统的实用主义科学和哲学，预示着在这轮人工智能大潮中崛起的必然

。3.海森堡的测不准原理，难道不是信息量在量子尺度的一种表现吗？

内容宽泛，写的不如吴军老师那本数学之美好

就是讲故事

讲故事性质的科普书 适合入门看看---好久前看过的了 其实现在记不得写些啥了:)

2019#182nd。比较全面的一本统计学入门科普，章节编排循序渐进，由数学原理（概率论）出发，一路深入到数据可视化、人工智能及大数据领域，都有着不错的讲解，与故事案例的配合相得益彰。其中开头部分对科学中存在的归纳及演绎的分析令人印象深刻，体现了作者深厚的科学及人文素养。

序言写得low到地心。毫无人文素养。

没看到美..算是一本统计相关的文章集锦吧。各方面的边边角角都说了一点，看起来也还算轻松

像是很多书拼凑起来的，某些章节还不错

入门还是可以的，特别是文科生。

不知道评分为什么这么低，每篇附带的案例非常有趣，全书把统计学原理串讲的也符合逻辑。

挺好，讲的很生动，作者有功力

前面部分还行，后面都是历史介绍了

3.5。根本没展现出“美”。内容还算扎实，但基本和教科书没两样

用故事讲统计、人工智能和故事，有些觉得太浅，有些又觉得不适合，过深。适合入门期看看，涨涨兴趣

案例都很不错，但是作者讲的比较差，看的稀里糊涂的。

[统计之美_下载链接1](#)

书评

代友转，以下是他发在票圈的书评一则：
用了几天时间，读完了《统计之美—人工智能时代的科学思维》，收获不小。
这是本统计学的入门书籍，是今年数据科学科普书中的爆款，今年3月出版，据说已经第4次印刷，对于非统计学专业又对大数据和人工智能感兴趣的同学们来说很值得读一..
.

当时没有认真了解就买了这本书，读了一章，发表下看法：东拼西凑，完全是用书名讨
噱头！一点都不美。想说的东西太多，没一个讲明白！这样的东西就是写出来骗钱的！
大家还是要睁大眼睛看清楚再买！数学想绕过公式是不太可能的！请作者写下一本书用
点心！如果是真有能力就表达出...

这本书主要从基础统计学、大数据时代、数据的陷阱三方面入手，结合互联网时代背景
，没有局限于通用统计学教材的绝对专业，也给出了时代背景下的大数据应用方法及分

析工具。
另外，也用一章的篇幅讲述了在应用统计及数据分析的过程中可能存在的误导信息。
此外，这本书还有一个...

[统计之美_下载链接1](#)