

“

统计学学习地图与计算机应用

”

数据挖掘，大数据，网络存储，数据库管理，人工智能，机器学习，模式识别，数据可视化等技术的基础学科。



大话 统计学

陶文贤 陈静枝 著

清华大学出版社



大话 统计学

陈文贤 陈静枝 著

清华大学出版社
北京

内 容 简 介

“统计学”是兼具“数学计算”与“图形显示”的课程,所有的统计软件(如 SPSS),并非计算机辅助教学(CAI),因为它们并非“教你学会统计”,而是应该在“学会了统计”以后,再来用它。本书就是这样一本让你从零开始接触统计学,并将其真正应用到工作中的一本书,稳步跟进大数据时代。

本书前后连贯,各章之间也是先后呼应。例如:从概率到抽样,从描述到推断,从检验到因果;每章也是连贯的,开头有引言、观念图,结尾有流程图、思维导图;书中有许多阶层图、分类图、关联图、步骤图、流程图,以及因果表、比较表、决策法则表等。

本书专门的配套软件(中文统计)是在 Excel (2003~2016 版本适用)环境下,安装一个“加载项”,输入统计资料,就可以得到统计结果。“中文统计”可以公开下载,仅提供给合法取得本书之读者使用。

本书适合所有想掌握统计学的读者,也可以作为高校教材,由于内容比较多,教师可自行选择教学内容。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

大话统计学/陈文贤,陈静枝著. — 北京:清华大学出版社,2016

ISBN 978-7-302-45018-4

I. ①大… II. ①陈… ②陈… III. ①统计学—基本知识 IV. ①C8

中国版本图书馆 CIP 数据核字(2016)第 218501 号

责任编辑:栾大成

装帧设计:杨玉芳

责任校对:徐俊伟

责任印制:何 芊

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:清华大学印刷厂

经 销:全国新华书店

开 本:188mm×260mm

印 张:26.5

字 数:535 千字

版 次:2016 年 12 月第 1 版

印 次:2016 年 12 月第 1 次印刷

印 数:1~4000

定 价:59.00 元

产品编号:070745-01



百闻不如一见，兵难隄度，臣愿驰至金城，图上方略。

——《汉书·赵充国传》

因为计算机的普及，管理学院和工学院有些课程，在教材编写与教学方法方面，有结构性的改变。尤其是有关数学计算或图形显示的课程。这些课程应该附有应用软件的使用，配合教学。

“统计学”就是有“数学计算”与“图形显示”的课程。本书作者 1990 年在中国台湾出版的《统计学》一书，就附有“中文统计”软件，以前是在 DOS 操作系统下操作，现在则是在 Excel 上安装“加载项”操作。“中文统计”软件仅被提供给教师和购买本书的读者使用。

所有的统计软件（包括 SPSS），并非计算机辅助教学（CAI），因为它们并非“教你学会统计”。而是应该在“学会了统计”以后，再来用它。有了统计软件以后，大多数的统计问题，只要会选择统计功能，输入数据，就可以得到答案。但是使用统计软件要注意下列应用：

- (1) 能够判断统计数据，应该（适合）用哪一种统计方法（模型）。
- (2) 了解并解释统计方法计算结果（答案）的意义。

学习要有地图，学习地图或者知识地图，地图告诉你：你在这里（here）、为什么在这里（why）、要往哪里去（where）、如何去（how）、会得到什么（what）。其中 why 是假定条件（为何适用这个方法 how），而 where 是目的，what 是答案。

1980 年，Wonnacott 统计学第一章的开头引用了一句话：“He uses statistics as a drunken man uses lampposts – for support rather than for illumination.”

“人们利用统计，就好像醉汉利用路灯，是为了支撑，而不是照明。”

还有一个醉汉与路灯的故事：一个醉汉在夜晚的路灯下找东西，有个路人问他在找什么，醉汉说：“钱包。”路人就帮他找，两个人找了很久，但就是找不到。路人问：“你确

定是掉在‘这里’吗？”醉汉说：“我不知道掉在‘哪里’。”路人问：“为什么要在‘这里’找？”醉汉说：“因为‘这里’有路灯比较亮。”

利用统计学时，要注意假定条件是否符合，不要削足适履，不要因为“这个”方法，比较熟悉、比较容易用，就要用它来找答案，结果找到的答案根本不对。

本书取名《大话统计学》是清华大学出版社责任编辑栾大成先生的建议，“大话统计学”的意思不是说大话的统计学，而是让读者可以“大声说话”的统计学。因为本书的图画很多，所以本书也是“大画统计学”。

本书特色：

(1) 本书前后连贯，有前言、总论、结语。结语有：统计问题分类、统计概念复习。

(2) 各章之间也是先后呼应。例如：从概率到抽样，从描述到推断，从检验到因果。

(3) 每章也是连贯的：开头有引言、观念图，结尾有流程图、思维导图。

(4) 书中有许多阶层图、分类图、关联图、步骤图、流程图，以及因果表、比较表、决策法则表等，所以本书希望为统计学的学习地图。

(5) “中文统计”软件是在 Excel (2003 ~ 2016 版本适用) 环境下，安装一个“加载项”，输入统计资料、就可以得到统计结果。“中文统计”可以公开下载，仅提供给合法取得本书的读者使用。

(6) 中文统计的功能列表，配合本书章节设计。输入原始数据（观测值），可做描述或推断统计的计算。如果没有原始数据，只有样本容量、样本平均数、方差、比例等数据，那么使用“快速估计”或“快速检验”，可以得到推断结果。

(7) 为了节省篇幅，每章的部分例题、习题，放在互联网中，以便读者下载。

(8) 每章的部分案例讨论也放在互联网中，可提供学生做学期报告。这些案例的数据也可以下载，以方便学生使用。

(9) 补充教材如：分组资料描述统计、自力 (Bootstrap) 估计法、非参数统计补充、多因素方差分析、多元回归等，也放在互联网中。

(10) 因为配合计算机程序，包括 Excel 函数和命令，所以，所有的统计公式和计算步骤，都被很清楚地一一列举出来。由于强调应用导向，所以多数公式没有证明。

有了地图和交通工具（计算机软件），就可以快速地到达目的地。但是，如果一路直达目的地，就会错过沿途美妙的风光。所以，初学者还是要先走过一趟（自己计算了解过程），再利用交通工具（计算机），检查结果的正确性。

本书内容很多，所以教师可自行选择教学内容。感谢协助本书编辑及计算机软件的中国台湾大学资管所研究生。

感谢协助编撰本书及计算机软件的台大资管所研究生，使本书能够顺利出版。

由于作者的水平有限，本书中难免有许多错误和疏漏之处，恳请各位学者专家和读者，提出批评和建议，以便进一步修订和改进。

陈文贤 陈静枝 谨识

2016 年 9 月于台北

资源下载

1. 微云下载

- 例题
- Excel 文档
- 习题
- 补充教材
- “中文统计” 安装程序（随时更新）

<http://share.weiyun.com/6989e41582e4c3018d11972acb5b7392>



2. 百度云下载教学资源（要输入百度账号登录）

- PPT
- 习题解答

<http://pan.baidu.com/disk/home#list/path=%2F>



“中文统计” 安装说明

(1) 请到清华大学出版社 www.tup.com.cn 或 cnstat.weebly.com 下载“中文统计.exe”。

(2) 先执行“中文统计”安装程序。

然后，单击“Next >” → “Next >” → “Install” → “Finish”，完成“中文统计”安装，如图 A-1 和图 A-2 所示。

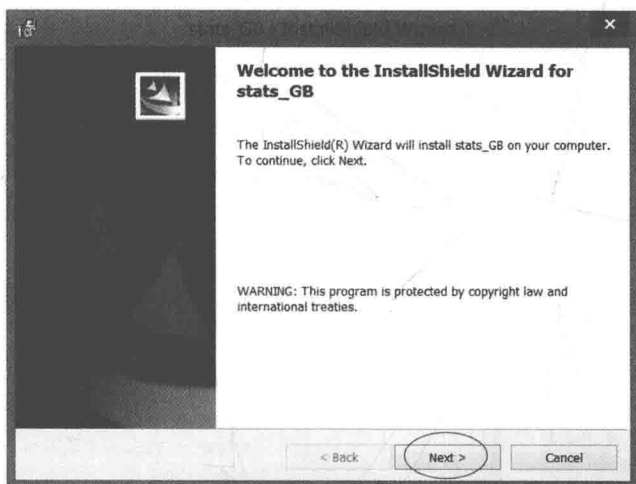


图 A-1 安装操作 1

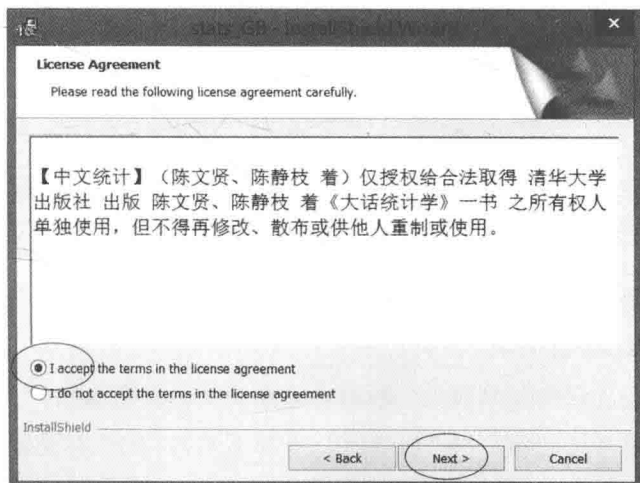


图 A-2 安装操作 2

(3) 请打开 Microsoft Excel (2013, 2016), Excel 2010 以下版本也可适用, 界面不同, 操作如图 A-3 所示。

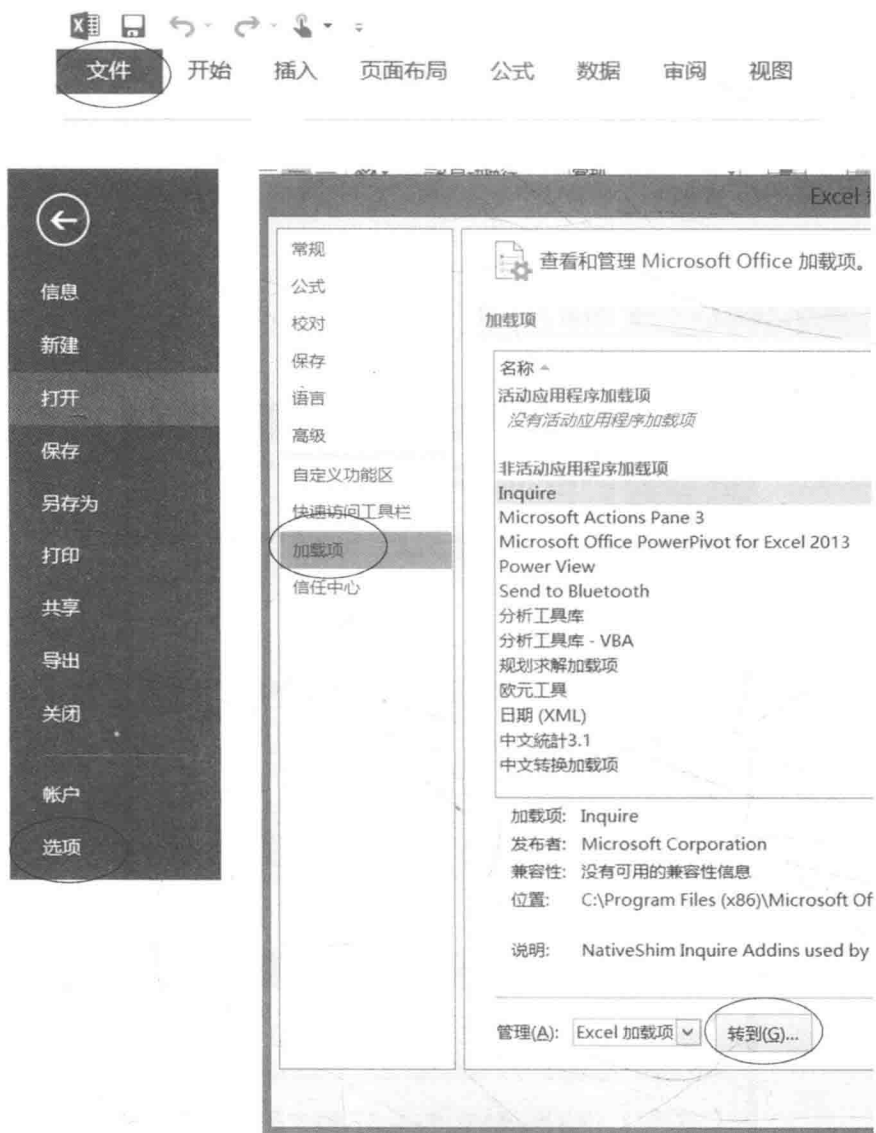


图 A-3 安装操作 3

(4) 然后, 弹出“加载宏”对话框, 勾选“中文统计 3.1”, 单击“确定”, 如图 A-4 所示, 然后出现图 A-5 ~ 图 A-7。

(5) 如果“可用加载宏”列表框中没有出现中文统计, 那么在“文件”→“选项”→“信任中心”→“信任中心设置”→“加载项”→都不要打勾→“宏设置”→单击“启用所有宏”。关闭 Excel, 再重新打开 Excel。

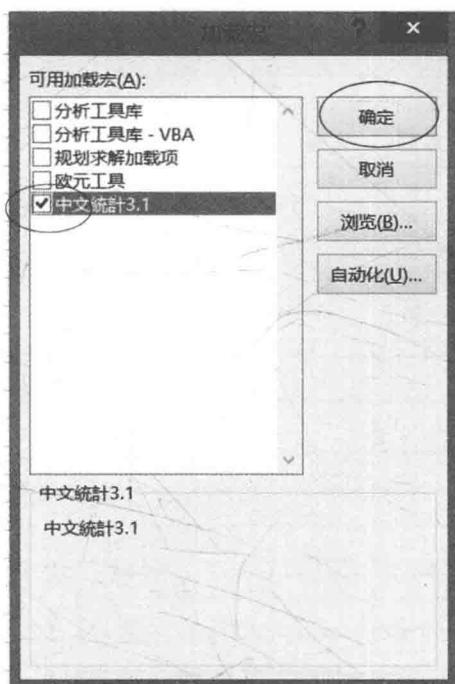


图 A-4 安装操作 4

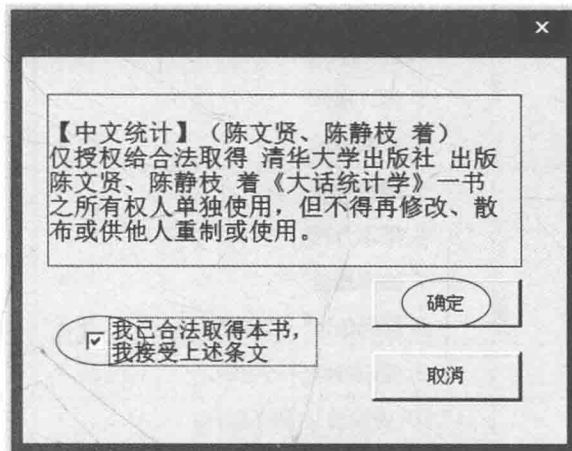


图 A-5 安装操作 5

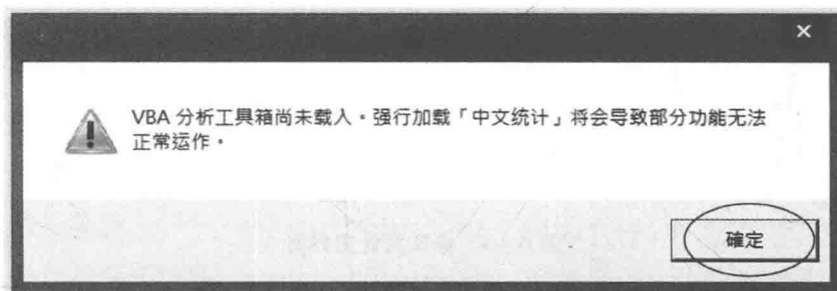


图 A-6 安装操作 6

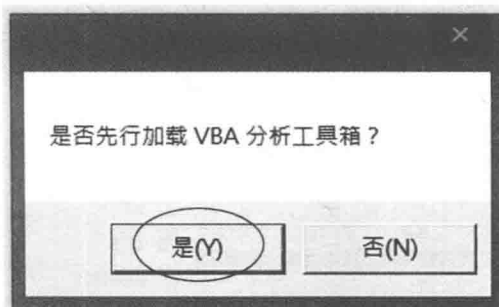


图 A-7 安装操作 7

(6) Excel 的菜单栏中出现“加载项”，单击“加载项”，就有“中文统计”菜单(选单)，如图 A-8 所示。



图 A-8 中文统计主界面

如果菜单栏中没出现“加载项”，则关闭 Excel，再重新打开 Excel，就会出现“加载项”，此时可以开始使用“中文统计”。

中文统计需要用到 Excel 的“数据分析”，请检查是否已安装“数据分析”，如图 A-9 所示。



图 A-9 Excel 数据分析



目 录

前言	i
“中文统计” 安装说明	v
 第 1 章 总论	1
1.1 统计是什么	3
1.2 统计学的基本概念	5
1.3 统计学的分类	8
1.4 抽样误差	10
1.5 统计数据的收集	12
1.6 变量与数据的衡量尺度	13
1.7 数据的类型	15
1.8 因果关系	16
1.9 统计的应用步骤	17
1.10 本章流程图	20
1.11 本章思维导图	21
习 题	22
 第 2 章 描述统计	25
2.1 描述统计	27
2.2 统计表	27
2.3 统计图	29
2.4 总量指标与相对指标	38
2.5 平均指标集中趋势量数	39

2.6	相对位置量数	46
2.7	离差量数	50
2.8	形态量数	55
2.9	中文统计应用	59
2.10	本章流程图	66
2.11	本章思维导图	67
	习 题	68
 第3章 时间序列		69
3.1	时间序列的分类	71
3.2	水平分析与速度分析	72
3.3	时间序列构成因素	74
3.4	平稳型序列预测	76
3.5	趋势型序列预测	79
3.6	季节指数分析	84
3.7	时间序列预测方法：趋势加季节	87
3.8	时间序列预测方法：趋势乘季节	88
3.9	预测误差	90
3.10	中文统计应用	92
3.11	本章流程图	98
3.12	本章思维导图	99
	习 题	100
 第4章 统计指数		103
4.1	指数的意义与分类	105
4.2	总指数的编制	109
4.3	指数的性质	115
4.4	指数体系与因素分析	119
4.5	指数应用	123
4.6	中文统计应用	123
4.7	本章流程图	125

4.8 本章思维导图	126
习 题	127
第5章 概率理论	129
5.1 试验与样本空间	131
5.2 事件概率	133
5.3 排列组合的公式	136
5.4 事件概率的计算	140
5.5 条件概率	142
5.6 独立事件与互斥事件	146
5.7 贝叶斯公式	150
5.8 中文统计应用	153
5.9 本章流程图	155
5.10 本章思维导图	156
习 题	157
第6章 随机变量	159
6.1 随机变量	161
6.2 概率分布函数与概率密度函数	163
6.3 期望与方差	166
6.4 双随机变量	169
6.5 中文统计应用	177
6.6 本章流程图	179
6.7 本章思维导图	180
习 题	181
第7章 概率分布	183
7.1 离散型随机变量的概率分布	185
7.2 连续型随机变量的概率分布	196
7.3 正态分布概率的计算	205
7.4 中文统计应用	207

7.5 本章思维导图	209
7.6 本章流程图	210
习 题	211
第8章 抽样理论	213
8.1 随机抽样	215
8.2 统计量	216
8.3 抽样平均与抽样方差的概率分布	217
8.4 中心极限定理	221
8.5 分层抽样	223
8.6 整群抽样	224
8.7 系统抽样	225
8.8 中文统计应用	226
8.9 本章流程图	229
8.10 本章思维导图	230
习 题	231
第9章 统计估计	233
9.1 估计量	235
9.2 正态分布平均数与方差的点估计	237
9.3 总体平均数的区间估计	237
9.4 总体比例的区间估计	240
9.5 总体方差的区间估计	240
9.6 抽样的样本量	241
9.7 标准误差	243
9.8 中文统计应用	244
9.9 本章流程图	248
9.10 本章思维导图	249
习 题	250

第 10 章 统计检验	251
10.1 假设检验	253
10.2 计算第一类错误与第二类错误	256
10.3 假设检验的步骤与方法	261
10.4 假设检验的样本量	263
10.5 总体平均数检验, 方差已知	265
10.6 总体平均数检验, 方差未知	266
10.7 总体比例检验	267
10.8 总体方差检验	268
10.9 中文统计应用	268
10.10 本章流程图	272
10.11 本章思维导图	273
习 题	273
第 11 章 两总体估计检验	275
11.1 因果与相关	277
11.2 两个总体参数的区间估计	278
11.3 两个总体平均数检验, 方差已知	282
11.4 两个总体平均数检验, 方差未知但相等	283
11.5 两个总体平均数检验, 方差未知且不等	284
11.6 两个总体平均数检验, 样本是配对数据	285
11.7 两个总体方差检验	286
11.8 两个总体比例检验	287
11.9 中文统计应用	288
11.10 本章流程图	290
11.11 本章思维导图	291
习 题	292
第 12 章 方差分析	295
12.1 方差分析介绍	297
12.2 单因素方差分析, 样本量相等	299

12.3	单因素方差分析, 样本量不等	304
12.4	多重比较法	305
12.5	检验方差是否相等	306
12.6	参数估计	307
12.7	双因素方差分析, 无交互作用	308
12.8	中文统计应用	311
12.9	本章流程图	314
12.10	本章思维导图	315
	习 题	316
第 13 章 回归与相关分析		317
13.1	回归与相关分析的区别	319
13.2	数学符号与关系式	321
13.3	一元线性回归分析参数的点估计	322
13.4	相关分析	324
13.5	检验自变量与因变量是否线性相关	328
13.6	回归与相关分析参数的区间估计与检验	329
13.7	中文统计应用	336
13.8	本章流程图	337
13.9	本章思维导图	338
	习 题	339
第 14 章 分类数据分析		341
14.1	卡方检验	343
14.2	多项分布卡方检验	343
14.3	拟合优度检验, 分布的参数已知	345
14.4	拟合优度检验, 分布的参数未知	347
14.5	卡方检验独立性与同构性	350
14.6	中文统计应用	355
14.7	本章流程图	357
14.8	本章思维导图	358
	习 题	358

第 15 章 非参数统计分析	361
15.1 非参数统计分析	363
15.2 符号检验	365
15.3 符号秩检验	369
15.4 游程检验, 检验随机性	371
15.5 Mann - Whitney 检验	373
15.6 Kruskal - Wallis 检验	375
15.7 Friedman 检验	376
15.8 Spearman 秩相关系数	377
15.9 中文统计应用	380
15.10 本章流程图	384
15.11 本章思维导图	388
习 题	389
第 16 章 结语	391
16.1 统计问题分类	393
16.2 误差名词说明	396
16.3 参数与统计量	398
16.4 统计概念复习	399
附录一 正态分布概率表	403
参考书目	405