

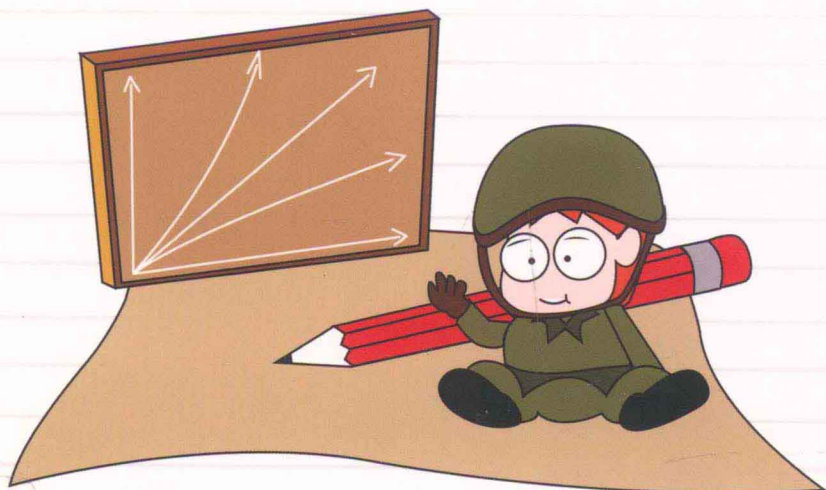
用Excel学习系列丛书



用Excel学 统计学

〔日〕田久浩志 著 冉剑玲 等 译

- 足球比赛的得分差是怎么预测出来的？
- 约会时，女人的高跟鞋多高说明对你有意思？
- 你的理想体重是多少？
- 你生病后，适合住哪家医院？
- 回归分析是一种怎样的分析工具？



科学出版社

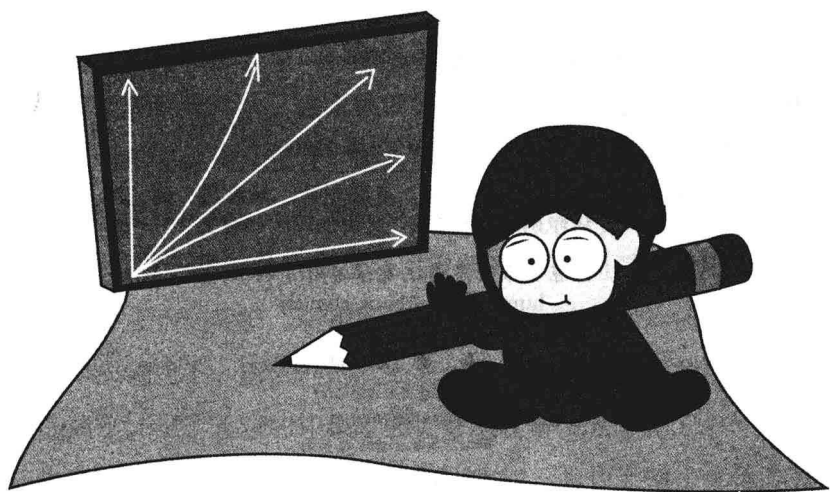
用Excel学习系列丛书



用Excel学

统计学

〔日〕田久浩志 著 冉剑玲 等 译



科学出版社

北京

图字：01-2011-7441号

内 容 简 介

你是不是正在学习统计学知识？是不是正在为一个项目做统计工作？是不是正为如何更好地做统计而头痛不已？那么，对你来说，这本书再合适不过了。它是将 Excel 与统计学完美结合的一本书，两者的结合可以让你熟练运用 Excel 知识解决统计学问题。只要你跟着作者的思路走，肯定能在最短的时间内学好统计学知识并漂亮地完成统计工作！

有趣的知识结合、细致的内容讲解定能给你留下深刻的印象。无论你是学生、上班族还是有一家属于自己的公司，活学活用统计学定会为你的学习、工作与生活增添更多便利。

图书在版编目(CIP)数据

用 Excel 学统计学 / (日) 田久浩志著；冉剑玲等译. —北京：科学出版社，2014

(用 Excel 学习系列丛书)

ISBN 978-7-03-038969-5

I. 用… II. 田… III. 表处理软件—应用—统计学
IV. C8-39

中国版本图书馆 CIP 数据核字 (2013) 第 253908 号

责任编辑：徐莹 唐璐 / 责任制作：刘素霞 魏谨

责任印制：赵德静 / 封面设计：郝建宝

北京东方科龙图文有限公司 制作

<http://www.okbook.com.cn>

科学出版社 出版

北京东黄城根北街16号

邮政编码：100717

<http://www.sciencep.com>

安泰印刷厂印刷

科学出版社发行 各地新华书店经销

*

2014年1月第 一 版 开本：720×1000 1/16

2014年1月第一次印刷 印张：17

字数：280 000

定价：42.00元

(如有印装质量问题，我社负责调换)

Original Japanese language edition
Excel de Manabu Yasashii Toukeigaku

By Hiroshi Takyu

Copyright © 2004 by Hiroshi Takyu

Published by Ohmsha, Ltd.

This Simplified Chinese version published by Science Press, Beijing

Under license from Ohmsha, Ltd.

Copyright © 2014

All rights reserved

Excelで学ぶやさしい統計学
田久浩志
オーム社 2004

著 者 簡 介

田久浩志

1978年3月 日本庆应义塾大学工学部电气学专业毕业

1980年3月 日本庆应义塾大学研究科电气专业毕业，获工学硕士学位

1980年4月 日本产业医科大学共同利用研究设施振动室

1989年9月 日本东海大学医学部医院管理学教室

1992年5月 日本东邦大学医学部医院管理学研究室

1993年6月 日本东邦大学医学博士

2000年4月 日本中部学院大学短期大学部经营学科教授

2001年4月 日本中部学院大学人间福祉学部健康福祉学科教授

其专业是医院管理学（医疗管理学）、医疗信息学。研究目标为保健、医疗、福利等领域的市场调查，并致力于研究医院和福利设施如何更好地经营和如何为大众更好地服务。曾于1999年获SAS User's Group International Japan功绩奖。

主要著作

《Excelによるナースドクターのための簡単データ整理法》

《ナースのためのかんたんWord》

《ナースのためのかんたんExcel》

《ナースのためのかんたんE-Mail》

《実力養成Word& Excel》（合著，羊土社）

《看護研究なんかこわくない》

《統計解析なんかこわくない》（合著，医学书院）

《JPMによる統計解析入門》（合著，欧姆社）

《転換期の福祉経営》（写作部分章节，日总研）

致中国读者

本书的写作目的是希望读者能够学会使用已知数据做成表格并进行分析。本书于2006年在日本出版发行，到目前为止（2013年11月）已经被200多所大学收藏并帮助了许多读者。

现在得知本书要在中国翻译出版，我感到非常高兴。我衷心希望本书也能对中国的读者起到帮助。

很多人都认为“统计学好难啊”，而且各种公式让数学没学好的人理解统计学更是难上加难。况且，就算是数学学得很好的人，如果让他们只靠公式和理论完全理解统计学，也是很困难的。那么，对于教育工作者来说，把枯燥难懂的内容变得浅显易懂，是其毕生的追求。

为了进行统计试验，笔者曾经记录过掷骰子时出现的各个数字。如今有了Excel软件，做各种汇总或者图表都很便利，Excel可以按照人们设定的条件来处理各种分布的数据。制作汇总表和图表，计算平均值和方差，在数据少和数据多的情况下，图表的形状会发生什么样的变化？我们应该自己动手进行这些操作，有了直观的印象才能更好地理解统计学。

与此同时，对于初学者的各种疑问，比如，为什么两个标准方差的差等于方差的和，看起来复杂的 t 检验公式是什么意思？泊松分布的 λ 代表什么？我们要尽量简明地说明这些问题，同时让大家自己处理分析数据从而加深理解。

本书的出发点是“不想增加人们对统计学的反感”。虽然本书不能把统计解说得如一般读物，但我们努力使本书尽可能地浅显易懂。

请读者把这本书当成“使用Excel、通过亲手操作来理解统计学的书”，并充分活用其中的知识。

田久浩志

注意事项

◆ Excel 版本

本书运用Microsoft Office Excel 2003（Windows版）进行数据处理。书中所出现的画面都是Windows XP Professional上Microsoft Office Excel 2003（Windows版）的截图。

Excel 2002（Windows版）也可以进行同样的操作，但是，部分画面或操作方法会有所不同。只要是高于Excel 2000（Windows版）的版本，除部分功能外，其他的操作都可以按照本书步骤进行。

本书不适用Macintosh版的Excel软件。

◆ 样本文件

读者可以用Excel软件实际操作并练习本书中所涉及的例题。通过自己在Excel上实际操作处理数据，能够更好地掌握统计知识。所以，笔者建议读者亲自动手制作各种各样的Excel工作表。

但是，有一部分练习题含有原始数据（例题数据）所需的内容。另外，上课的时候还要用到各种各样的工作簿。因此，包含原始数据的样本文件以及作为例题标准答案的文件都可以从日本欧姆社官方网站的“書籍連動ダウンロード（书籍互动/下载服务）”中下载。

样本文件的著作权归本书笔者田久浩志所有。虽然笔者保留著作权，但是在本书的学习中，读者可以自由地使用样本文件。

日本欧姆社官方网站为<http://www.ohmsha.co.jp/>

进入主页左下角“書籍連動/ダウンロードサービス(书籍互动/下载服务)”的“Excelで学ぶやさしい統計学(用Excel学统计学)”页面进行下载。

※下载服务可能会因为某些原因没有事先预告就中断或中止，请注意。

◆ “宏”的安全设置¹

本书中，笔者制作了一个Excel的“宏”，将其用于各种工作表。“宏”功能虽然简单，却为本书学习统计提供了不可或缺的功能。而且，这个“宏”也收录在下载的本文件中。设定和使用“宏”的功能并非难事，但是，在Excel的初期设定状态下是不能运行“宏”的。因此，在运行“宏”之前，必须进行以下设置。

(1) 从“工具”菜单中选择“宏”，然后选择“安全性”（图1）。



图1 设置“宏”安全性的步骤1

(2) 在“安全性”对话框里中，默认的设置是“非常高”或“高”，将其更改为“中”或“低”，单击“确定”（图2）。

¹ Excel 2007/2010/2013的“宏”安全性设置如下。

Excel 2013

• 打开Excel 2013→文件→Excel选项→信任中心→信任中心设置→宏设置→启用所有宏

Excel 2010

• 打开Excel 2010→文件→Excel选项→信任中心→信任中心设置→宏设置→启用所有宏

或者

打开Excel 2010→office按钮→Excel选项→常用→勾选在功能区显示“开发工具”选项卡，点击确定。重新打开Excel 2007后，在工具栏中会出现“开发工具”选项。点击开发工具→代码→宏安全性→信任中心→宏设置→启用所有宏

Excel 2007

• 打开Excel 2007→office按钮→Excel选项→信任中心→信任中心设置→宏设置→启用所有宏

或者

打开Excel 2007→office按钮→Excel选项→常用→勾选在功能区显示“开发工具”选项卡，点击确定。重新打开Excel 2007后，在工具栏中会出现“开发工具”选项。点击开发工具→代码→宏安全性→信任中心→宏设置→启用所有宏

经过上述设置，就可以运行“宏”了。

另外，按照本书所介绍的操作过程运行“宏”的时候，因读者的计算机配置和设定不同，可能会出现图表不能正常显示的情况。

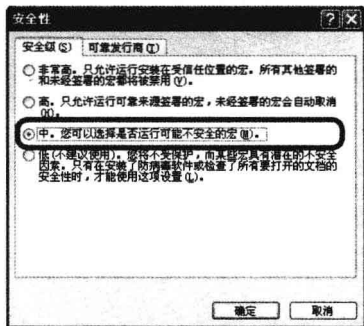


图 2 设置“宏”安全性的步骤 2

◆单元格的数值、图表的形状、随机变量的计算方法

本书以计算机随机生成的数值为例来制作Excel的工作表和图表。因此，就算读者进行同样的操作，得到的结果与书中介绍的Excel工作表的值和图表形状也可能并不相同。

关于单元格的格式设置和图表的格式设置等操作，会有一部分省略。

在书中介绍的例题工作表中，每次都会产生不同的随机变量，需要把Excel的“工具”→“选项”菜单中的“计算方法”设置为“手动”。因为在有些例题中，进行随机变量的计算需要很长时间，而且操作过程中数值也会发生变化，把“计算方法”设置为“手动”比较方便。读者可以根据情况自由切换这些设定。

◆免责声明

本书以及本书中所涉及的样本文件内容有些符合实际情况，有些则不符合，对于由此产生的直接或者间接影响，笔者、出版社概不负责。且关于软件的工作、运行环境、操作等问题，一概不做回应，请读者谅解。

原则上讲，本书的内容是执笔当时的成果。之后，根据情况会有一些变动，请读者注意。

序章 轻松学习统计学	1
------------------	---

引 言	2
-----------	---

第1章 汇总表和图表	5
------------------	---

1.1 汇总表和图表	6
------------------	---

1.1.1 引 言	6
-----------------	---

1.1.2 汇总表（透视表）	10
----------------------	----

1.1.3 柱形图的制作方法	14
----------------------	----

1.1.4 处理数据透视表的注意事项	17
--------------------------	----

1.1.5 结 语	18
-----------------	----

1.2 绘制折线图	18
-----------------	----

1.2.1 引 言	18
-----------------	----

1.2.2 生成符合标准正态分布的随机变量	19
-----------------------------	----

1.2.3 绘制折线图	23
-------------------	----

1.2.4 大量数据的汇总	24
---------------------	----

1.2.5 结 语	24
-----------------	----

1.3 调查数据的分析	25
-------------------	----

1.3.1 引 言	25
-----------------	----

1.3.2 数据的内容	25
-------------------	----

1.3.3 修正问卷调查的数据	26
-----------------------	----

1.3.4 整理问卷调查的数据	27
-----------------------	----

1.3.5 追加新的变量	29
1.3.6 调查数据的应用	31
1.3.7 从调查数据制作柱形图	33
1.4 调查数据的图表化——连续变量的图表	34
1.4.1 引 言	34
1.4.2 制作折线图	34
1.4.3 平均值的表示方法	37
1.4.4 散点图	38
1.5 平均和方差的计算方法	49
1.5.1 引 言	49
1.5.2 从原始数据求平均	49
1.5.3 从汇总表计算平均值	50
1.5.4 偏差平方和、方差	52
1.5.5 用基本方法计算方差	54
1.5.6 按照公式计算方差	55
1.5.7 利用汇总表计算方差	56
1.5.8 从一般汇总表计算方差	56
1.5.9 标准方差	57
1.5.10 总体方差和无偏方差	58
1.5.11 其他统计量	58
1.5.12 结 语	58
第2章 概 率	61
2.1 什么是概率	62
2.1.1 随机变量和概率分布	63
2.2 可能性——排列、组合、概率	63
2.2.1 引 言	63
2.2.2 排 列	64
2.2.3 组 合	65
2.2.4 用Excel函数计算阶乘、排列、组合的方法	66

2.2.5 概率的基础知识	67
2.2.6 结 语	68
2.3 概率分布	68
2.3.1 引 言	68
2.3.2 LOTO6的中奖号码	69
2.3.3 用Excel模拟LOTO6	72
2.3.4 概率的分布函数	76
2.3.5 结 语	78
第3章 分 布	81
3.1 分布的概念	82
3.2 制作能够直观感受统计的演习工作表	82
3.2.1 引 言	82
3.2.2 操作简介	83
3.2.3 定义单元格的名称	84
3.2.4 “宏”的设定与录制	85
3.2.5 确认“宏”的运行状况并绘制表格	87
3.2.6 利用均匀随机变量进行确认	90
3.2.7 均匀随机变量的和	91
3.2.8 结 语	91
3.3 体验二项分布——概率分布的基础是掷硬币	91
3.3.1 引 言	91
3.3.2 二项分布的概念	91
3.3.3 制作二项分布的演习工作表	95
3.3.4 结 语	97
3.4 泊松分布	98
3.4.1 引 言	98
3.4.2 什么是泊松分布	98
3.4.3 泊松分布的模拟演示	99
3.4.4 例1 交通事故和事故死亡人数的发生频率	101

3.4.5 例2 足球的得分差	102
3.4.6 结 语	104
3.5 分布之王——正态分布	106
3.5.1 引 言	106
3.5.2 制作演示工作表	106
3.5.3 正态分布的正式表现形式与标准化	108
3.5.4 标准正态分布	109
3.5.5 正态分布的图像（概率密度函数）及其累积（累积概率密度函数）	110
3.5.6 结 语	112
3.6 χ^2分布	113
3.6.1 引 言	113
3.6.2 什么是 χ^2 分布?	114
3.6.3 由正态分布推出 χ^2 分布	115
3.6.4 演习工作表	116
3.6.5 χ^2 分布和分割表的偏差	117
3.6.6 2×2 分割表	119
3.6.7 1×4 分割表	120
3.6.8 结 语	121
3.7 F分布	121
3.7.1 引 言	121
3.7.2 什么是F分布	122
3.7.3 演示工作表的准备	122
3.7.4 结 语	124
第4章 比较样本	125
4.1 两个变量的分布——假设检验法	126
4.1.1 引 言	126
4.1.2 假设检验法的基本思路	126
4.1.3 女人高跟鞋的高度	128
4.1.4 双侧检验和单侧检验	130

4.1.5 一般性假设检验法的操作方法	131
4.1.6 分布的位置关系以及第一种错误 (α)、第二种错误 (β)	131
4.1.7 结 语	134
4.2 数据的分布——你在众人当中的排名	134
4.2.1 引 言	134
4.2.2 准备服从正态分布的分数数据	135
4.2.3 将服从正态分布的分数标准化	136
4.2.4 从标准正态分布到偏差值、排名	137
4.2.5 关注偏差值时应当注意的问题	139
4.2.6 结 语	139
4.3 通过样本的离散计算整体的离散	140
4.3.1 引 言	140
4.3.2 准备数据	140
4.3.3 无偏方差和总体方差的关系	141
4.3.4 结 语	144
4.4 平均值的离散——平均值不是只有一个吗?	145
4.4.1 样本平均和总体平均的关系	145
4.4.2 大数法则	145
4.4.3 中心极限定理	148
4.4.4 潜藏在平均值中的误差	150
4.4.5 结 语	151
4.5 t分布——通过少量样本考察总体	152
4.5.1 引 言	152
4.5.2 体验 t 分布	152
4.5.3 准备工作表	153
4.5.4 结 语	155
4.6 两个平均值之和与之差的分布——合二为一	156
4.6.1 引 言	156
4.6.2 理解平均值之差的性质	157
4.6.3 准备工作表	158

4.6.4 结 语	160
4.7 考虑两个样本平均值——<i>t</i>检验	160
4.7.1 引 言	160
4.7.2 <i>t</i> 分布的工作表	164
4.7.3 结 语	166
第5章 思考差异	167
5.1 如何选择检验的方法	168
5.1.1 引 言	168
5.1.2 确定数据的种类和性质	168
5.1.3 变量的对应	170
5.1.4 检验方法的选择	171
5.1.5 结 语	171
5.2 等方差的检验	172
5.2.1 引 言	172
5.2.2 BMI指数的偏差	172
5.2.3 传统的等方差检验	173
5.2.4 通过FTEST函数进行等方差检验	175
5.2.5 结 语	175
5.3 <i>t</i>检验的例题	176
5.3.1 引 言	176
5.3.2 传统 <i>t</i> 检验的方法	177
5.3.3 通过TTEST函数进行 <i>t</i> 检验	179
5.3.4 成对数据的 <i>t</i> 检验	180
5.3.5 传统的Welch检验	181
5.3.6 通过TTEST函数进行Welch检验的方法	181
5.3.7 使用频率表来进行 <i>t</i> 检验的方法	182
5.3.8 结 语	183
5.4 卡方检验	184
5.4.1 引 言	184

5.4.2 独立性的检验——一口闷和年龄的关系	184
5.4.3 独立性的检验——联谊时一口闷的是与非	188
5.4.4 契合度的检验——Numbers3的中奖数字	191
5.4.5 拟合优度的检验——足球运动员的得分差	192
5.4.6 McNemar检验	194
5.4.7 McNemar检验的例题	197
5.4.8 结 语	198

第6章 方差分析和回归分析的实践

6.1 理解方差分析	200
6.1.1 引 言	200
6.1.2 方差分析的概念	200
6.1.3 计算组间波动的平方和	203
6.1.4 表示组内波动的平方和	203
6.1.5 平均值的多重比较	205
6.1.6 结 语	206
6.2 方差分析的例题	206
6.2.1 引 言	206
6.2.2 从原始数据进行方差分析	207
6.2.3 从汇总表进行方差分析	208
6.2.4 散点图	210
6.2.5 结 语	211
6.3 回归分析	212
6.3.1 引 言	212
6.3.2 准备数据	212
6.3.3 绘制散点图, 计算斜率和截距	213
6.3.4 回归分析的概念	215
6.3.5 计算回归系数和截距	216
6.3.6 相关系数	217
6.3.7 回归分析和方差分析的关系	218

6.3.8 结 语	221
第7章 U检验和Wilcoxon符号秩和检验	223
7.1 秩和检验	224
7.1.1 引 言	224
7.1.2 秩的性质	224
7.1.3 U检验的概述	225
7.1.4 U检验的详细内容	226
7.1.5 U的分布	228
7.1.6 Wilcoxon符号秩和检验	230
7.1.7 少量数据的解析——对同性和异性一口闷的印象	231
7.1.8 T_{cal} 分布	234
7.1.9 制作通用的工作表	236
7.1.10 结 语	238
7.2 无参数检验的例题	238
7.2.1 U检验的例题	239
7.2.2 Wilcoxon符号秩和检验的例题	241
7.2.3 结 语	243
7.3 综合练习——抽烟、性骚扰、一口闷等话题	243
7.3.1 抽烟、一口闷都是性骚扰的预备军?	243
7.3.2 对象和方法	244
7.3.3 抽烟与性骚扰拒绝度关系的 t 检验	246
7.3.4 一口闷与拒绝度	248
7.3.5 分析结果的应用	249
后 记	251
译后记	253