

写给所有人的极简统计学

作者：永野裕之

写给所有人的极简统计学

（日）永野裕之著
李俊译

北京时代华文书局

图书在版编目（CIP）数据

写给所有人的极简统计学 /（日）永野裕之著；李俊译.——北京：北京时代华文书局，2017.5

ISBN 978-7-5699-1438-2

I.①写...II.①永...②李...III.①统计学 IV.①C8

中国版本图书馆CIP数据核字（2017）第040751号

北京市版权局著作权合同登记号 图字：01-2016-9561

TOKEIGAKU NO TAME NO SUGAKU KYOSHITSU

by Hiroyuki Nagano

supervised by KENSUKE OKADA

illustrated by Ryuji Kitami

Copyright ©2015 by Hiroyuki Nagano

Chinese（in simplified character only）translation copyright©2017 by Sunnbook

Culture & Art Co., Ltd.

All rights reserved.

Original Japanese language edition published by Diamond, Inc.

Chinese（in simplified character only）translation rights arranged with Diamond, Inc.

through BARDON-CHINESE MEDIA AGENCY

写给所有人的极简统计学

著者 |（日）永野裕之

译者 | 李俊

出版人 | 王训海

选题策划 | 阳光博客

责任编辑 | 陈丽杰 袁思远

责任校对 | 陈丽杰 袁思远

装帧设计 | 阳光博客+李昆仑

责任印制 | 刘社涛

出版发行 | 北京时代华文书局 <http://www.bjsdsj.com.cn>

北京市东城区安定门外大街136号皇城国际大厦A座8楼

邮编：100011 电话：010-64267120 64267397

印刷 | 三河市华成印务有限公司 电话：0316-3521288

（如发现印装质量问题，请与印刷厂联系调换）

开本 | 710×930mm 1/16 印张 | 24.75字数 | 280千字

版次 | 2017年5月第1版 印次 | 2017年5月第1次印刷

书号 | ISBN 978-7-5699-1438-2

定价 | 56.00元

版权所有，侵权必究

[目录](#)
[扉页](#)
[版权信息](#)
[序言](#)
[第1章 数据整理的基础知识](#)
[第1章 前言](#)
[平均数](#)
[除法两个意义](#)
[\(A\) 除法的意义 \(1\) ——将一个整体平分](#)
[\(B\) 除法的意义 \(2\) ——将整体平均分成数个等份](#)
[比例](#)
[相同单位的比例为包含除](#)
[不同单位的比例为等分除](#)
[图表](#)
[\(1\) 柱状图：表示大小](#)
[\(2\) 折线图：表示变化](#)
[\(3\) 饼图：表示比例](#)
[\(4\) 带状图：比较比例](#)
[在统计学中的应用](#)
[数据与变量](#)
[实质型数据](#)
[量化数据](#)
[频数分布表](#)
[观察频数分布表时的注意事项](#)
[矩形图](#)
[绘制矩形图时的注意事项](#)
[代表值](#)
[研究数据的离散性](#)
[最大值与最小值](#)
[四分位数](#)
[箱形图](#)
[第2章 数据分析的基础知识](#)
[第2章 前言](#)
[平方根](#)
[根号 \(\)](#)
[平方根的计算](#)
[简化平方根](#)
[代数式的规则](#)
[分配律](#)
[分配律在心算中的实际运用](#)
[多项式展开](#)
[乘法公式](#)
[多项式展开的练习](#)
[在统计学中的应用](#)
[方差](#)
[标准差](#)
[偏差](#)
[第3章 用于研究关联性的数学](#)
[第3章 前言](#)
[函数](#)
[函数与图像的关系](#)
[函数与原因和结果的关系](#)
[一次函数](#)
[关于斜率的正负与图形](#)
[一次函数图像代数式的求法](#)
[二次函数的基础](#)
[图形的平移](#)
[配方法与二次函数的图像](#)
[配方法的要素](#)
[配方法](#)

[二次函数图像的画法](#)
[二次函数的最大值与最小值](#)
[二次函数与二次方程式](#)
[因式分解法](#)
[解的公式](#)
[图形与判别式的关系](#)
[二次不等式](#)
[在统计学中的应用](#)
[点位图](#)
[关于相关关系的注意点](#)
[相关系数](#)
[求导相关系数](#)
[相关系数](#)
[相关系数的理论背景](#)
[直观理解相关系数](#)
[相关系数取最大值与最小值](#)
[第4章 分析离散数据的数学](#)
[第4章 前言](#)
[阶乘](#)
[排列](#)
[关于0!](#)
[组合](#)
[关于的注意点](#)
[二项系数](#)
[集合](#)
[概率](#)
[和事件与积事件](#)
[独立试验](#)
[重复试验](#)
[等差数列](#)
[数列](#)
[等差数列的和](#)
[等比数列](#)
[等比数列的和](#)
 [\$\Sigma\$ 记号](#)
 [\$\Sigma\$ 记号的意义](#)
 [\$\Sigma\$ 记号的基本性质](#)
 [\$\Sigma\$ 记号在统计学中的应用](#)
[随机变数与概率分布](#)
[平均值](#)
 [\$aX+b\$ 的平均值](#)
[随机变数的方差与标准差](#)
 [\$aX+b\$ 的方差与标准差](#)
[随机变数的标准化](#)
[和的平均值](#)
[积的平均值](#)
[和的方差](#)
[二项分布](#)
[第5章 分析连续数据的数学](#)
[第5章 前言](#)
[对于“无限”的理解](#)
[什么是无限](#)
[极限](#)
[欧拉常数e](#)
[积分](#)
[阿基米德的求积法](#)
[积分的符号与意义](#)
[在统计学中的应用](#)
[连续型随机变数与概率密度函数](#)
[概率密度函数的性质](#)

[连续型随机变数的平均值与方差](#)

[正态分布](#)

[标准正态分布](#)

[正态分布表](#)

[什么是推测统计](#)

[使用标准正态分布进行“推算”](#)

[使用标准正态分布进行“检验”](#)

[t检验](#)

[练习题答案](#)

[第1章](#)

[第2章](#)

[第3章](#)

[第4章](#)

[第5章](#)

[结语](#)

序言

统计学能力的代沟

2015年1月份，日本被称作“非自主教育时代”的第一届高中生参加了全国统一大学入学考试。以下问题，是这些高中生们数学考试中的必答题（不必现在思考如何解析，可以快速阅览一遍，答案将在后面揭晓）。

一个有40人的高二班级，让每个学生分别扔出两次手球，之后统计每次手球飞出的距离。图2是以每个人第一次扔出的距离作横轴、第二次扔出的距离作纵轴绘制出的点位图。因一人缺席，现有39人的数据。

日

（协方差为第一组数值的偏差【指各数值与平均值的差】与第二组数值偏差的乘积的平均数）

（1）从①～⑩中选出最恰当的选项。欢迎加入书社,每日海量书籍,大师课精彩分享微信:dedao555

与两组数值的相关系数最接近的值是_____。

①0.67 ②0.71 ③0.75 ④0.79 ⑤0.83

⑥0.87 ⑦0.91 ⑧0.95 ⑨0.99 ⑩1.03

（2）从①～⑧中选出应填入空格中的不等式关系。

后来，也让之前缺席的那位学生做了相同的两次投掷，并记录了数据。第一次的数据为24.7m，第二次为26.9m。将此学生的数据加入之前全班的数据中重新计算。假设新的共同方差为A，之前的共同方差为B，新的相关系数为C，之前的相关系数为D。那么A与B, C与D的大小关系成立的是_____。

① $A > B, C > D$ ② $A > B, C = D$ ③ $A > B, C$

④ $A = B, C > D$ ⑤ $A = B, C = D$ ⑥ $A = B, C$

⑦AD ⑧A

“非自主教育时代”的高中生，无论是文科还是理科，在被称作“数据分析”的数学必修课中，都被要求掌握柱状图、箱形图、方差、标准差、相关系数等统计学的基础。

我觉得1974年之前高中毕业的人，能将这个问题顺利解析出来的应该不多。因为，1974年之前，包括我在内，大部分读者读书时，统计学只是一门选修课（且大部分人不会选），在学校学习统计学的人少之又少。

而对于2015年3月份以后高中毕业的一代人来说，这个问题根本不是什么难题。作为刚毕业不久的社会新人，在“统计学知识（应用统计学的能力）”这个问题上，与我们这代人之间存在巨大的代沟。

西内启的畅销书《看穿一切数字的统计学》（钻石出版社）的出版，让大家开始关注统计学，其中有一节是这么写的：

“如今，有了计算机这个强有力的伙伴，各个领域，世界上的每一个角落，以及人生中的每一个瞬间，所有待解的难题，都能从统计学中找到答案。”

在信息过剩与价值观多样化的现代，能够理解并表达出通过统计学运算得出的结论，已经成为在社会上行走必须具备的一项能力。毋庸置疑的是，统计学知识已经成为现代人不可或缺的一项技能。

社会人士无法理解统计学的原因

我在一家名为“永野数学私塾”的数学教育机构，给社会人士做数学方面的培训。学员们重新学习数学的原因各不相同，最近听到比较多的一个原因是“想学统计学”。我一开始觉得：“市面上都出了这么多统计学的书籍，还特地过来学习，一定是想了解统计学领域比较深奥的知识吧。”可是，讲了几节课之后，我发现难住很多学生的不是统计学知识，而是数学基础知识。由于不理解统计学书籍中出现的那些初、高中数学知识，因此连基础的统计学都看不懂。换句话说，只要学好数学，就可以参考各种介绍统计学知识的书籍自主学习了。

让人觉得不可思议的是，市面上介绍“用在统计学上的数学”的书寥寥无几。正因为如此，我才提笔写了这本书。读者朋友们即便在学校没有选修统计学，也可以轻松自学和统计学相关的数学基础知识。

本书的内容

本书精选了学习统计学所需的初、高中阶段的基础数学知识。从除法的意义与比值（第一章）这种小学水平的算术开始，到平方根、多项式的计算（第二章）、函数与表格（第三章）、概率、 Σ 记号（第四章）、极限、微积分（第五章）等，涵盖的内容相当广。每个章节都尽可能地以“易于理解”为第一主旨。北见龙子小姐的插画也帮了我很大的忙。同时，在篇幅允许的情况下，我列举了一些例题与练习题。这些题目，可以帮助大家确认是否掌握了书中介绍的知识点，请勿跳过，一定要试着做做看！

本书还介绍了相关数学知识在统计学中的应用范围与具体方法。这部分内容还特地邀请了心理统计学研究室的冈田谦介教授担任编辑顾问。冈田老师会在书中登场亮相，为大家解析哦！

本书将“非自主教育时代”的统计学必修内容归纳在前三章中，希望读者通过对第四章的学习，理解“离散型数据”（即零零散散的数据），通过第五章理解“连续型数据的概率密度函数”等相关内容。简单来说，此书将会为大家总结“对收集的数据加以分析”的统计方法，以及介绍“从部分数据中推算出整体情况”的推测统计的入门知识。

□

统计数学是社会人士必须掌握的数学能力

在写这本书的时候，我深深觉得，这本书中介绍的数学知识真的是社会人士必须要掌握的（虽然只是针对统计学选择的一小部分数学知识）。只要掌握了本书中的数学知识，即便是和数学能力很强的人交谈，也不用再胆怯、自卑了。除此之外，看到满是数字的资料和Excel中的函数，也能轻松应对了，甚至可以用图表做出更精彩、更成功的演讲。对思考能力的锻炼也更加不在话下！

写到这儿，也差不多该进入正题了。我将通过此书，带你通过最短的捷径抵达目标。所以，请相信我，一起来学习吧！

□

（开头数学问题的答案：（1）⑧ （2）⑧）

第1章 数据整理的基础知识

第1章 前言

总的来说，统计就是将收集到的数据（数值）进行整理并加以分析的学问。

注释：数据与数值的意思是相同的。在中小学通常称之为“数值”，到了高中以后大多称之为“数据”。本书将统一写作“数据”。

在这一章，我们先学习整理数据过程中必须要掌握的平均、比例以及图表等概念的相关知识。这些内容虽说都是小学数学水平，不过也别抱着“这种简单的知识我都知道”的心态掉以轻心。你还别不信，来看看这道题目吧。

问题 对某初中三年级100名学生的身高进行测量，计算出的平均值为163.5cm。以此结果为依据，对以下命题进行判断。正确的在前面的空格中画√，错误的画×。

- （1）身高在163.5cm以上和163.5cm以下的学生各有50名。
- （2）将这100名学生的身高相加求和，结果等于 $163.5 \times 100 = 16350\text{cm}$
- （3）将学生身高以10cm为单位分成若干区间，如身高在“130~140cm”“140~150cm”……那么身高在“160~170cm”这个区间的学生人数最多。

（选自：日本数学协会主页）

这是日本数学协会在2011年，针对全国约6000名大学生进行的“大学生数学基础知识调查”中的第一道题。结果显示，这道题的正确率为76%。当时，媒体以“大学生平均4人中就有一人不会”为热点进行了大肆报道。这道题，你有自信答对吗？（该题答案将在本章最后揭晓。）

在学习数学的过程中，“比例”这一知识点，是很让学生头疼的一个难点。国家教育政策研究所在2013年进行的“全国学生学习能力与状况调查”结果显示，有关除法的问题，正确率排在最后一位。要准确理解“比例”这一概念，就必须从根本上理解除法两个意义。但根据我的教学经验，即便是成年人，能正确理解“除法的两个意义”的也只有寥寥数人而已。“比例”这个概念，是学习统计学的基础，而且“比例”与“概率”之间存在千丝万缕的联系，只要在理解上稍有偏差，就可能为今后学习统计学埋下祸根。

接下来，我们再说图表。将数据加以整理，并以图表形式呈现出来，可以说图表是一种方便大家理解的有力工具。但是，若选错图表，导致你使用的图表与想呈现的数据或内容不相匹配，不但不能起到方便理解的作用，反而会让人越看越乱，甚至造成误导。相信某些读者朋友有过这种经历，公司开会时，上司愤怒地指责你：“用这种图表，你觉得谁能看懂？”被这样责骂，还不是因为没有选择恰当的图表。欢迎加入书社,每日海量书籍,大师课精彩分享微信:dedao555

综上所述，在这一章节，我们就重新学习一遍“平均”“比例”以及“图表”等整理数据所需的基础知识吧！

平均数

“平均”一词，顾名思义，是指将事物平等、均等分配的意思。

例如，有如下三个长方形，高度分别为2、7、3。要想让这三个长方形一样高（平均），要怎么做呢？将高度为“7”的长方形切断，然后分给其他两个长方形，如图所示：

□

使高度一致后，我们得到了一个大的长方形。这个长方形的纵向长度为平均数（长），横向长度为个数（宽），面积就是总数。即：

□

那么，由此可以得到

□

将以上内容，用代数整理。将所有数据列为：

$x_1, x_2, x_3, \dots, x_n$

数据总个数为 n 。用这些数据的总和除以 n 就是平均数。在数学中一般用“ $\bar{x}$ ”表示平均数，即在字母上方添加一横来表示。

平均数的定义

□

那么，马上来练习一下吧！

例题1-1 下表为两个小组的数学测验成绩，分别是有6人的A组与有5人的B组。求出两组各自的平均分。

□

【解析】

A组平均分

□

B组平均分

□

两组的平均分都是50分。

像这样数据个数（人数）不同的多组数据之间，也可以相互比较。

但是，仔细观察两组数据中每个人的分数，不难发现：A组的分数中，低于平均分、与平均分持平以及高于平均分三个领域的学生各有2名，分布较均匀；而B组的分数中，未达到平均分的有4人，高于平均分的仅有一人。由此可见，B组中是100分的学生拉高了整组的平均分。

□

在这些数据中，还有许多通过平均数无法发现的特点。在统计学中，作为显示数据特征的数值，除了平均数以外，还会使用中位数与众数（这些概念将在后面进行详述）。

接下来，我们要学习的是“除法两个意义”。这部分内容在我的另一本书《数学好的人是如何思考的》中提及过，但这个知识点是理解概率和比例这两个概念的前提，因此在这里再详细给大家介绍一遍。

除法的两个意义

我们先做一个实验。

如下图所示，有6个圆。

□

用这些圆，将

$$6 \div 3 = 2$$

表现出来。因为此题不存在唯一的正确答案，所以大家可以放开思维，任意遐想。如果可以，让家人、朋友也一起来做这道题吧。做完之后，你会发现一些有趣的答案。

欢迎访问：电子书学习和下载网站（<https://www.shgis.cn>）

文档名称：《写给所有人的极简统计学》永野裕之 著.epub

请登录 <https://shgis.cn/post/880.html> 下载完整文档。

手机端请扫码查看：

