

新世纪经济学、法学、管理学类专业统计学课程辅助教材

《统计学基础》 案例分析和习题解答

王庆石 霍红 编著

$$P(A_j|B) = \frac{P(A_j)P(B|A_j)}{\sum_{j=1}^n P(A_j)P(B|A_j)}$$



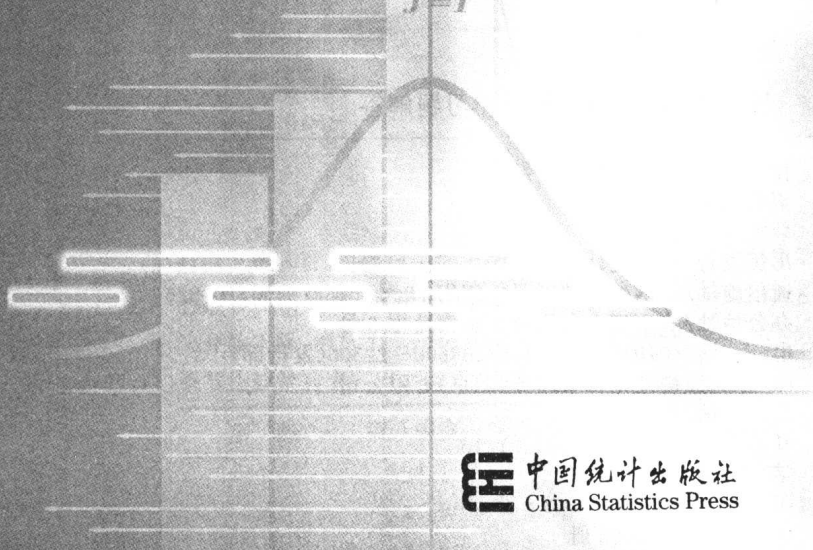
中国统计出版社
China Statistics Press

新世纪经济学、法学、管理学类专业统计学课程辅助教材

《统计学基础》 案例分析和习题解答

王庆石 霍红 编著

$$P(A_j|B) = \frac{P(A_j)P(B|A_j)}{\sum_{j=1}^n P(A_j)P(B|A_j)}$$



 中国统计出版社
China Statistics Press

(京)新登字 041 号

图书在版编目(CIP)数据

《统计学基础》案例分析和习题解答/王庆石 霍红编著.

—北京:中国统计出版社,2003.7

ISBN 7-5037-4025-4

I. 统…

Ⅱ. ①王… ②霍…

Ⅲ. 统计学—高等学校—教学参考资料

Ⅳ. C8

中国版本图书馆 CIP 数据核字(2003)第 058313 号

《统计学基础》案例分析和习题解答

作 者/王庆石 霍 红

责任编辑/姚 立

装帧设计/张建民

出版发行/中国统计出版社

通信地址/北京市西城区月坛南街 75 号 邮政编码/100826

办公地址/北京市丰台区西三环南路甲 6 号

电 话/(010)63459084、63266600—22500(发行部)

印 刷/科伦克三莱印务(北京)有限公司

经 销/新华书店

开 本/787×1092mm 1/18

字 数/140 千字

印 张/7.5

印 数/1—3000 册

版 别/2003 年 8 月第 1 版

版 次/2003 年 8 月北京第 1 次印刷

书 号/ISBN 7-5037-4205-4/C·2036

定 价/16.00 元

中国统计版图书,版权所有,侵权必究。

中国统计版图书,如有印装错误,本社发行部负责调换。

前 言

2001年,由香港城市大学管理科学系陈乃九教授牵头,由香港城市大学、东北财经大学、中国人民大学、上海财经大学多位教授共同参加编写的《统计学基础》一书由中国统计出版社正式出版发行。该书是面向财经类非统计学专业本科生编写的,突出统计学的思想性和应用性,在内容体系上力求反映统计学的本来面貌。

在该书的每一章后面,虽然给出了大量的思考题和练习题,但是许多教师还是希望每一章都能配以真实的案例,通过对一个较大案例的详细分析,使学生消理解各章所讲解的统计方法的应用原理,从应用中把握统计学的真谛。基于这种考虑,我和我的研究生霍红共同研究、策划、编写了这本《〈统计学基础〉案例分析及习题解答》。书中除了提供案例问题和案例分析之外,还对每一章后面的练习题提供了详细的解答过程。因此,本书作为与《统计学基础》相配套的教学参考书,既可以为教师的教学准备提供一个便捷的帮手,也可以为学生的课后学习提供咨询帮助。当然,本书也同样适合于讲解和学习其它版本统计学教材的参考,因为统计学教材的内容体系大同小异。

本书内容的安排基本上是这样的:与《基础统计学》教材中的内容完全对应,每一章对应一个或两个案例,通过对所给案例的求解和分析,可以把该章教材中介绍的大部分统计方法融会贯通起来。案例的求解是直接采用 SPSS 软件或 Excel 软件做出的,案例的分析也是直接根据 SPSS 或 Excel 的输出结果做出的,当然求解和分析我们只给出了一种思路和方法。

本书各个案例的原始数据均可以在以下两个网站下载:(1)<http://tj.dufe.edu.cn> 或 (2)<http://csp.stats.gov.cn>。关于《统计学基础》一书的试题库和其它教学资料,我们也将陆续通过该网站推出,与国内统计同仁们共同交流和分享。

本书的出版,应该首先感谢东北财经大学统计系主任蒋萍教授的积极推荐,其次应该感谢中国统计出版社副社长严建辉先生、编辑部主任刘国宁先生和责任编辑姚立女士,再次我要感谢教育部“新世纪经济学、法学、管理学类专

业教育改革工程”项目“统计学案例教学理论与方法研究及统计学基础课教科书与案例教材建设”子项目的资助(项目编号:教高司财函[2000]第 28 号文件,127201016 号),本书也是该项目的阶段性成果。最后我还要感谢陈乃九教授、卢兴普教授、易丹辉教授和项目组的其他各位教授对我工作的支持。

由于作者知识水平有限,书中错误在所难免,真诚希望各位统计同仁提出宝贵意见和建议。

电子信箱:wqingshi@dufe.edu.cn.

王庆石 霍 红

2003 年 7 月 6 日

统计学基础

目 录

第一部分 案例问题

案例一	降雨量的描述性统计分析	(3)
案例二	身高和体重间的数量关系分析	(5)
案例三	彩票游戏中的概率计算	(7)
案例四	关于哑铃重量的抽样分布分析	(11)
案例五	奶粉蛋白质含量的初步统计推断	(13)
案例六	两个样本数据的总体均值和比例的推断	(15)
案例七	类别型变量的相关分析——列联表	(17)
案例八	非参数统计方法的两个应用	(19)
案例九	某高中文理科学学生数学和英语成绩的方差分析	(21)
案例十	妇女生育率影响因素的回归分析	(23)
案例十一	美国活牛存栏数的时间序列分析	(25)

案例三分析	彩票游戏中的概率计算	(43)
案例四分析	关于哑铃重量的抽样分布分析	(48)
案例五分析	奶粉蛋白质含量的初步统计推断	(54)
案例六分析	两个样本数据的总体均值和比例的推断	(58)
案例七分析	类别型变量的相关分析——列联表	(64)
案例八分析	非参数统计方法的两个应用	(69)
案例九分析	某高中文理科学生数学和英语成绩的方差分析	(74)
案例十分析	妇女生育率影响因素的回归分析	(78)
案例十一分析	美国活牛存栏数的时间序列分析	(84)

第三部分 习题解答

《统计学基础》练习题解答	(93)
--------------------	------

第一部分 案例问题

降雨量的描述性统计分析

三、要求讨论和解决的问题

1. 分别用图示法和特征值法表现降雨量数据的分布；
2. 将降雨量数据按步长 100 进行分段,构造一个新的分类变量,分别用图示法和特征值法表现这个变量的分布；
3. 根据以上分析,写一篇分析报告。

案例二

✕ ✕ ✕ ✕ ✕ ✕ ✕

一、学习目的

通过对这个案例的学习,希望达到以下目的:

1. 熟练掌握描述变量之间数量关系的三种方法和工具；
2. 能够利用统计软件处理案例数据所包含的变量之间的数量关系；
3. 能够对统计处理的结果进行正确的解释和分析。

二、数据文件说明

本案例数据以 SPSS 格式保存, 文件名为 data2, 请到 <http://tj.dufe.edu.cn> 网站或 <http://csp.stats.gov.cn> 网站上下载数据。本案例数据来源于某小学对该校的 10 岁儿童进行的身高和体重的调查结果。其中包含的三个变量分别是: SEX——代表性别, 其中 1 表示男, 2 表示女; HIGH——代表身高, 单位是厘米; WEIGHT——代表体重, 单位是千克。

建议用 SPSS 统计软件分析本案例的数据。用 SPSS 软件读取数据的步骤是：(1)把在上述网站上本案例的数据文件下载到本地计算机磁盘的一个指定目录中，比如 CASEDATA。(2)用 SPSS 的菜单命令 Open——Data...，选择文件名 data2，文件类型选为 SPSS(*.sav)，就可以打开数据文件了。然后观察数据文件的格式，变量名、变量个数和观测值个数。

三、要求讨论和解决的问题

1. 用散点图描述变量 HIGH 和 WEIGHT 之间的关系,并用变量 SEX 加以标识;
2. 计算变量 HIGH 和 WEIGHT 之间的相关系数;
3. 运用二向交叉表分析变量 HIGH 和 SEX 以及 WEIGHT 和 SEX 之间的关系;
4. 对以上处理结果进行分析,然后写一篇分析报告。

彩票游戏中的概率计算

1. 理解概率论中相关概念的基本定义和广义概率准则;
2. 区分独立事件与相容事件,并能够计算简单的概率问题;
3. 掌握条件概率公式与贝叶斯公式的区别;
4. 能够运用相关的正态分布近似理论来求解大样本事件的概率问题。

概率论是研究随机现象统计规律的一门数学分支学科。对概率问题的研究最早起源于赌博,而赌博在实质上是一种机会游戏或概率游戏。在现实生活中,常见的赌博行为比如博彩,也就是通常所说的彩票。目前国内大众比较熟悉的彩票有福利彩票、体育彩票和足球彩票。这里我们将对这三种彩票予以简单介绍,并分析它们所蕴含的概率意义。

一种被称做“幸运 35 选 7”的福利彩票,它的投注规则是在数字 1—35 中选择 7 个数,投注人可以自己选择这 7 个号码,也可以由电脑随机抽取(不重复抽取,即每个数字在一组号码中最多出现一次)出下注的号码,在开奖日由

电脑随机抽取(不重复抽取)出 7 个数字,若投注人所选择的号码与这一组号码相同(不计顺序,如“10”在 7 个数中排在第几位对结果没有影响),则认为投注人赢,反之认为投注人输。下面的表格给出了该种彩票过去的 202 期的开奖号码统计结果。

表 3.1 福利彩票过去 202 期的号码统计频数表

号码	频数	号码	频数	号码	频数
1	44	13	53	25	26
2	38	14	51	26	43
3	34	15	38	27	32
4	49	16	35	28	48
5	47	17	38	29	40
6	40	18	42	30	36
7	40	19	45	31	35
8	43	20	32	32	34
9	37	21	45	33	40
10	46	22	40	34	42
11	37	23	56	35	44
12	30	24	34		

表 3.2 连号出现次数统计表

出现次数	0	1	2	3	4	合计
频数	44	82	62	12	2	202

注:“连号”是指在一组号码中出现相邻的两个数字。如“13”和“14”就是一次连号,“13”和“15”不是连号,“13”、“14”和“15”要计作两次连号。

请思考下列问题:

- (1) 根据表 3.1 和表 3.2 中所给出的统计结果,计算这 35 个数字出现的频率以及连号出现的频率;
- (2) 根据概率的定义分别求这 35 个数字出现的概率;
- (3) 求抽到的号码小于 10 的概率,号码为奇数的概率,号码介于 13 和 24 之间的概率。

(二)体育彩票

这里我们所介绍的体育彩票是传统型的“36 选 7”彩票,投注规则是从 36

个运动项目中选取 7 个项目号码组合为一注彩票,只选 7 项的各注彩票为单式票,选 8 项或 8 项以上(不超过 16 项)的多注彩票为复式票。36 个体育运动项目如下表:

表 3.3 36 种体育项目列表

01 田径	02 游泳	03 跳水	04 水球	05 体操	06 举重
07 射击	08 射箭	09 击剑	10 柔道	11 摔跤	12 拳击
13 马术	14 足球	15 篮球	16 排球	17 乒乓	18 羽毛
19 网球	20 手球	21 棒球	22 垒球	23 滑雪	24 冰球
25 帆船	26 帆板	27 皮艇	28 划艇	29 赛艇	30 技巧
31 武术	32 滑水	33 蹼泳	34 围棋	35 象棋	36 桥牌

中奖号码:每次开奖由专用开奖机从 36 个项目号码中产生 6 个正选项目号码,即基本号码,再从剩下的 30 个号码中产生 1 个号码作为特别项目号码。其中的中奖办法与前面的福利彩票基本一致。这里我们需要讨论和计算的问题是:

- (1) 投注者选中所有 7 个号码的概率;
- (2) 投注者只选中 4 个正选项目号码的概率;
- (3) 投注者没有选中任何号码的概率。

(三)足球彩票

足球彩票是一种特殊的体育彩票,由于其结果具有更大的随机性,所以其表现出的投机性也就更强。每一期的足彩都是对给定的 13 场比赛的结果进行竞猜,其中 3 代表主队胜,1 代表两队平,0 代表主队负。购票人在投注时,可以在每一场比赛的 3 种预测结果中,任意选择 1 种或 1 种以上(复式投注)预测结果。每一张投注单的投注数量不得高于 10 000 注。由于足球彩票号码的选择有很大的主观性,这里我们只考虑如何确定复式投注的数量。

假设某一投注者有 m 场选择了 1 种结果, n 场选择了 2 种结果,剩余场次选择了 3 种结果,那么这个复式投注的数量是多少?

三、一个补充说明性的例子

假设有甲、乙、丙三家工厂生产同种产品,经过随机试验,发现甲、乙、丙三厂产品的合格品和不合格品的比率分别是 49 : 1、19 : 1、24 : 1。此外,我们还

知道三家工厂的产量之比为 $3:2:1$ 。现在任意抽取一件该产品,问:

1. 该产品分别为甲、乙、丙三厂的概率;
2. 该产品为次品的概率;
3. 该产品为甲厂、乙厂或丙厂的不合格品的概率;
4. 若抽取的是次品,则该产品分别是甲、乙、丙三厂生产的概率;
5. 现在我们重新抽取 20 件该产品,其中只有 1 件不合格品的概率;
6. 某采购员已购买了 1 000 件该产品,问其中不合格品少于 30 件的概

率。