



卫生部“十一五”规划教材

全国高等医药教材建设研究会规划教材

全国高等学校配套教材 • 供药学类专业用

医药数理统计方法 学习指导与习题集

主 编 高祖新

副主编 刘艳杰 张丕德



人民卫生出版社

卫生部“十一五”规划教材
全国高等医药教材建设研究会规划教材
全国高等学校配套教材
供药学类专业用

医药数理统计方法

学习指导与习题集

主 编 高祖新

副主编 刘艳杰 张丕德

编 者 (以姓氏笔画为序)

刘 勇 (南昌大学)

刘启贵 (大连医科大学)

刘艳杰 (沈阳药科大学)

闵心畅 (四川大学)

张丕德 (广东药学院)

何俏军 (浙江大学)

高祖新 (中国药科大学)

袁兆康 (南昌大学)

韩可勤 (中国药科大学)

人 民 卫 生 出 版 社

· 图书在版编目 (CIP) 数据

医药数理统计方法学习指导与习题集/高祖新主编.
—北京: 人民卫生出版社, 2007. 8

ISBN 978-7-117-09008-7

I. 医… II. 高… III. 数理统计-应用-医药学-
高等学校-教学参考资料 IV. R311

中国版本图书馆 CIP 数据核字 (2007) 第 107465 号

医药数理统计方法学习指导与习题集

主 编: 高祖新

出版发行: 人民卫生出版社 (中继线 010-67616688)

地 址: 北京市丰台区方庄芳群园 3 区 3 号楼

邮 编: 100078

网 址: [http://www. pmph. com](http://www.pmph.com)

E - mail: [pmph @ pmph. com](mailto:pmph@pmph.com)

购书热线: 010-67605754 010-65264830

印 刷: 尚艺印装有限公司

经 销: 新华书店

开 本: 787×1092 1/16 印张: 12.75

字 数: 289 千字

版 次: 2007 年 8 月第 1 版 2007 年 8 月第 1 版第 1 次印刷

标准书号: ISBN 978-7-117-09008-7/R · 9009

定 价: 17.00 元

版权所有, 侵权必究, 打击盗版举报电话: 010-87613394

(凡属印装质量问题请与本社销售部联系退换)

前 言

本书作为卫生部“十一五”规划教材,是全国高等医药教材建设研究会、卫生部教材办公室规划的第六轮药学专业教材《医药数理统计方法》(第4版)的配套学习指导书,是该规划教材的细化和延伸。本书编写目的是为了使學生更扎实地掌握医药数理统计的知识和方法,培养其分析和解决问题的统计应用能力,也便于教师更好地实施医药统计课程的教学。

医药数理统计作为一门理论性和应用性较强的数学学科,其教学的顺利完成既需要教师的课堂教学和指导,又需要学生独立思考,认真做课后练习,以消化掌握课堂所学知识内容。学生必须亲自动手,独立完成一定数量的、不同形式的练习题,才能掌握概率和数理统计的基本概念、原理和方法,把握其知识要点和解题技巧,培养正确的统计思维方法,提高其统计分析和应用能力。

本书编委会由国内培养药学人才的主要医药院校统计学专家教授组成,包括了卫生部“十一五”规划教材《医药数理统计方法》(第4版)的全体编写人员,具有从事多年的医药统计教学和科研实践的丰富经验和坚实专业基础。

本书的特点是对《医药数理统计方法》(第4版)教材的知识体系进行了精心梳理,列出了各章的教学目的和要求,并将教材的核心内容高度概括为简明的内容提要简表,突出了各章的知识要点和重点;同时每章精选了较多综合性典型例题进行解析,对教材的全部习题进行详细解答,为读者提供了极为有效的兼具基础性和技巧性的解题指导;另外每章还增加了包括填充题、选择题和计算题等多种题型的思考与练习(及答案)供平时自测训练,书后还编有综合测试试卷及参考解答、上机训练题 Excel 题解等,便于帮助读者消化、巩固所学内容,加深对概率和数理统计知识的深刻理解,并能及时测试自己对所学知识的掌握程度,从而提高其学习效率和成绩。

本书由高祖新主编,刘艳杰、张丕德副主编,并由高祖新负责全书的整体设计、各章内容提要简表的主体编制和全书统稿工作等。本书编著时注意博采众长,参考了国内外多种教材和参考文献,同时还得到人民卫生出版社及编委所在单位等的大力支持,在此一并表示衷心的感谢。由于编写时间和水平所限,书中疏漏和不妥之处在所难免,恳请各位专家、读者批评指正,以便再版时改进修正。

编 者

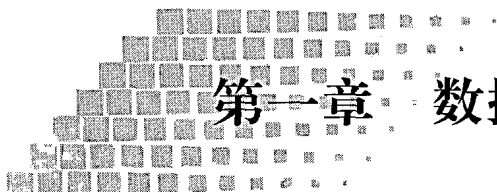
2007年7月

目 录

第一章 数据的描述和整理	1
一、学习目的和要求.....	1
二、内容提要.....	1
三、综合例题解析.....	3
四、习题一解答.....	3
五、思考与练习.....	7
六、思考与练习参考答案.....	8
第二章 随机事件与概率	9
一、学习目的和要求.....	9
二、内容提要.....	9
三、综合例题解析	11
四、习题二解答	15
五、思考与练习	21
六、思考与练习参考答案	22
第三章 随机变量及其分布	24
一、学习目的和要求	24
二、内容提要	24
三、综合例题解析	29
四、习题三解答	35
五、思考与练习	48
六、思考与练习参考答案	51
第四章 抽样分布	55
一、学习目的和要求	55
二、内容提要	55
三、综合例题解析	57
四、习题四解答	59
五、思考与练习	61
六、思考与练习参考答案	62

第五章 参数估计	63
一、学习目的和要求	63
二、内容提要	63
三、综合例题解析	64
四、习题五解答	67
五、思考与练习	72
六、思考与练习参考答案	73
第六章 参数假设检验	74
一、学习目的和要求	74
二、内容提要	74
三、综合例题解析	77
四、习题六解答	79
五、思考与练习	87
六、思考与练习参考答案	89
第七章 非参数假设检验	91
一、学习目的和要求	91
二、内容提要	91
三、综合例题解析	92
四、习题七解答	95
五、思考与练习	104
六、思考与练习参考答案	105
第八章 方差分析	106
一、学习目的和要求	106
二、内容提要	106
三、综合例题解析	109
四、习题八解答	112
五、思考与练习	120
六、思考与练习参考答案	121
第九章 相关分析与回归分析	123
一、学习目的和要求	123
二、内容提要	123
三、综合例题解析	126
四、习题九解答	130
五、思考与练习	137
六、思考与练习参考答案	139

第十章 试验设计	141
一、学习目的和要求	141
二、内容提要	141
三、综合例题解析	143
四、习题十解答	147
五、思考与练习	155
六、思考与练习参考答案	156
综合测试试卷	158
综合测试试卷参考答案	163
上机训练题 Excel 题解	166
常用统计表	182
附表 1 二项分布表	182
附表 2 泊松分布表	185
附表 3 标准正态分布表	187
附表 4 标准正态分布的双侧临界值表	188
附表 5 χ^2 分布表	189
附表 6 t 分布表	190
附表 7 F 分布表	191



第一章 数据的描述和整理

一、学习目的和要求

1. 掌握数据的类型及特性；
2. 掌握定性数据和定量数据的整理步骤、显示方法；
3. 掌握描述数据分布的集中趋势、离散程度和分布形状的常用统计量；
4. 理解并熟练掌握样本均值、样本方差的计算；
5. 了解统计图形和统计表的表示及意义；
6. 了解用 Excel 软件进行统计作图、频数分布表与直方图生成、统计量的计算。

二、内 容 提 要

(一) 数据的分类

数据类型	定性数据(品质数据)		定量数据
	定类数据(计数数据)	定序数据(等级数据)	数值数据(计量数据)
表现形式	类别(无序)	类别(有序)	数值(+ - × ÷)
对应变量	定类变量	定序变量	数值变量(离散变量、连续变量)
主要统计方法	计算各组频数,进行列联表分析、 χ^2 检验等非参数方法		计算各种统计量,进行参数估计和检验、回归分析、方差分析等参数方法
常用统计图形	条形图,圆形图(饼图)		直方图,折线图,散点图,茎叶图,箱形图

(二) 常用统计量

1. 描述集中趋势的统计量

名 称	公式(原始数据)	公式(分组数据)	意 义
均值 $\bar{x}$	$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$	$\bar{x} \approx \frac{1}{n} \sum_{i=1}^k m_i f_i$	反映数据取值的平均水平,是描述数据分布集中趋势的最主要测度值

续表

名 称	公式(原始数据)	公式(分组数据)	意 义
中位数 M_e	$M_e = \begin{cases} x_{(\frac{n+1}{2})}, & \text{当 } n \text{ 为奇数} \\ \frac{1}{2}(x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}), & \text{当 } n \text{ 为偶数} \end{cases}$	中位数所在组: 累积频数超过 $n/2$ 的那个最低组	是典型的位置平均数,不受极端值的影响
众数 M_o	数据中出现次数最多的观察值	众数所在组: 频数最大的组	测度定性数据集中趋势,对于定量数据意义不大

2. 描述离散程度的统计量

名 称	公式(原始数据)	公式(分组数据)	意 义
极差 R	$R = \text{最大值} - \text{最小值}$	$R \approx \text{最高组上限值} - \text{最低组下限值}$	反映离散程度的最简单测度值,不能反映中间数据的离散性
总体方差 σ^2	$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$	$\sigma^2 \approx \frac{1}{N} \sum_{i=1}^k (m_i - \bar{x})^2 f_i$	反映每个总体数据偏离其总体均值的平均程度,是离散程度的最重要测度值,其中标准差具有与观察值数据相同的量纲
总体标准差 σ	$\sigma = \sqrt{\sigma^2}$ $= \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$	$\sigma = \sqrt{\sigma^2}$ $\approx \sqrt{\frac{1}{N} \sum_{i=1}^k (m_i - \bar{x})^2 f_i}$	
样本方差 S^2	$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$	$S^2 = \frac{1}{n-1} \sum_{i=1}^k (m_i - \bar{x})^2 f_i$	反映每个样本数据偏离其样本均值的平均程度,是离散程度的最重要测度值,其中标准差具有与观察值数据相同的量纲
样本标准差 S	$S = \sqrt{S^2}$ $= \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$	$S = \sqrt{S^2}$ $= \sqrt{\frac{1}{n-1} \sum_{i=1}^k (m_i - \bar{x})^2 f_i}$	
变异系数 CV	$CV = \frac{S}{ \bar{x} } \times 100\%$		反映数据偏离其均值的相对偏差,是无量纲的相对变异性测度
样本标准误 $S_{\bar{x}}$	$S_{\bar{x}} = \frac{S}{\sqrt{n}}$		反映样本均值偏离总体均值的平均程度,在用样本均值估计总体均值时测度偏差

3. 描述分布形状的统计量

名 称	公式(原始数据)	公式(分组数据)	意 义
偏度 S_k	$S_k = \frac{n \sum (x_i - \bar{x})^3}{(n-1)(n-2)S^3}$	$S_k = \frac{\sum_{i=1}^k (m_i - \bar{x})^3 f_i}{nS^3}$	反映数据分布的非对称性 $S_k = 0$ 时为对称; $S_k > 0$ 时为正偏或右偏; $S_k < 0$ 时为负偏或左偏

续表

名 称	公式(原始数据)	公式(分组数据)	意 义
峰度 K_u	$K_u = \frac{n(n+1) \sum (x_i - \bar{x})^4 - 3 [\sum (x_i - \bar{x})^2]^2 (n-1)}{(n-1)(n-2)(n-3)S^4}$ (原始数据)	$K_u = \frac{\sum_{i=1}^k (m_i - \bar{x})^4 f_i}{nS^4} - 3$ (分组数据)	反映数据分布的平峰或尖峰程度 $K_u = 0$ 时为标准正态分布; $K_u > 0$ 时为尖峰分布; $K_u < 0$ 时为扁平分布

注:在分组数据公式中, m_i, f_i 分别为各组的组中值和观察值出现的频数。

三、综合例题解析

例 1 证明:各数据观察值与其均值之差的平方和(称为离差平方和)最小,即对任意常数 C ,有

$$\sum_{i=1}^n (x_i - \bar{x})^2 \leq \sum_{i=1}^n (x_i - C)^2$$

证一:设

$$f(C) = \sum_{i=1}^n (x_i - C)^2$$

由函数极值的求法,对上式求导数,得

$$f'(C) = -2 \sum_{i=1}^n (x_i - C) = -2 \sum_{i=1}^n x_i + 2nC, \quad f''(C) = 2n$$

令 $f'(C) = 0$, 得唯一驻点

$$C = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$$

由于 $f''(\bar{x}) = 2n > 0$, 故当 $C = \bar{x}$ 时 $f(C)$ 有最小值,其最小值为

$$f(\bar{x}) = \sum_{i=1}^n (x_i - \bar{x})^2.$$

证二:因为对任意常数 C 有

$$\begin{aligned}
 \sum_{i=1}^n (x_i - \bar{x})^2 - \sum_{i=1}^n (x_i - C)^2 &= \sum_{i=1}^n x_i^2 - n\bar{x}^2 - \left(\sum_{i=1}^n x_i^2 - 2C \sum_{i=1}^n x_i + nC^2 \right) \\
 &= -n\bar{x}^2 + 2C \sum_{i=1}^n x_i - nC^2 = -n(\bar{x}^2 - 2C\bar{x} + C^2) \\
 &= -n(\bar{x} - C)^2 \leq 0
 \end{aligned}$$

故有

$$\sum_{i=1}^n (x_i - \bar{x})^2 \leq \sum_{i=1}^n (x_i - C)^2.$$

四、习题一解答

1. 在某药合成过程中,测得的转化率(%)如下:

94.3 92.8 92.7 92.6 93.3 92.9 91.8 92.4 93.4 92.6
 92.2 93.0 92.9 92.2 92.4 92.2 92.8 92.4 93.9 92.0
 93.5 93.6 93.0 93.0 93.4 94.2 92.8 93.2 92.2 91.8
 92.5 93.6 93.9 92.4 91.8 93.8 93.6 92.1 92.0 90.8

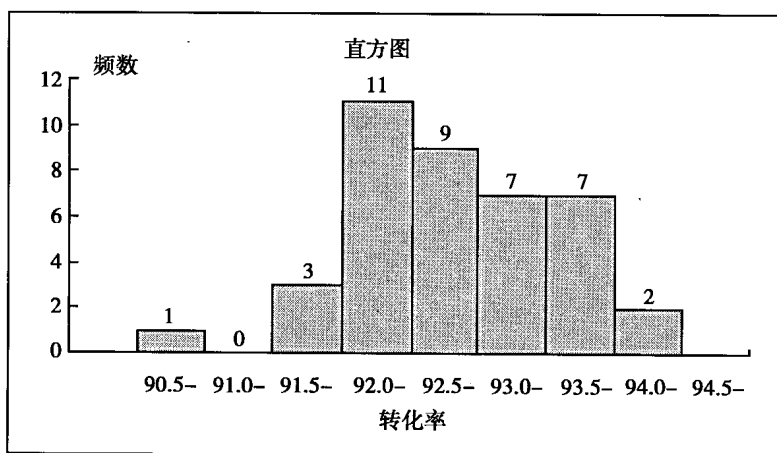
- (1) 取组距为 0.5, 最低组下限为 90.5, 试作出频数分布表;
 (2) 作频数直方图和频率折线图;
 (3) 根据频数分布表的分组数据, 计算样本均值和样本标准差。

解: (1) 所求频数分布表:

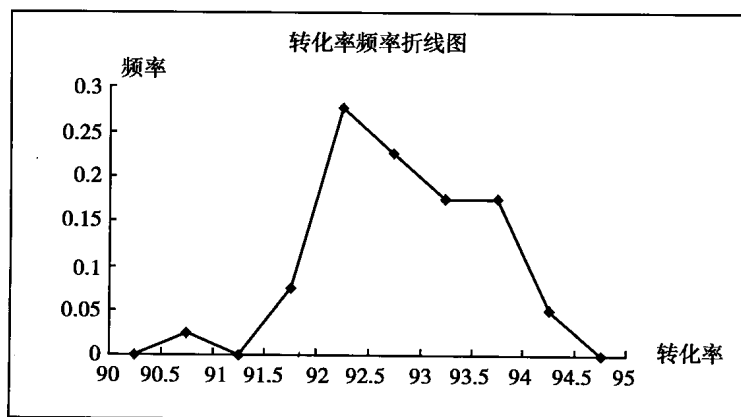
转化率的频数分布表

转化率分组	频数	频率	累积频率
90.5 ~	1	0.025	0.025
91.0 ~	0	0.00	0.025
91.5 ~	3	0.075	0.10
92.0 ~	11	0.275	0.375
92.5 ~	9	0.225	0.60
93.0 ~	7	0.175	0.775
93.5 ~	7	0.175	0.95
94.0 ~ 94.5	2	0.05	1.00

(2) 频数直方图:



频率折线图:



(3) 由频数分布表可得

转化率分组	组中值 m_i	频数
90.5 ~	90.75	1
91.0 ~	91.25	0
91.5 ~	91.75	3
92.0 ~	92.25	11
92.5 ~	92.75	9
93.0 ~	93.25	7
93.5 ~	93.75	7
94.0 ~ 94.5	94.25	2

$$\text{则 } \bar{x} \approx \frac{1}{n} \sum_{i=1}^8 m_i f_i = \frac{90.75 \times 1 + 91.25 \times 0 + \cdots + 94.25 \times 2}{40} = \frac{3\,713}{40} = 92.825$$

$$\begin{aligned} S^2 &\approx \frac{1}{n-1} \sum_{i=1}^8 (m_i - \bar{x})^2 f_i \\ &= \frac{1}{39} [(90.75 - 92.825)^2 \times 1 + (91.25 - 92.825)^2 \times 0 + \cdots + (94.25 - 92.825)^2 \times 2] \\ &= 0.584 \end{aligned}$$

$$\begin{aligned} \text{或者 } S^2 &\approx \frac{1}{n-1} \left(\sum_{i=1}^8 m_i^2 f_i - n\bar{x}^2 \right) \\ &= \frac{1}{39} (90.75^2 \times 1 + 91.25^2 \times 0 + \cdots + 94.25^2 \times 2 - 40 \times 92.825^2) = 0.584 \end{aligned}$$

$$S = \sqrt{S^2} = \sqrt{0.584} \approx 0.7642$$

2. 测得 10 名接触某种病毒的工人的白细胞 ($10^9/\text{L}$) 如下:

7.1, 6.5, 7.4, 6.35, 6.8, 7.25, 6.6, 7.8, 6.0, 5.95

(1) 计算其样本均值、方差、标准差、标准误和变异系数;

(2) 求出该组数据对应的标准化值;

(3) 计算其偏度。

$$\begin{aligned} \text{解: (1) } \sum_{i=1}^{10} x_i &= 7.1 + 6.5 + \cdots + 5.95 = 67.75, \quad n = 10 \\ \sum_{i=1}^{10} x_i^2 &= 7.1^2 + 6.5^2 + \cdots + 5.95^2 = 462.35 \end{aligned}$$

$$\text{样本均值: } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{67.75}{10} = 6.775$$

$$\text{方差: } S^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right) = \frac{1}{9} (462.35 - 10 \times 6.775^2) = 0.371$$

$$\text{标准差: } S = \sqrt{S^2} = \sqrt{0.371} \approx 0.609$$

$$\text{标准误: } S_{\bar{x}} = \frac{S}{\sqrt{n}} = \frac{0.609}{\sqrt{10}} = 0.193$$

$$\text{变异系数: } CV = \frac{S}{|\bar{x}|} \times 100\% = \frac{0.609}{6.775} \times 100\% = 8.99\%;$$

(2) 对应的标准化值公式为

$$u_i = \frac{x_i - \bar{x}}{S} = \frac{x_i - 6.775}{0.609}$$

对应的标准化值为

0.534, -0.452, 1.026, -0.698, 0.041, 0.78, -0.287, 1.683, -1.273, -1.355;

$$(3) S_k = \frac{n \sum (x_i - \bar{x})^3}{(n-1)(n-2)S^3} = 0.204。$$

3. 已知某年某城市居民家庭月人均支出分组数据如下表所示

按月人均支出分组(元)	家庭户数占总户数的比例(%)
200 以下	1.5
200 ~	18.2
500 ~	46.8
800 ~	25.3
1 000 以上	8.2
合计	100.0

试计算(1) 该市平均每户月人均支出的均值和标准差;

(2) 指出其月人均支出的中位数与众数所在组。

解:(1) 由原分组数据表可得

支出分组(元)	组中值	比例(%)
200 以下	100	1.5
200 ~	350	18.2
500 ~	650	46.8
800 ~	900	25.3
1 000 以上	1 100	8.2

$$\text{则 } \bar{x} \approx \frac{1}{n} \sum_{i=1}^5 m_i f_i = \frac{1}{100} (100 \times 1.5 + 350 \times 18.2 + \cdots + 1\,100 \times 8.2) = 687.3$$

$$\begin{aligned} S^2 &\approx \frac{1}{n-1} \left(\sum_{i=1}^5 m_i^2 f_i - n \bar{x}^2 \right) \\ &= \frac{1}{99} (100^2 \times 1.5 + 350^2 \times 18.2 + \cdots + 1\,100^2 \times 8.2 - 5 \times 687.3^2) \\ &= 52\,468.39 \end{aligned}$$

$$S = \sqrt{S^2} = \sqrt{52\,468.39} = 229.06;$$

(2) 由原分组数据表可得

支出分组(元)	比例(%)	累积比例(%)
200 以下	1.5	1.5
200 ~	18.2	19.7
500 ~	46.8	66.5
800 ~	25.3	91.8
1 000 以上	8.2	100.0

中位数所在组,即累积比例超过 50 的那个最低组,即为 500 ~ 组。

众数所在组是频数即比例最大的组,也是 500 ~ 组。

4. 设 $x_1, x_2, \dots, x_n$ 和 $y_1, y_2, \dots, y_n$ 为两组样本观察值,它们有下列关系:

$$y_i = \frac{x_i - a}{b}, \quad i = 1, 2, \dots, n$$

其中 a, b 为常数且 $b \neq 0$, 求样本均值 $\bar{x}$ 与 $\bar{y}$ 及样本方差 S_x^2 和 S_y^2 之间的关系。

$$\text{解: } \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - a}{b} \right) = \frac{1}{b} \left(\frac{1}{n} \sum_{i=1}^n x_i - \frac{na}{n} \right) = \frac{\bar{x} - a}{b}$$

$$\begin{aligned} S_y^2 &= \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{x_i - a}{b} - \frac{\bar{x} - a}{b} \right)^2 = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{b} \right)^2 \\ &= \frac{1}{b^2} \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{b^2} S_x^2. \end{aligned}$$

五、思考与练习

(一) 填空题

1. 统计数据可以分为_____数据、_____数据、_____数据等三类,其中_____数据、_____数据属于定性数据。
2. 常用于表示定性数据整理结果的统计图有_____;而_____,_____、_____、_____是专用于表示定量数据的特征和规律的统计图。
3. 用于数据整理和统计分析的常用统计软件有_____,_____,_____等。
4. 描述数据集中趋势的常用测度值主要有_____,_____和_____等,其中最重要的是_____;描述数据离散程度的常用测度值主要有_____,_____,_____等,其中最重要的是_____。

(二) 选择题

1. 各样本观察值均加同一常数 C 后()。
 - A. 样本均值不变,样本标准差改变
 - B. 样本均值改变,样本标准差不变
 - C. 两者均不变
 - D. 两者均改变
2. 关于样本标准差,以下哪项是错误的()。
 - A. 反映样本观察值的离散程度
 - B. 度量了数据偏离样本均值的大小
 - C. 反映了均值代表性的好坏
 - D. 不会小于样本均值
3. 比较腰围和体重两组数据变异度大小宜采用()。
 - A. 变异系数(CV)
 - B. 方差(S^2)
 - C. 极差(R)
 - D. 标准差(S)

(三) 计算题

1. 在某次实验中,用洋地黄溶液分别注入 10 只家鸽内,直至动物死亡。将致死量折算至原来洋地黄叶粉的重量。其数据记录为(单位:mg/kg)

97.3, 91.3, 102, 129, 92.8, 98.4, 96.3, 99.0, 89.2, 90.1

试计算该组数据的样本均值、方差、标准差、标准误和变异系数。

六、思考与练习参考答案

(一) 填充题

1. 定类,定序,数值,定类,定序
2. 条形图、圆形图;直方图、频数折线图、茎叶图、箱形图
3. SAS、SPSS、Excel
4. 均值、众数、中位数,均值,极差、方差、标准差、变异系数,方差、标准差

(二) 选择题

1. B; 2. D; 3. A

(三) 计算题

1. 均值 98.54、方差 132.27、标准差 11.501、标准误 3.637、变异系数 11.67%。

(高祖新)

第二章 随机事件与概率

一、学习目的和要求

1. 掌握事件等的基本概念及运算关系；
2. 熟练掌握古典概率及计算；
3. 理解统计概率、主观概率和概率的公理化定义；
4. 熟练掌握概率的加法公式、乘法公式及计算；
5. 理解并掌握条件概率与事件独立性的概念并进行计算；
6. 掌握并应用全概率和贝叶斯公式进行计算。

二、内 容 提 要

(一) 基本概念

概 念	符号	概率论的定义	集合论的含义
随机试验(试验)	E	具有以下特征的观测或试验： 1. 试验在相同的条件下可重复地进行 2. 试验的所有结果事先已知,且不止一个 3. 每次试验恰好出现其中之一,但试验前无法预知到底出现哪一个结果。	
样本空间	Ω	试验所有可能结果组成的集合,即所有基本事件的全体	全集
基本事件(样本点)	ω	试验的每个不可再分的可能结果,即样本空间的元素	元素
随机事件(事件)	A	试验中可能发生也可能不发生的结果,是由基本事件组成的样本空间的子集	子集
必然事件	Ω	在试验中一定发生的事件	全集
不可能事件	$\emptyset$	在试验中一定不发生的事件,不含任何基本事件	空集

(二) 事件间的关系

关系	符号	概率论的定义	集合论的含义
包含	$A \subset B$	事件 A 的发生必然导致事件 B 的发生	A 是 B 的子集
相等	$A = B$	$A \subset B$ 而且 $B \subset A$	A 与 B 相等
和(并)	$A + B (A \cup B)$	事件 A 与 B 中至少有一个事件发生	A 与 B 的并
积(交)	$AB (A \cap B)$	事件 A 与 B 同时发生	A 与 B 的交
差	$A - B$	事件 A 发生同时 B 不发生	A 与 B 的差
互不相容	$AB = \emptyset$	事件 A 与 B 不可能同时发生	A 与 B 不相交
对立	$\bar{A}$	事件 A 不发生	A 的补集(余集)

(三) 事件的运算规律

运 算 律	公 式
交换律	$A + B = B + A, AB = BA$
结合律	$(A + B) + C = A + (B + C), (AB)C = A(BC)$
分配律	$(A + B)C = AC + BC, A + (BC) = (A + B)(A + C)$
差积转换律	$A - B = A\bar{B} = A - AB$
对立律	$A\bar{A} = \emptyset, A + \bar{A} = \Omega$
德·摩根对偶律	$\overline{A + B} = \bar{A}\bar{B}, \overline{AB} = \bar{A} + \bar{B}$

(四) 概率的定义

类 型	定 义 公 式
古典概率	$P(A) = \frac{m}{n} = \frac{A \text{ 所含的基本事件数}}{\text{基本事件总数}}$
统计概率	$P(A) = p \left(\approx f_n(A) = \frac{n_A}{n} \right)$
公理化定义(基本性质)	对样本空间中任意事件 A 对应的一个实数 $P(A)$, 满足 公理 1(非负性): $0 \leq P(A) \leq 1$ 公理 2(规范性): $P(\Omega) = 1, P(\emptyset) = 0$ 公理 3(可加性): 若 $A_1, A_2, \dots, A_n, \dots$, 两两互不相容, $P(A_1 + A_2 + \dots + A_n + \dots) = P(A_1) + P(A_2) + \dots + P(A_n) + \dots$ 则称 $P(A)$ 为随机事件 A 的概率。

(五) 概率的计算公式

名 称	计 算 公 式
加法公式	$P(A + B) = P(A) + P(B) - P(AB)$
	若 A, B 互不相容($AB = \emptyset$): $P(A + B) = P(A) + P(B)$
对立事件公式	$P(A) = 1 - P(\bar{A}); P(\bar{A}) = 1 - P(A)$