

高等医药院校教材

(供药学专业用)

# 数理统计方法

黄志宏 方积乾

主编

人民卫生出版社

高等医药院校教材

(供药学专业用)

# 数理统计方法

**黄志宏** 方积乾 主编

编审小组

|            |              |
|------------|--------------|
| <b>黄志宏</b> | 宋学源 (沈阳药学院)  |
| 姚金华        | 方积乾 (北京医科大学) |
| 刘定远        | (华西医科大学)     |
| 王 珍        | (上海医科大学)     |
| 李柏新        | (南京药学院)      |

人 民 卫 生 出 版 社

**数理统计方法**

**黄志宏** 方积乾 主编

人民卫生出版社 出版  
(北京市崇文区天坛西里 10 号)

人民卫生出版社 印刷厂印刷  
新华书店北京发行所发行

787×1092毫米16开本 17 $\frac{1}{2}$ 印张 1 插页 413千字  
1987年10月第1版 1987年10月第1版第1次印刷  
印数：00,001—7,600

ISBN 7-117-00061-9/R·62 定价：2.45元

统一书号：14048·5498

## 序 言

数理统计是研究随机现象数量规律的一门数学学科。它在自然科学、社会科学和工农业生产中都有广泛的应用。特别是在医学、药学和药物生产中，数理统计已成为一门不可缺少的学问。例如，对于临床观察、计量诊断、处方选择、工艺改进、质量控制、药物评价、生物检定及试验设计等数理统计都是十分必要的。同时，它也是药理学、毒理学、药物动力学以及临床医学和预防医学的一门基础理论课。

数理统计的理论基础是概率论，因此本书在第一、二两章中简要地介绍数理统计所需的一些基本的概率论知识；第三章是用概率论的原理研究数理统计的几个重要概念和分布；第四章至第九章，较系统地讨论医药学中常用的统计方法，主要包括参数估计、假设检验、方差分析、回归分析和正交试验等。这部分内容以介绍统计方法为主，对于必要的基本理论问题，尽可能地给出证明，并通过实际例子示范性地说明方法的应用。

在教学上，本书有较大的适应性。一般说，若不讲星号所指的内容，约需 60 学时；如果从中删去较难的概率知识和部分比较不常用的统计方法，40 至 50 学时也能讲授，讲授全书约需 72 学时。

# 目 录

|                               |    |
|-------------------------------|----|
| <b>第一章 随机事件的概率及其计算</b> .....  | 1  |
| 第一节 随机事件的概率 .....             | 1  |
| 一、随机事件 .....                  | 1  |
| 二、频率和概率 .....                 | 1  |
| 三、古典概型 .....                  | 3  |
| 第二节 事件间的相互关系 .....            | 4  |
| 一、事件间的基本关系 .....              | 4  |
| 二、事件的运算规律 .....               | 6  |
| 练习 1·1 .....                  | 7  |
| 第三节 概率的加法公式 .....             | 8  |
| 第四节 概率的乘法公式 .....             | 10 |
| 一、条件概率 .....                  | 10 |
| 二、乘法公式 .....                  | 11 |
| 三、事件独立的条件 .....               | 12 |
| 练习 1·2 .....                  | 14 |
| 第五节 全概率公式与逆概率公式 .....         | 14 |
| 一、全概率公式 .....                 | 14 |
| 二、逆概率公式(贝叶斯公式) .....          | 16 |
| 三、贝叶斯公式在鉴别诊断中的应用 .....        | 18 |
| 练习 1·3 .....                  | 20 |
| 习题一 .....                     | 20 |
| <b>第二章 随机变量的分布和数字特征</b> ..... | 22 |
| 第一节 随机变量与离散型随机变量的分布 .....     | 22 |
| 一、随机变量 .....                  | 22 |
| 二、离散型随机变量与概率函数 .....          | 23 |
| 三、二项分布与泊松分布 .....             | 24 |
| 练习 2·1 .....                  | 30 |
| 第二节 分布函数与连续型随机变量的分布 .....     | 31 |
| 一、随机变量的分布函数 .....             | 31 |
| 二、连续型随机变量与概率密度函数 .....        | 32 |
| 三、正态分布、对数正态分布与威布尔分布 .....     | 33 |
| 第三节 随机向量 .....                | 37 |
| 一、随机向量的分布 .....               | 37 |
| 二、独立随机变量 .....                | 40 |
| 练习 2·2 .....                  | 41 |
| 第四节 数学期望 .....                | 42 |
| 一、数学期望的概念 .....               | 42 |
| 二、数学期望的性质 .....               | 43 |

|                                 |     |
|---------------------------------|-----|
| 三、常见的几种随机变量的数学期望                | 45  |
| 第五节 方差                          | 46  |
| 一、方差的概念                         | 46  |
| 二、方差的性质                         | 47  |
| 三、常见的几种随机变量的方差                  | 48  |
| 第六节 大数定理与中心极限定理                 | 52  |
| 一、大数定理                          | 52  |
| 二、中心极限定理                        | 53  |
| 练习 2·3                          | 55  |
| 习题二                             | 55  |
| 第三章 随机抽样与抽样分布                   | 57  |
| 第一节 总体与样本                       | 57  |
| 一、总体和个体                         | 57  |
| 二、简单随机样本                        | 57  |
| 第二节 经验分布与统计量                    | 58  |
| 一、经验分布                          | 58  |
| 二、统计量                           | 60  |
| 练习 3·1                          | 61  |
| 第三节 和的分布与均数的分布                  | 61  |
| 一、和的分布                          | 62  |
| 二、均数 $\bar{X}$ 的分布的证明           | 63  |
| 第四节 $\chi^2$ 分布、 $t$ 分布和 $F$ 分布 | 64  |
| 一、 $\chi^2$ 分布                  | 64  |
| 二、 $t$ 分布                       | 65  |
| 三、 $F$ 分布                       | 66  |
| 练习 3·2                          | 68  |
| 习题三                             | 68  |
| 第四章 连续型资料的分析                    | 69  |
| 第一节 点估计                         | 69  |
| 一、点估计及其性质                       | 69  |
| 二*、最大似然估计                       | 71  |
| 练习 4·1                          | 73  |
| 第二节 一个正态总体的区间估计和假设检验            | 74  |
| 一、区间估计                          | 74  |
| 二、假设检验                          | 79  |
| 练习 4·2                          | 87  |
| 第三节 两个正态总体的假设检验                 | 88  |
| 练习 4·3                          | 100 |
| 习题四                             | 101 |
| 第五章 离散型资料的分析                    | 104 |
| 第一节 区间估计和假设检验                   | 104 |
| 一、区间估计                          | 104 |

|                      |     |
|----------------------|-----|
| 二、假设检验               | 109 |
| 练习 5.1               | 115 |
| 第二节 离散型的 $\chi^2$ 检验 | 116 |
| 一、拟合优度检验             | 116 |
| 二、列联表中的独立性检验         | 119 |
| 练习 5.2               | 124 |
| 习题五                  | 125 |
| 第六章 非参数检验与概率纸        | 127 |
| 第一节 非参数检验            | 127 |
| 一、成对资料的符号检验          | 127 |
| 二、秩和检验               | 128 |
| 三、多组资料的秩和检验          | 128 |
| 练习 6.1               | 131 |
| 第二节 经验分布函数与三种概率纸     | 132 |
| 一、正态概率纸              | 132 |
| 二、对数正态概率纸            | 135 |
| 三、威布尔概率纸             | 137 |
| 练习 6.2               | 139 |
| 习题六                  | 139 |
| 第七章 方差分析             | 141 |
| 第一节 单因素试验的方差分析       | 141 |
| 一、方差分析的基本思想和步骤       | 141 |
| 二、单因素试验方差分析的计算       | 143 |
| 三、两两间多重比较的检验法        | 147 |
| 练习 7.1               | 151 |
| 第二节 两因素试验的方差分析       | 152 |
| 一、无重复试验的情形           | 152 |
| 二、有重复试验的情形           | 157 |
| 练习 7.2               | 161 |
| 习题七                  | 162 |
| 第八章 相关与回归            | 164 |
| 第一节 相关               | 164 |
| 一、散点图                | 164 |
| 二、相关系数的定义            | 165 |
| 三、相关系数的检验            | 166 |
| 四、相关系数的计算            | 167 |
| 练习 8.1               | 168 |
| 第二节 关于单个自变量的线性回归     | 169 |
| 一、单个自变量线性回归的计算       | 169 |
| 二*、参数估计量的分布          | 174 |
| 三、建立回归方程后进一步的统计分析    | 177 |
| 四*、关于回归的两个推广         | 181 |

|                                                   |     |
|---------------------------------------------------|-----|
| 练习 8.2                                            | 184 |
| 第三节* 失拟情况的检验与两条回归直线的比较                            | 185 |
| 一、回归方程拟合优度的检验                                     | 185 |
| 二、两条回归直线的比较                                       | 190 |
| 练习 8.3                                            | 192 |
| 第四节* $ED_{50}$ 或 $LD_{50}$ 估计的概率单位法               | 192 |
| 练习 8.4                                            | 197 |
| 第五节* 两个自变量的线性回归                                   | 197 |
| 习题八                                               | 201 |
| 第九章 正交试验设计                                        | 203 |
| 第一节 正交表与试验设计                                      | 203 |
| 第二节 多指标的试验                                        | 206 |
| 第三节 有交互作用的设计                                      | 210 |
| 第四节 正交试验的方差分析                                     | 213 |
| 习题九                                               | 216 |
| 附表                                                | 219 |
| 附表 1. 二项分布表                                       | 219 |
| 附表 2. 泊松 (Poisson) 分布表                            | 221 |
| 附表 3. 标准正态分布的密度函数表                                | 227 |
| 附表 4. 标准正态分布表                                     | 228 |
| 附表 5. 正态分布的双侧分位数 ( $u_{1-\frac{\alpha}{2}}$ ) 表   | 230 |
| 附表 6. $\Gamma(1 + \frac{1}{m})$ 的数值表              | 230 |
| 附表 7. $\chi^2$ 分布的上侧分位数 ( $\chi^2_{1-\alpha}$ ) 表 | 231 |
| 附表 8. $t$ 分布的双侧分位数 ( $t_{1-\frac{\alpha}{2}}$ ) 表 | 232 |
| 附表 9. 用极差作 $t$ 检验的临界值表                            | 233 |
| 附表 10. 符号检验表                                      | 234 |
| 附表 11. 秩和检验表                                      | 234 |
| 附表 12. $F$ 检验的临界值 ( $F_{1-\alpha}$ ) 表            | 235 |
| 附表 13. 二项分布参数 $p$ 的置信区间表                          | 241 |
| 附表 14. 泊松 (Poisson) 分布参数 $\lambda$ 的置信区间表         | 245 |
| 附表 15. 多重比较中的 $q$ 表                               | 246 |
| 附表 16. 多重比较中的 $s$ 表                               | 249 |
| 附表 17. 相关系数临界值表                                   | 250 |
| 附表 18. 概率单位和权重系数表                                 | 250 |
| 附表 19. 常用正交表                                      | 251 |
| 习题答案                                              | 272 |

# 第一章 随机事件的概率及其计算

## 第一节 随机事件的概率

### 一、随机事件

自然界里有各种现象. 有一类现象在一定条件下必然发生. 例如, 地球上, 在标准大气压下, 摄氏 100 度的水必然沸腾; 上抛的石子必然下落. 这类现象称为确定性现象. 过去我们学过的微积分知识就是研究这类现象的一种数学.

还有一类现象, 在同一条件下有不确定的结果. 例如, 投掷一枚硬币, 结果可能是正面 (图案面), 也可能是反面 (币额面). 对一次投掷来说, 究竟出现哪一种结果, 在投掷前是不能预先确定的; 在同一条件下生产的一批针剂中, 有的是合格品, 有的是不合格品; 某种疾病的患者, 服剂量相同的一种药物, 有的痊愈, 有的无效, 有的显效而未痊愈. 这类现象的共同特点是: 在相同条件下进行实验或观察, 每次结果不尽相同, 对一次实验或观察来说, 究竟出现哪一种结果, 事前是不能确定的. 这种在同一条件下进行实验观察具有不确定结果的现象, 就是所谓随机现象. 对随机现象进行实验或观察称为随机试验, 以后简称试验. 随机试验的各种可能结果称为随机事件, 简称事件. 在试验中必然发生的事件称为必然事件; 必然不发生的事件称为不可能事件. 这两种事件也可当作随机事件. 随机事件常用大写字母  $A, B, C, \dots$  表示.

概率论和数理统计就是研究随机现象的数量规律的学科.

### 二、频率和概率

随机事件是一种可能发生、也可能不发生的事件. 看起来似乎没有什么规律可循, 其实不然. 当我们在同一条件下进行大量试验时, 就会呈现某种规律性.

【例 1】有人作投掷硬币的试验, 结果列于表 1.1.

表 1.1 硬币投掷试验

| 试验者      | 投掷次数 ( $n$ ) | 出现正面次数 ( $m$ ) | 频率 ( $m/n$ ) |
|----------|--------------|----------------|--------------|
| Demorgen | 2046         | 1061           | 0.5186       |
| Buffon   | 4040         | 2048           | 0.5069       |
| Pearson  | 12000        | 6019           | 0.5016       |
| Pearson  | 24000        | 12012          | 0.5005       |

对单独一次投掷来说, 出现正面还是反面, 预先是不能确定的, 但是随着试验次数的大量增加, 可以看出, 出现正面的次数接近投掷次数的百分之五十.

若随机事件  $A$  在  $n$  次重复试验中出现了  $m$  次, 则比值  $m/n$  称为事件  $A$  出现的频率:

$$\varphi_n(A) = m/n \quad (1)$$

表 1·1 告诉我们, 随着投掷次数的增加, 出现正面的频率稳定地接近于 0.5.

【例 2】某地 1927 年至 1932 年婴儿出生数如表 1·2 所示.

表 1·2 某地婴儿出生数

| 出生年份  | 出生数       |           |             | 频率     |        |
|-------|-----------|-----------|-------------|--------|--------|
|       | 男孩( $m$ ) | 女孩( $k$ ) | 合计( $m+k$ ) | 男孩 $p$ | 女孩 $q$ |
| 1927  | 496544    | 462189    | 958733      | 0.518  | 0.482  |
| 1928  | 513654    | 477339    | 990993      | 0.518  | 0.482  |
| 1929  | 514765    | 479336    | 994101      | 0.518  | 0.482  |
| 1930  | 528072    | 494739    | 1022811     | 0.516  | 0.484  |
| 1931  | 496986    | 467587    | 964573      | 0.515  | 0.485  |
| 1932  | 482431    | 452232    | 934663      | 0.516  | 0.484  |
| 合计或平均 | 3032452   | 2833422   | 5865874     | 0.517  | 0.483  |

对单独一个生育事件, 生男生女是不能预先确定的, 因此, 这是一个随机事件. 但是大量的生育事件却有其规律性. 表 1·2 告诉我们, 生育事件中, 生男孩的频率在 0.517 附近摆动.

可以看出, 随机事件的频率总是围绕着一个稳定值摆动. 这个稳定值正好代表了随机事件出现的可能性大小. 这就是概率概念的客观背景.

**定义 1** 设在同一条件下, 重复进行  $n$  次试验, 随机事件  $A$  出现  $m$  次. 若试验次数  $n$  充分大以后, 频率  $m/n$  稳定地在某一确定值  $p$  的附近摆动, 则称  $p$  为事件  $A$  的概率, 记作  $P(A) = p$ .

这个定义通过频率的客观性不仅说明了概率的客观性, 而且还指出了求概率的近似值的方法, 即经过大量试验后, 可用事件  $A$  的频率作为所求概率的近似值.

$$P(A) = p \approx m/n. \quad (2)$$

【例 3】某医院用新药治疗老年性气管炎, 其疗效如下:

| 治疗结果         | 临床治愈  | 明显好转  | 症状缓解  | 无效    | 合 计        |
|--------------|-------|-------|-------|-------|------------|
| 例数 ( $m$ )   | 83    | 180   | 117   | 23    | 403( $n$ ) |
| 频率 ( $m/n$ ) | 0.206 | 0.477 | 0.290 | 0.057 | 1.00       |

这里, 感兴趣的问题应该是临床治愈率 (概率) 是多少, 而不是 403 例的治愈频率. 由概率的定义知道, 当  $n$  比较大时, 频率可以作为概率的近似值. 所以, 分别算出各治疗结果的频率, 以估计其相应的概率. 当然, 如果治疗例数非常多, 这个近似值就会更说明问题.

显然, 任何事件  $A$  的概率满足

$$0 \leq P(A) \leq 1$$

(3)

必然事件 $U$ 的概率为1, 不可能事件 $V$ 的概率为0, 即

$$P(U)=1 \quad P(V)=0.$$

### 三、古典概型

有些事件的概率不用进行大量重复的试验也能确定, 象例1中的事件. 因为硬币是比较均匀的, 可以认为投掷一次时, 出现正面和出现反面具有同等可能性, 谁也没有优先出现的理由, 而且每次只能出现其中的一个结果. 所以出现正面的可能性的概率是 $\frac{1}{2}$ . 即概率 $P(\text{正}) = \frac{1}{2}$ . 这就是说, 可以用划分等可能事件的个数的方法求得事件的概率.

【例4】 设一个口袋里装有大小相等、质量相同的各色球16个, 其中, 白球2个, 红球3个, 黄球5个, 黑球6个. 从中任摸一球, 问摸得白、红、黄、黑球的概率分别是多少?

解: 从袋内摸球, 共有16种可能结果, 每次只能出现一种结果, 且每一种结果都是等可能的, 所以总共可以划分出16种等可能事件, 每次只能出现其中的一种.

以 $A_1, A_2, A_3, A_4$ 分别表示摸得白球、红球、黄球、黑球等事件. 因为事件 $A_1$ 包含两个等可能事件, 所以它的概率是 $P(A_1) = 2/16 = 1/8$ . 同样, 由于 $A_2, A_3, A_4$ 分别包含3个、5个、6个等可能事件, 所以, 它们的概率分别为

$$P(A_2) = 3/16, \quad P(A_3) = 5/16, \quad P(A_4) = 6/16 = 3/8.$$

令 $B = \{\text{摸得有色球}\}$ , 则 $B$ 包含14个等可能事件. 于是

$$P(B) = 14/16 = 7/8.$$

依照例4的思路, 我们可以总结出一种求概率的方法. 为此, 先给出等概基本事件组的定义.

**定义2** 如果一组事件 $e_1, e_2, \dots, e_N$ 满足以下三个条件, 则称该事件组为等概基本事件组.

- (1)  $N$ 个事件中, 每一个出现的概率是相等的 (等可能性).
- (2) 经一次试验,  $N$ 个事件中只能出现一个 (互不相容性).
- (3) 经一次试验,  $N$ 个事件中必然出现一个 (完备性).

如果在等概基本事件组的 $N$ 个等概基本事件中, 事件 $A$ 包含 $M$  ( $M \leq N$ ) 个等概基本事件, 则事件 $A$ 发生的概率为

$$P(A) = \frac{M}{N} = \frac{A \text{ 所包含的等概基本事件个数}}{\text{等概基本事件总数}}. \quad (4)$$

这种用等概基本事件的个数来计算概率的模型称为古典概型.

【例5】 瓶中装有50片药, 其中有3片次品. 今自瓶中任取5片, 求所取5片中有2片是次品的概率.

解: 自50片中任取5片, 总共有 $C_{50}^5$ 种机会均等的取法, 每次只能是一种取法. 其中, 5片中有2片次品和3片正品的情况占 $C_3^2 \cdot C_{47}^3$ 种, 故所求概率为

$$P = \frac{C_3^2 \cdot C_{47}^3}{C_{50}^5} = \frac{9}{392} = 0.0230.$$

【例 6】男人的性染色体为  $(X, Y)$ ，女人的性染色体为  $(X, X)$ ，某种遗传病（如血友病）反映在染色体  $X$  上有问题，这种染色体记作  $X^*$ 。某人的性染色体对若为  $(X^*, Y)$  或  $(X^*, X^*)$  时，便是遗传病患者；性染色体为  $(X^*, X)$  时，遗传病虽不表现，但为遗传病携带者。精卵结合时，父母的各一个染色体被结合在一起。试讨论其子女分别患遗传病或为遗传病携带者的可能性。

解：假设每种结合方式是等可能的，按古典概型，共分六种情况讨论：

(1) 若父母均为遗传病患者，即他们的性染色体分别为  $(X^*, Y)$ ， $(X^*, X^*)$ ，按结合关系，其男性子代染色体为  $(X^*, Y)$ ，女性子代为  $(X^*, X^*)$ 。因此子代中无论男女都为遗传病患者。即男女子代为遗传病患者的概率都是 1。

(2) 若父亲的染色体为  $(X^*, Y)$ ，母亲为  $(X^*, X)$ ，按结合关系，则男性子代为  $(X, Y)$ ，或  $(X^*, Y)$ ，女性子代为  $(X^*, X)$  或  $(X^*, X^*)$ ，即男性子代中有  $\frac{1}{2}$  为遗传病患者；女性中有  $\frac{1}{2}$  为遗传病患者， $\frac{1}{2}$  为遗传病携带者。

(3) 若父亲的染色体为  $(X^*, Y)$ ，母亲为  $(X, X)$ ，则男性子代为  $(X, Y)$ ，女性子代为  $(X^*, X)$ ，即子代中男性无遗传病，女性必为遗传病携带者。

(4) 若父亲的染色体为  $(X, Y)$ ，母亲为  $(X^*, X^*)$ ，则男性子代为  $(X^*, Y)$ ，女性子代为  $(X^*, X)$ 。即子代中男性必为遗传病患者；女性必为遗传病携带者。

(5) 若父亲的染色体为  $(X, Y)$ ，母亲为  $(X^*, X)$ ，则男性子代为  $(X, Y)$ ， $(X^*, Y)$ ，女性子代为  $(X, X)$ ， $(X^*, X)$ ，即子代中男性有  $\frac{1}{2}$  为遗传病患者；女性有  $\frac{1}{2}$  为遗传病携带者。

(6) 若父母的染色体都正常，显然，其子代无论男女均正常。

从优生学的观点看，上述(1)、(2)、(4)三种情况的婚配是不合适的，而(3)、(5)两种情况的婚配在生育时应先作产前诊断，然后决定取舍。

最后指出两点：第一，上面的概率定义有较大的局限性，更一般的是概率的公理化定义。但是这已超出了本书的范围；第二，通常一个随机试验的所有可能结果只具有完备性和互不相容性，而不具备等可能性。这时我们称所有可能结果为结果空间或样本空间，记作  $\Omega$ 。结果空间内的每一个事件被称为基本事件。

## 第二节 事件间的相互关系

运用古典概型计算概率是比较难的，借以作较复杂事件的概率计算就更难了。但是我们已经看到，一个事件  $A$  常常包含若干个等概基本事件；一个比较复杂的事件常常可能包含若干简单事件。因此，弄清事件之间的相互关系对于把复杂事件的概率分解成若干简单事件的概率来算是很必要的。为此先定义事件间的一些基本关系，然后讨论事件的运算规律。

### 一、事件间的基本关系

#### 1 包含与相等

两个事件  $A$  和  $B$ ，如果事件  $A$  的发生必然导致事件  $B$  的发生，则称事件  $B$  包含事件

A. 记作  $B \supseteq A$  或  $A \subseteq B$ .

若事件  $A$  包含事件  $B$ , 同时事件  $B$  又包含事件  $A$ , 即  $A \supseteq B$  且  $B \supseteq A$ , 则称事件  $A$  与事件  $B$  相等. 记作  $A = B$ .

例如, 在第一节例 4 中, 摸得有色球的事件一定包含摸得红球这一事件. 即  $B \supseteq A_2$ .

## 2 事件的和

事件  $A$  与事件  $B$  中至少有一个发生, 这样的事件称为事件  $A$  与事件  $B$  的和. 记作  $A + B$  (或  $A \cup B$ ).

一般地, 若事件  $A_1, A_2, \dots, A_n$  中至少有一个发生, 这一事件被称为事件  $A_1, A_2, \dots, A_n$  的和, 记作  $A_1 + A_2 + \dots + A_n$ .

例如, 在第一节例 4 里, 摸得有色球就是摸得红、黄、黑三种球中任何一种球的事件,  $B = A_2 + A_3 + A_4$ .

事件  $A$  发生且事件  $B$  不发生, 这一事件称为事件  $A$  与事件  $B$  的差. 记作  $A - B$ .

例如, 在第一节例 4 里, 摸得红球可以说成是摸得有色球但不是黄球也不是黑球. 故  $A_2 = B - (A_3 + A_4)$ .

## 3 事件的积

事件  $A$  与事件  $B$  同时发生, 这一事件称为事件  $A$  与事件  $B$  的积. 记作  $AB$  或  $A \cdot B$  或  $A \cap B$ .

例如, 投掷两枚硬币, 若  $A$  表示事件“第一枚为正面”,  $B$  表示事件“第二枚为反面”,  $C$  表示事件“第一枚为正面且第二枚为反面”. 则  $C = AB$  或  $C = A \cap B$ .

类似可定义: 有限个事件  $A_1, A_2, \dots, A_n$  同时发生, 这一事件被称为  $n$  个事件的积, 记作  $A_1 A_2 \dots A_n$  或  $\prod_{i=1}^n A_i$  或  $A_1 \cap A_2 \cap \dots \cap A_n$ .

## 4 互不相容事件

若事件  $A$  与事件  $B$  不可能同时发生, 则称这两个事件互不相容. 如果  $n$  个事件是两两互不相容的, 则称这  $n$  个事件互不相容.

例如, 在第一节例 4 里, 摸得有色球的事件  $B$  所包含的三个事件: 摸得红球 ( $A_2$ ), 摸得黄球 ( $A_3$ ), 摸得黑球 ( $A_4$ ) 就是两两互不相容的. 显然, 摸得白球 ( $A_1$ ) 及  $A_2, A_3, A_4$  也是两两互不相容的.

一组事件  $A_1, A_2, \dots, A_n$  是两两互不相容, 且它们的和是必然事件, 则称该事件组为互不相容完备事件组. 特别, 当  $AB$  为不可能事件,  $A + B$  为必然事件时, 即  $AB = \emptyset$  且  $A + B = U$  时, 则称事件  $B$  为事件  $A$  的对立事件. 记作  $\bar{A}$ . 显然  $A$  与  $\bar{A}$  互为对立事件.

例如, 在第一节例 4 里, 摸得的不是白球就一定是有色球, 即  $\bar{A}_1 = B$ . 当然, 摸得的不是有色球就一定是白球, 即  $\bar{B} = A_1$ .

显然,  $\bar{\bar{A}} = A$ .

【例 1】一张有某种特殊细胞的片子, 在显微镜下观察时, 它出现在某一位置上 是随机的, 若在  $A$  圈内找到该种细胞用事件  $A$  表示, 在  $B$  圈内找到该种细胞用事件  $B$  表示. 试将事件  $A + B$ ,  $A - B$ ,  $\bar{A}$ ,  $AB$  和  $A \subseteq B$  用几何图形表示出来.

解 图 1.1 的 (a), (b), (c), (d), (e) 的阴影部分就分别表示了事件

$A + B$ ,  $A - B$ ,  $\bar{A}$ ,  $AB$  和  $A \subseteq B$ .

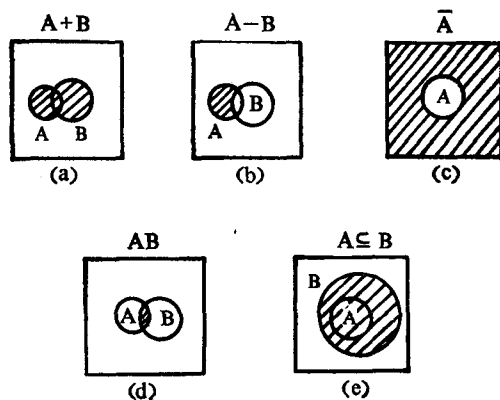


图 1.1

## 二、事件的运算规律

从集合的观点看，任何事件都可看作试验结果集合的子集。例如，生男、生女事件可看作生育集合的子集；四十岁前死亡的事件可看作寿命集合的子集。所以事件就是一种集合。它们的关系如表 1.3 所示。

表 1.3 事件与集合的关系

| 符 号                | 事 件                    | 集 合              |
|--------------------|------------------------|------------------|
| $U$                | 必然事件                   | 全集               |
| $V$                | 不可能事件                  | 空集               |
| $e$                | 基本事件                   | 元素               |
| $A$                | 事件                     | 子集               |
| $\bar{A}$          | $A$ 的对立事件              | $A$ 的余集          |
| $A \subseteq B$    | 事件 $A$ 发生必导致事件 $B$ 发生  | $A$ 是 $B$ 的子集    |
| $A = B$            | 事件 $A$ 与事件 $B$ 相等      | $A$ 与 $B$ 相等     |
| $A + B (A \cup B)$ | 事件“ $A$ 与 $B$ 至少有一个发生” | $A$ 与 $B$ 的和集    |
| $AB (A \cap B)$    | 事件“ $A$ 与 $B$ 同时发生”    | $A$ 与 $B$ 的交集    |
| $A - B$            | 事件“ $A$ 发生而 $B$ 不发生”   | $A$ 与 $B$ 的差集    |
| $AB = V$           | 事件 $A$ 与事件 $B$ 互不相容    | $A$ 与 $B$ 没有公共元素 |

既然事件是一种集合，事件的运算便服从集合运算的规律。常用的集合运算规律有：

- (1)  $A + B = B + A$ ;
- (2)  $(A + B) + C = A + (B + C)$ ;
- (3)  $A + A = A$ ;
- (4)  $A + \bar{A} = U$ ;
- (5)  $A + U = U$ ;
- (6)  $A + V = A$ ;

$$(7) AB = BA;$$

$$(8) (AB)C = A(BC)$$

$$(9) AA = A;$$

$$(10) A\bar{A} = V;$$

$$(11) AU = A;$$

$$(12) AV = V;$$

$$(13) A(B+C) = AB+AC;$$

$$(14) A+BC = (A+B)(A+C);$$

$$(15) \overline{A+B} = \bar{A}\bar{B};$$

$$(16) \overline{AB} = \bar{A} + \bar{B}.$$

读者还可通过几何图形验证这些规律对事件也是符合的。例如规律(14)左端的事件可用图 1·2 的阴影部分表示；而右端的事件则是由两个单影图形叠成的阴影图形（图 1·2 b）。比较 (a) (b) 两图的双影部分，它们完全相同，故

$$A+BC = (A+B)(A+C).$$

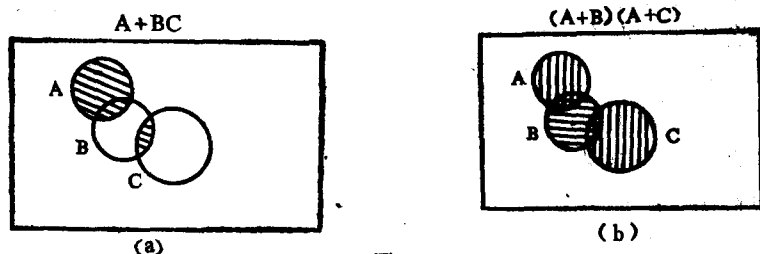


图 1·2

【例 2】 设  $A, B, C$  表示三个随机事件。试将下列事件用  $A, B, C$  表示：

- (1)  $A$  出现， $B, C$  不出现；
- (2) 三个事件都出现；
- (3) 三个事件都不出现；
- (4) 三个事件中至少一个出现；
- (5) 不多于两个事件出现。

解：按事件间的关系及事件运算规律，所列的五种事件分别用  $A, B, C$  表示于下：

$$(1) A\bar{B}\bar{C} = A - (B + C);$$

$$(2) ABC;$$

$$(3) \bar{A}\bar{B}\bar{C} = \overline{A+B+C};$$

$$(4) A+B+C = \overline{\bar{A}\bar{B}\bar{C}};$$

$$(5) \overline{ABC} = \bar{A} + \bar{B} + \bar{C} = \bar{A}\bar{B}\bar{C} + A\bar{B}\bar{C} + \bar{A}B\bar{C} + \bar{A}\bar{B}C + A\bar{B}C + \bar{A}BC + ABC.$$

### 练习 1·1

1. 某事件的频率和概率有何区别？
2. 某市在某年的第一季度出生婴儿的情况为
  - 一月份：男孩 145 个，女孩 135 个；
  - 二月份：男孩 125 个，女孩 136 个；
  - 三月份：男孩 152 个，女孩 140 个。

问该季度生男孩的频率是多少?

3. 一部四卷的文集, 按任意顺序放到书架上. 问各卷自左向右或自右向左的卷号恰好为 1, 2, 3, 4 的概率是多少?

4. 电话号码由 5 个数字组成, 每个数字可以是 0, 1, ..., 9 中的任意一个数. 求电话号码由完全不同数字组成的概率.

5. 袋中有 5 个白球, 3 个红球, 从中任取 4 个, 问恰好抽到 3 个白球的概率是多少?

6. 从 1, 2, 3, 4, 5 号小白鼠中任取两只作新药试验: (1) 两只都用新药; (2) 一只用新药, 一只用安慰剂. 试写出这两种试验的样本空间, 并计算所取两只中有一只是 4 号小白鼠的概率.

7. 何谓互不相容事件? 何谓对立事件? 事件“ $A$ 或 $B$ ”; “ $A$ 且 $B$ ”各是什么含义?

8. 设  $A, B, C$  表示三个随机事件, 试将下列事件用  $A, B, C$  表示出来.

- (1) 三个事件至少有两个出现;
- (2) 三个事件中不多于一个事件出现;
- (3) 恰好有一个事件出现;
- (4) 恰好有两个事件出现;
- (5)  $A, B$  都出现,  $C$  不出现.

### 第三节 概率的加法公式

把复杂事件的概率分解成简单事件的概率, 第一个基本公式是加法公式.

**定理 1** 若  $A, B$  是两个互不相容的事件, 则它们的和事件的概率等于各事件的概率之和, 即

$$P(A+B) = P(A) + P(B). \quad (1)$$

证: 按古典概型给予证明.

设等概基本事件组共有  $n$  个等概基本事件, 其中事件  $A$  由  $m_1$  个等概基本事件组成, 事件  $B$  由  $m_2$  个等概基本事件组成. 由于事件  $A$  与  $B$  互不相容, 因此, 组成事件  $A$  的等概基本事件与组成事件  $B$  的等概基本事件是不相同的. 所以, 事件  $A+B$  应该由  $m_1+m_2$  个等概基本事件组成. 于是,

$$P(A+B) = \frac{m_1+m_2}{n} = \frac{m_1}{n} + \frac{m_2}{n} = P(A) + P(B).$$

**推论 1** 有限个两两互不相容的事件  $A_1, A_2, \dots, A_n$ , 它们的和事件的概率等于各事件的概率之和, 即

$$P(A_1+A_2+\dots+A_n) = P(A_1) + P(A_2) + \dots + P(A_n). \quad (2)$$

**推论 2** 互不相容完备组的各事件的概率之和等于 1.

因为  $P(A_1+A_2+\dots+A_n) = P(U) = 1$ ,

所以,  $P(A_1) + P(A_2) + \dots + P(A_n) = 1$ .

**推论 3** 互相对立的事件的概率满足

$$P(A) = 1 - P(\bar{A}). \quad (3)$$

在第一节例 4 里已经看到, 摸得白、红、黄、黑球的概率分别是:

$$P(A_1) = \frac{1}{8}, \quad P(A_2) = \frac{3}{16}, \quad P(A_3) = \frac{5}{16}, \quad P(A_4) = \frac{3}{8}.$$

因为  $A_1, A_2, A_3, A_4$  互不相容, 可以用定理 1 求得摸有色球的概率:

$$\begin{aligned} P(B) &= P(A_2 + A_3 + A_4) = P(A_2) + P(A_3) + P(A_4) \\ &= \frac{3}{16} + \frac{5}{16} + \frac{3}{8} = \frac{7}{8}. \end{aligned}$$

又因为  $B = \bar{A}_1$ , 所以也可以用对立事件算  $P(B)$ :

$$P(B) = 1 - P(A_1) = 1 - \frac{1}{8} = \frac{7}{8}.$$

**定理 1** 给出的加法公式要求各事件具有互不相容性. 更一般的加法公式由下面的定理给出.

**定理 2** 对任意两个事件  $A$  和  $B$  均有

$$P(A + B) = P(A) + P(B) - P(AB). \quad (4)$$

证: 因为  $B = BU = B(A + \bar{A}) = BA + B\bar{A}$  且  $A \supseteq AB$ , 所以,  $A + B = A + BA + B\bar{A} = A + B\bar{A}$ .

因为  $A$  与  $B\bar{A}$  互不相容,

$$P(A + B) = P(A) + P(B\bar{A})$$

又因为  $BA$  与  $B\bar{A}$  互不相容, 所以

$$P(B) = P(AB) + P(B\bar{A}),$$

$$P(B\bar{A}) = P(B) - P(AB)$$

于是  $P(A + B) = P(A) + P(B) - P(AB)$ , 证毕.

定理 1 是定理 2 的特例. 因为当  $A, B$  互不相容时,  $P(AB) = 0$ .

不难把一般的加法公式推广到有限个事件的情形. 下面仅推出三个事件的加法公式, 三个以上事件的推导方法与此类似.

设  $A, B, C$  为任意三个事件, 则

$$\begin{aligned} P(A + B + C) &= P(A) + P(B) + P(C) - P(AB) \\ &\quad - P(AC) - P(BC) + P(ABC). \end{aligned} \quad (5)$$

证:

$$\begin{aligned} P(A + B + C) &= P[A + (B + C)] \\ &= P(A) + P(B + C) - P[A(B + C)] \\ &= P(A) + P(B) + P(C) - P(BC) - P(AB + AC) \\ &= P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC). \end{aligned}$$