



# 高 2023 届

# 数学课程讲义

# 八分册 (教师用书)

(排列组合与二项式定理、概率与分布、统计、成对数据的统计分析)

巴蜀中学数学教研组

# 目录

|                                      |    |
|--------------------------------------|----|
| 第十七章 排列组合与二项式定理 .....                | 1  |
| 17.1 分类加法计数原理与分步乘法计数原理 .....         | 1  |
| 17.1.1 分类加法计数原理与分步乘法计数原理（第一课时） ..... | 1  |
| 17.1.2 分类加法计数原理与分步乘法计数原理（第二课时） ..... | 5  |
| 17.2 排列 .....                        | 8  |
| 17.2.1 排列（第一课时） .....                | 8  |
| 17.2.2 排列（第二课时） .....                | 12 |
| 17.3 组合 .....                        | 16 |
| 17.3.1 组合（第一课时） .....                | 16 |
| 17.3.2 组合（第二课时） .....                | 20 |
| 17.4 排列与组合的应用 .....                  | 24 |
| 17.5 二项式定理 .....                     | 32 |
| 17.5.1 二项式定理的定义 .....                | 32 |
| 17.5.2 “杨辉三角”与二项式系数的性质 .....         | 35 |
| 17.6 习题课：排列与组合、二项式定理 .....           | 39 |
| 第十八章 概率与分布 .....                     | 42 |
| 18.1 有限样本空间与随机事件 .....               | 42 |
| 18.2 事件的关系和运算 .....                  | 42 |
| 18.3 古典概型 .....                      | 48 |
| 18.4 概率的基本性质 .....                   | 52 |
| 18.5 事件的相互独立性 .....                  | 55 |
| 18.6 频率的稳定性 .....                    | 59 |
| 18.7 随机模拟 .....                      | 59 |
| 18.8 条件概率 .....                      | 63 |
| 18.9 全概率公式 .....                     | 67 |
| 18.10 离散型随机变量及其分布列 .....             | 70 |
| 18.10.1 离散型随机变量 .....                | 70 |

|         |                          |     |
|---------|--------------------------|-----|
| 18.10.2 | 离散型随机变量的分布列 .....        | 73  |
| 18.11   | 离散型随机变量的数字特征 .....       | 78  |
| 18.11.1 | 离散型随机变量的均值 .....         | 78  |
| 18.11.2 | 离散型随机变量的方差 .....         | 82  |
| 18.12   | 二项分布 .....               | 86  |
| 18.13   | 超几何分布 .....              | 90  |
| 18.14   | 正态分布 .....               | 92  |
| 第十九章    | 统计.....                  | 98  |
| 19.1    | 随机抽样.....                | 98  |
| 19.1.1  | 简单随机抽样.....              | 98  |
| 19.1.2  | 分层随机抽样 .....             | 102 |
| 19.1.3  | 获取数据的途径 .....            | 105 |
| 19.2    | 用样本估计总体 .....            | 107 |
| 19.2.1  | 总体取值规律的估计（第一课时） .....    | 107 |
| 19.2.2  | 总体取值规律的估计（第二课时） .....    | 114 |
| 19.2.3  | 总体百分位数的估计 .....          | 118 |
| 19.2.4  | 总体集中趋势的估计 .....          | 121 |
| 19.2.5  | 总体离散程度的估计 .....          | 125 |
| 第二十章    | 成对数据的统计分析.....           | 130 |
| 20.1    | 变量间的相关关系与相关系数 .....      | 130 |
| 20.2    | 一元线性回归模型及其应用 .....       | 135 |
| 20.2.1  | 一元线性回归模型及其应用（第一课时） ..... | 135 |
| 20.2.2  | 一元线性回归模型及其应用（第二课时） ..... | 138 |
| 20.3    | 列联表与独立性检验 .....          | 142 |
| 20.3.1  | 分类变量与列联表 .....           | 142 |
| 20.3.2  | 独立性检验 .....              | 145 |

## 第十七章 排列组合与二项式定理

### 17.1 分类加法计数原理与分步乘法计数原理

#### 17.1.1 分类加法计数原理与分步乘法计数原理（第一课时）

##### 【学习目标】

1.归纳得出分类加法计数原理和分步乘法计数原理，能应用它们解决简单的实际问题.

2.正确地理解“完成一件事情”的含义.

3.根据实际问题的特征，正确地区分“分类”或“分步”.

重点：分类加法计数原理和分步乘法计数原理及其应用.

难点：“完成一件事情”的含义；根据具体问题的特征，选择分类加法计数原理或分步乘法计数原理.

##### 【要点整合】

1. 分类加法计数原理：完成一件事有两类不同方案，在第1类方案中有 $m$ 种不同的方法，在第2类方案中有 $n$ 种不同的方法，那么完成这件事共有 $N=m+n$ 种不同的方法.

2. 分步乘法计数原理：完成一件事需要两个步骤，做第1步有 $m$ 种不同的方法，做第2步有 $n$ 种不同的方法，那么完成这件事共有 $N=m \times n$ 种不同的方法.

3. 两个计数原理的区别和联系

|     | 分类加法计数原理                 | 分步乘法计数原理         |
|-----|--------------------------|------------------|
| 联系  | 回答的都是关于完成一件事情的不同方法的种数的问题 |                  |
| 区别一 | 针对的是“分类”问题               | 针对的是“分步”问题       |
| 区别二 | 各种方法相互独立                 | 各个步骤中的方法互相依存     |
| 区别三 | 任何一种方法都可以做完这件事           | 只有各个步骤都完成才算做完这件事 |

4. 分类加法计数原理的推广：完成一件事有 $n$ 类不同的方案，在第1类方案中有 $m_1$ 种不同的方法，在第2类方案中有 $m_2$ 种不同的方法，……，在第 $n$ 类方案中有 $m_n$ 种不同的方法，那么完成这件事共有 $N = m_1 + m_2 + \cdots + m_n$ 种不同的方法.

5. 分步乘法计数原理的推广：完成一件事需要分成 $n$ 个步骤，做第1步有 $m_1$ 种不同的方法，做第2步有 $m_2$ 种不同的方法，……，做第 $n$ 步有 $m_n$ 种不同的方法，那么完成这件事共有 $N = m_1 \times m_2 \times \cdots \times m_n$ 种不同的方法.

6. 注意的问题

1.应用分类加法计数原理应注意的问题

(1)明确题目中所指的“完成一件事”是什么事，完成这件事可以有哪些办法，怎样才算是完成这件事.

(2)完成这件事的 $n$ 类方法是相互独立的，无论哪种方案中的哪种方法都可以独立完成这件事，而不需要再用到其他的方法.

(3)确定恰当的分类标准，准确地对“这件事”进行分类，要求每一种方法必属于某一类方案，不同方案的任意两种方法是不同的方法，也就是分类时必须做到既“不重复”也“不遗漏”.

## 2. 应用分步乘法计数原理要注意的问题

(1)明确题目中所指的“完成一件事”是什么事，单独用题目中所给的某种方法是不是能完成这件事，也就是说是否必须要经过几步才能完成这件事.

(2)完成这件事分若干个步骤，只有每个步骤都完成了，才算完成这件事，缺少哪一步，这件事都不能完成.

(3)根据题意正确分步，要求各步之间必须连续，只有按照这几步逐步地去做，才能完成这件事，各步骤之间不能重复，也不能遗漏.

### 【典例讲练】

#### 题型一：分类加法计数原理

例1 某校高二(1)、(2)、(3)班，各班人数如下表：

|        | 男生人数 | 女生人数 | 总人数 |
|--------|------|------|-----|
| 高二(1)班 | 30   | 20   | 50  |
| 高二(2)班 | 30   | 30   | 60  |
| 高二(3)班 | 35   | 20   | 55  |

(1)从这三个班中选1名学生任学生会主席，有多少种不同的选法；

(2)从高二(1)班、(2)班男生中或从高二(3)班女生中选1名学生任学生会生活部部长，有多少种不同的选法？

[解析](1)从每个班选1名学生任学生会主席，共有3类不同的方案：

第1类，从高二(1)班中选出1名学生，有50种不同的选法；

第2类，从高二(2)班中选出1名学生，有60种不同的选法；

第3类，从高二(3)班中选出1名学生，有55种不同的选法.

根据分类加法计数原理知，从三个班中选1名学生任学生会主席，共有 $50+60+55=165$ 种不同的选法.

(2)从高二(1)班、(2)班男生或高二(3)班女生中选1名学生任学生会生活部部长，共有3类不同的方案：

第1类，从高二(1)班男生中选出1名学生，有30种不同的选法；

第2类，从高二(2)班男生中选出1名学生，有30种不同的选法；

第3类，从高二(3)班女生中选出1名学生，有20种不同的选法.

根据分类加法计数原理知，从高二(1)班、(2)班男生或高二(3)班女生中选1名学生任学生会生活部部长，共有 $30+30+20=80$ 种不同的选法.

应用分类加法计数原理的关键：

用分类加法计数原理计数，关键在于根据问题的特点确定一个适合它的分类标准，在这个分类标准下，完成这件事的任何一种方法只属于某一类，并且分别属于不同种类的两种方法是不同的.

【练习1】 $x, y \in \mathbb{N}^*$ ，且 $x+y \leq 6$ ，试求有序正整数对 $(x, y)$ 的个数.

[解析]按 $x$ 的取值进行分类： $x=1$ 时， $y=1, 2, 3, 4, 5$ ，共构成5个有序正整数对；

$x=2$ 时， $y=1, 2, 3, 4$ ，共构成4个有序正整数对； $x=3$ 时， $y=1, 2, 3$ ，共构成3个有序正整数对；

$x=4$ 时， $y=1, 2$ ，共构成2个有序正整数对； $x=5$ 时， $y=1$ ，共构成1个有序正整数对.

根据分类加法计数原理得：共有 $N=5+4+3+2+1=15$ (个)有序正整数对.

## 题型二：分步乘法计数原理

例2 设巴蜀中学高一某班级有任课教师11名，男生28名，女生25名。现要从中选出男、女生以及任课教师各一名代表班级参加比赛，共有多少种不同的选法？

[解析]  $28 \times 25 \times 11 = 7700$ （种）

利用分步乘法计数原理解决问题的注意事项：

(1) 仔细审题，抓住关键点确立分步标准，有特殊要求的先行安排。

(2) 分步要保证各步之间的连续性和相对独立性。

【练习2】春节期间，某电视台开展了“替你为父母送东西”的活动，在外地打工的小王要给家在农村的父母买一台冰箱和一台洗衣机，现有5种型号的冰箱和3种型号的洗衣机，那么小王共有多少种购买方案？

[解析] 小王可分两步完成：第一步，购买一台冰箱，有5种方法；

第二步，购买一台洗衣机，有3种方法。因此共有  $5 \times 3 = 15$  种不同的购买方案。

## 题型三 两个原理的综合应用

例3 一栋7层的楼房备有电梯，在一楼有甲、乙、丙三人进了电梯，则满足：有且只有一个人上7楼，且每人都在不同的楼层下电梯，且甲不去2楼的所有可能情况有多少种？

[解析] 分类为：若甲在七楼下电梯，分步为  $5 \times 4$ ；若乙在七楼下电梯，分步为  $4 \times 4$ ；

若丙在七楼下电梯，分步为  $4 \times 4$ ；故共有  $5 \times 4 + 4 \times 4 + 4 \times 4 = 52$  种情况。

求解较复杂的计数问题：

(1) 对于较复杂的问题，可以在分类方法中分步进行，或者在每步中进行分类。

(2) 对于一些比较复杂的既要运用分类计数原理又要运用分步计数原理的问题，我们可以恰当地画出示意图或列出表格，使问题更加直观、清晰。

(3) 涉及“多面手”的题型，关键是分清“多面手”可以“干什么活”。

【练习3】有一项活动，需在3名老师、8名男同学和5名女同学中选部分人员参加。

(1) 若只需一人参加，有多少种不同的选法？

(2) 若需老师、男同学、女同学各一人参加，有多少种不同的选法？

(3) 若需一名老师、一名同学参加，有多少种不同的选法？

[解析] (1) 有三类：3名老师中选一人，有3种方法；8名男同学中选一人，有8种方法；5名女同学中选一人，有5种方法。

由分类加法计数原理知，有  $3 + 8 + 5 = 16$  种选法。

(2) 分三步：第1步选老师，有3种方法；第2步选男同学，有8种方法；第3步选女同学，有5种方法。

由分步乘法计数原理知，共有  $3 \times 8 \times 5 = 120$  种选法。

(3) 可分两类，每一类又分两步。

第1类，选一名老师再选一名男同学，有  $3 \times 8 = 24$  种选法；

第2类，选一名老师再选一名女同学，共有  $3 \times 5 = 15$  种选法。

由分类加法计数原理知，共有  $24 + 15 = 39$  种选法。

#### 题型四：分类讨论思想解决排数问题

例 4 用 0,1,2,3,4,5 可以组成多少个无重复数字且比 2 015 大的四位偶数？

[解析]解法一：按末位是 0,2,4 分为三类：

第一类，末位是 0 的有  $4 \times 4 \times 3 = 48$  个；

第二类，末位是 2 的有  $3 \times 4 \times 3 = 36$  个；

第三类，末位是 4 的有  $3 \times 4 \times 3 = 36$  个。

其中 2 014 不合题意，应去除，

由分类加法计数原理，得  $N = 48 + 36 + 36 - 1 = 119$  个。

解法二：按千位是 2,3,4,5 分四类：

第一类，千位是 2 的有  $2 \times 4 \times 3 = 24$  个，其中 2 014 不合题意，应去除；

第二类，千位是 3 的有  $3 \times 4 \times 3 = 36$  个；

第三类，千位是 4 的有  $2 \times 4 \times 3 = 24$  个；

第四类，千位是 5 的有  $3 \times 4 \times 3 = 36$  个。由分类加法计数原理，得  $N = 24 + 36 + 24 + 36 - 1 = 119$  个。

【练习 4】从 0, 2 中选一个数字，从 1, 3, 5 中选两个数字，组成无重复数字的三位数，其中奇数的个数为（ ）

A. 24

B. 18

C. 12

D. 6

[答案]B

[解析]由于题目要求的是奇数，那么对于此三位数可以分成两种情况：奇偶奇，偶奇奇；

如果是第一种奇偶奇的情况，可以从个位开始分析（3 种选择），之后十位（2 种选择），最后百位（2 种选择），共 12 种；

如果是第二种偶奇奇的情况，分析同理：个位（3 种情况），十位（2 种情况），百位（不能是 0，一种情况），共 6 种；

综上：总共  $12 + 6 = 18$  种情况，选择 B。

【课后巩固】——完成课时作业（165）

## 17.1.2 分类加法计数原理与分步乘法计数原理（第二课时）

### 【学习目标】

熟练应用分类加法计数原理和分步乘法计数原理.

### 【要点整合】

#### 1. 解决计数问题的求解方法

(1)直接结合运用两个原理解决：首先要明确是先“分类”后“分步”，还是先“分步”后“分类”；其次在“分类”和“分步”的过程中，均要确定明确的分类标准和分步程序.

(2)利用一些非常规计数问题的解决方法

①枚举法 将各种情况通过树形图、表格等方法一一列举出来，它适用于计数种数较少时，分类计数时将问题分类实际也是将分类种数一一列举出来.

②间接法 若计数时分类较多或无法直接计数时，可用间接法，先求出没有限制条件的总数，再减去不满足条件的种数.

③转换法 转换问题的角度或转换成其他已知的问题，在实际应用中，应根据具体问题，灵活处理.

### 【典例讲练】

#### 题型一：组数问题

例1 用0,1,2,3,4这五个数字可以组成多少个无重复数字的(1)四位密码？(2)四位数？(3)四位奇数？

[解析](1)完成“组成无重复数字的四位密码”这件事，可以分为四步：第一步，选取左边第一个位置上的数字，有5种选取方法；第二步，选取左边第二个位置上的数字，有4种选取方法；第三步，选取左边第三个位置上的数字，有3种选取方法；第四步，选取左边第四个位置上的数字，有2种选取方法. 由分步乘法计数原理，可以组成不同的四位密码共有  $N=5 \times 4 \times 3 \times 2=120$  个.

(2)完成“组成无重复数字的四位数”这件事，可以分四步：第一步，从1,2,3,4中选取一个数字作千位数字，有4种不同的选取方法；第二步，从1,2,3,4中剩余的三个数字和0共四个数字中选取一个数字作百位数字，有4种不同的选取方法；第三步，从剩余的三个数字中选取一个数字作十位数字，有3种不同的选取方法；第四步，从剩余的两个数字中选取一个数字作个位数字，有2种不同的选取方法. 由分步乘法计数原理，可以组成不同的四位数共有  $N=4 \times 4 \times 3 \times 2=96$  个.

(3)完成“组成无重复数字的四位奇数”这件事，可以分四步：

第一步定个位，只能从1,3中任取一个有2种方法，第二步定首位，把1,2,3,4中除去用过还有一个还有3个可任取一个有3种方法，第三步，第四步把剩下的包括0在内的还有3个数字先排百位有3种方法，再排十位有2种方法. 由分步乘法计数原理共有  $2 \times 3 \times 3 \times 2=36$  个.

【练习1】由数字0,1,2,3,4,5能组成多少个没有重复数字的四位数？

[解析]第一步：千位上有1,2,3,4,5，共5种选法；第二步：百位上可以从剩余的5个数中选1个，共5种选法；第三步：十位上可从剩余的4个数中选1个，共4种选法；第四步：个位上可从剩余的3个数中选1个，共3种选法.

利用分步计数原理，共有  $5 \times 5 \times 4 \times 3=300$  个没有重复数字的四位数.

## 题型二：分配问题

例 2(1)8 本不同的书，任选 3 本分给 3 个同学，每人 1 本，有多少种不同的分法？

(2)3 位旅客投宿到 1 个旅馆的 4 个房间(每房间最多可住 3 人)有多少种不同的住宿方法？

(3)将 4 封信投入 3 个邮筒，有多少种不同的投法？

[解析](1)分三步，每位同学取书一本，第 1、2、3 个同学分别有 8、7、6 种取法，因而由分步乘法计数原理，不同分法共有  $N=8 \times 7 \times 6=336$  (种).

(2)分三步，每位旅客都有 4 种不同的住宿方法，因而共有不同的方法  $N=4 \times 4 \times 4=64$  (种).

(3)完成这件事情可以分四步，第一步，投第一封信，可以在 3 个邮筒中任选一个，因此有 3 种投法；第二步，投第二封信，同样有 3 种投法；第三步，投第三封信，也同样有 3 种投法；第四步，投第四封信，仍然有 3 种投法. 由分步乘法计数原理，可得出不同的投法共有  $N=3 \times 3 \times 3 \times 3=81$  (种). 或应用住店法：此题相当于 4 个人住三间店.

【练习 2】有四位学生参加三项不同的竞赛，

①每位学生必须参加一项竞赛，则有\_\_\_\_\_种不同的参赛方法；

②每项竞赛只许有一位学生参加，则有\_\_\_\_\_种不同的参赛方法.

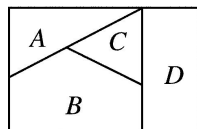
[答案]① $3^4$  ② $4^3$

## 题型三：涂色问题

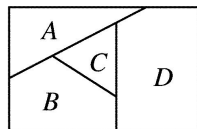
例 3 用  $n$  种不同颜色为下面两块广告牌着色(如下图①②)，要求在  $A$ 、 $B$ 、 $C$ 、 $D$  四个区域中相邻(有公共边的)区域不用同一种颜色.

(1)若  $n=6$ ，为①着色时共有多少种不同的方法？

(2)若为②着色时共有 120 种不同的方法，求  $n$ .



①



②

[解析](1)方法一 分步：先涂  $B$  区，有 6 种涂法，再涂  $C$  区，有 5 种涂法，最后涂  $A$ 、 $D$  区域，各有 4 种涂法， $\therefore$  共有  $6 \times 5 \times 4 \times 4=480$  种涂法.

方法二 以四个区域涂  $n$  种颜色为标准分类，可知至少用三种颜色，最多用四种颜色.

第一类：用三种颜色着色， $A$ 、 $D$  区域必须是同种颜色，有  $6 \times 5 \times 4=120$  种涂法.

第二类：用四种颜色着色，四个区域的颜色均不相同，

有  $6 \times 5 \times 4 \times 3=360$  种涂法. 所以共有  $120+360=480$  种不同方法.

(2)与(1)的区别在于与  $D$  相邻的区域由两块变成三块.

同理，不同的着色方法种数是  $n(n-1)(n-2)(n-3)=120=5 \times 4 \times 3 \times 2$ ，对照易知  $n=5$ .

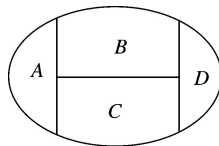
【练习 3】用 6 种不同的颜色把右图中  $A$ 、 $B$ 、 $C$ 、 $D$  四块区域分开，允许同一色涂不同的区域，但相邻的区域不能涂同一色，则不同的涂法共有( )

A. 400 种

B. 460 种

C. 480 种

D. 496 种



[答案]C

[解析] $6 \times 5 \times 4 \times 4=480$ .

#### 题型四：多面手问题

例4 某艺术小组有9人，每人至少会钢琴和小号中的一种乐器，其中7人会钢琴，3人会小号，从中选出会钢琴与会小号的各1人，有多少种不同的选法？

[解析]由题意可知，在艺术小组9人中，有且仅有1人既会钢琴又会小号(把该人称为“多面手”)，只会钢琴的有6人，只会小号的有2人，把选出会钢琴、小号各1人的方法分为两类：

第一类：多面手入选，另1人只需从其他8人中任选一个，故这类选法共有8种。

第二类：多面手不入选，则会钢琴者只能从6个只会钢琴的人中选出，会小号者也只能从只会小号的2人中选出，故这类选法共有 $6 \times 2 = 12$ 种。

因此共有  $N = 8 + 12 = 20$  种不同的选法。

【练习4】若将上题改为：其中7人会钢琴，4人会小号，结果如何？

[解析]可知既会钢琴，又会小号的有  $(7+4)-9=2$  人，以此两人是否入选进行分类。(其中只会钢琴5人，只会小号2人)

(1)2人中有1人入选，有2种选法，剩下的1人从其余7人中任选一个，有7种选法，

根据分步计数原理共有  $2 \times 7 = 14$  种选法；

(2)2人均入选，此类即1种选法；

(3)2人都不入选，则会钢琴的1人只能从5个只会钢琴者当中选出，会小号的1人也只能从只会小号的2人中选出，故这类选法共有  $5 \times 2 = 10$  种。

根据分类计数原理，共有  $14 + 1 + 10 = 25$  种不同选法。

【课后巩固】—————完成课时作业（166）

## 17.2 排列

### 17.2.1 排列（第一课时）

#### 【学习目标】

- 1.理解排列的概念.
- 2.了解排列数的概念.
- 3.掌握排列数公式的推导方法.
- 4.能用排列知识解决简单的实际问题.

重点：排列的概念；排列数公式；用排列知识解决简单的实际问题.

难点：排列数公式的推导方法.

#### 【要点整合】

##### 1. 排列的有关概念

(1)定义：一般地，从  $n$  个不同元素中取出  $m(m \leq n)$  个元素，按照一定的顺序排成一行，叫做从  $n$  个不同元素中取出  $m$  个元素的一个排列.

(2)相同排列：两个排列相同，当且仅当两个排列的元素完全相同，且元素的排列顺序也相同.

##### 2. 排列数与排列数公式

(1)排列数：从  $n$  个不同元素中取出  $m(m \leq n)$  个元素的所有不同排列的个数叫做从  $n$  个不同元素中取出  $m$  个元素的排列数，用符号  $A_n^m$  表示.

(2)排列数公式：
$$A_n^m = n(n-1)(n-2)\dots(n-m+1) = \frac{n!}{(n-m)!};$$

特别地， $A_n^n = n \times (n-1) \times (n-2) \times \dots \times 2 \times 1 = n!$  ( $m, n \in N^*$  且  $m \leq n$ )，规定： $0! = 1$  .

#### [注]1.关于排列的概念

给出的  $n$  个元素是互不相同的，且抽取的  $m$  个元素是没有重复抽取的；排列的定义中包含两个基本内容：一是“取元素”，二是“按照一定顺序排列”. 注意在解题时应细心观察：一“抽取”是否“重复”，二是否与顺序有关.

#### 2.排列数公式的特征

①  $m$  个连续自然数之积；②最大数是  $n$ ，最小的是  $(n-m+1)$  .

#### 【典例讲练】

##### 题型一：排列概念

例 1 判断下列问题是否是排列问题：

- (1)从 2,3,5,7,11 中任取两数相乘可得多少个不同的积？
- (2)从上面各数中任取两数相除，可得多少个不同的商？
- (3)某班共有 50 名同学，现要投票选举正副班长各一人，共有多少种可能的选举结果？
- (4)某商场有四个大门，若从一个门进去，购买商品后再从另一个门出来，不同的出入方式共有多少种？

[解析](1)乘法符合交换律与顺序无关,不是排列问题.

(2)上、下互换结果不一样,与顺序有关,是排列问题.

(3)请同学们记住“正”的就是“正”的,正副不同,是排列问题.

(4)“门”不同,先后也不一样,是排列问题.

探究 排列的核心是“顺序”,有“顺序”就是排列问题.那么如何判断是否有顺序呢?最常用的办法是把得到的结果变换元素的位置,如果结果变了,就是有“顺序”,若结果不变,就是无“顺序”.

【练习 1】判断下列问题是否是排列问题.

①从 2,3,5,7,9 中任取两数作为对数的底数与真数,可得多少个不同的对数值?

②空间有 10 个点,任何三点不共线,任何四点不共面,则这 10 个点共可组成多少个不同的四面体?

③某班有 10 名三好学生,5 名后进生,班委会决定选 5 名三好学生对 5 名后进生实行一帮一活动,共有多少种安排方式?

④若从 10 名三好学生中选出 5 名和 5 名后进生组成一个学习小组,共有多少种安排方式?

[答案]①是 ②不是 ③是 ④不是

## 题型二:排列数的运算

例 2 (1)计算:  $\frac{2A_8^5 + 7A_8^4}{A_8^8 - A_9^5}$ ;

(2)解方程:  $3A_x^3 = 2A_{x+1}^2 + 6A_x^2$ .

[解析](1)  $\frac{2A_8^5 + 7A_8^4}{A_8^8 - A_9^5} = \frac{2 \times 8 \times 7 \times 6 \times 5 \times 4 + 7 \times 8 \times 7 \times 6 \times 5}{8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 - 9 \times 8 \times 7 \times 6 \times 5} = \frac{8 \times 7 \times 6 \times 5 \times (8+7)}{8 \times 7 \times 6 \times 5 \times (24-9)} = 1$

(2)由  $3A_x^3 = 2A_{x+1}^2 + 6A_x^2$ , 得  $3x(x-1)(x-2) = 2(x+1)x + 6x(x-1)$ .

$\therefore x \geq 3, \therefore 3(x-1)(x-2) = 2(x+1) + 6(x-1)$ .

即  $3x^2 - 17x + 10 = 0$ . 解得  $x=5$  或  $x=\frac{2}{3}$  (舍去).  $\therefore x=5$ .

探究 2 上述类型题的处理方法是利用排列数公式  $A_n^m = n(n-1)(n-2)\dots(n-m+1)$

或  $A_n^m = \frac{n!}{(n-m)!}$ , 消掉式子中的  $A_n^m$ , 转化为关于  $n$  的代数方程再求解.

【练习 2】(1)解方程  $3A_8^x = 4A_9^{x-1}$ ; (2)已知  $\frac{A_n^7 - A_n^5}{A_n^5} = 89$ , 求  $n$  的值.

[解析](1)由  $3A_8^x = 4A_9^{x-1}$ , 得  $\frac{3 \times 8!}{(8-x)!} = \frac{4 \times 9!}{(10-x)!}$ , 化简得  $x^2 - 19x + 78 = 0$ , 解得  $x_1=6, x_2=13$ .

又  $\therefore x \leq 8$  且  $x-1 \leq 9$ ,  $\therefore$  原方程的解是  $x=6$ .

(2)  $\therefore \frac{A_n^7 - A_n^5}{A_n^5} = \frac{[(n-5)(n-6)-1]A_n^5}{A_n^5} = 89, \therefore (n-5)(n-6)=90, \therefore n=15$  或  $n=-4$  (舍).

### 题型三：证明排列数恒等式

例 3 (1)求证:  $A_{n+1}^m = A_n^m + mA_n^{m-1} (m \leq n)$ ;

(2)求证:  $\frac{1}{(n+1)!} = \frac{1}{n} \left( \frac{1}{n!} - \frac{1}{(n+1)!} \right)$ .

[证明](1)证法一:  $A_n^m + mA_n^{m-1} = n(n-1)\cdots(n-m+1) + m \cdot n(n-1)\cdots(n-m+2)$

$$= n(n-1)\cdots(n-m+2)[(n-m+1) + m] = (n+1)n(n-1)\cdots(n-m+2) = A_{n+1}^m.$$

$$\text{证法二: } A_n^m + mA_n^{m-1} = \frac{n!}{(n-m)!} + m \frac{n!}{(n-m+1)!} = \frac{n![(n-m+1) + m]}{(n-m+1)!} = \frac{(n+1)!}{[(n+1)-m]!} = A_{n+1}^m.$$

证法三: (构造法) 设从  $n+1$  个不同元素  $a_1, a_2, \dots, a_{n+1}$  中取出  $m$  个不同元素的排列数为  $A_{n+1}^m$ , 则其可分为两类, 一类是不含  $a_1$  的排列共有  $A_n^m$ , 另一类是含  $a_1$  的排列, 先排列  $a_1$  的位置有  $m$  种, 再从剩余的  $n$  个元素中选出  $m-1$  个元素排在剩余的  $m-1$  个位置有  $A_n^{m-1}$  种, 此类共有  $mA_n^{m-1}$  种, 由分类计数原理可知等式  $A_{n+1}^m = A_n^m + mA_n^{m-1}$  成立.

$$(2) \text{右边 } \frac{1}{n} \left[ \frac{1}{n!} - \frac{1}{(n+1)!} \right] = \frac{1}{n} \cdot \frac{(n+1)-1}{(n+1)!} = \frac{1}{(n+1)!} = \text{左边, 所以原等式成立.}$$

探究 3 (1)正确运用排列数公式是顺利进行推进运算的关键.

(2)类似等式  $n \cdot n! = (n+1)! - n!$  你会运用与证明吗?

【练习 3】 $1! + 2 \cdot 2! + 3 \cdot 3! + \cdots + n \cdot n! = (n+1)! - \underline{\hspace{2cm}}$ .

[答案]1

[解析]原式  $= (2! - 1!) + (3! - 2!) + \cdots + [(n+1)! - n!] = (n+1)! - 1$ .

### 题型四：排列的简单应用

例 4(1)10 个人走进只有 6 把不同椅子的屋子, 若每把椅子必须且只能坐一个人, 则共有多少种不同的坐法?

(2)6 个人走进放有 10 把不同椅子的屋子, 每个人必须且只能坐一把椅子, 则共有多少种不同的坐法?

[解析](1)坐在椅子上的 6 个人是走进屋子的 10 个人中的任意 6 个人. 若我们把人抽象地看成元素, 将 6 把椅子当成 6 个不同的位置, 则原问题便抽象为: 从 10 个元素中任取 6 个元素占据 6 个不同的位置, 显然是从 10 个元素中任取 6 个元素的排列问题, 可见, 共有  $A_{10}^6 = 151200$  种坐法.

(2) 同(1)恰好相反, 被坐的 6 把椅子是 10 把中的任意 6 把. 将椅子看成元素, 6 个人看成 6 个不同的位置, 本质上与(1)相同, 这种看法显然同生活中的看法截然相反, 它正说明了排列中“元素”和“位置”的相对性. 本题也有  $A_{10}^6 = 151200$  种坐法.

探究 本例为排列提供了一个易于操作的(占位)模型:  $n$  个不同的元素去占据  $m$  个不同的位置, 若  $n \geq m$  且每个位置只占(排)一个元素, 则有  $A_n^m$  种不同的占(排)法; 若  $n < m$  且每元素只占一个位置, 则有  $A_m^n$  种不同的占(排)法.

【练习 4】(1)①4 种不同蔬菜种植在 3 块不同土质的田地里，每块田地种一种蔬菜，共有\_\_\_\_\_种不同种法.

②3 种不同蔬菜种植在 4 块不同的田地里，每种蔬菜种一块地，共有\_\_\_\_\_种不同种法.

[答案]①  $A_4^3 = 24$     ②  $A_4^3 = 24$

(2)从 1,2,3,4 中，任取两个不同数字组成平面直角坐标系中一个点的坐标，则组成不同点的个数为(     )

A. 2

B. 4

C. 12

D. 24

[答案]C

【课后巩固】—————完成课时作业（167）

## 17.2.2 排列（第二课时）

### 【学习目标】

重点：用排列知识解决实际问题.

### 【要点整合】

1.某些元素要求相邻的问题，常用“捆绑”的办法：把相邻或要求在一起的元素捆在一起看成一个元素与其他元素排列；然后再松绑，即内部再排列.

2.某些元素要求不相邻的问题，常用“插空”的办法：即先排其他元素，然后在其形成的空位中选出空位排要求不相邻的元素.

3.部分元素顺序一定的问题可用“消序法”： $n$  个元素全排列有  $A_n^n$  种顺序，若其中  $m$  个元素之间顺序一定，则只有  $\frac{A_n^n}{A_m^m} = A_n^{n-m}$  顺序.

4.在复杂的排列问题中，一般不是一下子就将全部元素排好，而是分步排列有限制要求的元素或位置，再用分步计数原理求解.

### 【典例讲练】

#### 题型一：捆绑法

例 1(1)4 名男同学，3 名女同学站成一排.

①3 名女同学必须排在一起，有多少种不同的排法？

②甲、乙两同学之间恰有 3 人，有多少种不同的排法？

(2)用 1,2,3,4,5 这五个数字组成没有重复数字的五位数，其中 2 和 3 相邻的五位偶数有多少个？

(3)某单位有 7 个连在一起的停车位，现有 3 辆不同型号的车需要停放，如果要求剩余的 4 个空车位连在一起，那么不同的停放方法有( )

A. 16 种

B. 18 种

C. 24 种

D. 32 种

[答案](1)①720 ②720 (2)18 (3)C

[解析](1)①3 名女同学是特殊元素，我们先把她们排好，共有  $A_3^3$  种排法，由于 3 名女同学必须排在一起，我们可视排好的女同学为一整体再与男同学一起排队，这时是 5 个元素的全排列，应有  $A_5^5$  种排列，由分步计数原理可得，有  $A_3^3 A_5^5 = 720$  种不同排法.

②分步进行，首先将甲、乙两人排好，有  $A_2^2$  种排法，再从余下的 5 人中选 3 人排在甲、乙二人中间，有  $A_5^3$  种排法，这时把已排好的 5 人视为一个整体，与最后剩下的 2 人再排，即看作 3 个元素的全排列，又有  $A_3^3$  种排法，这样，由分步计数原理总共有  $A_2^2 A_5^3 A_3^3 = 720$  种不同排法.

(2)分两类：

①形如

|   |   |   |   |   |
|---|---|---|---|---|
| × | × | × | 3 | 2 |
|---|---|---|---|---|

的五位数有  $A_3^3 = 6$  个，

②形如

|   |   |   |   |   |
|---|---|---|---|---|
| × | × | × | × | 4 |
|---|---|---|---|---|

的五位数有  $A_3^3 \cdot A_2^2 = 12$  个, 由分类计数原理知共有  $6+12=18$  个.

(3) 4 个空位连在一起, 看成 1 个元素, 与 3 辆不同型号的车进行全排列有  $A_4^4 = 24$  种不同放车方法.

探究 1 元素相邻(或在一起)形成小集团就用捆绑法. 注意先捆后排!

【练习 1】(1) 四对夫妻坐一排照相, 每对夫妻都不能隔开坐, 则不同的坐法种数为\_\_\_\_\_.

[解析] 四对夫妻看成四个“元素”有  $A_4^4$  种排法, 每对夫妻有 2 种排法, 故其有  $A_4^4 \cdot 2^4 = 384$  种.

(2) 计划在某画展展出 10 幅不同的画, 其中 1 幅水彩画, 4 幅油画, 5 幅国画, 排成一行陈列, 要求同一品种的画必须连在一起, 并且水彩画不放在两端, 那么不同的陈列方式有( )

A.  $A_4^4 A_5^5$  种                      B.  $A_3^3 A_4^4 A_5^5$  种                      C.  $3A_4^4 A_5^5$  种                      D.  $A_2^2 A_4^4 A_5^5$  种

[解析] 先考虑特殊元素的特殊位置. 由题意 4 幅油画的不同陈列方式有  $A_4^4$  种, 5 幅国画的不同陈列方式有  $A_5^5$  种, 一幅水彩画只能陈列在中间位置, 这样油画和国画的陈列方式又有 4 种. 根据分步计数原理, 画展的不同陈列方式有  $A_2^2 A_4^4 A_5^5$  种, 故选 D.

## 题型二: 插空法

例 2 (1) 4 名男同学, 3 名女同学站成一排,

①任何两名女同学彼此不相邻, 有多少种不同的排法?

②甲、乙两人相邻, 但都不与丙相邻, 有多少种不同的排法?

(2) 用 1,2,3,4,5,6,7,8 组成没有重复数字的八位数, 要求 1 与 2 相邻, 3 与 4 相邻, 5 与 6 相邻, 而 7 与 8 不相邻, 这样的八位数共有\_\_\_\_\_个. (用数字作答)

(3) 有 6 个座位连成一排, 现有 3 人就坐, 恰有两个空位相邻的不同坐法有\_\_\_\_\_种.

[答案](1)①1 440    ②960    (2)576    (3)72

[解析](1)①先将男生排好, 共有  $A_4^4$  种排法, 再在这 4 名男生中间及两头的 5 个空位中插入 3 个女生有  $A_5^3$  种方案, 故符合条件的排法共有  $A_4^4 A_5^3 = 1440$  种不同排法.

②先排甲、乙、丙 3 人以外的其他 4 人, 有  $A_4^4$  种排法; 由于甲、乙要相邻, 故再将甲、乙连在一起排好, 有  $A_2^2$  种排法; 最后把甲、乙排好的这个整体及丙两个元素分别插入原先排好的 4 人的空格位置中有  $A_5^2$  种排法, 因此, 这样一共有  $A_4^4 A_2^2 A_5^2 = 960$  种不同排法.

(2) 相邻捆绑, 不相邻插空, 八位数共有  $A_3^3 A_4^2 A_2^2 A_2^2 = 576$ .

(3) 6 个座位 3 人就坐, 有 3 个空位. 3 人坐成一排有  $A_3^3$  种不同坐法, 且形成 4 个空, 将两个相连的空位看成一个大空位, 把这个大空位及另一个空位插入上述 4 个空中有  $A_4^2$  种, 所以共有  $A_3^3 A_4^2 = 72$  种.

探究 2 (1) 不相邻则插空.

(2) 要恰当的理解题意, 如恰有两个空位相邻, 其实质就是这相邻的两个空位与另一个空位不相邻!

【练习 2】(1)6 人排成一排, 甲、乙不相邻的排法有\_\_\_\_\_种.

[解析]另外 4 人排成一排有  $A_4^4$  种排法, 且形成 5 个空, 从中选 2 个排甲、乙有  $A_5^2$  种排法,

故共有  $A_4^4 A_5^2 = 480$  种.

(2)2 个男生和 4 个女生排成一排, 其中男生既不相邻也不排两端的不同的排法有( )

A.  $A_4^4 A_3^2$  种      B.  $A_4^2 A_6^6$  种      C.  $A_6^2 A_6^4$  种      D.  $A_2^2 A_4^4$  种

[解析]先排 4 个女生有  $A_4^4$  种排法, 然后让 2 个男生插入中间三个空中的两个位置有  $A_3^2$  种排法. 据分步计数原理知选 A.

### 题型三: 消序法

例 3 (1)男生 4 名, 女生 3 名排成一排, 若三名女生从左至右顺序一定 (不一定相邻), 则有\_\_\_\_\_种不同的排法.

(2) $A, B, C, D, E$  五人并排站成一排, 如果  $B$  必须站在  $A$  的右边( $A, B$  可以不相邻), 那么不同的排法共有\_\_\_\_\_种.

[答案](1)840 (2)60

[解析](1)方法一: 在 7 个位置上任意排列 7 名学生, 有排法  $A_7^7$  种, 由于女生的顺序一定, 而在  $A_7^7$  中每一种情况均以  $A_3^3$  计算, 故三名女生顺序一定的排法有  $\frac{A_7^7}{A_3^3} = 840$  种.

方法二: 在 7 个位置上任选 4 个位置排 4 名男生, 剩下的 3 个位置三个女生按规定顺序对号入座即可, 故三名女生顺序一定的排法有  $A_7^4$  种.

(2)五人并排站成一排有  $A_5^5$  种排法,  $A, B$  排列有  $A_2^2$  种排法, 故  $B$  必须站在  $A$  的右边的排法有  $\frac{A_5^5}{A_2^2} = 60$  种.

探究 3(1)此类题目也可采用元素分析法. 如①中把 4 名男生看成特殊元素, 从 7 个位置中选 4 人先排男生有  $A_7^4$  种, 再排女生只有 1 种方法. 共有  $A_7^4 \cdot 1 = 840$  种; ②中  $A_5^3 \cdot 1 = 60$  种.

(2)自己写出此类题目的解题规律.

【练习 3】八个人排一纵队, 甲在乙的前面(可以与乙不相邻), 乙在丙的前面(可以与丙不相邻), 则这样的排法共有\_\_\_\_\_种(用数字作答).

[解析]八个人去掉甲、乙、丙后的排法有  $A_8^5$  种, 再排甲、乙、丙只有一种排法, 所以满足题意的排法

共有  $A_8^5 = 8 \times 7 \times 6 \times 5 \times 4 = 6720$  种.

### 题型四: 一题多解

例 4 书架上原来摆放着 6 本书, 现要再插入 3 本书, 则不同插法的种数为( )

A.  $A_7^3$  种      B.  $A_4^4$  种      C.  $9 \times 8 \times 7$  种      D.  $2A_3^3$  种

[解析]方法 1: 9 本书排成一排, 相当于 9 个位置, 原来 6 本书顺序不变, 故此题相当于在 9 个位置中选 3 个排后来的 3 本书, 有  $A_9^3$  种排法, 而原来的 6 本书只有一种排法, 故共有  $A_9^3 \times 1 = A_9^3 = 9 \times 8 \times 7$  种方法.

方法 2: 因为要插入 3 本书, 故分三步, 第一步, 插第一本书有 7 种方法; 第二步, 插第二本书, 有 8 种方法; 第三步, 插第三本书, 有 9 种方法, 由分步计数原理, 共有  $7 \times 8 \times 9$  种方法.

方法 3: 9 本书排成一排有  $A_9^9$  种排法, 原 6 本书若排列则有  $A_6^6$  种排法,

故符合题意的插法有  $\frac{A_9^9}{A_6^6} = A_9^3 = 9 \times 8 \times 7$ .

探究 4 (1)相邻则捆绑: 先捆后松. (2)不相邻则插空: 先排后插. (3)顺序已定则消序: 先排后除.

【练习 4】(1)6 人站成一排, 其中甲、乙不相邻的站法有 \_\_\_\_\_ 种.

[解析]法 1: (插空法)  $A_4^4 \cdot A_5^2 = 480$ . 法 2: (去杂法)  $A_6^6 - A_2^2 \cdot A_5^5 = 480$ .

(2)设椭圆  $\frac{x^2}{a} + \frac{y^2}{b} = 1$ , 其中  $a, b \in \{1, 2, 3, 4, 5\}$ , 且椭圆焦点在  $x$  轴上, 则这样的椭圆有 \_\_\_\_\_ 个.

[解析]焦点在  $x$  轴上, 则  $a > b$ , 方法 1: 利用分类计数原理得共有  $4+3+2+1=10$  个.

方法 2: (消序法)  $\frac{A_5^2}{A_2^2} = 10$ .

【练习 5】用 0,1,2,3,4,5 这六个数字可以组成多少个无重复数字的

(1)六位奇数; (2)个位数字不是 5 的六位数.

[解析](1)解法一 (从特殊位置入手) 分三步完成, 第一步先填个位, 有  $A_3^1$  种填法, 第二步再填十万位, 有  $A_4^1$  种填法, 第三步填其他位, 有  $A_4^4$  种填法, 故共有  $A_3^1 A_4^1 A_4^4 = 288$  个六位奇数.

解法二 (从特殊元素入手) 0 不在两端有  $A_4^1$  种排法, 从 1,3,5 中任选一个排在个位有  $A_3^1$  种排法, 其他各位上用剩下的元素做全排列有  $A_4^4$  种排法, 故共有  $A_4^1 A_3^1 A_4^4 = 288$  个六位奇数.

解法三 (排除法) 6 个数字的全排列有  $A_6^6$  个, 0,2,4 在个位上的排列数为  $3A_5^5$  个, 1,3,5 在个位上, 0 在十万位上的排列数有  $3A_4^4$  个, 故对应的六位奇数的排列数为  $A_6^6 - 3A_5^5 - 3A_4^4 = 288$  个.

(2)解法一 (排除法) 0 在十万位和 5 在个位的排列都不对应符合题意的六位数.

故符合题意的六位数共有  $A_6^6 - 2A_5^5 + A_4^4 = 504$  个.

解法二 (直接法) 个位不排 5, 有  $A_5^1$  种排法, 但十万位数字的排法因个位上排 0 与不排 0 而有所不同. 因此需分两类. 第一类: 当个位排 0 时, 有  $A_5^5$  个. 第二类: 当个位不排 0 时, 有  $A_4^1 A_4^1 A_4^4$  个.

故共有符合题意的六位数  $A_5^5 + A_4^1 A_4^1 A_4^4 = 504$  个.

【课后巩固】 \_\_\_\_\_ 完成课时作业 (168)

## 17.3 组合

### 17.3.1 组合（第一课时）

#### 【学习目标】

- 1.理解组合的概念.
- 2.能根据两个计数原理推导组合数公式.
- 3.能用组合知识解决简单的实际问题.
- 4.根据实际问题的特征, 正确区分“排列”或“组合”.

重点: 组合的概念; 组合数公式的推导; 应用组合知识解决简单的实际问题.

难点: 组合数公式的推导, 根据实际问题的特征, 正确区分“排列”或“组合”.

#### 【要点整合】

##### 1. 组合及组合数的概念

(1)组合: 一般地, 从  $n$  个不同元素中取出  $m(m \leq n)$  个元素合成一组, 叫做从  $n$  个不同元素中取出  $m$  个元素的一个组合.

(2)组合数: 从  $n$  个不同元素中取出  $m(m \leq n)$  个元素的所有不同组合的个数, 叫做从  $n$  个不同元素中取出  $m$  个元素的组合数, 用符号  $C_n^m$  表示.

##### 2. 组合数公式及其性质

|    |      |                                                                     |
|----|------|---------------------------------------------------------------------|
| 公式 | 展开式  | $C_n^m = \frac{A_n^m}{A_m^m} = \frac{n(n-1)(n-2)\cdots(n-m+1)}{m!}$ |
|    | 阶乘式  | $C_n^m = \frac{n!}{m!(n-m)!}$                                       |
| 性质 | 性质 1 | $C_n^m = C_n^{n-m}$                                                 |
|    | 性质 2 | $C_{n+1}^m = C_n^m + C_n^{m-1}$                                     |
|    | 规定   | $C_n^0 = 1$                                                         |

[注]1.组合的概念

(1)如果两个组合中的元素完全相同, 不管它们的顺序如何都是相同的组合. 组合的定义中包含两个基本内容: 一是“取出元素”; 二是“合成一组”, “合成一组”即表示与顺序无关.

(2)若两个组合中的元素不完全相同(即使只有一个元素不同), 则是不同的组合. 例如从  $a$ 、 $b$ 、 $c$  三个不同的元素中取出两个元素的所有组合有 3 个, 它们分别是  $ab$ 、 $ac$ 、 $bc$ . 其中  $ba$ 、 $ab$  是相同的组合, 而  $ab$ 、 $ac$  是不同的组合.

(3)组合与排列问题的共同点是都要“从  $n$  个不同元素中, 任取  $m$  个元素”; 不同点是前者是“不管顺序合成一组”, 而后者都要“按照一定顺序排成一列”.

2.区分排列问题、组合问题: 前者是从  $n$  个不同元素中选取  $m$  个不同元素后, 还要按照一定的顺序排成一列, 而后者只要从  $n$  个不同元素中选取  $m$  个不同的元素合成一组, 所以区分某一问题是排列还是组合, 关

键看选出的元素与顺序是否有关,若交换某两个元素的位置对结果产生影响,则是排列问题;若交换任意两个元素的位置对结果都没有影响,则是组合问题.

### 3.组合数、组合数公式及组合数性质

(1)组合数符号表示为  $C_n^m$ ,如从 4 个不同元素取出 3 个元素的组合数为  $C_4^3$ ;

(2)“组合”与“组合数”是两个不同的概念,组合是指“从  $n$  个不同的元素中,任取  $m(m \leq n)$  个元素合成一组”,它不是一个数,而是具体的一件事;组合数是指从  $n$  个不同元素中取出  $m$  个元素的所有组合的个数,它是一个数;

(3)组合数公式是在  $m, n \in N^*$ , 且  $m \leq n$  时成立;

(4)当  $m, n$  数值较大或含有字母的组合数式进行变形论证时,利用公式  $C_n^m = \frac{n!}{m!(n-m)!}$  解题较方便;

(5) $C_n^0 = 1$  是一种规定,不能用组合数定义解释;

(6)计算组合数时,特别是  $m$  较大时,注意利用公式  $C_n^m = C_n^{n-m}$  转化.

### 【典例讲练】

#### 题型一:组合的概念

例 1 判断下列问题是不是组合问题?

- (1)从 10 人中选 4 人①参加座谈会; ②分赴四地搞调查.  
(2)从 1,2,3,4,5,6 中任取两数①构成对数或指数; ②相加或相乘.  
(3)三个人互相①握手; ②送礼品.  
(4)由正四面体 4 个顶点①可形成多少个向量; ②形成多少对异面直线.

[答案](1)中①是②不是 (2)中①不是②是 (3)中①是②不是 (4)中①不是②是

探究 1 对于概念题,绝不能含糊,要严扣定义,紧抓本质,有序排列,无序组合!

【练习 1】判断下列各事件是排列问题,还是组合问题,并求出相应的排列数或组合数.

- (1)10 个人相互各写一封信,共写了多少封信?  
(2)10 个人规定相互通一次电话,共通了多少次电话?  
(3)10 支球队以单循环进行比赛(每两队比赛一次),这次比赛需要进行多少场次?  
(4)10 支球队以单循环进行比赛,这次比赛冠亚军获得者有多少种可能?  
(5)从 10 个人里选 3 个代表去开会,有多少种选法?  
(6)从 10 个人里选出 3 个不同学科的课代表,有多少种选法?

[解析](1)是排列问题,因为发信人与收信人是有顺序区别的.排列数为  $A_{10}^2 = 90$ (种).

(2)是组合问题,因为甲与乙通一次电话即乙与甲通一次电话,没有顺序的区别.组合数为  $C_{10}^2 = 45$ (种).

(3)是组合问题,因为每两个队比赛一次,并不需要考虑谁先谁后,没有顺序的区别.组合数为  $C_{10}^2 = 45$ (种).

(4)是排列问题,因为甲队得冠军、乙队得亚军与甲队得亚军、乙队得冠军是不一样的,是有顺序区别的.排列数为  $A_{10}^2 = 90$ (种).

(5)是组合问题. 因为三个代表之间没有顺序的区别. 组合数为  $C_{10}^3 = 120$  (种).

(6)是排列问题. 因为三个人中, 担任哪一科的课代表是有顺序区别的. 排列数为  $A_{10}^3 = 720$  (种).

## 题型二: 组合数公式

例 2 (1)求函数  $f(n) = C_{3n}^{38-n} + C_{21+n}^{3n}$  的定义域; (2)解不等式:  $C_{10}^{n-3} < C_{10}^{n-2} < C_{10}^{n-1}$ .

[解析](1)由题意知, 原式中的自然数  $n$  必须满足不等式组  $\begin{cases} 3n \geq 38-n \geq 0, \\ 21+n \geq 3n \geq 0, \end{cases}$

由①得  $\frac{38}{4} \leq n \leq 38$ , 由②得  $0 \leq n \leq \frac{21}{2} \therefore \frac{38}{4} \leq n \leq \frac{21}{2}$ , 又  $n \in N^*$ ,  $\therefore n=10$ . 即  $f(n)$  定义域为  $\{10\}$ .

(2)由题设不等式, 得  $\frac{10!}{(n-3)!(13-n)!} < \frac{10!}{(n-2)!(12-n)!} < \frac{10!}{(n-1)!(11-n)!}$ .

化简得  $n-2 < 13-n$  且  $n-1 < 12-n$ , 解得  $n < \frac{13}{2}$ . 又由题设知  $3 \leq n \leq 11$ ,  $\therefore n=3, 4, 5, 6$ .

$\therefore$  原不等式的解集为  $\{3, 4, 5, 6\}$ .

探究 2 在涉及  $C_n^m$  时, 要充分运用(或要时刻注意)  $m \leq n$  且  $m, n \in N^*$  来解题.

【练习 2】(1)计算  $C_{10}^3$ ; (2)证明:  $C_n^m = \frac{n}{n-m} C_{n-1}^m$ ; (3)已知  $\frac{1}{C_5^n} - \frac{1}{C_6^n} = \frac{7}{10C_7^n}$ , 求  $C_8^n$  的值.

[解析](1)  $C_{10}^3 = \frac{10 \times 9 \times 8}{3 \times 2 \times 1} = 120$ .

(2)  $\therefore \frac{n}{n-m} C_{n-1}^m = \frac{n}{n-m} \cdot \frac{(n-1)!}{m!(n-m-1)!} = \frac{n!}{m!(n-m)!} = C_n^m$ ,  $\therefore C_n^m = \frac{n}{n-m} C_{n-1}^m$ .

(3)由组合数公式, 可得  $\frac{n!(5-n)!}{5!} - \frac{n!(6-n)!}{6!} = \frac{7}{10} \cdot \frac{n!(7-n)!}{7!}$ ,

化简得  $n^2 - 23n + 42 = 0$ .  $\therefore n=21$  或  $n=2$ .  $\therefore n \leq 5$ ,  $\therefore n=2$ .  $\therefore C_8^n = C_8^2 = 28$ .

探究 计算  $C_n^m$  时常公式  $C_n^m = \frac{n(n-1)(n-2)\cdots(n-m+1)}{m!}$ .

证明与  $C_n^m$  有关的问题时常公式  $C_n^m = \frac{n!}{m!(n-m)!}$ .

## 题型三: 组合数的性质 1 的应用

例 3 (1)解方程组  $\begin{cases} C_x^y = C_x^{2y}, \\ 3C_x^{y+1} = 11C_x^{y-1}. \end{cases}$  (2)解不等式  $2C_{x+1}^{x-2} < 3C_{x+1}^2$ .

[解析](1)  $\therefore C_x^y = C_x^{2y}$ ,  $\therefore y=2y$  或  $y=x-2y$ . 若  $y=2y$ , 则  $y=0$ ,  $y-1 < 0$ , 不合题意, 舍去.

$\therefore y=x-2y$ , 即  $x=3y$ , 代入  $3C_x^{y+1} = 11C_x^{y-1}$ , 得  $3C_{3y}^{y+1} = 11C_{3y}^{y-1}$ ,

即  $3 \cdot \frac{(3y)!}{(y+1)!(2y-1)!} = 11 \cdot \frac{(3y)!}{(y-1)!(2y+1)!}$ , 化简得  $y^2 - 5y = 0 \therefore y = 0$  (舍) 或  $y = 5, \therefore x = 15$ .

$\therefore$  方程组的解  $\begin{cases} x = 15 \\ y = 5 \end{cases}$

$$(2) \because 2C_{x+1}^{x-2} < 3C_{x+1}^2, \therefore 2C_{x+1}^3 < 3C_{x+1}^2, \text{ 即 } 2 \times \frac{(x+1)x(x-1)}{1 \times 2 \times 3} < 3 \times \frac{(x+1)x}{1 \times 2}. \quad ①$$

$\therefore x+1 \geq 3, x \geq 2, \therefore (x+1)x > 0$ .

①式两边同除以  $(x+1)x$ , 得  $x-1 < \frac{9}{2}, \therefore x < \frac{11}{2} \therefore x = 2, 3, 4, 5$ . 即不等式的解集为  $\{2, 3, 4, 5\}$ .

【练习 3】(1)  $C_9^7 + C_9^8 =$  \_\_\_\_\_; (2)  $C_{3n}^{n+21} + C_{n+21}^{3n-2} =$  \_\_\_\_\_.

[答案](1) 45 (2) 65

(3) 已知  $x, y \in N^*$ , 且  $C_n^x = C_n^y$ , 则  $x, y$  的关系是( )

A.  $x=y$                       B.  $y=n-x$                       C.  $x=y$  或  $x+y=n$                       D.  $x \geq y$

[答案]C

[解析] 当  $x=y$  时显然成立; 因为  $C_n^x = C_n^{n-x}$ , 故有  $n-x=y$ , 得  $x+y=n$ .

#### 题型四：组合数的性质 2 的应用

例 4 (1) 求和:  $C_2^2 + C_3^2 + C_4^2 + \cdots + C_{100}^2$ . (2) 证明:  $C_n^0 + C_{n+1}^1 + C_{n+2}^2 + \cdots + C_{n+m-1}^{m-1} = C_{n+m}^{m-1}$ .

[解析](1) 方法 1 原式  $= C_3^3 + (C_4^3 - C_3^3) + (C_5^3 - C_4^3) + \cdots + (C_{101}^3 - C_{100}^3) = C_{101}^3 = 166650$ .

方法 2 原式  $= C_3^0 + C_3^1 + C_4^2 + C_5^3 + \cdots + C_{100}^{98} = C_4^1 + C_4^2 + C_5^3 + \cdots + C_{100}^{98}$

$$= C_5^2 + C_5^3 + \cdots + C_{100}^{98} = \cdots = C_{101}^{98} = C_{101}^3 = 166650.$$

方法 3 原式  $= C_3^3 + C_3^2 + C_4^2 + C_5^2 + \cdots + C_{100}^2 = C_4^3 + C_4^2 + C_5^2 + \cdots + C_{100}^2$

$$= C_5^3 + C_5^2 + \cdots + C_{100}^2 = \cdots = C_{101}^3 = 166650.$$

(2) 左式  $= C_{n+1}^0 + C_{n+1}^1 + C_{n+2}^2 + \cdots + C_{n+m-1}^{m-1} = C_{n+2}^1 + C_{n+2}^2 + \cdots + C_{n+m-1}^{m-1} = \cdots = C_{n+m-1}^{m-2} + C_{n+m-1}^{m-1}$   
 $= C_{n+m}^{m-1} =$  右式.

探究 3 (1)  $C_{n+1}^m = C_n^m + C_n^{m-1} \Leftrightarrow C_n^{m-1} = C_{n+1}^m - C_n^m$ ; (2)  $C_3^0 = C_2^2 = C_3^3 = \cdots = C_n^n$ ;

(3) 公式的灵活运用, 体现了思维的灵活性.

【练习 4】(1) 计算 ①  $C_3^1 + C_3^2 + C_4^3 + C_5^4 + C_6^5$ ;                      ②  $C_5^5 + C_6^5 + C_7^5 + C_8^5 + C_9^5 + C_{10}^5$ ;

(2) 计算  $C_{201}^{198} + C_{200}^{196} + C_{200}^{197}$ .

[答案](1) ① 21    ② 462 (2) 原式  $= C_{201}^{198} + C_{201}^{197} = C_{202}^{198} = C_{202}^4 = 67\,331\,650$ .

【课后巩固】\_\_\_\_\_ 完成课时作业 (169)

## 17.3.2 组合（第二课时）

### 【要点整合】

#### 1. 解组合应用题的总体思路

- (1)考查顺序：区别排列与组合的重要标志是“有序”与“无序”，无序问题用组合解答，有序问题属排列问题.
- (2)整体分类：对事件进行整体分类，从集合的意义讲，分类要做到各类的并集等于全集. 计算结果时，使用分类计数原理.
- (3)局部分步：整体分类以后，对每一类进行局部分步，分步要做到步骤连续且独立，计算每一类相应结果时使用分步计数原理.

#### 2. 组合问题常见类型及解题思路

(1)无条件限制的组合应用题. 其解题步骤：①判断；②转化；③求值；④作答.

(2)有限制条件的组合应用题.

①“含”与“不含”问题，其解题思路是将限制条件视为②“至多”与“至少”问题  
这类问题通常采用排除法，也可以用直接法.

③几何中的计算问题

在处理几何问题中的组合应用问题时，应先明确几何中的点、线、面及构型，明确平面图形和立体图形中的点、线、面之间的关系，将几何问题抽象成组合问题来解决.

特殊元素和特殊位置. 一般来讲，特殊要先满足，其余则“一视同仁”. 若正面入手不易，则从反面入手，寻找问题的突破口，即采用排除法. 解题时要注意分清“有且仅有”、“至多”、“至少”、“全是”、“都不是”、“不都是”等词语的确切含义，准确把握分类标准.

### 【典例讲练】

#### 题型一：简单的组合问题

例1 现有 10 名教师，其中男教师 6 名，女教师 4 名.

(1)现要从中选 2 名去参加会议，有多少种不同的选法？

(2)现要从中选出男、女教师各 2 名去参加会议，有多少种不同的选法？

[解析]本问题中选出的教师不需要考虑顺序，因此是组合问题. 第(1)小题选 2 名教师不考虑男女，实质上是从 10 个不同的元素中取出 2 个元素的组合问题，可用直接法求解. 第(2)小题必须选男、女教师各 2 名，才算完成所做的事，因此需要分两步进行，先从 6 名男教师中选 2 名，再从 4 名女教师中选 2 名，由基本原理，可用直接法求解.

(1)从 10 名教师中选 2 名去参加会议的选法数，就是从 10 个不同的元素中取出 2 个元素的组合数，即

$$C_{10}^2 = \frac{10 \times 9}{2 \times 1} = 45 \text{ 种}.$$

(2)从 6 名男教师中选 2 名的选法有  $C_6^2$  种，从 4 名女教师中选 2 名的选法有  $C_4^2$  种，根据分步乘法计数原理，

$$\text{因此共有不同的选法 } C_6^2 \cdot C_4^2 = \frac{6 \times 5}{2 \times 1} \cdot \frac{4 \times 3}{2 \times 1} = 90 \text{ 种}.$$

探究 1 解简单的组合应用题时，首先要判断它是不是组合问题，组合问题与排列问题的根本区别在于：排列问题与取出元素之间的顺序有关，而组合问题与取出元素的顺序无关，只要元素相同即可，只有当它能构成组合模型时，才能运用组合数公式求出其种数；其次要注意两个基本原理的运用，即分类与分步的灵活运用，在运用分类和分步时，一定要注意有无重复和遗漏。

【练习 1】一位教练的足球队共有 17 名初级学员，他们中以前没有一人参加过比赛。按照足球比赛规则，比赛时一个足球队的上场队员是 11 人。问：

(1)这位教练从这 17 名学员中可以形成多少种学员上场方案？

(2)如果在选出 11 名上场队员时，还要确定其中的守门员，那么教练员有多少种方式做这件事情？

[解析](1)由于上场学员没有角色差异，所以可以形成的学员上场方案有  $C_{17}^{11} = 12376$  (种)。

(2)教练员可以分两步完成这件事情：第 1 步，从 17 名学员中选出 11 人组成上场小组，共有  $C_{17}^{11}$  种选法；

第 2 步，从选出的 11 人中选出 1 名守门员，共有  $C_{11}^1$  种选法。

所以教练员做这件事情的方法数有  $C_{17}^{11} \times C_{11}^1 = 136136$  种。

## 题型二：有限制条件的组合问题

例 2 在我国武汉抗击新冠疫情战斗中，某医院从 10 名医疗专家中抽调 6 名奔赴抗疫前线，其中这 10 名医疗专家中有 4 名是呼吸科专家。问：

(1)抽调的 6 名专家中恰有 2 名是呼吸科专家的抽调方法有多少种？

(2)至少有 2 名呼吸科专家的抽调方法有多少种？

(3)至多有 2 名呼吸科专家的抽调方法有多少种？

[解析](1)分步：首先从 4 名呼吸科专家中任选 2 名，有  $C_4^2$  种选法，再从除呼吸科专家的 6 人中选取 4 人，有  $C_6^4$  种选法，所以共有  $C_4^2 \cdot C_6^4 = 90$  种抽调方法。

(2)“至少”的含义是不低于，有两种解答方法，

方法一(直接法)：按选取的呼吸科专家的人数分类：

①选 2 名呼吸科专家，共有  $C_4^2 \cdot C_6^4$  种选法；②选 3 名呼吸科专家，共有  $C_4^3 \cdot C_6^3$  种选法；

③选 4 名呼吸科专家，共有  $C_4^4 \cdot C_6^2$  种选法；共有  $C_4^2 \cdot C_6^4 + C_4^3 \cdot C_6^3 + C_4^4 \cdot C_6^2 = 185$  种抽调方法。

方法二(间接法)：不考虑是否有呼吸科专家，共有  $C_{10}^6$  种选法，考虑选取 1 名呼吸科专家参加，有  $C_4^1 \cdot C_6^5$  种选法；没有呼吸科专家参加，有  $C_6^6$  种选法，所以共有： $C_{10}^6 - C_4^1 \cdot C_6^5 - C_6^6 = 185$  种抽调方法。

(3)“至多 2 名”包括“没有”、“有 1 名”、“有 2 名”三种情况，分类解答。

①没有呼吸科专家参加，有  $C_6^6$  种选法；②有 1 名呼吸科专家参加，有  $C_4^1 \cdot C_6^5$  种选法；

③有 2 名呼吸科专家参加，有  $C_4^2 \cdot C_6^4$  种选法。所以共有  $C_6^6 + C_4^1 \cdot C_6^5 + C_4^2 \cdot C_6^4 = 115$  种抽调方法。

探究 2 解答有限制条件的组合问题的基本方法是“直接法”和“间接法(排除法)”。其中用直接法求解时，则应坚持“特殊元素优先选取”的原则，优先安排特殊元素的选取，再安排其他元素的选取。而选择间接法的原则是“正难则反”，也就是若正面问题分类较多、较复杂或计算量较大，不妨从反面问题入手，试一试看

是否简捷，特别是涉及“至多”、“至少”等组合问题时更是如此. 此时正确理解“都不是”、“不都是”、“至多”、“至少”等词语的确切含义是解决这些组合问题的关键.

**【练习 2】** 课外活动小组共 13 人，其中男生 8 人，女生 5 人，并且男、女生各指定一名队长. 现从中选 5 人主持某项活动，依下列条件各有多少种选法？

- (1) 只有一名女生当选；
- (2) 两名队长当选；
- (3) 至少有一名队长当选；
- (4) 至多有两女生当选；
- (5) 既要有队长，又要有女生当选.

**【解析】**(1) 一名女生，四名男生. 故共有  $C_5^1 \cdot C_8^4 = 350$  种.

(2) 将两名队长作为一类，其他 11 人作为一类，故共有  $C_2^2 \cdot C_{11}^3 = 165$  种.

(3) 至少有一名队长含有两类：只有一名队长和两名队长.

故共有：  $C_2^1 \cdot C_{11}^4 + C_2^2 \cdot C_{11}^3 = 825$  种，或采用排除法：  $C_{13}^5 - C_{11}^5 = 825$  种.

(4) 至多有两女生含有三类：有两女生、只有一女生、没有女生.

故选法为  $C_5^2 \cdot C_8^3 + C_5^1 \cdot C_8^4 + C_8^5 = 686$  种.

(5) 分两类：第一类女队长当选：  $C_{12}^4$ ；第二类女队长不当选：  $C_4^1 \cdot C_7^3 + C_4^2 \cdot C_7^2 + C_4^3 \cdot C_7^1 + C_4^4$ .

故选法共有：  $C_{12}^4 + C_4^1 \cdot C_7^3 + C_4^2 \cdot C_7^2 + C_4^3 \cdot C_7^1 + C_4^4 = 790$  种.

### 题型三：与几何有关的组合应用题

**例 3** (1) 四面体的一个顶点为  $A$ ，从其他顶点和各棱中点中取 3 个点，使它们和点  $A$  在同一平面上，有多少种不同的取法？

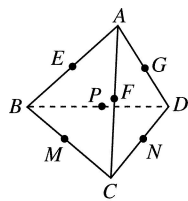
(2) 四面体的顶点和各棱中点共 10 个点，在其中取 4 个不共面的点，有多少种不同的取法.

**【解析】**(1)(直接法)如图，含顶点  $A$  的四面体的 3 个面上，除点  $A$  外都有 5 个点，

从中取出 3 点必与点  $A$  共面共有  $3C_5^3$  种取法；含顶点  $A$  的三条棱上各有三个点，

它们与所对的棱的中点共面，共有 3 种取法.

根据分类计数原理，与顶点  $A$  共面三点的取法有  $3C_5^3 + 3 = 33$  种.



(2)(间接法)如图，从 10 个点中取 4 个点的取法有  $C_{10}^4$  种，除去 4 点共面的取法种数可以得到结果. 从四面体同一个面上的 6 个点取出的 4 点必定共面. 有  $4C_6^4 = 60$  (种)，四面体的每一棱上 3 点与相对棱中点共面，共有 6 种共面情况，从 6 条棱的中点中取 4 个点时有 3 种共面情形(对棱中点连线两两相交且互相平分)，故 4 点不共面的取法为：  $C_{10}^4 - (60 + 6 + 3) = 141$  (种).

**探究 3** 利用组合知识解决与几何有关的问题时要注意：

- (1) 将已知条件中的元素的特征搞清，是用直接法还是用间接法；
- (2) 要使用分类方法，至于怎样确定分类标准，要具体问题具体分析；

(3)常用间接法解决该类问题.

【练习3】在 $\angle MON$ 的边 $OM$ 上有5个异于 $O$ 点的点, $ON$ 上有4个异于 $O$ 点的点,以这10个点(含 $O$ )为顶点,可以得到多少个三角形?

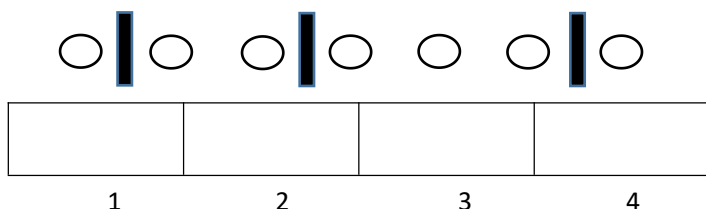
[解析]方法一 (直接法)分几种情况考虑:以 $O$ 为顶点的三角形中,另外两个顶点必须分别在 $OM$ 、 $ON$ 上,所以有 $C_5^1 \cdot C_4^1$ 个; $O$ 不为顶点的三角形中,两个顶点在 $OM$ 上,一个顶点在 $ON$ 上的有 $C_5^2 \cdot C_4^1$ 个,一个顶点在 $OM$ 上,两个顶点在 $ON$ 上的有 $C_5^1 \cdot C_4^2$ 个.因为这是分类问题,所以用分类计数原理,共有 $C_5^1 \cdot C_4^1 + C_5^2 \cdot C_4^1 + C_5^1 \cdot C_4^2 = 5 \times 4 + 10 \times 4 + 5 \times 6 = 90$ 个.

方法二 (间接法)先不考虑共线顶点的问题,从10个不同元素中任取3个点的组合数是 $C_{10}^3$ ,但其中 $OM$ 上的6个点(含 $O$ )中任取3个点不能得到三角形, $ON$ 上的5个点(含 $O$ )中任取3个点也不能得到三角形,所以共可以得到 $C_{10}^3 - C_6^3 - C_5^3$ ,即 $C_{10}^3 - C_6^3 - C_5^3 = 120 - 20 - 10 = 90$ 个.

方法三 也可以这样考虑,把 $O$ 看成是 $OM$ 边上的点,先从 $OM$ 上的6个点(含 $O$ )中取两点, $ON$ 上的4点(不含 $O$ )中取一点,有 $C_6^2 \cdot C_4^1$ 个三角形,再从 $OM$ 上的5点(不含 $O$ )中取一点,从 $ON$ 上的4点(不含 $O$ )中取两点,可得 $C_5^1 \cdot C_4^2$ 个三角形,所以共有 $C_6^2 \cdot C_4^1 + C_5^1 \cdot C_4^2 = 15 \times 4 + 5 \times 6 = 90$ 个.

例4 有7个参加数学竞赛的名额,要分给4所学校,每校至少一个名额,有多少种不同的名额分配方法?

[解析]可以看成7个相同的小球放入编号为1至4的4个盒子中,每个盒子至少有一个小球,共有多少种放法?现把7个小球从左至右一字排开,然后在小球之间的6个空中选出3个空插入隔板,如下图所示:7个小球被分成了从左至右1,2,3,1的4堆,然后把从左边数第一堆放入1号盒子,第二堆放入2号盒子,依此类推...,这样每有一种插入隔板的方法,就对应一种分配小球的方案.



所以共有 $C_6^3 = 20$ 种分配方法.

【练习4】已知方程: $a+b+c+d=10$ ,则该方程(1)有多少组正整数解?(2)有多少组自然数解?

[解析](1)可以看成把10个相同的小球放入编号为1至4的4个盒子中,每个盒子至少有一个小球.

所以共有 $C_9^3 = 84$ 组正整数解

(2)可以看成把10个相同的小球放入编号为1至4的4个盒子中,盒子可空.小球数和隔板数共13个,可以先从左至右一字排开放置13个小球,然后从这13个小球中选出3个小球换成隔板,共有 $C_{13}^3 = 286$ 种放法,这样每一种放置隔板的方法,就对应一种分配小球的方案,而且保证了盒子可空;

所以共有 $C_{13}^3 = 286$ 组自然数解.

【课后巩固】—————完成课时作业(170)

## 17.4 排列与组合的应用

### 【要点整合】

1.解排列组合的应用题，通常有以下几种途径

- (1)以元素为主体，即先满足特殊元素的要求，再考虑其他元素；
- (2)以位置为主体，即先满足特殊位置的要求，再考虑其他位置；
- (3)先不考虑附加条件，计算出排列或组合数，再减去不合要求的排列或组合数.

2.掌握典型题型的技巧解法

- (1)相邻问题——捆绑法；
- (2)相离问题——插空法；
- (3)多元问题——分类法；
- (4)标号排位问题——分步法；
- (5)定序问题——消序法；
- (6)多排问题——单排法；
- (7)至少问题——间接法；
- (8)选排问题——先取后排法；
- (9)组排问题——先组后排法.

3.处理排列组合问题的总原则

- (1)弄清事件的情景：首先搞清有无“顺序”要求，若有则用  $A_n^m$ ，反之用  $C_n^m$ ；其次弄清目标的实现，是分步达到的，还是分类达到的，从而正确选用计数原理，一个复杂问题往往是分类与分步交织在一起的；最后看一下元素可否重复.
- (2)掌握“双向”解题途径，即“正面凑”与“反过来剔”，一道题目“正面凑”繁，“反面剔”则简，反之亦然.
- (3)重视一题多解，通过一题多解积累解题方法，提高分析能力，同时它也是重要的检验方法.

### 【典例讲练】

#### 题型一：排列组合的综合问题

例1 从1,3,5,7中任取2个数字，从0,2,4,6,8中任取2个数字，组成没有重复数字的四位数，求其中能被5整除的四位数的个数.

[解析]符合条件的四位数的个位必须是0,5，但0不能排在首位，故0是其中的特殊元素，应优先安排，按照0排在个位，0排在十、百位和不含0为标准分为三类：

- ①0排在个位，能被5整除的四位数有  $A_1^1 C_4^1 C_4^2 A_3^3 = 144$  个；
- ②0排在十、百位，但5必须排在个位有  $A_2^2 A_1^1 C_4^1 C_3^2 A_2^2 = 48$  个；
- ③不含0，但5必须排在个位有  $A_1^1 C_4^2 C_3^1 A_3^3 = 108$  个.

由分类加法计数原理得所求的四位数共有300个.

探究1 本例采取先特殊(特殊元素、特殊位置)，再一般的解题思路.

【练习 1】用 0,1,2,3,4,5 这六个数字，

(1)可以组成多少个无重复数字的五位数；

(2)可以组成多少个无重复数字的五位奇数；

(3)可以组成多少个无重复数字的能被 5 整除的五位数.

[解析](1)方法 1: (直接法)从 1,2,3,4,5 这五个数字中任取一个作首位，有  $C_5^1$  种；余下的 5 个数字可排在后四位中的任何一个位置，有  $A_5^4$  种，由分步计数原理，共有  $C_5^1 A_5^4 = 600$  个.

方法 2: (间接法)不考虑任何限制，共有  $A_6^5$  种，而 0 作首位时，有  $A_5^4$  种，

故适合题意的数字个数为  $A_6^5 - A_5^4 = 600$  个.

(2)一个数是否为奇数取决于个位数字，所以个位为特殊位置，又 0 不能排在首位，所以 0 为特殊数字，应优先考虑，有  $C_3^1 C_4^1 A_4^3 = 288$  个.

(3)能被 5 整除的五位数，则个位数字是 0 或 5.

当个位数字是 0 时，共有  $A_5^4$  个；当个位数字是 5 时，共有  $C_4^1 A_4^3$  个，

由分类计数原理，适合题意的数字共有  $A_5^4 + C_4^1 A_4^3 = 216$  个.

例 2 从 6 名运动员中选出 4 名参加  $4 \times 100$  米接力赛，如果甲、乙两人都不跑第一棒，那么不同的参赛方案有( )

A. 180 种

B. 240 种

C. 300 种

D. 360 种

[解析]分三种：(1)甲、乙都不参加，有  $A_4^4 = 24$  (种). (2)甲、乙仅有 1 人参加，有  $C_2^1 C_3^1 A_4^3 = 144$  (种).

(3)甲、乙都参加有  $A_3^2 A_4^2 = 72$  (种)，由分类计数原理，所以共有  $24+144+72=240$  (种)，故 B 对.

探究 2 本例采取先选后排的解题思路.

【练习 2】从  $a$ 、 $b$ 、 $c$ 、 $d$ 、 $e$ 、 $f$  六个字母中每次取 4 个进行排列，若每个排列都含有  $a$ 、 $b$ ，且  $a$  在  $b$  前的排列有多少个？

[解析]方法一 第一步，在四个位置中任选两个作为  $a$ 、 $b$  的位置，因  $a$  在  $b$  前，顺序一定，故是组合，方法应有  $C_4^2 = 6$ . 第二步，在余下的两个位置由除  $a$ 、 $b$  外的另四个元素来排，因为有位置顺序的要求，故是排列问题，排法为  $A_4^2 = 12$  种，由乘法原理得排法应有  $C_4^2 \cdot A_4^2 = 72$  种.

方法二 第一步，从  $c$ 、 $d$ 、 $e$ 、 $f$  中任选出两个元素来，选法为  $C_4^2$ . 第二步，把选出的两个元素排在四个位置上的任意两个上，排法为  $A_4^2$ ，此时余下的两个位置即是  $a$  在前  $b$  在后的位置，只有一种选法，由乘法原理得  $C_4^2 \cdot A_4^2 = 72$  种.

方法三 第一步，从  $c$ 、 $d$ 、 $e$ 、 $f$  中选出两个元素来，选法为  $C_4^2$  种. 第二步，将  $a$ 、 $b$  与第一步选出的两个元素共四个元素排在四个不同的位置上，共有  $A_4^4 = 24$  种排法，又  $a$  应在  $b$  前，应除以 2，

故得排法为  $\frac{1}{2} C_4^2 \cdot A_4^4 = 72$  种.

## 题型二：排列、组合的综合应用

例3 有五张卡片，它们的正、反面分别写着0与1,2与3,4与5,6与7,8与9，将其中任意三张并排放在一起组成三位数，共可组成多少个不同的三位数？

[解析]方法一(直接法) 从0与1两个特殊值着眼，可分三类：

(1)取0不取1，可先从另四张卡片中选一张作百位，有 $C_4^1$ 种选法；0可在后两位，有 $C_2^1$ 种方法；最后剩下的三张中任取一张，有 $C_3^1$ 种方法；又除含0的那张外，其他两张都有正面或反面两种可能，故此时可得不同的三位数有 $C_4^1 C_2^1 C_3^1 2^2$ (个)。

(2)取1不取0，同上分析可得不同的三位数 $C_4^2 \cdot 2^2 \cdot A_3^3$ (个)。

(3)0和1都不取，有不同的三位数 $C_4^3 \cdot 2^3 \cdot A_3^3$ (个)。

综上所述，共有不同的三位数： $C_4^1 C_2^1 C_3^1 2^2 + C_4^2 \cdot 2^2 \cdot A_3^3 + C_4^3 \cdot 2^3 \cdot A_3^3 = 432$ (个)。

方法二 (间接法)任取三张卡片可以组成不同的三位数 $C_5^3 \cdot 2^3 \cdot A_3^3$ (个)，其中0在百位的有 $C_4^2 \cdot 2^2 \cdot A_2^2$ (个)，这是不合题意的，故共有不同的三位数： $C_5^3 \cdot 2^3 \cdot A_3^3 - C_4^2 \cdot 2^2 \cdot A_2^2 = 432$ (个)。

探究3 解排列组合的应用题，要注意三点：

(1)仔细审题，判断是排列问题还是组合问题；要按元素的性质分类，按事件发生的过程进行分步。

(2)深入分析，周密思考，分清是乘还是加，既不少也不多，多角度分析，全面考虑，提高逻辑推理能力。

(3)对有附加条件的比较复杂的排列组合应用题，要周密分析，设计出合理的方案，把复杂问题分解成若干简单的基本问题，然后再用分类计数原理或分步计数原理求解。

【练习3】(1)从5名男医生、4名女医生中选3名医生组成一个医疗小分队，要求其中男、女医生都有，则不同的组队方案共有( )

- A. 70种                      B. 80种                      C. 100种                      D. 140种

[解析]方法一 (间接法)当选择的3名医生都是男的或都是女的时候，共有 $C_5^3 + C_4^3 = 14$ 种方法，从9人中选3人一共有 $C_9^3 = 84$ 种方法，所以要求男、女都有，共有 $84 - 14 = 70$ 种组队方法。

方法二 (直接法)当队中有一名女医生时，有 $C_4^1 C_5^2 = 40$ 种组法，当队中有2名女医生时，有 $C_4^2 C_5^1 = 30$ 种组法，综上，共有70种组队方法。

(2)甲、乙两人从4门课程中各选修2门，则甲、乙所选的课程中至少有1门不相同的选法共有( )

- A. 6种                      B. 12种                      C. 30种                      D. 36种

[解析]甲、乙两人从4门课程中各选修2门的选法共有 $C_4^2 \times C_4^2 = 36$ 种，而甲、乙所选的课程中全部相同的选法共有 $C_4^2 = 6$ 种，所以甲、乙所选的课程中至少有1门不相同的选法共有 $36 - 6 = 30$ 种，选C。

## 题型三：较复杂的排列组合问题

例4 有4个不同的球，四个不同的盒子，把球全部放入盒内。

(1)共有多少种做法？

(2)恰有一个盒子不放球，有多少种放法？

(3)恰有一个盒内放 2 个球，有多少种放法？

(4)恰有两个盒子不放球，有多少种放法？

[解析](1)一个球一个球的放到盒子里去，每只球都可有 4 种独立的放法，

由分步乘法计数原理知，放法共有  $4^4=256$  (种)。

(2)为保证“恰有一个盒子不放球”，先从四个盒子中任意拿出去 1 个，即将 4 个球分成 2,1,1 的三组，有  $C_4^2$  种分法；然后再从三个盒子中选一个放两个球，其余两个球，两个盒子，全排列即可。由分步乘法计数原理知，共有放法  $C_4^1 \cdot C_4^2 \cdot C_3^1 \cdot A_2^2 = 144$  (种)。

(3)“恰有一个盒子内放 2 个球”，即另外的三个盒子放 2 个球，每个盒子至多放 1 个球，即另外三个盒子中恰有一个空盒。因此，“恰有一个盒子放 2 球”与“恰有一个盒子不放球”是一回事。故也有 144 种放法。

(4)先从四个盒子中任取两个有  $C_4^2$  种，问题转化为：“4 个球，两个盒子，每盒必放球，有几种放法？”从放球数目看，可分为(3,1)，(2,2)两类。第一类：可从 4 个球中先选 3 个，然后放入指定的一个盒子中即可，有  $C_4^3 \cdot C_2^1$  种放法；第二类：有  $C_4^2$  种放法。因此共有  $C_4^3 \cdot C_2^1 + C_4^2 = 14$  (种)。由分步乘法计数原理得“恰有两个盒子不放球”的放法有  $C_4^2 \cdot 14 = 84$  (种)。

探究 1 解排列组合问题的“16 字方针”是：有序排列、无序组合；分类为加，分步为乘。

【练习 4】某单位安排 7 位员工在 10 月 1 日至 7 日值班，每天安排 1 人，每人值班 1 天。若 7 位员工中的甲、乙排在相邻两天，丙不排在 10 月 1 日，丁不排在 10 月 7 日，则不同的安排方案共有( )

A. 504 种

B. 960 种

C. 1 008 种

D. 1 108 种

[解析]依题意，满足甲、乙两人值班安排在相邻两天的方法共有  $A_2^2 A_6^6 = 1440$  种，其中满足甲、乙两人值班安排在相邻两天且丙在 10 月 1 日值班的方法共有  $C_5^1 A_2^2 A_4^4 = 240$  种；满足甲、乙两人值班安排在相邻两天且丁在 10 月 7 日值班的方法有  $C_5^1 A_2^2 A_4^4 = 240$  种；满足甲、乙两人值班安排在相邻两天且丙在 10 月 1 日值班、丁在 10 月 7 日值班的方法共有  $C_4^1 A_2^2 A_3^3 = 48$  种，

因此满足题意的方法共有  $1440 - 2 \times 240 + 48 = 1008$  种，选 C。

#### 题型四：分组分配问题

例 5 六本不同的书，按照以下要求处理，各有几种分法？

(1)一堆一本，一堆两本，一堆三本；

(2)甲得一本，乙得两本，丙得三本；

(3)一人得一本，一人得二本，一人得三本；

(4)平均分给甲、乙、丙三人，每人两本；

(5)平均分成三堆，每堆两本。

[解析](1)先在六本书中任取一本，作为一堆，有  $C_6^1$  种取法；再从余下的五本书中任取两本，作为一堆，有  $C_5^2$  种取法；最后从余下的 3 本中取 3 本，作为一堆，有  $C_3^3$  种取法，故共有  $C_6^1 \cdot C_5^2 \cdot C_3^3 = 60$  种。

(2)由(1)知，分成三堆的方法有  $C_6^1 \cdot C_5^2 \cdot C_3^3$  种，而每种分组方法，就对应一种甲得一本，乙得两本，丙得三本的一种分配方法，故甲得一本，乙得两本，丙得三本的分法应为  $C_6^1 \cdot C_5^2 \cdot C_3^3 = 60$  种。

(3)由(1)知，分成三堆的方法有  $C_6^1 \cdot C_5^2 \cdot C_3^3$  种，但每一种分组方法，又有  $A_3^3$  种不同的分配方案，故一人得一本，一人得两本，一人得 3 本的分法有  $C_6^1 C_5^2 C_3^3 A_3^3 = 360$  种。

(4)三个人一个一个地来取书，甲从 6 本不同的书本中任取 2 本的方法有  $C_6^2$  种，甲不论用哪一种方法取得 2 本书后，乙再从余下的 4 本书中取 2 本有  $C_4^2$  种方法，而甲、乙不论用哪一种方法各取 2 本书后，丙从余下的两本书中取两本书，有  $C_2^2$  种方法，所以一共有  $C_6^2 \cdot C_4^2 \cdot C_2^2 = 90$  种。

(5)把六本不同的书分成三堆，每堆两本与把六本不同的书分给甲、乙、丙三人，每人两本的区别在于，后者相当于把 6 本不同的书，平均分成三堆后，再把每次分得的三堆书分给甲、乙、丙 3 人，因此，设把 6 本不同的书，平均分成三堆的方法有  $x$  种，那么把 6 本不同的书分给甲、乙、丙 3 人，每人 2 本的分法就应有  $xA_3^3$  种，由(4)知，把 6 本不同的书分给甲、乙、丙 3 人，每人 2 本的方法有  $C_6^2 C_4^2 C_2^2$  种。所以，

$$xA_3^3 = C_6^2 C_4^2 C_2^2, \text{ 则 } x = \frac{C_6^2 C_4^2 C_2^2}{A_3^3} = 15 \text{ 种.}$$

探究 2 对于非均匀“分组”“分配”问题[如(1)(2)(3)]可先“分组”再考虑“分配”，而对于均匀“分组”“分配”问题[如(4)(5)]，可先“分配”再考虑“分组”，无论使用先“分组”后“分配”，还是先“分配”后“分组”，都应因题而异，因人而异。

【练习 5】教委派 5 名教研员到 3 所学校去调研学生课业负担问题，每校至少 1 人，有多少种不同的选派方法？

[解析]先分组，再分配。先把 5 个人分成 3 组，有两种分法：

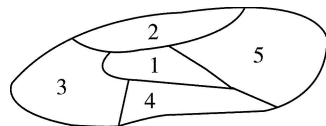
①一组 3 人，另两组各 1 人，有  $\frac{C_5^3 C_2^1 C_1^1}{A_2^2}$  种分法；

②一组 1 人另两组各 2 人，有  $\frac{C_5^1 C_4^2 C_2^2}{A_2^2}$  种分法。再分配到三所学校去有  $A_3^3$  种分法。

$$\therefore \text{共有} \left( \frac{C_5^3 C_2^1 C_1^1}{A_2^2} + \frac{C_5^1 C_4^2 C_2^2}{A_2^2} \right) A_3^3 = 150 \text{ 种方法.}$$

### 题型五：涂色问题

例 6 如下图所示，一个地区分为 5 个行政区域，现给地图着色，要求相邻区域不得使用同一颜色。现有 4 种颜色可供选择，则不同的着色方法共有 \_\_\_\_\_ 种。(以数字作答)



[解析]本小题在各类教辅资料上都能找到影子，但所给图形变化后，需要同学们有敏锐的观察力．本题能较深刻地测试逻辑思维能力．

因区域 1 与其他四个区域都相邻，宜先考虑．区域 1 有 4 种涂法．若区域 2、4 同色，有 3 种涂色，此时区域 3、5 均有两种涂法，涂法总数为  $4 \times 3 \times 2 \times 2 = 48$  种；若区域 2、4 不同色，先涂区域 2 有 3 种方法，再涂区域 4 有 2 种方法．此时区域 3、5 也都只有 1 种涂法，涂法总数为  $4 \times 3 \times 2 \times 1 \times 1 = 24$  种．因此涂法共有  $48 + 24 = 72$  种．

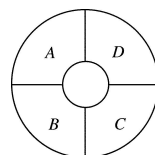
【练习 6】如图，一环形花坛分成 A、B、C、D 四块．现有 4 种不同的花供选种，要求在每块里种 1 种花，且相邻的 2 块种不同的花，则不同的种法总数为( )

A. 96

B. 84

C. 60

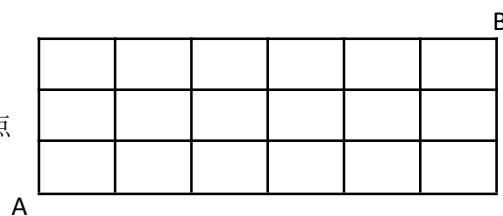
D. 48



[解析]根据题意可分情况讨论．取 2 种花来种，则有  $C_4^2 C_2^1 = 12$  种方法；取 3 种花来种，则有  $C_4^3 C_3^1 C_2^1 C_2^1 = 48$  种方法；取 4 种花来种，则有  $C_4^4 C_3^1 C_2^1 C_2^1 = 24$  种方法，故共有  $12 + 48 + 24 = 84$  种方法．

#### 题型六：网格问题

例 7 某小镇的街道是如图所示的矩形网格形状，现某人欲从图中 A 点位置沿街道步行到 B 点位置（走最短距离），共有多少种走法？



[解析]从 A 到 B 共有 8 步，其中 5 步向右走，3 步向上走，只需确定 8 步中哪 5 步向右走即可，所以共有：

$$C_8^5 = C_8^3 = 56 \text{ 种走法.}$$

【练习 7】例 7 图中共有多少个矩形？

[解析]图中每一个矩形都是由两条横线和两条竖线相交形成，所以只需先从 4 条横线中选出 2 条，再从 6 条竖线中选出 2 条即可产生一个矩形，共有： $C_4^2 \cdot C_6^2 = 90$  个矩形．

#### 题型七：错位排列问题

例 8 编号为 1、2、3 的 3 个人分别去坐编号为 1、2、3 的 3 个座位，所有人的编号与座位号都不一致的坐法有\_\_\_\_\_种．

[答案]2

[解析]易知 2 个人的错位排列共有 1 种做法，3 个人的错位排列可以用排除法，3 个人都对号入座有 1 种，没有恰好 2 个人对号入座的情况，恰好一个人对号入座的情况有  $C_3^1 \times 1$  种，所以共有： $A_3^3 - 1 - C_3^1 \times 1 = 2$  种坐法

【练习 8】编号为 1、2、3、4 的 4 个人分别去坐编号为 1、2、3、4 的 4 个座位，所有人的编号与座位号都不一致的坐法有\_\_\_\_\_种．

[答案]9

[解析] $A_4^4 - 1 - C_4^2 \times 1 - C_4^1 \times 2 = 9$  种

## 易错题型辨析

### 1. 两个原理混淆

两个原理的区别在于一个与分类有关，一个与分步有关。

例 9 某校高一有 6 个班，高二有 5 个班，高三有 8 个班，各年级分别举行班与班之间篮球单循环赛，则共需要进行比赛的场数为( )

- A.  $C_6^2 C_5^2 C_8^2$       B.  $C_6^2 + C_5^2 + C_8^2$       C.  $A_6^2 A_5^2 A_8^2$       D.  $C_{19}^2$

[错解]依题意，高一比赛有  $C_6^2$  场，高二比赛有  $C_5^2$  场，高三比赛有  $C_8^2$  场，由分步计数原理，得共需要进行比赛的场数为  $C_6^2 C_5^2 C_8^2$ ，选 A

[剖析]结合题意，各年级之间进行的比赛是分类计数，而不是分步计数。

[正解]依题意，高一比赛有  $C_6^2$  场，高二比赛有  $C_5^2$  场，高三比赛  $C_8^2$  场，由分类计数原理，得共需要进行比赛的场数为  $C_6^2 + C_5^2 + C_8^2$ ，选 B.

### 2. 排列组合混淆

怎样界定排列与组合问题？唯一的标准是“顺序”，“有序”是排列问题，“无序”是组合问题，排列与组合问题并存时，一般采用先组合后排列的方法。

例 10 七位身高各不相同的同学排成一排，要求正中间的最高，左右两边分别顺次一个比一个矮，这样的排法共有多少种？

[错解 1]最高的同学必须站在中间，再从其他 6 位同学中选取 3 位同学，有  $A_6^3$  种，剩下的 3 位同学也有  $A_3^3$  种，故共有  $A_6^3 A_3^3 = 720$  种。

[错解 2]最高的同学必须站在中间，再从其他 6 位同学中选取 3 位同学按从高到矮的顺序站在一边，有  $C_6^3$  种；剩下的 3 位同学也按从高到矮的顺序站在另一边，有  $C_3^3$  种。又两边可以交换，故共有  $C_6^3 A_2^2 = 40$  种。

[剖析]本题看似排列问题，其实是组合问题。

[正解 1]最高的同学必须站在中间，再从其他 6 位同学中选取 3 位同学按从高到矮的顺序站在一边，有  $C_6^3$  种，则剩下三位同学的位置已定。故共有  $C_6^3 = 20$  种。

### 3. 重复计数

例 11 七个人排成一排，甲不在排头，乙不在排尾的排法有几种？

[错解 1]排在排头的有除甲之外的  $A_6^1$  种情形，排在排尾的也有除乙之外的  $A_6^1$  种情形，两端排好后余下的排中间有  $A_5^5$  种情形，所以不同的排法有  $A_6^1 A_6^1 A_5^5 = 4320$  种。

[错解 2]头尾两个位置可从甲、乙之外的 5 人中选两人来排，有  $A_5^2$  种排法，余下的人排中间有  $A_5^5$  种排法，所以甲、乙不在排头、排尾的排法有  $A_5^2 A_5^5$  种；又甲、乙分别在排尾、排头的排法各有  $A_6^6$  种，因此不同的排法共有  $A_5^2 A_5^5 + 2A_6^6 = 3840$  种。

[剖析]对于错解 1 中排排头的 6 种情形也有乙不在排尾的情况，因此重复计算了  $5A_5^5$  种情形。对于错解 2

中甲在排尾且乙在排头已包含在甲在排尾或乙在排头的情形中, 因此重复计算了  $A_5^5$  种排法.

[正解 1]在错解 1 中, 减去重复数, 应为  $A_6^1 A_6^1 A_5^5 - 5A_5^5 = 3720$  种排法.

[正解 2]在错解 2 中, 减去重复数, 应为  $A_5^2 A_5^5 + 2A_6^5 - A_5^5 = 3720$  种排法.

#### 4. 遗漏计数

例 12  $A$ 、 $B$ 、 $C$ 、 $D$ 、 $E$  五人站成一排, 如果  $B$  必须站在  $A$  的右边( $A$ 、 $B$  可以不相邻), 那么不同的站法有( )种

A. 24 种

B. 60 种

C. 90 种

D. 120 种

[错解]把  $A$ 、 $B$ “捆绑”为一个元素( $B$  在  $A$  的右边), 与  $C$ 、 $D$ 、 $E$  一起全排列, 有  $A_4^4 = 24$  种站法, 故选 A.

[剖析]审题不严, 未注意到“ $A$ 、 $B$  可以不相邻”而漏解.

[正解 1]按  $A$  的位置分为四类:  $A$  排第一、二、三、四位时的排法数分别是  $A_4^4$ 、 $3A_3^3$ 、 $2A_3^3$ 、 $A_3^3$ ,

所以共有  $A_4^4 + 3A_3^3 + 2A_3^3 + A_3^3 = 60$  种排法, 选 B.

[正解 2]利用对称关系(注意到  $A$  在  $B$  左边与  $A$  在  $B$  右边的排列情形是对称相同的), 故有  $\frac{A_5^5}{2} = 60$  种,

选 B.

**【课后巩固】**—————完成课时作业 (171)

## 17.5 二项式定理

### 17.5.1 二项式定理的定义

#### 【学习目标】

- 1、能用计数原理证明二项式定理.
- 2、掌握二项式定理及其展开式的通项公式. (重点)
- 3、会用二项式定理解决与二项展开式有关的简单问题. (难点)

#### 【要点整合】

知识点一：二项式定理

$$(a+b)^n = C_n^0 a^n + C_n^1 a^{n-1}b + C_n^2 a^{n-2}b^2 + \cdots + C_n^k a^{n-k}b^k + \cdots + C_n^n b^n \quad (n \in N^*)$$

(1) 这个公式所表示的规律叫做二项式定理.

(2) 展开式：等号右边的多项式叫做  $(a+b)^n$  的二项展开式，展开式中一共有  $n+1$  项.

(3) 二项式系数：展开式各项的系数  $C_n^k$  ( $k \in \{0, 1, 2, \dots, n\}$ ) 叫做二项式系数.

知识点二：二项展开式的通项公式

$$(a+b)^n \text{ 展开式的第 } k+1 \text{ 项叫做二项展开式的通项, 记作: } T_{k+1} = C_n^k a^{n-k}b^k \quad (k=0, 1, 2, \dots, n).$$

注意：“二项式系数”与“系数”的区别：二项式系数与项的系数是不同的两个概念. 二项式系数是指  $C_n^0, C_n^1, \dots, C_n^n$ ，它只与各项的项数有关，而与  $a, b$  无关. 而项的系数是指该项中除变量外的常数部分，它不仅与各项的项数有关，而且也与  $a, b$  有关.

#### 【典例讲练】

题型一：二项式定理的正用、逆用

例 1 (1) 求  $(a+2b)^4$  的展开式；

(2) 求  $(2x - \frac{1}{3x})^3$  的展开式；

(3) 化简  $(x-1)^5 + 5(x-1)^4 + 10(x-1)^3 + 10(x-1)^2 + 5(x-1)$ .

[解析](1) 根据二项式定理  $(a+b)^n = C_n^0 a^n + C_n^1 a^{n-1}b + C_n^2 a^{n-2}b^2 + \cdots + C_n^k a^{n-k}b^k + \cdots + C_n^n b^n$ ，得：

$$(a+2b)^4 = C_4^0 a^4 + C_4^1 a^3(2b) + C_4^2 a^2(2b)^2 + C_4^3 a(2b)^3 + C_4^4 (2b)^4 = a^4 + 8a^3b + 24a^2b^2 + 32ab^3 + 16b^4.$$

$$(2) (2x - \frac{1}{3x})^3 = C_3^0 (2x)^3 + C_3^1 (2x)^2(-\frac{1}{3x}) + C_3^2 (2x)(-\frac{1}{3x})^2 + C_3^3 (-\frac{1}{3x})^3 = 8x^3 - 4x + \frac{2}{3x} - \frac{1}{27x^3}.$$

$$(3) \text{原式} = C_5^0 (x-1)^5 \cdot 1^0 + C_5^1 (x-1)^4 \cdot 1^1 + C_5^2 (x-1)^3 \cdot 1^2 + C_5^3 (x-1)^2 \cdot 1^3 + C_5^4 (x-1) \cdot 1^4 + C_5^5 (x-1)^0 \cdot 1^5 - 1 \\ = [(x-1)+1]^5 - 1 = x^5 - 1$$

规律方法 1.  $(a+b)^n$  的二项展开式有  $n+1$  项，是和的形式，各项的幂指数规律是：(1) 各项的次数等于  $n$ ；(2) 字母  $a$  按降幂排列，从第一项起，次数由  $n$  逐项减 1 直到 0；字母  $b$  按升幂排列，从第一项起，次数由 0 逐项加 1 直到  $n$ .

2.逆用二项式定理可以化简多项式,体现的是整体思想,注意分析已知多项式的特点,向二项展开式的形式靠拢.

【练习 1】化简:  $8-12(x+2)+6(x+2)^2-(x+2)^3=$ \_\_\_\_\_.

[解析]原式= $C_3^0 2^3 + C_3^1 2^2[-(x+2)] + C_3^2 2^1[-(x+2)]^2 + C_3^3 2^0[-(x+2)]^3 = [2-(x+2)]^3 = (-x)^3 = -x^3$

## 题型二: 二项展开式通项的应用

例 2 若  $(\sqrt{x} + \frac{1}{2\sqrt[4]{x}})^n$  展开式中前三项系数成等差数列,求:

(1)展开式中含  $x^{-\frac{1}{2}}$  的项; (2)展开式中所有的有理项; (3)展开式中有无常数项? 并说明理由.

[解析](1)由已知可得  $C_n^0 \left(\frac{1}{2}\right)^0 + C_n^2 \left(\frac{1}{2}\right)^2 = 2C_n^1 \left(\frac{1}{2}\right)^1$ , 即  $n^2 - 9n + 8 = 0$ , 解得  $n=8$  或  $n=1$  (舍去).

$$T_{k+1} = C_8^k (\sqrt{x})^{8-k} \left(\frac{1}{2\sqrt[4]{x}}\right)^k = C_8^k 2^{-k} \cdot x^{4-\frac{3}{4}k},$$

令  $4 - \frac{3}{4}k = -\frac{1}{2}$ , 得  $k=6$ . 所以含  $x^{-\frac{1}{2}}$  的项为  $T_7 = C_8^6 \cdot 2^{-6} \cdot x^{-\frac{1}{2}} = \frac{7}{16}x^{-\frac{1}{2}}$ .

(2)令  $4 - \frac{3}{4}k \in \mathbb{Z}$ , 且  $0 \leq k \leq 8$ , 则  $k=0, 4, 8$ , 所以有理项分别为  $T_1 = x^4$ ,  $T_5 = \frac{35}{8}x$ ,  $T_9 = \frac{1}{256x^2}$ .

(3) $T_{k+1} = C_8^k 2^{-k} \cdot x^{4-\frac{3}{4}k}$ , 令  $4 - \frac{3}{4}k = 0$ , 无自然数解, 所以展开式中没有常数项.

规律方法 1.利用二项式的通项求二项展开式的特定项的常见题型:

(1)求第  $k$  项,  $T_k = C_n^{k-1} a^{n-k+1} b^{k-1}$ ;

(2)求含  $x^k$  的项(或  $x^p y^q$  的项);

(3)求常数项;

(4)求有理项.

2.求二项展开式的特定项的常用方法

(1)对于常数项, 隐含条件是字母的指数为 0 (即 0 次项);

(2)对于有理项, 一般是先写出通项公式, 其所有的字母的指数恰好都是整数的项.

【练习 2】(1)在  $(2x^2 - \frac{1}{\sqrt[3]{x}})^7$  的展开式中, 求第 4 项的二项式系数及第 4 项的系数.

[解析](1)  $T_4 = C_7^3 (2x^2)^4 \left(-\frac{1}{\sqrt[3]{x}}\right)^3$ ,

第 4 项的二项式系数是  $C_7^3 = 35$ , 第 4 项的系数是  $C_7^3 \cdot 2^4 \cdot (-1)^3 = -560$ .

(2)  $(1+x)^6(1-x)^5$  展开式中含  $x^3$  的项为\_\_\_\_\_.

[答案]  $-5x^3$

### 题型三：二项式定理的应用

例 3 (1) 试求  $2^{51}$  除以 7 的余数；

(2) 求证： $3^{2n+2}-8n-9 (n \in \mathbf{N}^*)$  能被 64 整除；

(3) 求证： $x^{2n+1}+a^{2n+1}$  能被  $x+a$  整除.

[解析](1) 解： $2^{51}=8^{17}=(7+1)^{17}=C_{17}^0 7^{17}+C_{17}^1 7^{16}+\cdots+C_{17}^{16} 7+1$ ,

以上各项除了最后一项 1 之外，均为 7 的倍数，所以  $2^{51}$  除以 7 的余数为 1.

(2) 证明： $3^{2n+2}-8n-9=(8+1)^{n+1}-8n-9$

$$=C_{n+1}^0 8^{n+1} 1^0+\cdots+C_{n+1}^{n-1} 8^2 1^{n-1}+C_{n+1}^n 8^1 1^n+C_{n+1}^{n+1} 8^0 1^{n+1}-8n-9$$

$$=C_{n+1}^0 8^{n+1} 1^0+\cdots+C_{n+1}^{n-1} 8^2 1^{n-1}+8n+8+1-8n-9=C_{n+1}^0 8^{n+1} 1^0+\cdots+C_{n+1}^{n-1} 8^2 1^{n-1} \quad \blacklozenge$$

①式中的每一项都含有  $8^2$  这个因数，故原式能被 64 整除.

$$(3) x^{2n+1}+a^{2n+1}=[(x+a)-a]^{2n+1}+a^{2n+1}$$

$$=C_{2n+1}^0 (x+a)^{2n+1}-C_{2n+1}^1 (x+a)^{2n} a+\cdots+C_{2n+1}^{2n} (x+a) a^{2n}-a^{2n+1}+a^{2n+1}$$

$$=(x+a)^{2n+1}-(2n+1)(x+a)^{2n} a+\cdots+(2n+1)(x+a) a^{2n},$$

以上各项均含有因式  $x+a$ ，所以原式能被  $x+a$  整除.

规律方法 利用二项式定理可以解决求余数和整除的问题，通常需将底数化成两数的和与差的形式，且这种转化形式与除数有密切的关系.

【练习 3】已知  $n \in \mathbf{N}^*$ ，求证： $1+2+2^2+\cdots+2^{5n-1}$  能被 31 整除.

$$[\text{证明}] 1+2+2^2+\cdots+2^{5n-1}=\frac{1-2^{5n}}{1-2}=2^{5n}-1=32^n-1=(31+1)^n-1$$

$$=31^n+C_n^1 \times 31^{n-1}+\cdots+C_n^{n-1} \times 31+1-1=31 \times (31^{n-1}+C_n^1 31^{n-2}+\cdots+C_n^{n-1}),$$

显然括号内的数为正整数，故原式能被 31 整除.

【课后巩固】—————完成课时作业（172）

## 17.5.2 “杨辉三角”与二项式系数的性质

### 【学习目标】

- 1、了解杨辉三角，会用杨辉三角求二项式乘方次数不大时的各项的二项式系数.
- 2、理解二项式系数的性质并灵活运用.（重点）

### 【要点整合】

知识点一：杨辉三角的特点（杨辉：我国南宋数学家）

- 1、在同一行中每行两端都是 1，与这两个 1 等距离的项的系数相等.

$$\begin{array}{r}
 (a+b)^1 \text{-----} 1 \quad 1 \\
 (a+b)^2 \text{-----} 1 \quad 2 \quad 1 \\
 (a+b)^3 \text{-----} 1 \quad 3 \quad 3 \quad 1 \\
 (a+b)^4 \text{-----} 1 \quad 4 \quad 6 \quad 4 \quad 1 \\
 (a+b)^5 \text{-----} 1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1 \\
 (a+b)^6 \text{-----} 1 \quad 6 \quad 15 \quad 20 \quad 15 \quad 6
 \end{array}$$

- 2、在相邻的两行中，除 1 外的每一个数都等于它“肩上”两个数的和，即  $C_{n+1}^r = C_n^{r-1} + C_n^r$ .

知识点二：二项式系数的性质

|          |                                                                                                                                                                                                                                                                                                                       |
|----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 对称性      | 在 $(a+b)^n$ 展开式中，与首末两端“等距离”的两个二项式系数相等，即 $C_n^m = C_n^{n-m}$                                                                                                                                                                                                                                                           |
| 增减性与最大值  | <p>增减性：当 <math>k &lt; \frac{n+1}{2}</math> 时，二项式系数是逐渐增大；当 <math>k &gt; \frac{n+1}{2}</math> 时，二项式系数是逐渐减小.</p> <p>最大值：当 <math>n</math> 为偶数时，中间一项的二项式系数 <math>C_n^{\frac{n}{2}}</math> 最大；</p> <p>当 <math>n</math> 为奇数时，中间两项的二项式系数 <math>C_n^{\frac{n-1}{2}}</math> 和 <math>C_n^{\frac{n+1}{2}}</math> 相等，且同时取得最大值.</p> |
| 各二项式系数的和 | <p>① <math>C_n^0 + C_n^1 + C_n^2 + \cdots + C_n^n = 2^n</math>；</p> <p>② <math>C_n^0 + C_n^2 + C_n^4 + \cdots = +C_n^1 + C_n^3 + C_n^5 + \cdots = 2^{n-1}</math></p>                                                                                                                                                  |

### 【典例讲练】

#### 题型一 与杨辉三角有关的问题

例 1 如图，在由二项式系数所构成的杨辉三角中，第\_\_\_\_\_行中从左到右第 14 与第 15 个数的比为 2:3

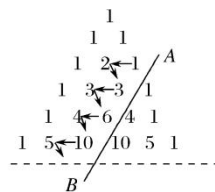
$$\begin{array}{c}
 \text{第 0 行 } 1 \\
 \text{第 1 行 } 1 \quad 1 \\
 \text{第 2 行 } 1 \quad 2 \quad 1 \\
 \text{第 3 行 } 1 \quad 3 \quad 3 \quad 1 \\
 \text{第 4 行 } 1 \quad 4 \quad 6 \quad 4 \quad 1 \\
 \text{第 5 行 } 1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1 \\
 \dots \quad \dots \quad \dots
 \end{array}$$

[答案]34

[解析]设第  $n$  行从左至右第 14 与第 15 个数之比为 2:3，则  $C_n^{13} : C_n^{14} = 2:3$ .

$$\therefore 3C_n^{13} = 2C_n^{14}, \text{ 即 } \frac{3n!}{13!(n-13)!} = \frac{2n!}{14!(n-14)!}, \text{ 得: } \frac{3}{n-13} = \frac{2}{14}, \therefore n=34.$$

【练习1】如图在“杨辉三角”中，斜线 $AB$ 的上方，从1开始箭头所示的数组成一个锯齿形数列：1, 2, 3, 3, 6, 4, 10, 5, ..., 记其前 $n$ 项和为 $S_n$ ，求 $S_{19}$ 的值.



[解析]由题图知，数列中的首项是 $C_2^2$ ，第2项是 $C_2^1$ ，第3项是 $C_3^2$ ，第4项是 $C_3^1$ ，...

第17项是 $C_{10}^2$ ，第18项是 $C_{10}^1$ ，第19项是 $C_{11}^2$ .

$$\begin{aligned} \therefore S_{19} &= (C_2^1 + C_2^2) + (C_3^1 + C_3^2) + (C_4^1 + C_4^2) + \cdots + (C_{10}^1 + C_{10}^2) + C_{11}^2 = C_3^2 + C_4^2 + C_5^2 + \cdots + C_{11}^2 + C_{11}^2 \\ &= C_3^3 + C_3^2 + C_4^2 + C_5^2 + \cdots + C_{11}^2 - 1 + C_{11}^2 = C_{12}^3 - 1 + C_{11}^2 = 274. \end{aligned}$$

规律方法 解决与杨辉三角有关问题的一般思路

(1) 观察找出每一行数据间的相互联系以及行与行间数据的相互联系.

(2) 将数据间的这种联系用数学式表达出来，使问题得解.

(3) 注意观察方向：横看、竖看、斜看、连续看、隔行看，从多角度观察.

## 题型二：二项展开式的系数和问题

例2 已知 $(2x-1)^5 = a_0x^5 + a_1x^4 + a_2x^3 + a_3x^2 + a_4x + a_5$ ，求：

(1)  $a_5$ ；

(2)  $a_0 + a_1 + a_2 + a_3 + a_4 + a_5$ ；

(3)  $|a_0| + |a_1| + |a_2| + |a_3| + |a_4| + |a_5|$ ；

(4)  $a_0 + a_2 + a_4$ .

[解析](1)  $a_5$ 为展开式的常数项， $\therefore a_5 = C_5^5 \cdot (-1)^5 = -1$ .

(2) 令 $x=1$ ，得： $(2 \times 1 - 1)^5 = a_0 + a_1 + a_2 + a_3 + a_4 + a_5$ ， $\therefore a_0 + a_1 + a_2 + a_3 + a_4 + a_5 = 1$ .

(3) 方法1： $\because (2x-1)^5$ 的展开式中偶数项的系数为负值，

$$\therefore |a_0| + |a_1| + |a_2| + |a_3| + |a_4| + |a_5| = a_0 - a_1 + a_2 - a_3 + a_4 - a_5.$$

令 $x=-1$ ，得： $[2 \times (-1) - 1]^5 = -a_0 + a_1 - a_2 + a_3 - a_4 + a_5$ ，

即 $a_0 - a_1 + a_2 - a_3 + a_4 - a_5 = -(-3)^5 = 3^5$ ， $\therefore |a_0| + |a_1| + |a_2| + |a_3| + |a_4| + |a_5| = 3^5 = 243$ .

方法2：所求即为 $(2x+1)^5$ 各项展开式系数之和，令 $x=1$ 得系数和为 $3^5 = 243$ ，为所求.

$$(4) \text{由上题得: } \begin{cases} a_0 + a_1 + a_2 + a_3 + a_4 + a_5 = 1, \\ a_0 - a_1 + a_2 - a_3 + a_4 - a_5 = 243, \end{cases} \text{ 两式相加得 } a_0 + a_2 + a_4 = \frac{1}{2} \times (1 + 243) = 122.$$

规律方法 (1) 赋值法是求二项展开式系数及有关问题的常用方法，注意取值要有利于问题的解决，可以取一个值或几个值，也可以取几组值，解决问题时要避免漏项.

(2) 一般地，对于多项式 $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ ，各项系数和为 $f(1)$ ，

奇次项系数和为 $\frac{1}{2}[f(1) - f(-1)]$ ，偶次项系数和为 $\frac{1}{2}[f(1) + f(-1)]$ ， $a_0 = f(0)$ .

【练习 2】已知  $(1-3x)^6 = a_0 + a_1x + a_2x^2 + \cdots + a_6x^6$ ，求：

(1)  $a_1 + a_2 + \cdots + a_6$ ； (2)  $a_0 + a_2 + a_4 + a_6$ ； (3)  $|a_0| + |a_1| + |a_2| + \cdots + |a_6|$ 。

[解析](1)令  $x=1$ ，得  $a_0 + a_1 + \cdots + a_6 = (-2)^6 = 64$  ①，而  $a_0 = 1$ ， $\therefore a_1 + \cdots + a_6 = 63$ 。

(2)令  $x=-1$ ，得  $a_0 - a_1 + a_2 - a_3 + a_4 - a_5 + a_6 = 4^6$  ②

①+②，得  $2(a_0 + a_2 + a_4 + a_6) = 2^6 + 4^6$ ， $\therefore a_0 + a_2 + a_4 + a_6 = \frac{1}{2} \times (2^6 + 4^6) = 2080$ ，

(3)同  $(1+3x)^6$  各项系数之和， $\therefore |a_0| + |a_1| + |a_2| + \cdots + |a_6| = (1+3)^6 = 4^6 = 4096$ 。

### 题型三：二项式系数性质的应用

例 3 已知  $f(x) = (\sqrt[3]{x^2} + 3x^2)^n$  展开式中各项的系数和比各项的二项式系数和大 992。

(1)求展开式中二项式系数最大的项；(2)求展开式中系数最大的项。

[解析](1)令  $x=1$ ，则展开式各项系数的和为  $f(1) = (1+3)^n = 4^n$ ，

又展开式中各项的二项式系数之和为  $2^n$ 。由题意知， $4^n - 2^n = 992$ 。

$$\therefore (2^n)^2 - 2^n - 992 = 0, \therefore (2^n + 31)(2^n - 32) = 0,$$

$$\therefore 2^n = -31 \text{ (舍)} \text{ 或 } 2^n = 32, \therefore n = 5.$$

由于  $n=5$  为奇数，所以展开式中二项式系数最大的项为中间两项，它们分别是

$$T_3 = C_5^2 (x^{\frac{2}{3}})^3 (3x^2)^2 = 90x^6, \quad T_4 = C_5^3 (x^{\frac{2}{3}})^2 (3x^2)^3 = 270x^{\frac{22}{3}}.$$

(2)展开式的通项公式为  $T_{r+1} = C_5^r 3^r x^{\frac{2}{3}(5+2r)}$ 。假设  $T_{r+1}$  项系数最大，则有 
$$\begin{cases} C_5^r 3^r \geq C_5^{r-1} 3^{r-1} \\ C_5^r 3^r \geq C_5^{r+1} 3^{r+1} \end{cases}$$

$$\therefore \begin{cases} \frac{5!}{r!(5-r)!} 3 \geq \frac{5!}{(r-1)!(6-r)!} \\ \frac{5!}{r!(5-r)!} \geq \frac{5!}{(r+1)!(4-r)!} 3 \end{cases} \therefore \begin{cases} \frac{1}{r} 3 \geq \frac{1}{(6-r)} \\ \frac{1}{(5-r)} \geq \frac{1}{(r+1)} 3 \end{cases} \therefore \frac{7}{2} \leq r \leq \frac{9}{2}, \therefore r \in \mathbf{N}, \therefore r = 4.$$

$$\therefore \text{展开式中系数最大的项为 } T_5 = C_5^4 3^4 x^{\frac{26}{3}} = 405x^{\frac{26}{3}}.$$

规律方法 (1)求二项式系数最大的项，要依据二项式系数的性质对  $(a+b)^n$  中的  $n$  进行讨论， $n$  为奇数时中间两项的二项式系数最大； $n$  为偶数时，中间一项的二项式系数最大。

(2)求展开式中系数最大项与求二项式系数最大项是不同的。求展开式系数最大的项，如求

$(a+bx)^n (a, b \in \mathbf{R})$  展开式中系数最大的项，一般是采用待定系数法。设展开式各项系数分别为

$A_1, A_2, \dots, A_{n+1}$ ，且第  $r+1$  项系数最大，应用 
$$\begin{cases} A_r \geq A_{r-1} \\ A_r \geq A_{r+1} \end{cases}$$
 解出  $r$  来，即得系数最大的项。

【练习3】已知  $(\sqrt{x} - \frac{2}{x^2})^n$  ( $n \in N^*$ ) 的展开式中第5项的系数与第3项的系数的比是10:1.

(1)求展开式中各项系数的和;

(2)求展开式中含  $x^{\frac{3}{2}}$  的项;

(3)求展开式中系数的绝对值最大的项.

[解析]  $\therefore (\sqrt{x} - \frac{2}{x^2})^n$  的展开式的通项是  $T_{r+1} = C_n^r (\sqrt{x})^{n-r} \cdot (-\frac{2}{x^2})^r = (-2)^r C_n^r x^{\frac{n-5r}{2}}$  ( $0 \leq r \leq n, r \in N$ )

$\therefore T_5 = (-2)^4 C_n^4 x^{\frac{n-10}{2}}$ ,  $T_3 = (-2)^2 C_n^2 x^{\frac{n-5}{2}}$ .  $\therefore \frac{2^4 C_n^4}{2^2 C_n^2} = \frac{10}{1}$ ,  $\therefore n^2 - 5n - 24 = 0$ , 得  $n=8$  或  $n=-3$  (舍去).

(1)令  $x=1$ , 则  $(\sqrt{x} - \frac{2}{x^2})^8 = (1-2)^8 = 1$ , 即所求各项系数的和为1.

(2)展开式的通项为  $T_{r+1} = (-2)^r C_8^r x^{\frac{8-5r}{2}}$  ( $0 \leq r \leq 8, r \in N$ ). 令  $\frac{8-5r}{2} = \frac{3}{2}$ , 得  $r=1$ ,

$\therefore$  展开式中含  $x^{\frac{3}{2}}$  的项为  $T_2 = (-2)^1 C_8^1 x^{\frac{3}{2}} = -16x^{\frac{3}{2}}$ .

(3)展开式的第  $r$  项、第  $r+1$  项、第  $r+2$  项的系数的绝对值分别为  $C_8^{r-1} 2^{r-1}$ ,  $C_8^r 2^r$ ,  $C_8^{r+1} 2^{r+1}$ . 若第  $r+1$  项

的系数绝对值最大, 则有  $\begin{cases} C_8^{r-1} 2^{r-1} \leq C_8^r 2^r \\ C_8^r 2^r \geq C_8^{r+1} 2^{r+1} \end{cases}$  解得  $5 \leq r \leq 6$ ,

故系数的绝对值最大的项为第6项和第7项, 即  $T_6 = -1792x^{-\frac{17}{2}}$ ,  $T_7 = 1792x^{-11}$ .

【课后巩固】—————完成课时作业(173)

## 17.6 习题课：排列与组合、二项式定理

### 【要点整合】

1. 两个计数原理：分类加法计数原理与分步乘法计数原理是排列组合中解决问题的重要手段，也是基础方法，正确地分类与分步是用好两个原理的关键。

2. 排列与组合：对于应用题，则首先要分清是否有序，即是排列问题还是组合问题。

排列数与组合数计算公式主要应用于求值和证明恒等式。在证明恒等式时，要注意观察恒等式左右两边的形式，基本遵循由繁到简的原则，有时也会从两边向中间靠拢。

3. 排列数与组合数公式及性质

|    | 排列与排列数                                                            | 组合与组合数                                                                                               |
|----|-------------------------------------------------------------------|------------------------------------------------------------------------------------------------------|
| 公式 | 排列数公式<br>$A_n^m = n(n-1)(n-2)\cdots(n-m+1)$ $= \frac{n!}{(n-m)!}$ | 组合数公式<br>$C_n^m = \frac{A_n^m}{A_m^m} = \frac{n(n-1)(n-2)\cdots(n-m+1)}{m!}$ $= \frac{n!}{m!(n-m)!}$ |
| 性质 | 当 $m=n$ 时， $A_n^m$ 为全排列 $A_n^n = n!$ ；<br>$0! = 1$                | $C_n^0 = C_n^n = 1$ ；<br>$C_n^m = C_n^{n-m}$ ；<br>$C_n^m + C_n^{m-1} = C_{n+1}^m$                    |
| 备注 | $A_n^m = C_n^m \cdot m!$ ， $n, m \in N^*$ 且 $m \leq n$            |                                                                                                      |

4. 二项式定理

(1)与二项式定理有关：包括定理的正向应用、逆向应用，题型如证明整除性、证明一些简单的组合恒等式等，此时主要是要构造二项式，合理应用展开式；

(2)与通项公式有关：主要是求特定项，比如常数项、有理项、 $x$  的某次幂等，此时要特别注意二项式展开式的通项是  $T_{k+1} = C_n^k a^{n-k} b^k$  ( $k=0,1,\dots,n$ )，其二项式系数是  $C_n^k$ ，而不是  $C_n^{k+1}$ ，这是一个易错点。

### 【典例讲练】

#### 题型一：排列与组合的应用

例 1 袋中装有带有编号的红球和白球共 16 个，现从中任取 2 球，若取出的 2 球是同色的取法和取出的 2 球是异色的取法相等，则取出 2 球都是红球的取法有多少种？

[解析]设红球有  $x$  个，则白球的个数为  $16-x$ 。

则取出的 2 球同色的取法为  $C_x^2 + C_{16-x}^2$  种，2 球异色的取法为  $C_x^1 C_{16-x}^1$  种。

根据题意，得  $C_x^2 + C_{16-x}^2 = C_x^1 C_{16-x}^1$ ，整理得  $x^2 - 16x + 60 = 0$ ，解得  $x=6$  或  $x=10$ ，符合题意。

(1)若  $x=6$ ，则取出的 2 球都是红球的取法共有  $C_6^2 = 15$  (种)；

(2)若  $x=10$ ，则取出的 2 球都是红球的取法共有  $C_{10}^2 = 45$  (种)。

## 题型二：二项式定理的应用

例2 求  $(x + \frac{1}{x} - 1)^5$  的展开式中的常数项.

[解析]解法一:  $\because \left(x + \frac{1}{x} - 1\right)^5 = \left[\left(x + \frac{1}{x}\right) - 1\right]^5$ ,

$\therefore$  展开式的通项为  $T_{r+1} = C_5^r \cdot \left(x + \frac{1}{x}\right)^{5-r} \cdot (-1)^r$  ( $r=0,1,2,\dots,5$ ). 当  $r=5$  时,  $T_6 = C_5^5 \cdot (-1)^5 = -1$ .

当  $0 \leq r < 5$  时,  $\left(x + \frac{1}{x}\right)^{5-r}$  的展开式的通项为  $T'_{k+1} = C_{5-r}^k \cdot x^{5-r-k} \cdot \left(\frac{1}{x}\right)^k$

$= C_{5-r}^k \cdot x^{5-r-2k}$  ( $k=0,1,2,\dots,5-r$ ). 欲求常数项, 令  $5-r-2k=0$ , 即  $r+2k=5$ .

$\because 0 \leq r < 5$ , 且  $k, r \in \mathbf{N}^*$ ,  $\therefore r$  只能取 1 或 3, 相应的  $k$  值分别为 2 或 1,

即  $\begin{cases} r=1, \\ k=2, \end{cases}$  或  $\begin{cases} r=3, \\ k=1. \end{cases} \therefore$  常数项为  $C_5^1 C_4^2 (-1)^1 + C_5^3 C_2^1 (-1)^3 + (-1) = -51$ .

解法二:  $(x + \frac{1}{x} - 1)^5 = (x + \frac{1}{x} - 1)(x + \frac{1}{x} - 1) \cdots (x + \frac{1}{x} - 1)$ , 共 5 个因式, 常数项为:

$$(-1)^5 + C_5^1 x \cdot C_4^1 \left(\frac{1}{x}\right) \cdot (-1)^3 + C_5^2 x^2 \cdot C_3^2 \left(\frac{1}{x}\right)^2 \cdot (-1) = -1 - 20 - 30 = -51.$$

(注: 常数项由三部分构成: 五个因式都取 -1 作积; 取一个因式抽取其中的  $x$ , 另取一个因式抽取其中的  $\frac{1}{x}$ ,

其余因式取出其中的 -1 作积; 取两个因式抽取其中的  $x$ , 另取两个因式抽取其中的  $\frac{1}{x}$ , 剩下一个因式取

出其中的 -1 作积)  $C_4^1 \times 2 - C_4^3 \times C_3^2 \times 2^2 = -40$

【练习 1】(1)  $(x + \frac{1}{x} - 2)^5$  展开式的常数项是\_\_\_\_\_.

[答案] -252

(2)  $(x^2 - x - 2)^4$  的展开式中,  $x^3$  的系数为\_\_\_\_\_.

[答案] -40

[解析]  $(x^2 - x - 2)^4 = [x^2 - (x+2)]^4$ , 展开后只有  $(x+2)^4$  与  $-C_4^3 x^2 (x+2)^3$  中含  $x^3$  项,

其系数和为  $C_4^1 \times 2 - C_4^3 \times C_3^2 \times 2^2 = -40$ .

例3 若  $(2x + \sqrt{3})^4 = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4$ , 则  $(a_0 + a_2 + a_4)^2 - (a_1 + a_3)^2$  的值为\_\_\_\_\_.

[答案] 1

[解析] 在  $(2x + \sqrt{3})^4 = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4$  中,

令  $x=1$ , 得  $(2+\sqrt{3})^4 = a_0 + a_1 + a_2 + a_3 + a_4$ ; 令  $x=-1$ , 得  $(-2+\sqrt{3})^4 = a_0 - a_1 + a_2 - a_3 + a_4$ .

两式相乘, 得:  $(2+\sqrt{3})^4 (-2+\sqrt{3})^4 = (a_0 + a_1 + a_2 + a_3 + a_4) \cdot (a_0 - a_1 + a_2 - a_3 + a_4)$ .

所以  $(a_0 + a_2 + a_4)^2 - (a_1 + a_3)^2 = (-4+3)^4 = 1$ .

【练习 2】若  $(x^2+1)(x-3)^9 = a_0 + a_1(x-2) + a_2(x-2)^2 + a_3(x-2)^3 + \cdots + a_{11}(x-2)^{11}$ , 则

$a_1 + a_2 + a_3 + \cdots + a_{11}$  的值为\_\_\_\_\_.

[答案]5

[解析]令  $x=2$ , 得  $a_0 = (2^2+1)(2-3)^9 = -5$ , 令  $x=3$ , 则  $a_0 + a_1 + a_2 + a_3 + \cdots + a_{11} = (3^2+1)(3-3)^9 = 0$ ,

所以  $a_1 + a_2 + a_3 + \cdots + a_{11} = -a_0 = 5$ .

例 4 若将函数  $f(x) = x^5$  表示为  $f(x) = a_0 + a_1(1+x) + a_2(1+x)^2 + \cdots + a_5(1+x)^5$ , 其中  $a_0, a_1, a_2, \cdots, a_5$  为实数, 则  $a_4 =$ \_\_\_\_\_.

[答案]-5

[解析]  $f(x) = x^5 = [-1 + (1+x)]^5$ ,  $\therefore a_4 = C_5^4(-1) = -5$ .

例 5 计算  $0.997^5$  的近似值, 精确到 0.001.

[解析]  $0.997^5 = (1-0.003)^5 = 1 - C_5^1 \times 0.003 + C_5^2 \times 0.003^2 - \cdots - C_5^5 \times 0.003^5 \approx 1 - 0.015 = 0.985$ .

$\therefore 0.997^5$  的近似值为 0.985.

【课后巩固】\_\_\_\_\_完成课时作业 (174)

## 第十八章 概率与分布

### 18.1 有限样本空间与随机事件

### 18.2 事件的关系和运算

#### 【学习目标】

1. 了解随机试验、样本空间的概念；
2. 通过实例，了解必然事件、不可能事件与随机事件的含义；
3. 理解并掌握事件的关系和运算；
4. 能够将事件的运算关系知识灵活运用 to 实际事件中。

#### 【要点整合】

##### 要点一：样本空间

1. 随机试验：我们把对随机现象的实现和对它的观察称为随机试验 (*random experiment*)，简称试验，常用字母  $E$  表示。
2. 随机试验的特点
  - (1) 试验可以在相同条件下重复进行；
  - (2) 试验的所有可能结果是明确可知的，并且不止一个；
  - (3) 每次试验总是恰好出现这些可能结果中的一个，但事先不能确定出现哪一个结果。
3. 样本空间：我们把随机试验  $E$  的每个可能的基本结果称为样本点，全体样本点的集合称为试验  $E$  的样本空间 (*sample space*)。一般地，我们用  $\Omega$  表示样本空间，用  $\omega$  表示样本点。在本书中，我们只讨论  $\Omega$  为有限集的情况。如果一个随机试验有  $n$  个可能结果  $\omega_1, \omega_2, \dots, \omega_n$ ，则称样本空间  $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$  为有限样本空间。

##### 要点二：随机事件

1. 随机事件：一般地，随机试验中的每个随机事件都可以用这个试验的样本空间的子集来表示。为了叙述方便，我们将样本空间  $\Omega$  的子集称为随机事件 (*random event*)，简称事件，并把只包含一个样本点的事件称为基本事件 (*elementary event*)。随机事件一般用大写字母  $A, B, C, \dots$  表示。在每次试验中，当且仅当  $A$  中某个样本点出现时，称为事件  $A$  发生。
2. 必然事件，不可能事件：在每次试验中总有一个样本点发生，所以  $\Omega$  总会发生，我们称  $\Omega$  为必然事件。而空集  $\emptyset$  不包含任何样本点，在每次试验中都不会发生，我们称  $\emptyset$  为不可能事件。

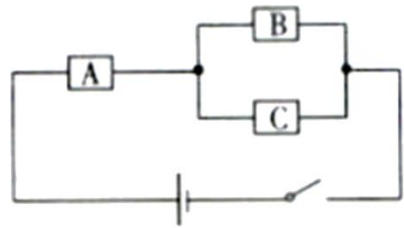
要点三：事件的关系和运算

|       |      | 定义                                                                                           | 表示法                                                   |
|-------|------|----------------------------------------------------------------------------------------------|-------------------------------------------------------|
| 事件的运算 | 包含关系 | 一般地，对于事件 A 与事件 B，如果事件 A 发生，则事件 B 一定发生，这时称事件 B 包含事件 A(或称事件 A 包含于事件 B)                         | $B \supseteq A$ (或 $A \subseteq B$ )                  |
|       | 并事件  | 若某事件发生当且仅当事件 A 发生或事件 B 发生，则称此事件为事件 A 与事件 B 的并事件(或和事件)                                        | $A \cup B$ (或 $A+B$ )                                 |
|       | 交事件  | 若某事件发生当且仅当事件 A 发生且事件 B 发生，则称此事件为事件 A 与事件 B 的交事件(或积事件)                                        | $A \cap B$ (或 $AB$ )                                  |
|       | 互斥关系 | 若 $A \cap B$ 为不可能事件，则称事件 A 与事件 B 互斥                                                          | 若 $A \cap B = \emptyset$ ，则 A 与 B 互斥                  |
|       | 对立关系 | 若 $A \cap B$ 为不可能事件， $A \cup B$ 为必然事件，那么称事件 A 与事件 B 互为对立事件，可记为 $B = \bar{A}$ 或 $A = \bar{B}$ | 若 $A \cap B = \emptyset$ ， $A \cup B = U$ ，则 A 与 B 对立 |

【典例讲练】

题型一：样本空间

例 1 如图，一个电路中有 A, B, C 三个电器元件，每个元件可能正常，也可能失效，把这个电路是否为通路看成是一个随机现象，观察这个电路中各元件是否正常.

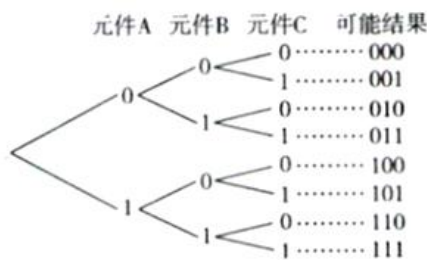


- (1) 写出试验的样本空间；
- (2) 用集合表示下列事件： $M$ ="恰好两个元件正常"； $N$ ="电路是通路"； $T$ ="电路是断路"

[解析]分别用  $x_1, x_2$  和  $x_3$  表示元件 A, B 和 C 的可能状态，则这个电路的工作状态可用  $(x_1, x_2, x_3)$  表示，进一步地，用 1 表示元件的“正常”状态，用 0 表示“失效”状态。

- (1) 则样本空间  $\Omega = \{(0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 0), (1, 0, 1), (0, 1, 1), (1, 1, 1)\}$

如图，还可以借助树状图帮助我们列出试验的所有可能结果



(2) “恰好两个元件正常”等价于  $(x_1, x_2, x_3) \in \Omega$ ，且  $x_1, x_2, x_3$  中恰有两个为 1，所以

$$M = \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}.$$

“电路是通路”等价于  $(x_1, x_2, x_3) \in \Omega$ ， $x_1 = 1$ ，且  $x_2, x_3$  中至少有一个是 1，所以

$$N = \{(1, 1, 0), (1, 0, 1), (1, 1, 1)\}.$$

同理，“电路是断路”等价于  $(x_1, x_2, x_3) \in \Omega$ ， $x_1 = 0$ ，或  $x_1 = 1, x_2 = x_3 = 0$ 。所以

$$N = \{(0, 0, 0), (0, 1, 0), (0, 0, 1), (0, 1, 1), (1, 0, 0)\}.$$

【练习 1】某人做试验，从一个装有标号为 1,2,3,4 的小球的盒子中，无放回地取两个小球，每次取一个，先取的小球的标号为  $x$ ，后取的小球的标号为  $y$ ，这样构成有序实数对  $(x, y)$ 。

(1) 写出这个试验的样本空间；

(2) 写出“第一次取出的小球上的标号为 2”这一事件。

[解析](1) 当  $x=1$  时， $y=2,3,4$ ；当  $x=2$  时， $y=1,3,4$ ；当  $x=3$  时， $y=1,2,4$ ；当  $x=4$  时， $y=1,2,3$ 。因此，这个试验的样本空间是  $\{(1,2), (1,3), (1,4), (2,1), (2,3), (2,4), (3,1), (3,2), (3,4), (4,1), (4,2), (4,3)\}$ 。

(2) 记“第一次取出的小球上的标号为 2”为事件  $A$ ，则  $A = \{(2,1), (2,3), (2,4)\}$ 。

## 题型二：必然事件、不可能事件与随机事件的判断

例 2 指出下列事件是必然事件、不可能事件还是随机事件：

(1) 中国体操运动员将在下届奥运会上获得全能冠军。

(2) 出租车司机小李驾车通过几个十字路口都将遇到绿灯。

(3) 若  $x \in \mathbb{R}$ ，则  $x^2 + 1 \geq 1$ 。

(4) 掷一枚骰子两次，朝上面的数字之和小于 2。

[答案](1)(2)是随机事件；(3)是必然事件；(4)是不可能事件。

[解析]由题意知(1)(2)中事件可能发生，也可能不发生，所以是随机事件；(3)中事件一定会发生，是必然事件；由于骰子朝上面的数字最小是 1，两次朝上面的数字之和最小是 2，不可能小于 2，所以(4)中事件不可能发生，是不可能事件。

【练习 2】下列事件：①一个口袋内装有 5 个红球，从中任取一球是红球；②掷两枚骰子，所得点数之和为 9；③ $x^2 \geq 0 (x \in \mathbb{R})$ ；④方程  $x^2 - 3x + 5 = 0$  有两个不相等的实数根；⑤巴西足球队会在下届世界杯足球赛中夺得冠军，其中随机事件的个数为( )

- A. 1                      B. 2                      C. 3                      D. 4

[答案]B

[解析]在所给条件下，①是必然事件；②是随机事件；③是必然事件；④是不可能事件；⑤是随机事件.

### 题型三：事件关系的判断

例 3 一个袋子中有大小和质地相同的 4 个球，其中有 2 个红色球（标号为 1 和 2），2 个绿色球（标号为 3 和 4），从袋中不放回地依次随机摸出 2 个球. 设事件  $R_1$  = “第一次摸到红球”， $R_2$  = “第二次摸到红球”， $R$  = “两次都摸到红球”， $G$  = “两次都摸到绿球”， $M$  = “两个球颜色相同”， $N$  = “两个球颜色不同”.

(1) 用集合的形式分别写出试验的样本空间以及上述各事件；

(2) 事件  $R$  与  $R_1$ ， $R$  与  $G$ ， $M$  与  $N$  之间各有什么关系？

(3) 事件  $R$  与事件  $G$  的并事件与事件  $M$  有什么关系？事件  $R_1$  与事件  $R_2$  的交事件与事件  $R$  有什么关系？

[解析] (1) 所有的试验结果如图所示，

|    |    |    |
|----|----|----|
| ①② | ①③ | ①④ |
| ②① | ②③ | ②④ |
| ③① | ③② | ③④ |
| ④① | ④② | ④③ |

用数组  $(x_1, x_2)$  表示可能的结果， $x_1$  是第一次摸到的球的标号， $x_2$  是第二次摸到的球的标号，则试验的样本空间

$$\Omega = \{(1, 2), (1, 3), (1, 4), (2, 1), (2, 3), (2, 4), (3, 1), (3, 2), (3, 4), (4, 1), (4, 2), (4, 3)\}$$

事件  $R_1$  = “第一次摸到红球”，即  $x_1 = 1$  或 2，于是  $R_1 = \{(1, 2), (1, 3), (1, 4), (2, 1), (2, 3), (2, 4)\}$ ；

事件  $R_2$  = “第二次摸到红球”，即  $x_2 = 1$  或 2，于是  $R_2 = \{(2, 1), (3, 1), (4, 1), (1, 2), (3, 2), (4, 2)\}$ 。

同理，有  $R = \{(1, 2), (2, 1)\}$ ， $G = \{(3, 4), (4, 3)\}$ ， $M = \{(1, 2), (2, 1), (3, 4), (4, 3)\}$ ，

$$N = \{(1, 3), (1, 4), (2, 3), (2, 4), (3, 1), (3, 2), (4, 1), (4, 2)\}.$$

(2) 因为  $R \subseteq R_1$ ，所以事件  $R_1$  包含事件  $R$ ；因为  $R \cap G = \emptyset$ ，所以事件  $R$  与事件  $G$  互斥；

因为  $M \cup N = \Omega$ ， $M \cap N = \emptyset$ ，所以事件  $M$  与事件  $N$  互为对立事件.

(3) 因为  $R \cup G = M$ ，所以事件  $M$  是事件  $R$  与事件  $G$  的并事件；

因为  $R_1 \cap R_2 = R$ ，所以事件  $R$  是事件  $R_1$  与事件  $R_2$  的交事件.

【练习3】判断下列各事件是否为互斥事件，是否为对立事件，并说明理由.

某小组有3名男生和2名女生，从中任选2名同学去参加演讲比赛，其中：

(1)恰有1名男生和恰有2名男生；

(2)至少有1名男生和至少有1名女生；

(3)至少有1名男生和全是男生；

(4)至少有1名男生和全是女生.

[答案](1)是互斥事件，不是对立事件. (2)不是互斥事件，也不是对立事件.

(3)不是互斥事件，也不是对立事件. (4)是互斥事件，也是对立事件. 理由见解析

[解析](1)是互斥事件，不是对立事件.

理由是：在所选的2名同学中，“恰有1名男生”实质是选出“1名男生和1名女生”，它与“恰有2名男生”不可能同时发生，所以是一对互斥事件. 但其并事件不是必然事件，所以不是对立事件.

(2)不是互斥事件，也不是对立事件.

理由是：“至少有1名男生”包括“1名男生、1名女生”和“2名都是男生”两种结果.

“至少有1名女生”包括“1名女生、1名男生”和“2名都是女生”两种结果，它们可同时发生.

(3)不是互斥事件，也不是对立事件.

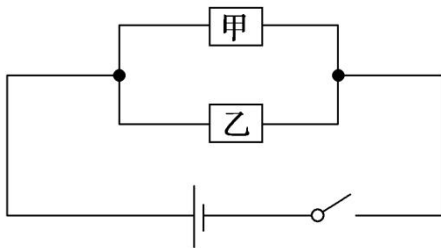
理由是：“至少有1名男生”包括“1名男生、1名女生”和“2名都是男生”，这与“全是男生”可同时发生.

(4)是互斥事件，也是对立事件.

理由是：“至少有1名男生”包括“1名男生、1名女生”和“2名都是男生”两种结果，它与“全是女生”不可能同时发生. 其并事件是必然事件，所以是对立事件.

#### 题型四：事件的运算

例4 如图，由甲、乙两个元件组成一个并联电路，每个元件可能正常或失效. 设事件 $A$  = “甲元件正常”， $B$  = “乙元件正常”.



(1)写出表示两个元件工作状态的样本空间；

(2)用集合的形式表示事件 $A$ ,  $B$  以及它们的对立事件；

(3)用集合的形式表示事件 $A \cup B$  和事件 $\overline{A} \cap \overline{B}$ ，并说明它们的含义及关系.

[答案](1)样本空间为 $\Omega = \{(0,0), (0,1), (1,0), (1,1)\}$ .

(2) $A = \{(1,0), (1,1)\}$ ,  $B = \{(0,1), (1,1)\}$ ,  $\bar{A} = \{(0,0), (0,1)\}$ ,  $\bar{B} = \{(0,0), (1,0)\}$ .

(3) $A \cup B = \{(0,1), (1,0), (1,1)\}$ ,  $\bar{A} \cap \bar{B} = \{(0,0)\}$ ;  $A \cup B$  表示电路工作正常,  $\bar{A} \cap \bar{B}$  表示电路工作不正常;  
 $A \cup B$  和  $\bar{A} \cap \bar{B}$  互为对立事件.

[解析](1)用  $x_1, x_2$  分别表示甲、乙两个元件的状态, 则可以用  $(x_1, x_2)$  表示这个并联电路的状态.

以 1 表示元件正常, 0 表示元件失效, 则样本空间为 $\Omega = \{(0,0), (0,1), (1,0), (1,1)\}$ .

(2)根据题意, 可得  $A = \{(1,0), (1,1)\}$ ,  $B = \{(0,1), (1,1)\}$ ,  $\bar{A} = \{(0,0), (0,1)\}$ ,  $\bar{B} = \{(0,0), (1,0)\}$ .

(3) $A \cup B = \{(0,1), (1,0), (1,1)\}$ ,  $\bar{A} \cap \bar{B} = \{(0,0)\}$ ;  $A \cup B$  表示电路工作正常,  $\bar{A} \cap \bar{B}$  表示电路工作不正常;  
 $A \cup B$  和  $\bar{A} \cap \bar{B}$  互为对立事件.

【练习 4】盒子里有 6 个红球, 4 个白球, 现从中任取 3 个球, 设事件  $A = \{3 \text{ 个球中有 1 个红球 2 个白球}\}$ , 事件  $B = \{3 \text{ 个球中有 2 个红球 1 个白球}\}$ , 事件  $C = \{3 \text{ 个球中至少有 1 个红球}\}$ , 事件  $D = \{3 \text{ 个球中既有红球又有白球}\}$ .

求: (1)事件  $D$  与  $A, B$  是什么样的运算关系?

(2)事件  $C$  与  $A$  的交事件是什么事件?

[答案](1)  $D = A \cup B$ . (2)  $C \cap A = A$ .

[解析](1)对于事件  $D$ , 可能的结果为 1 个红球, 2 个白球或 2 个红球, 1 个白球, 故  $D = A \cup B$ .

(2)对于事件  $C$ , 可能的结果为 1 个红球, 2 个白球或 2 个红球, 1 个白球或 3 个均为红球,

所以  $A \subseteq C$ , 故  $C \cap A = A$ .

【课后巩固】—————完成课时作业 (175)

## 18.3 古典概型

### 【学习目标】

1. 理解古典概型的特征和计算公式，会判断古典概型；
2. 会求古典概型中事件的概率.

### 【要点整合】

要点一：概率

对随机事件发生可能性大小的度量(数值)称为事件的概率(probability)，事件  $A$  的概率用  $P(A)$  表示.

要点二：古典概型

1、古典概型

考察试验的共同特征：

- ①有限性：样本空间的样本点只有有限个；
- ②等可能性：每个样本点发生的可能性相等.

我们将具有以上两个特征的试验称为古典概型试验，其数学模型称为古典概率模型(classical models of probability)，简称古典概型.

2、概率公式：一般地，设试验  $E$  是古典概型，样本空间  $\Omega$  包含  $n$  个样本点，事件  $A$  包含其中的  $k$  个样本点，

则定义事件  $A$  的概率  $P(A) = \frac{k}{n} = \frac{n(A)}{n(\Omega)}$ . 其中， $n(A)$  和  $n(\Omega)$  分别表示事件  $A$  和样本空间  $\Omega$  包含的样本点个数.

### 【典例讲练】

#### 题型一：简单古典概型的计算

例 1 抛掷两枚质地均匀的骰子（标记为  $I$  号和  $II$  号），观察两枚骰子分别可能出现的基本结果，

- (1) 写出这个试验的样本空间，并判断这个试验是否为古典概型；
- (2) 求下列事件的概率： $A$  = “两个点数之和是 5”； $B$  = “两个点数相等”； $C$  = “ $I$  号骰子的点数大于  $II$  号骰子的点数”.

[答案] (1)  $\Omega = \{(m, n) | m, n \in \{1, 2, 3, 4, 5, 6\}\}$ ，是古典概型 (2)  $\frac{1}{9}; \frac{1}{6}; \frac{5}{12}$

[解析] (1) 抛掷一枚骰子有 6 种等可能的结果， $I$  号骰子的每一个结果都可与  $II$  号骰子的任意一个结果配对，组成掷两枚骰子试验的一个结果. 用数字  $m$  表示  $I$  号骰子出现的点数是  $m$ ，数字  $n$  表示  $II$  号骰子出现的点数是  $n$ ，则数组  $(m, n)$  表示这个试验的一个样本点. 因此该试验的样本空间

$\Omega = \{(m, n) | m, n \in \{1, 2, 3, 4, 5, 6\}\}$ ，其中共有 36 个样本点.

由于骰子的质地均匀，所以各个样本点出现的可能性相等，因此这个试验是古典概型.

(2) 因为  $A = \{(1, 4), (2, 3), (3, 2), (4, 1)\}$ ，所以  $n(A) = 4$ ，从而  $P(A) = \frac{n(A)}{n(\Omega)} = \frac{4}{36} = \frac{1}{9}$ ；

因为  $B = \{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\}$ , 所以  $n(B) = 6$ , 从而  $P(B) = \frac{n(B)}{n(\Omega)} = \frac{6}{36} = \frac{1}{6}$ ;

因为  $C = \{(2,1), (3,1), (3,2), (4,1), (4,2), (4,3), (5,1), (5,2), (5,3), (5,4), (6,1), (6,2), (6,3), (6,4), (6,5)\}$ ,

所以  $n(C) = 15$ , 从而  $P(C) = \frac{n(C)}{n(\Omega)} = \frac{15}{36} = \frac{5}{12}$

解题技巧 (求古典概型的一般步骤)

(1)明确实验的条件及要观察的结果, 用适当的符号(字母/数字/数组等)表示实验的可能结果(可借助图表);

(2)根据实际问题情景判断样本点的等可能性;

(3)计算样本点总个数及事件包含的样本点个数, 求出事件 A 的概率.

【练习 1】某校夏令营有 3 名男同学 A, B, C 和 3 名女同学 X, Y, Z, 其年级情况如下表:

|     | 一年级 | 二年级 | 三年级 |
|-----|-----|-----|-----|
| 男同学 | A   | B   | C   |
| 女同学 | X   | Y   | Z   |

现从这 6 名同学中随机选出 2 人参加知识竞赛(每人被选到的可能性相同).

(1)用表中字母列举出所有可能的结果;

(2)设 M 为事件“选出的 2 人来自不同年级且恰有 1 名男同学和 1 名女同学”, 求事件 M 发生的概率.

[解析](1)从 6 名同学中随机选出 2 人参加知识竞赛的样本空间为  $\{A, B\}, \{A, C\}, \{A, X\}, \{A, Y\}, \{A, Z\}, \{B, C\}, \{B, X\}, \{B, Y\}, \{B, Z\}, \{C, X\}, \{C, Y\}, \{C, Z\}, \{X, Y\}, \{X, Z\}, \{Y, Z\}$ , 共 15 种.

(2)选出的 2 人来自不同年级且恰有 1 名男同学和 1 名女同学的样本空间为  $\{A, Y\}, \{A, Z\}, \{B, X\},$

$\{B, Z\}, \{C, X\}, \{C, Y\}$ , 共 6 种. 因此, 事件 M 发生的概率  $P(M) = \frac{6}{15} = \frac{2}{5}$ .

## 题型二: 较复杂的古典概型的计算

例 2 (1) 从集合  $A = \{-2, -1, 2\}$  中随机选取一个数记为 a, 从集合  $B = \{-1, 1, 3\}$  中随机选取一个数记为 b, 则直线  $ax - y + b = 0$  不经过第四象限的概率为( )

A.  $\frac{2}{9}$

B.  $\frac{1}{3}$

C.  $\frac{4}{9}$

D.  $\frac{1}{4}$

[答案]A

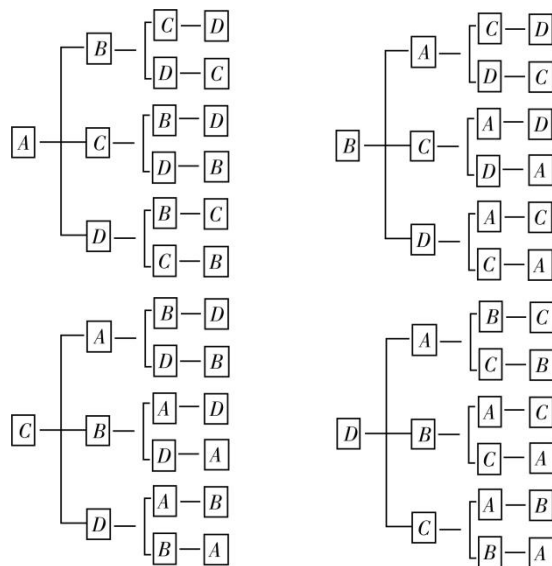
[解析]分别从集合 A, B 所取的数 a, b 表示为(a, b)的形式, 一共有 9 个结果:

$(-2, -1), (-2, 1), (-2, 3), (-1, -1), (-1, 1), (-1, 3), (2, -1), (2, 1), (2, 3)$

(2) 有  $A, B, C, D$  四位贵宾, 应分别坐在  $a, b, c, d$  四个席位上, 现在这四人均未留意, 在四个席位上随便就坐时,

- ①求这四人恰好都坐在自己席位上的概率;
- ②求这四人恰好都没坐在自己席位上的概率;
- ③求这四人恰好有 1 位坐在自己席位上的概率.

[解析]将  $A, B, C, D$  四位贵宾就座情况用下面图形表示出来:



如上图所示, 本题中的等可能基本事件共有 24 个.

(1) 设事件  $A$  为“这四人恰好都坐在自己的席位上”, 则事件  $A$  只包含 1 个基本事件, 所以  $P(A) = \frac{1}{24}$ .

(2) 设事件  $B$  为“这四人恰好都没坐在自己席位上”, 则事件  $B$  包含 9 个基本事件, 所以  $P(B) = \frac{9}{24} = \frac{3}{8}$ .

(3) 设事件  $C$  为“这四人恰好有 1 位坐在自己席位上”, 则事件  $C$  包含 8 个基本事件, 所以  $P(C) = \frac{8}{24} = \frac{1}{3}$ .

(3) 某校有  $A, B$  两个学生食堂, 若  $a, b, c$  三名学生各自随机选择其中的一个食堂用餐, 则三人不在同一个食堂用餐的概率为\_\_\_\_\_.

[答案]  $\frac{3}{4}$

[解析]  $a, b, c$  三名学生选择食堂的结果有:  $(A, A, A), (A, A, B), (A, B, A), (B, A, A), (A, B, B), (B, A, B), (B, B, A), (B, B, B)$  共 8 个, 三人在同一食堂用餐的结果有  $(A, A, A), (B, B, B)$ , 共 2 个, 所以“三人在同一食堂用餐”的概率为  $\frac{2}{8} = \frac{1}{4}$ , 而“三人不在同一食堂用餐”与“三人在同一食堂用餐”是对

立事件, 所以“三人不在同一食堂用餐”的概率为  $1 - \frac{1}{4} = \frac{3}{4}$ .

规律方法 解决有序和无序问题应注意两点

(1) 关于不放回抽样, 计算基本事件个数时, 既可以看作是有顺序的, 也可以看作是无顺序的, 其最后结果是一致的. 但不论选择哪一种方式, 观察的角度必须一致, 否则会产生错误.

(2) 关于有放回抽样, 应注意在连续取出两次的过程中, 因为先后顺序不同, 所以  $(a_1, b), (b, a_1)$  不是

同一个基本事件，解题的关键是要清楚无论是“不放回抽取”还是“有放回抽取”，每一件产品被取出的机会都是均等的.

【练习 2】(1) 我国数学家陈景润在哥德巴赫猜想的研究中取得了世界领先的成果. 哥德巴赫猜想是“每个大于 2 的偶数可以表示为两个素数的和”，如  $30 = 7 + 23$ . 在不超过 30 的素数中，随机选取两个不同的数，其和等于 30 的概率是 ( )

- A.  $\frac{1}{12}$                       B.  $\frac{1}{14}$                       C.  $\frac{1}{15}$                       D.  $\frac{1}{18}$

[答案]C

[解析]不超过 30 的素数有 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 共 10 个，从中随机选取两个不同的

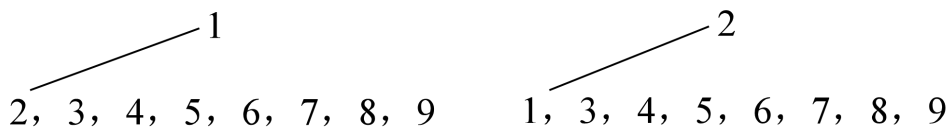
数有  $C_{10}^2$  种不同的取法，这 10 个数中两个不同的数的和等于 30 的有 3 对，所以所求概率  $P = \frac{3}{C_{10}^2} = \frac{1}{15}$ .

(2) 从分别标有 1, 2, ..., 9 的 9 张卡片中不放回地随机抽取 2 次，每次抽取 1 张. 则抽到的 2 张卡片上的数奇偶性不同的概率是 ( )

- A.  $\frac{5}{18}$                       B.  $\frac{4}{9}$                       C.  $\frac{5}{9}$                       D.  $\frac{7}{9}$

[答案]C

[解析]不放回的抽取 2 次有  $C_9^1 C_8^1 = 9 \times 8 = 72$ ，如图



可知 (1,2) 与 (2,1) 是不同，所以抽到的 2 张卡片上的数奇偶性不同有  $2C_5^1 C_4^1 = 40$ ，所求概率为  $\frac{40}{72} = \frac{5}{9}$ .

(3) 袋中共有 15 个除了颜色外完全相同的球，其中有 10 个白球，5 个红球. 从袋中任取 2 个球，所取的 2 个球中恰有 1 个白球，1 个红球的概率为 ( )

- A.  $\frac{5}{21}$                       B.  $\frac{10}{21}$                       C.  $\frac{11}{21}$                       D. 1

[答案]B

[解析]基本事件总数为  $C_{15}^2$ ，恰有 1 个白球与 1 个红球的基本事件为  $C_{10}^1 C_5^1$ ，所求概率为  $\frac{C_{10}^1 C_5^1}{C_{15}^2} = \frac{10}{21}$ .

【课后巩固】—————完成课时作业 (176)

## 18.4 概率的基本性质

### 【学习目标】

1. 理解并掌握概率的基本性质；
2. 能够运用概率的基本性质求一些简单事件的概率.

### 【要点整合】

概率的基本性质

一般地, 概率有如下性质:

性质 1: 对任意的事件  $A$ , 都有  $P(A) \geq 0$ .

性质 2: 必然事件的概率为 1, 不可能事件的概率为 0, 即  $P(\Omega) = 1$ ,  $P(\emptyset) = 0$ .

性质 3: 如果事件  $A$  与事件  $B$  互斥, 那么  $P(A \cup B) = P(A) + P(B)$ .

性质 4: 如果事件  $A$  与事件  $B$  互为对立事件, 那么  $P(B) = 1 - P(A)$ ,  $P(A) = 1 - P(B)$ .

性质 5: 如果  $A \subseteq B$ , 那么  $P(A) \leq P(B)$ .

性质 6: 设  $A, B$  是一个随机试验中的两个事件, 那么  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ .

### 【典例讲练】

#### 题型一: 概率的基本性质

例 1 从不包含大小王牌的 52 张扑克牌中随机抽取一张, 设事件  $A$  = “抽到红心”, 事件  $B$  = “抽到方片”,

$P(A) = P(B) = \frac{1}{4}$ , 那么

(1)  $C$  = “抽到红花色”, 求  $P(C)$ ;

(2)  $D$  = “抽到黑花色”, 求  $P(D)$ .

[答案] (1)  $\frac{1}{2}$  (2)  $\frac{1}{2}$

[解析] (1) 因为  $C = A \cup B$ , 且  $A$  与  $B$  不会同时发生, 所以  $A$  与  $B$  是互斥事件,

根据互斥事件的概率加法公式, 得  $P(C) = P(A) + P(B) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

(2) 因为  $C$  与  $D$  互斥, 又因为  $C \cup D$  是必然事件, 所以  $C$  与  $D$  互为对立事件,

因此  $P(D) = 1 - P(C) = 1 - \frac{1}{2} = \frac{1}{2}$

解题技巧 (概率性质公式)

(1) 运用概率加法公式解题的步骤:

① 确定诸事件彼此互斥; ② 先求诸事件分别发生的概率, 再求其和.

(2) 求复杂事件的概率通常有两种方法:

一是将所求事件转化成彼此互斥的事件的并; 二是先求对立事件的概率, 进而再求所求事件的概率.

【练习 1】袋中有 12 个小球，分别为红球、黑球、黄球、绿球，从中任取一球，已知得到红球的概率是  $\frac{1}{3}$ ，得到黑球或黄球的概率是  $\frac{5}{12}$ ，得到黄球或绿球的概率也是  $\frac{5}{12}$ ，试求得到黑球、黄球、绿球的概率分别是多少？

[答案]  $\frac{1}{4}, \frac{1}{6}, \frac{1}{4}$ .

[解析] 设得到黑球、黄球的概率分别为  $x, y$ ，由题意得 
$$\begin{cases} x+y=\frac{5}{12} \\ y+(1-\frac{1}{3}-x-y)=\frac{5}{12} \end{cases}$$

解得  $x=\frac{1}{4}, y=\frac{1}{6}$ ，所以得到绿球的概率为  $1-\frac{1}{3}-\frac{1}{4}-\frac{1}{6}=\frac{1}{4}$ ，所以得到黑球、黄球、绿球的概率分别是  $\frac{1}{4}, \frac{1}{6}, \frac{1}{4}$ .

## 题型二：概率的基本性质的应用

例 2 为了推广一种新饮料，某饮料生产企业开展了有奖促销活动：将 6 罐这种饮料装一箱，每箱中都放置 2 罐能够中奖的饮料. 若从一箱中随机抽出 2 罐，能中奖的概率为多少？



[答案]  $\frac{3}{5}$

[解析] 设事件  $A$  = “中奖”，事件  $A_1$  = “第一罐中奖”，事件  $A_2$  = “第二罐中奖”，

那么事件  $A_1 A_2$  = “两罐都中奖”， $A_1 \overline{A_2}$  = “第一罐中奖，第二罐不中奖”，

$\overline{A_1} A_2$  = “第一罐不中奖，第二罐中奖”，且  $A = A_1 A_2 \cup A_1 \overline{A_2} \cup \overline{A_1} A_2$ ，

因为  $A_1 A_2, A_1 \overline{A_2}, \overline{A_1} A_2$  两两互斥，所以根据互斥事件的概率加法公式，

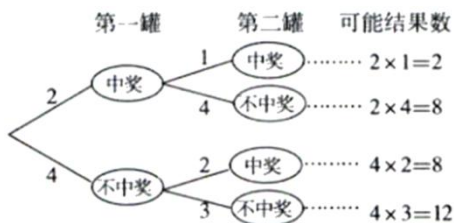
可得  $P(A) = P(A_1 A_2) + P(A_1 \overline{A_2}) + P(\overline{A_1} A_2)$ ，

我们借助树状图来求相应事件的样本点数，

可以得到，样本空间包含的样本点个数为  $n(\Omega) = 6 \times 5 = 30$ ，

且每个样本点都是等可能的，因为  $n(A_1 A_2) = 2, n(A_1 \overline{A_2}) = 8, n(\overline{A_1} A_2) = 8$ ，

所以  $P(A) = \frac{2}{30} + \frac{8}{30} + \frac{8}{30} = \frac{18}{30} = \frac{3}{5}$ ，故中奖的概率的为  $\frac{3}{5}$ 。



解题技巧 (概率性质的应用)

1. 对于一个较复杂的事件, 一般将其分解为几个简单的事件. 当这些事件彼此互斥时, 即可用概率加法公式.

2. 运用事件的概率加法公式解题的步骤: (1)确定题中哪些事件彼此互斥; (2)将待求事件拆分为几个互斥事件之和; (3)先求各互斥事件分别发生的概率, 再求和.

【练习 2】(1) 经统计, 在某储蓄所一个营业窗口等候的人数及相应的概率如下:

|      |     |      |     |     |     |            |
|------|-----|------|-----|-----|-----|------------|
| 排队人数 | 0   | 1    | 2   | 3   | 4   | 5 人及 5 人以上 |
| 概率   | 0.1 | 0.16 | 0.3 | 0.3 | 0.1 | 0.04       |

求: (1)至多 2 人排队等候的概率是多少?

(2)至少 3 人排队等候的概率是多少?

[答案](1) 0.56. (2)0.44.

[解析]记“无人排队等候”为事件  $A$ , “1 人排队等候”为事件  $B$ , “2 人排队等候”为事件  $C$ , “3 人排队等候”为事件  $D$ , “4 人排队等候”为事件  $E$ , “5 人及 5 人以上排队等候”为事件  $F$ , 则事件  $A$ 、 $B$ 、 $C$ 、 $D$ 、 $E$ 、 $F$  互斥.

(1)记“至多 2 人排队等候”为事件  $G$ , 则  $G=A \cup B \cup C$ ,

所以  $P(G)=P(A \cup B \cup C)=P(A)+P(B)+P(C)=0.1+0.16+0.3=0.56$ .

(2)解法一: 记“至少 3 人排队等候”为事件  $H$ , 则  $H=D \cup E \cup F$ ,

所以  $P(H)=P(D \cup E \cup F)=P(D)+P(E)+P(F)=0.3+0.1+0.04=0.44$ .

解法二: 记“至少 3 人排队等候”为事件  $H$ , 则其对立事件为事件  $G$ , 所以  $P(H)=1-P(G)=0.44$ .

(2) 甲、乙两人下棋, 和棋的概率是  $\frac{1}{2}$ , 乙获胜的概率为  $\frac{1}{3}$

求: (1)甲获胜的概率; (2)甲不输的概率.

[答案](1)  $\frac{1}{6}$ ; (2)  $\frac{2}{3}$

[解析](1)“甲获胜”可看成是“和棋或乙获胜”的对立事件, 所以“甲获胜”的概率为  $1-\frac{1}{2}-\frac{1}{3}=\frac{1}{6}$ .

(2)方法一 “甲不输”可看成是“甲获胜”“和棋”这两个互斥事件的并事件, 所以  $P(\text{甲不输})=\frac{1}{6}+\frac{1}{2}=\frac{2}{3}$ .

方法二 “甲不输”可看成是“乙获胜”的对立事件, 所以  $P(\text{甲不输})=1-\frac{1}{3}=\frac{2}{3}$ , 故甲不输的概率为  $\frac{2}{3}$ .

【课后巩固】——完成课时作业 (177)

## 18.5 事件的相互独立性

### 【学习目标】

1. 在具体情境中，理解两个事件相互独立的概念；
2. 能利用相互独立事件同时发生的概率公式解决一些简单的实际问题.

### 【要点整合】

要点一：相互独立事件的概念

设  $A$ 、 $B$  为两个事件，若  $P(AB)=P(A)P(B)$ ，则称事件  $A$  与事件  $B$  相互独立.

思考：(1) 不可能事件与任何一个事件相互独立吗？

(2) 必然事件与任何一个事件相互独立吗？

[答案] (1) 相互独立，不可能事件的发生与任何一个事件的发生没有影响；

(2) 相互独立，必然事件的发生与任何一个事件的发生没有影响.

要点二：相互独立事件的性质

如果事件  $A$  与  $B$  相互独立，那么  $A$  与  $\bar{B}$ ， $\bar{A}$  与  $B$ ， $\bar{A}$  与  $\bar{B}$  也都相互独立.

### 【典型讲练】

#### 题型一：事件独立性的判断

例 1 判断下列各对事件是不是相互独立事件.

(1) 甲组 3 名男生、2 名女生，乙组 2 名男生、3 名女生，现从甲、乙两组中各选 1 名同学参加演讲比赛，“从甲组中选出 1 名男生”与“从乙组中选出 1 名女生”；

(2) 容器内盛有 5 个白乒乓球和 3 个黄乒乓球，“从 8 个球中任意取出 1 个，取出的是白球”与“从剩下的 7 个球中任意取出 1 个，取出的还是白球”.

[解析] (1) “从甲组中选出 1 名男生”这一事件是否发生，对“从乙组中选出 1 名女生”这一事件发生的概率没有影响，所以它们是相互独立事件；

(2) “从 8 个球中任意取出 1 个，取出的是白球”的概率为  $\frac{5}{8}$ ，若这一事件发生了，则“从剩下的 7 个球中任意取出 1 个，取出的仍是白球”的概率为  $\frac{4}{7}$ ，若前一事件没有发生，则后一事件发生的概率为  $\frac{5}{7}$ .

可见，前一事件是否发生，对后一事件发生的概率有影响，所以两者不是相互独立事件.

【练习 1】(1) 甲、乙两名射手同时向一目标射击，设事件  $A$ ：“甲击中目标”，事件  $B$ ：“乙击中目标”，则事件  $A$  与事件  $B$ ( )

A. 相互独立但不互斥

B. 互斥但不相互独立

C. 相互独立且互斥

D. 既不相互独立也不互斥



【练习 2】甲、乙两人破译一密码，他们能破译的概率分别为 $\frac{1}{3}$ 和 $\frac{1}{4}$ ，求两人破译时，以下事件发生的概率：

- (1) 两人都能破译的概率；
- (2) 恰有一人能破译的概率；
- (3) 至多有一人能破译的概率.

[答案](1) $\frac{1}{12}$ ; (2) $\frac{5}{12}$ ; (3) $\frac{11}{12}$

[解析]记事件 A 为“甲独立地破译出密码”，事件 B 为“乙独立地破译出密码”，

(1) 两个人都破译出密码的概率为： $P(AB)=P(A)P(B)=\frac{1}{3}\times\frac{1}{4}=\frac{1}{12}$ ;

(2) 恰有一人破译出密码分为两类：甲破译出乙破译不出，乙破译出甲破译不出，即  $\overline{A}\overline{B} + \overline{A}B$ ，  
 $\therefore P(\overline{A}\overline{B} + \overline{A}B)=P(\overline{A}\overline{B})+P(\overline{A}B)=P(\overline{A})P(\overline{B})+P(\overline{A})P(B)=\frac{1}{3}\times(1-\frac{1}{4})+(1-\frac{1}{3})\times\frac{1}{4}=\frac{5}{12}$ .

(3) 至多有一人破译出密码的对立事件是两人都破译出密码， $\therefore$  其概率为  $1-P(AB)=1-\frac{1}{12}=\frac{11}{12}$ .

### 题型三：相互独立事件概率的综合应用

例 3 计算机考试分理论考试与实际操作两部分进行，每部分考试成绩只记“合格”与“不合格”，两部分考试都“合格”者，则计算机考试“合格”，并颁发合格证书. 甲、乙、丙三人在理论考试中“合格”的概率依次为 $\frac{4}{5}$ 、 $\frac{3}{4}$ 、 $\frac{2}{3}$ ，在实际操作考试中“合格”的概率依次为 $\frac{1}{2}$ 、 $\frac{2}{3}$ 、 $\frac{5}{6}$ ，所有考试是否合格相互之间没有影响.

- (1) 假设甲、乙、丙三人同时进行理论与实际操作两项考试，谁获得合格证书的可能性大？
- (2) 这三人进行理论与实际操作两项考试后，求恰有两人获得合格证书的概率.

[解析](1) 记“甲获得合格证书”为事件 A，“乙获得合格证书”为事件 B，“丙获得合格证书”为事件 C，

则  $P(A)=\frac{4}{5}\times\frac{1}{2}=\frac{2}{5}$ ,  $P(B)=\frac{3}{4}\times\frac{2}{3}=\frac{1}{2}$ ,  $P(C)=\frac{2}{3}\times\frac{5}{6}=\frac{5}{9}$ . 因为  $P(C)>P(B)>P(A)$ ，所以丙获得合格证书的可能性大.

(2) 设“三人考试后恰有两人获得合格证书”为事件 D，

则  $P(D)=P(\overline{A}B\overline{C})+P(A\overline{B}C)+P(\overline{A}BC)=\frac{2}{5}\times\frac{1}{2}\times\frac{4}{9}+\frac{2}{5}\times\frac{1}{2}\times\frac{5}{9}+\frac{3}{4}\times\frac{1}{2}\times\frac{5}{9}=\frac{11}{30}$ .

【练习 3】甲、乙两名篮球运动员互不影响地在同一位置投球，命中率分别为 $\frac{1}{2}$ 与 p，且乙投球 2 次均未命中的概率为 $\frac{1}{16}$ .

- (1) 求乙投球的命中率 p；
- (2) 求甲投球 2 次，至少命中 1 次的概率.

[答案] (1) $\frac{3}{4}$ ; (2) $\frac{3}{4}$

[解析] (1) 设“甲投一次球命中”为事件 A, “乙投一次球命中”为事件 B. 由题意得  $P(\bar{B}) = \frac{1}{16}$ ,

解得  $P(\bar{B}) = \frac{1}{4}$  或  $P(\bar{B}) = -\frac{1}{4}$  (舍去), 故  $p = 1 - P(\bar{B}) = \frac{3}{4}$ , 所以乙投球的命中率为  $\frac{3}{4}$ .

(2) 方法一: 由题设知,  $P(A) = \frac{1}{2}$ ,  $P(\bar{A}) = \frac{1}{2}$ , 故甲投球 2 次,

至少命中 1 次的概率为  $1 - P(\bar{A} \cdot \bar{A}) = 1 - P(\bar{A})P(\bar{A}) = \frac{3}{4}$ .

方法二: 由题设知  $P(A) = \frac{1}{2}$ ,  $P(\bar{A}) = \frac{1}{2}$ , 故甲投球 2 次, 至少命中 1 次的概率为  $2P(A)P(\bar{A}) + P(A)P(A) = \frac{3}{4}$ .

例 4 某大学开设甲、乙、丙三门选修课, 学生选修哪门课互不影响. 已知学生小张只选甲的概率为 0.08, 只选甲和乙的概率为 0.12, 至少选一门课的概率为 0.88, 用  $\xi$  表示小张选修的课程门数和没有选修的课程门数的乘积.

(1) 求学生小张选修甲的概率;

(2) 记“函数  $f(x) = x^2 + \xi x$  为  $\mathbf{R}$  上的偶函数”为事件 A, 求事件 A 的概率.

[解析] (1) 设学生小张选修甲、乙、丙的概率分别为  $x, y, z$ ,

$$\text{则} \begin{cases} x(1-y)(1-z) = 0.08, \\ xy(1-z) = 0.12, \\ (1-x)(1-y)(1-z) = 0.12, \end{cases} \quad \text{解得} \begin{cases} x = 0.4 \\ y = 0.6 \\ z = 0.5 \end{cases}; \text{所以学生小张选修甲的概率为 } 0.4.$$

(2) 若函数  $f(x) = x^2 + \xi x$  为  $\mathbf{R}$  上的偶函数, 则  $\xi = 0$ . 当  $\xi = 0$  时, 表示小张选修三门课或三门课都不选, 所以  $P(A) = P(\xi = 0) = xyz + (1-x)(1-y)(1-z) = 0.4 \times 0.6 \times 0.5 + (1-0.4)(1-0.6)(1-0.5) = 0.24$ .

【练习 4】甲、乙两名学生通过某种听力测试的概率分别为  $\frac{1}{2}$  和  $\frac{1}{3}$ , 两人同时参加测试, 求其中有且只有一人能通过的概率.

[解析] 设事件 A 表示“甲通过听力测试”, 事件 B 表示“乙通过听力测试”. 根据题意, 知事件 A 和 B 相互独立, 且  $P(A) = \frac{1}{2}$ ,  $P(B) = \frac{1}{3}$ .

记“有且只有一人通过听力测试”为事件 C, 则  $C = A\bar{B} \cup \bar{A}B$ , 且  $A\bar{B}$  和  $\bar{A}B$  互斥.

故  $P(C) = P(A\bar{B} \cup \bar{A}B) = P(A\bar{B}) + P(\bar{A}B) = P(A)P(\bar{B}) + P(\bar{A})P(B) = \frac{1}{2} \times (1 - \frac{1}{3}) + (1 - \frac{1}{2}) \times \frac{1}{3} = \frac{1}{2}$ .

【课后巩固】——完成课时作业 (178)

## 18.6 频率的稳定性

## 18.7 随机模拟

### 【学习目标】

1. 了解频率与概率的关系；
2. 结合实例，会用频率估计概率；
3. 理解随机模拟试验出现地意义；
4. 利用随机模拟试验求概率.

### 【要点整合】

要点一：概率的稳定性

一般地，随着试验次数  $n$  的增大，频率偏离概率的幅度会缩小，即事件  $A$  发生的频率  $f_n(A)$  会逐渐稳定于事件  $A$  发生的概率  $P(A)$ ，我们称这个性质为频率的稳定性.

要点二：概率与频率的区别与联系

区别：(1) 频率具有随机性，它反映随机事件出现的频繁程度；

(2) 概率是一个客观常数，它反映了随机事件发生可能性的大小.

联系：(1) 概率是频率的稳定值，频率是概率的估计值.

要点三：蒙特卡洛方法

(1) 我们知道，利用计算器或计算软件可以产生随机数.实际上，我们也可以根据不同的随机试验构建相应的随机数模拟实验，这样就可以快速地进行大量重复试验了，这么随机模拟方式叫做随机模拟.我们称利用随机模拟解决问题地方法为蒙特卡洛 (Monte Carlo) 方法.

### 【典型讲练】

#### 题型一：概率的稳定性

例 1 新生儿性别比是每 100 名女婴对应的男婴数，通过抽样调查得知，我国 2014 年、2015 年出生的婴儿性别比分别为 115.88 和 113.51.

(1) 分别估计我国 2014 年和 2015 年男婴的出生率 (新生儿中男婴的比率，精确到 0.001)；

(2) 根据估计结果，你认为“生男孩和女孩是等可能的”这个判断可靠吗？

[解析] (1) 根据“性别比”的定义和抽样调查结果，可以计算男婴出生的频率；由频率的稳定性，可以估计

男婴出生率，所以 2014 年男婴出生的频率为  $\frac{115.88}{100+115.88} \approx 0.537$ .

(2) 由于调查新生儿人数的样本非常大，根据频率的稳定性，上述对男婴出生率的估计具有较高的可信度.因此，我们有理由怀疑“生男孩和生女孩是等可能的”结论.

【练习 1】(多选) 给出下列四个命题, 其中正确的命题有 ( )

- A. 做 100 次抛硬币的试验, 结果 51 次出现正面朝上, 因此, 出现正面朝上的概率是  $\frac{51}{100}$ ;
- B. 随机事件发生的频率就是这个随机事件发生的概率;
- C. 抛掷骰子 100 次, 得点数是 1 的结果有 18 次, 则出现 1 点的频率是  $\frac{9}{50}$ ;
- D. 随机事件发生的频率不一定是这个随机事件发生的概率.

[答案]CD

[解析]对于 A, 混淆了频率与概率的区别, 故 A、B 错误; 对于 C, 抛掷骰子 100 次, 得点数是 1 的结果有 18 次, 则出现 1 点的频率是  $\frac{18}{100} = \frac{9}{50}$ , 符合频率的定义, 故 C 正确; 对于 D, 频率是概率的估计值, 故 D 正确.

## 题型二：概率的应用

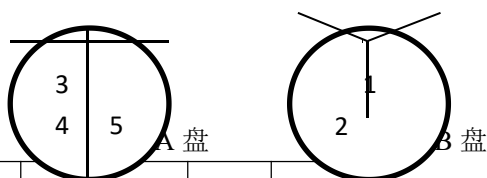
例 2 在游戏过程中甲发现: 玩了 10 次时, 双方各胜 5 次; 但玩到 1000 次时, 自己才 300 次, 而乙却胜了 700 次, 据此, 甲认为游戏不公平, 但乙认为游戏是公平的, 你更支持谁的结论? 为什么?

[解析]当游戏玩了 10 次时, 甲、乙获胜的频率都为 0.5; 当游戏玩了 1000 次时, 甲获胜的频率为 0.3, 乙获胜的频率为 0.7. 根据频率的稳定性, 随着试验次数的增加, 频率偏离概率很大的可能性会越来越小. 相对 10 次游戏, 1000 次游戏时的频率接近概率的可能性更大, 因此我们更愿意相信 1000 次时的频率离概率更近, 而游戏玩到 1000 次时, 甲、乙获胜的频率分别是 0.3 和 0.7, 存在很大差距, 所以有理由认为游戏是不公平的. 因此, 应该支持甲对游戏公平性的判断.

【练习 2】如图所示, 有两个可以自由转动的均匀转盘 A、B, 转盘 A 被平均分成 3 等份, 分别标上 1、2、3 三个数字; 转盘 B 被平均分成 4 等份, 分别标上 3、4、5、6 四个数字. 现为甲、乙两人设计游戏规则: 自由转动转盘 A 和转盘 B, 转盘停止后, 指针指上一个数字, 将指针所指的两个数字相加, 如果和是 6, 那么甲获胜, 否则乙获胜, 你认为这个规则公平吗?

[答案]不公平

[解析]列表如下:



|          |   |   |   |   |   |
|----------|---|---|---|---|---|
| <i>B</i> |   | 3 | 4 | 5 | 6 |
| <i>A</i> |   | 3 | 4 | 5 | 6 |
|          | 1 | 4 | 5 | 6 | 7 |

|   |   |   |   |   |
|---|---|---|---|---|
| 2 | 5 | 6 | 7 | 8 |
| 3 | 6 | 7 | 8 | 9 |

由表可知，可能的结果有 12 种，和为 6 的结果只有 3 种。

因此甲获胜的概率为  $\frac{3}{12} = \frac{1}{4}$ ，乙获胜的概率为  $\frac{9}{12} = \frac{3}{4}$ ，甲、乙获胜的概率不相等，所以这个游戏规则不公平。

### 题型三：利用随机模拟实验求概率

例 3 从你所在班级任意选出 6 名同学，调查他们的出生月份，假设出生在一月、二月……十二月是等可能的。设事件 A = “至少有两人出生月份相同”，设计一种试验方法，模拟 20 次，估计事件 A 发生的概率。

[解析]根据假设，每个人的出生月份在 12 个月中是等可能的，而且相互之间没有影响，所以观察 6 个人的出生月份可以看成可重复试验。因此，可以构建如下有放回摸球试验进行模拟：在袋子中装入编号为 1、2、…、12 的 12 个球，这些球除编号外没有什么差别。有放回地随机从袋中摸 6 次球，得到 6 个数代表 6 个人的出生月份，这就完成了一次模拟试验。如果这 6 个数中至少有 2 个相同，表示事件 A 发生了。重复以上模拟试验 20 次，就可以统计出事件 A 发生的频率。

例 4 在一次奥运会男子羽毛球单打比赛中，运动员甲和乙进入了决赛。假设每局比赛甲获胜的概率为 0.6，乙获胜的概率为 0.4。利用计算机模拟试验，估计甲获得冠军的概率。

[答案]0.65

[解析]设事件 A = “甲获得冠军”，事件 B = “单局比赛甲胜”，则  $P(B) = 0.6$ 。用计算器或计算机产生 1~5 之间的随机数，当出现随机数 1、2 或 3 时，表示一局比赛甲获胜，其概率为 0.6。由于要比赛 3 局，所以每 3 个随机数为一组。例如，产生 20 组随机数：

423 123 423 344 114 453 525 332 152 342

534 443 512 541 125 432 334 151 314 354

相当于做了 20 次重复试验。其中事件 A 发生了 13 次，对应的数组分别是 423、123、423、114、332、152、342、512、125、432、334、151、314，用频率估计事件 A 的概率的近似值为  $\frac{13}{20} = 0.65$ 。

[解题技巧]（利用随机模拟实验求概率）

用随机模拟来估计概率，一般有如下特点的事件可以用这种方法来估计：(1)对于满足“有限性”但不满足“等可能性”的概率问题，我们可采取随机模拟方法来估计概率。(2)对于一些基本事件的总数比较大而导致很难把它列举得不重复、不遗漏的概率问题或对于基本事件的等可能性难于验证的概率问题，可用随机模拟方法来估计概率。

【练习 3】袋子中有四个小球，分别写有“中、华、民、族”四个字，有放回地从中任取一个小球，直到“中”“华”两个字都取到才停止.用随机模拟的方法估计恰好抽取三次停止的概率，利用电脑随机产生 0 到 3 之间取整数值的随机数，分别用 0、1、2、3 代表“中、华、民、族”这四个字，以每三个随机数为一组，表示取球三次的结果，经随机模拟产生了以下 18 组随机数：

232 321 230 023 123 021 132 220 001  
231 130 133 231 031 320 122 103 233

由此可以估计，恰好抽取三次就停止的概率为（ ）

- A.  $\frac{1}{9}$                       B.  $\frac{3}{18}$                       C.  $\frac{2}{9}$                       D.  $\frac{5}{18}$

[答案]C

[解析]由随机产生的随机数可知恰好抽取三次就停止的有 021、001、130、031，共 4 组随机数，恰好抽取三次就停止的概率约为  $\frac{4}{18} = \frac{2}{9}$

【练习 4】一个袋中有 7 个大小、形状相同的小球，6 个白球、1 个红球。现任取 1 个，若为红球就停止，若为白球就放回，搅拌均匀后再接着取.试设计一个模拟试验，计算恰好第三次摸到红球的概率.

[答案]0.1

[解析]用 1、2、3、4、5、6 表示白球，7 表示红球，利用计算器或计算机产生 1 到 7 之间取整数值的随机数，因为要求恰好第三次摸到红球的概率，所以每三个随机数作为一组。例如，产生 20 组随机数：

666 743 671 464 571 561 156 567 732 375  
716 116 614 445 117 573 552 274 114 622

就相当于做了 20 次试验，在这组数中，前两个数字不是 7，第三个数字恰好是 7，就表示第一次、第二次摸的是白球，第三次恰好是红球，它们分别是 567 和 117 共两组，

因此恰好第三次摸到红球的概率约为  $\frac{2}{20} = 0.1$ .

【课后巩固】—————完成课时作业（179）

## 18.8 条件概率

### 【学习目标】

1. 理解条件概率的定义.
2. 掌握条件概率的计算方法.
3. 利用条件概率公式解决一些简单的实际问题.

### 【要点整合】

要点一 条件概率的概念

设  $A, B$  为两个事件, 且  $P(A) > 0$ , 称  $P(B|A) = \frac{P(AB)}{P(A)}$  为在事件  $A$  发生的条件下, 事件  $B$  发生的条

件概率. 简称“条件概率”.  $P(B|A)$  读作  $A$  发生的条件下  $B$  发生的概率。

思考 (1) 若事件  $A, B$  互斥, 则  $P(B|A)$  是多少?

(2) 如果  $P(B|A)$  与  $P(B)$  相等, 那么事件  $A, B$  应满足什么条件?

答案 (1)  $A, B$  互斥, 即  $A, B$  不同时发生.  $\therefore P(AB) = 0, \therefore P(B|A) = 0$

(2)  $\because P(B|A) = \frac{P(AB)}{P(A)} = P(B) \therefore P(AB) = P(A)P(B)$  即事件  $A$  与事件  $B$  独立。

要点二 概率的乘法公式

由条件概率的定义, 对任意两个事件  $A, B$ , 若  $P(A) > 0$ , 则  $P(AB) = P(A)P(B|A)$

要点三 条件概率的性质

1.  $P(B|A) \in [0, 1] \quad P(\Omega|A) = 1 (\Omega \text{ 为样本空间})$
2. 如果  $B$  与  $C$  是两个互斥事件, 则  $P(B \cup C|A) = P(B|A) + P(C|A)$
3. 如果  $B$  与  $\bar{B}$  互为对立事件, 则  $P(\bar{B}|A) = 1 - P(B|A)$

### 【典例讲练】

题型一: 条件概率的定义及计算

例 1 现有 6 个节目准备参加比赛, 其中 4 个舞蹈节目, 2 个语言类节目, 如果不放回地依次抽取 2 个节目, 求: (1) 第 1 次抽到舞蹈节目的概率;

(2) 第 1 次和第 2 次都抽到舞蹈节目的概率;

(3) 在第 1 次抽到舞蹈节目的条件下, 第 2 次抽到舞蹈节目的概率.

[解析] 设第 1 次抽到舞蹈节目为事件  $A$ , 第 2 次抽到舞蹈节目为事件  $B$ , 则第 1 次和第 2 次都抽到舞蹈节目为事件  $AB$ .

(1) 从 6 个节目中不放回地依次抽取 2 个, 总的事件数  $n(\Omega) = A_6^2 = 30$ .

根据分步乘法计数原理, 有  $n(A) = A_1 A_2 = 20$ , 所以  $P(A) = \frac{n(A)}{n(\Omega)} = \frac{20}{30} = \frac{2}{3}$ .

(2) 因为  $n(AB) = A_2 = 12$ , 所以  $P(AB) = \frac{n(AB)}{n(\Omega)} = \frac{12}{30} = \frac{2}{5}$ .

(3) 方法一 由(1)(2), 得在第 1 次抽到舞蹈节目的条件下, 第 2 次抽到舞蹈节目的概率

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{\frac{2}{5}}{\frac{2}{3}} = \frac{3}{5}.$$

方法二 因为  $n(AB) = 12$ ,  $n(A) = 20$ , 所以  $P(B|A) = \frac{n(AB)}{n(A)} = \frac{12}{20} = \frac{3}{5}$ .

**【练习 1】** 从混有 5 张假钞的 20 张百元钞票中任意抽取两张

(1) 已知第一次抽取的是假钞, 求第二次也是假钞的概率。

(2) 一次取一张, 第一次取后放回, 已知第一次抽取的是假钞, 求第二次也是假钞的概率。

(3) 求有一张是假钞, 求另一张也是假钞的概率。

[解析] (1) 法一:  $A$  表示“第一次抽到假钞”  $B$  表示“第二次抽到假钞”

$$P(A) = \frac{C_5^1}{C_{20}^1} = \frac{1}{4} \quad P(AB) = \frac{C_5^2}{C_{20}^2} = \frac{1}{19} \quad P(B|A) = \frac{P(AB)}{P(A)} = \frac{4}{19}$$

法二: 缩小基本事件范围

第一次抽到假钞后, 剩 19 张钞票, 其中假钞 4 张, 故第二次也是假钞的概率为  $\frac{4}{19}$

(2) 缩小基本事件范围

第二次还是从含有 5 张假钞的 20 张钞票里面抽一张, 故第二次也是假钞的概率为  $\frac{1}{4}$

(3) 注意: 其中有一张是假钞的意思是至少有一张是假钞。

$A$  表示“抽到的两张中至少有一张是假钞”,  $B$  表示“抽到的两张都是假钞”, 则所求概率为  $P(A|B)$ .

$$\because P(AB) = P(B) = \frac{C_5^2}{C_{20}^2}, P(A) = \frac{C_5^2 + C_5^1 C_{15}^1}{C_{20}^2}, \therefore P(B|A) = \frac{P(AB)}{P(A)} = \frac{C_5^2}{C_5^2 + C_5^1 C_{15}^1} = \frac{10}{85} = \frac{2}{17}.$$

例 2 集合  $A = \{1, 2, 3, 4, 5, 6\}$ , 甲、乙两人各从  $A$  中任取一个数, 若甲先取(不放回), 乙后取,

(1) 在甲抽到奇数的条件下, 求乙抽到的数比甲抽到的数大的概率。

(2) 在甲抽到奇数的条件下, 求乙抽到偶数概率。

(3) 若甲先取(放回), 乙后取, 若事件  $A$ : “甲抽到的数大于 4”; 事件  $B$ : “甲、乙抽到的两数之和等于 7”, 求  $P(B|A)$ .

[解析] (1) 将甲抽到数字  $a$ , 乙抽到数字  $b$ , 记作  $(a, b)$ , 甲抽到奇数的情形有  $(1, 2), (1, 3), (1, 4), (1, 5),$

$(1, 6), (3, 1), (3, 2), (3, 4), (3, 5), (3, 6), (5, 1), (5, 2), (5, 3), (5, 4), (5, 6)$ , 共 15 个. 在这 15 个中,

乙抽到的数比甲抽到的数大的有  $(1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (3, 4), (3, 5), (3, 6), (5, 6)$ , 共 9 个,

所以所求概率  $P = \frac{9}{15} = \frac{3}{5}$ .

(2) 在甲抽到奇数的情形中, 乙抽到偶数的有(1,2), (1,4), (1,6), (3,2), (3,4), (3,6), (5,2), (5,4), (5,6), 共 9 个, 所以所求概率  $P=\frac{9}{15}=\frac{3}{5}$ .

(3) 甲抽到的数大于 4 的情形有: (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6), 共 12 个, 其中甲、乙抽到的两数之和等于 7 的情形有: (5,2), (6,1), 共 2 个. 所以  $P(B|A)=\frac{2}{12}=\frac{1}{6}$ .

【练习 2】抛掷红、蓝两颗骰子, 记事件  $A$  为“蓝色骰子的点数为 4 或 6”, 事件  $B$  为“两颗骰子的点数之和大于 8”, 求:

(1) 事件  $A$  发生的条件下事件  $B$  发生的概率;

(2) 事件  $B$  发生的条件下事件  $A$  发生的概率.

[解析] (1)  $n(A)=6 \times 2=12$ .

由  $3+6=6+3=4+5=5+4>8, 4+6=6+4=5+5>8, 5+6=6+5>8, 6+6>8$  知  $n(B)=10$ ,

其中  $n(AB)=6$ . 所以  $(1) P(B|A)=\frac{n_{AB}}{n_A}=\frac{6}{12}=\frac{1}{2}$ .

$(2) P(A|B)=\frac{n_{AB}}{n_B}=\frac{6}{10}=\frac{3}{5}$ .

【练习 3】考虑恰有两个小孩的家庭. 若已知某家有男孩, 求这家有两个男孩的概率; 若已知某家第一个是男孩, 求这家有两个男孩(相当于第二个也是男孩)的概率(假定生男生女为等可能).

[解析]  $\Omega=\{(\text{男}, \text{男}), (\text{男}, \text{女}), (\text{女}, \text{男}), (\text{女}, \text{女})\}$ .

设  $B$  = “有男孩”, 则  $B=\{(\text{男}, \text{男}), (\text{男}, \text{女}), (\text{女}, \text{男})\}$ .

$A$  = “有两个男孩”, 则  $A=\{(\text{男}, \text{男})\}$ ,

$B_1$  = “第一个是男孩”, 则  $B_1=\{(\text{男}, \text{男}), (\text{男}, \text{女})\}$ ,

于是得  $P(B)=\frac{3}{4}$ ,  $P(BA)=P(A)=\frac{1}{4}$ ,  $\therefore P(A|B)=\frac{P(BA)}{P(B)}=\frac{1}{3}$ ,

$P(B_1)=\frac{1}{2}$ ,  $P(B_1A)=P(A)=\frac{1}{4}$ ,  $\therefore P(A|B_1)=\frac{P(B_1A)}{P(B_1)}=\frac{1}{2}$ .

## 题型二: 概率的乘法公式

例 3 已知 1 号箱中有 2 个白球和 4 个红球, 2 号箱中有 5 个白球和 3 个红球, 现随机地从 1 号箱中取出一球放入 2 号箱, 然后从 2 号箱随机取出一球, 求两次都取到红球的概率.

[解析] 设从 1 号箱取到红球为事件  $A$ , 从 2 号箱取到红球为事件  $B$ .

由题意,  $P(A)=\frac{4}{2+4}=\frac{2}{3}$ ,  $P(B|A)=\frac{3+1}{8+1}=\frac{4}{9}$ ,

$\therefore P(AB)=P(B|A) \cdot P(A)=\frac{2}{3} \times \frac{4}{9}=\frac{8}{27}$ , 所以两次都取到红球的概率为  $\frac{8}{27}$ .

【练习3】已知3张奖券中只有1张有奖，甲、乙、丙3名同学依次无放回地各抽一张，他们中奖的概率与抽奖得次序有关吗？

[解析]用  $A, B, C$  分别表示甲、乙、丙中奖的事件，则  $B = \overline{AB}, C = \overline{AB}$   $P(A) = \frac{1}{3}$

$$P(B) = P(\overline{AB}) = P(\overline{A})P(B|\overline{A}) = \frac{2}{3} \times \frac{1}{2} = \frac{1}{3} \quad P(C) = P(\overline{AB}) = P(\overline{A})P(\overline{B}|\overline{A}) = \frac{2}{3} \times \frac{1}{2} = \frac{1}{3}$$

$P(A) = P(B) = P(C)$ , 所以中奖的概率与抽奖的次序无关。

### 题型三：条件概率的性质与应用

例4 把外形相同的球分装在三个盒子中，每盒10个。其中，第一个盒子中有7个球标有字母A，3个球标有字母B；第二个盒子中有红球和白球各5个；第三个盒子中有红球8个，白球2个。试验按如下规则进行：先在第一个盒子中任取一个球，若取得标有字母A的球，则在第二个盒子中任取一个球；若第一次取得标有字母B的球，则在第三个盒子中任取一个球。如果第二次取出的是红球，则称试验成功，求试验成功的概率。

[解析]设  $A = \{\text{从第一个盒子中取得标有字母A的球}\}$ ,  $B = \{\text{从第一个盒子中取得标有字母B的球}\}$ ,  $R = \{\text{第二次取出的球是红球}\}$ ,  $W = \{\text{第二次取出的球是白球}\}$ ,

$$\text{则容易求得 } P(A) = \frac{7}{10}, P(B) = \frac{3}{10}, P(R|A) = \frac{1}{2}, P(W|A) = \frac{1}{2}, P(R|B) = \frac{4}{5}, P(W|B) = \frac{1}{5}.$$

事件“试验成功”表示为  $AR \cup BR$ ，又事件  $AR$  与事件  $BR$  互斥，故由概率的加法公式，得

$$P(AR \cup BR) = P(AR) + P(BR) = P(R|A)P(A) + P(R|B)P(B) = \frac{1}{2} \times \frac{7}{10} + \frac{4}{5} \times \frac{3}{10} = 0.59.$$

【练习4】在某次考试中，要从20道题中随机抽出6道题，若考生至少能答对其中4道题即可通过，至少能答对其中5道题就获得优秀。已知某考生能答对其中10道题，并且知道他在这次考试中已经通过，求他获得优秀成绩的概率。

[解析]记事件  $A$  为“该考生6道题全答对”，事件  $B$  为“该考生答对了其中5道题，另一道答错”，事件  $C$  为“该考生答对了其中4道题，另2道题答错”，事件  $D$  为“该考生在这次考试中通过”，事件  $E$  为“该考生在这次考试中获得优秀”，则  $A, B, C$  两两互斥，且  $D = A \cup B \cup C$ ,  $E = A \cup B$ ，可知  $P(D) = P(A \cup B \cup C) = P(A) +$

$$P(B) + P(C) = \frac{C_{10}^6}{C_{20}^6} + \frac{C_{10}^5 C_{10}^1}{C_{20}^6} + \frac{C_{10}^4 C_{10}^2}{C_{20}^6} = \frac{12180}{C_{20}^6}, P(AD) = P(A), P(BD) = P(B),$$

$$P(E|D) = P(A|D) + P(B|D) = \frac{P(A)}{P(D)} + \frac{P(B)}{P(D)} = \frac{\frac{C_{10}^6}{C_{20}^6}}{\frac{12180}{C_{20}^6}} + \frac{\frac{C_{10}^5 C_{10}^1}{C_{20}^6}}{\frac{12180}{C_{20}^6}} = \frac{13}{58}. \text{故获得优秀成绩的概率为 } \frac{13}{58}.$$

【课后巩固】——完成课时作业（180）

## 18.9 全概率公式

### 【学习目标】

1. 理解并掌握全概率公式
2. 了解贝叶斯公式
3. 会用全概率公式及贝叶斯公式解题

### 【要点整合】

#### 要点一 全概率公式

一般地, 设  $A_1, A_2, \dots, A_n$  是一组两两互斥的事件,  $A_1 \cup A_2 \cup \dots \cup A_n = \Omega$ , 且  $P(A_i) > 0, i = 1, 2, \dots, n$ , 则

对任意的事件  $B \subseteq \Omega$ , 有 
$$P(B) = \sum_{i=1}^n P(A_i)P(B|A_i)$$

对全概率公式的理解:

某一事件  $B$  的发生有各种可能的原因( $i=1, 2, \dots, n$ ), 如果  $B$  是由原因  $A_i$  所引起, 则  $B$  发生的概率是

$P(BA_i) = P(A_i)P(B|A_i)$ , 每一原因都可能导致  $B$  发生, 故  $B$  发生的概率是各原因引起  $B$  发生概率的总和, 即全概率公式. 由此可以形象地把全概率公式看成为“由原因推结果”, 每个原因对结果的发生有一定的“作用”, 即结果发生的可能性与各种原因的“作用”大小有关. 全概率公式表达了它们之间的关系

#### 要点二 贝叶斯公式

设  $A_1, A_2, \dots, A_n$  是一组两两互斥的事件,  $A_1 \cup A_2 \cup \dots \cup A_n = \Omega$ , 且  $P(A_i) > 0, i = 1, 2, \dots, n$ , 则对任意的

事件  $B \subseteq \Omega, P(B) > 0$ , 有 
$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{k=1}^n P(A_k)P(B|A_k)} \quad i = 1, 2, \dots, n$$

对贝叶斯公式的理解: 贝叶斯公式是在“结果”已经发生条件下, 寻找各“原因”发生的条件概率. 因此要先计算事件  $B$  发生的概率(这里导致  $B$  发生的原因有  $A_i$ , 因此要使用全概率公式), 再计算由某种原因(比如  $A_1$ )导致  $B$  发生的概率作为分子, 最终计算已知  $B$  发生的前提下, 是由  $A_1$  导致它发生的概率.

### 【典例讲练】

#### 题型一: 全概率公式的概念及计算

例 1 (1) 已知事件  $A, B$ , 且  $P(A) = \frac{1}{3}, P(B|A) = \frac{1}{5}, P(B|\bar{A}) = \frac{2}{5}$ , 则  $P(B)$  等于( )

A.  $\frac{3}{5}$

B.  $\frac{1}{5}$

C.  $\frac{1}{3}$

D.  $\frac{1}{15}$

[答案] C

[解析]  $P(B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A}) = \frac{1}{3}$

(2) 一袋中装有大小、形状均相同的 5 个球, 其中 2 个黑球, 3 个白球, 从中先后不放回地任取一球, 则第二次取到的是黑球的概率为\_\_\_\_\_.

[答案]  $\frac{2}{5}$

[解析] 设事件  $A, B$  分别表示第一、二次取到的是黑球, 由古典概型可知  $P(A)=\frac{2}{5}, P(B|A)=\frac{1}{4}, P(B|\bar{A})=\frac{1}{2}$ .

则  $P(B)=P(AB)+P(\bar{A}B)=P(A)P(B|A)+P(\bar{A})P(B|\bar{A})=\frac{2}{5}$

【练习 1】有三个箱子, 分别编号为 1, 2, 3, 其中 1 号箱中装有 1 个红球和 4 个黄球, 2 号箱中装有 2 个红球和 3 个黄球, 3 号箱中装有 3 个红球. 某人从三个箱中任取一箱, 从中任意摸出一个球, 求取得红球的概率.

[解析] 记  $A_i$  = “球取自  $i$  号箱”,  $i=1, 2, 3$ ;  $B$  = “取得红球”,

即  $\Omega=A_1 \cup A_2 \cup A_3$ , 且  $A_1, A_2, A_3$  两两互斥, 根据  $P(A_1)=P(A_2)=P(A_3)=\frac{1}{3}$

$P(B|A_1)=\frac{1}{5}, P(B|A_2)=\frac{2}{5}, P(B|A_3)=1$  根据全概率公式得  $P(B)=\sum_{i=1}^3 P(A_i)P(B|A_i)=\frac{8}{15}$

例 2 甲箱的产品中有 5 个正品和 3 个次品, 乙箱的产品中有 4 个正品和 3 个次品.

(1) 从甲箱中任取 2 个产品, 求这 2 个产品都是次品的概率;

(2) 若从甲箱中任取 2 个产品放入乙箱中, 然后再从乙箱中任取一个产品, 求取出的这个产品是正品的概率.

[解析] (1) 从甲箱中任取 2 个产品的事件数为  $C_8^2=\frac{8 \times 7}{2}=28$ ,

这 2 个产品都是次品的事件数为  $C_3^2=3$ .  $\therefore$  这 2 个产品都是次品的概率为  $\frac{3}{28}$ .

(2) 设事件  $A$  为“从乙箱中取出的一个产品是正品”, 事件  $B_1$  为“从甲箱中取出 2 个产品都是正品”, 事件  $B_2$  为“从甲箱中取出 1 个正品 1 个次品”, 事件  $B_3$  为“从甲箱中取出 2 个产品都是次品”, 则事件  $B_1$ 、事件  $B_2$ 、事件  $B_3$  彼此互斥.

$P(B_1)=\frac{C_5^2}{C_8^2}=\frac{5}{14}, P(B_2)=\frac{C_5^1 C_3^1}{C_8^2}=\frac{15}{28}, P(B_3)=\frac{C_3^2}{C_8^2}=\frac{3}{28}, P(A|B_1)=\frac{2}{3}, P(A|B_2)=\frac{5}{9}, P(A|B_3)=\frac{4}{9},$

$\therefore P(A)=P(B_1)P(A|B_1)+P(B_2)P(A|B_2)+P(B_3)P(A|B_3)=\frac{5}{14} \times \frac{2}{3} + \frac{15}{28} \times \frac{5}{9} + \frac{3}{28} \times \frac{4}{9} = \frac{7}{12}.$

【练习 2】甲、乙、丙三人向同一飞靶射击, 设甲、乙、丙射中的概率分别为 0.4, 0.5, 0.7, 又设只有一人射中, 飞靶坠落的概率为 0.2, 若有二人射中, 飞靶坠落的概率为 0.6, 若有三人射中, 飞靶必坠落. 求飞靶坠落的概率.

[解析] 记  $A$  = “飞靶坠落”,  $B_i$  = “ $i$  个人射中飞靶”,  $i=1, 2, 3$ .

$B = “甲射中, 乙、丙未射中” + “乙射中, 甲、丙未射中” + “丙射中, 甲、乙未射中”$

$\therefore P(B_1)=0.4 \times 0.5 \times 0.3 + 0.6 \times 0.5 \times 0.3 + 0.6 \times 0.5 \times 0.7 = 0.36.$

同理  $P(B_2)=0.6 \times 0.5 \times 0.7 + 0.4 \times 0.5 \times 0.7 + 0.4 \times 0.5 \times 0.3 = 0.41. P(B_3)=0.4 \times 0.5 \times 0.7 = 0.14.$

由已知得  $P(A|B_1)=0.2$ ,  $P(A|B_2)=0.6$ ,  $P(A|B_3)=1$ .

利用全概率公式,  $P(A)=\sum_{i=1}^3 P(B_i)P(A|B_i)=0.36\times 0.2+0.41\times 0.6+0.14\times 1=0.458$

通过本例我们发现, 当直接求事件  $A$  发生的概率不好求时, 可以采用化整为零的方式, 即把  $A$  事件分解, 然后借助全概率公式间接求出事件  $A$  发生的概率.

## 题型二 贝叶斯公式的概念及计算

例 3 有 3 台车床加工同一型号的零件, 第 1 台加工的次品率为 6%, 第 2, 3 台加工的次品率均为 5%, 加工出来的零件混放在一起, 已知第 1, 2, 3 台车床加工的零件数分别占总数的 25%, 30%, 45%

(1) 任取一个零件, 计算它是次品的概率

(2) 如果取到的零件是次品, 计算它是第 1 台车床加工的概率。

[解析] 设  $B$  = “任取一个零件为次品”,  $A_i$  = “零件为第  $i$  台车床加工 ( $i=1, 2, 3$ )”, 则  $\Omega=A_1 \cup A_2 \cup A_3$

且,  $A_1, A_2, A_3$  两两互斥, 根据题意  $P(A_1)=0.25, P(A_2)=0.3, P(A_3)=0.45$

$$P(B|A_1)=0.06, P(B|A_2)=P(B|A_3)=0.05$$

(1) 由全概率公式, 得  $P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+P(A_3)P(B|A_3)$

$$=0.25\times 0.06+0.3\times 0.05+0.45\times 0.05=0.0525$$

$$(2) P(A_1|B)=\frac{P(A_1B)}{P(B)}=\frac{P(A_1)P(B|A_1)}{P(B)}=\frac{0.25\times 0.06}{0.0525}=\frac{2}{7}$$

【练习 3】一项血液化验用来鉴别是否患有某种疾病. 在患有此种疾病的人群中, 通过化验有 95% 的人呈阳性反应, 而健康的人通过化验也会有 1% 的人呈阳性反应. 某地区此种病的患者仅占人口的 0.5%. 若某人化验结果为阳性, 问此人确实患有此病的概率是多大?

[解析] 设  $A$  = “呈阳性反应”,  $B$  = “患有此种疾病”, 则  $P(A)=P(B)\cdot P(A|B)+P(\bar{B})\cdot P(A|\bar{B})=0.5\%\times 95\%+$

$$99.5\%\times 1\%=1.47\%. \text{ 所以 } P(B|A)=\frac{P(AB)}{P(A)}=\frac{P(B)P(A|B)}{P(B)P(A|B)+P(\bar{B})P(A|\bar{B})}=\frac{0.5\%\times 95\%}{1.47\%}=32.3\%.$$

[注] 利用贝叶斯公式求概率的步骤

第一步: 利用全概率公式计算  $P(A)$ , 即  $P(A)=\sum_{i=1}^n P(B_i)P(A|B_i)$ ;

第二步: 计算  $P(AB)$ , 可利用  $P(AB)=P(B)P(A|B)$  求解;

第三步: 代入  $P(B|A)=\frac{P(AB)}{P(A)}$  求解.

【课后巩固】——完成课时作业 (181)

## 18.10 离散型随机变量及其分布列

### 18.10.1 离散型随机变量

#### 【学习目标】

1. 理解随机变量及离散型随机变量的含义；
2. 了解随机变量与函数的区别与联系；
3. 会用离散型随机变量描述随机现象。

#### 【要点整合】

##### 要点一 随机试验

对于任何一个随机试验，总可以把它的每个样本点与一个实数相对应，即引入一个取值依赖于样本点的变量  $X$ ，来刻画样本点和实数的对应关系，实现样本点的数量化。因为在随机试验中样本点的出现具有随机性，所以变量  $X$  的取值也具有随机性。且变量  $X$  有如下特征：

- (1) 取值依赖于样本点；
- (2) 所有可能取值是明确的。

##### 要点二 随机变量

(1) 定义：一般地，对于随机试验样本空间  $\Omega$  中的每个样本点  $\omega$ ，都有唯一的实数  $X(\omega)$  与之对应，我们称  $X$  为随机变量。

(2) 表示：通常用大写英文字母表示随机变量，例如  $X, Y, Z$ ；用小写英文字母表示随机变量的取值，例如  $x, y, z$ 。

##### 要点三 随机变量与函数的关系

- (1) 联系：两者定义类似，样本点  $\omega$  相当于自变量，样本空间  $\Omega$  相当于函数的定义域。
- (2) 不同：样本空间  $\Omega$  不一定是数集。

##### 要点四 离散型随机变量

对于可能取值为有限个或可以一一列举的随机变量，称之为离散型随机变量。

#### 【典例讲练】

##### 题型一：随机变量的概念

例 1 指出下列变量中，哪些是随机变量，哪些不是随机变量，并说明理由

- (1) 任意掷一枚均匀硬币 5 次，出现正面向上的次数；
- (2) 某人射击一次命中的环数；
- (3) 投一颗质地均匀的骰子出现的点数(最上面的数字)；
- (4) 某个人的属相随年龄的变化；

[解析] (1) 任意掷一枚硬币 1 次，可能出现正面向上也可能出现反面向上，投掷 5 次硬币，出现正面向上的次数可能是 0,1,2,3,4,5，而且出现的结果是随机的，是随机变量。

(2) 某人射击一次，可能命中的环数是 0 环、1 环，...，10 环结果中的一个而且出现哪一个结果是随机的，因此是随机变量。

(3) 投一颗骰子出现的结果是 1 点, 2 点, 3 点, 4 点, 5 点, 6 点中的一个且出现哪个结果是随机的, 因此是随机变量。

(4) 属相是出生时便确定的, 不随年龄的变化而变化, 不是随机变量。

**【练习 1】** 投掷均匀硬币一枚, 随机变量为( )

- (1) 出现正面的次数
- (2) 出现正面或反面的次数
- (3) 掷硬币的次数
- (4) 出现正、反面次数之和

**[答案]** (1)

**[解析]** 掷一枚硬币, 可能出现的结果是正面向上或反面向上, 以一个标准如正面向上的次数来描述这一随机试验, 那么正面向上的次数就是随机变量  $X$ ,  $X$  的取值是 0, 1, 故选 (1); 而 (2) 中标准模糊不清,

(3) 中掷硬币次数是确定的, 都不是随机变量; (4) 中对应的事件是必然事件。

## 题型二：离散型随机变量的判定

**例 2** 指出下列随机变量是不是离散型随机变量, 并说明理由

- (1) 从 10 张已编好号码的卡片(1 号到 10 号)中任取一张, 被取出的卡片的号数  $X$ ;
- (2) 一个袋中装有 5 个白球和 5 个黑球, 从中任取 3 个, 其中所含白球的个数  $Y$ ;
- (3) 嘉陵江某地水位监测站所测水位在  $(0, 29]$  这一范围内变化, 该水位站所测水位  $H$ ;
- (4) 某加工厂加工的某种铜管的外径与规定的外径尺寸之差  $X$ 。

**[解析]** (1) 只要取出一张, 便有一个号码, 因此被取出的卡片号数可以一一列出, 符合离散型随机变量的定义。

(2) 从 10 个球中取 3 个球, 所得的结果有以下几种: 3 个白球; 2 个白球和 1 个黑球; 1 个白球和 2 个黑球; 3 个黑球, 即其结果可以一一列出, 符合离散型随机变量的定义。

(3) 水位在  $(0, 29]$  这一范围内变化, 对水位值我们不能按一定次序一一列出, 所以不是离散型随机变量。

(4) 实际测量值与规定值之间的差值无法一一列出, 不是离散型随机变量。

**【练习 2】** 判断下列变量是不是随机变量, 如果是, 判断该随机变量是不是离散型随机变量。

- (1) 某高速公路收费站在一天内经过的车辆数;
- (2) 某品牌电视机的使用寿命;
- (3) 巴蜀中学每一名学生的身高 (单位: cm, 精确到 1cm);
- (4) 重庆市在国庆节这一天的温度数。

**[解析]** (1) 变量可以一一列出, 所以是离散型随机变量。

(2) 使用寿命的可能取值是任何一个非负实数, 而所有非负实数不能一一列出, 所以使用寿命不是离散型随机变量。

(3) 由于每一个学生对对应着唯一的身高, 并且身高取整数值, 所以是离散型随机变量

(4) 温度数可以是国庆节当天最低温度和最高温度组成的温度区间内的任何一个数值, 是不可以一一列出的, 故不是离散型随机变量。

### 题型三：随机变量的应用

例3 写出下列各随机变量可能的取值，并说明随机变量所取的值表示随机试验的结果

- (1) 从一个编号为1号到20号的20个球的袋中，任取一球，被取出的球的编号 $X$ ；
- (2) 一个袋中装有大小相同的2个白球和5个黑球，从中任取3个，其中所含白球的个数为 $Y$ ；
- (3) 投掷两枚骰子，所得点数之和为 $X$ ，所得点数之和是奇数 $Y$ 。

[解析] (1)  $X$ 的可能取值为1,2,3, ..., 20,  $X = k(k = 1, 2, \dots, 10)$ 表示取出编号为 $k$ 的球。

(2)  $Y$ 可取0,1,2。  $Y = i(i = 0, 1, 2)$ 表示取出的3个球中有 $i$ 个白球， $(3-i)$ 个黑球。

(3) 若以 $(i, j)$ 表示投掷甲乙两枚骰子后骰子甲得 $i$ 点且骰子乙得 $j$ 点， $X$ 的可能取值为2,3,4, ..., 12, 则 $X=2$ 表示(1,1)； $X=3$ 表示(1,2),(2,1)； $X=4$ 表示(1,3),(2,2),(3,1)；...  $X=12$ 表示(6,6)。

$Y$ 可能的取值为3,5,7,9,11。则 $Y=3$ 表示(1,2),(2,1)； $Y=5$ 表示(1,4),(2,3),(3,2),(4,1)；...  $Y=11$ 表示(5,6),(6,5)。

【练习3】一个木箱中装有6个大小相同的篮球，编号为1, 2, 3, 4, 5, 6，现随机抽取3个篮球，以 $X$ 表示取出的篮球的最大号码，则 $X$ 的试验结果有\_\_\_\_\_种。

[答案]20

[解析] 从6个球中选出3个球， $X$ 可取3, 4, 5, 6。

当 $X=3$ 时，另两个球从1, 2中选取，有一种抽法；

当 $X=4$ 时，另两个球从1, 2, 3中选取，有3(种)；

当 $X=5$ 时，另两个球从1, 2, 3, 4中选取，有6(种)；

当 $X=6$ 时，另两个球从1, 2, 3, 4, 5中选取，有10(种)。

所以， $X$ 的试验结果共有 $1+3+6+10=20$ (种)。

【课后巩固】—————完成课时作业（182）

## 18.10.2 离散型随机变量的分布列

### 【学习目标】

1. 理解取有限值的离散型随机变量及其分布列的概念；
2. 认识分布列对刻画随机现象的重要性；
3. 掌握离散型随机变量分布列的表示方法和性质；
4. 理解两点分布。

### 【要点整合】

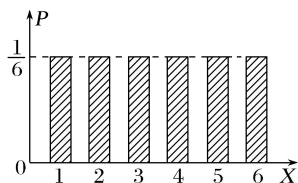
#### 要点一 离散型随机变量的分布列

(1) 分布列的定义：一般地，设离散型随机变量  $X$  的可能取值为  $x_1, x_2, \dots, x_n$ ，我们称  $X$  取每一个值  $x_i$  的概率  $P(X=x_i)=p_i, i=1, 2, \dots, n$  为  $X$  的概率分布列，简称分布列。

与函数的表示法类似，离散型随机变量的分布列也可以用表格表示：

|     |       |       |         |       |
|-----|-------|-------|---------|-------|
| $X$ | $x_1$ | $x_2$ | $\dots$ | $x_n$ |
| $P$ | $p_1$ | $p_2$ | $\dots$ | $p_n$ |

还可以用图形表示，如下图直观地表示了掷骰子试验中掷出的点数  $X$  的分布列，称为  $X$  的概率分布图。



(2) 分布列的性质：根据概率的性质，离散型随机变量分布列具有下述两个性质：

- (1)  $p_i \geq 0, i=1, 2, \dots, n$ ;
- (2)  $p_1 + p_2 + \dots + p_n = 1$ .

#### 要点二 两点分布

(1) 对于只有两个结果的随机试验，用  $A$  表示“成功”， $\bar{A}$  表示“失败”，定义

$$X = \begin{cases} 1, A \text{ 发生}, \\ 0, \bar{A} \text{ 发生}. \end{cases}$$

如果  $P(A) = p$ ，则  $P(\bar{A}) = 1 - p$ ，那么  $X$  的分布列如下表所示

|     |         |     |
|-----|---------|-----|
| $X$ | 0       | 1   |
| $P$ | $1 - p$ | $p$ |

我们称  $X$  服从两点分布或 0—1 分布。

### 【典例讲练】

#### 题型一：求离散型随机变量的分布列

例 1 某学校高二年级有 200 名学生，他们的体育综合测试成绩分 5 个等级，每个等级对应的分数和人数如下表所示

|    |     |    |    |    |    |
|----|-----|----|----|----|----|
| 等级 | 不及格 | 及格 | 中等 | 良  | 优  |
| 分数 | 1   | 2  | 3  | 4  | 5  |
| 人数 | 20  | 50 | 60 | 40 | 30 |

从这 200 名学生中任意选取 1 人, 求所选同学分数  $X$  的分布列, 以及  $P(X \geq 4)$ 。

[解析]由题意知,  $X$  是一个离散型随机变量, 其可能取值为 1,2,3,4,5, 且  $\{X=1\}$  =“不及格”,  $\{X=2\}$  =“及格”,  $\{X=3\}$  =“中等”,  $\{X=4\}$  =“良”,  $\{X=5\}$  =“优”。

根据古典概型的知识, 可得  $X$  的分布列, 如下表所示

|     |                |               |                |               |                |
|-----|----------------|---------------|----------------|---------------|----------------|
| $X$ | 1              | 2             | 3              | 4             | 5              |
| $P$ | $\frac{1}{10}$ | $\frac{1}{4}$ | $\frac{3}{10}$ | $\frac{1}{5}$ | $\frac{3}{20}$ |

$$P(X \geq 4) = P(X=4) + P(X=5) = \frac{1}{5} + \frac{3}{20} = \frac{7}{20}$$

【练习 1】一批笔记本电脑共有 10 台, 其中 A 品牌 3 台, B 品牌 7 台。如果从中随机挑选 2 台, 求这 2 台电脑中 A 品牌台数的分布列。

[解析]设挑选的 2 台电脑中 A 品牌的台数为  $X$ , 则  $X$  的可能取值为 0,1,2, 根据古典概型的知识, 可得  $X$

的分布列为  $P(X=0) = \frac{7}{15}, P(X=1) = \frac{7}{15}, P(X=2) = \frac{1}{15}$

用表格表示  $X$  的分布列, 如下表所示

|     |                |                |                |
|-----|----------------|----------------|----------------|
| $X$ | 0              | 1              | 2              |
| $P$ | $\frac{7}{15}$ | $\frac{7}{15}$ | $\frac{1}{15}$ |

例 2 已知随机变量  $X$  的分布列为

|     |                |               |               |                |               |                |
|-----|----------------|---------------|---------------|----------------|---------------|----------------|
| $X$ | -2             | -1            | 0             | 1              | 2             | 3              |
| $P$ | $\frac{1}{12}$ | $\frac{1}{4}$ | $\frac{1}{3}$ | $\frac{1}{12}$ | $\frac{1}{6}$ | $\frac{1}{12}$ |

分别求出随机变量  $Y_1 = \frac{1}{2}X$ ,  $Y_2 = X^2$  的分布列

[解析] 由  $Y_1 = \frac{1}{2}X$  知, 对于  $X$  取不同的值 -2, -1, 0, 1, 2, 3 时,  $Y_1$  的值分别为 -1,  $-\frac{1}{2}$ , 0,  $\frac{1}{2}$ , 1,  $\frac{3}{2}$ , 所以  $Y_1$

的分布列为

|       |                |                |               |                |               |                |
|-------|----------------|----------------|---------------|----------------|---------------|----------------|
| $Y_1$ | -1             | $-\frac{1}{2}$ | 0             | $\frac{1}{2}$  | 1             | $\frac{3}{2}$  |
| $P$   | $\frac{1}{12}$ | $\frac{1}{4}$  | $\frac{1}{3}$ | $\frac{1}{12}$ | $\frac{1}{6}$ | $\frac{1}{12}$ |

由  $Y_2 = X^2$  知, 对于  $X$  的不同取值 -2, 2 及 -1, 1,  $Y_2$  分别取相同的值 4 与 1, 即  $Y_2$  取 4 这个值的概率应是  $X$

取-2 与 2 的概率 $\frac{1}{12}$ 与 $\frac{1}{6}$ 的和,  $Y_2$  取 1 这个值的概率应是  $X$  取-1 与 1 的概率 $\frac{1}{4}$ 与 $\frac{1}{12}$ 的和,

所以  $Y_2$  的分布列为

|       |               |               |               |                |
|-------|---------------|---------------|---------------|----------------|
| $Y_2$ | 0             | 1             | 4             | 9              |
| $P$   | $\frac{1}{3}$ | $\frac{1}{3}$ | $\frac{1}{4}$ | $\frac{1}{12}$ |

【练习 2】已知随机变量  $X$  的分布列为

|     |                |               |               |                |               |                |
|-----|----------------|---------------|---------------|----------------|---------------|----------------|
| $X$ | -2             | -1            | 0             | 1              | 2             | 3              |
| $P$ | $\frac{1}{12}$ | $\frac{1}{4}$ | $\frac{1}{3}$ | $\frac{1}{12}$ | $\frac{1}{6}$ | $\frac{1}{12}$ |

求出随机变量  $Y_1 = -X + \frac{1}{2}$ ,  $Y_1 = X^2 - 2X$  的分布列

[解析]

|       |                |               |               |                |                |                |
|-------|----------------|---------------|---------------|----------------|----------------|----------------|
| $Y_1$ | $\frac{5}{2}$  | $\frac{3}{2}$ | $\frac{1}{2}$ | $-\frac{1}{2}$ | $-\frac{3}{2}$ | $-\frac{5}{2}$ |
| $P$   | $\frac{1}{12}$ | $\frac{1}{4}$ | $\frac{1}{3}$ | $\frac{1}{12}$ | $\frac{1}{6}$  | $\frac{1}{12}$ |

|       |                |               |               |                |
|-------|----------------|---------------|---------------|----------------|
| $Y_2$ | 8              | 3             | 0             | -1             |
| $P$   | $\frac{1}{12}$ | $\frac{1}{3}$ | $\frac{1}{2}$ | $\frac{1}{12}$ |

## 题型二：离散型随机变量分布列的性质及应用

例 3 设随机变量  $X$  的分布列  $P(X = \frac{k}{5}) = ak (k=1,2,3,4,5)$

(1)求常数  $a$  的值; (2)求  $P(X \geq \frac{3}{5})$ ; (3)求  $P(\frac{1}{10} < X < \frac{7}{10})$ .

[解析] (1)由分布列的性质得  $a+2a+3a+4a+5a=1$ ,  $a = \frac{1}{15}$

$$(2) P\left(X \geq \frac{3}{5}\right) = P\left(X = \frac{3}{5}\right) + P\left(X = \frac{4}{5}\right) + P\left(X = \frac{5}{5}\right) = \frac{3}{15} + \frac{4}{15} + \frac{5}{15} = \frac{4}{5}$$

$$\text{或 } P\left(X \geq \frac{3}{5}\right) = 1 - P\left(X \leq \frac{2}{5}\right) = 1 - \left(\frac{1}{15} + \frac{2}{15}\right) = \frac{4}{5}$$

$$(3) \because \frac{1}{10} < X < \frac{7}{10} \therefore P\left(\frac{1}{10} < X < \frac{7}{10}\right) = P\left(X = \frac{1}{5}\right) + P\left(X = \frac{2}{5}\right) + P\left(X = \frac{3}{5}\right) = \frac{1}{15} + \frac{2}{15} + \frac{3}{15} = \frac{2}{5}$$

【练习 3】若离散型随机变量  $X$  的分布列为

|     |        |          |
|-----|--------|----------|
| $X$ | -1     | 1        |
| $P$ | $4a-1$ | $3a^2+a$ |

试求出  $a$  的值

[答案]  $a = \frac{1}{3}$

[解析] 由离散型随机变量  $X$  的概率分布表知: 
$$\begin{cases} 0 \leq 4a-1 \leq 1 \\ 0 \leq 3a^2+a \leq 1 \\ 4a-1+3a^2+a=1 \end{cases} \text{ .解得 } a = \frac{1}{3} .$$

### 题型三：两点分布问题

例 4 一个箱子中装有红色玻璃球 3 个,白色玻璃球 5 个,从中摸出 2 个球,记随机变量  $X = \begin{cases} 0, \text{两球非全红} \\ 1, \text{两球全红} \end{cases}$ ,

求  $X$  的分布列。

[解析] 则  $X$  显然服从两点分布, 且  $P(X=1) = \frac{C_3^2}{C_8^2} = \frac{3}{28}$ ,  $\therefore P(X=0) = 1 - \frac{3}{28} = \frac{25}{28}$ ,  $\therefore X$  的分布列为

|     |                 |                |
|-----|-----------------|----------------|
| $X$ | 0               | 1              |
| $P$ | $\frac{25}{28}$ | $\frac{3}{28}$ |

【练习 4】若随机变量  $Y$  只能取两个值  $x_1$  与  $x_2$ , 且  $Y$  取  $x_1$  的概率是取  $x_2$  的概率的 3 倍, 写出  $Y$  的分布列, 并且说明是不是两点分布。

[解析]

|     |               |               |
|-----|---------------|---------------|
| $Y$ | $x_1$         | $x_2$         |
| $P$ | $\frac{3}{4}$ | $\frac{1}{4}$ |

不一定是两点分布, 当且仅当  $x_1$  与  $x_2$  的值为 0, 1 时为两点分布。

### 题型四：离散型随机变量的分布列的综合应用

例 5 某种资格证考试, 每位考生一年之内最多有 3 次考试机会, 一旦某次考试通过, 便可领取资格证书, 不再参加以后的考试, 否则就继续参加考试, 直到用完 3 此机会。小明决定参加考试, 如果他每次参加考试通过的概率依次为 0.6, 0.7, 0.8, 且每次考试是否通过相互独立, 试求:

- (1) 小明在一年内参加考试次数  $X$  的分布列;
- (2) 求小明在一年内领到资格证书的概率。

[解析] (1) 的取值分别为 1,2,3

$X = 1$ ，表明小明第一次参加考试就通过了，故  $P(X = 1) = 0.6$

$X = 2$ ，表明小明在第一次考试未通过，第二次通过了，故  $P(X = 2) = (1 - 0.6) \times 0.7 = 0.28$

$X = 3$ ，表明小明在第一、二次考试未通过，

故  $P(X = 3) = (1 - 0.6) \times (1 - 0.7) = 0.12$

|     |     |      |      |
|-----|-----|------|------|
| $X$ | 1   | 2    | 3    |
| $P$ | 0.6 | 0.28 | 0.12 |

(2) 小明在一年内通过考试的概率为  $1 - (1 - 0.6)(1 - 0.7)(1 - 0.8) = 0.976$

【练习 5】袋中装有黑球和白球共 7 个，从中任取 2 个球都是白球的概率为  $\frac{1}{7}$ ，现有甲、乙两人从袋中轮流摸取 1 球，甲先取，乙后取，然后甲再取，……，取后不放回，直到两人中有一人取到白球时终止，

每个球在每一次被取出的机会是等可能的，用  $X$  表示取球终止所需要的取球次数.

(1) 求袋中原有白球的个数；

(2) 求随机变量  $X$  的分布列；

(3) 求甲取到白球的概率.

[解析] (1) 设袋中原有  $n$  个白球，由题意知  $\frac{1}{7} = \frac{C_n^2}{C_7^2}$  可得  $n=3$  或  $n=-2$  (舍去)，即原有 3 个白球.

(2) 由题意， $X$  的可能取值为 1,2,3,4,5.

$$P(X=1) = \frac{3}{7}; \quad P(X=2) = \frac{4 \times 3}{7 \times 6} = \frac{2}{7}; \quad P(X=3) = \frac{4 \times 3 \times 3}{7 \times 6 \times 5} = \frac{6}{35};$$

$$P(X=4) = \frac{4 \times 3 \times 2 \times 3}{7 \times 6 \times 5 \times 4} = \frac{3}{35}; \quad P(X=5) = \frac{4 \times 3 \times 2 \times 1 \times 3}{7 \times 6 \times 5 \times 4 \times 3} = \frac{1}{35}.$$

所以  $X$  分布列为

|     |               |               |                |                |                |
|-----|---------------|---------------|----------------|----------------|----------------|
| $X$ | 1             | 2             | 3              | 4              | 5              |
| $P$ | $\frac{3}{7}$ | $\frac{2}{7}$ | $\frac{6}{35}$ | $\frac{3}{35}$ | $\frac{1}{35}$ |

(2) 因为甲先取，所以甲只有可能在第一次、第三次和第五次取到白球，记“甲取到白球”为事件  $A$ ,

(3) 则  $P(A) = P(X=1) + P(X=3) + P(X=5) = \frac{22}{35}$

【课后巩固】——完成课时作业 (183)

## 18.11 离散型随机变量的数字特征

### 18.11.1 离散型随机变量的均值

#### 【学习目标】

1. 通过实例理解离散型随机变量均值的概念，能计算简单离散型随机变量的均值；
2. 理解离散型随机变量均值的性质；
3. 掌握两点分布的均值；
4. 会利用离散型随机变量的均值，反映离散型随机变量取值水平，并解决一些实际问题。

#### 【要点整合】

要点一 离散型随机变量的均值或数学期望

一般地，若离散型随机变量  $X$  的分布列为

|     |       |       |          |       |
|-----|-------|-------|----------|-------|
| $X$ | $x_1$ | $x_2$ | $\cdots$ | $x_n$ |
| $P$ | $p_1$ | $p_2$ | $\cdots$ | $p_n$ |

则称  $E(X) = x_1p_1 + x_2p_2 + \cdots + x_np_n = \sum_{i=1}^n x_i p_i$  为随机变量  $X$  的均值或数学期望，数学期望简称期望。

均值是随机变量可能取值关于取值概率的加权平均数，它综合了随机变量的取值和取值的概率，反映了随机变量取值的平均水平。

要点二 离散型随机变量的均值的性质

若  $X$  是一个离散型随机变量，将  $X$  进行平行或伸缩后，即

$$(1) E(X+b) = E(X) + b \quad (2) E(aX) = aE(X) \quad (3) E(aX+b) = aE(X) + b$$

(1) 证明如下：设  $X$  的分布列为  $P(X=x_i) = p_i, i=1, 2, \cdots, n$ . 根据随机变量均值的定义，

$$\begin{aligned} E(X+b) &= (x_1+b)p_1 + (x_2+b)p_2 + \cdots + (x_n+b)p_n \\ &= (x_1p_1 + x_2p_2 + \cdots + x_np_n) + b(p_1 + p_2 + \cdots + p_n) \\ &= E(X) + b \end{aligned}$$

类似的可以证明 (2) (3)

要点三 两点分布的均值

一般地，如果随机变量  $X$  服从两点分布，那么  $E(X) = 0 \times (1-p) + 1 \times p = p$

#### 【典例讲练】

题型一：利用定义求均值

例 1 抛掷一枚质地均匀的骰子，设出现的点数为  $X$ ，求  $X$  的均值

[解析]  $X$  的分布列为  $P(X=k) = \frac{1}{6}, k=1, 2, 3, 4, 5, 6$ ；因此  $E(X) = \frac{1}{6}(1+2+3+4+5+6) = 3.5$

【练习 1】(1) 袋中有 4 只红球，3 只黑球，今从袋中随机取出 4 只球，设取到一只红球得 2 分，取到一只黑球得 1 分，试求得分  $X$  的均值

[解析]取出 4 只球颜色及得分分布情况是 4 红得 8 分, 3 红 1 黑得 7 分, 2 红 2 黑得 6 分, 1 红 3 黑得 5 分,

$$\text{因此 } P(X=5) = \frac{C_4^1 C_3^3}{C_7^4} = \frac{4}{35}, P(X=6) = \frac{C_4^2 C_3^2}{C_7^4} = \frac{18}{35}, P(X=7) = \frac{C_4^3 C_3^1}{C_7^4} = \frac{12}{35}, P(X=8) = \frac{C_4^4 C_3^0}{C_7^4} = \frac{1}{35}$$

故  $X$  的分布列如下:

|     |                |                 |                 |                |
|-----|----------------|-----------------|-----------------|----------------|
| $X$ | 5              | 6               | 7               | 8              |
| $P$ | $\frac{4}{35}$ | $\frac{18}{35}$ | $\frac{12}{35}$ | $\frac{1}{35}$ |

$$\therefore E(X) = 5 \times \frac{4}{35} + 6 \times \frac{18}{35} + 7 \times \frac{12}{35} + 8 \times \frac{1}{35} = \frac{44}{7}$$

(2) 在篮球比赛中, 罚球命中 1 次得 1 分, 不中得 0 分。如果某运动员罚球命中的概率为 0.8, 那么他罚球 1 次的得分  $X$  的均值是多少?

[解析]罚球有命中和不中两种可能结果, 命中时  $X=1$  不中时  $X=0$ , 因此随机变量  $X$  服从两点分布。

所以  $E(X) = 0.8$

## 题型二: 均值的性质

例 2 已知随机变量  $X$  的分布列为:

|     |               |               |               |     |                |
|-----|---------------|---------------|---------------|-----|----------------|
| $X$ | -2            | -1            | 0             | 1   | 2              |
| $P$ | $\frac{1}{4}$ | $\frac{1}{3}$ | $\frac{1}{5}$ | $m$ | $\frac{1}{20}$ |

(1) 求  $m$  的值;

(2) 求  $E(X)$ ;

(3) 若  $Y = 2X - 3$ , 求  $E(Y)$

$$[\text{答案}] (1) m = \frac{1}{6} \quad (2) E(X) = -\frac{17}{30} \quad (3) E(Y) = -\frac{62}{15}$$

[解析] (1) 由随机变量分布列的性质, 得  $\frac{1}{4} + \frac{1}{3} + \frac{1}{5} + m + \frac{1}{20} = 1$ , 解得  $m = \frac{1}{6}$

$$(2) \therefore E(X) = (-2) \times \frac{1}{4} + (-1) \times \frac{1}{3} + 0 \times \frac{1}{5} + 1 \times \frac{1}{6} + 2 \times \frac{1}{20} = -\frac{17}{30}.$$

$$(3) \text{由公式 } E(aX + b) = aE(X) + b, \text{ 得 } E(Y) = E(2X - 3) = 2E(X) - 3 = -\frac{62}{15}$$

【练习 2】(1) 已知随机变量  $X$  和  $Y$ , 其中  $Y = 12X + 7$ , 且  $E(Y) = 34$ , 若  $X$  的分布列如下表, 则  $m$  的值为\_\_\_\_\_.

|     |               |     |     |                |
|-----|---------------|-----|-----|----------------|
| $X$ | 1             | 2   | 3   | 4              |
| $P$ | $\frac{1}{4}$ | $m$ | $n$ | $\frac{1}{12}$ |

$$[\text{答案}] m = \frac{1}{3}$$

(2) 设离散型随机变量  $X$  可能的取值为  $1, 2, 3, 4$ ,  $P(X=k)=ak+b(k=1, 2, 3, 4)$ , 又  $X$  的数学期望为  $E(X)=3$ , 则  $a+b=$ \_\_\_\_\_

[答案]  $\frac{1}{10}$

[解析] 依题意可得  $X$  的分布列为

|     |       |        |        |        |
|-----|-------|--------|--------|--------|
| $X$ | 1     | 2      | 3      | 4      |
| $P$ | $a+b$ | $2a+b$ | $3a+b$ | $4a+b$ |

依题意得  $\begin{cases} a+b+2a+b+3a+b+4a+b=1 \\ (a+b)+2(2a+b)+3(3a+b)+4(4a+b)=3 \end{cases}$ , 解得  $a=\frac{1}{10}, b=0$ , 故  $a+b=\frac{1}{10}$ .

### 题型三：均值的应用

例 3 某商场举行有奖促销活动, 顾客购买每满 400 元的商品即可抽奖一次. 抽奖规则如下: 抽奖者掷各面标有 1~6 点数的正方体骰子 1 次, 若掷得点数不大于 4, 则可继续在抽奖箱中抽奖; 否则获得三等奖, 结束抽奖. 已知抽奖箱中装有 2 个红球与  $m(m \geq 2, m \in N^*)$  个白球, 抽奖者从箱中任意摸出 2 个球, 若 2 个球均为红球, 则获得一等奖, 若 2 个球为 1 个红球和 1 个白球, 则获得二等奖, 否则, 获得三等奖(抽奖箱中的所有小球, 除颜色外均相同). (1) 若  $m=4$ , 求顾客参加一次抽奖活动获得三等奖的概率;

(2) 若一等奖可获奖金 400 元, 二等奖可获奖金 300 元, 三等奖可获奖金 100 元, 记顾客一次抽奖所获得的奖金为  $X$ , 若商场希望  $X$  的数学期望不超过 150 元, 求  $m$  的最小值.

[解析] (1) 设顾客获得三等奖为事件  $A$ , 可分为两种情况: 顾客掷骰子掷得点数大于 4, 其概率为  $\frac{2}{6}=\frac{1}{3}$ ;

顾客掷骰子掷得点数不大于 4, 且摸出 2 个球均为白球, 其概率为  $\frac{4}{6} \times \frac{C_4^2}{C_6^2} = \frac{4}{15}$ ,

故当  $m=4$  时, 顾客参加一次抽奖活动获得三等奖的概率为  $P(A) = \frac{1}{3} + \frac{4}{15} = \frac{3}{5}$ .

(2) 由题意可得  $X$  的可能取值有 100, 300, 400,

且  $P(X=100) = \frac{1}{3} + \frac{2}{3} \times \frac{C_m^2}{C_{m+2}^2} = \frac{1}{3} + \frac{2m(m-1)}{3(m+2)(m+1)}$ ,

$P(X=300) = \frac{2}{3} \times \frac{C_2^1 C_m^1}{C_{m+2}^2} = \frac{8m}{3(m+2)(m+1)}$ ,  $P(X=400) = \frac{2}{3} \times \frac{C_2^2}{C_{m+2}^2} = \frac{4}{3(m+2)(m+1)}$ ,

则  $E(X) = 100 \times P(X=100) + 300 \times P(X=300) + 400 \times P(X=400)$ , 化简可得

$E(X) = \frac{100}{3} + \frac{200m^2 + 2200m + 1600}{3(m+2)(m+1)}$ , 由题意可得  $E(X) \leq 150$ , 即  $3m^2 - 23m - 18 \geq 0$ ,

又  $m \geq 2, m \in N^*$ , 则  $m \geq 9$ , 即  $m$  的最小值为 9.

【练习3】某迷宫有三个通道,进入迷宫的每个人都要经过一扇智能门。首次到达此门,系统会随机(即等可能)为你打开一个通道。若是1号通道,则需要1小时走出迷宫;若是2号、3号通道,则分别需要2小时、3小时返回智能门,再次到达智能门时,系统会随机打开一个你未到过的通道,直至走出迷宫为止。令 $X$ 表示走出迷宫所需的时间.

(1)求 $X$ 的分布列;

(2)求 $X$ 的数学期望.

[解析](1) $X$ 的所有可能取值为: 1, 3, 4, 6.  $X$ 的分布列为:

|     |               |               |               |               |
|-----|---------------|---------------|---------------|---------------|
| $X$ | 1             | 3             | 4             | 6             |
| $P$ | $\frac{1}{3}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{3}$ |

(2)  $E(X) = \frac{7}{2}$  (小时)

【课后巩固】—————完成课时作业(184)

## 18.11.2 离散型随机变量的方差

### 【学习目标】

1. 理解取有限个值的离散型随机变量的方差及标准差的概念；
2. 能计算简单离散型随机变量的方差，并能解决一些实际问题；
3. 掌握方差的性质，以及两点分布的方差的求法，会利用公式求方差。

### 【要点整合】

要点一 离散型随机变量的方差、标准差

(1) 设离散型随机变量  $X$  的分布列如下表所示

|     |       |       |          |       |
|-----|-------|-------|----------|-------|
| $X$ | $x_1$ | $x_2$ | $\cdots$ | $x_n$ |
| $P$ | $p_1$ | $p_2$ | $\cdots$ | $p_n$ |

考虑  $X$  所有可能取值  $x_i$  与  $E(X)$  的偏差的平方  $(x_1 - E(X))^2, (x_2 - E(X))^2, \dots, (x_n - E(X))^2$ 。因为  $X$  取每个值得概率不尽相同，所以我们用偏差平方关于取值概率的加权平均，来度量随机变量  $X$  取值与其均值  $E(X)$  的偏离程度。

我们称  $D(X) = (x_1 - E(X))^2 p_1 + (x_2 - E(X))^2 p_2 + \cdots + (x_n - E(X))^2 p_n = \sum_{i=1}^n (x_i - E(X))^2 p_i$

为随机变量  $X$  的方差，有时也记为  $Var(X)$ ，并称  $\sqrt{D(X)}$  为随机变量  $X$  的标准差，记为  $\sigma(X)$ 。

(2) 随机变量的方差和标准差都可以度量随机变量取值与其均值的偏离程度，反映了随机变量取值的离散程度。方差或标准差越小，随机变量的取值越集中；方差或标准差越大，随机变量的取值越分散。

(3) 在方差的计算中，利用下面的结论经常可以使计算简化

$$\begin{aligned} D(X) &= \sum_{i=1}^n (x_i - E(X))^2 p_i \\ &= \sum_{i=1}^n (x_i^2 - 2E(X)x_i + (E(X))^2) p_i \\ &= \sum_{i=1}^n x_i^2 p_i - 2E(X) \sum_{i=1}^n x_i p_i + (E(X))^2 \sum_{i=1}^n p_i \\ &= \sum_{i=1}^n x_i^2 p_i - (E(X))^2 \end{aligned}$$

要点二 离散型随机变量的方差的性质

(1) 若离散型随机变量  $X$  加上一个常数  $b$ ，仅仅使  $X$  的值产生一个平移，不改变  $X$  与其均值的离散程度，方差保持不变，即  $D(X+b) = D(X)$

(2) 而离散型随机变量  $X$  乘以一个常数  $a$ ，其方差变为原方差的  $a^2$  倍，即  $D(aX) = a^2 D(X)$

(3) 一般地，可以证明： $D(aX+b) = a^2 D(X)$

要点三 两点分布的均值

一般地，如果随机变量  $X$  服从两点分布，那么  $D(X) = p(1-p)$

## 【典例讲练】

### 题型一：求随机变量的方差与标准差

例 1 (1) 抛掷一枚质地均匀的骰子，设出现的点数为  $X$ ，求  $X$  的方差

[解析]  $X$  的分布列为  $P(X=k)=\frac{1}{6}, k=1,2,3,4,5,6$

$$\text{因此 } E(X)=\frac{7}{2}, \sum_{k=1}^6 (k^2 \times \frac{1}{6}) = \frac{1}{6}(1^2+2^2+3^2+4^2+5^2+6^2) = \frac{91}{6}$$

$$\text{所以 } D(X) = \sum_{k=1}^6 (k^2 \times \frac{1}{6}) - (\frac{7}{2})^2 = \frac{35}{12}$$

(2) 若某运动员投篮命中率  $P=0.8$ ，则该运动员在一次投篮中命中次数  $X$  的方差为\_\_\_\_\_.

[答案] 0.16

[解析] 依题意知： $X$  服从两点分布，所以  $D(X)=0.8 \times (1-0.8)=0.16$ .

【练习 1】设离散型随机变量  $X$  的分布列为

|     |               |               |               |               |
|-----|---------------|---------------|---------------|---------------|
| $X$ | 1             | 2             | 3             | 4             |
| $P$ | $\frac{1}{4}$ | $\frac{1}{3}$ | $\frac{1}{6}$ | $\frac{1}{4}$ |

则  $D(X)$  等于\_\_\_\_\_.

[解析] 由题意知， $E(X)=1 \times \frac{1}{4} + 2 \times \frac{1}{3} + 3 \times \frac{1}{6} + 4 \times \frac{1}{4} = \frac{29}{12}$ ，故  $D(X)=\frac{179}{144}$ .

### 题型二：方差的性质

例 2 已知随机变量  $X$  的分布列为：

|     |               |               |     |
|-----|---------------|---------------|-----|
| $X$ | 0             | 1             | $x$ |
| $P$ | $\frac{1}{2}$ | $\frac{1}{3}$ | $p$ |

$$\text{若 } E(X)=\frac{2}{3}$$

(1) 求  $D(X)$  的值；

(2) 若  $Y=3X-2$ ，求  $\sqrt{D(Y)}$  的值

[解析] 由分布列的性质，得  $\frac{1}{2} + \frac{1}{3} + p = 1$ ，解得  $p = \frac{1}{6}$ ，

$$\because E(X) = 0 \times \frac{1}{2} + 1 \times \frac{1}{3} + \frac{1}{6}x = \frac{2}{3}, \quad \therefore x = 2.$$

$$(1) D(X) = \frac{5}{9}.$$

$$(2) \because Y = 3X - 2, \quad \therefore D(Y) = D(3X - 2) = 9D(X) = 5, \quad \therefore \sqrt{D(Y)} = \sqrt{5}$$

【练习 2】随机变量  $X$  的分布列为

|     |               |               |               |
|-----|---------------|---------------|---------------|
| $X$ | -1            | 0             | 1             |
| $P$ | $\frac{1}{2}$ | $\frac{1}{3}$ | $\frac{1}{6}$ |

若  $Y=2X+2$ , 则  $D(Y)$  等于( )

A.  $-\frac{1}{3}$

B.  $\frac{5}{9}$

C.  $\frac{10}{9}$

D.  $\frac{20}{9}$

[答案]D

[解析]由题意知,  $E(X) = -1 \times \frac{1}{2} + 0 \times \frac{1}{3} + 1 \times \frac{1}{6} = -\frac{1}{3}$ , 故  $D(X) = \frac{5}{9}$ ,  $D(Y) = D(2X+2) = 4D(X) = \frac{20}{9}$

### 题型三：均值和方差的综合应用

例 3 设袋子中装有  $a$  个红球、 $b$  个黄球、 $c$  个蓝球, 且规定: 取出 1 个红球得 1 分, 取出 1 个黄球得 2 分, 取出 1 个蓝球得 3 分.

(1) 当  $a=3$ ,  $b=2$ ,  $c=1$  时, 从该袋子中依次任取 (有放回, 且每个球取到的机会均等) 2 个球, 记随机变量  $X$  为取出此 2 球所得分数之和, 求  $X$  的分布列;

(2) 从该袋子中任取 (每球取到的机会均等) 1 个球, 记随机变量  $Y$  为取出此球所得分数.

若  $E(Y) = \frac{5}{3}$ ,  $D(Y) = \frac{5}{9}$ , 求  $a:b:c$ .

[解析] (1) 依题意,  $X = 2, 3, 4, 5, 6$ , 且  $P(X=2) = \frac{3}{6} \times \frac{3}{6} = \frac{1}{4}$ ,  $P(X=3) = 2 \times \frac{3}{6} \times \frac{2}{6} = \frac{1}{3}$ ,

$$P(X=4) = 2 \times \frac{3}{6} \times \frac{1}{6} + \frac{2}{6} \times \frac{2}{6} = \frac{5}{18}, \quad P(X=5) = 2 \times \frac{2}{6} \times \frac{1}{6} = \frac{1}{9}, \quad P(X=6) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}.$$

所以  $X$  的分布列为

|     |               |               |                |               |                |
|-----|---------------|---------------|----------------|---------------|----------------|
| $X$ | 2             | 3             | 4              | 5             | 6              |
| $P$ | $\frac{1}{4}$ | $\frac{1}{3}$ | $\frac{5}{18}$ | $\frac{1}{9}$ | $\frac{1}{36}$ |

(2) 依题意可知  $Y$  的分布列为

|     |                   |                   |                   |
|-----|-------------------|-------------------|-------------------|
| $Y$ | 1                 | 2                 | 3                 |
| $P$ | $\frac{a}{a+b+c}$ | $\frac{b}{a+b+c}$ | $\frac{c}{a+b+c}$ |

$$E(Y) = \frac{a}{a+b+c} + \frac{2b}{a+b+c} + \frac{3c}{a+b+c} = \frac{5}{3} \text{ ①},$$

$$D(Y) = \left(1 - \frac{5}{3}\right)^2 \times \frac{a}{a+b+c} + \left(2 - \frac{5}{3}\right)^2 \times \frac{b}{a+b+c} + \left(3 - \frac{5}{3}\right)^2 \times \frac{c}{a+b+c} = \frac{5}{9} \text{ ②},$$

由①②得  $\begin{cases} 2a-b-4c=0 \\ a+4b-11c=0 \end{cases}$ , 解得  $a=3c, b=2c$ , 故  $a:b:c=3:2:1$

【练习 3】投资 A,B 两种股票, A 股票收益  $X$  元, B 股票收益  $Y$  元, 分布列如下表所示

|           |     |     |     |
|-----------|-----|-----|-----|
| 收益 $X$ /元 | -1  | 0   | 2   |
| 概率        | 0.1 | 0.3 | 0.6 |

|           |     |     |     |
|-----------|-----|-----|-----|
| 收益 $X$ /元 | 0   | 1   | 2   |
| 概率        | 0.3 | 0.4 | 0.3 |

(1) 投资哪种股票的期望收益大?

(2) 投资哪种股票的风险较高?

[解析] (1) 股票 A 和股票 B 投资收益的期望分别为  $E(X) = (-1) \times 0.1 + 0 \times 0.3 + 2 \times 0.6 = 1.1$

$E(Y) = 0 \times 0.3 + 1 \times 0.4 + 2 \times 0.3 = 1$ ; 因为  $E(X) > E(Y)$  所以投资股票 A 的期望收益较大。

(2)  $D(X) = (-1)^2 \times 0.1 + 0^2 \times 0.3 + 2^2 \times 0.6 - 1.1^2 = 1.29, D(Y) = 0^2 \times 0.3 + 1^2 \times 0.4 + 2^2 \times 0.3 - 1^2 = 0.6$ ,

因为  $E(X)$ 、 $E(Y)$  相差不大, 且  $D(X) > D(Y)$  所以投资股票 A 比投资股票 B 的风险高。

【课后巩固】—————完成课时作业 (185)

## 18.12 二项分布

### 【学习目标】

1. 理解  $n$  次独立重复试验的模型;
2. 理解二项分布;
3. 能利用独立重复试验的模型及二项分布解决一些简单的实际问题.

### 【要点整合】

#### 要点一 独立重复试验

- (1) 独立重复试验的定义: 一般地, 在相同条件下重复做的  $n$  次试验称为  $n$  次独立重复试验.
- (2) 独立重复试验中事件  $A$  恰好发生  $k$  次的概率: 一般地, 如果在 1 次试验中某事件发生的概率是  $p$ , 那么在  $n$  次独立重复试验中这个事件恰好发生  $k$  次的概率  $P(X = k) = C_n^k p^k (1 - p)^{n-k}, k = 0, 1, 2, \dots, n$ .

思考: 1. 有放回地抽样试验是独立重复试验吗?

2. 在  $n$  次独立重复试验中, 各次试验的结果相互有影响吗?

答案: 1. 是, 放回地抽样试验是相同条件下重复做的  $n$  次试验, 是独立重复试验.

2. 在  $n$  次独立重复试验中, 各次试验的结果相互之间无影响, 因为每次试验是在相同条件下独立进行的, 所以第  $i$  次试验的结果不受前  $i - 1$  次结果的影响 (其中  $i = 1, 2, \dots, n$ ).

#### 要点二 二项分布

(1) 一般地, 在  $n$  次独立重复试验中, 用  $X$  表示事件  $A$  发生的次数, 设每次试验中事件  $A$  发生的概率为  $p$ , 则  $P(X = k) = C_n^k p^k (1 - p)^{n-k}, k = 0, 1, 2, \dots, n$ . 此时称随机变量  $X$  服从二项分布, 记作  $X \sim B(n, p)$ , 并称  $p$  为成功概率.

### 【典例讲练】

#### 题型一: 独立重复试验的判断

例 1 判断下列试验是不是独立重复试验

- (1) 依次投掷四枚质地不同的硬币, 3 次正面向上;
- (2) 某人射击, 击中目标的概率是稳定的, 他连续射击了 10 次, 其中 6 次击中;
- (3) 口袋中装有 5 个白球, 3 个红球, 2 个黑球, 依次从中抽取 5 个球, 恰好抽出 4 个白球.

[解析] (1) 由于试验的条件不同(质地不同), 因此不是独立重复试验;

(2) 某人射击且击中的概率是稳定的, 因此是独立重复试验;

(3) 每次抽取, 试验的结果有三种不同的颜色, 且每种颜色出现的可能性不相等, 故不是独立重复试验.

【练习 1】下列事件: ①运动员甲射击一次, “射中 9 环”与“射中 8 环”; ②甲、乙两运动员各射击一次, “甲射中 10 环”与“乙射中 9 环”; ③甲、乙两运动员各射击一次, “甲、乙都射中目标”与“甲、乙都没射中目标”; ④在相同的条件下, 甲射击 10 次, 5 次击中目标. 其中是独立重复试验的是\_\_\_\_\_

[答案] ④

[解析] ①③符合互斥事件的概念, 是互斥事件; ②是相互独立事件; ④是独立重复试验.

## 题型二：独立重复试验的概率

例2 甲、乙两人各射击一次，击中目标的概率分别是 $\frac{2}{3}$ 和 $\frac{3}{4}$ ，假设每次射击是否击中目标，相互之间没有影响.(结果需用分数作答)

- (1) 求甲射击3次，至少有1次未击中目标的概率；  
(2) 求两人各射击2次，甲恰好击中目标2次且乙恰好击中目标1次的概率.

[解析] (1) 记“甲射击3次至少有1次未击中目标”为事件 $A_1$ ,

由题意，知射击3次，相当于3次独立重复试验，故 $P(A_1) = 1 - P(\bar{A}_1) = 1 - \left(\frac{2}{3}\right)^3 = \frac{19}{27}$ .

(2) 记“甲射击2次，恰有2次击中目标”为事件 $A_2$ ，“乙射击2次，恰有1次击中目标”为事件 $B_2$ ,

则 $P(A_2) = C_2^2 \times \left(\frac{2}{3}\right)^2 = \frac{4}{9}$ ,  $P(B_2) = C_2^1 \times \frac{3}{4} \times \left(1 - \frac{3}{4}\right) = \frac{3}{8}$ ; 由于甲、乙射击相互独立，

故 $P(A_2B_2) = \frac{4}{9} \times \frac{3}{8} = \frac{1}{6}$ .

[引申探究] 1. 在本例2的条件下，求甲、乙均击中目标1次的概率；

2. 在本例2的条件下，求甲未击中，乙击中2次的概率.

[解析] 1. 记“甲击中目标1次”为事件 $A_3$ ，“乙击中目标1次”为事件 $B_3$ ,

则 $P(A_3) = C_2^1 \times \frac{2}{3} \times \left(1 - \frac{2}{3}\right) = \frac{4}{9}$ ,  $P(B_3) = \frac{3}{8}$ , 所以甲、乙均击中目标1次的概率为 $P(A_3B_3) = \frac{4}{9} \times \frac{3}{8} = \frac{1}{6}$ .

2. 记“甲未击中目标”为事件 $A_4$ ，“乙击中2次”为事件 $B_4$ , 则 $P(A_4) = C_2^0 \times \left(1 - \frac{2}{3}\right)^2 = \frac{1}{9}$ ,  $P(B_4) = C_2^2 \times \left(\frac{3}{4}\right)^2 = \frac{9}{16}$ ,

所以甲未击中，乙击中2次的概率为 $P(A_4B_4) = \frac{1}{9} \times \frac{9}{16} = \frac{1}{16}$ .

【练习2】甲、乙两队进行排球比赛，已知在一局比赛中甲队胜的概率为 $\frac{2}{3}$ ，没有平局.

- (1) 若进行三局两胜制比赛，先胜两局者为胜，甲获胜的概率是多少？  
(2) 若进行五局三胜制比赛，甲获胜的概率为多少？

[解析] (1) 甲第一、二局胜，或第二、三局胜，或第一、三局胜，则 $P = \left(\frac{2}{3}\right)^2 + C_2^1 \times \frac{2}{3} \times \frac{1}{3} \times \frac{2}{3} = \frac{20}{27}$ .

(2) 甲前三局胜，或甲第四局胜，而前三局仅胜两局，或甲第五局胜，而前四局仅胜两局，

则 $P = \left(\frac{2}{3}\right)^3 + C_3^2 \times \left(\frac{2}{3}\right)^2 \times \frac{1}{3} \times \frac{2}{3} + C_4^2 \times \left(\frac{2}{3}\right)^2 \times \left(\frac{1}{3}\right)^2 \times \frac{2}{3} = \frac{64}{81}$ .

## 题型三：二项分布的应用

例3 一名学生每天骑自行车上学，从家到学校的途中有5个交通岗，假设他在各交通岗遇到红灯的事件是相互独立的，并且概率都是 $\frac{1}{3}$ .

- (1) 求这名学生在途中遇到红灯的次数 $\xi$ 的分布列；  
(2) 求这名学生在首次遇到红灯或到达目的地停车前经过的路口数 $\eta$ 的分布列；  
(3) 求这名学生在途中至少遇到一次红灯的概率.

[解析] (1) 由 $\xi \sim B\left(5, \frac{1}{3}\right)$ , 则 $P(\xi = k) = C_5^k \left(\frac{1}{3}\right)^k \left(1 - \frac{1}{3}\right)^{5-k}, k = 0, 1, 2, 3, 4, 5$

即 $P(\xi = 0) = C_5^0 \times \left(\frac{1}{3}\right)^0 \times \left(1 - \frac{1}{3}\right)^5 = \frac{32}{243}$   $P(\xi = 1) = C_5^1 \times \frac{1}{3} \times \left(1 - \frac{1}{3}\right)^4 = \frac{80}{243}$

$P(\xi = 2) = C_5^2 \times \left(\frac{1}{3}\right)^2 \times \left(1 - \frac{1}{3}\right)^3 = \frac{80}{243}$   $P(\xi = 3) = C_5^3 \times \left(\frac{1}{3}\right)^3 \times \left(1 - \frac{1}{3}\right)^2 = \frac{40}{243}$

$$P(\xi = 4) = C_5^4 \times (\frac{1}{3})^4 \times (1 - \frac{1}{3}) = \frac{10}{243} \quad P(\xi = 5) = C_5^5 \times (\frac{1}{3})^5 = \frac{1}{243}$$

故 $\xi$ 的分布列为

|       |                  |                  |                  |                  |                  |                 |
|-------|------------------|------------------|------------------|------------------|------------------|-----------------|
| $\xi$ | 0                | 1                | 2                | 3                | 4                | 5               |
| $P$   | $\frac{32}{243}$ | $\frac{80}{243}$ | $\frac{80}{243}$ | $\frac{40}{243}$ | $\frac{10}{243}$ | $\frac{1}{243}$ |

(2)  $\eta$ 的分布列为 $P(\eta = k) = P(\text{前 } k \text{ 个是绿灯, 第 } k+1 \text{ 个是红灯}) = (\frac{2}{3})^k \times \frac{1}{3}, k = 0, 1, 2, 3, 4$

$$\text{即 } P(\eta = 0) = (\frac{2}{3})^0 \times \frac{1}{3} = \frac{1}{3} \quad P(\eta = 1) = (\frac{2}{3})^1 \times \frac{1}{3} = \frac{2}{9}$$

$$P(\eta = 2) = (\frac{2}{3})^2 \times \frac{1}{3} = \frac{4}{27} \quad P(\eta = 3) = (\frac{2}{3})^3 \times \frac{1}{3} = \frac{8}{81}$$

$$P(\eta = 4) = (\frac{2}{3})^4 \times \frac{1}{3} = \frac{16}{243} \quad P(\eta = 5) = P(\text{前 } 5 \text{ 个是绿灯}) = (\frac{2}{3})^5 = \frac{4}{27}$$

故 $\eta$ 的分布列为

|       |               |               |                |                |                  |                |
|-------|---------------|---------------|----------------|----------------|------------------|----------------|
| $\xi$ | 0             | 1             | 2              | 3              | 4                | 5              |
| $P$   | $\frac{1}{3}$ | $\frac{2}{9}$ | $\frac{4}{27}$ | $\frac{8}{81}$ | $\frac{16}{243}$ | $\frac{4}{27}$ |

(3) 所求概率为 $P(\xi \geq 1) = 1 - P(\xi = 0) = 1 - (\frac{2}{3})^5 = \frac{211}{243}$ .

**【练习 3】**某一中学生心理咨询中心服务电话接通率为 $\frac{3}{4}$ , 某班 3 名同学商定明天分别就同一问题询问该服务中心, 且每人只拨打一次, 求他们中成功咨询的人数 $X$ 的分布列.

[解析]由题意可知 $X \sim B(3, \frac{3}{4})$ , 则 $P(X = k) = C_3^k (\frac{3}{4})^k (\frac{1}{4})^{3-k}, k = 0, 1, 2, 3$

$$\text{即 } P(X = 0) = C_3^0 \times (\frac{3}{4})^0 \times (\frac{1}{4})^3 = \frac{1}{64} \quad P(X = 1) = C_3^1 \times \frac{3}{4} \times (\frac{1}{4})^2 = \frac{1}{64}$$

$$P(X = 2) = C_3^2 \times (\frac{3}{4})^2 \times \frac{1}{4} = \frac{27}{64} \quad P(X = 3) = C_3^3 \times (\frac{3}{4})^3 = \frac{27}{64}$$

所以  $X$  的分布列为

|     |                |                |                 |                 |
|-----|----------------|----------------|-----------------|-----------------|
| $X$ | 0              | 1              | 2               | 3               |
| $P$ | $\frac{1}{64}$ | $\frac{9}{64}$ | $\frac{27}{64}$ | $\frac{27}{64}$ |

**例 4** 某公司安装了 3 台报警器, 它们彼此独立工作, 且发生险情时每台报警器报警的概率均为 0.9. 求发生险情时, 下列事件的概率:

(1) 3 台都未报警; (2) 恰有 1 台报警; (3) 恰有 2 台报警.

[解析]令 $X$ 为在发生险情时 3 台报警器中报警的台数, 那么 $X \sim B(3, 0.9)$ ,

则它的分布列为 $P(X = k) = C_3^k (0.9)^k (0.1)^{3-k}, k = 0, 1, 2, 3$

(1) 3 台都未报警的概率为:  $P(X = 0) = C_3^0 \times (0.9)^0 \times (0.1)^3 = 0.001$

(2) 恰有 1 台报警的概率为:  $P(X = 1) = C_3^1 \times 0.9 \times (0.1)^2 = 0.027$

(3) 恰有 2 台报警的概率为:  $P(X = 2) = C_3^2 \times 0.9^2 \times 0.1 = 0.243$

【练习 4】100 件产品中有 3 件不合格品，每次取一件，有放回地抽取 3 次，求取得不合格品的件数 $X$ 的分布列.

[解析] $X$ 的可能取值为 0,1,2,3.由于是有放回地每次取一件，连续取 3 次，所以相当于做了 3 次独立重复试验，每次抽取到不合格品的概率 $P = 0.03$ ，因此 $X \sim B(3, 0.03)$ .

$$\text{即 } P(X = 0) = C_3^0 \times (0.03)^0 \times (1 - 0.03)^3 = 0.912673$$

$$P(X = 1) = C_3^1 \times 0.03 \times (1 - 0.03)^2 = 0.084681$$

$$P(X = 2) = C_3^2 \times 0.03^2 \times (1 - 0.03) = 0.002619.$$

$$P(X = 3) = 0.03^3 = 0.000027$$

所以  $X$  的分布列为

|     |          |          |          |          |
|-----|----------|----------|----------|----------|
| $X$ | 0        | 1        | 2        | 3        |
| $P$ | 0.912673 | 0.084681 | 0.002619 | 0.000027 |

【课后巩固】—————完成课时作业（186）

## 18.13 超几何分布

### 【学习目标】

1. 理解超几何分布；
2. 能利用超几何分布解决一些简单的实际问题.

### 【要点整合】

一般地, 在含有  $M$  件次品产品中, 任取  $n$  件, 其中恰有  $X$  件次品, 则  $P(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$ ,  $k = 0, 1, 2, \dots, m$ , 其中  $m = \min\{M, n\}$ , 且  $n \leq N, M \leq N, n, M, N \in \mathbb{N}^*$ , 则称分布列

|     |                                     |                                     |     |                                     |
|-----|-------------------------------------|-------------------------------------|-----|-------------------------------------|
| $X$ | 0                                   | 1                                   | ... | $m$                                 |
| $P$ | $\frac{C_M^0 C_{N-M}^{n-0}}{C_N^n}$ | $\frac{C_M^1 C_{N-M}^{n-1}}{C_N^n}$ | ... | $\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}$ |

为超几何分布列. 如果随机变量  $X$  的分布列为超几何分布列, 则称随机变量  $X$  服从超几何分布.

### 【典例讲练】

#### 题型一: 超几何分布

例 1 某市 A, B 两所中学的学生组队参加辩论赛, A 中学推荐了 3 名男生、2 名女生, B 中学推荐了 3 名男生、4 名女生, 两校所推荐的学生一起参加集训. 由于集训后队员水平相当, 从参加集训的男生中随机抽取 3 人、女生中随机抽取 3 人组成代表队.

(1) 求 A 中学至少有 1 名学生入选代表队的概率;

(2) 某场比赛前, 从代表队的 6 名队员中随机抽取 4 人参赛, 设  $X$  表示参赛的男生人数, 求  $X$  的分布列.

[解析] (1) 由题意知, 参加集训的男生、女生各有 6 人. 代表队中的学生全从 B 中学抽取(等价于 A 中学没有学生入选代表队)的概率为  $\frac{C_3^3 C_4^3}{C_6^3 C_6^3} = \frac{1}{100}$ . 因此, A 中学至少有 1 名学生入选代表队的概率为  $1 - \frac{1}{100} = \frac{99}{100}$ .

(2) 根据题意, 知  $X$  的所有可能取值为 1, 2, 3.

$P(X = 1) = \frac{C_3^1 C_3^3}{C_6^4} = \frac{1}{5}$ ,  $P(X = 2) = \frac{C_3^2 C_3^2}{C_6^4} = \frac{3}{5}$ ,  $P(X = 3) = \frac{C_3^3 C_3^1}{C_6^4} = \frac{1}{5}$ . 所以  $X$  的分布列为

|     |               |               |               |
|-----|---------------|---------------|---------------|
| $X$ | 1             | 2             | 3             |
| $P$ | $\frac{1}{5}$ | $\frac{3}{5}$ | $\frac{1}{5}$ |

[反思感悟] 解决超几何分布问题的两个关键点:

①超几何分布是概率分布的一种形式, 一定要注意公式中字母的范围及其意义, 解决问题时可以直接利用公式求解, 但不能机械地记忆;

②超几何分布中, 只要知道  $M, N, n$  就可以利用公式求出  $X$  取不同  $k$  的概率  $P(X = k)$ , 从而求出  $X$  的分布列.

【练习 1】从某小组的 5 名女生和 4 名男生中任选 3 人去参加一项公益活动.

(1) 求所选 3 人中恰有一名男生的概率;

(2) 求所选 3 人中男生人数  $\xi$  的分布列.

[解析] (1) 所选 3 人中恰有一名男生的概率  $P = \frac{C_5^2 C_4^1}{C_9^3} = \frac{10}{21}$ .

(2)  $\xi$  的可能取值为 0, 1, 2, 3.

$$P(\xi = 0) = \frac{C_5^3}{C_9^3} = \frac{5}{42}, P(\xi = 1) = \frac{C_5^2 C_4^1}{C_9^3} = \frac{10}{21}, P(\xi = 2) = \frac{C_5^1 C_4^2}{C_9^3} = \frac{5}{14}, P(\xi = 3) = \frac{C_4^3}{C_9^3} = \frac{1}{21}.$$

所以 $\xi$ 的分布列为

|       |                |                 |                |                |
|-------|----------------|-----------------|----------------|----------------|
| $\xi$ | 0              | 1               | 2              | 3              |
| $P$   | $\frac{5}{42}$ | $\frac{10}{21}$ | $\frac{5}{14}$ | $\frac{1}{21}$ |

## 题型二：分布列的综合应用

例2 袋中装有标有数字1、2、3、4、5的小球各2个，从袋中任取3个小球，用 $X$ 表示取出的3个小球的最大数字.

(1) 取出的3个小球上的数字互不相同的概率;

(2) 随机变量 $X$ 的分布列.

[解析] (1) “一次取出的3个小球上的数字互不相同”的事件记为 $A$ ，则 $P(A) = \frac{C_5^3 C_2^1 C_2^1}{C_{10}^3} = \frac{2}{3}$ .

(2) 由题意知 $X$ 所有可能的取值是2、3、4、5， $P(X = 2) = \frac{C_2^2 C_2^1 + C_2^1 C_2^2}{C_{10}^3} = \frac{1}{30}$ ， $P(X = 3) = \frac{C_4^2 C_2^1 + C_4^1 C_2^2}{C_{10}^3} = \frac{2}{15}$ ，

$$P(X = 4) = \frac{C_6^2 C_2^1 + C_6^1 C_2^2}{C_{10}^3} = \frac{3}{10}, P(X = 5) = \frac{C_8^2 C_2^1 + C_8^1 C_2^2}{C_{10}^3} = \frac{8}{15}.$$

所以随机变量 $X$ 的分布列为

|     |                |                |                |                |
|-----|----------------|----------------|----------------|----------------|
| $X$ | 2              | 3              | 4              | 5              |
| $P$ | $\frac{1}{30}$ | $\frac{2}{15}$ | $\frac{3}{10}$ | $\frac{8}{15}$ |

[反思感悟] (1) 在求某些比较难计算的事件的概率时，我们可以先求随机变量取其他值时的概率，再根据概率之和为1的性质即可解决问题.

(2) 在解决含有“至少”“至多”的问题时，利用对立事件进行求解不失为一种好方法.

【练习2】某人有5把钥匙，其中只有一把能打开办公室的门，一次他醉酒后拿钥匙去开门.由于看不清是哪把钥匙，他只好逐一去试，若不能开门，则把钥匙扔到一边，记打开门时试开门的次数为 $\xi$ ，试求 $\xi$ 的分布列，并求他至多试开3次的概率.

[解析]  $\xi$ 的所有可能取值为1、2、3、4、5，且 $P(\xi = 1) = \frac{C_1^1}{C_5^1} = \frac{1}{5}$ ， $P(\xi = 2) = \frac{C_4^1 C_1^1}{C_5^2 C_4^1} = \frac{1}{5}$ ，

$$P(\xi = 3) = \frac{C_4^1 C_3^1 C_1^1}{C_5^3 C_4^2 C_3^1} = \frac{1}{5}, P(\xi = 4) = \frac{C_4^1 C_3^1 C_2^1 C_1^1}{C_5^4 C_4^3 C_3^2 C_2^1} = \frac{1}{5}, P(\xi = 5) = \frac{C_4^1 C_3^1 C_2^1 C_1^1 C_1^1}{C_5^5 C_4^4 C_3^3 C_2^2 C_1^1} = \frac{1}{5}.$$

因此 $\xi$ 的分布列为

|     |               |               |               |               |               |
|-----|---------------|---------------|---------------|---------------|---------------|
| $X$ | 1             | 2             | 3             | 4             | 5             |
| $P$ | $\frac{1}{5}$ | $\frac{1}{5}$ | $\frac{1}{5}$ | $\frac{1}{5}$ | $\frac{1}{5}$ |

由分布列知 $P(\xi \leq 3) = P(\xi = 1) + P(\xi = 2) + P(\xi = 3) = \frac{1}{5} + \frac{1}{5} + \frac{1}{5} = \frac{3}{5}$

【课后巩固】——完成课时作业（187）

## 18.14 正态分布

### 【学习目标】

1. 利用实际问题的直方图,了解正态分布曲线的特点及曲线所表示的意义.
2. 了解变量落在区间 $(\mu-\sigma, \mu+\sigma)$ ,  $(\mu-2\sigma, \mu+2\sigma)$ ,  $(\mu-3\sigma, \mu+3\sigma)$ 的概率大小.
3. 会用正态分布去解决实际问题.

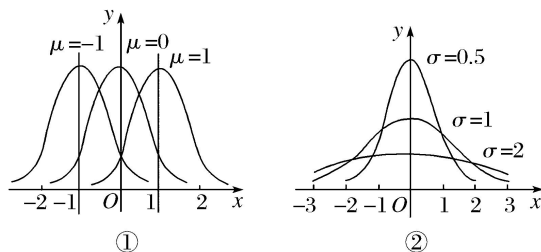
### 【要点整合】

#### 要点一 正态曲线

1. 正态曲线: 函数 $\varphi_{\mu, \sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ ,  $x \in \mathbf{R}$ , 其中实数 $\mu$ 和 $\sigma(\sigma > 0)$ 为参数,  $\varphi_{\mu, \sigma}(x)$ 的图象为正态分布密度曲线, 简称正态曲线.

2. 正态曲线 $\varphi_{\mu, \sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ ,  $x \in \mathbf{R}$ 有以下性质:

- (1) 曲线位于  $x$  轴上方, 与  $x$  轴不相交;
- (2) 曲线是单峰的, 它关于直线  $x = \mu$  对称;
- (3) 曲线在  $x = \mu$  处达到最大值  $\frac{1}{\sigma\sqrt{2\pi}}$ ;
- (4) 曲线与  $x$  轴之间的面积为 1;
- (5) 当  $\sigma$  一定时, 曲线的位置由  $\mu$  确定, 曲线随着  $\mu$  的变化而沿  $x$  轴平移, 如图①;
- (6) 当  $\mu$  一定时, 曲线的形状由  $\sigma$  确定,  $\sigma$  越小, 曲线越“瘦高”;  $\sigma$  越大, 曲线越“矮胖”, 如图②.



#### 要点二 正态分布

1. 正态分布: 一般地, 如果对于任何实数  $a, b(a < b)$ , 随机变量  $X$  满足  $P(a < X \leq b)$  等于  $x = a, x = b, x$  轴及正态曲线所围成的区域面积, 则称随机变量  $X$  服从正态分布. 正态分布完全由参数  $\mu$  和  $\sigma$  确定, 因此正态分布常记作  $N(\mu, \sigma^2)$ , 如果随机变量  $X$  服从正态分布, 则记为  $X \sim N(\mu, \sigma^2)$ . (此时,  $X$  不是离散型随机变量, 是连续型随机变量)

2.  $X \sim N(\mu, \sigma^2)$  中  $\mu, \sigma$  的统计意义

- (1)  $\mu$  可取任意实数, 表示平均水平的特征数,  $E(X) = \mu$ ;
- (2)  $\sigma > 0$  表示标准差,  $D(X) = \sigma^2$ .

当  $\mu$  一定时,  $\sigma$  越小, 正态曲线越“瘦高”, 表示总体的分布越集中;  $\sigma$  越大, 曲线越“矮胖”, 表示总体的分布越分散

#### 要点三 正态总体在三个特殊区间内取值的概率值及 $3\sigma$ 原则

1.  $P(\mu - \sigma < X \leq \mu + \sigma) = 0.6826$ ;  $P(\mu - 2\sigma < X \leq \mu + 2\sigma) = 0.9544$ ;  $P(\mu - 3\sigma < X \leq \mu + 3\sigma) = 0.9974$ .
2. 由  $P(\mu - 3\sigma < X \leq \mu + 3\sigma) = 0.9974$ , 知正态总体几乎总取值于区间  $(\mu - 3\sigma, \mu + 3\sigma)$  之内. 而在此区间以外取值的

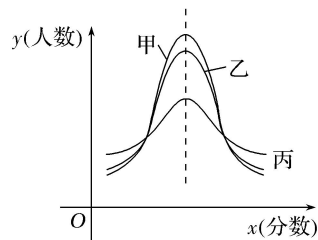
概率只有 0.0026,通常认为这种情况在一次试验中几乎不可能发生.

3. 在实际应用中,通常认为服从于正态分布  $N(\mu, \sigma^2)$  的随机变量  $X$  只取  $(\mu-3\sigma, \mu+3\sigma)$  之间的值,并简称之为  $3\sigma$  原则.

### 【典例讲练】

#### 题型一：正态曲线

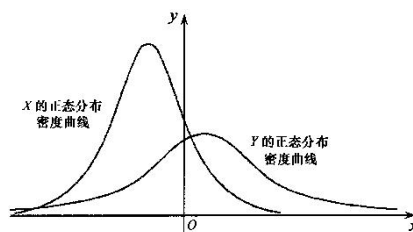
例 1 (1) 某市教学质量检测中,甲、乙、丙三科考试成绩的正态分布图如图所示(由于人数众多,成绩分布的直方图可视为正态分布),下列说法中正确的是( )



- A. 甲科总体的标准差最小
- B. 丙科总体的平均数最小
- C. 乙科总体的标准差及平均数都居中
- D. 甲、乙、丙总体的平均数不相同

[答案]A

(2) 设  $X \sim N(\mu_1, \sigma_1^2)$ ,  $Y \sim N(\mu_2, \sigma_2^2)$ , 这两个正态分布密度曲线如图所示. 下列结论中正确的是



- A.  $P(Y \geq \mu_2) \geq P(Y \geq \mu_1)$
- B.  $P(X \leq \sigma_2) \leq P(X \leq \sigma_1)$
- C. 对任意正数  $t$ ,  $P(X \leq t) \geq P(Y \leq t)$
- D. 对任意正数  $t$ ,  $P(X \geq t) \geq P(Y \geq t)$

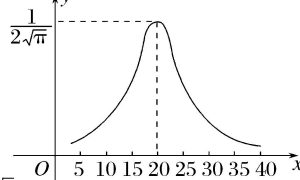
[答案]C

[解析]由正态分布密度曲线的性质可知,  $X \sim N(\mu_1, \sigma_1^2)$ ,  $Y \sim N(\mu_2, \sigma_2^2)$  的密度曲线分别关于直线  $x = \mu_1$ ,  $x = \mu_2$  对称, 因此结合题中所给图象可得,  $\mu_1 < \mu_2$ , 所以  $P(Y \geq \mu_2) < P(Y \geq \mu_1)$ , 故 A 错误. 又  $X \sim N(\mu_1, \sigma_1^2)$  得密度曲线较  $Y \sim N(\mu_2, \sigma_2^2)$  的密度曲线“瘦高”, 所以  $\sigma_1 < \sigma_2$ , 所以  $P(X \leq \sigma_2) > P(X \leq \sigma_1)$ , B 错误. 对任意正数  $t$ ,  $P(X \leq t) \geq P(Y \leq t)$ ,  $P(X \geq t) \geq P(Y \geq t)$ , C 正确, D 错误.

【练习 1】如图所示是一个正态分布的图象, 试根据该图象求出随机变量总体的均值和方差.

[答案]均值是  $\mu=20$ , 方差是  $\sigma^2=(\sqrt{2})^2=2$ .

[解析]从给出的正态曲线可知该正态曲线关于直线  $x=20$  对称, 最大值是  $\frac{1}{2\sqrt{\pi}}$ ,



所以  $\mu=20$ . 由  $\frac{1}{\sqrt{2\pi}\sigma} = \frac{1}{2\sqrt{\pi}}$ , 解得  $\sigma=\sqrt{2}$ . 随机变量总体的均值是  $\mu=20$ , 方差是  $\sigma^2=(\sqrt{2})^2=2$ .

#### 题型二：利用正态分布求概率

例 2 (1) 已知随机变量  $\xi$  服从正态分布  $N(2, \sigma^2)$ , 且  $P(\xi < 4) = 0.8$ , 则  $P(0 < \xi < 2) = ( )$

- A. 0.6
- B. 0.4
- C. 0.3
- D. 0.2

[答案]C

(2) 设  $\xi \sim N(1, 4)$ , 试求: ①  $P(\xi \geq 1)$ ; ②  $P(-1 < \xi \leq 3)$ ; ③  $P(3 < \xi \leq 5)$ ; ④  $P(\xi \geq 5)$ .

[解析]  $\because \xi \sim N(1, 4), \therefore \mu = 1, \sigma = 2,$

①  $P(\xi \geq 1) = 0.5$  ②  $P(-1 < \xi \leq 3) = P(1 - 2 < \xi \leq 1 + 2) = P(\mu - \sigma < \xi \leq \mu + \sigma) = 0.6826$ ;

③  $\because P(3 < \xi \leq 5) = P(-3 < \xi \leq -1),$

$$\therefore P(3 < \xi \leq 5) = \frac{1}{2}[P(-3 < \xi \leq 5) - P(-1 < \xi \leq 3)] = \frac{1}{2}[P(1 - 4 < \xi \leq 1 + 4) - P(1 - 2 < \xi \leq 1 + 2)]$$

$$= \frac{1}{2}[P(\mu - 2\sigma < \xi \leq \mu + 2\sigma) - P(\mu - \sigma < \xi \leq \mu + \sigma)] = \frac{1}{2}(0.9544 - 0.6826) = 0.1359.$$

$$\text{④ } P(\xi \geq 5) = P(\xi \leq -3) = \frac{1}{2}[1 - P(-3 < \xi \leq 5)] = \frac{1}{2}[1 - P(1 - 4 < \xi \leq 1 + 4)]$$

$$= \frac{1}{2}[1 - P(\mu - 2\sigma < \xi \leq \mu + 2\sigma)] = \frac{1}{2}(1 - 0.9544) = 0.0228.$$

[注] 利用正态分布求概率的两个方法

(1) 对称法: 由于正态曲线是关于直线  $x = \mu$  对称的, 且概率的和为 1, 故关于直线  $x = \mu$  对称的区间上概率相等.

如: ①  $P(X < a) = 1 - P(X \geq a)$ ; ②  $P(X < \mu - a) = P(X > \mu + a)$ ; ③  $P(X < \mu) = P(X \geq \mu) = 0.5$

(2) “3 $\sigma$ ”原则: 利用  $X$  落在区间  $(\mu - \sigma, \mu + \sigma]$ ,  $(\mu - 2\sigma, \mu + 2\sigma]$ ,  $(\mu - 3\sigma, \mu + 3\sigma]$  内的概率分别是 0.6826, 0.9544, 0.9974 求解.

【练习 2】(1) 设  $\xi \sim N(1, 1)$ , 试求: ①  $P(0 < \xi \leq 2)$ ; ②  $P(2 < \xi \leq 3)$ ; ③  $P(\xi \geq 3)$ .

[解析]  $\because \xi \sim N(1, 1), \therefore \mu = 1, \sigma = 1,$  ①  $P(0 < \xi \leq 2) = P(\mu - \sigma < \xi \leq \mu + \sigma) = 0.6826$

$$\text{② } P(2 < \xi \leq 3) = \frac{1}{2}[P(\mu - 2\sigma < \xi \leq \mu + 2\sigma) - P(\mu - \sigma < \xi \leq \mu + \sigma)] = \frac{1}{2}(0.9544 - 0.6826) = 0.1359.$$

$$\text{③ } P(\xi \geq 3) = \frac{1}{2}[1 - P(\mu - 2\sigma < \xi \leq \mu + 2\sigma)] = \frac{1}{2}(1 - 0.9544) = 0.0228$$

(2) 已知随机变量  $X$  服从正态分布  $N(m, \sigma^2)$ , 若  $P(X \leq -3) = P(X \geq 4)$ , 则  $m =$  \_\_\_\_\_

[答案] 0.5

### 题型三: 正态分布的实际应用

例 3 有一种精密零件, 其尺寸  $X$  (单位: mm) 服从正态分布  $N(20, 4)$ . 若这批零件共有 5 000 个, 试求:

(1) 这批零件中尺寸在 18~22 mm 间的零件所占的百分比;

(2) 若规定尺寸在 24~26 mm 间的零件不合格, 则这批零件中不合格的零件大约有多少个?

[解析] (1)  $\because X \sim N(20, 4), \therefore \mu = 20, \sigma = 2, \therefore \mu - \sigma = 18, \mu + \sigma = 22,$

于是尺寸在 18~22 mm 间的零件所占的百分比大约是 68.27%.

(2)  $\because \mu - 3\sigma = 14, \mu + 3\sigma = 26, \mu - 2\sigma = 16, \mu + 2\sigma = 24,$

$$\therefore \text{尺寸在 } 24 \sim 26 \text{ mm 间的零件所占的百分比大约是 } \frac{99.73\% - 95.45\%}{2} = 2.14\%.$$

$\therefore$  尺寸在 24~26 mm 间的零件大约有  $5\,000 \times 2.14\% \approx 107$  (个).

【练习 3】在某次大型考试中, 某班同学的成绩服从正态分布  $N(80, 5^2)$ , 现在已知该班同学中成绩在 80~85 分的有 17 人, 该班成绩在 90 分以上的同学有多少人?

[解析]  $\because$  成绩服从正态分布  $N(80, 5^2)$ ,  $\therefore \mu = 80, \sigma = 5$ , 则  $\mu - \sigma = 75, \mu + \sigma = 85$ .

$\therefore$  成绩在  $(75, 85]$  内的同学占全班同学的 68.27%, 成绩在  $(80, 85]$  内的同学占全班同学的 34.135%. 设该班有  $x$  名同学, 则  $x \times 34.135\% = 17$ , 解得  $x \approx 50$ .

$\because \mu - 2\sigma = 80 - 10 = 70, \mu + 2\sigma = 80 + 10 = 90$ ,

$\therefore$  成绩在  $(70, 90]$  内的同学占全班同学的 95.45%, 成绩在 90 分以上的同学占全班同学的 2.275%. 即有  $50 \times 2.275\% \approx 1$  (人), 即成绩在 90 分以上的仅有 1 人.

例 4 为了监控某种零件的一条生产线的生产过程, 检验员每天从该生产线上随机抽取 16 个零件, 并测量其尺寸 (单位:  $cm$ ). 根据长期生产经验, 可以认为这条生产线正常状态下生产的零件的尺寸服从正态分布  $N(\mu, \sigma^2)$ .

(1) 假设生产状态正常, 记  $X$  表示一天内抽取的 16 个零件中其尺寸在  $(\mu - 3\sigma, \mu + 3\sigma)$  之外的零件数, 求  $P(X \geq 1)$  及  $X$  的数学期望;

(2) 一天内抽检零件中, 如果出现了尺寸在  $(\mu - 3\sigma, \mu + 3\sigma)$  之外的零件, 就认为这条生产线在这一天的生产过程可能出现了异常情况, 需对当天的生产过程进行检查.

(i) 试说明上述监控生产过程方法的合理性;

(ii) 下面是检验员在一天内抽取的 16 个零件的尺寸:

|       |       |       |       |       |       |       |       |
|-------|-------|-------|-------|-------|-------|-------|-------|
| 9.95  | 10.12 | 9.96  | 9.96  | 10.01 | 9.92  | 9.98  | 10.04 |
| 10.26 | 9.91  | 10.13 | 10.02 | 9.22  | 10.04 | 10.05 | 9.95  |

经计算得  $\bar{x} = \frac{1}{16} \sum_{i=1}^{16} x_i = 9.97, s = \sqrt{\frac{1}{16} \sum_{i=1}^{16} (x_i - \bar{x})^2} = \sqrt{\frac{1}{16} \left( \sum_{i=1}^{16} x_i^2 - 16\bar{x}^2 \right)} \approx 0.212$ , 其中  $x_i$  为抽取的第  $i$

个零件的尺寸,  $i = 1, 2, \dots, 16$ .

用样本平均数  $\bar{x}$  作为  $\mu$  的估计值  $\hat{\mu}$ , 用样本标准差  $s$  作为  $\sigma$  的估计值  $\hat{\sigma}$ , 利用估计值判断是否需对当天的生产过程进行检查? 剔除  $(\hat{\mu} - 3\hat{\sigma}, \hat{\mu} + 3\hat{\sigma})$  之外的数据, 用剩下的数据估计  $\mu$  和  $\sigma$  (精确到 0.01).

附: 若随机变量  $Z$  服从正态分布  $N(\mu, \sigma^2)$ , 则  $P(\mu - 3\sigma, \mu + 3\sigma) = 0.9974$ ,

$0.9974^{16} \approx 0.9592, \sqrt{0.008} \approx 0.09$ .

[解析] (1) 抽取的一个零件的尺寸在  $(\mu - 3\sigma, \mu + 3\sigma)$  之内的概率为 0.9974, 从而零件的尺寸在

$(\mu - 3\sigma, \mu + 3\sigma)$  之外的概率为 0.0026, 故  $X \sim B(16, 0.0026)$ .

因此  $P(X \geq 1) = 1 - P(X = 0) = 1 - 0.9974^{16} \approx 0.0408$ .

$X$  的数学期望为  $EX = 16 \times 0.0026 = 0.0416$ .

(2) (i) 如果生产状态正常, 一个零件尺寸在  $(\mu - 3\sigma, \mu + 3\sigma)$  之外的概率只有 0.0026, 一天内抽取的 16 个零件

中,出现尺寸在 $(\mu-3\sigma, \mu+3\sigma)$ 之外的零件的概率只有 0.0408,发生的概率很小.因此一旦发生这种情况,就有理由认为这条生产线在这一天的生产过程可能出现了异常情况,需对当天的生产过程进行检查,可见上述监控生产过程的方法是合理的.

(ii)由 $\bar{x}=9.97, s \approx 0.212$ , 得 $\mu$ 的估计值为 $\hat{\mu}=9.97$ ,  $\sigma$ 的估计值为 $\hat{\sigma}=0.212$ ,由样本数据可以看出有一个零件的尺寸在 $(\hat{\mu}-3\hat{\sigma}, \hat{\mu}+3\hat{\sigma})$ 之外, 因此需对当天的生产过程进行检查.剔除 $(\hat{\mu}-3\hat{\sigma}, \hat{\mu}+3\hat{\sigma})$ 之外的数据 9.22, 剩下数据的平均数为 $\frac{1}{15}(16 \times 9.97 - 9.22) = 10.02$ , 因此 $\mu$ 的估计值为 10.02.

$\sum_{i=1}^{16} x_i^2 = 16 \times 0.212^2 + 16 \times 9.97^2 \approx 1591.134$ , 剔除 $(\hat{\mu}-3\hat{\sigma}, \hat{\mu}+3\hat{\sigma})$ 之外的数据 9.22, 剩下数据的样本方差为 $\frac{1}{15}(1591.134 - 9.22^2 - 15 \times 10.02^2) \approx 0.008$ ,

因此 $\sigma$ 的估计值为 $\sqrt{0.008} \approx 0.09$ .

【练习 4】从某企业生产的某种产品中抽取 500 件, 测量这些产品的一项质量指标值, 由测量结果得如下频率分布直方图:

(1)求这 500 件产品质量指标值的样本平均数 $\bar{x}$ 和样本方差 $s^2$

(同一组中的数据用该组区间的中点值作代表);

(2)由直方图可以认为, 这种产品的质量指标值 $Z$ 服从正态分布 $N(\mu, \sigma^2)$ , 其中 $\mu$ 近似为样本平均数 $\bar{x}$ ,  $\sigma^2$ 近似为样本方差 $s^2$ .

(i)利用该正态分布, 求 $P(187.8 < Z < 212.2)$ ;

(ii)某用户从该企业购买了 100 件这种产品, 记 $X$ 表示这 100 件产品中质量指标值位于区间 $(187.8, 212.2)$ 的产品件数, 利用①的结果, 求 $E(X)$ .

附:  $\sqrt{150} \approx 12.2$ .

若 $Z \sim N(\mu, \sigma^2)$ , 则 $P(\mu - \sigma < Z < \mu + \sigma) = 0.6826$ ,  $P(\mu - 2\sigma < Z < \mu + 2\sigma) = 0.9544$ .

[解析](1)抽取产品的质量指标值的样本平均数 $\bar{x}$ 和样本方差 $s^2$ 分别为

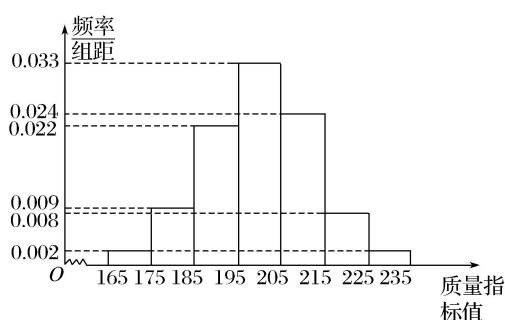
$$\bar{x} = 170 \times 0.02 + 180 \times 0.09 + 190 \times 0.22 + 200 \times 0.33 + 210 \times 0.24 + 220 \times 0.08 + 230 \times 0.02 = 200,$$

$$s^2 = (-30)^2 \times 0.02 + (-20)^2 \times 0.09 + (-10)^2 \times 0.22 + 0 \times 0.33 + 10^2 \times 0.24 + 20^2 \times 0.08 + 30^2 \times 0.02 = 150.$$

(2) (i)由(1)知,  $Z \sim N(200, 150)$ , 从而 $P(187.8 < Z < 212.2) = P(200 - 12.2 < Z < 200 + 12.2) = 0.6826$ .

(ii)由(i)知, 一件产品的质量指标值位于区间 $(187.8, 212.2)$ 的概率为 0.6826, 依题意知 $X \sim B(100, 0.6826)$ , 所以 $E(X) = 100 \times 0.6826 = 68.26$ .

(非标准正态分布转化为标准正态分布)



例 5 “公平正义”是社会主义和谐社会的重要特征, 是社会主义法治理念的价值追求.“考试”作为一种公平公正选拔人才的有效途径, 正被广泛采用.每次考试过后, 考生最关心的是:自己的考试名次是多少?自己能否被录取?能获得什么样的职位? 某单位准备通过考试(按照高分优先录取的原则)录用 300 名, 其中 275 个高薪职位和 25 个普薪职位.实际报名人数为 2000 名, 考试满分为 400 分.(一般地, 对于一次成功的

考试来说, 考试成绩应服从正态分布.) 考试后考试成绩的部分统计结果如下:

考试平均成绩是180分, 360分及其以上的高分考生30名.

(1)最低录取分数是多少?(结果保留为整数)

(2)考生甲的成绩为286分, 若甲被录取, 能否获得高薪职位?若不能被录取, 请说明理由. 参考资料:①

当  $X \sim N(\mu, \sigma^2)$  时, 令  $Y = \frac{X - \mu}{\sigma}$ , 则  $Y \sim N(0, 1)$ .

②当  $Y \sim N(0, 1)$  时,  $P(Y \leq 2.17) \approx 0.985$ ,  $P(Y \leq 1.28) \approx 0.900$ ,  $P(Y \leq 1.09) \approx 0.863$ ,  
 $P(Y \leq 1.04) \approx 0.85$ .

[解析] (1)设考生成绩为  $X$ , 则依题意  $X$  应服从正态分布, 即  $X \sim N(180, \sigma^2)$ .

令  $Y = \frac{X - 180}{\sigma}$ , 则  $Y \sim N(0, 1)$ .

由360分及其以上的高分考生30名可得  $P(X \geq 360) = \frac{30}{2000}$

即  $P(X < 360) = 1 - \frac{30}{2000} \approx 0.985$ , 亦即  $P\left(Y < \frac{360 - 180}{\sigma}\right) \approx 0.985$ .

则  $\frac{360 - 180}{\sigma} = 2.17$ , 解得  $\sigma \approx 83$ ,  $\therefore N(180, 83^2)$ ,

设最低录取分数线为  $x_0$ , 则  $P(X \geq x_0) = P\left(Y \geq \frac{x_0 - 180}{83}\right) = \frac{300}{2000}$

则  $P\left(Y < \frac{x_0 - 180}{83}\right) = 1 - \frac{300}{2000} \approx 0.85$ ,  $\therefore \frac{x_0 - 180}{83} = 1.04 \therefore x_0 \approx 266.32$ .

即最低录取分数线为266分或267分.

(2)考生甲的成绩  $286 > 267$ , 所以能被录取.  $P(X < 286) = P\left(Y < \frac{286 - 180}{83}\right) = P(Y < 1.28) \approx 0.90$ ,

表明不低于考生甲的成绩的人数约为总人数的  $1 - 0.90 = 0.10$ ,  $2000 \times 0.1 \approx 200$ ,

即考生甲大约排在第200名, 排在275名之前, 所以他能获得高薪职位.

**【课后巩固】** \_\_\_\_\_ **完成课时作业 (188)**

## 第十九章 统计

### 19.1 随机抽样

#### 19.1.1 简单随机抽样

##### 【学习目标】

- 1.通过实例，了解简单随机抽样的含义及其解决问题的过程。
- 2.掌握两种随机抽样的方法：抽签法和随机数法。
- 3.会计算样本均值，了解样本与总体的关系。

##### 【要点整合】

##### 1.全面调查与抽样调查以及相关概念

###### (1) 全面调查（普查）

对每一个调查对象都进行调查的方法，称为全面调查，又称普查。

###### (2) 总体、个体

总体：我们把调查对象的全体称为总体。

个体：组成总体的每一个调查对象称为个体。

###### (3) 抽样调查

根据一定目的，从总体中抽取一部分个体进行调查，并以此为依据对总体的情况作出估计和推断的调查方法，称为抽样调查。

###### (4) 样本、样本量、样本数据

样本：我们把从总体中抽取的那部分个体称为样本。

样本量：样本中包含的个体数称为样本量。

样本数据：调查样本获得的变量值称为样本的观测数据，简称样本数据。

##### 2.简单随机抽样的定义

###### (1) 放回简单随机抽样和不放回简单随机抽样

一般地，设一个总体含有有限个（ $N$ 个， $N$ 为正整数）个体，从中逐个抽取 $n$ （ $1 \leq n < N$ ）个个体作为样本，如果抽取是放回的，且每次抽取时总体内的各个个体被抽到的概率都相等，我们把这样的抽样方法叫做放回简单随机抽样；如果抽取是不放回的，且每次抽取时总体内未进入样本的各个个体被抽到的概率都相等，我们把这样的抽样方法叫做不放回简单随机抽样。放回简单随机抽样和不放回简单随机抽样统称为简单随机抽样。

###### (2) 方法：抽签法和随机数法。

##### 3.常用简单随机抽样及方法

###### (1) 抽签法：

用抽签法从容量为 $N$ 的总体中抽取一个容量为 $n$ 的样本的步骤：

①编号：给总体中的所有个体编号(号码可以从1到 $N$ )；

②制作号签：将1~ $N$ 这 $N$ 个号码写在形状、大小相同的号签上(号签可以用小球、卡片、纸条等制作)；

③均匀搅拌：将号签放在一个不透明的容器里，搅拌均匀；

④抽取号码：每次从容器中不放回地抽取一个号签，连续抽取  $n$  次；

⑤构成样本：从总体中将抽到的号签上的号码一致的个体抽取，就构成了一个容量为  $n$  的样本.

(2) 随机数法：用随机试验生成随机数；用信息技术生成随机数：①用计算器生成随机数；②用电子表格软件生成随机数；③用 **R** 统计软件生成随机数.

(3) 随机数表法

①编号：对总体的个体进行编号(每个号码位数一致)；

②选定初始值：在随机数表中任选一个数作为开始；

③选号：从选定的数开始按一定的方向读下去，得到的号码若不在编号中，则跳过，若在编号中，则取出，如果得到的号码前面已经取出，也跳过，如此继续下去，直到取满为止；

④确定样本：根据选定的号码抽取样本.

#### 4. 总体均值与样本均值

(1) 总体均值：一般地，总体中有  $N$  个个体，它们的变量值分别为  $Y_1, Y_2, \dots, Y_N$ ，则称

$$\bar{Y} = \frac{Y_1 + Y_2 + \dots + Y_N}{N} = \frac{1}{N} \sum_{i=1}^N Y_i \text{ 为总体均值，又称总体平均数.}$$

(2) 样本均值：如果从总体中抽取一个容量为  $n$  的样本，它们的变量值分别为  $y_1, y_2, \dots, y_n$ ，则称

$$\bar{y} = \frac{y_1 + y_2 + \dots + y_n}{n} = \frac{1}{n} \sum_{i=1}^n y_i \text{ 为样本均值，又称样本平均数.}$$

(3) 样本均值与总体均值的关系：我们常用样本均值  $\bar{y}$  估计总体均值  $\bar{Y}$ .

### 【典例讲练】

#### 题型一 简单随机抽样的判断

例1 (1) 下列 5 个抽样中，简单随机抽样的个数是( )

①从无数个个体中抽取 50 个个体作为样本；

②仓库中有 1 万支火炬，从中一次性抽取 100 支火炬进行质量检查；

③某连队从 200 名党员官兵中，挑选出 50 名最优秀的官兵赶赴青海参加抗震救灾工作；

④一彩民选号，从装有 36 个大小、形状都相同的号签的盒子中无放回地抽出 6 个号签.

A. 0

B. 1

C. 2

D. 3

[答案]B

(2) 在简单随机抽样中，某一个个体被抽到的可能性( )

A. 与第几次抽样有关，第一次抽到的可能性大一些

B. 与第几次抽样无关，每次抽到的可能性都相等

C. 与第几次抽样有关，最后一次抽到的可能性要大些

D. 与第几次抽样无关，每次都是等可能的抽取，但各次抽取的可能性不一定

[答案]B

归纳小结：简单随机抽样必须具备的特点

(1)被抽取样本的总体中的个体数是有限的.

(2)抽取的样本是从总体中逐个抽取的.

(3)简单随机抽样是一种等可能的抽样.

如果三个特征有一个不满足，就不是简单随机抽样.

【练习1】(1)判断下面的抽样方法是否为简单随机抽样，并说明理由.

①某班 45 名同学，指定个子最矮的 5 名同学参加学校组织的某项活动；

②从 20 个零件中一次性抽出 3 个进行质量检查.

[答案]①不是简单随机抽样. 因为指定个子最矮的 5 名同学，是在 45 名同学中特指的，不存在随机性，不是等可能抽样.

②不是简单随机抽样. 因为一次性抽取 3 个不是逐个抽取，不符合简单随机抽样的特征.

(2)从总体容量为  $N$  的一批零件中，通过简单随机抽样抽取一个容量为 30 的样本，若每个零件被抽到的可能性为 0.25，则  $N$  的值为( )

A.120

B.200

C.150

D.100

[答案]A

## 题型二 用抽签法抽取样本

例2 (1)上海某中学从 40 名学生中选 1 人作为上海男篮啦啦队的成员，采用下面两种选法，则抽签法的序号是\_\_\_\_\_.

①将这 40 名学生从 1~40 进行编号，相应地制作 1~40 的 40 个号签，把这 40 个号签放在一个暗箱中搅匀，然后随机地从中抽取 1 个号签，与这个号签编号一致的学生幸运入选；

②将 39 个白球与 1 个红球(球除颜色外，其他完全相同)混合放在一个暗箱中搅匀，让 40 名学生逐一从中摸取一球，摸到红球的学生成为啦啦队成员.

(2)在社区公益活动中，某单位共有 50 名志愿者参与了报名，现要从中随机抽出 6 人参加一项活动，请用抽签法进行抽样，并写出过程.

[答案] (1)① (2)第一步，将 50 名志愿者编号，号码依次为 1, 2, 3, ..., 50；

第二步，将号码分别写在大小、形状、质地都相同的纸条上，揉成团，制成号签；

第三步，将所有号签放入一个不透明的箱子中，搅拌均匀；

第四步，一次取出 1 个号签，连取 6 次(不放回抽取)，并记录其编号；

第五步，将对应编号的志愿者选出即可.

【练习2】(1)抽签法确保样本具有代表性的关键是( )

A.制签

B.搅拌均匀

C.逐一抽取

D.抽取不放回

[答案]B

(2)使用简单随机抽样从 1 000 件产品中抽出 50 件进行某项检查，合适的抽样方法是( )

A.抽签法

B.随机数法

C.随机抽样法

D.以上都不对

[答案]B

### 题型三 用随机数法抽取样本

例3 要考察某种品牌的850颗种子的发芽率，从中抽取50颗种子进行实验，利用随机数表法抽取种子，先将850颗种子按001,002，…，850进行编号，如果从随机数表第3行第6列的数开始向右读，请依次写出最先检验的4颗种子的编号：\_\_\_\_\_。(下面抽取了随机数表第1行至第8行)

|                |                |                |                |                |
|----------------|----------------|----------------|----------------|----------------|
| 03 47 43 73 86 | 36 96 47 36 61 | 46 98 63 71 62 | 33 26 16 80 45 | 60 11 14 10 95 |
| 97 74 24 67 62 | 42 81 14 57 20 | 42 53 32 37 32 | 27 07 36 07 51 | 24 51 79 89 73 |
| 16 76 62 27 66 | 56 50 26 71 07 | 32 90 79 78 53 | 13 55 38 58 59 | 88 97 54 14 10 |
| 12 56 85 99 26 | 96 96 68 27 31 | 05 03 72 93 15 | 57 12 10 14 21 | 88 26 49 81 76 |
| 55 59 56 35 64 | 38 54 82 46 22 | 31 62 43 09 90 | 06 18 44 32 53 | 23 83 01 30 30 |
| 16 22 77 94 39 | 49 54 43 54 82 | 17 37 93 23 78 | 87 35 20 96 43 | 84 26 34 91 64 |
| 84 42 17 53 31 | 57 24 55 06 88 | 77 04 74 47 67 | 21 76 33 50 25 | 83 92 12 06 76 |
| 63 01 63 78 59 | 16 95 55 67 19 | 98 10 50 71 75 | 12 86 73 58 07 | 44 39 52 38 79 |

[答案]227,665,650,

### 题型四 用样本的平均数估计总体的平均数

例4 为了调查某校高一学生每天午餐消费情况，从该校高一学生中抽查了20名学生，通过调查这20名学生每天午餐消费数据如下(单位：元)

8 10 6 6 8 12 15 6 8 6 10 8 8 15 6 8 10 8 8 10

试估计该校高一学生每天午餐的平均费用，以及午餐费用不低于10元的比例。

[答案]样本的平均数为 $\bar{y} = \frac{8+10+6+6+8+12+15+6+8+6+10+8+8+15+6+8+10+8+8+10}{20} = 8.8$ ,

样本中消费不低于10元的比例为 $\frac{7}{20} = 0.35$ ,

所以估计该校高一全体学生每天午餐的平均费用为8.8元。

在全体学生中，午餐费用不低于10元的比例约为0.35。

【练习3】为了了解某校高三学生每天的作业量，通过简单随机抽样从该校高三学生中抽取了60名学生，通过调查发现这60名学生每天完成作业平均用时2小时，则可以推测该校高三学生每天完成作业所需时间的平均数( )

A.一定为2小时 B.高于2小时 C.低于2小时 D.约为2小时

[答案]D

【课后巩固】\_\_\_\_\_完成课时作业(189)

## 19.1.2 分层随机抽样

### 【学习目标】

- 1.通过实例，了解分层随机抽样的特点和适用范围.
- 2.了解分层随机抽样的必要性，掌握各层样本量比例分配的方法.
- 3.结合具体实例，掌握分层随机抽样的样本均值.

### 【要点整合】

#### 1.分层随机抽样的相关概念

分层随机抽样总体是由差异明显的各层组成的

(1)分层随机抽样的定义：一般地，按一个或多个变量把总体划分成若干个子总体，每个个体属于且仅属于一个子总体，在每个子总体中独立地进行简单随机抽样，再把所有子总体中抽取的样本合在一起作为总样本，这样的抽样方法称为分层随机抽样，每一个子总体称为层.

(2)比例分配：在分层随机抽样中，如果每层样本量都与层的大小成比例，那么称这种样本量的分配方式为比例分配.

#### 2.样本平均数的计算公式

在分层随机抽样中，如果层数分为2层，第1层和第2层包含的个体数分别为 $M$ 和 $N$ ，抽取的样本量分别为 $m$ 和 $n$ ，第1层和第2层样本的平均数分别为 $\bar{x}$ 和 $\bar{y}$ ，

$$\text{则样本的平均数 } \bar{\omega} = \frac{m}{m+n} \bar{x} + \frac{n}{m+n} \bar{y} = \frac{M}{M+N} \bar{x} + \frac{N}{M+N} \bar{y}$$

### 【典例讲练】

#### 题型一 对分层随机抽样概念的理解

例1 (1)某政府机关在编人员共100人，其中副处级以上干部10人，一般干部70人，工人20人，上级部门为了了解该机关对政府机构改革的意见，要从中抽取20人，用下列哪种方法最合适( )

- A.抽签法  
B.随机数  
C.简单随机抽样  
D.分层随机抽样

(2)分层随机抽样又称类型抽样，即将相似的个体归入一类(层)，然后每类抽取若干个个体构成样本，所以分层随机抽样为保证每个个体被等可能抽取，必须进行( )

- A.每层等可能抽样  
B.每层可以不等可能抽样  
C.所有层按同一抽样比等可能抽样  
D.所有层抽取的个体数量相同

[答案](1)D (2)C

[解析](1)总体由差异明显的三部分构成，应选用分层随机抽样.

(2)为了保证每个个体等可能的被抽取，分层随机抽样时必须在所有层都按同一抽样比等可能抽取.

归纳小结 1.使用分层随机抽样的前提

分层随机抽样的总体按一个或多个变量划分成若干个子总体，并且每一个个体属于且仅属于一个子总体，而层内个体间差异较小.

## 2.使用分层随机抽样应遵循的原则

- (1)将相似的个体归入一类,即为一层,分层要求每层的各个个体互不交叉,即遵循不重复、不遗漏的原则;  
(2)分层随机抽样为保证每个个体等可能抽取,需遵循在各层中进行简单随机抽样,每层样本量与每层个体数量的比等于抽样比.

【练习1】下列问题中,最适合用分层随机抽样抽取样本的是( )

- A.从10名同学中抽取3人参加座谈会  
B.某社区有500个家庭,其中高收入的家庭125户,中等收入的家庭280户,低收入的家庭95户,为了了解生活购买力的某项指标,要从中抽取一个容量为100的样本  
C.从1 000名工人中,抽取100人调查上班途中所用时间  
D.从生产流水线上,抽取样本检查产品质量

[答案]B

[解析]A中总体所含个体无差异且个数较少,适合用简单随机抽样;C和D中总体所含个体无差异但个数较多,不适合用分层随机抽样;B中总体所含个体差异明显,适合用分层随机抽样.

## 题型二 分层随机抽样的应用

例2 某学校有在职人员160人,其中行政人员有16人,教师有112人,后勤人员有32人.教育部门为了了解在职人员对学校机构改革的意见,要从中抽取一个容量为20的样本,请利用分层随机抽样的方法抽取,写出抽样过程.

[解析]抽样过程如下:

第一步,确定抽样比,样本量与总体的个体数的比为 $\frac{20}{160} = \frac{1}{8}$ .

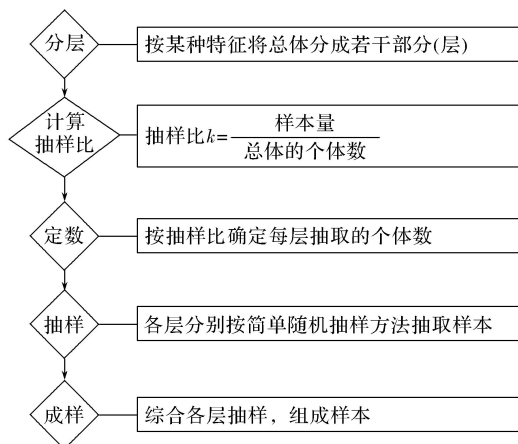
第二步,确定分别从三类人员中抽取的人数,从行政人员中抽取 $16 \times \frac{1}{8} = 2$ (人);

从教师中抽取 $112 \times \frac{1}{8} = 14$ (人);从后勤人员中抽取 $32 \times \frac{1}{8} = 4$ (人).

第三步,采用简单随机抽样的方法,抽取行政人员2人,教师14人,后勤人员4人.

第四步,把抽取的个体组合在一起构成所需样本.

归纳小结3: 分层随机抽样的步骤



**【练习2】** 一个地区共有5个乡镇，人口3万人，其人口比例为3:2:5:2:3，从3万人中抽取一个300人的样本，分析某种疾病的发病率，已知这种疾病与不同的地理位置及水土有关，问应采取什么样的方法？并写出具体过程.

[解析]因为疾病与地理位置和水土均有关系，所以不同乡镇的发病情况差异明显，因而采用分层随机抽样的方法.具体过程如下：第一步，将3万人分为5层，一个乡镇为一层.

第二步，按照抽样比求得各乡镇应抽取的人数分别为60，40，100，40，60.

第三步，采用简单随机抽样的方法，按照各层抽取的人数抽取各乡镇的样本.

第四步，将300人合到一起，即得到一个样本.

### 题型三 分层随机抽样中的计算问题

例3(1)交通管理部门为了解机动车驾驶员(简称驾驶员)对某新法规的知晓情况，对甲、乙、丙、丁四个社区做分层随机抽样调查，假设四个社区驾驶员的总人数为 $N$ ，其中甲社区有驾驶员96人.若在甲、乙、丙、丁四个社区抽取驾驶员的人数分别为12，21，25，43，则这四个社区驾驶员的总人数 $N$ 为( )

- A.101                      B.808                      C.1212                      D.2012

(2)将一个总体分为 $A$ ， $B$ ， $C$ 三层，其个体数之比为5:3:2.若用分层随机抽样方法抽取容量为100的样本，则应从 $C$ 中抽取\_\_\_\_\_个个体.

(3)分层随机抽样中，总体共分为2层，第1层的样本量为20，样本平均数为3，第2层的样本量为30，样本平均数为8，则该样本的平均数为\_\_\_\_\_.

[答案](1)B              (2)20              [(3)6

[解析](1)因为甲社区有驾驶员96人，并且在甲社区抽取的驾驶员的人数为12，

所以抽取驾驶员的抽样比为 $\frac{12}{96} = \frac{1}{8}$ ，所以驾驶员的总人数为 $(12+21+25+43) \div \frac{1}{8} = 808$ .

(2) $\because A$ ， $B$ ， $C$ 三层个体数之比为5:3:2，又由总体中每个个体被抽到的概率相等， $\therefore$ 分层随机抽样应从 $C$ 中抽取 $100 \times \frac{2}{10} = 20$ (个)个体.

(3) $\bar{w} = \frac{20}{20+30} \times 3 + \frac{30}{20+30} \times 8 = 6$ .

**【练习3】** 甲校有3 600名学生，乙校有5 400名学生，丙校有1 800名学生，为统计三校学生某方面的情况，计划采用分层随机抽样法抽取一个容量为90的样本，应在这三校分别抽取学生( )

- A.30人，30人，30人                      B.30人，45人，15人  
C.20人，30人，40人                      D.30人，50人，10人

[答案]B

[解析] 先求抽样比 $\frac{n}{N} = \frac{90}{3600+5400+1800} = \frac{1}{120}$ ，再各层按抽样比分别抽取，甲校抽取 $3\,600 \times \frac{1}{120} = 30$ (人)，

乙校抽取 $5\,400 \times \frac{1}{120} = 45$ (人)，丙校抽取 $1\,800 \times \frac{1}{120} = 15$ (人)，故选B.

**【课后巩固】** \_\_\_\_\_完成课时作业(190)

### 19.1.3 获取数据的途径

#### 【学习目标】

- 1.知道获取数据的基本途径包括：统计报表和年鉴、社会调查、试验设计、普查和抽样、互联网等.
- 2.了解总体、样本、样本量的概念，了解数据的随机性.

#### 【要点整合】

##### 1.获取数据的途径

统计学是通过收集数据和分析数据来认识未知现象的，因此如何收集数据，是统计学研究的重要内容。在实践中，获取数据的途径多种多样，如通过调查获得数据、通过实验获得数据，通过观察获得数据，通过查询获得数据。

##### 2.获得数据的途径的适用类型和注意问题

| 获取数据的基本途径 | 适用类型                          | 注意问题                                      |
|-----------|-------------------------------|-------------------------------------------|
| 通过调查获取数据  | 对于有限总体问题，我们一般通过抽样调查或普查的方法获取数据 | 要充分有效地利用背景信息选择或创建更好的抽样方法，并有效地避免抽样过程中的人为错误 |
| 通过试验获取数据  | 没有现存的数据可以查询                   | 严格控制实验环境，通过精心的设计安排试验，以提高数据质量              |
| 通过观察获取数据  | 自然现象                          | 要通过长久的持续观察获取数据                            |
| 通过查询获得数据  | 众多专家研究过，其收集的数据有所存储            | 必须根据问题背景知识“清洗数据”，去伪存真                     |

思考1：利用统计报表和年鉴属于那中获取数据的途径？属于通过查询获取数据的途径。

思考2：要了解一种新型灯管的寿命，能通过观察获取数据吗？不能，应该通过试验获取数据。

#### 【典例讲练】

##### 题型一 获取数据途径的选择

例1(1)下列哪些数据一般是通过试验获取的( )

- A. 1988年济南市的降雨量
- B. 2019年新生儿人口数量
- C. 某学校高一年级同学的数学测试成绩
- D. 某种特效中成药的配方

(2)“中国天眼”为500米口径球面射电望远镜(Five hundred meters Aperture Spherical Telescope, 简称FAST),是具有我国自主知识产权、世界最大单口径、最灵敏的射电望远镜.建造“中国天眼”的目的是( )

- A. 通过调查获取数据
- B. 通过试验获取数据
- C. 通过观察获取数据
- D. 通过查询获得数据



[答案](1)D (2)C (1)某种特效中成药的配方的数据只能通过试验获得.

(2)“中国天眼”主要是通过观察获取数据.

【练习1】要得到某乡镇的贫困人口数据,应采取的方法是( )

- A. 通过调查获取数据
- B. 通过试验获取数据
- C. 通过观察获取数据
- D. 通过查询获得数据

[答案]A

## 题型二 获取数据途径的方法的设计

例2 为了缓解城市的交通拥堵情况,某市准备出台限制私家车的政策,为此要进行民意调查.某个调查小组调查了一些拥有私家车的市民,你认为这样的调查结果能很好地反映该市市民的意愿吗?

[解析](1)一个城市的交通状况的好坏将直接影响着生活在这个城市中的每个人,关系到每个人的利益.

为了调查这个问题,在抽样时应当关注到各种人群,既要抽到拥有私家车的市民,也要抽到没有私家车的市民.

(2)调查时,如果只对拥有私家车的市民进行调查,结果一定是片面的,不能代表所有市民的意愿.因此,在调查时,要对生活在该城市的所有市民进行随机地抽样调查,不要只关注到拥有私家车的市民.

【练习2】一些期刊杂志社经常会请一些曾经高考落榜而在某方面的事业上取得成就的著名专家、学者,谈他们对高考落榜的看法,这些名人所讲的都是大同小异,不外乎“我也有过落榜的沮丧,但从长远看,它有益于我的人生”,“我是因祸得福,落榜使我走了另一条成功之路”等等.小明据此得出一条结论,上大学不如高考落榜,他的结论正确吗?

[解析]小明的结论是错误的,在众多的高考落榜生中,走出另外一条成功之路的是少数,小明通过研究一些期刊杂志社报道过的一些成功人士就得出结论是片面的,因为他的抽样不具有代表性.

**【课后巩固】**—————**完成课时作业(191)**

## 19.2 用样本估计总体

### 19.2.1 总体取值规律的估计（第一课时）

#### 【学习目标】

1. 理解用样本的频率分布估计总体分布的思想方法.
2. 会列频率分布表, 会画频率分布直方图.

#### 【要点整合】

- 1、制作频率分布表、画频率分布直方图的步骤

(1)求极差: 即一组数据中最大值与最小值的差.

(2)决定组距和组数: ①当样本容量不超过100时, 按照数据的多少常分为5~12组; ②组数  $k = \frac{\text{极差}}{\text{组距}}$  (一般取等长组距)

(3)将数据分组: 前面各组均为左闭右开区间, 最后一组是闭区间.

(4)列频率分布表: 一般分四列: 分组、频数累计、频数、频率, 最后一行是合计, 其中频数合计应是样本容量, 频率合计应是1.

(5)画频率分布直方图: 在频率分布直方图中, 纵轴表示  $\frac{\text{频率}}{\text{组距}}$ , 数据落在各小组内的频率用各小长方形的

面积来表示, 小长方形的面积  $= \text{组距} \times \frac{\text{频率}}{\text{组距}} = \text{频率}$ . 各小长方形的面积的总和等于1.

- 2、其他常见统计图

条形图、扇形图、折线图和频数分布直方图等.

扇形图主要用于直观描述各类数据占总数的比.

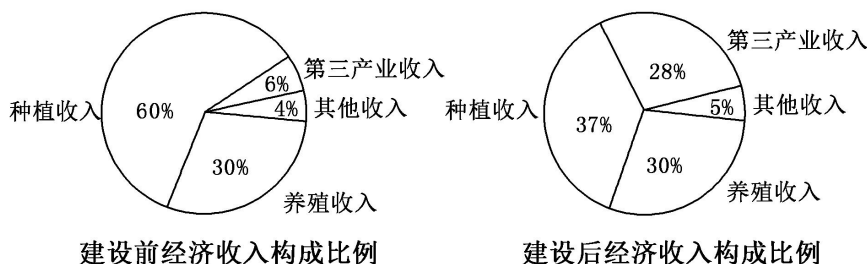
条形图(适用于描述离散型数据)和直方图(适用于描述连续型数据)主要用于描述直观描述不同类别或分组数据的频数和频率.

折线图主要用于描述数据随时间的变化趋势.

#### 【典例讲练】

##### 题型一 扇形图

例1 某地区经过一年的新农村建设, 农村的经济收入增加了一倍, 实现翻番. 为更好地了解该地区农村的经济收入变化情况, 统计了该地区新农村建设前后农村的经济收入构成比例, 得到如图所示的饼图, 则下面结论中不正确的是( )

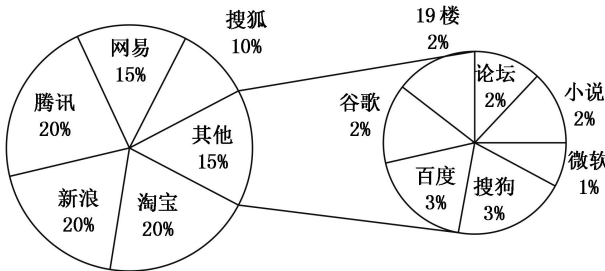


- A. 新农村建设后，种植收入减少
- B. 新农村建设后，其他收入增加了一倍以上
- C. 新农村建设后，养殖收入增加了一倍
- D. 新农村建设后，养殖收入与第三产业收入的总和超过了经济收入的一半

[答案]A

[解析]设新农村建设前的经济收入为 $M$ ，而新农村建设后的经济收入为 $2M$ ，则新农村建设前种植收入为 $0.6M$ ，而新农村建设后的种植收入为 $0.74M$ ，所以种植收入增加了，所以A项不正确。

【练习1】如图给出的是某小区居民一段时间内访问网站的比例图，则下列选项中不超过21%的为( )



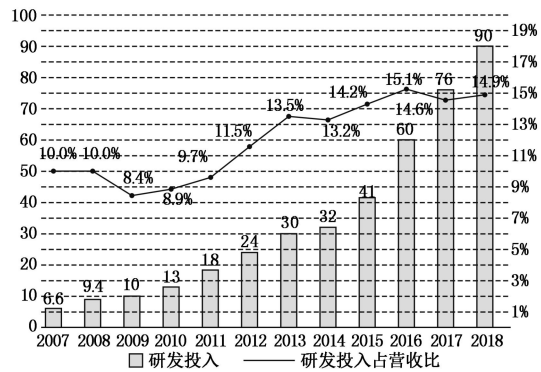
- A. 网易与搜狗的访问量所占比例之和
- B. 腾讯与百度的访问量所占比例之和
- C. 淘宝与论坛的访问量所占比例之和
- D. 新浪与小说的访问量所占比例之和

[答案]A

[解析]由题图得，网易与搜狗的访问量所占比例分别为15%和3%，总和为18%，不超过21%，满足题意。故选A。

## 题型二 折线图与条形图

例2(1)“科技引领，布局未来”科技研发是企业发展的驱动力量。2007年至2018年，某企业连续12年累计研发投入达4100亿元。我们将研发投入与经营收入的比值记为研发投入占营收比。这12年间的研发投入(单位：十亿元)用如图所示的条形图表示，研发投入占营收比用图中的折线图表示，根据折线图和条形图，下列结论错误的是( )

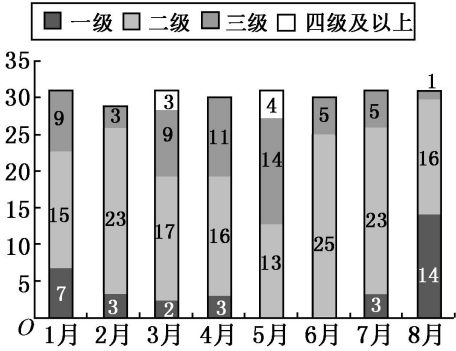


- A. 2012年至2013年研发投入占营收比增量相比2017年至2018年增量小
- B. 2013年至2014年研发投入增量相比2015年至2016年增量小
- C. 该企业连续12年研发投入逐年增加
- D. 该企业连续12年来研发投入占营收比逐年增加

[答案]D

[解析]对于 A, 2012 年至 2013 年研发投入占营收比增量为  $13.5\% - 11.5\% = 2\%$ , 2017 年至 2018 年研发投入占营收比增量为  $14.9\% - 14.6\% = 0.3\%$ , 正确; 对于 B, 2013 年至 2014 年研发投入增量为  $32 - 30 = 2$  (十亿元), 2015 年至 2016 年研发投入增量为  $60 - 41 = 19$  (十亿元), 正确; 对于 C, 由图易知该企业连续 12 年研发投入逐年增加, 正确; 对于 D, 由图知 2008 年至 2009 年研发投入占营收比是减少的, 错误, 故选 D。

(2) 随着人民生活水平的提高, 对城市空气质量的关注度也逐渐增高, 如图是某城市 1 月至 8 月的空气质量检测情况, 图中一、二、三、四级是空气质量等级, 一级空气质量最好, 一级和二级都是质量合格空气, 下面叙述不正确的是( )

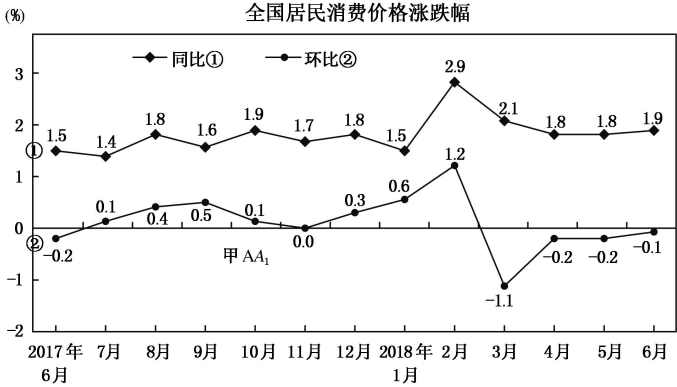


- A. 1 月至 8 月空气质量合格天数超过 20 天的月份有 5 个
- B. 第二季度与第一季度相比, 空气质量达标天数的比重下降了
- C. 8 月份是空气质量最好的一个月
- D. 6 月份的空气质量最差

[答案]D

[解析]一级和二级都是质量合格空气, 观察统计图可以看出, 1 月、2 月、6 月、7 月、8 月这 5 个月的空气质量合格天数均超过了 20 天, 故选项 A 叙述正确; 1 月、2 月、3 月相比于 4 月、5 月、6 月, 整体上空气质量较好, 故选项 B 叙述正确; 8 月份的空气质量合格天数为 30 天, 且一级达到了 14 天, 所以 8 月是 1 月至 8 月中空气质量最好的一个月, 故选项 C 叙述正确; 5 月份空气质量合格的只有 13 天, 四级及以上甚至有 4 天, 所以 5 月是 1 月至 8 月中空气质量最差的一个月, 所以选项 D 叙述错误。

【练习 2】(1) 如图为国家统计局发布的 2018 年上半年全国居民消费价格指数(CPI)数据折线图, (注: 同比是今年第  $n$  个月与去年第  $n$  个月之比, 环比是现在的统计周期和上一个统计周期之比。)



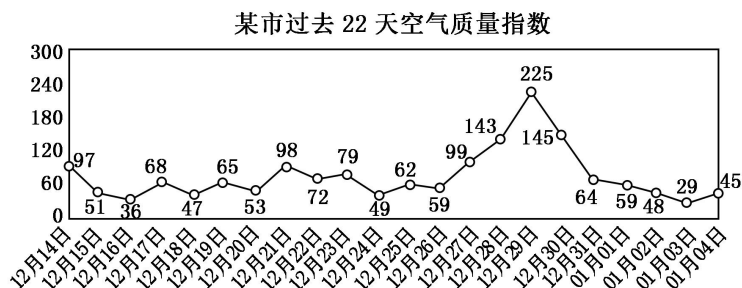
下列说法错误的是( )

- A. 2018 年 6 月 CPI 环比下降 0.1%，同比上涨 1.9%
- B. 2018 年 3 月 CPI 环比下降 1.1%，同比上涨 2.1%
- C. 2018 年 2 月 CPI 环比上涨 0.6%，同比上涨 1.4%
- D. 2018 年 6 月 CPI 同比涨幅比上月略微扩大 0.1 个百分点

[答案]C

[解析]观察题图中数据知 A, B, D 正确, 对于选项 C, 2018 年 2 月 CPI 环比上涨 1.2%，同比上涨 2.9%，故 C 错误。故选 C。

(2) 空气质量指数(简称: AQI)是定量描述空气质量状况的指数, 空气质量按照 AQI 大小分为六级:  $[0, 50]$  为优,  $(50, 100]$  为良,  $(100, 150]$  为轻度污染,  $(150, 200]$  为中度污染,  $(200, 300]$  为重度污染, 300 以上为严重污染。如图记录了某市过去 22 天的空气质量指数, 据此判断下列结论错误的是( )



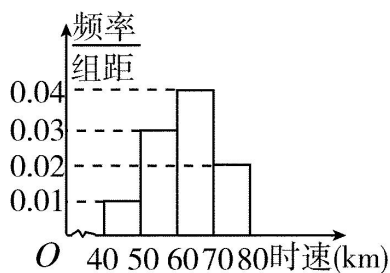
- A. 在该市这 22 天的空气质量中, 从空气质量指数的平均数来看, 最后 4 天的空气质量优于最前面 4 天的空气质量
- B. 在该市这 22 天的空气质量中, 有 3 天达到污染程度
- C. 在该市这 22 天的空气质量中, 12 月 29 日的空气质量最好
- D. 在该市这 22 天的空气质量中, 空气质量为优的天数为 6

[答案]C

[解析]因为该市最前面 4 天的空气质量指数的平均数为  $\frac{1}{4} \times (97 + 51 + 36 + 68) = 63$ , 最后 4 天的空气质量指数的平均数为  $\frac{1}{4} \times (59 + 48 + 29 + 45) = 45.25$ , 所以在该市这 22 天的空气质量中, 从空气质量指数的平均数来看, 最后 4 天的空气质量优于最前面 4 天的空气质量, 即选项 A 中结论正确; AQI 高于 100 的数据有 3 个: 143, 225, 145, 所以在该市这 22 天的空气质量中, 有 3 天达到污染程度, 即选项 B 中结论正确; 12 月 29 日的 AQI 为 225, 为重度污染, 该天的空气质量最差, 即选项 C 中结论错误; AQI 在  $[0, 50]$  内的数据有 6 个: 36, 47, 49, 48, 29, 45, 即空气质量为优的天数为 6, 即选项 D 中结论正确。

### 题型三 频率分布直方图

例 3 (1) 已知 200 辆汽车通过某一段公路时的时速的频率分布直方图如图所示, 求时速在  $[60, 70]$  的汽车大约有\_\_\_\_\_辆.



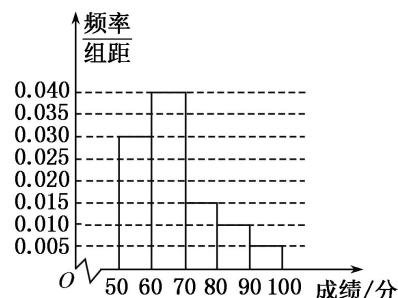
[答案]80

[解析]由图可知, 时速在 $[60, 70)$ 的汽车的频率为 $0.04 \times 10 = 0.4$ ,

$\therefore$  时速在 $[60, 70)$ 的汽车大约有  $200 \times 0.4 = 80$  辆.

(2)在某中学举行的环保知识竞赛中, 将三个年级参赛学生的成绩进行整理后分为5组, 绘制如图所示的频率分布直方图, 图中从左到右依次为第一、第二、第三、第四、第五小组, 已知第二小组的频数是40, 则成绩在80~100分的学生人数是( )

- A. 15                      B. 18                      C. 20                      D. 25



[答案]A

[解析]根据频率分布直方图, 得第二小组的频率是 $0.04 \times 10 = 0.4$ , 因为频数是40,

所以样本容量是 $\frac{40}{0.4} = 100$ , 又成绩在80~100分的频率是 $(0.010 + 0.005) \times 10 = 0.15$ , 所以成绩在80~100分的学生人数是 $100 \times 0.15 = 15$ . 故选 A.

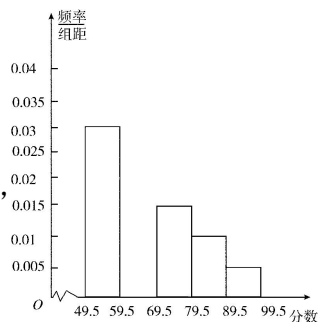
(3)在英才中学举行的信息知识竞赛中, 将高二年级两个班的参赛学生成绩(得分均为整数)进行整理后分成五组, 绘制如图所示的频率分布直方图. 已知图中从左到右的第一、第三、第四、第五小组的频率分别是0.30, 0.15, 0.10, 0.05, 第二小组的频数是40.

(1)求第二小组的频率, 并补全这个频率分布直方图;

(2)求这两个班参赛学生的总人数是多少?

[解析](1)因为各小组的频率之和为1, 第一、三、四、五小组的频率分别是0.30, 0.15, 0.10, 0.05,

所以第二小组的频率为 $1 - (0.30 + 0.15 + 0.10 + 0.05) = 0.40$ .

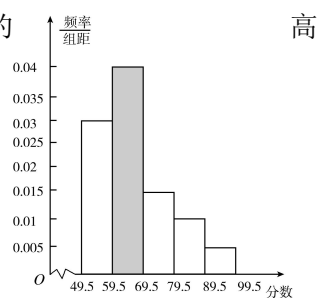


因为第二小组的频率为0.40, 所以落在59.5~69.5内的第二小组的小长方形的

为 $\frac{\text{频率}}{\text{组距}} = \frac{0.40}{10} = 0.04$ , 由此可补全频率分布直方图(如图阴影部分所示).

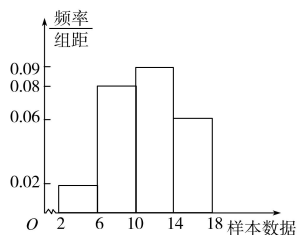
(2)设高二年级两个班参赛学生的总人数为 $x$ 人, 因为第二小组的频数为40,

频率为0.40, 所以 $\frac{40}{x} = 0.40$ , 解得 $x = 100$  (人).



所以高二年级两个班参赛学生的总人数为100人.

【练习 3】(1)如图是容量为100的样本的频率分布直方图，试根据图中的数据填空.



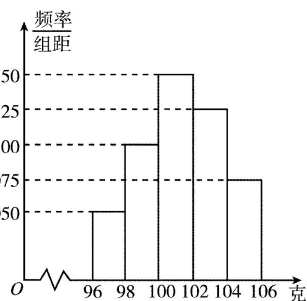
①样本数据落在范围 $[6,10)$ 内的频率为\_\_\_\_\_;

②样本数据落在范围 $[10,14)$ 内的频数为\_\_\_\_\_; ③总体在 $[2,6)$ 的概率约为\_\_\_\_\_.

[答案]①0.32 ②36 ③0.08

[解析] 在频率分布直方图中,用小矩形的面积表示频率,即 $4 \times 0.08 = 0.32$ , 频数 $= 4 \times 0.09 \times 100 = 36$ ,  $0.02 \times 4 = 0.08$ , 用样本的频率估计总体的概率.

(2)某工厂对一批产品进行了抽样检测. 下图是根据抽样检测后的产品净重(单位克)数据绘制的频率分布直方图, 其中产品净重的范围是 $[96,106]$ , 样本数据分组为 $[96,98)$ ,  $[98,100)$ ,  $[100,102)$ ,  $[102,104)$ ,  $[104,106]$ . 已知样本中产品净重小于100克的个数是36, 则样本中净重大于或等于98克并且小于104克的产品的个数是( )



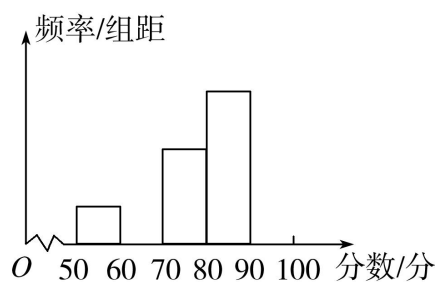
A. 90                      B. 75                      C. 60                      D. 45

[答案]A

[解析]产品净重小于100克的频率为 $(0.050 + 0.100) \times 2 = 0.300$ , 已知样本中产品净重小于100克的个数是36, 设样本容量为 $n$ , 则 $\frac{36}{n} = 0.300$ , 所以 $n = 120$ , 净重大于或等于98克并且小于104克的产品的个数是 $120 \times 0.75 = 90$ .

(3)为了让学生了解环保知识, 增强环保意识, 某中学举行了一次环保知识竞赛, 共有900名学生参加了这次竞赛. 为了了解本次竞赛的成绩情况, 从中抽取了部分学生的成绩(得分取正整数, 满分为100分)进行统计. 请你根据下面尚未完成的频率分布表和频率分布直方图, 解答下列问题:

| 分组         | 频数 | 频率   |
|------------|----|------|
| $[50,60)$  | 4  | 0.08 |
| $[60,70)$  | 8  | 0.16 |
| $[70,80)$  | 10 | 0.20 |
| $[80,90)$  | 16 | 0.32 |
| $[90,100]$ |    |      |
| 合计         |    |      |



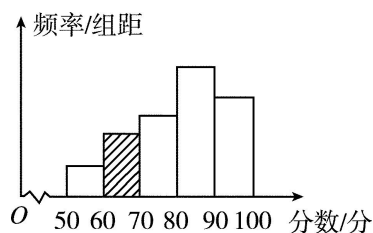
(1)填充频率分布表中的空格；(2)不具体计算频率/组距，补全频率分布直方图.

[解析](1) $\frac{4}{0.08} = 50$ ，即样本容量为50.第五小组的频数为 $50 - 4 - 8 - 10 - 16 = 12$ ，

第五小组的频率为 $\frac{12}{50} = 0.24$ .又各小组频率之和为1，

所以频率分布表中的四个空格应分别填12，0.24，50，1.

(2)根据小长方形的高与频数成正比，设第一个小长方形的高为 $h_1$ ，第二个小长方形的高为 $h_2$ ，第五个小长方形的高为 $h_5$ .由等量关系，得 $\frac{h_1}{h_2} = \frac{4}{8}$ ， $\frac{h_1}{h_5} = \frac{4}{12}$ .所以 $h_2 = 2h_1$ ， $h_5 = 3h_1$ .这样即可补全频率分布直方图如下：



**【课后巩固】** \_\_\_\_\_ **完成课时作业（192）**

## 19.2.2 总体取值规律的估计（第二课时）

### 【学习目标】

- 1、理解茎叶图的概念，会画茎叶图
- 2、通过茎叶图分析统计数据

### 【要点整合】

#### 1、茎叶图

统计中还有一种被用来表示数据的图叫做茎叶图，茎是指中间的一列数，叶是从茎的旁边生长出来的数．一般情况下茎按从小到大的顺序从上向下列出，共茎的叶同行列出．

2、制作茎叶图的步骤：将每个样本数据分为“茎”和“叶”两部分．若是一位数，则它是叶，它的茎为0；若是两位数，则一般个位上的数字为叶，十位上的数字为茎；若是三位数，则个位上的数字为叶，十位和百位上的两个数字为茎；若是小数的，通常把整数部分作为茎，小数部分作为叶，解题时要根据数据的特点合理选择茎和叶．

#### 3、用茎叶图刻画数据有两个优点：

一是所有的信息都可以从茎叶图中得到；二是茎叶图便于记录和表示，能够展示数据的分布情况．但当样本数据较多或数据位数较多时，茎叶图就显得不太方便了．

### 【典例讲练】

#### 题型一 识读茎叶图

例 1(1)以下茎叶图记录了甲、乙两组各五名学生在一次英语听力测试中的成绩(单位：分)．

已知甲组数据的中位数为15，乙组数据的平均数为16.8，则  $x$ ， $y$  的值分别为( )

| 甲组    |   |  | 乙组      |  |
|-------|---|--|---------|--|
| 9     | 0 |  | 9       |  |
| $x$ 2 | 1 |  | 5 $y$ 8 |  |
| 7 4   | 2 |  | 4       |  |

- A. 2，5                      B. 5，5                      C. 5，8                      D. 8，8

[答案]C

[解析]由甲组数据中位数为15，可得  $x=5$ ；而乙组数据的平均数  $16.8 = \frac{9+15+(10+y)+18+24}{5}$ ，

可解得  $y=8$ ．故选 C．

(2)甲、乙两组数据用茎叶图表示如图，中间一列的数字表示该数据的十位数，两边的数字表示该数据的个位数，则甲组数据的平均数是\_\_\_\_\_，乙组数据的中位数是\_\_\_\_\_．

| 甲 |   |   |   |   |   | 乙 |   |   |   |  |
|---|---|---|---|---|---|---|---|---|---|--|
|   |   |   |   |   | 1 | 9 | 7 | 1 |   |  |
| 0 | 1 | 3 | 2 | 0 | 2 | 1 | 4 | 2 | 4 |  |
|   |   |   |   |   | 3 | 0 | 2 | 0 |   |  |

[答案]24    23

[解析]  $\bar{x}_{\text{甲}} = \frac{18+19+20+22+23+21+20+35+31 \times 2}{10} = 24$ ，又乙组数据中间两位是 24，22，

故中位数是 23.

【练习 1】(1)如图是根据某校 10 名高一学生的身高(单位: cm)数据画出的茎叶图，其中左边的数字从左到右分别表示学生身高的百位数字和十位数字，右边的数字表示学生身高的个位数字，则这 10 名学生身高数据的中位数是( )

|    |         |
|----|---------|
| 15 | 5 5 7 8 |
| 16 | 1 3 3 5 |
| 17 | 1 2     |

- A. 161                      B. 162                      C. 163                      D. 164

[答案]B

[解析]由给定的茎叶图可知，这 10 名学生身高数据的中位数为  $\frac{161+163}{2} = 162$ .

(2)某苗圃基地为了解基地内甲、乙两块地种植的同一种树苗的长势情况，从两块地各随机抽取了 10 株树苗，用茎叶图表示上述两组数据，对两块地抽取树苗的高度的平均数  $\bar{x}_{\text{甲}}$ ， $\bar{x}_{\text{乙}}$  和中位数  $y_{\text{甲}}$ ， $y_{\text{乙}}$  进行比较，下面结论正确的是( )

|           |   |         |   |
|-----------|---|---------|---|
|           | 甲 |         | 乙 |
|           | 9 | 1       | 0 |
| 9 5 3 1 0 | 2 | 6       |   |
| 7 3 2     | 3 | 0 0 4 7 |   |
| 1         | 4 | 4 6 6 7 |   |

- A.  $\bar{x}_{\text{甲}} > \bar{x}_{\text{乙}}$ ， $y_{\text{甲}} > y_{\text{乙}}$                       B.  $\bar{x}_{\text{甲}} < \bar{x}_{\text{乙}}$ ， $y_{\text{甲}} < y_{\text{乙}}$   
C.  $\bar{x}_{\text{甲}} < \bar{x}_{\text{乙}}$ ， $y_{\text{甲}} > y_{\text{乙}}$                       D.  $\bar{x}_{\text{甲}} > \bar{x}_{\text{乙}}$ ， $y_{\text{甲}} < y_{\text{乙}}$

[答案]B

[解析]由茎叶图得  $\bar{x}_{\text{甲}} = \frac{19+20+21+23+25+29+32+33+37+41}{10} = 28$ ，

$$\bar{x}_{\text{乙}} = \frac{10+26+30+30+34+37+44+46+46+47}{10} = 35，$$

$$y_{\text{甲}} = \frac{25+29}{2} = 27，y_{\text{乙}} = \frac{34+37}{2} = 35.5，\therefore \bar{x}_{\text{甲}} < \bar{x}_{\text{乙}}，y_{\text{甲}} < y_{\text{乙}}，\text{故选 B.}$$

## 题型二 茎叶图的应用

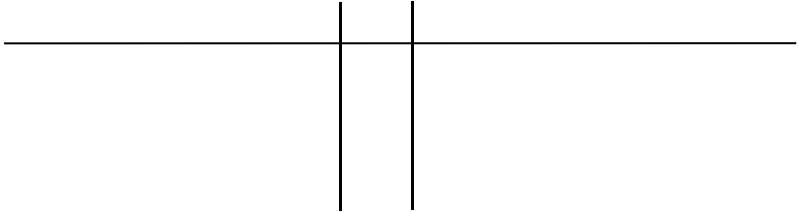
例 2 为了比较两种治疗失眠症的药(分别称为 A 药，B 药)的疗效，随机地选取 20 位患者服用 A 药，20 位患者服用 B 药，这 40 位患者服用一段时间后，记录他们日平均增加的睡眠时间(单位: h)，试验的观测结果如下: 服用 A 药的 20 位患者日平均增加的睡眠时间:

0.6    1.2    2.7    1.5    2.8    1.8    2.2    2.3    3.2    3.5  
2.5    2.6    1.2    2.7    1.5    2.9    3.0    3.1    2.3    2.4

服用  $B$  药的 20 位患者日平均增加的睡眠时间：

3.2    1.7    1.9    0.8    0.9    2.4    1.2    2.6    1.3    1.4  
1.6    0.5    1.8    0.6    2.1    1.1    2.5    1.2    2.7    0.5

- (1) 分别计算两组数据的平均数，从计算结果看，哪种药的疗效更好？  
(2) 根据两组数据完成下面茎叶图，从茎叶图看，哪种药的疗效更好？



[解析] (1) 设  $A$  药观测数据的平均数为  $\bar{x}$ ， $B$  药观测数据的平均数为  $\bar{y}$ ，又观测结果可得

$$\bar{x} = \frac{1}{20}(0.6+1.2+1.2+1.5+1.5+1.8+2.2+2.3+2.3+2.4+2.5+2.6+2.7+2.7+2.8+3.0+3.1+3.2+3.5) = 2.3,$$

$$\bar{y} = \frac{1}{20}(0.5+0.5+0.6+0.8+0.9+1.1+1.2+1.2+1.3+1.4+1.6+1.7+1.8+1.9+2.1+2.4+2.5+2.6+2.7+3.2) = 1.6, \quad \bar{x} > \bar{y},$$

因此可看出  $A$  药的疗效更好。

(2) 由观测结果可绘制如下茎叶图：

| $A$ 药               |    | $B$ 药             |
|---------------------|----|-------------------|
| 6                   | 0. | 5 5 6 8 9         |
| 8 5 5 2 2           | 1. | 1 2 2 3 4 6 7 8 9 |
| 9 8 7 7 6 5 4 3 3 2 | 2. | 1 4 5 6 7         |
| 5 2 1 0             | 3. | 2                 |

从以上茎叶图可以看出， $A$  药疗效的试验结果有的叶集中在茎 2.3 上，而  $B$  药疗效的试验结果有  $\frac{7}{10}$  的叶集中在茎 0，1 上，由此可看出  $A$  药的疗效更好。

【练习 2】下面茎叶图是甲、乙两名同学自高中以来的数学成绩．中间表示十位或百位数字，两边表示个位数字．回答下面问题：

| 甲       |    | 乙         |
|---------|----|-----------|
| 5       | 6  |           |
| 6 5 1   | 7  | 9         |
| 9 8 6 1 | 8  | 3 6 8     |
| 5 4 1   | 9  | 3 7 8 8 9 |
| 7       | 10 | 1 2 3     |
| 0       | 11 | 4         |

- ①写出甲、乙两组数据；②写出两组数据的中位数，并且分析两组数据的稳定性．

[解析]①甲的得分：65，71，75，76，81，86，88，89，91，94，95，107，110.

乙的得分：79，83，86，88，93，97，98，98，99，101，102，103，114.

②由茎叶图可以看出乙同学得分情况大致对称，中位数是98.较集中在90多分.甲同学得分也大致对称，中位数是88，较集中在80多分，因此乙同学成绩较好，发挥比较稳定.

**【课后巩固】**—————**完成课时作业（193）**

### 19.2.3 总体百分位数的估计

#### 【学习目标】

- 1、能用样本估计百分位数，理解百分位数的统计含义。
- 2、会把实际问题转化为数学问题，并进行计算，对数据分析。

#### 【要点整合】

##### 1.第 $p$ 百分位数的定义

一般地，一组数据的第  $p$  百分位数是这样一个值，它使得这组数据中至少有  $p\%$  的数据小于或等于这个值，且至少有  $(100-p)\%$  的数据大于或等于这个值。

第 50 百分位数就是中位数，中位数是百分位数的特例。

##### 2.计算一组 $n$ 个数据的第 $p$ 百分位数的步骤

第 1 步，按从小到大排列原始数据

第 2 步，计算  $i = n \times p\%$ 。

第 3 步，若  $i$  不是整数，而大于  $i$  的比邻整数为  $j$ ，则第  $p$  百分位数为第  $j$  项数据；若  $i$  是整数，则第  $p$  百分位数为第  $i$  项与第  $(i+1)$  项数据的平均数。

##### 3.四分位数

第 25 百分位数，第 50 百分位数，第 75 百分位数，这三个分位数把一组由小到大排列后的数据分成四等份，因此称为四分位数，其中第 25 百分位数也称为第一四分位数或下四分位数，第 75 百分位数也称为第三四分位数或上四分位数。

#### 【典例讲练】

##### 题型一 百分位数的计算

例 1 从某珍珠公司生产的产品中，任意抽取 12 颗珍珠，得到它们的质量(单位：g)如下：

7.9, 9.0, 8.9, 8.6, 8.4, 8.5, 8.5, 8.5, 9.9, 7.8, 8.3, 8.0. (1)分别求出这组数据的第 25, 50, 95 百分位数；(2)请你找出珍珠质量较小的前 15% 的珍珠质量。

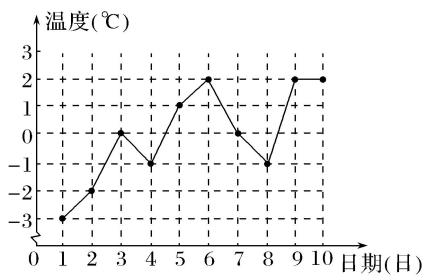
[答案] (1) 第 25 百分位数为 8.15，第 50 百分位数为 8.5，第 95 百分位数为 9.9.

(2) 质量较小的前 15% 的珍珠质量分别为 7.8, 7.9.

[解析] (1) 将所有数据从小到大排列，得 7.8, 7.9, 8.0, 8.3, 8.4, 8.5, 8.5, 8.5, 8.6, 8.9, 9.0, 9.9，因为共有 12 个数据，所以  $12 \times 25\% = 3$ ， $12 \times 50\% = 6$ ， $12 \times 95\% = 11.4$ ，则第 25 百分位数是  $\frac{8.0+8.3}{2} = 8.15$ ，第 50 百分位数是  $\frac{8.5+8.5}{2} = 8.5$ ，第 95 百分位数是第 12 个数据为 9.9.

(2) 因为共有 12 个数据，所以  $12 \times 15\% = 1.8$ ，则第 15 百分位数是第 2 个数据为 7.9. 即产品质量较小的前 15% 的产品有 2 个，它们的质量分别为 7.8, 7.9.

【练习 1】如图所示是某市 3 月 1 日至 3 月 10 日的最低气温(单位： $^{\circ}\text{C}$ )的情况绘制的折线统计图，由图可知这 10 天最低气温的第 80 百分位数是( )



A. -2

B. 0

C. 1

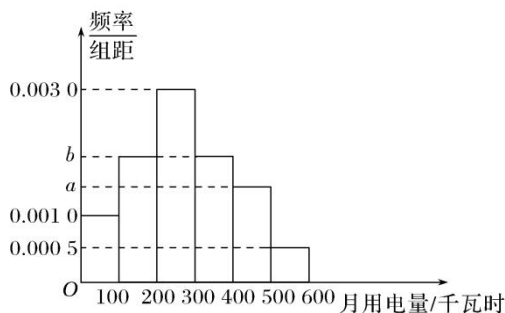
D. 2

[答案]D

[解析]由折线图可知,这10天的最低气温按照从小到大的排列为:  $-3, -2, -1, -1, 0, 0, 1, 2, 2, 2$ , 因为有10个数据,所以  $10 \times 80\% = 8$ , 是整数,则这10天最低气温的第80百分位数是  $\frac{2+2}{2} = 2$ .

## 题型二 百分位数的综合应用

例2 某市为了鼓励居民节约用电,实行“阶梯式”电价,将该市每户居民的月用电量划分为三档,月用电量不超过200千瓦时的部分按0.5元/千瓦时收费,超过200千瓦时但不超过400千瓦时的部分按0.8元/千瓦时收费,超过400千瓦时的部分按1.0元/千瓦时收费。



(1)求某户居民用电费用  $y$  (单位: 元)关于月用电量  $x$  (单位: 千瓦时)的函数解析式;

(2)为了了解居民的用电情况,通过抽样获得了今年1月份100户居民每户的用电量,统计分析后得到如图所示的频率分布直方图.若这100户居民中,今年1月份用电费用低于260元的占80%,求  $a, b$  的值;

(3)根据(2)中求得的数据计算用电量的75%分位数.

[解析](1)当  $0 \leq x \leq 200$  时,  $y = 0.5x$ ; 当  $200 < x \leq 400$  时,  $y = 0.5 \times 200 + 0.8 \times (x - 200) = 0.8x - 60$ ; 当  $x > 400$  时,  $y = 0.5 \times 200 + 0.8 \times 200 + 1.0 \times (x - 400) = x - 140$ .

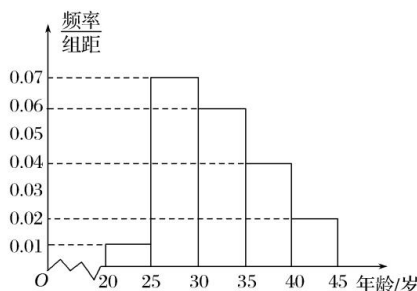
所以  $y$  与  $x$  之间的函数解析式为  $y = \begin{cases} 0.5x, & 0 \leq x \leq 200, \\ 0.8x - 60, & 200 < x \leq 400, \\ x - 140, & x > 400. \end{cases}$

(2)由(1)可知,当  $y = 260$  时,  $x = 400$ , 即用电量低于400千瓦时的占80%,

结合频率分布直方图可知  $\begin{cases} 0.001 \times 100 + 2 \times 100b + 0.003 \times 100 = 0.8, \\ 100a + 0.0005 \times 100 = 0.2, \end{cases}$  解得  $a = 0.0015, b = 0.0020$ .

(3) 设 75% 分位数为  $m$ ，因为用电量低于 300 千瓦时的所占比例为  $(0.001+0.002+0.003)\times 100=60\%$ ，用电量低于 400 千瓦时的占 80%，所以 75% 分位数  $m$  在  $[300,400)$  内，所以  $0.6+(m-300)\times 0.002=0.75$ ，解得  $m=375$  (千瓦时)，即用电量的 75% 分位数为 375 千瓦时。

【练习 2】某市为了了解人们对“中国梦”的伟大构想的认知程度，对不同年龄和不同职业的人举办了一次“一带一路”知识竞赛，满分 100 分(90 分及以上为认知程度高)，现从参赛者中抽取了  $x$  人，按年龄分成 5 组(第一组： $[20,25)$ ，第二组： $[25,30)$ ，第三组： $[30,35)$ ，第四组： $[35,40)$ ，第五组： $[40,45]$ )，得到如图所示的频率分布直方图，已知第一组有 5 人。



- (1) 求  $x$ ；
- (2) 求抽取的  $x$  人的年龄的 50% 分位数(结果保留整数)；
- (3) 以下是参赛的 10 人的成绩：90，96，97，95，92，92，98，88，96，99，求这 10 人成绩的 20% 分位数和平均数，以这两个数据为依据，评价参赛人员对“一带一路”的认知程度。

[解析](1) 第一组频率为  $0.01\times 5=0.05$ ，所以  $x=\frac{5}{0.05}=100$ 。

(2) 由图可知年龄低于 30 岁的所占比例为 40%，年龄低于 35 岁的所占比例为 70%，所以抽取的  $x$  人的年龄的 50% 分位数在  $[30,35)$  内，由  $30+5\times\frac{0.50-0.40}{0.70-0.40}=\frac{95}{3}\approx 32$ ，所以抽取的  $x$  人的年龄的 50% 分位数为 32。

(3) 把参赛的 10 人的成绩按从小到大的顺序排列：88，90，92，92，95，96，96，97，98，99，计算  $10\times 20\%=2$ ，所以这 10 人成绩的 20% 分位数为  $\frac{90+92}{2}=91$ ，

这 10 人成绩的平均数为  $\frac{1}{10}(88+90+92+92+95+96+96+97+98+99)=94.3$ 。

评价：从第 20 百分位数和平均数来看，参赛人员的认知程度很高。

【课后巩固】\_\_\_\_\_完成课时作业 (194)

## 19.2.4 总体集中趋势的估计

### 【学习目标】

1. 理解众数、中位数、平均数的概念。
2. 会求频率分布直方图的众数，中位数，平均数。
3. 了解平均数，中位数，众数在描述数据时的差异，能灵活运用这三个数据代表解决实际问题。

### 【要点整合】

#### 1、众数、中位数、平均数

(1)众数：在一组数据中，出现次数最多的数叫这组数据的众数。

(2)中位数：把一组数据按从小到大的顺序排列，把处于最中间位置的数叫中位数。

①当数据个数为奇数时，中位数是按从小到大顺序排列的中间那个数。

②当数据个数为偶数时，中位数是按从小到大顺序排列的中间两个数的平均数。

(3)平均数：如果有  $n$  个数  $x_1, x_2, \dots, x_n$ ，那么  $\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$ ，叫做这  $n$  个数的平均数。

#### 2、频率分布直方图中估计众数、中位数、平均数

(1)众数：众数通常是频率分布直方图中最高矩形的中点的横坐标。

(2)中位数：在频率分布直方图中，中位数左边和右边的直方图的面积相等。

(3)平均数：平均数=每个小矩形的面积与小矩形底边中点横坐标之积的和。

### 【典例讲练】

#### 题型一 求平均数、中位数、众数

例1 (1) 某地区某村的前3年的经济收入(单位：万元)分别为100, 200, 300，其统计数据的中位数为  $x$ ，平均数为  $y$ 。今年经过政府新农村建设工作，该村经济收入(单位：万元)在上年基础上翻番，则在这4年里经济收入的统计数据中，下列说法正确的是( )

- A. 中位数为  $x$ ，平均数为  $1.5y$                       B. 中位数为  $1.25x$ ，平均数为  $y$   
C. 中位数为  $1.25x$ ，平均数为  $1.5y$                       D. 中位数为  $1.5x$ ，平均数为  $2y$

[答案]C

[解析]由数据100, 200, 300可得，前3年统计数据的中位数  $x = 200$ ，平均数  $y = \frac{100 + 200 + 300}{3} = 200$ 。

根据题意得第4年该村的经济收入的统计数据为600，则由数据100, 200, 300, 600可得，这4年统计数据的中位数为  $\frac{200 + 300}{2} = 250 = 1.25x$ ，平均数为  $\frac{100 + 200 + 300 + 600}{4} = 300 = 1.5y$ 。故选C。

(2)在如图所示一组数据的茎叶图中，有一个数字被污染后模糊不清，但曾计算得该组数据的极差与中位数之和为61，则被污染的数字为( )

- A. 1                      B. 2                      C. 3                      D. 4

|   |   |   |   |   |   |
|---|---|---|---|---|---|
| 2 | 0 | 1 | 5 |   |   |
| 3 | 1 | 1 | ■ | 4 |   |
| 4 | 2 | 3 | 5 | 7 | 8 |

[答案]B

[解析]由图可知该组数据的极差为 $48-20=28$ ，则该组数据的中位数为 $61-28=33$ ，易得被污染的数字为2。故选B。

【练习1】(1) 某便利店记录了100天某商品的日需求量(单位：件)，整理得下表：

|             |     |     |     |     |     |
|-------------|-----|-----|-----|-----|-----|
| 日需求量 $n$ /件 | 14  | 15  | 16  | 18  | 20  |
| 频率          | 0.1 | 0.2 | 0.3 | 0.2 | 0.2 |

试估计该商品日平均需求量为( )

- A. 16件                  B. 16.2件                  C. 16.6件                  D. 16.8件

[答案]D

[解析]由题意可知，日平均需求量为 $14 \times 0.1 + 15 \times 0.2 + 16 \times 0.3 + 18 \times 0.2 + 20 \times 0.2 = 16.8$ (件)。

(2) 某校课外活动小组对该市做空气含尘量调查，下面是一天中每隔两小时测得的数据(单位： $\text{g}/\text{m}^3$ )：  
0.03，0.04，0.03，0.02，0.04，0.01，0.03，0.03，0.04，0.05，0.01，0.03.

①求出这组数据的众数和中位数；

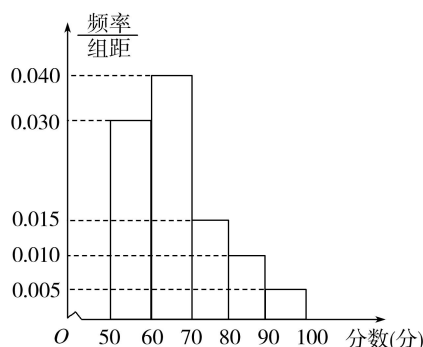
②若国标(国家环保总局的标准)是平均值不得超过 $0.025 \text{ g}/\text{m}^3$ ，则这一天该城市的空气是否符合国标？

[解析]①众数是0.03(出现5次)，将这12个数从小到大排列中间的两个数是0.03，故中位数是0.03.

②这组数据的平均数 $\bar{x} = \frac{1}{12} \times (0.01 \times 2 + 0.02 + 0.03 \times 5 + 0.04 \times 3 + 0.05) = 0.03(\text{g}/\text{m}^3)$ ，也就是说，这一天该城市的空气含尘量的平均值是 $0.03 \text{ g}/\text{m}^3$ ，超过了 $0.025 \text{ g}/\text{m}^3$ ，故不符合国标.

## 题型二 根据频率分布直方图求众数，平均数，中位数

例2 某中学举行电脑知识竞赛，现将高一参赛学生的成绩进行整理后分成五组绘制成如图所示的频率分布直方图，已知图中从左到右的第一、二、三、四、五小组的频率分别是0.30，0.40，0.15，0.10，0.05.

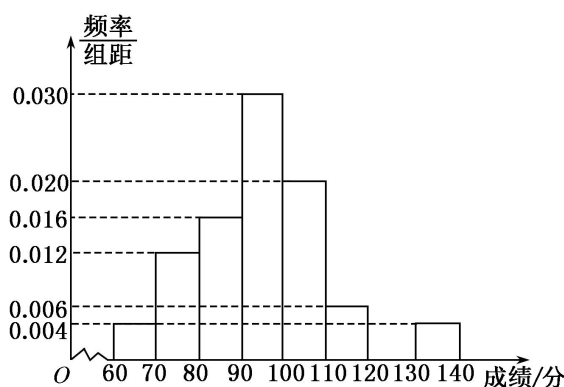


求：①成绩的众数、中位数的估计值；      ②平均成绩的估计值

[解析]①由图可知众数约为65.又 $\because$ 第一个小矩形的面积为0.3,  $\therefore$ 设第二个小矩形底边的一部分长为 $x$ , 则 $x \times 0.04 = 0.2$ , 得 $x = 5$ . $\therefore$ 中位数约为 $60 + 5 = 65$ .

②依题意, 平均成绩为 $55 \times 0.3 + 65 \times 0.4 + 75 \times 0.15 + 85 \times 0.1 + 95 \times 0.05 = 67$ ,  $\therefore$ 平均成绩约为67.

例3 某校高三年级的500名学生参加了一次数学测试, 已知这500名学生的成绩全部介于60分到140分之间, 为统计学生的这次考试情况, 从这500名学生中随机抽取50名学生的考试成绩作为样本进行统计. 将这50名学生的考试成绩的统计结果按如下方式分成八组: 第一组 $[60, 70)$ , 第二组 $[70, 80)$ , 第三组 $[80, 90)$ , ..., 第八组 $[130, 140]$ 。如图是按上述分组方法得到的频率分布直方图的一部分。



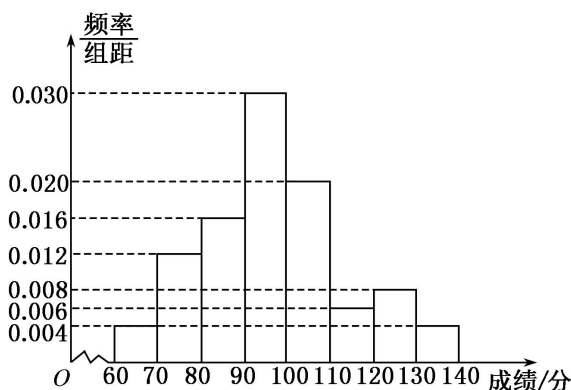
(1)求第七组的频率, 并完成频率分布直方图;

(2)估计该校高三年级的这500名学生这次考试成绩的中位数;

(3)若从样本中考试成绩属于第一组和第六组的所有学生中随机抽取2名, 求这2名学生的分数差的绝对值大于10分的概率。

[解析](1)由频率分布直方图得第七组的频率为:

$$1 - (0.004 + 0.012 + 0.016 + 0.03 + 0.02 + 0.006 + 0.004) \times 10 = 0.08, \text{ 频率分布直方图如图.}$$



(2)由频率分布直方图得考试成绩在 $[60, 90)$ 内的频率为 $(0.004 + 0.012 + 0.016) \times 10 = 0.32$ ,

考试成绩在 $[90, 100)$ 内的频率为 $0.03 \times 10 = 0.3$ , 所以估计该校高三年级的这500名学生这次考试成绩的

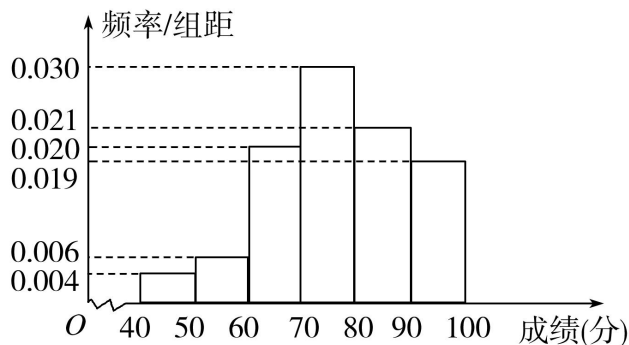
$$\text{中位数为 } 90 + \frac{0.5 - 0.32}{0.3} \times 10 = 96.$$

(3)样本中第一组有学生 $50 \times 0.004 \times 10 = 2$ (人), 第六组有学生 $50 \times 0.006 \times 10 = 3$ (人), 从样本中考试成

绩属于第一组和第六组的所有学生中随机抽取 2 名，基本事件总数  $n = C_5^2 = 10$ ，这 2 名学生的分数差的绝对值大于 10 分包含的基本事件个数  $m = C_2^1 C_3^1 = 6$ ，

所以这 2 名学生的分数差的绝对值大于 10 分的概率  $P = \frac{m}{n} = \frac{6}{10} = 0.6$ 。

【练习2】高三抽出 50 名学生参加数学竞赛，由成绩得到如下的频率分布直方图。由于一些数据丢失，试利用频率分布直方图求：



①这 50 名学生成绩的众数与中位数（结果保留 1 位小数）；

②这 50 名学生的平均成绩。

[解析]①由众数的概念可知，众数是出现次数最多的数。在频率分布直方图中高度最高的小长方形框的中间值的横坐标即为所求，所以众数应为 75. 由于中位数是所有数据中的中间值，故在频率分布直方图中体现的是中位数的左右两边频数应相等，即频率也相等，从而就是小矩形的面积和相等。因此在频率分布直方图中将频率分布直方图中所有小矩形的面积一分为二的直线所对应的成绩即为所求。

$$\because 0.004 \times 10 + 0.006 \times 10 + 0.02 \times 10 = 0.04 + 0.06 + 0.2 = 0.3,$$

$$\therefore \text{前三个小矩形面积的和为 } 0.3. \text{ 而第四个小矩形面积为 } 0.03 \times 10 = 0.3, \quad 0.3 + 0.3 > 0.5,$$

$$\therefore \text{中位数应位于第四个小矩形内。设其底边为 } x, \text{ 高为 } 0.03,$$

$$\therefore \text{令 } 0.03x = 0.2, \text{ 得 } x \approx 6.7. \text{ 故中位数应为 } 70 + 6.7 = 76.7.$$

②样本平均值，即每个小矩形底边的中点值乘以每个小矩形的面积。  $\therefore$  平均成绩为

$$45 \times (0.004 \times 10) + 55 \times (0.006 \times 10) + 65 \times (0.02 \times 10) + 75 \times (0.03 \times 10) + 85 \times (0.021 \times 10) + 95 \times (0.019 \times 10) = 76.5.$$

【课后巩固】——完成课时作业（195）

## 19.2.5 总体离散程度的估计

### 【学习目标】

- 1、理解标准差、方差的统计意义
- 2、会求标准差、方差

### 【要点整合】

- 1、方差、标准差

(1)方差的求法:

$$s^2 = \frac{1}{n} \left[ (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2 \right]$$

(2)标准差的求法:

方差的算术平均数即为标准差。

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n} \left[ (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2 \right]}$$

其中,  $x_n$  是样本数据,  $n$  是样本容量,  $\bar{x}$  是样本平均数.

(3)总体方差与样本方差的关系:

总体方差是反映总体波动大小的特征数, 样本方差是反映样本波动大小的特征数. 通常用样本方差估计总体方差, 当样本容量很大时, 样本方差很接近总体方差.

- 2、标准差(方差)的两个作用

(1)标准差(方差)较大, 数据的离散程度较大; 标准差(方差)较小, 数据的离散程度较小.

(2)在实际应用中, 常常把平均数与标准差(方差)结合起来进行决策. 在平均值相等的情况下, 比较方差或标准差以确定稳定性.

### 【典例讲练】

#### 题型一 求方差、标准差

例 1(1)抽样统计甲、乙两位射击运动员的 5 次训练成绩(单位: 环), 结果如下:

| 运动员 | 第 1 次 | 第 2 次 | 第 3 次 | 第 4 次 | 第 5 次 |
|-----|-------|-------|-------|-------|-------|
| 甲   | 87    | 91    | 90    | 89    | 93    |
| 乙   | 89    | 90    | 91    | 88    | 92    |

则成绩较为稳定(方差较小)的那位运动员成绩的方差为\_\_\_\_\_.

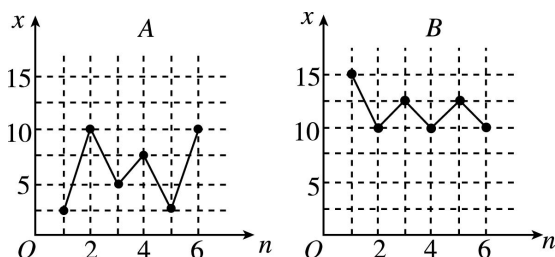
[答案]2

[解析]本题考查统计中的方差计算.

$$\bar{x}_{\text{甲}} = 90, \text{ 且 } s_{\text{甲}}^2 = \frac{1}{5} \left[ (87-90)^2 + (91-90)^2 + (90-90)^2 + (89-90)^2 + (93-90)^2 \right] = 4,$$

$$\bar{x}_{\text{乙}} = 90, \text{ 且 } s_{\text{乙}}^2 = \frac{1}{5} \left[ (89-90)^2 + (90-90)^2 + (91-90)^2 + (88-90)^2 + (92-90)^2 \right] = 2. \text{ 所求方差为 } 2.$$

(2)如图, 样本  $A$  和  $B$  分别取自两个不同的总体, 它们的样本平均数分别为  $\bar{x}_A$  和  $\bar{x}_B$ , 样本标准差分别为  $S_A$  和  $S_B$ , 则( )



- A.  $\bar{x}_A > \bar{x}_B$ ,  $S_A > S_B$     B.  $\bar{x}_A < \bar{x}_B$ ,  $S_A > S_B$     C.  $\bar{x}_A > \bar{x}_B$ ,  $S_A < S_B$     D.  $\bar{x}_A < \bar{x}_B$ ,  $S_A < S_B$

[答案]B

[解析]由图可知  $A$  组的 6 个数为 2.5, 10, 5, 7.5, 2.5, 10,  $B$  组的 6 个数为 15, 10, 12.5, 10, 12.5, 10, 所以  $\bar{x}_A = \frac{2.5+10+5+7.5+2.5+10}{6} = \frac{37.5}{6}$ ,  $\bar{x}_B = \frac{15+10+12.5+10+12.5+10}{6} = \frac{70}{6}$ .

显然  $\bar{x}_A < \bar{x}_B$ , 又由图形可知,  $B$  组的数据分布比  $A$  均匀, 变化幅度不大, 故  $B$  组数据比较稳定, 方差较小, 从而标准差较小, 所以  $S_A > S_B$ , 故选 B.

【练习 1】(1)某校甲、乙两个班级各有 5 名编号为 1, 2, 3, 4, 5 的学生进行投篮练习, 每人投 10 次, 投中的次数如下表:

| 学生 | 1 号 | 2 号 | 3 号 | 4 号 | 5 号 |
|----|-----|-----|-----|-----|-----|
| 甲班 | 6   | 7   | 7   | 8   | 7   |
| 乙班 | 6   | 7   | 6   | 7   | 9   |

则以上两组数据的方差中较小的一个  $s^2 =$  \_\_\_\_\_.

[答案] $\frac{2}{5}$

[解析] $\bar{x}_{\text{甲}} = 7, s_{\text{甲}}^2 = \frac{1}{5}(1^2 + 0^2 + 0^2 + 1^2 + 0^2) = \frac{2}{5}$ ,  $\bar{x}_{\text{乙}} = 7, s_{\text{乙}}^2 = \frac{1}{5}(1^2 + 0^2 + 1^2 + 0^2 + 2^2) = \frac{6}{5}$ ,

$\therefore s_{\text{甲}}^2 < s_{\text{乙}}^2$ ,  $\therefore$  方差中较小的一个为  $s_{\text{甲}}^2$ , 即  $s^2 = \frac{2}{5}$ .

(2)甲、乙两名运动员在某项测试中的 6 次成绩如下图所示,  $\bar{x}_1$ ,  $\bar{x}_2$  分别表示甲、乙两名运动员这项测试成绩的平均数,  $s_1$ ,  $s_2$  分别表示甲、乙两名运动员这项测试成绩的标准差, 则有( )

| 甲 |   |   |   |   | 乙 |   |   |   |
|---|---|---|---|---|---|---|---|---|
| 9 |   |   |   | 0 | 8 |   |   |   |
| 6 | 5 | 5 | 4 | 1 | 3 | 5 | 5 | 7 |
| 1 |   |   |   | 2 | 2 |   |   |   |

- A.  $\bar{x}_1 > \bar{x}_2$ ,  $s_1 < s_2$     B.  $\bar{x}_1 = \bar{x}_2$ ,  $s_1 < s_2$     C.  $\bar{x}_1 = \bar{x}_2$ ,  $s_1 = s_2$     D.  $\bar{x}_1 < \bar{x}_2$ ,  $s_1 > s_2$

[答案]B

[解析]由题图可得  $\bar{x}_1 = \frac{1}{6} \times (9+14+15+15+16+21) = 15$ ,  $\bar{x}_2 = \frac{1}{6} \times (8+13+15+15+17+22) = 15$ ,

代入标准差公式可得  $s_1 = \sqrt{\frac{37}{3}} < s_2 = \sqrt{\frac{53}{3}}$ .

## 题型二 方差、标准差的应用

例2 甲、乙两机床同时加工直径为100 cm的零件,为了检验质量,各从中抽取6件进行测量,分别记录数据为:

甲: 99 100 98 100 100 103.

乙: 99 100 102 99 100 100

(1)分别计算两组数据的平均数及标准差;

(2)根据计算结果判断哪台机床加工零件的质量更稳定.

[解析](1)  $\bar{x}_{\text{甲}} = \frac{1}{6} (99+100+98+100+100+103) = 100$

$\bar{x}_{\text{乙}} = \frac{1}{6} (99+100+102+99+100+100) = 100$

$s_{\text{甲}}^2 = \frac{1}{6} [(99-100)^2 + (100-100)^2 + (98-100)^2 + (100-100)^2 + (100-100)^2 + (103-100)^2] = \frac{7}{3}$ ,

$\therefore s_{\text{甲}} = \frac{\sqrt{21}}{3}$

$s_{\text{乙}}^2 = \frac{1}{6} [(99-100)^2 + (100-100)^2 + (102-100)^2 + (99-100)^2 + (100-100)^2 + (100-100)^2] = 1$ ,

$\therefore s_{\text{乙}} = 1$ .

(2)两台机床所加工零件的直径的平均数相同,又  $s_{\text{甲}} > s_{\text{乙}}$ ,所以乙机床加工零件的质量更稳定.

【练习2】甲、乙两种冬小麦试验品种连续5年的平均单位面积产量如下(单位: t/hm<sup>2</sup>):

| 品种 | 第一年 | 第二年  | 第三年  | 第四年 | 第五年  |
|----|-----|------|------|-----|------|
| 甲  | 9.8 | 9.9  | 10.1 | 10  | 10.2 |
| 乙  | 9.4 | 10.3 | 10.8 | 9.7 | 9.8  |

根据这组数据判断应该选择哪一种小麦进行推广.

[解析]甲种冬小麦的平均产量  $\bar{x}_{\text{甲}} = \frac{9.8+9.9+10.1+10+10.2}{5} = 10$ ,

乙种冬小麦的平均产量  $\bar{x}_{\text{乙}} = \frac{9.4+10.3+10.8+9.7+9.8}{5} = 10$ ,

则甲、乙两种冬小麦平均产量相同. 甲种冬小麦产量的方差为

$$s_{\text{甲}}^2 = \frac{1}{5} \times [(9.8-10)^2 + (9.9-10)^2 + (10.1-10)^2 + (10-10)^2 + (10.2-10)^2] = 0.02,$$

乙种冬小麦产量的方差为

$$s_{\text{乙}}^2 = \frac{1}{5} \times [(9.4-10)^2 + (10.3-10)^2 + (10.8-10)^2 + (9.7-10)^2 + (9.8-10)^2] = 0.244,$$

则  $s_{\text{甲}}^2 = 0.02 < s_{\text{乙}}^2 = 0.244$ ，所以甲种小麦的产量比较稳定。因此选择甲种小麦进行推广。

### 题型三 方差、标准差的性质

例 3(1)已知一组数据  $x_1, x_2, \dots, x_n$  的方差是  $a$ ，求另一组数据  $x_1-2, x_2-2, \dots, x_n-2$  的方差；

(2)设一组数据  $x_1, x_2, \dots, x_n$  的标准差为  $s_x$ ，另一组数据  $3x_1+a, 3x_2+a, \dots, 3x_n+a$  的标准差为  $s_y$ ，求  $s_x$  与  $s_y$  的关系。

[解析](1)设  $x_1, x_2, \dots, x_n$  的平均数为  $\bar{x}$ ，则有  $a = \frac{1}{n} [(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2]$ 。

$\because x_1-2, x_2-2, \dots, x_n-2$  的平均数为  $\bar{x}-2$ ，则这组数据的方差

$$s^2 = \frac{1}{n} [(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2] = \frac{1}{n} [(x_1 - 2 - \bar{x} + 2)^2 + (x_2 - 2 - \bar{x} + 2)^2 + \dots + (x_n - 2 - \bar{x} + 2)^2] = a.$$

(2)设  $x_1, x_2, \dots, x_n$  的平均数为  $\bar{x}$ ，则  $3x_1+a, 3x_2+a, \dots, 3x_n+a$  的平均数为  $3\bar{x}+a$ 。

$$\begin{aligned} s_y &= \sqrt{s_y^2} = \sqrt{\frac{1}{n} [(3\bar{x}+a-3x_1-a)^2 + (3\bar{x}+a-3x_2-a)^2 + \dots + (3\bar{x}+a-3x_n-a)^2]} \\ &= \sqrt{\frac{1}{n} \cdot 3^2 \cdot [(\bar{x}-x_1)^2 + (\bar{x}-x_2)^2 + \dots + (\bar{x}-x_n)^2]} = \sqrt{9 \cdot s_x^2} = 3s_x, \therefore s_y = 3s_x. \end{aligned}$$

【练习 3】(1)一组数据的标准差为  $s$ ，将这组数据中每一个数据都扩大到原来的 2 倍，所得到的一组数据的方差是( )

A.  $\frac{s^2}{2}$

B.  $4s^2$

C.  $2s^2$

D.  $s^2$

[答案]B

(2)若  $k_1, k_2, \dots, k_8$  的方差为 3，则  $2(k_1-3), 2(k_2-3), \dots, 2(k_8-3)$  的方差为\_\_\_\_\_。

[答案]12

(3)若 40 个数据的平方和是 56，平均数是  $\frac{\sqrt{2}}{2}$ ，则这组数据的方差是\_\_\_\_\_，标准差是\_\_\_\_\_。

[答案]  $0.9, \frac{3\sqrt{10}}{10}$

[解析]由方差公式  $s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n}$ ,

$$\text{得 } s^2 = \frac{(x_1^2 + x_2^2 + \cdots + x_n^2) - 2\bar{x}(x_1 + x_2 + \cdots + x_n) + n\bar{x}^2}{n} = \frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n} - \bar{x}^2.$$

$$\text{由已知 } n=40, \quad x_1^2 + x_2^2 + \cdots + x_{40}^2 = 56, \quad \bar{x} = \frac{\sqrt{2}}{2} \therefore s^2 = \frac{56}{40} - \left(\frac{\sqrt{2}}{2}\right)^2 = 0.9, \quad s = \frac{3\sqrt{10}}{10}.$$

**【课后巩固】** \_\_\_\_\_ 完成课时作业（196）

## 第二十章 成对数据的统计分析

### 20.1 变量间的相关关系与相关系数

#### 【学习目标】

1. 了解变量间的相关关系，会画散点图；
2. 根据散点图，能判断两个变量是否具有相关关系.

#### 【要点整合】

##### 1. 变量间的相关关系

###### (1) 变量之间常见的关系

|      |                       |
|------|-----------------------|
| 函数关系 | 变量之间的关系可以用函数表示        |
| 相关关系 | 变量之间有一定的联系，但不能完全用函数表示 |

###### (2) 相关关系与函数关系的区别与联系

| 类别   | 区别                                                                                                                | 联系                                                                                                                                 |
|------|-------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------|
| 函数关系 | ①函数关系中两个变量间是一种 <u>确定性关系</u> ；②函数是一种因果关系，有这样的因，必有这样的果. 例如，圆的半径由1增大为2，其面积必然由 $\pi$ 增大到 $4\pi$                       | ①在一定的条件下可以相互转化，对于具有线性相关关系的两个变量来说，当求得其线性回归方程后，可以用一种确定性的关系对这两个变量间的取值进行评估；<br>②相关关系在现实生活中大量存在，从某种意义上讲，函数关系是一种理想的关系模型，而相关关系是一种更为一般的情况. |
| 相关关系 | ①相关关系是一种 <u>非确定性关系</u> . 例如，吸烟与患肺癌之间的关系，两者之间虽然没有确定的函数关系，但吸烟多的人患肺癌的风险会大幅增加，两者之间即是一种非确定性的关系；②相关关系不一定是因果关系，也可能是伴随关系. |                                                                                                                                    |

##### 2. 散点图及正、负相关的概念

###### (1) 散点图

将样本中 $n$ 个数据点 $(x_i, y_i)(i = 1, 2, \dots, n)$ 描在平面直角坐标系中，以表示具有相关关系的两个变量的一组

数据的图形叫做散点图. 点 $(\bar{x}, \bar{y})$ 叫样本点中心. 其中 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ ,  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ .

###### (2) 正相关与负相关

- ①正相关：散点图中的点散布在从左下角到右上角的区域.
- ②负相关：散点图中的点散布在从左上角到右下角的区域.

### 3. 相关系数

散点图只是形象地描述点的分布情况，它的“线性”是否明显只能通过观察，要想把握其特征，必须进行定量的研究.

相关系数：对于变量  $x$  与  $y$  随机取到  $n$  对数据  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ ，则样本的线性相关系数

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sqrt{\sum_{i=1}^n x_i^2 - n\bar{x}^2} \sqrt{\sum_{i=1}^n y_i^2 - n\bar{y}^2}};$$

相关系数的性质：

- (1)  $|r| \leq 1$ ;
- (2)  $|r|$  越接近于1，相关程度越强； $|r|$  越接近于0，相关程度越弱；
- (3) 通常，当  $|r| \geq 0.75$  时，我们认为两变量具有很强的相关性。

注：若  $r > 0$ ，则  $b > 0$ ，两变量正相关；若  $r < 0$ ，则  $b < 0$ ，两变量负相关；若  $r = 0$ ，则两变量不相关。

#### 【典例讲练】

##### 题型一 相关关系的判断

例1 (1) 下列变量之间的关系不是相关关系的是 ( )

- A. 二次函数  $y = ax^2 + bx + c$  中， $a, c$  是常数，取  $b$  为自变量，因变量是判别式  $\Delta = b^2 - 4ac$
- B. 光照时间和果树亩产量
- C. 降雪量和交通事故发生率
- D. 每亩田施肥量和粮食亩产量

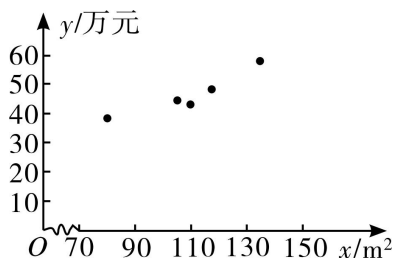
[答案] A

(2) 以下是在某地搜集到的不同楼盘房屋的销售价格  $y$  (单位：万元) 和房屋面积  $x$  (单位： $m^2$ ) 的数据：

|              |      |      |      |      |     |
|--------------|------|------|------|------|-----|
| 房屋面积 $x/m^2$ | 115  | 110  | 80   | 135  | 105 |
| 销售价格 $y$ /万元 | 49.6 | 43.2 | 38.8 | 58.4 | 44  |

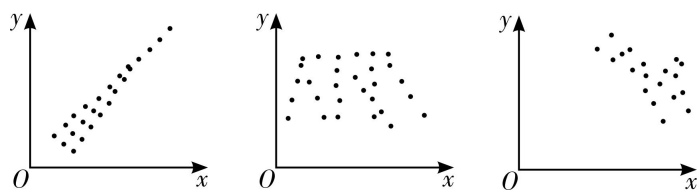
- ① 画出数据对应的散点图；
- ② 判断房屋的销售价格和房屋面积之间是否具有相关关系，如果有相关关系，是正相关还是负相关？

[答案] ① 数据对应的散点图如图所示.



② 通过以上数据对应的散点图可以判断，房屋的销售价格和房屋面积之间具有相关关系，并且是正相关。

【练习1】(1) 观察下列关于两个变量 $x$ 和 $y$ 的三个散点图，它们从左到右的对应关系为 ( )



- A. 正相关、负相关、不相关  
B. 负相关、不相关、正相关  
C. 负相关、正相关、不相关  
D. 正相关、不相关、负相关

[答案]D

(2) 下列说法正确的是 ( )

- A. 圆的面积与半径之间的关系是相关关系  
B. 粮食产量与施肥量之间的关系是函数关系  
C. 一定范围内，学生的成绩与学习时间成正相关关系  
D. 人的体重与视力成负相关关系

[答案]C

## 题型二 相关系数

例2 在一组样本数据为  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n) (n \geq 2, x_1, x_2, \dots, x_n \text{ 不全相等})$  的散点图中，若所有样本点  $(x_i, y_i) (i = 1, 2, \dots, n)$  都在直线  $y = -\frac{1}{3}x + 2$  上，则这组样本数据的相关系数为 ( )

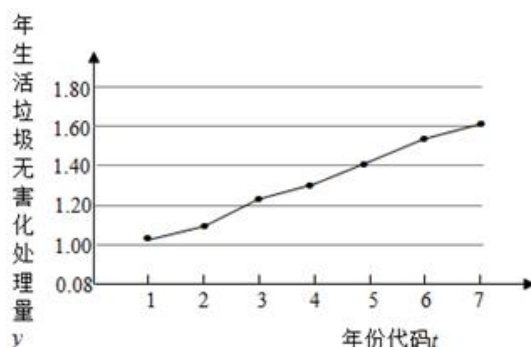
- A.  $-\frac{1}{3}$   
B.  $\frac{1}{3}$   
C. 1  
D. -1

[答案]D

[解析]根据回归直线方程是  $y = -\frac{1}{3}x + 2$ ，可得这两个变量是负相关，故这组样本数据的样本相关系数为负值，且所有样本点  $(x_i, y_i) (i = 1, 2, \dots, n)$  都在直线上，则有  $|r| = 1$ ， $\therefore$  相关系数  $r = -1$ 。故选D。

例3 如图是我国2014年至2020年生活垃圾无害化处理量（单位：亿吨）的折线图。注：年份代码1 - 7分别对应年份2014 - 2020。

由折线图看出，可用线性回归模型拟合  $y$  与  $t$  的关系，请用相关系数加以证明；



附注：参考数据： $\sum_{i=1}^7 y_i = 9.32$ ， $\sum_{i=1}^7 t_i y_i = 40.17$ ， $\sqrt{\sum_{i=1}^7 (y_i - \bar{y})^2} = 0.55$ ， $\sqrt{7} \approx 2.646$ 。

参考公式：相关系数  $r = \frac{\sum_{i=1}^n (t_i - \bar{t})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (t_i - \bar{t})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$ .

[解析]由折线图看出， $y$ 与 $t$ 之间存在较强的正相关关系，理由如下：

因为  $\sum_{i=1}^7 y_i = 9.32$ ， $\sum_{i=1}^7 t_i y_i = 40.17$ ， $\sqrt{\sum_{i=1}^7 (y_i - \bar{y})^2} = 0.55$ ， $\bar{t} = \frac{1+2+3+4+5+6+7}{7} = 4$ ，

所以  $r = \frac{\sum_{i=1}^7 (t_i - \bar{t})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^7 (t_i - \bar{t})^2 \sum_{i=1}^7 (y_i - \bar{y})^2}} = \frac{\sum_{i=1}^7 t_i y_i - 7\bar{t}\bar{y}}{\sqrt{\sum_{i=1}^7 (t_i - \bar{t})^2 \sum_{i=1}^7 (y_i - \bar{y})^2}} \approx \frac{40.17 - 4 \times 9.32}{2\sqrt{7} \cdot 0.55} \approx \frac{2.89}{2.9106} \approx 0.993$ ，

$\because 0.993 > 0.75$ ，故 $y$ 与 $t$ 之间存在较强的正相关关系.

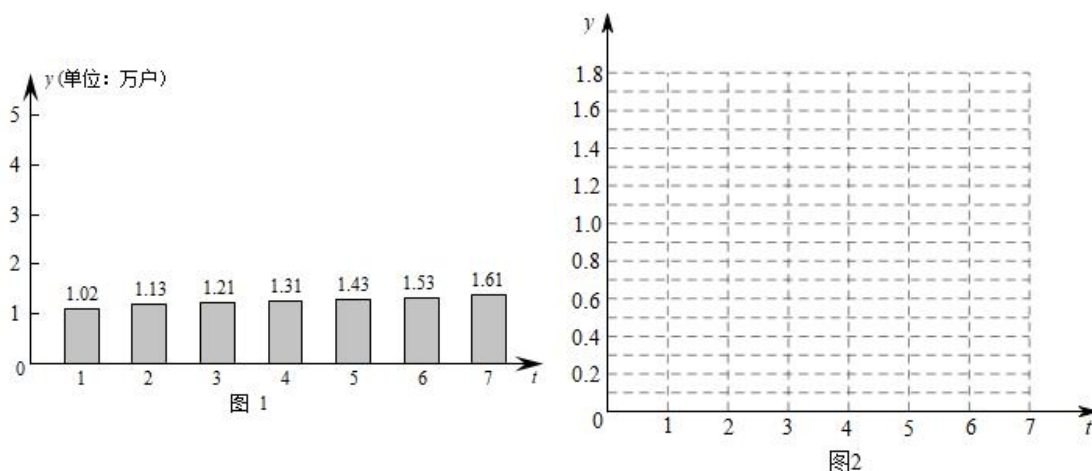
【练习2】在一组样本数据为 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n) (n \geq 2, x_1, x_2, \dots, x_n \text{ 不全相等})$ 的散点图中，若所有样本点 $(x_i, y_i) (i = 1, 2, \dots, n)$ 都在直线 $y = 0.4x + 1$ 上，则这组样本数据的样本相关系数为（ ）

- A. -1                      B. 0.4                      C. 0.6                      D. 1

[答案]D

[解析]因为所有样本点 $(x_i, y_i) (i = 1, 2, \dots, n)$ 都在直线 $y = 0.4x + 1$ 上，所以这组数据完全相关，即说明这组数据完全正相关，相关系数为1，故选D.

【练习3】我国北方广大农村地区、一些城镇以及部分大中城市的周边区域，还在大量采用分散燃煤和散烧煤取暖，既影响了居民基本生活的改善，也加重了北方地区冬季的雾霾天气.推进北方地区冬季清洁取暖，是重大民生工程、民心工程，关系北方地区广大群众温暖过冬，关系雾霾天能不能减少，是能源生产和消费革命、农村生活方式革命的重要内容.2017年9月国家发改委制定了煤改气、煤改电价格扶植新政策，从而使得煤改气、煤改电用户大幅度增加，下面条形图反映了某省2018年1~7月份煤改气、煤改电的用户数量.

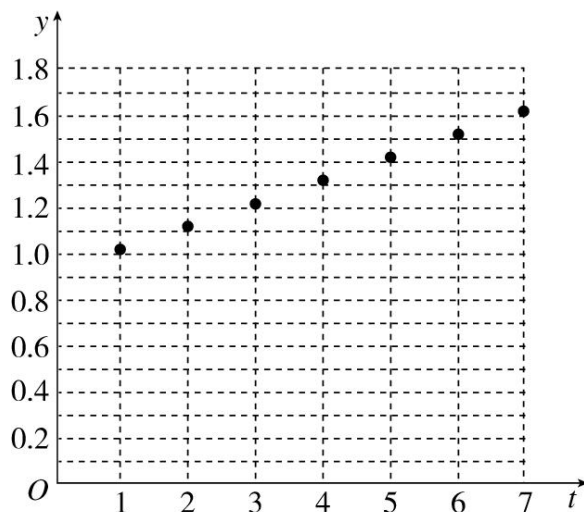


在给定坐标系中作出煤改气、煤改电用户数量 $y$ 随月份 $t$ 变化的散点图，并用散点图和相关系数说明 $y$ 与 $t$ 之间具有线性相关性；

参考数据:  $\sum_{i=1}^7 y_i = 9.24$ ,  $\sum_{i=1}^7 t_i y_i = 39.75$ ,  $\sqrt{\sum_{i=1}^7 (y_i - \bar{y})^2} \approx 0.53$ ,  $\sqrt{7} \approx 2.646$ .

参考公式: 相关系数  $r = \frac{\sum_{i=1}^n (t_i - \bar{t})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (t_i - \bar{t})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$ .

[解析]作出散点图如图所示:



由条形图数据和参考数据得:

$$\bar{t} = 4, \quad \sum_{i=1}^7 (t_i - \bar{t})^2 = 28, \quad \sqrt{\sum_{i=1}^7 (y_i - \bar{y})^2} \approx 0.53,$$

$$\sum_{i=1}^7 (t_i - \bar{t})(y_i - \bar{y}) = \sum_{i=1}^7 t_i y_i - \bar{t} \sum_{i=1}^7 y_i = 39.75 - 4 \times 9.24 = 2.79, \quad r \approx \frac{2.79}{0.53 \times 2 \times 2.646} \approx 0.99.$$

因为 $y$ 与 $t$ 的相关系数近似为0.99, 说明 $y$ 与 $t$ 的线性相关性相当高, 从而可以用线性回归模型拟合 $y$ 与 $t$ 的关系.

**【课后巩固】** ————— **完成课时作业 (197)**

## 20.2 一元线性回归模型及其应用

### 20.2.1 一元线性回归模型及其应用（第一课时）

#### 【学习目标】

- 1.了解线性回归思想，会求回归直线的方程；
- 2.将非线性回归模型转化为线性回归模型，并求出对应方程。

#### 【要点整合】

- 1.回归直线：如果散点图中点的分布从整体上看大致在一条直线附近，就称这两个变量之间具有线性相关关系，这条直线叫做回归直线。回归直线过样本点中心。
- 2.线性回归方程：回归直线对应的方程叫做回归直线的方程，简称回归方程。
- 3.最小二乘法：

求线性回归方程  $\hat{y} = \hat{b}x + \hat{a}$  时，使得样本数据的点到回归直线的距离的平方和最小的方法叫做最小二乘法。其中， $\hat{b}$  是线性回归方程的斜率， $\hat{a}$  是线性回归方程在  $y$  轴上的截距，求得的  $\hat{b}$ ， $\hat{a}$  叫做  $b, a$  的最小二乘估计。

- 4.用最小二乘法求回归方程中的  $\hat{a}$ ， $\hat{b}$  有下面的公式：

$$\hat{b} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}; \quad \hat{a} = \bar{y} - \hat{b}\bar{x}.$$

其中  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ ， $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ 。这样，回归方程的斜率为  $\hat{b}$ ，纵截距为  $\hat{a}$ ，即回归方程为  $\hat{y} = \hat{b}x + \hat{a}$ 。

- 5.残差：通过观测得到的数据称为观测值，通过回归方程得到的  $\hat{y}$  称为预测值（估计值），观测值（实际值） 减去 预测值（估计值） 称为残差，即  $e_i = y_i - \hat{y}_i = y_i - (\hat{b}x_i + \hat{a})$ ； $\sum_{i=1}^n (y_i - \hat{y}_i)^2$  称为残差平方和。

- 6.相关指数：用来刻画不同回归模型的拟合效果，记为  $R^2$ 。计算公式为  $R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$ ， $R^2$  越大，

意味着残差平方和越小，即模型的拟合效果越好。

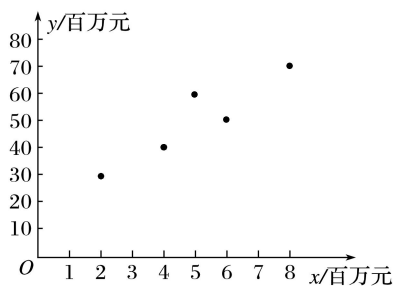
#### 【典例讲练】

例1 某种产品的广告费支出  $x$  (单位：百万元) 与销售额  $y$  (单位：百万元) 之间有如下对应数据：

|     |    |    |    |    |    |
|-----|----|----|----|----|----|
| $x$ | 2  | 4  | 5  | 6  | 8  |
| $y$ | 30 | 40 | 60 | 50 | 70 |

- (1) 画出散点图；
- (2) 求回归方程。

[解析] (1) 散点图如图所示.



(2) 列出下表, 并进行有关计算.

|                                                                            |    |     |     |     |     |
|----------------------------------------------------------------------------|----|-----|-----|-----|-----|
| $i$                                                                        | 1  | 2   | 3   | 4   | 5   |
| $x_i$                                                                      | 2  | 4   | 5   | 6   | 8   |
| $y_i$                                                                      | 30 | 40  | 60  | 50  | 70  |
| $x_i y_i$                                                                  | 60 | 160 | 300 | 300 | 560 |
| $x_i^2$                                                                    | 4  | 16  | 25  | 36  | 64  |
| $\bar{x}=5, \bar{y}=50, \sum_{i=1}^n x_i^2=145, \sum_{i=1}^n x_i y_i=1380$ |    |     |     |     |     |

$$\text{于是可得, } \hat{b} = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sum_{i=1}^n x_i^2 - n \bar{x}^2} = \frac{1380 - 5 \times 5 \times 50}{145 - 5 \times 5^2} = 6.5, \quad \hat{a} = \bar{y} - \hat{b} \bar{x} = 50 - 6.5 \times 5 = 17.5.$$

故所求的回归方程是  $y = 6.5x + 17.5$ .

【练习1】(1) 若例1的条件不变, 利用例1中所求得的回归方程, 计算若广告费支出增加一个单位, 销售额增加多少?

[解析] 由回归方程  $y = 6.5x + 17.5$  可知, 当  $x$  增加一个单位时,  $y$  大约增加6.5.

(2) 若例1的条件不变, 要使销售额提升到100(单位: 百万元), 则广告费至少要支出多少?

[解析] 由  $6.5x + 17.5 = 100$ , 解得  $x = 12.7$ , 即广告费至少要支出12.7(单位: 百万元).

(3) 已知变量  $x$  和  $y$  的统计数据如下表:

|     |     |   |   |     |   |
|-----|-----|---|---|-----|---|
| $x$ | 3   | 4 | 5 | 6   | 7 |
| $y$ | 2.5 | 3 | 4 | 4.5 | 6 |

根据上表可得回归直线方程为  $\hat{y} = \hat{b}x - 0.25$ , 据此可以预测当  $x = 8$  时,  $\hat{y} = (\quad)$

A.6.4

B.6.25

C.6.55

D.6.45

[答案]C

[解析] 由题意知  $\bar{x} = \frac{3+4+5+6+7}{5} = 5, \bar{y} = \frac{2.5+3+4+4.5+6}{5} = 4$ , 将点  $(5, 4)$  代入  $\hat{y} = \hat{b}x - 0.25$ ,

解得  $\hat{b} = 0.85$ , 则  $\hat{y} = 0.85x - 0.25$ , 所以当  $x = 8$  时,  $\hat{y} = 0.85 \times 8 - 0.25 = 6.55$ , 故选C.

例2 某公司为了预测下月产品销售情况，找出了近7个月的产品销售量 $y$ （单位：万件）的统计表：

|                 |       |       |       |       |       |       |       |
|-----------------|-------|-------|-------|-------|-------|-------|-------|
| 月份代码<br>$t$     | 1     | 2     | 3     | 4     | 5     | 6     | 7     |
| 销售量 $y$<br>(万件) | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ |

但其中数据污损不清，经查证  $\sum_{i=1}^7 y_i = 9.32$ ， $\sum_{i=1}^7 t_i y_i = 40.17$ ， $\sqrt{\sum_{i=1}^7 (y_i - \bar{y})^2} = 0.55$ 。

- (1) 请用相关系数说明销售量 $y$ 与月份代码 $t$ 有很强的线性相关关系；
- (2) 求 $y$ 关于 $t$ 的回归方程（系数精确到0.01）；
- (3) 公司经营期间的广告宣传费 $x_i = \sqrt{t_i}$ （单位：万元）（ $i=1, 2, \dots, 7$ ），每件产品的销售价为10元，预测第8个月的毛利润能否突破15万元，请说明理由。（毛利润等于销售金额减去广告宣传费）

参考数据： $\sqrt{7} \approx 2.646$ ，相关系数  $r = \frac{\sum_{i=1}^n (t_i - \bar{t})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (t_i - \bar{t})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$ ，当 $|r| > 0.75$ 时认为两个变量有很强的

线性相关关系。

[解析](1)由折线图中的数据 and 附注中的参考数据得  $\bar{t} = 4$ ， $\sum_{i=1}^7 (t_i - \bar{t})^2 = 28$ ， $\sqrt{\sum_{i=1}^7 (y_i - \bar{y})^2} = 0.55$ ，

$$\sum_{i=1}^7 (t_i - \bar{t})(y_i - \bar{y}) = \sum_{i=1}^7 t_i y_i - \bar{t} \sum_{i=1}^7 y_i = 40.17 - 4 \times 9.32 = 2.89$$

$$\therefore r = \frac{2.89}{2\sqrt{7} \times 0.55} \approx \frac{2.89}{2 \times 2.646 \times 0.55} \approx 0.99, \quad \text{因为 } 0.99 > 0.75$$

所以销售量 $y$ 与月份代码 $t$ 有很强的线性相关关系。

$$(2) \text{由 } \bar{y} = \frac{9.32}{7} \approx 1.331 \text{ 及 (1) 得 } \hat{b} = \frac{\sum_{i=1}^7 (t_i - \bar{t})(y_i - \bar{y})}{\sum_{i=1}^7 (t_i - \bar{t})^2} = \frac{2.89}{28} \approx 0.103$$

$$\hat{a} = \bar{y} - \hat{b}\bar{t} \approx 1.331 - 0.103 \times 4 \approx 0.92 \text{ 所以 } y \text{ 关于 } t \text{ 的回归方程为 } \hat{y} = 0.10t + 0.92$$

(3) 当 $t=8$ 时，代入回归方程得  $\hat{y} = 0.10 \times 8 + 0.92 = 1.72$ （万件）

$$\text{第8个月的毛利润为 } z = 10 \times 1.72 - \sqrt{8} = 17.2 - 2 \times 1.414 = 14.372 < 15, \quad \text{}$$

预测第8个月的毛利润不能突破15万元。

**【课后巩固】**—————完成课时作业（198）

## 20.2.2 一元线性回归模型及其应用（第二课时）

### 【典例讲练】

例1 为了研究某班男生身高和体重的关系，从该班男生中随机选取6名，得到他们的身高和体重的数据如下表所示：

|                   |     |     |     |     |     |     |
|-------------------|-----|-----|-----|-----|-----|-----|
| 编号                | 1   | 2   | 3   | 4   | 5   | 6   |
| 身高 $x(\text{cm})$ | 165 | 171 | 167 | 173 | 179 | 171 |
| 体重 $y(\text{kg})$ | 62  | $m$ | 64  | 74  | 74  | 66  |

在收集数据时，2号男生的体重数值因字迹模糊看不清，故利用其余5位男生的数值得到身高与体重的线性回归方程为  $\hat{y} = \hat{b}_1 x + \hat{a}_1$ . 后来得到2号男生的体重精准数值  $m$  后再次计算得到线性回归方程为  $\hat{y} = \hat{b}_2 x + \hat{a}_2$ .

(1) 求回归方程  $\hat{y} = \hat{b}_1 x + \hat{a}_1$ ；

(2) 若分别按照  $\hat{y} = \hat{b}_1 x + \hat{a}_1$  和  $\hat{y} = \hat{b}_2 x + \hat{a}_2$  来预测身高为180cm的男生的体重，得到的估计值分别为  $w_1, w_2$ ，且  $w_2 - w_1 = 2$ ，求  $m$  的值；

(3) BMI指数是目前国际上常用的衡量人体胖瘦程度以及是否健康的一个标准，其中BMI指数在24到27.9之间的定义为超重.通过计算可知这6人的BMI指数分别为：22.8，27.4，22.9，24.7，23.1，22.6，现从这6人中任选2人，求恰有1人体重为超重的概率.

[解析] (1)  $\bar{x} = 171, \bar{y} = 68$ ,

$$\text{所以 } \sum_{i=1}^5 (x_i - \bar{x})(y_i - \bar{y}) = 36 + 16 + 12 + 48 = 112, \sum_{i=1}^5 (x_i - \bar{x})^2 = 36 + 16 + 4 + 64 = 120$$

$$\text{所以 } \hat{b}_1 = \frac{\sum_{i=1}^5 (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^5 (x_i - \bar{x})^2} = \frac{112}{120} = \frac{14}{15}, \hat{a}_1 = \bar{y} - \hat{b}_1 \bar{x} = 68 - \frac{14}{15} \times 171 = -\frac{1374}{15}, \text{ 所以 } \hat{y} = \frac{14}{15}x - \frac{1374}{15}.$$

(2) 根据题意，将  $x = 180$  代入方程  $\hat{y} = \frac{14}{15}x - \frac{1374}{15}$  得  $w_1 = \frac{1146}{15}$ ,

$$\text{所以 } w_2 = w_1 + 2 = \frac{1176}{15} = \hat{b}_2 \times 180 + \hat{a}_2 \quad ①$$

另一方面，6名男生的身高的平均值为  $\bar{x}' = 171$ ，体重的平均值为  $\bar{y}' = \frac{340 + m}{6}$ ,

$$\text{所以 } \frac{340 + m}{6} = \hat{b}_2 \times 171 + \hat{a}_2 \quad ②$$

$$\sum_{i=1}^6 (x_i - \bar{x})(y_i - \bar{y}) = 36 + 16 + 12 + 48 = 112, \sum_{i=1}^6 (x_i - \bar{x})^2 = 36 + 16 + 4 + 64 = 120$$

$$\text{所以 } \hat{b}_2 = \frac{\sum_{i=1}^6 (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^6 (x_i - \bar{x})^2} = \frac{14}{15}, \quad \text{③ 综合①②③即可得: } \hat{a}_2 = -\frac{1344}{15}, m = 80.$$

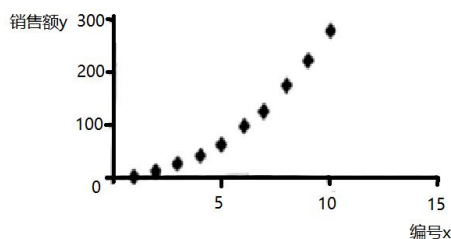
$$(3) P = \frac{C_4^1 C_2^1}{C_6^2} = \frac{8}{15}.$$

例2 2019年双十一落下帷幕,天猫交易额定格在268(单位:十亿元)人民币(下同),再创新高,比去年218(十亿元)多了50(十亿元),这些数字的背后,除了是消费者买买买的表现,更是购物车里中国新消费的奇迹,为了研究历年销售额的变化趋势,一机构统计了2010年到2019年天猫双十一的销售额数据 $y$ (单位:十亿元).绘制如下表1:

表1

| 年份      | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016  | 2017  | 2018 | 2019 |
|---------|------|------|------|------|------|------|-------|-------|------|------|
| 编号 $x$  | 1    | 2    | 3    | 4    | 5    | 6    | 7     | 8     | 9    | 10   |
| 销售额 $y$ | 0.9  | 8.7  | 22.4 | 41   | 65   | 94   | 132.5 | 172.5 | 218  | 268  |

根据以上数据绘制散点图,如图所示.



(1) 根据散点图判断, $y = a + bx$ 与 $y = cx^2 + d$ 哪一个适宜作为销售额 $y$ 关于 $x$ 的回归方程类型?(给出判断即可,不必说明理由)

(2) 根据(1)的判断结果及下表中的数据,建立 $y$ 关于 $x$ 的回归方程,并预测2020年天猫双十一销售额;(注:数据保留小数点后一位)

(3) 把销售额超过10(十亿元)的年份叫“畅销年”,把销售额超过100(十亿元)的年份叫“狂欢年”,从2010年到2019年这十年的“畅销年”中任取3个,求取到的“狂欢年”个数 $\xi$ 的分布列与期望.

参考数据: $t_i = x_i^2$ .

|                              |                                  |                             |                                       |                                         |                            |
|------------------------------|----------------------------------|-----------------------------|---------------------------------------|-----------------------------------------|----------------------------|
| $\sum_{i=1}^{10} y_i = 1020$ | $\sum_{i=1}^{10} x_i y_i = 8088$ | $\sum_{i=1}^{10} t_i = 385$ | $\sum_{i=1}^{10} t_i^2 \approx 25380$ | $\sum_{i=1}^{10} t_i y_i \approx 67770$ | $(\bar{t})^2 \approx 1483$ |
|------------------------------|----------------------------------|-----------------------------|---------------------------------------|-----------------------------------------|----------------------------|

[解析] (1) 根据其图像的形状可知,  $y = cx^2 + d$  更适宜.

$$(2) \quad \hat{b} = \frac{\sum_{i=1}^{10} t_i y_i - 10 \bar{t} \bar{y}}{\sum_{i=1}^{10} t_i^2 - 10 \bar{t}^2} = \frac{67770 - 10 \times 38.5 \times 102}{25380 - 14830} = \frac{28500}{10550} = \frac{570}{211} \approx 2.7,$$

$$\hat{a} = \bar{y} - \hat{b} \bar{t} = 102 - 2.7 \times 38.5 \approx -2.0,$$

$\therefore \hat{y} = 2.7x^2 - 2.0$ , 当  $x = 11$  时,  $\hat{y} = 324.7$  (十亿元),  $\therefore$  预测2020年双十一的销售额为324.7 十亿元.

(3) 畅销年个数为8, 狂欢年个数为4,  $\xi$  的可能取值为0, 1, 2, 3

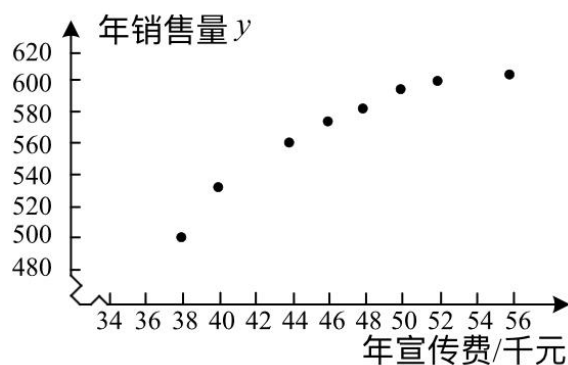
$$P(\xi = 0) = \frac{C_4^3}{C_8^3} = \frac{4}{56} = \frac{1}{14}, P(\xi = 1) = \frac{C_4^2 \cdot C_4^1}{C_8^3} = \frac{24}{56} = \frac{3}{7},$$

$$P(\xi = 2) = \frac{C_4^2 \cdot C_4^1}{C_8^3} = \frac{24}{56} = \frac{3}{7}, P(\xi = 3) = \frac{C_4^3}{C_8^3} = \frac{4}{56} = \frac{1}{14},$$

|       |                |               |               |                |
|-------|----------------|---------------|---------------|----------------|
| $\xi$ | 0              | 1             | 2             | 3              |
| $P$   | $\frac{1}{14}$ | $\frac{3}{7}$ | $\frac{3}{7}$ | $\frac{1}{14}$ |

$$\therefore E(\xi) = 0 \times \frac{1}{14} + 1 \times \frac{3}{7} + 2 \times \frac{3}{7} + 3 \times \frac{1}{14} = \frac{3}{2}.$$

【练习1】某公司为确定下一年度投入某种产品的宣传费, 需了解年宣传费  $x$  (单位: 千元) 对年销售量  $y$  (单位:  $t$ ) 和年利润  $z$  (单位: 千元) 的影响, 对近8年的年宣传费  $x_i$  和年销售量  $y_i$  ( $i = 1, 2, \dots, 8$ ) 数据作了初步处理, 得到下面的散点图及一些统计量的值.



|           |           |           |                                  |                                  |                                               |                                                        |
|-----------|-----------|-----------|----------------------------------|----------------------------------|-----------------------------------------------|--------------------------------------------------------|
| $\bar{x}$ | $\bar{y}$ | $\bar{w}$ | $\sum_{i=1}^s (x_i - \bar{x})^2$ | $\sum_{i=1}^s (w_i - \bar{w})^2$ | $\sum_{i=1}^s (x_i - \bar{x})(y_i - \bar{y})$ | $\sum_{i=1}^s (w_i - \bar{w})^2 \cdot (y_i - \bar{y})$ |
| 46.6      | 563       | 6.8       | 298.8                            | 1.6                              | 1469                                          | 108.8                                                  |

表中  $w_i = \sqrt{x_i}$ ,  $\bar{w} = \frac{1}{8} \sum_{i=1}^8 w_i$

(1) 根据散点图判断,  $y = a + bx$  与  $y = c + d\sqrt{x}$  哪一个适宜作为年销售量  $y$  关于年宣传费  $x$  的回归方程类型? (给出判断即可, 不必说明理由)

(2) 根据 (1) 的判断结果及表中数据, 建立  $y$  关于  $x$  的回归方程;

(3) 已知这种产品的年利润  $z$  与  $x$ 、 $y$  的关系为  $z = 0.2y - x$ . 根据 (2) 的结果求年宣传费  $x = 49$  时, 年销售量及年利润的预报值是多少?

[解析] (1) 由散点图可以判断,  $y = c + d\sqrt{x}$  适宜作为年销售量  $y$  关于年宣传费  $x$  的回归方程类型.

(2) 令  $w = \sqrt{x}$ , 先建立  $y$  关于  $w$  的线性回归方程. 由于

$$\hat{d} = \frac{\sum_{i=1}^n (w_i - \bar{w})(y_i - \bar{y})}{\sum_{i=1}^n (w_i - \bar{w})^2} = \frac{108.8}{1.6} = 68, \hat{c} = \bar{y} - \hat{d}\bar{w} = 563 - 68 \times 6.8 = 100.6$$

所以  $y$  关于  $w$  的线性回归方程为  $\hat{y} = 100.6 + 68w$ , 因此  $y$  关于  $x$  的回归方程为  $\hat{y} = 100.6 + 68\sqrt{x}$ .

(3) 由 (2) 知, 当  $x = 49$  时, 年销售量  $y$  的预报值  $\hat{y} = 100.6 + 68\sqrt{49} = 576.6$   
年利润  $z$  的预报值  $\hat{z} = 576.6 \times 0.2 - 49 = 66.32$ .

**【课后巩固】** ————— **完成课时作业 (199)**

## 20.3 列联表与独立性检验

### 20.3.1 分类变量与列联表

#### 【学习目标】

1. 通过实例,理解 $2 \times 2$ 列联表的统计意义,了解 $2 \times 2$ 列联表独立性检验及其应用.
2. 在学习分类变量和 $2 \times 2$ 列联表的过程中,提升数学抽象、数据分析、数学建模、数学运算的核心素养.

#### 【要点整合】

##### 1. 分类变量

- (1) 分类变量:用一种特殊的随机变量,以区别不同的现象或性质,这类随机变量称为分类变量.
- (2) 取值:分类变量的取值可以用实数表示.
- (3) 范围:本节主要讨论取值于 $\{0, 1\}$ 的分类变量的关联性问题.

##### 2. $2 \times 2$ 列联表

(1) 在实践中,由于保存原始数据的成本较高,人们经常按研究问题的需要,将数据分类统计,并做成表格加以保存.这种形式的数据统计表称为 $2 \times 2$ 列联表. $2 \times 2$ 列联表给出了成对分类变量数据的交叉分类频数.

##### (2) 制作 $2 \times 2$ 列联表的基本步骤

第一步,合理选取两个变量,且每一个变量都可以取两个值.

第二步,抽取样本,整理数据.

第三步,画出 $2 \times 2$ 列联表.

#### 【典例讲练】

##### 题型一 用 $2 \times 2$ 列联表分析两变量间的相关关系

例1 在对人们饮食习惯的一次调查中,共调查了124人,其中六十岁以上的70人,六十岁以下的54人.六十岁以上的人中有43人的饮食以蔬菜为主,另外27人则以肉类为主;六十岁以下的人中有21人饮食以蔬菜为主,另外33人则以肉类为主.请根据以上数据作出饮食习惯与年龄的列联表,并判断二者是否有关系.

[解析]用 $\Omega$ 表示124构成的集合.考虑以 $\Omega$ 为样本空间的古典概型.对于 $\Omega$ 中每一个人,定义分类变量 $X$ 和 $Y$

如下:  $X = \begin{cases} 0, & \text{饮食以蔬菜为主} \\ 1, & \text{饮食以肉类为主} \end{cases}$   $Y = \begin{cases} 0, & \text{年龄60岁以上} \\ 1, & \text{年龄60岁以下} \end{cases}$

将数据整理成如下 $2 \times 2$ 列联表.

| 饮食习惯             | 年龄                |                   | 合计  |
|------------------|-------------------|-------------------|-----|
|                  | 年龄在六十岁以上( $Y=0$ ) | 年龄在六十岁以下( $Y=1$ ) |     |
| 饮食以蔬菜为主( $X=0$ ) | 43                | 21                | 64  |
| 饮食以肉类为主( $X=1$ ) | 27                | 33                | 60  |
| 合计               | 70                | 54                | 124 |

$$P(X=0|Y=1) = \frac{n(X=0, Y=1)}{n(X=0)} = \frac{a}{a+b} = \frac{43}{64} = 0.671875$$

$$P(X=1|Y=0) = \frac{n(X=1, Y=0)}{n(X=1)} = \frac{c}{c+d} = \frac{27}{60} = 0.45$$

由 $P(X=0|Y=1)$ 大于 $P(X=0|Y=0)$ 可以判断年龄对饮食习惯有影响, 据此可以在某种程度上认为饮食习惯与年龄有关系.

[方法总结](1)作 $2 \times 2$ 列联表时，关键是对涉及的变量分清类别.

(2)利用 $2 \times 2$ 列联表分析两变量间的关系时,首先要根据题中数据获得 $2 \times 2$ 列联表,结合古典概型和条件概率

计算  $\frac{a}{a+b}$  与  $\frac{c}{c+d}$  (或  $\frac{b}{a+b}$  与  $\frac{d}{c+d}$ ) 的值, 以此来判断两个分类变量间是否相互影响.

**【练习1】** 假设有两个分类变量  $x$  与  $y$ ，它们的可能取值分别为  $\{x_1, x_2\}$  和  $\{y_1, y_2\}$ ，其  $2 \times 2$  列联表为：

|       |       |       |
|-------|-------|-------|
|       | $y_1$ | $y_2$ |
| $x_1$ | 10    | 18    |
| $x_2$ | m     | 26    |

则当 $m$ 取下面何值时， $x$ 与 $y$ 的关系最弱( )

- A. 8  
B. 9  
C. 14  
D. 19

[答案]C

[解析]由 $10 \times 26 = 18m$ , 解得 $m \approx 14.4$ , 所以当 $m = 14$ 时, X与Y的关系最弱.

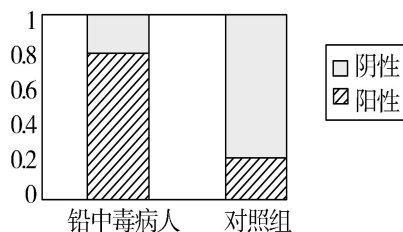
### 题型二 用等高条形图分析两变量间的关系

例2 为了解铅中毒病人与尿棕色素为阳性是否有关系，分别对病人组和对照组的尿液作尿棕色素定性检查，结果如下：

|                | 阳性数( $Y=0$ ) | 阴性数( $Y=1$ ) | 总计 |
|----------------|--------------|--------------|----|
| 铅中毒病人( $X=0$ ) | 29           | 7            | 36 |
| 对照组( $X=1$ )   | 9            | 28           | 37 |
| 总计             | 38           | 35           | 73 |

试画出列联表的等高条形图，分析铅中毒病人和对照组的尿棕色素阳性数有无差别，铅中毒病人与尿棕色素为阳性是否有关系？

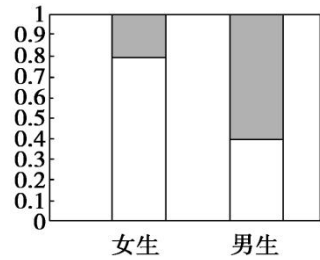
[解析]等高条形图如图所示:



其中两个深色条的高分别代表铅中毒病人和对照组样本中尿棕色素为阳性的频率。由图可以直观地看出铅中毒病人与对照组相比，尿棕色素为阳性的频率差异明显，因此铅中毒病人与尿棕色素为阳性有关系。

【练习2】如图所示的是调查某地区男、女中学生喜欢理科的等高条形图，阴影部分表示喜欢理科的百分比，从图中可以看出( )

- A. 性别与喜欢理科无关
- B. 女生中喜欢理科的比例约为80%
- C. 男生比女生喜欢理科的可能性大些
- D. 男生中不喜欢理科的比例约为60%



[答案]C

[解析]由题图可知女生中喜欢理科的比例约为20%，男生中喜欢理科的比例约为60%，因此男生比女生喜欢理科的可能性大些。

【课后巩固】—————完成课时作业（200）

## 20.3.2 独立性检验

### 【学习目标】

1. 理解列联表的意义，会根据列联表中数据大致判断两个变量是否独立.
2. 理解统计量  $\chi^2$  的意义和独立性检验的基本思想.

### 【要点整合】

#### 1.2×2列联表

| X   | Y   |     | 合计        |
|-----|-----|-----|-----------|
|     | Y=0 | Y=1 |           |
| X=0 | a   | b   | a+b       |
| X=1 | c   | d   | c+d       |
| 合计  | a+c | b+d | n=a+b+c+d |

上表是关于分类变量X和Y的抽样数据的2×2列联表:最后一行的前两个数分别是事件{Y=0}和{Y=1}的频数;最后一列的前两个数分别是事件{X=0}和{X=1}的频数;中间的四个数a,b,c,d是事件{X=x,Y=y}(x,y=0,1)的频数;右下角格中的数n是样本容量.

2.统计量  $\chi^2$ : 
$$\chi^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$$

3.独立性检验:要推断“Ⅰ与Ⅱ有关系”,可按下面的步骤进行:

(1)提出假设  $H_0$ : Ⅰ与Ⅱ没有关系;

(2)根据2×2列联表计算  $\chi^2$  的值;

(3)基于小概率值 $\alpha$ 的检验规则是:

| $\alpha$   | 0.1   | 0.05  | 0.01  | 0.005 | 0.001  |
|------------|-------|-------|-------|-------|--------|
| $x_\alpha$ | 2.706 | 3.841 | 6.635 | 7.879 | 10.828 |

当  $\chi^2 \geq x_\alpha$  时,我们就推断  $H_0$  不成立,即认为X和Y不独立,该推断犯错误的概率不超过 $\alpha$ ;

当  $\chi^2 < x_\alpha$  时,我们没有充分证据推断  $H_0$  不成立,可以认为X和Y独立.

### 【典例讲练】

#### 题型一 独立性检验

例1 某大学餐饮中心为了解新生的饮食习惯,在全校一年级学生中进行了抽样调查,调查结果如下表所示:

| 类型   | 喜欢甜品 | 不喜欢甜品 | 合计  |
|------|------|-------|-----|
| 南方学生 | 60   | 20    | 80  |
| 北方学生 | 10   | 10    | 20  |
| 合计   | 70   | 30    | 100 |

试根据小概率值 $\alpha=0.05$ 的独立性检验,分析南方学生和北方学生在选用甜品的饮食习惯方面是否有差异.

[解析]零假设为  $H_0$ :南方学生和北方学生在选用甜品的饮食习惯方面无差异.将 $2 \times 2$ 列联表中的数据代入公

$$\text{式计算,得 } \chi^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)} = \frac{100 \times (60 \times 10 - 20 \times 10)^2}{70 \times 30 \times 80 \times 20} = \frac{100}{21} \approx 4.762 > 3.841 = X_{0.05}$$

依据小概率值  $\alpha = 0.05$  的独立性检验,我们推断  $H_0$  不成立,即认为南方学生和北方学生在选用甜品的饮食习惯方面有差异.

【练习1】某省进行高中新课程改革已经四年了,为了解教师对新课程教学模式的使用情况,某一教育机构对某学校的教师关于新课程教学模式的使用情况进行了问卷调查,共调查了50人,其中有老教师20人,青年教师30人.老教师对新课程教学模式赞同的有10人,不赞同的有10人;青年教师对新课程教学模式赞同的有24人,不赞同的有6人.

(1)根据以上数据建立一个 $2 \times 2$ 列联表;

(2)试根据小概率值  $\alpha = 0.01$  的独立性检验,分析对新课程教学模式的赞同情况与教师年龄是否有关联.

[解析](1) $2 \times 2$ 列联表如下:

| 教师类型 | 赞同 | 不赞同 | 合计 |
|------|----|-----|----|
| 老教师  | 10 | 10  | 20 |
| 青年教师 | 24 | 6   | 30 |
| 合计   | 34 | 16  | 50 |

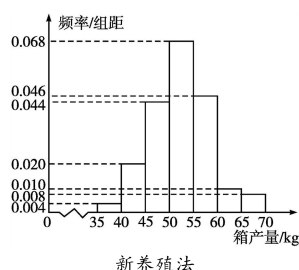
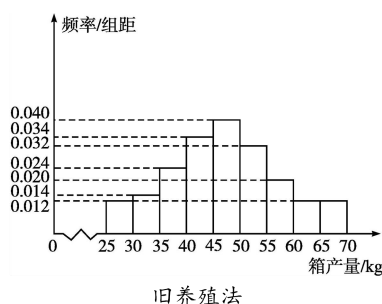
(2)零假设为  $H_0$ :对新课程教学模式的赞同情况与教师年龄无关联.

$$\text{由公式得 } \chi^2 = \frac{50 \times (10 \times 6 - 24 \times 10)^2}{34 \times 16 \times 20 \times 30} \approx 4.963 < 6.635 = X_{0.01},$$

依据小概率值  $\alpha = 0.01$  的独立性检验,没有充分证据推断  $H_0$  不成立,因此可以认为  $H_0$  成立,即认为对新课程教学模式的赞同情况与教师年龄无关联.

## 题型二 独立性检验的综合应用

例2 海水养殖场进行某水产品的新、旧网箱养殖方法的产量对比,收获时各随机抽取了100个网箱,测量各箱水产品的产量(单位:kg),其频率分布直方图如下:



(1)设两种养殖方法的箱产量相互独立,记A表示事件“旧养殖法的箱产量低于50 kg,新养殖法的箱产量不低于50 kg”,估计事件A的概率;

(2)填写下面列联表,并依据  $\alpha = 0.01$  的独立性检验,分析箱产量与养殖方法是否有关联;

| 类型   | 箱产量<50 kg | 箱产量≥50 kg | 合计 |
|------|-----------|-----------|----|
| 旧养殖法 |           |           |    |
| 新养殖法 |           |           |    |
| 合计   |           |           |    |

(3)根据箱产量的频率分布直方图,求新养殖法箱产量的中位数的估计值(精确到0.01).

附:  $\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$

|            |       |       |        |
|------------|-------|-------|--------|
| $\alpha$   | 0.050 | 0.010 | 0.001  |
| $x_\alpha$ | 3.841 | 6.635 | 10.828 |

[解析](1)记B表示事件“旧养殖法的箱产量低于50 kg”,C表示事件“新养殖法的箱产量不低于50 kg”.由题意知 $P(A)=P(BC)=P(B)P(C)$ .

旧养殖法的箱产量低于50 kg的频率为 $(0.012+0.014+0.024+0.034+0.040) \times 5 = 0.62$ ,故 $P(B)$ 的估计值为0.62

新养殖法的箱产量不低于50 kg的频率为 $(0.068+0.046+0.010+0.008) \times 5 = 0.66$ .

故 $P(C)$ 的估计值为0.66, 因此,事件A的概率估计值为 $0.62 \times 0.66 = 0.4092$

(2)零假设为 $H_0$ :箱产量与养殖方法无关联.根据箱产量的频率分布直方图得如下列联表:

| 类型   | 箱产量<50 kg | 箱产量≥50 kg | 合计  |
|------|-----------|-----------|-----|
| 旧养殖法 | 62        | 38        | 100 |
| 新养殖法 | 34        | 66        | 100 |
| 合计   | 96        | 104       | 200 |

$$\chi^2 = \frac{200 \times (62 \times 66 - 34 \times 38)^2}{100 \times 100 \times 96 \times 104} \approx 15.706 > 6.635 = X_{0.01},$$

依据 $\alpha = 0.01$ 的独立性检验,我们推断 $H_0$ 不成立,即认为箱产量与养殖方法有关联.

(3)因为在新养殖法的箱产量频率分布直方图中,箱产量低于50 kg的直方图面积为

$(0.004+0.020+0.044) \times 5 = 0.34 < 0.5$ ,箱产量低于55 kg的直方图面积为

$(0.004+0.020+0.044+0.068) \times 5 = 0.68 > 0.5$ ,故新养殖法箱产量的中位数的估计值为 $50 + \frac{0.5-0.34}{0.068} \approx 52.35$

【练习2】为了解某班学生喜爱打篮球是否与性别有关,对本班48人进行了问卷调查,并得到了如下的2×2列联表:

| 性别 | 喜爱打篮球 | 不喜爱打篮球 | 合计 |
|----|-------|--------|----|
| 男生 |       | 6      |    |
| 女生 | 10    |        |    |
| 合计 |       |        | 48 |

已知在全班48人中随机抽取1人,抽到喜爱打篮球的学生的概率为 $\frac{2}{3}$

(1)请将上面的2×2列联表补充完整(不用写计算过程).

(2)依据  $\alpha = 0.05$  的独立性检验,能否认为喜爱打篮球与性别有关联?说明你的理由.

(3)现从女生中抽取2人进一步调查,设其中喜爱打篮球的女生人数为X,求X的分布列与均值.

[解析](1)列联表补充如下:

| 性别 | 喜爱打篮球 | 不喜爱打篮球 | 合计 |
|----|-------|--------|----|
| 男生 | 22    | 6      | 28 |
| 女生 | 10    | 10     | 20 |
| 合计 | 32    | 16     | 48 |

(2)零假设为  $H_0$ :喜爱打篮球与性别无关联.根据列联表中的数据,经计算得到

$$\chi^2 = \frac{48 \times (22 \times 10 - 6 \times 10)^2}{28 \times 20 \times 32 \times 16} \approx 4.286 > 3.841 = X_{0.05}. \text{依据 } \alpha = 0.05 \text{ 的独立性检验,我们推断 } H_0 \text{ 不成立,即认为喜}$$

爱打篮球与性别有关联.

(3)喜爱打篮球的女生人数X的可能取值为0,1,2.

其概率分别为

$$P(X=0) = \frac{C_{10}^2}{C_{20}^2} = \frac{9}{38}, P(X=1) = \frac{C_{10}^1 C_{10}^1}{C_{20}^2} = \frac{10}{19}, P(X=2) = \frac{C_{10}^2}{C_{20}^2} = \frac{9}{38} \text{ 故X的分布列为}$$

| X | 0              | 1               | 2              |
|---|----------------|-----------------|----------------|
| P | $\frac{9}{38}$ | $\frac{10}{19}$ | $\frac{9}{38}$ |

$$X \text{ 的均值为 } E(X) = 0 + \frac{10}{19} + \frac{9}{19} = 1$$

### 题型三 独立性检验与统计的综合应用

例3 典例某校为了探索一种新的教学模式,进行了一项课题实验,乙班为实验班,甲班为对比班,甲、乙两班均有50人,一年后对两班进行测试,成绩如下表(总分:150分):

甲班

| 成绩 | [80,90) | [90,100) | [100,110) | [110,120) | [120,130) |
|----|---------|----------|-----------|-----------|-----------|
| 人数 | 4       | 20       | 15        | 10        | 1         |

乙班

| 成绩 | [80,90) | [90,100) | [100,110) | [110,120) | [120,130) |
|----|---------|----------|-----------|-----------|-----------|
| 人数 | 1       | 11       | 23        | 13        | 2         |

(1)现从甲班成绩位于[90,120)内的试卷中抽取9份进行试卷分析,请问用什么抽样方法更合理,并写出最后的抽样结果.

(2)根据所给数据可估计在这次测试中,甲班的平均分是101.8分,请你估计乙班的平均分,并计算两班平均分的差距.

(3)完成下面2×2列联表,并依据  $\alpha=0.05$  的独立性检验,分析这两个班在这次测试中成绩的差异与实施课题实验是否有关联,并请说明理由.

| 班别 | 成绩小于100分 | 成绩不小于100分 | 合计  |
|----|----------|-----------|-----|
| 甲班 | $a$      | 26        | 50  |
| 乙班 | 12       | $d$       | 50  |
| 合计 | 36       | 64        | 100 |

[解析](1)用分层随机抽样的方法更合理.甲班成绩位于 $[90,120)$ 内的试卷共有  $20+15+10=45$ ,

从中抽取9份,  $\frac{9}{45} = \frac{1}{5}$ ,

故在 $[90,100)$ , $[100,110)$ , $[110,120)$ 各分数段内分别抽取试卷  $20 \times \frac{1}{5} = 4$ ,  $15 \times \frac{1}{5} = 3$ ,  $10 \times \frac{1}{5} = 2$ .

(2)估计乙班的平均分为  $\bar{x}_乙 = 85 \times \frac{1}{50} + 95 \times \frac{11}{50} + 105 \times \frac{23}{50} + 115 \times \frac{13}{50} + 125 \times \frac{2}{50} = 105.8$ (分)

$105.8-101.8=4$ (分),即两班的平均分相差4分.

(3)补全列联表如下:

| 班别 | 成绩小于100分 | 成绩不小于100分 | 合计  |
|----|----------|-----------|-----|
| 甲班 | $a=24$   | 26        | 50  |
| 乙班 | 12       | $d=38$    | 50  |
| 合计 | 36       | 64        | 100 |

零假设为  $H_0$ :这两个班在这次测试中成绩的差异与实施课题实验无关联.由表中的数据,

$$\text{可得 } \chi^2 = \frac{100 \times (24 \times 38 - 26 \times 12)^2}{50 \times 50 \times 36 \times 64} = 6.25 > 3.841 = \chi_{0.05}^2$$

依据  $\alpha=0.05$  的独立性检验,我们推断 $H_0$ 不成立,即认为这两个班在这次测试中成绩的差异与实施课题实验有关联.

**【课后巩固】** \_\_\_\_\_ **完成课时作业（201）**