# 点金训练

# 教师用书

《点金训练》编写组　编

## 数学

选择性必修第三册

配人教A版

四川教育出版社

CONTENTS

# 目录

## 第六章　计数原理



## 第七章　随机变量及其分布

## 第八章　成对数据的统计分析

# 第六章 计数原理

## 单元概览

### 单元导航

计数是人类最基本、最原始、最古老的数学实践活动.从幼儿时期,我们就开始运用一个一个数的方法解决计数问题;在生活中,遇到复杂的计数问题时,我们也会自然而然地分类、分步计算.从这些直观经验出发,无论是分类加法与分步乘法的巧妙运用,还是排列与组合的精妙计算,亦或是二项式定理的广泛应用,都充分展现了数学的魅力与实用价值.本章系统安排了解决计数问题的原理和方法,包括两个计数原理——分类加法计数原理和分步乘法计数原理,两类特殊计数问题的计数公式——排列数公式和组合数公式,以及这些知识的一个应用——二项式定理.

### 学习目标

1.通过实例,引入计数原理、排列、组合和二项式定理等概念,用自己的方式构建单元知识结构,体会其逻辑关系.

2.通过逻辑推理和数学证明,掌握排列数、组合数的计算方法及性质,分析解决与计数有关的基本问题.

3.通过图形、表格等方式帮助理解二项展开式,能够运用二项式定理进行化简和求值.

### 核心概念

计数是数学中的一种基本技能,计数原理提供了一种直观且普适的思维方式.

在面对一个复杂的计数问题时,往往通过分类或分步将它分解为若干个简单计数问题,在解决这些简单计数问题的基础上,再将它们整合起来就得到原问题的答案.排列、组合、二项式定理是计数原理应用的直观体现,在学习的过程中,可以发展逻辑推理和数学运算素养.

### 学法指导

1.在学习计数原理时,从简单的计数问题入手,逐步深入到复杂的计数问题,在逐步深入的过程中,不断积累知识和经验,直至掌握计数方法,提高逻辑推理素养.

2.在学习排列与组合时,要结合具体实例,区别排列问题和组合问题,将实际计数问题抽象为排列或组合问题,并用排列数或组合数进行计算,提升数学抽象、数学运算、数学建模素养.

3.在学习二项式定理时,可以将其看作多项式乘积展开问题的延续,利用计数原理和排列与组合,探究二项展开式的规律,提升数学抽象、数学运算素养.

## 单元主题任务

某市车牌号由汉字、数字或者字母共七位构成，第一位是汉字，表示该车户口所在省的简称，第二位是字母，表示该市在省内的代码，后五位在 2000 年 1 月 1 日以前是五个数字，由于该市在 2000 年 1 月 1 日时汽车数量突破 100 000 辆，车牌号就不够用了，于是该市从 2000 年 1 月 1 日起在后五位的前两位中引入字母对车牌号进行扩容，由于字母 I，O 和数字 1，0 很像而不采用，所以这两位可以是数字 0～9 和除去 I，O 后的 24 个英文字母中的任何一个.

(1)如果该市的汽车数量将在 2027 年 1 月 1 日达到 150 万辆，该市需要对汽车号牌进一步扩容吗？

(2)2000 年 1 月 1 日这次扩容后的汽车号牌数量可以一个一个地数出来吗？

(3)你能举出一些生活中计数的例子吗？

# 探究构建

## 6.1　分类加法计数原理与分步乘法计数原理

### 第1课时　分类加法计数原理与分步乘法计数原理

**学习任务目标**

**1.**通过实例,了解分类加法计数原理、分步乘法计数原理及其意义.
**2.**能应用两个计数原理解决一些简单问题.

### 问题式预习

**【知识清单】**

**知识点一　分类加法计数原理**

(1)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 $N=\underline{m+n}$ 种不同的方法.

(2)分类加法计数原理的推广

如果完成一件事有 $n$ 类不同方案,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 $N=\underline{m_1+m_2+\cdots+m_n}$ 种不同的方法.

**知识点二　分步乘法计数原理**

(1)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 $N=\underline{m\times n}$ 种不同的方法.

(2)分步乘法计数原理的推广

如果完成一件事需要 $n$ 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 $N=\underline{m_1\times m_2\times\cdots\times m_n}$ 种不同的方法.

**【概念辨析】**

**1.**判断正误(正确的打"√",错误的打"×").

(1)在分类加法计数原理中,两类不同方案中的方法可以相同.　( × )

(2)在分类加法计数原理中,每类方案中的方法都能完成这件事.　( √ )

(3)在分步乘法计数原理中,一件事如果是分两步完成的,那么其中任何一个单独的步骤都不能完成这件事,只有两个步骤都完成后,这件事才算完成.　( √ )

(4)从甲地经丙地到乙地是分步问题.　( √ )

**2.**已知某校高二(1)班有54人,高二(2)班有56人.现从这两个班中任选一人去参加演讲比赛,则共有________种不同的选法.

110　**解析:**若这个人来自高二(1)班,则有54种不同的选法;若这个人来自高二(2)班,则有56种不同的选法.所以共有 $54+56=110$ 种不同的选法.

**3.**某商场共有4个门,若购物者可以从任意一个门进,从任意一个门出,则不同的进出方法的种数是________.

16　**解析:**进出商场可以看作是分两个步骤完成的:第1步,进门,共有4种不同的走法;第2步,出门,共有4种不同的走法.根据分步乘法计数原理知,共有 $4\times4=16$ 种不同的进出方法.

**4.**请思考并回答问题:

如何区分"完成一件事"是分类还是分步?

**提示:**区分"完成一件事"是分类还是分步,关键看一步能否完成这件事,若能完成,则是分类,否则是分步.

## ◦ 任务型课堂 ◦

### 任务1 分类加法计数原理

**1.** 某校举办教师演讲比赛,参赛的语文老师有 20 人,数学老师有 8 人,英语老师有 4 人.若从中评选出一个冠军,则可能的结果种数为 ( )

A.12　　B.28

C.32　　D.640

C　**解析**:由分类加法计数原理知,冠军可能的结果种数为 $20+8+4=32$.

**2.** 在中华传统文化里,建筑、器物、书法、诗歌、对联、绘画均讲究对称之美.数学也讲究对称之美,如 2020 年 02 月 02 日(20200202)被称为世界完全对称日(公历纪年日期中数字左右完全对称的日期).数学上把 20200202 这样的对称数叫回文数,如 11,242,5225 都是回文数,则用 0,1,2,3,4,5 这些数字构成的所有的三位数的回文数中,能被 3 整除的回文数的个数是 ( )

A.8　　B.10

C.11　　D.13

B　**解析**:当三位数的三个数位上的数字都相同时,能被 3 整除的有 111,222,333,444,555,共有 5 个;当三位数的三个数位上的数字有两个相同时,能被 3 整除的有 141,252,303,414,525,共有 5 个.

根据分类加法计数原理知,满足条件的回文数共有 10 个.

**3.** 设集合 $A=\{1,2,3,4\}$,$m,n\in A$,则方程 $\frac{x^2}{m}+\frac{y^2}{n}=1$ 表示的焦点位于 $x$ 轴上的椭圆有 ( )

A.6 个　　B.8 个

C.12 个　　D.16 个

A　**解析**:因为椭圆的焦点在 $x$ 轴上,所以 $m>n$.当 $m=4$ 时,$n=1,2,3$;当 $m=3$ 时,$n=1,2$;当 $m=2$ 时,$n=1$.故满足条件的椭圆共有 $3+2+1=6$ 个.

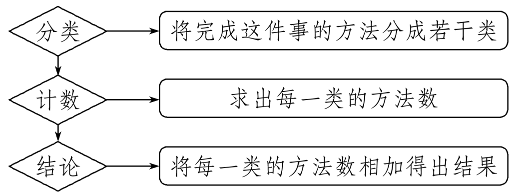
**【探究总结】**

1.使用分类加法计数原理计数的两个注意点:

(1)按照完成这件事的方法的类型合理分类,并做到“不重不漏”;

(2)每类方法都能独立完成这件事.

2.利用分类加法计数原理计数的流程:




### 任务2 分步乘法计数原理

#### 探究活动

**例**　某商店有甲型号电视机 10 台,乙型号电视机 8 台,丙型号电视机 12 台.从这三种型号的电视机中各选一台检验,有多少种不同的选法?

**解**:从这三种型号的电视机中各选一台检验可分三步完成:第 1 步,从甲型号中选一台,有 10 种不同的选法;第 2 步,从乙型号中选一台,有 8 种不同的选法;第 3 步,从丙型号中选一台,有 12 种不同的选法.根据分步乘法计数原理,有 $10\times8\times12=960$ 种不同的选法.

[一题多思]

**思考 1.** 从这三种型号的电视机中选 2 台检验,恰有一台甲型号电视机的选法有多少种?

**解**:根据分步乘法计数原理可知,共有 $10\times(12+8)=200$ 种选法.

**思考 2.** 从这三种型号的电视机中选 2 台检验,至少有一台甲型号电视机的选法有多少种?

**解**:完成这件事可分为两类:第一类,恰有一台甲型号电视机,有 200 种选法;第二类,两台都为甲型号电视机,有 $9+8+7+\cdots+2+1=45$ 种选法.所以共有 $200+45=245$ 种选法.

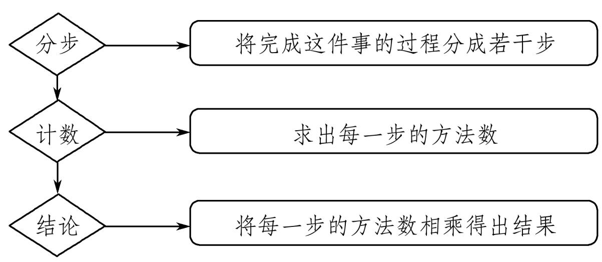
**【探究总结】**

1.使用分步乘法计数原理计数的两个注意点:

(1)按照事件发生的过程合理分步,即分步是有先后顺序的;

(2)只有各个步骤都完成才算完成这件事.

2.利用分步乘法计数原理计数的流程:



## 应用迁移

1.某市人民医院急诊科有 3 名男医生和 4 名女医生，内科有 4 名男医生和 4 名女医生，现从该医院急诊科和内科各选派 1 名男医生和 1 名女医生组成 4 人组，参加省人民医院组织的交流会，则所有不同的选派方案有　(　　)

A.192 种　　B.180 种

C.29 种　　D.15 种

A　**解析**：从急诊科选派 1 名男医生和 1 名女医生有 $3\times4=12$ 种方案，从内科选派 1 名男医生和 1 名女医生有 $4\times4=16$ 种方案，根据分步乘法计数原理，共有 $12\times16=192$ 种不同的选派方案.

2.若在登录某网站时会随机弹出一个 4 位的验证码 XXXX(如 $2a8t$)，第一位和第三位分别为0～9这 10 个数字中的一个，第二位和第四位分别为 $a\sim z$ 这 26 个英文字母中的一个，则这样的验证码共有________个.

67 600　**解析**：完成这件事可分四步：第 1 步，确定验证码的第一位，共有 10 种情况；第 2 步，确定验证码的第二位，共有 26 种情况；第 3 步，确定验证码的第三位，共有 10 种情况；第 4 步，确定验证码的第四位，共有 26 种情况.根据分步乘法计数原理知，这样的验证码共有 $10\times26\times10\times26=67\,600$ 个.

## 提质归纳

| 计数原理 | 分类加法计数原理 | 分步乘法计数原理 |
|---|---|---|
| 关键词 | 分类 | 分步 |
| 区别 | 每类方法都能独立完成这件事 | 每一个步骤都完成，才能完成这件事 |
| 联系 | 都是用来解决关于完成一件事的不同方法种数的问题 | |

# 课后素养评价（一）　分类加法计数原理与分步乘法计数原理

### A组　学习·理解

1.音乐播放器里有 15 首中文歌曲和 5 首英文歌曲，任选 1 首歌曲播放，则不同的选法共有　(　　)

A.30 种　　B.75 种

C.10 种　　D.20 种

D　**解析**：在 15 首中文歌曲和 5 首英文歌曲共 20 首歌曲中任选一首播放，不同的选法共有 20 种.

2.若 $x,y\in\mathbf{N}$，且 $1\leqslant x\leqslant3$，$x+y<7$，则满足条件的不同的有序自然数对 $(x,y)$ 的个数是　(　　)

A.15　　B.12

C.5　　D.4

A　**解析**：当 $x=1$ 时，$y=0,1,2,3,4,5$，有 6 个不同的有序自然数对；

当 $x=2$ 时，$y=0,1,2,3,4$，有 5 个不同的有序自然数对；

当 $x=3$ 时，$y=0,1,2,3$，有 4 个不同的有序自然数对.

根据分类加法计数原理，可得共有 $6+5+4=15$ 个不同的有序自然数对.

3.5 名同学报名参加两个课外活动小组，每名同学限报其中的一个小组，则不同的报名方法有　(　　)

A.10 种　　B.20 种　　C.25 种　　D.32 种

D　**解析**：由题意，每名同学有 2 种选择，故不同的报名方法有 $2^5=32$ 种.故选 D.

4.小明在某一天中有七个课间，为准备“小歌手”比赛，他想要选出至少一个课间来练习唱歌，且希望任意两个练习的课间之间至少有两个课间用来休息，则小明练习唱歌的方案一共有　(　　)

A.31 种　　B.18 种

C.21 种　　D.33 种

B　**解析**：七个课间编号为 1,2,3,4,5,6,7，

若仅有一个课间练习，则每个课间都可以，有 7 种方案；

若有两个课间练习，选法有 $\{1,4\}$，$\{1,5\}$，$\{1,6\}$，$\{1,7\}$，$\{2,5\}$，$\{2,6\}$，$\{2,7\}$，$\{3,6\}$，$\{3,7\}$，$\{4,7\}$，共 10 种方案；

若有三个课间练习，选法为 $\{1,4,7\}$，共 1 种方案.

故共有练习的方案 $7+10+1=18$ 种.

5.将 $(a_1+a_2)(b_1+b_2+b_3)(c_1+c_2+c_3+c_4)$ 展开后有________项.

24　**解析**：根据分步乘法计数原理，展开后的项数为 $2\times3\times4=24$.

6.已知 $a\in\{-1,2,3\}$，$b\in\{0,3,4,5\}$.方程 $(x-a)^2+(y-b)^2=4$ 表示的不同的圆共有________个；若 $r\in\{1,2\}$，则方程 $(x-a)^2+(y-b)^2=r^2$ 表示的不

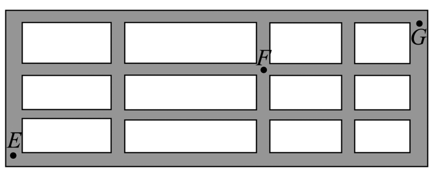
同的圆共有________个.

12　24　**解析**:分两步,第 1 步,$a$ 有 3 种选法;第 2 步,$b$ 有 4 种选法.故共有 $3\times4=12$ 种不同的选法.分三步,第 1 步,$a$ 有 3 种选法;第 2 步,$b$ 有 4 种选法;第 3 步,$r$ 有 2 种选法.故共有 $3\times4\times2=24$ 种不同的选法.

**7.**某城市的电话号码由七位改为八位(首位数字均不为零),则该城市可增加的电话号码个数是________.

$8.1\times10^7$　**解析**:由题意知,本题是一个分步计数问题.电话号码是七位数字时,该城市的电话号码有 $9\times10^6$ 个,同理,改为八位时,该城市的电话号码有 $9\times10^7$ 个.所以可增加的电话号码个数是 $9\times10^7-9\times10^6=81\times10^6=8.1\times10^7$.

**8.**如图,小明从街道的 $E$ 处出发,先到 $F$ 处与小红会合,再和她一起到位于 $G$ 处的老年公寓参加志愿者活动,则小明到老年公寓可以选择的最短路径条数为________.




18　**解析**:由题意可知,$E\to F$ 的最短路径有 6 种走法,$F\to G$ 的最短路径有 3 种走法.由分步乘法计数原理知,共有 $6\times3=18$ 种走法,即有 18 条最短路径.

## B组　应用·实践

**1.**已知三棱锥 $A\text{-}BCD$,现有质点 $Q$ 从点 $A$ 出发沿棱移动,规定质点 $Q$ 从一个顶点沿棱移动到另一个顶点为 1 次移动,则该质点经过 3 次移动后返回到点 $A$ 的不同路径的条数为　(　　)

A.3　　B.6

C.9　　D.12

B　**解析**:由题意可知,经过 3 次移动后返回到点 $A$ 的路径有 $ABCA$,$ABDA$,$ADBA$,$ADCA$,$ACBA$,$ACDA$,共有 6 条不同的路径.故选 B.

**2.**植树节这天,有 4 名同学植树,现有 3 棵不同种类的树,若一棵树只限 1 人植,则不同的分配方法有　(　　)

A.6 种　　B.3 种

C.81 种　　D.64 种

D　**解析**:完成这件事需分三步.第 1 步,植第一棵树有 4 种不同的分配方法;第 2 步,植第二棵树有 4 种不同的分配方法;第 3 步,植第三棵树也有 4 种不同的分配方法.由分步乘法计数原理得,不同的分配方法共有 $4^3=64$ 种.

**3.**从 6 名志愿者中选 4 人分别从事翻译、导游、导购、保洁四项不同的工作.若其中甲、乙两名志愿者不能从事翻译工作,则选派方案共有　(　　)

A.280 种　　B.240 种

C.180 种　　D.96 种

B　**解析**:由于甲、乙不能从事翻译工作,因此翻译工作从余下的 4 名志愿者中选 1 人,有 4 种选法.后面三项工作的选法有 $5\times4\times3$ 种,因此共有 $4\times5\times4\times3=240$ 种不同的选派方案.故选 B.

**4.**某运动会的百米决赛有 8 名男运动员参加.其中甲、乙、丙 3 人必须在 1,2,3,4,5,6,7,8 八条跑道的奇数号跑道上,则安排这 8 名运动员跑道的方式共有________种.

2 880　**解析**:分两步安排这 8 名男运动员.

第 1 步,安排甲、乙、丙 3 人,共有 1,3,5,7 四条跑道可安排,安排方式有 $4\times3\times2=24$ 种;

第 2 步,安排另外 5 人,可安排在 2,4,6,8 及余下的一条奇数号跑道上,安排方式有 $5\times4\times3\times2\times1=120$ 种.

所以安排这 8 名运动员跑道的方式有 $24\times120=2\,880$ 种.

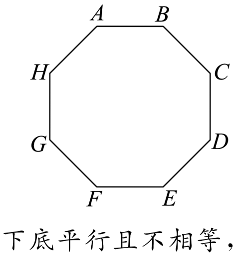
**5.**从 1,2,3,4,5,6,7,9 中,任取两个不同的数作对数的底数和真数,则所有不同的对数值有________个.

39　**解析**:当取 1 时,1 只能为真数,此时这个对数值为 0;当不取 1 时,根据分步乘法计数原理,共有 $7\times6=42$ 种取法,其中 $\log_2 4=\log_3 9=2$,$\log_4 2=\log_9 3=\dfrac{1}{2}$,$\log_2 3=\log_4 9$,$\log_3 2=\log_9 4$,

故此时有 $42-4=38$ 个不同的对数值.

所以共有不同的对数值 $38+1=39$ 个.

**6.**如图,从正八边形 $ABCDEFGH$ 的 8 个顶点中任意取出 4 个,以这 4 个点为梯形的顶点,一共能构成多少个梯形?




**解**:梯形的上、下底平行且不相等,

若以 $AB$ 为底边,则可构成 2 个梯形,根据对称性可知此类梯形有 $2\times8=16$ 个;

若以 $AC$ 为底边，则可构成 1 个梯形，此类梯形共有 $1\times8=8$ 个.

所以共能构成梯形 $16+8=24$ 个.

7.(1)4 名同学选报跑步、跳高、跳远三个项目，每人报一项，共有多少种不同的报名方法？

(2)4 名同学选报跑步、跳高、跳远三个项目，每项必须有人报名，且限报一人，每人至多报一项，共有多少种不同的报名方法？

(3)4 名同学争夺跑步、跳高、跳远三个项目的冠军，共有多少种可能的结果？

**解**：(1)要完成的是"4 名同学每人从三个项目中选一项报名"这件事，因为每人必报一项，4 人都报完才算完成，所以按人分步，且分为四步，又每人可在三项中选一项，选法为 3 种，所以共有 $3\times3\times3\times3=81$(种)不同的报名方法.

(2)因为每项必须有人报名，且限报一人，每人至多报一项，因此跑步项目有 4 种选法，跳高项目有 3 种选法，跳远项目只有 2 种选法.根据分步乘法计数原理，可得不同的报名方法有 $4\times3\times2=24$(种).

(3)要完成的是"三个项目的冠军的获取"这件事，因为每个项目的冠军只能有一人获得，三个项目的冠军都有得主，这件事才算完成，所以应以"确定三个项目的冠军得主"为线索进行分步，而每个项目的冠军的得主有 4 种可能的结果，所以共有 $4\times4\times4=64$(种)可能的结果.

## 第 2 课时　分类加法计数原理与分步乘法计数原理的应用

**学习任务目标**

1.进一步理解和掌握分类加法计数原理和分步乘法计数原理.

2.能应用两个计数原理解决一些实际问题.

## ◦ 问题式预习 ◦

**【知识清单】**

**知识点　两个计数原理的应用**

(1)应用两个计数原理的注意点

一是确定要完成的"一件事"是什么；二是分析需要分类还是需要分步.

(2)如何应用两个计数原理解决问题

分类要做到"不重不漏".分类后再分别对每一类进行计数，最后用分类加法计数原理求和，得到总数.

分步要做到"步骤完整"，即完成了所有步骤，才能完成任务.分步后再计算每一步的方法数，最后根据分步乘法计数原理，把完成每一步的方法数相乘，得到总数.

**【概念辨析】**

1.判断正误(正确的打"√"，错误的打"×").

(1)在一次运动会上有四项比赛，冠军均在甲、乙、丙三人中产生，那么不同的夺冠情况共有 $4^3$ 种.　(　　)

×　**提示**：因为每项比赛的冠军都有 3 种可能的情况，根据分步乘法计数原理，共有 $3^4$ 种不同的夺冠情况.

(2)三只口袋中装有若干小球，一只装有 5 个不同的白色小球，一只装有 6 个不同的黑色小球，一只装有 7 个不同的红色小球.若从中取 2 个不同颜色的小球，则共有 36 种不同的取法.　(　　)

×　**提示**：分为三类：第 1 类是取白球、黑球，有 $5\times6=30$ 种取法；第 2 类是取白球、红球，有 $5\times7=35$ 种取法；第 3 类是取黑球、红球，有 $6\times7=42$ 种取法.

由分类加法计数原理，共有 $30+35+42=107$ 种不同的取法.

2.在 1，2，3，4 四个数字中任取数(不重复)求和，则取出的这些数的不同的和共有　(　　)

A.8 个　　B.9 个　　C.10 个　　D.5 个

A　**解析**：第 1 类，两个数的和有 $1+2=3$，$1+3=4$，$1+4=5$，$2+3=5$，$2+4=6$，$3+4=7$；第 2 类，三个数的和有 $1+2+3=6$，$1+2+4=7$，$1+3+4=8$，$2+3+4=9$；第 3 类，四个数的和有 $1+2+3+4=10$.故得到的不同的和为 3，4，5，6，7，8，9，10，共有 8 个不同的数.

3.有 10 本不同的数学书，9 本不同的语文书，8 本不同的英语书，从中任取 2 本不同类的书，共有不同的取法________种.

242　**解析**：若取的两本书中，一本数学书、一本语

文书，根据分步乘法计数原理，有 $10\times9=90$ 种不同的取法；若取的两本书中，一本语文书、一本英语书，有 $9\times8=72$ 种不同的取法；若取的两本书中，一本数学书、一本英语书，有 $10\times8=80$ 种不同的取法.综上，共有不同的取法 $90+72+80=242$ 种.

4.请思考并回答下列问题：

(1)$(a_1+a_2+a_3)(b_1+b_2+b_3)(c_1+c_2+c_3+c_4+c_5)$展开后共有多少项？

**提示**：由于每一项都是 $a_ib_jc_k$($i\leqslant3,j\leqslant3,k\leqslant5$，且 $i,j,k\in\mathbf{N}^*$)的形式，所以可以分三步完成，第一步，取 $a_i$，有 3 种方法；第二步，取 $b_j$，也有 3 种方法；第三步，取 $c_k$，有 5 种方法.根据分步乘法计数原理，展开后共有 $3\times3\times5=45$ 项.

(2)综合应用两个计数原理的解题思路是什么？

**提示**：对于两个计数原理的综合应用问题，一般是先分类再分步，分类时要先定好分类标准，防止重复和遗漏；分步时要注意步与步之间的连续性，同时应合理设计步骤的顺序，使各步互不干扰.也可以根据题意画出示意图或列出表格，使问题的实质直观地显现出来，从而便于解题.

## ◦ 任务型课堂 ◦

### 任务1 选(抽)取与分配问题

#### 探究活动

**例1** 将3个不同的小球放入5个不同的盒子，每个盒子至多放一个小球，共有多少种不同的放法？

**解**：分三步来完成：

第1步，放第一个小球，有5种放法；

第2步，放第二个小球，有4种放法；

第3步，放第三个小球，有3种放法.

根据分步乘法计数原理，共有 $N=5\times4\times3=60$ 种不同的放法.

[一题多思]

**思考1.**将3个不同的小球放入5个不同的盒子，共有多少种不同的放法？

**解**：分三步来完成：

第1步，放第一个小球，有5种放法；

第2步，放第二个小球，有5种放法；

第3步，放第三个小球，有5种放法.

根据分步乘法计数原理，共有 $N=5\times5\times5=125$ 种不同的放法.

**思考2.**有5个不同的小球和3个不同的盒子，每个盒子中放入一个小球，共有多少种不同的放法？

**解**：分三步来完成：

第1步，向第一个盒子中放入小球，有5种放法；

第2步，向第二个盒子中放入小球，有4种放法；

第3步，向第三个盒子中放入小球，有3种放法.

根据分步乘法计数原理，共有 $N=5\times4\times3=60$ 种不同的放法.

**【探究总结】**

**解决抽取(分配)问题的方法**

(1)当涉及对象的数目不大时，一般选用列举法、树状图法、框图法或列表法.

(2)当涉及对象的数目很大时，一般有两种方法：①直接法.直接使用分类加法计数原理或分步乘法计数原理.一般地，若抽取是有顺序的，则按分步进行；若是按对象特征抽取的，则分类进行.②间接法.去掉限制条件，计算所有的抽取方法数，然后减去所有不符合条件的抽取方法数即可.

#### 应用迁移

1.(多选)现有四个兴趣小组，第一、二、三、四组分别有6人、7人、8人、9人，则下列说法正确的是 ( )

A.选1人为负责人的选法种数为30

B.每组选1名组长的选法种数为3 024

C.若推选2人发言，这2人需来自不同的小组，则不同的选法种数为335

D.若另有3名学生加入这4个小组，可自由选择小组，且第一组必有人选，则不同的选法有35种

ABC **解析**：对于A，选1人为负责人的选法种数为 $6+7+8+9=30$，故A正确；

对于B，每组选1名组长的选法种数为 $6\times7\times8\times9=3\ 024$，故B正确；

对于C，2人来自不同的小组的选法种数为 $6\times7+6\times8+6\times9+7\times8+7\times9+8\times9=335$，故C正确；

对于D，依题意，若不考虑限制，每个人有4种选法，共有 $4^3$ 种选法，若第一组没有人选，每个人有3种选法，共有 $3^3$ 种选法，所以不同的选法有 $4^3-3^3=37$ 种，故D错误.故选ABC.

2.在7名学生中，有3名会下象棋但不会下围棋，有2名会下围棋但不会下象棋，另2名既会下象棋又会下围棋.现在从这7人中选2人分别参加象棋比赛和围棋比赛，共有多少种不同的选法？

**解**：(方法一)第1类，从3名只会下象棋的学生中选1名参加象棋比赛，同时从2名只会下围棋的学生中选1名参加围棋比赛，有 $3\times2=6$ 种选法；

第2类，从3名只会下象棋的学生中选1名参加象棋比赛，同时从2名既会下象棋又会下围棋的学生

中选1名参加围棋比赛，有$3\times2=6$种选法；

第3类，从2名只会下围棋的学生中选1名参加围棋比赛，同时从2名既会下象棋又会下围棋的学生中选1名参加象棋比赛，有$2\times2=4$种选法；

第4类，从2名既会下象棋又会下围棋的学生中选2名分别参加象棋比赛和围棋比赛，有$2\times1=2$种选法.

故共有$6+6+4+2=18$种不同的选法.

(方法二)第1类，从3名只会下象棋的学生中选1名参加象棋比赛，这时7人中还有4人会下围棋，从中选1名参加围棋比赛，有$3\times4=12$种选法；

第2类，从2名既会下象棋又会下围棋的学生中选1名参加象棋比赛，这时7人中还有3人会下围棋，从中选1名参加围棋比赛，有$2\times3=6$种选法.

故共有$12+6=18$种不同的选法.

## 任务2 > 排数问题

### 探究活动

**例2**　用0，1，2，3，4五个数字完成下列排数问题.

(1)用这5个数字给某工厂的产品编号，产品编号由3个数字组成，则不同的编号有多少个？

(2)用这5个数字可以排成多少个三位数？

(3)用这5个数字可以排成多少个能被2整除的无重复数字的三位数？

**解**：(1)用3个数字给产品编号，首位可以是0，数字也可以重复，每个位置都有5种选择，共有$5\times5\times5=5^3=125$个不同的编号.

(2)三位数的首位不能为0，但可以有重复数字，首先考虑首位的排法，除0外共有4种方法，第二、三位可以排0，因此，共可以排成$4\times5\times5=100$个三位数.

(3)能被2整除的数即偶数，末位数字可取0，2，4，因此，可以分两类.第1类，末位数字是0，有$4\times3=12$种排法；第2类，末位数字不是0，则末位有2种排法，即2或4，再排首位，因为0不能在首位，所以有3种排法，中间位有3种排法，因此有$2\times3\times3=18$种排法.

所以共有$12+18=30$种排法，即可以排成30个能被2整除的无重复数字的三位数.

**【探究总结】**

**排数问题的解答策略**

(1)明确特殊位置或特殊数字，是确定采用“分类”还是“分步”的关键.一般按特殊位置(末位或首位)分类，各类中再按特殊数字(或特殊元素)优先的策略分步；如果直接分类较多，可采用间接法求解.

(2)要注意数字“0”不能排在两位数或两位以上的数的最高位.

### 应用迁移

**1.**在由0，1，2，3，4，5所组成的没有重复数字的四位数中，能被5整除的有　(　　)

A.512个　　B.192个

C.240个　　D.108个

D　**解析**：分两类，若末位数字是0，则有$5\times4\times3=60$个；若末位数字是5，则有$4\times4\times3=48$个.所以能被5整除的没有重复数字的四位数有$60+48=108$个.

**2.**用0，1，2，3，4能组成没有重复数字且比32 000小的数有　(　　)

A.212个　　B.213个

C.224个　　D.225个

D　**解析**：一位数有5个，两位数有$4\times4=16$个，三位数有$4\times4\times3=48$个，四位数有$4\times4\times3\times2=96$个，五位数分以下两种情况讨论：

①首位数字为1或2，此时共有$2\times4\times3\times2\times1=48$个；

②首位数字为3，则千位数从0或1中选择一个，其余三个数位任意排列，此时共有$2\times3\times2\times1=12$个.

综上所述，共有$5+16+48+96+48+12=225$个比32 000小的数.

故选D.

**3.**用1，2，3，4四个数字(可重复)排成三位数，并把这些三位数由小到大排成一个数列$\{a_n\}$.

(1)写出这个数列的前11项；

(2)这个数列共有多少项？

(3)若$a_n=341$，求$n$.

**解**：(1)111，112，113，114，121，122，123，124，131，132，133.

(2)这个数列的项数就是用1，2，3，4排成的三位数的个数，每个数位上都有4种排法，则共有$4\times4\times4=64$项.

(3)比$a_n=341$小的数有两类：

第1类如下：

| | | |
|---|---|---|
| 1 | × | × |
| 2 | × | × |

第2类如下：

| | | |
|---|---|---|
| 3 | 1 | × |
| 3 | 2 | × |
| 3 | 3 | × |

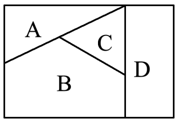
共有$2\times4\times4+3\times4=44$项.

所以$n=44+1=45$.

## 任务3 涂色与种植问题

### 探究活动

**例3** 用 $n$ 种不同的颜色为下面的广告牌着色，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 $n=6$，求共有多少种不同的着色方法；

(2)若共有180种不同的着色方法，求 $n$ 的值.

**解**：(1)分步进行，先为A着色，有6种不同的着色方法，再为B着色，有5种不同的着色方法，然后为C着色，有4种不同的着色方法，最后为D着色，也有4种不同的着色方法，

所以，共有 $6\times5\times4\times4=480$ 种不同的着色方法.

(2)根据分步乘法计数原理，不同的着色方法数是 $n(n-1)\cdot(n-2)(n-2)$.

因为 $n(n-1)(n-2)(n-2)=180$，

所以可用将自然数代入上式验证的方法，得 $n=5$ 时上式成立.

故 $n$ 的值为5.

**【探究总结】**

**解决涂色(种植)问题的一般思路**

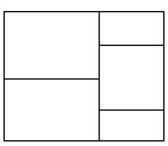
(1)以区域为主分步计数，对于不相邻区域，按照同色、不同色分类计数.

(2)以颜色为主分类计数，适用于为区域、点、线段涂色等问题，先确定使用颜色的种数，再分别涂色.

(3)将空间问题平面化，转化成平面区域的涂色问题.

### 应用迁移

1.某社区计划在如图所示的一块空地布置花卉，要求相邻区域布置的花卉种类不同，且每个区域只布置一种花卉.若有5种不同的花卉可供选择，则不同的布置方案有 ( )



A.360种

B.420种

C.480种

D.540种

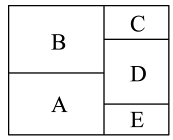
D **解析**：如图，先在区域A布置花卉，有5种不同的布置方案，再在区域E布置花卉，有4种不同的布置方案，再在区域D布置花卉，有3种不同的布置方案.

若区域B与区域E布置同一种花卉，则区域C有3种不同的布置方案；

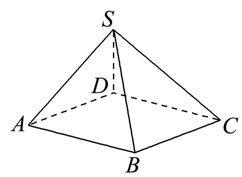
若区域B与区域E布置不同的花卉，则区域B有2种不同的布置方案，区域C有3种不同的布置方案.

故不同的布置方案有 $5\times4\times3\times(3+2\times3)=540$ 种.

故选D.




2.如图，将四棱锥 $S\text{-}ABCD$ 的每一个顶点染上一种颜色，并使同一条棱上的两个顶点异色.若只有5种颜色可供使用，则不同的染色方法总数为 ( )




A.180　　B.240

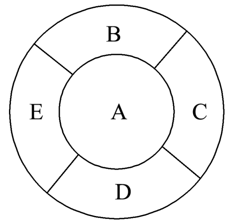
C.420　　D.480

C **解析**：分两步，先将四棱锥一侧面的三个顶点染色，然后再分类考虑另外两个顶点的染色情况.由题意，设5种颜色分别为1,2,3,4,5，四棱锥 $S\text{-}ABCD$ 的顶点 $S,A,B$ 所染的颜色互不相同，它们共有 $5\times4\times3=60$ 种染色方法.

当 $S,A,B$ 染好时，不妨设所染颜色依次为1,2,3，若 $C$ 染2，则 $D$ 可染3或4或5，有3种染法；若 $C$ 染4，则 $D$ 可染3或5，有2种染法；若 $C$ 染5，则 $D$ 可染3或4，有2种染法，即当 $S,A,B$ 染好时，$C,D$ 还有7种染法.

故不同的染色方法总数为 $60\times7=420$.故选C.

3.为迎接元宵节，某广场将一个圆形空地分成A,B,C,D,E五个区域(如图所示)，现用4种颜色的鲜花进行装扮(4种颜色的鲜花均用到)，每个区域用一种颜色的鲜花，相邻区域用不同颜色的鲜花，则不同的鲜花摆放方案共有 ( )




A.48种　　B.36种

C.24种　　D.12种

A　**解析**：满足条件的摆放方案可分为两类：

第一类，B，D区域同色，且和其他区域不同色的摆放方案，满足条件的方案可分四步完成，第一步，先摆区域A，有4种方法；第二步，摆放区域B，D，有3种方法；第三步，摆放区域C，有2种方法；第四步，考虑到区域A，B，C不同色，且4种颜色都要用到，摆放区域E，有1种方法.由分步乘法计数原理可得第一类中共有$4\times3\times2\times1=24$种方案.

第二类，C，E区域同色，且和其他区域不同色的摆放方案，满足条件的方案可分四步完成，第一步，先摆区域A，有4种方法；第二步，摆放区域B，有3种方法；第三步，摆放区域C，E，有2种方法；第四步，考虑到区域A，B，C不同色，且4种颜色都要用到，摆放区域D，有1种方法.由分步乘法计数原理可得第二类中共有$4\times3\times2\times1=24$种方案.

根据分类加法计数原理，可得不同的鲜花摆放方案共有$24+24=48$种.

故选A.

## 提质归纳

- 分类加法计数原理与分步乘法计数原理
  - 知识深化
    - 分类加法计数原理及其应用
    - 分步乘法计数原理及其应用
    - 两个原理的综合应用
  - 方法总结
    - 应用两个原理的步骤
      - 1.分析完成一件事是分类还是分步
      - 2.求出每一类(或每一步)的方法数
      - 3.将每一类(或每一步)的方法数相加(或相乘)
    - 关键点
      - 分类时要做到“不重不漏”
      - 分步时要做到“步骤完整”

# 课后素养评价（二）　分类加法计数原理与分步乘法计数原理的应用

## A组　学习·理解

**1.**某年级要从3名男生、2名女生中选派3人参加某次社区服务.如果要求至少有1名女生，那么不同的选派方案有　　(　　)

A.6种　　B.7种　　C.8种　　D.9种

D　**解析**：按女生人数分类：若选派1名女生，有$2\times3=6$种方案；若选派2名女生，则有$1\times3=3$种方案.由分类加法计数原理，共有$6+3=9$种不同的选派方案.

**2.**从集合$\{1,2,3,\cdots,10\}$中任意选出3个不同的数，使这3个数成等比数列，这样的等比数列的个数为　　(　　)

A.3　　B.4　　C.6　　D.8

D　**解析**：以1为首项的等比数列为1，2，4；1，3，9.以2为首项的等比数列为2，4，8.以4为首项的等比数列为4，6，9.把这4个数列的顺序颠倒，又得到4个等比数列，所以所求的数列共有$2\times(2+1+1)=8$个.

**3.**某外语组9人，每人至少会英语和日语中的一门，其中7人会英语，3人会日语，从中选出会英语和日语的各1人，则不同的选法有________种.

20　**解析**：依题意得，既会英语又会日语的有$7+3-9=1$人，6人只会英语，2人只会日语.

第1类：从只会英语的6人中选1人有6种选法，此时选会日语的有$2+1=3$种选法，由分步乘法计数原理得不同的选法有$6\times3=18$种；

第2类：从既会英语又会日语的人中选1人有1种选法，此时选会日语的有2种选法，由分步乘法计数原理得不同的选法有$1\times2=2$种.

综上，不同的选法共有$18+2=20$种.

**4.**甲、乙、丙3位志愿者被安排在周一至周五的5天中参加某项志愿者活动，要求每人参加一天且每天至多安排一人，并要求甲安排在另外两人前面，则不同的安排方法共有________种.

20　**解析**：分三类：若甲在周一，则乙、丙有$4\times3=12$种排法；

若甲在周二，则乙、丙有$3\times2=6$种排法；

若甲在周三，则乙、丙有$2\times1=2$种排法.

所以不同的安排方法共有$12+6+2=20$种.

**5.**在一个三位数中，若十位数字小于个位数字和百位数字，则称该数为“驼峰数”，比如102，546.由数字1，2，3，4构成的无重复数字的“驼峰数”有________个，其中偶数有________个.

8　5　**解析**：十位上的数为1时，有213，214，312，314，412，413，共6个；十位上的数为2时，有324，423，共2个.所以共有$6+2=8$个“驼峰数”.

其中的偶数有214，312，314，412，324，共5个.

**6.**某快递公司将一个快件从寄件人甲处揽收开始直至送达收件人乙处，需要经过5个转运环节，其中第1，2两个环节各有$a$，$b$两种运输方式，第3，4两个环节各有$b$，$c$两种运输方式，第5个环节有$d$，$e$两种运输方式.则快件从甲处送到乙处恰好用到4种运输方式的不同送达方式有________种.

16 **解析**：快件从甲处送到乙处恰用到 4 种运输方式，且第 5 个环节从 $d,e$ 两种运输方式中选一种，第 1,2,3,4 个环节必须包含 $a,b,c$ 三种不同的运输方式.

若第 1,2 个环节运输方式相同，则只能都选 $a$，第 3,4 个环节一个选 $b$，一个选 $c$，

故有 $2\times1\times2=4$ 种送达方式；

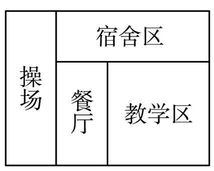
若第 1,2 个环节运输方式不相同，则已经包含 $a,b$ 两种运输方式，

因此，第 3,4 个环节一个选 $b$，一个选 $c$，或者都选 $c$，

故有 $2\times2\times2+2\times1\times2=8+4=12$ 种送达方式.

快件从甲处送到乙处恰用到 4 种运输方式的不同送达方式共有 $4+12=16$ 种.

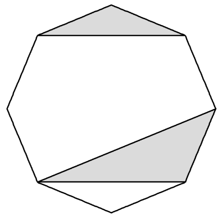
**7.** 如图是某校的校园设施平面示意图，现用不同的颜色为各区域着色，为了便于区分，要求相邻区域不能使用同一种颜色.若有 6 种不同的颜色可供使用，则有________种不同的着色方法.




480 **解析**：(方法一)操场可从 6 种颜色中任选 1 种着色；餐厅可从剩下的 5 种颜色中任选 1 种着色；宿舍区和操场、餐厅的颜色都不能相同，故可从其余的 4 种颜色中任选 1 种着色；教学区和宿舍区、餐厅的颜色都不能相同，故可从其余的 4 种颜色中任选 1 种着色.根据分步乘法计数原理，共有 $6\times5\times4\times4=480$ 种不同的着色方法.

(方法二)分两类：第一类，操场与教学区同色，有 $6\times5\times4=120$ 种着色方法；第二类，操场与教学区不同色，有 $6\times5\times4\times3=360$ 种着色方法.根据分类加法计数原理，共有 $120+360=480$ 种不同的着色方法.

**8.** 如图，由连接正八边形的三个顶点而组成的三角形中与正八边形有公共边的三角形有________个.



40 **解析**：满足条件的有两类：

第一类，与正八边形有两条公共边的三角形有 8 个；

第二类，与正八边形有一条公共边的三角形有 $8\times4=32$ 个，

所以满足条件的三角形共有 $8+32=40$ 个.

## B组 应用·实践

**1.** 现有 15 个不同的小球，其中红球 4 个、黄球 5 个、绿球 6 个，则下列说法不正确的是 ( )

A.从中任选 1 个球，有 15 种不同的选法

B.若每种颜色选出 1 个球，有 120 种不同的选法

C.若要选出不同颜色的 2 个球，有 31 种不同的选法

D.若要不放回地随机选出 2 个球，有 210 种不同的选法

C **解析**：对于 A，从中任选 1 个球，有 $4+5+6=15$ 种不同的选法，故选项 A 正确；对于 B，若每种颜色选出 1 个球，有 $4\times5\times6=120$ 种不同的选法，故选项 B 正确；对于 C，若要选出不同颜色的 2 个球，有 $4\times5+5\times6+4\times6=74$ 种不同的选法，故选项 C 错误；对于 D，若要不放回地随机选出 2 个球，有 $15\times14=210$ 种不同的选法，故选项 D 正确.故选 C.

**2.** 已知集合 $A=\{0,1,2,3,4\}$，且 $a,b,c\in A$，用 $a,b,c$ 组成一个三位数，这个三位数满足“十位上的数字比其他两个数位上的数字都大”，则这样的三位数的个数为 ( )

A.14　　B.17　　C.20　　D.23

C **解析**：由题意知，集合 $A=\{0,1,2,3,4\}$，且 $a,b,c\in A$，

则这个三位数满足“十位上的数字比其他两个数位上的数字都大”包含以下三种情况：

①十位数是 4，则百位数可以是 1,2,3 中的一个数，个位数可以是 0,1,2,3 中的一个数，即符合要求的三位数的个数为 $3\times4=12$；

②十位数是 3，则百位数可以是 1,2 中的一个数，个位数可以是 0,1,2 中的一个数，即符合要求的三位数的个数为 $2\times3=6$；

③十位数是 2，则百位数只能是 1，个位数可以是 0，1 中的一个数，即符合要求的三位数的个数为 2.

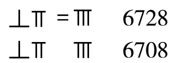
综上，符合要求的三位数共有 $12+6+2=20$ 个.

故选 C.

**3.** 算筹为我国古代数学的发展作出了很大贡献.在算筹计数法中，以“纵式”和“横式”两种方式来表示数字，如表：

| 数字 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| 纵式 | 𝍠 | 𝍡 | 𝍢 | 𝍣 | 𝍤 | 𝍥 | 𝍦 | 𝍧 | 𝍨 |
| 横式 | 𝍩 | 𝍪 | 𝍫 | 𝍬 | 𝍭 | 𝍮 | 𝍯 | 𝍰 | 𝍱 |

表示多位数时，个位用纵式，十位用横式，百位用纵式，千位用横式，以此类推，遇零则置空，如图：




如果把 5 根算筹以适当的方式全部放入下面的表格中，那么可以表示的三位数的个数为 ( )

| | | |
|---|---|---|
| | | |

A.46 B.44 C.42 D.40

B **解析**：按每一位算筹的根数分类一共有 15 种情况，如下：

(5,0,0)，(4,1,0)，(4,0,1)，(3,2,0)，(3,1,1)，(3,0,2)，(2,3,0)，(2,2,1)，(2,1,2)，(2,0,3)，(1,4,0)，(1,3,1)，(1,2,2)，(1,1,3)，(1,0,4).

2 根及 2 根以上的算筹可以表示两个数字，运用分步乘法计数原理，则上述情况能表示的三位数的个数分别为

2，2，2，4，2，4，4，4，4，4，2，2，4，2，2.

根据分类加法计数原理，5 根算筹能表示的三位数的个数为 $2+2+2+4+2+4+4+4+4+4+2+2+4+2+2=44$.故选 B.

**4.**有一密码为 802136 的手提保险箱，现在显示的号码为 721080，要打开箱子，至少旋转________次.(每个旋钮上转出一个新数字为一次旋转，逆转、顺转都可以)

14 **解析**：第一位最少旋转 1 次，其他位置至少旋转的次数依次为 2，1，1，5，4，故至少旋转 $1+2+1+1+5+4=14$ 次.

**5.**用 0，1，2，3，4，5 可以组成多少个无重复数字且比 2 000 大的四位偶数？

**解**：完成这件事可分为三类：

第一类是个位数字为 0 的比 2 000 大的四位偶数，可以分三步完成：

第一步，选取千位上的数字，只有 2，3，4，5 可以选择，有 4 种选法；

第二步，选取百位上的数字，除 0 和千位上已选定的数字以外，还有 4 个数字可以选择，有 4 种选法；

第三步，选取十位上的数字，有 3 种选法.

由分步乘法计数原理知，这类数的个数为 $4\times4\times3=48$.

第二类是个位数字为 2 的比 2 000 大的四位偶数，可以分三步完成：

第一步，选取千位上的数字，除去 2，1，0 只有 3 个数字可以选择，有 3 种选法；

第二步，选取百位上的数字，在去掉已经确定的首、尾 2 个数字之后，还有 4 个数字可以选择，有 4 种选法；

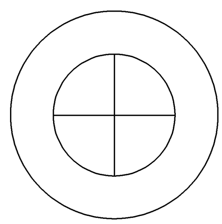
第三步，选取十位上的数字，有 3 种选法.

由分步乘法计数原理知，这类数的个数为 $3\times4\times3=36$.

第三类是个位数字为 4 的比 2 000 大的四位偶数，其方法步骤同第二类.

对以上三类用分类加法计数原理，得所求无重复数字且比 2 000 大的四位偶数有 $48+36+36=120$ 个.

**6.**某小区物业在该小区的一个广场布置了一个如图所示的圆形花坛，花坛分为 5 个区域.现有 6 种不同的花卉可供选择，要求相邻的区域(有公共边)不能布置相同的花卉，且每个区域只布置一种花卉，则不同的布置方案有多少种？

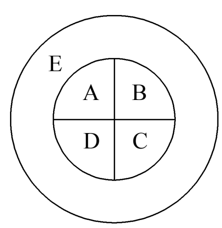


**解**：如图，不同的布置方案分两类：

当 A 与 C 布置相同的花卉时，先安排 E，有 6 种不同的选择；再安排 A 与 C，有 5 种不同的选择；再安排 B，有 4 种不同的选择；最后安排 D，有 4 种不同的选择.共有 $6\times5\times4\times4=480$ 种.

当 A 与 C 布置不同的花卉时，先安排 E，有 6 种不同的选择；再安排 A 与 C，有 $5\times4$ 种不同的选择；再安排 B，有 3 种不同的选择；最后安排 D，有 3 种不同的选择.共有 $6\times5\times4\times3\times3=1\ 080$ 种.

所以不同的布置方案有 $480+1080=1\ 560$ 种.



# 6.2 排列与组合

## 6.2.1 排列

**学习任务目标**

**1.**通过实例，理解排列的概念，并能准确判断一个问题是否是排列问题.

**2.**掌握常见的排列问题的处理方法，如列举法、树状图法.

**3.**会用排列的相关知识解决简单的排列问题.

## ○ 问题式预习 ○

**【知识清单】**

**知识点 排列的概念**

(1)定义：一般地，从 $n$ 个不同元素中取出 $m(m\leqslant n)$ 个元素，并按照一定的顺序排成一列，叫做从 $n$ 个不同元素中取出 $m$ 个元素的一个排列.

(2)两个排列相同的充要条件是：两个排列的元素完全相同，且元素的排列顺序也相同.

**【概念辨析】**

**1.**判断正误(正确的打"√"，错误的打"×").

(1)排列与所选出的元素的排列顺序有关. ( √ )

(2)若两个排列的元素相同，则这两个排列是相同的排列. ( × )

(3)"从 6 名学生中选 3 名学生参加数学、物理、化学竞赛，共有多少种选法?"属于排列问题. ( √ )

(4)"有 12 名学生参加植树活动，要求 3 人一组，共有多少种分组方案?"属于排列问题. ( × )

**2.**下面的问题中，是排列问题的是 ( )

A.由 1，2，3，4 四个数字组成无重复数字的四位数

B.从 60 人中选 11 人组成足球队

C.从 100 人中选 10 人调查消费习惯

D.从 1，2，3，4，5 中选 2 个数组成集合

A **解析**：选项 A 中组成的四位数与数字的排列顺序有关，选项 B，C，D 只需取出元素即可，与元素的排列顺序无关.故选 A.

**3.**请思考并回答下列问题：

(1)"班级选派同学甲参加周六的活动，同学乙参加周日的活动"与"班级选派同学乙参加周六的活动，同学甲参加周日的活动"是相同的安排吗?

**提示**：不是.

(2)排列有何特征?

**提示**：若干个元素按照一定的顺序排成一列，元素完全不同或元素部分相同或元素完全相同但排列顺序不同的排列都是不同的排列.只有当元素完全相同，并且元素的排列顺序也完全相同时，才是同一个排列.

## ○ 任务型课堂 ○

### 任务1 排列的概念及其判断

**1.**给出下列问题：

①用数字 1，2，3 可以组成多少个无重复数字的三位数?

②平面上有 5 个点，任意三点不共线，这 5 个点最多可以确定多少条射线?

③从 40 人中选 5 人组成篮球队，有多少种不同的选法?

其中是排列问题的是________.(填序号)

①② **解析**：①由 1，2，3 组成的无重复数字的三位数与数字的顺序有关，是排列问题；

②由平面上 5 个点确定的射线与端点的顺序有关，是排列问题；

③与顺序无关，不是排列问题.

**2.**判断下列问题是否是排列问题.

(1)从 3 个小组中选 2 个小组分别去植树和种菜；

(2)从 3 个小组中选 2 个小组去种菜；

(3)从 30 人中选 10 人组成一个学习小组；

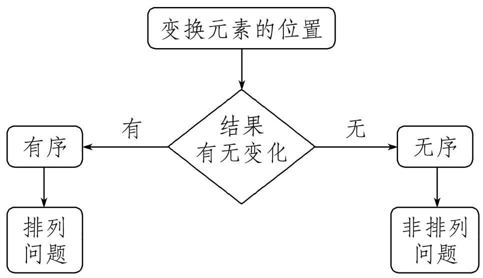
(4)从 30 人中选 3 人分别担任班长、学习委员、生活委员.

**解**：(1)植树和种菜是不同的，存在顺序问题，属于排列问题.

(2)(3)不存在顺序问题，不属于排列问题.

(4)每个人的职务不同,例如,甲当班长或当学习委员是不同的,存在顺序问题,属于排列问题.

【探究总结】




## 任务2 写出简单排列问题的所有排列

### 探究活动

**例1**　现有A,B,C,D四名同学照相留念.

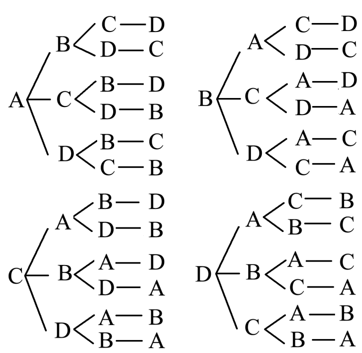
(1)若A,B,C三名同学站成一排照相,有多少种排法?

(2)若A,B,C,D站成一排照相,试将所有排法列出来.

(3)若A,B,C,D站成一排照相,要求自左向右,A不排在第一,B不排在第二,C不排在第三,D不排在第四,试将所有排法列出来.

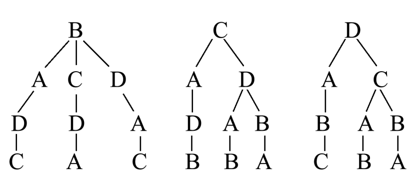
**解**:(1)所有的排法有ABC,ACB,BAC,BCA,CAB,CBA,共6种.

(2)画出树状图如下:



由树状图可知,所有排法为ABCD,ABDC,ACBD,ACDB,ADBC,ADCB,BACD,BADC,BCAD,BCDA,BDAC,BDCA,CABD,CADB,CBAD,CBDA,CDAB,CDBA,DACB,DABC,DBAC,DBCA,DCAB,DCBA,共24种.

(3)画出树状图如下:



由树状图知,所有排法有BADC,BCDA,BDAC,CADB,CDAB,CDBA,DABC,DCAB,DCBA,共9种.

【探究总结】

**树状图的画法**

(1)确定顶部元素,以哪个元素为分类标准,这个元素即为顶部元素.

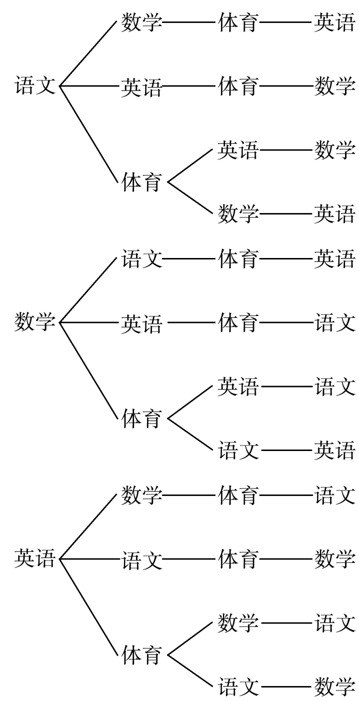
(2)确定分支元素,在每一个分支上将余下的元素按分类标准进行分类,并按顺序排列.

(3)重复以上步骤,直到写完所有排列为止.

### 应用迁移

**1.**某班上午要安排语文、数学、体育和英语4节课,而体育老师因故不能上第一节和第四节,试写出所有排课方案.

**解**:画出树状图如下:



由树状图可知,所有排课方案有:

语文、数学、体育、英语;语文、英语、体育、数学;

语文、体育、英语、数学;语文、体育、数学、英语;

数学、语文、体育、英语;数学、英语、体育、语文;

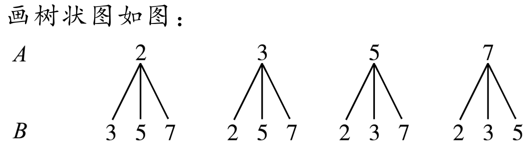
数学、体育、英语、语文;数学、体育、语文、英语;

英语、数学、体育、语文;英语、语文、体育、数学;

英语、体育、数学、语文;英语、体育、语文、数学.

**2.**若直线方程$Ax+By=0$的系数$A$,$B$可以从2,3,5,7中取不同的数值,可以表示多少条不同的直线?试全部列出.

**解**:画树状图如图:




所有不同的直线为$2x+3y=0$,$2x+5y=0$,$2x+7y=0$,

$3x+2y=0$,$3x+5y=0$,$3x+7y=0$,$5x+2y=0$,

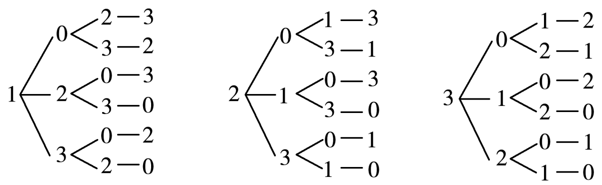
$5x+3y=0$，$5x+7y=0$，$7x+2y=0$，$7x+3y=0$，$7x+5y=0$，共 12 条.

## 任务3 解决简单的排列问题

### 探究活动

**例 2** 由 0，1，2，3 四个数字共能组成多少个没有重复数字的四位数？试全部列出.

**解**：画出树状图如下：



由树状图可知，所有没有重复数字的四位数为 1 023，1 032，1 203，1 230，1 302，1 320，2 013，2 031，2 103，2 130，2 301，2 310，3 012，3 021，3 102，3 120，3 201，3 210，共有 18 个.

［一题多思］

**思考 1.**能组成多少个没有重复数字的四位偶数？

**解**：可分为两类：第一类，个位为 0，有 $3\times2\times1=6$ 个；第二类，个位为 2，有 $2\times2\times1=4$ 个.

综上所述，没有重复数字的四位偶数共有 $6+4=10$ 个.

**思考 2.**能组成多少个四位偶数？

**解**：由题意可知，个位为 0 或 2，所以共有 $2\times3\times4\times4=96$ 个四位偶数.

**【探究总结】**

对于简单的排列问题，解题时可借助分步乘法计数原理或分类加法计数原理，采用元素分析法或位置分析法求解.

### 应用迁移

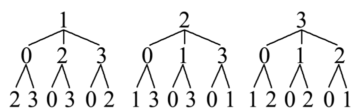
**1.**某信号兵用红、黄、蓝 3 面旗从上到下挂在竖直的旗杆上表示信号，每次可以任挂 1 面、2 面或 3 面，并且不同的顺序表示不同的信号，则一共可以表示________种不同的信号.

15 **解析**：第 1 类，挂 1 面旗表示信号，有 3 种不同的信号；第 2 类，挂 2 面旗表示信号，有 $3\times2=6$ 种不同的信号；第 3 类，挂 3 面旗表示信号，有 $3\times2\times1=6$ 种不同的信号.

根据分类加法计数原理，可以表示的信号共有 $3+6+6=15$ 种.

**2.**从 0，1，2，3 这四个数字中，每次取出三个不同的数字排成一个三位数，能组成多少个不同的三位数？写出这些三位数.

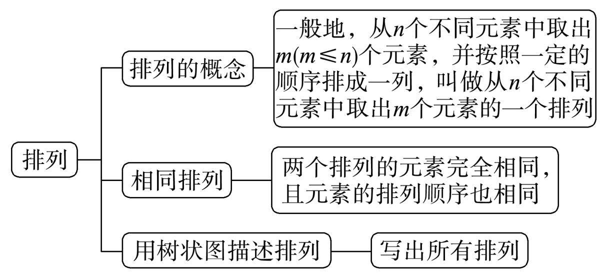
**解**：画出树状图，如图：



由树状图知，符合条件的三位数共有 18 个，

它们是 102，103，120，123，130，132，201，203，210，213，230，231，301，302，310，312，320，321.

### 提质归纳




# 课后素养评价（三） 排列

## A组 学习·理解

**1.**（多选）下列问题中，不是排列问题的是 （　　）

A.由 1，2，3 三个数字组成无重复数字的三位数

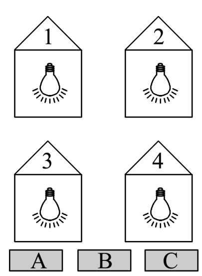
B.从 40 人中选 5 人组成篮球队

C.从 100 人中选 2 人抽样调查

D.从 1，2，3，4，5，6 中选 3 个数组成集合

BCD **解析**：选项 A 中组成的三位数与数字的排列顺序有关，选项 B，C，D 只需取出元素即可，与元素的排列顺序无关.

**2.**如图，A，B，C 三个开关控制着编号为 1，2，3，4 的四盏灯，其中开关 A 控制着编号为 2，3，4 的三盏灯，开关 B 控制着编号为 1，3，4 的三盏灯，开关 C 控制着编号为 1，2，4 的三盏灯.开始时，四盏灯都亮着.现先后按动 A，B，C 这三个开关中的两个开关，则使 1 号灯或 2 号灯亮的按动方法有 （　　）




A.3 种　　B.4 种　　C.5 种　　D.6 种

B **解析**：先后按动 A，B，C 中的两个不同的开关，有 $3\times2=6$ 种方法，分别记为(A，B)，(A，C)，(B，

A),(B,C),(C,A),(C,B).
若要 1 号灯亮,则按第一个开关时,1 号灯灭,按第二个开关时,1 号灯亮,
此时对应的方法有 2 种:(B,C),(C,B);
若要 2 号灯亮,同理可得有 2 种方法:(A,C),(C,A).
综上,要使 1 号灯或 2 号灯亮的按动方法有 $2+2=4$ 种.
故选 B.

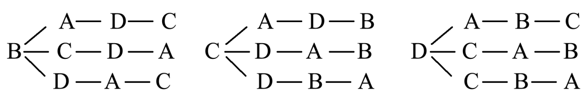
**3.** 6 名学生排成两排,每排 3 人,则不同的排法种数为（　　）

A.36　　B.120　　C.720　　D.240

C　**解析**:由于 6 人排两排,先排第一排,共有 $6\times5\times4=120$ 种排法;再排第二排,共有 $3\times2\times1=6$ 种排法.由分步乘法计数原理可知,共有 $120\times6=720$ 种不同的排法.

**4.** 元旦来临之际,某寝室四名同学各有一张贺卡,要求每位同学将自己的贺卡送给该寝室的另一名同学,且每人都必须得到一张,则不同的送法有________种.

9　**解析**:将 4 张贺卡分别记为 A,B,C,D,且按题意进行排列,用树状图表示为:

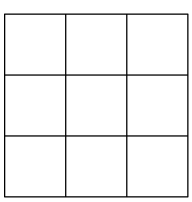



由此可知共有 9 种不同的送法.

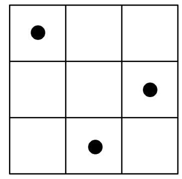
**5.** 从数字 1,3,5,7 中任选两个数做除法,可得不同的商________个.

12　**解析**:这是一个排列问题,可以先排分母,共有 4 种不同的排法,再排分子,共有 3 种不同的排法.由分步乘法计数原理可知,共有 $4\times3=12$ 种不同的排法.又 12 种排法得到的分数值均不相同,因此可得不同的商共有 12 个.

**6.** 将红、黄、蓝三种颜色的三颗棋子分别放入如图所示的 $3\times3$ 方格图中的三个方格内,要求任意两颗棋子不同行、不同列,且不在 $3\times3$ 方格图所在正方形的同一条对角线上,则不同放法共有________种.



24　**解析**:要想任意两颗棋子不在同一行、同一列和同一条对角线上,则三颗棋子必有一颗在正方形方格的顶点,另两颗在对角顶点的两侧,如图所示.由于正方形有四个顶点,故有四个不同的相对位置,又三颗棋子颜色不同,故不同的放法共有 $4\times3\times2\times1=24$ 种.



**7.** 学校乒乓球团体比赛采用 5 场 3 胜制(5 场单打),每支球队派 3 名运动员参赛,前 3 场比赛每名运动员各出场 1 次,其中第 1,2 位出场的运动员在后 2 场比赛中还各出场 1 次.

(1)从 5 名运动员中选 3 名参加比赛,前 3 场比赛有几种出场情况?

(2)甲、乙、丙 3 名运动员参加比赛,所有可能的出场情况有多少种?

**解**:(1)可看作是从 5 名运动员中选 3 名进行排列,则前 3 场单打比赛的出场情况有 $5\times4\times3=60$ 种.

(2)可分为三类:

第一类,3 场决胜负,有 $3\times2\times1=6$ 种出场情况,分别为甲乙丙,甲丙乙,乙甲丙,乙丙甲,丙甲乙,丙乙甲;

第二类,4 场决胜负,有 $3\times2\times1\times2=12$ 种出场情况,分别为甲乙丙甲,甲乙丙乙,甲丙乙甲,甲丙乙丙,乙甲丙乙,乙甲丙甲,乙丙甲丙,乙丙甲乙,丙甲乙丙,丙甲乙甲,丙乙甲乙,丙乙甲丙;

第三类,5 场决胜负,有 $3\times2\times1\times2\times1=12$ 种出场情况,分别为甲乙丙甲乙,甲乙丙乙甲,甲丙乙甲丙,甲丙乙丙甲,乙甲丙乙甲,乙甲丙甲乙,乙丙甲丙乙,乙丙甲乙丙,丙甲乙丙甲,丙甲乙甲丙,丙乙甲乙丙,丙乙甲丙乙.

因此,所有可能的出场情况共有 $6+12+12=30$ 种.

## B组　应用·实践

**1.** (多选)下列问题是排列问题的是（　　）

A.从甲、乙、丙三名同学中选出两名分别参加数学、物理兴趣小组

B.从甲、乙、丙三名同学中选出两名参加一项活动

C.从 $a,b,c,d$ 中选出 3 个字母

D.从 1,2,3,4,5 中取出 2 个数字组成两位数

AD　**解析**:对于 A,从甲、乙、丙三名同学中选出两名分别参加数学、物理兴趣小组,与顺序有关,是排列问题;对于 B,从甲、乙、丙三名同学中选出两名参加一项活动,只要求选出即可,不是排列问题;对于 C,从 $a,b,c,d$ 中选出 3 个字母,只要求选出即可,不是排列问题;对于 D,从 1,2,3,4,5 中取出 2 个数字组成两位数,需要先选出再排序,是排列问题.故选 AD.

**2.**小明、小红、小强 3 名同学随机排成一排照相，则小明站在小红左边的概率是 （ ）

A.$\frac{1}{6}$ B.$\frac{1}{3}$

C.$\frac{1}{2}$ D.$\frac{2}{3}$

C **解析**：记小明、小红、小强分别为 a，b，c，排成一排的情况有 6 种，分别为 abc，bac，acb，cab，bca，cba，小明站在小红左边共有 3 种情况，分别为 abc，acb，cab.所以小明站在小红左边的概率 $p=\frac{3}{6}=\frac{1}{2}$.

**3.**用数字 1，2，3，4，6 组成的无重复数字的五位偶数有 （ ）

A.48 个 B.64 个 C.72 个 D.90 个

C **解析**：先排个位，需要从 2，4，6 三个数字中任意选出一个数字，共有 3 种方法；再排前四位，共有 $4\times3\times2\times1=24$ 种方法.由分步乘法计数原理可知，共有 $3\times24=72$ 种方法.

**4.**甲、乙、丙三人踢毽子，互相传递，每人每次只能踢一下.由甲开始踢，经过 4 次传递后，毽子又被踢回甲，则不同的传递方式共有 （ ）

A.4 种 B.5 种 C.6 种 D.12 种

C **解析**：若甲先传给乙，则有甲→乙→甲→乙→甲，甲→乙→甲→丙→甲，甲→乙→丙→乙→甲 3 种不同的传递方式；同理，甲先传给丙也有 3 种不同的传递方式.故共有 6 种不同的传递方式.

**5.**从 3，5，7，11 这四个质数中，每次取出两个不同的数分别记为 $a$，$b$，共可得到 $\lg a-\lg b$ 的不同值的个数是______.

12 **解析**：由于 $\lg a-\lg b=\lg\frac{a}{b}$，所以从 3，5，7，11 中取出两个不同的数分别赋值给 $a$ 和 $b$，为排列问题，不同的取法共有 $4\times3=12$ 种，并且计算结果不会重复，所以得到不同的值有 12 个.

**6.**从集合$\{0,1,2,5,7,9,11\}$中任取 3 个元素分别作为直线方程 $Ax+By+C=0$ 中的参数 $A$，$B$，$C$ 的值，所得直线经过坐标原点的有________条.

30 **解析**：易知过原点的直线方程的常数项为 0，则 $C=0$；再从集合中任取两个非零元素作为系数 $A$，$B$ 的值，属于排列问题，所以符合条件的直线有 $6\times5=30$ 条.

**7.**在三位数中，如果十位上的数字比百位上的数字和个位上的数字都小，那么这个数为凹数，如 524，746 等都是凹数.那么用 0，1，2，3，4 这五个数字能组成多少个无重复数字的凹数？请列举出来.

**解**：符合要求的凹数可分为三类.

第 1 类，十位数字为 0 的凹数有 102，103，104，201，203，204，301，302，304，401，402，403，共 12 个；

第 2 类，十位数字为 1 的凹数有 213，214，312，314，412，413，共 6 个；

第 3 类，十位数字为 2 的凹数有 324，423，共 2 个.

所以由 0，1，2，3，4 可组成 $12+6+2=20$（个）无重复数字的凹数.

# 6.2.2　排列数

**学习任务目标**

1. 理解排列数的概念,能利用计数原理推导排列数公式.
2. 能熟练应用排列数公式及性质求与排列数有关的量,并能证明恒等式,求方程的解及不等式的解.
3. 能准确用排列数公式表示排列的关系,并能应用排列数公式解决与排列有关的实际问题.

## ○ 问题式预习 ○

### 【知识清单】

**知识点　排列数与排列数公式**

(1)排列数:我们把从 $n$ 个不同元素中取出 $m(m\leqslant n)$ 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:$A_n^m=n(n-1)(n-2)\cdots(n-m+1)$,其中 $m,n\in\mathbf{N}^*$,并且 $m\leqslant n$.

(3)全排列和阶乘:

①全排列:我们把 $n$ 个不同的元素全部取出的一个排列,叫做 $n$ 个元素的一个全排列.这时,排列数公式中 $m=n$,即有 $A_n^n=n\times(n-1)\times(n-2)\times\cdots\times3\times2\times1$.

②阶乘:正整数 1 到 $n$ 的连乘积,叫做 $n$ 的阶乘,用 $n!$ 表示.于是,$n$ 个元素的全排列数公式可以写成 $A_n^n=n!$.

规定 $0!=1$.

因此,排列数公式还可以写成阶乘式:$A_n^m=\dfrac{n!}{(n-m)!}$.

### 【概念辨析】

1. 判断正误(正确的打"√",错误的打"×").

(1)等式 $A_n^m=nA_{n-1}^{m-1}$ 成立.　( √ )

(2)排列数 $A_n^m$ 中,$n,m$ 满足 $m\leqslant n$.　( √ )

(3)$A_n^m=mA_{n-1}^{m-1}+A_{n-1}^m$.　( √ )

2. $90\times91\times92\times\cdots\times100$ 可以表示为　(　　)

A. $A_{100}^{10}$　　B. $A_{100}^{11}$

C. $A_{100}^{12}$　　D. $A_{100}^{13}$

B　**解析**:由排列数公式知可表示为 $A_{100}^{11}$.故选 B.

3. 把 5 本不同的课外读物分给 5 名同学,每人一本,则不同的分配方法有________种.

120　**解析**:利用排列的概念可知不同的分配方法有 $A_5^5=120$ 种.

4. 请思考并回答下列问题:

(1)排列与计数原理的关系是怎样的?

**提示**:①排列是利用分步乘法计数原理,从 $n$ 个不同元素中选出 $m$ 个,排在 $m$ 个位置上,分 $m$ 步完成.

②排列问题可以用分步乘法计数原理解决,但排列数公式将问题公式化,更简洁.

(2)排列与排列数有什么区别?

**提示**:排列与排列数是两个不同的概念,"排列"是指从 $n$ 个不同元素中取出 $m(m\leqslant n)$ 个元素按照一定顺序排成一列,是一种排法;"排列数"是指从 $n$ 个不同元素中取出 $m(m\leqslant n)$ 个元素所得不同排列的个数,是一个数值,用 $A_n^m$ 表示.

## ○ 任务型课堂 ○

### 任务 1 排列数的定义及排列数的计算

1. 计算:(1)$A_{12}^3=$________;

(2)$\dfrac{2A_8^5+7A_8^4}{A_8^8-A_9^5}=$________.

(1)1 320　(2)1　**解析**:(1)原式 $=12\times11\times10=1\ 320$.

(2)原式 $=\dfrac{2\times8\times7\times6\times5\times4+7\times8\times7\times6\times5}{8\times7\times6\times5\times4\times3\times2\times1-9\times8\times7\times6\times5}$

$=\dfrac{8\times7\times6\times5\times(8+7)}{8\times7\times6\times5\times(24-9)}=1$.

2. (1)已知 $A_x^2=20$,则 $x=$________;

(2)$(55-n)(56-n)\cdots(69-n)(n\in\mathbf{N}^*$ 且 $n<55)$ 可用排列数表示为________.

(1)5　(2)$A_{69-n}^{15}$　**解析**:(1)$A_x^2=x(x-1)=20$,解得 $x=5$ 或 $x=-4$.

由 $A_x^2$ 的意义知 $x\in\mathbf{N}^*$ 且 $x\geqslant2$,所以 $x=5$.

(2)因为 $55-n,56-n,\cdots,69-n$ 中最大的数为 $69-n$,且共有 $(69-n)-(55-n)+1=15$ 个数,

所以 $(55-n)(56-n)\cdots(69-n)=A_{69-n}^{15}$.

**【探究总结】**

公式 $A_n^m=n(n-1)(n-2)\cdots(n-m+1)$ $(n,m\in\mathbf{N}^*,m\leqslant n)$ 适用于具体计算及解含有排列数的方程和不等式.

## 任务2 排列数公式阶乘式的应用

**1.**化简：$\dfrac{(m-1)!}{A_{m-1}^{n-1}(m-n)!}=$____.

**2.**解方程：$3A_8^x=4A_9^{x-1}$.

**解**：原方程可化为 $3\times\dfrac{8!}{(8-x)!}=4\times\dfrac{9!}{(10-x)!}$，

化简得 $3=\dfrac{4\times9}{(10-x)(9-x)}$，

即 $x^2-19x+78=0$，解得 $x_1=6$，$x_2=13$.

又 $\begin{cases}0<x\leqslant8,\\0<x-1\leqslant9,\end{cases}$ 解得 $1<x\leqslant8$.

故原方程的解是 $x=6$.

**【探究总结】**

公式 $A_n^m=\dfrac{n!}{(n-m)!}$ 适用于与排列数有关的证明、解方程、解不等式等.在应用时，应注意先提取公因式再计算，同时要注意隐含条件（$m\leqslant n$ 且 $n\in\mathbf{N}^*$，$m\in\mathbf{N}^*$）的应用.

## 任务3 有限制条件的排列问题

### 探究活动

**例** 现有3个女生和5个男生排成一排.如果其中的甲、乙两人必须站两端，有多少种不同的排法？

**解**：甲、乙为特殊元素，先将他们排在两端位置，有 $A_2^2$ 种排法，其余6人全排列，有 $A_6^6$ 种排法，所以共有 $A_2^2A_6^6=1\ 440$ 种不同的排法.

［一题多思］

**思考1.**如果甲不站最左边，乙不站最右边，有多少种不同的排法？

**解**：甲、乙为特殊元素，左、右两端为特殊位置.

（方法一：特殊元素法）甲在最右边时，其他的人可全排列，有 $A_7^7$ 种排法；甲不在最右边时，可从余下6个位置中任选一个，有 $A_6^1$ 种排法，而乙可排在除去最右边位置和甲的位置后剩余的6个位置中的任一个上，有 $A_6^1$ 种排法，其余6人全排列，共有 $A_6^1A_6^1A_6^6$ 种排法.由分类加法计数原理知，共有 $A_7^7+A_6^1A_6^1A_6^6=30\ 960$ 种不同的排法.

（方法二：特殊位置法）先排最左边，除去甲外，有 $A_7^1$ 种排法，余下7个位置全排列，有 $A_7^7$ 种排法，但应剔除乙在最右边时的 $A_6^1A_6^6$ 种排法，所以共有 $A_7^1A_7^7-A_6^1A_6^6=30\ 960$ 种不同的排法.

（方法三：间接法）8人全排列，共 $A_8^8$ 种排法.其中，不符合条件的有甲在最左边时的 $A_7^7$ 种排法，乙在最右边时的 $A_7^7$ 种排法，其中都包含了甲在最左边，同时乙在最右边的情形，共 $A_6^6$ 种排法，所以共有 $A_8^8-2A_7^7+A_6^6=30\ 960$ 种不同的排法.

**思考2.**如果女生全排在一起，有多少种不同的排法？

**解**：（捆绑法）由于女生全排在一起，可把她们看成一个整体，这样同5个男生合在一起有6个元素，排成一排有 $A_6^6$ 种排法，而每一种排法中，3个女生间又有 $A_3^3$ 种排法，因此共有 $A_6^6A_3^3=4\ 320$ 种不同的排法.

**思考3.**如果女生不相邻，有多少种不同的排法？

**解**：（插空法）先将5名男生排好，共有 $A_5^5$ 种排法，在这5名男生中间以及两边的6个空位中插入3名女生，共有 $A_6^3$ 种排法，

由分步乘法计数原理，共有 $A_5^5\times A_6^3=120\times120=14\ 400$ 种不同的排法.

**思考4.**女生已站好，然后男生站入队伍，但不能改变女生的相对顺序，则不同的排法有多少种？

**解**：不同的排法有 $\dfrac{A_8^8}{A_3^3}=6\ 720$ 种.

**【探究总结】**

解决排列应用问题的5种主要方法

| | |
|---|---|
| 直接法 | 由符合条件的排列直接列式计算 |
| 优先法 | 优先排列特殊元素或特殊位置 |
| 捆绑法 | 把相邻元素看作一个整体与其他元素一起排列，同时注意捆绑元素的内部排列 |
| 插空法 | 对于不相邻问题，先考虑不受限制的元素的排列，再将不相邻的元素插在前面元素排列的空档中 |
| 间接法 | 正难则反，等价转化 |

### 应用迁移

**1.**（多选）A，B，C，D，E五个人并排站在一起，则下列说法正确的有（　　）

A.若A，B两人站在一起有48种排法

B.若A，B不相邻共有12种排法

C.若A在B左边有60种排法

D.若A不站在最左边，B不站在最右边，有72种排法

AC　**解析**：对于选项A，若A，B两人站在一起，则有 $A_2^2A_4^4=48$ 种排法，故A正确；

对于选项B，A，B，C，D，E五个人并排站在一起，则有 $A_5^5=120$ 种排法，所以A，B不相邻共有 $120-48=72$ 种排法，故B错误；

对于选项C，根据对称性可知A在B左边有 $\dfrac{120}{2}=60$ 种排法，故C正确；

对于选项D，A站在最左边，有 $A_4^4=24$ 种排法，B站在最右边，有 $A_4^4=24$ 种排法，A站在最左边，同时B站在最右边，有 $A_3^3=6$ 种排法，所以A不站在

最左边，B 不站在最右边，有 $120-24-24+6=78$ 种排法，故 D 错误.

故选 AC.

**2.** 将 0，1，2，3，4 这五个数字组成无重复数字的五位数，

(1)可以组成多少个偶数？

(2)可以组成多少个比 13 123 大的数？

**解**：(1)当个位数字为 0 时，可以组成 $A_4^4=24$ 个偶数；

当个位数字不为 0 时，可以组成 $2\times A_3^1A_3^3=36$ 个偶数.

所以一共可以组成 $24+36=60$ 个偶数.

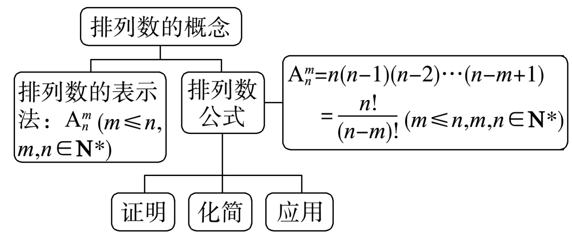
(2)所组成的比 13 123 大的五位数，可以分为以下 2 类：

第一类：形如 2□□□□，3□□□□，4□□□□，共有 $3A_4^4=72$ 个；

第二类：形如 13□□□，14□□□，共有 $2A_2^2+A_3^3=10$ 个.

所以一共可以组成 $72+10=82$ 个比 13 123 大的数.

### 提质归纳




## 课后素养评价（四） 排列数

### A组 学习·理解

**1.** $(n-1\ 998)(n-1\ 999)\cdots(n-2\ 023)(n-2\ 024)(n\in\mathbf{N},n>2\ 024)$可表示为 ( )

A. $A_{n-1\ 998}^{27}$ B. $A_{n-1\ 998}^{26}$

C. $A_{n-2\ 024}^{27}$ D. $A_{n-2\ 024}^{26}$

A **解析**：$(n-1\ 998)(n-1\ 999)\cdots(n-2\ 023)(n-2\ 024)$中总共有$(n-1\ 998)-(n-2\ 024)+1=27$ 个数连乘，故$(n-1\ 998)(n-1\ 999)\cdots(n-2\ 023)(n-2024)=A_{n-1\ 998}^{27}$.故选 A.

**2.** (2023·全国甲卷)有五名志愿者参加社区服务，共服务星期六、星期天两天，每天从中任选两人参加社区服务，则恰有 1 人连续参加两天社区服务的选择种数为 ( )

A. 120 B. 60 C. 30 D. 20

B **解析**：不妨记五名志愿者为 a，b，c，d，e，

假设 a 连续参加两天社区服务，再从剩余的 4 人中抽取 2 人各参加星期六与星期天的社区服务，共有 $A_4^2=12$ 种方法.

同理，若 b，c，d，e 连续参加两天社区服务，也各有 12 种方法.

所以恰有 1 人连续参加两天社区服务的选择种数为 $5\times12=60$.故选 B.

**3.** 某班上午有 5 节课，分别安排语文、数学、英语、物理、化学各 1 节课，要求语文与化学相邻，且数学不排在第一节课，则不同的排课法的种数是 ( )

A. 36 B. 24 C. 18 D. 12

A **解析**：将语文和化学捆绑，与英语、物理全排列有 $A_3^3A_2^2$ 种排法；数学不排在第一节课，将数学插空有 3 种排法.由分步乘法计数原理可得不同的排课法的种数是 $A_3^3A_2^2\times3=36$.故选 A.

**4.** 不等式 $A_8^{x+2}<6A_8^x$ 的解集为________.

{6} **解析**：由原不等式得

$$\begin{cases}\dfrac{8!}{[8-(x+2)]!}<6\times\dfrac{8!}{(8-x)!},\text{且 }x\in\mathbf{N}^*,\\ x+2\leqslant8\end{cases}$$

所以$\begin{cases}(8-x)(7-x)<6,\\ x\leqslant6,\end{cases}$即$\begin{cases}x^2-15x+50<0,\\ x\leqslant6,\end{cases}$解得 $5<x\leqslant6$ 且 $x\in\mathbf{N}^*$，

所以 $x=6$.

**5.** 两个家庭的 4 个大人与 2 个小孩一起到动物园游玩，购票后排队依次入园.为安全起见，首尾一定要排 2 个爸爸，另外，2 个小孩一定要排在一起，则这 6 人入园顺序有________种.

24 **解析**：第 1 步，将 2 个爸爸排在两端，有 2 种排法；第 2 步，将 2 个小孩视为一人与 2 个妈妈任意排在中间的三个位置上，有 $A_3^3$ 种排法；第 3 步，将 2 个小孩排序，有 2 种排法.故总的入园顺序有 $2\times2\times A_3^3=24$ 种.

**6.** (2023·全国乙卷改编)甲、乙两名同学从 6 种课外读物中各自选读 2 种，则这两人选读的课外读物中恰有 1 种相同的选法共有________种.

120 **解析**：首先确定一种相同的课外读物，共有 6 种情况；

然后两人各自的另外一种课外读物相当于在剩余的 5 种课外读物里，选出 2 种进行排列，共有 $A_5^2$ 种选法.

根据分步乘法计数原理，则共有 $6\times A_5^2=120$ 种选法.

**7.** 3 名男生、4 名女生排成一行.在下列要求下，分别求不同排列方法的种数：

(1)甲不在最左边，乙不在最右边；

(2)男生必须排在一起；

(3)男生和女生相间排列；

(4)在甲、乙两人中间必须有 3 人.

**解**：(1)依题意，先排最左边，除去甲外，有 $A_6^1$ 种排法，余下的 6 个位置全排列有 $A_6^6$ 种排法，

但应剔除其中乙在最右边的排法 $A_5^1A_5^5$ 种，

则符合条件的排法共有 $A_6^1A_6^6-A_5^1A_5^5=3\ 720$ 种.

(2)将男生看成一个整体，进行全排列，有 $A_3^3$ 种排法，

再与其他元素进行全排列，有 $A_5^5$ 种排法，

故共有 $A_3^3A_5^5=720$ 种不同的排法.

(3)先排好男生，然后将女生插入男生所形成的四个空位，共有 $A_3^3A_4^4=144$ 种不同的排法.

(4)从除甲、乙以外的 5 人中选 3 人排在甲、乙中间的排法有 $A_5^3$ 种，

将甲、乙看作一个整体，和其余 2 人排成一排的排法有 $A_2^2A_3^3$ 种，

最后再把选出的 3 人插入到甲、乙之间即可，

共有 $A_5^3A_2^2A_3^3=720$ 种不同的排法.

## B组 应用·实践

**1.** 下列各式中与排列数 $A_n^m$ 相等的是 （　　）

A. $\dfrac{n!}{(n-m+1)!}$

B. $n(n-1)(n-2)\cdots(n-m)$

C. $\dfrac{nA_{n-1}^{m-1}}{n-m+1}$

D. $A_n^1A_{n-1}^{m-1}$

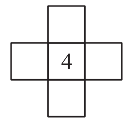
D　**解析**：$A_n^m=\dfrac{n!}{(n-m)!}=n(n-1)(n-2)\cdots(n-m+1)$，故 A，B 错误；而 $A_n^1A_{n-1}^{m-1}=nA_{n-1}^{m-1}=n\cdot\dfrac{(n-1)!}{(n-m)!}=\dfrac{n!}{(n-m)!}$，故 C 错误，D 正确.

**2.** 五声音阶是中国古乐的基本音阶，故有成语“五音不全”.中国古乐中的五声音阶依次为宫、商、角、徵、羽.如果用这五个音阶排成一个五音阶音序，且商、角不相邻，徵位于羽的左侧，那么可排成的不同音序有 （　　）

A. 18 种　B. 24 种　C. 36 种　D. 72 种

C　**解析**：先将宫、徵、羽三个音节进行排序，且徵位于羽的左侧，再将商、角插入 4 个空中.所以共有 $3A_4^2=36$ 种不同的音序.故选 C.

**3.** 如图，从 1，2，3，5，6，7，8，9 中选 4 个数填入空格，要求从左至右、从上至下所填的数字都是从大到小排列的，则不同的填法种数为 （　　）




A. 70　B. 120　C. 140　D. 144

B　**解析**：比 4 小的有 1，2，3，共 3 个，从中选出 2 个排在 4 的右边和下方，方法数有 $A_3^2$ 种；

比 4 大的有 5，6，7，8，9，共 5 个，从中选出 2 个排在 4 的左边和上方，方法数有 $A_5^2$ 种.

所以不同的填法为 $A_3^2A_5^2=120$ 种.

**4.** 一条铁路有 $n$ 个车站，为适应客运需要，新增了 $m$ 个车站，且知 $m>1$，客运车票增加了 62 种，则现在车站的个数为 （　　）

A. 15　B. 16　C. 17　D. 18

C　**解析**：原来 $n$ 个车站有 $A_n^2$ 种车票，新增了 $m$ 个车站，有 $A_{n+m}^2$ 种车票.

由题意得 $A_{n+m}^2-A_n^2=62$，即 $(n+m)(n+m-1)-n(n-1)=62$，

整理得 $2mn+m^2-m=62$，所以 $n=\dfrac{31}{m}-\dfrac{m-1}{2}$.

因为 $m>1$，$n>0$，所以 $\dfrac{31}{m}>\dfrac{m-1}{2}$，所以 $m^2-m-62<0$，解得 $1<m<\dfrac{1+\sqrt{249}}{2}$，即 $1<m\leqslant 8$.

当 $m=3,4,5,6,7,8$ 时，$n$ 均不为整数，只有当 $m=2$ 时，$n=15$ 符合题意.

所以 $m+n=17$，故现在有 17 个车站.故选 C.

**5.** 把 5 件不同的产品摆成一排.若产品 A 与产品 B 相邻，且产品 A 与产品 C 不相邻，则不同的摆法有________种.

36　**解析**：先考虑产品 A 与 B 相邻，把 A，B 作为一个元素有 $2A_4^4=48$ 种摆法.又 A，B 相邻且 A，C 相邻，有 $2A_3^3=12$ 种摆法，故满足条件的不同的摆法有 $48-12=36$ 种.

**6.** 用 1，2，3，4，5，6，7 组成没有重复数字的七位数.若 1，3，5，7 的顺序一定，则有________个七位数符合条件.

210　**解析**：因为 1，3，5，7 四个数有 $A_4^4=24$ 种排法，所以 1，3，5，7 的顺序一定的排法数只占总排法

数的$\frac{1}{24}$.故有$\frac{1}{24}\times A_7^7=210$个七位数符合条件.

**7.**甲、乙、丙、丁、戊5名学生进行劳动技术比赛,决出第1名到第5名.赛后,甲、乙两名参赛者去询问成绩,回答者对甲说:"很遗憾,你和乙都没有得到冠军."又对乙说:"你的成绩不是最差的."则这5人可能的排名情况有多少种?

**解**:根据题意,乙可能是第二或第三或第四名,故有$A_3^1$种可能的排名情况;

乙排好再排甲,共有$A_3^1$种可能的排名情况;

剩下3人全排列共有$A_3^3$种可能的排名情况.

故共有$A_3^1A_3^1A_3^3=3\times3\times6=54$种可能的排名情况.

**8.**从1～9这9个数字中取出5个数进行排列.

(1)奇数位置上是奇数的有多少种排法?

(2)取出的奇数必须排在奇数位置上有多少种排法?

**解**:(1)奇数共5个,奇数位置共有3个;偶数共有4个,偶数位置有2个.第一步先在奇数位置上排上奇数共有$A_5^3$种排法;第二步再排偶数位置,有4个偶数和余下的2个奇数可以排,排法为$A_6^2$种.由分步乘法计数原理知,排法种数为$A_5^3A_6^2=1\ 800$.

(2)因为偶数位置上不能排奇数,所以先排偶数位,排法为$A_4^2$种,余下的2个偶数与5个奇数都可以排在奇数位置上,排法为$A_7^3$种.由分步乘法计数原理知,排法种数为$A_4^2A_7^3=2\ 520$.

## 6.2.3　组合

**学习任务目标**

**1.**通过实例,理解组合的概念,正确认识组合与排列的区别与联系.

**2.**能用列举法、顺序后移法、树状图法列出简单的组合.

## ○ 问题式预习 ○

**【知识清单】**

**知识点　组合的概念**

(1)定义:一般地,从$n$个不同元素中取出$m(m\leqslant n)$个元素作为一组,叫做从$n$个不同元素中取出$m$个元素的一个组合.

(2)两个组合相同的条件:两个组合只要元素相同,不论元素的顺序如何,都是相同的.

**【概念辨析】**

**1.**判断正误(正确的打"√",错误的打"×").

(1)组合与所选出的元素的排列顺序有关.　(×)

(2)若两个组合的元素相同,则这两个组合是相同的.　(√)

(3)"从6名学生中选3名学生参加元旦晚会演出,共有多少种选法?"属于组合问题.　(√)

(4)"从30名学生中选2名学生分别担任班长、学习委员两个职位,共有多少种选法?"属于组合问题.　(×)

**2.**请思考并回答下列问题:

(1)你能说一说排列与组合之间的联系与区别吗?

**提示**:从排列与组合的定义可以知道,两者都是从$n$个不同元素中取出$m(m\leqslant n)$个元素,这是它们的共同点.但排列与元素的顺序有关,而组合与元素的顺序无关.只有元素相同且顺序也相同的两个排列才是相同的;而两个组合只要元素相同,不论元素的顺序如何,都是相同的.

(2)你能举一些生活中与组合相关的例子吗?

**提示**:班级选5名同学参加学校组织的活动;有10辆共享单车,随机选3辆进行安全检测等.

## ○ 任务型课堂 ○

### 任务1 组合的概念及其判断

**1.**从2,3,4,5四个数中任取2个数.若分别作为对数式$\log_a b$的底数与真数,求得到的对数的个数,则是________问题;若求两个数相乘得到的积有几种情况,则是________问题.(用"排列""组合"填空)

排列　组合　**解析**:对数式$\log_a b$的值,与$a,b$取值的顺序有关,属于排列问题;$ab$的值与$a,b$取值的顺序无关,属于组合问题.

**2.**判断下列各事件是排列问题还是组合问题.

(1)10人相互通一次电话,共通多少次电话?

(2)10支球队以单循环形式进行比赛(每两队比赛一次),共进行多少场次?

(3)10支球队以单循环形式进行比赛,比赛冠、亚军获得者有多少种可能?

(4)从 10 名职工中选出 3 名代表去开会，有多少种选法？

(5)从 10 名学生中选出 3 个不同学科的课代表，有多少种选法？

**解**：(1)因为甲与乙通了一次电话，也就是乙与甲通了一次电话，没有顺序的区别，所以这是一个组合问题.

(2)因为每两支球队比赛一次，并不需要考虑谁先谁后，没有顺序的区别，所以这是一个组合问题.

(3)因为甲队获得冠军、乙队获得亚军，与乙队获得冠军、甲队获得亚军是不一样的，与顺序有关，所以这是一个排列问题.

(4)因为 3 名代表之间没有顺序上的区别，所以这是一个组合问题.

(5)因为 3 名学生中，担任哪一科的课代表是有顺序上的区别的，所以这是一个排列问题.

**【探究总结】**

**排列、组合辨析切入点**

(1)组合的特点是"只选不排"，即组合只是从 $n$ 个不同的元素中取出 $m(m\leqslant n)$ 个元素即可.只要两个组合中的元素相同就是相同的组合，不需考虑顺序.

(2)判断组合与排列的依据是看是否与顺序有关，与顺序有关的是排列问题，与顺序无关的是组合问题.

## 任务 2 写出简单组合问题的所有组合

### 探究活动

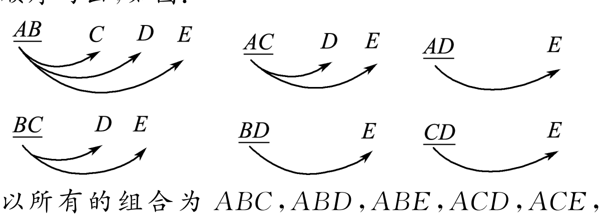
**例** (1)从 5 个不同的元素 $A,B,C,D,E$ 中取出 2 个，列出所有的组合；

(2)从 5 个不同的元素 $A,B,C,D,E$ 中取出 3 个，列出所有的组合.

**解**：(1)如图所示.

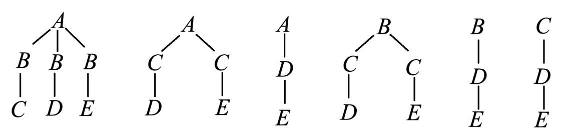

故所有的组合为 $AB,AC,AD,AE,BC,BD,BE,CD,CE,DE$.

(2)(方法一)可按 $AB\rightarrow AC\rightarrow AD\rightarrow BC\rightarrow BD\rightarrow CD$ 的顺序写出，如图.




所以所有的组合为 $ABC,ABD,ABE,ACD,ACE,ADE,BCD,BCE,BDE,CDE$.

(方法二)画出树状图，如图所示.



由此可以写出所有的组合为 $ABC,ABD,ABE,ACD,ACE,ADE,BCD,BCE,BDE,CDE$.

[一题多思]

**思考 1.**观察(1)(2)的结果，你有什么发现？

**提示**：(2)中取出组合 $ABC$ 后，剩下的 $DE$ 正好是(1)中的一个组合，其他组合也是一一对应的.

**思考 2.**这种发现给你什么启示？

**提示**：当取出的元素个数较多时，直接写取出元素的组合往往比较复杂，这时我们可以找它的对立面——"剩下的元素"有哪些，进而得到取出元素的组合.

**思考 3.**从 5 个不同的元素 $A,B,C,D,E$ 中取出 4 个，列出所有的组合.

**解**：(方法一)"顺序后移法".所有的组合为 $ABCD,ABCE,ABDE,ACDE,BCDE$.

(方法二)剩下 $A$，则取出 $BCDE$；剩下 $B$，则取出 $ACDE$……共有 $BCDE,ACDE,ABDE,ABCE,ABCD$ 这 5 个组合.

**【探究总结】**

**1.**写出从 $n$ 个不同元素中取出 $m(m\leqslant n)$ 个元素的所有组合的方法：从 $n$ 个不同元素中取出 $m$ 个元素的组合，可借助"顺序后移法"或"树状图法"直观地写出，做到不重复、不遗漏.

**2.**两个注意点：

(1)利用"顺序后移法"时，箭头向后逐步推进，且写出的一个组合不必再交换元素位置.如写出 $ab$ 后，不必再交换元素位置为 $ba$，因为它们是同一个组合.

(2)画"树状图"时，应注意顶部元素及分支元素的排列思路，防止重复或遗漏.

### 应用迁移

**1.**从 1,2,3,6,9 中任取两个不同的数相加，列出所有的取法，并求出不同的相加结果的个数.

**解**：从 1,2,3,6,9 中任取两个不同的数，不同的取法有{1,2}，{1,3}，{1,6}，{1,9}，{2,3}，{2,6}，{2,9}，{3,6}，{3,9}，{6,9}，

不同的相加结果有 3,4,5,7,8,9,10,11,12,15，共 10 个.

**2.**从 2 名教师和 5 名学生中，选出 3 人参加"我爱我的祖国"主题活动，要求入选的 3 人中至少有一名教师，不同的选取方案有多少种？

**解**：记 2 名教师为 a,b，5 名学生为 1,2,3,4,5.

所有选取方案为

a12,a13,a14,a15,a23,a24,a25,a34,a35,a45,

b12,b13,b14,b15,b23,b24,b25,b34,b35,b45,

ab1,ab2,ab3,ab4,ab5,共 25 种.

**3.**甲、乙、丙、丁 4 个足球队举行单循环赛,列出:

(1)所有各场比赛的双方;

(2)所有冠、亚军的可能情况.

**解**:(1)所有各场比赛的双方为甲乙、甲丙、甲丁、乙丙、乙丁、丙丁.

(2)所有冠、亚军的可能情况为甲乙、乙甲、甲丙、丙甲、甲丁、丁甲、乙丙、丙乙、乙丁、丁乙、丙丁、丁丙.

## 提质归纳

- 组合
  - 组合的概念:从$n$个不同元素中取出$m(m\leqslant n)$个元素作为一组,叫做从$n$个不同元素中取出$m$个元素的一个组合
  - 相同组合:两个组合只要元素相同,不论元素的顺序如何,都是相同的
  - 排列与组合的关系:相同点:从$n$个不同元素中取出$m(m\leqslant n)$个元素;不同点:排列有顺序,组合无顺序
  - 画树状图列举所有组合

# 课后素养评价(五) 组合

## A组 学习·理解

**1.**(多选)下列问题是组合问题的是 ( )

A.从 10 名学生中任选 5 名去参观一个展览会

B.从 1,2,3,4,5 这 5 个数字中任取 2 个数作为一个点的坐标

C.一个黄袋中装有四张分别写有 1,3,5,7 的卡片,另一个红袋中装有四张分别写有 2,8,16,32 的卡片,从红袋和黄袋中各任取一张卡片,计算这两张卡片上的数的和

D.将四本不同的书分别送给四个人,每人一本

AC **解析**:选项 A,从 10 名学生中任选 5 名去参观一个展览会,选出的学生不用排序,是组合问题;

选项 B,从 1,2,3,4,5 这 5 个数字中,每次任取 2 个不同的数作为一个点的坐标,由于坐标有横、纵坐标之分,所以选出的 2 个不同的数需要排序,是排列问题;

选项 C,从红袋和黄袋中各任取一张卡片,求这两张卡片上的数相加所得的和,因为加法满足交换律,故选出的卡片不用排序,是组合问题;

选项 D,因为四本不同的书送给四个人,要求每人一本,所以这四本书需要排序,是排列问题.

**2.**下列问题中,组合问题的个数是 ( )

①从全班 50 人中选出 5 人组成班委会;

②从全班 50 人中选出 5 人分别担任班长、副班长、团支部书记、学习委员、生活委员;

③从 1,2,3,…,9 中任取两个数求积;

④从 1,2,3,…,9 中任取两个数求差或商.

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

B **解析**:对于①,从 50 人中选出 5 人组成班委会,不考虑顺序,是组合问题.②为排列问题.对于③,从 1,2,3,…,9 中任取两个数求积是组合问题.因为乘法满足交换律,而减法和除法不满足交换律,故④为排列问题.所以组合问题的个数是 2.故选 B.

**3.**下列四个问题属于组合问题的是 ( )

A.从 4 名志愿者中选出 2 人分别担任导游和翻译的工作

B.从 1,2,3,4 这 4 个数字中选取 3 个数字组成三位数

C.从全班同学中选出 3 名同学参加学校运动会开幕式

D.将甲、乙两位同学安排到 A,B 两个座位

C **解析**:对于 A,从 4 名志愿者中选出 2 人分别担任导游和翻译的工作,将 2 人选出后,还要安排导游或翻译的工作,与顺序有关,这个问题为排列问题;

对于 B,从 1,2,3,4 这 4 个数字中选取 3 个数字组成一个三位数,选出 3 个数字之后,还要将这 3 个数字安排至个位、十位、百位这三个数位,与顺序有关,这个问题为排列问题;

对于 C,从全班同学中选出 3 名同学参加学校运动会开幕式,只需将 3 名同学选出,与顺序无关,这个问题为组合问题;

对于 D,两位同学的座位与顺序有关,这个问题为排列问题.

故选 C.

**4.**(1)写出从 $a,b,c,d,e$ 五个元素中任取两个元素的所有组合;

(2)写出从 $a,b,c,d,e$ 五个元素中任取两个元素的所有排列.

**解**：(1)从 $a,b,c,d,e$ 五个元素中任取两个元素的所有组合：

$ab,ac,ad,ae,bc,bd,be,cd,ce,de$.

(2)从 $a,b,c,d,e$ 五个元素中任取两个元素的所有排列：

$(a,b),(a,c),(a,d),(a,e),(b,a),(b,c),(b,d),(b,e),(c,a),(c,b),(c,d),(c,e),(d,a),(d,b),(d,c),(d,e),(e,a),(e,b),(e,c),(e,d)$.

**5.**平面内有 $A,B,C,D$ 四个不同的点，其中任意三个点不共线.

(1)试写出以其中任意两个点为端点的有向线段；

(2)试写出以其中任意两个点为端点的线段；

(3)试写出以其中任意三个点为顶点的三角形.

**解**：(1)以其中任意两个点为端点的有向线段为一个排列，所有有向线段为 $AB,AC,AD,BC,BD,CD,BA,CA,DA,CB,DB,DC$.

(2)以其中任意两个点为端点的线段为一个组合问题，所有线段为 $AB,AC,AD,BC,BD,CD$.

(3)以其中任意三个点为顶点的三角形是一个组合问题，所有三角形为 $\triangle ABC$，$\triangle ABD$，$\triangle BCD$，$\triangle ACD$.

## B组 应用·实践

**1.**给出下列问题：

①从甲、乙、丙 3 名同学中选出 2 名分别去两个乡镇进行社会调查，有多少种不同的选法？

②将 4 张同样的电影票分给 7 人中的 4 人，有多少种不同的分法？

③某人射击 8 枪，击中 4 枪，且击中的 4 枪均为 2 枪连中，则不同的结果有多少种？

其中组合问题的个数为 （　　）

A.0　　B.1

C.2　　D.3

C　**解析**：①与顺序有关，是排列问题；②③均与顺序无关，是组合问题.

**2.**（多选）下列问题是组合问题的是 （　　）

A.把 5 本不同的书分给 5 名学生，每人一本，求有多少种不同的分法

B.从 7 本不同的书中取出 5 本给某名同学，求有多少种不同的取法

C.10 个人相互发一条短信，求一共发了几条短信

D.10 个人互相通一次电话，求一共通了几次电话

BD　**解析**：对于 A，学生与书都不相同，故与顺序有关，是排列问题；

对于 B，取出 5 本书后，即确定了取法，与顺序无关，故是组合问题；

对于 C，因为是相互发一条短信，与顺序有关，故是排列问题；

对于 D，因为是互相通一次电话，与顺序无关，故是组合问题.

故选 BD.

**3.**平面上有 12 个点，其中没有 3 个点在同一条直线上，也没有 4 个点在同一个圆上，则由这 12 个点所确定的圆的个数相当于从 12 个不同的元素中任取________个元素的组合的个数.

3　**解析**：因为每 3 个点可确定一个圆，所以由这 12 个点所确定的圆的个数相当于从 12 个不同元素中任取 3 个元素的组合的个数.

**4.**5 个代表分 4 张同样的参观券，每人最多分一张，且全部分完，那么每一种分法相当于从 5 个不同元素中任取________个元素的一个组合.

4　**解析**：因为是 4 张同样的参观券，所以参观券没有顺序，即该问题是一个组合问题，即从 5 个不同元素中取出 4 个元素.

**5.**（2022·全国甲卷）从正方体的 8 个顶点中任选 4 个，则这 4 个点在同一个平面的概率为________.

$\frac{6}{35}$　**解析**：从正方体的 8 个顶点中任选 4 个，有 $n=70$ 个结果，这 4 个点在同一个平面的有 $m=6+6=12$ 个.故所求概率 $P=\frac{m}{n}=\frac{12}{70}=\frac{6}{35}$.

**6.**某人决定投资 8 种股票和 4 种债券，经纪人向他推荐了 12 种股票和 7 种债券.问：此人的投资方案可以怎样得到？

**解**：需分两步：

第 1 步，从经纪人推荐的 12 种股票中选 8 种，是一个组合问题；

第 2 步，从经纪人推荐的 7 种债券中选 4 种，也是一个组合问题.

最后将选中的 8 种股票与选中的 4 种债券合在一起就是一种投资方案.

# 6.2.4　组合数

**学习任务目标**

1.理解组合数的概念,能利用计数原理推导组合数公式.
2.能够运用组合数公式解决一些简单的实际问题.
3.能在实际问题中区分排列与组合,准确选择恰当的方法解决简单的排列与组合综合问题.

## ○ 问题式预习 ○

### 【知识清单】

**知识点一　组合数**

从 $n$ 个不同元素中取出 $m(m\leqslant n)$ 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$\underline{C_n^m}$ 表示.

**知识点二　组合数公式**

(1)$C_n^m=\underline{\dfrac{A_n^m}{A_m^m}}=\dfrac{n(n-1)(n-2)\cdots(n-m+1)}{m!}$,其中 $n,m\in\mathbf{N}^*$,并且 $m\leqslant n$.

(2)$C_n^m=\dfrac{n!}{m!(n-m)!}$.

规定 $C_n^0=\underline{1}$.

**知识点三　组合数的性质**

性质 1:$C_n^m=\underline{C_n^{n-m}}$.

性质 2:$C_{n+1}^m=C_n^m+C_n^{m-1}$.

### 【概念辨析】

1.判断正误(正确的打"√",错误的打"×").

(1)从 $a_1,a_2,a_3$三个不同元素中任取两个元素的所有不同组合的个数为 $C_3^2$.　( √ )

(2)现有 4 枚不同的纪念币,送给 10 人中的 4 人留念,有 $C_{10}^4$种送法.　( × )

(3)$C_6^3=6\times5\times4=120$.　( × )

2.若 $C_n^2=28$,则 $n=$　(　　)

A.9　B.8　C.7　D.6

B　**解析**:$C_n^2=\dfrac{n(n-1)}{2}=28$,解得 $n=8$.

3.$C_3^3+C_4^3+C_5^3=$　(　　)

A.$C_5^4$　B.$C_6^5$　C.$C_6^3$　D.$C_6^4$

D　**解析**:$C_3^3+C_4^3+C_5^3=C_4^4+C_4^3+C_5^3=C_5^4+C_5^3=C_6^4$.

4.请思考并回答下列问题:

(1)某班级将在月底参加一场篮球比赛,包括体育委员在内,班上篮球运动员共有 8 人,按照篮球比赛规则,比赛时一个球队的上场队员是 5 人.该班级有多少种队员上场方案?又有多少种队员不上场方案?这两种方案有什么关系?

**提示**:队员上场的方案有 $C_8^5$ 种;队员不上场的方案有 $C_8^3$ 种;$C_8^5=C_8^3=56$.

(2)从问题(1)中的这 8 名篮球运动员中选择 5 人的时候,可以按照体育委员是否入选进行分类:当体育委员入选时,有 $C_7^4$ 种选法;当体育委员未入选时,有 $C_7^5$ 种选法.这与直接选 5 人的选法数一样吗?你能得出什么结论?

**提示**:一样,$C_8^5=C_7^4+C_7^5$.

## ○ 任务型课堂 ○

### 任务1 > 组合数公式与组合数的性质

1.(1)$C_4^0+C_4^1+C_5^2+C_6^3+\cdots+C_{2\,025}^{2\,022}$等于　(　　)

A.$C_{2\,023}^2$　B.$C_{2\,024}^3$　C.$C_{2\,025}^3$　D.$C_{2\,026}^4$

D　**解析**:原式$=C_4^0+C_4^1+C_5^2+C_6^3+\cdots+C_{2\,025}^{2\,022}=C_5^1+C_5^2+C_6^3+\cdots+C_{2\,025}^{2\,022}=C_6^2+C_6^3+\cdots+C_{2\,025}^{2\,022}=\cdots=C_{2\,025}^{2\,021}+C_{2\,025}^{2\,022}=C_{2\,026}^{2\,022}=C_{2\,026}^4$.

(2)若 $C_{18}^{3n+6}=C_{18}^{4n-2}$,则 $C_8^n=$________.

28　**解析**:由 $C_{18}^{3n+6}=C_{18}^{4n-2}$,

得 $3n+6+4n-2=18$ 或 $3n+6=4n-2$,

解得 $n=2$ 或 $n=8$(舍去),

故 $C_8^2=28$.

2.求下列各式的值.

(1)$C_5^2+C_5^4$;(2)$C_{10}^2\times C_{10}^0-C_{10}^{10}$;

(3)$C_6^2-C_6^5$;(4)$C_6^3\div C_8^4$.

**解**:(1)$C_5^2+C_5^4=\dfrac{5\times4}{2\times1}+\dfrac{5\times4\times3\times2}{4\times3\times2\times1}=10+5=15$.

(2)$C_{10}^2\times C_{10}^0-C_{10}^{10}=\dfrac{10\times9}{2\times1}\times1-1=45-1=44$.

(3)$C_6^2-C_6^5=\dfrac{6\times5}{2\times1}-\dfrac{6\times5\times4\times3\times2}{5\times4\times3\times2\times1}=15-6=9$.

(4)$C_6^3\div C_8^4=\dfrac{6\times5\times4}{3\times2\times1}\div\dfrac{8\times7\times6\times5}{4\times3\times2\times1}=20\div70=\dfrac{2}{7}$.

3.证明:$mC_n^m=nC_{n-1}^{m-1}$.

**证明**:$mC_n^m=m\cdot\dfrac{n!}{m!(n-m)!}$

$=\dfrac{n\cdot(n-1)!}{(m-1)!(n-m)!}$

$=nC_{n-1}^{m-1}$.

**【探究总结】**

1.组合数公式 $C_n^m=\dfrac{n(n-1)(n-2)\cdots(n-m+1)}{m!}$ 体现了组合数与相应排列数的关系,一般在计算具体的组合数时会用到.

2.组合数公式 $C_n^m=\dfrac{n!}{m!(n-m)!}$ 的主要作用:一是计算 $m,n$ 较大时的组合数;二是对含有字母的组合数的式子进行变形和证明.另外,当 $m>\dfrac{n}{2}$ 时,计算 $C_n^m$ 可用性质 $C_n^m=C_n^{n-m}$ 转化,减少运算量.

## 任务2 有限制条件的组合问题

### 探究活动

**例1** 课外活动小组共13人,其中男生8人、女生5人,并且男、女生各有一名队长.现从中选5人主持某项活动,至少有一名队长当选,有多少种不同的选法?

**解**:至少有一名队长有两种情况:有一名队长和有两名队长,故共有 $C_2^1C_{11}^4+C_2^2C_{11}^3=825$ 种选法.

[一题多思]

**思考1**.将"至少有一名队长当选"改为"至多有两名女生当选",有多少种不同的选法?

**解**:至多有两名女生有三种情况:有两名女生、只有一名女生、没有女生,故共有 $C_5^2C_8^3+C_5^1C_8^4+C_8^5=966$ 种不同的选法.

**思考2**.既要有队长,又要有女生当选,有多少种不同的选法?

**解**:分两种情况:

第一种:女队长当选,有 $C_{12}^4$ 种不同的选法;

第二种:女队长不当选,有 $(C_4^1C_7^3+C_4^2C_7^2+C_4^3C_7^1+C_4^4)$ 种不同的选法.

故共有 $C_{12}^4+C_4^1C_7^3+C_4^2C_7^2+C_4^3C_7^1+C_4^4=790$ 种不同的选法.

**【探究总结】**

**"含有"或"不含有"与"至少"或"至多"组合问题的处理策略**

1."含有"或"不含有"某些元素的组合问题:"含",则先将这些元素取出,再由另外元素补足;"不含",则先将这些元素剔除,再从剩余元素中选取.

2."至少"或"至多"含有几个元素的组合问题的常用解题方法有两种:

(1)直接分类法,注意分类要细、要全;

(2)间接法,注意找准对立面,确保不重不漏.

### 应用迁移

1.某校组织校庆活动,负责人将任务分解为编号为 $A,B,C,D$ 的四个子任务,并将任务分配给甲、乙、丙3人,且每人至少分得一个子任务,则甲没有分到编号为 $A$ 的子任务的分配方法共有 ( )

A.12种 B.18种 C.24种 D.36种

C **解析**:不考虑限制条件则共有 $C_4^2A_3^3=36$ 种分配方法.

若甲分到 $A$ 编号子任务,有两种情况:

甲分到一个子任务(即只有 $A$ 编号子任务),此时共有 $C_3^2A_2^2=6$ 种分配方法;

甲分到两个子任务(即包含 $A$ 编号子任务),此时共有 $A_3^3=6$ 种分配方法.

故所求的分配方法共有 $36-6-6=24$ 种.

故选C.

2.在一次数学竞赛中,某学校有12人通过了初试,学校要从中选出5人去参加市级培训,在下列条件下,分别有多少种不同的选法?

(1)任意选5人;

(2)甲、乙、丙三人必须都参加;

(3)甲、乙、丙三人都不能参加;

(4)甲、乙、丙三人中只能有1人参加;

(5)甲、乙、丙三人中至少有1人参加;

(6)甲、乙、丙三人中至多有2人参加.

**解**:(1)任意选5人,有 $C_{12}^5=792$ 种不同的选法.

(2)甲、乙、丙三人必须都参加,只需从另外的9人中选2人,共有 $C_9^2=36$ 种不同的选法.

(3)甲、乙、丙三人都不能参加,只需从另外的9人中选5人,共有 $C_9^5=126$ 种不同的选法.

(4)甲、乙、丙三人中只能有1人参加,分两步,先从甲、乙、丙中选1人,有 $C_3^1$ 种选法,再从另外的9人中选4人,有 $C_9^4$ 种选法,共有 $C_3^1C_9^4=378$ 种不同的选法.

(5)(方法一:直接法)

甲、乙、丙三人中至少有1人参加,可分为三类:

第一类,甲、乙、丙中有1人参加,共有 $C_3^1C_9^4$ 种选法;

第二类,甲、乙、丙中有2人参加,共有 $C_3^2C_9^3$ 种选法;

第三类,甲、乙、丙中有3人参加,共有 $C_3^3C_9^2$ 种选法.

共有 $C_3^1C_9^4+C_3^2C_9^3+C_3^3C_9^2=666$ 种不同的选法.

(方法二:间接法)

12人中任意选5人,共有 $C_{12}^5$ 种选法,甲、乙、丙三人都不能参加的有 $C_9^5$ 种选法,所以共有 $C_{12}^5-C_9^5=666$ 种不同的选法.

(6)(方法一:直接法)

甲、乙、丙三人中至多有2人参加，可分为三类：

第一类，甲、乙、丙都不参加，共有$C_9^5$种选法；

第二类，甲、乙、丙中有1人参加，共有$C_3^1C_9^4$种选法；

第三类，甲、乙、丙中有2人参加，共有$C_3^2C_9^3$种选法.

共有$C_9^5+C_3^1C_9^4+C_3^2C_9^3=756$种不同的选法.

(方法二：间接法)

12人中任意选5人，共有$C_{12}^5$种选法，甲、乙、丙三人都参加的有$C_9^2$种，所以共有$C_{12}^5-C_9^2=756$种不同的选法.

## 任务3 分组分配问题

### 探究活动

**例2**　将6个志愿者的名额分给3个班，每班至少一个名额，则有________种不同的分配方法.(用数字回答)

10　**解析**：将6个志愿者的名额分配给3个班，每班至少一个名额，采用隔板法，6个名额中间有5个空，在5个空中插入2个隔板，共有$C_5^2=10$种不同的分配方法.

**例3**　现在有6本不同的书，试根据条件完成下列分组分配问题.

(1)将这6本不同的书分给甲、乙、丙三人，每人2本，有多少种分法？

(2)将这6本不同的书分为三份，每份2本，有多少种分法？

(3)将这6本不同的书分给甲、乙、丙三人，一人1本，一人2本，一人3本，有多少种分法？

**解**：(1)先从6本书中选2本给甲，有$C_6^2$种选法；再从其余的4本书中选2本给乙，有$C_4^2$种选法；最后从余下的2本书中选2本给丙，有$C_2^2$种选法.所以分给甲、乙、丙三人，每人2本，共有$C_6^2C_4^2C_2^2=90$种分法.

(2)分给甲、乙、丙三人，每人2本，有$C_6^2C_4^2C_2^2$种分法，这个过程可以分两步完成：第一步，分为三份，每份2本，设有$x$种分法；第二步，再将这三份分给甲、乙、丙三人，有$A_3^3$种分法.根据分步乘法计数原理，可得$C_6^2C_4^2C_2^2=xA_3^3$，所以$x=\frac{C_6^2C_4^2C_2^2}{A_3^3}=15$.因此分为三份，每份2本，一共有15种分法.

(3)可以分两步完成：第一步，分为三份，分别有1本、2本、3本，有$C_6^1C_5^2C_3^3$种分法；第二步，再将这三份分给甲、乙、丙三人，有$A_3^3$种分法.根据分步乘法计数原理，一共有$C_6^1C_5^2C_3^3A_3^3=360$种分法.

**【探究总结】**

1.组合问题中常见的分组问题：

(1)完全均匀分组，每组的元素个数均相等；

(2)部分均匀分组，应注意不要重复，若有$n$组均分，最后必须除以$n!$；

(3)完全非均匀分组，这种分组不考虑重复现象.

2.不同元素分配问题属于排列问题，可以按要求逐个分配，也可以分组后再分配.

3.相同元素分配问题可采用"隔板法"：

将$n$个相同元素分给$m$ $(m\leqslant n)$个不同的对象，有$C_{n-1}^{m-1}$种分法，可描述为$(n-1)$个空中插入$(m-1)$块隔板.

### 应用迁移

1.某校为了丰富学生们的课外生活，分别成立了绘画、象棋和篮球三个兴趣小组.现有甲、乙、丙、丁、戊五名同学报名参加，每人仅参加一个兴趣小组，每个兴趣小组至少有一人报名，则不同的报名方法有(　　)

A.72种　　B.100种　　C.240种　　D.150种

D　**解析**：三个小组的人数可能是3，1，1或2，2，1.若是3，1，1的情况，则报名方法共有$\frac{C_5^3C_2^1}{A_2^2}A_3^3=60$种；若是2，2，1的情况，则报名方法共有$\frac{C_5^2C_3^2}{A_2^2}A_3^3=90$种.所以共有不同的报名方法$60+90=150$种.故选D.

2.现有13个数学竞赛参赛名额分给五个班，其中一班和二班每班至少3个名额，三班和四班每班至少2个名额，五班可以不分配名额，则名额分配方式共有(　　)

A.15种　　B.35种

C.70种　　D.125种

B　**解析**：根据题意，先将13个名额分配给一班、二班每班2个，三班、四班每班1个，

由于五班可以不分配名额，则将剩下的7个名额加上1个空名额，再分成5组，每组至少1个名额.

利用"隔板法"，有$C_7^4=35$种分配方式.故选B.

3.将6个相同的小球放入4个编号为1，2，3，4的盒子，求下列放法的种数.

(1)每个盒子都不空；

(2)恰有一个空盒子；

(3)恰有两个空盒子.

**解**：(1)先把6个相同的小球排成一行，在首尾两球外侧放置一块隔板，然后在小球之间的5个空隙中任选3个空隙各插一块隔板，即每个盒子都不空有$C_5^3=10$种放法.

(2)恰有一个空盒子，插板分两步进行.

第1步，在首尾两球外侧放置一块隔板，并在小球之间的5个空隙中任选2个空隙各插一块隔板，如|0|000|00|(用"0"表示小球)，有$C_5^2$种插法；

第2步，将剩下的一块隔板与前面任意一块并放形成空盒，如|0|000||00|，有 $C_4^1$ 种插法.

故共有 $C_5^2C_4^1=40$ 种放法.

(3)恰有两个空盒子，插板分两步进行.

第1步，在首尾两球外侧放置一块隔板，并在小球之间的5个空隙中任选1个空隙插一块隔板，有 $C_5^1$ 种插法，如|00|0000|；

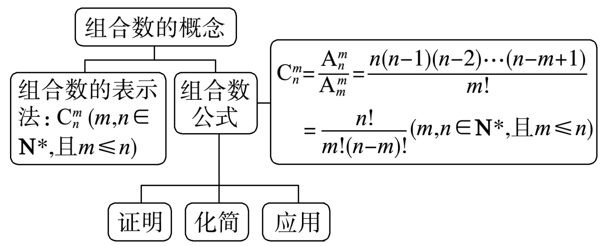
第2步，将剩下的两块隔板插入形成空盒，此时含有两种情况：

①这两块板与前面三块板形成不相邻的两个空盒，如||00||0000|，有 $C_3^2$ 种插法；

②将两块板与前面三块板之一并放，形成相邻的两个空盒，如|00|||0000|，有 $C_3^1$ 种插法.

故共有 $C_5^1(C_3^2+C_3^1)=30$ 种放法.

### 提质归纳




## 课后素养评价（六） 组合数

### A组 学习·理解

**1.** 若 $A_n^3=12C_n^2$，则 $n$ 等于（　　）

A.8　　B.5或6

C.3或4　　D.4

A　**解析**：因为 $A_n^3=n(n-1)(n-2)$，$C_n^2=\frac{1}{2}n(n-1)$，

所以 $n(n-1)(n-2)=12\times\frac{1}{2}n(n-1)$.

又 $n\in\mathbf{N}^*$，且 $n\geq3$，所以 $n=8$.

**2.** 若从1，2，3，…，9这9个整数中取4个不同的数，使其和为奇数，则不同的取法共有（　　）

A.60种　　B.63种

C.65种　　D.66种

A　**解析**：若4个数之和为奇数，则有1个奇数、3个偶数或者3个奇数、1个偶数.若是1个奇数、3个偶数，则有 $C_5^1C_4^3=20$ 种取法；若是3个奇数、1个偶数，则有 $C_5^3C_4^1=40$ 种取法.所以共有 $20+40=60$ 种不同的取法.

**3.** 某龙舟队有9名队员，其中3人只会划左舷，4人只会划右舷，2人既会划左舷又会划右舷.现要选派划左舷的3人、划右舷的3人共6人去参加比赛，则不同的选派方法共有（　　）

A.56种　B.68种　C.74种　D.92种

D　**解析**：根据会划左舷的人中"多面手"（既会划左舷又会划右舷）的人数进行分类：会划左舷的人中没有"多面手"的选派方法有 $C_3^3C_6^3$ 种；有一个"多面手"的选派方法有 $C_2^1C_3^2C_5^3$ 种；有两个"多面手"的选派方法有 $C_3^1C_4^3$ 种.故共有 $C_3^3C_6^3+C_2^1C_3^2C_5^3+C_3^1C_4^3=92$ 种不同的选派方法.

**4.** 从4名男生、2名女生中选3人组队参加"弘扬传统文化，增强文化自信"答题比赛，且至少有1名女生入选，则不同的选法种数为（　　）

A.20　B.16　C.12　D.8

B　**解析**：由题意知不同的选法可分两种情况：第一种情况，只有1名女生入选，不同的选法有 $C_2^1C_4^2=12$ 种；第二种情况，有2名女生入选，不同的选法有 $C_2^2C_4^1=4$ 种.根据分类加法计数原理知，至少有1名女生入选的不同的选法有 $12+4=16$ 种.故选B.

**5.** 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"任何一个大于2的偶数都可以写成两个素数的和"，如 $40=3+37$.在不超过40的素数中，随机选取2个不同的数，其和等于40的概率是(注：若一个大于1的整数除了1和它本身外无其他因数，则称这个整数为素数)（　　）

A.$\frac{1}{15}$　B.$\frac{1}{17}$　C.$\frac{1}{22}$　D.$\frac{1}{26}$

C　**解析**：因为不超过40的素数有2，3，5，7，11，13，17，19，23，29，31，37，共12个，

所以从中随机选取2个不同的数，共有 $C_{12}^2=\frac{12\times11}{2}=66$ 个样本点.

其中两个素数的和为40的情况包含3和37，11和29，17和23，共3个样本点，

所以所求概率 $P=\frac{3}{66}=\frac{1}{22}$.故选C.

**6.** (1)求 $3C_8^3+\frac{1}{3}A_8^3$ 的值；

(2)求 $C_3^3+C_4^3+\cdots+C_{10}^3$ 的值；

(3)解关于 $n$ 的不等式：$C_{10}^{n-3}<C_{10}^{n-2}$.

**解**：(1)$3C_8^3+\frac{1}{3}A_8^3=3\times\frac{8\times7\times6}{3\times2\times1}+\frac{1}{3}\times8\times7\times6=280$.

(2)$C_3^3+C_4^3+\cdots+C_{10}^3=C_4^4+C_4^3+\cdots+C_{10}^3$

$=C_5^4+C_5^3+\cdots+C_{10}^3=C_{11}^4=\frac{11\times10\times9\times8}{4\times3\times2\times1}=330$.

(3)由题意可得 $\begin{cases}0\leqslant n-3\leqslant10,\\0\leqslant n-2\leqslant10,\end{cases}$ 解得 $3\leqslant n\leqslant12$，且 $n\in\mathbf{N}^*$.

因为 $C_{10}^{n-3}<C_{10}^{n-2}$，

所以 $\frac{10!}{(n-3)!\,(13-n)!}<\frac{10!}{(n-2)!\,(12-n)!}$，

解得 $n<7.5$.

又因为 $n\in\mathbf{N}^*$，

所以 $n=3,4,5,6,7$.故不等式的解集为 $\{3,4,5,6,7\}$.

**7.** 将 4 个编号为 1，2，3，4 的小球放入 4 个编号为 1，2，3，4 的盒子中.

(1)有多少种放法？

(2)每盒至多一球，有多少种放法？

(3)把 4 个不同的小球换成 4 个相同的小球，恰有一个空盒，有多少种放法？

**解**：(1)每个小球都可能放入 4 个盒子中的任何一个，将小球一个一个放入盒子，共有 $4\times4\times4\times4=4^4=256$ 种放法.

(2)这是全排列问题，共有 $A_4^4=24$ 种放法.

(3)(方法一)先从四个盒子中选出三个盒子，再从三个盒子中选出一个盒子放入两个球，余下两个盒子各放一个.由于球是相同的即没有顺序，

故共有 $C_4^3C_3^1=12$ 种放法.

(方法二)恰有一个空盒子，第一步先选出一个盒子，有 $C_4^1$ 种选法，

第二步在 4 个小球之间的 3 个空隙中任选 2 个空隙各插一块隔板，有 $C_3^2$ 种方法.

由分步乘法计数原理得，共有 $C_4^1C_3^2=12$ 种放法.

## B组　应用·实践

**1.** 某学校安排小明和小李等 5 名志愿者将两个吉祥物安装在学校的体育广场，每人参与且只参与一个吉祥物的安装，每个吉祥物都至少由两名志愿者安装.若小明和小李必须安装不同的吉祥物，则不同的安排方案有　　(　　)

A.6 种　　B.12 种

C.18 种　　D.24 种

B　**解析**：由题意可知，应将志愿者分为三人组和两人组.先将小李、小明之外的三人分为两组，有 $C_3^1C_2^2=3$ 种分法；再将小李、小明分进两组，有 $A_2^2=2$ 种分法；最后安排两个组安装两个吉祥物，有 $A_2^2=2$ 种方案.所以共有不同的安排方案 $3\times2\times2=12$ 种.故选 B.

**2.** (2023·新高考全国Ⅱ卷)某学校为了解学生参加体育运动的情况，用比例分配的分层随机抽样方法做抽样调查，拟从初中部和高中部两层共抽取 60 名学生.已知该校初中部和高中部分别有 400 名和 200 名学生，则不同的抽样结果共有　　(　　)

A.$C_{400}^{45}\cdot C_{200}^{15}$ 种　　B.$C_{400}^{20}\cdot C_{200}^{40}$ 种

C.$C_{400}^{30}\cdot C_{200}^{30}$ 种　　D.$C_{400}^{40}\cdot C_{200}^{20}$ 种

D　**解析**：根据分层随机抽样的定义知，初中部抽取的人数为 $60\times\frac{400}{600}=40$，高中部抽取的人数为 $60\times\frac{200}{600}=20$.

根据组合公式和分步乘法计数原理，得不同的抽样结果共有 $C_{400}^{40}\cdot C_{200}^{20}$ 种.故选 D.

**3.** 6 名同学到甲、乙、丙三个场馆做志愿者，每名同学只去 1 个场馆，甲场馆安排 1 名同学，乙场馆安排 2 名同学，丙场馆安排 3 名同学，则不同的安排方法共有　　(　　)

A.120 种　　B.90 种

C.60 种　　D.30 种

C　**解析**：甲场馆安排 1 名同学有 $C_6^1$ 种方法，乙场馆安排 2 名同学有 $C_5^2$ 种方法，丙场馆安排 3 名同学有 $C_3^3$ 种方法.由分步乘法计数原理，得不同的安排方法共有 $C_6^1C_5^2C_3^3=60$ 种.

**4.** 某校将 12 个优秀团员名额分配给 4 个不同的班级，要求每个班级至少一个，则不同的分配方案有________种.

165　**解析**：将 12 个优秀团员名额分配给 4 个不同的班级，要求每个班级至少一个，应用隔板法.在 12 个名额之间的 11 个空中，用 3 个隔板分成四组(非空)，则不同的分配方案有 $C_{11}^3=\frac{11\times10\times9}{3\times2\times1}=165$ 种.

**5.** 为贯彻落实“立德树人”的根本任务，探索德、智、体、美、劳“五育并举”的实施路径，某校统筹推进以“五育并举＋教师教育”为特色的第二课堂养成体系，引导学生崇尚劳动、尊重劳动者、提高劳动素养，以劳动周的形式开展劳育工作的创新实践.若学生可以参加“民俗文化”“茶艺文化”“茶壶制作”“水果培育”“蔬菜种植”“3D 打印”这六门劳动课中的

一门，则甲、乙、丙、丁这4名学生中至少有3名所选劳动课互不相同的选课方法共有______种.

1 080 **解析**：分两种情况讨论：

若4名学生选的课互不相同，则有$C_6^4A_4^4=360$种方法；

若有3名学生选的课全不同，即只有2名学生选的课相同，则有$C_6^3C_4^2A_3^3=720$种方法.

故共有$360+720=1\ 080$种方法.

**6.**10件产品中有合格品8件，次品2件，从中抽取4件，则：

(1)都不是次品的取法共有多少种？

(2)至少有1件次品的取法共有多少种？

**解**：(1)根据题意，抽取的4件都不是次品，即4件都为合格品，

故有$C_8^4=\frac{8\times7\times6\times5}{4\times3\times2\times1}=70$种取法.

(2)根据题意，抽取的4件里面至少有1件次品，

可以先考虑对立情况，即4件都为合格品，共有$C_8^4=70$种取法，

总共有$C_{10}^4=\frac{10\times9\times8\times7}{4\times3\times2\times1}=210$种取法，

则至少有1件次品有$C_{10}^4-C_8^4=210-70=140$种取法.

**7.**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**：依据0与1两个特殊值分析，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$C_4^1$种方法；0可在后两位，有$C_2^1$种方法；最后需从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$C_4^1C_2^1C_3^1\times2^2=96$个.

②取1不取0，同上分析可得不同的三位数有$C_4^2\times2^2\times A_3^3=144$个.

③0和1都不取，有不同的三位数$C_4^3\times2^3\times A_3^3=192$个.

综上所述，不同的三位数共有$96+144+192=432$个.

# 6.3 二项式定理

## 6.3.1 二项式定理

**学习任务目标**

**1.**能运用二项式定理求二项展开式.

**2.**会用通项公式求展开式中的二项式系数、特殊项及特殊项系数.

## ◦ 问题式预习 ◦

**【知识清单】**

**知识点一　二项式定理**

| | |
|---|---|
| 二项式定理 | $(a+b)^n=C_n^0a^n+C_n^1a^{n-1}b^1+\cdots+C_n^ka^{n-k}b^k+\cdots+C_n^nb^n$，$n\in\mathbf{N}^*$ |
| 二项展开式 | 等号右边的多项式叫做$(a+b)^n$的二项展开式，展开式一共有$\underline{n+1}$项 |
| 二项式系数 | 各项的系数$C_n^k(k=0,1,2,\cdots,n)$叫做二项式系数 |
| 二项展开式的通项 | $T_{k+1}=\underline{C_n^ka^{n-k}b^k}$，通项为展开式的第$k+1$项 |

**知识点二　某一项的二项式系数与项的系数**

(1)二项式系数与项的系数是两个不同的概念.

(2)二项式系数$C_n^k$只与二项式的指数及项数有关，它是一个组合数；而项的系数是指通项中除字母之外的部分，它与二项式和二项式系数都有关.

**【概念辨析】**

**1.**判断正误(正确的打"√"，错误的打"×").

(1)$(a+b)^n$的展开式中共有$n$项.（×）

(2)在二项式定理中，交换$a$，$b$的顺序对各项没有影响.（×）

(3)$C_n^ka^{n-k}b^k$是$(a+b)^n$的展开式中的第$k$项.（×）

**2.**若$\left(x-\frac{2}{\sqrt{x}}\right)^n$的展开式共有11项，则$n$等于（　）

A.9　　B.10

C.11　　D.8

B **解析**：因为$\left(x-\frac{2}{\sqrt{x}}\right)^n$的展开式共有$(n+1)$项，所以$n+1=11$，故$n=10$.

3.$(x-\sqrt{2})^{10}$的展开式中含 $x^6$ 的项的二项式系数为（　　）

A.$-C_{10}^4$　　B.$C_{10}^4$　　C.$-4C_{10}^4$　　D.$4C_{10}^4$

B　**解析**：含 $x^6$ 的项为展开式中的第 5 项，所以二项式系数为 $C_{10}^4$.

4.$(\sqrt{2}x-3)^5$ 的展开式中第 4 项的系数是（　　）

A.10　　B.$-10$

C.540　　D.$-540$

D　**解析**：由展开式的通项，得 $T_4=C_5^3\cdot(\sqrt{2}x)^2\times(-3)^3=10\times2x^2\times(-27)=-540x^2$，所以第 4 项的系数为$-540$.

5.请思考并回答下列问题：

(1)在初中，我们用多项式乘法法则得到了$(a+b)^2$的展开式：$(a+b)^2=(a+b)(a+b)=a\times a+a\times b+b\times a+b\times b=a^2+2ab+b^2$.如何利用分步乘法计数原理解释上述展开过程呢？

**提示**：从展开过程可以看到，$(a+b)^2$是 2 个$(a+b)$相乘，根据多项式乘法法则，每个$(a+b)$在相乘时有两种选择，选 $a$ 或选 $b$，而且每个$(a+b)$中的 $a$ 或 $b$ 都选定后，才能得到展开式的一项.于是，由分步乘法计数原理，在合并同类项之前，$(a+b)^2$的展开式共有 $C_2^1\times C_2^1=2^2$ 项，而且每一项都是 $a^{2-k}\times b^k(k=0,1,2)$的形式，$a^{2-k}b^k$ 出现的次数相当于从 2 个$(a+b)$中取 $k$ 个 $b$ 的组合数 $C_2^k$.

(2)$(a+b)^n$的二项展开式有 $n+1$ 项，是和的形式，各项的幂指数具有什么规律？

**提示**：①各项的幂指数和等于 $n$.②字母 $a$ 按降幂排列，从第一项起，幂指数由 $n$ 逐项减 1 直到 0；字母 $b$ 按升幂排列，从第一项起，幂指数由 0 逐项加 1 直到 $n$.

## ○ 任务型课堂 ○

### 任务1＞二项式定理的正用与逆用

#### 探究活动

**例 1**　化简：$C_n^0(x+1)^n-C_n^1(x+1)^{n-1}+C_n^2(x+1)^{n-2}-\cdots+(-1)^kC_n^k(x+1)^{n-k}+\cdots+(-1)^nC_n^n=$________.

$x^n$　**解析**：原式$=C_n^0(x+1)^n+C_n^1(x+1)^{n-1}(-1)+C_n^2(x+1)^{n-2}(-1)^2+\cdots+C_n^k(x+1)^{n-k}(-1)^k+\cdots+C_n^n(-1)^n=[(x+1)+(-1)]^n=x^n$.

**【探究总结】**

1.正用：求形式简单的二项展开式时可直接由二项式定理展开，展开时注意二项展开式的特点：前一个字母是降幂，后一个字母是升幂.$(a-b)^n$ 的展开式中各项会出现正负相间的情况.对较繁杂的式子，先化简，再用二项式定理展开.

2.逆用：逆用二项式定理可将多项式化简，对于这类问题，要熟悉公式的特点、项数、各项幂指数的规律以及各项的系数.

**提醒**：逆用二项式定理时，若项的系数是正负相间的，则化简结果是$(a-b)^n$ 的形式.

#### 应用迁移

1.第 14 届国际数学教育大会(ICME-14)在我国上海华东师范大学举行.如图是本次大会的会标，会标中“ICME-14”的下方展示的是八卦中的三卦.这些卦可以看成八进制数 $3\ 744_{(8)}$，换算成十进制数是 $3\times8^3+7\times8^2+4\times8^1+4\times8^0=2\ 020$，正是会议计划召开的年份(实际推迟至 2021 年)，那么八进制数$\underbrace{77\cdots7}_{10个7}{}_{(8)}$换算成十进制数后，这个数的末位数字是（　　）

A.1　　B.3

C.5　　D.7

B　**解析**：由进位制的换算方法可知，八进制数$\underbrace{77\cdots7}_{10个7}{}_{(8)}$换算成十进制数得 $7\times8^9+7\times8^8+\cdots+7\times8^1+7\times8^0=7\times\dfrac{1-8^{10}}{1-8}=8^{10}-1$，

$8^{10}-1=(10-2)^{10}-1=C_{10}^0\times10^{10}+C_{10}^1\times10^9\times(-2)^1+\cdots+C_{10}^9\times10^1\times(-2)^9+C_{10}^{10}\times2^{10}-1$.

因为 $C_{10}^0\times10^{10}+C_{10}^1\times10^9\times(-2)^1+\cdots+C_{10}^9\times10^1\times(-2)^9$ 是 10 的倍数，

所以这个数的末位数字即为 $C_{10}^{10}\times2^{10}-1$ 的末位数字.

由 $C_{10}^{10}\times2^{10}-1=1\ 023$，得末位数字为 3.

2.化简：$3C_n^1+6C_n^2+12C_n^3+\cdots+3\times2^{n-1}C_n^n=$________.

$\dfrac{3}{2}(3^n-1)$　**解析**：$3C_n^1+6C_n^2+12C_n^3+\cdots+3\times2^{n-1}C_n^n=\dfrac{3}{2}(2C_n^1+2^2C_n^2+2^3C_n^3+\cdots+2^nC_n^n)=\dfrac{3}{2}[(1+2)^n-C_n^0]=\dfrac{3}{2}(3^n-1)$.

## 任务2 求展开式中的特定项

### 探究活动

**例 2** 根据二项式定理回答下列关于$\left(\sqrt[3]{x}-\dfrac{3}{\sqrt[3]{x}}\right)^n$的展开式的问题.

(1)求展开式的通项；

(2)若展开式的第 6 项为常数项，求 $n$ 的值.

**解**：(1)通项为 $T_{k+1}=C_n^k x^{\frac{n-k}{3}}(-3)^k x^{-\frac{k}{3}}=C_n^k\cdot(-3)^k x^{\frac{n-2k}{3}}$.

(2)若第 6 项为常数项，则$\dfrac{n-2\times5}{3}=0$，得 $n=10$.

［一题多思］

**思考 1.** 在(2)的条件下，求含 $x^2$ 的项的系数，并指出该项的二项式系数.

**解**：由题意知 $T_{k+1}=C_{10}^k(-3)^k x^{\frac{10-2k}{3}}$.

令$\dfrac{10-2k}{3}=2$，得 $k=2$，

所以含 $x^2$ 的项的系数为 $C_{10}^2\times(-3)^2=405$.

该项的二项式系数为 $C_{10}^2=45$.

**思考 2.** 在(2)的条件下，求此展开式中所有的有理项.

**解**：由题意得$\begin{cases}\dfrac{10-2k}{3}\in\mathbf{Z},\\0\leqslant k\leqslant10,\\k\in\mathbf{Z},\end{cases}$

令$\dfrac{10-2k}{3}=r(r\in\mathbf{Z})$，则 $10-2k=3r$，即 $k=5-\dfrac{3}{2}r$.

因为 $k\in\mathbf{Z}$，所以 $r$ 应为偶数，$r=2,0,-2$，即 $k=2,5,8$，

故二项展开式的所有有理项为

$T_3=C_{10}^2(-3)^2x^2=405x^2$，

$T_6=C_{10}^5(-3)^5=-61\ 236$，

$T_9=C_{10}^8(-3)^8x^{-2}=295\ 245x^{-2}$.

**【探究总结】**

**求二项展开式中的特定项的常见题型及解题策略**

(1)求$(a+b)^n$的展开式的第 $k$ 项.对于第 $k$ 项，利用公式 $T_k=C_n^{k-1}a^{n-k+1}b^{k-1}$ 可直接写出.

(2)求常数项.对于常数项，隐含条件是字母的指数为 0(即 0 次项)，令通项中字母的指数为 0 可得常数项.

(3)求有理项.对于有理项，其所有的字母的指数恰好都是整数，令通项中字母的指数为整数，再根据数的整除性来求解.

(4)求整式项.对于整式项，其字母的指数应是非负整数，求解方法与求有理项的方法类似.

### 应用迁移

**1.**(2024·天津卷)在$\left(\dfrac{3}{x^3}+\dfrac{x^3}{3}\right)^6$的展开式中，常数项为________.

20 **解析**：因为$\left(\dfrac{3}{x^3}+\dfrac{x^3}{3}\right)^6$的展开式的通项为

$T_{k+1}=C_6^k\left(\dfrac{3}{x^3}\right)^{6-k}\left(\dfrac{x^3}{3}\right)^k=3^{6-2k}C_6^k x^{6(k-3)}$，$k=0,1,\cdots,6$，

令 $6(k-3)=0$，可得 $k=3$，

所以常数项为 $3^0\times C_6^3=20$.

**2.** 若$\left(x+\dfrac{a}{\sqrt[3]{x}}\right)^8$的展开式中 $x^4$ 的系数为 7，则实数 $a=$________.

$\dfrac{1}{2}$ **解析**：根据二项展开式的通项可得

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

令 $8-\dfrac{4k}{3}=4$，可得 $k=3$，

此时 $C_8^k a^k=C_8^3a^3=7$，解得 $a=\dfrac{1}{2}$.

**3.** 已知 $n\in\mathbf{Z}$，且 $3\leqslant n\leqslant6$，若$\left(x-\dfrac{2}{x^3}\right)^n$的展开式中存在常数项，则该常数项为________.

$-8$ **解析**：由题意可知，$\left(x-\dfrac{2}{x^3}\right)^n$的展开式的通项为 $T_{k+1}=C_n^k\cdot x^{n-k}\cdot\left(-\dfrac{2}{x^3}\right)^k=(-2)^kC_n^k\cdot x^{n-4k}$.

令 $n-4k=0$，解得 $k=\dfrac{n}{4}$.

又因为 $n\in\mathbf{Z}$，且 $3\leqslant n\leqslant6$，所以 $n=4$，$k=\dfrac{n}{4}=1$，所以$\left(x-\dfrac{2}{x^3}\right)^n$的展开式中的常数项为 $T_2=(-2)^1C_4^1\cdot x^0=(-2)\times4=-8$.

**4.** 在$\left(2x^2-\dfrac{1}{\sqrt[3]{x}}\right)^8$的展开式中，求：

(1)第 5 项的二项式系数及第 5 项的系数；

(2)含 $x^2$ 项的系数.

**解**：(1)因为$\left(2x^2-\dfrac{1}{\sqrt[3]{x}}\right)^8$的展开式的通项是 $T_{k+1}=C_8^k(2x^2)^{8-k}\left(-\dfrac{1}{\sqrt[3]{x}}\right)^k=(-1)^kC_8^k\cdot2^{8-k}\cdot x^{16-\frac{7}{3}k}$，

$T_5=(-1)^4\times C_8^4\times 2^4\cdot x^{\frac{20}{3}}$，

所以第 5 项的二项式系数是 $C_8^4=70$，

第 5 项的系数是 $C_8^4\times 2^4=1\ 120$.

(2)由题意，令 $16-\dfrac{7}{3}k=2$，解得 $k=6$.

因此，含 $x^2$ 项的系数是 $(-1)^6\times C_8^6\times 2^{8-6}=112$.

## 任务3 二项式定理的灵活运用

### 探究活动

**例 3**　(1)已知 $4^{2\ 024}+a$ 能被 9 整除，则整数 $a$ 的值可以是（　　）

A. $-12$　　B. $-7$

C. 9　　D. 13

B　**解析**：因为 $4^{2\ 024}+a=2^{4\ 048}+a=2\times 8^{1\ 349}+a=2\times(9-1)^{1\ 349}+a=2\sum\limits_{k=0}^{1\ 348}[C_{1\ 349}^k\times 9^{1\ 349-k}\times(-1)^k]-2+a$，

又 $2\sum\limits_{k=0}^{1\ 348}[C_{1\ 349}^k\times 9^{1\ 349-k}\times(-1)^k]$ 能被 9 整除，

所以 $-2+a$ 能被 9 整除.

由选项知当 $a=-7$ 时符合，

故选 B.

(2)$(2x^2-3x)\left(2-\dfrac{1}{\sqrt{x}}\right)^7$ 的展开式中含 $x$ 项的系数为________.

960　**解析**：$\left(2-\dfrac{1}{\sqrt{x}}\right)^7$ 的展开式的通项为 $T_{k+1}=C_7^k 2^{7-k}(-1)^k x^{-\frac{k}{2}}$，当 $k=0,2$ 时，$T_1=C_7^0\times 2^7x^0$，$T_3=C_7^2\times 2^5x^{-1}$，

因此 $(2x^2-3x)\left(2-\dfrac{1}{\sqrt{x}}\right)^7$ 的展开式中含 $x$ 项的系数为 $(-3)\times C_7^0\times 2^7+2\times C_7^2\times 2^5=-384+1\ 344=960$.

(3)$(x-2y+1)^5$ 的展开式中含 $x^2y$ 项的系数为________.

$-60$　**解析**：$(x-2y+1)^5=[1+(x-2y)]^5$，

其展开式的通项为 $T_{r+1}=C_5^r\cdot 1^{5-r}\cdot(x-2y)^r=C_5^r\cdot(x-2y)^r$.

因为 $x^2y$ 的次数为 3，所以 $r=3$.

又 $(x-2y)^3$ 的通项为 $T'_{r'+1}=C_3^{r'}\cdot x^{3-r'}\cdot(-2y)^{r'}$.

令 $r'=1$，得含 $x^2y$ 项的系数为 $C_5^3\times C_3^1\times(-2)=-60$.

**【探究总结】**

1. $(a+b+c)^n$ 的展开式中特定项的求解方法：

| 方法技巧 | |
|---|---|
| 因式分解法 | 将三项式利用因式分解转化为两个二项式的积，然后再用二项式定理求解 |
| 逐层展开法 | 将三项式分成两组，用二项式定理展开，再把其中含二项式的部分展开，从而求解 |
| 组合知识法 | 根据 $(a+b+c)^n$ 是 $n$ 个 $(a+b+c)$ 的积，利用组合知识分析项的构成，从而求解 |

2. 求两个(或多个)二项式的正整数幂乘积的展开式的常用方法：(1)利用二项式定理分别展开，于是问题转化为求多项式与多项式乘积的展开式，此时只需利用多项式乘法法则对其展开即可(即用一个多项式的每一项分别乘另一个多项式的每一项)；(2)先利用运算性质对其进行化简，再利用二项式定理进行展开.

3. 利用二项式定理可以解决余数或整除的问题，通常需将底数化成两数的和或差的形式，且这种转化形式与除数有密切的关系.

### 应用迁移

1. $\left(x^2+2x-\dfrac{1}{x}\right)^6$ 的展开式中的常数项为（　　）

A. $-160$　　B. 15　　C. $-145$　　D. $-40$

C　**解析**：$\left(x^2+2x-\dfrac{1}{x}\right)^6$ 可写成 $\left[(x^2+2x)-\dfrac{1}{x}\right]^6$，故其展开式的通项为 $T_{k+1}=C_6^k\cdot(x^2+2x)^{6-k}\cdot\left(-\dfrac{1}{x}\right)^k$，$k=0,1,2,\cdots,6$，

即 $T_{k+1}=C_6^kC_{6-k}^m(x^2)^{6-k-m}(2x)^m\cdot\left(-\dfrac{1}{x}\right)^k=(-1)^k\cdot 2^m\cdot C_6^k\cdot C_{6-k}^m x^{12-3k-m}$，$0\leqslant k\leqslant 6$，$0\leqslant m\leqslant 6-k$，$k,m\in\mathbf{N}$.

要求展开式中的常数项，需要 $x$ 的指数为 0，即需使 $12-3k-m=0$，即 $3k+m=12$. 当 $k=4$ 时，$m=0$；当 $k=3$ 时，$m=3$.

故二项展开式中的常数项为 $C_6^4\times C_2^0-2^3\times C_6^3\times C_3^3=15-160=-145$.

故选 C.

2. $(1-x)^5(1+2x)^4$ 的展开式中含 $x^2$ 项的系数为（　　）

A. $-14$　　B. $-6$　　C. 34　　D. 74

B　**解析**：$(1-x)^5$ 的展开式的通项为 $T_{r+1}=C_5^r\cdot(-1)^r\cdot x^r(r=0,1,2,3,4,5)$，

$(1+2x)^4$ 的展开式的通项为 $T_{k+1}=C_4^k\cdot 2^k\cdot x^k(k=0,1,2,3,4)$.

当 $r=0,k=2$ 时，$x^2$ 的系数为 $C_4^2\times 2^2=24$；

当 $r=1,k=1$ 时，$x^2$ 的系数为 $-5\times 4\times 2=-40$；

当 $r=2,k=0$ 时，$x^2$ 的系数为 $C_5^2=10$.

故含 $x^2$ 项的系数为 $24+10-40=-6$.

故选 B.

3.(1)$1.05^6$ 的计算结果精确到 0.01 的近似值是______；

(2)求证：$1+2+2^2+\cdots+2^{5n-1}(n\in\mathbf{N}^*)$能被 31 整除.

(1)1.34　**解析**：$1.05^6=(1+0.05)^6=1+C_6^1\times0.05+C_6^2\times0.05^2+\cdots\approx1+0.3+0.0375=1.3375\approx1.34$.

(2)**证明**：因为 $1+2+2^2+\cdots+2^{5n-1}=\dfrac{1-2^{5n}}{1-2}$

$=2^{5n}-1$

$=32^n-1$

$=(31+1)^n-1$

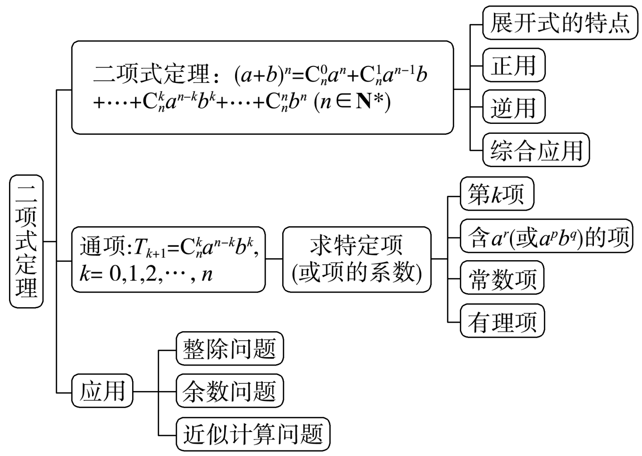
$=C_n^0\times31^n+C_n^1\times31^{n-1}+\cdots+C_n^{n-1}\times31+C_n^n\times31^0-1$

$=31(C_n^0\times31^{n-1}+C_n^1\times31^{n-2}+\cdots+C_n^{n-1})$，

显然 $C_n^0\times31^{n-1}+C_n^1\times31^{n-2}+\cdots+C_n^{n-1}$ 为整数，

所以 $1+2+2^2+\cdots+2^{5n-1}(n\in\mathbf{N}^*)$能被 31 整除.

## 提质归纳




# 课后素养评价（七）　二项式定理

## A组　学习·理解

1.(2024·北京卷)在$(x-\sqrt{x})^4$ 的展开式中，$x^3$ 的系数为　(　　)

A.6　　B.$-6$　　C.12　　D.$-12$

A　**解析**：(方法一：公式法)$(x-\sqrt{x})^4$ 的展开式的通项为 $T_{k+1}=C_4^kx^{4-k}(-\sqrt{x})^k=(-1)^kC_4^kx^{4-\frac{k}{2}}(k=0,1,2,3,4)$.令 $4-\dfrac{k}{2}=3$，得 $k=2$，所以$(x-\sqrt{x})^4$ 的展开式中 $x^3$ 的系数为$(-1)^2\times C_4^2=6$.

(方法二：组合数法)$(x-\sqrt{x})^4$ 的展开式中含 $x^3$ 的项是由$(x-\sqrt{x})(x-\sqrt{x})(x-\sqrt{x})(x-\sqrt{x})$中任意取 2 个括号内的 $x$ 与剩余的 2 个括号内的$(-\sqrt{x})$相乘得到的，所以$(x-\sqrt{x})^4$ 的展开式中含 $x^3$ 的项为 $C_4^2x^2\cdot C_2^2(-\sqrt{x})^2=6x^3$，所以$(x-\sqrt{x})^4$ 的展开式中 $x^3$ 的系数为 6.

2.$\left(x^3-\dfrac{2}{x^2}\right)^5$ 的展开式中的常数项为　(　　)

A.80　　B.$-80$

C.40　　D.$-40$

B　**解析**：$\left(x^3-\dfrac{2}{x^2}\right)^5$ 的展开式的通项为 $T_{k+1}=C_5^k\cdot(x^3)^{5-k}\left(-\dfrac{2}{x^2}\right)^k=(-2)^kC_5^kx^{15-5k}$.令 $15-5k=0$，得 $k=3$，所以常数项为 $T_4=(-2)^3\times C_5^3=-80$.

3.化简多项式$(2x+1)^5-5(2x+1)^4+10(2x+1)^3-10(2x+1)^2+5(2x+1)-1$ 的结果是　(　　)

A.$(2x+2)^5$　　B.$2x^5$

C.$(2x-1)^5$　　D.$32x^5$

D　**解析**：依题意，可知多项式为$[(2x+1)-1]^5$ 的展开式，则$[(2x+1)-1]^5=(2x)^5=32x^5$.故选 D.

4.用二项式定理展开：$(2x-1)^4=$____________________.

$16x^4-32x^3+24x^2-8x+1$　**解析**：$(2x-1)^4=C_4^0(2x)^4(-1)^0+C_4^1(2x)^3(-1)^1+C_4^2(2x)^2(-1)^2+C_4^3(2x)^1(-1)^3+C_4^4(2x)^0(-1)^4=16x^4-32x^3+24x^2-8x+1$.

5.已知$(1-x)^9+m(x+1)^{10}=a_0+a_1x+a_2x^2+\cdots+a_9x^9+a_{10}x^{10}$.若 $a_9=a_{10}$，则 $a_2=$________.

41　**解析**：由题意可得 $a_9=C_9^9(-1)^9+mC_{10}^1=10m-1$，$a_{10}=m$.

因为 $a_9=a_{10}$，所以 $10m-1=m$，解得 $m=\dfrac{1}{9}$，所以 $a_2=C_9^2+\dfrac{1}{9}C_{10}^8=41$.

6.已知 $n$ 为等差数列$-4,-2,0,\cdots$的第 6 项，则$\left(x+\dfrac{2}{x}\right)^n$ 的二项展开式的常数项是________.

160　**解析**：由题意得 $n=6$，所以 $T_{k+1}=2^kC_6^kx^{6-2k}$.令 $6-2k=0$，得 $k=3$.所以常数项为 $C_6^3\times2^3=160$.

7.求$\left(x^3+\frac{2}{3x^2}\right)^5$的展开式的第3项的系数和常数项.

**解**:由题意得通项为 $T_{k+1}=C_5^k(x^3)^{5-k}\left(\frac{2}{3x^2}\right)^k=\left(\frac{2}{3}\right)^k\cdot C_5^kx^{15-5k}$,则 $T_3=\left(\frac{2}{3}\right)^2\times C_5^2x^5=\frac{4}{9}\times C_5^2x^5$,

所以第3项的系数为$\frac{4}{9}\times C_5^2=\frac{40}{9}$.

令 $15-5k=0$,得 $k=3$,

所以常数项为 $T_4=\left(\frac{2}{3}\right)^3\times C_5^3=\frac{80}{27}$.

## B组　应用·实践

1.已知$(1+x)^5=a_0+a_1(1-x)+a_2(1-x)^2+\cdots+a_5(1-x)^5$,则 $a_3=$　　(　　)

A.$-40$　　B.40

C.10　　D.$-10$

A　**解析**:因为$(1+x)^5=-[-2+(1-x)]^5$,通项为 $T_{k+1}=-C_5^k(-2)^{5-k}(1-x)^k$,所以 $a_3=-C_5^3(-2)^2=-40$.

2.$\left(x+\frac{y^2}{x}\right)(x+y)^5$ 的展开式中 $x^3y^3$ 的系数为　　(　　)

A.5　　B.10　　C.15　　D.20

C　**解析**:$(x+y)^5$ 的展开式的通项为 $T_{k+1}=C_5^kx^{5-k}y^k(k\in\mathbf{N}$ 且 $k\leqslant5)$,

所以$\left(x+\frac{y^2}{x}\right)(x+y)^5$ 的展开式的通项可表示为 $xT_{k+1}=xC_5^kx^{5-k}y^k=C_5^kx^{6-k}y^k$ 与 $\frac{y^2}{x}T_{k+1}=\frac{y^2}{x}C_5^kx^{5-k}y^k=C_5^kx^{4-k}y^{k+2}$的和.

在 $xT_{k+1}=C_5^kx^{6-k}y^k$ 中,令 $k=3$,

可得 $xT_4=C_5^3x^3y^3$,该项中 $x^3y^3$ 的系数为10.

在$\frac{y^2}{x}T_{k+1}=C_5^kx^{4-k}y^{k+2}$ 中,令 $k=1$,可得$\frac{y^2}{x}T_2=C_5^1x^3y^3$,该项中 $x^3y^3$ 的系数为5.

所以 $x^3y^3$ 的系数为 $10+5=15$.

3.若 $n$ 为一组从小到大排列的数1,2,4,8,9,10的第60百分位数,则$\left(\sqrt[3]{x}+\frac{1}{2x}\right)^n$ 的展开式的常数项是　　(　　)

A.6　　B.7　　C.8　　D.10

B　**解析**:由 $6\times60\%=3.6$,可知 $n=8$.所以$\left(\sqrt[3]{x}+\frac{1}{2x}\right)^8$ 的展开式的通项为 $T_{k+1}=C_8^k\cdot(\sqrt[3]{x})^{8-k}\cdot\left(\frac{1}{2}x^{-1}\right)^k=C_8^k\cdot\left(\frac{1}{2}\right)^k\cdot x^{\frac{8-4k}{3}}$.

令$\frac{8-4k}{3}=0$,得 $k=2$,所以常数项为 $C_8^2\times\left(\frac{1}{2}\right)^2=7$.

故选B.

4.设 $a,b,m(m>0)$为整数,若 $a$ 和 $b$ 被 $m$ 除得的余数相同,则称 $a$ 和 $b$ 对模 $m$ 同余,记为 $a\equiv b(\mathrm{mod}\ m)$.若 $a=C_{20}^0+C_{20}^1\times3+C_{20}^2\times3^2+\cdots+C_{20}^{20}\times3^{20}$,$a\equiv b(\mathrm{mod}\ 5)$,则 $b$ 的值可以是　　(　　)

A.2 004　　B.2 005

C.2 025　　D.2 026

D　**解析**:易知 $a=C_{20}^0+C_{20}^1\times3+C_{20}^2\times3^2+\cdots+C_{20}^{20}\times3^{20}=(1+3)^{20}=4^{20}=(5-1)^{20}=C_{20}^0\times5^{20}-C_{20}^1\times5^{19}+\cdots-C_{20}^{19}\times5+C_{20}^{20}$.

因为 $C_{20}^0\times5^{20}-C_{20}^1\times5^{19}+\cdots-C_{20}^{19}\times5$ 能被5整除,所以 $a$ 除以5余 $C_{20}^{20}=1$.

又因为 $a\equiv b(\mathrm{mod}\ 5)$,选项中2 026除以5余1,所以 $b$ 的值可以是2 026.故选D.

5.$\left(x^2+\frac{1}{x}+1\right)^6$ 的展开式中 $x^3$ 的系数为______.

80　**解析**:因为$\left(x^2+\frac{1}{x}+1\right)^6$ 可以看成6个相同因式$\left(x^2+\frac{1}{x}+1\right)$相乘,

所以$\left(x^2+\frac{1}{x}+1\right)^6$ 的展开式中含 $x^3$ 的项为3个因式取 $x^2$、3个因式取$\frac{1}{x}$或2个因式取 $x^2$、1个因式取$\frac{1}{x}$、3个因式取1所得,

所以$\left(x^2+\frac{1}{x}+1\right)^6$ 的展开式中 $x^3$ 的系数为 $C_6^3C_3^3+C_6^2C_4^1C_3^3=80$.

6.若$\left(ax-\frac{b}{x}\right)^6$ 的展开式中常数项为$-160$,求$\frac{a^2+b^2+2}{ab}$的最小值.

**解**:$\left(ax-\frac{b}{x}\right)^6$ 的展开式的通项为 $T_{k+1}=C_6^k(ax)^{6-k}\left(-\frac{b}{x}\right)^k=C_6^ka^{6-k}(-b)^kx^{6-2k}$,

令 $6-2k=0$,得 $k=3$.

由题意得 $T_4=C_6^3a^{6-3}(-b)^3=-20a^3b^3=-160$,

即 $a^3b^3=8$,即 $ab=2$,因此$\frac{a^2+b^2+2}{ab}=\frac{a^2+b^2+2}{2}\geqslant\frac{2\sqrt{a^2b^2}}{2}+1=3$,

当且仅当 $a=b=\sqrt{2}$ 或 $a=b=-\sqrt{2}$ 时,等号成立.

故$\frac{a^2+b^2+2}{ab}$的最小值为3.

# 6.3.2 二项式系数的性质

**学习任务目标**

**1.** 掌握展开式中二项式系数的对称性、增减性与最大值.

**2.** 会用赋值法求各二项式系数的和或某些项的系数的和.

## 问题式预习

**【知识清单】**

**知识点一 二项式系数的性质**

(1)对称性：在$(a+b)^n$的展开式中，与首末两端"等距离"的两个二项式系数相等，即$C_n^0=C_n^n$，$C_n^1=C_n^{n-1}$，…，$C_n^r=C_n^{n-r}$.

(2)增减性与最大值：当$k<\frac{n+1}{2}$时，$C_n^k$随$k$的增加而增大；由对称性知，二项式系数的后半部分，$C_n^k$随$k$的增加而减小.当$n$是偶数时，中间的一项$C_n^{\frac{n}{2}}$取得最大值；当$n$是奇数时，中间的两项$C_n^{\frac{n-1}{2}}$与$C_n^{\frac{n+1}{2}}$相等，且同时取得最大值.

**知识点二 二项式系数的和**

(1)$C_n^0+C_n^1+C_n^2+\cdots+C_n^n=2^n$；

(2)$C_n^0+C_n^2+C_n^4+\cdots=C_n^1+C_n^3+C_n^5+\cdots=2^{n-1}$.

**【概念辨析】**

**1.** 判断正误(正确的打"√"，错误的打"×").

(1)二项展开式的各二项式系数的和为$C_n^1+C_n^2+\cdots+C_n^n$. ( × )

(2)二项展开式中系数最大项与二项式系数最大项相同. ( × )

(3)二项展开式的项的系数是先增后减的. ( × )

**2.** $\left(x-\frac{1}{x}\right)^{11}$的展开式中二项式系数最大的项是( )

A.第 6 项　　B.第 8 项

C.第 5 项与第 6 项　　D.第 6 项与第 7 项

D **解析**：因为$n=11$为奇数，所以展开式中二项式系数$C_{11}^5$与$C_{11}^6$相等，且同时取得最大值，即第 6 项和第 7 项的二项式系数相等，且最大.

**3.** $(1-2x)^{15}$的展开式中的各项系数的和是( )

A.1　　B.$-1$

C.$2^{15}$　　D.$3^{15}$

B **解析**：令$x=1$，得各项系数的和为$-1$.

**4.** 若$(1+x)^n(3-x)$的展开式中各项系数的和为 1 024，则$n$的值为( )

A.8　　B.9

C.10　　D.11

B **解析**：由题意知$(1+1)^n\times(3-1)=1\ 024$，即$2^{n+1}=1\ 024$，所以$n=9$.故选 B.

**5.** 请思考并回答下列问题：

(1)如何理解二项式定理中的$a$，$b$？

**提示**：实际上，$a$，$b$既可以取任意实数，也可以取任意多项式，还可以是别的.我们可以根据具体问题的需要灵活选取$a$，$b$的值.

(2)你能用组合的意义解释一下组合等式"$C_n^m=C_n^{n-m}$"吗？

**提示**：从$n$个不同元素中取出$m(m\leqslant n)$个元素，则剩余$(n-m)$个元素，故从$n$个不同元素中取出$m$个元素的组合数与取出$(n-m)$个元素的组合数相等，即$C_n^m=C_n^{n-m}$.

## 任务型课堂

### 任务1 二项式系数的性质的简单应用

**探究活动**

**例 1** (1)在$\left(\sqrt{x}-\frac{1}{2x}\right)^6$的展开式中，二项式系数最大的是( )

A.第 3 项　　B.第 4 项

C.第 5 项　　D.第 3 项和第 4 项

B **解析**：$\left(\sqrt{x}-\frac{1}{2x}\right)^6$的展开式共有 7 项，则二项式系数最大的项是第 4 项.故选 B.

(2)若$\left(x-\frac{2}{x^2}\right)^n$的二项展开式中所有二项式系数的和等于 32，则在展开式中含$x^2$项的系数是________.

$-10$ **解析**：因为$\left(x-\frac{2}{x^2}\right)^n$的二项展开式中所有二项式系数的和等于 32，

所以$2^n=32$，得$n=5$.

故$\left(x-\frac{2}{x^2}\right)^5$的展开式的通项为$T_{k+1}=C_5^k\cdot x^{5-k}\cdot$

$(-2)^k x^{-2k}=C_5^k(-2)^k x^{5-3k}$(其中 $0\leqslant k\leqslant 5$,且 $k\in\mathbf{N}$),令 $5-3k=2$,解得 $k=1$,

所以展开式中含 $x^2$ 项的系数为 $C_5^1\times(-2)^1=-10$.

【探究总结】

**二项式系数的性质**

(1)对称性:二项展开式中,与首尾两端"等距离"的两个二项式系数相等,即 $C_n^k=C_n^{n-k}$.

(2)增减性:当 $k<\dfrac{n+1}{2}$ 时,二项式系数递增;后半部分,二项式系数递减.

(3)最大值:当 $n$ 为奇数时,最中间两项的二项式系数相同且最大;当 $n$ 为偶数时,最中间一项的二项式系数最大.

(4)各二项式系数的和:$C_n^0+C_n^1+\cdots+C_n^k+\cdots+C_n^n=2^n$.

## 应用迁移

**1.** 若 $(a+b)^{2n}(n\in\mathbf{N}^*)$ 的展开式中第 6 项的二项式系数最大,则 $\left(\sqrt{x}-\dfrac{2}{x}\right)^n$ 的展开式中含 $x$ 项的系数为 (　　)

A. $-10$　　B. 10　　C. 5　　D. $-5$

A　**解析**:因为 $(a+b)^{2n}(n\in\mathbf{N}^*)$ 的展开式中第 6 项的二项式系数最大,且 $(a+b)^{2n}(n\in\mathbf{N}^*)$ 共有 $(2n+1)$ 项,则 $(a+b)^{2n}(n\in\mathbf{N}^*)$ 的展开式共 11 项,所以 $2n+1=11$,解得 $n=5$.

所以 $\left(\sqrt{x}-\dfrac{2}{x}\right)^5$ 的展开式的通项为 $T_{k+1}=C_5^k\cdot(\sqrt{x})^{5-k}\cdot\left(-\dfrac{2}{x}\right)^k=C_5^k(-2)^k x^{\frac{5-3k}{2}}(k=0,1,2,\cdots,5)$.

令 $\dfrac{5-3k}{2}=1$,可得 $k=1$,因此,$\left(\sqrt{x}-\dfrac{2}{x}\right)^5$ 的展开式中含 $x$ 项的系数为 $C_5^1\times(-2)=-10$.

**2.** 若 $\left(x^2+\dfrac{1}{x}\right)^n$ 的展开式的各二项式系数之和为 64,则展开式的常数项为________.

15　**解析**:因为 $\left(x^2+\dfrac{1}{x}\right)^n$ 的展开式的各二项式系数之和为 64,所以 $2^n=64$,所以 $n=6$.

所以 $\left(x^2+\dfrac{1}{x}\right)^6$ 的展开式的通项为 $T_{k+1}=C_6^k(x^2)^{6-k}\cdot\left(\dfrac{1}{x}\right)^k=C_6^k x^{12-3k}$.

令 $12-3k=0$,得 $k=4$,

所以 $T_5=C_6^4=15$,即常数项为 15.

**3.** 若 $\left(\dfrac{1}{2}+2x\right)^n$ 的展开式中前 3 项的二项式系数和等于 79,则展开式中二项式系数最大的项为________.

$924x^6$　**解析**:依题意,$C_n^0+C_n^1+C_n^2=1+n+\dfrac{n(n-1)}{2}=79$,解得 $n=12$ 或 $n=-13$(舍去),

即 $\left(\dfrac{1}{2}+2x\right)^{12}$ 的展开式有 13 项,最中间的项为第 7 项,也是二项式系数最大的项,

即 $T_7=C_{12}^6\left(\dfrac{1}{2}\right)^6(2x)^6=924x^6$.

## 任务 2 求二项展开式的系数和

## 探究活动

**例 2**　已知 $(1-2x)^{2\,024}=a_0+a_1x+a_2x^2+\cdots+a_{2\,024}x^{2\,024}(x\in\mathbf{R})$,求 $a_0+a_1+a_2+\cdots+a_{2\,024}$ 的值.

**解**:令 $x=1$,可得 $a_0+a_1+a_2+\cdots+a_{2\,024}=1$.

[一题多思]

**思考 1.** 如何求 $a_0-a_1+a_2-\cdots-a_{2\,023}+a_{2\,024}$ 的值?

**解**:令 $x=-1$,可得 $a_0-a_1+a_2-\cdots-a_{2\,023}+a_{2\,024}=3^{2\,024}$.

**思考 2.** 将 $(1-2x)^{2\,024}=a_0+a_1x+a_2x^2+\cdots+a_{2\,024}x^{2\,024}$ 与 $(1+2x)^{2\,024}=b_0+b_1x+b_2x^2+\cdots+b_{2\,024}\cdot x^{2\,024}$ 对比,思考 $|a_0|+|a_1|+|a_2|+\cdots+|a_{2\,024}|$ 与 $b_0+b_1+b_2+\cdots+b_{2\,024}$ 有什么关系?

**解**:$|a_0|+|a_1|+|a_2|+\cdots+|a_{2\,024}|$ 与 $b_0+b_1+b_2+\cdots+b_{2\,024}$ 相等.

**思考 3.** 如何求 $a_1+2a_2+\cdots+2\,023a_{2\,023}+2\,024a_{2\,024}$ 的值?

**解**:对 $(1-2x)^{2\,024}=a_0+a_1x+a_2x^2+\cdots+a_{2\,024}\cdot x^{2\,024}$ 两边求导数,得 $2\,024(1-2x)^{2\,023}\cdot(-2)=a_1+2a_2x+\cdots+2\,023a_{2\,023}\cdot x^{2\,022}+2\,024a_{2\,024}\cdot x^{2\,023}$.令 $x=1$,得 $a_1+2a_2+\cdots+2\,023a_{2\,023}+2\,024a_{2\,024}=4\,048$.

**思考 4.** 如何求 $a_0$ 的值?

**解**:令 $x=0$,得 $a_0=1$.

**思考 5.** 如何求 $a_0+a_2+a_4+\cdots+a_{2\,024}$ 的值?

**解**:由例题与思考 1 可得,$a_0+a_2+a_4+\cdots+a_{2\,024}=\dfrac{(a_0+a_1+\cdots+a_{2\,024})+(a_0-a_1+a_2-\cdots+a_{2\,024})}{2}=\dfrac{3^{2\,024}+1}{2}$.

【探究总结】

**1.** 对形如 $(ax+b)^n$,$(ax^2+bx+c)^m(a,b,c\in\mathbf{R},m,n\in\mathbf{N}^*)$ 的式子求其展开式的各项系数之和,常用赋值法,只需令 $x=1$ 即可;对形如 $(ax+by)^n(a,b$

$\in \mathbf{R}, n\in \mathbf{N}^*$）的式子求其展开式的各项系数之和，只需令 $x=y=1$ 即可.

**2.** 一般地，已知二项展开式 $a_0+a_1x+a_2x^2+\cdots+a_nx^n=f(x)$，则展开式中各项系数之和为 $a_0+a_1+a_2+\cdots+a_n=f(1)$，奇数项系数之和为 $a_0+a_2+a_4+\cdots=\dfrac{f(1)+f(-1)}{2}$，偶数项系数之和为 $a_1+a_3+a_5+\cdots=\dfrac{f(1)-f(-1)}{2}$.

### 应用迁移

**1.** 若 $(1+x)(1-2x)^8=a_0+a_1x+\cdots+a_9x^9, x\in \mathbf{R}$，则 $a_1\cdot 2+a_2\cdot 2^2+\cdots+a_9\cdot 2^9$ 的值为（　　）

A. $2^9$　　B. $2^9-1$　　C. $3^9$　　D. $3^9-1$

D　**解析**：令 $x=0$，得 $a_0=1$. 令 $x=2$，则 $a_0+2a_1+2^2a_2+\cdots+2^9a_9=3^9$，所以 $2a_1+2^2a_2+\cdots+2^9a_9=3^9-1$.

**2.** 已知 $A_n^5=56C_n^7$，且 $(1-3x)^n=a_0+a_1x+a_2x^2+a_3x^3+\cdots+a_nx^n$.

(1) 求 $n$ 的值；

(2) 求 $\dfrac{a_1}{3}+\dfrac{a_2}{3^2}+\cdots+\dfrac{a_n}{3^n}$ 的值.

**解**：(1) 易知 $n\geqslant 7, n\in \mathbf{N}$. 因为 $A_n^5=56C_n^7$，

所以 $n(n-1)(n-2)(n-3)(n-4)$

$=56\times\dfrac{n(n-1)(n-2)(n-3)(n-4)(n-5)(n-6)}{7\times6\times5\times4\times3\times2\times1}$，

整理可得 $\dfrac{(n-5)(n-6)}{90}=1$，即 $n^2-11n-60=0$，

解得 $n=15$ 或 $n=-4$（舍去）. 故 $n$ 的值为 15.

(2) 由 (1) 得 $n=15$，

所以 $(1-3x)^n=(1-3x)^{15}=a_0+a_1x+a_2x^2+a_3x^3+\cdots+a_{15}x^{15}$.

令 $x=0$，可得 $a_0=1$；

令 $x=\dfrac{1}{3}$，可得 $\left(1-3\times\dfrac{1}{3}\right)^{15}=a_0+\dfrac{a_1}{3}+\dfrac{a_2}{3^2}+\cdots+\dfrac{a_{15}}{3^{15}}=0$.

所以 $\dfrac{a_1}{3}+\dfrac{a_2}{3^2}+\cdots+\dfrac{a_{15}}{3^{15}}=-1$.

**3.** 在 $(2x-y)^9$ 的展开式中，求：

(1) 各二项式系数之和；

(2) 各项系数之和；

(3) 所有偶数项系数之和；

(4) 各项系数绝对值之和.

**解**：(1) 由题意得各二项式系数之和为 $C_9^0+C_9^1+C_9^2+\cdots+C_9^9=2^9=512$.

(2) 设 $(2x-y)^9=a_0x^9+a_1x^8y+a_2x^7y^2+\cdots+a_9y^9$，

则各项系数之和为 $a_0+a_1+a_2+\cdots+a_9$，

令 $x=1, y=1$，得 $a_0+a_1+a_2+\cdots+a_9=(2-1)^9=1$.

(3) 由 (2) 知 $a_0+a_1+a_2+\cdots+a_9=1$，令 $x=1, y=-1$ 可得 $a_0-a_1+a_2-\cdots-a_9=3^9$，

将两式相减，可得 $a_1+a_3+a_5+a_7+a_9=\dfrac{1-3^9}{2}=-9\ 841$.

故所有偶数项系数之和为 $-9\ 841$.

(4)（方法一）$|a_0|+|a_1|+|a_2|+\cdots+|a_9|=a_0-a_1+a_2-a_3+\cdots-a_9$，

令 $x=1, y=-1$，则 $|a_0|+|a_1|+|a_2|+\cdots+|a_9|=a_0-a_1+a_2-a_3+\cdots-a_9=3^9=19\ 683$.

（方法二）$|a_0|+|a_1|+|a_2|+\cdots+|a_9|$ 即为 $(2x+y)^9$ 展开式中各项系数和，

令 $x=1, y=1$，得 $|a_0|+|a_1|+|a_2|+\cdots+|a_9|=3^9=19\ 683$.

故各项系数绝对值之和为 19 683.

## 任务 3 二项展开式中的系数的最大值问题

### 探究活动

**例 3**　在 $\left(\dfrac{1}{\sqrt{x}}-2x\right)^n$ 的展开式中，第 3 项的二项式系数是第 2 项的二项式系数的 4 倍.

(1) 求 $n$ 的值；

(2) 求 $\left(\dfrac{1}{\sqrt{x}}-2x\right)^n$ 的展开式中的第 6 项的系数及常数项；

(3) 展开式中系数绝对值最大的项是第几项？

**解**：(1) 由题意可得 $C_n^2=4\cdot C_n^1$，即 $\dfrac{n(n-1)}{2}=4n$，整理得 $n^2-9n=0$，

解得 $n=9$ 或 $n=0$. 又 $n\geqslant 2$，所以 $n=9$.

(2) 由 $n=9$，得 $\left(\dfrac{1}{\sqrt{x}}-2x\right)^9$ 的展开式的通项为 $T_{k+1}=C_9^k\cdot\left(\dfrac{1}{\sqrt{x}}\right)^{9-k}(-2x)^k=(-2)^k\cdot C_9^kx^{\frac{3k-9}{2}}$，

则 $T_6=(-2)^5\cdot C_9^5x^{\frac{3\times5-9}{2}}=-32\times126x^3=-4\ 032x^3$.

令 $k=3$，则 $T_4=(-2)^3\cdot C_9^3x^0=-8\times84=-672$.

故展开式中的第 6 项的系数及常数项分别为 $-4\ 032$，$-672$.

(3)设第$(k+1)$项的系数绝对值最大，

则有$\begin{cases}2^k\cdot C_9^k\geqslant 2^{k+1}\cdot C_9^{k+1},\\ 2^k\cdot C_9^k\geqslant 2^{k-1}\cdot C_9^{k-1},\end{cases}$

即$\begin{cases}\dfrac{9!}{k!(9-k)!}\geqslant 2\times\dfrac{9!}{(k+1)!(8-k)!},\\ 2\times\dfrac{9!}{k!(9-k)!}\geqslant\dfrac{9!}{(k-1)!(10-k)!},\end{cases}$

即$\begin{cases}k+1\geqslant 2(9-k),\\ 2(10-k)\geqslant k,\end{cases}$解得$\dfrac{17}{3}\leqslant k\leqslant\dfrac{20}{3}$.

又$k\in\mathbf{N}$，所以$k=6$，即展开式中系数绝对值最大的项是第7项.

**【探究总结】**

**二项展开式中系数最大的项的求法**

求展开式中系数最大的项与求二项式系数最大的项是不同的，需要根据各项系数的正、负变化情况进行分析.如求$(a+bx)^n(a,b\in\mathbf{R})$的展开式中系数最大的项，一般采用待定系数法.设展开式中各项的系数分别为$A_0,A_1,A_2,\cdots,A_n$，且第$(r+1)$项的系数最大，根据$\begin{cases}A_r\geqslant A_{r-1}\\ A_r\geqslant A_{r+1}\end{cases}$解出$r$，即得出系数最大的项.

## 应用迁移

(2024·全国甲卷)$\left(\dfrac{1}{3}+x\right)^{10}$的展开式中，各项系数中的最大值为________.

5　**解析**：$\left(\dfrac{1}{3}+x\right)^{10}$的展开式的通项为$T_{k+1}=C_{10}^k\left(\dfrac{1}{3}\right)^{10-k}x^k$，$0\leqslant k\leqslant 10$且$k\in\mathbf{Z}$，

设展开式中第$(k+1)$项的系数最大，

则$\begin{cases}C_{10}^k\left(\dfrac{1}{3}\right)^{10-k}\geqslant C_{10}^{k+1}\left(\dfrac{1}{3}\right)^{9-k},\\ C_{10}^k\left(\dfrac{1}{3}\right)^{10-k}\geqslant C_{10}^{k-1}\left(\dfrac{1}{3}\right)^{11-k},\end{cases}$

解得$\begin{cases}k\geqslant\dfrac{29}{4},\\ k\leqslant\dfrac{33}{4},\end{cases}$即$\dfrac{29}{4}\leqslant k\leqslant\dfrac{33}{4}$.又$k\in\mathbf{Z}$，故$k=8$，

所以展开式中系数最大的项是第9项，且该项系数为$C_{10}^8\times\left(\dfrac{1}{3}\right)^2=5$.

## 提质归纳

- 二项式系数的性质
  - 性质
    - 对称性：与首末两端“等距离”的两个二项式系数相等，即$C_n^k=C_n^{n-k}$
    - 增减性与最大值：
      - 当$k<\dfrac{n+1}{2}$时，二项式系数是递增的；后半部分，二项式系数是递减的
      - 当$n$为偶数时，中间一项的二项式系数$C_n^{\frac{n}{2}}$最大；当$n$为奇数时，中间两项的二项式系数$C_n^{\frac{n-1}{2}}$与$C_n^{\frac{n+1}{2}}$相等且最大
    - 二项式系数的和：$C_n^0+C_n^1+C_n^2+\cdots+C_n^n=2^n$；$C_n^0+C_n^2+C_n^4+\cdots=C_n^1+C_n^3+C_n^5+\cdots=2^{n-1}$
  - 二项式定理的应用

# 课后素养评价（八）　二项式系数的性质

## A组　学习·理解

**1.**已知$(3-x)^n=a_0+a_1x+a_2x^2+\cdots+a_nx^n$.若展开式中第2项的二项式系数与第4项的二项式系数相等，则$a_0-a_1+a_2-\cdots+(-1)^na_n$等于　(　　)

A.32　　B.64　　C.128　　D.256

D　**解析**：因为$C_n^1=C_n^3$，所以$n=4$.令二项式中$x=-1$，可得$a_0-a_1+a_2-a_3+a_4=4^4=256$.故选D.

**2.**(多选)已知$(x-1)^n$的展开式中奇数项的二项式系数之和是64，则　(　　)

A.$n=7$

B.所有项的系数和为0

C.偶数项的系数和为$-64$

D.展开式的中间项为$-35x^3$和$35x^4$

ABC　**解析**：由已知可得$2^{n-1}=64$，解得$n=7$，故A正确；$(x-1)^7$的展开式中共有8项，令$x=1$，得所有项的系数和为0，故B正确；偶数项的系数和为$-64$，故C正确；展开式的中间项为第4项与第5项，$T_4=C_7^3x^4\cdot(-1)^3=-35x^4$，$T_5=C_7^4x^3(-1)^4=35x^3$，故D错误.故选ABC.

**3.**(多选)在$\left(3x-\dfrac{1}{\sqrt{x}}\right)^n$的展开式中，各项系数的和与各二项式系数的和之和为128，则　(　　)

A.各二项式系数的和为64

B.各项系数的和为64

C.常数项为$-135$

D.常数项为135

ABD　**解析**：在$\left(3x-\dfrac{1}{\sqrt{x}}\right)^n$的展开式中，各项系数

的和与各二项式系数的和之和为 128.令 $x=1$,得各项系数的和为 $2^n$,各二项式系数的和为 $2^n$,则 $2\times2^n=128$,得 $n=6$,即各二项式系数的和为 64,各项系数的和也为 64.$\left(3x-\frac{1}{\sqrt{x}}\right)^6$ 的展开式的通项为

$T_{k+1}=C_6^k\cdot(3x)^{6-k}\cdot\left(-\frac{1}{\sqrt{x}}\right)^k=C_6^k\cdot(-1)^k3^{6-k}\cdot x^{6-\frac{3}{2}k}$.令 $6-\frac{3}{2}k=0$,得 $k=4$.因此,展开式中的常数项为 $T_5=C_6^4\times(-1)^4\times3^2=135$.

**4.** 已知 $m$ 为正整数,$(x+y)^{2m}$ 的展开式的二项式系数的最大值为 $a$,$(x+y)^{2m+1}$ 的展开式的二项式系数的最大值为 $b$,且 $13a=7b$,则 $m$ 的值为 ( )

A.4　　B.5　　C.6　　D.7

C　**解析**:由题意可知 $C_{2m}^m=a$,$C_{2m+1}^m=C_{2m+1}^{m+1}=b$.因为 $13a=7b$,

所以 $13C_{2m}^m=7C_{2m+1}^m$,即 $13\times\frac{(2m)!}{m!\ m!}=7\times\frac{(2m+1)!}{m!\ (m+1)!}$,

所以 $13=7\times\frac{2m+1}{m+1}$,解得 $m=6$.故选 C.

**5.** $\left(\sqrt{x}+\frac{1}{x}\right)^n$ 的展开式中只有第 6 项的二项式系数最大,则第 4 项为________.

$120x^{\frac{1}{2}}$　**解析**:因为展开式中只有第 6 项的二项式系数最大,即 $\frac{n}{2}+1=6$,所以 $n=10$,

所以 $T_4=C_{10}^3(\sqrt{x})^7\left(\frac{1}{x}\right)^3=120x^{\frac{1}{2}}$.

**6.** 已知 $\left(x+\frac{m}{x}\right)^n$ 的展开式的各二项式系数的和为 256.

(1)求 $n$ 的值;

(2)若展开式中常数项为 $\frac{35}{8}$,求 $m$ 的值;

(3)若 $(x+m)^n$ 的展开式中系数最大的项只有第 6 项和第 7 项,求 $m$ 的值.

**解**:(1)由各二项式系数的和为 $2^n=256$,可得 $n=8$.

(2)设常数项为第 $(r+1)$ 项,则

$T_{r+1}=C_8^rx^{8-r}\left(\frac{m}{x}\right)^r=C_8^rm^rx^{8-2r}$.

令 $8-2r=0$,得 $r=4$,则 $C_8^4m^4=\frac{35}{8}$,

解得 $m=\pm\frac{1}{2}$.

(3)易知 $m>0$,设第 $(r+1)$ 项的系数最大.

则 $\begin{cases}C_8^rm^r\geqslant C_8^{r-1}m^{r-1},\\C_8^rm^r\geqslant C_8^{r+1}m^{r+1},\end{cases}$ 化简可得 $\frac{8m-1}{m+1}\leqslant r\leqslant\frac{9m}{m+1}$.

由于只有第 6 项和第 7 项的系数最大,

所以 $\begin{cases}4<\frac{8m-1}{m+1}\leqslant5,\\6\leqslant\frac{9m}{m+1}<7,\end{cases}$ 解得 $\begin{cases}\frac{5}{4}<m\leqslant2,\\2\leqslant m<\frac{7}{2}.\end{cases}$

所以 $m$ 的值为 2.

## B组　应用·实践

**1.** 已知 $(x^2+1)(x-2)^{10}=a_0(x-1)^{12}+a_1(x-1)^{11}+\cdots+a_{11}(x-1)+a_{12}$,则 $a_0+a_1+\cdots+a_{11}$ 的值为 ( )

A.2　　B.0

C.$-2$　　D.$-4$

C　**解析**:在展开式中,令 $x=2$,得 $a_0+a_1+\cdots+a_{11}+a_{12}=0$.令 $x=1$,得 $a_{12}=2$.所以 $a_0+a_1+\cdots+a_{11}=-2$.故选 C.

**2.** 已知 $(1-2x)^6=a_0+a_1x+a_2x^2+\cdots+a_6x^6$,则 $|a_0|+|a_1|+|a_2|+\cdots+|a_6|=$ ( )

A.1　　B.$-1$

C.$3^6$　　D.$2^6$

C　**解析**:由已知得展开式中 $a_0,a_2,a_4,a_6$ 大于零,$a_1,a_3,a_5$ 小于零,所以 $|a_0|+|a_1|+|a_2|+\cdots+|a_6|=a_0-a_1+a_2-a_3+a_4-a_5+a_6$.

令 $x=-1$,得 $a_0-a_1+a_2-a_3+a_4-a_5+a_6=3^6$.

所以 $|a_0|+|a_1|+|a_2|+\cdots+|a_6|=3^6$.

**3.** 在 $(a+b)^{10}$ 的二项展开式中,与第 3 项的二项式系数相同的项是 ( )

A.第 8 项　　B.第 7 项

C.第 9 项　　D.第 10 项

C　**解析**:由二项式系数的性质,与首末两端"等距离"的两项的二项式系数相等,可知选 C.

**4.** (多选)已知 $\left(ax^2+\frac{1}{\sqrt{x}}\right)^{10}$ $(a>0)$ 的展开式的各项系数之和为 1 024,则下列说法正确的是 ( )

A.展开式中奇数项的二项式系数之和为 256

B.展开式中第 6 项的系数最大

C.展开式中存在含 $x^6$ 的项

D.展开式中含 $x^{15}$ 的项的系数为 45

BD　**解析**:因为展开式的各项系数之和为 1 024,所以令 $x=1$,得 $(a+1)^{10}=1\ 024$.因为 $a>0$,所以 $a=1$,则 $\left(x^2+\frac{1}{\sqrt{x}}\right)^{10}$ 的展开式的通项为 $T_{k+1}=C_{10}^k(x^2)^{10-k}\left(\frac{1}{\sqrt{x}}\right)^k=C_{10}^kx^{20-\frac{5}{2}k}$.

展开式中奇数项的二项式系数之和为 $\frac{1}{2}\times2^{10}=512$，故 A 错误；

由展开式的通项可知，项的系数与其二项式系数相同，且展开式有 11 项，故展开式中第 6 项的系数最大，故 B 正确；

令 $20-\frac{5}{2}k=6$，可得 $k=\frac{28}{5}$ 不是自然数，则展开式中不存在含 $x^6$ 的项，故 C 错误；

令 $20-\frac{5}{2}k=15$，解得 $k=2$，所以展开式中含 $x^{15}$ 的项的系数为 $C_{10}^{2}=45$，故 D 正确.

故选 BD.

**5.** $(2x^2-3x+a)^5$ 的展开式的各项系数之和为 1，则该展开式中含 $x^7$ 项的系数是　　(　　)

A. $-600$　　B. $-840$

C. $-1\ 080$　　D. $-2\ 040$

D　**解析**：因为 $(2x^2-3x+a)^5$ 的展开式的各项系数之和为 1，

令 $x=1$，得 $(-1+a)^5=1$，解得 $a=2$，

所以 $(2x^2-3x+2)^5$ 的展开式中含 $x^7$ 项为 $C_5^3\cdot(2x^2)^3C_2^1\cdot(-3x)\times2+C_5^2(2x^2)^2C_3^3(-3x)^3=-2\ 040x^7$，

所以该展开式中含 $x^7$ 项的系数是 $-2\ 040$.

故选 D.

**6.**（2022·新高考全国Ⅰ卷）$\left(1-\frac{y}{x}\right)(x+y)^8$ 的展开式中 $x^2y^6$ 的系数为________.（用数字作答）

$-28$　**解析**：因为 $\left(1-\frac{y}{x}\right)(x+y)^8=(x+y)^8-\frac{y}{x}(x+y)^8$，所以 $\left(1-\frac{y}{x}\right)(x+y)^8$ 的展开式中含 $x^2y^6$ 的项为 $C_8^6x^2y^6-\frac{y}{x}C_8^5x^3y^5=-28x^2y^6$.

所以 $\left(1-\frac{y}{x}\right)(x+y)^8$ 的展开式中 $x^2y^6$ 的系数为 $-28$.

**7.** $\left(\frac{1}{\sqrt{x}}-2x\right)^n$ 的展开式中第 3 项与第 7 项的二项式系数相等，则 $\left(\frac{1}{\sqrt{x}}-2x\right)^n$ 的展开式中各项系数的绝对值之和为________.

$3^8$（或者写成 6 561）　**解析**：因为展开式中第 3 项与第 7 项的二项式系数相等，所以 $C_n^2=C_n^6$. 由组合数的性质可得 $n=6+2=8$，即 $\left(\frac{1}{\sqrt{x}}-2x\right)^n=\left(\frac{1}{\sqrt{x}}-2x\right)^8$.

因为 $\left(\frac{1}{\sqrt{x}}-2x\right)^8$ 的展开式中各项系数的绝对值之和等于 $\left(\frac{1}{\sqrt{x}}+|-2x|\right)^8$ 的展开式中各项系数和，

所以，令 $x=1$ 可得 $\left(\frac{1}{\sqrt{1}}+2\right)^8=3^8=6\ 561$.

故答案为 $3^8$（或者写成 6 561）.

**8.** 已知 $C_n^2=C_n^3$，且 $(1-2x)^n=a_0+a_1x+a_2x^2+a_3x^3+\cdots+a_nx^n$.

(1) 求 $(a_0+a_1)+(a_0+a_2)+(a_0+a_3)+\cdots+(a_0+a_n)$ 的值；

(2) 求 $\frac{a_1}{2}+\frac{a_2}{2^2}+\cdots+\frac{a_n}{2^n}$ 的值；

(3) 求 $a_1+2a_2+3a_3+\cdots+na_n$ 的值.

**解**：(1) 因为 $C_n^2=C_n^3$，所以 $n=5$.

则 $(1-2x)^5=a_0+a_1x+a_2x^2+a_3x^3+a_4x^4+a_5x^5$.

令 $x=0$，则 $a_0=1$.

令 $x=1$，则 $a_0+a_1+a_2+a_3+a_4+a_5=-1$.

所以 $(a_0+a_1)+(a_0+a_2)+(a_0+a_3)+\cdots+(a_0+a_n)=3$.

(2) 令 $x=\frac{1}{2}$，则 $0=a_0+\frac{a_1}{2}+\frac{a_2}{2^2}+\cdots+\frac{a_5}{2^5}$，所以 $\frac{a_1}{2}+\frac{a_2}{2^2}+\cdots+\frac{a_5}{2^5}=-1$.

(3) 对 $(1-2x)^5=a_0+a_1x+a_2x^2+a_3x^3+a_4x^4+a_5x^5$ 两边关于 $x$ 求导，则 $5(1-2x)^4(-2)=a_1+2a_2x+3a_3x^2+4a_4x^3+5a_5x^4$.

再令 $x=1$，可得 $a_1+2a_2+3a_3+\cdots+na_n=-10$.

★★★

# 迁移应用

**学习目标**

1.能够运用排列与组合相关知识解决实际生活中的问题,熟练掌握两个计数原理,能够准确地利用计数原理进行运算.

2.能根据通项公式求出二项展开式中指定项的系数、特定项的次数等,掌握运用二项式定理解决各类问题的方法.

## 类型一 排列、组合的综合应用

**例 1** 小赵、小钱、小孙、小李、小周 5 位同学报名参加 3 个项目,每人只报名 1 个项目,每个项目至少 1 人报名,小赵和小钱不参加同一个项目,则不同的报名方法共有 ( )

A.72 种 B.114 种 C.120 种 D.144 种

B **解析**:(方法一)5 位同学报名参加 3 个项目,人数构成分为 2+2+1 与 3+1+1 两种情况.

先分组,再将不同组分配去参加项目:

当人数构成为 2+2+1 时,

小赵和小钱分别在两个 2 人组:$C_3^1C_2^1A_3^3=36$;

小赵和小钱分别在 2 人组和 1 人组:$C_2^1C_3^1A_3^3=36$;

当人数构成为 3+1+1 时,

小赵和小钱分别在两个 1 人组:$A_3^3=6$;

小赵和小钱分别在 1 人组和 3 人组:$C_2^1C_3^2A_3^3=36$.

所以共有 36+36+6+36=114 种不同的报名方法.

(方法二)不考虑小赵与小钱的特殊要求,5 位同学报名参加 3 个项目,人数构成分为 2+2+1 与 3+1+1 两种情况:

①2+2+1:$\dfrac{C_5^2C_3^2}{A_2^2}A_3^3=90$;②3+1+1:$\dfrac{C_5^3C_2^1}{A_2^2}A_3^3=60$.故共有 90+60=150 种不同的报名方法.

假如小赵与小钱参加同一个项目,分为他们都在同一个 2 人组和都在 3 人组两种情况,

①都在同一个 2 人组:$C_3^2\times A_3^3=18$;②都在 3 人组:$\dfrac{C_3^1C_2^1}{A_2^2}A_3^3=18$.

考虑两人的特殊要求之后,共有 150−18−18=114 种不同的报名方法.故选 B.

**【总结升华】**

排列与组合是两类特殊的计数方式,在计数原理的应用中起着举足轻重的作用,解决排列与组合的综合问题要树立先选后排,特殊元素(特殊位置)优先的原则.解决相关问题时要注意以下几点:

(1)首先要分清该问题是排列问题还是组合问题.

(2)对于含有多个限制条件的复杂问题,应认真分析每个限制条件,再考虑是分类还是分步,分类时要不重不漏,分步时要步步相接.

(3)对于含有“至多”“至少”的问题,常采用间接法,此时要考虑全面.

## 类型二 二项式定理的应用

**例 2** 已知在$\left(\sqrt{x}-\dfrac{2}{\sqrt[3]{x}}\right)^n$的展开式中,第 5 项的系数与第 3 项的系数之比是 56∶3.

(1)求展开式中的所有有理项;

(2)求展开式中系数的绝对值最大的项;

(3)求 $n+9C_n^2+81C_n^3+\cdots+9^{n-1}C_n^n$ 的值.

**解**:(1)由题意得 $C_n^4(-2)^4:C_n^2(-2)^2=56:3$,解得 $n=10$(负值舍去),此时$\left(\sqrt{x}-\dfrac{2}{\sqrt[3]{x}}\right)^{10}$的展开式的通项为 $T_{k+1}=C_{10}^k(\sqrt{x})^{10-k}\left(-\dfrac{2}{\sqrt[3]{x}}\right)^k=(-2)^kC_{10}^kx^{5-\frac{5k}{6}}$.

当 $5-\dfrac{5k}{6}$ 为整数时,$k$ 可取 0,6,于是有理项为 $T_1=x^5$ 和 $T_7=13\ 440$.

(2)设第$(k+1)$项系数的绝对值最大,

则$\begin{cases}C_{10}^k2^k\geqslant C_{10}^{k-1}2^{k-1},\\C_{10}^k2^k\geqslant C_{10}^{k+1}2^{k+1},\end{cases}$解得$\begin{cases}k\leqslant\dfrac{22}{3},\\k\geqslant\dfrac{19}{3}.\end{cases}$

又因为 $k\in\{1,2,3,\cdots,9\}$,所以 $k=7$,

当 $k=7$ 时,$T_8=-15\ 360x^{-\frac{5}{6}}$.

又因为当 $k=0$ 时,$T_1=x^5$,

当 $k=10$ 时,$T_{11}=(-2)^{10}x^{-\frac{10}{3}}=1\ 024x^{-\frac{10}{3}}$,

所以系数的绝对值最大的项为 $T_8=-15\ 360x^{-\frac{5}{6}}$.

(3)原式$=10+9\times C_{10}^2+81\times C_{10}^3+\cdots+9^{10-1}\times C_{10}^{10}$

$$=\frac{9\times C_{10}^{1}+9^{2}\times C_{10}^{2}+9^{3}\times C_{10}^{3}+\cdots+9^{10}\times C_{10}^{10}}{9}$$

$$=\frac{C_{10}^{0}+9\times C_{10}^{1}+9^{2}\times C_{10}^{2}+9^{3}\times C_{10}^{3}+\cdots+9^{10}\times C_{10}^{10}-1}{9}$$

$$=\frac{(1+9)^{10}-1}{9}=\frac{10^{10}-1}{9}.$$

**【总结升华】**

求形如$(a+b)^n(n\in \mathbf{N}^*)$的展开式中与特定项相关的量(常数项、参数值、有理项等)的步骤：

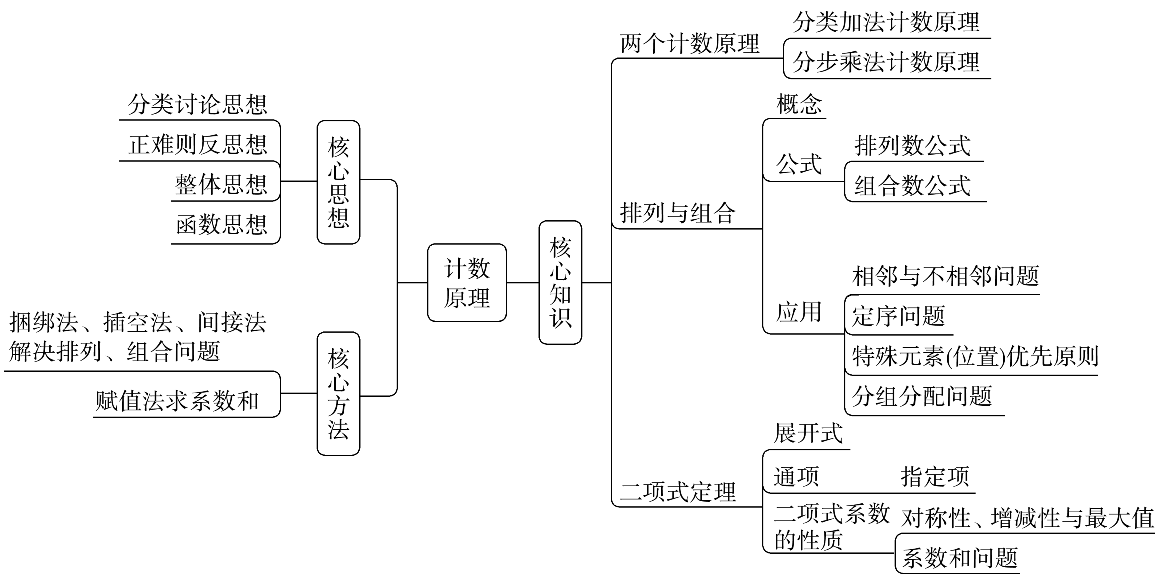
第一步,利用二项式定理写出二项展开式的通项公式$T_{k+1}=C_n^k a^{n-k}b^k$,常把字母和系数分离开来(注意符号不要出错)；

第二步,根据题目中的相关条件(如常数项要求指数为零,有理项要求指数为整数)列出相应方程(组)或不等式(组),解出$k$；

第三步,把$k$代入通项公式中,即可求出$T_{k+1}$,有时还需要先求$n$,再求$k$,才能求出$T_{k+1}$或者其他量.

# ★★★ 重构拓展

## 多维体系构建




**【问题探究】**

**1.**应用分类加法计数原理和分步乘法计数原理时,需要注意什么问题？举例说明.

**2.**在解决实际问题中的计数问题时,往往需要区分该问题是排列问题还是组合问题,你能举例说明排列与组合的区别与联系吗？

**3.**排列数公式是如何推导的？推导过程体现了什么数学思想与方法？

**4.**二项式系数与项的系数的区别是什么？如何用组合的意义解释二项式系数的对称性？

## 学科视野拓展

### 拓展一　组合数公式与杨辉三角的交汇问题

**【拓展总结】**

杨辉三角是中国古代数学的杰出研究成果之一,它把二项式系数图形化,把组合数内在的一些代数性质直观地在图形中体现出来,可以说二项式定理与杨辉三角形是一对天然的数形趣遇.因此二者的交汇命题是高考考查的热点之一.

**应用1**　杨辉三角在我国南宋数学家杨辉1261年所著的《详解九章算法》一书中被记载,它包含了很多有趣的组合数性质.如图,将杨辉三角从第1行开始的每一个数$C_n^r$都换成分数$\frac{1}{(n+1)C_n^r}$,得到的三角形称为“莱布尼茨三角形”,则“莱布尼茨三角形”第8行第

5 个数是________；若 $S_n=\frac{1}{4}+\frac{1}{20}+\frac{1}{60}+\frac{1}{140}+\cdots+\frac{1}{nC_{n-1}^{n-4}}+\frac{1}{(n+1)C_n^{n-3}}(n\geqslant 3)$，则 $S_n=$______（用含 $n$ 的代数式作答）.

| | 杨辉三角 | 莱布尼茨三角形 | |
|---|---|---|---|
| 第 0 行 | 1 | 1 | 第 0 行 |
| 第 1 行 | 1　1 | $\frac{1}{2}$　$\frac{1}{2}$ | 第 1 行 |
| 第 2 行 | 1　2　1 | $\frac{1}{3}$　$\frac{1}{6}$　$\frac{1}{3}$ | 第 2 行 |
| 第 3 行 | 1　3　3　1 | $\frac{1}{4}$　$\frac{1}{12}$　$\frac{1}{12}$　$\frac{1}{4}$ | 第 3 行 |
| ⋮ | ⋮ | ⋮ | ⋮ |

$\frac{1}{630}$　$\frac{1}{3}-\frac{2}{n(n-1)(n+1)}$　**解析**：观察杨辉三角中各数，要求第 8 行第 5 个数，所以 $n=8$，$r=4$，所以第 8 行第 5 个数为 $\frac{1}{(8+1)C_8^4}=\frac{1}{630}$.

由"莱布尼茨三角形"的特点可知，每个数均等于其"脚下"两个数之和，

所以 $\frac{1}{(n+1)C_n^{n-2}}+\frac{1}{(n+1)C_n^{n-3}}=\frac{1}{nC_{n-1}^{n-3}}$，

$\frac{1}{nC_{n-1}^{n-3}}+\frac{1}{nC_{n-1}^{n-4}}=\frac{1}{(n-1)C_{n-2}^{n-4}}$，

$\frac{1}{(n-1)C_{n-2}^{n-4}}+\frac{1}{(n-1)C_{n-2}^{n-5}}=\frac{1}{(n-2)C_{n-3}^{n-5}}$，…，

$\frac{1}{5C_4^1}+\frac{1}{5C_4^2}=\frac{1}{4C_3^1}$，$\frac{1}{4C_3^1}+\frac{1}{4C_3^0}=\frac{1}{3C_2^0}$，

将上述各式相加，得 $\frac{1}{(n+1)C_n^{n-2}}+\frac{1}{(n+1)C_n^{n-3}}+\frac{1}{nC_{n-1}^{n-4}}+\cdots+\frac{1}{20}+\frac{1}{4}=\frac{1}{3}$，

所以 $\frac{1}{(n+1)C_n^{n-2}}+S_n=\frac{1}{3}$，

即 $S_n=\frac{1}{3}-\frac{2}{n(n-1)(n+1)}$.

## 拓展二 新定义、新概念问题

**【拓展总结】**

**解决"新定义"问题，主要分如下几步：**

(1)对新定义进行信息提取，明确新定义的名称和符号.

(2)对提取的信息进行加工，与熟悉的相近的知识点对比，明确它们的相同点和相似点.

(3)若新定义的是运算、法则，直接按照法则计算即可；若新定义的是性质，一般要先判断性质的适用性，再利用性质解决问题.

注意：有关新定义的解答题一般在试卷的压轴位置，往往设置三问，第一问的难度并不大，所以基础差的考生也不要轻易放弃.

**应用 2**　组合数学研究的内容之一是计数，母函数是重要的计数工具之一.其定义如下：对于序列 $a_0,a_1,a_2,\cdots$，定义 $G(x)=a_0+a_1x+a_2x^2+\cdots$ 为序列 $a_0,a_1,a_2,\cdots$ 的母函数.母函数的计数方法与二项式定理的原理相似：假设有红、黄、蓝小球各一个，计算由它们组成的所有组合的个数，可分三步完成，即考虑每个小球是否参与组合.我们用 $x^0$ 即 1 代表小球不参与，$x$ 代表小球参与，根据分类加法计数原理，$1+x$ 代表一个小球是否参与组合的两种情况，根据分步乘法计数原理，用代数式 $(1+x)\cdot(1+x)(1+x)$ 表示三个小球是否参与组合的情况，所以母函数为 $G(x)=(1+x)(1+x)(1+x)=1+3x+3x^2+x^3$.例如，其中 $3x^2$ 中的系数 3 就是由两个小球构成的所有组合个数，而总的组合个数就是 $1+3+3+1=8$.

(1)假设有四个不同的小球，令 $a_n$ 为由它们组成的含有 $n$ 个小球的所有组合个数，试写出 $a_0,a_1,a_2,a_3,a_4$ 的母函数 $G(x)$；

(2)已知 $rC_{n+r}^r=(n+1)C_{n+r}^{n+1}$，其中 $1\leqslant r\leqslant n$.现有一序列 $a_0,a_1,a_2,\cdots,a_{2n}$ 的母函数 $G(x)=\sum_{i=1}^{n}i(1+x)^{n+i}=a_0+a_1x+\cdots+a_nx^n+\cdots+a_{2n}x^{2n}$，其中 $n\in\mathbf{N}^*$，求 $a_n$；

(3)在某班的 8 位男同学和 5 位女同学中，组一个由偶数个男生和不少于两个女生的小组，令 $a_n$ 为从 8 位男同学中选取 $n$ 位的所有组合个数，令 $b_n$ 为从 5 位女同学中选取 $n$ 位的所有组合个数，分别写出 $a_0,a_1,a_2,\cdots,a_n$ 和 $b_0,b_1,b_2,\cdots,b_n$ 的母函数 $A(x)$ 和 $B(x)$，并求总的组合个数.

**解**：(1)$G(x)=(1+x)(1+x)(1+x)(1+x)$

$=1+4x+6x^2+4x^3+x^4$.

(2)因为 $G(x)=(1+x)^{n+1}+2(1+x)^{n+2}+3(1+x)^{n+3}+\cdots+n(1+x)^{2n}$，

所以展开式中 $x^n$ 的系数为 $C_{n+1}^n+2C_{n+2}^n+3C_{n+3}^n+\cdots+nC_{2n}^n=C_{n+1}^1+2C_{n+2}^2+3C_{n+3}^3+\cdots+nC_{2n}^n$.

因为 $rC_{n+r}^r=(n+1)C_{n+r}^{n+1}$，

所以 $x^n$ 的系数为 $a_n=(n+1)(C_{n+1}^{n+1}+C_{n+2}^{n+1}+C_{n+3}^{n+1}+\cdots+C_{2n}^{n+1})=(n+1)(C_{n+2}^{n+2}+C_{n+2}^{n+1}+C_{n+3}^{n+1}+\cdots+C_{2n}^{n+1})=(n+1)(C_{n+3}^{n+2}+C_{n+3}^{n+1}+\cdots+C_{2n}^{n+1})=(n+1)C_{2n+1}^{n+2}$.

(3)显然 $a_1=a_3=a_5=a_7=0$，$a_0=1$，$a_2=C_8^2=28$，$a_4=C_8^4=70$，$a_6=C_8^6=28$，$a_8=1$，

故 $A(x)=1+28x^2+70x^4+28x^6+x^8$.

同样 $b_0=b_1=0$，$b_2=b_3=10$，$b_4=5$，$b_5=1$，

故 $B(x)=10x^2+10x^3+5x^4+x^5$，

令 $C(x)=A(x)\cdot B(x)=(1+28x^2+70x^4+28x^6+x^8)\cdot(10x^2+10x^3+5x^4+x^5)$

$=10x^2+10x^3+285x^4+281x^5+840x^6+728x^7+630x^8+350x^9+150x^{10}+38x^{11}+5x^{12}+x^{13}$，

$C(x)$ 中 $x^k$ 的系数 $c_k$ 为符合要求的 $k$ 个人组成的小组的个数，所有组合的个数为 $10+10+285+281+840+728+630+350+150+38+5+1=3\ 328$.

# 质量评估(一)

(范围:第六章)

## 一、单项选择题

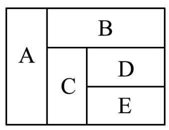
1.$4A_5^2+5C_4^2=$ (　　)

A.110　　B.98

C.124　　D.148

A　**解析**:$4A_5^2+5C_4^2=4\times5\times4+5\times\frac{4\times3}{2\times1}=110.$

2.如图所示,用五种不同的颜色分别为 A,B,C,D,E 五部分着色,相邻部分不能用同一种颜色,但同一种颜色可以反复使用,也可以不使用,则符合这种要求的不同的着色方法共有 (　　)




A.120 种　　B.240 种

C.480 种　　D.540 种

D　**解析**:先涂 A,有 5 种涂法,再涂 B,因为 B 与 A 相邻,所以 B 的颜色只要与 A 不同即可,有 4 种涂法;同理 C 有 3 种涂法,D 有 3 种涂法,E 有 3 种涂法.由分步乘法计数原理可知,共有 $5\times4\times3\times3\times3=540$ 种不同的着色方法.故选 D.

3.在$\left(2\sqrt{x}-\frac{1}{\sqrt{x}}\right)^6$的展开式中,含 $x^{-2}$ 的项的系数为 (　　)

A.12　　B.$-12$

C.$-2$　　D.2

B　**解析**:展开式的通项为 $T_{k+1}=(-1)^kC_6^k2^{6-k}\cdot x^{\frac{6-k}{2}}\cdot x^{-\frac{k}{2}}=(-1)^kC_6^k2^{6-k}x^{3-k}(k=0,1,2,\cdots,6)$,令 $3-k=-2$,得 $k=5$,所以 $T_6=-C_6^5\times2^{6-5}x^{-2}=-12x^{-2}$.故选 B.

4.将 A,B,C,D,E 排成一排,要求 A,B,C 在排列中的顺序为 A,B,C 或 C,B,A(可以不相邻),则不同的排列方法有 (　　)

A.12 种　　B.20 种

C.40 种　　D.60 种

C　**解析**:五个元素没有限制条件,全排列数为 $A_5^5$,若 A,B,C 的顺序为 A,B,C 或 C,B,A(可以不相邻),则不同的排列方法有 $2\cdot\frac{A_5^5}{A_3^3}=40$ 种.

5.已知 $S=(x-1)^4+4(x-1)^3+6(x-1)^2+4x-3$,则 $S$ 可化简为 (　　)

A.$x^4$　　B.$x^4+1$

C.$(x-2)^4$　　D.$x^4+4$

A　**解析**:$S=C_4^0(x-1)^4+C_4^1(x-1)^3+C_4^2(x-1)^2+C_4^3(x-1)+C_4^4=[(x-1)+1]^4=x^4$.故选 A.

6.某部队计划将 5 艘不同的军舰全部安排到甲、乙、丙三个海上区域进行军事演习,要求每个区域至少安排一艘军舰,且其中的军舰 A 必须安排在甲区域,则甲区域还有其他军舰的安排方案共有 (　　)

A.14 种　　B.24 种　　C.36 种　　D.50 种

C　**解析**:依题意,甲区域除军舰 A 外至少还有一艘军舰,至多还有两艘军舰.

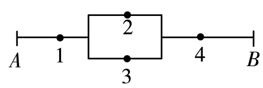
若甲区域除军舰 A 外还有一艘军舰,则安排方案共有 $C_4^1\times C_3^1\times C_2^2\times A_2^2=24$ 种;

若甲区域除军舰 A 外还有两艘军舰,则安排方案共有 $C_4^2\times A_2^2=12$ 种.

所以甲区域还有其他军舰的安排方案共有 $24+12=36$ 种.

故选 C.

7.如图所示,在 $A,B$ 间有四个焊接点 1,2,3,4.若焊接点脱落导致断路,则电路不通.现发现 $A,B$ 之间电路不通,则焊接点脱落的不同情况有 (　　)




A.9 种　　B.11 种　　C.13 种　　D.15 种

C　**解析**:按照焊接点脱落的个数进行分类:第 1 类,脱落 1 个,有 1,4,共 2 种;第 2 类,脱落 2 个,有(1,4),(2,3),(1,2),(1,3),(4,2),(4,3),共 6 种;第 3 类,脱落 3 个,有(1,2,3),(1,2,4),(2,3,4),(1,3,4),共 4 种;第 4 类,脱落 4 个,有(1,2,3,4),共 1 种.根据分类加法计数原理,共有 $2+6+4+1=13$ 种焊接点脱落的不同情况.故选 C.

8.为了全面推进乡村振兴,加快农村、农业现代化建设,某市准备派 6 名乡村振兴指导员到 A,B,C 三地指导工作;每地上午和下午各安排一名乡村振兴指导员,且每名乡村振兴指导员只能被安排一次,其中张指导员不安排到 C 地,李指导员不安排在下午,则不同的安排方案共有 (　　)

A.180 种　B.240 种　　C.480 种　　D.540 种

B　**解析**:李指导员安排在 C 地上午时,张指导员有

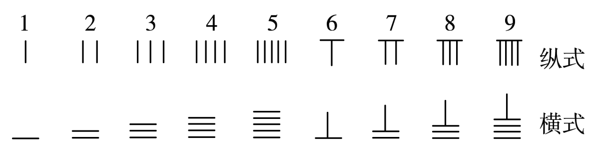
$C_4^1$ 种安排方案，其余 4 位指导员有 $A_4^4$ 种安排方案，则共有 $C_4^1A_4^4=96$ 种安排方案；

李指导员不安排在 C 地上午时，李指导员有 $C_2^1$ 种安排方案，张指导员有 $C_3^1$ 种安排方案，其余 4 位指导员有 $A_4^4$ 种安排方案，则共有 $C_2^1C_3^1A_4^4=144$ 种安排方案.

综上，共有 $96+144=240$ 种不同的安排方案.故选 B.

## 二、多项选择题

**9.**我国古代在珠算发明之前多是用算筹为工具来记数、列式和计算的.算筹实际上是一根根相同长度的小木棍.如图，算筹表示数 1～9 的方法有两种，即“纵式”和“横式”，规定个位数用纵式，十位数用横式，百位数用纵式，千位数用横式，万位数用纵式……依此类推，交替使用纵、横两式.例如：27 可以表示为“𝍪𝍦”.如果用算筹表示一个不含 0 的两位数，那么下列说法正确的是（　　）




A.36 可以表示为𝍫𝍮

B.用七根算筹表示不同的两位数，若十位为 1，则可以表示 9 个这样的两位数

C.用七根算筹表示不同的两位数，若十位为 4，则可以表示 41，42，43，46，47

D.用七根算筹表示不同的两位数，共可以表示 48 个这样的两位数

BC　**解析**：36 可以表示为𝍫𝍥，A 错误；

当十位为 1 时，个位可以是 1，2，3，4，5，6，7，8，9，共 9 个，B 正确；

当十位为 4 时，个位可以是 1，2，3，6，7，即可以表示 41，42，43，46，47，共 5 个，C 正确；

当十位为 2 时，个位可以是 1，2，3，4，5，6，7，8，9，共 9 个；

当十位为 3 时，个位可以是 1，2，3，4，6，7，8，共 7 个；

当十位为 5 时，个位可以是 1，2，6，共 3 个；

当十位为 6 时，个位可以是 1，2，3，4，5，6，7，8，9，共 9 个；

当十位为 7 时，个位可以是 1，2，3，4，6，7，8，共 7 个；

当十位为 8 时，个位可以是 1，2，3，6，7，共 5 个；

当十位为 9 时，个位可以是 1，2，6，共 3 个.

所以总共可表示 $9+5+9+7+3+9+7+5+3=57$ 个不同的两位数，D 错误.

**10.**身穿红、黄两种颜色衣服的各有两人，身穿蓝色衣服的有一人.现将这五人排成一行，则（　　）

A.穿黄色衣服的人不相邻的排法种数为 48

B.穿红色衣服的人相邻的排法种数为 48

C.穿红色衣服的人与穿黄色衣服的人同时相邻的排法种数为 36

D.穿相同颜色衣服的人不相邻的排法种数为 48

BD　**解析**：穿黄色衣服的人不相邻的排法有 $A_3^3A_4^2=72$ 种，穿红色衣服的人相邻的排法有 $A_2^2A_4^4=48$ 种，同理，穿黄色衣服的人相邻的排法也有 48 种.而穿红色、黄色衣服的人同时相邻的排法有 $A_2^2\cdot A_2^2\cdot A_3^3=24$ 种.故穿相同颜色衣服的人不相邻的排法有 $A_5^5-2\times48+24=48$ 种.

**11.**若 $(2x+\sqrt{3})^4=a_0+a_1x+a_2x^2+a_3x^3+a_4x^4$，则（　　）

A.$a_0=3$

B.$a_0+a_1+a_2+a_3+a_4=88+56\sqrt{3}$

C.$a_0+a_2+a_4=97$

D.$a_1+a_3=56\sqrt{3}$

CD　**解析**：令 $x=0$，得 $(\sqrt{3})^4=a_0$，即 $a_0=9$.

令 $x=1$，得 $(2+\sqrt{3})^4=a_0+a_1+a_2+a_3+a_4=97+56\sqrt{3}$.

令 $x=-1$，得 $(-2+\sqrt{3})^4=a_0-a_1+a_2-a_3+a_4=97-56\sqrt{3}$.

所以 $a_0+a_2+a_4=97$，$a_1+a_3=56\sqrt{3}$.

## 三、填空题

**12.**若 $(x+a)^2\left(\frac{1}{x}-1\right)^5$ 的展开式中常数项为 $-1$，则 $a$ 的值为________.

1 或 9　**解析**：由于 $(x+a)^2=x^2+2ax+a^2$，而 $\left(\frac{1}{x}-1\right)^5$ 的展开式通项为 $T_{k+1}=(-1)^kC_5^k\cdot x^{k-5}$，其中 $k=0,1,2,\cdots,5$.

于是 $\left(\frac{1}{x}-1\right)^5$ 的展开式中 $x^{-2}$ 的系数为 $(-1)^3\times C_5^3=-10$，

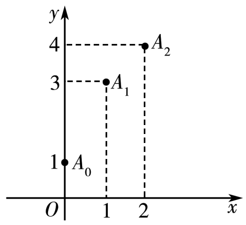
$x^{-1}$ 的系数为 $(-1)^4\times C_5^4=5$，常数项为 $-1$.

因此 $(x+a)^2\left(\frac{1}{x}-1\right)^5$ 的展开式中常数项为 $1\times(-10)+2a\times5+a^2\times(-1)=-a^2+10a-10$.

由题意得 $-a^2+10a-10=-1$，即 $a^2-10a+9=0$，解得 $a=1$ 或 $a=9$.

**13.**设 $a\neq0$，$n$ 是大于 1 的自然数，$\left(1+\frac{x}{a}\right)^n$ 的展开式

为 $a_0+a_1x+a_2x^2+\cdots+a_nx^n$.若点 $A_i(i,a_i)(i=0,1,2)$ 的位置如图所示,则 $a_2=$________,$a=$________.




4　3　**解析**:根据题意知 $a_0=1,a_1=3,a_2=4$.结合二项式定理得 $\begin{cases}C_n^1\cdot\dfrac{1}{a}=3,\\C_n^2\cdot\dfrac{1}{a^2}=4,\end{cases}$ 即 $\begin{cases}n=3a,\\n(n-1)=8a^2,\end{cases}$ 解得 $a=3$.

**14.** 艺术节期间,主办方派甲、乙、丙、丁四名工作人员分别到A,B,C三个不同的演出场馆工作,每个演出场馆至少派一人.若要求甲、乙两人不能到同一演出场馆工作,则不同的分派方案有________种.

30　**解析**:四个人分别到三个不同的演出场馆工作,每个演出场馆至少派一人的方法总数为 $C_4^2A_3^3=36$,甲、乙两人在同一演出场馆工作的方法数为 $A_3^3=6$,故甲、乙两人不在同一演出场馆工作的不同分派方案有 $36-6=30$ 种.

**四、解答题**

**15.** (1)解不等式 $A_6^x<4A_6^{x-2}$;

(2)若 $C_3^2+C_4^2+C_5^2+\cdots+C_n^2=55$,求正整数 $n$.

**解**:(1)由 $A_6^x<4A_6^{x-2}$,可得 $\begin{cases}0\leqslant x\leqslant 6,\\0\leqslant x-2\leqslant 6,\\x\in\mathbf{N},\end{cases}$ 可得 $2\leqslant x\leqslant 6,x\in\mathbf{N}$.

可得 $\dfrac{6!}{(6-x)!}<4\times\dfrac{6!}{(8-x)!}$,所以 $\dfrac{4}{(8-x)(7-x)}>1$,即 $x^2-15x+52<0$.

因为 $2^2-15\times2+52>0,3^2-15\times3+52>0,4^2-15\times4+52>0,5^2-15\times5+52>0,6^2-15\times6+52<0$,所以 $x=6$.

(2)因为 $C_3^2+C_4^2+C_5^2+\cdots+C_n^2=C_3^3+C_3^2+C_4^2+C_5^2+\cdots+C_n^2-1=C_4^3+C_4^2+C_5^2+\cdots+C_n^2-1=C_5^3+C_5^2+\cdots+C_n^2-1=C_{n+1}^3-1=55$,

所以 $C_{n+1}^3=56=C_8^3$,解得 $n=7$.

**16.** 用数字0,1,2,3,4,5组成没有重复数字的自然数,问:

(1)能够组成多少个五位偶数?

(2)能够组成多少个小于2 018的正整数?

(3)可以组成多少个大于3 000且小于5 421的四位数?

**解**:(1)依题意,当0在个位时,组成五位偶数的个数为 $A_5^4=120$;

当2或4在个位时,组成首位不为0的五位偶数的个数为 $2\times C_4^1\times A_4^3=192$.

所以共计组成的五位偶数的个数为 $120+192=312$.

(2)小于2 018的正整数包含:

一位数有5个;

两位数有 $C_5^1\times C_5^1=25$ 个;

三位数有 $C_5^1\times A_5^2=100$ 个;

四位数中千位为1时有 $A_5^3=60$ 个,

千位为2时有2 013,2 014,2 015,3个.

综上,可以组成小于2 018的正整数共有 $5+25+100+60+3=193$ 个.

(3)分4类:

①千位数字为3或4时,后面三个数位上可随便选择,此时共有 $2\times5\times4\times3=120$ 个;

②千位数字为5,百位数字为0,1,2,3之一时,共有 $4\times4\times3=48$ 个;

③千位数字为5,百位数字是4,十位数字为0,1之一时,共有 $2\times3=6$ 个;

④5 420也满足条件.

故所求四位数共有 $120+48+6+1=175$ 个.

**17.** 已知 $(1-2x)^n=a_0+a_1x+a_2x^2+\cdots+a_nx^n(n\in\mathbf{N}^*)$,且 $a_2=60$,求:

(1)$n$ 的值;

(2)$-\dfrac{a_1}{2}+\dfrac{a_2}{2^2}-\dfrac{a_3}{2^3}+\cdots+(-1)^n\dfrac{a_n}{2^n}$ 的值.

**解**:(1)展开式的通项为 $T_{k+1}=C_n^k(-2x)^k=(-2)^kC_n^kx^k$.

因为 $T_3=C_n^2(-2x)^2=a_2x^2$,

所以 $a_2=C_n^2(-2)^2=60$,

化简可得 $n(n-1)=30$,且 $n\in\mathbf{N}^*$,

解得 $n=6$.

(2)因为 $T_{k+1}=C_6^k(-2x)^k=a_kx^k$,

所以 $a_k=C_6^k(-2)^k$,

所以 $(-1)^k\cdot\dfrac{a_k}{2^k}=C_6^k$,

$-\dfrac{a_1}{2}+\dfrac{a_2}{2^2}-\dfrac{a_3}{2^3}+\cdots+(-1)^6\cdot\dfrac{a_6}{2^6}=C_6^1+C_6^2+\cdots+C_6^6=2^6-1=63$.

**18.** 已知在 $\left(\sqrt{x}+\dfrac{1}{2\sqrt[3]{x}}\right)^n$ 的展开式中,前3项的系数分别为 $a_1,a_2,a_3$,且满足 $2a_2=a_1+a_3$.

(1)求展开式中各项的二项式系数的和；

(2)求展开式中系数最大的项；

(3)求展开式中所有有理项.

**解**：(1)由题意知，$\left(\sqrt{x}+\frac{1}{2\sqrt[3]{x}}\right)^n$的展开式的通项为$T_{k+1}=C_n^k(\sqrt{x})^{n-k}\left(\frac{1}{2\sqrt[3]{x}}\right)^k=\frac{1}{2^k}C_n^kx^{\frac{3n-5k}{6}}$，$k=0,1,2,\cdots,n(n\geqslant2)$，

则$a_1=\frac{1}{2^0}C_n^0=1$，$a_2=\frac{1}{2^1}C_n^1=\frac{1}{2}n$，

$a_3=\frac{1}{2^2}C_n^2=\frac{n(n-1)}{8}$.

因为$2a_2=a_1+a_3$，即$2\times\frac{1}{2}n=1+\frac{n(n-1)}{8}$，解得$n=8$或$n=1$(舍去)，

所以$\left(\sqrt{x}+\frac{1}{2\sqrt[3]{x}}\right)^8$的展开式中各项的二项式系数的和为$2^8=256$.

(2)由(1)知$\left(\sqrt{x}+\frac{1}{2\sqrt[3]{x}}\right)^8$的展开式的通项为$T_{k+1}=C_8^k(\sqrt{x})^{8-k}\left(\frac{1}{2\sqrt[3]{x}}\right)^k=\frac{1}{2^k}C_8^kx^{\frac{24-5k}{6}}$($0\leqslant k\leqslant8$且$k\in\mathbf{N}$)，记第$k$项的系数最大，则有$T_k\geqslant T_{k+1}$，且$T_k\geqslant T_{k-1}$，

即$\begin{cases}C_8^{k-1}2^{-k+1}\geqslant C_8^k2^{-k},\\C_8^{k-1}2^{-k+1}\geqslant C_8^{k-2}2^{-k+2},\end{cases}$解得$3\leqslant k\leqslant4$.

又$k\in\mathbf{N}$，所以$k=3$或$k=4$，

所以系数最大的项为第3项$T_3=7x^{\frac{7}{3}}$和第4项$T_4=7x^{\frac{3}{2}}$.

(3)因为$\left(\sqrt{x}+\frac{1}{2\sqrt[3]{x}}\right)^8$的展开式的通项为$T_{k+1}=\frac{1}{2^k}C_8^kx^{\frac{24-5k}{6}}$($0\leqslant k\leqslant8$且$k\in\mathbf{N}$)，

令$\frac{24-5k}{6}\in\mathbf{Z}$，$0\leqslant k\leqslant8$且$k\in\mathbf{N}$，则$k=0$或$k=6$，

所以展开式中有理项为$T_1=x^4$和$T_7=\frac{7}{16x}$.

**19.**设数列$\{a_n\}$是等比数列，$a_1=C_{2m+3}^{3m}\cdot A_{m-2}^1$，公比$q$是$\left(x+\frac{1}{4x^2}\right)^4$的展开式中的第2项，前$n$项和为$S_n$.

(1)用$n,x$表示$a_n$与$S_n$；

(2)若$A_n=C_n^1S_1+C_n^2S_2+\cdots+C_n^nS_n$，用$n,x$表示$A_n$.

**解**：(1)因为$a_1=C_{2m+3}^{3m}\cdot A_{m-2}^1$，

所以$\begin{cases}2m+3\geqslant3m,\\m-2\geqslant1,\end{cases}$即$\begin{cases}m\leqslant3,\\m\geqslant3,\end{cases}$

所以$m=3$，所以$a_1=C_9^9\cdot A_1^1=1$.

又由$\left(x+\frac{1}{4x^2}\right)^4$的展开式的通项为$T_{k+1}=C_4^kx^{x-k}\cdot\left(\frac{1}{4x^2}\right)^k$，知$T_2=C_4^1x^3\cdot\frac{1}{4x^2}=x$，所以$q=x$.

所以$a_n=x^{n-1}(n\geqslant1$，且$n\in\mathbf{N}^*)$，

$S_n=\begin{cases}n,x=1,\\\frac{1-x^n}{1-x},x\neq1.\end{cases}$

(2)当$x=1$时，$S_n=n$，

所以$A_n=0C_n^0+C_n^1+2C_n^2+\cdots+(n-1)C_n^{n-1}+nC_n^n$①.

又$A_n=nC_n^0+(n-1)C_n^1+\cdots+C_n^{n-1}+0C_n^n$②，

由①+②，得$2A_n=n(C_n^0+C_n^1+\cdots+C_n^n)=n\cdot2^n$，

所以$A_n=n\cdot2^{n-1}$.

当$x\neq1$时，$S_n=\frac{1-x^n}{1-x}$，

所以$A_n=\frac{1-x}{1-x}C_n^1+\frac{1-x^2}{1-x}C_n^2+\frac{1-x^3}{1-x}C_n^3+\cdots+\frac{1-x^n}{1-x}C_n^n$

$=\frac{1}{1-x}[(C_n^1+C_n^2+C_n^3+\cdots+C_n^n)-(xC_n^1+x^2C_n^2+x^3C_n^3+\cdots+x^nC_n^n)]$

$=\frac{1}{1-x}[(2^n-1)-(1+xC_n^1+x^2C_n^2+\cdots+x^nC_n^n-1)]$

$=\frac{1}{1-x}[2^n-(1+x)^n]$.

所以$A_n=\begin{cases}n\cdot2^{n-1},x=1,\\\frac{2^n-(1+x)^n}{1-x},x\neq1.\end{cases}$

# 第七章 随机变量及其分布

## 单元概览

### 单元导航

本章以概率论为核心，通过数学工具系统分析随机现象的统计规律.本章通过学习条件概率，建立概率的乘法公式和全概率公式来计算较复杂事件的概率，结合独立性的判定深化对复杂随机事件的理解；从随机变量的基本概念入手，分别以分布列和概率密度函数刻画随机变量的概率分布规律，重点解析二项分布、超几何分布及正态分布；以随机变量的数学期望和方差量化随机变量的平均结果与波动性，最终构建从概率建模到实际分析的完整框架，培养运用分布模型解决预测与决策问题的核心能力.

### 学习目标

**1.**结合古典概型，采用归纳的方法建立条件概率的概念，导出概率的乘法公式和全概率公式，分析解决与概率有关的问题.

**2.**通过具体实例，了解离散型随机变量的概念，理解离散型随机变量分布列及其数字特征，分析解决有关分布列的问题.

**3.**通过具体实例，了解伯努利试验，掌握二项分布及其均值与方差，并能解决简单的实际问题.

**4.**通过具体实例，了解超几何分布及其均值，并能解决简单的实际问题.

**5.**通过误差模型，了解服从正态分布的随机变量.通过具体实例，借助频率分布直方图的几何直观，了解正态分布的特征、均值、方差，分析解决正态分布的 $3\sigma$ 原则在生产生活中的应用.

### 核心概念

随机变量是对随机试验可能结果的量化表示，是样本空间到实数集上的映射.

随机变量的分布列、期望与方差都是对随机现象的刻画，例如用均值刻画随机变量取值的平均水平，用方差刻画随机变量取值相对其均值的离散程度.在学习的过程中，可以进一步强化利用概率模型解决实际问题的能力，发展数学抽象、逻辑推理、数学运算、数学建模的核心素养.

### 学法指导

**1.**在学习条件概率与全概率公式时，从熟悉的问题情境入手，逐步深入到复杂的概率计算问题，在学习的过程中，掌握条件概率与全概率公式的概念及应用，提高逻辑推理素养.

**2.**在学习离散型随机变量的分布列及数字特征时，要结合具体实例，理解分布列的性质，将其应用到实际问题中，并用期望和方差进行决策性的判断，提升数学抽象、数学运算素养.

**3.**在学习二项分布、超几何分布时，通过比较放回和不放回随机抽样中次品数的分布，了解二者的区别与联系，提升数学抽象、数学运算、数学建模素养.

4.在学习正态分布时，借助误差频率直方图描述误差分布，描述样本数据的分布规律，根据频率与概率的关系，加以直观想象，建立正态分布模型，提升数学抽象、数学运算、数学建模素养.

## 单元主题任务

在学校组织的高二篮球比赛中，通过小组循环赛，甲、乙两班顺利进入最后的决赛.在两班的每一场比赛中，甲班取胜的概率为0.6，乙班取胜的概率为0.4.

(1)如果比赛采用三局两胜制，甲班获胜的可能性是多少？

(2)如果比赛采用五局三胜制，甲班获胜的可能性是多少？

(3)如果你是甲班的学生，你认为采用哪种赛制对你班更有利？

# 探究构建

## 7.1　条件概率与全概率公式

### 7.1.1　条件概率

| 学习任务目标 |
| --- |
| 1.结合古典概型，了解条件概率，能计算简单随机事件的条件概率. |
| 2.结合古典概型，了解条件概率与独立性的关系，会利用乘法公式计算概率. |
| 3.能利用条件概率的性质解决简单的实际问题. |

#### 问题式预习

**【知识清单】**

**知识点一　条件概率**

(1)定义：一般地，设 $A$，$B$ 为两个随机事件，且 $P(A)>0$，我们称 $P(B|A)=\dfrac{P(AB)}{P(A)}$ 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.

(2)概率的乘法公式：由条件概率的定义，对任意两个事件 $A$ 与 $B$，若 $P(A)>0$，则 $\underline{P(AB)=P(A)P(B|A)}$.我们称上式为概率的乘法公式.

**知识点二　条件概率的性质**

条件概率只是缩小了样本空间，因此条件概率同样具有概率的性质.设 $P(A)>0$，则

(1)$P(\Omega|A)=\underline{1}$；

(2)如果 $B$ 和 $C$ 是两个互斥事件，则

$P((B\cup C)|A)=\underline{P(B|A)+P(C|A)}$；

(3)设 $\overline{B}$ 和 $B$ 互为对立事件，则

$P(\overline{B}|A)=\underline{1-P(B|A)}$.

**【概念辨析】**

1.判断正误(正确的打“√”，错误的打“×”).

(1)$P(B|A)<P(AB)$.　　( × )

(2)事件 $B$ 在“事件 $A$ 已发生”这个附加条件下的概率与没有这个附加条件的概率一般是不同的.　　( √ )

(3)$0<P(B|A)<1$.　　( × )

(4)若事件 $A$ 等于事件 $B$，则 $P(B|A)=1$.　　( √ )

(5)$P(B|A)$与 $P(A|B)$相同.　　( × )

2.已知 $P(AB)=\dfrac{5}{13}$，$P(A)=\dfrac{5}{7}$，则 $P(B|A)$等于　　(　　)

A.$\dfrac{25}{91}$　　B.$\dfrac{7}{13}$　　C.$\dfrac{10}{13}$　　D.$\dfrac{7}{91}$

B　**解析**：$P(B|A)=\dfrac{P(AB)}{P(A)}=\dfrac{\frac{5}{13}}{\frac{5}{7}}=\dfrac{7}{13}$.

3.在 100 件产品中有 95 件合格产品，5 件不合格产品.现从中不放回地取两次，每次任取一件，则在第一次取到不合格产品后，第二次取到不合格产品的概率为________.

$\dfrac{4}{99}$　**解析**：(方法一)在第一次取到不合格产品后，由于不放回，故还有 99 件产品，其中 4 件不合格产品，故第二次取到不合格产品的概率为$\dfrac{4}{99}$.

(方法二)第一次取到不合格产品的概率 $p_1=\dfrac{5}{100}=\dfrac{1}{20}$，两次都取到不合格产品的概率 $p_2=\dfrac{C_5^2}{C_{100}^2}=$

$\frac{1}{495}$，所以所求概率 $p=\frac{p_2}{p_1}=\frac{\frac{1}{495}}{\frac{1}{20}}=\frac{4}{99}$.

**4.** 请思考并回答下列问题：

(1) $P(B|A)$ 与 $P(AB)$，$P(A)$ 有何区别与联系？

**提示**：$P(B|A)$ 是在事件 $A$ 发生的条件下，事件 $B$ 发生的概率；$P(AB)$ 是事件 $A$ 与 $B$ 同时发生的概率，无附加条件；$P(A)$ 是事件 $A$ 发生的概率，无附加条件.它们的联系是 $P(B|A)=\frac{P(AB)}{P(A)}$.

(2) $P(B|A)$ 与 $P(A|B)$ 表示的意思相同吗？

**提示**：不相同.$P(B|A)$ 表示在事件 $A$ 发生的条件下，事件 $B$ 发生的概率；而 $P(A|B)$ 表示在事件 $B$ 发生的条件下，事件 $A$ 发生的概率.另外，从计算公式上看，$P(B|A)=\frac{P(AB)}{P(A)}$，$P(A|B)=\frac{P(AB)}{P(B)}$.

## ○ 任务型课堂 ○

### 任务1 条件概率的求法

#### 探究活动

**例1** 校运会组委会将甲、乙、丙、丁4名志愿者随机派往铅球、跳远、跳高三个比赛区域，每个区域至少派1名志愿者，每名志愿者只能去一个区域.$A$ 表示事件“志愿者甲被派往铅球区域”；$B$ 表示事件“志愿者乙被派往铅球区域”；$C$ 表示事件“志愿者乙被派往跳远区域”，则（　　）

A.事件 $A$ 与 $B$ 相互独立

B.事件 $A$ 与 $C$ 为互斥事件

C.$P(C|A)=\frac{1}{3}$

D.$P(B|A)=\frac{1}{6}$

D　**解析**：由题意易知人员分组情况为：2，1，1，即所有安排方案有 $C_4^2C_3^1\times2=36$（种），铅球区域可能安排2人或1人，所以 $P(A)=\frac{C_3^1\times2+C_3^2C_2^1}{36}=\frac{1}{3}$，同理 $P(B)=\frac{1}{3}$，$P(C)=\frac{1}{3}$.而 $P(AB)=\frac{2}{36}=\frac{1}{18}\neq P(A)\cdot P(B)$，$P(AC)=\frac{C_2^1C_2^1+1}{36}=\frac{5}{36}$，由相互独立事件的充要条件可知，事件 $A$ 与 $B$ 不相互独立，故A错误；

显然，事件 $A$ 与 $C$ 能同时发生，不为互斥事件，故B错误；

由条件概率公式知 $P(C|A)=\frac{P(AC)}{P(A)}=\frac{\frac{5}{36}}{\frac{1}{3}}=\frac{5}{12}$，故C错误；

$P(B|A)=\frac{P(AB)}{P(A)}=\frac{\frac{1}{18}}{\frac{1}{3}}=\frac{1}{6}$，故D正确.

故选D.

**【探究总结】**

**条件概率的三种求法**

| | |
|---|---|
| 定义法 | 先求 $P(A)$ 和 $P(AB)$，再由 $P(B|A)=\frac{P(AB)}{P(A)}$ 求 $P(B|A)$ |
| 样本点法 | 先求事件 $A$ 包含的样本点个数 $n(A)$，再求事件 $AB$ 包含的样本点个数 $n(AB)$，得 $P(B|A)=\frac{n(AB)}{n(A)}$ |
| 缩样法 | 缩小样本空间，就是去掉第一次抽到的情况，只研究剩下的情况，用古典概型求解，该求法能化繁为简 |

#### 应用迁移

**1.**（2024·天津卷）有 $A$，$B$，$C$，$D$，$E$ 五个活动，甲、乙都要选择三个活动参加.甲选到 $A$ 活动的概率为________；已知乙选了 $A$ 活动，他再选择 $B$ 活动的概率为________.

$\frac{3}{5}$　$\frac{1}{2}$　**解析**：(方法一)从五个活动中选三个的情况有 $ABC$，$ABD$，$ABE$，$ACD$，$ACE$，$ADE$，$BCD$，$BCE$，$BDE$，$CDE$，共10种，

其中甲选到 $A$ 活动的情况有 $ABC$，$ABD$，$ABE$，$ACD$，$ACE$，$ADE$，共6种，

所以甲选到 $A$ 活动的概率为 $\frac{6}{10}=\frac{3}{5}$.

乙选 $A$ 活动的情况有 $ABC,ABD,ABE,ACD,ACE,ADE$，共 6 种，

乙同时选了 $A,B$ 两个活动的情况有 $ABC,ABD,ABE$，共 3 种，

所以乙选了 $A$ 活动，他再选择 $B$ 活动的概率为 $\frac{3}{6}=\frac{1}{2}$.

(方法二)设甲、乙选到 $A$ 活动分别为事件 $M,N$，乙选到 $B$ 活动为事件 $Q$，

则甲选到 $A$ 活动的概率为 $P(M)=\frac{C_4^2}{C_5^3}=\frac{3}{5}$；

乙选了 $A$ 活动，他再选择 $B$ 活动的概率为 $P(Q|N)=\frac{P(NQ)}{P(N)}=\frac{\frac{C_3^1}{C_5^3}}{\frac{C_4^2}{C_5^3}}=\frac{1}{2}$.

2.一个盒子里装有除颜色外完全相同的 9 个球，其中黄球 4 个、蓝球 3 个、绿球 2 个，现从盒子中随机取出两个球，记事件 $A$ 为“取出的两个球颜色不同”，记事件 $B$ 为“取出一个蓝球，一个绿球”，则 $P(B|A)=$________.

$\frac{3}{13}$　**解析**：事件 $A$ 为“取出的两个球颜色不同”，包括一个黄球一个蓝球，一个黄球一个绿球以及一个蓝球一个绿球，共三种情况，则 $n(A)=C_4^1C_3^1+C_4^1C_2^1+C_3^1C_2^1=26$，

事件 $B$ 为“取出一个蓝球，一个绿球”，则 $n(AB)=C_3^1C_2^1=6$，

所以 $P(B|A)=\frac{n(AB)}{n(A)}=\frac{6}{26}=\frac{3}{13}$.

## 任务2 利用乘法公式求概率

### 探究活动

**例 2**　一袋中有 10 个球，其中 3 个黑球、7 个白球，先后两次从中任意取一球(不放回).设 $A_i$ 表示事件“第 $i$ 次取到的是黑球”($i=1,2$)，求两次取到的均为黑球的概率.

**解**：由题设知 $P(A_1)=\frac{3}{10}$，$P(A_2|A_1)=\frac{2}{9}$.

根据概率的乘法公式，有 $P(A_1A_2)=P(A_1)P(A_2|A_1)=\frac{3}{10}\times\frac{2}{9}=\frac{1}{15}$.

[一题多思]

**思考 1.** 求第一次取到黑球，第二次取到白球的概率.

**解**：由题意知 $P(A_1)=\frac{3}{10}$，$P(\overline{A_2}|A_1)=\frac{7}{9}$.

根据概率的乘法公式，有 $P(A_1\overline{A_2})=P(A_1)P(\overline{A_2}|A_1)=\frac{3}{10}\times\frac{7}{9}=\frac{7}{30}$.

**思考 2.** 将“不放回”改为“有放回”，重新研究“例 2”与“思考 1”.

**解**：由题意知 $P(A_1)=\frac{3}{10}$，$P(A_2)=\frac{3}{10}$，$P(\overline{A_2})=\frac{7}{10}$，则两次取到的均是黑球的概率为 $P(A_1A_2)=\frac{3}{10}\times\frac{3}{10}=\frac{9}{100}$，第一次取到黑球，第二次取到白球的概率为 $P(A_1\overline{A_2})=\frac{3}{10}\times\frac{7}{10}=\frac{21}{100}$.

**【探究总结】**

**非相互独立事件同时发生的概率的求法**

若 $A,B$ 不是相互独立事件，则 $P(AB)$ 可用乘法公式 $P(AB)=P(A)\cdot P(B|A)$ 来求解.$P(B|A)$ 可采用样本点法来计算，$P(B|A)=\frac{n(AB)}{n(A)}$，其中 $n(AB)$ 表示事件 $AB$ 包含的样本点个数，$n(A)$ 表示事件 $A$ 包含的样本点个数.

### 应用迁移

1.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(白球与红球大小、质地相同)，现从 1 号箱中随机取出一球放入 2 号箱，再从 2 号箱中随机取出一球，则两次都取到红球的概率是　　(　　)

A.$\frac{11}{27}$　　B.$\frac{11}{24}$　　C.$\frac{8}{27}$　　D.$\frac{3}{8}$

C　**解析**：设“从 1 号箱中取到红球放入 2 号箱”为事件 $A$，“从 2 号箱中取到红球”为事件 $B$.

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

所以 $P(AB)=P(A)P(B|A)=\frac{2}{3}\times\frac{4}{9}=\frac{8}{27}$，

所以两次都取到红球的概率为 $\frac{8}{27}$.

2.已知某品牌的手机从 1 m 高的地方掉落时，屏幕第一次未碎掉的概率为 0.5，当第一次未碎掉时第二次也未碎掉的概率为 0.3，试求这样的手机从 1 m 高的地方掉落两次后屏幕仍未碎掉的概率.

**解**：设 $A_i$＝“第 $i$ 次掉落手机屏幕未碎掉”，$i=1,2$，则由已知可得 $P(A_1)=0.5$，$P(A_2|A_1)=0.3$.

由概率的乘法公式可得 $P(A_1A_2)=P(A_1)P(A_2|A_1)=0.5\times0.3=0.15$，

即这样的手机从 1 m 高的地方掉落两次后屏幕仍未碎掉的概率为 0.15.

## 任务3 利用条件概率的性质求互斥事件的条件概率

### 探究活动

**例 3** 在某次考试中，要从 20 道题中随机抽出 6 道题，若考生能答对其中至少 4 道题即可通过，能答对其中至少 5 道题就获得优秀.已知某考生能答对其中 10 道题，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 1 道题答错”，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”.试完成下列问题：

(1)判断事件 $A,B,C$ 之间的关系；

(2)求该考生在这次考试中通过的概率；

(3)已知该考生通过这次考试，求该考生在这次考试中获得优秀的概率.

**解**：(1)$A,B,C$ 之间两两互斥.

(2)因为 $A,B,C$ 两两互斥，且 $D=A\cup B\cup C$，

所以 $P(D)=P(A\cup B\cup C)=P(A)+P(B)+P(C)=\dfrac{C_{10}^6}{C_{20}^6}+\dfrac{C_{10}^5C_{10}^1}{C_{20}^6}+\dfrac{C_{10}^4C_{10}^2}{C_{20}^6}=\dfrac{203}{646}$.

(3)$E=A\cup B$，因为 $P(AD)=P(A)$，$P(BD)=P(B)$，

所以 $P(E|D)=P(A|D)+P(B|D)=\dfrac{P(A)}{P(D)}+\dfrac{P(B)}{P(D)}=\dfrac{\dfrac{C_{10}^6}{C_{20}^6}}{\dfrac{12\ 180}{C_{20}^6}}+\dfrac{\dfrac{C_{10}^5C_{10}^1}{C_{20}^6}}{\dfrac{12\ 180}{C_{20}^6}}=\dfrac{13}{58}$.

故该考生在这次考试中获得优秀的概率为 $\dfrac{13}{58}$.

**【探究总结】**

当求一个较复杂事件的概率时，往往把该事件分成两个(或多个)互斥的较简单的事件之和，求出这些较简单事件的概率，再利用 $P(B\cup C|A)=P(B|A)+P(C|A)$ 便可求得原事件的概率，但应注意这个公式在“$B$ 与 $C$ 互斥”这一前提下才成立.

### 应用迁移

1.已知事件 $A,B,C$ 满足 $A,B$ 是互斥事件，且 $P(A\cup B|C)=\dfrac{1}{2}$，$P(BC)=\dfrac{1}{12}$，$P(C)=\dfrac{1}{4}$，则 $P(A|C)=$ (　　)

A.$\dfrac{1}{6}$　　B.$\dfrac{1}{12}$　　C.$\dfrac{1}{4}$　　D.$\dfrac{1}{3}$

A　**解析**：由题意可得 $P(B|C)=\dfrac{P(BC)}{P(C)}=\dfrac{1}{3}$，由 $A,B$ 是互斥事件知，$P(A\cup B|C)=P(A|C)+P(B|C)$，所以 $P(A|C)=P(A\cup B|C)-P(B|C)=\dfrac{1}{2}-\dfrac{1}{3}=\dfrac{1}{6}$.

2.将大小相同的球分装在三个盒子中，每盒 10 个.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，那么称试验成功.求试验成功的概率.

**解**：设 $A$＝“从第一个盒子中取得标有字母 A 的球”，

$B$＝“从第一个盒子中取得标有字母 B 的球”，

$R$＝“第二次取出的球是红球”，

$W$＝“第二次取出的球是白球”，

则 $P(A)=\dfrac{7}{10}$，$P(B)=\dfrac{3}{10}$，

所以 $P(R|A)=\dfrac{1}{2}$，$P(R|B)=\dfrac{4}{5}$.

事件“试验成功”表示为 $RA\cup RB$，

又事件 $RA$ 与事件 $RB$ 互斥，

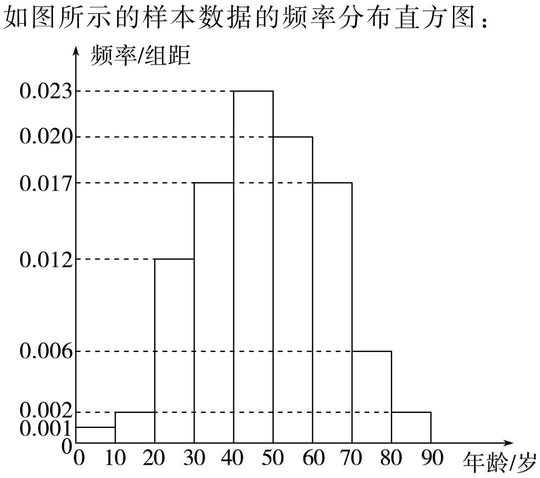
所以由概率加法公式与乘法公式得

$P(RA\cup RB)=P(RA)+P(RB)$

$=P(R|A)P(A)+P(R|B)P(B)$

$=\dfrac{1}{2}\times\dfrac{7}{10}+\dfrac{4}{5}\times\dfrac{3}{10}=\dfrac{59}{100}$.

3.(2022·新高考全国Ⅱ卷)在某地区进行流行病学调查，随机调查了 100 位某种疾病患者的年龄，得到如图所示的样本数据的频率分布直方图：



(1)估计该地区这种疾病患者的平均年龄(同一组中的数据用该组区间的中点值代表);

(2)估计该地区一位这种疾病患者的年龄位于区间[20,70)的概率;

(3)已知该地区这种疾病的患病率为0.1%,该地区年龄位于区间[40,50)的人口占该地区总人口的16%.从该地区中任选一人,若此人的年龄位于区间[40,50),求此人患这种疾病的概率.(以样本数据中患者的年龄位于各区间的频率作为患者的年龄位于该区间的概率,精确到0.000 1)

**解**:(1)平均年龄为$(5\times0.001+15\times0.002+25\times0.012+35\times0.017+45\times0.023+55\times0.020+65\times0.017+75\times0.006+85\times0.002)\times10=47.9$(岁).

(2)设$A$="一位这种疾病患者的年龄在区间[20,70)",所以$P(A)=1-P(\overline{A})=1-(0.001+0.002+0.006+0.002)\times10=1-0.11=0.89$.

(3)设$B$="任选一人年龄位于区间[40,50)",$C$="从该地区中任选一人患这种疾病".

由已知得$P(B)=16\%=0.16$,$P(C)=0.1\%=0.001$,$P(B|C)=0.023\times10=0.23$.

由条件概率公式,可得从该地区中任选一人,若此人的年龄位于区间[40,50),此人患这种疾病的概率为

$P(C|B)=\frac{P(BC)}{P(B)}=\frac{P(C)P(B|C)}{P(B)}=\frac{0.001\times0.23}{0.16}=0.001\ 437\ 5\approx0.001\ 4$.

### 提质归纳

条件概率
- 条件概率的定义:$P(B|A)=\frac{P(AB)}{P(A)}$
- 条件概率的性质
  - (1)$P(\Omega|A)=1$;
  - (2)如果$B$和$C$是两个互斥事件,则$P(B\cup C|A)=P(B|A)+P(C|A)$;
  - (3)设$\overline{B}$和$B$互为对立事件,则$P(\overline{B}|A)=1-P(B|A)$
- 概率的乘法公式:$P(AB)=P(A)P(B|A)$

## 课后素养评价(九)　条件概率

### A组　学习·理解

**1.**下列概率是条件概率的是　(　　)

A.甲、乙二人投篮命中率分别为0.6,0.7,甲、乙各投篮一次都投中的概率

B.甲、乙二人投篮命中率分别为0.6,0.7,在甲投中的条件下乙投篮一次命中的概率

C.有10件产品,其中4件次品,抽2件产品进行检验,恰好抽到1件次品的概率

D.小明上学路上要过四个路口,每个路口遇到红灯的概率都是$\frac{2}{5}$,小明在一次上学途中遇到红灯的概率

B　**解析**:条件概率为在某一事件发生的条件下,另一事件发生的概率.

对于A,甲、乙各投篮一次都投中的概率,不是条件概率;

对于B,在甲投中的条件下乙投篮一次命中的概率,是条件概率;

对于C,抽2件产品恰好抽到1件次品的概率,不是条件概率;

对于D,一次上学途中遇到红灯的概率,不是条件概率.

故选B.

**2.**下列说法正确的是　(　　)

A.事件$A$,$B$中至少有一个发生的概率一定比$A$,$B$中恰有一个发生的概率大

B.若$P(A\cup B)=P(A)+P(B)=1$,则事件$A$,$B$为对立事件

C.若$A$,$B$为互斥事件,则$P(A)+P(B)\leqslant1$

D.若事件$A$,$B$,$C$满足条件$P(B)>0$,$A$和$C$为互斥事件,则$P(A\cup C|B)<P(A|B)+P(C|B)$

C　**解析**:对于A,若事件$A$和$B$都为不可能事件,此时两个概率相等,所以A错误;

对于B,若在不同试验下,虽然有$P(A\cup B)=P(A)+P(B)=1$,但事件$A$与$B$不对立,若在同一试验下,则事件$A$与$B$对立,所以B错误;

对于C,若$A$,$B$互斥,且$A$,$B$对立,则$P(A)+P(B)=1$,

若$A$,$B$互斥,且$A$,$B$不对立,则$P(A)+P(B)<1$,所以C正确;

对于D,若事件$A$,$B$,$C$满足条件$P(B)>0$,$A$和$C$为互斥事件,

则$P(A\cup C|B)=P(A|B)+P(C|B)$,所以D错误.

故选C.

**3.**袋中有除颜色外完全相同的4个黑球和2个白球.如果不放回地随机取出2个球,那么在第一次取到

的是黑球的条件下，第二次取到黑球的概率为 （　　）

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

C　**解析**：设事件 $A$ 表示"第一次取出黑球"，事件 $B$ 表示"第二次取出黑球"，

$P(A)=\frac{4}{6}=\frac{2}{3}$，$P(AB)=\frac{4\times3}{6\times5}=\frac{2}{5}$，

所以在第一次取到的是黑球的条件下，第二次取到黑球的概率为 $P(B|A)=\frac{P(AB)}{P(A)}=\frac{3}{5}$.

**4.** 随着社会的发展，越来越多的共享资源陆续出现，它们不可避免地与我们每个人产生密切的关联，逐渐改变着每个人的生活方式.已知某种型号的共享充电宝循环充电超过 500 次的概率为 $\frac{3}{4}$，超过 1 000 次的概率为 $\frac{1}{2}$，现有一个该型号的充电宝已经循环充电超过 500 次，则其能够循环充电超过 1 000 次的概率是 （　　）

A. $\frac{3}{4}$　　B. $\frac{2}{3}$　　C. $\frac{1}{2}$　　D. $\frac{1}{3}$

B　**解析**：记事件 $A$ 为"该充电宝循环充电超过 500 次"，则 $P(A)=\frac{3}{4}$，记事件 $B$ 为"该充电宝循环充电超过 1 000 次"，则 $P(B)=\frac{1}{2}$. 易知 $P(AB)=P(B)=\frac{1}{2}$，所以 $P(B|A)=\frac{P(AB)}{P(A)}=\frac{\frac{1}{2}}{\frac{3}{4}}=\frac{1}{2}\times\frac{4}{3}=\frac{2}{3}$. 故选 B.

**5.** 若 $B,C$ 是互斥事件且 $P(B|A)=\frac{1}{3}$，$P(C|A)=\frac{1}{4}$，则 $P(B\cup C|A)=$ （　　）

A. $\frac{1}{2}$　　B. $\frac{1}{3}$　　C. $\frac{3}{10}$　　D. $\frac{7}{12}$

D　**解析**：因为 $B,C$ 是互斥事件，

所以 $P(B\cup C|A)=P(B|A)+P(C|A)=\frac{1}{3}+\frac{1}{4}=\frac{7}{12}$.

**6.** 数学上把相差为 2 的两个素数叫做"孪生素数".在不超过 30 的自然数中，随机选取两个不同的数，记事件 $A$="这两个数都是素数"，事件 $B$="这两个数不是孪生素数"，则 $P(B|A)=$ （　　）

A. $\frac{11}{15}$　　B. $\frac{37}{45}$　　C. $\frac{41}{45}$　　D. $\frac{43}{45}$

C　**解析**：不超过 30 的自然数有 31 个，其中素数有 2，3，5，7，11，13，17，19，23，29，共 10 个，孪生素数有 3 和 5，5 和 7，11 和 13，17 和 29，共 4 组.

所以 $P(A)=\frac{C_{10}^2}{C_{31}^2}=\frac{45}{465}$，$P(AB)=\frac{C_{10}^2-4}{C_{31}^2}=\frac{41}{465}$，

所以 $P(B|A)=\frac{P(AB)}{P(A)}=\frac{\frac{41}{465}}{\frac{45}{465}}=\frac{41}{45}$.

故选 C.

**7.** 已知甲在上班途中要经过两个路口，在第一个路口遇到红灯的概率为 0.5，两个路口连续遇到红灯的概率为 0.4，则甲在第一个路口遇到红灯的条件下，第二个路口遇到红灯的概率为 （　　）

A. 0.6　　B. 0.7　　C. 0.8　　D. 0.9

C　**解析**：设"第一个路口遇到红灯"为事件 $A$，"第二个路口遇到红灯"为事件 $B$，则 $P(A)=0.5$，$P(AB)=0.4$，所以 $P(B|A)=\frac{P(AB)}{P(A)}=0.8$.

**8.** 数学家高斯在著作《算术研究》中首次引入了二次剩余的概念.二次剩余理论在噪声工程学、密码学及大数分解等领域都有广泛的应用.已知对于正整数 $a,n(n\geqslant2)$，若存在一个整数 $x$，使得 $n$ 整除 $x^2-a$，则称 $a$ 是 $n$ 的一个二次剩余，否则为二次非剩余.从 1 到 20 这 20 个整数中随机抽取 1 个整数 $a$，记事件 $A$="$a$ 与 12 互质"，$B$="$a$ 是 12 的二次非剩余"，则 $P(A)=$________，$P(B|A)=$________.

$\frac{7}{20}$　$\frac{5}{7}$　**解析**：在 1 到 20 内与 12 互质的整数有 1，5，7，11，13，17，19，所以 $P(A)=\frac{7}{20}$.

根据定义，若对于 $\frac{x^2-a}{12}=m(m\in\mathbf{Z})$ 的 $x$ 不存在，则 $a$ 是 12 的二次非剩余，

显然，当 $a=1$ 时，$x$ 可取 11；当 $a=13$ 时，$x$ 可取 7；当 $a=5,7,11,17,19$ 时，$x$ 不存在.

所以 $P(AB)=\frac{1}{4}$.

所以 $P(B|A)=\frac{P(AB)}{P(A)}=\frac{5}{7}$.

## B组　应用·实践

**1.**（多选）已知事件 $A,B$ 满足 $P(A)=0.4$，$P(B)=0.3$，则下列结论正确的是 （　　）

A. 若 $B\subseteq A$，则 $P(AB)=0.4$

B. 若 $A$ 与 $B$ 互斥，则 $P(A+B)=0.7$

C. 若 $A$ 与 $B$ 相互独立，则 $P(AB)=0.4$

D. 若 $P(B|A)=0.3$，则 $A$ 与 $B$ 相互独立

BD　**解析**：对于 A，因为 $P(A)=0.4$，$P(B)=0.3$，

$B\subseteq A$，所以 $P(AB)=P(B)=0.3$，故 A 错误；
对于 B，因为 $A$ 与 $B$ 互斥，所以 $P(A+B)=P(A)+P(B)=0.7$，故 B 正确；
对于 C，因为 $A$ 与 $B$ 相互独立，所以 $P(AB)=0.4\times0.3=0.12$，故 C 错误；
对于 D，因为 $P(B|A)=0.3$，即 $\dfrac{P(AB)}{P(A)}=0.3$，所以 $P(AB)=0.3\times0.4=0.12$，
又因为 $P(A)P(B)=0.12$，所以 $P(AB)=P(A)\cdot P(B)$，所以 $A$ 与 $B$ 相互独立，故 D 正确.
故选 BD.

**2.**甲、乙、丙、丁 4 名同学报名参加假期社区服务活动，社区服务活动共有关怀老人、环境监测、教育咨询、交通宣传四个项目，每人限报其中一项.记事件 $A$ 为“4 名同学所报项目各不相同”，事件 $B$ 为“只有甲同学一人报关怀老人项目”，则 $P(A|B)=$ （　　）

A.$\dfrac{1}{4}$　　B.$\dfrac{3}{4}$　　C.$\dfrac{2}{9}$　　D.$\dfrac{5}{9}$

C　**解析**：由已知有 $P(B)=\dfrac{3^3}{4^4}=\dfrac{27}{256}$，$P(AB)=\dfrac{A_3^3}{4^4}=\dfrac{3}{128}$，所以 $P(A|B)=\dfrac{P(AB)}{P(B)}=\dfrac{2}{9}$.

**3.**将两枚骰子各掷一次，设事件 $A=$“两枚骰子出现的点数不相同”，$B=$“至少出现一个 6 点”，则 $P(A|B)=$ （　　）

A.$\dfrac{10}{11}$　　B.$\dfrac{5}{11}$　　C.$\dfrac{5}{18}$　　D.$\dfrac{5}{36}$

A　**解析**：给两枚骰子编号：1 号与 2 号，至少出现一个 6 点的情况分三类：1 号是 6 点，2 号不是 6 点，有 5 种；2 号是 6 点，1 号不是 6 点，有 5 种；1 号是 6 点，2 号也是 6 点，有 1 种.故事件 $B$ 包含的样本点数为 11，即 $n(B)=11$，事件 $AB$ 包含的样本点数为 10，即 $n(AB)=10$，所以 $P(A|B)=\dfrac{n(AB)}{n(B)}=\dfrac{10}{11}$.

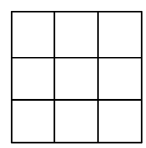
**4.**加工某种零件需要两道工序，第一道工序出废品的概率为 0.4，两道工序都出废品的概率为 0.2，则在第一道工序出废品的条件下，第二道工序又出废品的概率为________.

0.5　**解析**：设“第一道工序出废品”为事件 $A$，则 $P(A)=0.4$，“第二道工序出废品”为事件 $B$.根据题意可得 $P(AB)=0.2$，故在第一道工序出废品的条件下，第二道工序又出废品的概率 $P(B|A)=\dfrac{P(AB)}{P(A)}=0.5$.

**5.**某校组织甲、乙、丙等 6 名学生参加演讲比赛，采用抽签法决定演讲顺序，在“学生甲和乙都不是第一个出场，且甲不是最后一个出场”的前提下，学生丙第一个出场的概率为________.

$\dfrac{1}{4}$　**解析**：设事件 $A$：“学生甲和乙都不是第一个出场，且甲不是最后一个出场”，事件 $B$：“学生丙第一个出场”.对事件 $A$，甲和乙都不是第 1 个出场，第一类：乙在最后，则优先从中间 4 个位置中选 1 个给甲，再将余下的 4 个人全排列，有 $C_4^1\cdot A_4^4$ 种排法；第二类：乙没有在最后，则优先从中间 4 个位置中选 2 个给甲、乙，再将余下的 4 个人全排列，有 $A_4^2\cdot A_4^4$ 种排法.故总的样本点数 $n(A)=C_4^1\cdot A_4^4+A_4^2\cdot A_4^4$.对事件 $AB$，此时丙第一个出场，优先从除了甲以外的 4 人中选 1 人安排在最后，再将余下的 4 人全排列，有 $C_4^1\cdot A_4^4$ 种排法，则 $n(AB)=C_4^1\cdot A_4^4$，所以 $P(B|A)=\dfrac{n(AB)}{n(A)}=\dfrac{C_4^1\cdot A_4^4}{C_4^1\cdot A_4^4+A_4^2\cdot A_4^4}=\dfrac{1}{4}$.

**6.**如图，一个大正方形被平均分成 9 个部分，向大正方形区域随机地投掷一个点（每次都能投中）.将投中最左侧 3 个小正方形区域记为事件 $A$，投中最上面 3 个小正方形或正中间的 1 个小正方形区域记为事件 $B$，求 $P(AB)$，$P(A|B)$.



**解**：由题意可知 $n(\Omega)=9$，$n(A)=3$，$n(B)=4$，$n(AB)=1$，所以 $P(AB)=\dfrac{1}{9}$，$P(A|B)=\dfrac{n(AB)}{n(B)}=\dfrac{1}{4}$.

**7.**一袋中共有 10 个除颜色外完全相同的黑球和白球.若从袋中任意摸出 2 个球，至少有 1 个白球的概率为 $\dfrac{7}{9}$.

(1)求白球的个数；

(2)现从中不放回地摸球，每次随机摸出 1 个球，摸两次，已知第一次摸出白球，求第二次摸出黑球的概率.

**解**：(1)记“从袋中任意摸出 2 个球，至少有 1 个白球”为事件 $A$，记袋中白球的个数为 $x$.则 $P(A)=1-\dfrac{C_{10-x}^2}{C_{10}^2}=\dfrac{7}{9}$，解得 $x=5$，即白球的个数为 5.

(2)记“第一次摸出白球”为事件 $B$，“第二次摸出黑球”为事件 $C$，则 $P(BC)=\dfrac{5\times5}{10\times9}=\dfrac{25}{90}=\dfrac{5}{18}$，$P(B)=\dfrac{5}{10}=\dfrac{1}{2}$.

故 $P(C|B)=\dfrac{P(BC)}{P(B)}=\dfrac{\dfrac{5}{18}}{\dfrac{1}{2}}=\dfrac{5}{9}$.

# 7.1.2 全概率公式

**学习任务目标**

1.结合古典概型，了解利用概率的加法公式与乘法公式推导出全概率公式的过程，为解决一类概率问题奠定基础.

2.理解全概率公式，并能利用全概率公式进行相关的概率计算.

3.了解贝叶斯公式，并能利用贝叶斯公式进行简单的计算.

## 问题式预习

**【知识清单】**

**知识点一 全概率公式**

一般地，设 $A_1,A_2,\cdots,A_n$ 是一组两两互斥的事件，$A_1\cup A_2\cup\cdots\cup A_n=\Omega$，且 $\underline{P(A_i)>0}$，$i=1,2,\cdots,n$，则对任意的事件 $B\subseteq\Omega$，有 $P(B)=\underline{\sum\limits_{i=1}^{n}P(A_i)P(B|A_i)}$.

我们称上面的公式为全概率公式.全概率公式是概率论中最基本的公式之一.

**知识点二 贝叶斯公式**

设 $A_1,A_2,\cdots,A_n$ 是一组两两互斥的事件，$A_1\cup A_2\cup\cdots\cup A_n=\Omega$，且 $\underline{P(A_i)>0}$，$i=1,2,\cdots,n$，则对任意的事件 $B\subseteq\Omega$，$P(B)>0$，有 $P(A_i|B)=\underline{\dfrac{P(A_i)P(B|A_i)}{P(B)}=\dfrac{P(A_i)P(B|A_i)}{\sum\limits_{k=1}^{n}P(A_k)P(B|A_k)}}$，$i=1,2,\cdots,n$.

**【概念辨析】**

1.判断正误(正确的打“√”，错误的打“×”).

(1)若 $P(A)>0$，$P(\overline{A})>0$，则 $P(B)=P(A)P(B|A)+P(\overline{A})P(B|\overline{A})$. ( √ )

(2)若事件 $A_1,A_2,A_3$ 两两互斥且 $P(A_1)>0$，$P(A_2)>0$，$P(A_3)>0$，则 $P(B)=\sum\limits_{i=1}^{3}P(A_i)P(B|A_i)$.( × )

2.设袋中共有10个除颜色外其余均相同的球，其中2个红球，其余为白球.两人分别从袋中任取一球(第一个人取出的球不放回)，则第二个人取得红球的概率为________.

$\dfrac{1}{5}$ **解析**：设 $A$=“第二个人取得红球”，$B$=“第一个人取得红球”，则 $P(B)=\dfrac{2}{10}$，$P(\overline{B})=\dfrac{8}{10}$，$P(A|B)=\dfrac{1}{9}$，$P(A|\overline{B})=\dfrac{2}{9}$，所以 $P(A)=P(B)P(A|B)+P(\overline{B})P(A|\overline{B})=\dfrac{2}{10}\times\dfrac{1}{9}+\dfrac{8}{10}\times\dfrac{2}{9}=\dfrac{1}{5}$.

3.设某公路上经过的货车与客车的数量之比为2∶1，货车中途停车修理的概率为0.02，客车中途停车修理的概率为0.01.现有一辆汽车中途停车修理，则该汽车是货车的概率为________.

0.8 **解析**：设 $B$=“中途停车修理”，$A_1$=“经过的是货车”，$A_2$=“经过的是客车”，则 $B=A_1B\cup A_2B$.

由贝叶斯公式有

$$P(A_1|B)=\frac{P(A_1)P(B|A_1)}{P(A_1)P(B|A_1)+P(A_2)P(B|A_2)}=\frac{\frac{2}{3}\times0.02}{\frac{2}{3}\times0.02+\frac{1}{3}\times0.01}=0.8.$$

4.请思考并回答下列问题：

(1)你能否用自己的语言解释全概率公式？

**提示**：事件 $A$ 发生的概率等于所有可能的状态 $B_i$ 和事件 $A$ 同时发生的概率之和，其中每个状态 $B_i$ 和事件 $A$ 同时发生的概率由该状态下 $A$ 发生的概率和状态 $B_i$ 发生的概率决定.

(2)在使用全概率公式时，需要注意什么？

**提示**：在使用全概率公式时，需要注意以下两点：

①必须确定所有可能的状态 $B_i$，并给出它们对应的概率 $P(B_i)$；

②必须确定每个状态 $B_i$ 下事件 $A$ 发生的概率 $P(A|B_i)$.

## ◦ 任务型课堂 ◦

### 任务1 利用全概率公式求概率

#### 探究活动

**例 1**　甲盒中有 2 个白球、5 个红球；乙盒中有 3 个白球、4 个红球.用 $A$ 表示"从乙盒中取到白球"，用 $B$ 表示"从甲盒中取到白球".

(1)从甲盒中任取一个球放入乙盒，再从乙盒中任取一个球，求取到白球的概率；

(2)从乙盒中任取一个球放入甲盒，再从甲盒中任取一个球，求取到红球的概率.

**解**：(1)$A$ 发生当且仅当 $AB\cup A\overline{B}$ 发生，即 $A=AB\cup A\overline{B}=AB+A\overline{B}$，则 $P(A)=P(AB)+P(A\overline{B})=P(B)P(A|B)+P(\overline{B})P(A|\overline{B})=\frac{2}{7}\times\frac{4}{8}+\frac{5}{7}\times\frac{3}{8}=\frac{23}{56}$.

(2)$\overline{B}$ 发生当且仅当 $\overline{B}A\cup\overline{B}\ \overline{A}$ 发生，即 $\overline{B}=\overline{B}A+\overline{B}\ \overline{A}$，

则 $P(\overline{B})=P(\overline{B}A)+P(\overline{B}\ \overline{A})=P(A)P(\overline{B}|A)+P(\overline{A})\cdot P(\overline{B}|\overline{A})=\frac{3}{7}\times\frac{5}{8}+\frac{4}{7}\times\frac{6}{8}=\frac{39}{56}$.

**例 2**　有一批同一型号的产品，已知其中一厂生产的占 30%，二厂生产的占 50%，三厂生产的占 20%.又知这三个厂的产品的次品率分别为 2%，1%，1%.求从这批产品中任取的一件是次品的概率.

**解**：设事件 $A$="任取的一件为次品"，事件 $B_i$="任取的一件为 $i$ 厂的产品"，$i=1,2,3$.

由全概率公式得 $P(A)=P(A|B_1)P(B_1)+P(A|B_2)\cdot P(B_2)+P(A|B_3)P(B_3)$.

又 $P(B_1)=0.3$，$P(B_2)=0.5$，$P(B_3)=0.2$，$P(A|B_1)=0.02$，$P(A|B_2)=0.01$，$P(A|B_3)=0.01$，

所以 $P(A)=0.02\times0.3+0.01\times0.5+0.01\times0.2=0.013$.

**【探究总结】**

**1.**全概率公式的主要用处在于它可以将一个复杂事件的概率计算问题分解为若干个简单事件的概率计算问题，最后应用概率的加法公式求出最终结果.

**2.**利用全概率公式求概率的步骤：

(1)设事件：把事件 $B$(结果事件)看作某一过程的结果，把 $A_1,A_2,\cdots,A_n$ 看作导致结果的若干个原因；

(2)写概率：由已知，写出每一个原因发生的概率，即 $P(A_i)(i=1,2,\cdots,n)$及每一个原因对结果的影响程度，即 $P(B|A_i)$；

(3)代公式：用全概率公式计算结果发生的概率，即 $P(B)=\sum_{i=1}^{n}P(A_i)P(B|A_i)$.

#### 应用迁移

**1.**已知 $P(B)=\frac{3}{10}$，$P(B|A)=\frac{9}{10}$，$P(B|\overline{A})=\frac{1}{5}$，则 $P(A)=$________.

$\frac{1}{7}$　**解析**：由 $P(B)=P(A)P(B|A)+P(\overline{A})P(B|\overline{A})$，得 $\frac{3}{10}=P(A)\times\frac{9}{10}+[1-P(A)]\times\frac{1}{5}$，解得 $P(A)=\frac{1}{7}$.

**2.**某训练小组有 20 名射手，其中一、二、三级射手分别有 6 名、9 名、5 名.已知一、二、三级射手在比赛中击中目标的概率分别为 0.9，0.8，0.6.现从该小组随机选一人参加比赛，则在比赛中击中目标的概率为________.

0.78　**解析**：由题意知一、二、三级射手占比分别为 $\frac{3}{10}$，$\frac{9}{20}$，$\frac{1}{4}$，且一、二、三级射手在比赛中击中目标的概率分别为 0.9，0.8，0.6，所以从该小组随机选一人参加比赛，在比赛中击中目标的概率为 $\frac{3}{10}\times0.9+\frac{9}{20}\times0.8+\frac{1}{4}\times0.6=0.78$.

**3.**某芯片制造企业采用流水线的方式生产芯片.原有生产线生产某型号的芯片需要经过三道工序，这三道工序互不影响.已知三道工序产生不合格产品的概率分别为 $\frac{1}{50}$，$\frac{1}{49}$，$\frac{1}{48}$，三道工序均合格的产品成为合格品，否则成为次品.

(1)求该企业原有生产线的次品率；

(2)为了提高产量，该企业又引进一条新生产线加工同一型号的芯片，将两条生产线生产出的芯片随机混放在一起.已知新生产线的次品率为 $\frac{1}{25}$，且新生产线的产量是原生产线产量的两倍.从混放的芯片中任取一个，计算它是次品的概率.

**解**：(1)该企业原有生产线的正品率为 $P_1=\Big(1-$

$\left.\frac{1}{50}\right)\times\left(1-\frac{1}{49}\right)\times\left(1-\frac{1}{48}\right)=\frac{47}{50}$,所以该企业原有生产线的次品率为 $P=1-P_1=1-\frac{47}{50}=\frac{3}{50}$.

(2)记"任取一个芯片来自原生产线"为事件 $A$,"任取一个芯片来自新生产线"为事件 $\overline{A}$,

记"任取一个芯片是次品"为事件 $B$,

则 $P(A)=\frac{1}{3}$,$P(\overline{A})=\frac{2}{3}$,且 $P(B|A)=\frac{3}{50}$,$P(B|\overline{A})=\frac{1}{25}$,

所以 $P(B)=P(A)P(B|A)+P(\overline{A})P(B|\overline{A})=\frac{1}{3}\times\frac{3}{50}+\frac{2}{3}\times\frac{1}{25}=\frac{7}{150}$,即从混放的芯片中任取一个,它是次品的概率为 $\frac{7}{150}$.

## 任务2 贝叶斯公式的应用

### 探究活动

**例 3** 某商业银行对在校贫困大学生提供助学贷款,某贷款的大学生承诺毕业三年内还清助学贷款,若未还清,则视该生不遵守承诺.假设贷款学生中可信的学生占比为80%,可信的学生不遵守承诺的概率为0.1,不可信的学生不遵守承诺的概率为0.95.用 $A$ 表示事件"该生不遵守承诺",$B$ 表示事件"该生可信".

(1)若该生在毕业三年内未还清贷款,则该生是可信的学生的概率是多少?

(2)若该生在毕业三年内还清贷款,则该生是可信的学生的概率是多少?(注:结果均精确到0.01)

**解**:(1)由题意知,$P(B)=0.8$,$P(\overline{B})=0.2$,

$P(A|B)=0.1$,$P(A|\overline{B})=0.95$.

由贝叶斯公式,得该生在毕业三年内未还清贷款,该生是可信的学生的概率为

$$P(B|A)=\frac{P(B)P(A|B)}{P(B)P(A|B)+P(\overline{B})P(A|\overline{B})}$$

$$=\frac{0.8\times0.1}{0.8\times0.1+0.2\times0.95}\approx0.30.$$

(2)由题意知,$P(B)=0.8$,$P(\overline{B})=0.2$,$P(\overline{A}|B)=0.9$,

$P(\overline{A}|\overline{B})=0.05$.

由贝叶斯公式,得该生在毕业三年内还清贷款,该生是可信的学生的概率为

$$P(B|\overline{A})=\frac{P(B)P(\overline{A}|B)}{P(B)P(\overline{A}|B)+P(\overline{B})P(\overline{A}|\overline{B})}$$

$$=\frac{0.8\times0.9}{0.8\times0.9+0.2\times0.05}\approx0.99.$$

**【探究总结】**

1.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 $P(A)=\sum\limits_{i=1}^{n}P(B_i)P(A|B_i)$;

第二步:计算 $P(AB_i)$,可利用 $P(AB_i)=P(B_i)\cdot P(A|B_i)$求解;

第三步:利用 $P(B_i|A)=\frac{P(AB_i)}{P(A)}$求解.

2.贝叶斯公式实质上是条件概率公式 $P(B_i|A)=\frac{P(AB_i)}{P(A)}$,$P(AB_i)=P(B_i)\cdot P(A|B_i)$,全概率公式 $P(A)=\sum\limits_{i=1}^{n}P(B_i)P(A|B_i)$的综合.

3.若随机试验可以看成分两个阶段进行,且第一阶段的各试验结果已知,随机试验结果未知,则:(1)如果要求的是第二阶段某一个结果发生的概率,那么用全概率公式;(2)如果第二阶段的某一个结果是已知的,要求的是此结果为第一阶段某一个结果所引起的概率,一般用贝叶斯公式,类似于求条件概率.熟记这个特征,在遇到相关的题目时,可以准确地选择方法进行计算,保证解题正确高效.

### 应用迁移

1.某地举办了一场地区性的中国象棋比赛,小明作为参赛选手参加.除小明外的其他参赛选手中,一、二、三类棋手的人数之比为5∶7∶8,小明与一、二、三类棋手比赛获胜的概率分别是0.6,0.5,0.4.

(1)从参赛选手中随机抽取一位棋手与小明比赛,求小明获胜的概率;

(2)如果小明获胜,求与小明比赛的棋手为一类棋手的概率.

**解**:(1)记事件 $B$="小明获胜",记事件 $A_i$="小明与 $i(i=1,2,3)$类棋手相遇",

由题意得,$P(A_1)=\frac{5}{20}=0.25$,$P(A_2)=\frac{7}{20}=0.35$,$P(A_3)=\frac{8}{20}=0.4$,$P(B|A_1)=0.6$,$P(B|A_2)=0.5$,$P(B|A_3)=0.4$.

由全概率公式可得 $P(B)=P(A_1)P(B|A_1)+P(A_2)\cdot P(B|A_2)+P(A_3)P(B|A_3)$

$=0.25\times0.6+0.35\times0.5+0.4\times0.4=0.485$.

(2)由(1)及贝叶斯公式可得 $P(A_1|B)=\frac{P(A_1B)}{P(B)}$

$=\frac{P(A_1)P(B|A_1)}{P(B)}=\frac{0.25\times0.6}{0.485}=\frac{30}{97}$,

即小明获胜,对手为一类棋手的概率为 $\frac{30}{97}$.

**2.** 试卷中的一道选择题有 4 个选项可供选择，其中只有 1 个选项是正确的.若考生会做这道题，则一定能选出正确选项；若考生不会做这道题，则会随机选择一个选项.设考生会做这道题的概率为 0.85.

(1)求任一考生选出此题正确选项的概率；

(2)已知某考生做对了此题，求该考生会做这道题的概率.

**解**：设 $A$ 表示“该考生会做这道题”，$B$ 表示“该考生选出正确选项”，则 $P(A)=0.85$，$P(\overline{A})=0.15$，$P(B|A)=1$，$P(B|\overline{A})=0.25$.

(1)由全概率公式得 $P(B)=P(A)P(B|A)+P(\overline{A})\cdot P(B|\overline{A})=0.85\times1+0.15\times0.25=0.887\ 5$.

(2)由(1)及贝叶斯公式得 $P(A|B)=\dfrac{P(A)P(B|A)}{P(B)}=\dfrac{0.85\times1}{0.887\ 5}\approx0.958$.

## 提质归纳

条件概率：$P(B|A)=\dfrac{P(AB)}{P(A)}$ → 乘法公式：$P(AB)=P(A)P(B|A)$

↓

全概率公式(由因求果)：$P(B)=\sum\limits_{i=1}^{n}P(A_i)P(B|A_i)$，$i=1,2,\cdots,n$ $\xrightarrow{\text{步骤}}$ 1.设事件 2.写概率 3.代公式

↓

*贝叶斯公式(执果寻因)：$P(A_i|B)=\dfrac{P(A_i)P(B|A_i)}{P(B)}=\dfrac{P(A_i)P(B|A_i)}{\sum\limits_{k=1}^{n}P(A_k)P(B|A_k)}$，$i=1,2,\cdots,n$

# 课后素养评价（十）　全概率公式

## A组　学习·理解

**1.** 某校篮球运动员进行投篮练习.若他前一球投进，则后一球也投进的概率为 $\dfrac{3}{4}$；若他前一球投不进，则后一球投进的概率为 $\dfrac{1}{4}$.若他第 1 球投进的概率为 $\dfrac{3}{4}$，则他第 2 球投进的概率为　　（ B ）

A.$\dfrac{3}{4}$　　B.$\dfrac{5}{8}$　　C.$\dfrac{7}{16}$　　D.$\dfrac{9}{16}$

**2.** 已知 $A$，$B$ 为样本空间 $\Omega$ 中的事件，$AB$ 与 $\overline{A}B$ 是互斥的，$B=AB+\overline{A}B$，且 $P(AB)=\dfrac{1}{2}$，$P(\overline{A}B)=\dfrac{1}{3}$，则 $P(B)=$　　（　　）

A.$\dfrac{1}{2}$　　B.$\dfrac{1}{3}$

C.$\dfrac{1}{6}$　　D.$\dfrac{5}{6}$

D　**解析**：由互斥事件概率的加法公式，得 $P(B)=P(AB)+P(\overline{A}B)=\dfrac{1}{2}+\dfrac{1}{3}=\dfrac{5}{6}$.

**3.** 某人外出，委托邻居给家里盆栽浇一次水，若不浇水，盆栽枯萎的概率为 0.8；若浇水，盆栽枯萎的概率为 0.2.若邻居浇水的概率为 $P$，该人回来盆栽没有枯萎的概率为 0.74，则实数 $P$ 的值为　　（　　）

A.0.9　　B.0.85　　C.0.8　　D.0.75

A　**解析**：记事件 $A$ 为“盆栽没有枯萎”，事件 $W$ 为“邻居给盆栽浇水”，

由题意可得 $P(W)=P$，$P(\overline{W})=1-P$，$P(\overline{A}|\overline{W})=0.8$，$P(\overline{A}|W)=0.2$.

由对立事件的概率公式可得 $P(\overline{A})=1-P(A)=1-0.74=0.26$.

由全概率公式可得 $P(\overline{A})=P(W)P(\overline{A}|W)+P(\overline{W})\cdot P(\overline{A}|\overline{W})=P\times0.2+(1-P)\times0.8=0.26$，解得 $P=0.9$.故选 A.

**4.** 已知甲箱中有 2 个白球和 4 个红球，乙箱中有 4 个白球和 2 个红球(两箱中的球除颜色外，没有其他区别).质点从原点出发，每次等可能地向左或向右移动一个单位长度，记事件 $A=$“质点移动 6 次，最终在 2 的位置”.

(1)求事件 $A$ 发生的概率；

(2)若事件 $A$ 发生，则从甲箱中取一球，否则从乙箱中取一球.求取出的球是红球的概率.

**解**：(1)要使质点移动 6 次，最终在 2 的位置，则需质点向左移动 2 次，向右移动 4 次，所以 $P(A)=\dfrac{n(A)}{n(\Omega)}=\dfrac{C_6^2}{2^6}=\dfrac{15}{64}$ 或 $P(A)=C_6^2\times\left(\dfrac{1}{2}\right)^2\times\left(\dfrac{1}{2}\right)^4=\dfrac{15}{64}$.

(2)设“取出的球是红球”为事件 $B$，则 $B=AB\cup\overline{A}B$.又由题知，$P(B|A)=\dfrac{2}{3}$，$P(B|\overline{A})=\dfrac{1}{3}$，故根

据全概率公式得 $P(B)=P(AB)+P(\overline{A}B)=P(A)P(B|A)+P(\overline{A})\cdot P(B|\overline{A})=\frac{15}{64}\times\frac{2}{3}+\left(1-\frac{15}{64}\right)\times\frac{1}{3}=\frac{79}{192}$.

## B组 应用·实践

1.(多选)甲箱中有 4 个红球和 3 个白球,乙箱中有 2 个红球和 3 个白球(两箱中的球除颜色外没有其他区别).先从甲箱中随机取出一球放入乙箱,分别用事件 $A_1$ 和 $A_2$ 表示从甲箱中取出的球是红球和白球;再从乙箱中随机取出两球,用事件 $B$ 表示从乙箱中取出的两球都是红球,则 ( )

A.$P(B|A_1)=\frac{1}{15}$　　B.$P(A_1)=\frac{4}{7}$

C.$P(B)=\frac{1}{7}$　　D.$P(A_2B)=\frac{1}{15}$

BC **解析**:甲箱中有 4 个红球和 3 个白球,则 $P(A_1)=\frac{4}{7}$,

若事件 $A_1$ 发生,则乙箱中有 3 个红球和 3 个白球,所以 $P(B|A_1)=\frac{P(A_1B)}{P(A_1)}=\frac{C_3^2}{C_6^2}=\frac{3}{15}=\frac{1}{5}$,故 A 错误,B 正确;

因为若事件 $A_2$ 发生,则乙箱中有 2 个红球和 4 个白球,所以 $P(B|A_2)=\frac{C_2^2}{C_6^2}=\frac{1}{15}$,

所以 $P(A_2B)=P(B|A_2)\times P(A_2)=\frac{1}{15}\times\frac{3}{7}=\frac{1}{35}$.

所以 $P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)=\frac{4}{7}\times\frac{1}{5}+\frac{3}{7}\times\frac{1}{15}=\frac{1}{7}$,故 C 正确,D 错误.

故选 BC.

2.1 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,这些球除颜色外没有区别.现随机从 1 号箱中取出 1 个球放入 2 号箱,然后从 2 号箱中随机取出 1 个球,则从 2 号箱取出红球的概率是________.

$\frac{11}{27}$ **解析**:设 $A$=“从 2 号箱中取出的是红球”,$B$=“从 1 号箱中取出的是红球”,

则 $P(B)=\frac{4}{2+4}=\frac{2}{3}$,$P(\overline{B})=1-P(B)=\frac{1}{3}$,

$P(A|B)=\frac{3+1}{8+1}=\frac{4}{9}$,$P(A|\overline{B})=\frac{3}{8+1}=\frac{1}{3}$.

所以 $P(A)=P(AB\cup A\overline{B})$

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

$=P(A|B)P(B)+P(A|\overline{B})P(\overline{B})$

$=\frac{4}{9}\times\frac{2}{3}+\frac{1}{3}\times\frac{1}{3}$

$=\frac{11}{27}$.

3.设某工厂有两个车间生产同型号家用电器,第一车间的次品率为 0.15,第二车间的次品率为 0.12,两个车间的成品都混合堆放在一个仓库,假设第一、二车间生产的成品比例为 2∶3.今有一客户从成品仓库中随机提一台产品,则该产品合格的概率为________.

0.868 **解析**:设 $B$=“从成品仓库中随机提一台产品是合格品”,$A_i$=“提出的产品是第 $i$ 车间生产的产品”,$i=1,2$,则 $B=A_1B\cup A_2B$.

因为第一、二车间生产的成品比例为 2∶3,

所以 $P(A_1)=0.4$,$P(A_2)=0.6$.

因为第一车间的次品率为 0.15,第二车间的次品率为0.12,所以 $P(B|A_1)=1-0.15=0.85$,$P(B|A_2)=1-0.12=0.88$.

所以 $P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)=0.4\times0.85+0.6\times0.88=0.868$.

4.8 支步枪中有 5 支已校准过,3 支未校准.一名射手用校准过的枪射击时,中靶的概率为 $\frac{4}{5}$;用未校准的枪射击时,中靶的概率为 $\frac{3}{10}$.现该射手从 8 支枪中任取 1 支用于射击,结果中靶,则他所用的枪是校准过的概率为________.

$\frac{40}{49}$ **解析**:设事件 $B_1$ 表示“使用的枪校准过”,事件 $B_2$ 表示“使用的枪未校准”,事件 $A$ 表示“射击时中靶”,

则 $P(B_1)=\frac{5}{8}$,$P(B_2)=\frac{3}{8}$,$P(A|B_1)=\frac{4}{5}$,$P(A|B_2)=\frac{3}{10}$.

由贝叶斯公式,得 $P(B_1|A)=\frac{P(A|B_1)P(B_1)}{P(A|B_1)P(B_1)+P(A|B_2)P(B_2)}=\frac{\frac{4}{5}\times\frac{5}{8}}{\frac{4}{5}\times\frac{5}{8}+\frac{3}{10}\times\frac{3}{8}}=\frac{40}{49}$.

所以所用的枪是校准过的概率为 $\frac{40}{49}$.

5.现有甲、乙两个箱子装有大小、外观均相同的青团,已知甲箱中有 5 个蛋黄馅的青团和 3 个肉松馅的青团,乙箱中有 4 个蛋黄馅的青团和 3 个肉松馅的青团.

(1)若从甲箱中任取 2 个青团,求这 2 个青团都是

肉松馅的概率.

(2)若先从甲箱中任取2个青团放入乙箱中,然后再从乙箱中任取1个青团.求取出的这个青团是蛋黄馅的概率.

**解**:(1)从甲箱中任取2个青团的样本点数为$C_8^2=28$,这2个青团都是肉松馅的样本点数为$C_3^2=3$,所以这2个青团都是肉松馅的概率$P=\frac{3}{28}$.

(2)设事件$A$为"从乙箱中任取1个青团,取出的这个青团是蛋黄馅",事件$B_1$为"从甲箱中取出的2个青团都是蛋黄馅",事件$B_2$为"从甲箱中取出的2个青团为1个蛋黄馅1个肉松馅",事件$B_3$为"从甲箱中取出的2个青团都是肉松馅",则事件$B_1$,$B_2$,$B_3$彼此互斥.

由题意得$P(B_1)=\frac{C_5^2}{C_8^2}=\frac{5}{14}$,$P(B_2)=\frac{C_5^1C_3^1}{C_8^2}=\frac{15}{28}$,$P(B_3)=\frac{C_3^2}{C_8^2}=\frac{3}{28}$,$P(A|B_1)=\frac{2}{3}$,$P(A|B_2)=\frac{5}{9}$,$P(A|B_3)=\frac{4}{9}$,所以$P(A)=P(B_1)P(A|B_1)+P(B_2)P(A|B_2)+P(B_3)P(A|B_3)=\frac{5}{14}\times\frac{2}{3}+\frac{15}{28}\times\frac{5}{9}+\frac{3}{28}\times\frac{4}{9}=\frac{7}{12}$.

所以取出的这个青团是蛋黄馅的概率为$\frac{7}{12}$.

## 7.2　离散型随机变量及其分布列

**学习任务目标**

**1.**通过具体实例,了解离散型随机变量的概念.

**2.**掌握离散型随机变量的分布列的性质.

**3.**理解两点分布,并能进行简单应用.

**4.**会求简单的离散型随机变量的概率分布.

### ○ 问题式预习 ○

**【知识清单】**

**知识点一　离散型随机变量**

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.可能取值为有限个或可以一一列举的随机变量,我们称为离散型随机变量.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$x,y,z$.

**知识点二　概率分布列**

(1)定义:一般地,设离散型随机变量$X$的可能取值为$x_1,x_2,\cdots,x_n$,我们称$X$取每一个值$x_i$的概率$P(X=x_i)=p_i,i=1,2,\cdots,n$为$X$的概率分布列,简称分布列.

(2)表示方法:表格法、图形法.

(3)性质:①$p_i\geqslant0,i=1,2,\cdots,n$;

②$p_1+p_2+\cdots+p_n=1$.

**知识点三　两点分布**

若离散型随机变量$X$的分布列为(其中$0\leqslant p\leqslant1$)

| $X$ | 0 | 1 |
|---|---|---|
| $P$ | $1-p$ | $p$ |

则称$X$服从两点分布,或0—1分布.

**【概念辨析】**

**1.**判断正误(正确的打"√",错误的打"×").

(1)离散型随机变量的取值是任意的实数.　( × )

(2)随机变量的取值可以是有限个,也可以是无限个.　( √ )

(3)杭州第19届亚运会上中国取得的金牌数是随机变量.　( × )

**2.**已知随机变量$X$的分布列是

| $X$ | 1 | 2 | 3 |
|---|---|---|---|
| $P$ | $\frac{1}{3}$ | $a$ | $b$ |

则$a+b=$　(　　)

A.$\frac{2}{3}$　　B.$\frac{3}{2}$　　C.1　　D.$\frac{3}{4}$

A　**解析**:由随机变量$X$的分布列的性质得$\frac{1}{3}+a+b=1$,解得$a+b=\frac{2}{3}$.

**3.**(多选)下列问题中的随机变量服从两点分布的是　(　　)

A.抛掷一枚骰子,所得点数为随机变量$X$

B.某射手射击一次,击中目标的次数为随机变量$X$

C.从装有5个红球、3个白球的袋中任取1个球，令随机变量 $X=\begin{cases}1,取出白球,\\0,取出红球\end{cases}$

D.某医生做一次手术，手术成功的次数为随机变量 $X$

BCD　**解析**：两点分布又叫0－1分布，所有的试验结果有两个，B，C，D满足定义.而抛掷一枚骰子，所得点数为随机变量 $X$，则 $X$ 的所有可能的结果有6种，不服从两点分布.

**4.**请思考并回答下列问题：

(1)随机变量的定义与函数的定义有何区别与联系？

**提示**：随机变量的定义与函数的定义类似，样本点 $\omega$ 相当于函数定义中的自变量，而样本空间 $\Omega$ 相当于函数的定义域，不同之处在于 $\Omega$ 不一定是数集.

(2)一次试验中随机变量 $X$ 只有1和2两个取值，随机变量 $X$ 是否服从两点分布？

**提示**：不服从，因为两点分布中随机变量的取值只能是0和1，因此两点分布又称0－1分布.

## ◦ 任务型课堂 ◦

### 任务1 离散型随机变量的判定及取值

**1.**下列变量中，哪些是随机变量？哪些是离散型随机变量？并说明理由.

(1)某机场一年中每天服务乘客的数量；

(2)某单位办公室一天中接到电话的次数；

(3)某地明年5月1日到10月1日期间所查酒驾的人数；

(4)一瓶标注净含量为(500±2) mL的果汁的净含量.

**解**：(1)某机场一年中每天服务乘客的数量可能为0，1，2，3，…，是随机变化的，因此是随机变量，也是离散型随机变量.

(2)某单位办公室一天中接到电话的次数可能为0，1，2，3，…，是随机变化的，因此是随机变量，也是离散型随机变量.

(3)某地明年5月1日到10月1日期间所查酒驾的人数可能为0，1，2，3，…，是随机变化的，因此是随机变量，也是离散型随机变量.

(4)果汁的净含量在498 mL～502 mL之间波动，是随机变量，但不是离散型随机变量.

**2.**写出下列随机变量 $X$ 可能的取值，并说明随机变量所取的值表示的随机试验的结果.

(1)从一个装有编号为1～10的10个球的袋中，任取1个球，被取出的球的编号为 $X$；

(2)一个袋中装有10个红球、5个白球，从中任取4个球，其中所含红球的个数为 $X$；

(3)投掷甲、乙两枚骰子，所得点数之和为 $X$.

**解**：(1)$X$ 的可能取值为1，2，3，…，10.

$X=k(k=1,2,\cdots,10)$表示“取出编号为 $k$ 的球”.

(2)$X$ 的可能取值为0，1，2，3，4.

$X=k$ 表示“取出 $k$ 个红球，$(4-k)$个白球”，其中 $k=0,1,2,3,4$.

(3)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.

$X=2$ 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).

**【探究总结】**

**用离散型随机变量表示随机试验结果的关键点和注意点**

**1.**关键点：明确随机变量的所有可能取值，以及取每一个值时对应的意义，即一个随机变量的取值对应一个或多个随机试验的结果.

**2.**注意点：解答过程中不要漏掉某些试验结果.

### 任务2 离散型随机变量分布列的性质

#### 探究活动

**例1**　随机变量 $X$ 的分布列如下表所示：

| $X$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $P$ | 0.1 | $m$ | 0.3 | $2m$ |

则 $P(X>2)=$________.

0.7　**解析**：由分布列的性质可得，$0.1+m+0.3+2m=1$，解得 $m=0.2$，

所以 $P(X>2)=P(X=3)+P(X=4)=0.3+2\times0.2=0.7$.

**【探究总结】**

**离散型随机变量分布列的性质的应用**

(1)利用离散型随机变量的分布列的性质可以求与概率有关的参数的值或取值范围，还可以检验所求分布列是否正确.

(2)由于离散型随机变量的各个可能取值表示的事件是两两互斥的，所以离散型随机变量在某一范围内取值的概率等于它取这个范围内各个值的概率之和.

#### 应用迁移

**1.**若随机变量 $X$ 的分布列为

| $X$ | $-2$ | $-1$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|
| $P$ | 0.1 | 0.2 | 0.1 | 0.3 | 0.1 | 0.2 |

则当 $P(X<a)=0.7$ 时，实数 $a$ 的取值范围是（　　）

A.$(-\infty,2]$　　B.$[1,2]$

C.$(1,2]$　　D.$(1,2)$

C　**解析**：由随机变量 $X$ 的分布列知，

$P(X<-1)=0.1$，$P(X<0)=0.3$，$P(X<1)=0.4$，$P(X<2)=0.7$，

则当 $P(X<a)=0.7$ 时，实数 $a$ 的取值范围是 $(1,2]$.

**2.** 离散型随机变量 $X$ 的分布列为

| $X$ | 0 | 1 |
|---|---|---|
| $P$ | $9C^2-C$ | $3-8C$ |

则常数 $C$ 的值为（　　）

A.$\frac{2}{3}$　　B.$\frac{1}{3}$

C.$\frac{2}{3}$或$\frac{1}{3}$　　D.以上都不对

B　**解析**：由离散型随机变量 $X$ 的分布列，

得 $\begin{cases}0\leqslant 9C^2-C\leqslant 1,\\ 0\leqslant 3-8C\leqslant 1,\\ 9C^2-C+3-8C=1,\end{cases}$　解得 $C=\frac{1}{3}$.

**3.** 设随机变量 $X$ 的分布列为

| $X$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $P$ | $\frac{1}{3}$ | $m$ | $\frac{1}{4}$ | $\frac{1}{6}$ |

则 $P(|X-3|=1)=$（　　）

A.$\frac{7}{12}$　　B.$\frac{5}{12}$

C.$\frac{1}{4}$　　D.$\frac{1}{6}$

B　**解析**：根据分布列的性质，有 $\frac{1}{3}+m+\frac{1}{4}+\frac{1}{6}=1$，解得 $m=\frac{1}{4}$.

所以 $P(|X-3|=1)=P(X=4)+P(X=2)=\frac{1}{6}+\frac{1}{4}=\frac{5}{12}$.

**4.** 设随机变量 $X$ 的分布列为 $P(X=k)=\frac{C}{k(k+1)}$，$k=1,2,3$，$C$ 为常数，则 $P(0.5<X<2.5)=$________.

$\frac{8}{9}$　**解析**：由 $C\times\left(\frac{1}{1\times 2}+\frac{1}{2\times 3}+\frac{1}{3\times 4}\right)=1$，得 $C=\frac{4}{3}$.

所以 $P(0.5<X<2.5)=P(X=1)+P(X=2)=\frac{2}{3}+\frac{2}{9}=\frac{8}{9}$.

## 任务3　两点分布

### 探究活动

**例2**　一个袋中装有除颜色外其他都相同的3个白球和4个红球.

(1)从此袋中任意摸出1个球，用 $X=0$ 表示摸出白球，用 $X=1$ 表示摸出红球，即 $X=\begin{cases}0,\text{摸出白球},\\ 1,\text{摸出红球},\end{cases}$ 求 $X$ 的分布列；

(2)从此袋中任意摸出2个球，用 $X=0$ 表示“2个球全是白球”，用 $X=1$ 表示“2个球不全是白球”，求 $X$ 的分布列.

**解**：(1)由题意知 $P(X=0)=\frac{3}{7}$，$P(X=1)=\frac{4}{7}$.

所以 $X$ 的分布列为

| $X$ | 0 | 1 |
|---|---|---|
| $P$ | $\frac{3}{7}$ | $\frac{4}{7}$ |

(2)由题意知 $P(X=0)=\frac{C_3^2}{C_7^2}=\frac{1}{7}$，$P(X=1)=1-P(X=0)=\frac{6}{7}$，

所以 $X$ 的分布列为

| $X$ | 0 | 1 |
|---|---|---|
| $P$ | $\frac{1}{7}$ | $\frac{6}{7}$ |

**【探究总结】**

**1.** 两点分布的特点：

(1)两点分布中只有两个对应结果，且两个结果是对立的.

(2)由对立事件的概率求法可知，若 $X$ 服从两点分布，则 $P(X=0)+P(X=1)=1$.

**2.** 两点分布的适用范围：

(1)研究只有两个结果的随机试验的概率分布规律.

(2)研究某一随机事件是否发生的概率分布规律.

如抽取的彩券是否中奖、买回的一件产品是不是正品、新生婴儿的性别、投篮是否命中等，都可以用两点分布来研究.

### 应用迁移

**1.** 设某试验的成功率是失败率的2倍.用随机变量 $\xi$ 去描述一次试验的成功与否($\xi=1$ 表示“试验成功”，$\xi=0$ 表示“试验失败”)，则 $P(\xi=0)$ 等于（　　）

A.0　B.$\frac{1}{2}$　C.$\frac{1}{3}$　D.$\frac{2}{3}$

C　**解析**：由题意知 $\xi=0$ 表示“试验失败”，$\xi=1$ 表示“试验成功”.

设失败率为 $p$，则成功率为 $2p$，$\xi$ 的分布列为

| $\xi$ | 0 | 1 |
|---|---|---|
| $P$ | $p$ | $2p$ |

由 $p+2p=1$，得 $p=\frac{1}{3}$. 故 $P(\xi=0)=\frac{1}{3}$.

**2.** 设随机变量 $X$ 服从两点分布.若 $P(X=1)-P(X=0)=0.2$，则 $P(X=1)=$ （　　）

A.0.2　　B.0.4　　C.0.6　　D.0.8

C　**解析**：根据题意和两点分布的性质可知

$$\begin{cases}P(X=1)-P(X=0)=0.2,\\P(X=1)+P(X=0)=1,\end{cases}$$

解得 $P(X=1)=0.6$.

## 任务4 > 离散型随机变量的分布列

### 探究活动

**例 3**　一个箱子里装有 5 个大小相同的球，有 3 个白球、2 个红球.从此箱中随机摸出 2 个球，用 $X$ 表示摸出的 2 个球中的白球个数，求 $X$ 的分布列.

**解**：用 $X$ 表示摸出的 2 个球中的白球个数，$X$ 的所有可能取值为 0，1，2.

$P(X=0)=\frac{C_2^2}{C_5^2}=\frac{1}{10}$，$P(X=1)=\frac{C_3^1C_2^1}{C_5^2}=\frac{3}{5}$，

$P(X=2)=\frac{C_3^2}{C_5^2}=\frac{3}{10}$.

故 $X$ 的分布列为

| $X$ | 0 | 1 | 2 |
|---|---|---|---|
| $P$ | $\frac{1}{10}$ | $\frac{3}{5}$ | $\frac{3}{10}$ |

［**一题多思**］

**思考 1.** 每次从此箱中任意摸出 1 个球，每次摸出的红球不再放回，直到摸出白球为止，求摸球次数 $X$ 的分布列.

**解**：由题意知，$X$ 的可能取值为 1，2，3，

第 1 次摸到白球的概率为 $P(X=1)=\frac{3}{5}$，

第 2 次摸到白球的概率为 $P(X=2)=\frac{2}{5}\times\frac{3}{4}=\frac{3}{10}$，

第 3 次摸到白球的概率为 $P(X=3)=\frac{2}{5}\times\frac{1}{4}\times1=\frac{1}{10}$.

所以 $X$ 的分布列为

| $X$ | 1 | 2 | 3 |
|---|---|---|---|
| $P$ | $\frac{3}{5}$ | $\frac{3}{10}$ | $\frac{1}{10}$ |

**思考 2.** 每次从此箱中任意摸出 1 个球，摸出的球不再放回，直到摸出所有白球为止，求摸球次数 $X$ 的分布列.

**解**：由题意知，$X$ 的所有可能取值为 3，4，5，

3 次摸到所有白球的概率为 $P(X=3)=\frac{3}{5}\times\frac{2}{4}\times\frac{1}{3}=\frac{1}{10}$，

4 次摸到所有白球的概率为 $P(X=4)=\frac{2}{5}\times\frac{3}{4}\times\frac{2}{3}\times\frac{1}{2}+\frac{3}{5}\times\frac{2}{4}\times\frac{2}{3}\times\frac{1}{2}+\frac{3}{5}\times\frac{2}{4}\times\frac{2}{3}\times\frac{1}{2}=\frac{3}{10}$，

5 次摸到所有白球的概率为 $P(X=5)=1-\frac{1}{10}-\frac{3}{10}=\frac{3}{5}$，所以 $X$ 的分布列为

| $X$ | 3 | 4 | 5 |
|---|---|---|---|
| $P$ | $\frac{1}{10}$ | $\frac{3}{10}$ | $\frac{3}{5}$ |

**【探究总结】**

**求离散型随机变量分布列的步骤**

明取值—明确随机变量的可能取值有哪些，且每一个取值都有意义

↓

求概率—要弄清楚随机变量的概率类型，利用相关公式求出变量取每个值的概率

↓

画表格—按规范形式写出分布列

↓

做检验—利用分布列的性质检验分布列是否正确

### 应用迁移

**1.** 甲、乙两名射击运动员，根据历史统计数据，甲射击一次命中 10 环、9 环、8 环的概率分别为 $\frac{2}{5}$，$\frac{2}{5}$，$\frac{1}{5}$；乙射击一次命中 10 环、9 环的概率分别为 $\frac{1}{6}$，$\frac{5}{6}$.一轮射击中，甲、乙各射击一次，甲、乙射击相互独立，每次射击也互不影响.

(1)在一轮射击中，求甲命中的环数不高于乙命中的环数的概率；

(2)记一轮射击中，甲、乙命中的环数之和为 $X$，求 $X$ 的分布列；

(3)进行三轮射击，求甲、乙命中的环数之和不低于 52 的概率.

**解**：(1)当甲命中的环数高于乙命中的环数时，只有一种情况：甲命中 10 环，且乙命中 9 环，这时概率 $p_0=\frac{2}{5}\times\frac{5}{6}=\frac{1}{3}$.所以甲命中的环数不高于乙命中的环数的概率 $p=1-p_0=\frac{2}{3}$.

(2)甲、乙命中的环数之和 $X$ 的可能取值为 17，18，19，20，

$P(X=17)=\frac{1}{5}\times\frac{5}{6}=\frac{1}{6}$，$P(X=18)=\frac{1}{5}\times\frac{1}{6}+\frac{2}{5}\times\frac{5}{6}=\frac{11}{30}$，$P(X=19)=\frac{2}{5}\times\frac{5}{6}+\frac{2}{5}\times\frac{1}{6}=\frac{2}{5}$，$P(X=20)=\frac{2}{5}\times\frac{1}{6}=\frac{1}{15}$.

所以随机变量 $X$ 的分布列为

| $X$ | 17 | 18 | 19 | 20 |
|---|---|---|---|---|
| $P$ | $\frac{1}{6}$ | $\frac{11}{30}$ | $\frac{2}{5}$ | $\frac{1}{15}$ |

(3)三轮射击后，甲、乙命中的环数之和低于 52 时，甲、乙每轮命中环数之和都是 17，其概率 $p_1=\left(\frac{1}{6}\right)^3=\frac{1}{216}$，

所以甲、乙命中的环数之和不低于 52 的概率 $p=1-p_1=1-\frac{1}{216}=\frac{215}{216}$.

**2.**学校羽毛球社团中的甲、乙、丙三名社员进行羽毛球比赛，约定如下：先从甲、乙、丙三人中随机选择两人打第一局，获胜者与第三人进行下一局的比赛，率先获胜两局者为优胜者，比赛结束，且每局比赛均无平局.已知甲赢乙的概率为 0.3，乙赢丙的概率为 0.5，丙赢甲的概率为 0.7.

(1)若甲、乙二人率先开局比赛，求比赛局数 $X$ 的概率分布列；

(2)求甲成为优胜者的概率.

**解**：(1)比赛局数 $X$ 的可能取值为 2，3，4.

若比赛两局结束，则甲连胜两局或乙连胜两局，所以 $P(X=2)=0.3\times0.3+0.7\times0.5=0.44$；

若比赛三局结束，则第二局、第三局丙连胜，所以 $P(X=3)=0.3\times0.7\times0.5+0.7\times0.5\times0.7=0.35$；

若比赛四局结束，此时 $P(X=4)=1-P(X=2)-P(X=3)=0.21$.

所以 $X$ 的分布列为

| $X$ | 2 | 3 | 4 |
|---|---|---|---|
| $P$ | 0.44 | 0.35 | 0.21 |

(2)记甲、乙比赛第一局为事件 $A$，甲、丙比赛第一局为事件 $B$，乙、丙比赛第一局为事件 $C$，甲成为优胜者为事件 $D$.第一局比赛双方可能是甲乙、甲丙、乙丙共三种情况，

则 $P(A)=P(B)=P(C)=\frac{1}{3}$，

所以 $P(D|A)=0.3\times0.3+0.3\times0.7\times0.5\times0.3+0.7\times0.5\times0.3\times0.3=0.153$，

$P(D|B)=0.3\times0.3+0.3\times0.7\times0.5\times0.3+0.7\times0.5\times0.3\times0.3=0.153$，

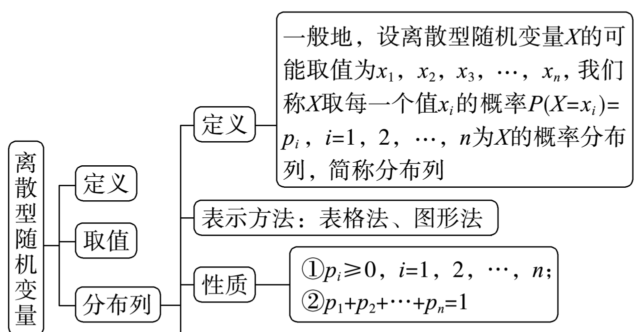
$P(D|C)=0.5\times0.3\times0.3+0.5\times0.3\times0.3=0.09$.

所以 $P(D)=P(D|A)P(A)+P(D|B)P(B)+P(D|C)P(C)$

$=0.153\times\frac{1}{3}+0.153\times\frac{1}{3}+0.09\times\frac{1}{3}$

$=0.132$.

## 提质归纳




# 课后素养评价（十一）　离散型随机变量及其分布列

## A组　学习·理解

**1.**给出下列各量：

①某机场候机室中一天的游客数量；

②某寻呼台一天内收到的寻呼次数；

③某同学离开自己学校的距离；

④将要举行的绘画比赛中某同学获得的名次；

⑤体积为 8 $m^3$ 的正方体的棱长.

其中是离散型随机变量的是　　(　　)

A.①②④　B.①②③　C.③④⑤　D.②③④

A　**解析**：由题意知，①②④是离散型随机变量，③是连续型随机变量，⑤中体积为 8 $m^3$ 的正方体的棱长是一个常量，不是随机变量.故选 A.

**2.** 袋中装有大小相同的5个球，分别标有1，2，3，4，5五个号码，现在在有放回的条件下随机取出2个球，设取出的2个球的号码之和为随机变量$\xi$，则$\xi$所有可能取值的个数是（　　）

A.25　B.10　C.15　D.9

D　**解析**：由题意得取出的2个球的号码之和可能为2，3，4，5，6，7，8，9，10，共9个.故选D.

**3.** 已知离散型随机变量$X$的分布列服从两点分布，且$P(X=0)=2-5P(X=1)=a$，则$a=$（　　）

A.$\frac{3}{4}$　B.$\frac{1}{2}$　C.$\frac{1}{3}$　D.$\frac{2}{3}$

A　**解析**：因为$X$的分布列服从两点分布，

所以$P(X=0)+P(X=1)=1$.

又$P(X=0)=2-5P(X=1)=a$，

所以$P(X=0)=2-5[1-P(X=0)]$，

解得$P(X=0)=\frac{3}{4}$，所以$a=\frac{3}{4}$.

**4.** 设随机变量$X$等可能地从1，2，3，4，…，10中取值.又设随机变量$Y=2X-1$，则$P(Y<6)$的值为（　　）

A.0.3　B.0.5　C.0.1　D.0.2

A　**解析**：因为$Y<6$，即$2X-1<6$，所以$X<3.5$.所以$X=1,2,3$，故$P(Y<6)=0.3$.

**5.** 袋中装有大小相同的10个红球，5个黑球.从中每次随机抽取1个球，若取到黑球，则另换1个红球放回袋中，直到取到红球为止.若抽取的次数为$X$，则事件"放回5个红球"可表示为（　　）

A.$X=4$　B.$X=5$

C.$X=6$　D.$X\leqslant4$

C　**解析**：第一次取到黑球，则放回1个红球；第二次取到黑球，则共放回2个红球……第五次取到黑球，则共放回5个红球；第六次取到了红球，停止取球.故$X=6$.故选C.

**6.** 设随机变量$X$的分布列为$P(X=i)=ai(i=1,2,\cdots,8)$，则常数$a=$________.

$\frac{1}{36}$　**解析**：因为$P(X=i)=ai(i=1,2,\cdots,8)$，所以$a(1+2+3+\cdots+8)=36a=1$，解得$a=\frac{1}{36}$.

**7.** 已知随机变量$X$的可能取值有3个，且取每个值的概率$p_1,p_2,p_3$成等差数列，则公差$d$的取值范围是________.

$\left[-\frac{1}{3},\frac{1}{3}\right]$　**解析**：由分布列的性质及等差数列的性质得$p_1+p_2+p_3=3p_2=1$，$p_2=\frac{1}{3}$.

又$\begin{cases}p_1\geqslant0,\\p_3\geqslant0,\end{cases}$即$\begin{cases}\frac{1}{3}-d\geqslant0,\\\frac{1}{3}+d\geqslant0,\end{cases}$得$-\frac{1}{3}\leqslant d\leqslant\frac{1}{3}$，

所以公差$d$的取值范围是$\left[-\frac{1}{3},\frac{1}{3}\right]$.

**8.** 设离散型随机变量$X$的分布列为

| $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P$ | 0.2 | 0.1 | 0.1 | 0.3 | $m$ |

求随机变量$\eta=|X-1|$的分布列.

**解**：由题意可知$m=1-0.2-0.1-0.1-0.3=0.3$，列表为

| $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $\lvert X-1\rvert$ | 1 | 0 | 1 | 2 | 3 |
| $P$ | 0.2 | 0.1 | 0.1 | 0.3 | 0.3 |

所以$P(\eta=0)=P(X=1)=0.1$，

$P(\eta=1)=P(X=0)+P(X=2)=0.2+0.1=0.3$，

$P(\eta=2)=P(X=3)=0.3$，

$P(\eta=3)=P(X=4)=0.3$.

故$\eta=|X-1|$的分布列为

| $\eta$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P$ | 0.1 | 0.3 | 0.3 | 0.3 |

## B组　应用·实践

**1.**（多选）甲、乙两人下象棋，赢了得3分，平局得1分，输了得0分，共下三局.用$\xi$表示甲的得分，则事件$\{\xi=3\}$表示的可能结果为（　　）

A.甲赢三局

B.甲赢一局输两局

C.甲、乙平局三次

D.甲赢一局

BC　**解析**：甲赢一局输两局得3分，甲与乙平局三次得3分.

**2.** 已知随机变量$\xi$的分布列为

| $\xi$ | 0 | 1 | 2 |
|---|---|---|---|
| $P$ | $a$ | $b$ | $c$ |

其中$a,b,c$成等差数列，则函数$f(x)=x^2+2x+\xi$有且只有一个零点的概率为（　　）

A.$\frac{1}{6}$　B.$\frac{1}{3}$　C.$\frac{1}{2}$　D.$\frac{5}{6}$

B　**解析**：由题意知$a,b,c\in[0,1]$，且$\begin{cases}2b=a+c,\\a+b+c=1,\end{cases}$解得$b=\frac{1}{3}$.又由函数$f(x)=x^2+$

$2x+\xi$ 有且只有一个零点，即方程 $x^2+2x+\xi=0$ 只有一个根，可得 $\Delta=4-4\xi=0$，解得 $\xi=1$.所以 $P(\xi=1)=\frac{1}{3}$.故选 B.

**3.**若随机变量 $\xi$ 满足 $P(\xi\leqslant n)=1-a$，$P(\xi\geqslant m)=1-b$，其中 $m<n$，则 $P(m\leqslant\xi\leqslant n)$ 等于　(　　)

A.$(1-a)(1-b)$　　B.$1-a(1-b)$

C.$1-(a+b)$　　D.$1-b(1-a)$

C　**解析**：$P(m\leqslant\xi\leqslant n)=1-P(\xi>n)-P(\xi<m)=1-[1-(1-a)]-[1-(1-b)]=1-(a+b)$.

**4.**某商场经销某商品，根据以往资料统计，顾客采用的付款期数 $\xi$ 的分布列为

| $\xi$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $P$ | 0.4 | 0.2 | 0.2 | 0.1 | 0.1 |

商场经销一件该商品，顾客采用 1 期付款，其利润为 200 元；分 2 期或 3 期付款，其利润为 250 元；分 4 期或 5 期付款，其利润为 300 元.若 $\eta$ 表示商场经销一件该商品的利润，求 $\eta$ 的分布列.

**解**：由题意知，$\eta$ 的可能取值为 200，250，300，

$P(\eta=200)=P(\xi=1)=0.4$，

$P(\eta=250)=P(\xi=2)+P(\xi=3)=0.2+0.2=0.4$，

$P(\eta=300)=P(\xi=4)+P(\xi=5)=0.1+0.1=0.2$.

故 $\eta$ 的分布列为

| $\eta$ | 200 | 250 | 300 |
|---|---|---|---|
| $P$ | 0.4 | 0.4 | 0.2 |

**5.**甲、乙两人准备进行羽毛球比赛，比赛规定：一回合中赢球的一方作为下一回合的发球方.若甲发球，则本回合甲赢的概率为 $\frac{2}{3}$，若乙发球，则本回合甲赢的概率为 $\frac{1}{3}$，每回合比赛的结果相互独立.经抽签决定，第 1 回合由甲发球.

(1)求第 4 回合甲发球的概率；

(2)设前 4 个回合中，甲发球的次数为 $X$，求 $X$ 的分布列.

**解**：(1)由题意可知，第 2 回合甲发球的概率为 $\frac{2}{3}$，乙发球的概率为 $\frac{1}{3}$，

所以第 3 回合甲发球的概率为 $\frac{2}{3}\times\frac{2}{3}+\frac{1}{3}\times\frac{1}{3}=\frac{5}{9}$，乙发球的概率为 $\frac{2}{3}\times\frac{1}{3}+\frac{1}{3}\times\frac{2}{3}=\frac{4}{9}$.

因此第 4 回合甲发球的概率为 $\frac{5}{9}\times\frac{2}{3}+\frac{4}{9}\times\frac{1}{3}=\frac{14}{27}$.

故第 4 回合甲发球的概率为 $\frac{14}{27}$.

(2)由题意可知，$X$ 可以取 1，2，3，4.

当 $X=1$ 时，$P_1=\frac{1}{3}\times\frac{2}{3}\times\frac{2}{3}=\frac{4}{27}$；当 $X=4$ 时，$P_4=\left(\frac{2}{3}\right)^3=\frac{8}{27}$.

当 $X=2$ 时，前 4 个回合甲发球两次的情况分以下三种：

第一种情况，甲第 1，2 回合发球，乙第 3，4 回合发球，其概率为 $\frac{2}{3}\times\frac{1}{3}\times\frac{2}{3}=\frac{4}{27}$；

第二种情况，甲第 1，3 回合发球，乙第 2，4 回合发球，其概率为 $\frac{1}{3}\times\frac{1}{3}\times\frac{1}{3}=\frac{1}{27}$；

第三种情况，甲第 1，4 回合发球，乙第 2，3 回合发球，其概率为 $\frac{1}{3}\times\frac{2}{3}\times\frac{1}{3}=\frac{2}{27}$.

故前 4 个回合甲发球两次的概率为 $P_2=\frac{4}{27}+\frac{1}{27}+\frac{2}{27}=\frac{7}{27}$；

当 $X=3$ 时，$P_3=1-P_1-P_2-P_4=\frac{8}{27}$.

故 $X$ 的分布列为

| $X$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $P$ | $\frac{4}{27}$ | $\frac{7}{27}$ | $\frac{8}{27}$ | $\frac{8}{27}$ |

# 7.3 离散型随机变量的数字特征

## 7.3.1 离散型随机变量的均值

**学习任务目标**

1. 理解离散型随机变量的均值的概念和性质.
2. 会根据离散型随机变量的分布列求出均值.
3. 会利用离散型随机变量的均值解决一些实际问题.

### ○ 问题式预习 ○

**【知识清单】**

**知识点 均值的定义及性质**

(1)一般地，若离散型随机变量 $X$ 的分布列为

| $X$ | $x_1$ | $x_2$ | … | $x_n$ |
|---|---|---|---|---|
| $P$ | $p_1$ | $p_2$ | … | $p_n$ |

则称 $E(X)=\underline{x_1p_1+x_2p_2+\cdots+x_np_n=\sum_{i=1}^{n}x_ip_i}$ 为随机变量 $X$ 的均值或数学期望，数学期望简称期望.

(2)意义：均值是随机变量可能取值关于取值概率的<u>加权平均数</u>，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<u>平均水平</u>.

(3)一般地，如果随机变量 $X$ 服从两点分布，那么 $E(X)=\underline{p}$.

(4)性质：$E(X+b)=\underline{E(X)+b}$，$E(aX)=\underline{aE(X)}$，$E(aX+b)=\underline{aE(X)+b}$.

**【概念辨析】**

1. 判断正误（正确的打“√”，错误的打“×”）.

(1)随机变量的均值与样本的均值是同一个概念.（ × ）

(2)$E(aX+bY)=(a+b)[E(X)+E(Y)]$.（ × ）

(3)随机变量的均值与随机变量本身的单位不相同.（ × ）

(4)常数的均值就是这个常数本身.（ √ ）

2. 若随机变量 $X$ 的分布列为

| $X$ | $-1$ | $0$ | $1$ |
|---|---|---|---|
| $P$ | $\frac{1}{2}$ | $\frac{1}{6}$ | $\frac{1}{3}$ |

则 $E(X)=$（　　）

A. $0$　　B. $-1$

C. $-\frac{1}{6}$　　D. $-\frac{1}{2}$

C　**解析**：$E(X)=(-1)\times\frac{1}{2}+0\times\frac{1}{6}+1\times\frac{1}{3}=-\frac{1}{6}$.

3. 请思考并回答下列问题：

(1)如何理解随机变量的均值？

**提示**：随机变量的均值也称数学期望（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.

(2)随机变量的均值与样本的均值有何区别与联系？

**提示**：随机变量的均值是一个确定的数，而样本的均值具有随机性，它围绕着随机变量的均值波动.随着重复试验次数的增加，样本的均值波动幅度一般会越来越小.因此，我们常用随机变量的观测值的均值（样本均值）去估计随机变量的均值.

### ○ 任务型课堂 ○

**任务1 > 离散型随机变量的均值及性质**

**探究活动**

**例1** 已知随机变量 $X$ 的分布列为

| $X$ | $-2$ | $-1$ | $0$ | $1$ | $2$ |
|---|---|---|---|---|---|
| $P$ | $\frac{1}{4}$ | $\frac{1}{3}$ | $\frac{1}{5}$ | $m$ | $\frac{1}{20}$ |

若 $Y=-2X$，则 $E(Y)=$________.

$\frac{17}{15}$　**解析**：由分布列的性质，得

$\frac{1}{4}+\frac{1}{3}+\frac{1}{5}+m+\frac{1}{20}=1$，解得 $m=\frac{1}{6}$.

所以 $E(X)=(-2)\times\frac{1}{4}+(-1)\times\frac{1}{3}+0\times\frac{1}{5}+1\times\frac{1}{6}+2\times\frac{1}{20}=-\frac{17}{30}$.

由 $Y=-2X$，得 $E(Y)=-2E(X)$，

即 $E(Y)=-2\times\left(-\frac{17}{30}\right)=\frac{17}{15}$.

［一题多思］

思考 1.若 $Y=2X-3$，求$E(Y)$.

**解**：由例题知 $E(X)=-\frac{17}{30}$，则 $E(Y)=E(2X-3)=2E(X)-3=2\times\left(-\frac{17}{30}\right)-3=-\frac{62}{15}$.

思考 2.若 $Y=aX+3$，且$E(Y)=-\frac{11}{2}$，求实数 $a$ 的值.

**解**：由例题知 $E(X)=-\frac{17}{30}$，

则 $E(Y)=E(aX+3)=aE(X)+3=-\frac{17}{30}a+3=-\frac{11}{2}$，

解得 $a=15$.

**【探究总结】**

1.已知离散型随机变量的分布列求均值，可直接套用公式$E(X)=x_1p_1+x_2p_2+\cdots+x_np_n$ 来求解.

2.$aX+b$ 型随机变量的均值，可利用均值的性质求解，即 $E(aX+b)=aE(X)+b$；也可以先列出$aX+b$ 的分布列，再用均值公式求解.

## 应用迁移

1.已知随机变量 $\xi$ 的分布列为 $P(\xi=k)=\frac{1}{4}$，$k=0$，1，2，3，则 $E(2\xi+4)=$　　(　　)

A.6　　B.7

C.8　　D.9

B　**解析**：因为随机变量 $\xi$ 的分布列为 $P(\xi=k)=\frac{1}{4}$，$k=0$，1，2，3，所以 $E(\xi)=(0+1+2+3)\times\frac{1}{4}=\frac{3}{2}$，$E(2\xi+4)=2E(\xi)+4=7$.

2.(2022·浙江)现有 7 张卡片，分别写上数字 1，2，2，3，4，5，6.从这 7 张卡片中随机抽取3 张，记所抽取卡片上数字的最小值为 $\xi$，则$P(\xi=2)=$________，$E(\xi)=$________.

$\frac{16}{35}$　$\frac{12}{7}$　**解析**：从写有数字 1，2，2，3，4，5，6 的 7 张卡片中任取 3 张，共有 $C_7^3$ 种取法，其中所抽取的卡片上的数字的最小值为 2 的取法有$(C_2^2C_4^1+C_2^1C_4^2)$种，所以 $P(\xi=2)=\frac{C_2^2C_4^1+C_2^1C_4^2}{C_7^3}=\frac{16}{35}$.

由已知可得 $\xi$ 的可能取值有 1，2，3，4，所以 $P(\xi=1)=\frac{C_6^2}{C_7^3}=\frac{15}{35}$，$P(\xi=2)=\frac{16}{35}$，$P(\xi=3)=\frac{C_3^2}{C_7^3}=\frac{3}{35}$，$P(\xi=4)=\frac{1}{C_7^3}=\frac{1}{35}$，所以 $E(\xi)=1\times\frac{15}{35}+2\times\frac{16}{35}+3\times\frac{3}{35}+4\times\frac{1}{35}=\frac{12}{7}$.

# 任务2 求离散型随机变量的均值

## 探究活动

**例 2**　某超市计划销售一种酸奶，根据往年销售经验，每天的需求量与当天的最高气温(单位：℃)有关.如果最高气温不低于 25，需求量为 500 瓶；如果最高气温位于区间[20，25)，需求量为 300 瓶；如果最高气温低于 20，需求量为 200 瓶.该超市统计了前三年六月份各天的最高气温数据，得到下面的频数分布表：

| 最高气温/℃ | [10，15) | [15，20) | [20，25) | [25，30) | [30，35) | [35，40) |
|---|---|---|---|---|---|---|
| 天数 | 2 | 16 | 36 | 25 | 7 | 4 |

以最高气温位于各区间的频率代替最高气温位于该区间的概率.

(1)设六月份这种酸奶一天的需求量为 $X$(单位：瓶)，求 $X$ 的数学期望；

(2)设六月份一天销售这种酸奶的利润为 $Y$(单位：元)，且 $Y=1.2X$，求 $Y$ 的数学期望.

**解**：(1)由题意知，$X$ 的所有可能取值为 200，300，500.

由题表可知 $P(X=200)=\frac{2+16}{90}=0.2$，

$P(X=300)=\frac{36}{90}=0.4$，$P(X=500)=\frac{25+7+4}{90}=0.4$.

因此 $X$ 的分布列为

| $X$ | 200 | 300 | 500 |
|---|---|---|---|
| $P$ | 0.2 | 0.4 | 0.4 |

$E(X)=200\times0.2+300\times0.4+500\times0.4=360$.

(2)由 $Y=1.2X$，得 $E(Y)=E(1.2X)=1.2E(X)=1.2\times360=432$.

**【探究总结】**

**求离散型随机变量的均值的步骤**

(1)确定取值：根据随机变量 $X$ 的意义，写出 $X$ 可能取得的全部值.

(2)求概率：求出 $X$ 取每个值的概率.

(3)写分布列：写出 $X$ 的分布列.

(4)求均值：由均值的定义 $E(X)=x_1p_1+x_2p_2+\cdots+x_np_n$ 求出 $E(X)$.

## 应用迁移

**1.** 甲、乙两人分别独立参加某高校招生面试.若甲、乙能通过面试的概率都是$\frac{1}{3}$，则面试结束后通过的人数$\xi$的数学期望是 （　　）

A.$\frac{2}{3}$　　B.$\frac{11}{9}$　　C.1　　D.$\frac{8}{9}$

A　**解析**：由题意知，$\xi$的可能取值为0，1，2，

且$P(\xi=0)=\left(1-\frac{1}{3}\right)\times\left(1-\frac{1}{3}\right)=\frac{4}{9}$，

$P(\xi=1)=\frac{1}{3}\times\left(1-\frac{1}{3}\right)+\left(1-\frac{1}{3}\right)\times\frac{1}{3}=\frac{4}{9}$，

$P(\xi=2)=\frac{1}{3}\times\frac{1}{3}=\frac{1}{9}$.

所以$\xi$的分布列为

| $\xi$ | 0 | 1 | 2 |
|---|---|---|---|
| $P$ | $\frac{4}{9}$ | $\frac{4}{9}$ | $\frac{1}{9}$ |

故$\xi$的数学期望$E(\xi)=0\times\frac{4}{9}+1\times\frac{4}{9}+2\times\frac{1}{9}=\frac{2}{3}$.

**2.** 某袋中装有大小相同、质地均匀的黑球和白球共5个.从袋中随机取出3个球，已知取出的3个球全为黑球的概率为$\frac{1}{10}$，若记取出的3个球中黑球的个数为$X$，则$E(X)=$________.

$\frac{9}{5}$　**解析**：依题意，设黑球的个数为$n$，由$\frac{C_n^3}{C_5^3}=\frac{1}{10}$，得$C_n^3=1$，则$n=3$.

记取出的3个球中黑球的个数为$X$，则$X$的取值可以为1，2，3，

$P(X=1)=\frac{C_2^2C_3^1}{C_5^3}=\frac{3}{10}$，$P(X=2)=\frac{C_2^1C_3^2}{C_5^3}=\frac{6}{10}=\frac{3}{5}$，$P(X=3)=\frac{C_2^0C_3^3}{C_5^3}=\frac{1}{10}$.

所以$X$的分布列为

| $X$ | 1 | 2 | 3 |
|---|---|---|---|
| $P$ | $\frac{3}{10}$ | $\frac{3}{5}$ | $\frac{1}{10}$ |

因此$E(X)=\frac{3}{10}\times1+\frac{3}{5}\times2+\frac{1}{10}\times3=\frac{9}{5}$.

**3.** 现有某品种杂交水稻，从中随机抽取15株作为样本进行观测，并记录每株水稻的生长周期(单位：天)，按从小到大排序结果如下：

93　97　98　101　103　104　107　108　109
110　112　116　121　124　126

已知这组样本数据的10%分位数、80%分位数分别为$a$，$b$.

(1)求$a$，$b$；

(2)在某科研任务中，把该品种所有生长周期位于区间$(a,b)$的稻株记为“高产稻株”，其余记为“低产稻株”.现从该品种水稻中随机抽取3株，设其中高产稻株有$X$株，求$X$的分布列与数学期望(以样本中高产稻株的频率作为该品种水稻的一株稻株属于高产稻株的概率).

**解**：(1)由题意知，样本数据的个数$n=15$，

因为$n\times10\%=1.5$，所以10%分位数为第2项数据，即$a=97$.

因为$n\times80\%=12$，所以80%分位数为第12项与第13项数据的平均数，即$b=\frac{116+121}{2}=118.5$.

(2)因为区间$(a,b)=(97,118.5)$，样本数据中共有10个数据位于该区间，

所以由题意得该品种水稻的一株稻株属于高产稻株的概率为$P=\frac{10}{15}=\frac{2}{3}$.

由题意知，随机变量$X$的所有可能取值为0，1，2，3，

$P(X=0)=\left(1-\frac{2}{3}\right)^3=\frac{1}{27}$，$P(X=1)=\frac{2}{3}\times\left(1-\frac{2}{3}\right)^2+\left(1-\frac{2}{3}\right)\times\frac{2}{3}\times\left(1-\frac{2}{3}\right)+\left(1-\frac{2}{3}\right)^2\times\frac{2}{3}=\frac{2}{9}$，

$P(X=2)=\left(\frac{2}{3}\right)^2\times\left(1-\frac{2}{3}\right)+\frac{2}{3}\times\left(1-\frac{2}{3}\right)\times\frac{2}{3}+\left(1-\frac{2}{3}\right)\times\left(\frac{2}{3}\right)^2=\frac{4}{9}$，$P(X=3)=\left(\frac{2}{3}\right)^3=\frac{8}{27}$，

所以$X$的分布列为

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P$ | $\frac{1}{27}$ | $\frac{2}{9}$ | $\frac{4}{9}$ | $\frac{8}{27}$ |

故$X$的数学期望$E(X)=0\times\frac{1}{27}+1\times\frac{2}{9}+2\times\frac{4}{9}+3\times\frac{8}{27}=2$.

## 任务3 两点分布的均值

### 探究活动

**例3**　(1)已知随机变量$X$满足$P(X=1)=0.3$，$P(X=0)=0.7$，则$E(X)$等于 （　　）

A.0.3　　B.0.7　　C.0.21　　D.1

A　**解析**：根据题意知随机变量 $X$ 服从两点分布，所以 $E(X)=0.3$.

(2)若离散型随机变量 $X$ 的分布列为

| $X$ | 0 | 1 |
|---|---|---|
| $P$ | $\frac{a}{2}$ | $\frac{a^2}{2}$ |

则 $X$ 的均值 $E(X)$ 等于　　(　　)

A.2　　B.2 或 $\frac{1}{2}$

C.$\frac{1}{2}$　　D.1

C　**解析**：由分布列的性质知，$\frac{a}{2}+\frac{a^2}{2}=1$，解得 $a=1$ 或 $a=-2$(舍去).

所以 $E(X)=0\times\frac{1}{2}+1\times\frac{1}{2}=\frac{1}{2}$.

**【探究总结】**

**两点分布的均值**

(1)若随机变量服从两点分布，则其均值为成功概率 $p$，即 $E(X)=0\times(1-p)+1\times p=p$.

(2)两点分布的均值直接反映事件成功的平均概率.

## 应用迁移

**1.**设随机变量 $X$ 服从两点分布，若 $P(X=1)-P(X=0)=0.4$，则 $E(X)=$　　(　　)

A.0.3　B.0.4　C.0.6　D.0.7

D　**解析**：由题意得 $P(X=1)+P(X=0)=1$，又因为 $P(X=1)-P(X=0)=0.4$，所以 $P(X=1)=0.7$，$P(X=0)=0.3$，

所以 $E(X)=1\times0.7+0\times0.3=0.7$.

**2.**已知离散型随机变量 $X$ 服从两点分布，且 $P(X=0)=3-4P(X=1)$，则随机变量 $Y=3X-1$ 的期望为________.

1　**解析**：因为随机变量 $X$ 服从两点分布，所以 $P(X=0)+P(X=1)=1$.

又 $P(X=0)=3-4P(X=1)$，得到 $P(X=0)=\frac{1}{3}$，$P(X=1)=\frac{2}{3}$，

所以 $E(X)=0\times\frac{1}{3}+1\times\frac{2}{3}=\frac{2}{3}$，故 $E(Y)=E(3X-1)=3E(X)-1=2-1=1$.

# 任务4 均值的实际应用

## 探究活动

**例4**　甲、乙两名射箭运动员射中目标箭靶的环数的分布列如表所示.

| 环数 $X$ | 7 | 8 | 9 | 10 |
|---|---|---|---|---|
| 甲射中的概率 | 0.1 | 0.2 | 0.3 | 0.4 |
| 乙射中的概率 | 0.15 | 0.25 | 0.4 | 0.2 |

(1)分别计算甲射中环数和乙射中环数的均值；

(2)从均值角度，比较甲、乙射箭水平的高低.

**解**：(1)甲射中环数的均值为

$7\times0.1+8\times0.2+9\times0.3+10\times0.4=9$.

乙射中环数的均值为

$7\times0.15+8\times0.25+9\times0.4+10\times0.2=8.65$.

(2)从均值角度看，$9>8.65$，故甲的射箭水平比乙高.

**【探究总结】**

**1.实际问题中的均值问题**

均值在实际生活中有着广泛的应用，如体育比赛的安排、成绩预测、消费预测、工程方案的预测、产品合格率的预测和投资收益的预测等，都可以通过随机变量的均值来进行估计.

**2.概率模型的解答步骤**

(1)审题，确定实际问题是哪一种概率模型，可能用到的事件类型，所用的公式有哪些；

(2)确定随机变量的分布列，计算随机变量的均值；

(3)对照实际意义，给出由概率、均值等所得出的结论.

## 应用迁移

**1.**受轿车在保修期内维修费用等因素的影响，企业生产每辆轿车的利润与该轿车首次出现故障的时间有关.某轿车制造厂生产甲、乙两种品牌轿车，保修期均为2年，现从该厂已售出的两种品牌轿车中各随机抽取50辆，统计数据如下：

| 品牌 | 甲 | | | 乙 | |
|---|---|---|---|---|---|
| 首次出现故障时间 $x$/年 | $0<x\leqslant1$ | $1<x\leqslant2$ | $x>2$ | $0<x\leqslant2$ | $x>2$ |
| 轿车数量/辆 | 2 | 3 | 45 | 5 | 45 |
| 每辆利润/万元 | 1 | 2 | 3 | 1.8 | 2.9 |

将频率视为概率，解答下列问题：

(1)从该厂生产的甲品牌轿车中随机抽取一辆，求其首次出现故障发生在保修期内的概率.

(2)若该厂生产的轿车均能售出，记生产一辆甲品牌轿车的利润为 $X_1$，生产一辆乙品牌轿车的利润为 $X_2$，分别求 $X_1$，$X_2$ 的分布列.

(3)该厂预计今后这两种品牌轿车销量相当，由于资金限制，只能生产其中一种品牌轿车，若从经济效益的角度考虑，你认为应该生产哪种品牌的轿车？说明理由.

**解**：(1)设“甲品牌轿车首次出现故障发生在保修期内”为事件 $A$，则 $P(A)=\frac{2+3}{50}=\frac{1}{10}$.

(2)依题意得，$X_1$ 的分布列为

| $X_1$ | 1 | 2 | 3 |
|---|---|---|---|
| $P$ | $\frac{1}{25}$ | $\frac{3}{50}$ | $\frac{9}{10}$ |

$X_2$ 的分布列为

| $X_2$ | 1.8 | 2.9 |
|---|---|---|
| $P$ | $\frac{1}{10}$ | $\frac{9}{10}$ |

(3)由(2)得 $E(X_1)=1\times\frac{1}{25}+2\times\frac{3}{50}+3\times\frac{9}{10}=2.86$(万元)，

$E(X_2)=1.8\times\frac{1}{10}+2.9\times\frac{9}{10}=2.79$(万元).

因为 $E(X_1)>E(X_2)$，所以应生产甲品牌轿车.

**2.** 某乡镇计划引进 A，B 两种矮化果树，已知 A 种矮化果树种植成功率为 $\frac{2}{3}$，成功后每公顷收益为 7.5 万元；B 种矮化果树种植成功率为 $\frac{3}{5}$，成功后每公顷收益为 9 万元.假设种植不成功时，种植 A，B 两种矮化果树每公顷均损失 1.5 万元，每公顷是否种植成功相互独立.

(1)甲种植户试种两种矮化果树各 1 公顷，总收益为 $X$ 万元，求 $X$ 的分布列及数学期望；

(2)乙种植户有良田 6 公顷，本计划全部种植 A 种矮化果树，但是甲劝说乙应该种植两种矮化果树各 3 公顷，请按照总收益的角度分析一下，乙应选择哪一种方案？

**解**：(1)由题意知，当 A，B 两种矮化果树均种植成功时，$X=9+7.5=16.5$，此时 $P(X=16.5)=\frac{3}{5}\times\frac{2}{3}=\frac{2}{5}$；

当 A 种矮化果树种植不成功，B 种矮化果树种植成功时，$X=9-1.5=7.5$，此时 $P(X=7.5)=\left(1-\frac{2}{3}\right)\times\frac{3}{5}=\frac{1}{5}$；

当 A 种矮化果树种植成功，B 种矮化果树种植不成功时，$X=7.5-1.5=6$，此时 $P(X=6)=\frac{2}{3}\times\left(1-\frac{3}{5}\right)=\frac{4}{15}$；

当 A，B 两种矮化果树均种植不成功时，$X=2\times(-1.5)=-3$，此时 $P(X=-3)=\left(1-\frac{2}{3}\right)\times\left(1-\frac{3}{5}\right)=\frac{2}{15}$.

所以 $X$ 的分布列为

| $X$ | 16.5 | 7.5 | 6 | $-3$ |
|---|---|---|---|---|
| $P$ | $\frac{2}{5}$ | $\frac{1}{5}$ | $\frac{4}{15}$ | $\frac{2}{15}$ |

数学期望为 $E(X)=16.5\times\frac{2}{5}+7.5\times\frac{1}{5}+6\times\frac{4}{15}+(-3)\times\frac{2}{15}=9.3$(万元).

(2)全种植 A 种矮化果树的收益为 $W_1=\frac{2}{3}\times6\times7.5-\frac{1}{3}\times6\times1.5=27$(万元)，

由(1)得，各种植 3 公顷的收益为 $W_2=3\times9.3=27.9$(万元).

因为 $W_2>W_1$，

所以乙应选择两种果树各种植 3 公顷.

## 提质归纳

离散型随机变量的均值
- 定义：$E(X)=x_1p_1+x_2p_2+\cdots+x_np_n=\sum\limits_{i=1}^{n}x_ip_i$
- 两点分布的均值：$E(X)=p$
- 性质：$E(aX+b)=aE(X)+b$
- 应用

# 课后素养评价（十二）　离散型随机变量的均值

## A组　学习·理解

1.下列说法正确的是　（　　）

A.随机变量 $X$ 的数学期望 $E(X)$ 是个变量，其随 $X$ 的变化而变化

B.随机变量的均值反映样本的平均水平

C.若随机变量 $X$ 的数学期望 $E(X)=2$，则 $E(2X)=4$

D.随机变量 $X$ 的均值 $E(X)=\dfrac{x_1+x_2+\cdots+x_n}{n}$

C　**解析**：A 错误，随机变量 $X$ 的数学期望 $E(X)$ 是个常量，是随机变量 $X$ 本身固有的一个数字特征.B 错误，随机变量的均值反映随机变量取值的平均水平.C 正确，由均值的性质可知.D 错误，$E(X)=x_1p_1+x_2p_2+\cdots+x_np_n$.

2.（多选）已知随机变量 $X$ 的分布列为

| $X$ | 4 | $a$ | 9 | 10 |
|---|---|---|---|---|
| $P$ | 0.3 | 0.1 | $b$ | 0.2 |

若 $E(X)=7.5$，则下列结论正确的是　（　　）

A.$a=7$　　B.$b=0.4$

C.$E(aX)=52.5$　　D.$E(X+b)=7.9$

ABCD　**解析**：由 $0.3+0.1+b+0.2=1$，得 $b=0.4$.因为 $E(X)=4\times0.3+a\times0.1+9\times0.4+10\times0.2=6.8+a\times0.1=7.5$，所以 $0.1a=0.7$，得 $a=7$.所以 $E(aX)=aE(X)=7\times7.5=52.5$，$E(X+b)=7.5+0.4=7.9$.

3.已知随机变量 $X$ 服从两点分布，满足 $P(X=0)=\dfrac{2}{9P(X=1)}$，且 $P(X=0)<P(X=1)$，则 $E(X)=$　（　　）

A.$\dfrac{1}{3}$　B.$\dfrac{1}{2}$　C.$\dfrac{2}{3}$　D.$\dfrac{1}{4}$

C　**解析**：因为随机变量 $X$ 的分布列服从两点分布，所以 $P(X=0)+P(X=1)=1$，

则 $P(X=1)+\dfrac{2}{9P(X=1)}=1$，解得 $P(X=1)=\dfrac{1}{3}$ 或 $P(X=1)=\dfrac{2}{3}$.又因为 $P(X=0)<P(X=1)$，

所以 $P(X=1)=\dfrac{2}{3}$，则 $P(X=0)=\dfrac{1}{3}$，所以 $E(X)=\dfrac{2}{3}$.故选 C.

4.有 10 件产品，其中 3 件是次品，从中任取 2 件.若 $\xi$ 表示取到次品的件数，则 $E(\xi)$ 等于　（　　）

A.$\dfrac{3}{5}$　B.$\dfrac{8}{15}$　C.$\dfrac{14}{15}$　D.1

A　**解析**：由题意知，$\xi$ 的所有可能取值为 0，1，2，$P(\xi=0)=\dfrac{C_7^2}{C_{10}^2}=\dfrac{7}{15}$，$P(\xi=1)=\dfrac{C_7^1C_3^1}{C_{10}^2}=\dfrac{7}{15}$，$P(\xi=2)=\dfrac{C_3^2}{C_{10}^2}=\dfrac{1}{15}$.

所以 $\xi$ 的分布列为

| $\xi$ | 0 | 1 | 2 |
|---|---|---|---|
| $P$ | $\dfrac{7}{15}$ | $\dfrac{7}{15}$ | $\dfrac{1}{15}$ |

所以 $E(\xi)=0\times\dfrac{7}{15}+1\times\dfrac{7}{15}+2\times\dfrac{1}{15}=\dfrac{3}{5}$.

5.某人进行一项试验，若试验成功，则停止试验；若试验失败，则重新试验一次；若试验 3 次均失败，则放弃试验.若此人每次试验成功的概率为 $\dfrac{2}{3}$，则此人试验次数 $\xi$ 的均值是　（　　）

A.$\dfrac{4}{3}$　B.$\dfrac{13}{9}$　C.$\dfrac{5}{3}$　D.$\dfrac{13}{7}$

B　**解析**：由题意知，试验次数 $\xi$ 的可能取值为 1，2，3，$P(\xi=1)=\dfrac{2}{3}$，$P(\xi=2)=\dfrac{1}{3}\times\dfrac{2}{3}=\dfrac{2}{9}$，$P(\xi=3)=\dfrac{1}{3}\times\dfrac{1}{3}\times\left(\dfrac{2}{3}+\dfrac{1}{3}\right)=\dfrac{1}{9}$.

所以 $\xi$ 的分布列为

| $\xi$ | 1 | 2 | 3 |
|---|---|---|---|
| $P$ | $\dfrac{2}{3}$ | $\dfrac{2}{9}$ | $\dfrac{1}{9}$ |

所以 $E(\xi)=1\times\dfrac{2}{3}+2\times\dfrac{2}{9}+3\times\dfrac{1}{9}=\dfrac{13}{9}$.

6.甲、乙两名射击运动员一次射击得分（分别用 $X_1$，$X_2$ 表示）的分布列如下：

| $X_1$ | 1 | 2 | 3 |
|---|---|---|---|
| $P$ | 0.4 | 0.1 | 0.5 |

| $X_2$ | 1 | 2 | 3 |
|---|---|---|---|
| $P$ | 0.1 | 0.6 | 0.3 |

则从得分的均值比较甲、乙两人的射击技术，下列结论正确的是　（　　）

A.甲更好　　B.乙更好

C.甲、乙一样好　　D.不可比较

B　**解析**：由题意得 $E(X_1)=1\times0.4+2\times0.1+3\times0.5=2.1$，$E(X_2)=1\times0.1+2\times0.6+3\times0.3=2.2$，所以$E(X_2)>E(X_1)$，故乙的射击技术更好.

**7.**某班举行了一次“心有灵犀”的活动，教师把一张写有成语的纸条出示给A组的某个同学，这个同学再用身体语言把成语的意思传递给本组其他同学.若小组内同学甲猜对成语的概率是0.4，同学乙猜对成语的概率是0.5，且规定猜对得1分，猜错得0分，则这两个同学各猜1次，得分之和 $X$ 的均值为（　　）

A.0.9　　B.0.8

C.1.2　　D.1.1

A　**解析**：由题意得，$X$ 的所有可能取值为0，1，2，

$P(X=0)=(1-0.4)\times(1-0.5)=0.3$，

$P(X=1)=0.4\times(1-0.5)+(1-0.4)\times0.5=0.5$，

$P(X=2)=0.4\times0.5=0.2$.

所以 $X$ 的分布列为

| $X$ | 0 | 1 | 2 |
|---|---|---|---|
| $P$ | 0.3 | 0.5 | 0.2 |

所以 $E(X)=0\times0.3+1\times0.5+2\times0.2=0.9$.

**8.**某城市有甲、乙、丙三个旅游景点，一位游客游览这三个景点的概率分别为0.4，0.5，0.6，且此人是否游览某个景点互不影响.设 $\xi$ 表示该客人离开该城市时游览的景点数与没有游览的景点数之差的绝对值，则 $E(\xi)=$________.

1.48　**解析**：由题意知，随机变量 $\xi$ 的所有可能取值为1，3，

$\xi=3$ 表示三个景点都游览了或都没有游览，

所以 $P(\xi=3)=0.4\times0.5\times0.6+0.6\times0.5\times0.4=0.24$，$P(\xi=1)=1-P(\xi=3)=0.76$，

所以随机变量 $\xi$ 的分布列为

| $\xi$ | 1 | 3 |
|---|---|---|
| $P$ | 0.76 | 0.24 |

所以 $E(\xi)=1\times0.76+3\times0.24=1.48$.

## B组　应用·实践

**1.**某射击运动员射击所得环数 $\xi$ 的分布列为

| $\xi$ | 7 | 8 | 9 | 10 |
|---|---|---|---|---|
| $P$ | $x$ | 0.1 | 0.3 | $y$ |

已知 $\xi$ 的期望 $E(\xi)=8.9$，则 $y$ 的值为（　　）

A.0.1　B.0.2　C.0.3　D.0.4

D　**解析**：由 $\begin{cases}x+0.1+0.3+y=1,\\7x+8\times0.1+9\times0.3+10y=8.9,\end{cases}$ 解得 $y=0.4$.

**2.**设离散型随机变量 $X$ 的可能取值为1，2，3，4，$P(X=k)=ak+b(k=1,2,3,4)$.又 $X$ 的均值 $E(X)=3$，则 $a+b=$（　　）

A.$\frac{1}{10}$　B.$\frac{1}{5}$　C.$\frac{3}{10}$　D.$\frac{2}{5}$

A　**解析**：因为 $P(X=1)=a+b$，$P(X=2)=2a+b$，$P(X=3)=3a+b$，$P(X=4)=4a+b$，

所以 $E(X)=1\times(a+b)+2\times(2a+b)+3\times(3a+b)+4\times(4a+b)=3$，所以 $30a+10b=3$①.

又因为$(a+b)+(2a+b)+(3a+b)+(4a+b)=1$，所以 $10a+4b=1$②.

由①②可得 $a=\frac{1}{10}$，$b=0$，所以 $a+b=\frac{1}{10}$.

**3.**家住福田区、罗湖区、盐田区、南山区的4位志愿者被随机派到福田区、罗湖区、盐田区、南山区这四个区工作，每人只去一个区，每个人去的区均不相同.记 $\xi$ 为4人中没有去到自家所在区工作的人数，则 $E(\xi)=$（　　）

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

D　**解析**：由题意得 $\xi$ 的取值不可能为1，可能的取值为4，3，2，0，

$P(\xi=4)=\frac{3\times(1+2)}{A_4^4}=\frac{9}{24}$，$P(\xi=3)=\frac{C_4^3\times2}{A_4^4}=\frac{8}{24}$，$P(\xi=2)=\frac{C_4^2\times1}{A_4^4}=\frac{6}{24}$，$P(\xi=0)=\frac{1}{A_4^4}=\frac{1}{24}$.

故 $E(\xi)=4\times\frac{9}{24}+3\times\frac{8}{24}+2\times\frac{6}{24}+0\times\frac{1}{24}=3$.故选D.

**4.**现有两台独立工作的雷达，这两台雷达发现飞行目标的概率分别为0.9和0.85.设发现飞行目标的雷达台数为 $\xi$，则 $E(\xi)=$（　　）

A.0.765　　B.1.75

C.1.765　　D.0.22

B　**解析**：设事件 $A$，$B$ 分别表示这两台雷达发现飞行目标，且 $A$，$B$ 相互独立，$\xi$ 的可能取值为0，1，2，

$P(\xi=0)=P(\overline{A}\,\overline{B})=P(\overline{A})P(\overline{B})=(1-0.9)\times(1-0.85)=0.015$，

$P(\xi=1)=P(A\overline{B})+P(\overline{A}B)=P(A)P(\overline{B})+P(\overline{A})P(B)=0.9\times0.15+0.1\times0.85=0.22$，

$P(\xi=2)=P(AB)=P(A)P(B)=0.9\times0.85=$

0.765.

所以 $E(\xi)=0\times0.015+1\times0.22+2\times0.765=1.75$.

**5.** 某天 A,B 两个沿海城市受台风袭击(相互独立)的概率相同,已知 A 城市或 B 城市受台风袭击的概率为0.36.若用 $X$ 表示这一天 A,B 两个城市中受台风袭击的城市个数,则 $E(X)=$________.

0.4　**解析**:设 A,B 两城市受台风袭击的概率均为 $p$,则 A 城市和 B 城市均不受台风袭击的概率为$(1-p)^2=1-0.36$,解得 $p=0.2$ 或 $p=1.8$(舍去).则 $P(X=0)=1-0.36=0.64$,$P(X=1)=2\times0.8\times0.2=0.32$,$P(X=2)=0.2\times0.2=0.04$,所以 $E(X)=0\times0.64+1\times0.32+2\times0.04=0.4$.

**6.** 某种考试规定:每位考试者一年之内最多有 4 次参加考试的机会,一旦某次考试通过,即可领取证书,不再参加以后的考试,否则就一直考到第 4 次为止.如果某人决定参加考试,设他每次参加考试通过的概率依次为 0.6,0.7,0.8,0.9,求在一年内此人参加考试的次数 $X$ 的分布列和均值.

**解**:由题意可知,$X$ 的所有可能取值为 1,2,3,4,

$X=1$ 表示此人第一次参加考试就通过了,

故 $P(X=1)=0.6$.

$X=2$ 表示此人第一次考试未通过,第二次考试通过了,

故 $P(X=2)=(1-0.6)\times0.7=0.28$.

$X=3$ 表示此人第一、二次考试均未通过,第三次考试通过了,

故 $P(X=3)=(1-0.6)\times(1-0.7)\times0.8=0.096$.

$X=4$ 表示此人第一、二、三次考试都未通过,

故 $P(X=4)=(1-0.6)\times(1-0.7)\times(1-0.8)=0.024$.

所以 $X$ 的分布列为

| $X$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $P$ | 0.6 | 0.28 | 0.096 | 0.024 |

所以 $E(X)=1\times0.6+2\times0.28+3\times0.096+4\times0.024=1.544$.

**7.** 飞行棋是一种竞技游戏,玩家在图纸上按线路行棋,通过掷骰子决定行棋步数.为增加游戏乐趣,往往在线路格子中设置一些“前进”“后退”等奖惩环节,当骰子点数大于或等于到达终点的格数时,玩家顺利通关.已知甲、乙两人的棋子已接近终点,位置如图所示.


(1)求乙还需抛掷 2 次骰子才顺利通关的概率.

(2)若甲、乙每人最多再投掷 3 次骰子,且第 3 次无论是否通关,该玩家游戏结束.设甲、乙两人再投掷骰子的次数分别为 $X$,$Y$,求 $X$,$Y$ 的分布列和期望.

**解**:(1)因为乙还需抛掷 2 次骰子才顺利通关,则第一次不能通关,所以第一次只能掷 1,2,3.

若第一次为 1 或 2,则第二次为 2,3,4,5,6;若第一次为 3,则第二次为 4,5,6.

所以乙还需抛掷 2 次骰子才顺利通关的概率为 $2\times\frac{1}{6}\times\frac{5}{6}+\frac{1}{6}\times\frac{3}{6}=\frac{13}{36}$.

(2)$X$,$Y$ 的可能取值均为 1,2,3,

由(1)得,$P(Y=1)=\frac{3}{6}=\frac{1}{2}$,$P(Y=2)=\frac{13}{36}$,

$P(Y=3)=1-P(Y=1)-P(Y=2)=\frac{5}{36}$,

所以 $Y$ 的分布列为

| $Y$ | 1 | 2 | 3 |
|---|---|---|---|
| $P$ | $\frac{1}{2}$ | $\frac{13}{36}$ | $\frac{5}{36}$ |

$E(Y)=1\times\frac{1}{2}+2\times\frac{13}{36}+3\times\frac{5}{36}=\frac{59}{36}$.

若甲还需抛掷 2 次通关,当第一次为 1 时,第二次为 4,5,6;

当第一次为 2 或 3 时,第二次为 2,3,4,5,6;

当第一次为 4 时,第二次为 4,5,6.

故 $P(X=2)=\frac{1}{6}\times\frac{3}{6}+2\times\frac{1}{6}\times\frac{5}{6}+\frac{1}{6}\times\frac{3}{6}=\frac{16}{36}=\frac{4}{9}$,

易得 $P(X=1)=\frac{2}{6}=\frac{1}{3}$,

$P(X=3)=1-P(X=1)-P(X=2)=\frac{2}{9}$,

所以 $X$ 的分布列为

| $X$ | 1 | 2 | 3 |
|---|---|---|---|
| $P$ | $\frac{1}{3}$ | $\frac{4}{9}$ | $\frac{2}{9}$ |

$E(X)=1\times\frac{1}{3}+2\times\frac{4}{9}+3\times\frac{2}{9}=\frac{17}{9}$.

# 7.3.2　离散型随机变量的方差

**学习任务目标**

**1.**理解离散型随机变量的方差及标准差的概念.
**2.**掌握方差的性质,会利用公式求离散型随机变量的方差.
**3.**会利用离散型随机变量的方差解决一些实际问题.

## ◦ 问题式预习 ◦

### 【知识清单】

**知识点　方差的定义及性质**

(1)设离散型随机变量 $X$ 的分布列为

| $X$ | $x_1$ | $x_2$ | … | $x_n$ |
|---|---|---|---|---|
| $P$ | $p_1$ | $p_2$ | … | $p_n$ |

我们称 $D(X)=\underline{(x_1-E(X))^2p_1+(x_2-E(X))^2p_2+\cdots+(x_n-E(X))^2p_n}=\underline{\sum_{i=1}^{n}(x_i-E(X))^2p_i}$ 为随机变量 $X$ 的方差,
并称 $\sqrt{D(X)}$ 为随机变量 $X$ 的标准差,记为 $\sigma(X)$.
(2)方差的性质:$D(X+b)=\underline{D(X)}$;$D(aX)=\underline{a^2D(X)}$;$D(aX+b)=\underline{a^2D(X)}$.
(3)两点分布的方差:$D(X)=\underline{p(1-p)}$.

### 【概念辨析】

**1.**判断正误(正确的打“√”,错误的打“×”).
(1)离散型随机变量的方差越大,随机变量越稳定.　( × )
(2)若 $a$ 是常数,则 $D(a)=0$.　( √ )
(3)离散型随机变量的方差反映了随机变量取值偏离于均值的平均程度.　( √ )
(4)若 $a,b$ 为常数,则 $\sqrt{D(aX+b)}=a\sqrt{D(X)}$.　( × )

**2.**已知随机变量 $X$ 的分布列为

| $X$ | $-1$ | 0 | 1 |
|---|---|---|---|
| $P$ | 0.5 | 0.3 | 0.2 |

则 $D(X)$ 等于　(　　)
A.0.7　B.0.61　C.$-0.3$　D.0

B　**解析**:$E(X)=(-1)\times0.5+0\times0.3+1\times0.2=-0.3$,$D(X)=(-1+0.3)^2\times0.5+(0+0.3)^2\times0.3+(1+0.3)^2\times0.2=0.61$.

**3.**请思考并回答下列问题:
(1)样本方差的计算公式有何意义?
**提示**:样本方差是刻画数据偏离样本均值程度的指标.样本方差越大,说明样本数据偏离样本均值的程度越大;样本方差越小,说明样本数据越集中于样本均值的附近.
(2)随机变量的方差与样本的方差有何关系?
**提示**:随机变量的方差即为总体的方差,它是一个客观存在的常数,样本的方差随着样本容量的不同而不同.对于简单随机抽样,随着样本容量的增大,样本的方差越来越接近总体的方差,即越来越接近随机变量的方差.
(3)如何证明方差的性质:$D(aX+b)=a^2D(X)$?
**提示**:设离散型随机变量 $X$ 的分布列为

| $X$ | $x_1$ | $x_2$ | … | $x_i$ | … | $x_n$ |
|---|---|---|---|---|---|---|
| $P$ | $p_1$ | $p_2$ | … | $p_i$ | … | $p_n$ |

由 $Y=aX+b$($a,b$ 为常数)知 $Y$ 也是离散型随机变量,
且 $Y$ 的分布列为

| $Y$ | $ax_1+b$ | $ax_2+b$ | … | $ax_i+b$ | … | $ax_n+b$ |
|---|---|---|---|---|---|---|
| $P$ | $p_1$ | $p_2$ | … | $p_i$ | … | $p_n$ |

由均值的性质得 $E(Y)=aE(X)+b$,于是
$$D(Y)=D(aX+b)$$
$$=\sum_{i=1}^{n}(ax_i+b-E(Y))^2p_i$$
$$=\sum_{i=1}^{n}(ax_i+b-aE(X)-b)^2p_i$$
$$=\sum_{i=1}^{n}(ax_i-aE(X))^2p_i$$
$$=a^2\sum_{i=1}^{n}(x_i-E(X))^2p_i$$
$$=a^2D(X).$$

# ◦ 任务型课堂 ◦

## 任务1 离散型随机变量的方差及性质

### 探究活动

**例 1**　袋中有大小相同的三个球,编号分别为 1,2,3,从袋中每次不放回地任取一个球.若取到的球的编号为奇数,则停止取球,用 $X$ 表示所有被取到的球的编号之和,求 $X$ 的方差.

**解**:由题意可知,$X$ 的所有可能取值为 1,3,5,$P(X=1)=\frac{1}{3}$,$P(X=3)=\frac{1}{3}+\frac{1}{3}\times\frac{1}{2}=\frac{1}{2}$,$P(X=5)=\frac{1}{3}\times\frac{1}{2}=\frac{1}{6}$,所以 $X$ 的分布列为

| $X$ | 1 | 3 | 5 |
|---|---|---|---|
| $P$ | $\frac{1}{3}$ | $\frac{1}{2}$ | $\frac{1}{6}$ |

故 $E(X)=1\times\frac{1}{3}+3\times\frac{1}{2}+5\times\frac{1}{6}=\frac{8}{3}$,

$D(X)=\left(1-\frac{8}{3}\right)^2\times\frac{1}{3}+\left(3-\frac{8}{3}\right)^2\times\frac{1}{2}+\left(5-\frac{8}{3}\right)^2\times\frac{1}{6}=\frac{17}{9}$.

[一题多思]

**思考**.将"若取到的球的编号为奇数,则停止取球,用 $X$ 表示所有被取到的球的编号之和"改为"若取到的球的编号为偶数,则停止取球,用 $X$ 表示取到的球的个数",求 $X$ 的方差.

**解**:由题意知,$X$ 的可能取值为 1,2,3,

$P(X=1)=\frac{A_1^1}{A_3^1}=\frac{1}{3}$,$P(X=2)=\frac{A_2^1A_1^1}{A_3^2}=\frac{1}{3}$,$P(X=3)=\frac{A_2^2A_1^1}{A_3^3}=\frac{1}{3}$.

则 $X$ 的分布列为

| $X$ | 1 | 2 | 3 |
|---|---|---|---|
| $P$ | $\frac{1}{3}$ | $\frac{1}{3}$ | $\frac{1}{3}$ |

所以 $E(X)=1\times\frac{1}{3}+2\times\frac{1}{3}+3\times\frac{1}{3}=2$,

$D(X)=(1-2)^2\times\frac{1}{3}+(2-2)^2\times\frac{1}{3}+(3-2)^2\times\frac{1}{3}=\frac{2}{3}$.

**【探究总结】**

**1.** 当分布列已知时,先利用定义求出均值 $E(X)$,再由公式 $D(X)=\sum\limits_{i=1}^{n}(x_i-E(X))^2p_i$ 求出方差.另外注意方差性质的应用,即 $D(aX+b)=a^2D(X)$.

**2.** 当分布列未知时,先根据已知条件和概率知识列出分布列,再根据上述方法来求解.

### 应用迁移

**1.**(多选)已知随机变量 $X,Y$,且 $Y=3X+1$,$X$ 的分布列如下:

| $X$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $P$ | $m$ | $\frac{1}{10}$ | $\frac{1}{5}$ | $n$ | $\frac{3}{10}$ |

若 $E(Y)=10$,则　　(　　)

A.$m=\frac{3}{10}$　　B.$n=\frac{1}{5}$

C.$E(X)=3$　　D.$D(Y)=\frac{7}{3}$

AC　**解析**:由 $m+\frac{1}{10}+\frac{1}{5}+n+\frac{3}{10}=1$,可得 $m+n=\frac{2}{5}$①,

又因为 $E(Y)=E(3X+1)=3E(X)+1=10$,解得 $E(X)=3$,

所以 $E(X)=m+2\times\frac{1}{10}+3\times\frac{1}{5}+4n+5\times\frac{3}{10}=3$,

则 $m+4n=\frac{7}{10}$②,所以由①②可得 $n=\frac{1}{10}$,$m=\frac{3}{10}$,

故 A 正确,B 错误,C 正确;

$D(X)=(1-3)^2\times\frac{3}{10}+(2-3)^2\times\frac{1}{10}+(3-3)^2\times\frac{1}{5}+(4-3)^2\times\frac{1}{10}+(5-3)^2\times\frac{3}{10}$

$=4\times\frac{3}{10}+1\times\frac{1}{10}+1\times\frac{1}{10}+4\times\frac{3}{10}=\frac{13}{5}$,

$D(Y)=D(3X+1)=9D(X)=9\times\frac{13}{5}=\frac{117}{5}$,故 D 错误.

故选 AC.

**2.** 某袋中装有除颜色外完全相同的 6 个球,其中 4 个黑球和 2 个白球.从袋中随机取出 2 个球,记取出白球的个数为 $X$.

(1)写出 $X$ 的分布列,并求出 $E(X)$ 和 $D(X)$ 的值;

(2)若取出一个白球得 1 分,取出一个黑球得 2 分,总得分为 $Z$,求出 $E(Z)$ 和 $D(Z)$ 的值.

**解**:(1)由题意,得 $X$ 的取值可能为 0,1,2,

$P(X=0)=\frac{C_4^2}{C_6^2}=\frac{2}{5}$,$P(X=1)=\frac{C_4^1C_2^1}{C_6^2}=\frac{8}{15}$,$P(X$

$=2)=\frac{C_2^2}{C_6^2}=\frac{1}{15}$，所以随机变量 $X$ 的分布列为

| $X$ | 0 | 1 | 2 |
|---|---|---|---|
| $P$ | $\frac{2}{5}$ | $\frac{8}{15}$ | $\frac{1}{15}$ |

$E(X)=0\times\frac{2}{5}+1\times\frac{8}{15}+2\times\frac{1}{15}=\frac{2}{3}$，

$D(X)=\left(0-\frac{2}{3}\right)^2\times\frac{2}{5}+\left(1-\frac{2}{3}\right)^2\times\frac{8}{15}+\left(2-\frac{2}{3}\right)^2\times\frac{1}{15}=\frac{16}{45}$.

(2)依题意，得 $Z=X+2(2-X)=4-X$，

则 $E(Z)=4-E(X)=4-\frac{2}{3}=\frac{10}{3}$，$D(Z)=D(X)=\frac{16}{45}$.

## 任务2 方差的实际应用

### 探究活动

**例2** 甲、乙两名工人加工同一种零件，两人每天加工的零件数相同，所得次品数分别为 $X$，$Y$，$X$ 和 $Y$ 的分布列如表所示.

| $X$ | 0 | 1 | 2 |
|---|---|---|---|
| $P$ | $\frac{3}{5}$ | $\frac{1}{10}$ | $\frac{3}{10}$ |

| $Y$ | 0 | 1 | 2 |
|---|---|---|---|
| $P$ | $\frac{1}{2}$ | $\frac{3}{10}$ | $\frac{1}{5}$ |

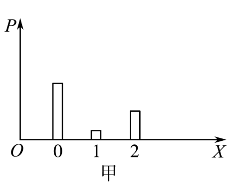
(1)试通过均值对这两名工人的技术水平进行比较；
(2)甲、乙两名工人技术的稳定性如何？

**解**：(1)$E(X)=0\times\frac{3}{5}+1\times\frac{1}{10}+2\times\frac{3}{10}=0.7$，

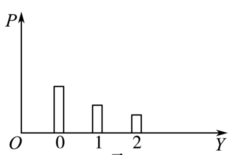
$E(Y)=0\times\frac{1}{2}+1\times\frac{3}{10}+2\times\frac{1}{5}=0.7$.

由 $E(X)=E(Y)$ 知，两人出次品的均值相同，技术水平相当.

(2)(方法一)作出两名工人加工所得次品数的概率分布图如图，比较发现乙工人的技术更稳定.



甲



乙

(方法二)可通过方差来刻画两名工人技术的稳定性.计算可得 $D(X)=0.81$，$D(Y)=0.61$，$D(Y)<D(X)$，所以乙工人的技术更稳定.

**【探究总结】**

**利用均值和方差的意义分析解决实际问题的步骤**

(1)比较均值.离散型随机变量的均值反映了离散型随机变量取值的平均水平，因此，在实际决策问题中，需先计算均值，看一下谁的平均水平高.

(2)在均值相等的情况下计算方差.方差反映了离散型随机变量取值的稳定与波动、集中与离散的程度.通过计算方差，分析一下谁的发挥相对稳定.

(3)下结论.依据均值与方差的意义得出结论.

### 应用迁移

**1.**(多选)投资甲、乙两种股票，每股收益的分布列分别如表1和表2所示.

表1 股票甲收益的分布列

| 收益 $X$/元 | $-1$ | 0 | 2 |
|---|---|---|---|
| 概率 | 0.1 | 0.3 | 0.6 |

表2 股票乙收益的分布列

| 收益 $Y$/元 | 0 | 1 | 2 |
|---|---|---|---|
| 概率 | 0.3 | 0.4 | 0.3 |

关于两种股票，下列结论正确的是（　　）

A.$E(2X+1)=3.2$

B.$D(2Y+1)=2.2$

C.投资股票甲的期望收益较大

D.投资股票甲比投资股票乙风险高

ACD **解析**：$E(X)=-0.1+1.2=1.1$，$E(Y)=0.4+0.6=1$，$E(X)>E(Y)$，

$D(X)=(-1-1.1)^2\times0.1+(0-1.1)^2\times0.3+(2-1.1)^2\times0.6=1.29$，

$D(Y)=(0-1)^2\times0.3+(1-1)^2\times0.4+(2-1)^2\times0.3=0.6$，$D(X)>D(Y)$，

则投资股票甲的期望收益较大，投资股票甲比投资股票乙风险高.

$E(2X+1)=2E(X)+1=3.2$，$D(2Y+1)=4D(Y)=2.4$.故选ACD.

**2.**某地区拟建立一个艺术博物馆，采取竞标的方式从多家建筑公司中选取一家，经过层层筛选，甲、乙两家建筑公司进入最后的招标环节.现从建筑设计院聘请专家设计了一个招标方案：两家公司从6个招标问题中随机抽取3个问题，已知这6个招标问题中，甲公司可正确回答其中4个问题，而乙公司能正确回答每个问题的概率均为 $\frac{2}{3}$，甲、乙两家公司对每个问题的回答都是相互独立，互不影响的.

(1)求甲、乙两家公司共答对2个问题的概率；

(2)设甲公司答对题数为随机变量 $X$,求 $X$ 的分布列、数学期望和方差；

(3)请从期望和方差的角度分析,哪家公司竞标成功的可能性更大?

**解**:(1)记“甲、乙两家公司共答对 2 个问题”为事件 $A$,由题意可知,事件 $A$ 是甲、乙各答对 1 个问题与甲答对 2 题乙没答对题目的和事件,它们互斥,

则有 $P(A)=\frac{C_4^1C_2^2}{C_6^3}\times3\times\frac{2}{3}\times\left(1-\frac{2}{3}\right)^2+\frac{C_4^2C_2^1}{C_6^3}\times\left(1-\frac{2}{3}\right)^3=\frac{1}{15}$,

所以甲、乙两家公司共答对 2 个问题的概率是 $\frac{1}{15}$.

(2)设甲公司答对题目数为 $X$,则 $X$ 的取值可能为 1,2,3,

$P(X=1)=\frac{C_4^1C_2^2}{C_6^3}=\frac{1}{5}$,$P(X=2)=\frac{C_4^2C_2^1}{C_6^3}=\frac{3}{5}$,

$P(X=3)=\frac{C_4^3C_2^0}{C_6^3}=\frac{1}{5}$,

则 $X$ 的分布列为

| $X$ | 1 | 2 | 3 |
|---|---|---|---|
| $P$ | $\frac{1}{5}$ | $\frac{3}{5}$ | $\frac{1}{5}$ |

期望 $E(X)=1\times\frac{1}{5}+2\times\frac{3}{5}+3\times\frac{1}{5}=2$,方差 $D(X)=(1-2)^2\times\frac{1}{5}+(2-2)^2\times\frac{3}{5}+(3-2)^2\times\frac{1}{5}=\frac{2}{5}$.

(3)设乙公司答对题目数为 $Y$,则 $Y$ 的取值可能为 0,1,2,3,

$P(Y=0)=\left(\frac{1}{3}\right)^3=\frac{1}{27}$,$P(Y=1)=3\times\frac{2}{3}\times\left(\frac{1}{3}\right)^2=\frac{2}{9}$,

$P(Y=2)=3\times\left(\frac{2}{3}\right)^2\times\frac{1}{3}=\frac{4}{9}$,$P(Y=3)=\left(\frac{2}{3}\right)^3=\frac{8}{27}$,

则 $Y$ 的分布列为

| $Y$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P$ | $\frac{1}{27}$ | $\frac{2}{9}$ | $\frac{4}{9}$ | $\frac{8}{27}$ |

期望 $E(Y)=0\times\frac{1}{27}+1\times\frac{2}{9}+2\times\frac{4}{9}+3\times\frac{8}{27}=2$,

方差 $D(Y)=(0-2)^2\times\frac{1}{27}+(1-2)^2\times\frac{2}{9}+(2-2)^2\times\frac{4}{9}+(3-2)^2\times\frac{8}{27}=\frac{2}{3}$.

显然 $E(X)=E(Y)$,$D(X)<D(Y)$,

所以甲公司竞标成功的可能性更大.

### 提质归纳

- 离散型随机变量的方差
  - 方差 $D(X)$
    - 定义：$D(X)=\sum_{i=1}^{n}(x_i-E(X))^2p_i=\sum_{i=1}^{n}x_i^2p_i-(E(X))^2$
    - 性质：$D(aX+b)=a^2D(X)$
  - 标准差：$\sigma(X)=\sqrt{D(X)}$
  - 应用

## 课后素养评价(十三)　离散型随机变量的方差

### A组　学习·理解

**1.**已知随机变量 $X$ 的分布列为

| $X$ | $-1$ | 0 | 1 |
|---|---|---|---|
| $P$ | $a$ | $b$ | $\frac{1}{2}$ |

若 $E(X)=\frac{1}{3}$,则 $D(X)$的值是　(　　)

A.$\frac{1}{3}$　　B.$\frac{2}{3}$　　C.$\frac{5}{9}$　　D.$\frac{7}{9}$

C　**解析**:由分布列的性质可知 $a+b+\frac{1}{2}=1$,

所以 $a+b=\frac{1}{2}$.

又由 $E(X)=-a+\frac{1}{2}=\frac{1}{3}$,解得 $a=\frac{1}{6}$,所以 $b=\frac{1}{3}$.

所以 $D(X)=\left(-1-\frac{1}{3}\right)^2\times\frac{1}{6}+\left(0-\frac{1}{3}\right)^2\times\frac{1}{3}+\left(1-\frac{1}{3}\right)^2\times\frac{1}{2}=\frac{5}{9}$.

**2.**若随机变量 $\xi$ 的分布列如表所示,则$D(1-3\xi)=$　(　　)

| $\xi$ | $-1$ | 0 | 1 |
|---|---|---|---|
| $P$ | $\frac{1}{3}a$ | $\frac{1}{3}$ | $a^2$ |

A.$\frac{50}{27}$　　B.2　　C.$\frac{41}{9}$　　D.$\frac{50}{9}$

D　**解析**：由已知可得$\frac{1}{3}a+\frac{1}{3}+a^2=1$，$0\leqslant\frac{1}{3}a\leqslant1$，$0\leqslant a^2\leqslant1$，所以$a=\frac{2}{3}$.

所以$E(\xi)=-1\times\frac{2}{9}+0\times\frac{1}{3}+1\times\frac{4}{9}=\frac{2}{9}$，

所以$D(\xi)=\left(-1-\frac{2}{9}\right)^2\times\frac{2}{9}+\left(0-\frac{2}{9}\right)^2\times\frac{1}{3}+\left(1-\frac{2}{9}\right)^2\times\frac{4}{9}=\frac{50}{81}$，

所以$D(1-3\xi)=9D(\xi)=\frac{50}{9}$.

**3.** 如果$X$是离散型随机变量，$E(X)=6$，$D(X)=0.5$，$X_1=2X-5$，那么$E(X_1)$和$D(X_1)$的值分别是（　　）

A.$E(X_1)=12$，$D(X_1)=1$

B.$E(X_1)=7$，$D(X_1)=1$

C.$E(X_1)=12$，$D(X_1)=2$

D.$E(X_1)=7$，$D(X_1)=2$

D　**解析**：$E(X_1)=2E(X)-5=12-5=7$，$D(X_1)=4D(X)=4\times0.5=2$.

**4.** 一道试题，同学甲解出的概率为$\frac{2}{3}$，同学乙解出的概率为$\frac{4}{5}$.设两人中解出该题的人数为$X$，则$D(X)$等于（　　）

A.$\frac{22}{15}$　　B.$\frac{86}{225}$

C.$\frac{225}{484}$　　D.$\frac{225}{85}$

B　**解析**：由题意知，$X$的可能取值为0，1，2，

且$P(X=0)=\frac{1}{3}\times\frac{1}{5}=\frac{1}{15}$，$P(X=1)=\frac{2}{3}\times\frac{1}{5}+\frac{1}{3}\times\frac{4}{5}=\frac{6}{15}=\frac{2}{5}$，$P(X=2)=\frac{2}{3}\times\frac{4}{5}=\frac{8}{15}$.

所以$X$的分布列为

| $X$ | 0 | 1 | 2 |
| --- | --- | --- | --- |
| $P$ | $\frac{1}{15}$ | $\frac{2}{5}$ | $\frac{8}{15}$ |

所以$E(X)=0\times\frac{1}{15}+1\times\frac{2}{5}+2\times\frac{8}{15}=\frac{22}{15}$，

$D(X)=\frac{1}{15}\times\left(0-\frac{22}{15}\right)^2+\frac{2}{5}\times\left(1-\frac{22}{15}\right)^2+\frac{8}{15}\times\left(2-\frac{22}{15}\right)^2=\frac{86}{225}$.

**5.** 设随机变量$X$的分布列如下（其中$0<p<1$），$D(X)$表示$X$的方差，则当$p$从0增大到1时，$D(X)$的变化情况为（　　）

| $X$ | 0 | 1 | 2 |
| --- | --- | --- | --- |
| $P$ | $\frac{1-p}{2}$ | $\frac{1}{2}$ | $\frac{p}{2}$ |

A.$D(X)$增大

B.$D(X)$减小

C.$D(X)$先减小后增大

D.$D(X)$先增大后减小

D　**解析**：由题意可得$E(X)=0\times\frac{1-p}{2}+1\times\frac{1}{2}+2\times\frac{p}{2}=\frac{1}{2}+p$，

则$D(X)=\frac{1-p}{2}\left(\frac{1}{2}+p\right)^2+\frac{1}{2}\left(\frac{1}{2}+p-1\right)^2+\frac{p}{2}\left(\frac{1}{2}+p-2\right)^2=-p^2+p+\frac{1}{4}=-\left(p-\frac{1}{2}\right)^2+\frac{1}{2}$，

因为$0<p<1$，所以$D(X)$先增大后减小.

故选D.

**6.**（多选）编号为1，2，3的三名学生随意入座编号为1，2，3的三个座位，每名学生坐一个座位，设坐在与自己编号相同的座位上的学生的人数是$\xi$，则（　　）

A.$\xi$的所有可能取值是1，2，3　　B.$P(\xi=1)=\frac{1}{2}$

C.$E(\xi)=1$　　D.$D(\xi)=1$

BCD　**解析**：由题意知，$\xi$的所有可能取值为0，1，3，$\xi=0$表示三名学生全坐在与自己编号不同的座位上，有2种情况，即编号为1，2，3的座位上分别坐了编号为2，3，1或3，1，2的学生，则$P(\xi=0)=\frac{2}{A_3^3}=\frac{1}{3}$；$\xi=1$表示三名学生只有一名学生坐在了与自己编号相同的座位上，则$P(\xi=1)=\frac{C_3^1}{A_3^3}=\frac{1}{2}$；$\xi=3$表示三名同学全坐在了与自己编号相同的座位上，即对号入座，则$P(\xi=3)=\frac{1}{A_3^3}=\frac{1}{6}$.

所以$\xi$的分布列为

| $\xi$ | 0 | 1 | 3 |
| --- | --- | --- | --- |
| $P$ | $\frac{1}{3}$ | $\frac{1}{2}$ | $\frac{1}{6}$ |

$E(\xi)=0\times\frac{1}{3}+1\times\frac{1}{2}+3\times\frac{1}{6}=1$，

$D(\xi)=\frac{1}{3}\times(0-1)^2+\frac{1}{2}\times(1-1)^2+\frac{1}{6}\times(3-1)^2=1$.

## B组　应用·实践

1.(多选)已知随机变量 $\xi$ 的分布列如表所示,且满足 $E(\xi)=0$,则下列结论正确的是　　(　　)

| $\xi$ | $-1$ | $0$ | $2$ |
|---|---|---|---|
| $P$ | $a$ | $\frac{1}{2}$ | $b$ |

A.$D(\xi)=1$　　B.$D(|\xi|)=1$

C.$D(2\xi+1)=4$　　D.$D(3|\xi|-2)=6$

AC　**解析**:由题意得 $\begin{cases}a+b+\frac{1}{2}=1,\\-1\times a+0\times\frac{1}{2}+2\times b=0,\end{cases}$

解得 $\begin{cases}a=\frac{1}{3},\\b=\frac{1}{6},\end{cases}$

所以 $\xi$ 的分布列为

| $\xi$ | $-1$ | $0$ | $2$ |
|---|---|---|---|
| $P$ | $\frac{1}{3}$ | $\frac{1}{2}$ | $\frac{1}{6}$ |

则 $D(\xi)=\frac{1}{3}\times(-1-0)^2+\frac{1}{2}\times(0-0)^2+\frac{1}{6}\times(2-0)^2=1$,故A正确;

则 $D(2\xi+1)=2^2D(\xi)=4$,故C正确;

所以 $|\xi|$ 的分布列为

| $\|\xi\|$ | $1$ | $0$ | $2$ |
|---|---|---|---|
| $P$ | $\frac{1}{3}$ | $\frac{1}{2}$ | $\frac{1}{6}$ |

则 $E(|\xi|)=1\times\frac{1}{3}+2\times\frac{1}{6}=\frac{2}{3}$,

$D(|\xi|)=\frac{1}{3}\times\left(1-\frac{2}{3}\right)^2+\frac{1}{2}\times\left(0-\frac{2}{3}\right)^2+\frac{1}{6}\times\left(2-\frac{2}{3}\right)^2=\frac{5}{9}$,故B错误;

所以 $D(3|\xi|-2)=3^2D(|\xi|)=5$,故D错误.

故选AC.

2.已知随机变量 $\xi$ 的分布列为 $P(\xi=k)=\frac{1}{3}$,$k=1,2,3$,则 $D(3\xi+5)=$　　(　　)

A.6　　B.9　　C.3　　D.4

A　**解析**:由题意得 $E(\xi)=\frac{1}{3}\times(1+2+3)=2$,

所以 $D(\xi)=\frac{1}{3}\times[(1-2)^2+(2-2)^2+(3-2)^2]=\frac{2}{3}$.

所以 $D(3\xi+5)=3^2\times D(\xi)=6$.故选A.

3.(多选)设 $0<p<1$,已知随机变量 $\xi$ 的分布列如表所示:

| $\xi$ | $0$ | $1$ | $2$ |
|---|---|---|---|
| $P$ | $2p-p^2$ | $p^2$ | $1-2p$ |

则下列结论正确的是　　(　　)

A.$P(\xi=2)>P(\xi=1)$

B.$P(\xi=0)>P(\xi=1)$

C.$E(\xi)$随着 $p$ 的增大而减小

D.当 $p=\frac{1}{3}$ 时,$D(\xi)=\frac{68}{81}$

BCD　**解析**:当 $p=\frac{3}{7}$ 时,$P(\xi=2)=\frac{1}{7}$,$P(\xi=1)=\frac{9}{49}$,$P(\xi=1)>P(\xi=2)$,故A错误;

因为 $0<p<1$,所以 $P(\xi=0)-P(\xi=1)=2p-p^2-p^2=2p-2p^2>0$,

所以 $P(\xi=0)>P(\xi=1)$,故B正确;

因为 $E(\xi)=p^2+2-4p$,$0<p<1$,

所以 $E(\xi)$随着 $p$ 的增大而减小,故C正确;

当 $p=\frac{1}{3}$ 时,$\xi$ 的分布列为

| $\xi$ | $0$ | $1$ | $2$ |
|---|---|---|---|
| $P$ | $\frac{5}{9}$ | $\frac{1}{9}$ | $\frac{1}{3}$ |

$E(\xi)=0\times\frac{5}{9}+1\times\frac{1}{9}+2\times\frac{1}{3}=\frac{7}{9}$,

$D(\xi)=\left(0-\frac{7}{9}\right)^2\times\frac{5}{9}+\left(1-\frac{7}{9}\right)^2\times\frac{1}{9}+\left(2-\frac{7}{9}\right)^2\times\frac{1}{3}=\frac{68}{81}$,故D正确.

故选BCD.

4.若随机变量 $X$ 服从两点分布,且 $P(X=1)=0.7$,则 $D(X)=$______.

0.21　**解析**:因为随机变量 $X$ 服从两点分布,且 $P(X=1)=0.7$,

所以 $P(X=0)=1-0.7=0.3$.

所以 $E(X)=0\times0.3+1\times0.7=0.7$,

所以 $D(X)=(0-0.7)^2\times0.3+(1-0.7)^2\times0.7=0.21$.

5.随机变量 $\xi$ 的分布列如表所示,其中 $a,b,c$ 成等差数列.若 $E(\xi)=\frac{5}{3}$,则 $D(\xi)$的值为________.

| $\xi$ | $1$ | $2$ | $3$ |
|---|---|---|---|
| $P$ | $a$ | $b$ | $c$ |

$\frac{5}{9}$ **解析**：因为 $a,b,c$ 成等差数列，

所以 $a+c=2b$.

又因为 $a+b+c=1$，

所以 $b=\frac{1}{3}$.

又因为 $E(\xi)=a+2b+3c=\frac{5}{3}$，

所以 $a=\frac{1}{2},b=\frac{1}{3},c=\frac{1}{6}$.

所以 $\xi$ 的分布列为

| $\xi$ | 1 | 2 | 3 |
|---|---|---|---|
| $P$ | $\frac{1}{2}$ | $\frac{1}{3}$ | $\frac{1}{6}$ |

所以 $D(\xi)=\left(1-\frac{5}{3}\right)^2\times\frac{1}{2}+\left(2-\frac{5}{3}\right)^2\times\frac{1}{3}+\left(3-\frac{5}{3}\right)^2\times\frac{1}{6}=\frac{5}{9}$.

**6.**盒中有 4 个球，其中 1 个红球、1 个绿球、2 个黄球.从盒中随机取球，每次取 1 个，不放回，直到取出红球为止.设此过程中取到黄球的个数为 $\xi$，则 $P(\xi=0)$ =________，$D(\xi)$=________.

$\frac{1}{3}$　$\frac{2}{3}$ **解析**：随机变量 $\xi$ 的所有可能取值为 0，1，2，$P(\xi=0)=\frac{1}{4}+\frac{1}{4}\times\frac{1}{3}=\frac{1}{3}$，$P(\xi=1)=\frac{2}{4}\times\frac{1}{3}+\frac{2}{4}\times\frac{1}{3}\times\frac{1}{2}+\frac{1}{4}\times\frac{2}{3}\times\frac{1}{2}=\frac{1}{3}$，$P(\xi=2)=\frac{2}{4}\times\frac{1}{3}\times\frac{1}{2}+\frac{1}{4}\times\frac{2}{3}\times\frac{1}{2}+\frac{2}{4}\times\frac{1}{3}\times\frac{1}{2}+\frac{2}{4}\times\frac{1}{3}\times\frac{1}{2}=\frac{1}{3}$，所以 $E(\xi)=0\times\frac{1}{3}+1\times\frac{1}{3}+2\times\frac{1}{3}=1$，$D(\xi)=\frac{1}{3}\times(0-1)^2+\frac{1}{3}\times(1-1)^2+\frac{1}{3}\times(2-1)^2=\frac{2}{3}$.

**7.**某投资公司准备明年年初将 1 000 万元投资到“低碳”项目上，现有两个项目供选择：

项目一：新能源汽车.据市场调研，投资到该项目上，到年底可能获利 30%，也可能亏损 15%，且这两种情况发生的概率分别为 $\frac{7}{9}$ 和 $\frac{2}{9}$；

项目二：通信设备.据市场调研，投资到该项目上，到年底可能获利 50%，可能损失 30%，也可能不赔不赚，且这三种情况发生的概率分别为 $\frac{3}{5}$，$\frac{1}{3}$ 和 $\frac{1}{15}$.

针对以上两个投资项目，请你为投资公司选择一个合理的项目，并说明理由.

**解**：若投资项目一，设获利为 $X_1$ 万元，$X_1$ 的所有可能取值为 300，$-150$，则 $X_1$ 的分布列为

| $X_1$ | 300 | $-150$ |
|---|---|---|
| $P$ | $\frac{7}{9}$ | $\frac{2}{9}$ |

所以 $E(X_1)=300\times\frac{7}{9}+(-150)\times\frac{2}{9}=200$，

$D(X_1)=(300-200)^2\times\frac{7}{9}+(-150-200)^2\times\frac{2}{9}=35\ 000$.

若投资项目二，设获利 $X_2$ 万元，$X_2$ 的所有可能取值为 500，$-300$，0，则 $X_2$ 的分布列为

| $X_2$ | 500 | $-300$ | 0 |
|---|---|---|---|
| $P$ | $\frac{3}{5}$ | $\frac{1}{3}$ | $\frac{1}{15}$ |

所以 $E(X_2)=500\times\frac{3}{5}+(-300)\times\frac{1}{3}+0\times\frac{1}{15}=200$，

$D(X_2)=(500-200)^2\times\frac{3}{5}+(-300-200)^2\times\frac{1}{3}+(0-200)^2\times\frac{1}{15}=140\ 000$.

由上可知 $E(X_1)=E(X_2)$，$D(X_1)<D(X_2)$，这说明虽然项目一、项目二获利相同，但项目一更稳妥.

综上所述，建议该投资公司选择项目一投资.

# 7.4　二项分布与超几何分布

## 7.4.1　二项分布

### 第1课时　二项分布及其分布列

**学习任务目标**

1.通过具体实例,了解伯努利试验,理解二项分布的概念.
2.能利用二项分布概率模型解决简单的实际问题.

## ◦ 问题式预习 ◦

**【知识清单】**

**知识点一　$n$ 重伯努利试验**

(1)定义:我们把只包含两个可能结果的试验叫做伯努利试验,将一个伯努利试验独立地重复进行 $n$ 次所组成的随机试验称为 $n$ 重伯努利试验.

(2)特征:①同一个伯努利试验重复做 $n$ 次;

②各次试验的结果相互独立.

**知识点二　二项分布**

一般地,在 $n$ 重伯努利试验中,设每次试验中事件 $A$ 发生的概率为 $p(0<p<1)$,用 $X$ 表示事件 $A$ 发生的次数,则 $X$ 的分布列为 $P(X=k)=C_n^k p^k(1-p)^{n-k}$,$k=0,1,2,\cdots,n$.如果随机变量 $X$ 的分布列具有上式的形式,则称随机变量 $X$ 服从二项分布,记作 $X\sim B(n,p)$.

由二项式定理,容易得到 $\sum\limits_{k=0}^{n}P(X=k)=\sum\limits_{k=0}^{n}C_n^k p^k(1-p)^{n-k}=[p+(1-p)]^n=1$.

**【概念辨析】**

1.判断正误(正确的打"√",错误的打"×").

(1)两点分布就是二项分布.　(×)

(2)判断一个随机变量是否服从二项分布,关键是判断其是否满足独立性和重复性.　(√)

(3)将一枚质地均匀的硬币连续抛掷5次,正面向上的次数为 $X$,则 $X\sim B(5,0.5)$.　(√)

2.有以下试验:

①掷一枚质地均匀的硬币5次;

②连续投篮3次(每次命中率相同);

③袋中装有除颜色外其他都相同的3个红球、2个白球,不放回地从中随机抽取3个球;

④袋中装有除颜色外其他都相同的3个红球、2个白球,有放回地从中随机抽取3个球.

其中为 $n$ 重伯努利试验的是________.(填序号)

①②④　**解析**:③中不放回地取球每次结果是相互影响的,不是相互独立事件,因此③不是 $n$ 重伯努利试验.

3.请思考并回答下列问题:

(1)掷一枚骰子的试验是伯努利试验吗?

**提示**:不一定,要看研究的结果是什么.如果要研究出现的点数是多少,它就有六个结果,不是伯努利试验;如果要研究出现的点数是奇数还是偶数,它只有两个结果,就是伯努利试验.

(2)伯努利试验和 $n$ 重伯努利试验的关注点有何不同?

**提示**:伯努利试验是一个有两个结果的试验,只能关注某个事件 $A$ 发生或不发生;$n$ 重伯努利试验是对一个有两个结果的试验重复进行了 $n$ 次,所以关注点是这 $n$ 次重复试验中事件 $A$ 发生的次数.

(3)对比二项分布与二项式定理,你能看出它们之间的联系吗?

**提示**:如果把 $p$ 看成 $b$,$1-p$ 看成 $a$,那么 $C_n^k p^k(1-p)^{n-k}$ 就是二项式 $[(1-p)+p]^n$ 的展开式的通项.

## ◦ 任务型课堂 ◦

### 任务1 > $n$ 重伯努利试验

判断下列试验是不是 $n$ 重伯努利试验:

①依次投掷四枚质地不同的硬币;

②某人射击,击中目标的概率是稳定的,他连续射击10次;

③口袋中装有5个白球,3个红球,2个黑球,不放回地随机从中抽取5个球.

**解**:①由于试验的条件不同(质地不同),因此不是 $n$ 重伯努利试验.

②某人射击且击中的概率是稳定的,因此是 $n$ 重伯努利试验.

③每次抽取,试验的结果有三种不同的情况,且每种情况出现的可能性不相等,因此不是 $n$ 重伯努利试验.

【探究总结】

$n$ 重伯努利试验的判断依据

(1)试验在相同的条件下重复进行.

(2)每次试验相互独立,互不影响.

(3)每次试验都只有两种结果,即事件发生、不发生.

## 任务2 $n$ 重伯努利试验的概率计算

### 探究活动

**例1** 位于坐标原点的一个质点 $P$ 按下述规则移动:质点每次移动一个单位长度,移动的方向为向左或向右,并且向左移动的概率为 $\frac{1}{3}$,向右移动的概率为 $\frac{2}{3}$.质点 $P$ 移动五次后位于点(1,0)的概率是 ( )

A.$\frac{4}{243}$ B.$\frac{8}{243}$ C.$\frac{40}{243}$ D.$\frac{80}{243}$

D **解析**:由题意可知,五次中质点 $P$ 向左移动了两次,向右移动了三次,因此质点 $P$ 移动五次后位于点(1,0)的概率是 $C_5^3\times\left(\frac{2}{3}\right)^3\times\left(\frac{1}{3}\right)^2=\frac{80}{243}$.

**例2** 甲射击一次,击中目标的概率是 $\frac{2}{3}$,假设每次射击是否击中目标相互之间没有影响.求甲射击3次至少有1次未击中目标的概率.

**解**:记"甲射击3次至少有1次未击中目标"为事件 $A$.由题意知,射击3次,相当于3重伯努利试验,由对立事件的概率计算公式得 $P(A)=1-P(\overline{A})=1-\left(\frac{2}{3}\right)^3=\frac{19}{27}$.

[一题多思]

**思考1.**求甲射击3次恰有2次击中目标的概率.

**解**:记"甲射击3次,恰有2次击中目标"为事件 $B$,则 $P(B)=C_3^2\times\left(\frac{2}{3}\right)^2\times\left(1-\frac{2}{3}\right)=\frac{4}{9}$.

**思考2.**乙射击一次,击中目标的概率为 $\frac{3}{4}$,每次射击是否击中目标相互之间没有影响.求两人各射击2次,甲恰有2次击中目标且乙恰有1次击中目标的概率.

**解**:记"甲射击2次,恰有2次击中目标"为事件 $C_1$,"乙射击2次,恰有1次击中目标"为事件 $C_2$,则 $P(C_1)=C_2^2\times\left(\frac{2}{3}\right)^2=\frac{4}{9}$,$P(C_2)=C_2^1\times\left(\frac{3}{4}\right)\times\left(1-\frac{3}{4}\right)=\frac{3}{8}$.由于甲、乙射击相互独立,所以 $P(C_1C_2)=\frac{4}{9}\times\frac{3}{8}=\frac{1}{6}$.

【探究总结】

$n$ 重伯努利试验的概率计算步骤

判断 —— 依据$n$重伯努利试验的特征,判断所给试验是否为$n$重伯努利试验

↓

分析 —— 判断所求事件是否需要拆分

↓

计算 —— 依据$n$重伯努利试验的概率公式求解每个事件的概率,最后利用互斥事件的概率加法公式计算

### 应用迁移

1.已知甲、乙两名跳高运动员每次试跳2 m高度成功的概率分别是0.7,0.6,且每次试跳成功与否相互之间没有影响.

(1)求甲、乙两人在一次试跳中至少有一人成功的概率;

(2)若甲、乙各试跳两次,求甲比乙多成功一次的概率.

**解**:(1)记"甲在一次试跳中成功"为事件 $A$,"乙在一次试跳中成功"为事件 $B$,"甲、乙两人在一次试跳中至少有一人成功"为事件 $C$.

由对立事件的概率计算公式得

$P(C)=1-P(\overline{A}\,\overline{B})=1-P(\overline{A})P(\overline{B})=1-0.3\times0.4=0.88$.

(2)设"甲在两次试跳中成功 $i$ 次"为事件 $M_i$,"乙在两次试跳中成功 $i$ 次"为事件 $N_i$,$i=0,1,2$,

所求概率 $p=P(M_1N_0)+P(M_2N_1)=P(M_1)\cdot P(N_0)+P(M_2)P(N_1)=C_2^1\times0.7\times0.3\times0.4^2+0.7^2\times C_2^1\times0.6\times0.4=0.302\ 4$.

2.甲、乙两人进行乒乓球比赛,已知在一局比赛中,甲胜的概率为 $\frac{2}{3}$,没有平局.

(1)若进行三局两胜制比赛,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**:(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

$P=\left(\frac{2}{3}\right)^3+C_3^2\times\left(\frac{2}{3}\right)^2\times\frac{1}{3}\times\frac{2}{3}+C_4^2\times\left(\frac{2}{3}\right)^2\times\left(\frac{1}{3}\right)^2\times\frac{2}{3}=\frac{64}{81}$.

## 任务3 二项分布及其应用

### 探究活动

**例3** 假设某种人寿保险规定:若投保人未达65周岁身故,则保险公司要赔偿10万元;若投保人生存至

65周岁，则保险公司不赔偿，但要给投保人一次性支付4万元.已知购买此种人寿保险的每个投保人生存至65周岁的概率都为0.9，随机抽取其中的4个投保人，设其中生存至65周岁的人数为$X$.

(1)求$X$的分布列；(参考数据：$0.9^4=0.656\ 1$)

(2)设保险公司支付给这4人的总金额为$Y$万元，写出$Y$与$X$的关系，并求$P(Y\geqslant 22)$.

**解**：(1)由于每个投保人生存至65周岁的概率都为0.9，因此$X$服从二项分布，即$X\sim B(4,0.9)$，

则$P(X=k)=C_4^k 0.9^k\times(1-0.9)^{4-k}(k=0,1,2,3,4)$，

故随机变量$X$的分布列为

| $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P$ | 0.000 1 | 0.003 6 | 0.048 6 | 0.291 6 | 0.656 1 |

(2)因为4个投保人中，生存至65周岁的人数为$X$，所以未达65周岁身故的人数为$4-X$，

因此$Y=10(4-X)+4X$，

即$Y=40-6X(X=0,1,2,3,4)$.

由$Y\geqslant 22$，即$40-6X\geqslant 22$，得$X\leqslant 3$.

所以$P(Y\geqslant 22)=P(X\leqslant 3)=P(X=0)+P(X=1)+P(X=2)+P(X=3)=1-P(X=4)=1-0.656\ 1=0.343\ 9$.

**【探究总结】**

1.当$X$服从二项分布时，应弄清$B(n,p)$中的试验次数$n$与事件发生的概率$p$.

2.解决二项分布问题的两个关注点：

(1)对于公式$P(X=k)=C_n^k p^k(1-p)^{n-k}(k=0,1,2,\cdots,n)$，必须在满足“伯努利试验”时才能应用，否则不能应用该公式.

(2)判断一个随机变量是否服从二项分布，关键有两点：一是对立性，即一次试验中，事件发生与否两者必有其一；二是重复性，即试验是独立重复地进行了$n$次.

## 应用迁移

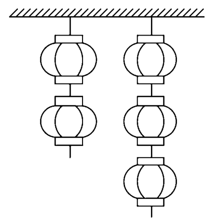
1.箱子中有标号为1,2,3,4,5,6且大小、质地完全相同的6个球，从箱子中一次随机摸出两个球，记下号码并放回，如果两球号码之积是4的倍数，那么获奖.若有4人参与摸奖，则恰好有3人获奖的概率为　(　　)

A.$\frac{16}{625}$　　B.$\frac{96}{625}$

C.$\frac{624}{625}$　　D.$\frac{4}{625}$

B　**解析**：获奖的概率为$p=\frac{6}{C_6^2}=\frac{2}{5}$，记获奖的人数为$\xi$，则$\xi\sim B\left(4,\frac{2}{5}\right)$，所以4人中恰好有3人获奖的概率为$P=C_4^3\times\left(\frac{2}{5}\right)^3\times\frac{3}{5}=\frac{96}{625}$.故选B.

2.现挂有如图所示的两串灯笼，每次随机选取其中一串并摘下其最下方的一个灯笼，直至某一串灯笼被摘完为止，则右边灯笼先被摘完的概率为　(　　)



A.$\frac{1}{4}$　　B.$\frac{1}{8}$　　C.$\frac{3}{16}$　　D.$\frac{5}{16}$

D　**解析**：由题意，右边灯笼先被摘完，摘灯笼的次数$X$可能为3,4，

$X=3$时，3次均摘下右边灯笼，

故$P(X=3)=\left(\frac{1}{2}\right)^3=\frac{1}{8}$；

$X=4$时，前3次中有2次摘下右边灯笼，1次摘下左边灯笼，第4次摘下右边灯笼，

故$P(X=4)=C_3^2\times\left(\frac{1}{2}\right)^4=\frac{3}{16}$.

所以右边灯笼先被摘完的概率为$\frac{1}{8}+\frac{3}{16}=\frac{5}{16}$.故选D.

3.某大学学生宿舍4人都需要从网上购物.大家约定：每个人通过掷一枚质地均匀的骰子决定自己去哪家网站购物，掷出点数为5或6的人去A网站购物，掷出点数小于5的人去B网站购物，且参加者必须从A网站和B网站中选择一家购物.

(1)求这4个人中恰有1人去A网站购物的概率；

(2)用$\xi,\eta$分别表示这4个人中去A网站和B网站购物的人数，令$X=\xi\eta$，求随机变量$X$的分布列.

**解**：依题意，这4个人中，每个人去A网站购物的概率为$\frac{1}{3}$，去B网站购物的概率为$\frac{2}{3}$.

设“这4个人中恰有$i$个人去A网站购物”为事件$A_i(i=0,1,2,3,4)$，

则$P(A_i)=C_4^i\left(\frac{1}{3}\right)^i\left(\frac{2}{3}\right)^{4-i}(i=0,1,2,3,4)$.

(1)这4个人中恰有1人去A网站购物的概率为

$P(A_1)=C_4^1\times\left(\frac{1}{3}\right)^1\times\left(\frac{2}{3}\right)^3=\frac{32}{81}$.

(2)$X$ 的所有可能取值为 0,3,4,

则 $P(X=0)=P(A_0)+P(A_4)=C_4^0\times\left(\frac{1}{3}\right)^0\times\left(\frac{2}{3}\right)^4+C_4^4\times\left(\frac{1}{3}\right)^4\times\left(\frac{2}{3}\right)^0=\frac{16}{81}+\frac{1}{81}=\frac{17}{81}$,

$P(X=3)=P(A_1)+P(A_3)$

$=C_4^1\times\left(\frac{1}{3}\right)^1\times\left(\frac{2}{3}\right)^3+C_4^3\times\left(\frac{1}{3}\right)^3\times\left(\frac{2}{3}\right)^1$

$=\frac{32}{81}+\frac{8}{81}=\frac{40}{81}$,

$P(X=4)=P(A_2)=C_4^2\times\left(\frac{1}{3}\right)^2\times\left(\frac{2}{3}\right)^2=\frac{8}{27}$.

所以随机变量 $X$ 的分布列为

| $X$ | 0 | 3 | 4 |
|---|---|---|---|
| $P$ | $\frac{17}{81}$ | $\frac{40}{81}$ | $\frac{8}{27}$ |

### 提质归纳

- 伯努利试验与二项分布
  - 伯努利试验 — $n$重伯努利试验
    - 应用
    - 特征
  - 二项分布
    - 分布列：$P(X=k)=C_n^k p^k(1-p)^{n-k},k=0,1,2,\cdots,n$
    - 符号表示：$X\sim B(n,p)$
    - 应用

## 课后素养评价（十四） 二项分布及其分布列

### A组 学习·理解

**1.**某高三学生进行心理素质测试，场景相同的条件下每次通过测试的概率为$\frac{4}{5}$，则连续测试 4 次，至少有 3 次通过的概率为（　　）

A.$\frac{512}{625}$　B.$\frac{256}{625}$　C.$\frac{64}{625}$　D.$\frac{64}{125}$

A　**解析**：连续测试 4 次，即 4 次独立重复试验，故所求概率为 $C_4^3\times\left(\frac{4}{5}\right)^3\times\frac{1}{5}+C_4^4\times\left(\frac{4}{5}\right)^4=\frac{512}{625}$.

**2.**一名射击运动员对同一目标独立地射击 4 次，已知他至少命中 1 次的概率为$\frac{80}{81}$，则此射击运动员每次射击命中的概率为（　　）

A.$\frac{1}{3}$　B.$\frac{2}{3}$　C.$\frac{1}{4}$　D.$\frac{2}{5}$

B　**解析**：设此射击运动员射击 4 次命中的次数为 $\xi$，每次射击命中的概率为 $p$，则 $\xi\sim B(4,p)$.

依题意可知，$P(\xi\geqslant1)=\frac{80}{81}$，所以 $1-P(\xi=0)=1-C_4^0(1-p)^4=\frac{80}{81}$，则 $(1-p)^4=\frac{1}{81}$，解得 $p=\frac{2}{3}$.

**3.**已知随机变量 $X\sim B\left(6,\frac{1}{3}\right)$，则 $P(X=2)$等于（　　）

A.$\frac{3}{16}$　B.$\frac{42}{43}$　C.$\frac{13}{243}$　D.$\frac{80}{243}$

D　**解析**：因为 $X\sim B\left(6,\frac{1}{3}\right)$，所以 $P(X=2)=C_6^2\times\left(\frac{1}{3}\right)^2\times\left(\frac{2}{3}\right)^4=\frac{80}{243}$.

**4.**某市公租房的房源位于甲、乙、丙三个片区，设每位申请人只申请其中一个片区的房源，且申请其中任一个片区的房源是等可能的，则该市的 4 位申请人中恰有 2 人申请甲片区房源的概率为________.

$\frac{8}{27}$　**解析**：每位申请人申请一个片区的房源为一次试验，这是 4 重伯努利试验.设 $A$＝“申请甲片区房源”，则 $P(A)=\frac{1}{3}$，恰有 2 人申请甲片区房源的概率 $p=C_4^2\times\left(\frac{1}{3}\right)^2\times\left(\frac{2}{3}\right)^2=\frac{8}{27}$.

**5.**假设一架飞机的每一个引擎在飞行中出现故障的概率为 $1-p$，且各引擎是否出现故障是相互独立的.已知四引擎飞机中至少有 3 个引擎正常运行，飞机才可成功飞行；两引擎飞机要 2 个引擎全部正常运行，飞机才可成功飞行.要使四引擎飞机比两引擎飞机更安全，则 $p$ 的取值范围是________.

$\left(\frac{1}{3},1\right)$　**解析**：四引擎飞机成功飞行的概率为 $C_4^3p^3(1-p)+p^4$，两引擎飞机成功飞行的概率为 $p^2$.由 $C_4^3p^3(1-p)+p^4>p^2$，得$\frac{1}{3}<p<1$.

**6.**甲、乙两人进行乒乓球比赛，比赛规则：每一局比赛中，胜者得 1 分，负者得 0 分，且比赛中没有平局.根据以往战绩，每局比赛甲获胜的概率为$\frac{2}{3}$，每局比赛的结果互不影响.

(1)经过 3 局比赛，记甲的得分为 $X$，求 $X$ 的分

布列；

(2)若比赛采取 3 局制，试计算 3 局比赛后，甲的累计得分高于乙的累计得分的概率.

**解**：(1)由题意得，$X\sim B\left(3,\frac{2}{3}\right)$，$X$ 的可能取值为 0，1，2，3，

则 $P(X=0)=\left(1-\frac{2}{3}\right)^3=\frac{1}{27}$，

$P(X=1)=C_3^1\times\frac{2}{3}\times\left(1-\frac{2}{3}\right)^2=\frac{2}{9}$，

$P(X=2)=C_3^2\times\left(\frac{2}{3}\right)^2\times\left(1-\frac{2}{3}\right)=\frac{4}{9}$，

$P(X=3)=\left(\frac{2}{3}\right)^3=\frac{8}{27}$.

所以 $X$ 的分布列为

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P$ | $\frac{1}{27}$ | $\frac{2}{9}$ | $\frac{4}{9}$ | $\frac{8}{27}$ |

(2)3 局比赛后，甲的累计得分高于乙的累计得分有两种情况：

甲获胜 2 局，甲获胜 3 局，

所以所求概率为 $P=P(X=2)+P(X=3)=\frac{4}{9}+\frac{8}{27}=\frac{20}{27}$.

## B组　应用·实践

**1.** 有 8 件产品，其中 4 件是次品，从中有放回地取 3 次(每次取 1 件)，若 $X$ 表示取得次品的次数，则 $P(X\leqslant 2)=$　(　　)

A. $\frac{3}{8}$　　B. $\frac{13}{14}$　　C. $\frac{4}{5}$　　D. $\frac{7}{8}$

D　**解析**：因为是有放回地取产品，所以每次取产品取到次品的概率为 $\frac{4}{8}=\frac{1}{2}$.从中取 3 次，$X$ 为取得次品的次数，则 $X\sim B\left(3,\frac{1}{2}\right)$，

所以 $P(X\leqslant 2)=P(X=2)+P(X=1)+P(X=0)=C_3^2\times\left(\frac{1}{2}\right)^2\times\frac{1}{2}+C_3^1\times\left(\frac{1}{2}\right)\times\left(\frac{1}{2}\right)^2+C_3^0\times\left(\frac{1}{2}\right)^3=\frac{7}{8}$.

**2.** 一名学生每次通过某种英语听力测试的概率是 $\frac{1}{2}$，他连续测试 $n$ 次，要保证他至少有一次通过的概率大于 0.9，那么 $n$ 的最小值为　(　　)

A. 6　　B. 5　　C. 4　　D. 3

C　**解析**：由题意可得 $1-C_n^0\left(\frac{1}{2}\right)^n>0.9$，即 $\left(\frac{1}{2}\right)^n<0.1$，所以 $n\geqslant 4$.

**3.** 一个袋中装有除颜色外其他都相同的 5 个白球和 3 个红球，现从袋中往外取球，每次任取 1 个记下颜色后放回，直到红球出现 10 次时停止.设停止时共取了 $\xi$ 次球，则 $P(\xi=12)$等于　(　　)

A. $C_{12}^{10}\times\left(\frac{3}{8}\right)^{10}\times\left(\frac{5}{8}\right)^2$

B. $C_{11}^9\times\left(\frac{3}{8}\right)^9\times\left(\frac{5}{8}\right)^2\times\frac{3}{8}$

C. $C_{11}^9\times\left(\frac{5}{8}\right)^9\times\left(\frac{3}{8}\right)^2$

D. $C_{11}^9\times\left(\frac{3}{8}\right)^9\times\left(\frac{5}{8}\right)^2$

B　**解析**：由题意可知，每次取出红球的概率为 $\frac{3}{8}$.“$\xi=12$”的含义是前 11 次中红球出现 9 次，第 12 次取出的球是红球，故 $P(\xi=12)=C_{11}^9\times\left(\frac{3}{8}\right)^9\times\left(\frac{5}{8}\right)^2\times\frac{3}{8}$.

**4.** 泊松分布是一种描述随机现象的概率分布，泊松分布的概率分布列为 $P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}(k=0,1,2,\cdots)$，其中 e 为自然对数的底数，$\lambda$ 是泊松分布的均值.当 $n$ 很大且 $p$ 很小时，二项分布 $B(n,p)$近似于泊松分布，其中 $\lambda=np$.一般地，当 $n\geqslant 20$ 且 $p\leqslant 0.05$时，泊松分布可作为二项分布的近似.若随机变量 $X\sim B(1\ 000,0.001)$，则 $P(X\geqslant 2)$的近似值为　(　　)

A. $1-\frac{1}{e}$　　B. $1-\frac{2}{e}$

C. $1-\frac{e}{4}$　　D. $1-\frac{1}{e^2}$

B　**解析**：由题可知，$n=1\ 000>20$，$p=0.001<0.05$，

所以泊松分布可作为二项分布的近似，

此时 $\lambda=1\ 000\times 0.001=1$，

所以 $P(X=k)=\frac{1}{k!}e^{-1}$.

所以 $P(X=0)=\frac{1}{0!}e^{-1}=\frac{1}{e}$，$P(X=1)=\frac{1}{1!}e^{-1}=\frac{1}{e}$，则 $P(X\geqslant 2)=1-P(X=0)-P(X=1)=1-\frac{2}{e}$.

**5.** 张师傅驾车从公司开往火车站，途经 4 个路口，这 4 个路口将公司到火车站分成 5 个路段，每个路段的

驾车时间都是3 min.如果遇到红灯要停留1 min，假设他在各个路口是否遇到红灯是相互独立的，并且遇到红灯概率都是$\frac{1}{3}$，那么张师傅此行所需时间不少于16 min的概率为________.

$\frac{65}{81}$　**解析**：如果不遇到红灯，全程需要15 min，否则至少需要16 min，所以张师傅此行所需时间不少于16 min的概率 $p=1-\left(1-\frac{1}{3}\right)^4=\frac{65}{81}$.

**6.** 羽毛球比赛的计分规则：采用21分制，3局2胜.每回合中，取胜的一方加1分.每局中，先得21分且领先至少2分的一方该局获胜，否则继续比赛；若双方打成29平，则一方领先1分，即算该局取胜.某次羽毛球比赛中，甲选手在每回合中得分的概率为$\frac{3}{4}$，乙选手在每回合中得分的概率为$\frac{1}{4}$.

(1)在一局比赛中，若甲、乙两名选手的得分均为18，求再经过4回合，甲选手获胜的概率；

(2)在一局比赛中，记前4回合甲选手的得分为$X$，求$X$的分布列.

**解**：(1)记"再经过4回合，甲选手获胜"为事件$A$，可知甲在第4回合胜，前3回合胜2个回合，

所以$P(A)=\frac{3}{4}\times C_3^2\times\left(\frac{3}{4}\right)^2\times\frac{1}{4}=\frac{81}{256}$.

(2)易知$X$的可能取值为0，1，2，3，4，且$X\sim B\left(4,\frac{3}{4}\right)$，

$P(X=0)=C_4^0\times\left(\frac{1}{4}\right)^4=\frac{1}{256}$，

$P(X=1)=C_4^1\times\frac{3}{4}\times\left(\frac{1}{4}\right)^3=\frac{3}{64}$，

$P(X=2)=C_4^2\times\left(\frac{3}{4}\right)^2\times\left(\frac{1}{4}\right)^2=\frac{27}{128}$，

$P(X=3)=C_4^3\times\left(\frac{3}{4}\right)^3\times\frac{1}{4}=\frac{27}{64}$，

$P(X=4)=C_4^4\times\left(\frac{3}{4}\right)^4=\frac{81}{256}$.

所以$X$的分布列为

| $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P$ | $\frac{1}{256}$ | $\frac{3}{64}$ | $\frac{27}{128}$ | $\frac{27}{64}$ | $\frac{81}{256}$ |

# 第2课时　二项分布的均值与方差

**学习任务目标**

**1.** 能熟练计算二项分布的均值与方差.

**2.** 能利用二项分布的均值与方差解决简单的实际问题.

## ○ 问题式预习 ○

### 【知识清单】

**知识点　二项分布的均值与方差**

若随机变量$X$服从二项分布$B(n,p)$，则$E(X)=$__$np$__，$D(X)=$__$np(1-p)$__.

特别地，当$n=1$时，$X$服从两点分布，此时$E(X)=$__$p$__，$D(X)=$__$p(1-p)$__.

### 【概念辨析】

**1.** 判断正误(正确的打"√"，错误的打"×").

(1)离散型随机变量$X$的均值$E(X)$是一个随机数值.　(×)

(2)若两个随机变量的均值相同，则这两个随机变量的分布也一定相同.　(×)

(3)若$X$服从两点分布，则$E(X)=\frac{1}{2}$.　(×)

**2.** 已知$\xi\sim B\left(n,\frac{1}{2}\right)$，$\eta\sim B\left(n,\frac{1}{3}\right)$，且$E(\xi)=15$，则$E(\eta)$等于　(　)

A.5　　B.10　　C.15　　D.20

B　**解析**：由$E(\xi)=\frac{1}{2}n=15$，解得$n=30$，

所以$\eta\sim B\left(30,\frac{1}{3}\right)$，所以$E(\eta)=30\times\frac{1}{3}=10$.

**3.** 请思考并回答下列问题：

(1)二项分布与两点分布有何关系？

**提示**：两点分布是一种特殊的二项分布，即$n=1$时的二项分布；二项分布中每次试验的结果都服从两点分布.

(2)请举出两个服从二项分布的随机变量的例子.

**提示**：①某射击运动员命中10环的概率为0.8，他在10次射击中命中10环的次数$X$是一个随机变量，$X\sim B(10,0.8)$.

②经过某路口碰到红灯的概率为$p$，某人经过该路口10次，碰到红灯的次数$X$是一个随机变量，$X\sim B(10,p)$.

## ◦ 任务型课堂 ◦

### 任务1 二项分布的均值与方差

**1.**珠算是以算盘为工具进行数字计算的一种方法.算盘每个档(挂珠的杆)上有 7 个算珠,用梁隔开,梁上面 2 颗叫上珠,梁下面 5 颗叫下珠,如图所示.若一个算盘共有 13 档,从每档中的 7 颗算珠中任取 1 颗,设 $X$ 为取得上珠的颗数,则 $D(X)=$　(　　)


A.$\frac{100}{49}$　　B.$\frac{65}{7}$　　C.$\frac{26}{7}$　　D.$\frac{130}{49}$

D　**解析**:由题意知,$X\sim B\left(13,\frac{2}{7}\right)$,所以 $D(X)=13\times\frac{2}{7}\times\frac{5}{7}=\frac{130}{49}$.

**2.**已知随机变量 $X\sim B(n,p)$,若 $E(X)=1$,$D(X)=\frac{4}{5}$,则 $P(X=3)=$　(　　)

A.$\frac{64}{3\ 125}$　　B.$\frac{128}{625}$　　C.$\frac{1}{25}$　　D.$\frac{32}{625}$

D　**解析**:由 $E(X)=1$,$D(X)=\frac{4}{5}$,得 $np=1$,$np(1-p)=\frac{4}{5}$,解得 $n=5$,$p=\frac{1}{5}$,

所以 $P(X=3)=C_5^3\times\left(\frac{1}{5}\right)^3\times\left(1-\frac{1}{5}\right)^2=\frac{32}{625}$.故选 D.

**3.**某运动员投篮命中率为 0.6.

(1)求投篮 1 次时,命中次数 $X$ 的数学期望;

(2)求重复 5 次投篮时,命中次数 $Y$ 的数学期望和方差.

**解**:(1)投篮 1 次,命中次数 $X$ 的分布列如表:

| $X$ | 0 | 1 |
|---|---|---|
| $P$ | 0.4 | 0.6 |

则 $E(X)=0.6$.

(2)由题意得,重复 5 次投篮,命中的次数 $Y$ 服从二项分布,即 $Y\sim B(5,0.6)$.

则 $E(Y)=5\times0.6=3$,$D(Y)=5\times0.6\times0.4=1.2$.

**【探究总结】**

如果随机变量 $X$ 服从二项分布,即 $X\sim B(n,p)$,那么 $E(X)=np$,$D(X)=np(1-p)$.

计算二项分布的均值与方差时,可直接代入求解,从而避免繁杂的计算过程.

### 任务2 二项分布的均值与方差的实际应用

#### 探究活动

**例 1**　在新高考方案"3+1+2"模式中,"3"指统考科目语文、数学、外语 3 门,不分文理;"1"指在物理、历史 2 门科目中选择 1 门;"2"指在思想政治、地理、化学、生物 4 门科目中选择 2 门.学生根据高校的要求,结合自身特长兴趣自主选择.某校统计发现选择物理的学生占全体学生的 $\frac{3}{4}$,并且在选择物理的条件下,选择地理的概率为 $\frac{2}{3}$,在选择历史的条件下,选择地理的概率为 $\frac{4}{5}$.设该校学生甲、乙、丙三人中选择地理的人数为随机变量 $X$.

(1)求 $X=2$ 的概率;

(2)求 $X$ 的分布列以及数学期望.

**解**:(1)设"该校学生选择地理"为事件 $A$,

则 $P(A)=\frac{3}{4}\times\frac{2}{3}+\frac{1}{4}\times\frac{4}{5}=\frac{7}{10}$,

因此 $X\sim B\left(3,\frac{7}{10}\right)$,

所以 $P(X=2)=C_3^2\times\left(\frac{7}{10}\right)^2\times\frac{3}{10}=\frac{441}{1\ 000}$.

(2)由于 $X\sim B\left(3,\frac{7}{10}\right)$,

则 $P(X=0)=C_3^0\times\left(\frac{7}{10}\right)^0\times\left(\frac{3}{10}\right)^3=\frac{27}{1\ 000}$,

$P(X=1)=C_3^1\times\left(\frac{7}{10}\right)^1\times\left(\frac{3}{10}\right)^2=\frac{189}{1\ 000}$,

$P(X=2)=C_3^2\times\left(\frac{7}{10}\right)^2\times\frac{3}{10}=\frac{441}{1\ 000}$,

$P(X=3)=C_3^3\times\left(\frac{7}{10}\right)^3\times\left(\frac{3}{10}\right)^0=\frac{343}{1\ 000}$.

所以随机变量 $X$ 的分布列为

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P$ | $\frac{27}{1\ 000}$ | $\frac{189}{1\ 000}$ | $\frac{441}{1\ 000}$ | $\frac{343}{1\ 000}$ |

所以 $E(X)=3\times\frac{7}{10}=\frac{21}{10}$.

**例 2**　一出租车司机从某饭店到火车站途中有 6 个路口,假设他在各路口是否遇到红灯是相互独立的,并且在各路口遇到红灯的概率均是 $\frac{1}{3}$.

(1)求这位司机遇到红灯的次数 $\xi$ 的期望与方差;

(2)若遇上红灯,则需等待 30 s,求这位司机总共等待的时间 $\eta$(单位:s)的期望与方差.

**解**：(1)易知司机遇到红灯的次数 $\xi \sim B\left(6,\frac{1}{3}\right)$，

故 $E(\xi)=6\times\frac{1}{3}=2$，$D(\xi)=6\times\frac{1}{3}\times\left(1-\frac{1}{3}\right)=\frac{4}{3}$.

(2)由已知得 $\eta=30\xi$，故 $E(\eta)=30E(\xi)=60$，
$D(\eta)=900D(\xi)=1\ 200$.

［一题多思］

**思考 1.**求这名司机在首次遇到红灯或到达目的地停车前经过的路口数 $X$ 的分布列.

**解**：设前 $k$ 个是绿灯，第 $(k+1)$ 个是红灯，

则 $X$ 的分布列为 $P(X=k)=\left(\frac{2}{3}\right)^{k}\times\frac{1}{3}$，$k=0,1,2,3,4,5$，

所以 $P(X=0)=\left(\frac{2}{3}\right)^{0}\times\frac{1}{3}=\frac{1}{3}$，

$P(X=1)=\frac{2}{3}\times\frac{1}{3}=\frac{2}{9}$，

$P(X=2)=\left(\frac{2}{3}\right)^{2}\times\frac{1}{3}=\frac{4}{27}$，

$P(X=3)=\left(\frac{2}{3}\right)^{3}\times\frac{1}{3}=\frac{8}{81}$，

$P(X=4)=\left(\frac{2}{3}\right)^{4}\times\frac{1}{3}=\frac{16}{243}$，

$P(X=5)=\left(\frac{2}{3}\right)^{5}\times\frac{1}{3}=\frac{32}{729}$，

若全为绿灯，则 $P(X=6)=\left(\frac{2}{3}\right)^{6}=\frac{64}{729}$.

故 $X$ 的分布列为

| $X$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $P$ | $\frac{1}{3}$ | $\frac{2}{9}$ | $\frac{4}{27}$ | $\frac{8}{81}$ | $\frac{16}{243}$ | $\frac{32}{729}$ | $\frac{64}{729}$ |

**思考 2.**求这名司机在途中至少遇到一次红灯的概率.

**解**：所求概率为 $P(\xi\geqslant 1)=1-P(\xi=0)=1-\left(\frac{2}{3}\right)^{6}=\frac{665}{729}$.

**【探究总结】**

**1.用二项分布求解实际应用问题的步骤**

(1)判定随机变量 $X$ 服从二项分布，即 $X\sim B(n,p)$.

(2)根据 $P(X=k)=C_n^k p^k(1-p)^{n-k}(k=0,1,\cdots,n)$求出 $X$ 的分布列.

(3)随机变量 $X$ 的均值可用公式 $E(X)=np$ 求解，方差可用公式 $D(X)=np(1-p)$求解.

(4)根据 $X$ 的均值与方差解决实际问题.

**2.**有些随机变量虽不服从二项分布，但与之具有线性关系的另一随机变量服从二项分布，这时，可以综合应用 $E(a\xi+b)=aE(\xi)+b$，$D(a\xi+b)=a^2D(\xi)$以及二项分布的性质求解.

## 应用迁移

某商场为刺激消费，拟按以下方案进行促销：顾客每消费 500 元便得到抽奖券一张，每张抽奖券的中奖概率为$\frac{1}{2}$.若中奖，商场返还顾客现金 100 元.某顾客现购买价格为2 300 元的台式电脑一台，得到抽奖券四张，每次抽奖互不影响.

(1)设该顾客抽奖后中奖的抽奖券张数为 $X$，求随机变量 $X$ 的分布列；

(2)设该顾客购买台式电脑的实际支出为 $Y$ 元，用 $X$ 表示 $Y$，并求随机变量 $Y$ 的均值.

**解**：(1)因为每张抽奖券是否中奖是相互独立的，

因此 $X\sim B\left(4,\frac{1}{2}\right)$.

所以 $P(X=0)=C_4^0\times\left(\frac{1}{2}\right)^{4}=\frac{1}{16}$，

$P(X=1)=C_4^1\times\left(\frac{1}{2}\right)^{4}=\frac{1}{4}$，

$P(X=2)=C_4^2\times\left(\frac{1}{2}\right)^{4}=\frac{3}{8}$，

$P(X=3)=C_4^3\times\left(\frac{1}{2}\right)^{4}=\frac{1}{4}$，

$P(X=4)=C_4^4\times\left(\frac{1}{2}\right)^{4}=\frac{1}{16}$.

所以随机变量 $X$ 的分布列为

| $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P$ | $\frac{1}{16}$ | $\frac{1}{4}$ | $\frac{3}{8}$ | $\frac{1}{4}$ | $\frac{1}{16}$ |

(2)因为 $X\sim B\left(4,\frac{1}{2}\right)$，所以 $E(X)=4\times\frac{1}{2}=2$.

又由题意可知 $Y=2\ 300-100X$，

所以 $E(Y)=E(2\ 300-100X)=2\ 300-100E(X)=2\ 300-100\times 2=2\ 100$.

因此随机变量 $Y$ 的均值为 2 100.

## 提质归纳

二项分布

二项分布的均值：$E(X)=np$

二项分布的方差：$D(X)=np(1-p)$

实际应用

# 课后素养评价（十五）　二项分布的均值与方差

## A组　学习·理解

1.已知随机变量 $X$ 服从二项分布 $B(n,p)$.若 $E(X)=2$，$D(X)=\frac{4}{3}$，则 $p=$（　　）

A.$\frac{3}{4}$　　B.$\frac{2}{3}$　　C.$\frac{1}{3}$　　D.$\frac{1}{4}$

C　**解析**：由随机变量 $X$ 服从二项分布 $B(n,p)$，又 $E(X)=2$，$D(X)=\frac{4}{3}$，所以 $np=2$，$np(1-p)=\frac{4}{3}$，解得 $p=\frac{1}{3}$.故选 C.

2.同时抛掷 5 枚质地均匀的硬币 80 次.设 5 枚硬币正好出现 2 枚正面向上，3 枚反面向上的次数为 $X$，则 $X$ 的均值是（　　）

A.20　　B.25　　C.30　　D.40

B　**解析**：抛掷一次正好出现 2 枚正面向上，3 枚反面向上的概率为 $\frac{C_5^2}{2^5}=\frac{5}{16}$，所以 $X\sim B\left(80,\frac{5}{16}\right)$.故 $E(X)=80\times\frac{5}{16}=25$.

3.已知随机变量 $X\sim B(10,0.5)$，$Y=2X-8$，则 $E(Y)=$（　　）

A.6　　B.2

C.4　　D.3

B　**解析**：由题意，随机变量 $X\sim B(10,0.5)$，所以 $E(X)=10\times0.5=5$.因为 $Y=2X-8$，所以 $E(Y)=2E(X)-8=2\times5-8=2$.故选 B.

4.某人从家乘车到单位，途中经过 3 个路口.假设在各路口是否遇到红灯是相互独立的，且遇到红灯的概率都是0.4，则此人上班途中遇到红灯的次数的方差为（　　）

A.0.48　　B.1.2　　C.0.72　　D.0.6

C　**解析**：设此人上班途中遇到红灯的次数为 $X$，则 $X\sim B(3,0.4)$，

所以 $D(X)=3\times0.4\times0.6=0.72$.

5.（多选）一次数学测验的试卷由 25 道选择题构成，每道选择题有四个选项，其中有且仅有一个选项是正确的，每道题选择正确得 4 分，不作出选择或选错不得分，满分 100 分.某学生每道题选对的概率均为0.6，则（　　）

A.该学生在这次数学测验中做对的题目的个数的均值为 15

B.该学生在这次数学测验中做对的题目的个数的方差为 6

C.该学生在这次测验中的成绩的均值为 60

D.该学生在这次测验中的成绩的方差为 24

ABC　**解析**：设该学生做对题目的个数为 $X$，则 $X\sim B(25,0.6)$，所以 $E(X)=25\times0.6=15$，A 正确；

$D(X)=25\times0.6\times(1-0.6)=6$，B 正确；

设该学生的得分为 $Y$，则 $Y=4X$，所以 $E(Y)=4E(X)=60$，C 正确；

$D(Y)=4^2D(X)=16\times6=96$，D 错误.

故选 ABC.

6.某校为举办甲、乙两项不同活动，分别设计了相应的活动方案：方案一、方案二.为了解该校学生对活动方案是否支持，对学生进行简单随机抽样，获得数据如下表：

| 方案 | 男生 | | 女生 | |
|---|---|---|---|---|
| | 支持 | 不支持 | 支持 | 不支持 |
| 方案一 | 200 人 | 400 人 | 300 人 | 100 人 |
| 方案二 | 350 人 | 250 人 | 150 人 | 250 人 |

假设所有学生对活动方案是否支持相互独立.

(1)分别估计该校男生支持方案一的概率、该校女生支持方案一的概率；

(2)从该校全体男生中随机抽取 2 人，全体女生中随机抽取 1 人，估计这 3 人中恰有 2 人支持方案一的概率；

(3)将该校学生支持方案二的概率估计值记为 $p_0$，假设该校高一年级有 500 名男生和 300 名女生，除高一年级外其他年级学生支持方案二的概率估计值记为 $p_1$，试比较 $p_0$ 与 $p_1$ 的大小.（结论不要求证明）

**解**：(1)该校男生支持方案一的概率为 $\frac{200}{200+400}=\frac{1}{3}$，该校女生支持方案一的概率为 $\frac{300}{300+100}=\frac{3}{4}$.

(2)3 人中恰有 2 人支持方案一分两种情况：仅有 2 名男生支持方案一；仅有 1 名男生和 1 名女生支持方案一.

所以 3 人中恰有 2 人支持方案一的概率为 $\left(\frac{1}{3}\right)^2\times\left(1-\frac{3}{4}\right)+C_2^1\times\frac{1}{3}\times\left(1-\frac{1}{3}\right)\times\frac{3}{4}=\frac{13}{36}$.

(3) $p_1 < p_0$.

**7.** 某咖啡店，男会员占60%，女会员占40%.现对会员进行服务质量满意度调查.根据调查结果得知，男会员对服务质量满意的概率为$\frac{5}{6}$，女会员对服务质量满意的概率为$\frac{5}{8}$.

(1)随机选取一名会员，求其对服务质量满意的概率；

(2)从会员中随机抽取3人，记抽取的3人中对服务质量满意的人数为$X$，求$X$的分布列和数学期望.

**解**：(1)记事件$A_1$为"会员为男会员"，$A_2$为"会员为女会员"，事件$B$为"对服务质量满意"，则由题意可知，$P(A_1)=\frac{3}{5}$，$P(A_2)=\frac{2}{5}$，$P(B|A_1)=\frac{5}{6}$，$P(B|A_2)=\frac{5}{8}$，

所以$P(B)=P(A_1)\cdot P(B|A_1)+P(A_2)\cdot P(B|A_2)=\frac{3}{5}\times\frac{5}{6}+\frac{2}{5}\times\frac{5}{8}=\frac{3}{4}$.

(2)由题意及(1)知，$X\sim B\left(3,\frac{3}{4}\right)$，

则$P(X=0)=C_3^0\times\left(\frac{3}{4}\right)^0\times\left(\frac{1}{4}\right)^3=\frac{1}{64}$，

$P(X=1)=C_3^1\times\left(\frac{3}{4}\right)^1\times\left(\frac{1}{4}\right)^2=\frac{9}{64}$，

$P(X=2)=C_3^2\times\left(\frac{3}{4}\right)^2\times\left(\frac{1}{4}\right)^1=\frac{27}{64}$，

$P(X=3)=C_3^3\times\left(\frac{3}{4}\right)^3\times\left(\frac{1}{4}\right)^0=\frac{27}{64}$，

所以$X$的分布列为

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P$ | $\frac{1}{64}$ | $\frac{9}{64}$ | $\frac{27}{64}$ | $\frac{27}{64}$ |

所以$E(X)=0\times\frac{1}{64}+1\times\frac{9}{64}+2\times\frac{27}{64}+3\times\frac{27}{64}=\frac{9}{4}$.

## B组 应用·实践

**1.** 若随机变量$\xi\sim B(n,0.6)$，且$E(\xi)=3$，则$P(\xi=1)$的值为 ( )

A. $2\times0.4^4$　　B. $2\times0.4^5$

C. $3\times0.4^4$　　D. $3\times0.6^4$

C　**解析**：因为$\xi\sim B(n,0.6)$，所以$E(\xi)=n\times0.6=3$，解得$n=5$.故$P(\xi=1)=C_5^1\times0.6\times0.4^4=3\times0.4^4$.

**2.** 设随机变量$\xi$的分布列为$P(\xi=k)=C_n^k\left(\frac{2}{3}\right)^k\cdot\left(\frac{1}{3}\right)^{n-k}$，$k=0,1,2,\cdots,n$，且$E(\xi)=24$，则$D(\xi)$的值为 ( )

A. 8　　B. 12

C. $\frac{2}{9}$　　D. $\frac{1}{6}$

A　**解析**：由题意可得$\xi\sim B\left(n,\frac{2}{3}\right)$，所以$E(\xi)=\frac{2}{3}n=24$，解得$n=36$.所以$D(\xi)=n\times\frac{2}{3}\times\frac{1}{3}=36\times\frac{2}{9}=8$.

**3.** 某综艺节目中，有一个盲拧魔方游戏，就是玩家先观察魔方状态并进行记忆，记住后蒙住眼睛快速还原魔方.为了解某市盲拧魔方爱好者的水平状况，某兴趣小组在全市范围内随机抽取了100名盲拧魔方爱好者进行调查，得到的情况如表所示.

| 用时/s | [5,10] | (10,15] | (15,20] | (20,25] |
|---|---|---|---|---|
| 男性人数 | 15 | 22 | 14 | 9 |
| 女性人数 | 5 | 11 | 17 | 7 |

以这100名盲拧魔方爱好者用时不超过10 s的频率，代替全市所有盲拧魔方爱好者用时不超过10 s的概率，每位盲拧魔方爱好者用时是否超过10 s相互独立.若该兴趣小组在全市范围内再随机抽取20名盲拧魔方爱好者进行测试，其中用时不超过10 s的人数最有可能(即概率最大)是 ( )

A. 2　　B. 3　　C. 4　　D. 5

C　**解析**：根据题意得，每位盲拧魔方爱好者用时不超过10 s的概率为$\frac{20}{100}=\frac{1}{5}$.

设随机抽取的20名盲拧魔方爱好者中用时不超过10 s的人数为$\xi$，

则$\xi\sim B\left(20,\frac{1}{5}\right)$，其中$P(\xi=k)=C_{20}^k\left(\frac{1}{5}\right)^k\cdot\left(\frac{4}{5}\right)^{20-k}$，$k=0,1,2,\cdots,20$.

由$\begin{cases}P(\xi=k)\geqslant P(\xi=k+1),\\P(\xi=k)\geqslant P(\xi=k-1),\end{cases}$

得$\begin{cases}C_{20}^k\left(\frac{1}{5}\right)^k\left(\frac{4}{5}\right)^{20-k}\geqslant C_{20}^{k+1}\left(\frac{1}{5}\right)^{k+1}\left(\frac{4}{5}\right)^{19-k},\\C_{20}^k\left(\frac{1}{5}\right)^k\left(\frac{4}{5}\right)^{20-k}\geqslant C_{20}^{k-1}\left(\frac{1}{5}\right)^{k-1}\left(\frac{4}{5}\right)^{21-k},\end{cases}$

化简得$\begin{cases}4(k+1)\geqslant20-k,\\21-k\geqslant4k,\end{cases}$解得$\frac{16}{5}\leqslant k\leqslant\frac{21}{5}$.

又 $k\in\mathbf{N}$，所以 $k=4$.

所以这 20 名盲拧魔方爱好者中用时不超过 10 s 的人数最有可能是 4.故选 C.

**4.**已知某同学投篮投中的概率为 $\frac{2}{3}$，现该同学要投篮 3 次，且每次投篮结果相互独立，则恰好投中 2 次的概率为________；记 $X$ 为该同学在这 3 次投篮中投中的次数，则随机变量 $X$ 的数学期望为________.

$\frac{4}{9}$　2　**解析**：由独立重复试验的概率公式，可得恰好投中 2 次的概率为 $C_3^2\times\left(\frac{2}{3}\right)^2\times\frac{1}{3}=\frac{4}{9}$.

由题意知随机变量 $X\sim B\left(3,\frac{2}{3}\right)$，

所以 $E(X)=3\times\frac{2}{3}=2$.

**5.**某电视台开展有奖答题活动，要求每位选手答 30 道选择题，每道选择题有 4 个选项，其中有且只有一个正确选项，每一题选对得 5 分，选错或不选得 0 分，满分 150 分.规定满 100 分得三等奖，满 120 分得二等奖，满 140 分得一等奖.有一个选手选对任一题的概率都是 0.8，则该选手最可能得到________等奖.

二　**解析**：设答对题的个数为 $X$，则 $X\sim B(30,0.8)$，所以 $E(X)=30\times0.8=24$.因为 $24\times5=120$(分)，所以最可能得到二等奖.

**6.**设随机变量 $\xi$ 服从二项分布 $B\left(5,\frac{1}{2}\right)$，则函数 $f(x)=x^2+4x+\xi$ 存在零点的概率是________.

$\frac{31}{32}$　**解析**：由函数 $f(x)=x^2+4x+\xi$ 存在零点，得方程 $x^2+4x+\xi=0$ 有解，所以 $\Delta=16-4\xi\geqslant0$，得 $\xi\leqslant4$.

又因为随机变量 $\xi\sim B\left(5,\frac{1}{2}\right)$，

所以所求概率 $P(\xi\leqslant4)=1-P(\xi=5)=1-C_5^5\times\left(\frac{1}{2}\right)^5=\frac{31}{32}$.

## 7.4.2　超几何分布

**学习任务目标**

1. 通过具体实例，了解超几何分布的概念.
2. 会利用公式求服从超几何分布的随机变量的分布列及均值.
3. 了解超几何分布与二项分布的关系，能利用超几何分布的概率模型解决简单的实际问题.

## ◦ 问题式预习 ◦

### 【知识清单】

**知识点　超几何分布**

(1)概念：一般地，假设一批产品共有 $N$ 件，其中有 $M$ 件次品.从 $N$ 件产品中随机抽取 $n$ 件(不放回)，用 $X$ 表示抽取的 $n$ 件产品中的次品数，则 $X$ 的分布列为 $P(X=k)=\frac{C_M^kC_{N-M}^{n-k}}{C_N^n}$，$k=m,m+1,m+2,\cdots,r$.其中 $n,N,M\in\mathbf{N}^*$，$M\leqslant N$，$n\leqslant N$，$m=\max\{0,n-N+M\}$，$r=\underline{\min\{n,M\}}$.如果随机变量 $X$ 的分布列具有上式的形式，那么称随机变量 $X$ 服从超几何分布.

(2)均值：$E(X)=\frac{nM}{N}=\underline{np}$，其中 $P=\frac{M}{N}$，是 $N$ 件产品的次品率.

### 【概念辨析】

**1.**判断正误(正确的打“√”，错误的打“×”).

(1)将一枚硬币连抛 3 次，正面朝上的次数 $X$ 服从超几何分布.　( × )

(2)盒中有 4 个白球和 3 个黑球，有放回地随机摸取 3 个球，取出黑球的个数 $X$ 服从超几何分布.　( × )

(3)某射手的命中率为 0.8，现对目标射击 3 次，命中目标的次数 $X$ 服从超几何分布.　( × )

**2.**某导游团有外语导游 10 人，其中 6 人会说日语.现要选出 4 人去完成一项任务，则其中有 2 人会说日语的概率为________.

$\frac{3}{7}$　**解析**：设选出的 4 人中，会说日语的人数为 $X$，则 $X$ 服从 $N=10,M=6,n=4$ 的超几何分布.所以有 2 人会说日语的概率为 $P(X=2)=\frac{C_6^2C_4^2}{C_{10}^4}=\frac{3}{7}$.

**3.**请思考并回答问题：

如何区别二项分布与超几何分布？

**提示**：一般地，超几何分布的模型是“取次品”，是不放回抽样，而二项分布的模型是“独立重复试验”，是有放回抽样.

## 任务型课堂

### 任务1 超几何分布的概念

盒中共有9个球,其中有4个红球、3个黄球和2个白球,这些球除颜色外完全相同.

(1)若用随机变量$X$表示任取的4个球中红球的个数,则$X$服从超几何分布,其参数的值为 ( )

A.$N=9,M=4,n=4$

B.$N=9,M=5,n=5$

C.$N=13,M=4,n=4$

D.$N=14,M=5,n=5$

(2)若用随机变量$Y$表示任取的3个球中红球的个数,则$Y$的可能取值为________.

(3)若用随机变量$Z$表示任取的5个球中白球的个数,则$P(Z=2)=$________.

(1)A (2)0,1,2,3 (3)$\frac{5}{18}$ **解析**:(1)根据超几何分布的定义知,$N=9,M=4,n=4$.

(2)由于选取了3个球,因此随机变量$Y$的所有可能取值为0,1,2,3.

(3)$P(Z=2)=\frac{C_2^2C_7^3}{C_9^5}=\frac{5}{18}$.

**【探究总结】**

**超几何分布的三个注意点**

(1)超几何分布的模型是不放回抽样.

(2)超几何分布中的参数是$N,M,n$.

(3)超几何分布可解决产品中的正品和次品、盒中的白球和黑球等问题,研究对象往往由差异明显的两部分组成.

### 任务2 超几何分布的分布列

#### 探究活动

**例1** 已知外形完全一样的某品牌电子笔6支装一盒,每盒电子笔最多一支次品,每盒电子笔有次品的概率是$\frac{1}{10}$.

(1)现有一盒电子笔,随机抽出2支来检测.

①求抽出的2支均是正品的概率;

②已知抽出的2支是正品,求剩余产品有次品的概率.

(2)已知甲、乙两盒电子笔均有次品,由于某种原因将两盒电子笔完全随机混合在了一起,现随机选3支电子笔进行检测,记$\xi$为选出的3支电子笔中次品的数量,求$\xi$的分布列.

**解**:(1)①记事件$A$:该盒有次品,事件$B$:抽出的2支均是正品,

则$P(A)=\frac{1}{10}$,$P(B|A)=\frac{C_5^2}{C_6^2}=\frac{10}{15}=\frac{2}{3}$,$P(B|\overline{A})=1$,

所以$P(B)=P(A)P(B|A)+P(\overline{A})\cdot P(B|\overline{A})=\frac{1}{10}\times\frac{2}{3}+\frac{9}{10}\times1=\frac{29}{30}$.

②$P(A|B)=\frac{P(A)P(B|A)}{P(B)}=\frac{\frac{1}{10}\times\frac{2}{3}}{\frac{29}{30}}=\frac{2}{29}$.

(2)由题意知,两盒电子笔中共有10支正品,2支次品,所以$\xi$的所有可能取值为0,1,2,

$P(\xi=0)=\frac{C_{10}^3}{C_{12}^3}=\frac{120}{220}=\frac{6}{11}$,

$P(\xi=1)=\frac{C_{10}^2C_2^1}{C_{12}^3}=\frac{90}{220}=\frac{9}{22}$,

$P(\xi=2)=\frac{C_{10}^1C_2^2}{C_{12}^3}=\frac{10}{220}=\frac{1}{22}$.

所以$\xi$的分布列为

| $\xi$ | 0 | 1 | 2 |
|---|---|---|---|
| $P$ | $\frac{6}{11}$ | $\frac{9}{22}$ | $\frac{1}{22}$ |

**例2** 在10件产品中有2件次品,连续不放回地抽3次,每次抽1件,求抽到的次品数$X$的均值与方差.

**解**:(方法一)由题意知$X$的可能取值为0,1,2,

$P(X=0)=\frac{C_8^3}{C_{10}^3}=\frac{7}{15}$,

$P(X=1)=\frac{C_2^1C_8^2}{C_{10}^3}=\frac{7}{15}$,

$P(X=2)=\frac{C_2^2C_8^1}{C_{10}^3}=\frac{1}{15}$.

所以随机变量$X$的分布列为

| $X$ | 0 | 1 | 2 |
|---|---|---|---|
| $P$ | $\frac{7}{15}$ | $\frac{7}{15}$ | $\frac{1}{15}$ |

$E(X)=0\times\frac{7}{15}+1\times\frac{7}{15}+2\times\frac{1}{15}=\frac{3}{5}$,

$D(X)=\left(0-\frac{3}{5}\right)^2\times\frac{7}{15}+\left(1-\frac{3}{5}\right)^2\times\frac{7}{15}+\left(2-\frac{3}{5}\right)^2\times\frac{1}{15}=\frac{28}{75}$.

(方法二)由题意知$P(X=k)=\frac{C_2^kC_8^{3-k}}{C_{10}^3}$,$k=0,1,2$,

所以随机变量$X$服从超几何分布,

$n=3, M=2, N=10$,

所以 $E(X)=\frac{nM}{N}=\frac{3\times2}{10}=\frac{3}{5}$,

$D(X)=\left(0-\frac{3}{5}\right)^2\times\frac{7}{15}+\left(1-\frac{3}{5}\right)^2\times\frac{7}{15}+\left(2-\frac{3}{5}\right)^2\times\frac{1}{15}=\frac{28}{75}$.

[一题多思]

**思考.**将“不放回”改为“有放回”,求抽到的次品数 $X$ 的均值与方差.

**解:**由题意知,抽取 1 次抽到次品的概率为 $\frac{2}{10}=\frac{1}{5}$,

所以随机变量 $X$ 服从二项分布 $B\left(3,\frac{1}{5}\right)$,

所以 $E(X)=3\times\frac{1}{5}=\frac{3}{5}$,

$D(X)=3\times\frac{1}{5}\times\left(1-\frac{1}{5}\right)=\frac{12}{25}$.

**【探究总结】**

**1.**超几何分布问题的求解步骤:

(1)辨模型:结合实际情境分析所求概率分布问题中的研究对象是否由具有明显差异的两部分组成,如“正品、次品”“优、劣”等,或可转化为具有明显差异的两部分.并检查抽样方法是否为不放回抽样.

(2)算概率:可以直接借助公式 $P(X=k)=\frac{C_M^kC_{N-M}^{n-k}}{C_N^n}$ 求解,也可以利用排列、组合及概率的知识求解,需注意借助公式求解时应理解参数 $M,N,n$ 的含义.

(3)列分布列:把求得的概率值通过表格表示出来.

(4)根据分布列解决问题.

**2.**对于多次抽取,求抽取到“次品”数的相关问题,若“不放回地抽取”一般为超几何分布,若“有放回地抽取”一般为二项分布.

### 应用迁移

网民对一电商平台的某种特色农产品的销售服务质量进行评价,每位参加购物的网民在“好评”“中评”“差评”中选择一个进行评价,在参与评价的网民中随机抽取 2 万人,按年龄分为“50 岁以下”和“50 岁以上(含 50 岁)”两类进行了统计,得到给予“好评”“中评”“差评”评价的人数如表所示.

| 网民年龄 | 好评人数 | 中评人数 | 差评人数 |
| --- | --- | --- | --- |
| 50 岁以下 | 9 000 | 3 000 | 2 000 |
| 50 岁以上(含 50 岁) | 1 000 | 2 000 | 3 000 |

(1)根据这 2 万人的样本估计总体,从参与评价的网民中每次随机抽取 1 人,若抽取到“好评”,则终止抽取,否则继续抽取,直到抽取到“好评”,但抽取次数最多不超过 5 次,求抽取了 5 次的概率;

(2)从给予“中评”评价的网民中,用分层随机抽样的方法抽取 10 人,再从这 10 人中随机抽取 3 人,记抽取的 3 人中年龄在 50 岁以下的人数为 $X$,求 $X$ 的分布列.

**解:**(1)从参与评价的网民中随机抽取 1 人,抽取到“好评”的概率为 $\frac{9\ 000+1\ 000}{20\ 000}=\frac{1}{2}$,

则抽取了 5 次的概率为 $\left(1-\frac{1}{2}\right)^4\times\frac{1}{2}+\left(1-\frac{1}{2}\right)^5=\frac{1}{16}$.

(2)在给予“中评”评价的网民中,50 岁以下与 50 岁以上(含 50 岁)的人数之比为 3∶2,

因此在抽取的 10 人中,50 岁以下与 50 岁以上(含 50 岁)的人数分别为 6 和 4.

由题意知,$X$ 服从参数 $N=10, n=3, M=6$ 的超几何分布,

所以 $P(X=k)=\frac{C_6^kC_4^{3-k}}{C_{10}^3}, k=0,1,2,3$.

于是 $X$ 的分布列为

| $X$ | 0 | 1 | 2 | 3 |
| --- | --- | --- | --- | --- |
| $P$ | $\frac{1}{30}$ | $\frac{3}{10}$ | $\frac{1}{2}$ | $\frac{1}{6}$ |

## 任务3 超几何分布的应用

### 探究活动

**例 3**　为营造浓厚的全国文明城市创建氛围,积极响应创建全国文明城市号召,提高对创城行动的责任感和参与度,某学校号召师生利用周末参与创城志愿活动.高二(1)班某小组有男生 4 人,女生 2 人,现从中随机选取 2 人作为志愿者参加活动.

(1)求在有女生参加活动的条件下,恰有一名女生参加活动的概率;

(2)记参加活动的女生人数为 $X$,求 $X$ 的分布列、期望及方差.

**解:**(1)设“有女生参加活动”为事件 $A$,“恰有一名女生参加活动”为事件 $B$,

则 $P(AB)=\frac{C_4^1C_2^1}{C_6^2}=\frac{8}{15}, P(A)=\frac{C_4^1C_2^1+C_2^2}{C_6^2}=\frac{3}{5}$,

所以 $P(B|A)=\frac{P(AB)}{P(A)}=\frac{\frac{8}{15}}{\frac{3}{5}}=\frac{8}{9}$.

(2)由题意知，$X$ 服从超几何分布，且 $P(X=k)=\frac{C_2^kC_4^{2-k}}{C_6^2}(k=0,1,2)$，

则 $P(X=0)=\frac{C_4^2}{C_6^2}=\frac{2}{5}$，$P(X=1)=\frac{C_4^1C_2^1}{C_6^2}=\frac{8}{15}$，$P(X=2)=\frac{C_2^2}{C_6^2}=\frac{1}{15}$，

所以 $X$ 的分布列为

| $X$ | 0 | 1 | 2 |
|---|---|---|---|
| $P$ | $\frac{2}{5}$ | $\frac{8}{15}$ | $\frac{1}{15}$ |

$E(X)=0\times\frac{2}{5}+1\times\frac{8}{15}+2\times\frac{1}{15}=\frac{2}{3}$，$D(X)=\frac{2}{5}\times\left(0-\frac{2}{3}\right)^2+\frac{8}{15}\times\left(1-\frac{2}{3}\right)^2+\frac{1}{15}\times\left(2-\frac{2}{3}\right)^2=\frac{16}{45}$.

**【探究总结】**

**1.超几何分布与二项分布的关系**

(1)超几何分布的均值与二项分布的均值都是 $np$，但是 $n,p$ 的意义不同.超几何分布中 $n$ 是不放回地随机抽取的产品数，$p$ 是次品率.

(2)对于不放回抽样，当 $n$ 远远小于 $N$ 时，每抽取一次后，对 $N$ 的影响很小，此时，超几何分布可以近似看作二项分布.

**2.超几何分布的应用**

对于服从超几何分布的随机变量，通过公式可以计算概率、列分布列、求期望等，利用这些可以解决实际生活中与之相关的问题.

## 应用迁移

**1.**已知高一某班共有学生 21 人，其中男生 12 人，女生 9 人.现采用分层随机抽样的方法从中抽取 7 人，测试他们学习某课程的效果，效果分为优秀和良好两种，优秀得 2 分，良好得 1 分.

(1)应抽取男生、女生各多少人？

(2)若抽取的 7 人中，4 人的测试效果为优秀，3 人为良好，现从这 7 人中随机抽取 3 人.

①用 $X$ 表示抽取的 3 人的得分之和，求随机变量 $X$ 的分布列及数学期望；

②设事件 $A$ 为“抽取的 3 人中，既有测试效果为优秀的，也有测试效果为良好的”，求事件 $A$ 发生的概率.

**解**：(1)因为采用分层随机抽样的方法进行抽样，所以应抽取女生 $7\times\frac{9}{21}=3$(人)，抽取男生 $7\times\frac{12}{21}=4$(人).

(2)①由题意知，随机变量 $X$ 的所有可能取值为 3，4，5，6，

$P(X=3)=\frac{C_4^0C_3^3}{C_7^3}=\frac{1}{35}$，$P(X=4)=\frac{C_4^1C_3^2}{C_7^3}=\frac{12}{35}$，

$P(X=5)=\frac{C_4^2C_3^1}{C_7^3}=\frac{18}{35}$，$P(X=6)=\frac{C_4^3C_3^0}{C_7^3}=\frac{4}{35}$.

所以随机变量 $X$ 的分布列为

| $X$ | 3 | 4 | 5 | 6 |
|---|---|---|---|---|
| $P$ | $\frac{1}{35}$ | $\frac{12}{35}$ | $\frac{18}{35}$ | $\frac{4}{35}$ |

$E(X)=3\times\frac{1}{35}+4\times\frac{12}{35}+5\times\frac{18}{35}+6\times\frac{4}{35}=\frac{165}{35}=\frac{33}{7}$.

②由①知 $P(A)=P(X=4)+P(X=5)=\frac{12}{35}+\frac{18}{35}=\frac{6}{7}$，所以事件 $A$ 发生的概率为 $\frac{6}{7}$.

**2.**某学校高一、高二、高三三个年级的学生人数之比为 3∶2∶2，该校用分层随机抽样的方法抽取 7 名学生来了解学生的睡眠情况.

(1)应从高一、高二、高三三个年级的学生中分别抽取多少人？

(2)若抽出的 7 人中有 4 人睡眠不足，3 人睡眠充足.

①从这 7 人中随机抽取 3 人做进一步的身体健康检查，用 $X$ 表示抽取的 3 人中“睡眠不足”的学生人数，求随机变量 $X$ 的分布列；

②将这 7 名学生中“睡眠不足”的频率视为该学校学生中“睡眠不足”的概率，从该学校全体学生(人数较多)中随机抽取 3 人做进一步的身体健康检查，记 $Y$ 表示抽到“睡眠不足”学生的人数，求 $Y$ 的期望和方差.

**解**：(1)由已知，三个年级的人数之比为 3∶2∶2，由于采用分层随机抽样的方法从中抽取 7 人，因此应从高一、高二、高三三个年级的学生中分别抽取 3 人、2 人、2 人.

(2)①由题可知随机变量 $X$ 服从超几何分布，随机变量 $X$ 的所有可能取值为 0，1，2，3，

则 $P(X=k)=\frac{C_4^k\cdot C_3^{3-k}}{C_7^3}(k=0,1,2,3)$.

所以随机变量 $X$ 的分布列为

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P$ | $\frac{1}{35}$ | $\frac{12}{35}$ | $\frac{18}{35}$ | $\frac{4}{35}$ |

②由题意知 $Y\sim B\left(3,\frac{4}{7}\right)$，

所以 $E(Y)=3\times\frac{4}{7}=\frac{12}{7}$，

$D(Y)=3\times\frac{4}{7}\times\frac{3}{7}=\frac{36}{49}$.

**提质归纳**

超几何分布
- 分布列：$P(X=k)=\frac{C_M^kC_{N-M}^{n-k}}{C_N^n}$，$k=m,m+1,m+2,\cdots,r$.其中 $n,N,M\in\mathbf{N}^*,M\leqslant N,n\leqslant N,m=\max\{0,n-N+M\},r=\min\{n,M\}$
- 均值：$E(X)=\frac{nM}{N}=np$
- 应用

## 课后素养评价（十六）　超几何分布

### A组　学习·理解

**1.**（多选）下列随机变量 $X$ 不服从超几何分布的是（　　）

A.$X$ 表示 $n$ 次重复抛掷一枚骰子出现点数是3的倍数的次数

B.$X$ 表示连续抛掷两枚骰子，所得的2个骰子的点数之和

C.有一批产品共有 $N$ 件，其中次品有 $M$ 件（$N>M>0$），有放回随机抽取 $n$ 件（$n>N$），抽出的次品件数为 $X$

D.有一批产品共有 $N$ 件，其中次品有 $M$ 件（$N>M>0$），不放回随机抽取 $n$ 件（$n\leqslant M$），抽出的次品件数为 $X$

ABC　**解析**：对于A，设事件 $E$ 为“抛掷一枚骰子出现的点数是3的倍数”，则 $P(E)=\frac{1}{3}$，

而在 $n$ 次独立重复试验中事件 $E$ 恰好发生了 $k(k=0,1,2,\cdots,n)$ 次的概率 $P(\xi=k)=C_n^k\left(\frac{1}{3}\right)^k\cdot\left(\frac{2}{3}\right)^{n-k}$，符合二项分布的定义，不是超几何分布，故A符合题意；

对于B，$X$ 的可能取值是 $2,3,\cdots,12$，

且 $P(X=2)=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}$，$P(X=3)=2\times\frac{1}{6}\times\frac{1}{6}=\frac{1}{18}$，…，显然不符合超几何分布的定义，因此 $X$ 不服从超几何分布，故B符合题意；

C和D的区别：C是“有放回”抽取，而D是“不放回”抽取，显然D中 $n$ 次试验是不独立的，因此D服从超几何分布，对于C，有 $X\sim B\left(n,\frac{M}{N}\right)$，故C符合题意，D不符合题意.

故选ABC.

**2.**设袋中有80个红球、20个白球.若从袋中任取10个球，则其中恰有6个红球的概率为（　　）

A.$\frac{C_{80}^4C_{10}^6}{C_{100}^{10}}$　　B.$\frac{C_{80}^6C_{10}^4}{C_{100}^{10}}$

C.$\frac{C_{80}^4C_{20}^6}{C_{100}^{10}}$　　D.$\frac{C_{80}^6C_{20}^4}{C_{100}^{10}}$

D　**解析**：若随机变量 $X$ 表示任取10个球中红球的个数，则 $X$ 服从参数为 $N=100,M=80,n=10$ 的超几何分布.取到的10个球中恰有6个红球，即 $X=6$，$P(X=6)=\frac{C_{80}^6C_{20}^4}{C_{100}^{10}}$.

**3.**已知6件产品中有2件次品，4件正品，检验员从中随机抽取3件进行检测，记取到的正品数为 $X$，则下列结论正确的是（　　）

A.$E(2X-1)=\frac{4}{3}$　　B.$D(X)=\frac{1}{5}$

C.$E(X)=1$　　D.$D(2X-1)=\frac{8}{5}$

D　**解析**：根据题意可知，$X$ 可能取1，2，3，且服从超几何分布，

则 $P(X=1)=\frac{C_2^2C_4^1}{C_6^3}=\frac{1}{5}$，

$P(X=2)=\frac{C_2^1C_4^2}{C_6^3}=\frac{3}{5}$，

$P(X=3)=\frac{C_4^3}{C_6^3}=\frac{1}{5}$，

所以 $E(X)=1\times\frac{1}{5}+2\times\frac{3}{5}+3\times\frac{1}{5}=2$，

$D(X)=(1-2)^2\times\frac{1}{5}+(2-2)^2\times\frac{3}{5}+(3-2)^2\times\frac{1}{5}=\frac{2}{5}$，

$E(2X-1)=2E(X)-1=2\times2-1=3$，

$D(2X-1)=4D(X)=\frac{8}{5}$.

故选D.

**4.**在10个排球中有6个正品、4个次品.从中任取4个，则正品数比次品数少的概率为（　　）

A.$\frac{5}{42}$　　B.$\frac{4}{35}$　　C.$\frac{19}{42}$　　D.$\frac{8}{21}$

A　**解析**：正品数比次品数少，有两种情况：0 个正品 4 个次品，1 个正品 3 个次品.

由超几何分布的概率可知，当取到 0 个正品 4 个次品时，$p_1=\frac{C_4^4}{C_{10}^4}=\frac{1}{210}$.

当取到 1 个正品 3 个次品时，$p_2=\frac{C_6^1C_4^3}{C_{10}^4}=\frac{24}{210}=\frac{4}{35}$.

所以正品数比次品数少的概率为 $p_1+p_2=\frac{5}{42}$.

**5.**盒中有 10 个螺丝钉，其中 3 个是坏的.现从盒中随机抽取 4 个，则下列事件中概率是$\frac{3}{10}$的为　　（　　）

A.恰有 1 个是坏的

B.4 个全是好的

C.恰有 2 个是好的

D.至多有 2 个是坏的

C　**解析**：对于 A，事件的概率为$\frac{C_3^1C_7^3}{C_{10}^4}=\frac{1}{2}$；

对于 B，事件的概率为$\frac{C_7^4}{C_{10}^4}=\frac{1}{6}$；

对于 C，事件的概率为$\frac{C_3^2C_7^2}{C_{10}^4}=\frac{3}{10}$；

对于 D，事件的概率为$\frac{C_7^4+C_3^1C_7^3+C_3^2C_7^2}{C_{10}^4}=\frac{29}{30}$.

**6.**（多选）一个袋中装有除颜色外其余完全相同的 6 个黑球和 4 个白球，现从中任取 4 个小球，设取出的 4 个小球中白球的个数为 $X$，则　　（　　）

A.随机变量 $X$ 服从二项分布

B.随机变量 $X$ 服从超几何分布

C.$P(X=2)=\frac{3}{7}$

D.$E(X)=\frac{8}{5}$

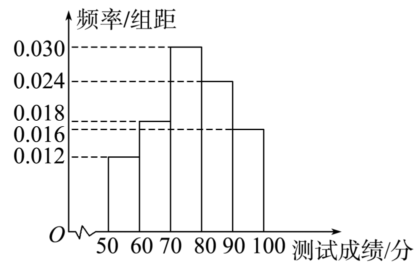
BCD　**解析**：由题意可知随机变量 $X$ 服从参数为 $N=10,M=4,n=4$ 的超几何分布，故 A 错误，B 正确；随机变量 $X$ 的可能取值为 0，1，2，3，4，所以 $P(X=0)=\frac{C_6^4}{C_{10}^4}=\frac{1}{14}$，$P(X=1)=\frac{C_4^1C_6^3}{C_{10}^4}=\frac{8}{21}$，$P(X=2)=\frac{C_4^2C_6^2}{C_{10}^4}=\frac{3}{7}$，$P(X=3)=\frac{C_4^3C_6^1}{C_{10}^4}=\frac{4}{35}$，$P(X=4)=\frac{C_4^4}{C_{10}^4}=\frac{1}{210}$，故 $E(X)=0\times\frac{1}{14}+1\times\frac{8}{21}+2\times\frac{3}{7}+3\times\frac{4}{35}+4\times\frac{1}{210}=\frac{8}{5}\left(或 E(X)=4\times\frac{4}{10}=\frac{8}{5}\right)$，故 C，D 正确.故选 BCD.

**7.**现有 10 件商品，其中 3 件瑕疵品、7 件合格品.若从这 10 件商品中任取 2 件，设取到 $X$ 件瑕疵品，则 $X$ 的数学期望是________.

$\frac{3}{5}$　**解析**：由题意知，$X$ 的可能取值是 0，1，2，$P(X=0)=\frac{C_7^2}{C_{10}^2}=\frac{7}{15}$，$P(X=1)=\frac{C_7^1C_3^1}{C_{10}^2}=\frac{7}{15}$，$P(X=2)=\frac{C_3^2}{C_{10}^2}=\frac{1}{15}$，

故 $E(X)=0\times\frac{7}{15}+1\times\frac{7}{15}+2\times\frac{1}{15}=\frac{3}{5}$.所以 $X$ 的数学期望是$\frac{3}{5}$.

**8.**某校为了解高三学生身体素质情况，从某项体育测试成绩中随机抽取 $n$ 名学生的成绩进行分析，得到成绩频率分布直方图如图所示.已知成绩在[90，100]内的学生人数为 8，且有 4 名女生的成绩在[50，60)内，则 $n=$________；现从成绩在[50，60)内的学生中随机抽取 2 名学生，记所抽取学生中女生的人数为 $\xi$，则 $\xi$ 的均值是________.




50　$\frac{4}{3}$　**解析**：依题意得 $0.016\times10n=8$，则 $n=50$.

可得成绩在[50，60)内的人数为 $0.012\times10\times50=6$，

其中 4 名女生，2 名男生.

所以 $\xi$ 的可能取值为 0，1，2，

$P(\xi=0)=\frac{C_2^2}{C_6^2}=\frac{1}{15}$，

$P(\xi=1)=\frac{C_2^1C_4^1}{C_6^2}=\frac{8}{15}$，

$P(\xi=2)=\frac{C_4^2}{C_6^2}=\frac{6}{15}=\frac{2}{5}$.

故 $E(\xi)=0\times\frac{1}{15}+1\times\frac{8}{15}+2\times\frac{2}{5}=\frac{4}{3}$.

**9.**一袋中有质地、大小、颜色都相同的 10 个球，其中 8 个标有数字 1，2 个标有数字 5.从中任取 2 个球，记所取的 2 个球上所标数字之和为 $x$，求 $x$ 的分布列.

**解**：$X$ 的可能取值为 2，6，10，

$P(X=2)=\frac{C_8^2}{C_{10}^2}=\frac{28}{45}$，$P(X=6)=\frac{C_8^1C_2^1}{C_{10}^2}=\frac{16}{45}$，$P(X=10)=\frac{C_2^2}{C_{10}^2}=\frac{1}{45}$.

故 $X$ 的分布列为

| $X$ | 2 | 6 | 10 |
|---|---|---|---|
| $P$ | $\frac{28}{45}$ | $\frac{16}{45}$ | $\frac{1}{45}$ |

## B组　应用·实践

1.一个小组有甲、乙、丙等共6人，从中任选2名代表，则其中甲当选的概率是　(　　)

A.$\frac{1}{2}$　B.$\frac{1}{3}$　C.$\frac{1}{4}$　D.$\frac{1}{5}$

B　**解析**：设$X$表示2名代表中甲的个数，$X$的可能取值为0，1.由题意知$X$服从超几何分布，其中参数$N=6$，$M=1$，$n=2$，则$P(X=1)=\frac{C_1^1C_5^1}{C_6^2}=\frac{1}{3}$.

2.孪生素数猜想是希尔伯特在1900年提出的23个数学问题之一，可以直观地描述为：存在无穷多个素数$p$，使得$p+2$是素数.素数对$(p,p+2)$称为孪生素数对.从8个数对：(3，5)，(5，7)，(7，9)，(9，11)，(11，13)，(13，15)，(15，17)，(17，19)中任取3个，设取出的孪生素数对的个数为$X$，则$E(X)=$　(　　)

A.$\frac{3}{8}$　B.$\frac{1}{2}$　C.$\frac{3}{2}$　D.3

C　**解析**：由题意知，8个数对中有(3，5)，(5，7)，(11，13)，(17，19)，共4个孪生素数对，

所以$X$的可能取值为0，1，2，3.

$P(X=0)=\frac{C_4^0C_4^3}{C_8^3}=\frac{1}{14}$，$P(X=1)=\frac{C_4^1C_4^2}{C_8^3}=\frac{3}{7}$，

$P(X=2)=\frac{C_4^2C_4^1}{C_8^3}=\frac{3}{7}$，$P(X=3)=\frac{C_4^3C_4^0}{C_8^3}=\frac{1}{14}$，

所以$E(X)=0\times\frac{1}{14}+1\times\frac{3}{7}+2\times\frac{3}{7}+3\times\frac{1}{14}=\frac{3}{2}$.

3.(多选)袋中有10个除颜色外完全相同的球，其中6个黑球、4个白球，现从中任取4个球，记随机变量$X$为其中白球的个数，随机变量$Y$为其中黑球的个数.若取出一个白球得2分，取出一个黑球得1分，随机变量$Z$为取出4个球的总得分，则下列结论中正确的是　(　　)

A.$P(|Z-6|\leqslant1)=\frac{97}{105}$

B.$E(X)>E(Y)$

C.$D(X)=D(Y)$

D.$E(Z)=\frac{28}{5}$

ACD　**解析**：由题意知，$X$，$Y$均服从超几何分布，且$X+Y=4$，$Z=2X+Y$，故$P(X=k)=\frac{C_4^kC_6^{4-k}}{C_{10}^4}(k=0,1,2,3,4)$.$P(|Z-6|\leqslant1)=1-P(Z=4)-P(Z=8)=1-P(X=0)-P(X=4)=\frac{97}{105}$，故选项A正确；$E(X)=4\times\frac{4}{10}=\frac{8}{5}$，$E(Y)=4-E(X)=\frac{12}{5}$，$D(X)=D(4-Y)=D(Y)$，故选项B错误，选项C正确；$E(Z)=2E(X)+E(Y)=\frac{28}{5}$，故选项D正确.故选ACD.

4.在30瓶饮料中，有3瓶已过保质期.从这30瓶饮料中任取2瓶，则至少取到1瓶已过保质期的饮料的概率为________.(结果用最简分数表示)

$\frac{28}{145}$　**解析**：从这30瓶饮料中任取2瓶，设“至少取到1瓶已过保质期的饮料”为事件$A$，则$P(A)=\frac{C_{27}^1C_3^1}{C_{30}^2}+\frac{C_3^2}{C_{30}^2}=\frac{28}{145}$.

5.某中学举办校庆知识竞赛，参赛的同学需要从10道题中随机抽取4道来回答.竞赛规则：每题回答正确得10分，回答不正确得$-5$分.

(1)已知甲同学每题回答正确的概率均为0.5，且各题回答正确与否之间没有影响，记甲的总得分为$X$，求$X$的期望和方差；

(2)已知乙同学能正确回答10道题中的6道，记乙的总得分为$Y$，求$Y$的分布列.

**解**：(1)设甲答对的题目数为$\xi$，则$\xi\sim B(4,0.5)$，

所以$E(\xi)=4\times0.5=2$，$D(\xi)=4\times0.5\times(1-0.5)=1$.

又因为$X=10\xi-5(4-\xi)=15\xi-20$，

所以$E(X)=15E(\xi)-20=15\times2-20=10$，

$D(X)=15^2\times D(\xi)=15^2\times1=225$.

(2)设乙答对的题目数为$\eta$，可知$\eta$的可能取值为0，1，2，3，4，

则$Y=10\eta-5(4-\eta)=15\eta-20$，则有

$P(Y=-20)=P(\eta=0)=\frac{C_4^4C_6^0}{C_{10}^4}=\frac{1}{210}$，

$P(Y=-5)=P(\eta=1)=\frac{C_4^3C_6^1}{C_{10}^4}=\frac{4}{35}$，

$P(Y=10)=P(\eta=2)=\frac{C_4^2C_6^2}{C_{10}^4}=\frac{3}{7}$，

$P(Y=25)=P(\eta=3)=\frac{C_4^1C_6^3}{C_{10}^4}=\frac{8}{21}$，

$P(Y=40)=P(\eta=4)=\frac{C_4^0C_6^4}{C_{10}^4}=\frac{1}{14}$，

所以$Y$的分布列为

| $Y$ | $-20$ | $-5$ | 10 | 25 | 40 |
|---|---|---|---|---|---|
| $P$ | $\frac{1}{210}$ | $\frac{4}{35}$ | $\frac{3}{7}$ | $\frac{8}{21}$ | $\frac{1}{14}$ |

# 7.5　正态分布

**学习任务目标**

**1.**理解正态分布与标准正态分布的概念.

**2.**了解正态密度函数,理解正态曲线的性质.

**3.**会求正态分布的随机变量在给定区间取值的概率,能利用正态分布知识解决实际问题.

## ◦ 问题式预习 ◦

### 【知识清单】

**知识点一　正态分布**

对于刻画随机误差分布的解析式:$f(x)=\dfrac{1}{\sigma\sqrt{2\pi}}\cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$,$x\in\mathbf{R}$,其中$\mu\in\mathbf{R}$,$\sigma>0$为参数.显然,对任意的$x\in\mathbf{R}$,$f(x)>0$,它的图象在$x$轴的上方.可以证明$x$轴和曲线之间的区域的面积为1.我们称$f(x)$为正态密度函数,称它的图象为正态密度曲线,简称正态曲线.若随机变量$X$的概率分布密度函数为$f(x)$,则称随机变量$X$服从正态分布,记为$X\sim N(\mu,\sigma^2)$.特别地,当$\mu=0$,$\sigma=1$时,称随机变量$X$服从标准正态分布.

**知识点二　正态曲线的特点及性质**

(1)曲线是单峰的,它关于直线$x=\mu$对称.

(2)曲线在$x=\mu$处达到峰值$\dfrac{1}{\sigma\sqrt{2\pi}}$.

(3)当$|x|$无限增大时,曲线无限接近$x$轴.

(4)当参数$\sigma$取固定值时,正态曲线的位置由$\mu$确定,且随着$\mu$的变化而沿$x$轴平移,如图1所示.

(5)当$\mu$取定值时,曲线的形状由$\sigma$确定:当$\sigma$较小时,峰值高,正态曲线"瘦高",表示随机变量$X$的分布比较集中;当$\sigma$较大时,峰值低,正态曲线"矮胖",表示随机变量$X$的分布比较分散,如图2所示.

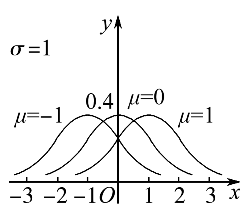



图1

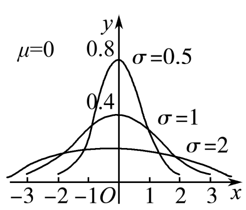



图2

**知识点三　正态变量在三个特殊区间内取值的概率值及3σ原则**

假设$X\sim N(\mu,\sigma^2)$,可以证明:对给定的$k\in\mathbf{N}^*$,$P(\mu-k\sigma\leqslant X\leqslant\mu+k\sigma)$是一个只与$k$有关的定值.特别地,

(1)$P(\mu-\sigma\leqslant X\leqslant\mu+\sigma)\approx0.682\ 7$;

(2)$P(\mu-2\sigma\leqslant X\leqslant\mu+2\sigma)\approx0.954\ 5$;

(3)$P(\mu-3\sigma\leqslant X\leqslant\mu+3\sigma)\approx0.997\ 3$.

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$[\mu-3\sigma,\mu+3\sigma]$中的值,这在统计学中称为$3\sigma$原则.

### 【概念辨析】

**1.**判断正误(正确的打"√",错误的打"×").

(1)正态密度函数中参数$\mu$,$\sigma$的意义分别是样本的均值与方差.　(×)

(2)若$X\sim N(\mu,\sigma^2)$,则$P(X<\mu)=\dfrac{1}{2}$.　(√)

(3)在正态分布中,参数$\sigma$越小,表示随机变量$X$的分布越集中;$\sigma$越大,表示随机变量$X$的分布越分散.　(√)

(4)正态曲线与$x$轴可能相交.　(×)

**2.**已知随机变量$\xi$服从正态分布$N(1,\sigma^2)$,且$P(0<\xi<1)=0.3$,则$P(\xi<2)=$　(　)

A.0.8　B.0.75　C.0.7　D.0.6

A　**解析**:因为$\xi\sim N(1,\sigma^2)$,且$P(0<\xi<1)=0.3$,所以$P(\xi\geqslant2)=P(\xi\leqslant0)=P(\xi<1)-P(0<\xi<1)=0.5-0.3=0.2$.

所以$P(\xi<2)=1-P(\xi\geqslant2)=1-0.2=0.8$.故选A.

**3.**已知随机变量$X$服从正态分布$N(2,1)$,则$P(X<1)\approx$________.

0.158 65　**解析**:因为$\mu=2$,$\sigma=1$,所以$P(1\leqslant X\leqslant3)\approx0.682\ 7$.

所以$P(X<1)=P(X>3)\approx\dfrac{1-0.682\ 7}{2}=0.158\ 65$.

**4.**请思考并回答问题:

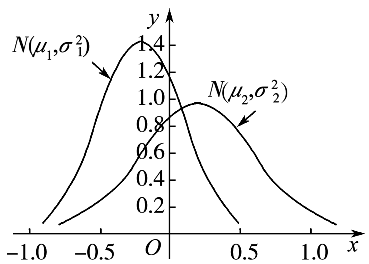
一个正态分布由参数$\mu$和$\sigma$完全确定,这两个参数可以由什么近似估计?

**提示**:参数$\mu$是反映随机变量取值的平均水平的特征数,可以用样本均值去估计;$\sigma$是衡量随机变量总体波动大小的特征数,可以用样本标准差去估计.

# ◦ 任务型课堂 ◦

## 任务1 正态曲线及其性质

1.设两个正态分布 $N(\mu_1,\sigma_1^2)(\sigma_1>0)$ 和 $N(\mu_2,\sigma_2^2)(\sigma_2>0)$ 对应的正态密度曲线如图所示，则有（　　）




A.$\mu_1<\mu_2,\sigma_1<\sigma_2$

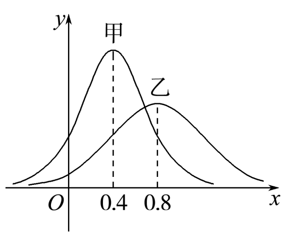
B.$\mu_1<\mu_2,\sigma_1>\sigma_2$

C.$\mu_1>\mu_2,\sigma_1<\sigma_2$

D.$\mu_1>\mu_2,\sigma_1>\sigma_2$

A　**解析**：$\mu$ 反映的是随机变量的平均水平，直线 $x=\mu$ 是正态密度曲线的对称轴，由题图可知 $\mu_1<\mu_2$.$\sigma$ 反映的是随机变量的离散程度，$\sigma$ 越大，越分散，曲线越“矮胖”；$\sigma$ 越小，越集中，曲线越“瘦高”，由题图可知 $\sigma_1<\sigma_2$.

2.甲、乙两类水果的质量(单位：kg)分别服从正态分布 $N_1(\mu_1,\sigma_1^2)$，$N_2(\mu_2,\sigma_2^2)$，相应的正态密度曲线如图所示，则下列说法正确的是（　　）




A.甲类水果的平均质量比乙类水果的平均质量大

B.乙类水果的质量比甲类水果的质量更集中于均值左右

C.$\sigma_1>\sigma_2$

D.$\mu_1=0.4$

D　**解析**：由题图可知甲类水果的平均质量为 0.4 kg，D 正确；

乙类水果的平均质量为 0.8 kg，故甲类水果的平均质量比乙类水果的平均质量小，A 错误；

由于甲曲线比乙曲线更“瘦高”，则 $\sigma_1<\sigma_2$，故甲类水果的质量比乙类水果的质量更集中于均值左右，B，C 错误.

故选 D.

**【探究总结】**

**利用正态曲线的特点求参数 $\mu,\sigma$**

(1)正态曲线是单峰的，它关于直线 $x=\mu$ 对称，由此特点结合图象可求出 $\mu$.

(2)正态曲线在 $x=\mu$ 处达到峰值 $\frac{1}{\sigma\sqrt{2\pi}}$，由此特点结合图象可求出 $\sigma$.

## 任务2 正态分布的概率计算

### 探究活动

**例1**　设 $X\sim N(10,1)$.

(1)求证：$P(1<X<2)=P(18<X<19)$；

(2)若 $P(X\leqslant 2)=a$，求 $P(10<X<18)$.

(1)**证明**：因为 $X\sim N(10,1)$，

所以正态密度曲线关于直线 $x=10$ 对称.

而区间(1，2)和(18，19)关于直线 $x=10$ 对称，

所以 $P(1<X<2)=P(18<X<19)$.

(2)**解**：因为 $P(X\leqslant 2)+P(2<X\leqslant 10)+P(10<X<18)+P(X\geqslant 18)=1$，$\mu=10$，

所以 $P(X\leqslant 2)=P(X\geqslant 18)=a$，

$P(2<X\leqslant 10)=P(10<X<18)$.

所以 $2a+2P(10<X<18)=1$，

即 $P(10<X<18)=\frac{1-2a}{2}=\frac{1}{2}-a$.

**【探究总结】**

**正态分布的概率计算方法**

(1)充分利用正态曲线的对称性和曲线与 $x$ 轴之间的区域的面积为 1.

(2)熟悉并参考 $P(\mu-\sigma\leqslant X\leqslant\mu+\sigma)$，$P(\mu-2\sigma\leqslant X\leqslant\mu+2\sigma)$，$P(\mu-3\sigma\leqslant X\leqslant\mu+3\sigma)$的值.

(3)注意概率的转化：

①$P(X<a)=1-P(X\geqslant a)$；

②$P(X<\mu-a)=P(X>\mu+a)$；

③$P(a<X<b)=P(X<b)-P(X\leqslant a)$；

④若 $b<\mu$，则 $P(X<b)=\frac{1}{2}-P(b\leqslant X\leqslant\mu)$.

**提醒**：正态曲线并非都关于 $y$ 轴对称，标准正态分布的正态曲线关于 $y$ 轴对称.

### 应用迁移

1.若随机变量 $\xi\sim N(3,\sigma^2)$，且 $P(\xi<6)=0.86$，则 $P(3<\xi<6)=$（　　）

A.0.26　B.0.34　C.0.36　D.0.42

C　**解析**：因为随机变量 $\xi\sim N(3,\sigma^2)$，且 $P(\xi<6)=0.86$，

所以 $P(3<\xi<6)=P(\xi<6)-P(\xi\leqslant 3)=0.86-0.5=0.36$.

**2.** 已知某批产品的质量指标 $X$ 服从正态分布 $N(25,0.16)$，其中 $X\in[24.6,26.2]$ 的产品为“可用产品”，则在这批产品中任取1件，抽到“可用产品”的概率约为________.

0.84 **解析**：由题意知，$X\sim N(25,0.16)$，则 $\mu=25$，$\sigma=0.4$，

所以 $P(24.6\leqslant X\leqslant 26.2)=P(25-0.4\leqslant X\leqslant 25+3\times 0.4)=P(\mu-\sigma\leqslant X\leqslant\mu+3\sigma)\approx\frac{0.682\ 7}{2}+\frac{0.997\ 3}{2}=0.84$，

即抽到“可用产品”的概率约为0.84.

**3.**（2022·新高考全国Ⅱ卷）已知随机变量 $X$ 服从正态分布 $N(2,\sigma^2)$，且 $P(2<X\leqslant 2.5)=0.36$，则 $P(X>2.5)=$________.

0.14 **解析**：因为 $X\sim N(2,\sigma^2)$，所以 $P(X<2)=P(X>2)=0.5$，$P(X>2.5)=P(X>2)-P(2<X\leqslant 2.5)=0.5-0.36=0.14$.

## 任务3 正态分布的实际应用

### 探究活动

**例2** 已知某地外出务工人员的月平均收入服从 $\mu=8\ 000$，$\sigma=500$ 的正态分布.求此地外出务工人员的月平均收入在8 000～8 500元之间的人数所占的百分比.

**解**：设 $X$ 表示此地外出务工人员的月平均收入，

则 $X\sim N(8\ 000,500^2)$，

所以 $P(8\ 000\leqslant X\leqslant 8\ 500)$

$=\frac{1}{2}P(8\ 000-500\leqslant X\leqslant 8\ 000+500)$

$\approx 0.341\ 35$.

所以此地外出务工人员的月平均收入在8 000～8 500元之间的人数所占的百分比为34.135%.

［一题多思］

**思考**.求此地外出务工人员的月平均收入低于7 000元的人数所占的百分比.

**解**：因为 $P(X<7\ 000)=\frac{1}{2}[1-P(8\ 000-2\times 500\leqslant X\leqslant 8\ 000+2\times 500)]\approx\frac{1}{2}\times(1-0.954\ 5)=0.022\ 75$，

所以此地外出务工人员的月平均收入低于7 000元的人数所占的百分比为2.275%.

**【探究总结】**

**正态分布问题的解题策略**

解答此类问题的关键在于将随机变量的取值范围向 $[\mu-\sigma,\mu+\sigma]$，$[\mu-2\sigma,\mu+2\sigma]$，$[\mu-3\sigma,\mu+3\sigma]$ 这三个区间进行转化，然后利用随机变量在上述区间取值的概率求出相应概率，在此过程中一般会用到化归思想及数形结合思想.

### 应用迁移

**1.** 某班有45名学生，最近一次的市联考数学成绩 $X$ 服从正态分布 $N(95,\sigma^2)$.若 $X\in[75,115]$ 的学生人数为18，则 $P(X>115)=$（　　）

A.0.2　　B.0.25

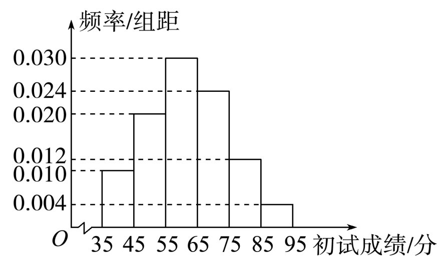
C.0.3　　D.0.35

C **解析**：由题意可设 $P(X>115)=P(X<75)=m$，则 $P(75\leqslant X\leqslant 115)=1-2m$.

又 $X\in[75,115]$ 的学生人数为 $45(1-2m)=18$，解得 $m=0.3$.

故选C.

**2.** 某地区举行专业技能考试，共有8 000人参加，分为初试和复试，初试通过后，才能参加复试.为了解考生的考试情况，随机抽取了100名考生的初试成绩（单位：分），并以此为样本，绘制了样本的频率分布直方图，如图所示.




(1)若所有考生的初试成绩 $X$ 近似服从正态分布 $N(\mu,\sigma^2)$，其中 $\mu$ 为样本平均数的估计值，$\sigma=11.5$，试利用正态分布估计所有考生中初试成绩不低于85分的人数；

(2)复试共四道题，前两道题考生每题答对得5分，答错得0分，后两道题考生每题答对得10分，答错得0分，四道题的总得分为考生的复试成绩.已知某考生进入复试，他在复试中，前两道题每题能答对的概率均为 $\frac{3}{4}$，后两道题每题能答对的概率均为 $\frac{3}{5}$，且每道题回答正确与否互不影响.规定复试成绩在20分以上（含20分）的考生能进入面试，请问该考生进入面试的概率有多大？

附：若随机变量 $X$ 服从正态分布 $N(\mu,\sigma^2)$，则 $P(\mu-\sigma\leqslant X\leqslant\mu+\sigma)\approx 0.682\ 7$，$P(\mu-2\sigma\leqslant X\leqslant\mu+2\sigma)$

$\approx 0.9545$，$P(\mu-3\sigma\leqslant X\leqslant\mu+3\sigma)\approx 0.9973$.

**解**：(1)由题意得，样本平均数的估计值为

$(40\times0.010+50\times0.020+60\times0.030+70\times0.024+80\times0.012+90\times0.004)\times10=62$.

因为考生的初试成绩 $X$ 服从正态分布 $N(\mu,\sigma^2)$，其中 $\mu=62$，$\sigma=11.5$，则 $\mu+2\sigma=85$，

所以 $P(X\geqslant85)=P(X\geqslant\mu+2\sigma)\approx\dfrac{1-0.9545}{2}=0.02275$，

所以估计初试成绩不低于 85 分的人数为 $0.02275\times8000=182$.

(2)记该考生的复试成绩为 $Y$，则能进入面试的复试成绩为 20 分，25 分，30 分，

$P(Y=20)=\left(\dfrac{1}{4}\right)^2\times\left(\dfrac{3}{5}\right)^2+\left(\dfrac{3}{4}\right)^2\times C_2^1\times\dfrac{3}{5}\times\dfrac{2}{5}=\dfrac{117}{400}$，

$P(Y=25)=C_2^1\times\dfrac{3}{4}\times\dfrac{1}{4}\times\left(\dfrac{3}{5}\right)^2=\dfrac{54}{400}$，

$P(Y=30)=\left(\dfrac{3}{4}\right)^2\times\left(\dfrac{3}{5}\right)^2=\dfrac{81}{400}$，

所以该考生进入面试的概率为 $P(Y=20)+P(Y=25)+P(Y=30)=\dfrac{117}{400}+\dfrac{54}{400}+\dfrac{81}{400}=\dfrac{63}{100}$.

### 提质归纳

- 正态分布
  - 概念
  - 正态曲线的特点及性质
    - 范围
    - 对称性
    - 最高点
    - 曲线与$x$轴之间的区域的面积为1
    - 集中、离散程度
  - 标准正态分布
  - $3\sigma$原则

## 课后素养评价（十七）　正态分布

### A组　学习·理解

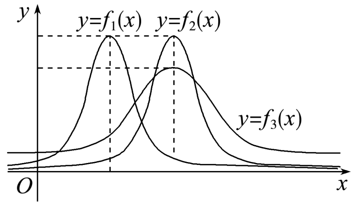
**1.**已知正态密度函数的解析式是 $f(x)=\dfrac{1}{\sigma\sqrt{2\pi}}\cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.设随机变量 $X$ 的概率分布密度函数为 $f(x)=\dfrac{1}{\sqrt{8\pi}}\cdot e^{-\frac{(x-10)^2}{8}}(x\in\mathbf{R})$，则随机变量 $X$ 的均值 $\mu$ 与标准差 $\sigma$ 分别是　(　　)

A.10 与 8　　B.10 与 2

C.8 与 10　　D.2 与 10

B　**解析**：因为 $f(x)=\dfrac{1}{\sqrt{8\pi}}e^{-\frac{(x-10)^2}{8}}=\dfrac{1}{2\sqrt{2\pi}}\cdot e^{-\frac{(x-10)^2}{2\times2^2}}$，所以均值 $\mu=10$，标准差 $\sigma=2$.故选 B.

**2.**已知三个正态密度函数 $f_i(x)=\dfrac{1}{\sqrt{2\pi}\sigma_i}\cdot e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}}(x\in\mathbf{R},i=1,2,3)$的图象如图所示，则下列结论正确的是　(　　)




A.$\mu_1=\mu_2>\mu_3$，$\sigma_1=\sigma_2>\sigma_3$

B.$\mu_1<\mu_2=\mu_3$，$\sigma_1=\sigma_2<\sigma_3$

C.$\mu_1<\mu_2=\mu_3$，$\sigma_1=\sigma_2>\sigma_3$

D.$\mu_1=\mu_2>\mu_3$，$\sigma_1=\sigma_2<\sigma_3$

B　**解析**：根据正态分布密度函数中参数 $\mu$，$\sigma$ 的意义，

由题图可知 $f_2(x)$，$f_3(x)$图象的对称轴位置相同，所以可得 $\mu_2=\mu_3$，

且都在 $f_1(x)$图象的右侧，所以 $\mu_1<\mu_2=\mu_3$；

比较 $f_1(x)$和 $f_2(x)$的图象可得，其形状相同，即 $\sigma_1=\sigma_2$，

又 $f_3(x)$图象的离散程度比 $f_1(x)$和 $f_2(x)$大，所以可得 $\sigma_1=\sigma_2<\sigma_3$.故选 B.

**3.**已知随机变量 $\xi$ 服从正态分布 $N(1,\sigma^2)$.若 $P(\xi>3)=0.012$，则 $P(-1\leqslant\xi\leqslant1)=$　(　　)

A.0.976　　B.0.024

C.0.488　　D.0.048

C　**解析**：因为随机变量 $\xi$ 服从正态分布 $N(1,\sigma^2)$，所以其正态曲线关于直线 $x=1$ 对称.又 $P(\xi>3)=0.012$，所以 $P(\xi<-1)=0.012$.因此 $P(-1\leqslant\xi\leqslant1)=0.5-P(\xi<-1)=0.5-0.012=0.488$.

**4.**设随机变量 $\xi$ 服从正态分布 $N(3,\sigma^2)$.若 $P(\xi\geqslant m)=a$，则 $P(\xi\geqslant6-m)=$　(　　)

A.$a$　　B.$1-2a$

C.$2a$　　D.$1-a$

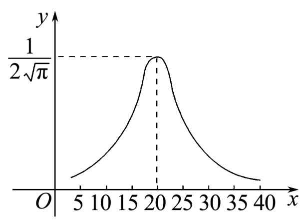
D　**解析**：由直线 $\xi=m$ 与直线 $\xi=6-m$ 关于直线 $\xi=3$ 对称，得 $P(\xi\geqslant m)=P(\xi\leqslant6-m)=a$，则 $P(\xi\geqslant6-m)=1-a$.

**5.** 某早餐店发现加入网络平台后，每天小笼包的销售量 $X \sim N(1\,000, 2\,500)$（单位：个），估计 300 天内小笼包的销售量在 950 到 1 100 个的天数是（　　）

A.236　　B.246　　C.270　　D.275

B　**解析**：由题意可知，$\mu=1\,000$，$\sigma=50$，则 $P(950 \leqslant X \leqslant 1\,100)=P(\mu-\sigma \leqslant X \leqslant \mu+2\sigma)=P(\mu-\sigma \leqslant X \leqslant \mu)+P(\mu \leqslant X \leqslant \mu+2\sigma) \approx \frac{0.682\,7}{2}+\frac{0.954\,5}{2}=0.818\,6$，所以 300 天内小笼包的销售量在 950 到 1 100个的天数大约是 $300 \times 0.818\,6=245.58 \approx 246$. 故选 B.

**6.** 已知随机变量 $X$ 服从正态分布 $N(\mu, \sigma^2)$，其正态曲线如图所示，则 $\mu=$__________，$\sigma^2=$__________.




20　2　**解析**：由题图可知，该正态曲线关于直线 $x=20$ 对称，峰值是 $\frac{1}{2\sqrt{\pi}}$，所以 $\mu=20$，$\frac{1}{\sigma\sqrt{2\pi}}=\frac{1}{2\sqrt{\pi}}$，解得 $\sigma=\sqrt{2}$，因此 $\mu=20$，$\sigma^2=(\sqrt{2})^2=2$.

**7.** 某机器生产的产品质量误差 $X \sim N(1,4)$，$t$ 是 1，2，4，5，7，8，12，15，18，23 的第 60 百分位数，则 $P(-3 \leqslant X \leqslant t-5) \approx$__________.

0.954 5　**解析**：因为 $10 \times 60\%=6$，所以 $t=\frac{8+12}{2}=10$.

由 $X \sim N(1,4)$，可知 $\mu=1$，$\sigma=2$，

所以 $P(-3 \leqslant X \leqslant t-5)=P(-3 \leqslant X \leqslant 5)=P(\mu-2\sigma \leqslant X \leqslant \mu+2\sigma) \approx 0.954\,5$.

**8.** 某种品牌摄像头的使用寿命 $\xi$（单位：年）服从正态分布，且使用寿命不少于 2 年的概率为 0.8，使用寿命不少于 6 年的概率为 0.2.某校在大门口同时安装了两个该品牌的摄像头，则在 4 年内这两个摄像头都能正常工作的概率为__________.

0.25　**解析**：由题意知 $P(\xi \geqslant 2)=0.8$，$P(\xi \geqslant 6)=0.2$，所以 $P(\xi<2)=P(\xi \geqslant 6)=0.2$.所以正态曲线的对称轴为直线 $\xi=4$，即 $P(\xi \geqslant 4)=0.5$，即每个摄像头在 4 年内能正常工作的概率为 0.5.所以两个该品牌的摄像头在 4 年内都能正常工作的概率为 $0.5 \times 0.5=0.25$.

## B组　应用·实践

**1.**（多选）（2024·新高考全国Ⅰ卷）随着“一带一路”国际合作的深入，某茶叶种植区多措并举推动茶叶出口.为了了解推动出口后的亩收入（单位：万元）情况，从该种植区抽取样本，得到推动出口后亩收入的样本均值 $\overline{x}=2.1$，样本方差 $s^2=0.01$.已知该种植区以往的亩收入 $X$ 服从正态分布 $N(1.8, 0.1^2)$，假设推动出口后的亩收入 $Y$ 服从正态分布 $N(\overline{x}, s^2)$，则（若随机变量 $Z$ 服从正态分布 $N(\mu, \sigma^2)$，则 $P(Z<\mu+\sigma) \approx 0.841\,3$）（　　）

A.$P(X>2)>0.2$

B.$P(X>2)<0.5$

C.$P(Y>2)>0.5$

D.$P(Y>2)<0.8$

BC　**解析**：由题意可知，$\overline{x}=2.1$，$s^2=0.01$，

所以 $Y \sim N(2.1, 0.1^2)$.

故 $P(Y>2)=P(Y>2.1-0.1)=P(Y<2.1+0.1) \approx 0.841\,3>0.5$，C 正确，D 错误.

因为 $X \sim N(1.8, 0.1^2)$，

所以 $P(X>2)=P(X>1.8+2 \times 0.1)$.

因为 $P(X<1.8+0.1) \approx 0.841\,3$，

所以 $P(X>1.8+0.1) \approx 1-0.841\,3=0.158\,7<0.2$，

而 $P(X>2)=P(X>1.8+2 \times 0.1)<P(X>1.8+0.1)<0.2$，故 B 正确，A 错误.故选 BC.

**2.** 某校进行了一次体能测试，这次体能测试满分为 100 分，从高三年级抽取 1 000 名学生的测试成绩，已知测试成绩 $\xi$ 服从正态分布 $N(70, \sigma^2)$.若 $\xi$ 在 $(50,70)$ 内取值的概率为 0.4，则 $\xi$ 的值大于 90 的概率为（　　）

A.0.05　　B.0.1　　C.0.2　　D.0.4

B　**解析**：因为 $\xi$ 服从正态分布 $N(70, \sigma^2)$，所以正态曲线的对称轴是直线 $x=70$，所以 $\xi$ 在 $(70,100)$ 内取值的概率为 0.5.因为 $\xi$ 在 $(50,70)$ 内取值的概率为 0.4，所以 $\xi$ 在 $(70,90)$ 内取值的概率也为 0.4，则 $\xi$ 的值大于 90 的概率为 $0.5-0.4=0.1$.故选 B.

**3.**（多选）某市组织了一次高三调研考试，考试后统计的数学成绩服从正态分布，其正态密度函数为 $f(x)=\frac{1}{\sqrt{2\pi} \times 10} e^{-\frac{(x-80)^2}{200}}$，则下列说法正确的是（　　）

A.该市这次考试的数学平均成绩为 80

B.分数在 120 以上的人数与分数在 60 以下的人数相同

C.分数在 110 以上的人数与分数在 50 以下的人数相同

D.该市这次考试的数学成绩的标准差为 10

ACD　**解析**:因为其正态密度函数为 $f(x)=\frac{1}{\sqrt{2\pi}\times10}\cdot e^{-\frac{(x-80)^2}{200}}(x\in\mathbf{R})$,

所以该市这次考试的数学平均成绩为 80,

该市这次考试的数学成绩的标准差为 10.

易知其正态密度曲线关于直线 $x=80$ 对称,

又直线 $x=50$ 与直线 $x=110$ 也关于直线 $x=80$ 对称,

故分数在 110 以上的人数与分数在 50 以下的人数相同.故选 ACD.

**4.**设随机变量 $X\sim N(2,9)$,$P(X>1+c)=P(X<c-1)$.

(1)$c=$______;

(2)$P(-4\leqslant X\leqslant 8)\approx$________.

(1)2　(2)0.954 5　**解析**:(1)由 $X\sim N(2,9)$,可知 $\mu=2$,$\sigma=3$,正态密度曲线关于直线 $x=2$ 对称.

又 $P(X>1+c)=P(X<c-1)$,故有 $\frac{c-1+c+1}{2}=2$,解得 $c=2$.

(2)$P(-4\leqslant X\leqslant 8)=P(2-2\times3\leqslant X\leqslant 2+2\times3)\approx0.954\ 5$.

**5.**某商场开展了一项有奖闯关活动,并对每一关根据难度进行赋分,闯关活动共五关,规定:上一关不通过则不进入下一关,本关第一次未通过有再挑战一次的机会,两次均未通过,则闯关失败,且各关能否通过相互独立.已知甲、乙、丙三人都参加了该项闯关活动.

(1)若甲第一关通过的概率为 $\frac{2}{3}$,第二关通过的概率为 $\frac{5}{6}$,求甲可以进入第三关的概率.

(2)已知所有闯关者的得分服从正态分布,且满分为 450 分,现要根据得分给 2 500 名闯关者中得分前 400 名的闯关者发放奖励.

①假设该闯关活动平均得分为 171 分,351 分以上共有 57 人,已知甲的得分为 270 分,则甲能否获得奖励?请说明理由.

②丙得知他的得分为 430 分,而乙告诉丙:"这次闯关活动平均得分为 201 分,351 分以上共有 57 人."请结合统计学知识帮助丙辨别乙所说信息的真伪.

附:若随机变量 $X\sim N(\mu,\sigma^2)$,则 $P(\mu-\sigma\leqslant X\leqslant\mu+\sigma)\approx0.682\ 7$;$P(\mu-2\sigma\leqslant X\leqslant\mu+2\sigma)\approx0.954\ 5$;$P(\mu-3\sigma\leqslant X\leqslant\mu+3\sigma)\approx0.997\ 3$.

**解**:(1)设 $A_i$:第 $i$ 次通过第一关,$B_i$:第 $i$ 次通过第二关,甲可以进入第三关的概率为 $P$,由题意知 $P=P(A_1B_1)+P(\overline{A_1}A_2B_1)+P(A_1\overline{B_1}B_2)+P(\overline{A_1}A_2\overline{B_1}B_2)$

$=P(A_1)P(B_1)+P(\overline{A_1})P(A_2)P(B_1)+P(A_1)\cdot P(\overline{B_1})P(B_2)+P(\overline{A_1})P(A_2)P(\overline{B_1})P(B_2)$

$=\frac{2}{3}\times\frac{5}{6}+\left(1-\frac{2}{3}\right)\times\frac{2}{3}\times\frac{5}{6}+\frac{2}{3}\times\left(1-\frac{5}{6}\right)\times\frac{5}{6}+\left(1-\frac{2}{3}\right)\times\frac{2}{3}\times\left(1-\frac{5}{6}\right)\times\frac{5}{6}=\frac{70}{81}$.

(2)若闯关者的分数记为 $X$,则 $X\sim N(\mu,\sigma^2)$.

①由题意可知 $\mu=171$,因为 $\frac{57}{2\ 500}=0.022\ 8$,且 $P(X>\mu+2\sigma)=\frac{1-P(\mu-2\sigma\leqslant X\leqslant\mu+2\sigma)}{2}\approx\frac{1-0.954\ 5}{2}\approx0.022\ 8$,

所以 $\mu+2\sigma=351$,则 $\sigma=\frac{351-171}{2}=90$.

而 $\frac{400}{2\ 500}=0.16$,

且 $P(X>\mu+\sigma)=\frac{1-P(\mu-\sigma\leqslant X\leqslant\mu+\sigma)}{2}\approx\frac{1-0.682\ 7}{2}\approx0.158\ 7<0.16$,

所以前 400 名参赛者的最低得分低于 $\mu+\sigma=261$.

因为甲的得分为 270 分,所以甲能够获得奖励.

②假设乙所说为真,则 $\mu=201$,

$P(X\geqslant\mu+2\sigma)=\frac{1-P(\mu-2\sigma\leqslant X\leqslant\mu+2\sigma)}{2}\approx\frac{1-0.954\ 5}{2}\approx0.022\ 8$.

而 $\frac{57}{2\ 500}=0.022\ 8$,所以 $\mu+2\sigma=351$,得 $\sigma=\frac{351-201}{2}=75$,从而 $\mu+3\sigma=201+3\times75=426<430$.

而 $P(X\geqslant\mu+3\sigma)=\frac{1-P(\mu-3\sigma\leqslant X\leqslant\mu+3\sigma)}{2}\approx\frac{1-0.997\ 3}{2}=0.001\ 35<0.005$,

所以 $X\geqslant\mu+3\sigma$ 为小概率事件,即丙的分数为 430 分是小概率事件,可认为其一般不可能发生,但却又发生了,所以可认为乙所说为假.

★★★

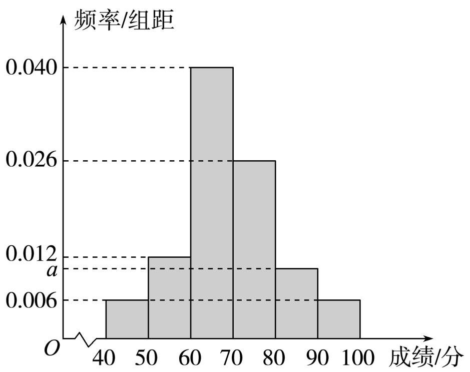
# 迁移应用

**学习目标**

1.明确频率分布直方图中横坐标、纵坐标的含义，理解其反映的频率分布特征.
2.能够根据频率分布直方图确定随机变量的取值及对应的概率，从而构建分布列.
3.掌握二项分布与超几何分布的联系与区别.

## 类型一　频率分布直方图与分布列的综合问题

**例1**　为了让学生了解毒品的危害，加强禁毒教育，某校组织了全体学生参加禁毒知识竞赛，现随机抽取50名学生的成绩(满分100分)进行分析，把他们的成绩分成以下6组：[40,50)，[50,60)，[60,70)，[70,80)，[80,90)，[90,100]，整理得到如图所示的频率分布直方图.




(1)求图中$a$的值并估计全校学生的平均成绩$\mu$；(同一组中的数据用该组区间的中点值代表)

(2)在(1)的条件下，若此次知识竞赛的得分$X\sim N(\mu,12^2)$，为了激发学生学习禁毒知识的兴趣，对参赛学生制定如下奖励方案：得分不超过57分的不予奖励，得分超过57分但不超过81分的可获得学校食堂消费券5元，得分超过81分但不超过93分的可获得学校食堂消费券10元，超过93分的可获得学校食堂消费券15元.试估计全校1 000名学生参加知识竞赛共可获得食堂消费券多少元.(结果四舍五入保留整数)

附：若$X\sim N(\mu,\sigma^2)$，则$P(\mu-\sigma\leqslant X\leqslant\mu+\sigma)\approx0.682\ 7$，$P(\mu-2\sigma\leqslant X\leqslant\mu+2\sigma)\approx0.954\ 5$，$P(\mu-3\sigma\leqslant X\leqslant\mu+3\sigma)\approx0.997\ 3$.

**解**：(1)由题意可知，$(0.006\times2+a+0.012+0.026+0.040)\times10=1$，

解得$a=0.010$.

$\mu=(45+95)\times0.06+55\times0.12+65\times0.40+75\times0.26+85\times0.10=69$.

(2)随机抽取一名学生，设获得的学校食堂消费券为$Y$元，

$P(Y=0)=P(X\leqslant57)\approx0.5-\frac{0.682\ 7}{2}=0.158\ 65$，

$P(Y=5)=P(57<X\leqslant81)\approx0.682\ 7$，

$P(Y=10)=P(81<X\leqslant93)\approx\frac{0.954\ 5-0.682\ 7}{2}=0.135\ 9$，

$P(Y=15)=P(X>93)\approx\frac{1-0.954\ 5}{2}=0.022\ 75$，

所以$Y$的分布列为

| $Y$ | 0 | 5 | 10 | 15 |
|---|---|---|---|---|
| $P$ | 0.158 65 | 0.682 7 | 0.135 9 | 0.022 75 |

即一名学生获得的学校食堂消费券的期望为$E(Y)=0\times0.158\ 65+5\times0.682\ 7+10\times0.135\ 9+15\times0.022\ 75=5.113\ 75$，

所以全校学生可获得食堂消费券$1\ 000\times5.113\ 75=5\ 113.75\approx5\ 114$(元).

故估计全校1 000名学生参加知识竞赛共可获得食堂消费券5 114元.

**【总结升华】**

**频率分布直方图与分布列的综合问题的解题策略**

解题时要正确理解频率分布直方图，能利用频率分布直方图正确计算出各组数据.概率问题以计算为主，往往和实际问题相结合，要注意理解实际问题的意义，使之与相应的概率计算对应起来.

## 类型二 二项分布与超几何分布的综合问题

**例 2** 写出下列离散型随机变量的分布列，并指出其中服从二项分布的是哪些，服从超几何分布的是哪些.

(1)随机变量 $X_1$ 表示 $n$ 次重复抛掷 1 枚骰子出现点数是 3 的倍数的次数；

(2)有一批产品共有 $N$ 件，其中次品有 $M$ 件($N>M>0$)，采用有放回的抽取方法抽取 $n$ 件($n>N$)，抽出的次品件数为随机变量 $X_2$；

(3)有一批产品共有 $N$ 件，其中 $M$ 件为次品，采用不放回的抽取方法抽取 $n$ 件，出现次品的件数为随机变量 $X_3$.($N-M\geqslant n>0$，且 $M\geqslant n$)

**解**：(1)$X_1$ 的分布列为

| $X_1$ | 0 | 1 | 2 | … | $n$ |
|---|---|---|---|---|---|
| $P$ | $C_n^0\left(\frac{1}{3}\right)^0\cdot\left(\frac{2}{3}\right)^n$ | $C_n^1\left(\frac{1}{3}\right)^1\cdot\left(\frac{2}{3}\right)^{n-1}$ | $C_n^2\left(\frac{1}{3}\right)^2\cdot\left(\frac{2}{3}\right)^{n-2}$ | … | $C_n^n\cdot\left(\frac{1}{3}\right)^n$ |

$X_1$ 服从二项分布，即 $X_1\sim B\left(n,\frac{1}{3}\right)$.

(2)$X_2$ 的分布列为

| $X_2$ | 0 | 1 | 2 | … | $n$ |
|---|---|---|---|---|---|
| $P$ | $\left(1-\frac{M}{N}\right)^n$ | $C_n^1\frac{M}{N}\cdot\left(1-\frac{M}{N}\right)^{n-1}$ | $C_n^2\left(\frac{M}{N}\right)^2\cdot\left(1-\frac{M}{N}\right)^{n-2}$ | … | $\left(\frac{M}{N}\right)^n$ |

$X_2$ 服从二项分布，即 $X_2\sim B\left(n,\frac{M}{N}\right)$.

(3)$X_3$ 的分布列为

| $X_3$ | 0 | 1 | … | $n$ |
|---|---|---|---|---|
| $P$ | $\frac{C_{N-M}^n}{C_N^n}$ | $\frac{C_M^1C_{N-M}^{n-1}}{C_N^n}$ | … | $\frac{C_M^n}{C_N^n}$ |

$X_3$ 服从超几何分布.

**【总结升华】**

**超几何分布和二项分布的区别与联系**

| | 超几何分布 | 二项分布 |
|---|---|---|
| 区别 | 需要知道总体的容量 | 不需要知道总体的容量 |
| | 描述的是不放回抽样问题，总体在变化 | 描述的是有放回抽样问题，总体不变化 |
| | 各次抽取不相互独立 | 各次抽取相互独立 |
| | 考察对象分为两类，并且已知各类对象的个数 | 每一次试验都是独立重复试验，两种结果(发生与不发生)、事件发生的概率不变 |
| 联系 | 当总体容量很大时，超几何分布可近似看作二项分布 | |

## 类型三 正态分布模型中的决策问题

**例 3** 随着网络技术的迅速发展，各种购物群成为网络销售的新渠道.在丑橘销售旺季，某丑橘基地随机抽查了 100 个购物群的销售情况，各购物群销售丑橘的数量(单位：箱，取值都在 100 到 600 之间)情况如下：

| 丑橘数量/箱 | [100，200) | [200，300) | [300，400) | [400，500) | [500，600] |
|---|---|---|---|---|---|
| 购物群数量/个 | $a$ | 18 | $a+8$ | $a+20$ | 18 |

(1)求实数 $a$ 的值，并用组中值估计这 100 个购物群销售丑橘数量的平均数.

(2)假设所有购物群销售丑橘的数量 $X$ 服从正态分布 $N(\mu,\sigma^2)$，其中 $\mu$ 为(1)中的平均数，$\sigma^2=12\ 100$.若参与销售该基地丑橘的购物群约有 2 000 个，销售丑橘的数量在[266，596]内的购物群为“一级群”，销售数量小于 266 的购物群为“二级群”，销售数量大于 596 的购物群为“优质群”.该丑橘基地对每个“优质群”奖励 1 000 元，每个“一级群”奖励 200 元，“二级群”不奖励，则该丑橘基地大约需要准备多少元用于奖励？

附：若 $X$ 服从正态分布 $N(\mu,\sigma^2)$，则 $P(\mu-\sigma\leqslant X\leqslant\mu+\sigma)\approx0.682\ 7$，$P(\mu-2\sigma\leqslant X\leqslant\mu+2\sigma)\approx0.954\ 5$，$P(\mu-3\sigma\leqslant X\leqslant\mu+3\sigma)\approx0.997\ 3$.

**解**：(1)由题意得 $a+18+a+8+a+20+18=100$，解得 $a=12$.

故平均数为 $\frac{1}{100}\times(150\times12+250\times18+350\times20+450\times32+550\times18)=376$(箱).

(2)由(1)及题意知，$\mu=376$，$\sigma=110$，且 $266=376-110=\mu-\sigma$，$596=376+220=\mu+2\sigma$，

则 $P(X>596)=P(X>\mu+2\sigma)\approx\frac{1}{2}\times(1-0.954\ 5)=0.022\ 75$，

所以“优质群”约有 $2\ 000\times0.022\ 75\approx46$(个);

$P(266\leqslant X\leqslant596)=P(\mu-\sigma\leqslant X\leqslant\mu+2\sigma)\approx\frac{1}{2}\times0.682\ 7+\frac{1}{2}\times0.954\ 5=0.818\ 6$,

所以“一级群”约有 $2\ 000\times0.818\ 6\approx1\ 637$(个).

所以需要资金约 $46\times1\ 000+1\ 637\times200=373\ 400$(元).

故大约需要准备 373 400 元用于奖励.

**【总结升华】**

**正态分布下两类常见的概率计算**

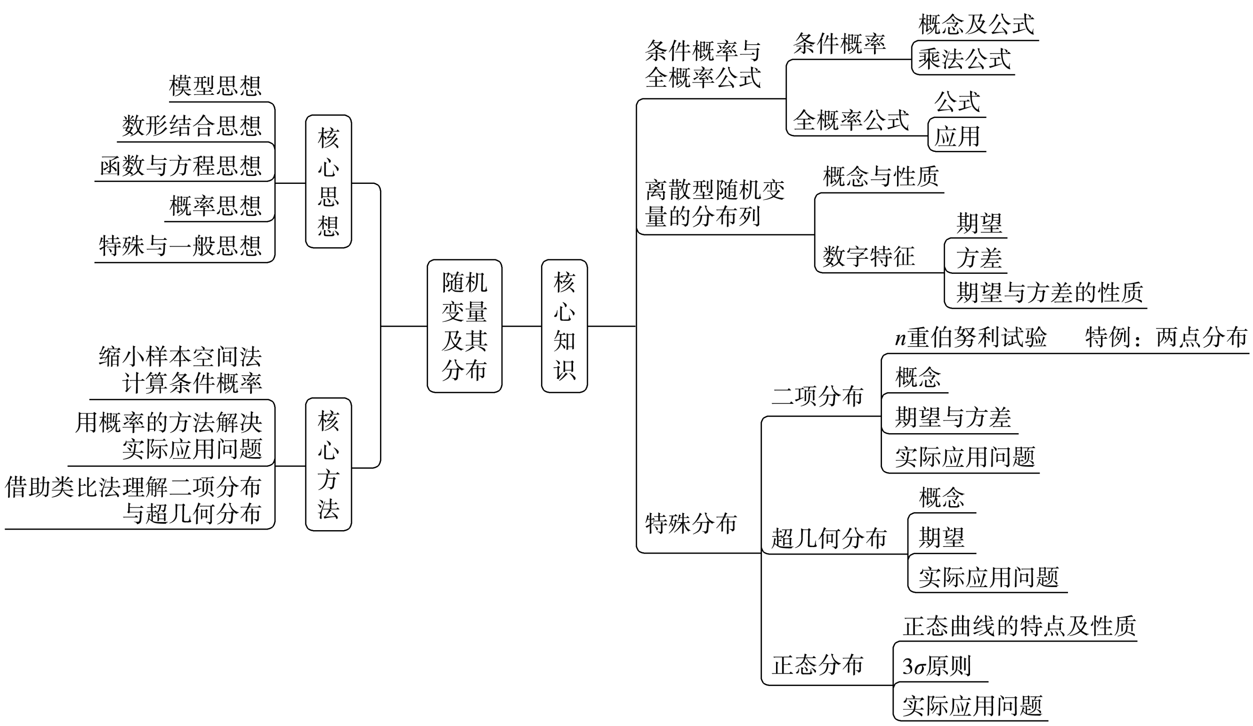
(1)利用正态密度曲线的对称性研究相关概率问题,涉及的知识主要是正态曲线关于直线 $x=\mu$ 对称,曲线与 $x$ 轴之间的区域的面积为 1.

(2)利用 $3\sigma$ 原则求概率时,要注意把给出的区间的端点与 $\mu,\sigma$ 进行对比联系,确定它们属于 $[\mu-\sigma,\mu+\sigma]$,$[\mu-2\sigma,\mu+2\sigma]$,$[\mu-3\sigma,\mu+3\sigma]$ 中的哪一个.

★★★

# 重构拓展

## 多维体系构建




**【问题探究】**

**1.** 如何从条件概率的角度,理解相互独立事件的判断及概率计算公式?

**2.** 当 $P(AB)>0$ 时,$P(ABC)=P(A)P(B|A)\cdot P(C|AB)$.据此你能得出计算 $P(A_1A_2\cdots A_n)$ 的公式吗?

**3.** 在计算离散型随机变量的方差时,如何选择公式简化运算?如何利用方差和标准差分析、解决生活中的实际问题?

**4.** 类比函数性质的研究,你能发现二项分布的哪些性质?提出你的猜想.

## 学科视野拓展

### 拓展一　概率与函数有关的综合问题

**【拓展总结】**

在概率与函数的问题中，决策的工具是样本的数字特征或有关概率.决策方案的最佳选择是将概率最大(最小)或均值最大(最小)的方案作为最佳方案，解题时通常先结合概率、方差、均值的公式列出函数解析式，再利用函数的性质(单调性、最值等)求解.

**应用1**　为降低废气排放量，某工厂生产一种减排器，每件减排器的质量是一等品的概率为$\frac{1}{2}$，是二等品的概率为$\frac{2}{5}$.若达不到一、二等品，则为不合格品.

(1)若工厂已生产3件减排器，设$X$为其中二等品的件数，求$X$的分布列和数学期望.

(2)已知一件减排器的利润如表：

| 等级 | 一等品 | 二等品 | 不合格品 |
|---|---|---|---|
| 利润/万元 | 1 | 0.5 | −0.3 |

①求工厂生产的2件减排器的利润不少于1万元的概率；

②若工厂要增加产量，需引入设备和更新技术，产量增加$n$件，成本相应增加$(n-\ln n)$万元，假设你是工厂的决策者，你觉得目前应不应该增加产量？如果应该增加产量，增加多少件最好？如果不应该增加产量，请说明理由.(参考数据：$\ln 2\approx0.69$，$\ln 3\approx1.1$)

**解**：(1)由题可知$X$服从二项分布，$X$的所有可能取值为0，1，2，3，

$P(X=0)=\left(1-\frac{2}{5}\right)^3=\frac{27}{125}$，

$P(X=1)=C_3^1\times\frac{2}{5}\times\left(1-\frac{2}{5}\right)^2=\frac{54}{125}$，

$P(X=2)=C_3^2\times\left(\frac{2}{5}\right)^2\times\left(1-\frac{2}{5}\right)=\frac{36}{125}$，

$P(X=3)=\left(\frac{2}{5}\right)^3=\frac{8}{125}$，

所以$X$的分布列为

| X | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P$ | $\frac{27}{125}$ | $\frac{54}{125}$ | $\frac{36}{125}$ | $\frac{8}{125}$ |

$E(X)=0\times\frac{27}{125}+1\times\frac{54}{125}+2\times\frac{36}{125}+3\times\frac{8}{125}=\frac{6}{5}$.

(2)①设2件减排器的利润为$Y$万元，

$P(Y\geqslant1)=P(Y=1)+P(Y=1.5)+P(Y=2)=\left(\frac{2}{5}\right)^2+\frac{1}{2}\times\frac{2}{5}\times2+\left(\frac{1}{2}\right)^2=\frac{81}{100}$，

所以2件减排器的利润不少于1万元的概率为$\frac{81}{100}$.

②一件减排器的平均利润为$\frac{1}{2}\times1+\frac{2}{5}\times0.5+\frac{1}{10}\times(-0.3)=0.67$(万元)，

则产量增加$n$件，利润增加$0.67n$万元，成本也相应提高$(n-\ln n)$万元，

所以净利润为$0.67n-n+\ln n=\ln n-0.33n(n\in\mathbf{N}^*)$.

设$f(x)=\ln x-0.33x(x\geqslant1)$，则$f'(x)=\frac{1}{x}-0.33$，

当$1\leqslant x<\frac{100}{33}$时，$f'(x)>0$，当$x>\frac{100}{33}$时，$f'(x)<0$，

所以函数$f(x)$在$\left[1,\frac{100}{33}\right)$上单调递增，在$\left(\frac{100}{33},+\infty\right)$上单调递减，

所以当$x=\frac{100}{33}$时，$f(x)$取得最大值.

因为$n$只能取正整数，又$3<\frac{100}{33}<4$，

$f(3)=\ln 3-0.33\times3\approx1.1-0.99=0.11$，

$f(4)=\ln 4-0.33\times4\approx2\times0.69-1.32=0.06<f(3)$，

所以应该增加产量，增加3件最好.

### 拓展二　概率与马尔科夫链的综合问题

**【拓展总结】**

马尔科夫链具备“无记忆”的性质，即第$(n+1)$次状态的概率分布只跟第$n$次的状态有关，与第$n-1$，$n-2$，$n-3$，…次的状态是“没有任何关系的”.

**应用2**　从甲、乙、丙、丁4人中随机抽取3人去做传球训练.训练规则是确定一人第一次将球传出，每次传球时，传球者都等可能地将球传给另外两人中的任何一人，每次必须将球传出.

(1)记甲、乙、丙三人中被抽到的人数为随机变量$X$，求$X$的分布列.

(2)若刚好抽到甲、乙、丙三人相互做传球训练，且第1次由甲将球传出，记第$n$次传球后球在甲手中的概率为$p_n$，$n=1,2,3,\cdots$.

①直接写出$p_1$，$p_2$，$p_3$的值；

②求 $p_{n+1}$ 与 $p_n(n\in\mathbf{N}^*)$ 的关系式,并求 $p_n$.

**应用 2** **解**:(1)由题意知,$X$ 的可能取值为 2 和 3,

$P(X=2)=\frac{C_3^2}{C_4^3}=\frac{3}{4}$,$P(X=3)=\frac{C_3^3}{C_4^3}=\frac{1}{4}$,

所以随机变量 $X$ 的分布列为

| $X$ | 2 | 3 |
|---|---|---|
| $P$ | $\frac{3}{4}$ | $\frac{1}{4}$ |

(2)①若刚好抽到甲、乙、丙三人相互做传球训练,且第 1 次由甲将球传出,第 $n$ 次传球后球在甲手中的概率为 $p_n$,$n=1,2,3,\cdots$,

则有 $p_1=0$,$p_2=2\times\frac{1}{2}\times\frac{1}{2}=\frac{2}{2^2}=\frac{1}{2}$,$p_3=2\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}=\frac{2}{2^3}=\frac{1}{4}$.

②记事件 $A_n$ 表示"第 $n$ 次传球后,球在甲手中",

$A_{n+1}=\overline{A_n}\cdot A_{n+1}+A_n\cdot A_{n+1}$,

所以 $p_{n+1}=P(\overline{A_n}\cdot A_{n+1}+A_n\cdot A_{n+1})$

$=P(\overline{A_n}\cdot A_{n+1})+P(A_n\cdot A_{n+1})$

$=P(\overline{A_n})\cdot P(A_{n+1}|\overline{A_n})+P(A_n)\cdot P(A_{n+1}|A_n)$

$=(1-p_n)\cdot\frac{1}{2}+p_n\cdot 0=\frac{1}{2}(1-p_n)$,

即 $p_{n+1}=-\frac{1}{2}p_n+\frac{1}{2}$,$n=1,2,3,\cdots$.

所以 $p_{n+1}-\frac{1}{3}=-\frac{1}{2}\left(p_n-\frac{1}{3}\right)$,且 $p_1-\frac{1}{3}=-\frac{1}{3}\neq 0$.

所以数列 $\left\{p_n-\frac{1}{3}\right\}$ 为以 $-\frac{1}{3}$ 为首项,$-\frac{1}{2}$ 为公比的等比数列,

所以 $p_n-\frac{1}{3}=-\frac{1}{3}\times\left(-\frac{1}{2}\right)^{n-1}$,所以 $p_n=-\frac{1}{3}\times\left(-\frac{1}{2}\right)^{n-1}+\frac{1}{3}=\frac{1}{3}\left[1-\left(-\frac{1}{2}\right)^{n-1}\right]=\frac{1}{3}\left[1+\frac{(-1)^n}{2^{n-1}}\right]$,

即第 $n$ 次传球后球在甲手中的概率是 $\frac{1}{3}\left[1+\frac{(-1)^n}{2^{n-1}}\right]$.

## 拓展三 概率中的新定义题

**【拓展总结】**

若 $X\sim B(n,p)$,则 $E(X)=np$,$D(X)=np(1-p)$.

**应用 3** 若随机变量 $X$ 具有数学期望 $E(X)=\mu$,方差 $D(X)=\sigma^2$,则为对任意正数 $\varepsilon$,不等式 $P(|X-\mu|<\varepsilon)\geqslant 1-\frac{\sigma^2}{\varepsilon^2}$ 成立.已知在某通信设备中,信号是由密文"A"和"B"组成的序列,现连续发射信号 $n$ 次,记发射信号"A"的次数为 $X$.

(1)若每次发射信号"A"和"B"的概率是相等的,

①当 $n=5$ 时,求 $P(X\leqslant 2)$;

②为了至少有 98% 的把握使发射信号"A"的频率在 0.4 与 0.6 之间,试估计信号发射次数 $n$ 的最小值.

(2)若每次发射信号"A"和"B"的概率之比是 7∶3,已知在 2 024 次发射中,信号"A"发射 $m$ 次的概率最大,求 $m$ 的值.

**解**:(1)①由题意知,$X\sim B\left(5,\frac{1}{2}\right)$,

所以 $P(X\leqslant 2)=P(X=0)+P(X=1)+P(X=2)$

$=C_5^0\times\left(\frac{1}{2}\right)^5+C_5^1\times\left(\frac{1}{2}\right)^5+C_5^2\times\left(\frac{1}{2}\right)^5=\frac{1}{2}$.

②由题意知,$X\sim B\left(n,\frac{1}{2}\right)$,则 $E(X)=0.5n$,$D(X)=\frac{1}{2}n\times\left(1-\frac{1}{2}\right)=0.25n$.

若 $0.4n\leqslant X\leqslant 0.6n$,则 $-0.1n\leqslant X-0.5n\leqslant 0.1n$,

所以 $P(|X-\mu|<\varepsilon)=P(|X-0.5n|<0.1n)\geqslant 1-\frac{\sigma^2}{\varepsilon^2}=1-\frac{0.25n}{(0.1n)^2}\geqslant 0.98$.

又 $n>0$,解得 $n\geqslant 1\ 250$,即发射次数 $n$ 的最小值为 1 250.

(2)由题意知,$X\sim B(2\ 024,0.7)$,

则 $P(X=m)=C_{2\ 024}^m\times 0.7^m\times 0.3^{2\ 024-m}=\frac{2\ 024!}{(2\ 024-m)!\ m!}\times 0.7^m\times 0.3^{2\ 024-m}$,

$P(X=m+1)=C_{2\ 024}^{m+1}\times 0.7^{m+1}\times 0.3^{2\ 023-m}=\frac{2\ 024!}{(2\ 023-m)!\ (m+1)!}\times 0.7^{m+1}\times 0.3^{2\ 023-m}$,

所以 $\frac{P(X=m+1)}{P(X=m)}$

$$=\frac{\frac{2\ 024!}{(2\ 023-m)!\ (m+1)!}\times 0.7^{m+1}\times 0.3^{2\ 023-m}}{\frac{2\ 024!}{(2\ 024-m)!\ m!}\times 0.7^m\times 0.3^{2\ 024-m}}$$

$=\frac{0.7(2\ 024-m)}{0.3(m+1)}\leqslant 1$,解得 $m\geqslant 1\ 416.5$.

又 $m\in\mathbf{N}^*$,所以当 $m=1\ 417$ 时,$P(X=m)$ 最大.

# 质量评估(二)

(范围:第七章)

## 一、单项选择题

**1.** 下表是离散型随机变量 $X$ 的分布列,则常数 $a$ 的值是 ( )

| $X$ | 3 | 4 | 5 | 9 |
|---|---|---|---|---|
| $P$ | $\frac{a}{2}$ | $\frac{1}{6}+a$ | $\frac{1}{2}$ | $\frac{1}{6}$ |

A.$\frac{1}{6}$　B.$\frac{1}{12}$　C.$\frac{1}{9}$　D.$\frac{1}{2}$

C　**解析**:由题意可得 $\frac{a}{2}+\frac{1}{6}+a+\frac{1}{2}+\frac{1}{6}=1$,解得 $a=\frac{1}{9}$.

**2.** 甲、乙两颗卫星同时独立地监测台风.在同一时刻,甲、乙两颗卫星准确预报台风的概率分别为 0.8 和 0.75,则在同一时刻至少有一颗卫星准确预报的概率为 ( )

A.0.95　B.0.6　C.0.05　D.0.4

A　**解析**:(方法一)在同一时刻至少有一颗卫星准确预报可分为:①甲准确预报,乙没有准确预报;②甲没有准备预报,乙准确预报;③甲准确预报,乙准确预报.这三个事件彼此互斥,故至少有一颗卫星准确预报的概率为 $0.8\times(1-0.75)+(1-0.8)\times0.75+0.8\times0.75=0.95$.

(方法二)"在同一时刻至少有一颗卫星准确预报"的对立事件是"在同一时刻两颗卫星都没有准确预报",故至少有一颗卫星准确预报的概率为 $1-(1-0.8)\times(1-0.75)=0.95$.

**3.** 某地的中学生中有 60% 的同学爱好滑冰,50% 的同学爱好滑雪,70% 的同学爱好滑冰或爱好滑雪.在该地的中学生中随机调查一名同学,若该同学爱好滑雪,则该同学也爱好滑冰的概率为 ( )

A.0.8　B.0.6　C.0.5　D.0.4

A　**解析**:记"该同学爱好滑雪"为事件 $A$,"该同学爱好滑冰"为事件 $B$,

则 $P(A)=0.5$,$P(AB)=0.5+0.6-0.7=0.4$,

所以 $P(B|A)=\frac{P(AB)}{P(A)}=\frac{0.4}{0.5}=0.8$.故选 A.

**4.** 已知随机变量 $X\sim B\left(6,\frac{1}{2}\right)$,则 $D(2X+1)$ 等于 ( )

A.6　B.4　C.3　D.9

A　**解析**:因为 $D(2X+1)=4D(X)$,$D(X)=6\times\frac{1}{2}\times\left(1-\frac{1}{2}\right)=\frac{3}{2}$,所以 $D(2X+1)=4\times\frac{3}{2}=6$.

**5.** 同时掷红、蓝两枚质地均匀的骰子,事件 $A$ 表示"两枚骰子的点数之和为 5",事件 $B$ 表示"红色骰子的点数是偶数",事件 $C$ 表示"两枚骰子的点数相同",事件 $D$ 表示"至少一枚骰子的点数是奇数".给出下列说法:①$A$ 与 $C$ 互斥;②$B$ 与 $D$ 对立;③$A$ 与 $D$ 相互独立;④$B$ 与 $C$ 相互独立.其中正确的是 ( )

A.①③　B.①④　C.②③　D.②④

B　**解析**:①因为两枚骰子的点数相同,所以两枚骰子的点数之和不能为 5,

所以 $A$ 与 $C$ 互斥,因此①说法正确;

②当红色骰子的点数是偶数,蓝色骰子的点数是奇数时,$B$ 与 $D$ 同时发生,

因此这两个事件可以同时发生,所以②说法不正确;

③$P(A)=\frac{4}{36}=\frac{1}{9}$,$P(D)=1-\frac{9}{36}=\frac{3}{4}$,$P(AD)=\frac{4}{36}=\frac{1}{9}$,

显然 $P(A)P(D)\neq P(AD)$,所以 $A$ 与 $D$ 不相互独立,所以③说法不正确;

④$P(B)=\frac{1}{2}$,$P(C)=\frac{1}{6}$,$P(BC)=\frac{3}{36}=\frac{1}{12}$,

显然 $P(B)P(C)=P(BC)$,所以 $B$ 与 $C$ 相互独立,所以④说法正确.

故选 B.

**6.** 若随机变量 $\eta$ 的概率分布列如下:

| $\eta$ | $-2$ | $-1$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|
| $P$ | 0.1 | 0.2 | 0.2 | 0.3 | 0.1 | 0.1 |

则当 $P(\eta<x)=0.8$ 时,实数 $x$ 的取值范围是 ( )

A.$x\leqslant1$　B.$1\leqslant x\leqslant2$

C.$1<x\leqslant2$　D.$1\leqslant x<2$

C　**解析**:因为 $P(\eta\leqslant1)=0.1+0.2+0.2+0.3=0.8$,$P(\eta\leqslant2)=0.1+0.2+0.2+0.3+0.1=0.9$,所以当 $1<x\leqslant2$ 时,$P(\eta<x)=0.8$.故选 C.

**7.** 设某医院仓库中有 10 盒同样规格的 X 光片,已知其中有 5 盒、3 盒、2 盒分别是甲厂、乙厂、丙厂生产

的，且甲、乙、丙三厂生产的该种X光片的次品率分别为$\frac{1}{10}$，$\frac{1}{15}$，$\frac{1}{20}$.现从这10盒中任取1盒，再从这盒中任取1张X光片，则取得的X光片是次品的概率为 ( )

A.0.08 B.0.1 C.0.15 D.0.2

A **解析**：设事件$A_1$，$A_2$，$A_3$分别表示取得的这盒X光片是由甲厂、乙厂、丙厂生产的，$B$表示取得的X光片为次品，

$P(A_1)=\frac{5}{10}$，$P(A_2)=\frac{3}{10}$，$P(A_3)=\frac{2}{10}$，

$P(B|A_1)=\frac{1}{10}$，$P(B|A_2)=\frac{1}{15}$，$P(B|A_3)=\frac{1}{20}$.

由全概率公式，得$P(B)=P(A_1)P(B|A_1)+P(A_2)\cdot P(B|A_2)+P(A_3)P(B|A_3)=\frac{5}{10}\times\frac{1}{10}+\frac{3}{10}\times\frac{1}{15}+\frac{2}{10}\times\frac{1}{20}=0.08$.故选A.

**8.**某次排球比赛的规则是5局3胜制（无平局），甲队与乙队进行比赛，甲队在每局比赛中获胜的概率都为$\frac{2}{3}$，前两局中乙队以2∶0领先，则最后乙队获胜的概率是 ( )

A.$\frac{4}{9}$ B.$\frac{8}{27}$ C.$\frac{19}{27}$ D.$\frac{40}{81}$

C **解析**：最后乙队获胜含3种情况：①第三局乙胜；②第三局甲胜，第四局乙胜；③第三局和第四局都是甲胜，第五局乙胜.故最后乙队获胜的概率$p=\frac{1}{3}+\frac{2}{3}\times\frac{1}{3}+\left(\frac{2}{3}\right)^2\times\frac{1}{3}=\frac{19}{27}$.故选C.

## 二、多项选择题

**9.**设随机变量$X\sim N(\mu,\sigma^2)$且$P(X<1)=\frac{1}{2}$，$P(X>2)=p$，则 ( )

A.正态曲线关于直线$x=1$对称

B.$P(X<0)=P(X>1)$

C.$\sigma=1$

D.$P(0<X<1)=\frac{1}{2}-p$

AD **解析**：由正态曲线的对称性及$P(X<1)=\frac{1}{2}$，知$\mu=1$，即正态曲线关于直线$x=1$对称，于是$P(X<0)=P(X>2)$，所以$P(0<X<1)=P(X<1)-P(X<0)=P(X<1)-P(X>2)=\frac{1}{2}-p$.

**10.**节日期间，某种鲜花进货价是每束2.5元，销售价是每束5元；节日卖不出去的鲜花以每束1.6元处理.根据前五年的销售情况预测，节日期间这种鲜花的需求量$X$的分布列如表所示，则 ( )

| $X$ | 200 | 300 | 400 | 500 |
|---|---|---|---|---|
| $P$ | 0.20 | 0.35 | 0.30 | 0.15 |

A.$P(X>400)=0.45$

B.$E(X)=340$

C.$D(X)=754$

D.进这种鲜花500束，利润的均值为706元

BD **解析**：$P(X>400)=P(X=500)=0.15$.因为$E(X)=200\times0.20+300\times0.35+400\times0.30+500\times0.15=340$，所以$D(X)=(200-340)^2\times0.20+(300-340)^2\times0.35+(400-340)^2\times0.30+(500-340)^2\times0.15=9\ 400$.

若进这种鲜花500束，则利润的均值为$340\times(5-2.5)-(500-340)\times(2.5-1.6)=706$（元）.故选BD.

**11.**有3台车床加工同一型号的零件，第1，2，3台车床加工零件的次品率分别为6%，5%，4%，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的比为5∶6∶9，现任取一个零件，记事件$A_i$="零件为第$i$台车床加工"$(i=1,2,3)$，事件$B$="零件为次品"，则 ( )

A.$P(A_1)=0.25$ B.$P(B|A_2)=\frac{1}{6}$

C.$P(B)=0.048$ D.$P(A_1|B)=\frac{5}{16}$

ACD **解析**：A，B选项，事件$A_i$="零件为第$i$台车床加工"$(i=1,2,3)$，事件$B$="零件为次品"，

则$P(A_1)=\frac{5}{5+6+9}=\frac{1}{4}$，$P(A_2)=\frac{6}{5+6+9}=\frac{3}{10}$，$P(A_3)=\frac{9}{5+6+9}=\frac{9}{20}$，

$P(B|A_1)=6\%$，$P(B|A_2)=5\%$，$P(B|A_3)=4\%$，故A正确，B错误；

C选项，$P(B)=P(A_1B)+P(A_2B)+P(A_3B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+P(A_3)P(B|A_3)=\frac{1}{4}\times6\%+\frac{3}{10}\times5\%+\frac{9}{20}\times4\%=0.048$，故C正确；

D选项，$P(A_1|B)=\frac{P(A_1B)}{P(B)}=\frac{P(B|A_1)P(A_1)}{P(B)}=\frac{6\%\times\frac{1}{4}}{0.048}=\frac{5}{16}$，故D正确.故选ACD.

## 三、填空题

**12.** 设 $P(A|B)=P(B|A)=\frac{1}{2}$，$P(A)=\frac{1}{3}$，则$P(B)$等于________.

$\frac{1}{3}$　**解析**：因为 $P(B|A)=\frac{P(AB)}{P(A)}$，

所以 $P(AB)=P(B|A)\cdot P(A)=\frac{1}{2}\times\frac{1}{3}=\frac{1}{6}$，

所以 $P(B)=\frac{P(AB)}{P(A|B)}=\frac{\frac{1}{6}}{\frac{1}{2}}=\frac{1}{3}$.

**13.** 袋中有除颜色外其余均相同的 4 个红球、3 个黑球，从袋中任取 4 个球，取到 1 个红球得 1 分，取到 1 个黑球得 3 分，设得分为随机变量 $X$，则 $P(X\leqslant 6)=$________.

$\frac{13}{35}$　**解析**：$P(X\leqslant 6)=P(X=4)+P(X=6)$

$=\frac{C_4^4+C_4^3C_3^1}{C_7^4}=\frac{13}{35}$.

**14.** 在等差数列$\{a_n\}$中，$a_4=2$，$a_7=-4$，现从$\{a_n\}$的前 10 项中随机取数，每次取出一个数，取后放回，连续取数 3 次.假设每次取数互不影响，那么在这 3 次取数中，取出的数恰好为两个正数和一个负数的概率为________.(用数字作答)

$\frac{6}{25}$　**解析**：由 $a_4=2$，$a_7=-4$ 可得等差数列$\{a_n\}$的通项公式为 $a_n=10-2n(n\in\mathbf{N}^*)$.$\{a_n\}$的前 10 项分别为 8，6，4，2，0，−2，−4，−6，−8，−10.由题意知 3 次取数相当于 3 次独立重复试验，在每次试验中取得正数的概率为 $\frac{2}{5}$，取得负数的概率为 $\frac{1}{2}$，在 3 次取数中，取出的数恰好为两个正数和一个负数的概率为$C_3^2\times\left(\frac{2}{5}\right)^2\times\left(\frac{1}{2}\right)^1=\frac{6}{25}$.

## 四、解答题

**15.** 在 10 件产品中，有 3 件一等品，4 件二等品，3 件三等品.从这 10 件产品中任取 3 件.求：

(1)取出的 3 件产品中一等品件数 $X$ 的分布列；

(2)取出的 3 件产品中一等品件数多于二等品件数的概率.

**解**：(1)由题意知 $X$ 的所有可能取值为 0，1，2，3，且 $X$ 服从参数为 $N=10$，$M=3$，$n=3$ 的超几何分布，

因此 $P(X=k)=\frac{C_3^kC_7^{3-k}}{C_{10}^3}(k=0,1,2,3)$.

所以 $P(X=0)=\frac{C_3^0C_7^3}{C_{10}^3}=\frac{35}{120}=\frac{7}{24}$，

$P(X=1)=\frac{C_3^1C_7^2}{C_{10}^3}=\frac{63}{120}=\frac{21}{40}$，

$P(X=2)=\frac{C_3^2C_7^1}{C_{10}^3}=\frac{21}{120}=\frac{7}{40}$，

$P(X=3)=\frac{C_3^3C_7^0}{C_{10}^3}=\frac{1}{120}$.

故 $X$ 的分布列为

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P$ | $\frac{7}{24}$ | $\frac{21}{40}$ | $\frac{7}{40}$ | $\frac{1}{120}$ |

(2)设“取出的 3 件产品中一等品件数多于二等品件数”为事件 $A$，“恰好取出 1 件一等品和 2 件三等品”为事件 $A_1$，“恰好取出 2 件一等品”为事件 $A_2$，“恰好取出 3 件一等品”为事件 $A_3$.

由于事件 $A_1$，$A_2$，$A_3$ 彼此互斥，且 $A=A_1+A_2+A_3$，

而 $P(A_1)=\frac{C_3^1C_3^2}{C_{10}^3}=\frac{3}{40}$，$P(A_2)=P(X=2)=\frac{7}{40}$，

$P(A_3)=P(X=3)=\frac{1}{120}$，

所以取出的 3 件产品中一等品件数多于二等品件数的概率 $P(A)=P(A_1)+P(A_2)+P(A_3)=\frac{3}{40}+\frac{7}{40}+\frac{1}{120}=\frac{31}{120}$.

**16.** 一名学生每天骑车上学，从他家到学校的途中有 5 个路口，假设他在各个路口遇到红灯的事件是相互独立的，并且概率都是 $\frac{1}{3}$.

(1)设 $X$ 为这名学生在途中遇到红灯的次数，求 $X$ 的分布列、期望、方差；

(2)设 $Y$ 为这名学生在首次遇到红灯前经过的路口数，求 $Y$ 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**：(1)由题意可知，$X$ 的可能取值为 0，1，2，3，4，5，且服从二项分布 $B\left(5,\frac{1}{3}\right)$，则

$P(X=0)=C_5^0\times\left(\frac{1}{3}\right)^0\times\left(\frac{2}{3}\right)^5=\frac{32}{243}$，

$P(X=1)=C_5^1\times\left(\frac{1}{3}\right)^1\times\left(\frac{2}{3}\right)^4=\frac{80}{243}$，

$P(X=2)=C_5^2\times\left(\frac{1}{3}\right)^2\times\left(\frac{2}{3}\right)^3=\frac{80}{243}$，

$P(X=3)=C_5^3\times\left(\frac{1}{3}\right)^3\times\left(\frac{2}{3}\right)^2=\frac{40}{243}$，

$P(X=4)=C_5^4\times\left(\frac{1}{3}\right)^4\times\left(\frac{2}{3}\right)^1=\frac{10}{243}$,

$P(X=5)=C_5^5\times\left(\frac{1}{3}\right)^5\times\left(\frac{2}{3}\right)^0=\frac{1}{243}$.

由此得 $X$ 的分布列为

| $X$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $P$ | $\frac{32}{243}$ | $\frac{80}{243}$ | $\frac{80}{243}$ | $\frac{40}{243}$ | $\frac{10}{243}$ | $\frac{1}{243}$ |

$E(X)=5\times\frac{1}{3}=\frac{5}{3}$，$D(X)=5\times\frac{1}{3}\times\frac{2}{3}=\frac{10}{9}$.

(2)由于 $Y$ 为这名学生在首次遇到红灯前经过的路口数，显然 $Y$ 是随机变量，其取值为 0，1，2，3，4，5，且

$P(Y=0)=\frac{1}{3}$,

$P(Y=1)=\left(\frac{2}{3}\right)^1\times\frac{1}{3}=\frac{2}{9}$,

$P(Y=2)=\left(\frac{2}{3}\right)^2\times\frac{1}{3}=\frac{4}{27}$,

$P(Y=3)=\left(\frac{2}{3}\right)^3\times\frac{1}{3}=\frac{8}{81}$,

$P(Y=4)=\left(\frac{2}{3}\right)^4\times\frac{1}{3}=\frac{16}{243}$,

$P(Y=5)=\left(\frac{2}{3}\right)^5=\frac{32}{243}$.

由此得 $Y$ 的分布列为

| $Y$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $P$ | $\frac{1}{3}$ | $\frac{2}{9}$ | $\frac{4}{27}$ | $\frac{8}{81}$ | $\frac{16}{243}$ | $\frac{32}{243}$ |

(3)设“这名学生在途中至少遇到一次红灯”为事件 $A$，

则 $P(A)=1-P(X=0)=1-\frac{32}{243}=\frac{211}{243}$.

**17.**(2024·北京卷)某保险公司为了解该公司某种保险产品的索赔情况，从合同保险期限届满的保单中随机抽取 1 000 份，记录并整理这些保单的索赔情况，获得数据如下表：

| 索赔次数 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| 保单份数 | 800 | 100 | 60 | 30 | 10 |

假设：一份保单的保费为 0.4 万元；前三次索赔时，保险公司每次赔偿 0.8 万元；第四次索赔时，保险公司赔偿 0.6 万元.

假设不同保单的索赔次数相互独立，用频率估计概率.

(1)估计一份保单索赔次数不少于 2 的概率.

(2)一份保单的毛利润定义为这份保单的保费与赔偿总金额之差.

①记 $X$ 为一份保单的毛利润，估计 $X$ 的数学期望 $E(X)$；

②如果无索赔的保单的保费减少 4%，有索赔的保单的保费增加 20%，试比较这种情况下一份保单毛利润的数学期望估计值与①中 $E(X)$ 估计值的大小.(结论不要求证明)

**解**：(1)(方法一：正面计算)记“随机抽取一份保单，索赔次数不少于2”为事件 $A$，

由索赔次数不少于 2 知，索赔次数为 2，3，4，

所以 $P(A)=\frac{60+30+10}{1\ 000}=\frac{100}{1\ 000}=\frac{1}{10}$.

(方法二：反面计算)记“随机抽取一份保单，索赔次数不少于2”为事件 $A$，

则 $P(A)=1-\frac{800+100}{1\ 000}=\frac{1}{10}$.

(2)①由题意知 $X$ 的所有可能取值为 0.4，$-0.4$，$-1.2$，$-2.0$，$-2.6$，

$P(X=0.4)=\frac{800}{1\ 000}=0.8$,

$P(X=-0.4)=\frac{100}{1\ 000}=0.1$,

$P(X=-1.2)=\frac{60}{1\ 000}=0.06$,

$P(X=-2.0)=\frac{30}{1\ 000}=0.03$,

$P(X=-2.6)=\frac{10}{1\ 000}=0.01$,

故 $E(X)=0.4\times0.8-0.4\times0.1-1.2\times0.06-2.0\times0.03-2.6\times0.01=0.122$.

②如果无索赔的保单的保费减少 4%，有索赔的保单的保费增加 20%，这种情况下一份保单毛利润的数学期望估计值比①中 $E(X)$ 估计值大.

**18.**(2024·新课标全国Ⅱ卷)某投篮比赛分为两个阶段，每个参赛队由两名队员组成.比赛具体规则如下：第一阶段由参赛队中一名队员投篮 3 次，若 3 次都未投中，则该队被淘汰，比赛成绩为 0 分；若至少投中一次，则该队进入第二阶段.第二阶段由该队的另一名队员投篮 3 次，每次投篮投中得 5 分，未投中得 0 分，该队的比赛成绩为第二阶段的得分总和.

某参赛队由甲、乙两名队员组成，设甲每次投中的概率为 $p$，乙每次投中的概率为 $q$，各次投中与否相互独立.

(1)若 $p=0.4$，$q=0.5$，甲参加第一阶段比赛，求甲、乙所在队的比赛成绩不少于 5 分的概率.

(2)假设 $0<p<q$.

①为使得甲、乙所在队的比赛成绩为 15 分的概率

最大，应该由谁参加第一阶段比赛？

②为使得甲、乙所在队的比赛成绩的数学期望最大，应该由谁参加第一阶段比赛？

**解**：(1)甲、乙所在队的比赛成绩不少于5分，则甲第一阶段至少投中1次，乙第二阶段也至少投中1次，

所以比赛成绩不少于5分的概率为

$P=(1-0.6^3)\times(1-0.5^3)=0.686$.

(2)①若甲先参加第一阶段比赛，则甲、乙所在队的比赛成绩为15分的概率为 $P_{甲}=[1-(1-p)^3]q^3$；

若乙先参加第一阶段比赛，则甲、乙所在队的比赛成绩为15分的概率为 $P_{乙}=[1-(1-q)^3]p^3$.

所以 $P_{甲}-P_{乙}=q^3-(q-pq)^3-p^3+(p-pq)^3$

$=(q-p)(q^2+pq+p^2)+(p-q)[(p-pq)^2+(q-pq)^2+(p-pq)(q-pq)]$

$=(p-q)(3p^2q^2-3p^2q-3pq^2)$

$=3pq(p-q)(pq-p-q)$

$=3pq(p-q)[(1-p)(1-q)-1]$.

因为 $0<p<q$，又 $p<1,q<1$，所以 $p-q<0$，$(1-p)(1-q)-1<0$，

所以 $pq(p-q)[(1-p)(1-q)-1]>0$，

所以 $P_{甲}>P_{乙}$，应该由甲参加第一阶段比赛.

②若甲先参加第一阶段比赛，比赛成绩 $X$ 的所有可能取值为0，5，10，15，

$P(X=0)=(1-p)^3+[1-(1-p)^3]\cdot(1-q)^3$，

$P(X=5)=[1-(1-p)^3]\cdot C_3^1q(1-q)^2$，

$P(X=10)=[1-(1-p)^3]\cdot C_3^2q^2(1-q)$，

$P(X=15)=[1-(1-p)^3]\cdot q^3$.

所以 $E(X)=15[1-(1-p)^3]q$

$=15(p^3-3p^2+3p)\cdot q$.

若乙先参加第一阶段比赛，比赛成绩 $Y$ 的所有可能取值为0，5，10，15，

同理 $E(Y)=15(q^3-3q^2+3q)\cdot p$，

所以 $E(X)-E(Y)=15[pq(p+q)(p-q)-3pq(p-q)]=15pq(p-q)(p+q-3)$.

因为 $0<p<q$，$p<1$，$q<1$，所以 $p-q<0$，$p+q-3<1+1-3<0$，

所以 $pq(p-q)(p+q-3)>0$，所以 $E(X)>E(Y)$，

所以应该由甲参加第一阶段比赛.

**19.** 投掷四枚不同的硬币A，B，C，D，假定A，B两枚正面向上的概率均为 $\frac{1}{2}$，另两枚C，D为非均匀硬币，正面向上的概率均为 $a(0<a<1)$，把这四枚硬币各投掷一次，设 $X$ 表示正面向上的枚数.

(1)若A，B出现一枚正面向上、一枚反面向上与C，D出现两枚正面向上的概率相等，求 $a$ 的值；

(2)求 $X$ 的分布列及数学期望(用 $a$ 表示).

**解**：(1)由题意得 $2\times\frac{1}{2}\times\left(1-\frac{1}{2}\right)=a^2$，$0<a<1$，

所以 $a=\frac{\sqrt{2}}{2}$.

(2) $X$ 的可能取值为0，1，2，3，4.

$P(X=0)=C_2^0\times\left(1-\frac{1}{2}\right)^2\times C_2^0(1-a)^2=\frac{1}{4}(1-a)^2$，

$P(X=1)=C_2^1\times\left(\frac{1}{2}\right)^1\times\left(1-\frac{1}{2}\right)^1\times C_2^0(1-a)^2+C_2^0\times\left(1-\frac{1}{2}\right)^2\times C_2^1a(1-a)=\frac{1}{2}(1-a)$，

$P(X=2)=C_2^2\times\left(\frac{1}{2}\right)^2\times C_2^0(1-a)^2+C_2^1\times\left(\frac{1}{2}\right)^1\times\left(1-\frac{1}{2}\right)^1\times C_2^1a(1-a)+C_2^0\times\left(1-\frac{1}{2}\right)^2\times C_2^2a^2=\frac{1}{4}(1+2a-2a^2)$，

$P(X=3)=C_2^2\times\left(\frac{1}{2}\right)^2\times C_2^1a(1-a)+C_2^1\times\left(\frac{1}{2}\right)^1\times\left(1-\frac{1}{2}\right)^1\times C_2^2a^2=\frac{a}{2}$，

$P(X=4)=C_2^2\times\left(\frac{1}{2}\right)^2\times C_2^2a^2=\frac{1}{4}a^2$.

所以 $X$ 的分布列为

| $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P$ | $\frac{1}{4}(1-a)^2$ | $\frac{1}{2}(1-a)$ | $\frac{1}{4}(1+2a-2a^2)$ | $\frac{a}{2}$ | $\frac{1}{4}a^2$ |

$E(X)=1\times\frac{1}{2}(1-a)+2\times\frac{1}{4}(1+2a-2a^2)+3\times\frac{a}{2}+4\times\frac{1}{4}a^2=2a+1$.

# 第八章 成对数据的统计分析

## 单元概览

### 单元导航

本章在必修课程统计内容的基础上，通过成对样本数据研究两个随机变量之间的关系，在本章的学习过程中，将进一步感悟根据实际问题进行科学决策的必要性和可能性；体会统计思维与确定性思维的差异、归纳推断与演绎证明的差异；积累数据分析的经验，培养数据分析、数学建模、逻辑推理等素养.

### 学习目标

**1.**掌握成对样本数据的直观表示方法及线性相关统计特征的刻画方法，能够根据成对数据的统计相关性推断两个随机变量的相关性，分析解决与统计相关的简单实际问题.

**2.**结合具体实例，理解一元线性回归分析的方法，会用一元线性回归模型刻画两个变量之间的随机关系，并进行预测.

**3.**理解 $2\times2$ 列联表的统计意义，会用 $2\times2$ 列联表的方法解决两个随机变量独立性检验的简单实际问题.

### 核心概念

成对数据的统计相关性是对两个随机变量之间相关关系进行分析的一种统计方法，从数量上准确刻画了两个变量的相关程度.

当两个变量之间具有显著的线性相关关系时，可建立一元线性回归模型刻画两个变量之间的相关关系；通过计算 $\chi^2$ 值，进而分析相关性结论的可信程度，提升数学运算、数据分析素养.

### 学法指导

**1.**在学习变量的相关关系时，从散点图入手，认知相关系数的作用，逐步深入到理论层次，在逐步深入的过程中，不断积累知识和经验，直至掌握一元线性回归模型及其应用，提高逻辑推理素养.

**2.**在学习列联表时，要结合具体实例，利用等高堆积条形图，判断两个分类变量的相关性，提升数学抽象、数学运算、数学建模素养.

**3.**在学习独立性检验时，学会用随机变量 $\chi^2$ 取值的大小作为判断零假设 $H_0$ 是否成立的依据，提升数学抽象、数学运算素养.

## 单元主题任务

某校提倡学生进行体育锻炼，为了有针对性地提高学生体育锻炼的积极性，该校对学生是否经常锻炼的情况进行了调查.

(1)全校学生的普查数据为：523 名女生中有 331 名经常锻炼；601 名男生中有 473 名经常锻炼.你能利用这些数据，说明该校女生和男生在体育锻炼的经常性方面是否存在差异吗？

(2)调查学生每天课后锻炼时间 $x$(单位：min)和他们的数学成绩 $y$(单位：分)之间的关系，得到以下数据.

| 锻炼时间 $x$ | 20 | 30 | 40 | 50 | 60 |
| --- | --- | --- | --- | --- | --- |
| 数学成绩 $y$ | 59 | 82 | 72 | 110 | 97 |

对上述数据进行分析，$y$ 与 $x$ 之间是否具有相关关系？

★★★

# 探究构建

## 8.1 成对数据的统计相关性

| 学习任务目标 |
|---|
| 1.理解两个变量的相关关系的概念. |
| 2.会作散点图,并利用散点图判断两个变量之间是否具有相关关系. |
| 3.会根据样本相关系数判断两个变量的线性相关程度. |

## ◦ 问题式预习 ◦

### 【知识清单】

**知识点一 变量的相关关系**

(1)两个变量有关系,但又没有确切到可由其中的一个去精确地决定另一个的程度,这种关系称为相关关系.

(2)散点图:将样本中 $n$ 个数据点 $(x_i,y_i)(i=1,2,\cdots,n)$ 描在直角坐标系中得到的统计图.

(3)从整体上看,当一个变量的值增加时,另一个变量的相应值也呈现增加的趋势,我们就称这两个变量正相关;当一个变量的值增加时,另一个变量的相应值呈现减小的趋势,则称这两个变量负相关.

(4)一般地,如果两个变量的取值呈现正相关或负相关,而且散点落在一条直线附近,我们就称这两个变量线性相关;如果两个变量具有相关性,但不是线性相关,那么我们就称这两个变量非线性相关或曲线相关.

**知识点二 样本相关系数**

| | | |
|---|---|---|
| 计算 | | $r=\dfrac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}}=\dfrac{\sum\limits_{i=1}^{n}x_iy_i-n\overline{x}\,\overline{y}}{\sqrt{\left(\sum\limits_{i=1}^{n}x_i^2-n\overline{x}^2\right)\left(\sum\limits_{i=1}^{n}y_i^2-n\overline{y}^2\right)}}$ |
| 性质 | 范围 | $-1\leqslant r\leqslant 1$ |
| | 线性相关程度 | (1)成对样本数据正相关的充要条件是 $r>0$,成对样本数据负相关的充要条件是 $r<0$.<br>(2)当 $\lvert r\rvert$ 越接近 1 时,成对样本数据的线性相关程度越强;<br>当 $\lvert r\rvert$ 越接近 0 时,成对样本数据的线性相关程度越弱.<br>(3) $\lvert r\rvert=1$ 的充要条件是成对样本数据确定的点都在一条直线上.<br>(4)当 $r=0$ 时,表明成对样本数据间没有线性相关关系 |

### 【概念辨析】

1.判断正误(正确的打"√",错误的打"×").

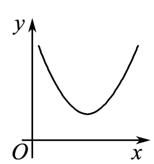
(1)两个变量之间产生相关关系的原因受许多不确定的随机因素的影响. ( √ )

(2)正方体的棱长和体积是相关关系. ( × )

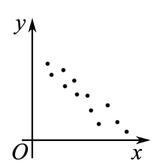
(3)两个变量的样本相关系数越大,它们的相关程度越强. ( × )

(4)若样本相关系数 $r=0$,则两变量之间没有关系. ( × )

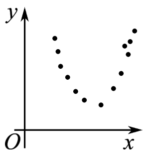
2.(多选)下列各图中所示的两个变量具有相关关系的是 ( )

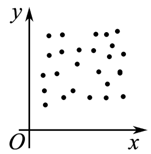


A



B





C　　D

BC　**解析**：A 为函数关系；B，C 为相关关系；D 中，因为点分布得比较分散，所以两者之间无相关关系.

**3.**已知两个变量负相关，且相关程度很强，则它们的样本相关系数可能是　（　　）

A.$-0.95$　　B.$-0.13$

C.$0.15$　　D.$0.96$

A　**解析**：样本相关系数 $r<0$ 时，成对样本数据负相关，且$|r|$越大，成对样本数据间的线性相关程度越强.

**4.**请思考并回答下列问题：

(1)相关关系与函数关系有何区别与联系？

**提示**：区别：

①函数关系中两个变量间是一种确定性关系.函数值由自变量的值唯一确定；

②相关关系中两个变量间是一种不确定性关系.例如，身高与体重之间的关系，两者之间虽然没有确定的函数关系，但身高高的人往往体重会更重些，两者之间是一种非确定性关系.

联系：

①两种关系在现实生活中均存在.客观上讲，函数关系是一种理想的关系模型，而相关关系是一种更为实际的情况；

②在一定条件下两种关系可以相互转化.有些相关关系可以用函数关系进行估计或推断.

(2)散点图有何特点？在数据分析中有何作用？

**提示**：具有直观、简明的特点，它能形象地体现成对数据的分布情况，并且可以根据散点图来大致推断两个变量之间是否存在相关关系、是正相关还是负相关、是线性相关还是非线性相关等.

## ○ 任务型课堂 ○

### 任务1 变量间相关关系的推断

**1.**下列变量间的关系，不是相关关系的是　（　　）

A.农田的水稻产量与施肥量之间的关系

B.正方形的面积与边长之间的关系

C.商品销售收入与其广告费支出之间的关系

D.人体内的脂肪含量与年龄之间的关系

B　**解析**：A 选项，水稻产量与施肥量之间没有明确的等量关系，是相关关系；

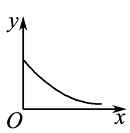
B 选项，正方形的面积与边长之间有着明确的等量关系，不是相关关系；

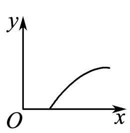
C 选项，商品销售收入与其广告费支出之间没有明确的等量关系，是相关关系；

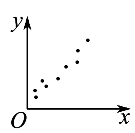
D 选项，人体内的脂肪含量与年龄之间没有明确的等量关系，是相关关系.

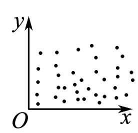
故选 B.

**2.**下列图形中的两个变量具有相关关系的是　（　　）









A　B　C　D

C　**解析**：A，B 为函数关系，D 无相关关系.

**3.**下列关系是相关关系的是______.(填序号)

①曲线上的点与该点的坐标之间的关系；

②苹果的产量与气候之间的关系；

③森林中同一种树木，其胸径与高度之间的关系；

④学生与其学号之间的关系.

②③　**解析**：利用相关关系的概念进行判断.①④中两个变量之间的关系是一种确定性关系，而②③中的两个变量之间是不确定性关系，它们具有相关关系.

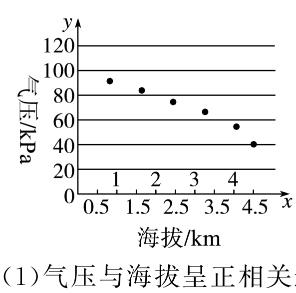
**【探究总结】**

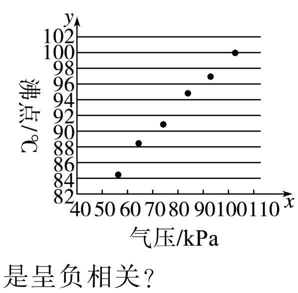
具体问题中，可借助积累的生活经验进行分析，判断两个变量是否具有相关关系.

### 任务2 变量正负相关的推断

**探究活动**

**例 1**　某校地理学习兴趣小组根据在某座山上测得的海拔、气压和沸点的六组数据绘制了散点图，如图所示.




(1)气压与海拔呈正相关还是呈负相关？

(2)沸点与气压呈正相关还是呈负相关？

(3)沸点与海拔呈正相关还是呈负相关？

**解**：(1)由题左图知气压随海拔的增加而降低，所以气压与海拔呈负相关.

(2)由题右图知沸点随气压的升高而升高，所以沸点与气压呈正相关.

(3)气压随海拔的增加而降低，沸点随气压的升高而

升高，所以沸点随海拔的增加而降低，所以沸点与海拔呈负相关.

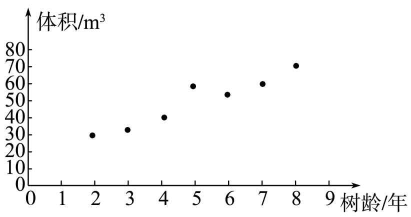
**例 2** 某种树木树干的体积（单位：$m^3$）与树龄（单位：年）之间有如下的对应关系：

| 树龄/年 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|
| 体积/$m^3$ | 30 | 34 | 40 | 60 | 55 | 62 | 70 |

(1)请作出这些数据的散点图；

(2)你能由散点图发现树干体积与树龄呈什么关系吗？

**解**：(1)以 $x$ 轴表示树龄，$y$ 轴表示树干的体积，可得相应的散点图如图所示.

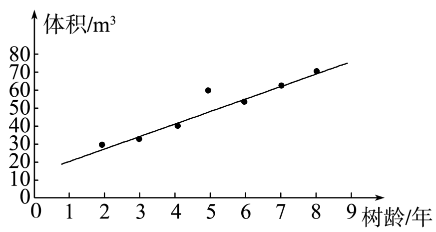



(2)由散点图发现树干体积随着树龄的增加呈现增加的趋势，且散点大致落在一条直线附近，所以树干的体积与树龄呈线性相关关系，且为正相关.

［一题多思］

**思考 1.** 若树干的体积与树龄呈线性相关关系，请画出一条直线来近似地表示这种线性相关关系.

**解**：如图所示.

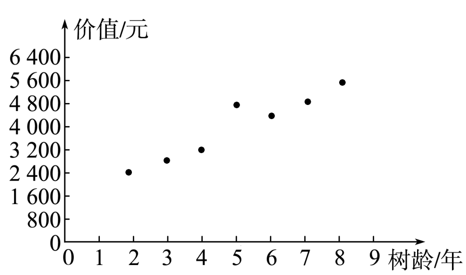



**思考 2.** 若该种树木的树干每立方米的价值是 80 元，作出树干的价值与树龄之间关系的散点图.

**解**：树干的价值与树龄之间的关系如表所示.

| 树龄/年 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|
| 体积/$m^3$ | 30 | 34 | 40 | 60 | 55 | 62 | 70 |
| 价值/元 | 2 400 | 2 720 | 3 200 | 4 800 | 4 400 | 4 960 | 5 600 |

以 $x$ 轴表示树龄，$y$ 轴表示树干的价值，可得相应的散点图如图所示.




**【探究总结】**

**1.** 判断两个变量 $x$ 和 $y$ 之间具有哪种相关关系，最简便的方法是绘制散点图.变量之间可能是线性相关的，也可能是非线性相关的，还可能不相关.

**2.** 画散点图时应注意合理选择单位长度，避免图形偏大或偏小，或者是点的坐标在坐标系中画不准，使图形失真，导致得出错误结论.

## 应用迁移

**1.** 某商场五天内某种衬衫的销售情况如下表：

| 日期编号 $x$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 销售量 $y$/件 | 19 | 39 | 59 | 79 | 104 |

则下列说法正确的是 （　　）

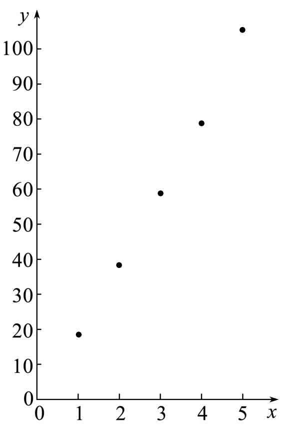
A. $y$ 与 $x$ 负相关

B. $y$ 与 $x$ 正相关

C. $y$ 与 $x$ 不相关

D. $y$ 与 $x$ 成正比例关系

B **解析**：根据题表中的数据作出散点图如图所示，



可知所有点都在一条直线附近波动，所以 $y$ 与 $x$ 是线性相关的，且 $y$ 的值随着 $x$ 的值的增大而增大，即 $y$ 与 $x$ 正相关.故选 B.

**2.** 对于变量 $x,y$，由观测数据 $(x_i,y_i)(i=1,2,3,\cdots,10)$ 得散点图如图 1；对于变量 $u,v$ 由观测数据 $(u_i,v_i)(i=1,2,3,\cdots,10)$ 得散点图如图 2.由这两个散点图可以推断 （　　）

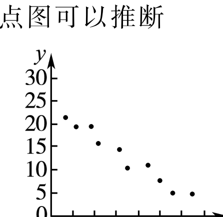


图1

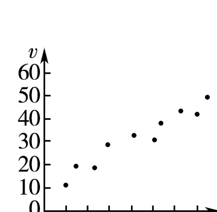


图2

A. $x$ 与 $y$ 正相关，$u$ 与 $v$ 正相关

B. $x$ 与 $y$ 正相关，$u$ 与 $v$ 负相关

C. $x$ 与 $y$ 负相关，$u$ 与 $v$ 正相关

D. $x$ 与 $y$ 负相关，$u$ 与 $v$ 负相关

C **解析**：由题图 1 可知，点散布在从左上角到右下

角的区域，各点整体呈下降趋势，故 $x$ 与 $y$ 负相关；由题图 2 可知，点散布在从左下角到右上角的区域，各点整体呈上升趋势，故 $u$ 与 $v$ 正相关.

## 任务3 样本相关系数与线性相关程度

### 探究活动

**例 3**　甲、乙、丙、丁四位同学各自对 $a,b$ 两变量的线性相关性做试验，并分别求得样本相关系数 $r$ 如下表：

| 甲 | 乙 | 丙 | 丁 |
|---|---|---|---|
| $-0.82$ | $-0.78$ | $-0.69$ | $-0.85$ |

则________同学的试验结果体现 $a,b$ 两变量有更强的线性相关性.

丁　**解析**：因为 $0.85>0.82>0.78>0.69$，已知相关系数的绝对值越接近 1，则两个变量的线性相关性越强，所以能体现出 $a,b$ 两变量有更强的线性相关性的是丁同学的试验结果.

**例 4**　直播带货是一种直播和电商相结合的销售手段，目前已被广大消费者所接受.针对这种现状，某公司决定逐月加大直播带货的投入，直播带货金额稳步提升，以下是该公司 2025 年前 5 个月的带货金额(单位:万元)统计表.

| 月份 | 1月 | 2月 | 3月 | 4月 | 5月 |
|---|---|---|---|---|---|
| 月份编号 $x$ | 1 | 2 | 3 | 4 | 5 |
| 带货金额 $y$/万元 | 7 | 12 | 13 | 19 | 24 |

(1)求该公司带货金额的平均数 $\overline{y}$；

(2)求该公司带货金额 $y$ 与月份编号 $x$ 的样本相关系数(精确到 0.01)，并判断它们是否具有线性相关关系($0.75\leqslant|r|\leqslant1$，则认为 $y$ 与 $x$ 有很强的线性相关关系；$|r|<0.75$，则认为 $y$ 与 $x$ 的线性相关性较弱).

参考公式：$r=\dfrac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}}$.

参考数据：$\sqrt{1\ 740}\approx41.7$，$\sum\limits_{i=1}^{5}(x_i-\overline{x})(y_i-\overline{y})=41$，$\sqrt{\sum\limits_{i=1}^{5}(x_i-\overline{x})^2}=\sqrt{10}$，$\sqrt{\sum\limits_{i=1}^{5}(y_i-\overline{y})^2}=\sqrt{174}$.

**解**：(1)由题表中数据可得 $\overline{y}=\dfrac{7+12+13+19+24}{5}=15$.

(2)由于 $\sum\limits_{i=1}^{5}(x_i-\overline{x})(y_i-\overline{y})=41$，$\sqrt{\sum\limits_{i=1}^{5}(y_i-\overline{y})^2}=\sqrt{174}$，$\sqrt{\sum\limits_{i=1}^{5}(x_i-\overline{x})^2}=\sqrt{10}$，

所以样本相关系数 $r=\dfrac{41}{\sqrt{1\ 740}}\approx\dfrac{41}{41.7}\approx0.98>0.75$，

因此，两个变量有很强的线性相关关系.

**【探究总结】**

1.样本相关系数是从数值上来判断变量间的线性相关程度，是定量的方法.与散点图相比较，样本相关系数要精细得多.需要注意的是线性相关系数 $r$ 的绝对值小，只是说明样本相关程度低，但不一定不相关，可能是非线性相关.

2.利用样本相关系数 $r$ 来检验线性相关显著性水平时，通常与 0.75 作比较.若 $|r|\geqslant0.75$，则线性相关较为显著，否则不显著.

### 应用迁移

1.(多选)某旅游公司设计了一款冰雪文创产品.试营销以来，这款冰雪文创产品的定价 $x$(单位:元)与销量 $y$(单位:万件)的数据如下表所示：

| 产品定价 $x$/元 | 9 | 9.5 | 10 | 10.5 | 11 |
|---|---|---|---|---|---|
| 销量 $y$/万件 | 11 | 10 | 8 | 6 | 5 |

则下列结论正确的是　(　　)

参考公式：$r=\dfrac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}}$，

参考数据：$\sum\limits_{i=1}^{5}(x_i-\overline{x})^2=2.5$，$\sum\limits_{i=1}^{5}(y_i-\overline{y})^2=26$，$\sum\limits_{i=1}^{5}(x_i-\overline{x})(y_i-\overline{y})=-8$，$\sqrt{65}\approx8.06$.

A.产品定价 $x$ 的平均数是 10 元

B.产品定价 $x$ 与销量 $y$ 存在正相关关系

C.产品定价 $x$ 与销量 $y$ 的线性相关程度很强

D.产品定价 $x$ 与销量 $y$ 的样本相关系数 $r\approx-0.99$

ACD　**解析**：由题可得 $\overline{x}=\dfrac{1}{5}\times(9+9.5+10+10.5+11)=10$，故 A 正确；

而 $r=\dfrac{\sum\limits_{i=1}^{5}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{5}(x_i-\overline{x})^2\sum\limits_{i=1}^{5}(y_i-\overline{y})^2}}=\dfrac{-8}{\sqrt{2.5\times26}}=\dfrac{-8}{\sqrt{65}}\approx-0.99$.

由于 $y$ 与 $x$ 的样本相关系数近似为 $-0.99$，故 $y$ 与 $x$ 的线性相关程度很强，同时，$y$ 与 $x$ 负相关，故 B 错误，C，D 正确.

故选 ACD.

2.某中学往届高三年级数学学科的复习采用的是“刷题—讲题—再刷题”的模式，效果不理想，于是改为采用“记题型—刷题—检测效果”的模式，并记录了某学生的记题型时间 $x$(单位:h)与检测效果 $y$ 的数据如表所示.

| $x$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $y$ | 2.9 | 3.3 | 3.6 | 4.4 | 4.8 | 5.2 | 5.9 |

据统计表明，$y$ 与 $x$ 之间具有线性相关关系，请用样本相关系数 $r$ 加以说明.(若$|r|\geqslant 0.75$，则认为 $y$ 与 $x$ 有很强的线性相关关系，否则认为 $y$ 与 $x$ 的线性相关性较弱)

参考公式：

$$r=\frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\overline{y})^2}}.$$

参考数据：$\overline{y}=4.3$，$\sum_{i=1}^{7}(y_i-\overline{y})^2=7.08$，

$\sum_{i=1}^{7}(x_i-\overline{x})(y_i-\overline{y})=14$，$\sqrt{198.24}\approx 14.08$.

**解**：由题得 $\overline{x}=\frac{1+2+3+4+5+6+7}{7}=4$，

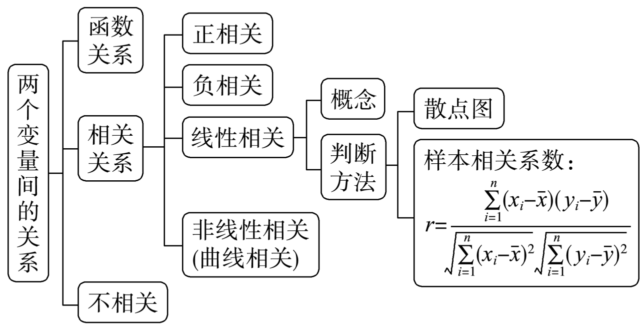
$\sum_{i=1}^{7}(x_i-\overline{x})^2=9+4+1+0+1+4+9=28$，

所以 $r=\frac{\sum_{i=1}^{7}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{7}(x_i-\overline{x})^2}\sqrt{\sum_{i=1}^{7}(y_i-\overline{y})^2}}$

$=\frac{14}{\sqrt{28}\times\sqrt{7.08}}\approx 0.99>0.75$，

所以 $y$ 与 $x$ 有很强的线性相关关系.

## 提质归纳




# 课后素养评价（十八） 成对数据的统计相关性

## A组 学习·理解

**1.**下列语句中的两个变量不具有相关关系的是（ ）

A.瑞雪兆丰年

B.读书破万卷，下笔如有神

C.吸烟有害健康

D.喜鹊叫喜

D **解析**："瑞雪兆丰年"和"读书破万卷，下笔如有神"是根据经验总结归纳出来的，"吸烟有害健康"具有科学根据，所以它们中的两个变量都是相关关系；喜鹊发出叫声是它们自身的生理反应，与事情的好坏无关，故它们不具有相关关系.

**2.**对两个变量 $x$，$y$ 的几组观测数据统计如表，则这两个变量的相关关系是（ ）

| $x$ | 10 | 9 | 8 | 7 | 6 | 5 |
|---|---|---|---|---|---|---|
| $y$ | 2 | 3 | 3.5 | 4 | 4.8 | 5 |

A.负相关　　B.正相关

C.先正相关后负相关　D.先负相关后正相关

A **解析**：根据题表知 $y$ 随 $x$ 的增大而减小，所以这两个变量负相关.

**3.**下列关于样本相关系数 $r$ 的说法错误的是（ ）

A.$r$ 的取值范围为$[0,1]$

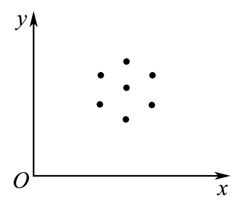
B.$r$ 为正时，两个变量正相关；$r$ 为负时，两个变量负相关

C.$|r|$越接近于 1，两个变量的线性相关程度越强；$|r|$越接近于 0，两个变量的线性相关程度越弱

D.当$|r|=1$ 时，所有样本点都在一条直线上

A **解析**：$r$ 的取值范围为$[-1,1]$.

**4.**变量 $x$，$y$ 的散点图如图所示，那么 $x$，$y$ 之间的样本相关系数 $r$ 最接近的值为（ ）

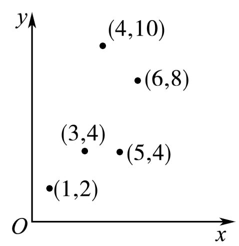



A.1　B.$-0.5$

C.0　D.0.5

C **解析**：根据题图，得 $x$，$y$ 之间的线性相关关系非常不明显，所以样本相关系数 $r$ 最接近的值应为 0.

**5.**如图所示，变量 $x$，$y$ 的五组数据中，去掉______后，剩下的四组数据线性相关程度增强.




(4,10) **解析**：去掉点(4,10)后，其余四点大致在一条直线附近，线性相关程度增强.

**6.**某老师很喜欢"学习强国"中"挑战答题"模块，他记录了自己连续七天每天一次最多答对的题数如

下表：

| 天数 $x$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 一次最多答对题数 $y$ | 12 | 15 | 16 | 18 | 21 | 24 | 27 |

附：$\overline{x}=4$，$\overline{y}=19$，$\sum\limits_{i=1}^{7}x_i^2=140$，

$\sum\limits_{i=1}^{7}y_i^2=2\ 695$，$\sum\limits_{i=1}^{7}x_iy_i=600$，$\sqrt{6}\approx2.45$，

样本相关系数

$$r=\frac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}\cdot\sqrt{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}}=\frac{\sum\limits_{i=1}^{n}x_iy_i-n\overline{x}\ \overline{y}}{\sqrt{\sum\limits_{i=1}^{n}x_i^2-n\overline{x}^2}\sqrt{\sum\limits_{i=1}^{n}y_i^2-n\overline{y}^2}}.$$

由表中数据可知该老师每天一次最多答对题数 $y$ 与天数 $x$ 之间是______相关(填“正”或“负”)，其样本相关系数 $r\approx$______(结果保留两位小数).

正　0.99　**解析**：由题表中数据得 $y$ 随 $x$ 的增大而增大，

所以该老师每天一次最多答对题数 $y$ 与天数 $x$ 之间是正相关，$r=\dfrac{\sum\limits_{i=1}^{7}x_iy_i-7\overline{x}\ \overline{y}}{\sqrt{\sum\limits_{i=1}^{7}x_i^2-7\overline{x}^2}\sqrt{\sum\limits_{i=1}^{7}y_i^2-7\overline{y}^2}}$

$$=\frac{600-7\times4\times19}{\sqrt{140-7\times4^2}\sqrt{2\ 695-7\times19^2}}=\frac{17}{7\sqrt{6}}\approx\frac{17}{7\times2.45}$$

$\approx0.99$.

**7.** 已知两个变量 $x$ 和 $y$ 的7组数据如表所示.

| $x$ | 21 | 23 | 25 | 27 | 29 | 32 | 35 |
|---|---|---|---|---|---|---|---|
| $y$ | 7 | 11 | 21 | 24 | 66 | 115 | 325 |

试判断 $y$ 与 $x$ 是否线性相关，并刻画它们的相关程度.

**解**：画散点图(图略)，观察散点图，可以看出散点都集中在一条直线附近，由此判断 $y$ 与 $x$ 线性相关.

又 $\overline{x}=\dfrac{1}{7}\times(21+23+25+27+29+32+35)\approx27.4$，

$\overline{y}=\dfrac{1}{7}\times(7+11+21+24+66+115+325)\approx81.3$，

$\sum\limits_{i=1}^{7}x_i^2=21^2+23^2+25^2+27^2+29^2+32^2+35^2=5\ 414$，

$\sum\limits_{i=1}^{7}x_iy_i=21\times7+23\times11+25\times21+27\times24+29\times66+32\times115+35\times325=18\ 542$，

$\sum\limits_{i=1}^{7}y_i^2=7^2+11^2+21^2+24^2+66^2+115^2+325^2=124\ 393$，

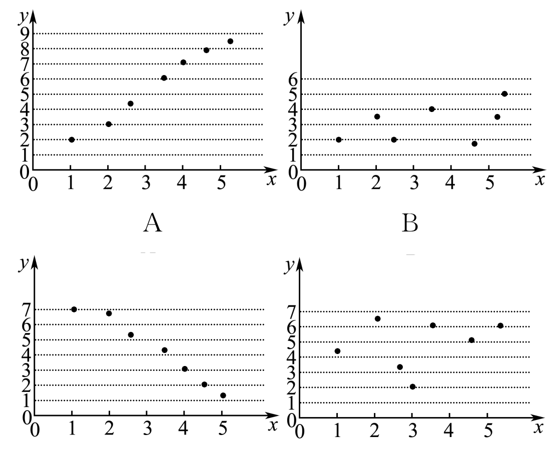
所以 $r=\dfrac{\sum\limits_{i=1}^{7}x_iy_i-7\overline{x}\ \overline{y}}{\sqrt{(\sum\limits_{i=1}^{7}x_i^2-7\overline{x}^2)(\sum\limits_{i=1}^{7}y_i^2-7\overline{y}^2)}}\approx$

$$\frac{18\ 542-7\times27.4\times81.3}{\sqrt{(5\ 414-7\times27.4^2)\times(124\ 393-7\times81.3^2)}}\approx\frac{2\ 948.66}{3\ 520.92}\approx0.837\ 5.$$

所以 $x$ 与 $y$ 具有线性相关关系，且线性相关程度较强.

## B组　应用·实践

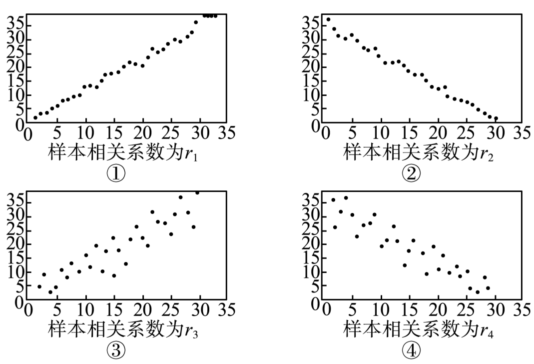
**1.** 已知 $x,y$ 是两个变量，下列四个散点图中，$x,y$ 呈负相关关系的是　　(　　)



A　　B　　C　　D

C　**解析**：对于A，散点图中的点从左向右是上升的，且在一条直线附近，是正相关关系；对于B，两个变量没有明显的相关关系；对于C，散点图中的点从左向右是下降的，且在一条直线附近，是负相关关系；对于D，两个变量没有明显的相关关系.

**2.** 某统计部门对四组数据进行统计分析后，获得如图所示的散点图，关于样本相关系数的比较，其中正确的是　　(　　)




A. $r_4<r_2<0<r_1<r_3$

B. $r_2<r_4<0<r_1<r_3$

C.$r_2<r_4<0<r_3<r_1$

D.$r_4<r_2<0<r_3<r_1$

C **解析**：根据散点图的特征，散点从左向右呈上升趋势的是正相关，呈下降趋势的是负相关；散点越集中在一条直线附近，说明线性相关性越强.

由题图可知①③为正相关，②④为负相关，

故 $r_1>0$，$r_3>0$，$r_2<0$，$r_4<0$；

又①与②中散点图更集中在一条直线附近，故 $r_1>r_3$，$r_2<r_4$，因此 $r_2<r_4<0<r_3<r_1$.

**3.** 对于样本相关系数 $r$，下列叙述正确的是（　　）

A.$|r|\in(0,+\infty)$，$|r|$越大，线性相关程度越强，反之，线性相关程度越弱

B.$r\in(-\infty,+\infty)$，$r$ 越大，线性相关程度越强，反之，线性相关程度越弱

C.$|r|\leqslant 1$，且$|r|$越接近 1，线性相关程度越强，$|r|$越接近 0，线性相关程度越弱

D.以上说法都不对

C **解析**：用样本相关系数 $r$ 可以衡量两个变量之间的线性相关程度的强弱，$r$ 的绝对值越接近于 1，表示两个变量的线性相关程度越强，$r$ 的绝对值越接近于 0，表示两个变量之间的线性相关程度越弱.

**4.** 下列说法中正确的是________.(填序号)

①变量间的样本相关系数 $r$ 的取值范围为$[-1,1]$；

②变量间的样本相关系数 $r$ 的绝对值越接近 0，变量间的线性相关程度越弱；

③变量间的样本相关系数越小，变量间的线性相关程度越弱.

①② **解析**：根据题意，依次分析.对于①，样本相关系数 $r$ 满足$|r|\leqslant 1$，即样本相关系数 $r$ 的取值范围为$[-1,1]$，①正确；

对于②，根据样本相关系数的性质知，$|r|$越接近 1，变量间的线性相关程度越强；$|r|$越接近 0，变量间的线性相关程度越弱，②正确；

对于③，当 $r$ 接近$-1$ 时，变量间的线性相关程度比 $r$ 接近 0 时的强，③错误.

**5.** 某专营店统计了最近 5 天到该店购物的人数，列表如下：

| 日期编号 $x$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 人数 $y$ | 75 | 84 | 93 | 98 | 100 |

(1)由表中给出的数据，判断是否可用线性回归模型拟合 $y$ 与 $x$ 之间的关系.(若$|r|>0.75$，则认为线性相关程度较强，可用线性回归模型拟合；否则，不可用线性回归模型拟合.计算 $r$ 时精确到0.01)

(2)该专营店为了吸引顾客，推出两种促销方案：方案一，购物金额每满 100 元可减 10 元；方案二，购物金额超过 800 元可抽奖三次，每次中奖的概率均为$\frac{1}{3}$，且每次抽奖互不影响，中奖一次打 9 折，中奖两次打 8 折，中奖三次打 6 折.某顾客计划在此专营店购买一件价值 1 000 元的商品，请从实际付款金额的数学期望的角度分析，选哪种方案更优惠？

附：$\sqrt{4\ 340}\approx 65.88$，

样本相关系数

$$r=\frac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}}.$$

**解**：(1)$\overline{x}=\frac{1+2+3+4+5}{5}=3$，

$\overline{y}=\frac{75+84+93+98+100}{5}=90$，

所以$\sum\limits_{i=1}^{5}(x_i-\overline{x})(y_i-\overline{y})=-2\times(-15)-1\times(-6)+0+1\times 8+2\times 10=64$，

$\sum\limits_{i=1}^{5}(x_i-\overline{x})^2=4+1+0+1+4=10$，$\sum\limits_{i=1}^{5}(y_i-\overline{y})^2=(-15)^2+(-6)^2+3^2+8^2+10^2=434$，

所以 $r=\frac{\sum\limits_{i=1}^{5}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{5}(x_i-\overline{x})^2\cdot\sum\limits_{i=1}^{5}(y_i-\overline{y})^2}}=\frac{64}{\sqrt{10\times 434}}$

$\approx\frac{64}{65.88}\approx 0.97>0.75$，

所以，$y$ 与 $x$ 的线性相关性很强，故可用线性回归模型拟合 $y$ 与 $x$ 之间的关系.

(2)设方案一的实际付款金额为 $X$ 元，方案二的实际付款金额为 $Y$ 元.

由题意可知，$E(X)=1\ 000-10\times 10=900$ 元.

$Y$ 的可能取值有 600，800，900，1 000，

$P(Y=600)=\left(\frac{1}{3}\right)^3=\frac{1}{27}$，

$P(Y=800)=C_3^2\times\left(\frac{1}{3}\right)^2\times\frac{2}{3}=\frac{2}{9}$，

$P(Y=900)=C_3^1\times\frac{1}{3}\times\left(\frac{2}{3}\right)^2=\frac{4}{9}$，

$P(Y=1\ 000)=\left(\frac{2}{3}\right)^3=\frac{8}{27}$，

所以 $E(Y)=600\times\frac{1}{27}+800\times\frac{2}{9}+900\times\frac{4}{9}+1\ 000\times\frac{8}{27}=\frac{24\ 200}{27}<\frac{24\ 300}{27}=E(X)$，

所以，方案二更优惠.

# 8.2　一元线性回归模型及其应用

## 第1课时　一元线性回归模型

**学习任务目标**

1.结合具体实例,了解一元线性回归模型的含义,了解模型参数的统计意义.
2.了解最小二乘原理,会用最小二乘法求经验回归方程并进行预测.
3.会通过分析残差判断一元线性回归模型的拟合效果.

## ◦ 问题式预习 ◦

### 【知识清单】

**知识点一　一元线性回归模型**

一元线性回归模型:$\begin{cases}Y=bx+a+e,\\E(e)=0,D(e)=\sigma^2,\end{cases}$

其中,$Y$ 称为因变量或响应变量,$x$ 称为自变量或解释变量;$a$ 和 $b$ 为模型的未知参数,$a$ 称为截距参数,$b$ 称为斜率参数;$e$ 是 $Y$ 与 $bx+a$ 之间的随机误差.

**知识点二　经验回归方程与最小二乘法**

(1)我们将 $\hat{y}=\hat{b}x+\hat{a}$ 称为 $Y$ 关于 $x$ 的经验回归方程,也称经验回归函数或经验回归公式,其图形称为经验回归直线.

经验回归方程的参数:

$$\hat{b}=\frac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}=\frac{\sum\limits_{i=1}^{n}x_iy_i-n\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{n}x_i^2-n\overline{x}^2},$$

$\hat{a}=\overline{y}-\hat{b}\overline{x}$.

(2)求经验回归方程的方法叫做最小二乘法,求得的 $\hat{b}$,$\hat{a}$ 叫做 $b$,$a$ 的最小二乘估计.

**知识点三　残差**

对于响应变量 $Y$,通过观测得到的数据称为观测值,通过经验回归方程得到的 $\hat{y}$ 称为预测值,观测值减去预测值所得的差称为残差.

### 【概念辨析】

1.判断正误(正确的打"√",错误的打"×").

(1)设 $y$ 关于 $x$ 的经验回归方程为 $\hat{y}=2-1.5x$,则当变量 $x$ 增加一个单位时,$y$ 平均增加 1.5 个单位.　( × )

(2)经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$ 表示最接近 $y$ 与 $x$ 之间真实关系的一条直线.　( √ )

(3)若经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$ 是由一组样本数据 $(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$ 得到的,则直线 $\hat{y}=\hat{b}x+\hat{a}$ 必经过点 $(x_1,y_1)$.　( × )

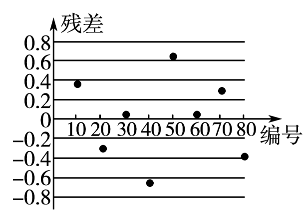
(4)若 $y$ 关于 $x$ 的经验回归方程为 $\hat{y}=\hat{b}x+\hat{a}$,则当 $x=x_0$ 时,$y$ 的值一定是 $\hat{b}x_0+\hat{a}$.　( × )

2.关于一元线性回归模型 $\begin{cases}Y=bx+a+e,\\E(e)=0,D(e)=\sigma^2,\end{cases}$ 下列说法正确的是　(　　)

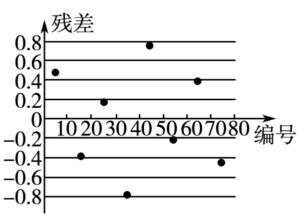
A.$Y=bx+a+e$ 是一次函数
B.因变量 $Y$ 是由自变量 $x$ 唯一确定的
C.因变量 $Y$ 除了受自变量 $x$ 的影响外,可能还受到其他因素的影响,这些因素会导致随机误差 $e$ 的产生
D.随机误差 $e$ 是由于计算不准确造成的,可通过精确计算避免随机误差 $e$ 的产生

C　**解析**:在一元线性回归模型中,表达式 $Y=bx+a+e$ 中,方程表示的不是确定性关系,因此不是一次函数,A 错误;选项 B 中,因变量 $Y$ 不是由自变量 $x$ 唯一确定的,B 错误;选项 D 中,随机误差是不能避免的,只能将误差缩小,但是不可能没有误差,D 错误.

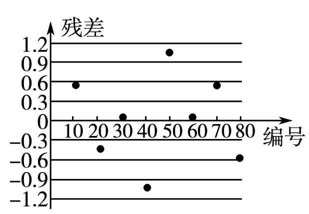
3.对变量 $x$,$y$ 进行回归分析时,依据得到的 4 个不同的线性回归模型画出残差图,则下列模型拟合精度最高的是　(　　)



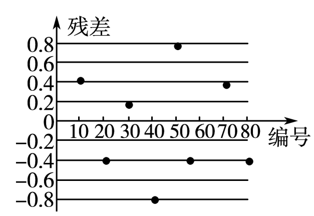

A




B




C




D

A　**解析**:用残差图判断模型的拟合效果时,残差点比较均匀地落在以横轴为对称轴的水平带状区域中,说明这样的模型比较合适.带状区域的宽度越窄,说明模型的拟合精度越高.故选 A.

4.请思考并回答下列问题:

(1)随机误差 $e$ 的主要来源是什么?

**提示**：①用线性回归模型模拟真实模型所引起的误差.可能存在非线性的回归模型能更好地描述两变量间的关系，但是却用线性回归模型来表述这种关系，结果就会产生误差.这种由于模型近似所引起的误差含在 $e$ 中.②忽略了某些因素的影响.影响 $Y$ 的因素不只变量 $x$ 一个，可能还包含其他许多因素，它们的影响都体现在 $e$ 中.③观测误差.由于测量工具、人为测量误差等原因得到的 $Y$ 值一般存在误差.

(2)如何通过散点图或样本相关系数推断两个变量是否存在相关关系？是正相关还是负相关？

**提示**：散点图：如果散点图中表示成对样本数据的点分布在一条直线（曲线）附近，那么两个变量之间具有线性（非线性）相关关系.如果散点落在从左下（上）角到右上（下）角的区域，那么两个变量之间正（负）相关.

样本相关系数 $r$：$|r|$ 越接近于1，两个变量之间的线性相关程度越强；$|r|$ 越接近于0，两个变量之间的线性相关程度越弱.$r>0$，两个变量正相关；$r<0$，两个变量负相关.

## ○ 任务型课堂 ○

### 任务1 一元线性回归模型

1.关于一元线性回归模型 $\begin{cases}Y=bx+a+e,\\E(e)=0,D(e)=\sigma^2,\end{cases}$ 给出下列说法：

①表达式 $Y=bx+a+e$ 刻画的是变量 $Y$ 与变量 $x$ 之间的线性相关关系；

②$bx+a$ 反映了由于 $x$ 的变化而引起的 $Y$ 的线性变化；

③随机误差 $e$ 是一个期望值为0的随机变量；

④对于所有的 $x$ 值，$e$ 的方差 $\sigma^2$ 都相同.

其中正确的是________.（填序号）

①②③④ **解析**：根据一元线性回归模型的含义可知，以上说法均正确.

2.已知某地区每年的财政收入 $x$（单位：亿元）与支出 $Y$（单位：亿元）满足一元线性回归模型 $\begin{cases}Y=bx+a+e,\\E(e)=0,D(e)=\sigma^2,\end{cases}$ 其中 $b=0.7,a=3,|e|\leqslant0.5$.如果今年该地区财政收入为10亿元，那么年支出预计不会超过________.

10.5亿元 **解析**：因为财政收入 $x$ 与支出 $Y$ 满足一元线性回归模型，表达式 $Y=bx+a+e$ 中 $b=0.7,a=3$，

所以 $Y=0.7x+3+e$.

当 $x=10$ 时，得 $Y=0.7\times10+3+e=10+e$.

又 $|e|\leqslant0.5$，即 $-0.5\leqslant e\leqslant0.5$，

所以 $9.5\leqslant Y\leqslant10.5$，

所以年支出预计不会超过10.5亿元.

### 任务2 经验回归方程参数的意义

1.（多选）某工厂某产品产量 $x$（单位：千件）与单位成本 $y$（单位：元）满足经验回归方程 $\hat{y}=77.36-1.82x$，则下列说法中正确的是（　　）

A.产品产量与单位成本正相关

B.产品产量与单位成本负相关

C.产量每增加1 000件，单位成本约减少1.82元

D.产量每增加1 000件，单位成本约增加1.82元

BC **解析**：由经验回归方程的参数 $\hat{b}$ 的意义，可知产品产量与单位成本负相关，且产量每增加1 000件，单位成本约减少1.82元.

2.已知变量 $x,y$ 的一组样本数据为 $(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$（$x_1,x_2,\cdots,x_n$ 不全相等），若由这组数据得到的样本相关系数为 $-1$，则 $y$ 关于 $x$ 的经验回归方程可能是（　　）

A.$\hat{y}=-\frac{1}{2}x+1$　　B.$\hat{y}=x-1$

C.$\hat{y}=x+1$　　D.$\hat{y}=-x^2$

A **解析**：因为变量 $x$ 和变量 $y$ 的样本相关系数为 $-1$，所以这组数据 $(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$ 线性相关，且是负相关，所以经验回归方程的系数 $\hat{b}<0$，可排除B,C,D.故选A.

**【探究总结】**

在经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$ 中，斜率参数 $\hat{b}$ 的意义可以解释为观测值 $x$ 每增加1个单位，预测值 $\hat{y}$ 平均增加 $\hat{b}$ 个单位，且当 $\hat{b}>0$ 时，变量 $x$ 与变量 $y$ 正相关，当 $\hat{b}<0$ 时，变量 $x$ 与变量 $y$ 负相关.

### 任务3 经验回归方程的应用

**探究活动**

**例1** 某种产品的广告费用支出 $x$（单位：万元）与销售额 $y$（单位：万元）之间有如下的对应数据表：

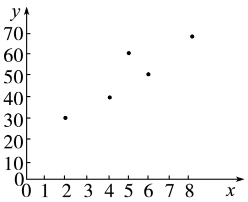
| $x$ | 2 | 4 | 5 | 6 | 8 |
| --- | --- | --- | --- | --- | --- |
| $y$ | 30 | 40 | 60 | 50 | 70 |

(1)画出散点图；

(2)求 $y$ 关于 $x$ 的经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$；

(3)试预测广告费用支出为10万元时的销售额.

**解**：(1)散点图如图所示.




(2)由已知数据得

$\overline{x}=\frac{25}{5}=5,\overline{y}=\frac{250}{5}=50,\sum\limits_{i=1}^{5}x_i^2=145,\sum\limits_{i=1}^{5}x_iy_i=1\ 380.$

于是可得 $\hat{b}=\frac{\sum\limits_{i=1}^{5}x_iy_i-5\overline{x}\ \overline{y}}{\sum\limits_{i=1}^{5}x_i^2-5\ \overline{x}^2}=\frac{1\ 380-5\times5\times50}{145-5\times25}=$ 6.5,

$\hat{a}=\overline{y}-\hat{b}\ \overline{x}=50-6.5\times5=17.5.$

所以所求的经验回归方程为 $\hat{y}=6.5x+17.5.$

(3)根据(2)中求得的经验回归方程,当 $x=10$ 时,$\hat{y}=6.5\times10+17.5=82.5$,即广告费用支出为 10 万元时,销售额大约为 82.5 万元.

[一题多思]

**思考 1.**要使销售额达到 180 万元,预估需要投入的广告费用是多少?

**解:**由例题可得 $\hat{y}=6.5x+17.5$,令 $6.5x+17.5=180$,得 $x=25$.所以预估需要投入的广告费用为 25 万元.

**思考 2.**由经验回归方程计算出结果,在作答时应注意什么?

**提示:**由经验回归方程计算出来的结果是估计值,所以在作答时结果(数据)前要有"估计""大约"等限定词语.

**【探究总结】**

**经验回归方程的应用问题的解题步骤**

(1)画出散点图,从直观上判断变量 $x,y$ 之间是否存在线性相关关系;只有在散点图大致呈直线状时,求出的经验回归方程才有实际意义,否则求出的回归方程毫无意义.

(2)若变量 $x,y$ 线性相关,则计算:$\overline{x},\overline{y},\sum\limits_{i=1}^{n}x_i^2,\sum\limits_{i=1}^{n}x_iy_i$.

(3)将(2)中求出的值代入公式求出 $\hat{y}=\hat{b}x+\hat{a}$ 中参数 $\hat{b},\hat{a}$ 的值.

(4)写出经验回归方程并根据方程解决实际问题.

## 应用迁移

**1.**某公司一种型号的产品近期销售情况如表:

| 月份 $x$ | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| 销售额 $y$/万元 | 15.1 | 16.3 | 17.0 | 17.2 | 18.4 |

根据上表可得到经验回归方程 $\hat{y}=0.75x+\hat{a}$,据此估计,该公司 7 月份这种型号产品的销售额为 (　　)

A.18.85 万元　　B.19.3 万元

C.19.25 万元　　D.19.05 万元

D　**解析:**由题表中数据可得 $\overline{x}=\frac{1}{5}\times(2+3+4+5+6)=4,\overline{y}=\frac{1}{5}\times(15.1+16.3+17+17.2+18.4)=16.8$,

因为经验回归直线过样本点的中心,所以 $16.8=0.75\times4+\hat{a}$,解得 $\hat{a}=13.8$,

所以经验回归方程为 $\hat{y}=0.75x+13.8$,

则估计该公司 7 月份这种型号产品的销售额为 $\hat{y}=0.75\times7+13.8=19.05$(万元).

故选 D.

**2.**已知变量 $x,y$ 呈线性相关关系,经验回归方程为 $\hat{y}=-x+\hat{a}$,且变量 $x,y$ 的样本数据如下表所示:

| $x$ | $-2$ | $-1$ | 0 | 1 | 2 |
|---|---|---|---|---|---|
| $y$ | 5 | 4 | $m$ | 2 | 1 |

据此计算出在 $x=3$ 时,预测值为$-0.2$,则 $m$ 的值为 (　　)

A.3　　B.2.8　　C.2　　D.1

C　**解析:**由题意知经验回归直线 $\hat{y}=-x+\hat{a}$ 过点 $(3,-0.2)$,则 $\hat{a}=2.8$,即 $\hat{y}=-x+2.8$.

又 $\overline{x}=\frac{1}{5}\times(-2-1+0+1+2)=0,\overline{y}=\frac{1}{5}(5+4+m+2+1)=\frac{1}{5}(12+m)$,

由于经验回归直线为 $\hat{y}=-x+\hat{a}$ 必过样本点的中心$(\overline{x},\overline{y})$,所以$\frac{1}{5}(12+m)=-0+2.8$,所以 $m=2$.

故选 C.

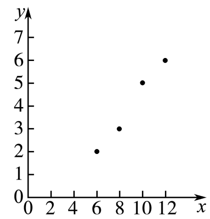
**3.**某研究机构对高三学生的记忆力 $x$ 和判断力 $y$ 进行统计分析,得下表数据.

| $x$ | 6 | 8 | 10 | 12 |
|---|---|---|---|---|
| $y$ | 2 | 3 | 5 | 6 |

(1)画出上表数据的散点图,判断两个变量是否线性相关;

(2)请根据上表提供的数据,用最小二乘法求出 $y$ 关于 $x$ 的经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$,并预测记忆力为 9 的同学的判断力.

**解:**(1)画出散点图(如下图),可判断两个变量线性相关.



(2)由题可得 $\overline{x}=\dfrac{6+8+10+12}{4}=9$，

$\overline{y}=\dfrac{2+3+5+6}{4}=4$，

$\sum\limits_{i=1}^{4}x_i^2=6^2+8^2+10^2+12^2=344$，

$\sum\limits_{i=1}^{4}x_iy_i=6\times2+8\times3+10\times5+12\times6=158$，

$\hat{b}=\dfrac{158-4\times9\times4}{344-4\times9^2}=\dfrac{14}{20}=0.7$，

$\hat{a}=\overline{y}-\hat{b}\,\overline{x}=4-0.7\times9=-2.3$，

故经验回归方程为 $\hat{y}=0.7x-2.3$.

当 $x=9$ 时，$\hat{y}=0.7\times9-2.3=4$，

即预测记忆力为 9 的同学的判断力约为 4.

## 任务4 残差计算

### 探究活动

**例 2** 耐盐碱水稻俗称“海水稻”，是一种可以长在滩涂和盐碱地的水稻.海水稻的灌溉方式是将海水稀释后进行灌溉.某试验基地为了研究海水浓度 $x$(‰)对亩产量 $y$(单位：t)的影响，通过在试验田的种植试验，测得了某种海水稻的亩产量与海水浓度的数据如下表：

| $x$ | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|
| $y$ | 0.62 | 0.58 | 0.49 | 0.4 | 0.31 |

绘制散点图发现，可用线性回归模型拟合亩产量 $y$ 与海水浓度 $x$ 之间的相关关系，用最小二乘法计算得 $y$ 与 $x$ 之间的经验回归方程为 $\hat{y}=\hat{b}x+0.88$.

(1)求出 $\hat{b}$ 的值，并估算当海水浓度为 8‰时该品种水稻的亩产量；

(2)完成下列残差表，并计算残差平方和 $\sum\limits_{i=1}^{n}(y_i-\hat{y}_i)^2$.

| $x$ | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|
| $y$ | 0.62 | 0.58 | 0.49 | 0.4 | 0.31 |
| $\hat{y}$ | | | | | |
| $y-\hat{y}$ | | | | | |

**解**：(1)经计算得 $\overline{x}=5$，$\overline{y}=0.48$.

由 $0.48=5\hat{b}+0.88$ 可得 $\hat{b}=-0.08$.

所以 $\hat{y}=-0.08x+0.88$.

当 $x=8$ 时，$\hat{y}=-0.08\times8+0.88=0.24$.

所以当海水浓度为 8‰时，该品种水稻的亩产量约为0.24 t.

(2)由(1)知 $\hat{y}=-0.08x+0.88$，从而有残差表如下：

| $x$ | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|
| $y$ | 0.62 | 0.58 | 0.49 | 0.4 | 0.31 |
| $\hat{y}$ | 0.64 | 0.56 | 0.48 | 0.4 | 0.32 |
| $y-\hat{y}$ | −0.02 | 0.02 | 0.01 | 0 | −0.01 |

残差平方和 $\sum\limits_{i=1}^{5}(y_i-\hat{y}_i)^2=(-0.02)^2+0.02^2+0.01^2+0^2+(-0.01)^2=0.001$.

**【探究总结】**

1. 利用残差分析可以判断模型刻画数据的效果，以及判断数据中是否存在可疑数据.
2. 当残差比较均匀地分布在以取值为 0 的横轴为对称轴的水平带状区域内时，说明模型的拟合效果较好.

### 应用迁移

1. 关于残差图的描述错误的是 ( )

A.残差图的横坐标可以是样本编号

B.残差图的横坐标可以是解释变量或响应变量

C.残差点分布的带状区域的宽度越窄，残差平方和越大

D.残差点分布的带状区域的宽度越窄，残差平方和越小

C **解析**：残差点分布的带状区域的宽度越窄，说明模型拟合精度越高，则残差平方和越小.

2. 随机选取变量 $x$ 和变量 $Y$ 的 5 对观测数据，选取的第 $i(i=1,2,3,4,5)$对观测数据记为$(x_i,y_i)$，其数值对应如下表所示：

| 编号 $i$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $x$ | 9 | 8 | 7 | 6 | 5 |
| $y$ | 75 | 95 | 110 | 135 | 150 |

计算得：$\overline{x}=\dfrac{1}{5}\sum\limits_{i=1}^{5}x_i=7$，$\overline{y}=\dfrac{1}{5}\sum\limits_{i=1}^{5}y_i=113$，$\sum\limits_{i=1}^{5}x_i^2-5\overline{x}^2=10$，$\sum\limits_{i=1}^{5}y_i^2-5\overline{y}^2=3\ 630$，$\sum\limits_{i=1}^{5}x_iy_i=3\ 765$.

(1)求变量 $x$ 和变量 $Y$ 的样本相关系数(小数点后保留 4 位)，判断这两个变量是正相关还是负相关，并推断它们的线性相关程度.

(2)假设变量 $Y$ 关于 $x$ 的一元线性回归模型为$\begin{cases}Y=bx+a+e,\\E(e)=0,D(e)=\sigma^2.\end{cases}$

①求 $Y$ 关于 $x$ 的经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$，并预测当 $x=10$ 时 $Y$ 的值；

②设 $\hat{e}_i$ 为 $x=x_i(i=1,2,3,4,5)$时该回归模型的残差，求 $\hat{e}_1,\hat{e}_2,\hat{e}_3,\hat{e}_4,\hat{e}_5$ 的方差.

参考公式：$r=\dfrac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}}$，$\hat{b}=\dfrac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}$，$\hat{a}=\overline{y}-\hat{b}\,\overline{x}$.

**解**：(1)$r=\dfrac{\sum\limits_{i=1}^{5}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{5}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{5}(y_i-\overline{y})^2}}$

$=\dfrac{\sum\limits_{i=1}^{5}x_iy_i-5\,\overline{x}\,\overline{y}}{\sqrt{\sum\limits_{i=1}^{5}x_i^2-5\,\overline{x}^2}\cdot\sqrt{\sum\limits_{i=1}^{5}y_i^2-5\,\overline{y}^2}}$

$=\dfrac{3\ 765-5\times7\times113}{\sqrt{10\times3\ 630}}\approx-0.997\ 2$，

所以，这两个变量负相关，且有较强的线性相关性.

(2)①$\hat{b}=\dfrac{\sum\limits_{i=1}^{5}x_iy_i-5\,\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{5}x_i^2-5\,\overline{x}^2}=\dfrac{-190}{10}=-19$，

则 $\hat{a}=\overline{y}-\hat{b}\,\overline{x}=113+19\times7=246$，

所以 $Y$ 关于 $x$ 的经验回归方程为 $\hat{y}=-19x+246$，

当 $x=10$ 时，$\hat{y}=-19\times10+246=56$.

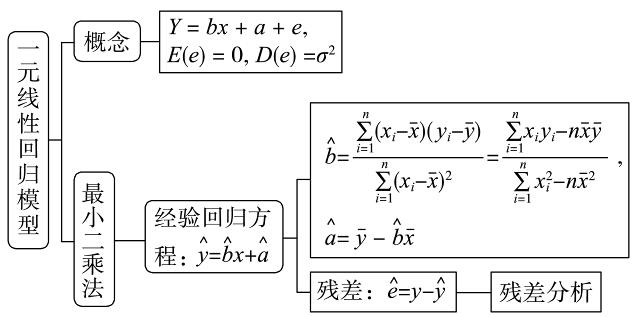
所以，当 $x=10$ 时，$Y$ 的预测值为 56.

②由 $\hat{e}_i=y_i-\hat{y}_i=y_i+19x_i-246$，计算得该回归模型的残差如下表所示：

| 编号 $i$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $x$ | 9 | 8 | 7 | 6 | 5 |
| $\hat{e}$ | 0 | 1 | $-3$ | 3 | $-1$ |

残差的方差为 $\dfrac{0^2+1^2+(-3)^2+3^2+(-1)^2}{5}=4$.

## 提质归纳




# 课后素养评价（十九）　一元线性回归模型

## A组　学习·理解

**1.** 废品率 $x\%$ 和每吨生铁成本 $y$(单位：元)之间的经验回归方程为 $\hat{y}=256+3x$，表明　(　　)

A.废品率每增加 1%，生铁成本增加 259 元

B.废品率每增加 1%，生铁成本增加 3 元

C.废品率每增加 1%，生铁成本平均每吨增加 3 元

D.废品率不变，生铁成本为 256 元

C　**解析**：经验回归方程的斜率参数 $\hat{b}$ 表示 $x$ 每增加一个单位，$\hat{y}$ 平均增加 $\hat{b}$. 当 $x$ 为 1 时，废品率应为 1%，故当废品率增加 1% 时，生铁成本平均每吨增加 3 元.

**2.** 已知 $x$ 与 $y$ 之间的一组数据：

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $y$ | $m$ | 3 | 5.5 | 7 |

已求得 $y$ 关于 $x$ 的经验回归方程为 $\hat{y}=2.2x+0.7$，则 $m$ 的值为　(　　)

A.1　　B.0.85　　C.0.7　　D.0.5

D　**解析**：$\overline{x}=\dfrac{0+1+2+3}{4}=1.5$，$\overline{y}=\dfrac{m+3+5.5+7}{4}$，将 $(\overline{x},\overline{y})$ 代入 $\hat{y}=2.2x+0.7$，解得 $m=0.5$.

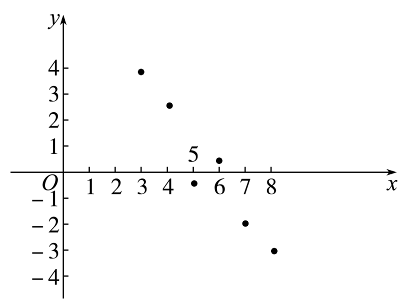
**3.** 根据如下样本数据得到的经验回归方程为 $\hat{y}=\hat{b}x+\hat{a}$，则　(　　)

| $x$ | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|
| $y$ | 4.0 | 2.5 | $-0.5$ | 0.5 | $-2.0$ | $-3.0$ |

A.$\hat{a}>0$，$\hat{b}>0$　　B.$\hat{a}>0$，$\hat{b}<0$

C.$\hat{a}<0$，$\hat{b}>0$　　D.$\hat{a}<0$，$\hat{b}<0$

B　**解析**：画出散点图如图所示.



由图可知，$\hat{a}>0$，$\hat{b}<0$.

**4.**（多选）为调研加工零件效率，调研员通过试验获得加工零件个数 $x$ 与所用时间 $y$(单位：min)的 5 组

数据：(10，52)，(20，67)，(30，70)，(40，75)，(50，86)，根据以上数据可得经验回归方程为 $\hat{y}=0.76x+\hat{a}$，则 ( )

A. $\hat{a}=47.3$

B. 经验回归直线 $\hat{y}=0.76x+\hat{a}$ 必过点(30，70)

C. 加工 60 个零件的时间大约为 92.8 min

D. 若去掉(30，70)，剩下 4 组数据的经验回归方程会有变化

BC **解析**：$\overline{x}=\frac{1}{5}\times(10+20+30+40+50)=30$，$\overline{y}=\frac{1}{5}\times(52+67+70+75+86)=70$，

所以经验回归直线 $\hat{y}=0.76x+\hat{a}$ 恒过点(30，70)，

所以 $70=0.76\times30+\hat{a}$，

解得 $\hat{a}=47.2$，故 A 错误，B 正确；

所以 $\hat{y}=0.76x+47.2$，令 $x=60$，则 $\hat{y}=0.76\times60+47.2=92.8$，

故加工 60 个零件的时间大约为 92.8 min，故 C 正确；

因为经验回归直线 $\hat{y}=0.76x+47.2$ 恒过点(30，70)，所以去掉(30，70)剩下 4 组数据的经验回归方程不会有变化，故 D 错误.

故选 BC.

**5.** 某化工厂为预测某产品的回收率，需要研究其回收率 $y$ 和原料有效成分含量 $x$ 之间的线性相关关系. 现取 8 对观测值，计算得 $\sum\limits_{i=1}^{8}x_i=52$，$\sum\limits_{i=1}^{8}y_i=228$，$\sum\limits_{i=1}^{8}x_i^2=478$，$\sum\limits_{i=1}^{8}x_iy_i=1\ 849$，则 $y$ 关于 $x$ 的经验回归方程是 ( )

A. $\hat{y}=11.47+2.62x$

B. $\hat{y}=-11.47+2.62x$

C. $\hat{y}=2.61x+11.47$

D. $\hat{y}=11.47-2.62x$

A **解析**：由题意可得 $\overline{x}=6.5$，$\overline{y}=28.5$，则 $\hat{b}=\frac{\sum\limits_{i=1}^{8}x_iy_i-8\overline{x}\ \overline{y}}{\sum\limits_{i=1}^{8}x_i^2-8\overline{x}^2}\approx2.62$，可得 $\hat{a}=\overline{y}-\hat{b}\ \overline{x}=11.47$. 故 $y$ 关于 $x$ 的经验回归方程是 $\hat{y}=11.47+2.62x$. 故选 A.

**6.** (多选)根据 $x$，$y$ 的一组样本数据 $(x_1，y_1)$，$(x_2，y_2)$，…，$(x_n，y_n)$，求得经验回归方程为 $\hat{y}=1.5x+0.5$，且 $\overline{x}=3$. 其中有两对样本数据(1.2，2.2)和(4.8，7.8)误差较大，移除后重新求得的经验回归方程的斜率参数为 1.2，则 ( )

A. 变量 $x$ 与 $y$ 具有正相关关系

B. 移除两对误差较大的样本数据后重新求得的经验回归方程为 $\hat{y}=1.2x+1.6$

C. 移除两对误差较大的样本数据后，$y$ 的估计值增加速度变快

D. 移除两对误差较大的样本数据后，$y$ 的估计值增加速度变慢

AD **解析**：因为经验回归方程为 $\hat{y}=1.5x+0.5$，$1.5>0$，所以变量 $x$ 与 $y$ 具有正相关关系，故 A 正确. 当 $\overline{x}=3$ 时，$\overline{y}=1.5\times3+0.5=5$，样本点中心为(3，5). 去掉(1.2，2.2)和(4.8，7.8)后，样本点中心还是(3，5). 又因为移除两对误差较大的样本点后重新求得的经验回归方程的斜率参数为 1.2，所以 $5=1.2\times3+\hat{a}$，解得 $\hat{a}=1.4$，故移除两对误差较大的样本数据后的经验回归方程为 $\hat{y}=1.2x+1.4$，故 B 错误. 因为 $1.5>1.2$，所以移除两对误差较大的样本数据后 $y$ 的估计值增加速度变慢，故 C 错误，D 正确.

**7.** 某种产品的广告支出 $x$(单位：万元)与销售额 $y$(单位：万元)之间的关系如表所示，

| $x$ | 2 | 4 | 5 | 6 | 8 |
|---|---|---|---|---|---|
| $y$ | 30 | 40 | 60 | 50 | 70 |

$y$ 关于 $x$ 的经验回归方程为 $\hat{y}=6.5x+17.5$，当广告支出为 5 万元时，残差为 ( )

A. $-10$　　B. $-20$

C. 20　　D. 10

D **解析**：当广告支出为 5 万元时，观测值为 60，预测值为 $\hat{y}=6.5\times5+17.5=50$，则残差为 $60-50=10$. 故选 D.

**8.** 已知 $y$ 关于 $x$ 的经验回归方程为 $\hat{y}=2x+a$. 若方程在样本点 $(r，1)$ 与 $(1，s)$ 处的残差相同，则有 ( )

A. $r=s$　　B. $s=2r$

C. $s=-2r+3$　　D. $s=2r+1$

C **解析**：样本点 $(r，1)$ 处的残差为 $1-(2r+a)$，样本点 $(1，s)$ 处的残差为 $s-(2+a)$，依题意 $1-(2r+a)=s-(2+a)$，故 $s=-2r+3$. 故选 C.

## B组 应用·实践

**1.** (多选)某电商平台为了对某一产品进行合理定价，采用不同的单价在平台试销，得到的数据如下表所示：

| 单价 $x$/元 | 8 | 8.5 | 9 | 9.5 | 10 |
|---|---|---|---|---|---|
| 销量 $y$/万件 | 89 | 85 | 80 | 78 | 68 |

根据以上数据得到 $y$ 与 $x$ 具有较强的线性关系，若用最小二乘法得到的经验回归方程为 $\hat{y}=-19.8x$

$+\hat{a}$，则（　　）

A.样本相关系数 $r>0$

B.点$(9,80)$一定在经验回归直线上

C.$\hat{a}=258.2$

D.$x=9.5$ 时，对应销量的残差为 $-7.9$

BC　**解析**：由题表中数据可得 $\overline{x}=\frac{8+8.5+9+9.5+10}{5}=9$，$\overline{y}=\frac{89+85+80+78+68}{5}=80$，

所以样本点中心为$(9,80)$，故点$(9,80)$一定在经验回归直线上，B 正确；

由 $\hat{y}=-19.8x+\hat{a}$ 可得 $y$ 与 $x$ 呈负相关，A 错误；

将$(9,80)$代入 $\hat{y}=-19.8x+\hat{a}$ 可得 $80=-19.8\times9+\hat{a}$，解得 $\hat{a}=258.2$，C 正确；

当 $x=9.5$ 时，$\hat{y}=-19.8\times9.5+258.2=70.1$，所以残差为 $78-70.1=7.9$，D 错误.

故选 BC.

**2.**某工厂对某产品的产量与成本的资料分析后得到如下数据：

| 产量 $x$/千件 | 2 | 3 | 5 | 6 |
|---|---|---|---|---|
| 成本 $y$/万元 | 7 | 8 | 9 | 12 |

由表中数据得到的经验回归方程为 $\hat{y}=\hat{b}x+\hat{a}$，且 $\hat{b}=1.1$，预测当产量为 9 千件时，成本约为________万元.

14.5　**解析**：由题表中数据得 $\overline{x}=4$，$\overline{y}=9$，代入经验回归方程得 $\hat{a}=4.6$. 所以 $\hat{y}=1.1x+4.6$，当 $x=9$ 时，$\hat{y}=1.1\times9+4.6=14.5$.

**3.**已知 $\hat{y}=0.85x-85.7$ 是根据女大学生的身高预报体重的经验回归方程(其中 $x$，$\hat{y}$ 的单位分别是 cm，kg)，则该方程在样本点$(165，57)$处的残差是________.

2.45　**解析**：当 $x=165$ 时，$\hat{y}=0.85\times165-85.7=54.55$，所以方程在样本点$(165,57)$处的残差是 $57-54.55=2.45$.

**4.**某中学物理兴趣小组通过试验对其中一道竞赛题的两个物理量 $u$，$v$ 进行测量，得到 10 组数据：$(u_1,v_1)$，$(u_2,v_2)$，…，$(u_{10},v_{10})$，通过散点图发现它们具有较强的线性相关关系，并且利用最小二乘法求得经验回归方程为 $\hat{v}=1.5u+1$. 由于数据保存失误导致 $\sum\limits_{i=1}^{10}v_i$ 的值丢失，已知 $\sum\limits_{i=1}^{10}u_i=50$，可以求得 $\sum\limits_{i=1}^{10}v_i=$________.

85　**解析**：由 $\sum\limits_{i=1}^{10}u_i=50$，得 $\overline{u}=50\times\frac{1}{10}=5$. 又经验回归直线恒过样本点的中心，所以 $\overline{v}=1.5\overline{u}+1=8.5$. 所以 $\sum\limits_{i=1}^{10}v_i=10\overline{v}=85$.

**5.**随着科技发展的日新月异，人工智能融入了各个行业，促进了社会的快速发展.其中利用人工智能生成的虚拟角色因为拥有更低的人工成本，正逐步取代传统的真人直播带货.某公司使用虚拟角色直播带货后销售金额得到逐步提升，以下为该公司自 2025 年 8 月使用虚拟角色直播带货后的销售金额情况统计.

| 时间 | 2025 年 8 月 | 2025 年 9 月 | 2025 年 10 月 | 2025 年 11 月 | 2025 年 12 月 | 2026 年 1 月 |
|---|---|---|---|---|---|---|
| 月份编号 $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
| 销售金额 $y$/万元 | 15.4 | 25.4 | 35.4 | 85.4 | 155.4 | 195.4 |

回答如下问题：

(1)求变量 $y$ 与 $x$ 的样本相关系数 $r$(结果精确到 0.01)；

(2)求 $y$ 关于 $x$ 的经验回归方程，并据此预测 2027 年 2 月份该公司的销售金额.

附：$\sum\limits_{i=1}^{6}x_iy_i=2\ 463.4$，

$\sqrt{\sum\limits_{i=1}^{6}(y_i-\overline{y})^2}=20\sqrt{70}$.

**解**：(1)$\overline{x}=\frac{1+2+3+4+5+6}{6}=3.5$，

$\overline{y}=\frac{15.4+25.4+35.4+85.4+155.4+195.4}{6}=85.4$，

$\sum\limits_{i=1}^{6}x_i^2-6\overline{x}^2=1+4+9+16+25+36-6\times3.5^2=17.5$，

所以 $r=\frac{\sum\limits_{i=1}^{6}x_iy_i-6\overline{x}\,\overline{y}}{\sqrt{\sum\limits_{i=1}^{6}x_i^2-6\overline{x}^2}\sqrt{\sum\limits_{i=1}^{6}y_i^2-6\overline{y}^2}}$

$=\frac{2\ 463.4-6\times3.5\times85.4}{\sqrt{17.5}\times20\sqrt{70}}=\frac{670}{20\times35}\approx0.96$.

(2)由题意 $\hat{b}=\frac{\sum\limits_{i=1}^{6}x_iy_i-6\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{6}x_i^2-6\overline{x}^2}=\frac{2\ 463.4-6\times3.5\times85.4}{17.5}\approx38.3$，

所以 $\hat{a}=85.4-3.5\times38.3=-48.65$，

所以 $y$ 关于 $x$ 的经验回归方程为 $\hat{y}=38.3x-48.65$.

所以预测 2027 年 2 月份该公司的销售金额为 $\hat{y}=38.3\times19-48.65=679.05$(万元).

# 第2课时　一元线性回归模型的应用

**学习任务目标**

**1.** 会求决定系数 $R^2$，并能用 $R^2$ 来比较两个模型的拟合效果.

**2.** 会利用线性回归模型解决非线性回归问题，并能用非线性回归模型解决相关问题.

## ○ 问题式预习 ○

**【知识清单】**

**知识点一　比较两个模型拟合效果的方法**

(1)比较两个模型的残差；

(2)比较两个模型的残差平方和；

(3)用决定系数 $R^2=1-\dfrac{\sum\limits_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}$ 来比较两个模型的拟合效果，$R^2$ 越大，模型的拟合效果越好，$R^2$ 越小，模型的拟合效果越差.

**知识点二　非线性回归分析问题**

对不具有线性相关关系的两个变量做统计分析时，可通过变量代换，将非线性回归模型转化为线性回归模型.

(1)方程 $y=ax^b$ 可以通过 $c=\ln a, v=\ln x, u=\ln y$ 进行变换，得到线性经验回归方程 $\hat{u}=\hat{c}+\hat{b}v$.

(2)方程 $y=a\mathrm{e}^{bx}$ 可以通过 $c=\ln a, u=\ln y$ 进行变换，得到线性经验回归方程 $\hat{u}=\hat{c}+\hat{b}x$.

(3)方程 $y=a+b\ln x$ 可以通过 $v=\ln x, u=y$ 进行变换，得到线性经验回归方程 $\hat{u}=\hat{a}+\hat{b}v$.

**【概念辨析】**

**1.** 判断正误(正确的打"√"，错误的打"×").

(1)决定系数 $R^2$ 越小，模型的拟合效果越好.( × )

(2)在线性回归模型中，$e$ 是 $bx+a$ 预报真实值 $y$ 的随机误差，它是一个可观测的量.　( × )

(3)残差平方和越小，模型的拟合效果越好.( √ )

**2.** 以模型 $y=c\mathrm{e}^{kx}(c>0)$ 去拟合一组数据时，通过 $z=\ln y$ 将其变换后得到经验回归方程 $\hat{z}=2x-1$，则 $k,c$ 的值分别是　(　　)

A. $-2,\mathrm{e}$　　　B. $2,\dfrac{1}{\mathrm{e}}$

C. $-2,\dfrac{1}{\mathrm{e}}$　　　D. $2,\mathrm{e}$

B　**解析**：由题意得 $\ln y=\ln(c\mathrm{e}^{kx})=\ln c+kx$. 又 $z=\ln y$，可得 $z=\ln c+kx$. 又经验回归方程为 $\hat{z}=2x-1$，所以 $k=2, \ln c=-1$，即 $k=2, c=\dfrac{1}{\mathrm{e}}$. 故选 B.

**3.** 请思考并回答下列问题：

(1)非线性相关问题中怎样构造模型对应的曲线方程？

**提示**：①画出散点图，确定曲线的大致形状；
②根据曲线的大致形状，确定基本初等函数特征；
③构造曲线方程.

(2)如何理解 $R^2$？

**提示**：在 $R^2$ 的表达式中，$\sum\limits_{i=1}^{n}(y_i-\overline{y})^2$ 与经验回归方程无关，残差平方和 $\sum\limits_{i=1}^{n}(y_i-\hat{y}_i)^2$ 与经验回归方程有关. 因此 $R^2$ 越大，表示残差平方和越小，即模型的拟合效果越好；$R^2$ 越小，表示残差平方和越大，即模型的拟合效果越差.

## ○ 任务型课堂 ○

### 任务1 > 线性回归分析

**探究活动**

**例1**　为研究悬挂物体的质量 $x$(单位:g)对弹簧长度 $y$(单位:cm)的影响，对6个不同质量的物体进行测量，数据如表所示.

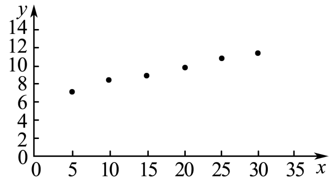
| $x$ | 5 | 10 | 15 | 20 | 25 | 30 |
|---|---|---|---|---|---|---|
| $y$ | 7.25 | 8.12 | 8.95 | 9.9 | 10.9 | 11.8 |

(1)作出散点图并求 $y$ 关于 $x$ 的经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$；($\hat{b},\hat{a}$ 的值精确到0.001)

(2)求出 $R^2$；

(3)进行残差分析.

**解**：(1)作出散点图如图所示.



$\overline{x}=\frac{1}{6}\times(5+10+15+20+25+30)=17.5$，

$\overline{y}=\frac{1}{6}\times(7.25+8.12+8.95+9.9+10.9+11.8)\approx 9.487$，

$\sum_{i=1}^{6}x_i^2=2\ 275$，$\sum_{i=1}^{6}x_iy_i=1\ 076.2$，

计算得 $\hat{b}=\frac{1\ 076.2-6\times17.5\times9.487}{2\ 275-6\times17.5^2}\approx0.183$，

$\hat{a}=9.487-0.183\times17.5\approx6.285$，

所求经验回归方程为 $\hat{y}=6.285+0.183x$.

(2)列表如下：

| 编号 $i$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $y-\hat{y}$ | 0.05 | 0.005 | −0.08 | −0.045 | 0.04 | 0.025 |
| $y-\overline{y}$ | −2.237 | −1.367 | −0.537 | 0.413 | 1.413 | 2.313 |

所以 $\sum_{i=1}^{6}(y_i-\hat{y}_i)^2=0.013\ 175$，$\sum_{i=1}^{6}(y_i-\overline{y})^2=14.678\ 334$.所以 $R^2=1-\frac{0.013\ 175}{14.678\ 334}\approx0.999\ 1$.

所以线性回归模型的拟合效果较好.

(3)由残差表中的数值可以看出第 3 个样本点的残差绝对值比较大，需要确认在采集这个数据的时候是否有人为的错误，如果有的话，需要纠正数据，重新建立线性回归模型；由表中数据作出残差图(图略)，观察表中数据及残差图，可得所有残差点比较均匀地分布在以横轴为对称轴、宽度小于 0.1 的狭窄的水平带状区域中，说明选用的线性回归模型的精度较高.由以上分析可知，弹簧长度与质量有线性相关关系.

[一题多思]

**思考.**进行残差分析时，一般从哪些方面分析？

**解：**利用残差分析研究两个变量间的关系时，首先要根据散点图来粗略判断它们是否线性相关，是否可以用线性回归模型来拟合数据.然后通过残差图来分析残差特性，用残差来判断原始数据中是否存在可疑数据，用 $R^2$ 来刻画模型拟合的效果.

**【探究总结】**

**刻画回归效果的三种方法**

(1)残差图法：残差点比较均匀地落在以横轴为对称轴的水平带状区域内，说明选用的模型比较合适.

(2)残差平方和法：残差平方和 $\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$ 越小，模型的拟合效果越好.

(3)决定系数法：$R^2=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\overline{y})^2}$ 越大，表示模型的拟合效果越好.

## 应用迁移

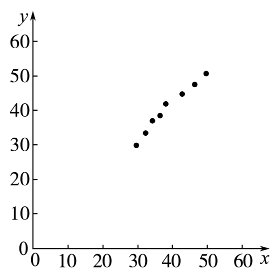
**1.**某运动员的训练次数 $x$ 与训练成绩 $y$ 的数据如表：

| $x$ | 30 | 33 | 35 | 37 | 39 | 44 | 46 | 50 |
|---|---|---|---|---|---|---|---|---|
| $y$ | 30 | 34 | 37 | 39 | 42 | 46 | 48 | 51 |

(1)作出散点图，求出 $y$ 关于 $x$ 的经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$；

(2)作出残差图，计算 $R^2$.

**解：**(1)作出该运动员的训练次数 $x$ 与训练成绩 $y$ 的散点图，如图所示.由散点图可知，它们之间有相关关系.




由题表中数据可得 $\overline{x}=39.25$，$\overline{y}=40.875$，$\sum_{i=1}^{8}x_i^2=12\ 656$，$\sum_{i=1}^{8}x_iy_i=13\ 180$，

所以 $\hat{b}=\frac{\sum_{i=1}^{8}x_iy_i-8\overline{x}\ \overline{y}}{\sum_{i=1}^{8}x_i^2-8\overline{x}^2}\approx1.041\ 5$，

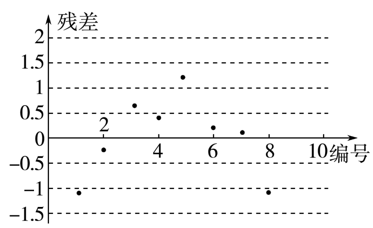
$\hat{a}=\overline{y}-\hat{b}\ \overline{x}\approx-0.003\ 9$.

所以经验回归方程为 $\hat{y}=1.041\ 5x-0.003\ 9$.

(2)残差分析：下面的表格列出了该运动员的训练次数 $x$ 和训练成绩 $y$ 的原始数据以及相应的残差数据.

| $x$ | 30 | 33 | 35 | 37 |
|---|---|---|---|---|
| $y$ | 30 | 34 | 37 | 39 |
| $\hat{e}$ | −1.241 1 | −0.365 6 | 0.551 4 | 0.468 4 |
| $x$ | 39 | 44 | 46 | 50 |
| $y$ | 42 | 46 | 48 | 51 |
| $\hat{e}$ | 1.385 4 | 0.177 9 | 0.094 9 | −1.071 1 |

作残差图如图所示.




由图可知，残差点比较均匀地分布在以横轴为对称轴的水平带状区域内，说明选择的模型比较合适.

由表中数据得，$R^2=1-\dfrac{\sum\limits_{i=1}^{8}(y_i-\hat{y}_i)^2}{\sum\limits_{i=1}^{8}(y_i-\overline{y})^2}\approx 0.9855$.

**2.** 某公司负责生产的A型材料应用前景十分广泛.该公司为了将A型材料更好地投入商用，拟对A型材料进行应用改造.根据市场调研与模拟，得到应用改造投入$x$(单位：亿元)与产品的直接收益$y$(单位：亿元)的数据统计如下表：

| 序号 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $x$ | 2 | 3 | 4 | 6 | 8 | 10 | 13 |
| $y$ | 15 | 22 | 27 | 40 | 48 | 54 | 60 |

根据表中数据，建立了$y$与$x$的两个回归模型，并分别计算出了经验回归方程.模型①：$\hat{y}=4.1x+10.9$，模型②：$\hat{y}=21.3\sqrt{x}-14.4$.

(1)比较两个回归模型的决定系数$R^2$的大小；

(2)据(1)选择拟合精度更高、更可靠的模型，预测对A型材料进行应用改造的投入为17亿元时的直接收益.

附：决定系数$R^2=1-\dfrac{\sum\limits_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}$，且当$R^2$越大时，模型的拟合效果越好；$\sqrt{17}\approx 4.1$.

| 回归模型 | 模型① | 模型② |
|---|---|---|
| $\sum\limits_{i=1}^{7}(y_i-\hat{y}_i)^2$ | 79.31 | 20.2 |

**解**：(1)对于模型①，

$\overline{y}=\dfrac{15+22+27+40+48+54+60}{7}=38$，

$\sum\limits_{i=1}^{7}(y_i-\overline{y})^2=\sum\limits_{i=1}^{7}y_i^2-7\overline{y}^2=1750$，

故对应的决定系数$R_1^2=1-\dfrac{79.13}{1750}\approx 0.955$.

对于模型②，同理可得对应的决定系数$R_2^2=1-\dfrac{20.2}{1750}\approx 0.988$，$R_2^2>R_1^2$.

(2)由(1)可知，模型②的拟合精度更高、更可靠.

故估计对A型材料进行应用改造的投入为17亿元时的直接收益为$\hat{y}=21.3\sqrt{17}-14.4\approx 72.93$亿元.

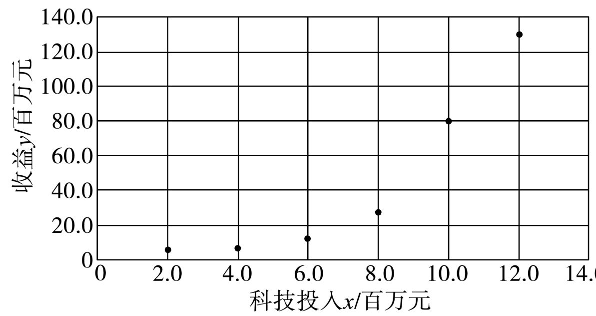
## 任务2 非线性回归分析

### 探究活动

**例2** 中国国际进口博览会是全球首个以进口为主题的国家级展会，旨在扩大开放中国市场，促进国际贸易合作.某企业为了参加这场盛会，提升行业竞争力，加大了科技投入.该企业连续6年来的科技投入$x$(单位：百万元)与收益$y$(单位：百万元)的数据统计如下：

| $x$ | 2 | 4 | 6 | 8 | 10 | 12 |
|---|---|---|---|---|---|---|
| $y$ | 5.6 | 6.5 | 12 | 27.5 | 80 | 129.2 |

并根据数据绘制散点图如图所示.




根据散点图的特点，甲认为样本点分布在指数曲线$y=c\cdot 2^{bx}$的附近，据此他对数据进行了初步处理，得到：

$\overline{y}=43.5$，$\overline{z}=4.5$，$\sum\limits_{i=1}^{6}(x_i-\overline{x})\cdot(y_i-\overline{y})=854$，$\sum\limits_{i=1}^{6}(x_i-\overline{x})(z_i-\overline{z})=34.7$，$\sum\limits_{i=1}^{6}(y_i-\overline{y})^2=12730.4$，$\sum\limits_{i=1}^{6}(x_i-\overline{x})^2=70$，其中$z_i=\log_2 y_i$，$\overline{z}=\dfrac{1}{6}\sum\limits_{i=1}^{6}z_i$.

(1)①请根据表中数据，建立$y$关于$x$的经验回归方程；($x$的系数精确到0.1)

②根据①中所建立的经验回归方程，若该企业想使下一年收益不低于2亿元，则科技投入的费用至少为多少？($\log_2 5\approx 2.322$)

(2)乙认为样本点分布在二次曲线$y=mx^2+n$的附近，并计算得到经验回归方程$\hat{y}=0.92x^2-12$，以及该模型的决定系数$R^2=0.94$，试比较甲、乙两人所建立的模型，谁的拟合效果更好.

附：对于一组数据$(u_1,v_1)$，$(u_2,v_2)$，$(u_3,v_3)$，…，$(u_n,v_n)$，其经验回归方程$\hat{v}=\hat{\alpha}+\hat{\beta}u$中，$\hat{\beta}=\dfrac{\sum\limits_{i=1}^{n}(u_i-\overline{u})(v_i-\overline{v})}{\sum\limits_{i=1}^{n}(u_i-\overline{u})^2}$，$\hat{\alpha}=\overline{v}-\hat{\beta}\,\overline{u}$.

决定系数$R^2=1-\dfrac{\sum\limits_{i=1}^{n}(v_i-\hat{v}_i)^2}{\sum\limits_{i=1}^{n}(v_i-\overline{v})^2}$.

**解**：(1)①由题表可得$\overline{x}=\dfrac{2+4+6+8+10+12}{6}=7$，

令$z=\log_2 y=bx+\log_2 c$，$a=\log_2 c$，则$z=bx+a$.

可知$\hat{b}=\dfrac{\sum\limits_{i=1}^{6}(x_i-\overline{x})(z_i-\overline{z})}{\sum\limits_{i=1}^{6}(x_i-\overline{x})^2}=\dfrac{34.7}{70}\approx 0.5$，

从而$\hat{a}=\overline{z}-\hat{b}\,\overline{x}=4.5-0.5\times 7=1$.

故经验回归方程为 $\hat{z}=0.5x+1$，即 $\hat{y}=2^{0.5x+1}$.

②令 $2^{0.5x+1}\geqslant 200$，得 $0.5x+1\geqslant \log_2 200$，即 $x\geqslant 4+4\log_2 5\approx 13.288$.

故科技投入的费用至少为 13.288 百万元时，下一年的收益才能不低于 2 亿元.

(2)甲建立的回归模型的残差如下：

| | | | | | | |
|---|---|---|---|---|---|---|
| $y$ | 5.6 | 6.5 | 12 | 27.5 | 80 | 129.2 |
| $\hat{y}$ | 4 | 8 | 16 | 32 | 64 | 128 |
| $y-\hat{y}$ | 1.6 | $-1.5$ | $-4$ | $-4.5$ | 16 | 1.2 |

则 $\sum\limits_{i=1}^{6}(y_i-\hat{y}_i)^2=298.5$.

从而 $R^2=1-\dfrac{298.5}{12\ 730.4}\approx 1-0.02=0.98>0.94$.

所以甲建立的模型拟合效果更好.

**【探究总结】**

**非线性回归问题的处理方法**

一般地，有些非线性回归模型通过变换可以转化为线性回归模型，即借助线性回归模型研究呈非线性相关关系的两个变量之间的关系.

(1)如果散点图中的点分布在一条直线附近，可以选用线性回归模型来建模.

(2)如果散点图中的点分布在一条曲线附近，要先对变量作适当变换，再利用线性回归模型来建模.

(3)非线性经验回归方程的求法：

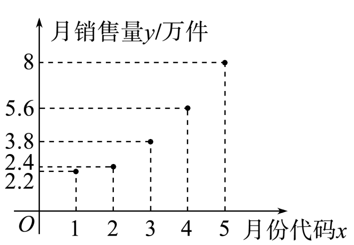
①根据原始数据作出散点图；

②根据散点图，选择恰当的拟合函数；

③作恰当的变换，将其转化成线性函数，求线性经验回归方程，从而得到非线性经验回归方程.

## 应用迁移

随着移动互联网的发展，直播带货已经成为一种热门的销售方式，商家通过直播展示产品，使顾客对产品有更全面的了解.下面统计了某新手开启直播带货后从 6 月份至 10 月份每个月的销售量 $y_i$($i=1,2,3,4,5$，单位：万件)的数据，得到如图所示的散点图.其中 6 月份至 10 月份相应的代码为 $x_i$($i=1,2,3,4,5$)，如：$x_1=1$ 表示 6 月份.




(1)根据散点图判断，模型 $y=a+bx$ 与模型 $y=c+dx^2$ 哪一个更适宜作为月销售量 $y$ 关于月份代码 $x$ 的回归模型？(给出判断即可，不必说明理由)

(2)①根据(1)的判断结果，建立 $y$ 关于 $x$ 的经验回归方程；(计算结果精确到 0.01)

②根据经验回归方程预测 12 月份的销售量.

参考公式：$\hat{b}=\dfrac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}=\dfrac{\sum\limits_{i=1}^{n}x_iy_i-n\,\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{n}x_i^2-n\,\overline{x}^2}$，$\hat{a}=\overline{y}-\hat{b}\,\overline{x}$.

参考数据：$\sum\limits_{i=1}^{5}x_i^2=55$，$\sum\limits_{i=1}^{5}t_i^2=979$，$\sum\limits_{i=1}^{5}x_iy_i=80.8$，$\sum\limits_{i=1}^{5}t_iy_i=335.6$，其中 $t_i=x_i^2$.

**解**：(1)由散点图可知模型 $y=c+dx^2$ 更适宜作为月销售量 $y$ 关于月份代码 $x$ 的回归模型.

(2)①令 $t=x^2$，则 $y=c+dt$，

可得 $\overline{t}=\dfrac{1}{5}\sum\limits_{i=1}^{5}x_i^2=11$，

$\overline{y}=\dfrac{1}{5}\times(2.2+2.4+3.8+5.6+8)=4.4$，

则 $\hat{d}=\dfrac{\sum\limits_{i=1}^{5}t_iy_i-5\,\overline{t}\,\overline{y}}{\sum\limits_{i=1}^{5}t_i^2-5\,\overline{t}^2}=\dfrac{335.6-5\times 11\times 4.4}{979-5\times 11^2}\approx 0.25$，

$\hat{c}=\overline{y}-\hat{d}\,\overline{t}=4.4-0.25\times 11=1.65$，

所以 $y$ 关于 $t$ 的经验回归方程为 $\hat{y}=1.65+0.25t$，

即 $y$ 关于 $x$ 的经验回归方程为 $\hat{y}=1.65+0.25x^2$.

②令 $x=7$，可得 $\hat{y}=1.65+0.25\times 7^2=13.9$，

预测 12 月份的销售量大约是 13.9 万件.

# 任务 3 回归分析的综合应用

## 探究活动

**例 3**　某大学生参加社会实践活动，对某公司 1 月份至 6 月份销售某种配件的销售量及销售单价进行了调查，销售单价 $x$(单位：元)和销售量 $y$(单位：件)之间的一组数据如下表所示：

| 月份 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $x$ | 11 | 9.5 | 12 | 10.5 | 9 | 10 |
| $y$ | 11 | 10 | 8 | 6 | 15 | 14.2 |

(1)根据 1 月份至 5 月份的数据，求出 $y$ 关于 $x$ 的经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$.

(2)若由经验回归方程得到的估计数据与剩下的检验数据的误差不超过 0.5，则认为所得到的经验回归方程是理想的，试问(1)中所得到的经验回归方程是否理想？

(3)预计在今后的销售中,销售量与销售单价服从(1)中的关系,若该种机器配件的成本是2.5元/件,那么该配件的销售单价应定为多少元才能获得最大利润?(注:利润=销售收入-成本)

参考公式:经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$ 中,$\hat{b}=\dfrac{\sum\limits_{i=1}^{n}x_iy_i-n\,\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{n}x_i^2-n\,\overline{x}^2}$,$\hat{a}=\hat{y}-\hat{b}\,\overline{x}$.

参考数据:$\sum\limits_{i=1}^{5}x_iy_i=510$,$\sum\limits_{i=1}^{5}x_i^2=546.5$.

**解**:(1)因为 $\overline{x}=\dfrac{11+9.5+12+10.5+9}{5}=10.4$,$\overline{y}=\dfrac{11+10+8+6+15}{5}=10$,

所以 $\hat{b}=\dfrac{\sum\limits_{i=1}^{5}x_iy_i-5\,\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{5}x_i^2-5\,\overline{x}^2}=\dfrac{510-5\times10.4\times10}{546.5-5\times10.4^2}\approx-1.75$,$\hat{a}=10-(-1.75)\times10.4=28.2$,

于是 $y$ 关于 $x$ 的经验回归方程为 $\hat{y}=-1.75x+28.2$.

(2)当 $x=10$ 时,$\hat{y}=-1.75\times10+28.2=10.7$,

因为 $|14.2-10.7|=3.5>0.5$,

所以认为所得到的经验回归方程是不理想的.

(3)令销售利润为 $W$,

则 $W=(x-2.5)(-1.75x+28.2)=-1.75x^2+32.575x-70.5(2.5<x<16.1)$.

因为 $W\approx1.75x(-x+18.6)-70.5\leqslant1.75\times\left(\dfrac{x-x+18.6}{2}\right)^2-70.5\approx80.86$,

当且仅当 $x=-x+18.6$,即 $x=9.3$ 时,$W$ 取最大值.

所以该配件的销售单价应定为9.3元才能获得最大利润.

**【探究总结】**

回归分析是对具有相关关系的两个变量进行统计分析的一种常用方法.其基本步骤为:通过散点图和经验选择经验回归方程的类型,然后通过一定的规则确定出相应的经验回归方程,通过一定的方法进行检验,最后应用于实际,如对预报变量进行预测等.

## 应用迁移

肥胖不仅影响形体美,而且给生活带来不便,此外还可能引发关节软组织损伤、心脏病、糖尿病、脂肪肝、痛风等.小王通过运动和控制饮食进行减肥,并根据时间 $x$(单位:周)和体重 $y$(单位:kg)记录制作如下统计表:

| $x$ | 1 | 2 | 3 | 4 | 6 | 8 |
|---|---|---|---|---|---|---|
| $y$ | 90.1 | 87.6 | 87.2 | 86.2 | 84.2 | 84.3 |

(1)若 $x$ 和 $y$ 满足经验回归模型 $\hat{y}=\hat{b}\log_2x+\hat{a}$,求 $\hat{a}$,$\hat{b}$.

(2)求该模型的决定系数 $R^2$,并判断该经验回归方程是否有价值.($R^2\geqslant0.9$ 时认为有价值)

(3)当某组数据残差的绝对值不超过0.3时,称该组数据为“身材有效管理数据”.现从这六组数据中任意抽取两组,设抽取的“身材有效管理数据”的个数为 $Y$,求 $Y$ 的分布列和期望.

参考公式:经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$ 中,$\hat{b}=\dfrac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}=\dfrac{\sum\limits_{i=1}^{n}x_iy_i-n\,\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{n}x_i^2-n\,\overline{x}^2}$,$\hat{a}=\overline{y}-\hat{b}\,\overline{x}$;决定系数 $R^2=1-\dfrac{\sum\limits_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}$.

参考数据:$t=\log_2x$,$\overline{t}=1.7$,$\overline{y}=86.6$,$\sum\limits_{i=1}^{6}t_iy_i=871.36$,$\lg2\approx0.3$,$\lg3\approx0.48$.

**解**:(1)因为 $\sum\limits_{i=1}^{6}(t_i-\overline{t})^2\approx(-1.7)^2+(1-1.7)^2+\left(\dfrac{0.48}{0.3}-1.7\right)^2+(2-1.7)^2+\left(1+\dfrac{0.48}{0.3}-1.7\right)^2+(3-1.7)^2=5.98$,

所以 $\hat{b}=\dfrac{\sum\limits_{i=1}^{6}t_iy_i-6\overline{t}\,\overline{y}}{\sum\limits_{i=1}^{6}(t_i-\overline{t})^2}=\dfrac{871.36-6\times1.7\times86.6}{5.98}=-2$.

又 $\overline{y}=86.6$,且经验回归直线 $\hat{y}=\hat{b}t+\hat{a}$ 过点(1.7,86.6),

所以 $\hat{a}=\overline{y}-\hat{b}\,\overline{t}=86.6-(-2)\times1.7=90$.

(2)由(1)可得 $\hat{y}=-2\log_2x+90$,$\overline{y}=86.6$,

| $x$ | 1 | 2 | 3 | 4 | 6 | 8 |
|---|---|---|---|---|---|---|
| $y$ | 90.1 | 87.6 | 87.2 | 86.2 | 84.2 | 84.3 |
| $\hat{y}$ | 90 | 88 | 86.8 | 86 | 84.8 | 84 |
| $(y-\hat{y})^2$ | 0.01 | 0.16 | 0.16 | 0.04 | 0.36 | 0.09 |
| $(y-\overline{y})^2$ | 12.25 | 1 | 0.36 | 0.16 | 5.76 | 5.29 |

则 $R^2=1-\dfrac{\sum\limits_{i=1}^{6}(y_i-\hat{y}_i)^2}{\sum\limits_{i=1}^{6}(y_i-\overline{y})^2}$

$=1-\dfrac{0.01+0.16+0.16+0.04+0.36+0.09}{12.25+1+0.36+0.16+5.76+5.29}$

$=1-\dfrac{0.82}{24.82}\approx0.967$.

因为 $R^2>0.9$,所以该经验回归方程有价值.

(3)经计算,这六组数据中,残差的绝对值不超过0.3

的有三组，分别是第一组、第四组和第六组，故从这六组数据中任意抽取两组，$Y$ 的可能取值有 0，1，2，

$P(Y=0)=\frac{C_3^2}{C_6^2}=\frac{1}{5}$，$P(Y=1)=\frac{C_3^1C_3^1}{C_6^2}=\frac{3}{5}$，$P(Y=2)=\frac{C_3^2}{C_6^2}=\frac{1}{5}$，

则 $Y$ 的分布列为

| $Y$ | 0 | 1 | 2 |
|---|---|---|---|
| $P$ | $\frac{1}{5}$ | $\frac{3}{5}$ | $\frac{1}{5}$ |

故数学期望为 $E(Y)=1\times\frac{3}{5}+2\times\frac{1}{5}=1$.

## 提质归纳

一元线性回归模型的应用
- 决定系数 $R^2=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$ — 应用
  - $R^2$越大，模型拟合效果越好
  - $R^2$越小，模型拟合效果越差
- 非线性回归分析 — 作适当变换，转化为线性回归模型

# 课后素养评价（二十）　一元线性回归模型的应用

## A组　学习·理解

**1.** 对于给定的样本点所建立的模型 $A$ 和模型 $B$，它们的残差平方和分别是 $a_1$，$a_2$，$R^2$ 的值分别为 $b_1$，$b_2$，下列说法正确的是　（ C ）

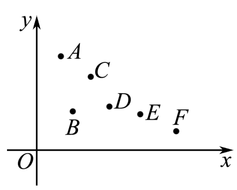
A. 若 $a_1<a_2$，则 $b_1<b_2$，$A$ 的拟合效果更好

B. 若 $a_1<a_2$，则 $b_1<b_2$，$B$ 的拟合效果更好

C. 若 $a_1<a_2$，则 $b_1>b_2$，$A$ 的拟合效果更好

D. 若 $a_1<a_2$，则 $b_1>b_2$，$B$ 的拟合效果更好

**2.**（多选）某同学根据 $x$，$y$ 的六组数据 $(x_i,y_i)(i=1,2,\cdots,6)$ 绘制了如下散点图，在这六个点中去掉点 $B$ 后重新进行回归分析，则下列说法正确的是　（　　）




A. 决定系数 $R^2$ 变小

B. 样本相关系数 $r$ 的绝对值更趋近于 1

C. 残差平方和变小

D. 解释变量 $x$ 与响应变量 $y$ 的相关性变弱

BC　**解析**：从题图中可以看出，点 $B$ 较其他点偏离直线较远，故去掉点 $B$ 后，回归效果更好，决定系数 $R^2$ 更接近于 1，所以去掉点 $B$ 后，$R^2$ 变大，故 A 错误；

去掉点 $B$ 后，变量间的线性相关性变强，

所以 $|r|$ 更趋近于 1，故 B 正确；

去掉点 $B$ 后，残差平方和变小，故 C 正确；

去掉点 $B$ 后，解释变量 $x$ 与响应变量 $y$ 的相关性增强，故 D 错误.

故选 BC.

**3.** 由一组成对数据 $(x_i,y_i)(i=1,2,\cdots,6)$ 得到 $y$ 关于 $x$ 的一元非线性回归方程为 $\hat{y}=bx^2+1$，且 $\sum_{i=1}^{6}x_i^2=12$，$\sum_{i=1}^{6}x_i=4$，$\sum_{i=1}^{6}y_i=18$，则 $b=$　（　　）

A. $-1$　　B. 1

C. $-\frac{9}{2}$　　D. $\frac{9}{2}$

B　**解析**：因为 $y$ 关于 $x$ 的一元非线性回归方程为 $\hat{y}=bx^2+1$，

设 $t=x^2$，则线性回归方程 $\hat{y}=bt+1$.

又因为 $\sum_{i=1}^{6}x_i^2=12$，$\sum_{i=1}^{6}y_i=18$，可得 $\frac{1}{6}\sum_{i=1}^{6}x_i^2=2$，$\frac{1}{6}\sum_{i=1}^{6}y_i=3$，即样本点中心为 $(2,3)$.

将样本点中心坐标 $(2,3)$ 代入线性回归方程 $\hat{y}=bt+1$，可得 $3=2b+1$，解得 $b=1$.

**4.** 关于 $x$ 与 $y$ 的数据如表所示.

| $x$ | 2 | 4 | 5 | 6 | 8 |
|---|---|---|---|---|---|
| $y$ | 30 | 40 | 60 | 50 | 70 |

有如下的两个模型：①$\hat{y}=6.5x+17.5$；②$\hat{y}=7x+17$. 通过残差分析，发现模型①比模型②拟合效果好，则 $R_1^2$ ________ $R_2^2$，$Q_1$ ________ $Q_2$.（用“＞”或“＜”填空，$R^2$，$Q$ 分别是决定系数和残差平方和）

＞　＜　**解析**：由 $R^2$ 的性质可得，$R^2$ 越大，模型的拟合效果越好，所以 $R_1^2>R_2^2$.

由残差的性质可得，残差平方和越小，模型的拟合效果越好，所以 $Q_1<Q_2$.

**5.** 在研究两个变量的相关关系时，观察散点图发现样本点集中于某一条指数曲线 $y=e^{bx+a}$ 的附近，令 $z=\ln y$，求得经验回归方程为 $\hat{z}=0.25x-2.58$，则 $y$

关于 $x$ 的经验回归方程为________.

$\hat{y}=e^{0.25x-2.58}$ **解析**：由 $z=\ln y$，$\hat{z}=0.25x-2.58$，

得 $\ln\hat{y}=0.25x-2.58$，

所以 $\hat{y}=e^{0.25x-2.58}$.

故 $y$ 关于 $x$ 的经验回归方程为 $\hat{y}=e^{0.25x-2.58}$.

**6.** 已知 $x$ 与 $y$ 之间的数据如下表：

| $x$ | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| $y$ | 2.2 | 3.8 | 5.5 | 6.5 | 7.0 |

(1)求 $y$ 关于 $x$ 的经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$；

(2)完成下面的残差表：

| $x$ | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| $y-\hat{y}$ | | | | | |

并判断(1)中经验回归方程的回归效果是否良好(若 $R^2>0.9$，则认为回归效果良好).

附：$\hat{b}=\dfrac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}=\dfrac{\sum\limits_{i=1}^{n}x_iy_i-n\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{n}x_i^2-n\overline{x}^2}$，$\hat{a}=\overline{y}-\hat{b}\overline{x}$，$R^2=1-\dfrac{\sum\limits_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}$，$\sum\limits_{i=1}^{5}(y_i-\hat{y}_i)^2=0.651$.

**解**：(1)由已知图表可得 $\overline{x}=4$，$\overline{y}=5$，$\sum\limits_{i=1}^{5}x_i^2=90$，$\sum\limits_{i=1}^{5}x_iy_i=112.3$，

则 $\hat{b}=\dfrac{112.3-5\times4\times5}{90-5\times4^2}=1.23$，$\hat{a}=\overline{y}-\hat{b}\,\overline{x}=0.08$，

故 $\hat{y}=1.23x+0.08$.

(2)因为 $\hat{e}=y-\hat{y}$，所以 $\hat{e}_1=-0.34$，$\hat{e}_2=0.03$，$\hat{e}_3=0.5$，$\hat{e}_4=0.27$，$\hat{e}_5=-0.46$，则残差表如下表所示.

| $x$ | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| $y-\hat{y}$ | −0.34 | 0.03 | 0.5 | 0.27 | −0.46 |

因为 $\sum\limits_{i=1}^{5}(y_i-\overline{y})^2=(2.2-5)^2+(3.8-5)^2+(5.5-5)^2+(6.5-5)^2+(7-5)^2=15.78$，

所以 $R^2=1-\dfrac{0.651}{15.78}\approx0.96>0.9$，

所以该经验回归方程的回归效果良好.

## B组 应用·实践

**1.** 在回归分析中，$R^2$ 的值越小，说明残差平方和 ( )

A.越小

B.越大

C.可能越大也可能越小

D.以上都不对

B **解析**：由 $R^2$ 的表达式可知，$R^2$ 越大，表示残差平方和越小，即模型的拟合效果越好，$R^2$ 越小，表示残差平方和越大，即模型的拟合效果越差.故选 B.

**2.** 已知变量 $y$ 关于 $x$ 的经验回归方程为 $\hat{y}=e^{bx-0.5}$，其一组数据如下表：

| $x$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $y$ | e | $e^3$ | $e^4$ | $e^6$ |

若 $x=5$，则 $y$ 的预测值为 ( )

A.$e^5$　　B.$e^{5.5}$

C.$e^7$　　D.$e^{7.5}$

D **解析**：由 $\hat{y}=e^{bx-0.5}$，得 $\ln\hat{y}=bx-0.5$.

令 $z=\ln y$，则 $\hat{z}=bx-0.5$.

可得下表：

| $x$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $z$ | 1 | 3 | 4 | 6 |

由表可得 $\overline{x}=\dfrac{1+2+3+4}{4}=2.5$，$\overline{z}=\dfrac{1+3+4+6}{4}=3.5$，

所以 $3.5=b\times2.5-0.5$，解得 $b=1.6$.

所以 $\hat{z}=1.6x-0.5$，所以 $\hat{y}=e^{1.6x-0.5}$.

当 $x=5$ 时，$\hat{y}=e^{1.6\times5-0.5}=e^{7.5}$.故选 D.

**3.** 已知对某组数据采用了四种不同的经验回归方程进行回归分析，$R^2$ 的值分别为 0.97，0.83，0.32，0.17，则拟合效果最好的回归模型对应的 $R^2$ 的值是 ( )

A.0.97　　B.0.83

C.0.32　　D.0.17

A **解析**：对于两个变量 $y$ 与 $x$ 的回归模型，$R^2$ 越大，说明模型的拟合效果越好.在所给的四个选项中，0.97 是四个决定系数中最大的值，所以 0.97 对应的回归模型拟合效果最好.故选 A.

**4.** 对 $y=ae^{bx}$ 进行线性变换后得到的经验回归方程为 $\hat{u}=1-0.6x$，则函数 $y=x^2+bx+a$ 的单调递增区间为 ( )

A.$(0,+\infty)$　　B.$\left(\dfrac{3}{10},+\infty\right)$

C.$\left(\dfrac{1}{2},+\infty\right)$　　D.$(1,+\infty)$

B **解析**：因为 $y=ae^{bx}$，所以两边取自然对数，作线性变换得 $\ln y=\ln(ae^{bx})=\ln a+\ln e^{bx}=\ln a+bx$.又因为对 $y=ae^{bx}$ 进行线性变换后得到的经验回归方程为 $\hat{u}=1-0.6x$，所以 $u=\ln y$，$\ln a=1$，$b=-0.6$，所以 $a=e$.由于函数 $y=x^2+bx+a=x^2-0.6x+e$ 为二次函数，图象开口向上，对称轴为直线

$x=\frac{3}{10}$，所以函数 $y=x^2+bx+a$ 的单调递增区间为$\left(\frac{3}{10},+\infty\right)$.故选 B.

**5.**为了解某市电动汽车的销售情况，调查了该市某电动汽车企业近 6 年的产值情况，数据如下表所示.

| 年份 | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 |
| --- | --- | --- | --- | --- | --- | --- |
| 编号 $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
| 产值 $y$/百万 | 9 | 18 | 30 | 51 | 59 | 80 |

(1)若用模型 $y=a\cdot e^{bx}$ 拟合 $y$ 与 $x$ 的关系，根据提供的数据，求出 $y$ 关于 $x$ 的经验回归方程(参数的值精确到0.01)；

(2)为了进一步了解车主对电动汽车的看法，从某品牌汽车销售店当日 5 位购买电动汽车和 3 位购买燃油汽车的车主中随机选取 4 位车主进行采访，记选取的 4 位车主中购买电动汽车的车主人数为 $X$，求随机变量 $X$ 的分布列与数学期望.

参考数据：$\sum\limits_{i=1}^{6}u_i=20.88$，$\sum\limits_{i=1}^{6}x_iu_i=80.58$，其中 $u=\ln y$.

参考公式：对于一组数据$(x_i,y_i)(i=1,2,3,\cdots,n)$，其经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$ 中，$\hat{b}=\frac{\sum\limits_{i=1}^{n}x_iy_i-n\overline{x}\cdot\overline{y}}{\sum\limits_{i=1}^{n}x_i^2-n\overline{x}^2}$，$\hat{a}=\overline{y}-\hat{b}\,\overline{x}$.

**解**：(1)令 $u=\ln y=\ln(ae^{bx})=bx+\ln a$，

$\overline{x}=\frac{1+2+3+4+5+6}{6}=3.5$，$\overline{u}=\frac{20.88}{6}=3.48$，

则 $\hat{b}=\frac{\sum\limits_{i=1}^{6}x_iu_i-6\,\overline{x}\cdot\overline{u}}{\sum\limits_{i=1}^{6}x_i^2-6\,\overline{x}^2}$

$=\frac{80.58-6\times3.5\times3.48}{1^2+2^2+3^2+4^2+5^2+6^2-6\times3.5^2}\approx0.43$，

$\ln\hat{a}=3.48-0.43\times3.5\approx1.98$，

所以 $\hat{a}=e^{1.98}$，

所以 $\hat{y}=a\cdot e^{bx}=e^{1.98}\times e^{0.43x}=e^{1.98+0.43x}$.

(2)由题意得 $X$ 的所有可能取值为 1,2,3,4，

$P(X=1)=\frac{C_5^1C_3^3}{C_8^4}=\frac{5}{70}=\frac{1}{14}$，

$P(X=2)=\frac{C_5^2C_3^2}{C_8^4}=\frac{30}{70}=\frac{3}{7}$，

$P(X=3)=\frac{C_5^3C_3^1}{C_8^4}=\frac{30}{70}=\frac{3}{7}$，

$P(X=4)=\frac{C_5^4C_3^0}{C_8^4}=\frac{5}{70}=\frac{1}{14}$，

故 $X$ 的分布列为

| $X$ | 1 | 2 | 3 | 4 |
| --- | --- | --- | --- | --- |
| $P$ | $\frac{1}{14}$ | $\frac{3}{7}$ | $\frac{3}{7}$ | $\frac{1}{14}$ |

数学期望 $E(X)=1\times\frac{1}{14}+2\times\frac{3}{7}+3\times\frac{3}{7}+4\times\frac{1}{14}=\frac{5}{2}$.

# 8.3　列联表与独立性检验

## 8.3.1　分类变量与列联表

**学习任务目标**

**1.**掌握分类变量的含义.
**2.**通过实例，理解 2×2 列联表的统计意义.
**3.**能通过等高堆积条形图分析两个分类变量之间的关系.

## ○ 问题式预习 ○

**【知识清单】**

**知识点一　分类变量和列联表**

(1)为了表述方便，我们经常会使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)一般地，假设有两个成对分类变量 $X$ 和 $Y$，它们分别取值于$\{x_1,x_2\}$和$\{y_1,y_2\}$，将数据分类统计，并制作成如下表格，我们将这种形式的数据统计表称为 2×2 列联表.

| $X$ | $Y$ | | 合计 |
| --- | --- | --- | --- |
| | $Y=y_1$ | $Y=y_2$ | |
| $X=x_1$ | $a$ | $b$ | $a+b$ |
| $X=x_2$ | $c$ | $d$ | $c+d$ |
| 合计 | $a+c$ | $b+d$ | $n=a+b+c+d$ |

其中，$a,b,c,d$ 分别为事件$\{X=x_1,Y=y_1\}$，$\{X=x_1,Y=y_2\}$，$\{X=x_2,Y=y_1\}$，$\{X=x_2,Y=y_2\}$的频数.

**知识点二　等高堆积条形图**

与列联表相比,等高堆积条形图能更直观地反映出两个分类变量间是否相互影响,常用等高堆积条形图展示列联表数据的频率特征.

**【概念辨析】**

1.判断正误(正确的打"√",错误的打"×").

(1)分类变量中的变量与函数中的变量是同一概念. ( × )

(2)$2\times2$ 列联表是借助两个分类变量之间的频率大小差异说明两个变量之间是否有关联. ( √ )

(3)等高堆积条形图可初步分析两个分类变量是否有关联. ( √ )

2.已知一个 $2\times2$ 列联表如下,则表中 $m$,$n$ 的值分别为 ( )

| $X$ | $Y$ | | 合计 |
|---|---|---|---|
| | $y_1$ | $y_2$ | |
| $x_1$ | $a$ | 35 | 45 |
| $x_2$ | 7 | $b$ | $n$ |
| 合计 | $m$ | 73 | $s$ |

A.10,38　　B.17,45

C.10,45　　D.17,38

B　**解析**:根据题表可知 $a+35=45$,解得 $a=10$.则 $m=10+7=17$.

又由 $35+b=73$,解得 $b=38$.则 $n=7+38=45$.故选 B.

3.在统计中,研究两个分类变量是否存在关联性时,常用的图表有 ( )

A.散点图和残差图

B.残差图和列联表

C.散点图和等高堆积条形图

D.等高堆积条形图和列联表

D　**解析**:散点图是研究两个变量间的相关关系,列联表是研究两个分类变量是否有关联,残差图体现预测值与观测值间的差距,等高堆积条形图能直观地反映两个分类变量间是否有关系.故选 D.

4.请思考并回答下列问题:

(1)如何利用等高堆积条形图判断两个分类变量是否具有关联性?

**提示**:等高堆积条形图中有两个高度相同的矩形,每一个矩形中都有两种颜色,观察下方颜色区域的高度,如果两个高度相差比较明显,就推断两个分类变量之间具有关联性.

(2)教材例 1 中的随机抽样数据是否能够确定与 $X$ 和 $Y$ 有关的所有概率?为什么?

**提示**:不能.因为随机抽样得到的样本具有随机性,根据样本数据计算出来的频率也具有随机性.在统计推断中,依据频率稳定于概率,可以用频率推断与 $X$ 和 $Y$ 有关的概率,但由于频率具有随机性,这种推断可能犯错误,因此,随机抽样数据不能够确定与 $X$ 和 $Y$ 有关的所有概率.

## ○ 任务型课堂 ○

### 任务1 $2\times2$ 列联表及其应用

**探究活动**

**例 1**　某大学一学院甲、乙两个专业的报考和录取情况如下表:

| 性别 | 甲专业报考人数/人 | 乙专业报考人数/人 | 甲专业录取率 | 乙专业录取率 |
|---|---|---|---|---|
| 男 | 100 | 400 | 25% | 45% |
| 女 | 300 | 100 | 30% | 50% |

根据表格中的数据分析,哪个专业录取率高?

**解**:由题表可得,甲专业录取了男生 25 人,女生 90 人,乙专业录取了男生 180 人,女生 50 人,得到 $2\times2$ 列联表如下:

单位:人

| 性别 | 专业 | | 合计 |
|---|---|---|---|
| | 甲 | 乙 | |
| 男 | 25 | 180 | 205 |
| 女 | 90 | 50 | 140 |
| 合计 | 115 | 230 | 345 |

甲专业的录取率为 $\dfrac{25+90}{100+300}=28.75\%$,

乙专业的录取率为 $\dfrac{180+50}{400+100}=46\%$,

所以乙专业的录取率比甲专业的录取率高.

[一题多思]

**思考 1.**女生的录取率是多少?

**解**:女生的录取率为 $\dfrac{90+50}{300+100}=35\%$.

**思考 2.**男生的录取率比女生的录取率高吗?

**解**:男生的录取率为 $\dfrac{25+180}{100+400}=41\%$,

又由思考 1 知女生的录取率为 35%,

所以男生的录取率比女生的录取率高.

**【探究总结】**

(1)列 $2\times2$ 列联表时,关键是对涉及的变量分清类别,计算时要准确无误.

(2)利用 $2\times2$ 列联表分析两个分类变量间的关系时,首先要根据题中数据列出 $2\times2$ 列联表,然后根据频率特征,即将 $\dfrac{a}{a+b}$ 与 $\dfrac{c}{c+d}$ 的值相比较,直接判断出两个分类变量间是否相互影响.

## 应用迁移

在下列关于吸烟情况与患肺癌情况的 $2\times2$ 列联表中，$d$ 的值为（　　）

单位：人

| 吸烟情况 | 患肺癌情况 | | 合计 |
|---|---|---|---|
| | 未患肺癌 | 患肺癌 | |
| 不吸烟 | 7 775 | 42 | 7 817 |
| 吸烟 | | $d$ | |
| 合计 | 9 874 | | 9 965 |

A.48　　B.49　　C.50　　D.51

B　**解析**：由题表中数据可得，总计患肺癌的人数为 $9\ 965-9\ 874=91$，则吸烟且患肺癌的人数 $d=91-42=49$.

## 任务2 等高堆积条形图的实际应用

### 探究活动

**例2**　为了了解网络对中学生学习成绩的影响，某地区教育主管部门从辖区内的初中生中随机抽取了1 000人进行调查，发现其中经常上网的有200人，这200人中有80人期末考试不及格，而另外800人中有120人期末考试不及格.

(1)根据所给数据作出 $2\times2$ 列联表和等高堆积条形图；

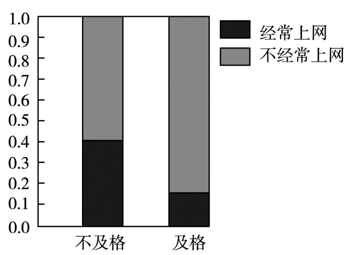
(2)利用(1)中的等高堆积条形图，判断学生经常上网对学习成绩的好坏是否有影响.

**解**：(1)根据题中所给的数据得到如下 $2\times2$ 列联表：

单位：人

| 学习成绩 | 上网情况 | | 合计 |
|---|---|---|---|
| | 经常上网 | 不经常上网 | |
| 不及格 | 80 | 120 | 200 |
| 及格 | 120 | 680 | 800 |
| 合计 | 200 | 800 | 1 000 |

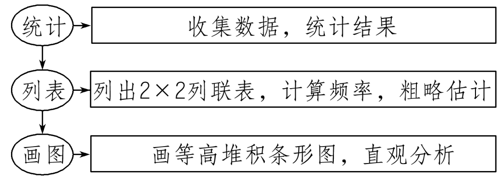
得出等高堆积条形图如图所示.




(2)比较(1)中等高堆积条形图中深色条的高，可以发现经常上网不及格的频率明显高于经常上网及格的频率，因此可以认为经常上网对学习成绩的好坏有影响.

**【探究总结】**

1.利用等高堆积条形图判断两个分类变量是否相关的步骤：




2.等高堆积条形图能形象直观地反映两个分类变量之间的差异，进而推断它们之间是否有关联.

### 应用迁移

1.某地高考采用“3＋1＋2”模式，其中“1”为首选科目，即物理与历史二选一.某校为了解学生的首选意愿，对部分高一学生进行了抽样调查，制作出如下两个等高堆积条形图，根据条形图信息，下列结论正确的是（　　）

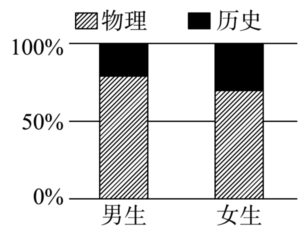


图1

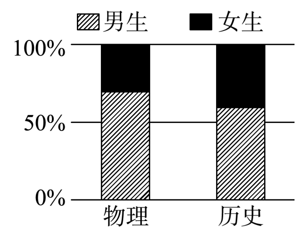


图2

A.样本中选择物理的男生人数少于选择历史的女生人数

B.样本中选择历史的女生人数多于选择历史的男生人数

C.样本中选择物理的人数多于选择历史的人数

D.样本中男生人数少于女生人数

C　**解析**：根据题图1可知样本中选择物理的人数多于选择历史的人数，故C正确；

根据题图2可知样本中男生人数多于女生人数，故D错误；

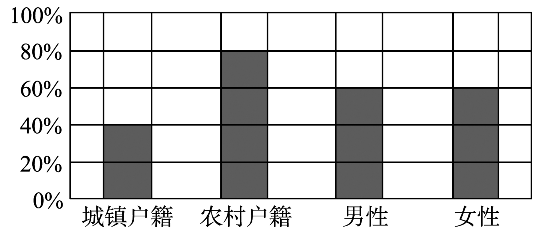
由题图可知样本中选择物理的人数多于选择历史的人数，而选择物理的男生比例高，选择历史的女生比例低，所以样本中选择物理的男生人数多于选择历史的女生人数，故A错误；

样本中选择历史的女生人数不一定多于选择历史的男生人数，故B错误.

故选C.

2.(多选)2021年5月，我国进一步优化生育政策，实施一对夫妻可以生育三个子女政策及配套支持措施.为了了解户籍与性别对生育三孩选择倾向的影响，某地从育龄人群中随机抽取了容量为100的调查样本，其中城镇户籍与农村户籍各50人；男性60

人，女性 40 人.绘制不同群体中倾向选择生育三孩与倾向选择不生育三孩的人数比例图(如图所示)，其中阴影部分表示倾向选择生育三孩的对应比例，则下列叙述中正确的是（　　）




A.是否倾向选择生育三孩与户籍有关

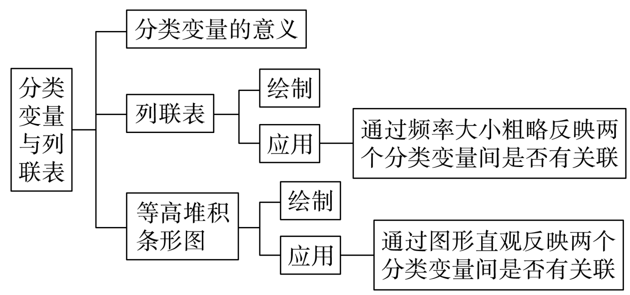
B.是否倾向选择生育三孩与性别有关

C.倾向选择生育三孩的群体中，男性人数与女性人数相同

D.倾向选择不生育三孩的群体中，农村户籍人数少于城镇户籍人数

AD　**解析**：对于 A，城镇户籍倾向选择生育三孩的比例为 40%，农村户籍倾向选择生育三孩的比例为 80%，两者相差较大，所以是否倾向选择生育三孩与户籍有关，故 A 正确；对于 B，男性倾向选择生育三孩的比例为 60%，女性倾向选择生育三孩的比例为 60%，所以是否倾向选择生育三孩与性别无关，故 B 错误；对于 C，男性倾向选择生育三孩的比例为 60%，人数为 $60\times60\%=36$，女性倾向选择生育三孩的比例为 60%，人数为 $40\times60\%=24$，所以倾向选择生育三孩的群体中，男性人数比女性人数多，故 C 错误；对于 D，倾向选择不生育三孩的群体中，农村户籍人数为 $50\times(1-80\%)=10$，城镇户籍人数为 $50\times(1-40\%)=30$，所以倾向选择不生育三孩的群体中，农村户籍人数少于城镇户籍人数，故 D 正确.

### 提质归纳




## 课后素养评价（二十一）　分类变量与列联表

### A组　学习·理解

**1.**下面是 $X$ 与 $Y$ 的 $2\times2$ 列联表：

| X | Y |  | 合计 |
|---|---|---|---|
|  | $y_1$ | $y_2$ |  |
| $x_1$ | $a$ | 21 | 73 |
| $x_2$ | 22 | 25 | 47 |
| 合计 | $b$ | 46 | 120 |

则表中 $a,b$ 的值分别为（　　）

A.94,72　　B.52,50

C.52,74　　D.74,52

C　**解析**：根据列联表的特点，可知 $\begin{cases}a+21=73,\\a+22=b,\end{cases}$ 解得 $\begin{cases}a=52,\\b=74.\end{cases}$

**2.**在 $2\times2$ 列联表中，两个比值相差越大，两个分类变量有关系的可能性就越大，那么这两个比值为（　　）

A.$\dfrac{a}{a+b}$与$\dfrac{c}{c+d}$

B.$\dfrac{a}{c+d}$与$\dfrac{c}{a+b}$

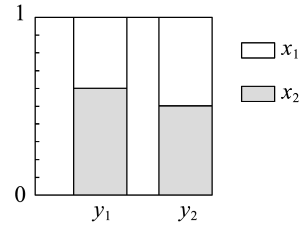
C.$\dfrac{a}{a+d}$与$\dfrac{c}{b+c}$

D.$\dfrac{a}{b+d}$与$\dfrac{c}{a+c}$

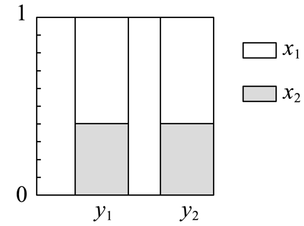
A　**解析**：由题意，得 $\left|\dfrac{a}{a+b}-\dfrac{c}{c+d}\right|=\left|\dfrac{ac+ad-ac-bc}{(a+b)(c+d)}\right|=\left|\dfrac{ad-bc}{(a+b)(c+d)}\right|$，

当$\dfrac{a}{a+b}$与$\dfrac{c}{c+d}$相差越大时，$|ad-bc|$的值越大，两个分类变量有关系的可能性就越大.故选 A.

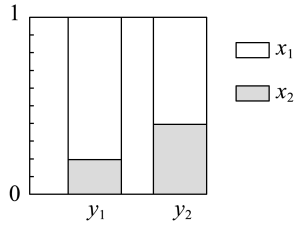
**3.**观察下列各图，其中两个分类变量 $X,Y$ 之间关系最强的是（　　）



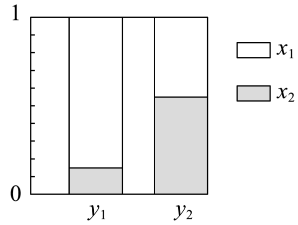

A




B



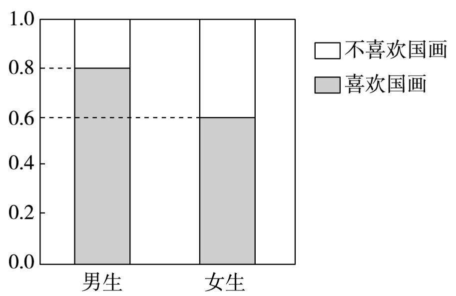
C



D

D　**解析**:在四个选项中,两个深色条的高相差最明显,说明两个分类变量之间关系最强.故选 D.

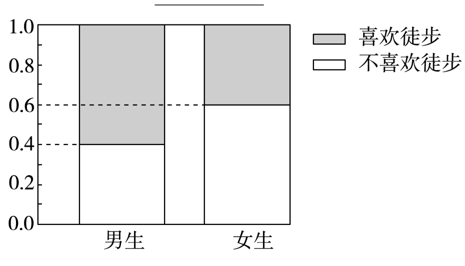
**4.**某艺术馆为了研究学生性别和是否喜欢国画之间的联系,随机抽取 80 名学生(其中男生 50 人,女生 30 人)进行调查,并绘制等高堆积条形图如图所示,则这 80 名学生中喜欢国画的人数为　(　　)




A.24　　B.32　　C.48　　D.58

D　**解析**:由题图可知,男生中喜欢国画的占 80%,女生中喜欢国画的占 60%,则这 80 名学生中喜欢国画的人数为 50×80%+30×60%=58.故选 D.

**5.**如图是调查某学校高一年级男、女学生是否喜欢徒步运动而得到的等高堆积条形图,阴影部分表示喜欢徒步的频率.已知该年级有男生 500 人、女生 400 人(假设所有学生都参加了调查),现从所有喜欢徒步的学生中按分层随机抽样的方法抽取 23 人,则抽取的男生人数为________.




15　**解析**:根据题图可知,喜欢徒步的男生人数为 0.6×500=300,喜欢徒步的女生人数为 0.4×400=160,

所以喜欢徒步的总人数为 300+160=460.

按分层随机抽样的方法抽取 23 人,则抽取的男生人数为$\dfrac{300}{460}\times 23=15$.

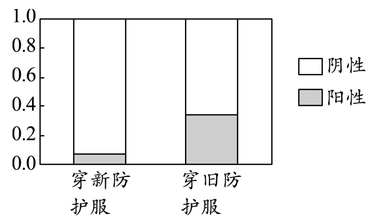
**6.**当某矿石粉厂生产一种矿石粉时,在数天内就有部分工人患职业性皮肤炎.在生产期间,随机抽取车间工人抽血化验,75 名穿新防护服的车间工人中有 5 例阳性、70 例阴性,28 名穿旧防护服的车间工人中有 10 例阳性、18 例阴性,请用等高堆积条形图判断这种新防护服对预防工人职业性皮肤炎是否有效.(注:呈阴性即未患皮肤炎)

**解**:由题中所给的数据得如下 2×2 列联表:

单位:例

| 防护服 | 皮肤炎 | | 合计 |
|---|---|---|---|
| | 阳性 | 阴性 | |
| 新防护服 | 5 | 70 | 75 |
| 旧防护服 | 10 | 18 | 28 |
| 合计 | 15 | 88 | 103 |

相应的等高堆积条形图如图所示.




图中两个深色条的高分别表示穿新、旧防护服样本中呈阳性的频率,从图中可以看出,穿旧防护服呈阳性的频率高于穿新防护服呈阳性的频率.因此,可以认为新防护服对预防工人职业性皮肤炎有效.

**7.**某学校对高三学生做了一项调查,发现在平时的模拟考试中,性格内向的 426 名学生中有 332 人在考前心情紧张,性格外向的 594 名学生中有 213 人在考前心情紧张.

(1)根据以上数据,作出考前心情是否紧张与性格类型的 2×2 列联表,并估计性格外向的学生中考前心情紧张的概率;

(2)作出等高堆积条形图,利用图形判断性格类型对考前心情是否有影响.

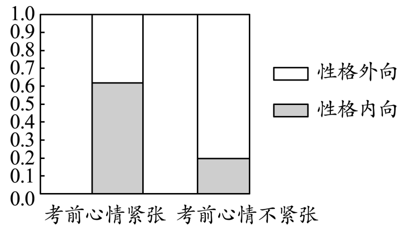
**解**:(1)2×2 列联表如下:

单位:人

| 考前心情 | 性格类型 | | 合计 |
|---|---|---|---|
| | 内向 | 外向 | |
| 紧张 | 332 | 213 | 545 |
| 不紧张 | 94 | 381 | 475 |
| 合计 | 426 | 594 | 1 020 |

由列联表中数据,估计性格外向的学生中考前心情紧张的概率为$\dfrac{213}{594}=\dfrac{71}{198}$.

(2)相应的等高堆积条形图如图所示.




图中阴影部分表示考前心情紧张与考前心情不紧张的学生中性格内向的学生所占的比例，从图中可以看出，考前心情紧张的样本中性格内向的学生占的比例比考前心情不紧张的样本中性格内向的学生占的比例高，所以可以认为性格类型对考前心情有影响.

## B组 应用·实践

**1.** 有甲、乙两个班级进行数学考试，按照大于等于85分为优秀，85分以下为非优秀，得到 $2\times2$ 列联表如下：

单位：人

| 班级 | 优秀情况 | | 合计 |
|---|---|---|---|
| | 优秀 | 非优秀 | |
| 甲班 | 10 | $b$ | |
| 乙班 | $c$ | 30 | |
| 合计 | | | 105 |

已知在105人中随机抽取1人，成绩优秀的概率为 $\frac{2}{7}$，则下列说法正确的是 （　　）

A.列联表中 $c$ 的值为30，$b$ 的值为35
B.列联表中 $c$ 的值为15，$b$ 的值为50
C.列联表中 $c$ 的值为20，$b$ 的值为50
D.由列联表可看出成绩与班级有关系

D　**解析**：依题意 $\frac{10+c}{105}=\frac{2}{7}$，解得 $c=20$，由 $10+20+b+30=105$，解得 $b=45$.

补全 $2\times2$ 列联表如下：

单位：人

| 班级 | 优秀情况 | | 合计 |
|---|---|---|---|
| | 优秀 | 非优秀 | |
| 甲班 | 10 | 45 | 55 |
| 乙班 | 20 | 30 | 50 |
| 合计 | 30 | 75 | 105 |

甲班的优秀率为 $\frac{10}{55}=\frac{2}{11}$，乙班的优秀率为 $\frac{20}{50}=\frac{2}{5}$，$\frac{2}{11}<\frac{2}{5}$，所以成绩与班级有关.所以选项D正确，选项A，B，C错误.

故选D.

**2.** （多选）已知两个分类变量 $X$，$Y$，它们分别取值于 $\{x_1,x_2\}$ 和 $\{y_1,y_2\}$，其 $2\times2$ 列联表为：

| $X$ | $Y$ | | 合计 |
|---|---|---|---|
| | $y_1$ | $y_2$ | |
| $x_1$ | $a$ | $b$ | $a+b$ |
| $x_2$ | $c$ | $d$ | $c+d$ |
| 合计 | $a+c$ | $b+d$ | $a+b+c+d$ |

若两个分类变量 $X$，$Y$ 没有关系，则下列结论正确的是 （　　）

A.$ad\approx bc$

B.$\frac{a}{a+b}\approx\frac{c}{c+d}$

C.$\frac{c+d}{a+b+c+d}\approx\frac{b+d}{a+b+c+d}$

D.$\frac{c+a}{a+b+c+d}\approx\frac{b+d}{a+b+c+d}$

AB　**解析**：因为分类变量 $X$，$Y$ 没有关系，

所以 $\frac{a}{a+b}\approx\frac{c}{c+d}$，化简得 $ad\approx bc$，

所以A，B正确，C，D显然不正确.

**3.** 在对人们休闲方式的一次调查中，共调查了110人，其中女性50人，男性60人.女性中有30人主要的休闲方式是看电视，另外20人主要的休闲方式是运动；男性中有20人主要的休闲方式是看电视，另外40人主要的休闲方式是运动.

(1)根据以上数据建立一个 $2\times2$ 列联表；

(2)由 $2\times2$ 列联表判断性别对休闲方式是否有影响.

**解**：(1) $2\times2$ 列联表如下：

单位：人

| 性别 | 休闲方式 | | 合计 |
|---|---|---|---|
| | 看电视 | 运动 | |
| 女 | 30 | 20 | 50 |
| 男 | 20 | 40 | 60 |
| 合计 | 50 | 60 | 110 |

(2)根据(1)中 $2\times2$ 列联表中的数据，可得女性中休闲方式为看电视的频率为 $\frac{30}{50}=0.6$，男性中休闲方式为看电视的频率为 $\frac{20}{60}\approx0.333$，二者差别较大，所以认为性别对休闲方式有影响.

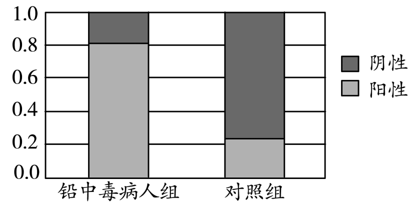
**4.** 为了解铅的毒素对人的尿棕色素呈阴性或阳性是否有影响，分别对铅中毒病人组和对照组的尿液作尿棕色素定性检查，结果如下：

单位：人

| 组别 | 尿棕色素 | | 合计 |
|---|---|---|---|
| | 阳性 | 阴性 | |
| 铅中毒病人组 | 29 | 7 | 36 |
| 对照组 | 9 | 28 | 37 |
| 合计 | 38 | 35 | 73 |

试画出等高堆积条形图，分析铅中毒病人组和对照组的尿棕色素阳性数有无差别，铅的毒素对人的尿棕色素呈阴性或阳性是否有影响？

**解**：等高堆积条形图如图所示.




其中两个浅色条的高分别代表铅中毒病人组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，铅中毒病人组与对照组中，尿棕色素为阳性的频率有明显差异，因此可以认为铅的毒素对人的尿棕色素呈阴性或阳性有影响.

## 8.3.2　独立性检验

**学习任务目标**

**1.** 了解随机变量 $\chi^2$ 的意义.

**2.** 通过实例，了解 $\chi^2$ 独立性检验及其应用.

## ◦ 问题式预习 ◦

**【知识清单】**

**知识点　独立性检验的基本思想**

(1)定义：利用 $\chi^2$ 的取值推断分类变量 $X$ 和 $Y$ 是否独立的方法称为 $\chi^2$ 独立性检验，读作“卡方独立性检验”，简称独立性检验.

(2)公式：$\chi^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$，其中 $n=a+b+c+d$.

(3)应用独立性检验解决实际问题大致应包括以下几个主要环节：

①提出零假设 $H_0$：$X$ 和 $Y$ 相互独立，并给出在问题中的解释.

②根据抽样数据整理出 $2\times2$ 列联表，计算 $\chi^2$ 的值，并与临界值 $x_\alpha$ 比较.

③根据检验规则得出推断结论.

④在 $X$ 和 $Y$ 不独立的情况下，根据需要，通过比较相应的频率，分析 $X$ 和 $Y$ 间的影响规律.

注意，上述几个环节的内容可以根据不同情况进行调整.例如，在有些时候，分类变量的抽样数据列联表是问题中给定的.

**【概念辨析】**

**1.** 判断正误(正确的打“√”，错误的打“×”).

(1)应用独立性检验的基本思想对两个变量间的关系作出的推断一定是正确的.　( × )

(2) $\chi^2$ 是判断两个变量是否相关的统计量.　( √ )

(3)在独立性检验中，若 $\chi^2$ 越大，则两个分类变量有关联的可能性越大.　( √ )

**2.** 下列关于 $\chi^2$ 的说法正确的是　(　　)

A. $\chi^2$ 越大，“两个变量有关联”的可信度越低

B. $\chi^2$ 越大，“两个变量无关”的可信度越高

C. $\chi^2$ 越小，“两个变量有关联”的可信度越低

D. $\chi^2$ 越小，“两个变量无关”的可信度越低

C　**解析**：$\chi^2$ 越大，“两个变量有关联”的可信度越高，“两个变量无关”的可信度越低；相反，$\chi^2$ 越小，“两个变量有关联”的可信度越低，“两个变量无关”的可信度越高.

**3.** 对于变量 $A$，$B$ 的独立性检验，根据下列条件可依据小概率值 $\alpha=0.05$ 的独立性检验，推断“$A$ 与 $B$ 有关联且犯错误的概率不超过 0.05”的是　( D )

A. $\chi^2=2.700$

B. $\chi^2=2.710$

C. $\chi^2=3.765$

D. $\chi^2=5.014$

**4.** 请思考并回答下列问题：

(1)如何理解临界值？

**提示**：①小概率值 $\alpha$ 的临界值 $x_\alpha$ 是一个正实数.

② $P(\chi^2\geqslant x_\alpha)=\alpha$ 时，$x_\alpha$ 称为 $\alpha$ 的临界值，即对于不同的小概率值 $\alpha$，有不同的临界值 $x_\alpha$.

③基于小概率值 $\alpha$ 的检验规则：当 $\chi^2 \geqslant x_\alpha$ 时，我们就推断 $H_0$ 不成立，即此时小概率事件不大可能发生，认为两变量之间不独立，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，可以认为两变量之间独立.

(2)对于已经获取的成对样本数据，检验结论“两个变量之间有关联”的实际含义是什么？检验结论“两个变量之间没有关联”的实际含义又是什么？

**提示**：检验结论“两个变量之间有关联”是“两个变量不独立”的另一种说法，指在零假设“两个变量独立”之下，成对样本数据显示在一次试验中某个不利于这个假设的小概率事件发生了，由此推断零假设不成立，从而得出“两个变量不独立”的检验结论.检验结论“两个变量之间没有关联”是“两个变量独立”的另一种说法，指在零假设“两个变量独立”之下，成对样本数据显示在一次试验中某个不利于这个假设的小概率事件没有发生，因此不能推断零假设不成立，按照通常的习惯接受零假设，即得出“两个变量独立”的检验结论.

## ○ 任务型课堂 ○

### 任务1 独立性检验的基本思想

#### 探究活动

**例1** 在某医院，因为患心脏病而住院的600名男性病人中，有200人秃顶，而另外750名不是因为患心脏病而住院的男性病人中有150人秃顶.

(1)填写下列2×2列联表：

单位：人

| 秃顶情况 | 患病情况 | | 合计 |
|---|---|---|---|
| | 患心脏病 | 患其他病 | |
| 秃顶 | | | |
| 不秃顶 | | | |
| 合计 | | | |

(2)根据表中数据估计秃顶病患中患心脏病的概率 $P_1$ 和不秃顶病患中患心脏病的概率 $P_2$，并用两个估计概率判断秃顶与患心脏病是否有关.

**解**：(1)2×2列联表如下：

单位：人

| 秃顶情况 | 患病情况 | | 合计 |
|---|---|---|---|
| | 患心脏病 | 患其他病 | |
| 秃顶 | 200 | 150 | 350 |
| 不秃顶 | 400 | 600 | 1 000 |
| 合计 | 600 | 750 | 1 350 |

(2) $P_1=\dfrac{200}{350}=\dfrac{4}{7}$，$P_2=\dfrac{400}{1\ 000}=\dfrac{2}{5}$.

由于 $P_1$ 与 $P_2$ 相差较大，所以判断秃顶与患心脏病有关.

[一题多思]

**思考1.**依据小概率值 $\alpha=0.001$ 的独立性检验，能否认为秃顶与患心脏病有关？请说明理由.

参考数据：

| $\alpha$ | 0.1 | 0.05 | 0.025 | 0.01 | 0.005 | 0.001 |
|---|---|---|---|---|---|---|
| $x_\alpha$ | 2.706 | 3.841 | 5.024 | 6.635 | 7.879 | 10.828 |

**解**：零假设为 $H_0$：秃顶与患心脏病无关.

由例1(1)中列联表中数据计算可得 $\chi^2=\dfrac{1\ 350\times(200\times600-150\times400)^2}{350\times1\ 000\times600\times750}=\dfrac{216}{7}\approx30.86>10.828=x_{0.001}$，

所以依据小概率值 $\alpha=0.001$ 的 $\chi^2$ 独立性检验，我们推断 $H_0$ 不成立，即认为秃顶与患心脏病有关，此推断犯错误的概率不大于0.001.

**思考2.**例1与思考1根据不同的分析方法得出了相同的结论，但哪种方法更科学合理些？

**解**：根据 $\chi^2$ 独立性检验得到的结果更理性、更全面，理论依据也更充分.

**【探究总结】**

**用独立性检验解决实际问题的基本步骤**

(1)根据相关数据列出2×2列联表.

(2)提出零假设 $H_0$：$X$ 和 $Y$ 相互独立.

(3)计算：将2×2列联表中的数据代入公式求出 $\chi^2$ 的值，并与临界值 $x_\alpha$ 比较.

(4)根据检验规则得出结论：如果 $\chi^2\geqslant x_\alpha$，那么“$X$ 与 $Y$ 有关(或不独立)”；如果 $\chi^2<x_\alpha$，那么“$X$ 与 $Y$ 无关(或独立)”.

#### 应用迁移

1.在某病毒疫苗的研发过程中，需要利用基因编辑小鼠进行动物试验.现随机抽取100只基因编辑小鼠对该病毒疫苗进行试验，得到如下2×2列联表(部分数据缺失)：

单位:只

| 注射疫苗情况 | 被感染情况 | | 合计 |
|---|---|---|---|
| | 被感染 | 未被感染 | |
| 注射疫苗 | 10 | | 50 |
| 未注射疫苗 | | 30 | 50 |
| 合计 | 30 | | 100 |

若根据小概率值 $\alpha=m$ 的独立性检验,判断给基因编辑小鼠注射该种疫苗能起到预防被该病毒感染的效果,则 $m$ 的值为 (　　)

附:$\chi^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$,$n=a+b+c+d$.

| $\alpha$ | 0.1 | 0.05 | 0.01 | 0.005 | 0.001 |
|---|---|---|---|---|---|
| $x_\alpha$ | 2.706 | 3.841 | 6.635 | 7.879 | 10.828 |

A.0.001　　B.0.05　　C.0.01　　D.0.005

B　**解析**:完善 2×2 列联表如下:

单位:只

| 注射疫苗情况 | 被感染情况 | | 合计 |
|---|---|---|---|
| | 被感染 | 未被感染 | |
| 注射疫苗 | 10 | 40 | 50 |
| 未注射疫苗 | 20 | 30 | 50 |
| 合计 | 30 | 70 | 100 |

零假设为 $H_0$:给基因编辑小鼠注射该种疫苗不能起到预防被该病毒感染的效果.

因为 $\chi^2=\dfrac{100\times(10\times30-40\times20)^2}{50\times50\times30\times70}\approx4.762$,$3.841<4.762<6.635$,

所以根据小概率值 $\alpha=0.05$ 的独立性检验,推断 $H_0$不成立,

即认为“给基因编辑小鼠注射该种疫苗能起到预防被该病毒感染的效果”,该推断犯错误的概率不超过 0.05.故选 B.

**2.** 为考察棉花种子是否生病跟是否经过处理之间的关系,进行试验后得到下表数据:

单位:粒

| 生病情况 | 种子是否经过处理 | | 合计 |
|---|---|---|---|
| | 经过处理 | 未经过处理 | |
| 生病 | 32 | 101 | 133 |
| 不生病 | 61 | 213 | 274 |
| 合计 | 93 | 314 | 407 |

依据小概率值 $\alpha=0.1$ 的独立性检验,可得出 (　　)

A.种子是否生病跟是否经过处理有关

B.种子是否生病跟是否经过处理无关

C.种子是否经过处理决定是否生病

D.以上结论都是错误的

B　**解析**:由题表中数据可得 $\chi^2=\dfrac{407\times(32\times213-101\times61)^2}{133\times274\times93\times314}\approx0.164<2.706=x_{0.1}$,

即没有充分证据推断 $H_0$ 不成立,因此可以认为种子是否生病跟是否经过处理无关.

## 任务2 独立性检验的综合应用

### 探究活动

**例 2**　甲、乙两机床加工同一种零件,抽检得到它们加工后的零件尺寸 $x$(单位:cm)及个数 $y$,如下表:

| 零件尺寸 $x$/cm | | 1.01 | 1.02 | 1.03 | 1.04 | 1.05 |
|---|---|---|---|---|---|---|
| 零件个数 $y$ | 甲 | 3 | 7 | 8 | 9 | 3 |
| | 乙 | 7 | 4 | 4 | 4 | $a$ |

由表中数据得 $y$ 关于 $x$ 的经验回归方程为$\hat{y}=-91+100x(1.01\leqslant x\leqslant1.05)$.已知合格零件的尺寸为$(1.03\pm0.01)$cm.

(1)根据已知数据完成下面的 2×2 列联表:

单位:个

| 机床 | 零件的质量 | | 合计 |
|---|---|---|---|
| | 合格 | 不合格 | |
| 甲 | | | |
| 乙 | | | |
| 合计 | | | |

(2)依据小概率值 $\alpha=0.01$ 的独立性检验,分析加工零件的质量与加工机床是否有关.

**解**:(1)由题表可得 $\overline{x}=1.03$,$\overline{y}=\dfrac{a+49}{5}$.

由 $\hat{y}=-91+100x$,知$\dfrac{a+49}{5}=-91+100\times1.03$.

所以 $a=11$.

由于合格零件的尺寸为$(1.03\pm0.01)$cm,

故 2×2 列联表为

单位:个

| 机床 | 零件的质量 | | 合计 |
|---|---|---|---|
| | 合格 | 不合格 | |
| 甲 | 24 | 6 | 30 |
| 乙 | 12 | 18 | 30 |
| 合计 | 36 | 24 | 60 |

(2)零假设为 $H_0$:加工零件的质量与加工机床无关.

由（1）中列联表数据计算得 $\chi^2=\dfrac{60\times(24\times18-6\times12)^2}{30\times30\times36\times24}=10>6.635=x_{0.01}$.

根据小概率值 $\alpha=0.01$ 的独立性检验，我们推断 $H_0$ 不成立，即认为加工零件的质量与加工机床有关，此推断犯错误的概率不大于 0.01.

**【探究总结】**

**1.** 解答独立性检验问题的关键在于正确计算 $\chi^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$ 的值，再将它与临界值 $x_\alpha$ 的大小作比较来判断零假设是否成立，从而使问题得到解决.

**2.** 独立性检验问题规律性强，解题方法比较格式化，填表、计算、比较、分析即可，要熟悉解题流程，不难解决问题.

## 应用迁移

**1.** 某学校举办了“传承中华优秀传统文化”宣传活动，学校从全体学生中抽取了 100 人对该宣传活动的了解情况进行问卷调查，统计结果如下：

单位：人

| 了解情况 | 性别 | | 合计 |
|---|---|---|---|
| | 男 | 女 | |
| 了解 | | 20 | |
| 不了解 | 20 | | 40 |
| 合计 | | | |

(1)将 2×2 列联表补充完整.

(2)根据 $\alpha=0.05$ 的独立性检验，能否认为该校学生对该宣传活动的了解情况与性别有关联？

(3)若把上表中的频率视作概率，现从了解该活动的学生中随机抽取 3 人参加传统文化知识竞赛.记抽取的 3 人中女生人数为 $X$，求随机变量 $X$ 的分布列和数学期望.

附：$\chi^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$，其中 $n=a+b+c+d$.

| $\alpha$ | 0.1 | 0.05 | 0.01 | 0.001 |
|---|---|---|---|---|
| $x_\alpha$ | 2.706 | 3.841 | 6.635 | 10.828 |

**解**：(1)由题得 2×2 列联表如下：

单位：人

| 了解情况 | 性别 | | 合计 |
|---|---|---|---|
| | 男 | 女 | |
| 了解 | 40 | 20 | 60 |
| 不了解 | 20 | 20 | 40 |
| 合计 | 60 | 40 | 100 |

(2)零假设 $H_0$：该校学生对该宣传活动的了解情况与性别无关.

由(1)可得 $\chi^2=\dfrac{100\times(40\times20-20\times20)^2}{60\times40\times60\times40}=\dfrac{25}{9}\approx2.778<3.841$，

则根据小概率值 $\alpha=0.05$ 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，即认为该校学生对该宣传活动的了解情况与性别无关.

(3)由(1)可知抽取的 100 名学生中了解该活动的学生中男生和女生分别为 40 人和 20 人，

所以从了解该活动的学生中随机抽取 1 人参加传统文化知识竞赛，抽取的是女生的概率为 $\dfrac{20}{40+20}=\dfrac{1}{3}$，

则由题意可知 $X=0,1,2,3$，且 $X\sim B\left(3,\dfrac{1}{3}\right)$，

所以 $P(X=0)=C_3^0\times\left(1-\dfrac{1}{3}\right)^3=\dfrac{8}{27}$，

$P(X=1)=C_3^1\times\left(1-\dfrac{1}{3}\right)^2\times\dfrac{1}{3}=\dfrac{12}{27}=\dfrac{4}{9}$，

$P(X=2)=C_3^2\times\left(1-\dfrac{1}{3}\right)\times\left(\dfrac{1}{3}\right)^2=\dfrac{6}{27}=\dfrac{2}{9}$，

$P(X=3)=C_3^3\times\left(\dfrac{1}{3}\right)^3=\dfrac{1}{27}$，

所以随机变量 $X$ 的分布列为

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P$ | $\dfrac{8}{27}$ | $\dfrac{4}{9}$ | $\dfrac{2}{9}$ | $\dfrac{1}{27}$ |

$E(X)=np=3\times\dfrac{1}{3}=1$.

**2.** 羽毛球正式比赛时的规则为：若发球方胜，则发球方得 1 分，且继续在下一回合发球；若接球方胜，则接球方得 1 分，且成为下一回合发球方.而在训练中，可以不遵循胜方发球的规则.已知甲、乙共进行了 60 回合的羽毛球比赛，得到如下待完善的 2×2 列联表：

单位：分

| 发球方 | 得分情况 | | 合计 |
|---|---|---|---|
| | 甲得分 | 乙得分 | |
| 甲 | 18 | | |
| 乙 | | 24 | |
| 合计 | 24 | | 60 |

(1)完成 2×2 列联表，依据小概率值 $\alpha=0.01$ 的独立性检验，能否认为获胜与接、发球有关？

(2)以上述 2×2 列联表中甲、乙各自接、发球的得分频率分别作为正式比赛中每一回合甲、乙各自接、发球的得分概率.

①若在正式比赛中,第1回合是甲先发球,设第 $i$ 回合是甲发球的概率为 $p_i$,证明:$\left\{p_i-\frac{1}{3}\right\}$是等比数列;

②已知:若 $X,Y$ 是随机变量,则有 $E(X+Y)=E(X)+E(Y)$.若在正式比赛中,第1回合是甲先发球,求甲、乙连续进行60回合比赛后甲的总得分 $X$ 的期望,并据此判断 $2\times2$ 列联表中的数据是正式比赛数据还是训练数据.

附:$\chi^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$,其中 $n=a+b+c+d$.

| $\alpha$ | 0.15 | 0.1 | 0.05 | 0.01 | 0.001 |
|---|---|---|---|---|---|
| $x_\alpha$ | 2.072 | 2.706 | 3.841 | 6.635 | 10.828 |

(1)**解**:完善 $2\times2$ 列联表如表所示.

单位:分

| 发球方 | 得分情况 | | 合计 |
|---|---|---|---|
| | 甲得分 | 乙得分 | |
| 甲 | 18 | 12 | 30 |
| 乙 | 6 | 24 | 30 |
| 合计 | 24 | 36 | 60 |

零假设为 $H_0$:获胜与接、发球无关.

由列联表中数据计算得 $\chi^2=\frac{60\times(18\times24-12\times6)^2}{30\times30\times24\times36}=10>6.635=x_{0.01}$,

根据小概率值 $\alpha=0.01$ 的独立性检验,我们推断 $H_0$ 不成立,即认为获胜与接、发球有关,此推断犯错误的概率不超过0.01.

(2)①**证明**:在甲发球的情况下,甲得分的概率为 $\frac{18}{30}=\frac{3}{5}$,乙得分的概率为 $1-\frac{3}{5}=\frac{2}{5}$;

在乙发球的情况下,甲得分的概率为 $\frac{6}{30}=\frac{1}{5}$,乙得分的概率为 $1-\frac{1}{5}=\frac{4}{5}$.

若第 $i(i\geqslant2)$ 回合是甲发球,则要分两种情况讨论:第 $(i-1)$ 回合是甲发球且甲得分,或第 $(i-1)$ 回合是乙发球且甲得分,

所以 $p_i=p_{i-1}\times\frac{3}{5}+(1-p_{i-1})\times\frac{1}{5}$,

即 $p_i=\frac{2}{5}p_{i-1}+\frac{1}{5}$,

故 $p_i-\frac{1}{3}=\frac{2}{5}p_{i-1}-\frac{2}{15}=\frac{2}{5}\left(p_{i-1}-\frac{1}{3}\right)(i\geqslant2)$.

又 $p_1=1$,所以 $p_1-\frac{1}{3}=\frac{2}{3}\neq0$,

故$\left\{p_i-\frac{1}{3}\right\}$是以$\frac{2}{3}$为首项,$\frac{2}{5}$为公比的等比数列.

②**解**:由①可得 $p_i-\frac{1}{3}=\frac{2}{3}\times\left(\frac{2}{5}\right)^{i-1}$,故 $p_i=\frac{1}{3}+\frac{2}{3}\times\left(\frac{2}{5}\right)^{i-1}$.

设第 $i$ 回合甲得分为 $X_i$,显然 $X_i$ 服从两点分布,且事件“$X_i=1$”等价于“第 $(i+1)$ 回合是甲发球”,故 $E(X_i)=p_{i+1}$.

又甲、乙连续进行60回合比赛后,甲的总得分 $X=X_1+X_2+\cdots+X_{60}$,

故 $E(X)=\sum\limits_{i=1}^{60}E(X_i)=\sum\limits_{i=1}^{60}p_{i+1}=\sum\limits_{i=1}^{60}\left[\frac{1}{3}+\frac{2}{3}\times\left(\frac{2}{5}\right)^i\right]$

$$=\frac{1}{3}\times60+\frac{\frac{2}{3}\times\frac{2}{5}\times\left[1-\left(\frac{2}{5}\right)^{60}\right]}{1-\frac{2}{5}}=20+\frac{4}{9}\times\left[1-\left(\frac{2}{5}\right)^{60}\right]=\frac{184}{9}-\frac{4}{9}\times\left(\frac{2}{5}\right)^{60}.$$

因为 $E(X)\approx\frac{184}{9}\neq24$,故列联表中的数据是训练数据.

## 提质归纳

独立性检验

- 统计量:$\chi^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$
- 基于小概率值 $\alpha$ 的检验规则
  - 当 $\chi^2\geqslant x_\alpha$ 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
- 基本步骤
  - (1)提出零假设 $H_0$: $X$ 和 $Y$ 相互独立,并给出在问题中的解释.
  - (2)根据抽样数据整理出 $2\times2$ 列联表,计算 $\chi^2$ 的值,并与临界值 $x_\alpha$ 比较.
  - (3)根据检验规则得出推断结论.
  - (4)在 $X$ 和 $Y$ 不独立的情况下,根据需要,通过比较相应的频率,分析 $X$ 和 $Y$ 间的影响规律

# 课后素养评价（二十二） 独立性检验

## A组 学习·理解

1.下列关于独立性检验的说法正确的是（　　）

A.独立性检验是对两个变量是否具有线性相关关系的一种检验

B.独立性检验可以100%确定两个变量之间是否具有某种关系

C.利用$\chi^2$独立性检验推断吸烟与患肺病的关联，根据小概率值$\alpha=0.01$的独立性检验，认为吸烟与患肺病有关系时，我们可以说在100个吸烟的人中，有99人患肺病

D.对于独立性检验，随机变量$\chi^2$的值越小，判定"两变量有关系"犯错误的概率越大

D　**解析**：对于A，独立性检验是通过计算$\chi^2$的值来判断两个变量存在关联的可能性的一种方法，并非检验二者是否是线性相关，故错误；

对于B，独立性检验并不能100%确定两个变量相关，故错误；

对于C，99%是指"吸烟"和"患肺病"存在关联的可能性，故错误；

对于D，根据$\chi^2$的定义可知该选项正确.

故选D.

2.某校随机调查了100名高中生是否喜欢篮球，按照性别区分得到2×2列联表，经计算得$\chi^2=8.133$.根据独立性检验的相关知识，对照下表，可以认为"是否喜欢篮球与性别有关"犯错误的概率不超过（　　）

| $\alpha$ | 0.05 | 0.01 | 0.005 | 0.001 |
|---|---|---|---|---|
| $x_\alpha$ | 3.841 | 6.635 | 7.879 | 10.828 |

A.0.05　　B.0.005

C.0.01　　D.0.001

B　**解析**：因为$\chi^2=8.133>7.879=x_{0.005}$，

所以犯错误的概率不超过0.005.

故选B.

3.某班主任对全班50名学生进行了作业量的评价调查，所得数据如下表所示.

单位：名

| 性别 | 学生对作业量的评价 | | 合计 |
|---|---|---|---|
| | 认为作业量大 | 认为作业量不大 | |
| 男 | 18 | 9 | 27 |
| 女 | 8 | 15 | 23 |
| 合计 | 26 | 24 | 50 |

已知$x_{0.05}=3.841$，$x_{0.025}=5.024$，则认为"对作业量的评价与学生的性别有关"犯错误的概率（　　）

A.不超过0.01　　B.不超过0.025

C.不超过0.10　　D.无法判断

B　**解析**：由题表中数据可得$\chi^2=\dfrac{50\times(18\times15-9\times8)^2}{27\times23\times26\times24}\approx5.059>5.024=x_{0.025}$，

所以认为"评价作业量的大小与学生的性别有关"犯错误的概率不超过0.025.故选B.

4.（多选）千百年来，我国劳动人民在生产实践中根据云的形状、走向、速度、厚度、颜色等的变化，总结了丰富的"看云识天气"的经验，并将这些经验编成谚语，如"天上钩钩云，地上雨淋淋""日落云里走，雨在半夜后"……小波同学为了验证"日落云里走，雨在半夜后"，随机观察了他所在地区的100天日落情况和后半夜天气，得到如下2×2列联表：

单位：天

| 日落云里走 | 后半夜天气 | | 合计 |
|---|---|---|---|
| | 下雨 | 未下雨 | |
| 出现 | 25 | 5 | 30 |
| 未出现 | 25 | 45 | 70 |
| 合计 | 50 | 50 | 100 |

并计算得到$\chi^2\approx19.05$，下列小波对该地区天气的判断正确的是（　　）

A.后半夜下雨的概率约为$\dfrac{1}{2}$

B.未出现"日落云里走"时，后半夜下雨的概率约为$\dfrac{5}{9}$

C.根据$\alpha=0.001$的独立性检验，可以推断"日落云里走"是否出现与当晚后半夜是否下雨有关

D.根据小概率值$\alpha=0.001$的独立性检验，若出现

“日落云里走”,则后半夜有99.9%的可能会下雨

AC　**解析**:由题意,把频率看作概率,可得后半夜下雨的概率约为$\frac{50}{100}=\frac{1}{2}$,故A判断正确;未出现“日落云里走”时,后半夜下雨的概率约为$\frac{25}{25+45}=\frac{5}{14}$,故B判断错误;$\chi^2\approx19.05>10.828=x_{0.001}$,根据小概率值$\alpha=0.001$的独立性检验,认为“日落云里走”是否出现与当晚后半夜是否下雨有关,故C判断正确,D判断错误.故选AC.

**5.**“3+1+2”新高考模式中,“1”表示考生从物理、历史两门首选科目中选择一门,“2”表示考生从思想政治、地理、化学、生物学四门再选科目中选择两门.某中学为调查高一年级学生的选科倾向,随机抽取了300人,其中首选物理的有220人,首选历史的有80人,统计选择各科人数如表所示,则下列说法正确的是　(　　)

单位:人

| 首选科目 | 再选科目 | | | |
|---|---|---|---|---|
| | 思想政治 | 地理 | 化学 | 生物学 |
| 物理 | 80 | 100 | 145 | 115 |
| 历史 | 50 | 45 | 30 | 35 |

A.首选物理的学生中选择思想政治的比例比首选历史的学生中选择思想政治的比例高

B.首选物理的学生中选择地理的比例比首选历史的学生中选择地理的比例高

C.根据小概率值$\alpha=0.1$的独立性检验,我们认为是否选择生物学与首选科目无关

D.根据小概率值$\alpha=0.1$的独立性检验,我们认为是否选择生物学与首选科目有关

C　**解析**:对于A项,$\frac{80}{220}=\frac{4}{11}<\frac{50}{80}=\frac{5}{8}$,故A项错误;

对于B项,$\frac{100}{220}=\frac{5}{11}<\frac{45}{80}=\frac{9}{16}$,故B项错误;

对于C,D项,根据已知,可列出如下2×2列联表:

单位:人

| 首选科目 | 生物学 | | 合计 |
|---|---|---|---|
| | 选择 | 不选择 | |
| 物理 | 115 | 105 | 220 |
| 历史 | 35 | 45 | 80 |
| 合计 | 150 | 150 | 300 |

零假设为$H_0$:是否选择生物学与首选科目无关.

由表中数据得$\chi^2=\frac{300\times(115\times45-105\times35)^2}{220\times80\times150\times150}=\frac{75}{44}\approx1.705<2.706=x_{0.1}$,

所以根据小概率值$\alpha=0.1$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,即认为是否选择生物学与首选科目无关,故C项正确,D项错误.故选C.

**6.**下表是某校对随机选取的304名新入校的学生进行调查所得的2×2列联表:

单位:人

| 性别 | 想学专业 | | 合计 |
|---|---|---|---|
| | 知道 | 不知道 | |
| 男 | 63 | 117 | 180 |
| 女 | 42 | 82 | 124 |
| 合计 | 105 | 199 | 304 |

根据表中数据,下列说法正确的是________.(填序号)

①性别与是否知道想学专业有关;

②性别与是否知道想学专业无关;

③女生知道想学专业的概率比男生大.

②　**解析**:由题表中数据计算可得$\chi^2=\frac{304\times(63\times82-117\times42)^2}{180\times124\times105\times199}\approx0.041<2.706=x_{0.1}$,所以性别与是否知道想学专业无关.

**7.**现给出某零售店在某日上午购买两种颜色玩偶的人数统计表(假定每人限购一个玩偶):

单位:人

| 顾客性别 | 玩偶颜色 | | 合计 |
|---|---|---|---|
| | 蓝色 | 粉色 | |
| 男 | $\frac{5}{6}a$ | $\frac{1}{6}a$ | $a$ |
| 女 | $\frac{2}{3}a$ | $\frac{4}{3}a$ | $2a$ |
| 合计 | $\frac{3}{2}a$ | $\frac{3}{2}a$ | $3a$ |

(1)若认为“顾客购买的玩偶颜色与顾客性别有关”犯错误的概率不超过0.01,求$a$的最小值;

(2)在(1)中$a$取得最小值的条件下,现从所有顾客中选出9人,记选到的人中女顾客人数为$X$,求$X$的分布列及数学期望.

附:$\chi^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$.

| $\alpha$ | 0.05 | 0.01 | 0.001 |
|---|---|---|---|
| $x_\alpha$ | 3.841 | 6.635 | 10.828 |

**解**:(1)不妨给出零假设$H_0$:顾客购买的玩偶颜色与顾客性别无关.

由题意知该假设成立的概率小于等于0.01,

又$P(\chi^2\geqslant6.635)=0.01$,

所以 $\chi^2=\dfrac{3a\left(\dfrac{5a}{6}\times\dfrac{4a}{3}-\dfrac{a}{6}\times\dfrac{2a}{3}\right)^2}{a\times 2a\times\dfrac{3a}{2}\times\dfrac{3a}{2}}=\dfrac{2a}{3}\geqslant 6.635$，解得 $a\geqslant 9.952\ 5$.

又 $a\in\mathbf{Z}$，$\dfrac{a}{6}\in\mathbf{Z}$，所以 $a$ 的最小值为 12.

(2)由(1)知 $a$ 的最小值为 12，

此时女顾客一共有 24 人，男顾客一共有 12 人.

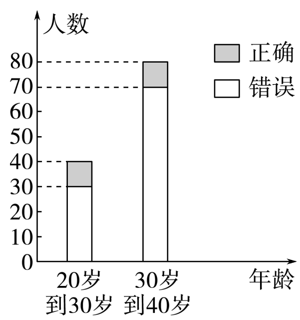
从所有顾客中选出 9 人，所以 $X$ 的所有可能取值是 0，1，2，3，4，5，6，7，8，9，

则 $X$ 服从超几何分布，且 $N=36$，$M=24$，$n=9$.

所以 $X$ 的分布列为 $P(X=i)=\dfrac{C_{24}^{i}C_{12}^{9-i}}{C_{36}^{9}}$，$i=0,1,2,3,4,5,6,7,8,9$，

所以 $E(X)=\dfrac{nM}{N}=\dfrac{216}{36}=6$.

**8.**某电视台推出了一个游戏节目：选手面对 1～8 号 8 扇大门，依次按响门上的门铃，门铃会播放一段音乐（将一首经典流行歌曲以单音色旋律的方式演绎），选手需正确回答出这首歌的名字，方可获得该扇门对应的奖金.该电视台经过调查，发现参赛选手可分为两个年龄段：20 岁到 30 岁，30 岁到 40 岁，这些选手猜对歌曲名称与否的人数如图所示.




(1)列出 2×2 列联表.

(2)依据小概率值 $\alpha=0.1$ 的独立性检验，能否认为猜对歌曲名称与年龄有关系？说明你的理由.

**解**：(1)根据题图得到 2×2 列联表如下：

单位：人

| 年龄 | 歌曲名称 | | 合计 |
|---|---|---|---|
| | 正确 | 错误 | |
| 20 岁到 30 岁 | 10 | 30 | 40 |
| 30 岁到 40 岁 | 10 | 70 | 80 |
| 合计 | 20 | 100 | 120 |

(2)有关系.理由如下：

零假设为 $H_0$：猜对歌曲名称与年龄无关.

根据(1)中列联表的数据计算得

$\chi^2=\dfrac{120\times(10\times 70-30\times 10)^2}{40\times 80\times 20\times 100}=3>2.706=x_{0.1}$.

依据 $\alpha=0.1$ 的独立性检验，推断 $H_0$ 不成立，即认为猜对歌曲名称与年龄有关系，此推断犯错误的概率不超过 0.1.

## B组 应用·实践

**1.**利用独立性检验判断喜欢参加体育活动是否与性别有关，零假设为 （　　）

A.$H_0$：男性喜欢参加体育活动

B.$H_0$：女性不喜欢参加体育活动

C.$H_0$：喜欢参加体育活动与性别有关

D.$H_0$：喜欢参加体育活动与性别无关

D　**解析**：独立性检验应先假设两个分类变量无关.

**2.**高二第二学期期中考试后，对甲、乙两个班级学生的数学考试成绩按照优秀和不优秀统计人数后，得到如下 2×2 列联表，则随机变量 $\chi^2$ 的值约为 （　　）

单位：人

| 班级 | 考试成绩 | | 合计 |
|---|---|---|---|
| | 优秀 | 不优秀 | |
| 甲班 | 11 | 34 | 45 |
| 乙班 | 8 | 37 | 45 |
| 合计 | 19 | 71 | 90 |

A.0.600　　B.0.828

C.2.712　　D.6.004

A　**解析**：随机变量 $\chi^2=\dfrac{90\times(11\times 37-34\times 8)^2}{45\times 45\times 19\times 71}\approx 0.600$.

**3.**针对时下的“短视频热”，某高校团委对学生性别和喜欢短视频是否有关联进行了一次调查，其中被调查的男生、女生人数均为 $5m(m\in\mathbf{N}^*)$，男生中喜欢短视频的人数占男生人数的 $\dfrac{4}{5}$，女生中喜欢短视频的人数占女生人数的 $\dfrac{3}{5}$.零假设为 $H_0$：喜欢短视频和性别相互独立.若依据 $\alpha=0.05$的独立性检验认为喜欢短视频和性别不相互独立，则 $m$ 的最小值为 （　　）

附：$\chi^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$，

| $\alpha$ | 0.05 | 0.01 |
|---|---|---|
| $x_\alpha$ | 3.841 | 6.635 |

A.7　B.8　C.9　D.10

C　**解析**：根据题意，不妨设 $a=4m$，$b=m$，$c=3m$，$d=2m$，

则 $\chi^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$

$=\frac{10m\cdot(5m^2)^2}{5m\cdot5m\cdot7m\cdot3m}=\frac{10m}{21}$.

由于依据 $\alpha=0.05$ 的独立性检验认为喜欢短视频和性别不相互独立，

所以根据表格可知 $\frac{10m}{21}\geqslant3.841$，解得 $m\geqslant8.066\ 1$，

所以 $m$ 最小值为 9.

故选 C.

**4.** 全国“村 BA”篮球赛点燃了全民的运动激情，深受广大球迷的喜爱.每支球队都有一个或几个主力队员，现有一支“村 BA”球队，其中甲球员是其主力队员.经统计该球队在某个赛季的所有比赛中，甲球员是否上场时该球队的胜负情况如表.

| 甲球员 | 球队的胜负情况 | | 合计 |
|---|---|---|---|
| | 胜 | 负 | |
| 上场 | 40 | | 45 |
| 未上场 | | 3 | |
| 合计 | 42 | | |

(1)完成 $2\times2$ 列联表，并判断依据小概率值 $\alpha=0.01$的独立性检验，能否认为球队的胜负与甲球员是否上场有关.

(2)由于队员的不同，甲球员主打的位置会进行调整，根据以往的数据统计，甲球员上场时，打前锋、中锋、后卫的概率分别为 0.3，0.5，0.2，相应球队赢球的概率分别为 0.7，0.8，0.6.

①当甲球员上场参加比赛时，求球队赢球的概率；

②当甲球员上场参加比赛时，在球队赢了某场比赛的条件下，求甲球员打中锋的概率.(精确到 0.01)

附：$\chi^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$，

$n=a+b+c+d$.

| $\alpha$ | 0.15 | 0.10 | 0.05 | 0.025 | 0.010 | 0.001 |
|---|---|---|---|---|---|---|
| $x_\alpha$ | 2.072 | 2.706 | 3.841 | 5.024 | 6.635 | 10.828 |

**解**：(1)根据题意，可得 $2\times2$ 列联表：

| 甲球员 | 球队的胜负情况 | | 合计 |
|---|---|---|---|
| | 胜 | 负 | |
| 上场 | 40 | 5 | 45 |
| 未上场 | 2 | 3 | 5 |
| 合计 | 42 | 8 | 50 |

零假设为 $H_0$：球队的胜负与甲球员是否上场无关.

计算得 $\chi^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$

$=\frac{50\times(40\times3-5\times2)^2}{45\times5\times42\times8}\approx8.003>6.635$，

依据小概率值 $\alpha=0.01$ 的独立性检验，推断 $H_0$ 不成立，即认为球队的胜负与甲球员是否上场有关，此推断犯错误的概率不大于 0.01.

(2)甲球员上场时，打前锋、中锋、后卫的概率分别为0.3，0.5，0.2，相应球队赢球的概率分别为 0.7，0.8，0.6.

①设事件 $A$：甲球员上场打前锋，事件 $B$：甲球员上场打中锋，事件 $C$：甲球员上场打后卫，事件 $D$：球队赢球，

则 $P(A)=0.3,P(B)=0.5,P(C)=0.2,P(D|A)=0.7,P(D|B)=0.8,P(D|C)=0.6$，

所以，当甲球员上场参加比赛时，球队赢球的概率为

$P(D)=P(A)P(D|A)+P(B)P(D|B)+P(C)P(D|C)=0.3\times0.7+0.5\times0.8+0.2\times0.6=0.73$.

②当甲球员上场参加比赛时，在球队赢了某场比赛的条件下，甲球员打中锋的概率为 $P(B|D)=\frac{P(DB)}{P(D)}=\frac{P(B)P(D|B)}{P(D)}=\frac{0.5\times0.8}{0.73}\approx0.55$.

# 迁移应用

**学习目标**

**1.**利用分布列的性质和特点，对回归模型进行更深入的分析和优化.
**2.**借助分布列的性质和特点，对独立性检验的结果进行更深入的分析和解释.

## 类型一 回归模型与分布列的综合问题

**例1** 中国新能源汽车销售火爆，A省相关部门调查了该省2024年1月至10月的新能源汽车销量情况，得到一组样本数据$(x_i,y_i)(i=1,2,\cdots,10)$，其中$x_i$即月份代码$i$，$y_i$表示第$i$个月A省新能源汽车的销量(单位:万辆)，由样本数据的散点图可知，$y$与$x$具有线性相关关系，并将这10个月的数据作了初步处理，得到下面一些统计量的值：

| $\overline{y}$ | $\sum\limits_{i=1}^{10}x_iy_i$ | $\sum\limits_{i=1}^{10}x_i^2$ | $\sum\limits_{i=1}^{10}y_i$ |
|---|---|---|---|
| 1.5 | 89.1 | 385 | 15 |

(1)建立$y$关于$x$的经验回归方程，并估计A省12月份新能源汽车的销量.

(2)为鼓励新能源汽车销售商积极参与调查，A省汽车行业协会针对新能源汽车销售商开展抽奖活动，所有费用由某新能源汽车厂商赞助.奖项共设一、二、三等奖共三个奖项，其中一、二、三等奖分别奖励2万元、1万元、0.5万元，抽中一、二、三等奖的概率分别为$\frac{1}{6}$，$\frac{1}{3}$，$\frac{1}{2}$.现有甲、乙两家汽车销售商参加了抽奖活动，假设他们是否中奖相互独立，求这两家汽车销售商所获奖金总额$X$(单位:万元)的分布列及均值.

附:对于一组数据$(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$，其经验回归方程$\hat{y}=\hat{b}x+\hat{a}$中，$\hat{b}=\frac{\sum\limits_{i=1}^{n}x_iy_i-n\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{n}x_i^2-n\overline{x}^2}$，$\hat{a}=\overline{y}-\hat{b}\,\overline{x}$.

**解**:(1)由题意得，$\overline{x}=\frac{1+2+3+\cdots+9+10}{10}=5.5$，

又$\overline{y}=1.5$，$\sum\limits_{i=1}^{10}x_iy_i=89.1$，$\sum\limits_{i=1}^{10}x_i^2=385$，

所以$\hat{b}=\frac{\sum\limits_{i=1}^{10}x_iy_i-10\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{10}x_i^2-10\overline{x}^2}$

$=\frac{89.1-10\times5.5\times1.5}{385-10\times5.5^2}=0.08$，

$\hat{a}=1.5-0.08\times5.5=1.06$，

所以$y$关于$x$的经验回归方程为$\hat{y}=0.08x+1.06$.

当$x=12$时，$\hat{y}=2.02$，

故A省12月份新能源汽车的销量约为2.02万辆.

(2)这两家汽车销售商所获得的奖金总额$X$的所有可能取值为4，3，2.5，2，1.5，1，

$P(X=4)=\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}$，

$P(X=3)=2\times\frac{1}{6}\times\frac{1}{3}=\frac{1}{9}$，

$P(X=2.5)=2\times\frac{1}{6}\times\frac{1}{2}=\frac{1}{6}$，

$P(X=2)=\frac{1}{3}\times\frac{1}{3}=\frac{1}{9}$，

$P(X=1.5)=2\times\frac{1}{3}\times\frac{1}{2}=\frac{1}{3}$，

$P(X=1)=\frac{1}{2}\times\frac{1}{2}=\frac{1}{4}$，

则$X$的分布列为

| $X$ | 4 | 3 | 2.5 | 2 | 1.5 | 1 |
|---|---|---|---|---|---|---|
| $P$ | $\frac{1}{36}$ | $\frac{1}{9}$ | $\frac{1}{6}$ | $\frac{1}{9}$ | $\frac{1}{3}$ | $\frac{1}{4}$ |

$E(X)=4\times\frac{1}{36}+3\times\frac{1}{9}+2.5\times\frac{1}{6}+2\times\frac{1}{9}+1.5\times\frac{1}{3}+1\times\frac{1}{4}=\frac{11}{6}$.

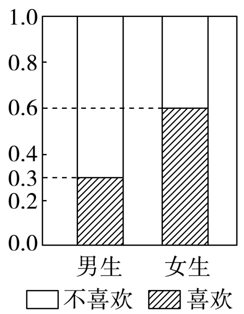
**【总结升华】**

**回归模型与分布列的综合问题的解题策略**

求经验回归方程时要充分利用已知数据，合理利用公式减少运算.求解概率问题时要注意概率模型的应用，明确所求问题对应的事件是关键.

## 类型二　独立性检验与分布列的综合问题

**例 2**　某高校为调查学生性别与是否喜欢排球运动的关系，在全校范围内采用简单随机抽样的方法，分别抽取了男生和女生各 100 人作为样本，经统计，得到了如图所示的等高堆积条形图.




(1)根据等高堆积条形图，填写下列 $2\times2$ 列联表，依据小概率值 $\alpha=0.001$ 的独立性检验，是否可以认为该校学生的性别与是否喜欢排球运动有关联？

单位：人

| 性别 | 排球运动 | | 合计 |
|---|---|---|---|
| | 喜欢 | 不喜欢 | |
| 男 | | | |
| 女 | | | |
| 合计 | | | 200 |

(2)将样本的频率视为概率，现从全校的学生中随机抽取 50 人，设其中喜欢排球运动的学生的人数为 $X$，求使得 $P(X=k)$ 取得最大值时的 $k$ 值.

附：$\chi^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$，其中 $n=a+b+c+d$，$x_{0.001}=10.828$.

**解**：(1)由题中等高堆积条形图得 $2\times2$ 列联表如下：

单位：人

| 性别 | 排球运动 | | 合计 |
|---|---|---|---|
| | 喜欢 | 不喜欢 | |
| 男 | 30 | 70 | 100 |
| 女 | 60 | 40 | 100 |
| 合计 | 90 | 110 | 200 |

零假设为 $H_0$：学生的性别与是否喜欢排球运动无关，根据列联表中的数据，

得 $\chi^2=\dfrac{200\times(30\times40-70\times60)^2}{100\times100\times90\times110}\approx18.182>10.828=x_{0.001}$，

依据小概率值 $\alpha=0.001$ 的独立性检验，可以推断 $H_0$ 不成立，即认为该校学生的性别与是否喜欢排球运动有关联，此推断犯错误的概率不大于 0.001.

(2)由(1)知，喜欢排球运动的频率为 $\dfrac{90}{200}=\dfrac{9}{20}$，

所以随机变量 $X\sim B\left(50,\dfrac{9}{20}\right)$，

则 $P(X=k)=C_{50}^{k}\left(\dfrac{9}{20}\right)^k\left(1-\dfrac{9}{20}\right)^{50-k}$ $(0\leqslant k\leqslant50,k\in\mathbf{N})$.

令 $\begin{cases}C_{50}^{k}\left(\dfrac{9}{20}\right)^k\left(1-\dfrac{9}{20}\right)^{50-k}\geqslant C_{50}^{k-1}\left(\dfrac{9}{20}\right)^{k-1}\left(1-\dfrac{9}{20}\right)^{51-k},\\ C_{50}^{k}\left(\dfrac{9}{20}\right)^k\left(1-\dfrac{9}{20}\right)^{50-k}\geqslant C_{50}^{k+1}\left(\dfrac{9}{20}\right)^{k+1}\left(1-\dfrac{9}{20}\right)^{49-k},\end{cases}$

解得 $\dfrac{439}{20}\leqslant k\leqslant\dfrac{459}{20}$.

因为 $k\in\mathbf{N}$，所以 $k=22$，即当 $k=22$ 时，$P(X=k)$ 取得最大值.

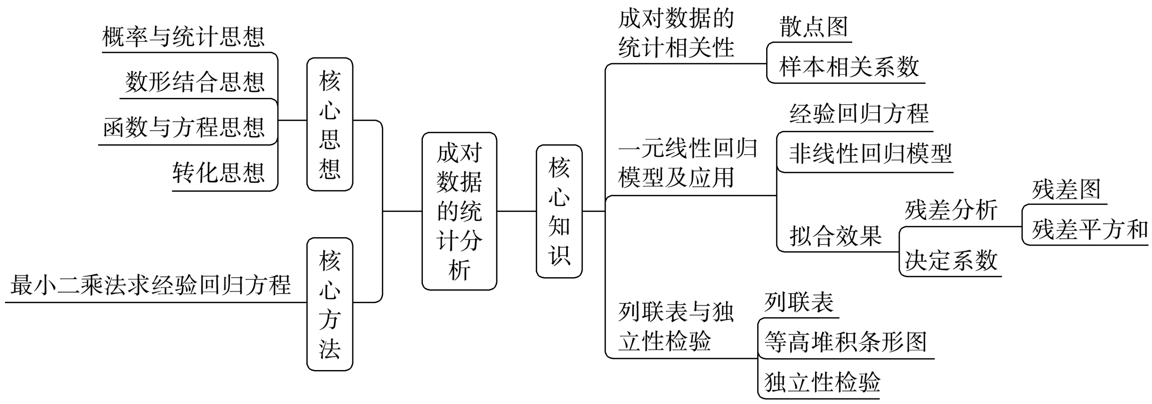
**【总结升华】**

**独立性检验与分布列的综合问题的解题策略**

解决独立性检验问题，要注意过好“三关”：假设关、公式关、对比关.解决概率问题要准确地把握题中所涉及的事件，明确所求问题对应的事件.

# 重构拓展

## 多维体系构建




**【问题探究】**

**1.**举例说明什么叫相关关系.相关关系与函数关系有什么区别?

**2.**在一元线性回归模型中,参数 $b$ 的含义是什么?一元线性回归模型有何作用?随机误差 $e$ 有哪些特征?

**3.**用随机变量 $\chi^2$ 取值的大小作为判断零假设 $H_0$ 是否成立的依据,当它比较大时推断 $H_0$ 不成立,否则认为 $H_0$ 成立.那么,究竟 $\chi^2$ 大到什么程度,可以推断 $H_0$ 不成立呢?或者说,怎样确定判断 $\chi^2$ 大小的标准呢?

**4.**根据有关规定,香烟盒上必须印上"吸烟有害健康"的警示语.

(1)吸烟是否一定会引发健康问题?

(2)有人认为吸烟不一定引发健康问题,因此可以吸烟.这种观点对吗?

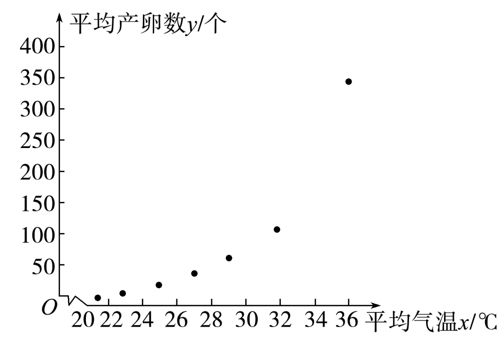
## 学科视野拓展

### 拓展一 非线性经验回归分析问题

**【拓展总结】**

对于非线性经验回归分析问题,可以先画出已知数据的散点图,再选择跟散点图拟合得最好的函数模型进行变量代换,最后作出变换后样本点的散点图,用线性回归模型拟合.

**应用1** 红蜘蛛是柚子树的主要害虫之一,能对柚子树造成严重伤害,每只红蜘蛛的平均产卵数 $y$(单位:个)和平均气温 $x$(单位:℃)有关,现收集了7组数据,得到下面的散点图.




(1)根据散点图,判断 $y=bx+a$ 与 $y=ce^{dx}$(其中 e=

2.718…为自然对数的底数)哪一个更适合作为平均产卵数 $y$ 关于平均气温 $x$ 的回归方程类型?(给出判断即可,不必说明理由)

(2)由(1)的判断结果及表中数据,求出 $y$ 关于 $x$ 的经验回归方程.

| 参考数据($z=\ln y$) | | | | | |
|---|---|---|---|---|---|
| $\sum\limits_{i=1}^{7}x_i^2$ | $\sum\limits_{i=1}^{7}x_iy_i$ | $\sum\limits_{i=1}^{7}x_iz_i$ | $\overline{x}$ | $\overline{y}$ | $\overline{z}$ |
| 5 215 | 17 713 | 714 | 27 | 81.3 | 3.6 |

(3)根据以往每年平均气温以及果园年产值的统计,得到以下数据:平均气温在 22 ℃以下的年数占 60%,对柚子产量影响不大,不需要采取防虫措施;平均气温在 22～28 ℃的年数占 30%,柚子产量会下降 20%;平均气温在 28 ℃以上的年数占 10%,柚子产量会下降 50%.为了更好地防治红蜘蛛虫害,农科所研发出各种防虫害措施供果农选择.

在每年价格不变,无虫害的情况下,某果园年产值为 200 万元.根据以上数据,以得到最高收益(收益=产值－防虫害费用)为目标,请为果农从以下三个方案中推荐最佳防虫害方案,并说明理由.

方案 1:可以防治各种气温的红蜘蛛虫害,费用是 18 万元;

方案 2:可以防治 22～28 ℃ 的红蜘蛛虫害,但无法防治 28 ℃以上的红蜘蛛虫害,费用是 10 万元;

方案 3:不采取防虫害措施.

附:经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$ 中,$\hat{b}=\dfrac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}=\dfrac{\sum\limits_{i=1}^{n}x_iy_i-n\,\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{n}x_i^2-n\,\overline{x}^2}$,$\hat{a}=\overline{y}-\hat{b}\,\overline{x}$.

**解**:(1)由题中散点图可以判断,$y=ce^{dx}$ 更适合作为平均产卵数 $y$ 关于平均气温 $x$ 的回归方程类型.

(2)由 $z=\ln y$,可得 $z=\ln y=\ln c+dx$,

由题中的数据可得,$\sum\limits_{i=1}^{7}x_iz_i-7\,\overline{x}\,\overline{z}=33.6$,

$\sum\limits_{i=1}^{7}(x_i-\overline{x})^2=\sum\limits_{i=1}^{7}x_i^2-7\overline{x}^2=112$,

所以 $\hat{d}=\dfrac{\sum\limits_{i=1}^{7}x_iz_i-7\,\overline{x}\,\overline{z}}{\sum\limits_{i=1}^{7}x_i^2-7\,\overline{x}^2}=\dfrac{33.6}{112}=0.3$,

则 $\ln\hat{c}=\overline{z}-\hat{d}\,\overline{x}=3.6-0.3\times27=-4.5$,

所以 $z$ 关于 $x$ 的经验回归方程为 $\hat{z}=0.3x-4.5$,

故 $y$ 关于 $x$ 的经验回归方程为 $\hat{y}=e^{0.3x-4.5}$.

(3)用 $X_1$,$X_2$ 和 $X_3$ 分别表示选择三种方案的收益.

选择方案 1,无论气温如何,产值不受影响,收益为 $200-18=182$ 万元,即 $X_1=182$.

选择方案 2,当不发生 28 ℃以上的红蜘蛛虫害时,收益为 $200-10=190$ 万元,

如果发生 28 ℃以上的红蜘蛛虫害,那么收益为 $100-10=90$ 万元,

即 $X_2=\begin{cases}190,\text{不发生 28 ℃以上的红蜘蛛虫害},\\90,\text{发生 28 ℃以上的红蜘蛛虫害}.\end{cases}$

同样,选择方案 3,有

$X_3=\begin{cases}200,\text{不发生虫害},\\160,\text{只发生 22～28℃虫害},\\100,\text{发生 28 ℃以上虫害},\end{cases}$

所以 $E(X_1)=182$,

$E(X_2)=190\times P(X_2=190)+90\times P(X_2=90)=190\times0.9+90\times0.1=171+9=180$,

$E(X_3)=200\times P(X_3=200)+160\times P(X_3=160)+100\times P(X_3=100)=200\times0.6+160\times0.3+100\times0.1=178$.

显然,$E(X_1)$最大,所以选择方案 1 最佳.

## 拓展二　统计案例中的新定义问题

**【拓展总结】**

常见类型:(1)相关性强弱分析中的新定义问题;(2)独立性检验中的新定义问题.

**应用 2**　某校 20 名学生的数学成绩 $x_i(i=1,2,\cdots,20)$ 和知识竞赛成绩 $y_i(i=1,2,\cdots,20)$ 如下表:

| 学生编号 $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 数学成绩 $x_i$ | 100 | 99 | 96 | 93 | 90 | 88 | 85 | 83 | 80 | 77 |
| 知识竞赛成绩 $y_i$ | 290 | 160 | 220 | 200 | 65 | 70 | 90 | 100 | 60 | 270 |
| 学生编号 $i$ | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 数学成绩 $x_i$ | 75 | 74 | 72 | 70 | 68 | 66 | 60 | 50 | 39 | 35 |
| 知识竞赛成绩 $y_i$ | 45 | 35 | 40 | 50 | 25 | 30 | 20 | 15 | 10 | 5 |

计算可得数学成绩的平均数是 $\overline{x}=75$,知识竞赛成绩的平均数是 $\overline{y}=90$,并且 $\sum\limits_{i=1}^{20}(x_i-\overline{x})^2=6\ 464$,$\sum\limits_{i=1}^{20}(y_i-\overline{y})^2=149\ 450$,$\sum\limits_{i=1}^{20}(x_i-\overline{x})\cdot(y_i-\overline{y})=21\ 650$.

(1)求这组学生的数学成绩和知识竞赛成绩的样本相关系数(精确到 0.01).

(2)设 $N\in\mathbf{N}^*$,变量 $x$ 的一组样本数据为 $x_1,x_2,\cdots,x_N$,变量 $y$ 的一组样本数据为 $y_1,y_2,\cdots,y_N$,其中 $x_i$

$(i=1,2,\cdots,N)$两两不相同，$y_i(i=1,2,\cdots,N)$两两不相同.按照从大到小的顺序，记$x_i$在样本数据中的排名是第$R_i$位，$y_i$在样本数据中的排名是第$S_i$位，$i=1,2,\cdots,N$.定义变量$x$和变量$y$的“斯皮尔曼相关系数”(记为$\rho$)为$x_i$的排名和$y_i$的排名的样本相关系数.

①记$d_i=R_i-S_i$，$i=1,2,\cdots,N$.证明：$\rho=1-\dfrac{6}{N(N^2-1)}\sum\limits_{i=1}^{N}d_i^2$；

②用①中的公式求得这组学生的数学成绩和知识竞赛成绩的“斯皮尔曼相关系数”约为0.91，简述“斯皮尔曼相关系数”在分析线性相关性时的优势.

附：$r=\dfrac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}}$；$\sum\limits_{k=1}^{n}k^2=\dfrac{n(n+1)(2n+1)}{6}$；$\sqrt{6\ 464\times149\ 450}\approx31\ 000$.

(1)**解**：由题意，这组学生的数学成绩和知识竞赛成绩的样本相关系数为

$$r=\frac{\sum\limits_{i=1}^{20}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{20}(x_i-\overline{x})^2\sum\limits_{i=1}^{20}(y_i-\overline{y})^2}}=\frac{21\ 650}{\sqrt{6\ 464\times149\ 450}}$$

$$\approx\frac{21\ 650}{31\ 000}\approx0.70.$$

(2)①**证明**：因为$\{R_i\}$和$\{S_i\}$都是$1,2,\cdots,N$的一个排列，所以$\sum\limits_{i=1}^{N}R_i=\sum\limits_{i=1}^{N}S_i=\dfrac{N(N+1)}{2}$，$\sum\limits_{i=1}^{N}R_i^2=\sum\limits_{i=1}^{N}S_i^2=\dfrac{N(N+1)(2N+1)}{6}$，

从而$\{R_i\}$和$\{S_i\}$的平均数都是$\overline{R}=\overline{S}=\dfrac{N+1}{2}$.

因此，$\sum\limits_{i=1}^{N}(R_i-\overline{R})^2=\sum\limits_{i=1}^{N}R_i^2-2\overline{R}\sum\limits_{i=1}^{N}R_i+N\overline{R}^2=\sum\limits_{i=1}^{N}R_i^2-N\overline{R}^2=\dfrac{N(N+1)(2N+1)}{6}-\dfrac{N(N+1)^2}{4}$

$=\dfrac{N(N+1)(N-1)}{12}$，

同理可得$\sum\limits_{i=1}^{N}(S_i-\overline{S})^2=\dfrac{N(N+1)(N-1)}{12}$.

由于$\sum\limits_{i=1}^{N}d_i^2=\sum\limits_{i=1}^{N}(R_i-S_i)^2$

$=\sum\limits_{i=1}^{N}[(R_i-\overline{R})-(S_i-\overline{S})]^2$

$=\sum\limits_{i=1}^{N}(R_i-\overline{R})^2+\sum\limits_{i=1}^{N}(S_i-\overline{S})^2-2\sum\limits_{i=1}^{N}(R_i-\overline{R})(S_i-\overline{S})=2\cdot\dfrac{N(N+1)(N-1)}{12}-2\sum\limits_{i=1}^{N}(R_i-\overline{R})(S_i-\overline{S})$，

所以$\rho=\dfrac{\sum\limits_{i=1}^{N}(R_i-\overline{R})(S_i-\overline{S})}{\sqrt{\sum\limits_{i=1}^{N}(R_i-\overline{R})^2}\sqrt{\sum\limits_{i=1}^{N}(S_i-\overline{S})^2}}$

$=\dfrac{\dfrac{N(N+1)(N-1)}{12}-\dfrac{1}{2}\sum\limits_{i=1}^{N}d_i^2}{\dfrac{N(N+1)(N-1)}{12}}$

$=1-\dfrac{6}{N(N^2-1)}\sum\limits_{i=1}^{N}d_i^2$.

②**解**：“斯皮尔曼相关系数”刻画的是样本数据排名的样本相关系数，与具体的数值无关，只与排名有关.如果一组数据有异常值，但排名依然符合一定的线性关系，那么可以采用“斯皮尔曼相关系数”刻画线性关系.

# 质量评估(三)

(范围:第八章)

| $\alpha$ | 0.1 | 0.05 | 0.01 | 0.005 | 0.001 |
|---|---|---|---|---|---|
| $x_{\alpha}$ | 2.706 | 3.841 | 6.635 | 7.879 | 10.828 |

## 一、单项选择题

1.对两个变量 $y$ 和 $x$ 进行回归分析,得到一组样本数据:$(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$,则下列说法不正确的是 (　　)

A.用决定系数 $R^2$ 来刻画拟合效果,$R^2$ 的值越小,说明模型的拟合效果越好

B.由样本数据得到的经验回归直线 $\hat{y}=\hat{b}x+\hat{a}$ 必过样本点的中心$(\overline{x},\overline{y})$

C.残差平方和越小的模型,拟合的效果越好

D.若变量 $y$ 和 $x$ 之间的样本相关系数 $r=-0.936\ 2$,则变量 $y$ 与 $x$ 之间具有较强的线性相关关系

A　**解析**:对于A,用决定系数 $R^2$ 来刻画拟合效果,$R^2$ 的值越接近1,说明模型的拟合效果越好,错误.对于B,由样本数据得到的经验回归直线 $\hat{y}=\hat{b}x+\hat{a}$ 必过样本点的中心$(\overline{x},\overline{y})$,正确.对于C,残差平方和越小的模型,拟合效果越好,正确.对于D,$|r|$越接近1,变量 $y$ 与 $x$ 之间的线性相关程度越强,正确.故选A.

2.已知呈现线性相关关系的变量 $x,y$ 之间的一组数据如下表所示,则 $y$ 关于 $x$ 的经验回归直线一定过点 (　　)

| $x$ | 0.1 | 0.2 | 0.3 | 0.5 |
|---|---|---|---|---|
| $y$ | 2.11 | 2.85 | 4.08 | 10.15 |

A.(0.1,2.11)　　B.(0.2,2.85)

C.(0.3,4.08)　　D.(0.275,4.797 5)

D　**解析**:经验回归直线一定过点$(\overline{x},\overline{y})$,通过题表中的数据计算得 $\overline{x}=0.275$,$\overline{y}=4.797\ 5$,易知选D.

3.小明利用课余时间参与科学探究活动——观察蒜苗的生长,下表记录了大蒜发芽后第4天至第8天的蒜苗高度.若用最小二乘法求得蒜苗高度 $y$(单位:cm)关于时间 $x$(单位:天)的经验回归方程为 $\hat{y}=\hat{b}x-4.4$,小明根据经验回归方程预测,从第 $n$ 天开始蒜苗高度大于20 cm,则 $n$ 的值为 (　　)

| 时间 $x$/天 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|
| 蒜苗高度 $y$/cm | 1 | 2.4 | 4.6 | 5.6 | 6.4 |

A.15　　B.16　　C.17　　D.18

D　**解析**:由题表中数据得 $\overline{x}=6$,$\overline{y}=4$,代入方程 $\overline{y}=\hat{b}\,\overline{x}-4.4$,解得 $\hat{b}=1.4$,

则经验回归方程为 $\hat{y}=1.4x-4.4$.

令 $1.4x-4.4>20$,解得 $x>17.4$.因为 $x\in\mathbf{N}^*$,所以 $x\geqslant18$.

故选D.

4.若 $y$ 关于 $x$ 的经验回归方程为 $\hat{y}=2-3.5x$,则变量 $x$ 增加1个单位,变量 $y$ 平均 (　　)

A.减少3.5个单位　　B.增加2个单位

C.增加3.5个单位　　D.减少2个单位

A　**解析**:由经验回归方程可知 $\hat{b}=-3.5$,则变量 $x$ 增加1个单位,变量 $y$ 平均减少3.5个单位.

5.下表是某厂1～4月份用水量(单位:百吨)的一组数据:

| 月份 $x$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 用水量 $y$ | 4.5 | 4 | 3 | 2.5 |

已知用水量 $y$ 与月份 $x$ 之间有较强的线性相关关系,其经验回归方程是 $\hat{y}=-0.7x+\hat{a}$,则 $\hat{a}$ 等于 (　　)

A.10.5　　B.5.15　　C.5.2　　D.5.25

D　**解析**:因为 $\overline{x}=\dfrac{1+2+3+4}{4}=2.5$,$\overline{y}=\dfrac{4.5+4+3+2.5}{4}=3.5$,所以 $3.5=-0.7\times2.5+\hat{a}$,解得 $\hat{a}=5.25$.

6.针对中学生追星问题,某校团委对“中学生性别和中学生追星是否有关”做了一次调查,调查样本中女生人数是男生人数的 $\dfrac{1}{2}$,男生追星的人数占男生人数的 $\dfrac{1}{6}$,女生追星的人数占女生人数的 $\dfrac{2}{3}$.若依据小概率值 $\alpha=0.05$ 的独立性检验,可以认为中学生性别和中学生追星有关,则调查样本中男生至少有 (　　)

附:$\chi^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$.

A.12人　　B.11人　　C.10人　　D.18人

A　**解析**:设男生人数为 $x$,依题意可得如下2×2列联表:

单位：人

| 性别 | 追星 | | 合计 |
|---|---|---|---|
| | 是 | 否 | |
| 男 | $\frac{x}{6}$ | $\frac{5x}{6}$ | $x$ |
| 女 | $\frac{x}{3}$ | $\frac{x}{6}$ | $\frac{x}{2}$ |
| 合计 | $\frac{x}{2}$ | $x$ | $\frac{3x}{2}$ |

若依据小概率值 $\alpha=0.05$ 的独立性检验，可以认为中学生性别和中学生追星有关，则 $\chi^2\geqslant3.841$.由 $\chi^2$

$=\frac{\frac{3x}{2}\left(\frac{x^2}{36}-\frac{5x^2}{18}\right)^2}{x\cdot\frac{x}{2}\cdot\frac{x}{2}\cdot x}=\frac{3}{8}x\geqslant3.841$ 及 $\frac{x}{2},\frac{x}{6}$ 为整数，可知 $x$ 最小可取 12，所以男生至少有 12 人.故选 A.

**7.**某种微生物的繁殖速度 $y$ 与生长环境中的营养物质浓度 $x$ 相关，在一定条件下可用模型 $y=2\lg x$进行拟合.在这个条件下，要使 $y$ 增加 2 个单位，则应该（　　）

A.使 $x$ 增加 1 个单位

B.使 $x$ 增加 2 个单位

C.使 $x$ 增加到原来的 2 倍

D.使 $x$ 增加到原来的 10 倍

D　**解析**：设 $y$ 的增加量为 $\Delta y=y_1-y_2$，$x$ 的增加量为$\Delta x=x_1-x_2$，

故可得 $\Delta y=2\lg x_1-2\lg x_2=2\lg\frac{x_1}{x_2}=2$，解得$\frac{x_1}{x_2}=10$.

故要使得 $y$ 增加 2 个单位，$x$ 应增加到原来的 10 倍.故选 D.

**8.**关于变量 $x,y$，下列说法正确的有（　　）

①若样本相关系数 $r>0$，则 $x$ 增大时，$y$ 也相应增大；

②若样本相关系数 $r<0$，则 $x$ 增大时，$y$ 也相应增大；

③若样本相关系数 $r=1$ 或 $-1$，则 $x$ 与 $y$ 的关系完全对应（函数关系），散点图中的各点均在一条直线上.

A.①②　　B.②③　　C.①③　　D.①②③

C　**解析**：若 $r>0$，则两个变量正相关，$x$ 增大时，$y$ 也相应增大，故①正确.若 $r<0$，则两个变量负相关，$x$ 增大时，$y$ 相应减小，故②错误.$|r|$越接近 1，表示两个变量相关性越高，$|r|=1$ 表示两个变量有确定的关系（即函数关系），故③正确.

**二、多项选择题**

**9.**已知在最小二乘法原理下，求得 $y$ 关于 $x$ 的经验回归方程为 $\hat{y}=-0.7x+10.3$，且变量 $x,y$ 之间的相关数据如表所示，则下列说法错误的是（　　）

| $x$ | 6 | 8 | 10 | 12 |
|---|---|---|---|---|
| $y$ | 6 | $m$ | 3 | 2 |

A.变量 $x,y$ 之间呈正相关关系

B.可以预测，当 $x=20$ 时，$y=3.7$

C.$m=4.7$

D.经验回归直线 $\hat{y}=-0.7x+10.3$ 必过点(9,4)

ABC　**解析**：对于 A，由 $x$ 与 $y$ 的经验回归方程，可知 $\hat{b}=-0.7<0$，

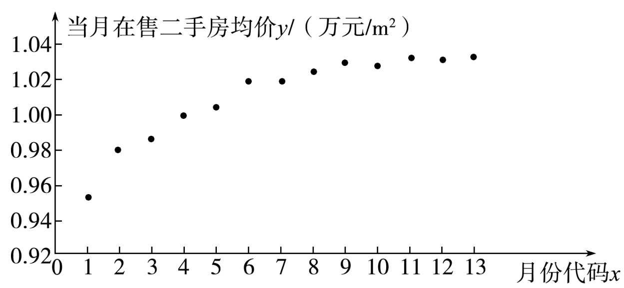
所以变量 $x,y$ 之间呈负相关关系，故 A 错误；

对于 B，当 $x=20$ 时，$\hat{y}=-0.7\times20+10.3=-3.7$，故 B 错误；

对于 C，由题表中数据可知$\overline{x}=9$，$\overline{y}=\frac{6+m+3+2}{4}=\frac{11+m}{4}$，由点$(\overline{x},\overline{y})$必在经验回归直线上，得$\frac{11+m}{4}=-0.7\times9+10.3$，解得 $m=5$，故 C 错误；

对于 D，因为 $m=5$，所以$\overline{y}=\frac{11+m}{4}=4$，所以经验回归直线必过点(9,4)，故 D 正确.故选 ABC.

**10.**下图是某小区 2023 年 8 月至 2024 年 8 月间，当月在售二手房均价（单位：万元/m²）的散点图.（图中月份代码 1～13 分别对应 2023 年 8 月～2024 年 8 月）




根据散点图选择 $y=a+b\sqrt{x}$ 和 $y=c+d\ln x$ 两个模型进行拟合，经过数据处理得到的两个经验回归方程分别为 $\hat{y}=0.936\ 9+0.028\ 5\sqrt{x}$ 和 $\hat{y}=0.955\ 4+0.030\ 6\ln x$，并得到统计量 $R^2$ 的值：

| 经验回归方程 | $\hat{y}=0.936\ 9+0.028\ 5\sqrt{x}$ | $\hat{y}=0.955\ 4+0.030\ 6\ln x$ |
|---|---|---|
| $R^2$ | 0.923 | 0.973 |

则下列说法正确的是（　　）

A.当月在售二手房均价 $y$ 与月份代码 $x$ 呈负相关关系

B.由 $\hat{y}=0.936\ 9+0.028\ 5\sqrt{x}$ 预测 2027 年 8 月在售二手房均价约为 1.136 4 万元/m²

C.曲线 $\hat{y}=0.936\ 9+0.028\ 5\sqrt{x}$ 与 $\hat{y}=0.955\ 4+$

$0.030\ 6\ln x$ 都经过点$(\overline{x},\overline{y})$

D.模型 $\hat{y}=0.955\ 4+0.030\ 6\ln x$ 的拟合效果比模型 $\hat{y}=0.936\ 9+0.028\ 5\sqrt{x}$ 的拟合效果好

BD　**解析**:对于A,由题图可知散点从左下到右上分布,所以当月在售二手房均价 $y$ 与月份代码 $x$ 呈正相关关系,故A不正确;对于B,令 $x=49$,得 $\hat{y}=0.936\ 9+0.028\ 5\times\sqrt{49}=1.136\ 4$,所以可以预测2027年8月在售二手房均价约为1.136 4万元/$m^2$,故B正确;对于C,非线性回归曲线不一定经过点$(\overline{x},\overline{y})$,故C错误;对于D,$R^2$ 越大,模型拟合效果越好,$0.923<0.973$,故D正确.

**11.**某市通过随机询问100名居民能否做到"光盘"行动,得到如下的 $2\times2$ 列联表,并进行独立性检验,则　　(　　)

单位:名

| 性别 | 光盘 | | 合计 |
|---|---|---|---|
| | 不能做到 | 能做到 | |
| 男 | 45 | $b$ | 55 |
| 女 | $c$ | 15 | |
| 合计 | | | 100 |

A.$c=30$

B.零假设为 $H_0$:能否做到"光盘"行动与性别有关

C.$\chi^2>3.841$

D.依据小概率值 $\alpha=0.1$ 的独立性检验,可以认为该市居民能否做到"光盘"行动与性别有关

AD　**解析**:由题可得 $a=45,b=10,c=30,d=15$,则 $a+b=55,c+d=45,a+c=75,b+d=25,ad=675,bc=300,n=100$.

零假设为 $H_0$:能否做到"光盘"行动与性别无关.

计算可得 $\chi^2=\dfrac{100\times(675-300)^2}{55\times45\times75\times25}\approx3.030$.

因为 $x_{0.1}=2.706<3.030$,

所以依据小概率 $\alpha=0.1$ 的独立性检验,可以认为该市居民能否做到"光盘"行动与性别有关,此推断犯错误的概率不大于0.1.

## 三、填空题

**12.**对于经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$,当 $x=3$ 时,对应的 $y$ 的估计值是17,当 $x=8$ 时,对应的 $y$ 的估计值是22,那么,该经验回归方程为________;根据经验回归方程,当 $x=$______时,$y$ 的估计值是38.

$\hat{y}=x+14$　24　**解析**:把$(3,17)$,$(8,22)$代入经验回归方程得 $\begin{cases}3\hat{b}+\hat{a}=17,\\8\hat{b}+\hat{a}=22,\end{cases}$ 解得 $\begin{cases}\hat{b}=1,\\\hat{a}=14.\end{cases}$

所以经验回归方程为 $\hat{y}=x+14$.

令 $x+14=38$,得 $x=24$.

**13.**某旅行社为调查市民对人文景观的态度(喜欢或不喜欢)是否与年龄有关,随机调查了55名市民,所得数据如下表所示.

单位:名

| 年龄 | 态度 | | 合计 |
|---|---|---|---|
| | 喜欢 | 不喜欢 | |
| 大于40岁 | 20 | 5 | 25 |
| 20岁至40岁 | 10 | 20 | 30 |
| 合计 | 30 | 25 | 55 |

根据小概率值 $\alpha=0.005$ 的独立性检验,______推断出市民对人文景观的态度与年龄有关.(填"能"或"不能")

能　**解析**:零假设为 $H_0$:市民对人文景观的态度与年龄无关.由题表中数据计算可得 $\chi^2=\dfrac{55\times(20\times20-5\times10)^2}{25\times30\times30\times25}\approx11.978>7.879=x_{0.005}$,根据小概率值 $\alpha=0.005$ 的独立性检验,我们推断 $H_0$ 不成立,即认为市民对人文景观的态度与年龄有关,此推断犯错误的概率不大于0.005.

**14.**对于数据组$(x_i,y_i)(i=1,2,\cdots,n)$,如果由经验回归方程得到的对应自变量 $x_i$ 的估计值是 $\hat{y}_i$,那么将 $y_i-\hat{y}_i$ 称为对应点$(x_i,y_i)$的残差.某商场为了给一种新商品进行合理定价,将该商品按事先拟定的价格进行试销,得到如下所示数据.

| 单价 $x$/元 | 8.2 | 8.4 | 8.6 | 8.8 |
|---|---|---|---|---|
| 销量 $y$/件 | 84 | 83 | 78 | $m$ |

根据表中的数据,得到销量 $y$(单位:件)与单价 $x$(单位:元)之间的经验回归方程为 $\hat{y}=-20x+\hat{a}$,据计算,样本点$(8.4,83)$处的残差为1,则 $m=$________.

75　**解析**:根据样本点$(8.4,83)$处的残差为1,得 $83-(-20\times8.4+\hat{a})=1$,得 $\hat{a}=250$,

所以 $\hat{y}=-20x+250$.

$\overline{x}=\dfrac{8.2+8.4+8.6+8.8}{4}=8.5$,$\overline{y}=\dfrac{84+83+78+m}{4}=\dfrac{245+m}{4}$,由 $\overline{y}=-20\overline{x}+250$,得 $m=75$.

## 四、解答题

**15.**偏差是指个别测定值与测定的平均值之差.在成绩统计中,我们把某个同学的某科考试成绩与该科平均分的差叫某科偏差(实际成绩－平均分＝偏差).在某次考试成绩统计中,某老师为了对学生数学偏

差 $x$（单位：分）与物理偏差 $y$（单位：分）之间的关系进行分析，随机挑选了 8 位同学，得到他们的两科成绩偏差数据如下表所示.

| 学生序号 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 数学偏差 $x$ | 20 | 15 | 13 | 3 | 2 | $-5$ | $-10$ | $-18$ |
| 物理偏差 $y$ | 6.5 | 3.5 | 3.5 | 1.5 | 0.5 | $-0.5$ | $-2.5$ | $-3.5$ |

(1)若 $x$ 与 $y$ 之间具有线性相关关系，求 $y$ 关于 $x$ 的经验回归方程；

(2)若该次考试数学平均分为 120 分，物理平均分为 91.5 分，试由(1)的结论预测数学成绩为 128 分的同学的物理成绩.

附：$\sum\limits_{i=1}^{8}x_iy_i=324$，$\sum\limits_{i=1}^{8}x_i^2=1\ 256$，

经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$ 中，$\hat{b}=\dfrac{\sum\limits_{i=1}^{n}x_iy_i-n\,\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{n}x_i^2-n\,\overline{x}^2}$，$\hat{a}=\overline{y}-\hat{b}\,\overline{x}$.

**解**：(1)由题意可得，

$\overline{x}=\dfrac{1}{8}\times[20+15+13+3+2+(-5)+(-10)+(-18)]=\dfrac{5}{2}$，

$\overline{y}=\dfrac{1}{8}\times[6.5+3.5+3.5+1.5+0.5+(-0.5)+(-2.5)+(-3.5)]=\dfrac{9}{8}$，

$\hat{b}=\dfrac{\sum\limits_{i=1}^{8}x_iy_i-8\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{8}x_i^2-8\,\overline{x}^2}=\dfrac{324-8\times\dfrac{5}{2}\times\dfrac{9}{8}}{1\ 256-8\times\left(\dfrac{5}{2}\right)^2}=\dfrac{1}{4}$，

所以 $\hat{a}=\overline{y}-\hat{b}\overline{x}=\dfrac{9}{8}-\dfrac{1}{4}\times\dfrac{5}{2}=\dfrac{1}{2}$.

故经验回归方程为 $\hat{y}=\dfrac{1}{4}x+\dfrac{1}{2}$.

(2)由题意，设该同学的物理成绩为 $\omega$，则物理偏差为 $\omega-91.5$.

而数学偏差为 $128-120=8$，

所以 $\omega-91.5=\dfrac{1}{4}\times8+\dfrac{1}{2}$，解得 $\omega=94$.

所以，可以预测这位同学的物理成绩为 94 分.

**16.** 长跑对于培养人们克服困难、磨炼刻苦耐劳的顽强意志具有良好的作用.某校开展冬季长跑活动，为了解学生对冬季长跑活动是否感兴趣与性别是否有关，某调查小组随机抽取该校 100 名学生进行问卷调查，根据所得数据制成如下 2×2 列联表：

单位：人

| 性别 | 冬季长跑活动 | | 合计 |
|---|---|---|---|
| | 感兴趣 | 不感兴趣 | |
| 男 | | 8 | |
| 女 | 32 | | |
| 合计 | 80 | | 100 |

(1)完成上面的 2×2 列联表，依据小概率值 $\alpha=0.1$ 的独立性检验，能否认为学生对冬季长跑活动是否感兴趣与性别有关联？

(2)若不感兴趣的男生中恰有 3 名高三学生，现从不感兴趣的男生中随机选出 3 名学生进行二次调查，记选出的学生中高三学生的人数为 $X$，求 $X$ 的分布列与数学期望.

附：$\chi^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$，其中 $n=a+b+c+d$.

**解**：(1)根据已知数据可补全 2×2 列联表如下：

单位：人

| 性别 | 冬季长跑活动 | | 合计 |
|---|---|---|---|
| | 感兴趣 | 不感兴趣 | |
| 男 | 48 | 8 | 56 |
| 女 | 32 | 12 | 44 |
| 合计 | 80 | 20 | 100 |

零假设为 $H_0$：学生对冬季长跑活动是否感兴趣与性别无关.

由表中数据计算得 $\chi^2=\dfrac{100\times(48\times12-8\times32)^2}{56\times44\times80\times20}\approx2.597<2.706=x_{0.1}$，

所以依据小概率值 $\alpha=0.1$ 的独立性检验，没有充分证据推断 $H_0$ 不成立，即可以认为学生对冬季长跑活动是否感兴趣与性别无关.

(2)由题知 $X$ 的所有可能取值为 0，1，2，3，

因为 $P(X=0)=\dfrac{C_5^3}{C_8^3}=\dfrac{10}{56}=\dfrac{5}{28}$，$P(X=1)=\dfrac{C_5^2C_3^1}{C_8^3}=\dfrac{30}{56}=\dfrac{15}{28}$，$P(X=2)=\dfrac{C_5^1C_3^2}{C_8^3}=\dfrac{15}{56}$，$P(X=3)=\dfrac{C_3^3}{C_8^3}=\dfrac{1}{56}$，所以 $X$ 的分布列为

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P$ | $\dfrac{5}{28}$ | $\dfrac{15}{28}$ | $\dfrac{15}{56}$ | $\dfrac{1}{56}$ |

$E(X)=0\times\frac{5}{28}+1\times\frac{15}{28}+2\times\frac{15}{56}+3\times\frac{1}{56}=\frac{9}{8}$.

**17.**(2024·全国甲卷)某工厂进行生产线智能化升级改造,升级改造后,从该工厂甲、乙两个车间的产品中随机抽取150件进行检验,数据如下:

单位:件

| 车间 | 品级 | | | 合计 |
|---|---|---|---|---|
| | 优级品 | 合格品 | 不合格品 | |
| 甲车间 | 26 | 24 | 0 | 50 |
| 乙车间 | 70 | 28 | 2 | 100 |
| 合计 | 96 | 52 | 2 | 150 |

(1)填写如下列联表:

单位:件

| 车间 | 品级 | |
|---|---|---|
| | 优级品 | 非优级品 |
| 甲车间 | | |
| 乙车间 | | |

依据小概率值 $\alpha=0.05$ 的独立性检验,能否认为甲、乙两车间产品的优级品率存在差异?依据小概率值 $\alpha=0.01$ 的独立性检验,能否认为甲、乙两车间产品的优级品率存在差异?

(2)已知升级改造前该工厂产品的优级品率 $p=0.5$,设 $\overline{p}$ 为升级改造后抽取的 $n$ 件产品的优级品率.如果 $\overline{p}>p+1.65\sqrt{\frac{p(1-p)}{n}}$,那么认为该工厂产品的优级品率提高了.根据抽取的150件产品的数据,能否认为生产线智能化升级改造后,该工厂产品的优级品率提高了?($\sqrt{150}\approx12.247$)

附:$\chi^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$.

| $\alpha$ | 0.050 | 0.010 | 0.001 |
|---|---|---|---|
| $x_\alpha$ | 3.841 | 6.635 | 10.828 |

**解**:(1)根据题意可得 $2\times2$ 列联表:

单位:件

| 车间 | 品级 | |
|---|---|---|
| | 优级品 | 非优级品 |
| 甲车间 | 26 | 24 |
| 乙车间 | 70 | 30 |

零假设为 $H_0$:甲、乙两车间产品的优级品率无差异.

可得 $\chi^2=\frac{150\times(26\times30-24\times70)^2}{50\times100\times96\times54}=4.6875$.

因为 $x_{0.05}=3.841<4.6875<6.635=x_{0.01}$,

所以依据小概率值 $\alpha=0.05$ 的独立性检验,推断 $H_0$ 不成立,即可以认为甲、乙两车间产品的优级品率存在差异,此推断犯错误的概率不超过0.05.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,即认为甲、乙两车间产品的优级品率无差异.

(2)由题意可知,生产线智能化升级改造后,该工厂产品的优级品的频率为 $\frac{96}{150}=0.64$,

用频率估计概率可得 $\overline{p}=0.64$.

又因为升级改造前该工厂产品的优级品率 $p=0.5$,

则 $p+1.65\sqrt{\frac{p(1-p)}{n}}=0.5+1.65\times\sqrt{\frac{0.5\times(1-0.5)}{150}}\approx0.5+1.65\times\frac{0.5}{12.247}\approx0.567$,

可知 $\overline{p}>p+1.65\sqrt{\frac{p(1-p)}{n}}$,

所以可以认为生产线智能化升级改造后,该工厂产品的优级品率提高了.

**18.**某田径协会组织开展竞走的步长和步频之间的关系的课题研究,得到相应的试验数据:

| 步频 $x$/s | 0.28 | 0.29 | 0.30 | 0.31 | 0.32 |
|---|---|---|---|---|---|
| 步长 $y$/cm | 90 | 95 | 99 | 103 | 117 |

(1)根据表中数据,得到步频和步长近似为线性相关关系,求出 $y$ 关于 $x$ 的经验回归方程,并利用经验回归方程预测,当步长为80 cm时,步频约是多少?

(2)记 $\hat{e}_i=y_i-\hat{y}_i=y_i-\hat{b}x_i-\hat{a}$,其中 $y_i$ 为观测值,$\hat{y}_i$ 为预测值,$\hat{e}_i$ 为对应 $(x_i,y_i)$ 的残差,求(1)中步长的残差的和,并探究这个结果是否对任意具有线性相关关系的两个变量都成立?若成立,请证明;若不成立,请说明理由.

参考数据:$\sum\limits_{i=1}^{5}x_i^2=0.451$,$\sum\limits_{i=1}^{5}x_iy_i=151.82$.

参考公式:$\hat{b}=\frac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}$,$\hat{a}=\overline{y}-\hat{b}\,\overline{x}$.

**解**:(1)$\overline{x}=\frac{1}{5}\sum\limits_{i=1}^{5}x_i=0.3$,$\overline{y}=\frac{1}{5}\sum\limits_{i=1}^{5}y_i=100.8$,

$\hat{b}=\frac{\sum\limits_{i=1}^{5}x_iy_i-5\times\overline{x}\times\overline{y}}{\sum\limits_{i=1}^{5}x_i^2-5\overline{x}^2}=620$,$\hat{a}=100.8-620\times0.3=-$

85.2，

所以经验回归方程为 $\hat{y}=620x-85.2$.

将 $y=80$ 代入得 $80=620x-85.2$，解得 $x\approx0.27$，所以当步长为 80 cm 时，步频约是 0.27 s.

(2)由(1)可知 $\hat{y}_1=620\times0.28-85.2=88.4$，$\hat{e}_1=90-88.4=1.6$；

$\hat{y}_2=620\times0.29-85.2=94.6$，$\hat{e}_2=95-94.6=0.4$；

$\hat{y}_3=620\times0.30-85.2=100.8$，$\hat{e}_3=99-100.8=-1.8$；

$\hat{y}_4=620\times0.31-85.2=107$，$\hat{e}_4=103-107=-4$；

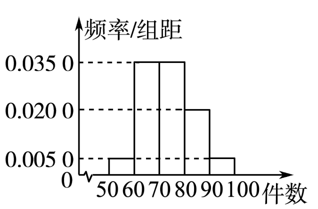
$\hat{y}_5=620\times0.32-85.2=113.2$，$\hat{e}_5=117-113.2=3.8$，

所以 $\sum\limits_{i=1}^{5}\hat{e}_i=1.6+0.4-1.8-4+3.8=0$，即步长残差和为 0.

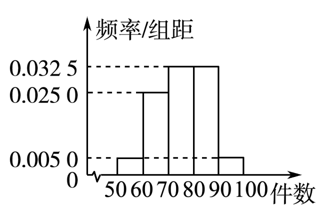
对任意具有线性相关关系的两个变量都成立，证明如下：

$$\sum_{i=1}^{n}\hat{e}_i=\sum_{i=1}^{n}(y_i-\hat{y}_i)=\sum_{i=1}^{n}(y_i-\hat{b}x_i-\hat{a})=\sum_{i=1}^{n}y_i-\hat{b}\sum_{i=1}^{n}x_i-n\hat{a}=n\overline{y}-n\hat{b}\overline{x}-n(\overline{y}-\hat{b}\overline{x})=0.$$

**19.** 某工厂有 25 周岁以上(含 25 周岁)工人 300 名，25 周岁以下工人 200 名. 为研究工人的日平均生产量是否与年龄有关，现采用分层随机抽样的方法，从中抽取了 100 名工人，先统计了他们某月的日平均生产件数，然后按工人年龄分为“25 周岁以上(含 25 周岁)”和“25 周岁以下”两组，再将两组工人的日平均生产件数分为 5 组：[50，60)，[60，70)，[70，80)，[80，90)，[90，100]分别加以统计，得到如图所示的频率分布直方图.




25周岁以上(含25周岁)组




25周岁以下组

(1)从样本中日平均生产件数不足 60 件的工人中随机抽取 2 名，求至少抽到 1 名“25 周岁以下组”工人的概率；

(2)规定日平均生产件数不少于 80 件者为“生产能手”，请你根据已知条件列出 2×2 列联表，并依据小概率值 $\alpha=0.1$ 的独立性检验，分析生产能手与工人所在的年龄组是否有关.

**解**：(1)由已知得，样本中有“25 周岁以上(含 25 周岁)”组工人 60 名，“25 周岁以下”组工人 40 名，

所以样本中日平均生产件数不足 60 件的工人中，“25 周岁以上(含 25 周岁)”组工人有 $60\times0.05=3$ 名，

记为 $A_1,A_2,A_3$；

“25 周岁以下”组工人有 $40\times0.05=2$ 名，

记为 $B_1,B_2$.

从中随机抽取 2 名工人，所有的可能结果共有 10 种：

$(A_1,A_2)$，$(A_1,A_3)$，$(A_2,A_3)$，$(A_1,B_1)$，$(A_1,B_2)$，$(A_2,B_1)$，$(A_2,B_2)$，$(A_3,B_1)$，$(A_3,B_2)$，$(B_1,B_2)$.

其中，至少有 1 名“25 周岁以下组”工人的可能结果共有 7 种：$(A_1,B_1)$，$(A_1,B_2)$，$(A_2,B_1)$，$(A_2,B_2)$，$(A_3,B_1)$，$(A_3,B_2)$，$(B_1,B_2)$.

故所求的概率 $p=\dfrac{7}{10}$.

(2)由题图可知，在抽取的 100 名工人中，

“25 周岁以上(含 25 周岁)”组中的生产能手有 $60\times(0.2+0.05)=15$ 名，

“25 周岁以下”组中的生产能手有 $40\times(0.325+0.05)=15$ 名.

据此可得 2×2 列联表如下：

单位：名

| 年龄组 | 生产能手情况 | | 合计 |
|---|---|---|---|
| | 生产能手 | 非生产能手 | |
| 25 周岁以上(含 25 周岁)组 | 15 | 45 | 60 |
| 25 周岁以下组 | 15 | 25 | 40 |
| 合计 | 30 | 70 | 100 |

零假设为 $H_0$：生产能手与工人所在的年龄组无关.

由表中数据得 $\chi^2=\dfrac{100\times(15\times25-45\times15)^2}{60\times40\times30\times70}=\dfrac{25}{14}\approx1.79<2.706=x_{0.1}$.

依据小概率值 $\alpha=0.1$ 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，即生产能手与工人所在的年龄组无关.

# 综合质量评估

（范围：选择性必修　第三册）

## 一、单项选择题

1.若 $3C_{2n}^{3}=5A_{n}^{3}$，则正整数 $n=$　（　　）

A.7　B.8　C.9　D.10

B　**解析**：因为 $3C_{2n}^{3}=5A_{n}^{3}$，

所以 $3\times\frac{2n(2n-1)(2n-2)}{3\times2\times1}=5n(n-1)(n-2)$，

解得 $n=8$（$n=1$ 不合题意，舍去）.

2.下列说法错误的是　（　　）

A.决定系数 $R^2$ 越大，模型的拟合效果越好

B.若变量 $x$ 和 $y$ 之间的样本相关系数 $r=-0.982$，则变量 $x$ 和 $y$ 之间的负相关程度很强

C.残差平方和越小的模型，拟合效果越好

D.对于经验回归方程 $\hat{y}=-3x+0.8$，当解释变量 $x$ 每增加 1 个单位时，响应变量 $y$ 平均增加 3 个单位

D　**解析**：用决定系数 $R^2$ 来刻画回归效果，$R^2$ 越大，表示残差平方和越小，即模型的拟合效果越好，故 A 正确；

若变量 $x$ 和 $y$ 之间的样本相关系数 $r=-0.982$，$r$ 接近 $-1$，则变量 $x$ 和 $y$ 之间的负相关程度很强，故 B 正确；

比较两个模型的拟合效果，可以比较残差平方和的大小，残差平方和越小的模型，拟合效果越好，故 C 正确；

对于经验回归方程 $\hat{y}=-3x+0.8$，当解释变量 $x$ 每增加 1 个单位时，响应变量 $y$ 平均减少 3 个单位，故 D 错误.

故选 D.

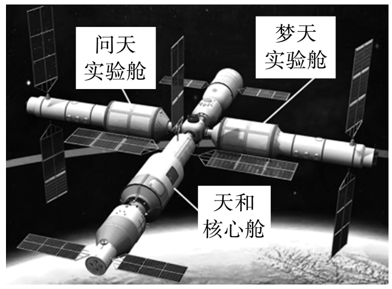
3.某种种子每粒发芽的概率都为 0.9，现播种了 1 000 粒，对于没有发芽的种子，每粒需再补种 2 粒，补种的种子数记为 $X$，则 $X$ 的数学期望为　（　　）

A.100　B.200　C.300　D.400

B　**解析**：记不发芽的种子数为 $\xi$，则 $\xi\sim B(1\ 000,0.1)$，所以 $E(\xi)=1\ 000\times0.1=100$.

又因为 $X=2\xi$，所以 $E(X)=E(2\xi)=2E(\xi)=200$.

4.如图，中国空间站的主体结构包括天和核心舱、问天实验舱和梦天实验舱.假设中国空间站要安排甲、乙、丙、丁、戊 5 名航天员开展实验，其中天和核心舱安排 3 人，问天实验舱与梦天实验舱各安排 1 人.若甲、乙两人不能同时在一个舱内做实验，则不同的安排方案共有　（　　）




A.8 种　B.14 种　C.20 种　D.16 种

B　**解析**：第一类，甲、乙都不在天和核心舱的安排方案共有 $A_2^2=2$ 种；

第二类，甲、乙恰好有一人在天和核心舱，先安排天和核心舱有 $C_2^1C_3^2=6$ 种安排方案，

然后安排问天实验舱与梦天实验舱有 $A_2^2=2$ 种安排方案，

所以，甲、乙恰好有一人在天和核心舱的安排方案共有 $6\times2=12$ 种.

综上，甲、乙两人不能同时在一个舱内做实验的安排方案共有 $2+12=14$ 种.

5.若实数 $a=2-\sqrt{2}$，则 $a^{10}-2C_{10}^{1}a^{9}+2^{2}C_{10}^{2}a^{8}-\cdots+2^{10}$ 等于　（　　）

A.32　B.$-32$　C.1 024　D.512

A　**解析**：由二项式定理，得 $a^{10}-2C_{10}^{1}a^{9}+2^{2}C_{10}^{2}a^{8}-\cdots+2^{10}=C_{10}^{0}(-2)^{0}a^{10}+C_{10}^{1}(-2)^{1}a^{9}+C_{10}^{2}(-2)^{2}a^{8}+\cdots+C_{10}^{10}(-2)^{10}=(a-2)^{10}=(-\sqrt{2})^{10}=2^{5}=32$.

6.4 个高尔夫球中有 3 个合格、1 个不合格，每次任取 1 个，不放回地取两次.若第一次取到合格的高尔夫球，则第二次取到合格高尔夫球的概率为　（　　）

A.$\frac{1}{2}$　B.$\frac{2}{3}$　C.$\frac{3}{4}$　D.$\frac{4}{5}$

B　**解析**：（方法一）记事件 $A=$“第一次取到的是合格高尔夫球”，事件 $B=$“第二次取到的是合格高尔夫球”.

由题意可得 $P(AB)=\frac{3\times2}{4\times3}=\frac{1}{2}$，$P(A)=\frac{3}{4}$，

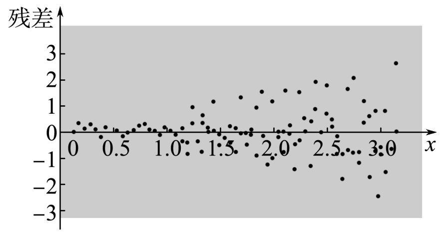
所以 $P(B|A)=\frac{P(AB)}{P(A)}=\frac{\frac{1}{2}}{\frac{3}{4}}=\frac{2}{3}$.

（方法二）记事件 $A=$“第一次取到的是合格高尔夫

球”，事件 $B$＝“第二次取到的是合格高尔夫球”.

由题意可得，事件 $A\cap B$ 所包含的样本点数 $n(A\cap B)=3\times2=6$，事件 $A$ 所包含的样本点数 $n(A)=3\times3=9$，所以 $P(B|A)=\dfrac{n(A\cap B)}{n(A)}=\dfrac{6}{9}=\dfrac{2}{3}$.

**7.** 对于变量 $Y$ 和变量 $x$ 的成对样本观测数据，利用一元线性回归模型 $\begin{cases}Y=bx+a+e,\\E(e)=0,D(e)=\sigma^2\end{cases}$ 得到经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$，对应的残差如图所示，则模型中的随机误差（　　）




A.满足一元线性回归模型的所有假设

B.不满足一元线性回归模型的 $E(e)=0$ 的假设

C.不满足一元线性回归模型的 $D(e)=\sigma^2$ 的假设

D.不满足一元线性回归模型的 $E(e)=0$ 和 $D(e)=\sigma^2$ 的假设

C　**解析**：利用一元线性回归模型 $\begin{cases}Y=bx+a+e,\\E(e)=0,D(e)=\sigma^2\end{cases}$ 得到经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$，根据题图，易知残差的方差不是一个常数，随变量 $x$ 的变大而变大，不满足一元线性回归模型的 $D(e)=\sigma^2$ 的假设.故选 C.

**8.** 在篮球比赛颁奖仪式上，某队队员 12 人（其中 1 人为队长）、教练组 3 人站成一排照相，要求队长必须站中间，教练组 3 人相邻并站在边上，满足要求的站法共有（　　）

A.$A_3^3A_{11}^{11}$ 种

B.$2A_3^3A_{11}^{11}$ 种

C.$A_3^3A_4^4A_7^7$ 种

D.$2A_3^3A_4^4A_7^7$ 种

B　**解析**：选择左、右两边其中一边将教练组 3 人捆绑看作一个整体安排，共有 $2A_3^3$ 种站法，

将剩余的 11 名队员全排列，共有 $A_{11}^{11}$ 种站法，

由分步乘法计数原理可得总的站法有 $2A_3^3A_{11}^{11}$ 种.

故选 B.

## 二、多项选择题

**9.** 已知在 $\left(\sqrt[3]{x}-\dfrac{1}{2\sqrt[3]{x}}\right)^n$ 的展开式中，第 6 项为常数项，则（　　）

A.$n=11$

B.第 4 项为 $-15x^{\frac{4}{3}}$

C.展开式中所有的有理项共有 3 项

D.第 5 项的二项式系数最大

BC　**解析**：由题可知，展开式的通项为 $T_{k+1}=C_n^kx^{\frac{n-k}{3}}\cdot\left(-\dfrac{1}{2\sqrt[3]{x}}\right)^k=C_n^k\left(-\dfrac{1}{2}\right)^k\cdot x^{\frac{n-2k}{3}}$.由第 6 项为常数项，得当 $k=5$ 时，$\dfrac{n-2k}{3}=0$，得 $n=10$.令 $\dfrac{10-2k}{3}=m\in\mathbf{Z}$，则 $10-2k=3m$，即 $k=5-\dfrac{3}{2}m$，故 $m$ 应为偶数.又 $0\leqslant k\leqslant10$，故 $m$ 可取 $2,0,-2$，即 $k$ 可取 $2,5,8$.故第 3 项、第 6 项与第 9 项为有理项.又 $n=10$，故第 6 项的二项式系数最大.将 $n=10$，$k=3$ 代入 $T_{k+1}$ 可得第 4 项为 $-15x^{\frac{4}{3}}$.

**10.** 盒中有 10 颗螺丝钉，其中有 3 颗是坏的，现从盒中随机地抽取 4 颗，则（　　）

A.恰有 1 颗是坏的的概率为 $\dfrac{1}{30}$

B.4 颗全是好的的概率为 $\dfrac{1}{6}$

C.恰有 2 颗是好的的概率为 $\dfrac{3}{10}$

D.至多有 2 颗是坏的的概率为 $\dfrac{1}{3}$

BC　**解析**：设 $X=k$ 表示取出的螺丝钉中恰有 $k$ 颗为好的，则 $P(X=k)=\dfrac{C_7^kC_3^{4-k}}{C_{10}^4}(k=1,2,3,4)$.

所以 $P(X=1)=\dfrac{1}{30}$，$P(X=2)=\dfrac{3}{10}$，$P(X=3)=\dfrac{1}{2}$，$P(X=4)=\dfrac{1}{6}$.

恰有 1 颗是坏的的概率为 $P(X=3)=\dfrac{1}{2}$，

至多有 2 颗是坏的的概率为 $P(X=4)+P(X=3)+P(X=2)=\dfrac{29}{30}$.

**11.** 已知随机变量 $X$ 服从正态分布 $N(2,\sigma^2)$，且 $P(X<4)=0.8$，则（　　）

A.正态曲线关于直线 $x=2$ 对称

B.$\sigma=2$

C.$P(X>4)=P(X<0)$

D.$P(0<X<2)=0.2$

AC　**解析**：由 $\mu=2$，可知正态曲线关于直线 $x=2$ 对称.由 $P(X<4)=0.8$，知 $P(X>4)=P(X<0)=0.2$，故 $P(0<X<2)=0.3$.故选 AC.

## 三、填空题

**12.** 随着现代科技的不断发展,使用手机支付越来越普遍,其中某群体的每位成员使用手机支付的概率都为 $p$,各成员的支付方式相互独立.设 $X$ 为该群体的 10 位成员中使用手机支付的人数,已知 $D(X)=2.4$ 且 $P(X=4)>P(X=6)$,则 $E(X)=$________.

4 **解析**:由题易知,$X\sim B(10,p)$.

因为 $D(X)=2.4,P(X=4)>P(X=6)$,

所以 $\begin{cases}10p(1-p)=2.4,\\ C_{10}^{4}p^{4}(1-p)^{6}>C_{10}^{6}p^{6}(1-p)^{4},\end{cases}$ 解得 $p=0.4$.

所以 $E(X)=10p=4$.

**13.** 某考察团对全国十大城市进行职工人均工资 $x$(单位:千元)与居民人均消费 $y$(单位:千元)统计调查,发现 $y$ 与 $x$ 具有相关关系,且经验回归方程为 $\hat{y}=0.66x+1.562$.当人均工资增加 1 千元时,人均消费平均增加________千元.若某城市居民人均消费为 7.675 千元,估计该城市人均消费占人均工资的百分比为________.

0.66 83% **解析**:由题可知 $\hat{b}=0.66$,故 $x$ 增加 1 时,$y$ 平均增加 0.66. 当 $\hat{y}=7.675$ 时,$x=\dfrac{7.675-1.562}{0.66}\approx 9.262$,所以 $\dfrac{7.675}{9.262}\times 100\ \%\approx 83\%$.

**14.** 某员工上班选择自驾、坐公交车、骑共享单车的概率分别为 $\dfrac{1}{6},\dfrac{1}{3},\dfrac{1}{2}$,而他自驾、坐公交车、骑共享单车迟到的概率分别为 $\dfrac{1}{4},\dfrac{1}{5},\dfrac{1}{6}$,结果今天他迟到了,在此条件下,他自驾去上班的概率为________.

$\dfrac{5}{23}$ **解析**:设该员工迟到为事件 $A$,该员工自驾为事件 $B$,

则 $P(A)=\dfrac{1}{6}\times\dfrac{1}{4}+\dfrac{1}{3}\times\dfrac{1}{5}+\dfrac{1}{2}\times\dfrac{1}{6}=\dfrac{23}{120}$,

$P(AB)=P(B)P(A|B)=\dfrac{1}{6}\times\dfrac{1}{4}=\dfrac{1}{24}$,

所以在该员工迟到的条件下,他自驾去上班的概率为 $P(B|A)=\dfrac{P(AB)}{P(A)}=\dfrac{\frac{1}{24}}{\frac{23}{120}}=\dfrac{5}{23}$.

## 四、解答题

**15.** 某校高中部,高一有 6 个班,高二有 7 个班,高三有 8 个班,学校利用周末组织学生到某厂进行社会实践活动.

(1)任选 1 个班的学生参加社会实践,有多少种不同的选法?

(2)三个年级各选 1 个班的学生参加社会实践,有多少种不同的选法?

(3)选 2 个班的学生参加社会实践,要求这 2 个班不同年级,有多少种不同的选法?

**解**:(1)分三类:第 1 类,从高一年级选 1 个班,有 6 种不同的选法;第 2 类,从高二年级选 1 个班,有 7 种不同的选法;第 3 类,从高三年级选 1 个班,有 8 种不同的选法.由分类加法计数原理,得共有 $6+7+8=21$ 种不同的选法.

(2)分三步:第 1 步,从高一年级选 1 个班,有 6 种不同的选法;第 2 步,从高二年级选 1 个班,有 7 种不同的选法;第 3 步,从高三年级选 1 个班,有 8 种不同的选法.由分步乘法计数原理,得共有 $6\times 7\times 8=336$ 种不同的选法.

(3)分三类,每类又分两步.第 1 类,从高一、高二两个年级各选 1 个班,有 $6\times 7$ 种不同的选法;第 2 类,从高一、高三两个年级各选 1 个班,有 $6\times 8$ 种不同的选法;第 3 类,从高二、高三年级各选 1 个班,有 $7\times 8$ 种不同的选法.故共有 $6\times 7+6\times 8+7\times 8=146$ 种不同的选法.

**16.** 区教育局准备组织一次安全知识竞赛.某校为了选拔学生参赛,按性别采用分层随机抽样的方法抽取 200 名学生进行安全知识测试,记 $A=$“性别为男”,$B=$“得分超过 85 分”,且 $P(A|\overline{B})=\dfrac{2}{5}$,$P(B|\overline{A})=\dfrac{5}{8}$,$P(B)=\dfrac{3}{4}$.

(1)完成下列 $2\times 2$ 列联表,根据小概率值 $\alpha=0.001$ 的独立性检验,能否推断该校学生了解安全知识的程度与性别有关?

单位:人

| 性别 | 了解安全知识的程度 | | 合计 |
|---|---|---|---|
| | 得分不超过 85 分 | 得分超过 85 分 | |
| 男 | | | |
| 女 | | | |
| 合计 | | | |

(2)学校准备选取参与测试的男生和女生各前两名学生代表学校参加区级别的竞赛,已知男生获奖的概率为 $\dfrac{3}{4}$,女生获奖的概率为 $\dfrac{2}{3}$,记该校获奖的人数为 $X$,求 $X$ 的分布列与数学期望.

附:

| $\alpha$ | 0.1 | 0.05 | 0.01 | 0.005 | 0.001 |
|---|---|---|---|---|---|
| $x_\alpha$ | 2.706 | 3.841 | 6.635 | 7.879 | 10.828 |

**解**:(1)由 $P(B)=\frac{3}{4}$,可知得分超过 85 分的人数为 $200\times\frac{3}{4}=150$,得分不超过 85 分的人数为 $200-150=50$.

因为 $P(\overline{A}|\overline{B})=1-P(A|\overline{B})=\frac{3}{5}$,$P(\overline{B}|\overline{A})=1-P(B|\overline{A})=\frac{3}{8}$,$P(B)=\frac{3}{4}$,$P(\overline{B})=\frac{1}{4}$,

所以 $P(\overline{A}|\overline{B})\cdot P(\overline{B})=P(\overline{B}|\overline{A})\cdot P(\overline{A})$,即 $\frac{3}{5}\times\frac{1}{4}=\frac{3}{8}P(\overline{A})$,得 $P(\overline{A})=\frac{2}{5}$,$P(A)=1-P(\overline{A})=\frac{3}{5}$,

故 200 人中男性人数为 $200\times\frac{3}{5}=120$,女性人数为 $200-120=80$.

又 $P(A|\overline{B})=\frac{2}{5}$,即得分不超过 85 分的人中,男性人数为 $50\times\frac{2}{5}=20$,女性人数为 $50-20=30$,

故在得分超过 85 分的人中,男性人数为 $120-20=100$,女性人数为 $80-30=50$.

$2\times2$ 列联表如下:

单位:人

| 性别 | 了解安全知识的程度 | | 合计 |
|---|---|---|---|
| | 得分不超过 85 分 | 得分超过 85 分 | |
| 男 | 20 | 100 | 120 |
| 女 | 30 | 50 | 80 |
| 合计 | 50 | 150 | 200 |

零假设为 $H_0$:该校学生了解安全知识的程度与性别没有关联.

经计算得到 $\chi^2=\frac{200\times(20\times50-100\times30)^2}{120\times80\times50\times150}=\frac{100}{9}\approx$

$11.11>10.828=x_{0.001}$.

根据小概率值 $\alpha=0.001$ 的独立性检验,可以推断 $H_0$ 不成立,即认为了解安全知识的程度与性别有关,此推断犯错误的概率不大于 0.001.

(2)$X$ 可能取 0,1,2,3,4.

$P(X=0)=\left(\frac{1}{4}\right)^2\times\left(\frac{1}{3}\right)^2=\frac{1}{144}$,

$P(X=1)=C_2^1\times\frac{3}{4}\times\frac{1}{4}\times\left(\frac{1}{3}\right)^2+C_2^1\times\frac{2}{3}\times\frac{1}{3}\times\left(\frac{1}{4}\right)^2=\frac{10}{144}$;

$P(X=2)=\left(\frac{3}{4}\right)^2\times\left(\frac{1}{3}\right)^2+\left(\frac{1}{4}\right)^2\times\left(\frac{2}{3}\right)^2+C_2^1\times\frac{2}{3}\times\frac{1}{3}\times C_2^1\times\frac{3}{4}\times\frac{1}{4}=\frac{37}{144}$;

$P(X=3)=\left(\frac{3}{4}\right)^2\times C_2^1\times\frac{2}{3}\times\frac{1}{3}+C_2^1\times\frac{3}{4}\times\frac{1}{4}\times\left(\frac{2}{3}\right)^2=\frac{60}{144}$;

$P(X=4)=\left(\frac{3}{4}\right)^2\times\left(\frac{2}{3}\right)^2=\frac{36}{144}$.

所以 $X$ 的分布列为

| $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P$ | $\frac{1}{144}$ | $\frac{10}{144}$ | $\frac{37}{144}$ | $\frac{60}{144}$ | $\frac{36}{144}$ |

所以 $E(X)=0\times\frac{1}{144}+1\times\frac{10}{144}+2\times\frac{37}{144}+3\times\frac{60}{144}+4\times\frac{36}{144}=\frac{17}{6}$.

**17.** 生物学家认为,睡眠中的恒温动物依然会消耗体内能量,主要是为了保持体温.脉搏率 $f$ 是单位时间心跳的次数,医学研究发现,动物的体重 $W$(单位:g)与脉搏率 $f$ 存在着一定的关系.下表中给出一些动物体重与脉搏率对应的数据,图 1 画出了体重 $W$ 与脉搏率 $f$ 的散点图,图 2 画出了 $\lg W$ 与 $\lg f$ 的散点图.

| 动物名 | 体重 | 脉搏率 |
|---|---|---|
| 小鼠 | 25 | 670 |
| 大鼠 | 200 | 420 |
| 豚鼠 | 300 | 300 |
| 兔 | 2 000 | 200 |
| 小狗 | 5 000 | 120 |
| 大狗 | 30 000 | 85 |
| 羊 | 50 000 | 70 |

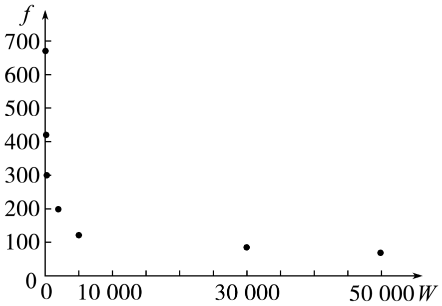


图 1

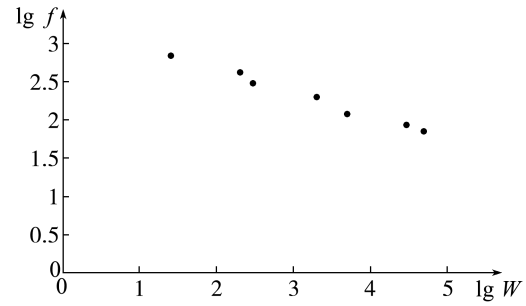



图 2

为了较好地描述体重和脉搏率的关系，现有以下两种模型供选择：

①$f=kW+b$；

②$\lg f=k\lg W+b$.

(1)选出你认为更符合实际的函数模型，并说明理由；

(2)不妨取表中豚鼠和兔的体重、脉搏率数据代入所选函数模型，求出 $f$ 关于 $W$ 的函数解析式；

(3)若马的体重是兔的 256 倍，根据(2)的结论，预测马的脉搏率.

附：$\lg 2\approx0.3$，$\lg 3\approx0.5$.

**解**：(1)模型②$\lg f=k\lg W+b$ 更符合实际.理由如下：根据题中散点图的特征，题图 2 基本上呈直线形式，所以可以选择一次函数来刻画 $\lg W$ 和 $\lg f$ 的关系.

(2)由题意可得$\begin{cases}\lg 300=k\lg 300+b,\\ \lg 200=k\lg 2\,000+b,\end{cases}$

又 $\lg 200=\lg 2+2\approx2.3$，$\lg 2\,000=\lg 2+3\approx3.3$，$\lg 300=\lg 3+2\approx2.5$，

故解得$\begin{cases}k=-\dfrac{1}{4},\\ b=\dfrac{25}{8},\end{cases}$

所以 $\lg f=-\dfrac{\lg W}{4}+\dfrac{25}{8}$.

所以 $f$ 关于 $W$ 的函数解析式为 $f=10^{\frac{25}{8}}\cdot W^{-\frac{1}{4}}$.

(3)设马的体重和脉搏率分别为 $W_1$，$f_1$，兔的体重和脉搏率分别为 $W_2$，$f_2$.由题意可知$\dfrac{W_1}{W_2}=256$，则

$\dfrac{f_1}{f_2}=\left(\dfrac{W_1}{W_2}\right)^{-\frac{1}{4}}=256^{-\frac{1}{4}}=(2^8)^{-\frac{1}{4}}=\dfrac{1}{4}$.

又因为 $f_2=200$，所以 $f_1=50$.

即马的脉搏率约为 50.

**18.**某种可能遭受污染的海产品在进入市场前必须进行两轮检测，只有两轮都合格才能进行销售，否则不能销售.已知该海产品第一轮检测不合格的概率为$\dfrac{1}{6}$，第二轮检测不合格的概率为$\dfrac{1}{10}$，两轮检测是否合格相互没有影响.

(1)求该海产品不能销售的概率.

(2)若该海产品可以销售，则每件产品可获利 40 元；若该海产品不能销售，则每件产品亏损 80 元(即获利$-80$元).已知一箱该海产品有 4 件，记一箱该海产品获利 $\xi$ 元，求 $\xi$ 的分布列，并求出均值 $E(\xi)$.

**解**：(1)设“该海产品不能销售”为事件 $A$，

则 $P(A)=1-\left(1-\dfrac{1}{6}\right)\times\left(1-\dfrac{1}{10}\right)=\dfrac{1}{4}$.

所以该海产品不能销售的概率为$\dfrac{1}{4}$.

(2)由已知得 $\xi$ 的所有可能取值为$-320$，$-200$，$-80$，40，160.

$P(\xi=-320)=\left(\dfrac{1}{4}\right)^4=\dfrac{1}{256}$，

$P(\xi=-200)=C_4^1\times\left(\dfrac{1}{4}\right)^3\times\dfrac{3}{4}=\dfrac{3}{64}$，

$P(\xi=-80)=C_4^2\times\left(\dfrac{1}{4}\right)^2\times\left(\dfrac{3}{4}\right)^2=\dfrac{27}{128}$，

$P(\xi=40)=C_4^3\times\dfrac{1}{4}\times\left(\dfrac{3}{4}\right)^3=\dfrac{27}{64}$，

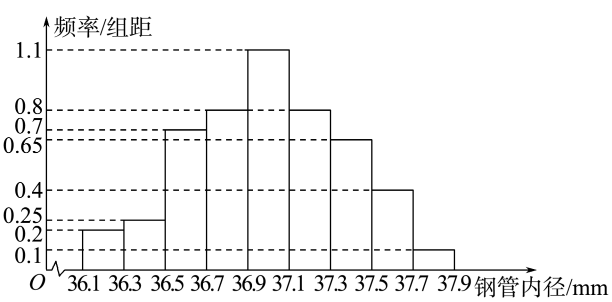
$P(\xi=160)=\left(\dfrac{3}{4}\right)^4=\dfrac{81}{256}$.

所以 $\xi$ 的分布列为

| $\xi$ | $-320$ | $-200$ | $-80$ | 40 | 160 |
|---|---|---|---|---|---|
| $P$ | $\dfrac{1}{256}$ | $\dfrac{3}{64}$ | $\dfrac{27}{128}$ | $\dfrac{27}{64}$ | $\dfrac{81}{256}$ |

$E(\xi)=(-320)\times\dfrac{1}{256}-200\times\dfrac{3}{64}-80\times\dfrac{27}{128}+40\times\dfrac{27}{64}+160\times\dfrac{81}{256}=40$.

**19.**根据以往大量的测量统计，某企业生产的钢管内径尺寸 $X$(单位：mm)服从正态分布 $N(\mu,\sigma^2)$，并把钢管内径在$(\mu-\sigma,\mu+\sigma)$内的产品称为一等品，钢管内径在$[\mu+\sigma,\mu+2\sigma)$内的产品称为二等品，一等品与二等品统称为正品，其余范围内的产品作为废品回收.现从该企业生产的产品中随机抽取 1 000件，测得钢管内径的样本数据的频率分布直方图如图所示.



(1)通过检测得到样本数据的标准差 $s=0.3$，用样本平均数$\overline{x}$作为$\mu$的近似值，用样本标准差 $s$ 作为$\sigma$的估计值，根据所给数据求该企业生产的产品为正品的钢管内径尺寸范围.(同一组中的数据用该组区间的中点值代表)

(2)假如企业包装时要求把 2 件一等品和 $n(n\geqslant 2,n\in\mathbf{N})$件二等品装在同一个箱子中，质检员从某箱子中摸出 2 件产品进行检验，若抽取到的 2 件产品等级相同，则该箱产品记为 $A$，否则该箱产品记为 $B$.

①试用含 $n$ 的代数式表示某箱产品抽检被记为 $B$ 的概率 $p$；

②设抽检 5 箱产品恰有 3 箱被记为 $B$ 的概率为$f(p)$，求当 $n$ 为何值时，$f(p)$取得最大值，并求出最大值.

附：$36.2\times0.2+36.4\times0.25+36.6\times0.7+36.8\times0.8+37\times1.1+37.2\times0.8+37.4\times0.65+37.6\times0.4+37.8\times0.1\approx185$.

**解**：(1)由题意，得样本平均数$\overline{x}\approx185\times0.2=37$，所以估计 $\mu=\overline{x}=37$，$\sigma=s=0.3$，

则 $\mu-\sigma=37-0.3=36.7$，$\mu+\sigma=37+0.3=37.3$，$\mu+2\sigma=37+0.6=37.6$，

则一等品钢管内径在$(\mu-\sigma,\mu+\sigma)$，即$(36.7,37.3)$内，

二等品钢管内径在$[\mu+\sigma,\mu+2\sigma)$，即$[37.3,37.6)$内，

所以该企业生产的产品为正品的钢管内径尺寸范围为$(36.7,37.6)$.

(2)①从$(n+2)$件正品中任选 2 件，有 $C_{n+2}^{2}$ 种选法，其中等级相同的有 $C_{n}^{2}+C_{2}^{2}$ 种选法，

所以某箱产品抽检被记为 $B$ 的概率为 $p=1-\dfrac{C_{n}^{2}+C_{2}^{2}}{C_{n+2}^{2}}=1-\dfrac{n^{2}-n+2}{n^{2}+3n+2}=\dfrac{4n}{n^{2}+3n+2}(n\geqslant2,n\in\mathbf{N})$.

②由题意，一箱产品抽检被记为 $B$ 的概率为 $p$，

则抽检 5 箱产品恰有 3 箱被记为 $B$ 的概率为

$f(p)=C_{5}^{3}p^{3}(1-p)^{2}=10p^{3}(1-2p+p^{2})=10(p^{3}-2p^{4}+p^{5})$，

$f'(p)=10(3p^{2}-8p^{3}+5p^{4})=10p^{2}(3-8p+5p^{2})=10p^{2}(p-1)(5p-3)$，

所以当 $p\in\left(0,\dfrac{3}{5}\right)$时，$f'(p)>0$，函数 $f(p)$单调递增；

当 $p\in\left(\dfrac{3}{5},1\right)$时，$f'(p)<0$，函数 $f(p)$单调递减.

所以当 $p=\dfrac{3}{5}$时，$f(p)$取得最大值 $f\left(\dfrac{3}{5}\right)=C_{5}^{3}\times\left(\dfrac{3}{5}\right)^{3}\times\left(1-\dfrac{3}{5}\right)^{2}=\dfrac{216}{625}$.

此时，$p=\dfrac{4n}{n^{2}+3n+2}=\dfrac{3}{5}$，解得 $n=3$ 或 $n=\dfrac{2}{3}$(舍去).

所以当 $n=3$ 时，$f(p)$取得最大值 $\dfrac{216}{625}$.