点金训练

教师用书

数学

选择性必修第三册

配人教A版

《点金训练》编写组　编

DIANJIN XUNLIAN

—— SHUXUE ——

JIAOSHI YONGSHU

四川教育出版社

教师用书　点金训练　数学　选择性必修第三册　配人教A版

四川教育出版社

winshare文轩教育
四川教育出版社

赠　品

6662025015019

扫码查看本书
配套资源包

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

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(正确的打"√",错误的打"×").

(1)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.　(√)

(2)在分步乘法计数原理中,事情是分步完成的,其中任何一个单独的步骤都能完成这件事.　(×)

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种方法;第2步,出门,共有4种方法.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.请思考并回答下列问题:

你能举一些生活中利用分类加法计数原理计数的例子吗?

提示:①某班有男生23人,女生26人,从中选出一人当班长,有多少种选法?

②有3名女同学,5名男同学,从中选出一人主持元旦晚会,则不同的选法有多少种?

任务型课堂

学习任务一　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. 数学中把 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，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，有 141，252，303，414，525，共有 5 个.

根据分类加法计数原理可得，满足条件的回文数共有 10 个.

3. 若 $x,y\in\mathbf{N}^*$，且 $x+y\leqslant 6$，试求有序自然数对 (x,y) 的个数.

解：按 x 的取值进行分类：

当 $x=1$ 时，$y=1,2,3,4,5$，共构成 5 个有序自然数对；

当 $x=2$ 时，$y=1,2,3,4$，共构成 4 个有序自然数对；

当 $x=3$ 时，$y=1,2,3$，共构成 3 个有序自然数对；

当 $x=4$ 时，$y=1,2$，共构成 2 个有序自然数对；

当 $x=5$ 时，$y=1$，共构成 1 个有序自然数对.

根据分类加法计数原理，共有 $N=5+4+3+2+1=15$(个)有序自然数对.

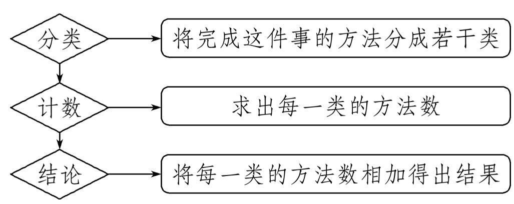
反思提炼

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



学习任务二　分步乘法计数原理

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，共有 200 种；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.

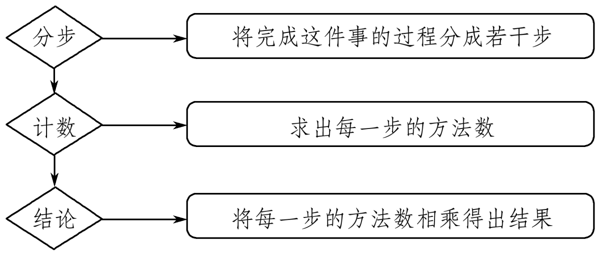
反思提炼

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;

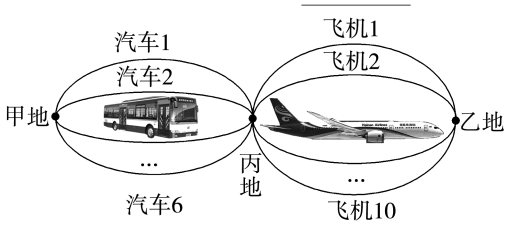
(2)只有各个步骤都完成才算完成这件事.

2. 利用分步乘法计数原理计数的流程



探究训练

1. 如图,从甲地到乙地的途径有________种.



60　**解析**:从甲地到乙地这件事可分两步完成:第1步,从甲地到丙地,有6种不同的途径;第2步,从丙地到乙地,有10种不同的途径.所以从甲地到乙地共有 $6\times10=60$(种)途径.

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到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**: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).

体系构建

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	各步按一定顺序都完成,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)

基础性·能力运用

1. 从甲地到乙地一天有汽车8班、火车3班、轮船2班.某人从甲地到乙地,则不同的方法有　(　　)

A.13种　　B.16种

C.24种　　D.48种

A　**解析**:应用分类加法计数原理,不同的方法有 $8+3+2=13$(种).故选A.

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,故不同报名方法有 $2^5=32$(种).故选D.

4. 由A村去B村的道路有4条,由B村去C村的道路有3条,从A村经B村去C村不同的走法有　(　　)

A.7种　　B.9种

C.11种　　D.12种

D　**解析**:由分步乘法计数原理知,有 $4\times3=12$(种)不同的走法.故选D.

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 $2\times3\times4=24$.

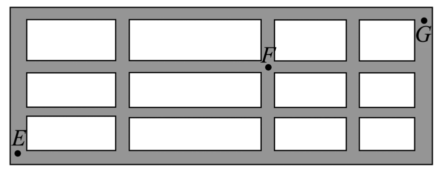
6.已知 $a\in\{-1,2,3\}$,$b\in\{0,3,4,5\}$.方程 $(x-a)^2+(y-b)^2=4$ 表示的不同的圆共有________个;若 $r\in\{1,2\}$,则方程 $(x-a)^2+(y-b)^2=r^2$ 表示的不同的圆共有________个.

12 24 **解析**:分两步,第 1 步,a 的值有 3 种选法;第 2 步,b 的值有 4 种选法,表示的不同的圆共有 $3\times4=12$(个).分三步,第 1 步,a 的值有 3 种选法;第 2 步,b 的值有 4 种选法;第 3 步,r 的值有 2 种选法.故表示的不同的圆共有 $3\times4\times2=24$(个).

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________.

8.1×10^7 **解析**:由题意知,本题是一个分步计数问题.电话号码是七位时,该城市电话号码有 9×10^6 个,同理,改为八位时,有 9×10^7 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的条数为________.



18 **解析**:由题意可知 $E\to F$ 有 6 种走法,$F\to G$ 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.

综合性·创新提升

1.已知三棱锥 A-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,不同路径有 $ABCA$,$ABDA$,$ADBA$,$ADCA$,$ACBA$,$ACDA$,共有 6 条不同路径.故选 B.

2.植树节那天,有 4 名同学参与植树,现将 3 棵不同种类的树分配给 4 名同学.若每棵树均由 1 人完成种植,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,共有不同的分配方法 $4^3=64$(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.其余三项工作的选法有 $5\times4\times3$ 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 人必须在编号为 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).

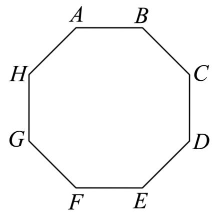
第 2 步:安排另外 5 人,可安排在 2,4,6,8 及余下的一条奇数号跑道上,安排方式有 $5\times4\times3\times2\times1=120$(种).

所以安排这 8 人的方式有 $24\times120=2\ 880$(种).

5.已知直线方程 $Ax+By=0$,若从 0,1,2,3,5,7 这六个数字中每次取两个不同的数作为系数 A,B 的值,则方程可表示不同直线的条数是________.

22 **解析**:若 $A=0$,则 B 的值从 1,2,3,5,7 中任取一个,均表示直线 $y=0$;同理,当 $B=0$ 时,均表示直线 $x=0$;当 $A\neq0$ 且 $B\neq0$ 时,能表示 $5\times4=20$(条)不同的直线.故方程可表示不同直线的条数是 $1+1+20=22$.

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.现有标号为 A,B,C 的三个口袋,A 袋中有 1 个红色小球,B 袋中有 2 个不同的白色小球,C 袋中有 3 个不同的黄色小球.从三个口袋中取出 2 个小球.

(1)若取出的两个小球的颜色不同,有多少种取法?

(2)若取出的两个小球颜色相同,有多少种取法?

解:(1)若两个小球颜色不同,则应从 A,B 袋中各取 1 个,或从 A,C 袋中各取 1 个,或从 B,C 袋中各取 1 个,共有 $1\times2+1\times3+2\times3=11$(种)取法.

(2)若两个小球颜色相同,则应从 B 袋中取出 2 个,或从 C 袋中取出 2 个,共有 $1+3=4$(种)取法.

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实际问题.

问题式预习

知识清单

知识点　计数原理的应用

(1)应用计数原理的注意点

一是要确定完成的“一件事”是什么;二是注意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,恰好完成任务.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)有三只口袋装有小球,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.若从口袋中取 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)作和,则取出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.

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？

(2)综合应用两个计数原理的解题思路是什么？

(1)**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)**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，从而便于我们解题.

任务型课堂

学习任务一　选（抽）取与分配问题

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）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）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）放法.

反思提炼

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，然后减去不符合条件的抽取方法数即可.

探究训练

1.高三年级的四个班到甲、乙、丙、丁、戊五个工厂进行社会实践，其中工厂甲必须有班级去，则不同的分配方案有（　　）

A.360 种　　B.420 种

C.369 种　　D.396 种

C　**解析**：方法一（直接法）

以甲工厂分配班级情况进行分类，共分为四类：

第一类，四个班级都去甲工厂，此时分配方案只有 1 种情况；

第二类，有三个班级去甲工厂，剩下的一个班级去另外四个工厂，其分配方案共有 $4\times4=16$（种）；

第三类，有两个班级去甲工厂，另外两个班级去其他四个工厂，其分配方案共有 $6\times4\times4=96$（种）；

第四类，有一个班级去甲工厂，其他三个班级去另

外四个工厂，其分配方案有 $4\times4\times4\times4=256$(种).

综上所述，不同的分配方案有 $1+16+96+256=369$(种).

方法二（间接法）

先计算四个班到五个工厂的总分配方案数，再扣除甲工厂无人去的情况，即有 $5\times5\times5\times5-4\times4\times4\times4=369$(种)分配方案.

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.

学习任务二　排数问题

例 2　(1)用 0，1，2，3，4 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)由 0，1，2，3，4 这 5 个数字可以组成多少个三位数?

(3)由 0，1，2，3，4 这 5 个数字可以组成多少个能被 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 种排法，共有 $5\times5\times5=5^3=125$(个)不同编号.

(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 种排法，第二、三位可以排 0，因此，共有 $4\times5\times5=100$(个)三位数.

(3)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，则有 $4\times3=12$(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.

因而共有 $12+18=30$(种)排法.

即可以组成 30 个能被 2 整除的无重复数字的三位数.

反思提炼

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.

探究训练

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，第 1 类，若末位数字是 0，则有 $5\times4\times3=60$(个)；第 2 类，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 0～9 十个数字，可以组成有重复数字的三位数的个数为　　(　　)

A.243　　B.252　　C.261　　D.279

B　**解析**：由分步乘法计数原理知，用 0，1，…，9 十个数字组成的三位数共有 $9\times10\times10=900$(个)，组成无重复数字的三位数共有 $9\times9\times8=648$(个).因此组成有重复数字的三位数共有 $900-648=252$(个).

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:

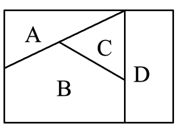
3	1	×
3	2	×
3	3	×

共有 $2\times4\times4+3\times4=44$(项).

所以 $n=44+1=45$.

学习任务 三　涂色与种植问题

例 3　用 n 种不同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 种着色方法,再为 B 着色,有 5 种着色方法,然后为 C 着色,有 4 种着色方法,最后为 D 着色,也有 4 种着色方法,

所以,共有着色方法 $6\times5\times4\times4=480$(种).

(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 $n(n-1)(n-2)(n-2)=180$,

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.

反思提炼

解决涂色(种植)问题的一般思路

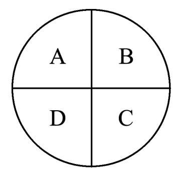
(1)以区域为主分步计数,对于不相邻区域,按照同色、不同色分类计数.

(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.

探究训练

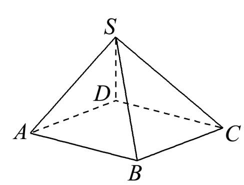
1.如图,一环形花坛分成 A,B,C,D 四块,现有四种不同的花供选种,要求在每块花坛里种一种花,且相邻的两块花坛种不同的花,则不同的种法数为　(　　)



A.96　　B.84　　C.60　　D.48

B　**解析**:当 C 与 A 种同一种花时,有 $4\times3\times1\times3=36$(种)种法;当 C 与 A 所种的花不同时,有 $4\times3\times2\times2=48$(种)种法.由分类加法计数原理知,不同的种法数为 $36+48=84$.

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. 小张正在玩一款种菜的游戏，他计划从仓库里的玉米、土豆、茄子、辣椒、胡萝卜这 5 种作物中选出 4 种分别种植在四块不同的空地上(一块空地只能种植一种作物).若小张已决定在第一块空地上种茄子或辣椒，则不同的种植方案共有________种.

48　**解析**：当第一块空地种茄子时，有 $4\times3\times2=24$(种)不同的种法；当第一块空地种辣椒时，有 $4\times3\times2=24$(种)不同的种法.故共有 48 种不同的种植方案.

体系构建

- 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.分析完成一件事应分类还是分步
      - 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)

基础性·能力运用

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，则有 $2\times3=6$(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. 某外语小组有 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 $N_1=6\times3=18$(种)；

第 2 类：从既会英语又会日语的人中选 1 人会英语的有 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 $N_2=1\times2=2$(种).

综上，不同的选法共有 $N=N_1+N_2=18+2=20$(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位前面.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.(新定义)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”.

偶数有 214,312,314,412,324，共 5 个.

6. 某超市的 6 个收费通道排成一排，每个通道处有 1 号、2 号两个收费点，根据每天的人流量，超市准备周一选择开放其中的 3 个通道，要求 3 个通道互不相邻，且每个通道至少开通一个收费点，则周一这天超市开通收费点的安排方式共有________种.

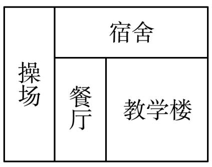
108　**解析**：设 6 个收费通道依次编号为 1,2,3,4,5,6，从中选择 3 个互不相邻的通道，有 135,136,146,246，共 4 种不同的选法.

对于每个通道，至少开通一个收费点，即可以开通 1

号收费点，开通 2 号收费点，同时开通两个收费点，共 3 种不同的安排方式.

由分步乘法计数原理，可得超市开通收费点的安排方式共有 $4\times3^3=108$（种）.

7. 如图是某校的校园设施平面图，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 种不同的颜色可选，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；教学楼和宿舍、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）着色方法.

方法二：分两类.第一类，操场与教学楼同色，有 $6\times5\times4=120$（种）着色方法；第二类，操场与教学楼不同色，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. 回文数是指从左到右读与从右到左读都一样的正整数，如 22，121，3 443，94 249 等.显然 2 位回文数有 9 个：11，22，33，…，99，3 位回文数有 90 个：101，111，121，…，191，202，…，999.

(1)5 位回文数有________个；

(2)$2n(n\in\mathbf{N}^*)$位回文数有________个.

(1)900　(2)$9\times10^{n-1}$　**解析**：(1)5 位回文数的个数等于三位正整数的个数，首位不能为零，首位有 9 种选法，十位和个位各有 10 种选法，由分步乘法计数原理可知，5 位回文数的个数为 $9\times10^2=900$.

(2)$2n(n\in\mathbf{N}^*)$位回文数的个数等于 $n(n\in\mathbf{N}^*)$位正整数的个数，

首位不能为零，首位有 9 种选法，其余各数位各有 10 种选法，

由分步乘法计数原理可知，$2n(n\in\mathbf{N}^*)$位回文数的个数为 $9\times10^{n-1}$.

综合性·创新提升

1. 现有不同的红球 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

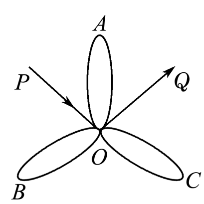
B.若每种颜色选出 1 个球，有 120 种不同的选法

C.若要选出不同颜色的 2 个球，有 31 种不同的选法

D.若要不放回地依次选出 2 个球，有 210 种不同的选法

C　**解析**：A 项，从中任选 1 个球，有 $4+5+6=15$（种）不同的选法，所以该选项正确；B 项，若每种颜色选出 1 个球，有 $4\times5\times6=120$（种）不同的选法，所以该选项正确；C 项，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，所以该选项错误；D 项，若要不放回地依次选出 2 个球，有 $15\times14=210$（种）不同的选法，所以该选项正确.故选 C.

2. 一个旅游景区的游览线路如图所示，某人从点 P 处进，点 Q 处出，沿图中线路游览 A,B,C 三个景点及沿途风景，则不重复（除交汇点 O 外）的不同的游览线路有　　（　　）



A.6 种　　B.8 种

C.12 种　　D.48 种

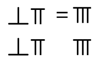
D　**解析**：每个景区都有 2 条线路，所以游览第一个景点有 6 种选法，游览第二个景点有 4 种选法，游览第三个景点有 2 种选法，故共有 $6\times4\times2=48$（种）不同的游览线路.

3. 算筹是在珠算发明以前我国独创并且有效的计算工具，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：



6728

6708

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**：按每一位的算筹的根数分类，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.将 3 种作物种植在如图所示的 5 块试验田里，每块试验田种植一种作物且相邻的试验田不能种植同一种作物，有多少种不同的种植方法？

解：分别用 a,b,c 代表 3 种作物，先安排第一块试验田，有 3 种方法.不妨设种植 a，再安排第二块试验田，有 b 或 c，共 2 种方法.不妨设种植 b，第三块试验田也有 a 或 c，共 2 种方法.

(1)若第三块试验田种植 c，则第四、五块试验田分别有 2 种方法，共有 $2\times2=4$(种)方法.

(2)若第三块试验田种植 a，则第四块试验田种植 b 或 c，共 2 种方法：

①若第四块试验田种植 c：

a	b	a	c

第五块试验田有 2 种种植方法；

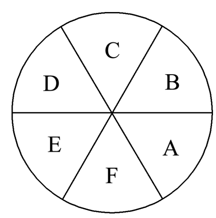
②若第四块试验田种植 b：

a	b	a	b

第五块试验田只能种植 c，共 1 种方法.

综上，共有 $3\times2\times(2\times2+2+1)=42$(种)方法.

6.如图，将圆分成 6 个扇形区域 A,B,C,D,E,F，现给这 6 个区域涂色，要求同一个区域涂同一种颜色，相邻的两个区域不得使用同一种颜色，现有 4 种不同的颜色可用，那么一共有多少种不同的涂色方法？



解：(1)当相间区域 A,C,E 涂同一种颜色时，有 4 种涂色方法，此时，B,D,F 各有 3 种涂色方法，故有 $4\times3\times3\times3=108$(种)涂色方法；

(2)当相间区域 A,C,E 涂 2 种不同的颜色时，有 $4\times3\times3=36$(种)涂色方法，此时 B,D,F 有 $3\times2\times2=12$(种)涂色方法，故共有 $36\times12=432$(种)涂色方法；

(3)当相间区域 A,C,E 涂 3 种不同的颜色时，有 $4\times3\times2=24$(种)涂色方法，此时 B,D,F 各有 2 种涂色方法，故共有 $24\times2\times2\times2=192$(种)涂色方法.

故总计有 $108+432+192=732$(种)涂色方法.

6.2　排列与组合

6.2.1　排列

学习任务目标

1.通过实例理解并掌握排列的概念.

2.能用列举法、树状图法列出简单的排列.

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)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 名学生分别参加数学、物理、化学竞赛，求共有多少种选法，属于排列问题.　(√)

(4)有 12 名学生参加植树活动，要求 3 人一组，求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.求由 1,2,3,4 组成无重复数字的四位数的个数

B.求从 60 人中选 11 人组成足球队的不同选法

C.求从 100 人中选 10 人调查消费习惯的不同选法

D.求从 1,2,3,4,5 中选 2 个数组成的集合的个数

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相同但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

学习任务一　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 个点确定的射线与点的顺序有关，是排列问题；

③与元素的排列顺序无关，不是排列问题.

2.判断下列问题是不是排列问题.

(1)从 3 个小组中选 2 个小组分别去植树和种菜，共有多少种选法？

(2)从 3 个小组中选 2 个小组去种菜,共有多少种选法?

(3)从 20 人中选 10 人组成一个学习小组,共有多少种选法?

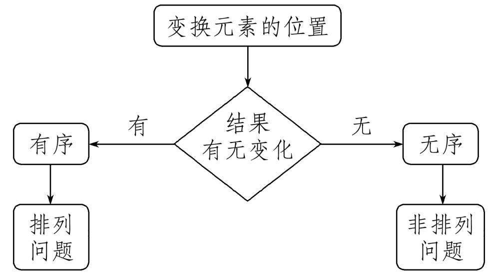
(4)从全班 40 人中选 3 人分别担任班长、学习委员、生活委员,共有多少种选法?

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.

反思提炼



学习任务二　写出简单排列问题的所有排列

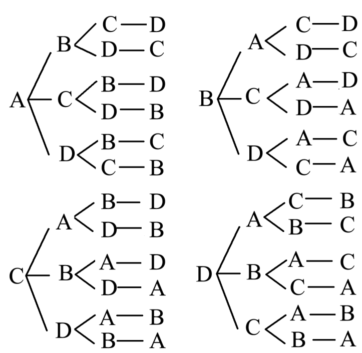
例 1　(1)若 A,B,C 三名同学排成一排照相,有多少种排法?

(2)若 A,B,C,D 四名同学排成一排照相,试将所有排法列出来.

(3)若 A,B,C,D 四名同学排成一排照相,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.

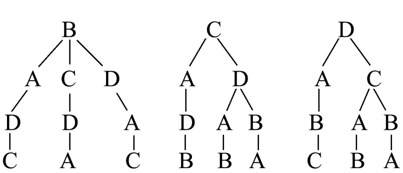
解:(1)所有的排法有 ABC,ACB,BAC,BCA,CAB,CBA,共 6 种.

(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.

反思提炼

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.

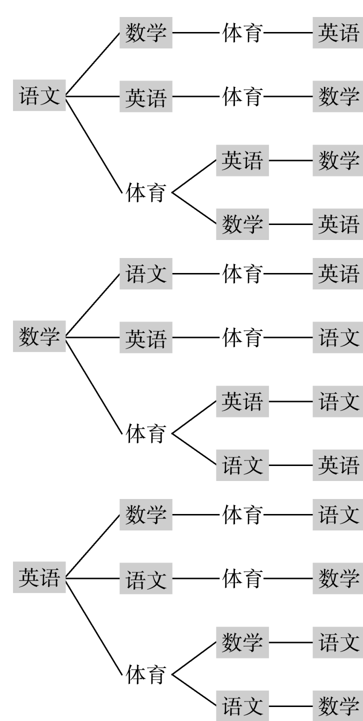
(2)确定分支元素,在每一个分支上将余下的元素按分类标准进行分类,并将各元素按顺序排列.

(3)重复以上步骤,直到写完一个排列为止.

探究训练

1.某班上午要上语文、数学、体育和英语 4 门课,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:



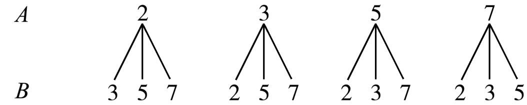
由树状图可知,所有排课方案有:

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，则可以确定多少条不同的直线？试全部列出.

解：画树状图如图：

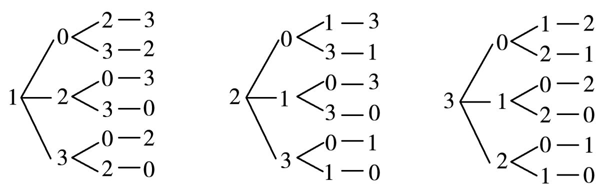


所有不同的直线为 $2x+3y=0$，$2x+5y=0$，$2x+7y=0$，$3x+2y=0$，$3x+5y=0$，$3x+7y=0$，$5x+2y=0$，$5x+3y=0$，$5x+7y=0$，$7x+2y=0$，$7x+3y=0$，$7x+5y=0$，共 12 条.

学习任务三　简单的排列问题

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：由题可知，个位为 0 或 2，所以共有 $2\times3\times4\times4=96$（个）四位偶数.

反思提炼

对于简单的排列问题，解题时首先要确定是利用分类加法计数原理还是分步乘法计数原理，可采用元素分析法或位置分析法求解.

探究训练

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.

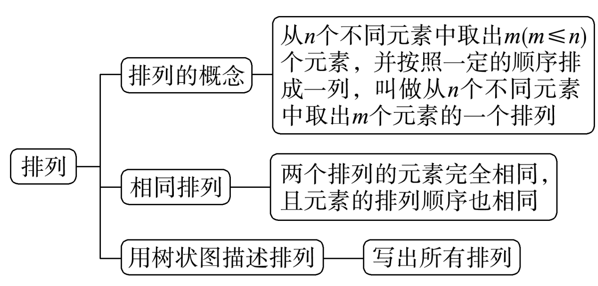
15　**解析**：第 1 类，挂 1 面旗表示信号，有 3 种不同的信号；第 2 类，挂 2 面旗表示信号，有 $3\times2=6$（种）不同的信号；第 3 类，挂 3 面旗表示信号，有 $3\times2\times1=6$（种）不同的信号.

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. 从分别写有 1，2，3，4 的 4 张卡片中任意选取 3 张，构成的无重复数字的三位数有哪些？一共有多少个？如果用分步乘法计数原理，应怎样计算该排列问题的排列的个数？

解：若选取的三张卡片是 1，2，3，则构成的三位数有 123，132，213，231，312，321；若选取的卡片是 1，2，4，则构成的三位数有 124，142，214，241，412，421；若选取的卡片是 1，3，4，则构成的三位数有 134，143，314，341，413，431；若选取的卡片是 2，3，4，则构成的三位数有 234，243，324，342，423，432. 共 24 个. 由分步乘法计数原理得，该排列问题的排列个数为 $4\times3\times2=24$.

体系构建



课后素养评价(三)

基础性·能力运用

1.(多选)完成下列事情,不涉及排列问题的有(　　)

A.由 1,2,3 三个数字组成无重复数字的三位数

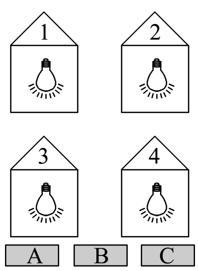
B.从 40 人中选 5 人组成篮球队

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 号灯亮有 $2+2=4$(种)方法.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)方法.

4.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,并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:

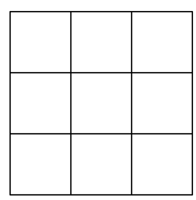
B—A—D—C, B—C—D—A, B—D—A—C
C—A—D—B, C—D—A—B, C—D—B—A
D—A—B—C, D—C—A—B, D—C—B—A

由此可知共有 9 种送法.

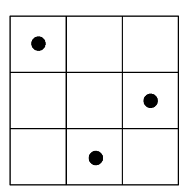
5.从数字 1,3,5,7 中任选两个数做除法,可得不同的商________个.

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.

6.将红、黄、蓝三种颜色的三颗棋子分别放入如图所示的 3×3 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 3×3 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).

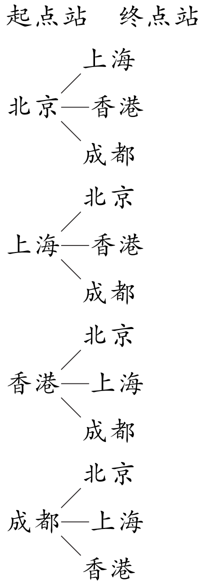


7.(1)北京、上海、香港、成都四个民航站之间的直达航线,需要准备多少种不同的飞机票?将它们列出来.

(2)A,B,C,D 四名同学排成一排照相,要求自左向右,A 不排第一,B 不排第四,共有多少种不同的排列方法?

解:(1)先确定起点,有 4 种情况,再确定终点,有 3 种情况.由分步乘法计数原理知,共需要 $4\times3=12$(种)不同的机票.

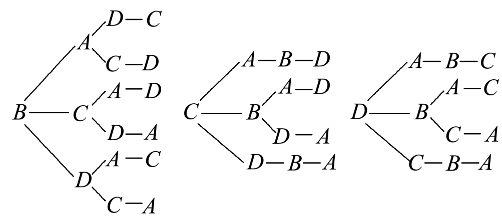
列举如下：



飞机票

北京⟶上海
北京⟶香港
北京⟶成都
上海⟶北京
上海⟶香港
上海⟶成都
香港⟶北京
香港⟶上海
香港⟶成都
成都⟶北京
成都⟶上海
成都⟶香港

(2)因为 A 不排第一,排第一位的情况有 3 类(可从 B,C,D 中任选一人排),而此时兼顾分析 B 的排法,列树状图如图.



所以符合题意的所有排列是 $BADC$,$BACD$,$BCAD$,$BCDA$,$BDAC$,$BDCA$,$CABD$,$CBAD$,$CBDA$,$CDBA$,$DABC$,$DBAC$,$DBCA$,$DCBA$,共 14 种.

综合性·创新提升

1.(多选)下列问题是排列问题的是　(　　)

A.求从甲、乙、丙三名同学中选出两名分别参加数学、物理兴趣小组的方法种数

B.求从甲、乙、丙三名同学中选出两名参加一项活动的方法种数

C.求从 a,b,c,d 中选出 3 个字母的方法种数

D.求从 1,2,3,4,5 中取出 2 个数字组成两位数的个数

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,毽子又被踢回甲处,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，属于排列问题，所以符合条件的直线有 $6\times5=30$(条).

7.(新定义)一个三位数，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.能用计数原理推导排列数公式.

2.能用排列数公式解决简单的实际问题.

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<u>个数</u>，叫做从 n 个不同元素中取出 m 个元素的<u>排列数</u>，用符号 <u>A_n^m</u> 表示.

(2)排列数公式：

①展开式：$A_n^m=\underline{n(n-1)(n-2)\cdots(n-m+1)}$，其中 $m,n\in\mathbf{N}^*$，并且 $m\leqslant n$.

②阶乘式：$A_n^m=\underline{\dfrac{n!}{(n-m)!}}$.

(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<u>全排列</u>.这时，排列数公式中 $m=n$，即有 $A_n^n=\underline{n\times(n-1)\times(n-2)\times\cdots\times3\times2\times1}$.

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 个元素的全排列公式可以写成 $A_n^n=\underline{n!}$.

规定 $0!=\underline{1}$.

概念辨析

1.$90\times91\times92\times\cdots\times100$ 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.

2.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).

3.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)证明下列排列数的性质：

性质 1：$A_n^m=nA_{n-1}^{m-1}$；

性质 2：$A_n^m=mA_{n-1}^{m-1}+A_{n-1}^m$.

证明：性质 1：从 n 个不同元素中，任取 $m(m\leqslant n)$ 个元素，分两步完成：

第 1 步：从 n 个不同元素中选出 1 个元素排在第一个位置，有 n 种方法；

第 2 步：从余下的 $(n-1)$ 个元素中选出 $(m-1)$ 个元素排在余下的 $(m-1)$ 个位置上，有 A_{n-1}^{m-1} 种方法.

由分步乘法计数原理得 $A_n^m=nA_{n-1}^{m-1}$.

(同理还可得 $A_{n-1}^{m-1}=(n-1)A_{n-2}^{m-2}=(n-1)(n-2)\cdot A_{n-3}^{m-3}$)

性质 2：从 n 个不同元素中，任取 $m(m\leqslant n)$ 个元素分两类完成：

第一类：m 个元素中含有 a(a 为 n 个不同元素中的一个)，分两步完成：

①将 a 排在某一个位置上，有 m 种方法；

②从其余的 $(n-1)$ 个元素中选出 $(m-1)$ 个元素排在余下的 $(m-1)$ 个位置上，有 A_{n-1}^{m-1} 种方法.

即有 mA_{n-1}^{m-1} 种不同的方法.

第二类：m 个元素中不含有 a，即从 $(n-1)$ 个元素中选出 m 个元素排在 m 个位置上，有 A_{n-1}^m 种方法.

由分类加法计数原理得 $A_n^m=mA_{n-1}^{m-1}+A_{n-1}^m$.

任务型课堂

学习任务一　排列数的定义及展开式的应用

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

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

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

2.(1)已知$A_x^2=20$，则$x=$________.

(2)用排列数表示$(55-n)(56-n)\cdots(69-n)(n\in\mathbf{N}^*$且$n<55)$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$55-n,56-n,\cdots,69-n$中的最大数为$69-n$，且共有$(69-n)-(55-n)+1=15$(个)数，

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

反思提炼

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.

学习任务二　排列数公式阶乘式的应用

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

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$x=6$.

反思提炼

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，同时要注意其中的隐含条件($m\leqslant n$且$n\in\mathbf{N}^*$，$m\in\mathbf{N}^*$).

学习任务三　排列数公式在实际问题中的应用

例　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.如果其中的甲不站最左边，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；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，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，其余人全排列，共有$A_6^1A_6^1A_6^6$种.由分类加法计数原理得，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，余下7个位置全排列，有A_7^7种，但应剔除乙在最右边时的$A_6^1A_6^6$种，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.其中，不符合条件的有甲在最左边时的A_7^7种，乙在最右边时的A_7^7种，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，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. 如果女生全排在一起且男生全排在一起，有多少种不同的排法？

解： 不同的排法有 $A_3^3A_5^5A_2^2=6\times120\times2=1\ 440$（种）.

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 $\frac{A_8^8}{A_3^3}=6\ 720$（种）.

反思提炼

排队问题的解题策略

排队问题除涉及特殊元素、特殊位置外，还往往涉及相邻、不相邻、定序等问题.

(1)对于"在"与"不在"问题，可采用"特殊元素优先考虑，特殊位置优先安排"的原则解决.

(2)对于相邻问题，可采用"捆绑法"解决，即将相邻的元素视为一个整体进行排列.

(3)对于不相邻问题，可采用"插空法"解决，即先排其余的元素，再将不相邻的元素插入空中.

(4)对于定序问题，可采用"除阶乘法"解决，即用不限制的排列数除以顺序一定元素的全排列数.

探究训练

1. 6 名同学排成一排，其中的甲、乙两人必须在一起的不同的排法共有　　（　　）

A. 720 种　　B. 360 种

C. 240 种　　D. 120 种

A　**解析：** C　**解析：** 甲、乙两人要排在一起，故将甲、乙两人"捆"在一起视作一人，与其余四人全排列共有 A_5^5 种排法，甲、乙两人之间有 A_2^2 种排法.

由分步乘法计数原理知，共有 $A_5^5A_2^2=240$（种）不同的排法.

2. 若 5 名成人带 2 个小孩排队上山，小孩不排在一起也不排在头尾，则不同的排法种数为　　（　　）

A. $A_5^5A_4^2$　　B. $A_5^5A_5^2$

C. $A_5^5A_6^2$　　D. $A_7^7-4A_6^6$

A　**解析：** 首先 5 名成人先排队，共有 A_5^5 种排法，然后把 2 个小孩插进中间的 4 个空中，共有 A_4^2 种排法，根据分步乘法计数原理，共有 $A_5^5A_4^2$ 种排法.

3. 已知 7 人站成一排.

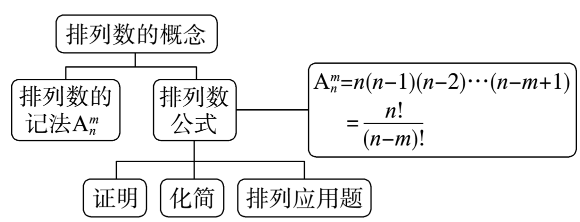
(1)其中的甲必须在乙的左边（不一定相邻），则有多少种不同的排列方法？

(2)其中的甲、乙、丙三人自左向右的顺序不变（不一定相邻），则有多少种不同的排列方法？

解：(1)甲、乙两人的顺序有 A_2^2 种情况，甲在乙左边的排法种数占全排列种数的 $\frac{1}{A_2^2}$，故有 $\frac{A_7^7}{A_2^2}=2\ 520$（种）不同的排法.

(2)甲、乙、丙三人的顺序有 A_3^3 种情况，即甲、乙、丙自左向右的顺序不变的排法种数占全排列种数的 $\frac{1}{A_3^3}$. 故有 $\frac{A_7^7}{A_3^3}=840$（种）不同的排法.

体系构建



课后素养评价（四）

基础性·能力运用

1. 6 位选手依次演讲，其中选手甲不排在第一个也不排在最后一个，则不同的演讲次序共有　　（　　）

A. 240 种　　B. 360 种

C. 480 种　　D. 720 种

C　**解析：** 先排甲，有 4 种方法，剩余 5 人全排列，有 $A_5^5=120$（种）方法，所以不同的演讲次序有 $4\times120=480$（种）.

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 连续参加了两天社区服务，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 连续参加了两天社区服务，也各有 12 种方法.

所以恰有 1 人连续参加了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、外语、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，则不同的排课方法的种数是（　　）

A.36　　B.24　　C.18　　D.12

A　**解析**：将语文和化学捆绑，与外语、物理全排列有 $A_3^3A_2^2$ 种排法；数学不排在第一节课，则数学有 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.

4.5 人站成一排，其中甲、乙两人之间恰有 1 人的不同站法的种数为（　　）

A.18　　B.24

C.36　　D.48

C　**解析**：5 人站成一排，甲、乙两人之间恰有 1 人的不同站法有 $A_3^1A_3^3A_2^2=36$(种).

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 人入园顺序的排法种数为________.

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.故总的排法有 $2\times2\times A_3^3=24$(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.用 0，1，2，3，4，5 这六个数字组数.

(1)能组成多少个无重复数字的四位偶数？

(2)能组成多少个无重复数字且为 5 的倍数的五位数？

(3)能组成多少个比 1 325 大的无重复数字的四位数？

解：(1)符合要求的四位偶数可分为三类：

第 1 类：0 在个位时有 A_5^3 个；

第 2 类：2 在个位时，首位从 1，3，4，5 中选定 1 个有 A_4^1 种选法，十位和百位从余下的数字中选，有 A_4^2 种选法，于是共有 $A_4^1A_4^2$ 个；

第 3 类：4 在个位时，与第 2 类同理，也有 $A_4^1A_4^2$ 个.

由分类加法计数原理知，共有四位偶数 $A_5^3+A_4^1A_4^2+A_4^1A_4^2=156$(个).

(2)五位数中是 5 的倍数的数可分为两类：个位数上的数字是 0 的五位数有 A_5^4 个；个位数上的数字是 5 的五位数有 $A_4^1A_4^3$ 个.

故满足条件的五位数共有 $A_5^4+A_4^1A_4^3=216$(个).

(3)比 1 325 大的无重复数字的四位数可分为三类：第 1 类：形如 2$\square\square\square$，3$\square\square\square$，4$\square\square\square$，5$\square\square\square$的数，共 $A_4^1A_5^3$ 个；

第 2 类：形如 14$\square\square$，15$\square\square$的数，共 $A_2^1A_4^2$ 个；

第 3 类：形如 134$\square$，135$\square$的数，共 $A_2^1A_3^1$ 个.

由分类加法计数原理，知比 1 325 大的无重复数字的四位数共有 $A_4^1A_5^3+A_2^1A_4^2+A_2^1A_3^1=270$(个).

综合性・创新提升

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，则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、羽三个音阶进行排序，且徵位

于羽的左侧，有$\frac{A_3^3}{2}=3$(种)排法；再将商、角插入 4 个空中.所以共有 $3A_4^2=36$(种)排法.故选 C.

3. 3 张卡片正反面分别标有数字 1 和 2，3 和 4，5 和 7.若将 3 张卡片进行排列组成一个三位数，可以得到不同的三位数的个数为　　(　　)

A.30　　B.48　　C.60　　D.96

B　**解析**："组成三位数"这件事，分 2 步完成：第 1 步，确定排在百位、十位、个位上的卡片，即为 3 个元素的全排列，共有 A_3^3 种排法；第 2 步，分别确定百位、十位、个位上的数字，各有 2 种情况.根据分步乘法计数原理，可以得到 $A_3^3\times2\times2\times2=48$(个)不同的三位数.

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 $n=\frac{31}{m}-\frac{m-1}{2}$.

因为 $m>1,n>0$，所以$\frac{31}{m}>\frac{m-1}{2}$，所以 $m^2-m-62<0$，解得 $1<m<\frac{1+\sqrt{249}}{2}$，即 $1<m\leqslant8$.

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.若其中产品 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 A_4^4 种摆法，而 A，B 可交换位置，所以有 $2A_4^4=48$(种)摆法.又当 A，B 相邻且 A，C 相邻，有 $2A_3^3=12$(种)摆法.故满足条件的摆法有 $48-12=36$(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**：1，3，5，7 四个数有 $A_4^4=24$(种)排法，故 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."对乙说："你当然不会是最差的."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.

剩下三人全排列共有 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.

8. 从 1～9 这 9 个数字中取出 5 个不同的数字进行排列.

(1)奇数位置上是奇数的排法有多少种？

(2)取出的奇数必须排在奇数位置上的排法有多少种？

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，故先排偶数位，排法为 A_4^2 种，余下的 2 个偶数与 5 个奇数全可排在奇数位置上，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、树状图法列出简单的组合.

问题式预习

知识清单

知识点　组合的相关概念

(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，求共有多少种选法，属于组合问题.　(√)

(4)从全班同学中选 2 名同学担任班长、学习委员两个职位，求共有多少种选法，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

学习任务一　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；若求这 2 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；两个数的积的大小与两个数的顺序无关，属于组合问题.

2.判断下列各问题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 个人中选出 3 个代表去开会，有多少种选法？

(5)从 10 个人中选出 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，所以是组合问题.

(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，所以是组合问题.

(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，所以是一个排列问题.

(4)因为 3 个代表之间没有顺序的区别，所以这是一个组合问题.

(5)因为 3 个人中，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.

反思提炼

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 $m(m\leqslant n)$ 个不同的元素即可.只要两个组合中的元素完全相同就是相同的组合，不用管元素的顺序.

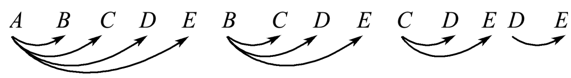
(2)判断组合与排列的依据是看是否与顺序有关，与顺序有关的是排列问题，与顺序无关的是组合问题.

学习任务二 写出简单问题的所有组合

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.

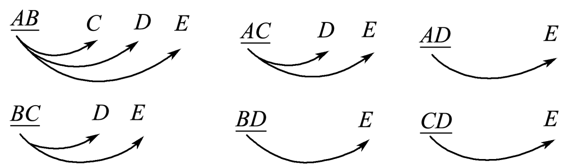
(2)从 5 个不同的元素 A,B,C,D,E 中取出 3 个,列出所有的组合.

解:(1)如图所示.



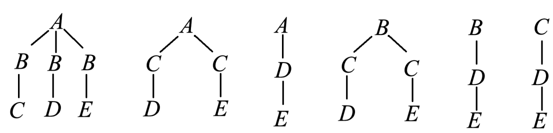
故所有的组合为 $AB,AC,AD,AE,BC,BD,BE,CD,CE,DE$.

(2)方法一(顺序后移法):可按 $AB\to AC\to AD\to BC\to BD\to CD$ 的顺序写出,即



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:$ABC,ABD,ABE,ACD,ACE,ADE,BCD,BCE,BDE,CDE$.

[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.

反思提炼

1.写出从 n 个不同元素中选出 $m(m\leqslant n)$ 个元素的所有组合的方法:从 n 个不同元素中选出 m 个元素的组合,可借助“顺序后移法”或“树状图法”,直观地写出组合,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.

探究训练

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:

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

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

$ab1,ab2,ab3,ab4,ab5$,共 25 种.

3.甲、乙、丙、丁 4 支足球队举行单循环赛.

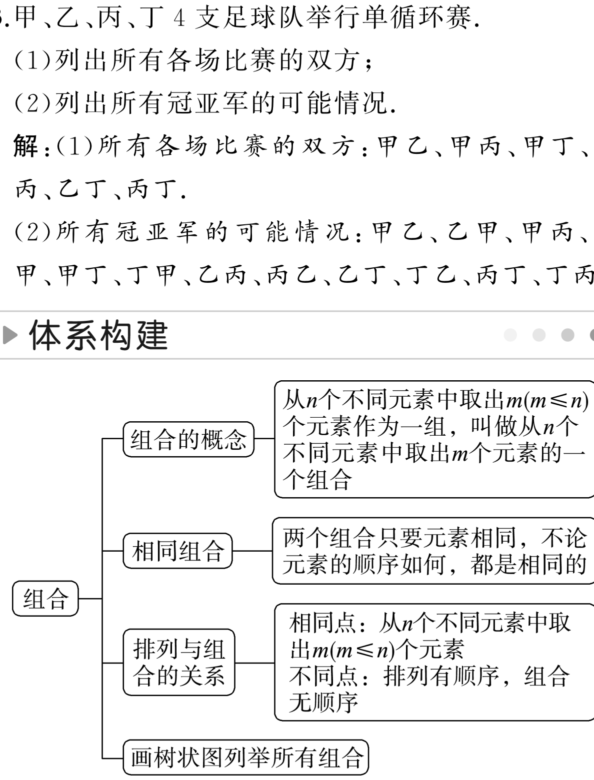
(1)列出所有各场比赛的双方;

(2)列出所有冠亚军的可能情况.

解:(1)所有各场比赛的双方:甲乙、甲丙、甲丁、乙丙、乙丁、丙丁.

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

体系构建



课后素养评价（五）

基础性·能力运用

1. 判断(正确的打"√",错误的打"×").

(1)组合的特性是元素的无序性,即取出的 m 个元素不讲究顺序,亦即元素没有位置的要求.（√）

(2)两个组合中的元素完全相同,顺序也完全一样时,两个组合才是相同的组合.（　）

×　**解析**:组合与顺序无关,只要两个组合的元素完全相同,就是相同的组合.

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,共有多少种不同的选法?

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,共有多少种不同的选法?

③从 1,2,3,…,9 中任取两个数求积,共有多少种不同的结果?

④从 1,2,3,…,9 中任取两个数求差或商,共有多少种不同的结果?

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.因为乘法满足交换律,是组合问题.对于④,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 人分别参加导游和翻译的工作,求所有选法的种数

B.从 1,2,3,4 这 4 个数字中选取 3 个不同的数字组成三位数,求组成的三位数的个数

C.从全班同学中选出 3 名同学参加学校运动会开幕式,求所有选法的种数

D.将甲、乙两位同学安排到 A,B 两个座位,求安排方法的种数

C　**解析**:对于 A 选项,从 4 名志愿者中选出 2 人分别参加导游和翻译的工作,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 选项,从 1,2,3,4 这 4 个数字中选取 3 个不同的数字排成一个三位数,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 选项,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 选项,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不同元素的所有组合;

(2)写出从 a,b,c,d,e 五个元素中任取两个不同元素的所有排列.

解:(1)从 5 个元素 a,b,c,d,e 中任取两个不同元素的所有组合:

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

(2)从 5 个元素 a,b,c,d,e 中任取两个不同元素的所有排列:

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

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点为端点的有向线段;

(2)试写出以其中任意两点为端点的线段;

(3)试写出以其中任意三点为顶点的三角形.

解:(1)以其中任意两点为端点的有向线段有 $\overrightarrow{AB}$, $\overrightarrow{AC}$, $\overrightarrow{AD}$, $\overrightarrow{BC}$, $\overrightarrow{BD}$, $\overrightarrow{CD}$, $\overrightarrow{BA}$, $\overrightarrow{CA}$, $\overrightarrow{DA}$, $\overrightarrow{CB}$, $\overrightarrow{DB}$, $\overrightarrow{DC}$.

(2)以其中任意两点为端点的线段有 AB,AC,AD,BC,BD,CD.

(3)以其中任意三点为顶点的三角形有 $\triangle ABC$, $\triangle ABD$, $\triangle BCD$, $\triangle ACD$.

综合性·创新提升

1. 给出下列问题:

①从甲、乙、丙 3 名同学中选出 2 名分别去参加两个乡镇的社会调查,有多少种不同的选法?

②有 4 张同样的电影票,要分给 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,且命中的 4 枪均为 2 枪

连中，则不同的情况有多少种？

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，故 A 错误；

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，故 B 正确；

对于 C，因为是相互发一条短信，与顺序有关，故是排列问题，故 C 错误；

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 D 正确.

故选 BD.

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

$\frac{6}{35}$　**解析**：从正方体的 8 个顶点中任选 4 个，有 $n=70$ 个结果，

这 4 个点在同一个平面内的有 $m=6+6=12$(个).

故所求概率 $P=\frac{m}{n}=\frac{12}{70}=\frac{6}{35}$.

5.判断下列问题分别是排列问题还是组合问题.

(1)从 10 名学生中任选 5 名去参观一个展览会，求有多少种不同的选法；

(2)从 1，2，3，4，5 这 5 个数中，每次任取 2 个不同的数作为一个点的坐标，求所有不同点的个数；

(3)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，求这两张卡片上的数相加所得的和有多少种；

(4)有四本不同的书要分别送给四个人，每人一本，求一共有多少种不同的送法.

解：(1)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，故这是组合问题.

(2)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，故这是排列问题.

(3)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，所以选出的卡片不用排序，故这是组合问题.

(4)因为四本不同的书分别送给四个人，要求每人一本，

所以这四本书需要排序，故这是排列问题.

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应用问题.

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) $C_n^m=\dfrac{\underline{A_n^m}}{\underline{A_m^m}}=\dfrac{n(n-1)(n-2)\cdots(n-m+1)}{m!}$，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. 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 $n=8$.

2. $C_3^3+C_4^3+C_5^3=$ （　　）

A.C_5^4　　B.C_6^5

C.C_6^3　　D.C_6^4

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

3. 设集合 $A=\{a_1,a_2,a_3,a_4,a_5\}$，则集合 A 含有3个元素的子集共有________个.

10　**解析**：从集合 A 的5个元素中选出3个元素构成一个子集，是组合问题，其个数为 $C_5^3=\dfrac{5\times4\times3}{3\times2\times1}=10$.

4. 若 $C_8^{2x-1}=C_8^{x+3}$，则 x 的值为________.

2或4　**解析**：由 $C_8^{2x-1}=C_8^{x+3}$ 得 $2x-1=x+3$ 或 $2x-1+x+3=8$，解得 $x=4$ 或 $x=2$.

5. 请思考并回答下列问题：

(1)某班级将在月底参加行一场篮球比赛，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

学习任务一　组合数的计算与证明

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

2. 证明：$mC_n^m=nC_{n-1}^{m-1}$.

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

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

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

反思提炼

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.

学习任务二　有限制条件的组合问题

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，有多少种选法？

解：至少有一名队长含有两种情况：有一名队长和两名队长，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”，有多少种选法？

解：至多有两名女生含有三种情况：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)选法.

思考 2. 既要有队长当选，又要有女生当选，有多少种选法？

解：分两种情况：

第一种：女队长当选，有 C_{12}^4 种选法；

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)选法.

反思提炼

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”含有几个元素的问题

“至多”“至少”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.

探究训练

1. (2023·新高考全国Ⅰ卷)某学校开设了 4 门体育类选修课和 4 门艺术类选修课，学生需从这 8 门课中选修 2 门或 3 门课，并且每类选修课至少选修 1 门，则不同的选课方案共有______种.(用数字作答)

64　**解析**：因为每类选修课至少选 1 门，所以若从 8 门课中选修 2 门，则不同的选课方案共有 $C_4^1C_4^1=16$(种).

当从 8 门课中选修 3 门时，

①若体育类选修课选修 1 门，则不同的选课方案共有 $C_4^1C_4^2=24$(种)；

②若体育类选修课选修 2 门，则不同的选课方案共有 $C_4^2C_4^1=24$(种).

综上所述，不同的选课方案共有 $16+24+24=64$(种).

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)其中甲、乙、丙 3 人必须参加；

(3)其中甲、乙、丙 3 人不能参加；

(4)其中甲、乙、丙 3 人只能有 1 人参加.

解：(1)有 $C_{12}^5=792$(种)不同的选法.

(2)甲、乙、丙 3 人必须参加，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 3 人不能参加，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 3 人只能有 1 人参加，分两步：第 1 步，从甲、乙、丙中选 1 人，有 C_3^1 种选法；第 2 步，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.

学习任务三　分组分配问题

例 2　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，有多少种不同的分法？

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、丙三名同学，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 $x=\dfrac{C_6^2C_4^2C_2^2}{A_3^3}=15$.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 1 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 2 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.

反思提炼

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 组均匀，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 个不同的对象($n\geqslant m$)，有 C_{n-1}^{m-1} 种分法，可描述为($n-1$)个空中插入($m-1$)块板.

探究训练

1.学校要安排甲、乙等 6 名志愿者去 A，B，C，D 四个比赛场地服务，要求每个比赛场地都有人去，每人只能去一个比赛场地，则甲、乙两人被分在同一个比赛场地的安排方法种数为　(　　)

A.120　B.240　C.360　D.480

B　**解析**：将 6 人按 3，1，1，1 分成四组，且甲、乙在同一组的安排方法有 C_4^1 种，

将 6 人按 2，2，1，1 分成四组，且甲、乙在同一组的安排方法有 C_4^2 种，

则甲、乙两人被分在同一个比赛场地的安排方法种数为 $(C_4^1+C_4^2)A_4^4=240$.

故选 B.

2.某宾馆安排 A，B，C，D，E 5 人入住 3 个不同的房间，每个房间至少住 1 人，且 A，B 不能住同一房间，则不同的安排方法有　(　　)

A.24 种　B.48 种

C.96 种　D.114 种

D　**解析**：5 个人住 3 个房间，每个房间至少住 1 人，则有(3，1，1)和(2，2，1)两种情况.当为(3，1，1)时，有 $C_5^3A_3^3=60$(种)，A，B 住同一房间的情况有 $C_3^1A_3^3=18$(种)，故有 $60-18=42$(种)；当为(2，2，1)时，有 $\dfrac{C_5^2C_3^2}{A_2^2}\cdot A_3^3=90$(种)，A，B 住同一房间的情况有 $C_3^1C_3^2A_2^2=18$(种)，故有 $90-18=72$(种).根据分类加法计数原理，共有 $42+72=114$(种)不同的安排方法.

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 5 个空隙中任选 3 个空隙各插一块隔板，故有 $C_5^3=10$(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|;

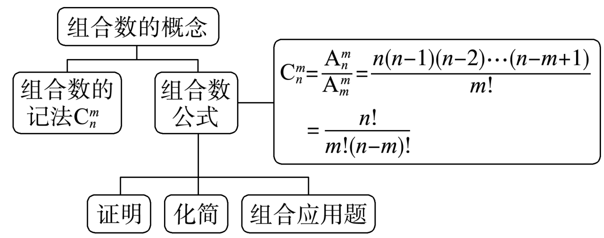
第 2 步,将剩下的两块隔板插入形成空盒,此时含有两种情况.

①这两块隔板与前面三块形成不相邻的两个空盒,如||00||0000|,有 C_3^2 种插法.

②将两块隔板与前面三块中的一块并放,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.

体系构建



课后素养评价(六)

基础性·能力运用

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 $n\geqslant 3$,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.即共有 $C_3^3C_6^3+C_2^1C_3^2C_5^3+C_3^1C_4^3=92$(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**:由题易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. (新定义)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 $C_{12}^2=\frac{12\times 11}{2}=66$(个)样本点,

其中两个素数和为 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. 方程 $C_n^2-n=5$ 的解为________；不等式 $C_n^2-n<5$ 的解集为________.

5　$\{2,3,4\}$　**解析**：由 $C_n^2-n=5$，得 $\frac{n(n-1)}{2}-n=5$，所以 $n^2-3n-10=0$，解得 $n=5$ 或 $n=-2$（舍去）.

由 $C_n^2-n<5$ 得 $n^2-3n-10<0$，解得 $-2<n<5$.

由题设条件知 $n\geqslant2$，且 $n\in\mathbf{N}^*$，

所以 $n=2,3,4$.故原不等式的解集为 $\{2,3,4\}$.

7. (1)计算 C_{100}^{97}；

(2)求满足等式 $C_{n+3}^{n+1}=C_{n+1}^{n-1}+C_n^{n-2}+C_{n+1}^n$ 的正整数 n.

解：(1)由性质 $C_n^m=C_n^{n-m}$，得 $C_{100}^{97}=C_{100}^3=\frac{100\times99\times98}{3\times2\times1}=161\ 700$.

(2)由性质 $C_n^m=C_n^{n-m}$ 和 $C_{n+1}^m=C_n^m+C_n^{m-1}$，

得 $C_{n+3}^{n+1}=C_{n+3}^2$，$C_{n+1}^{n-1}+C_n^{n-2}+C_{n+1}^n=C_{n+1}^2+C_n^2+C_{n+1}^1=C_{n+2}^2+C_n^2$.

所以已知等式可化为 $C_{n+3}^2=C_{n+2}^2+C_n^2$.

又由性质 $C_{n+1}^m=C_n^m+C_n^{m-1}$，得 $C_{n+3}^2=C_{n+2}^2+C_{n+2}^1$，

所以 $C_{n+2}^2+C_{n+2}^1=C_{n+2}^2+C_n^2$，则 $C_{n+2}^1=C_n^2$.

所以 $n+2=\frac{n(n-1)}{2}$，即 $n^2-3n-4=0$，解得 $n=4$ 或 $n=-1$（舍去）.

故 n 的值为 4.

8. 已知平面 α∥平面 β，在平面 α 内有 4 个点，在平面 β 内有 6 个点.

(1)过这 10 个点中的 3 个点作一平面，最多可作多少个不同的平面？

(2)以这些点为顶点，最多可作多少个三棱锥？

(3)(2)中最多可以有多少个不同体积的三棱锥？

解：(1)所作出的平面有三类.

①α 内 1 点，β 内 2 点确定的平面，最多有 $C_4^1C_6^2$ 个；

②α 内 2 点，β 内 1 点确定的平面，最多有 $C_4^2C_6^1$ 个；

③α，β 本身，有 2 个.

故所作的平面最多有 $C_4^1C_6^2+C_4^2C_6^1+2=98$（个）.

(2)所作的三棱锥有三类.

①α 内 1 点，β 内 3 点确定的三棱锥，最多有 $C_4^1C_6^3$ 个；

②α 内 2 点，β 内 2 点确定的三棱锥，最多有 $C_4^2C_6^2$ 个；

③α 内 3 点，β 内 1 点确定的三棱锥，最多有 $C_4^3C_6^1$ 个.

故最多可作出的三棱锥有 $C_4^1C_6^3+C_4^2C_6^2+C_4^3C_6^1=194$（个）.

(3)当等底面积、等高时，三棱锥的体积相等.所以体积不相同的三棱锥最多有 $C_6^3+C_6^2C_4^2+C_4^3=114$（个）.

故最多有 114 个不同体积的三棱锥.

综合性·创新提升

1. 某学校安排小明和小李等 5 名志愿者将两个不同的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 $3\times2\times2=12$（种）方案.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，则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. 从6男2女共8名学生中选出队长1人，副队长1人，普通队员2人组成4人服务队，要求服务队中至少有1名女生，共有________种不同的选法.(用数字作答)

660　**解析**：由题意可知，只选1名女生的选法有$C_2^1\cdot C_6^3\cdot C_4^1\cdot C_3^1=480$(种)，选2名女生的选法有$C_6^2\cdot C_4^1\cdot C_3^1=180$(种).所以选法共有$480+180=660$(种).

5. 在创建文明卫生城市期间，某居委会从辖区内甲、乙、丙三个小区中选取6人做志愿者，协助宣传工作.若每个小区至少选取1人做志愿者，则不同的选取方法有________种.

540　**解析**：①当6个人按各小区人数分为2,2,2三个小组时，共有$C_6^2C_4^2C_2^2=90$(种)选法；

②当6个人分为4,1,1三个小组时，共有$C_3^1C_6^4C_2^1C_1^1=90$(种)选法；

③当6个人分为3,2,1三个小组时，共有$C_6^3C_3^2C_1^1A_3^3=360$(种)选法.

综上，不同的选法共有$90+90+360=540$(种).

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$C_8^4=\dfrac{8\times7\times6\times5}{4\times3\times2\times1}=70$(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$C_{10}^4=\dfrac{10\times9\times8\times7}{4\times3\times2\times1}=210$(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、反面分别写有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$C_4^1C_2^1C_3^1\times2^2=96$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.能用计数原理证明二项式定理.

2.掌握二项式定理及二项展开式的通项公式.

3.能解决与二项式定理有关的简单问题.

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<u>$n+1$</u>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　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)$(a+b)^n$的展开式中，二项式系数C_n^k只与指数n及项数k有关，与a，b无关，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，它与指数n，项数k及a，b都有关.

概念辨析

1.判断(正确的打“√”，错误的打“×”).

(1)$(a+b)^n$的展开式共有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$(a+b)^n$的展开式的第k项.　(×)

2.若$\left(x-\dfrac{2}{\sqrt{x}}\right)^n$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$\left(x-\dfrac{2}{\sqrt{x}}\right)^n$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.$(\sqrt{2}x-3)^5$的展开式的第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各项的次数和等于n.②字母a按降幂排列，从第一项起，次数由n逐项减1直到0；字母b按升幂排列，从第一项起，次数由0逐项加1直到n.

任务型课堂

学习任务一　二项式定理的正用与逆用

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

反思提炼

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：$(a+b)^n$的展开式中，a按降幂排列，b按升幂排列.$(a-b)^n$的展开式中各项会出现正负间隔的情况.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.

探究训练

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

x^5-1　**解析**：原式$=C_5^0(x-1)^5+C_5^1(x-1)^4+C_5^2(x-1)^3+C_5^3(x-1)^2+C_5^4(x-1)^1+C_5^5(x-1)^0-1=[(x-1)+1]^5-1=x^5-1$.

2. 求$\left(3\sqrt{x}+\dfrac{1}{\sqrt{x}}\right)^4$的展开式.

解：方法一：$\left(3\sqrt{x}+\dfrac{1}{\sqrt{x}}\right)^4=(3\sqrt{x})^4+C_4^1(3\sqrt{x})^3\cdot\left(\dfrac{1}{\sqrt{x}}\right)+C_4^2(3\sqrt{x})^2\left(\dfrac{1}{\sqrt{x}}\right)^2+C_4^3(3\sqrt{x})\left(\dfrac{1}{\sqrt{x}}\right)^3+C_4^4\left(\dfrac{1}{\sqrt{x}}\right)^4=81x^2+108x+54+\dfrac{12}{x}+\dfrac{1}{x^2}$.

方法二：$\left(3\sqrt{x}+\dfrac{1}{\sqrt{x}}\right)^4=\left(\dfrac{3x+1}{\sqrt{x}}\right)^4=\dfrac{1}{x^2}(1+3x)^4$

$=\dfrac{1}{x^2}[1+C_4^1\cdot 3x+C_4^2(3x)^2+C_4^3(3x)^3+C_4^4(3x)^4]$

$=\dfrac{1}{x^2}(1+12x+54x^2+108x^3+81x^4)$

$=\dfrac{1}{x^2}+\dfrac{12}{x}+54+108x+81x^2$.

学习任务二　求展开式中的特定项

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 $T_{k+1}=C_n^kx^{\frac{n-k}{3}}(-3)^kx^{-\frac{k}{3}}=C_n^k\cdot(-3)^kx^{\frac{n-2k}{3}}$.

(2)若第 6 项为常数项，则 $k=5$ 时，有$\dfrac{n-2k}{3}=0$，即$n=10$.

[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：由例 2(2)知 $T_{k+1}=C_{10}^k(-3)^kx^{\frac{10-2k}{3}}$.

令$\dfrac{10-2k}{3}=2$，得 $k=2$，

所以含 x^2 的项的系数为 $C_{10}^2(-3)^2=405$.

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$\begin{cases}\dfrac{10-2k}{3}\in\mathbf{Z},\\ 0\leqslant k\leqslant 10,\\ k\in\mathbf{Z}.\end{cases}$

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 $k=2,5,8$，

故二项展开式中的所有有理项为

$T_3=C_{10}^2(-3)^2x^2=405x^2$，

$T_6=C_{10}^5(-3)^5=-61\ 236$，

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

反思提炼

求二项展开式中的特定项的常见题型及解题策略

(1)求第 k 项.利用二项展开式的通项可得 $T_k=C_n^{k-1}a^{n-k+1}b^{k-1}$.

(2)求常数项.常数项的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.有理项中字母的指数是整数.求有理项时,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.整式项中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.

探究训练

1. 若$\left(x^6+\frac{1}{x\sqrt{x}}\right)^n(n\in\mathbf{N}^*)$的展开式中含有常数项,则 n 的最小值等于　(　　)

A.3　　B.4　　C.5　　D.6

C　**解析**:$\left(x^6+\frac{1}{x\sqrt{x}}\right)^n$的展开式的通项为 $T_{k+1}=C_n^k(x^6)^{n-k}\cdot\left(\frac{1}{x\sqrt{x}}\right)^k=C_n^kx^{6n-6k-\frac{3}{2}k}=C_n^kx^{6n-\frac{15}{2}k}$.

令 $6n-\frac{15}{2}k=0$,得 $n=\frac{5}{4}k$.又 $n\in\mathbf{N}^*$,所以当 $k=4$ 时,n 取得最小值 5.

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

$\frac{1}{2}$　**解析**:根据二项展开式的通项可得

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

令 $8-\frac{4k}{3}=4$,可得 $k=3$,$C_8^ka^k=C_8^3a^3=7$,

解得 $a=\frac{1}{2}$.

3. 已知$\left(2x-\frac{1}{\sqrt{x}}\right)^n(n\in\mathbf{N}^*)$的展开式的第 2 项与第 3 项的二项式系数之比是 2∶5.

(1)求 n 的值;

(2)求展开式的常数项.

解:(1)由展开式的第 2 项与第 3 项的二项式系数之比是 2∶5,可得 $C_n^1∶C_n^2=2∶5$,解得 $n=6$.

(2)$\left(2x-\frac{1}{\sqrt{x}}\right)^n$的展开式的通项为

$T_{k+1}=C_n^k(2x)^{n-k}\left(-\frac{1}{\sqrt{x}}\right)^k=(-1)^k2^{n-k}C_n^kx^{n-\frac{3}{2}k}$.

令 $6-\frac{3}{2}k=0$,解得 $k=4$,

所以展开式的常数项为 $(-1)^4\times2^{6-4}\times C_6^4=60$.

4. (1)求$\left(2x^2-\frac{1}{\sqrt[3]{x}}\right)^8$的展开式中,第 5 项的二项式系数及第 5 项的系数;

(2)求$\left(2x^2-\frac{1}{\sqrt[3]{x}}\right)^8$的展开式中,$x^2$ 的系数.

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

(1)$T_5=(-1)^4C_8^4\cdot2^4\cdot x^{\frac{20}{3}}$,

所以第 5 项的二项式系数是 $C_8^4=70$,

第 5 项的系数是 $C_8^4\cdot2^4=1\ 120$.

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

因此,x^2 的系数是 $(-1)^6C_8^6\cdot2^{8-6}=112$.

学习任务三　二项式定理的灵活运用

例 3　利用二项式定理解决下列关于 $(a+b+c)^n$ 的展开式的问题.

(1)用两种方法求 $(x^2+3x+2)^5$ 的展开式中 x 的系数;

(2)求$\left(x+\frac{1}{x^2}-1\right)^4$的展开式中的常数项.

解:(1)方法一:$(x^2+3x+2)^5$ 即 5 个 (x^2+3x+2) 相乘,每个因式各提供一项相乘得到展开式中的一项,含 x 的项可由 1 个因式提供 $3x$,4 个因式提供 2 得到,即 $C_5^1 3x\cdot C_4^4\cdot2^4=240x$,所以 $(x^2+3x+2)^5$ 的展开式中 x 的系数为 240.

方法二:$(x^2+3x+2)^5=(1+x)^5(2+x)^5$.

因为 $(1+x)^5$ 的展开式的通项为 $T_{r+1}=C_5^rx^r$,$(2+x)^5$ 的展开式的通项为 $T_{k+1}=C_5^k2^{5-k}x^k$,所以展开式中 x 的系数是 $C_5^0C_5^12^4+C_5^1C_5^02^5=240$.

(2)将 $x+\frac{1}{x^2}$ 看成一个整体,$\left(x+\frac{1}{x^2}-1\right)^4$的展开式的通项为 $T_{k+1}=C_4^k\left(x+\frac{1}{x^2}\right)^{4-k}(-1)^k$.

$\left(x+\frac{1}{x^2}\right)^{4-k}$的展开式的通项为 $T_{m+1}=C_{4-k}^mx^{4-k-m}\cdot x^{-2m}=C_{4-k}^mx^{4-k-3m}(m\leqslant4-k)$.

取 $4-k-3m=0$,

当 $m=0$ 时,$k=4$,$C_4^4\times C_0^0\times(-1)^4=1$.

当 $m=1$ 时，$k=1$，$C_4^1\times C_3^1\times(-1)^1=-12$.

因此，常数项为 $1-12=-11$.

反思提炼

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：由 $(a+b+c)^n$ 为 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)二项式的正整数幂的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.

探究训练

1. $(x+2y+z)^5$ 的展开式中 x^2y^2z 的系数为(　　)

A.30　　B.40

C.60　　D.120

D　**解析**：展开式中含 x^2y^2z 的项可以看成 5 个 $(x+2y+z)$ 中，2 个提供 x，2 个提供 $2y$，一个提供 z，即得到 $C_5^2\cdot x^2\cdot C_3^2\cdot(2y)^2z=120x^2y^2z$. 所以展开式中 x^2y^2z 的系数为 120.

2. 设 $a\in\mathbf{Z}$，且 $0\leqslant a<13$. 若 $51^{2\,022}+a$ 能被 13 整除，则 $a=$(　　)

A.0　　B.1　　C.11　　D.12

D　**解析**：由于 $51=52-1$，$(52-1)^{2\,022}=C_{2\,022}^0\cdot 52^{2\,022}-C_{2\,022}^1 52^{2\,021}+\cdots-C_{2\,022}^{2\,021}52+1$，又 13 能整除 52，所以只需 13 能整除 $1+a$，$0\leqslant a<13$，$a\in\mathbf{Z}$，所以 $a=12$.

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

A.15　　B.20

C.30　　D.35

C　**解析**：$(1+x)^6$ 的展开式的通项为 $T_{k+1}=C_6^k x^k$，所以 $\left(1+\dfrac{1}{x^2}\right)(1+x)^6$ 的展开式中 x^2 的系数为 $1\times C_6^2+1\times C_6^4=30$. 故选 C.

4. 71^{72} 除以 100 的余数是________.

41　**解析**：$71^{72}=(70+1)^{72}=C_{72}^0 70^{72}+C_{72}^1 70^{71}+\cdots+C_{72}^{71}70+C_{72}^{72}=m\cdot 10^2+72\times 70+1=m\cdot 10^2+5\,041(m\in\mathbf{N})$，即 71^{72} 除以 100 的余数为 41.

体系构建

二项式定理
- 二项式定理：$(a+b)^n=C_n^0a^n+C_n^1a^{n-1}b+\cdots+C_n^ka^{n-k}b^k+\cdots+C_n^nb^n\ (n\in\mathbf{N}^*)$
  - 展开式的特点
  - 正用
  - 逆用
  - 综合应用
- 通项 $T_{k+1}=C_n^ka^{n-k}b^k$，$k=0,1,2,\cdots,n$ — 求特定项(或项的系数)
  - 第 k 项
  - 含 a^r(或 a^pb^q)的项
  - 常数项
  - 有理项
- 应用
  - 整除问题
  - 余数问题
  - 近似计算问题

课后素养评价（七）

基础性·能力运用

1.$(x+1)^5$ 的展开式中 x 的系数为　（　）

A.1　B.5　C.10　D.15

B　**解析**：$(x+1)^5$ 的展开式的通项为 $T_{k+1}=C_5^k\cdot x^{5-k}$.令 $5-k=1$，得 $k=4$.因此，$(x+1)^5$ 的展开式中 x 的系数为 $C_5^4=5$.故选B.

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，可知多项式的每一项都可看作 $C_5^k(2x+1)^{5-k}(-1)^k(k=0,\cdots,5)$，为 $[(2x+1)-1]^5$ 的展开式中各项，化简 $[(2x+1)-1]^5=(2x)^5=32x^5$.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 $\left(x^3+\dfrac{2}{3x^2}\right)^5$ 的展开式的第3项的系数和常数项.

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

则 $T_3=C_5^2(x^3)^3\left(\dfrac{2}{3x^2}\right)^2=C_5^2\cdot\dfrac{4}{9}x^5$，

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

令 $15-5k=0$，得 $k=3$，

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

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

证明：$1+2+2^2+2^3+\cdots+2^{5n-1}=\dfrac{1-2^{5n}}{1-2}=2^{5n}-1=32^n-1=(31+1)^n-1=31^n+C_n^1\times31^{n-1}+\cdots+C_n^{n-1}\times31+1-1=31\times(31^{n-1}+C_n^1\times31^{n-2}+\cdots+C_n^{n-1})$.

显然括号内的数为正整数，故原式能被31整除.

综合性·创新提升

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 $T_{k+1}=-C_5^k(-2)^{5-k}(1-x)^k$，$a_3=-C_5^3(-2)^2=-40$.

2.$\left(x+\dfrac{y^2}{x}\right)(x+y)^5$ 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$\left(3x^3-\frac{1}{2x^4}\right)^n$的展开式中存在常数项，则正整数$n$的最小值为　　（　　）

A.5　　B.6　　C.7　　D.14

C　**解析**：展开式的通项$T_{k+1}=C_n^k(3x^3)^{n-k}\cdot\left(-\frac{1}{2x^4}\right)^k=C_n^k3^{n-k}\cdot\left(-\frac{1}{2}\right)^kx^{3n-7k}$.令$3n-7k=0$，得$n=\frac{7k}{3}$.当$k=3$时，$n$有最小值为7.

4.(新定义)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.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**：$\left(x^2+\frac{1}{x}+1\right)^6$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.已知$m,n\in\mathbf{N}^*$，$(1+x)^m+(1+x)^n$的展开式中x的系数为19，求展开式中x^2的系数的最小值及此时展开式中x^7的系数.

解：由题设知$C_m^1+C_n^1=m+n=19$.

又$m,n\in\mathbf{N}^*$，所以$1\leqslant m\leqslant18$.

x^2的系数为$C_m^2+C_n^2=\frac{1}{2}(m^2-m)+\frac{1}{2}(n^2-n)$

$=m^2-19m+171$.

所以当$m=9$或10时，x^2的系数取得最小值81.

此时x^7的系数为$C_9^7+C_{10}^7=156$.

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(正确的打"√",错误的打"×").

(1)二项展开式中所有二项式系数的和为$C_n^1+C_n^2+\cdots+C_n^n$.　(×)

(2)二项展开式中系数最大的项与二项式系数最大的项相同.　(×)

(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:由于$n=11$为奇数,则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$x=1$,即得各项系数的和,所以各项系数的和为-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

学习任务一　二项式系数性质的简单应用

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

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)(多选)已知$\left(x-\frac{2}{\sqrt[3]{x}}\right)^n$的展开式中,有且只有第5项的二项式系数最大,则　　(　　)

A.$n=8$

B.二项展开式的各项系数和为1

C.二项展开式的二项式系数和为512

D.二项展开式中的常数项是第7项

ABD　**解析**:因为$\left(x-\frac{2}{\sqrt[3]{x}}\right)^n$的展开式中有且只有第5项的二项式系数最大,所以$n=8$,故A正确;

令$x=1$,得二项展开式的各项系数和为$(1-2)^8=1$,故B正确;

展开式的二项式系数和为$2^8=256$,故C错误;

二项展开式的通项为$T_{k+1}=C_8^k x^{8-k}\left(-\frac{2}{\sqrt[3]{x}}\right)^k=(-2)^k C_8^k x^{8-\frac{4k}{3}}$,令$8-\frac{4k}{3}=0$,解得$k=6$,所以第7项为常数项,故D正确.故选ABD.

反思提炼

二项式系数最大的项的求法

求二项式系数最大的项,根据二项式系数的性质对$(a+b)^n$中的n进行讨论.

①当n为奇数时,中间两项的二项式系数最大;

②当n为偶数时,中间一项的二项式系数最大.

探究训练

如图是与杨辉三角有类似性质的三角形数垒,a,b是某行的前两个数,当$a=7$时,$b=$　　(　　)

1

2　2

3　4　3

4　7　7　4

5　11　14　11　5

……

A.20　　B.21　　C.22　　D.23

C　**解析**:由$a=7$,可知b左肩上的数为6,右肩上的数为$11+5=16$,所以$b=6+16=22$.

学习任务二　求二项展开式的系数和

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.如何得到$a_0-a_1+a_2-\cdots-a_{2\,023}+a_{2\,024}$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$(1+2x)^{2\,024}=b_0+b_1x+b_2x^2+\cdots+b_{2\,024}x^{2\,024}$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.为了得到$a_1+2a_2+\cdots+2\,023a_{2\,023}+2\,024a_{2\,024}$的值,应如何操作?

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中的式子同侧相加,同除以2,即可得$a_0+a_2+a_4+\cdots+a_{2\,024}=\frac{3^{2\,024}+1}{2}$.

反思提炼

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$a_0+a_2+a_4+\cdots=\frac{f(1)+f(-1)}{2}$,偶数项系数之和为$a_1+a_3+a_5+\cdots=\frac{f(1)-f(-1)}{2}$.

探究训练

1.若$(1+x)(1-2x)^8=a_0+a_1x+\cdots+a_9x^9$,$x\in\mathbf{R}$,则$a_1\cdot2+a_2\cdot2^2+\cdots+a_9\cdot2^9$的值为　　(　　)

A.2^9　　B.2^9-1

C.3^9　　D.3^9-1

D　**解析**:令$x=0$,则$a_0=1$.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$\frac{a_1}{3}+\frac{a_2}{3^2}+\cdots+\frac{a_n}{3^n}$的值.

解:(1)易知$n\geqslant7$,$n\in\mathbf{N}$.因为$A_n^5=56C_n^7$,

所以$n(n-1)(n-2)(n-3)(n-4)=56\times$

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

整理可得$\frac{(n-5)(n-6)}{90}=1$，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 $x=\frac{1}{3}$，可得$\left(1-3\times\frac{1}{3}\right)^{15}=a_0+\frac{a_1}{3}+\frac{a_2}{3^2}+\cdots+\frac{a_{15}}{3^{15}}=0$.

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

3.已知$(2x-3)^4=a_0+a_1x+a_2x^2+a_3x^3+a_4x^4$，求：

(1)$a_0+a_1+a_2+a_3+a_4$ 的值；

(2)$(a_0+a_2+a_4)^2-(a_1+a_3)^2$ 的值.

解：(1)由$(2x-3)^4=a_0+a_1x+a_2x^2+a_3x^3+a_4x^4$，

令 $x=1$，得$(2-3)^4=a_0+a_1+a_2+a_3+a_4$，

所以 $a_0+a_1+a_2+a_3+a_4=1$.

(2)在$(2x-3)^4=a_0+a_1x+a_2x^2+a_3x^3+a_4x^4$ 中，

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

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

所以$(a_0+a_2+a_4)^2-(a_1+a_3)^2$

$=(a_0-a_1+a_2-a_3+a_4)(a_0+a_1+a_2+a_3+a_4)$

$=(-2-3)^4\times(2-3)^4=625$.

学习任务三 二项展开式中的系数的最值问题

例 3 已知$\left(x^{\frac{2}{3}}+3x^2\right)^n$的展开式中各项系数的和比各项的二项式系数的和大 992.

(1)求 n 的值；

(2)求展开式中二项式系数最大的项；

(3)求展开式中系数最大的项.

解：(1)令 $x=1$，可得展开式中各项系数的和为 4^n，又二项式系数的和为 2^n，各项系数的和比各项的二项式系数的和大 992，

所以 $4^n-2^n=992$，解得 $n=5$.

(2)因为 $n=5$，所以第 3 项、第 4 项的二项式系数最大.

由二项展开式的通项为 $T_{k+1}=C_5^k\cdot(x^{\frac{2}{3}})^{5-k}\cdot(3x^2)^k=C_5^k\cdot3^k\cdot x^{\frac{10+4k}{3}}$，所以 $T_3=C_5^2\cdot3^2\cdot x^{\frac{10+8}{3}}=90x^6$，$T_4=C_5^3\cdot3^3\cdot x^{\frac{10+12}{3}}=270x^{\frac{22}{3}}$.

(3)设展开式中系数最大的项是第 $k+1$ 项，

则$\begin{cases}C_5^k\cdot3^k\geqslant C_5^{k-1}\cdot3^{k-1},\\C_5^k\cdot3^k\geqslant C_5^{k+1}\cdot3^{k+1},\end{cases}$解得$\frac{7}{2}\leqslant k\leqslant\frac{9}{2}$.

又 $k\in\mathbf{N}$，所以 $k=4$，

所以展开式中系数最大的项为第 5 项，$T_5=C_5^4\cdot3^4\cdot x^{\frac{10+16}{3}}=405x^{\frac{26}{3}}$.

反思提炼

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.

探究训练

1.在$\left(x-\frac{1}{\sqrt{x}}\right)^n$的展开式中，只有第 5 项的二项式系数最大，则展开式中系数最小的项的系数为（　　）

A.-126　　B.-70

C.-56　　D.-28

C　**解析**：由题意可得 $n=8$，所以二项展开式的通项为 $T_{k+1}=C_8^kx^{8-k}\left(-\frac{1}{\sqrt{x}}\right)^k=C_8^k(-1)^kx^{8-\frac{3}{2}k}$，当 $k=3$ 或 5 时，系数 $C_8^k(-1)^k$ 最小，所以展开式中系数最小的项的系数为-56.

2.已知在$\left(\sqrt{x}-\frac{2}{\sqrt[3]{x}}\right)^n$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$[C_n^4(-2)^4]:[C_n^2(-2)^2]=56:3$，解得 $n=10$.

通项为 $T_{k+1}=C_{10}^k(\sqrt{x})^{10-k}\left(-\frac{2}{\sqrt[3]{x}}\right)^k=(-2)^kC_{10}^kx^{5-\frac{5k}{6}}$，

当 $5-\frac{5k}{6}$ 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 $r+1$ 项的系数的绝对值最大，

则$\begin{cases}C_{10}^{r}2^{r}\geqslant C_{10}^{r-1}2^{r-1},\\C_{10}^{r}2^{r}\geqslant C_{10}^{r+1}2^{r+1},\end{cases}$解得$\frac{19}{3}\leqslant r\leqslant\frac{22}{3}$.

又因为 $r\in\{1,2,3,\cdots,9\}$,

所以 $r=7$.当 $r=7$ 时,$T_8=-15\ 360x^{-\frac{5}{6}}$.

所以系数的绝对值最大的项为 $T_8=-15\ 360x^{-\frac{5}{6}}$.

(3)原式$=10+9C_{10}^{2}+81C_{10}^{3}+\cdots+9^{10-1}C_{10}^{10}$

$=\frac{9C_{10}^{1}+9^2C_{10}^{2}+9^3C_{10}^{3}+\cdots+9^{10}C_{10}^{10}}{9}$

$=\frac{C_{10}^{0}+9C_{10}^{1}+9^2C_{10}^{2}+9^3C_{10}^{3}+\cdots+9^{10}C_{10}^{10}-1}{9}$

$=\frac{(1+9)^{10}-1}{9}$

$=\frac{10^{10}-1}{9}$.

体系构建

- 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$k<\frac{n+1}{2}$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)

基础性·能力运用

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$x=-1$,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$n=7$.$(x-1)^7$的展开式中共有8项.所有项的系数和为0.偶数项的系数和为-64.展开式的中间项为第4项与第5项,$T_4=C_7^3x^4\cdot(-1)^3=-35x^4$,$T_5=C_7^4x^3(-1)^4=35x^3$.故选ABC.

3.(多选)在$\left(3x-\frac{1}{\sqrt{x}}\right)^n$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$\left(3x-\frac{1}{\sqrt{x}}\right)^n$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$T_5=C_6^4\cdot(-1)^4\cdot3^2=135$.

4.已知m为正整数,$(x+y)^{2m}$展开式的二项式系数的最大值为a,$(x+y)^{2m+1}$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$13\frac{(2m)!}{m!\ m!}=$

$7\dfrac{(2m+1)!}{m!(m+1)!}$，

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

5. 在 $(a+b)^8$ 的展开式中，二项式系数最大的项为________；在 $(a+b)^9$ 的展开式中，二项式系数最大的项为________________.

$70a^4b^4$　$126a^5b^4$ 与 $126a^4b^5$　**解析**：因为 $(a+b)^8$ 的展开式有 9 项，所以中间一项的二项式系数最大，该项为 $C_8^4a^4b^4=70a^4b^4$.

因为 $(a+b)^9$ 的展开式有 10 项，所以中间两项的二项式系数最大，这两项分别为 $C_9^4a^5b^4=126a^5b^4$，$C_9^5a^4b^5=126a^4b^5$.

6. 已知 $\left(x+\dfrac{m}{x}\right)^n$ 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 $\dfrac{35}{8}$，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

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

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

解得 $m=\pm\dfrac{1}{2}$.

(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 $\dfrac{8m-1}{m+1}\leqslant r\leqslant\dfrac{9m}{m+1}$.

由于只有第 6 项和第 7 项的系数最大，

所以 $\begin{cases}4<\dfrac{8m-1}{m+1}\leqslant 5,\\ 6\leqslant\dfrac{9m}{m+1}<7,\end{cases}$ 即 $\begin{cases}\dfrac{5}{4}<m\leqslant 2,\\ 2\leqslant m<\dfrac{7}{2}.\end{cases}$

所以 m 只能等于 2.

综合性·创新提升

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”的两个二项式系数相等，可知选 C.

4. (多选)已知 $\left(ax^2+\dfrac{1}{\sqrt{x}}\right)^{10}$ $(a>0)$ 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 $\left(x^2+\dfrac{1}{\sqrt{x}}\right)^{10}$ 的展开式的通项为 $T_{k+1}=$

$C_{10}^{k}(x^{2})^{10-k}\left(\frac{1}{\sqrt{x}}\right)^{k}=C_{10}^{k}x^{20-\frac{5}{2}k}$，

展开式中奇数项的二项式系数之和为 $\frac{1}{2}\times 1\ 024=512$，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. 若 $\left(2\sqrt{x}+\frac{a}{x}\right)^{n}$ 的展开式的各项系数的和为 1，各二项式系数的和为 128，则 $a=$________，展开式中 x^{2} 的系数为________.

-1　-448　**解析**：由题意得 $\left(2\sqrt{1}+\frac{a}{1}\right)^{n}=1$ 且 $2^{n}=128$，所以 $n=7$，$a=-1$.

所以 $\left(2\sqrt{x}+\frac{-1}{x}\right)^{7}$ 的展开式的通项为

$T_{k+1}=C_{7}^{k}(2\sqrt{x})^{7-k}\left(\frac{-1}{x}\right)^{k}=C_{7}^{k}2^{7-k}(-1)^{k}x^{\frac{7-3k}{2}}$.

令 $\frac{7-3k}{2}=2$，得 $k=1$.

所以 x^{2} 的系数为 $C_{7}^{1}2^{6}(-1)^{1}=-448$.

6. (2022·新高考全国Ⅰ卷) $\left(1-\frac{y}{x}\right)(x+y)^{8}$ 的展开式中 $x^{2}y^{6}$ 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 $x^{2}y^{6}$ 的项为 $C_{8}^{6}x^{2}y^{6}-\frac{y}{x}C_{8}^{5}x^{3}y^{5}=-28x^{2}y^{6}$.

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

7. 若 $(\sqrt{2}-x)^{10}=a_{0}+a_{1}x+a_{2}x^{2}+\cdots+a_{10}x^{10}$，则 $(a_{0}+a_{2}+\cdots+a_{10})^{2}-(a_{1}+a_{3}+\cdots+a_{9})^{2}=$________.

1　**解析**：令 $x=1$，

得 $a_{0}+a_{1}+a_{2}+\cdots+a_{10}=(\sqrt{2}-1)^{10}$；

令 $x=-1$，

得 $a_{0}-a_{1}+a_{2}-a_{3}+\cdots+a_{10}=(\sqrt{2}+1)^{10}$.

故 $(a_{0}+a_{2}+\cdots+a_{10})^{2}-(a_{1}+a_{3}+\cdots+a_{9})^{2}$

$=(a_{0}+a_{1}+a_{2}+\cdots+a_{10})(a_{0}-a_{1}+a_{2}-a_{3}+\cdots+a_{10})$

$=(\sqrt{2}-1)^{10}(\sqrt{2}+1)^{10}=1$.

8. 已知 $C_{n}^{2}=C_{n}^{3}$，且 $(1-2x)^{n}=a_{0}+a_{1}x+a_{2}x^{2}+a_{3}x^{3}+\cdots+a_{n}x^{n}$.

(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 $(1-2x)^{5}=a_{0}+a_{1}x+a_{2}x^{2}+a_{3}x^{3}+a_{4}x^{4}+a_{5}x^{5}$.

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 $(a_{0}+a_{1})+(a_{0}+a_{2})+(a_{0}+a_{3})+\cdots+(a_{0}+a_{n})=3$.

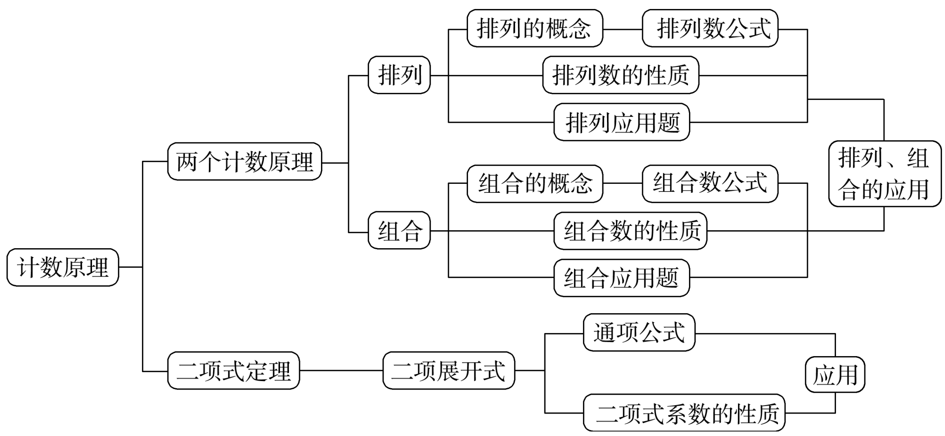
(2) 令 $x=\frac{1}{2}$，则 $0=a_{0}+\frac{a_{1}}{2}+\frac{a_{2}}{2^{2}}+\cdots+\frac{a_{5}}{2^{5}}$，所以 $\frac{a_{1}}{2}+\frac{a_{2}}{2^{2}}+\cdots+\frac{a_{5}}{2^{5}}=-1$.

(3) 对 $(1-2x)^{5}=a_{0}+a_{1}x+a_{2}x^{2}+a_{3}x^{3}+a_{4}x^{4}+a_{5}x^{5}$ 两边关于 x 求导，则 $5(1-2x)^{4}(-2)=a_{1}+2a_{2}x+3a_{3}x^{2}+4a_{4}x^{3}+5a_{5}x^{4}$. 再令 $x=1$，

可得 $a_{1}+2a_{2}+3a_{3}+\cdots+na_{n}=-10$.

单元活动研习

单元知识重构



专题任务探究

探究点一　两个计数原理的应用

例1　现有3幅不同的油画，4幅不同的国画，3幅不同的水彩画，从这些画中选一幅布置房间，则不同的选法共有（　　）

A.10种　　B.12种　　C.20种　　D.36种

A　**解析**：依题意，不同的选法共有 $3+4+3=10$（种）.

［一题多思］

思考1.将"从这些画中选一幅布置房间"改为"从这些画中选两幅布置房间，至少有一幅国画"，则不同的选法共有多少种？

解：依题意，分为两类，第一类，全是国画，有 $C_4^2=\frac{4\times3}{2\times1}=6$（种）选法；第二类，一幅国画和一幅水彩画或油画，有 $C_4^1C_6^1=24$（种）选法.由分类加法计数原理，可知共有 $6+24=30$（种）选法.

思考2.将"从这些画中选一幅布置房间"改为"从这些画中选两幅布置甲、乙两个房间，每个房间布置一幅画"，则不同的布置方法共有多少种？

解：依题意，分为两步，第一步，布置甲房间，有10种方法；第二步，布置乙房间，有9种方法.由分步乘法计数原理，可知共有 $10\times9=90$（种）布置方法.

例2　若一个三位数中，有且仅有两个数字一样，我们就把这样的三位数定义为"单重数"，例如232，114等，则不超过200的"单重数"中，从小到大排列的第22个"单重数"是（　　）

A.166　　B.171　　C.181　　D.188

B　**解析**：不超过200的三位自然数中，

两个相同数字为0时，有100，200，共2个；

两个相同数字为1时，有110，101，112，121，113，131，…，119，191，共18个；

两个相同数字为2时，有122，共1个，

同理，两个相同数字分别为3，4，5，6，7，8，9时各1个.

综上，不超过200的"单重数"共有 $2+18+8=28$（个）.

其中最大的是200，较小的依次为199，191，188，181，177，171，

故从小到大排列的第22个"单重数"为171.

反思提炼

利用两个计数原理解决应用问题的一般思路

(1)弄清完成的一件事是什么；

(2)确定是先分类后分步，还是先分步后分类；

(3)弄清分步、分类的标准是什么；

(4)利用两个计数原理求解.

探究训练

1.某市汽车牌照号码可以上网自编，但规定从左数第2个号码只能从字母B，C，D中选择，其他四个号码可以从0～9这10个数字中选择（数字可以重复）.若某车主第1个号码（从左到右）只想从数字3，5，6，8，9中选择，其他号码只想从1，3，6，9中选择，则他可选的车牌号码的所有可能情况有（　　）

A.180种　　B.360种

C.720种　　D.960种

D　**解析**：按照车主的要求，从左到右第 1 个号码有 5 种选法，第 2 个号码有 3 种选法，其余 3 个号码各有 4 种选法，因此共有 $5\times3\times4\times4\times4=960$（种）情况.

2. 某电视台连续播放 6 个广告，其中有 3 个不同的商业广告、2 个不同的宣传广告、1 个公益广告，要求最后播放的不能是商业广告，且宣传广告与公益广告不能相邻播放，2 个宣传广告也不能相邻播放，则有多少种不同的播放顺序？

解：用 1，2，3，4，5，6 表示广告的播放顺序，则完成这件事有三类方法：

第 1 类，宣传广告与公益广告的播放顺序是 2，4，6，分 6 步完成这件事，共有 $3\times3\times2\times2\times1\times1=36$（种）不同的播放顺序；

第 2 类，宣传广告与公益广告的播放顺序是 1，4，6，分 6 步完成这件事，共有 $3\times3\times2\times2\times1\times1=36$（种）不同的播放顺序；

第 3 类，宣传广告与公益广告的播放顺序是 1，3，6，同样分 6 步完成这件事，共有 $3\times3\times2\times2\times1\times1=36$（种）不同的播放顺序.

由分类加法计数原理得，6 个广告有 $36+36+36=108$（种）不同的播放顺序.

探究点二　排列与组合应用题

例 3　某部门将 5 个安保小组全部安排到指定的三个区域内工作，且每个区域至少有一个安保小组，则这样的安排方法共有________种.（用数字作答）

150　**解析**：由题意可知，将 5 个安保小组分成三队，每队的安保小组个数分别为 1，1，3 或 2，2，1.

若为 1，1，3，不同的安排方法共有 $\dfrac{C_5^1C_4^1}{A_2^2}\cdot A_3^3=60$（种）；

若为 2，2，1，不同的安排方法共有 $\dfrac{C_5^2C_3^2}{A_2^2}\cdot A_3^3=90$（种）.

所以共有 $60+90=150$（种）不同的安排方法.

例 4　5 个男同学和 4 个女同学站成一排.

（1）4 个女同学必须站在一起，有多少种不同的排法？

（2）任何 2 个女同学彼此不相邻，有多少种不同的排法？

（3）其中甲、乙两个同学之间必须有 3 人，有多少种不同的排法？

（4）男同学和女同学相间的排列方法有多少种？

解：（1）4 个女同学必须站在一起，则视 4 个女同学为一个整体，可得有 $A_6^6A_4^4=17\,280$（种）不同的排法.

（2）先排 5 个男同学，再插入女同学即可，所以有 $A_5^5A_6^4=43\,200$（种）不同的排法.

（3）根据题意有 $C_7^3A_3^3A_2^2A_5^5=50\,400$（种）不同的排法.

（4）5 个男同学中间有 4 个空，插入女同学即可，故有 $A_5^5A_4^4=2\,880$（种）不同的排法.

反思提炼

1. 常见排队问题的求解方法：（1）相邻问题采用“捆绑法”；（2）不相邻问题采用“插空法”；（3）有限制元素采用“优先法”；（4）特殊顺序问题，先让所有元素全排列，然后除以有限制元素的全排列数.

2. 对不同元素的分配问题

（1）对于整体均分，解题时要注意分组后，不管它们的顺序如何，都是一种情况，所以分组后一定要除以 A_n^n（n 为均分的组数），避免重复计数；

（2）对于部分均分，解题时注意重复的次数是均匀分组的阶乘数，即若有 m 组元素个数相等，则分组时应除以 $m!$，分组过程中有几个这样的均匀分组，就要除以几个这样的全排列数；

（3）对于不等分组，只需先分组，后排列.由于分组时任何组中元素的个数都不相等，所以不需要除以全排列数.

探究训练

1. 有 9 名歌舞演员，其中 7 名会唱歌，5 名会跳舞，从中选出 2 人，并指派一人唱歌，另一人跳舞，则不同的选派方法有　（　　）

A.19 种　　B.26 种

C.32 种　　D.72 种

C　**解析**：根据题意，有 9 名歌舞演员，其中 7 名会唱歌，5 名会跳舞，

则既会跳舞又会唱歌的有 $5+7-9=3$（人），

只会唱歌的有 $7-3=4$（人），只会跳舞的有 $5-3=2$（人）.

若选出的 2 人中没有既会跳舞又会唱歌的，则有 $C_4^1\times C_2^1=8$（种）选法；

若选出的 2 人中有 1 人既会跳舞又会唱歌，则有 $C_3^1\times C_2^1+C_3^1\times C_4^1=18$（种）选法；

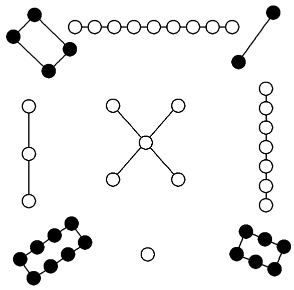
若选出的 2 人全部是既会跳舞又会唱歌的，则有 $A_3^2=6$（种）选法.

综上所述，共有 $8+18+6=32$（种）选派方法.

故选 C.

2. 如图，洛书（古称龟书）是阴阳五行术数之源.在古代传说中有神龟出于洛水，其甲壳上有此图形，结构是戴九履一，左三右七，二四为肩，六八为足，以五居中，五方白圈皆阳数，四隅黑点皆阴数.若从四个阴数和五个阳数中随机选取 4 个数，则选取的 4

个数之和为偶数的方法数为　　　　（　　）



A.60　　B.61　　C.65　　D.66

D　**解析**：由题意可知，阴数为 2，4，6，8，阳数为 1，3，5，7，9.

若选取的 4 个数的和为偶数，则

①4 个数都为阴数，共有 $C_4^4=1$(种)方法；

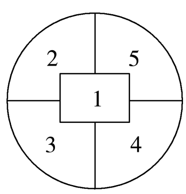
②2 个阳数，2 个阴数，共有 $C_4^2C_5^2=60$(种)方法；

③4 个数都为阳数，共有 $C_5^4=5$(种)方法.

综上，共有 $1+60+5=66$(种)方法.

故选 D.

3. 如图所示，花坛内有五个花池，有五种不同颜色的花卉可供栽种，每个花池内只能种同种颜色的花卉，相邻两个花池的花色需不同，则所有的栽种方案有　　　　（　　）



A.180 种　　B.240 种

C.360 种　　D.420 种

D　**解析**：由题意知，最少用三种颜色的花卉.按照选种的花卉颜色可分为三类，即用三种颜色、四种颜色、五种颜色.

①当用三种颜色时，花池 2，4 同色，花池 3，5 同色，此时共有 A_5^3 种方案；

②当用四种颜色时，花池 2，4 同色或花池 3，5 同色，故共有 $2A_5^4$ 种方案；

③当用五种颜色时，每个花池都不同色，故有 A_5^5 种方案.

因此所有栽种方案有 $A_5^3+2A_5^4+A_5^5=420$(种).

探究点三　二项式定理的应用

例 5　在 $\left(x^3-2x+\frac{1}{x}\right)^4$ 的展开式中，常数项为　　　　（　　）

A.28　　B.-28　　C.-56　　D.56

A　**解析**：因为 $x^3-2x+\frac{1}{x}=\frac{x^4-2x^2+1}{x}=\frac{(x^2-1)^2}{x}$，所以 $\left(x^3-2x+\frac{1}{x}\right)^4=\frac{(x^2-1)^8}{x^4}$. 又 $(x^2-1)^8$ 的展开式中 x^4 的系数为 $C_8^6(-1)^6=28$，所以 $\left(x^3-2x+\frac{1}{x}\right)^4$ 的展开式中的常数项为 28. 故选 A.

例 6　已知 $(1-2x)^2(1-3x)^m=a_0+a_1x+a_2x^2+\cdots+a_nx^n$，$a_1=-19$，求 $m+n$ 的值.

解：因为 $(1-2x)^2(1-3x)^m=(4x^2-4x+1)(1-3x)^m$，

$(1-3x)^m$ 的展开式的通项公式为 $T_{k+1}=C_m^k(-3x)^k=(-3)^kC_m^kx^k$，

所以 $a_1=-4+(-3)C_m^1=-4-3m=-19$，解得 $m=5$，

则 $n=2+m=2+5=7$，所以 $m+n=12$.

［一题多思］

思考 1. 求 $a_1+a_2+\cdots+a_n$ 的值.

解：令 $x=0$ 得 $a_0=1$；令 $x=1$，得 $(1-2)^2(1-3)^5=-32=a_0+a_1+a_2+\cdots+a_7$，

所以 $a_1+a_2+a_3+\cdots+a_7=-32-1=-33$.

思考 2. 求 a_2 的值.

解：$a_2=4\times1-4\times(-3)C_5^1+1\times(-3)^2C_5^2=154$.

思考 3. 求 $a_1+2a_2+3a_3+\cdots+na_n$ 的值.

解：由 $(1-2x)^2(1-3x)^5=a_0+a_1x+a_2x^2+\cdots+a_7x^7$ 两边对 x 求导，得

$2(1-2x)\cdot(-2)\cdot(1-3x)^5+(1-2x)^2\cdot5(1-3x)^4\cdot(-3)=a_1+2a_2x+3a_3x^2+\cdots+7a_7x^6$，

令 $x=1$，得 $2\times(-1)\times(-2)\times(-2)^5+(-1)^2\times5\times(-2)^4\times(-3)=a_1+2a_2+3a_3+\cdots+7a_7$，

所以 $a_1+2a_2+3a_3+\cdots+7a_7=-368$.

反思提炼

1. 求 $(a+b)^n(n\in\mathbf{N}^*)$ 的展开式中的特定项(常数项、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 $T_{k+1}=C_n^ka^{n-k}b^k$，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(如常数项要求字母的指数为零，有理项要求字母的指数为整数)列出相应方程(组)或不等式(组)，解出 k 的值；

第三步，把 k 的值代入通项中，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}.

2. 赋值法的应用

二项式定理给出的是一个恒等式 $(x+y)^n=C_n^0x^n+C_n^1x^{n-1}y^1+\cdots+C_n^kx^{n-k}y^k+\cdots+C_n^ny^n$，$n\in\mathbf{N}^*$，

对于 x,y 的一切值都成立.因此,可令 x,y 等于一些特殊的值.令 x,y 等于多少,应视具体情况而定,一般取 $1,-1$ 或 0,有时也取其他值.如:

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

(2)形如 $(ax+by)^n(a,b\in\mathbf{R})$ 的式子,求其展开式各项系数之和,只需令 $x=y=1$ 即可.

探究训练

1.(多选)若对任意实数 x,有 $(2x-3)^9=a_0+a_1(x-1)+a_2(x-1)^2+a_3(x-1)^3+\cdots+a_9(x-1)^9$,则下列结论正确的是　　(　　)

A.$a_2=-144$

B.$a_0=1$

C.$a_0+a_1+a_2+\cdots+a_9=1$

D.$a_0-a_1+a_2-a_3+\cdots-a_9=-3^9$

ACD　**解析**:因为对任意实数 x,有 $(2x-3)^9=a_0+a_1(x-1)+a_2(x-1)^2+a_3(x-1)^3+\cdots+a_9(x-1)^9=[-1+2(x-1)]^9$,

所以 $a_2=C_9^2\times(-1)^7\times2^2=-144$,故 A 正确;

令 $x=1$,可得 $a_0=(2-3)^9=-1$,故 B 错误;

令 $x=2$,可得 $a_0+a_1+a_2+\cdots+a_9=(2\times2-3)^9=1$,故 C 正确;

令 $x=0$,可得 $a_0-a_1+a_2-\cdots-a_9=(0-3)^9=-3^9$,故 D 正确.故选 ACD.

2.在①只有第 8 项的二项式系数最大;②奇数项的二项式系数之和为 4^7;③各项系数之和为 4^{14} 这三个条件中任选一个,补充在下面的问题中.若问题中的 k 存在,求出 k 的值;若 k 不存在,请说明理由.

在 $\left(\sqrt{x}+\dfrac{3}{x^3}\right)^n$ 的展开式中,________,是否存在整数 k,使得 T_{k+1} 是展开式中的常数项?

解:若选填条件①只有第 8 项的二项式系数最大,即 C_n^7 最大,由二项式系数的性质可得 $n=14$.

若选填条件③各项系数之和为 4^{14},令 $x=1$,得 $4^n=4^{14}$,

即 $n=14$.

$\left(\sqrt{x}+\dfrac{3}{x^3}\right)^{14}$ 的展开式中 $T_{k+1}=C_{14}^k(\sqrt{x})^{14-k}\left(\dfrac{3}{x^3}\right)^k=3^kC_{14}^kx^{\frac{14-7k}{2}}$.

令 $\dfrac{14-7k}{2}=0$,得 $k=2$.

即存在整数 $k=2$,使得 T_3 是展开式中的常数项.

若选填条件②奇数项二项式系数之和为 4^7,

则 $2^{n-1}=4^7=2^{14}$,所以 $n=15$.

$\left(\sqrt{x}+\dfrac{3}{x^3}\right)^{15}$ 的展开式中 $T_{k+1}=C_{15}^k(\sqrt{x})^{15-k}\cdot\left(\dfrac{3}{x^3}\right)^k=3^kC_{15}^kx^{\frac{15-7k}{2}}$.

令 $\dfrac{15-7k}{2}=0$,得 $k=\dfrac{15}{7}\notin\mathbf{Z}$.

即不存在整数 k,使得 T_{k+1} 是展开式中的常数项.

3.设 $(2-\sqrt{3}x)^{100}=a_0+a_1x+a_2x^2+\cdots+a_{100}x^{100}$,求下列各式的值.

(1)a_0;

(2)$a_1+a_2+\cdots+a_{100}$;

(3)$a_1+a_3+\cdots+a_{99}$;

(4)$(a_0+a_2+\cdots+a_{100})^2-(a_1+a_3+\cdots+a_{99})^2$;

(5)$|a_0|+|a_1|+|a_2|+\cdots+|a_{100}|$.

解:(1)令 $x=0$,得 $a_0=2^{100}$.

(2)令 $x=1$,可得 $a_0+a_1+a_2+\cdots+a_{100}=(2-\sqrt{3})^{100}$ (*),

所以 $a_1+a_2+\cdots+a_{100}=(2-\sqrt{3})^{100}-2^{100}$.

(3)令 $x=-1$,可得 $a_0-a_1+a_2-a_3+\cdots+a_{100}=(2+\sqrt{3})^{100}$,

与(2)中(*)式相减得

$a_1+a_3+\cdots+a_{99}=\dfrac{(2-\sqrt{3})^{100}-(2+\sqrt{3})^{100}}{2}$.

(4)原式 $=[(a_0+a_2+\cdots+a_{100})+(a_1+a_3+\cdots+a_{99})]\cdot[(a_0+a_2+\cdots+a_{100})-(a_1+a_3+\cdots+a_{99})]=(a_0+a_1+a_2+\cdots+a_{100})\cdot(a_0-a_1+a_2-a_3+\cdots+a_{98}-a_{99}+a_{100})=[(2-\sqrt{3})(2+\sqrt{3})]^{100}=1^{100}=1$.

(5)因为 $(2-\sqrt{3}x)^{100}$ 的展开式的通项为 $T_{k+1}=(-1)^kC_{100}^k2^{100-k}(\sqrt{3})^kx^k$,

所以 $a_{2n-1}<0(n\in\mathbf{N}^*)$.

所以 $|a_0|+|a_1|+|a_2|+\cdots+|a_{100}|=a_0-a_1+a_2-a_3+\cdots+a_{100}=(2+\sqrt{3})^{100}$.

第六章综合检测

（时间：120 分钟，分值：150 分）

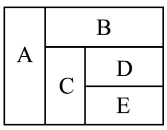
一、单项选择题（本题共 8 小题，每小题 5 分，共40 分）

1. $C_4^2+C_4^3=$　　（B）

A.8　　B.10

C.12　　D.16

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着色的方法共有　　（　　）



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)涂法.故选 D.

3. 已知集合 $A=\{1,2,3,4\}$，$B=\{5,6,7\}$，$C=\{8,9\}$.现在从这三个集合中取出两个集合，再从这两个集合中各取出一个元素，组成一个含有两个元素的集合，则一共可以组成集合　　（　　）

A.24 个　　B.36 个　　C.26 个　　D.27 个

C　**解析**：从三个集合中取出两个集合，有 $C_3^2=3$(种)取法.分别是集合 A，B；集合 A，C；集合 B，C.

当取出 A，B 时，从这两个集合各取一个元素，有 $C_4^1\times C_3^1=12$(种)取法；

当取出 A，C 时，从这两个集合各取一个元素，有 $C_4^1\times C_2^1=8$(种)取法；

当取出 B，C 时，从这两个集合各取一个元素，有 $C_3^1\times C_2^1=6$(种)取法.

一共可以组成 $12+8+6=26$(个)集合.

4. 将 A，B，C，D，E 排成一排，要求 A，B，C 在排列中顺序为 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 $2\cdot\dfrac{A_5^5}{A_3^3}=40$(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.

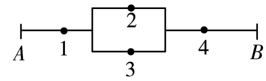
若甲区域除军舰 A 外还有一艘军舰，则安排方案共有 $C_4^1\times C_3^1\times C_2^2\times A_2^2=24$(种)；

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.今发现 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. 设集合 $A=\{1,2,3,4,5\}$，m，$n\in A$，则方程 $\dfrac{x^2}{m}+\dfrac{y^2}{n}=1$ 表示的焦点位于 x 轴上的椭圆有　　（　　）

A.8 个　　　　B.10 个

C.12 个　　　　D.16 个

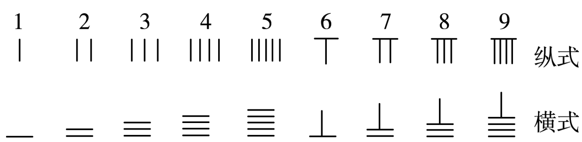
B　**解析**:因为椭圆的焦点在 x 轴上,所以 $m>n$.当 $m=5$ 时,有 $n=1,2,3,4$ 四种结果;

当 $m=4$ 时,有 $n=1,2,3$ 三种结果;当 $m=3$ 时,有 $n=1,2$ 两种结果;当 $m=2$ 时,有 $n=1$ 一种结果.

故所求的椭圆共有 $4+3+2+1=10$(个).故选 B.

二、多项选择题(本题共 3 小题,每小题 6 分,共18 分)

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棍.如图,用算筹表示数 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……以此类推,交替使用纵横两式.例如:27 可以表示为“═ ╥”.如果用算筹表示不含 0 的两位数,则下列说法正确的是　(　　)



A.36 可以表示为≡ ⊥

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≡ ⊤,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(本题共 3 小题,每小题 5 分,共 15 分)

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 $k=0,1,2,\cdots,5$.

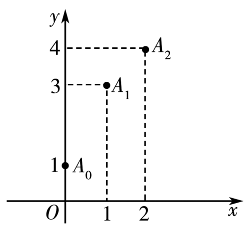
于是 $\left(\frac{1}{x}-1\right)^5$ 的展开式中 x^{-2} 的系数为 $(-1)^3C_5^3=-10$,

x^{-1} 的系数为 $(-1)^4C_5^4=5$,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 $1\times(-10)+2a\times5+a^2\times(-1)=-a^2+10a-10$.

依题意$-a^2+10a-10=-1$，即$a^2-10a+9=0$，解得$a=1$或$a=9$.

13. 设$a\neq 0$，n是大于1的自然数，$\left(1+\dfrac{x}{a}\right)^n$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、乙两人到同一演出场馆工作的方法数为$A_3^3=6$，故甲、乙两人不到同一演出场馆工作的不同分派方案有$36-6=30$（种）.

四、解答题（本题共5小题，共77分）

15.（13分）3名男生、4名女生站成一排.

（1）任意2名女生都不相邻，有多少种排法？

（2）其中的男生甲、乙相邻，有多少种排法？

解：（1）先把3名男生全排列，再把4名女生插在男生之间的4个空中，即可得有$A_3^3A_4^4=144$（种）排法.

（2）甲、乙相邻，看成一个“元素集团”，与其余5个元素全排列有A_6^6种排法，“元素集团”内部排法有A_2^2种，

所以共有$A_6^6A_2^2=1\ 440$（种）排法.

16.（15分）从集合$\{-3,-2,-1,0,1,2,3\}$中任取3个不同的元素作为抛物线方程$y=ax^2+bx+c(a\neq 0)$中a,b,c的值.设抛物线过原点，且顶点在第一象限.这样的抛物线共有多少条？

解：若抛物线$y=ax^2+bx+c$过原点，且顶点$\left(-\dfrac{b}{2a},\dfrac{4ac-b^2}{4a}\right)$在第一象限，则$a,b,c$应满足

$$\begin{cases}0=a\times 0^2+b\times 0+c,\\-\dfrac{b}{2a}>0,\\\dfrac{4ac-b^2}{4a}>0,\end{cases}\text{即}\begin{cases}c=0,\\a<0,\\b>0.\end{cases}$$

分三步：a可以取$-3,-2,-1$；b可以取$1,2,3$；c取0.

所以满足条件的抛物线的条数为$3\times 3\times 1=9$.

17.（15分）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$T_{k+1}=C_n^k(-2x)^k=(-2)^kC_n^k\cdot x^k$.

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）$T_{k+1}=C_n^k(-2x)^k=a_kx^k$，

所以$a_k=C_6^k(-2)^k$，

所以$(-1)^k\cdot\dfrac{a_k}{2^k}=C_6^k$，

$-\dfrac{a_1}{2}+\dfrac{a_2}{2^2}-\dfrac{a_3}{2^3}+\cdots+(-1)^n\cdot\dfrac{a_n}{2^n}=C_6^1+C_6^2+\cdots+C_6^6=2^6-1=63$.

18.（17分）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）$\left(\sqrt{x}+\dfrac{1}{2\sqrt[3]{x}}\right)^n$的展开式的通项为$T_{r+1}=C_n^r(\sqrt{x})^{n-r}\left(\dfrac{1}{2\sqrt[3]{x}}\right)^r=\dfrac{1}{2^r}C_n^rx^{\frac{3n-5r}{6}}$，$r=0,1,2,\cdots,n(n\geqslant 2)$，

则$a_1=\dfrac{1}{2^0}C_n^0=1$，$a_2=\dfrac{1}{2^1}C_n^1=\dfrac{1}{2}n$，

$a_3=\frac{1}{2^2}C_n^2=\frac{n(n-1)}{8}$.

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 $T_{r+1}=C_8^r(\sqrt{x})^{8-r}\left(\frac{1}{2\sqrt[3]{x}}\right)^r=\frac{1}{2^r}C_8^r x^{\frac{24-5r}{6}}$($0\leqslant r\leqslant 8$ 且 $r\in\mathbf{N}$),

记第 k 项系数最大,则有 $T_k\geqslant T_{k+1}$,且 $T_k\geqslant T_{k-1}$,

即 $\begin{cases}C_8^{k-1}2^{-k+1}\geqslant C_8^k 2^{-k},\\ C_8^{k-1}2^{-k+1}\geqslant C_8^{k-2}2^{-k+2},\end{cases}$

解得 $3\leqslant k\leqslant 4$.

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 $T_{r+1}=\frac{1}{2^r}C_8^r x^{\frac{24-5r}{6}}$($0\leqslant r\leqslant 8$ 且 $r\in\mathbf{N}$),

令 $\frac{24-5r}{6}\in\mathbf{Z}$,$0\leqslant r\leqslant 8$ 且 $r\in\mathbf{N}$,则 $r=0$ 或 $r=6$,

所以展开式中有理项为 $T_1=x^4$ 和 $T_7=\frac{7}{16x}$.

19.(17 分)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.

(1)用 n,x 表示 a_n 与数列$\{a_n\}$的前 n 项和 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 $\begin{cases}2m+3\geqslant 3m,\\ m-2\geqslant 1,\end{cases}$ 即 $\begin{cases}m\leqslant 3,\\ m\geqslant 3,\end{cases}$

所以 $m=3$,所以 $a_1=C_9^9\cdot A_1^1=1$.

又由 $\left(x+\frac{1}{4x^2}\right)^4$,知 $T_2=C_4^1x^3\cdot\frac{1}{4x^2}=x$,所以 $q=x$.

所以 $a_n=x^{n-1}$($n\geqslant 1$,且 $n\in\mathbf{N}^*$),$S_n=\begin{cases}n,x=1,\\ \frac{1-x^n}{1-x},x\neq 1.\end{cases}$

(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 $2A_n=n(C_n^0+C_n^1+\cdots+C_n^n)=n\cdot 2^n$,

所以 $A_n=n\cdot 2^{n-1}$.

当 $x\neq 1$ 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 $A_n=\begin{cases}n\cdot 2^{n-1},x=1,\\ \frac{2^n-(1+x)^n}{1-x},x\neq 1.\end{cases}$

第七章　随机变量及其分布

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，了解条件概率的概念，能计算简单随机事件的条件概率.
2.结合古典概型，了解条件概率与事件独立性的关系，会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 $P(AB)=P(A)P(B|A)$.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)$P(\Omega|A)=1$；

(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
$P(\overline{B}|A)=1-P(B|A)$.

概念辨析

1.判断(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以后，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 $p_2=\dfrac{C_5^2}{C_{100}^2}=\dfrac{1}{495}$，

所以所求概率 $p=\dfrac{p_2}{p_1}=\dfrac{\frac{1}{495}}{\frac{1}{20}}=\dfrac{4}{99}$.

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 $P(B|A)=\dfrac{P(AB)}{P(A)}$.

(2)$P(B|A)$与 $P(A|B)$表示的意思相同吗？

提示：不同.$P(B|A)$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，$P(B|A)=\dfrac{P(AB)}{P(A)}$，$P(A|B)=\dfrac{P(AB)}{P(B)}$.

任务型课堂

学习任务一　利用条件概率公式求条件概率

例 1　在 5 道题中有 3 道数学题和 2 道语文题，每次从中不放回地随机抽取 1 道题，抽取 2 次.设第 1 次抽到数学题为事件 A，第 2 次抽到数学题为事件 B，则第 1 次和第 2 次都抽到数学题为事件 AB.

(1)求在第 1 次抽到数学题的条件下，第 2 次抽到数学题的概率；

(2)求在第 1 次抽到语文题的条件下，第 2 次抽到数学题的概率.

解：(1)方法一：因为 $P(A)=\dfrac{3}{5}$，$P(AB)=\dfrac{3\times2}{5\times4}=\dfrac{3}{10}$，

所以在第 1 次抽到数学题的条件下，第 2 次抽到数学题的概率为 $P(B|A)=\dfrac{P(AB)}{P(A)}=\dfrac{\frac{3}{10}}{\frac{3}{5}}=\dfrac{1}{2}$.

方法二：因为 $n(AB)=3\times2=6$，$n(A)=3\times4=12$，

所以 $P(B|A)=\dfrac{n(AB)}{n(A)}=\dfrac{6}{12}=\dfrac{1}{2}$.

(2)方法一：因为 $P(\overline{A})=\dfrac{2}{5}$，$P(\overline{A}B)=\dfrac{2\times3}{5\times4}=\dfrac{3}{10}$，

所以在第 1 次抽到语文题的条件下，第 2 次抽到数学题的概率为 $P(B|\overline{A})=\dfrac{P(\overline{A}B)}{P(\overline{A})}=\dfrac{\frac{3}{10}}{\frac{2}{5}}=\dfrac{3}{4}$.

方法二：因为 $n(\overline{A}B)=2\times3=6$，$n(\overline{A})=2\times4=8$，

所以 $P(B|\overline{A})=\dfrac{n(\overline{A}B)}{n(\overline{A})}=\dfrac{6}{8}=\dfrac{3}{4}$.

反思提炼

条件概率的 3 种求法

续表

定义法	先求 $P(A)$和 $P(AB)$，再由 $P(B|A)=\dfrac{P(AB)}{P(A)}$求 $P(B|A)$
基本事件法	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 $P(B|A)=\dfrac{n(AB)}{n(A)}$
缩样法	缩小样本空间，就是去掉第一次抽取的情况，只研究剩下的情况，用古典概型求解，该求法能化繁为简

探究训练

1. 甲、乙两名游客准备分别从三峡大坝、三峡人家、三峡大瀑布和清江画廊四个景区中随机选择一个游玩.记事件 A 为"和乙至少一人选择三峡大坝景区"，事件 B 为"甲和乙选择的景点不同"，则 $P(B|A)=$(　　)

A.$\dfrac{6}{7}$　　B.$\dfrac{1}{7}$

C.$\dfrac{3}{4}$　　D.$\dfrac{1}{4}$

A　**解析**：$P(A)=1-\dfrac{3}{4}\times\dfrac{3}{4}=\dfrac{7}{16}$，$P(AB)=\dfrac{C_2^1C_3^1}{4\times4}=\dfrac{3}{8}$，所以 $P(B|A)=\dfrac{P(AB)}{P(A)}=\dfrac{3}{8}\div\dfrac{7}{16}=\dfrac{6}{7}$.故选 A.

2. 据某地区气象台统计，在某季节该地区下雨的概率是$\dfrac{4}{15}$，刮四级以上风的概率为$\dfrac{2}{15}$，既刮四级以上风又下雨的概率为$\dfrac{1}{10}$.设事件 A 为"下雨"，事件 B 为"刮四级以上的风"，那么$P(B|A)=$________.

$\dfrac{3}{8}$　**解析**：由题意知 $P(A)=\dfrac{4}{15}$，$P(AB)=\dfrac{1}{10}$，

所以 $P(B|A)=\dfrac{P(AB)}{P(A)}=\dfrac{3}{8}$.

学习任务 二　利用乘法公式求概率

例 2　一袋中有 10 个球，其中 3 个黑球、7 个白球，先后两次从中任取一个球（不放回）.设事件 A_i 表示"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：由题设知 $P(A_1)=\frac{3}{10}$，$P(\overline{A}_2|A_1)=\frac{7}{9}$.

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：例 2　$P(A_1)=\frac{3}{10}$，$P(A_2)=\frac{3}{10}$，则 $P(A_1A_2)=\frac{3}{10}\times\frac{3}{10}=\frac{9}{100}$.

思考 1　$P(A_1)=\frac{3}{10}$，$P(\overline{A}_2)=\frac{7}{10}$，则 $P(A_1\overline{A}_2)=\frac{3}{10}\times\frac{7}{10}=\frac{21}{100}$.

反思提炼

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 $P(AB)$ 可由条件概率的变形式 $P(AB)=P(A)\cdot P(B|A)$ 求得. $P(B|A)$ 可采用缩小样本空间的方法来计算，$P(B|A)=\frac{n(AB)}{n(A)}$，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.

探究训练

1. 已知某产品的次品率为 4%，其合格品中 75% 为一级品，则从该产品中任选一件为一级品的概率为（　　）

A. 75%　　　B. 96%

C. 72%　　　D. 78.125%

C　**解析**：记"任选一件产品是合格品"为事件 A，则 $P(A)=1-P(\overline{A})=1-4\%=96\%$. 记"任选一件产品是一级品"为事件 B. 由于一级品必是合格品，所以事件 A 包含事件 B，故 $P(AB)=P(B)$. 由合格品中 75% 为一级品知 $P(B|A)=75\%$，故 $P(B)=P(AB)=P(A)\cdot P(B|A)=96\%\times75\%=72\%$.

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.

学习任务 三　利用条件概率的性质求互斥事件的条件概率

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 $D=A\cup B\cup C$，可知 $P(D)=P(A\cup B\cup C)=P(A)+P(B)+P(C)=\frac{C_{10}^6}{C_{20}^6}+\frac{C_{10}^5C_{10}^1}{C_{20}^6}+\frac{C_{10}^4C_{10}^2}{C_{20}^6}=\frac{12\ 180}{C_{20}^6}$.

(3)因为 A，B 互斥，$E=A\cup B$，$P(AD)=P(A)$，$P(BD)=P(B)$，

所以 $P(E|D)=P(A|D)+P(B|D)=\frac{P(A)}{P(D)}+\frac{P(B)}{P(D)}=\frac{\frac{C_{10}^6}{C_{20}^6}}{\frac{12\ 180}{C_{20}^6}}+\frac{\frac{C_{10}^5C_{10}^1}{C_{20}^6}}{\frac{12\ 180}{C_{20}^6}}=\frac{13}{58}$.

故他获得优秀的概率为 $\frac{13}{58}$.

反思提炼

当所求事件的概率较复杂时，往往把该事件分成两个(或多个)互不相容的较简单的事件之和，求出这些较简单事件的概率，再利用概率的加法公式和乘法公式便可得出所求事件的概率.

探究训练

1. 在一个袋子中装有除颜色外完全相同的 10 个球，其中有 1 个红球、2 个黄球、3 个黑球、4 个白球.从中不放回地摸出 2 个球，求在摸出的第一个球是红球的条件下，第二个球是黄球或黑球的概率.

解：设"摸出的第一个球为红球"为事件 A，"摸出的第二个球为黄球"为事件 B，"摸出的第二个球为黑球"为事件 C.

方法一：$P(A)=\frac{1}{10}$，$P(AB)=\frac{1\times2}{10\times9}=\frac{1}{45}$，$P(AC)=\frac{1\times3}{10\times9}=\frac{1}{30}$，

所以 $P(B|A)=\frac{P(AB)}{P(A)}=\frac{\frac{1}{45}}{\frac{1}{10}}=\frac{10}{45}=\frac{2}{9}$，

$P(C|A)=\frac{P(AC)}{P(A)}=\frac{\frac{1}{30}}{\frac{1}{10}}=\frac{1}{3}$.

所以 $P(B\cup C|A)=P(B|A)+P(C|A)=\frac{2}{9}+\frac{1}{3}=\frac{5}{9}$.

故在摸出的第一个球是红球的条件下，第二个球是黄球或黑球的概率为 $\frac{5}{9}$.

方法二：因为 $n(A)=1\times C_9^1=9$，$n[(B\cup C)\cap A]=C_2^1+C_3^1=5$，所以 $P(B\cup C|A)=\frac{5}{9}$.

故在摸出的第一个球是红球的条件下，第二个球是黄球或黑球的概率为 $\frac{5}{9}$.

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.若第二次取出的是红球，则称试验成功.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 $P(A)=\frac{7}{10}$，$P(B)=\frac{3}{10}$，

所以 $P(R|A)=\frac{1}{2}$，$P(R|B)=\frac{4}{5}$.

事件"试验成功"表示为 $RA\cup RB$.

又事件 RA 与事件 RB 互斥，

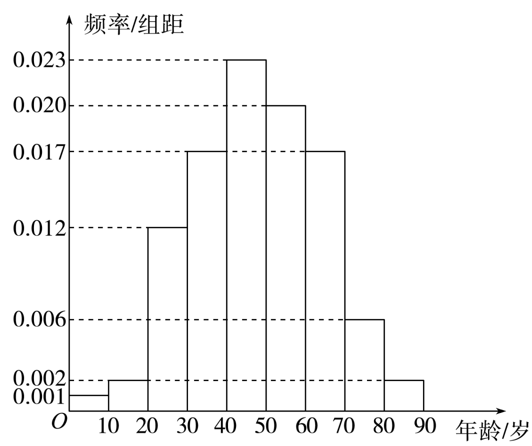
故由概率加法公式与乘法公式得

$P(RA\cup RB)=P(RA)+P(RB)$

$=P(R|A)P(A)+P(R|B)P(B)$

$=\frac{1}{2}\times\frac{7}{10}+\frac{4}{5}\times\frac{3}{10}=0.59$.

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，得到如下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

$P(C|B)=\dfrac{P(BC)}{P(B)}=\dfrac{P(C)P(B|C)}{P(B)}=\dfrac{0.001\times0.23}{0.16}=0.001\ 437\ 5\approx0.001\ 4$.

体系构建

条件概率
- 条件概率的定义 $P(B|A)=\dfrac{P(AB)}{P(A)}$ — 计算公式 — 计算方法
- 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 $\bar{B}$ 和 B 互为对立事件，则 $P(\bar{B}|A)=1-P(B|A)$
- 概率的乘法公式 $P(AB)=P(A)P(B|A)$

课后素养评价(九)

基础性·能力运用

1. 下面的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 $\dfrac{2}{5}$，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，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. 从 1，2，3，4，5，6，7，8，9 中不放回地随机取 2 个数，事件 A＝“第一次取到的是奇数”，B＝“第二次取到的是奇数”，则 $P(B|A)=$　　(　　)

A.$\dfrac{1}{5}$　　B.$\dfrac{3}{10}$

C.$\dfrac{2}{5}$　　D.$\dfrac{1}{2}$

D　**解析**：由题意，$P(AB)=\dfrac{C_5^2}{C_9^2}=\dfrac{5}{18}$，$P(A)=\dfrac{C_5^1}{C_9^1}=\dfrac{5}{9}$，所以 $P(B|A)=\dfrac{P(AB)}{P(A)}=\dfrac{1}{2}$.

3. 袋中有大小相同的 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 1 次取到的是黑球的条件下，第 2 次取到黑球的概率为　　(　　)

A.$\dfrac{1}{2}$　　B.$\dfrac{2}{5}$　　C.$\dfrac{3}{5}$　　D.$\dfrac{2}{3}$

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”，

$P(A)=\dfrac{4}{6}=\dfrac{2}{3}$，$P(AB)=\dfrac{4\times3}{6\times5}=\dfrac{2}{5}$，

所以在第 1 次取到的是黑球的条件下，第 2 次取到黑球的概率为 $P(B|A)=\dfrac{P(AB)}{P(A)}=\dfrac{3}{5}$.

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.已知某种型号的共享充电宝循环充电超过 500 次的概率为 $\dfrac{3}{4}$，超过 1 000 次的概率为 $\dfrac{1}{2}$，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.$\dfrac{3}{4}$　　B.$\dfrac{2}{3}$　　C.$\dfrac{1}{2}$　　D.$\dfrac{1}{3}$

B　**解析**：记事件 A 为“该充电宝循环充电超过 500 次”，则 $P(A)=\dfrac{3}{4}$.记事件 B 为“该充电宝循环充电超过1 000 次”，则 $P(B)=\dfrac{1}{2}$.易知 $P(AB)=P(B)=\dfrac{1}{2}$，所以 $P(B|A)=\dfrac{P(AB)}{P(A)}=\dfrac{\frac{1}{2}}{\frac{3}{4}}=\dfrac{1}{2}\times\dfrac{4}{3}=\dfrac{2}{3}$.

故选 B.

5.若 B,C 是互斥事件且事件 A 满足 $P(B|A)=\frac{1}{3}$，$P(C|A)=\frac{1}{4}$，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.某班学生的考试成绩中，数学不及格的占 15%，语文不及格的占 5%，两门都不及格的占 3%.已知一名学生数学不及格，则他语文也不及格的概率是 （　　）

A.$\frac{1}{5}$　　B.$\frac{3}{10}$　　C.$\frac{1}{2}$　　D.$\frac{3}{5}$

A　**解析**：设事件 A 为"数学不及格"，事件 B 为"语文不及格"，则 $P(B|A)=\frac{P(AB)}{P(A)}=\frac{0.03}{0.15}=\frac{1}{5}$.

所以该学生数学不及格时，语文也不及格的概率为 $\frac{1}{5}$.

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.(新定义)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、密码学以及大数分解等领域都有广泛的应用.已知对于正整数 $a,n(n\geqslant 2)$，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 个整数中随机抽取一个整数 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，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 $P(B|A)=\frac{P(AB)}{P(A)}=\frac{5}{7}$.

综合性·创新提升

1.甲、乙两人向同一目标各射击 1 次，已知甲命中目标的概率为 $\frac{1}{3}$，乙命中目标的概率为 $\frac{1}{2}$，已知目标至少被命中 1 次，则甲命中目标的概率为 （　　）

A.$\frac{1}{4}$　　B.$\frac{1}{3}$　　C.$\frac{1}{2}$　　D.$\frac{2}{3}$

C　**解析**：设事件 A 为"目标至少被命中 1 次"，事件 B 为"甲命中目标".

则 $P(A)=\frac{1}{3}\times\frac{1}{2}+\left(1-\frac{1}{3}\right)\times\frac{1}{2}+\frac{1}{3}\times\left(1-\frac{1}{2}\right)=\frac{2}{3}$，$P(AB)=\frac{1}{3}\times\frac{1}{2}+\frac{1}{3}\times\left(1-\frac{1}{2}\right)=\frac{1}{3}$，

所以 $P(B|A)=\frac{\frac{1}{3}}{\frac{2}{3}}=\frac{1}{2}$.故选 C.

2.甲、乙、丙、丁四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，每人限报其中一个项目.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.$\frac{1}{4}$　　B.$\frac{3}{4}$　　C.$\frac{2}{9}$　　D.$\frac{5}{9}$

C　**解析**：由已知有 $P(B)=\frac{3^3}{4^4}=\frac{27}{256}$，$P(AB)=\frac{A_3^3}{4^4}=\frac{3}{128}$，所以 $P(A|B)=\frac{P(AB)}{P(B)}=\frac{2}{9}$.

3.将两枚骰子各掷一次，设事件 $A=$"两枚骰子点数不相同"，$B=$"至少出现一个 6 点"，则概率 $P(A|B)=$ （　　）

A.$\frac{10}{11}$　　B.$\frac{5}{11}$　　C.$\frac{5}{18}$　　D.$\frac{5}{36}$

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$P(A|B)=\dfrac{n(AB)}{n(B)}=\dfrac{10}{11}$.

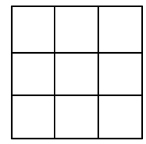
4.加工某种零件需要两道工序,第一道工序出废品的概率为0.4,两道工序都出废品的概率为0.2,则在第一道工序出废品的条件下,第二道工序又出废品的概率为________.

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$P(B|A)=\dfrac{P(AB)}{P(A)}=\dfrac{1}{2}=0.5$.

5.某校组织甲、乙、丙、丁、戊、己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为"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为"学生丙第一个出场".对事件A,甲和乙都不是第一个出场,第一类:乙在最后,则优先从中间4个位置中选一个给甲,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个位置中选两个给甲、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人中选一人安排在最后,再将余下的4人全排列,有$C_4^1\cdot A_4^4$种排法,故$n(AB)=C_4^1\cdot A_4^4$.故$P(B|A)=\dfrac{n(AB)}{n(A)}=\dfrac{C_4^1\cdot A_4^4}{C_4^1\cdot A_4^4+A_4^2\cdot A_4^4}=\dfrac{1}{4}$.

6.如图,一个大正方形被平均分成9个小正方形,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个大小相同的黑球和白球.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)现从中不放回地取球,每次任取1个球,取2次.已知第1次取得白球,求第2次取得黑球的概率.

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1次取得白球"为事件B,"第2次取得黑球"为事件C,

则$P(BC)=\dfrac{C_5^1C_5^1}{C_{10}^1C_9^1}=\dfrac{25}{90}=\dfrac{5}{18}$,$P(B)=\dfrac{5}{10}=\dfrac{1}{2}$.

故$P(C|B)=\dfrac{P(BC)}{P(B)}=\dfrac{\frac{5}{18}}{\frac{1}{2}}=\dfrac{5}{9}$.

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和乘法公式推导出全概率公式的过程.

2.会利用全概率公式计算概率.

3.了解贝叶斯公式以及其简单应用.

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 $P(A_i|B)=\dfrac{P(A_i)P(B|A_i)}{P(B)}=\underline{\dfrac{P(A_i)P(B|A_i)}{\sum\limits_{k=1}^{n}P(A_k)P(B|A_k)}}$,$i=1,2,\cdots,n$.

概念辨析

1.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,则第二个人取得红球的概率为________.(第一个人取出的球不放回)

$\dfrac{1}{5}$　**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 $P(A)=P(B)P(A|B)+P(\overline{B})P(A|\overline{B})=\dfrac{2}{10}\times\dfrac{1}{9}+\dfrac{8}{10}\times\dfrac{2}{9}=\dfrac{1}{5}$.

2.设某公路上经过的货车与客车的数量之比为 2∶1,货车与客车中途停车修理的概率分别为 0.02,0.01.今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

$$P(A_1|B)=\frac{P(A_1)P(B|A_1)}{P(A_1)P(B|A_1)+P(A_2)P(B|A_2)}$$

$$=\frac{\frac{2}{3}\times0.02}{\frac{2}{3}\times0.02+\frac{1}{3}\times0.01}=0.8.$$

3.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,需要注意以下几点:

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

学习任务一　利用全概率公式求概率

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%.从这批产品中任取的一件是次品的概率是多少？

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 $P(B_1)=0.3$，$P(B_2)=0.5$，$P(B_3)=0.2$，$P(A|B_1)=0.02$，$P(A|B_2)=0.01$，$P(A|B_3)=0.01$，

故 $P(A)=0.02\times0.3+0.01\times0.5+0.01\times0.2=0.013$.

［一题多思］

思考.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%.又知一厂、二厂的产品的次品率分别为2%，1%.从这批产品中任取的一件是次品的概率是0.013，求三厂的产品的次品率.

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 $P(A)=P(A|B_1)P(B_1)+P(A|B_2)P(B_2)+P(A|B_3)P(B_3)$.

又 $P(B_1)=0.3$，$P(B_2)=0.5$，$P(B_3)=0.2$，$P(A|B_1)=0.02$，$P(A|B_2)=0.01$，

设 $P(A|B_3)=m$，

故 $P(A)=0.02\times0.3+0.01\times0.5+m\times0.2=0.013$，

得 $m=0.01$.

所以三厂的产品的次品率为1%.

反思提炼

1.全概率公式的主要用处在于它可以将一个复杂的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 $P(A_i)(i=1,2,\cdots,n)$ 以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 $P(B)=\sum\limits_{i=1}^{n}P(B|A_i)P(A_i)$.

探究训练

1.已知某种传染疾病的患病率为5%，通过验血诊断该病的误诊率为2%，即非患者中有2%的人诊断为阳性，患者中有2%的人诊断为阴性.随机抽取一人进行验血，则其诊断结果为阳性的概率为（　　）

A.0.46　　B.0.046

C.0.68　　D.0.068

D　**解析**：设“随机抽取一人进行验血，其诊断结果为阳性”为事件 A，“随机抽取一人，其患病”为事件 B，则“随机抽取一人，其没有患病”为事件 $\overline{B}$，

则 $P(A)=P(B)P(A|B)+P(\overline{B})P(A|\overline{B})=0.05\times0.98+0.95\times0.02=0.068$.

故选D.

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，则该射手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，该射手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. 设有两箱同种商品，第一箱内装有 50 件，其中 10 件优质品；第二箱内装有 30 件，其中 18 件优质品．现在随意打开一箱，然后从箱中随意取出一件，求取到优质品的概率．

解：设 A＝“取到的是优质品”，B_i＝“打开的是第 i 箱”($i=1,2$)．

$P(B_1)=P(B_2)=\frac{1}{2}$，$P(A|B_1)=\frac{1}{5}$，$P(A|B_2)=\frac{3}{5}$．

由全概率公式，得 $P(A)=P(B_1)P(A|B_1)+P(B_2)\cdot P(A|B_2)=\frac{1}{2}\times\frac{1}{5}+\frac{1}{2}\times\frac{3}{5}=\frac{2}{5}$．

学习任务二　贝叶斯公式的应用

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．用事件 A 表示“该生不遵守承诺”，事件 B 表示“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)依据题意知 $P(B)=0.8$，$P(\overline{B})=0.2$，$P(A|B)=0.1$，$P(A|\overline{B})=0.95$．

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

$$P(B|A)=\frac{P(B)P(A|B)}{P(B)P(A|B)+P(\overline{B})P(A|\overline{B})}$$

$$=\frac{0.8\times0.1}{0.8\times0.1+0.2\times0.95}$$

≈0.30．

(2)依据题意知 $P(B)=0.8$，$P(\overline{B})=0.2$，$P(\overline{A}|B)=0.9$，$P(\overline{A}|\overline{B})=0.05$．

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

$$P(B|\overline{A})=\frac{P(B)P(\overline{A}|B)}{P(B)P(\overline{A}|B)+P(\overline{B})P(\overline{A}|\overline{B})}$$

$$=\frac{0.8\times0.9}{0.8\times0.9+0.2\times0.05}$$

≈0.99．

反思提炼

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 $P(B_i|A)=\frac{P(B_iA)}{P(A)}$，$P(B_iA)=P(B_i)\cdot P(A|B_i)$，全概率公式 $P(A)=\sum\limits_{i=1}^{n}P(B_i)P(A|B_i)$ 的综合应用．

3. 如果某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，那么：(1)若要求的是第二阶段某一个结果发生的概率，则用全概率公式；(2)若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．

探究训练

1. 某商店从甲、乙、丙三个厂购买了一批灯泡，甲厂占 25%，乙厂占 35%，丙厂占 40%，各厂的次品率分别为 5%，4%，2%．某消费者从该商店购买了一只灯泡．

(1)求消费者买到一只次品灯泡的概率；

(2)若消费者买到一只次品灯泡，则它是哪个厂生产的可能性最大？

解：记事件 B 表示“消费者买到一只次品灯泡”，A_1，A_2，A_3 分别表示买到的灯泡是甲、乙、丙厂生产的灯泡，根据题意，得

$P(A_1)=25\%$，$P(A_2)=35\%$，$P(A_3)=40\%$，

$P(B|A_1)=5\%$，$P(B|A_2)=4\%$，$P(B|A_3)=2\%$．

(1)$P(B)=\sum\limits_{i=1}^{3}P(A_i)P(B|A_i)=25\%\times5\%+35\%\times4\%+40\%\times2\%=0.034\ 5$．

(2)$P(A_1|B)=\frac{P(A_1B)}{P(B)}=\frac{P(A_1)P(B|A_1)}{P(B)}=\frac{25\%\times5\%}{0.034\ 5}\approx0.362\ 3$，

$P(A_2|B)=\frac{P(A_2B)}{P(B)}=\frac{P(A_2)P(B|A_2)}{P(B)}=\frac{35\%\times4\%}{0.034\ 5}\approx0.405\ 8$，

$P(A_3|B)=\frac{P(A_3B)}{P(B)}=\frac{P(A_3)P(B|A_3)}{P(B)}=$

$\frac{40\%\times2\%}{0.0345}\approx0.2319$.

所以它是乙厂生产的可能性最大.

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，则会随机选取一个选项.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“该考生选出正确答案”，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 $P(B)=P(A)P(B|A)+P(\overline{A})\cdot P(B|\overline{A})=0.85\times1+0.15\times0.25=0.8875$.

(2)由贝叶斯公式得 $P(A|B)=\frac{P(A)P(B|A)}{P(B)}=\frac{0.85\times1}{0.8875}\approx0.9577$.

体系构建

条件概率 $P(B|A)=\frac{P(AB)}{P(A)}$ → 乘法公式 $P(AB)=P(A)P(B|A)$

↓

全概率公式（由因求果）

$P(B)=P(BA_1)+P(BA_2)+\cdots+P(BA_n)$

$=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+\cdots+P(A_n)P(B|A_n)$

$\xrightarrow{\text{步骤}}$ 1.设事件　2.写概率　3.代公式

↓

*贝叶斯公式（执果寻因）

$P(A_i|B)=\frac{P(A_i)P(B|A_i)}{P(B)}=\frac{P(A_i)P(B|A_i)}{\sum_{k=1}^{n}P(A_k)P(B|A_k)}$，$i=1,2,\cdots,n$

课后素养评价（十）

基础性·能力运用

1. 某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 $\frac{3}{4}$；若他前一球投不进，则后一球投进的概率为 $\frac{1}{4}$.若他第 1 球投进的概率为 $\frac{3}{4}$，则他第 2 球投进的概率为　（B）

A. $\frac{3}{4}$　　B. $\frac{5}{8}$

C. $\frac{7}{16}$　　D. $\frac{9}{16}$

2. 已知 A,B 为样本空间 Ω 中的事件，BA 与 $B\overline{A}$ 是互斥的，$B=BA+B\overline{A}$，且 $P(AB)=\frac{1}{2}$，$P(\overline{A}B)=\frac{1}{3}$，则 $P(B)=$　（　）

A. $\frac{1}{2}$　　B. $\frac{1}{3}$

C. $\frac{1}{6}$　　D. $\frac{5}{6}$

D　**解析**：由互斥事件概率的加法公式，得 $P(B)=P(AB)+P(\overline{A}B)=\frac{1}{2}+\frac{1}{3}=\frac{5}{6}$.

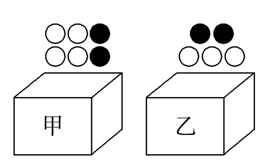
3. 已知甲袋中有 6 个红球、4 个白球；乙袋中有 8 个红球、6 个白球.如果随机取一个袋子，再从该袋中随机取一球，该球是红球的概率为 ________.

$\frac{41}{70}$　**解析**：设 B=“该球是红球”，A_1=“取自甲袋”，A_2=“取自乙袋”.

由题意知 $P(B|A_1)=\frac{6}{10}$，$P(B|A_2)=\frac{8}{14}$，

所以 $P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)=\frac{1}{2}\times\frac{6}{10}+\frac{1}{2}\times\frac{8}{14}=\frac{41}{70}$.

4. 如图，现有完全相同的甲、乙两个箱子，其中甲箱装有 2 个黑球和 4 个白球，乙箱装有 2 个黑球和 3 个白球，这些球除颜色外完全相同.某人先从两个箱子中任取一个箱子，再从中随机摸出一球.



(1)求摸出的球是黑球的概率；

(2)若已知摸出的球是黑球，请用概率公式判断该球取自哪个箱子的可能性更大.

解：(1)记事件 A 表示“球取自甲箱”，事件$\overline{A}$表示“球取自乙箱”，事件 B 表示“取得黑球”，

则 $P(A)=P(\overline{A})=\frac{1}{2}$，$P(B|A)=\frac{2}{6}=\frac{1}{3}$，

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

由全概率公式得 $P(B)=P(A)P(B|A)+P(\overline{A})\cdot P(B|\overline{A})=\frac{1}{2}\times\frac{1}{3}+\frac{1}{2}\times\frac{2}{5}=\frac{11}{30}$.

(2)该球取自乙箱的可能性更大.理由如下：

已知摸出的球是黑球，由条件概率得该球取自甲箱的概率为 $P(A|B)=\frac{P(A)P(B|A)}{P(B)}=\frac{\frac{1}{2}\times\frac{1}{3}}{\frac{11}{30}}=\frac{5}{11}$.

已知摸出的球是黑球，由条件概率得该球取自乙箱的概率为 $P(\overline{A}|B)=\frac{P(\overline{A})P(B|\overline{A})}{P(B)}=\frac{\frac{1}{2}\times\frac{2}{5}}{\frac{11}{30}}=\frac{6}{11}$.因为 $P(A|B)<P(\overline{A}|B)$，所以该球取自乙箱的可能性更大.

综合性·创新提升

1.有三个箱子，编号分别为 1，2，3.已知 1 号箱装有 1 个红球、4 个白球，2 号箱装有 2 个红球、3 个白球，3 号箱装有 3 个红球.某人从三个箱子中任取一个，从中任意摸出一球，取得红球的概率为 ( )

A.$\frac{8}{15}$ B.$\frac{7}{15}$

C.$\frac{5}{7}$ D.$\frac{7}{9}$

A **解析**：记 A_i＝“球取自 i 号箱”$(i=1,2,3)$，B＝“取得红球”.

B 发生总是伴随着 A_1，A_2，A_3 之一同时发生，

即 $B=A_1B+A_2B+A_3B$，且 A_1B，A_2B，A_3B 两两互斥，

则 $P(B)=P(A_1B)+P(A_2B)+P(A_3B)$，

即 $P(B)=\sum\limits_{i=1}^{3}P(A_i)P(B|A_i)$.

代入数据计算得 $P(B)=\frac{8}{15}$.

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.现随机地从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 $P(B)=\frac{4}{2+4}=\frac{2}{3}$，$P(\overline{B})=1-P(B)=\frac{1}{3}$.

$P(A|B)=\frac{3+1}{8+1}=\frac{4}{9}$，$P(A|\overline{B})=\frac{3}{8+1}=\frac{1}{3}$.

$P(A)=P(AB\cup A\overline{B})$

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

$=P(A|B)P(B)+P(A|\overline{B})P(\overline{B})$

$=\frac{4}{9}\times\frac{2}{3}+\frac{1}{3}\times\frac{1}{3}$

$=\frac{11}{27}$.

3.设某工厂有两个车间生产同一种产品，第一车间的次品率为 0.15，第二车间的次品率为 0.12，两个车间生产的产品都混合堆放在一个仓库，假设第一、二车间生产的产品数量之比为 2∶3.今有一客户从仓库中随机取一件产品，则该产品合格的概率为________.

0.868 **解析**：设 B＝“从仓库中随机取一件产品是合格品”，A_i＝“取出的产品是第 i 车间生产的产品”，$i=1,2$，则 $B=A_1B\cup A_2B$.

因为第一、二车间生产的产品数量之比为 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 0.8；用未校准的枪射击时，中靶的概率为 0.3.现该射手从 8 支枪中任取一支用于射击，结果中靶，则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 $P(B_1)=\frac{5}{8}$，$P(B_2)=\frac{3}{8}$，$P(A|B_1)=0.8$，

$P(A|B_2)=0.3$.

由贝叶斯公式，

得 $P(B_1|A)$

$$=\frac{P(A|B_1)P(B_1)}{P(A|B_1)P(B_1)+P(A|B_2)P(B_2)}$$

$$=\frac{0.8\times\frac{5}{8}}{0.8\times\frac{5}{8}+0.3\times\frac{3}{8}}=\frac{40}{49}.$$

所以所用的枪是校准过的概率为$\frac{40}{49}$.

5. 青团是江南地区的传统特色小吃.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.

$P(B_1)=\frac{C_5^2}{C_8^2}=\frac{5}{14}$，$P(B_2)=\frac{C_5^1C_3^1}{C_8^2}=\frac{15}{28}$，$P(B_3)=\frac{C_3^2}{C_8^2}=\frac{3}{28}$，$P(A|B_1)=\frac{2}{3}$，$P(A|B_2)=\frac{5}{9}$，$P(A|B_3)=\frac{4}{9}$.

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.了解随机变量的意义，理解随机变量的概念.(数学抽象)

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.(数学运算)

问题式预习

知识清单

知识点一　随机变量

一般地，对于随机试验样本空间 Ω 中的每个样本点 ω，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$p_i\geqslant 0$，$i=1,2,\cdots,n$；

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 $0\leqslant p\leqslant 1$)

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 $X=\begin{cases}1,\text{取出白球},\\0,\text{取出红球}\end{cases}$

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个,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 ω 相当于函数定义中的自变量,而样本空间 Ω 相当于函数的定义域,不同之处在于 Ω 不一定是数集.

(2)一次试验中变量 X 只有1和2两个取值,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

学习任务一　离散型随机变量的判定及取值

1. (多选)抛掷两枚骰子一次,记第一枚骰子掷出的点数减去第二枚骰子掷出的点数之差为 X,那么"$X\leqslant -4$"表示的随机试验的结果是　(　　)

A.第一枚1点、第二枚4点

B.第一枚2点、第二枚6点

C.第一枚1点、第二枚5点

D.第一枚1点、第二枚6点

BCD　**解析**:抛掷两枚骰子,点数之差满足小于等于-4的只有三种情况,即第一枚为1点、第二枚为6点,第一枚为1点、第二枚为5点,第一枚为2点、第二枚为6点.

2. 写出下列随机变量可能的取值,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 $1,2,3,\cdots,10$.

$X=k(k=1,2,\cdots,10)$表示"取出 k 号球".

(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,骰子甲得 i 点且骰子乙得 j 点".

X 的可能取值为 $2,3,4,\cdots,12$.

$X=2$ 表示$(1,1)$;$X=3$ 表示$(1,2),(2,1)$;$X=4$ 表示$(1,3),(2,2),(3,1)$;$\cdots$;$X=12$ 表示$(6,6)$.

反思提炼

用随机变量表示随机试验结果的问题的关键点和注意点

(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.

学习任务二　离散型随机变量分布列的性质

例 1　已知离散型随机变量 ξ 的分布列如下：

ξ	-1	0	1
P	$\frac{1}{2}$	$1-2q$	q^2

(1)上表中的 q 可以取任意实数吗？

(2)$\{\xi=-1\}$，$\{\xi=0\}$，$\{\xi=1\}$是两两互斥的事件，且$\{\xi=-1\}\cup\{\xi=0\}\cup\{\xi=1\}$是必然事件.由此思考，如何求 q 的值？

解：(1)由 $0\leqslant 1-2q\leqslant 1$，$0\leqslant q^2\leqslant 1$，解得 $0\leqslant q\leqslant \frac{1}{2}$.

(2)由$\frac{1}{2}+(1-2q)+q^2=1$，解得 $q=1-\frac{\sqrt{2}}{2}$ 或 $q=1+\frac{\sqrt{2}}{2}$(舍去).

反思提炼

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.

探究训练

1.若随机变量 X 的分布列如下表，则 a 的值为　　(　　)

X	1	2	3	4
P	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{4}$	a

A.1　　B.$\frac{1}{12}$　　C.$\frac{1}{3}$　　D.$\frac{1}{6}$

B　**解析**：由分布列的性质，有$\frac{1}{2}+\frac{1}{6}+\frac{1}{4}+a=1$，解得 $a=\frac{1}{12}$.

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**：根据概率分布列的性质，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 $P(0.5<X<2.5)=P(X=1)+P(X=2)=\frac{2}{3}+\frac{2}{9}=\frac{8}{9}$.

学习任务三　两点分布

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)从此袋中任意摸出两个球，用$X=0$表示“两个球全是白球”，用 $X=1$ 表示“两个球不全是白球”，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	$\frac{1}{7}$	$\frac{6}{7}$

反思提炼

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.

探究训练

1. 设某试验的成功率是失败率的 2 倍.用随机变量 ξ 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 $2p$，ξ 的分布列为

ξ	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 $P(X=1)=0.6$.

学习任务 四　离散型随机变量的分布列

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. 每次从此箱中任取一个球，若取出红球，则不再放回，直到取出白球为止，求取球次数 X 的分布列.

解：X 的可能取值为 1，2，3，

第 1 次取到白球的概率为 $P(X=1)=\frac{3}{5}$，

第 2 次取到白球的概率为 $P(X=2)=\frac{2}{5}\times\frac{3}{4}=\frac{3}{10}$，

第 3 次取到白球的概率为 $P(X=3)=\frac{2}{5}\times\frac{1}{4}\times1=\frac{1}{10}$.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. 每次从此箱中任取一个球，取出的球不再放回，直到取出所有白球为止，求取球次数 X 的分布列.

解：X 的所有可能取值为 3，4，5.

3 次取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 次取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 次取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	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{3}{5}$

反思提炼

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——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所对应的概率

↓

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

探究训练

1. 根据统计数据，甲、乙两名射击运动员，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环的概率.

解：(1)当甲命中环数高于乙命中环数时，只有一种情况：甲命中10环，且乙命中9环，这时概率$p'=\frac{2}{5}\times\frac{5}{6}=\frac{1}{3}$.所以甲命中的环数不高于乙命中的环数的概率$p=1-p'=\frac{2}{3}$.

(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)甲、乙命中的环数之和低于52环时，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环的概率$p=1-p_1=1-\frac{1}{216}=\frac{215}{216}$.

2. 设S是不等式$x^2-x-6\leqslant0$的解集，整数m，$n\in S$.

(1)记"有序数组(m,n)使得$m+n=0$成立"为事件A，试列举A包含的样本点；

(2)设$X=m^2$，求X的分布列.

解：(1)由$x^2-x-6\leqslant0$，得$-2\leqslant x\leqslant3$，

即$S=\{x|-2\leqslant x\leqslant3\}$.

因为$m,n\in\mathbf{Z}$，$m,n\in S$且$m+n=0$，

所以A包含的样本点为$(-2,2)$，$(2,-2)$，$(-1,1)$，$(1,-1)$，$(0,0)$.

(2)因为m的所有不同的取值为-2，-1，0，1，2，3，

所以$X=m^2$的所有不同的取值为0，1，4，9，

且有$P(X=0)=\frac{1}{6}$，$P(X=1)=\frac{2}{6}=\frac{1}{3}$，$P(X=4)=\frac{2}{6}=\frac{1}{3}$，$P(X=9)=\frac{1}{6}$.

故X的分布列为

X	0	1	4	9
P	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$

体系构建

离散型随机变量
- 定义
- 取值
- 分布列
  - 定义——设离散型随机变量X的可能取值为x_1，x_2，x_3，…，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，…，n为X的概率分布列，简称分布列
  - 表示方法：表格法、图形法
  - 性质——①$p_i\geqslant0$，$i=1,2,\cdots,n$　②$p_1+p_2+\cdots+p_n=1$
  - 特殊分布列——两点分布(或0–1分布)

课后素养评价(十一)

基础性·能力运用

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,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,现在有放回地随机取出 2 个球,设 2 个球的号码之和为随机变量 ξ,则 ξ 所有可能取值的个数是　(　　)

A.25　B.10　C.15　D.9

D　**解析**:由题意得两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.随机变量 X 的所有可能取值是 $-2,0,3,5$,且 $P(X=-2)=\frac{1}{4}$,$P(X=3)=\frac{1}{2}$,$P(X=5)=\frac{1}{12}$,则 $P(X=0)$的值为　(　　)

A.0　B.$\frac{1}{4}$　C.$\frac{1}{6}$　D.$\frac{1}{8}$

C　**解析**:因为 $P(X=-2)+P(X=0)+P(X=3)+P(X=5)=1$,即$\frac{1}{4}+P(X=0)+\frac{1}{2}+\frac{1}{12}=1$,所以$P(X=0)=\frac{2}{12}=\frac{1}{6}$.

4.设随机变量 X 等可能地从 1,2,3,4,…,10 中取值.若随机变量 $Y=2X-1$,则 $P(Y<6)$的值为(　　)

A.0.3　B.0.5　C.0.1　D.0.2

A　**解析**:$Y<6$,即 $2X-1<6$,所以 $X<3.5$.所以 $X=1,2,3$,故 $P(Y<6)=0.3$.

5.某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.

6.一袋中装有 5 个球,编号分别为 1,2,3,4,5,从袋中一次随机取出 3 个球,以 ξ 表示取出的 3 个球中的最小编号,则随机变量 ξ 的分布列为　(　　)

A.

ξ	1	2	3
P	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

B.

ξ	1	2	3	4
P	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{3}{10}$	$\frac{2}{5}$

C.

ξ	1	2	3
P	$\frac{3}{5}$	$\frac{3}{10}$	$\frac{1}{10}$

D.

ξ	1	2	3
P	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{3}{5}$

C　**解析**:随机变量 ξ 的可能取值为 1,2,3,

$P(\xi=1)=\frac{C_4^2}{C_5^3}=\frac{3}{5}$,$P(\xi=2)=\frac{C_3^2}{C_5^3}=\frac{3}{10}$,$P(\xi=3)=\frac{C_2^2}{C_5^3}=\frac{1}{10}$.故选 C.

7.已知随机变量 X 的可能取值有 3 个,取这 3 个值的概率分别为 p_1,p_2,p_3,且 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$-\frac{1}{3}\leqslant d\leqslant\frac{1}{3}$.

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:由题可知 $m=1-0.2-0.1-0.1-0.3=0.3$,列

表为

X	0	1	2	3	4
$\|X-1\|$	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

η	0	1	2	3
P	0.1	0.3	0.3	0.3

综合性·创新提升

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 ξ 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 ξ 的分布列为

ξ	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.已知随机变量 ξ 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.(新定义)泊松分布是一种离散概率分布，适合于描述单位时间内随机事件发生的次数.泊松分布的概率分布列为 $P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}(k=0,1,2,\cdots)$，e 是自然对数的底数，$\lambda$ 是泊松分布的均值.已知钚 239 是钚的同位素，一纳克钚 239 每秒平均发生 2.3 次放射性衰变，假设衰变次数服从泊松分布，则 2 s 内一纳克钚 239 恰好发生 4 次放射性衰变的概率约为(参考数据：$e^{4.6}\approx99.48$，$2.3^4\approx27.98$) ( )

A.0.342　B.0.188　C.0.658　D.0.812

B **解析**：因为一纳克钚 239 每秒平均发生 2.3 次放射性衰变，所以 2 s 内一纳克钚 239 发生放射性衰变的均值为 4.6 次.

因为衰变次数服从泊松分布，所以 $\lambda=4.6$，

所以 $P(X=4)=\frac{4.6^4}{4!}\times e^{-4.6}=\frac{2^4\times2.3^4}{4!}\times\frac{1}{e^{4.6}}\approx\frac{2}{3}\times\frac{27.98}{99.48}\approx0.188$.故选 B.

5.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 ξ 的分布列为

ξ	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(不分期)，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 η 表示商场经销一件该商品的利润，求 η 的分布列.

解：η 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 η 的分布列为

η	200	250	300
P	0.4	0.4	0.2

6.从一批含有 10 个合格品与 3 个次品的产品中，一个一个地抽取，设每个产品被抽到的可能性相同.在下列两种情况下，分别求出取到合格品所需抽取次数 X 的分布列.

(1)每次取出的产品都不放回到该批产品中；

(2)每次取出的产品都立即放回到该批产品中，然后抽取下一个产品.

解：(1)X 的所有可能取值为 1，2，3，4.

$P(X=1)=\frac{10}{13}$，

$P(X=2)=\frac{A_3^1C_{10}^1}{A_{13}^2}=\frac{5}{26}$，

$P(X=3)=\frac{A_3^2C_{10}^1}{A_{13}^3}=\frac{5}{143}$,

$P(X=4)=\frac{A_3^3C_{10}^1}{A_{13}^4}=\frac{1}{286}$.

故 X 的分布列为

X	1	2	3	4
P	$\frac{10}{13}$	$\frac{5}{26}$	$\frac{5}{143}$	$\frac{1}{286}$

(2)$X=1$,即第一次取出合格品,故 $P(X=1)=\frac{10}{13}$,

$X=2$,即第 2 次取到合格品,第 1 次取到不合格品,故 $P(X=2)=\frac{3\times10}{13\times13}=\frac{3}{13}\times\frac{10}{13}$,

……

$X=n(n\in\mathbf{N}^*,n\leqslant13)$,

即第 n 次取到合格品,前 $(n-1)$ 次取到的产品均不合格,故 $P(X=n)=\frac{3^{n-1}\times10}{13^n}=\left(\frac{3}{13}\right)^{n-1}\times\frac{10}{13}$.

故 X 的分布列为

X	1	2	…	n	…
P	$\frac{10}{13}$	$\frac{3}{13}\times\frac{10}{13}$	…	$\left(\frac{3}{13}\right)^{n-1}\times\frac{10}{13}$	…

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意义和性质.

2.会根据离散型随机变量的分布列求出均值.(数学抽象、数学运算)

3.会利用离散型随机变量的均值解决一些实际问题.(数学建模)

问题式预习

知识清单

知识点　离散型随机变量的均值及其性质

(1)定义: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 $E(X)=\underline{x_1p_1+x_2p_2+\cdots+x_np_n}=\underline{\sum\limits_{i=1}^{n}x_ip_i}$ 为随机变量 X 的均值或数学期望.

(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(正确的打"√",错误的打"×").

(1)随机变量的均值与样本的平均值是同一个概念.　(×)

(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,样本均值的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去估计随机变量的均值.

(3)通过教材例4的学习你有怎样的收获?

提示:决策问题首先要根据实际问题明确决策的目标:期望损失最小(期望收益最大等),然后计算每个方案的结果,再从期望损失最小(期望收益最大等)的角度做出合理的选择.决策问题的答案往往不唯一,只要有理有据即可,但我们多数从期望入手.

任务型课堂

学习任务一　离散型随机变量的均值公式及性质

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 $a=15$.

反思提炼

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. 求 $aX+b$ 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.

探究训练

1. 已知随机变量 ξ 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 ξ 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 ξ,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 ξ 的可能取值有1,2,3,4,$P(\xi=1)=\frac{C_6^2}{C_7^3}=\frac{15}{35}$,$P(\xi=2)=\frac{16}{35}$,$P(\xi=3)=\frac{C_3^2}{C_7^3}=\frac{3}{35}$,$P(\xi=4)=\frac{1}{C_7^3}=\frac{1}{35}$,所以 $E(\xi)=1\times\frac{15}{35}+2\times\frac{16}{35}+3\times\frac{3}{35}+4\times\frac{1}{35}=\frac{12}{7}$.

学习任务二　求离散型随机变量的均值

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)
天数	2	16	36
最高气温/℃	[25,30)	[30,35)	[35,40)
天数	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 $Y=1.2X$，试求 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)$E(Y)=E(1.2X)=1.2E(X)=1.2\times360=432$.

反思提炼

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 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 $E(X)$.

探究训练

1. 甲、乙两人分别独立参加招生面试.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 ξ 的数学期望是　　(　　)

A. $\frac{2}{3}$　　B. $\frac{11}{9}$

C. 1　　D. $\frac{8}{9}$

A　**解析**：依题意，ξ 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 ξ 的分布列为

ξ	0	1	2
P	$\frac{4}{9}$	$\frac{4}{9}$	$\frac{1}{9}$

故 ξ 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. 甲、乙两人对同一目标各射击一次，甲命中的概率为 $\frac{2}{3}$，乙命中的概率为 $\frac{4}{5}$.若命中目标的人数为 X，则 $E(X)=$________.

$\frac{22}{15}$　**解析**：由题意可知，X 的所有可能取值为 0，1，2.

因为 $P(X=0)=\left(1-\frac{2}{3}\right)\times\left(1-\frac{4}{5}\right)=\frac{1}{3}\times\frac{1}{5}=\frac{1}{15}$，

$P(X=1)=\left(1-\frac{2}{3}\right)\times\frac{4}{5}+\frac{2}{3}\times\left(1-\frac{4}{5}\right)=\frac{1}{3}\times\frac{4}{5}+\frac{2}{3}\times\frac{1}{5}=\frac{2}{5}$，

$P(X=2)=\frac{2}{3}\times\frac{4}{5}=\frac{8}{15}$.

所以 X 的分布列为

X	0	1	2
P	$\frac{1}{15}$	$\frac{2}{5}$	$\frac{8}{15}$

所以 $E(X)=0\times\frac{1}{15}+1\times\frac{2}{5}+2\times\frac{8}{15}=\frac{22}{15}$.

3. 一个盒子里有 1 个红球、1 个绿球、2 个黄球，每次从中随机取出一个，不放回，取出红球即停.设取出黄球的个数为 ξ，则 $P(\xi=0)=$________，$E(\xi)=$________.

$\frac{1}{3}$　1　**解析**：由题知，ξ 的可能取值为 0，1，2，

$P(\xi=0)=\frac{1}{4}+\frac{1}{4}\times\frac{1}{3}=\frac{1}{3}$，

$P(\xi=1)=\frac{2}{4}\times\frac{1}{3}+\frac{2}{4}\times\frac{1}{3}\times\frac{1}{2}+\frac{1}{4}\times\frac{2}{3}\times\frac{1}{2}$

$=\frac{1}{3}$，

$P(\xi=2)=1-\frac{1}{3}-\frac{1}{3}=\frac{1}{3}$.

所以 $E(\xi)=0\times\frac{1}{3}+1\times\frac{1}{3}+2\times\frac{1}{3}=1$.

学习任务 三　两点分布的均值

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 $E(X)=0\times\frac{1}{2}+1\times\frac{1}{2}=\frac{1}{2}$.

反思提炼

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 $P(X=0)+P(X=1)=1$.

探究训练

在掷一枚图钉的随机试验中，令 $X=\begin{cases}1,针尖向上,\\0,针尖向下.\end{cases}$ 如果针尖向上的概率为 $\frac{2}{5}$，那么试写出随机变量 X 的分布列并求其均值.

解：根据分布列的性质，针尖向下的概率是 $1-\frac{2}{5}=\frac{3}{5}$.

则随机变量 X 的分布列为

X	0	1
P	$\frac{3}{5}$	$\frac{2}{5}$

$E(X)=0\times\frac{3}{5}+1\times\frac{2}{5}=\frac{2}{5}$.

学习任务 四　离散型随机变量的均值的应用

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，甲的射箭水平比乙高.

反思提炼

1.实际生活中的均值问题

均值在实际生活中有着广泛的应用，如体育比赛的安排和成绩预测，消费预测，工程方案的预测，产品合格率的预测，投资收益的预测等，都可以通过计算随机变量的均值来进行.

2.概率模型的解题步骤

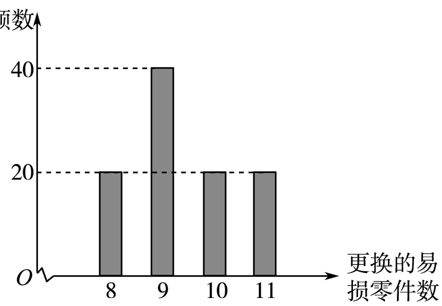
(1)审题，确定实际问题是哪一种概率模型，可能用到的事件类型，所用的公式有哪些；

(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，得出概率、均值等所表示的结论.

探究训练

1.某公司计划购买 2 台机器，该种机器使用三年后即被淘汰.机器有一易损零件，在购进机器时，可以额外购买这种零件作为备件，每个 200 元.在机器使用期间，如果备件不足再购买，则每个 500 元.现需对购买机器时应同时购买几个易损零件作出决策，为此搜集并整理了 100 台这种机器在三年使用期内更换的易损零件数，得到下面的统计图.



以这 100 台机器更换的易损零件数的频率代替每台机器更换的易损零件数的概率，记 X 表示 2 台机器三年内共需更换的易损零件数，n 表示购买 2 台机器时应同时购买的易损零件数.

(1)求 X 的分布列；

(2)若要求 $P(X\leqslant n)\geqslant 0.5$，确定 n 的最小值；

(3)以购买易损零件所需费用的期望值为决策依据，在 $n=19$ 与 $n=20$ 之中选择一个，应选择哪个？

解：(1)由题意可得，1 台机器在三年内需更换的易损零件数为 8，9，10，11 的概率分别为 0.2，0.4，0.2，0.2.

从而 $P(X=16)=0.2\times0.2=0.04$，

$P(X=17)=2\times0.2\times0.4=0.16$，

$P(X=18)=2\times0.2\times0.2+0.4\times0.4=0.24$，

$P(X=19)=2\times0.2\times0.2+2\times0.4\times0.2=0.24$，

$P(X=20)=2\times0.2\times0.4+0.2\times0.2=0.2$，

$P(X=21)=2\times0.2\times0.2=0.08$，

$P(X=22)=0.2\times0.2=0.04$.

所以 X 的分布列为

X	16	17	18	19	20	21	22
P	0.04	0.16	0.24	0.24	0.2	0.08	0.04

(2)由(1)知 $P(X\leqslant 18)=0.44$，$P(X\leqslant 19)=0.68$，故 n 的最小值为 19.

(3)记 Y 表示 2 台机器在购买易损零件上所需的费用(单位：元).

当 $n=19$ 时，

$E(Y)=19\times200\times0.68+(19\times200+500)\times0.2+(19\times200+2\times500)\times0.08+(19\times200+3\times500)\times0.04=4\ 040$.

当 $n=20$ 时，

$E(Y)=20\times200\times0.88+(20\times200+500)\times0.08+(20\times200+2\times500)\times0.04=4\ 080$.

可知当 $n=19$ 时所需费用的期望值小于当 $n=20$ 时所需费用的期望值，故应选 $n=19$.

2.《周易》包括《经》和《传》两个部分，《经》主要是六十四卦和三百八十四爻，它反映了中国古代的二进制计数的思想方法.若把阳爻“⚊”当作数字“1”，把阴爻“⚋”当作数字“0”，则六十四卦表示的数如下：

卦名	符号	二进制数	十进制数
坤	䷁	000000	0
剥	䷖	000001	1
比	䷇	000010	2
观	䷓	000011	3
…	…	…	…

(1)成语“否极泰来”包含了“否”卦䷋和“泰”卦䷊，试分别写出这两个卦所表示的十进制数.

(2)求由四个阳爻和两个阴爻构成的卦表示的十进制数的和.

(3)在由三个阳爻和三个阴爻构成的卦中任取一卦.若三个阳爻均相邻，则记 5 分；若只有两个阳爻相邻，则记 2 分；若三个阳爻均不相邻，则记 1 分.设任取一卦后的得分为随机变量 X，求 X 的分布列和数学期望.

解：(1)“否”卦所表示的二进制数为 000111，转化为十进制数是 $1\times2^0+1\times2^1+1\times2^2=7$，

“泰”卦所表示的二进制数为 111000，转化为十进制数是 $0\times2^0+0\times2^1+0\times2^2+1\times2^3+1\times2^4+1\times2^5=56$.

(2)因为卦由四个阳爻和两个阴爻构成，

所以这些卦所表示的二进制数共有 $C_6^2=15$(个)，分别为

001111，010111，011011，011101，011110，

100111，101011，101101，101110，110011，

110101,110110,111001,111010,111100.

因为这 15 个数中,每个位置都出现了 5 次 0,10 次 1,

所以这些卦表示的十进制数的和为 $10\times(2^0+2^1+2^2+2^3+2^4+2^5)=630$.

(3)依题意可得 X 的可能取值为 1,2,5,

$P(X=1)=\dfrac{C_4^3}{C_6^3}=\dfrac{4}{20}=\dfrac{1}{5}$,

$P(X=2)=\dfrac{A_4^2}{C_6^3}=\dfrac{12}{20}=\dfrac{3}{5}$,

$P(X=5)=\dfrac{C_4^1}{C_6^3}=\dfrac{4}{20}=\dfrac{1}{5}$,

所以 X 的分布列为

X	1	2	5
P	$\frac{1}{5}$	$\frac{3}{5}$	$\frac{1}{5}$

所以 $E(X)=1\times\dfrac{1}{5}+2\times\dfrac{3}{5}+5\times\dfrac{1}{5}=\dfrac{12}{5}$.

体系构建

离散型随机变量的均值
- 概念 $E(X)=x_1p_1+x_2p_2+\cdots+x_np_n=\sum\limits_{i=1}^{n}x_ip_i$
- 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)

基础性·能力运用

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. 某船队若出海后天气好,可获得 5 000 元;若出海后天气坏,将损失 2 000 元.根据预测知天气好的概率为 0.6,则出海的期望效益是　(　　)

A.2 000 元　　B.2 200 元

C.2 400 元　　D.2 600 元

B　**解析**:由题意,出海的期望效益 $E(X)=5\,000\times0.6+(-2\,000)\times(1-0.6)=3\,000-800=2\,200$(元).

4. 有 10 件产品,其中 3 件是次品,从中任取 2 件.若 ξ 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**:ξ 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 ξ 的分布列为

ξ	0	1	2
P	$\frac{7}{15}$	$\frac{7}{15}$	$\frac{1}{15}$

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 ξ 的均值是　(　　)

A.$\dfrac{4}{3}$　　B.$\dfrac{13}{9}$　　C.$\dfrac{5}{3}$　　D.$\dfrac{13}{7}$

B　**解析**:试验次数ξ的可能取值为1,2,3,

$P(\xi=1)=\frac{2}{3}$,$P(\xi=2)=\frac{1}{3}\times\frac{2}{3}=\frac{2}{9}$,

$P(\xi=3)=\frac{1}{3}\times\frac{1}{3}\times\left(\frac{2}{3}+\frac{1}{3}\right)=\frac{1}{9}$.

所以ξ的分布列为

ξ	1	2	3
P	$\frac{2}{3}$	$\frac{2}{9}$	$\frac{1}{9}$

所以$E(\xi)=1\times\frac{2}{3}+2\times\frac{2}{9}+3\times\frac{1}{9}=\frac{13}{9}$.

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$E(X_1)=1\times0.4+2\times0.1+3\times0.5=2.1$,$E(X_2)=1\times0.1+2\times0.6+3\times0.3=2.2$,得$E(X_2)>E(X_1)$,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**:依题意得,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,一位客人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ξ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**:随机变量ξ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ξ的分布列为

ξ	1	3
P	0.76	0.24

所以$E(\xi)=1\times0.76+3\times0.24=1.48$.

综合性·创新提升

1.某射击运动员射击所中环数ξ的分布列为

ξ	7	8	9	10
P	x	0.1	0.3	y

若$E(\xi)=8.9$,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,每人去的区均不相同.记 ξ 为 4 人中没有去到自家所在区工作的人数,则 $E(\xi)$ 等于 (　　)

A. $\frac{3}{2}$　　B. $\frac{5}{2}$

C. 1　　D. 3

D　**解析**:由题意得 ξ 的可能取值为 4,3,2,0,

$P(\xi=4)=\frac{3\times(1+2)}{A_4^4}=\frac{9}{24}$,

$P(\xi=3)=\frac{C_4^3\times 2}{A_4^4}=\frac{8}{24}$,

$P(\xi=2)=\frac{C_4^2\times 1}{A_4^4}=\frac{6}{24}$,$P(\xi=0)=\frac{1}{A_4^4}=\frac{1}{24}$.

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 ξ,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,ξ 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.如果李明决定参加考试,设他每次参加考试通过的概率依次为 0.6,0.7,0.8,0.9,求在一年内李明参加考试次数 X 的分布列和均值.

解:由题可知,X 的所有可能取值为 1,2,3,4,

$X=1$ 表示李明第一次参加考试就通过了,

故 $P(X=1)=0.6$.

$X=2$ 表示李明第一次考试未通过,第二次考试通过了,

故 $P(X=2)=(1-0.6)\times0.7=0.28$.

$X=3$ 表示李明第一、二次考试均未通过,第三次考试通过了,

故 $P(X=3)=(1-0.6)\times(1-0.7)\times0.8=0.096$.

$X=4$ 表示李明第一、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. 某理财公司有 A 和 B 两种理财产品,每种理财产品的不同投资结果之间相互独立,这两种理财产品一年后盈亏的情况如下(注:$p>0$,$q>0$):

产品 A

投资结果	获利 40%	不赔不赚	亏损 20%
概率	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

产品 B

投资结果	获利 20%	不赔不赚	亏损 10%
概率	p	$\frac{1}{3}$	q

(1)已知甲、乙两人分别选择了产品 A 和产品 B 投资,如果一年后他们中至少有一人获利的概率大于 $\frac{3}{5}$,求 p 的取值范围;

(2)若丙要将 10 万元进行投资,以一年后投资收益的期望值为决策依据,则选择哪种产品投资较理想?

解:(1)记事件 A 为"甲选择产品 A 盈利",事件 B 为"乙选择产品 B 盈利",事件 C 为"一年后甲、乙两人中至少有一人投资获利",

则 $P(\overline{A})=\frac{2}{3}$,$P(\overline{B})=1-p$.

所以 $P(C)=1-P(\overline{A}\ \overline{B})=1-\frac{2}{3}(1-p)=\frac{1}{3}+\frac{2p}{3}>\frac{3}{5}$,

解得 $p>\frac{2}{5}$.

又因为 $p+\frac{1}{3}+q=1,q>0$,

所以 $p<\frac{2}{3}$.

所以 $\frac{2}{5}<p<\frac{2}{3}$.

故实数 p 的取值范围是 $\left(\frac{2}{5},\frac{2}{3}\right)$.

(2)假设丙选择产品 A 进行投资,且记 X(单位:万元)为获利金额,则随机变量 X 的分布列为

X	4	0	-2
P	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

则 $E(X)=4\times\frac{1}{3}+0\times\frac{1}{2}+(-2)\times\frac{1}{6}=1$.

假设丙选择产品 B 进行投资,且记 Y(单位:万元)为获利金额,则随机变量 Y 的分布列为

Y	2	0	-1
P	p	$\frac{1}{3}$	q

则 $E(Y)=2\times p+0\times\frac{1}{3}+(-1)\times q=2p-q=2p-\left(\frac{2}{3}-p\right)=3p-\frac{2}{3}\left(0<p<\frac{2}{3}\right)$.

当 $p=\frac{5}{9}$ 时,$E(X)=E(Y)$,选择产品 A 和产品 B 一年后投资收益的数学期望相同,可以在产品 A 和产品 B 中任选一个;

当 $0<p<\frac{5}{9}$ 时,$E(X)>E(Y)$,选择产品 A 一年后投资收益的数学期望较大,应选产品 A;

当 $\frac{5}{9}<p<\frac{2}{3}$ 时,$E(X)<E(Y)$,选择产品 B 一年后投资收益的数学期望较大,应选产品 B.

综上所述,当 $p=\frac{5}{9}$ 时,选择产品 A 或产品 B 都可以;当 $0<p<\frac{5}{9}$ 时,应选产品 A;当 $\frac{5}{9}<p<\frac{2}{3}$ 时,应选产品 B.

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.(数学抽象)

2.掌握方差的性质,会利用公式求离散型随机变量的方差.(数学运算)

3.会计算离散型随机变量的方差,并能解决一些实际问题.(数学建模、数学运算)

问题式预习

知识清单

知识点　方差、标准差的定义及方差的性质

(1)方差及标准差的定义:

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

①方差 $D(X)=\underline{(x_1-E(X))^2p_1+(x_2-E(X))^2p_2+\cdots+(x_n-E(X))^2p_n}=\underline{\sum_{i=1}^{n}(x_i-E(X))^2p_i}$.

②称 $\sqrt{D(X)}$ 为随机变量 X 的标准差,记为 $\sigma(X)$.

(2)方差的性质:$D(X+b)=\underline{D(X)}$;$D(aX)=\underline{a^2D(X)}$;$D(aX+b)=\underline{a^2D(X)}$.

概念辨析

1.判断(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)离散型随机变量的方差与标准差的单位是相同的.　(×)

(3)若 a 是常数,则 $D(a)=0$.　(√)

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，

则 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

$D(Y)=D(aX+b)$

$=\sum\limits_{i=1}^{n}(ax_i+b-E(Y))^2p_i$

$=\sum\limits_{i=1}^{n}(ax_i+b-aE(X)-b)^2p_i$

$=\sum\limits_{i=1}^{n}(ax_i-aE(X))^2p_i$

$=a^2\sum\limits_{i=1}^{n}(x_i-E(X))^2p_i$

$=a^2D(X)$.

任务型课堂

学习任务一 计算离散型随机变量的方差

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	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

则 $E(X)=1\times\frac{1}{3}+3\times\frac{1}{2}+5\times\frac{1}{6}=\frac{8}{3}$，

$D(X)=\left(1-\frac{8}{3}\right)^2\times\frac{1}{3}+\left(3-\frac{8}{3}\right)^2\times\frac{1}{2}+\left(5-\frac{8}{3}\right)^2\times\frac{1}{6}=\frac{17}{9}$.

[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：X 的可能取值为1，2，3.

$P(X=1)=\frac{A_1^1}{A_3^1}=\frac{1}{3}$；$P(X=2)=\frac{A_2^1A_1^1}{A_3^2}=\frac{1}{3}$；$P(X=3)=\frac{A_2^2}{A_3^3}=\frac{1}{3}$.

则 X 的分布列为

X	1	2	3
P	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

$E(X)=1\times\frac{1}{3}+2\times\frac{1}{3}+3\times\frac{1}{3}=2$，

$D(X)=(1-2)^2\times\frac{1}{3}+(2-2)^2\times\frac{1}{3}+(3-2)^2\times\frac{1}{3}=\frac{2}{3}$.

反思提炼

求随机变量 X 的均值与方差的方法

(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.

探究训练

1.已知随机变量 ξ 满足 $P(\xi=1)=0.3$，$P(\xi=2)=0.7$，

则 $E(\xi)$ 和 $D(\xi)$ 的值分别为　　(　　)

A.0.6 和 0.7　　B.1.7 和 0.09

C.0.3 和 0.7　　D.1.7 和 0.21

D　**解析**:$E(\xi)=1\times0.3+2\times0.7=1.7$,

$D(\xi)=(1-1.7)^2\times0.3+(2-1.7)^2\times0.7=0.21$.

2.已知随机变量 X 的分布列为

X	0	1	2
P	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

设 $Y=2X+3$,则 $D(Y)=$　　(　　)

A.$\frac{8}{3}$　　B.$\frac{5}{3}$　　C.$\frac{2}{3}$　　D.$\frac{1}{3}$

A　**解析**:因为 $E(X)=0\times\frac{1}{3}+1\times\frac{1}{3}+2\times\frac{1}{3}=1$,所以 $D(X)=(0-1)^2\times\frac{1}{3}+(1-1)^2\times\frac{1}{3}+(2-1)^2\times\frac{1}{3}=\frac{2}{3}$.所以 $D(Y)=D(2X+3)=2^2D(X)=\frac{8}{3}$.

学习任务二　方差的实际应用

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?

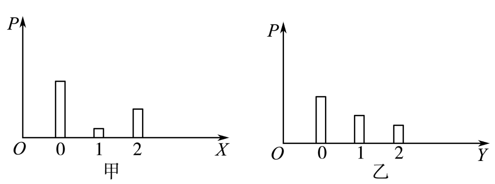
(3)如何定量刻画两名工人技术的稳定性?

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)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.



(3)可通过离散型随机变量的方差来定量刻画两名工人技术的稳定性.

计算可得 $D(X)\approx0.72$,$D(Y)=0.6$,$D(Y)<D(X)$,所以乙工人的技术更稳定.

反思提炼

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.

探究训练

1.(多选)猜歌名游戏是根据歌曲的主旋律制成的铃声来猜歌名.某嘉宾参加猜歌名游戏,每次猜歌名的结果相互独立,其猜对 A,B,C 三首歌曲歌名的概率及猜对时获得的奖金的情况是:猜对歌曲 A 的概率为 0.8,可获得奖金 1 万元;猜对歌曲 B 的概率为 0.5,可获得奖金 2 万元;猜对歌曲 C 的概率为0.5,可获得奖金 3 万元.规则如下:按照 A,B,C 的顺序猜,只有猜对当前歌曲的歌名才有资格猜下一首.记该嘉宾获得的奖金总额为 X 万元,则　　(　　)

A.$P(X=3)=0.4$

B.$E(2X+1)=5.4$

C.$D(2X+1)=18.24$

D.获得奖金的期望值与猜歌顺序无关

BC　**解析**:由题意,可分别用事件 A,B,C 表示猜对 A,B,C 三首歌曲的歌名,则事件 A,B,C 相互独立.

按照 A,B,C 的顺序猜,则 X 的取值可能为 0,1,3,6,

$P(X=0)=P(\overline{A})=0.2$,$P(X=1)=P(A\overline{B})=0.8\times0.5=0.4$,$P(X=3)=P(AB\overline{C})=0.8\times0.5\times0.5=0.2$,$P(X=6)=P(ABC)=0.8\times0.5\times0.5=0.2$,

则 $E(X)=0\times0.2+1\times0.4+3\times0.2+6\times0.2=$

2.2，

$D(X)=(0-2.2)^2\times0.2+(1-2.2)^2\times0.4+(3-2.2)^2\times0.2+(6-2.2)^2\times0.2=4.56$，

故 $E(2X+1)=2\times2.2+1=5.4$，$D(2X+1)=4\times4.56=18.24$，

由此可知 A 错误，B 正确，C 正确.

假设按照 A，C，B 的顺序猜，设 Y 表示此时获得的奖金总额，则 Y 的取值可能为 0，1，4，6，

$P(Y=0)=P(\overline{A})=0.2$，$P(Y=1)=P(A\overline{C})=0.8\times0.5=0.4$，$P(Y=4)=P(AC\overline{B})=0.8\times0.5\times0.5=0.2$，$P(Y=6)=P(ACB)=0.8\times0.5\times0.5=0.2$，

则 $E(Y)=0\times0.2+1\times0.4+4\times0.2+6\times0.2=2.4$，与按照 A，B，C 的顺序猜的期望值不同，故 D 错误.

故选 BC.

2. 甲、乙两名射击运动员在一次射击中射中的环数分别为两个相互独立的随机变量 ξ,η. 已知甲、乙两名射击运动员在每次射击中射中的环数均大于 6，且甲射中的环数为 10，9，8，7 的概率分别为 0.5，$3a$，a，0.1，乙射中的环数为 10，9，8 的概率分别为 0.3，0.3，0.2.

(1)求 ξ,η 的分布列；

(2)求 ξ,η 的均值与方差，并以此比较甲、乙的射击技术.

解：(1)依据题意得 $0.5+3a+a+0.1=1$，

解得 $a=0.1$.

因为乙射中的环数为 10，9，8 的概率分别为 0.3，0.3，0.2，

所以乙射中 7 环的概率为 $1-(0.3+0.3+0.2)=0.2$.

所以 ξ,η 的分布列分别为

ξ	10	9	8	7
P	0.5	0.3	0.1	0.1

η	10	9	8	7
P	0.3	0.3	0.2	0.2

(2)结合(1)中 ξ,η 的分布列，可得 $E(\xi)=10\times0.5+9\times0.3+8\times0.1+7\times0.1=9.2$，

$E(\eta)=10\times0.3+9\times0.3+8\times0.2+7\times0.2=8.7$，

$D(\xi)=(10-9.2)^2\times0.5+(9-9.2)^2\times0.3+(8-9.2)^2\times0.1+(7-9.2)^2\times0.1=0.96$，

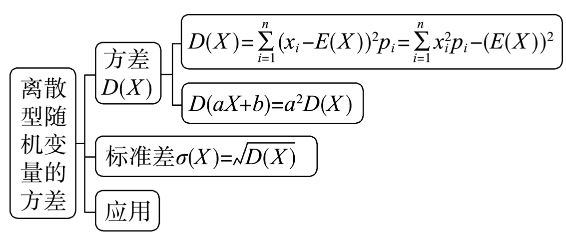
$D(\eta)=(10-8.7)^2\times0.3+(9-8.7)^2\times0.3+(8-8.7)^2\times0.2+(7-8.7)^2\times0.2=1.21$.

由于 $E(\xi)>E(\eta)$，说明甲平均射中的环数比乙大.

又 $D(\xi)<D(\eta)$，说明甲射中的环数比乙集中，比较稳定.

所以甲的射击技术比乙好.

体系构建



课后素养评价(十三)

基础性·能力运用

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. 已知随机变量 X 的可能取值为 0，1，2，若 $P(X=0)=\frac{1}{5}$，$E(X)=1$，则 X 的标准差为 (　　)

A. $\frac{2}{5}$　　B. $\frac{4}{5}$　　C. $\frac{\sqrt{10}}{5}$　　D. $\frac{4\sqrt{5}}{5}$

C　**解析**:设 $P(X=1)=p$,则 $P(X=2)=\frac{4}{5}-p$.

由 $E(X)=p+2\left(\frac{4}{5}-p\right)=1$,解得 $p=\frac{3}{5}$.

所以 $D(X)=1\times\frac{1}{5}+0\times\frac{3}{5}+1\times\frac{1}{5}=\frac{2}{5}$,

则标准差 $\sqrt{D(X)}=\frac{\sqrt{10}}{5}$.

3. 如果 X 是离散型随机变量,$E(X)=6$,$D(X)=0.5$,$X_1=2X-5$,那么下列结论正确的是　(　　)

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**: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. (多选)已知投资 A,B 两种项目获得的收益分别为 X,Y(单位:百万),X,Y 的分布列如表所示,则　(　　)

X	-1	0	2
P	0.2	m	0.6

Y	0	1	2
P	0.3	0.4	n

A. $m+n=0.5$

B. $E(2X+1)=4$

C. 投资两种项目的期望收益一样多

D. 投资 A 项目的风险比 B 项目高

ACD　**解析**:依题意可得 $0.2+m+0.6=1$,所以 $m=0.2.0.3+0.4+n=1$,所以 $n=0.3$.

$m+n=0.5$,故 A 正确;

$E(X)=(-1)\times0.2+0\times0.2+2\times0.6=1$,则 $E(2X+1)=2E(X)+1=3$,故 B 错误;

因为 $E(Y)=0\times0.3+1\times0.4+2\times0.3=1$,所以 $E(X)=E(Y)$,故 C 正确;

因为 $D(X)=(-1-1)^2\times0.2+(0-1)^2\times0.2+(2-1)^2\times0.6=1.6$,$D(Y)=(0-1)^2\times0.3+(1-1)^2\times0.4+(2-1)^2\times0.3=0.6$,即 $D(X)>D(Y)$,所以投资 A 项目的风险比 B 项目高,故 D 正确.故选 ACD.

6. (多选)编号为 1,2,3 的三位学生随意入座编号为 1,2,3 的三个座位,每位学生坐一个座位,设坐在与自己编号相同的座位上的学生的人数是 ξ,则　(　　)

A. ξ 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**:ξ 的所有可能取值为 0,1,3,$\xi=0$ 表示三位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 $P(\xi=0)=\frac{2}{A_3^3}=\frac{1}{3}$;$\xi=1$ 表示三位学生只有 1 位学生坐在了与自己编号相同的座位上,则 $P(\xi=1)=\frac{C_3^1}{A_3^3}=\frac{1}{2}$;$\xi=3$ 表示三位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 ξ 的分布列为

ξ	0	1	3
P	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

$E(\xi)=0\times\frac{1}{3}+1\times\frac{1}{2}+3\times\frac{1}{6}=1$.

$D(\xi)=\frac{1}{3}\times(0-1)^2+\frac{1}{2}\times(1-1)^2+\frac{1}{6}\times(3-1)^2=1$.

综合性·创新提升

1.已知随机变量 X 的分布列如下表：

X	1	2	3
P	$\frac{1}{2}$	x	y

若 $E(X)=\frac{15}{8}$，则 $D(X)$ 等于（　　）

A.$\frac{7}{32}$　　B.$\frac{9}{32}$　　C.$\frac{33}{64}$　　D.$\frac{55}{64}$

D　**解析**：由题知 $\begin{cases}1\times\frac{1}{2}+2x+3y=\frac{15}{8},\\ \frac{1}{2}+x+y=1,\end{cases}$ 得 $\begin{cases}x=\frac{1}{8},\\ y=\frac{3}{8}.\end{cases}$

所以 $D(X)=\left(1-\frac{15}{8}\right)^2\times\frac{1}{2}+\left(2-\frac{15}{8}\right)^2\times\frac{1}{8}+\left(3-\frac{15}{8}\right)^2\times\frac{3}{8}=\frac{55}{64}$.

2.已知随机变量 ξ 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 ξ 的分布列如下表：

ξ	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，ξ 的分布列为

ξ	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 ξ 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________.

ξ	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 ξ 的分布列为

ξ	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 ξ，则 $P(\xi=0)$ $=$________，$D(\xi)=$________.

$\frac{1}{3}$　$\frac{2}{3}$　**解析**：$\xi=0$ 表示停止取球时没有取到黄球，所以 $P(\xi=0)=\frac{1}{4}+\frac{1}{4}\times\frac{1}{3}=\frac{1}{3}$.随机变量 ξ 的所有可能取值为 0，1，2，$P(\xi=1)=\frac{2}{4}\times\frac{1}{3}+\frac{2}{4}\times\frac{1}{3}\times\frac{1}{2}+\frac{1}{4}\times\frac{2}{3}\times\frac{1}{2}=\frac{1}{3}$，$P(\xi=2)=\frac{2}{4}\times\frac{1}{3}\times\frac{1}{2}+\frac{1}{4}\times\frac{2}{3}\times\frac{1}{2}+\frac{2}{4}\times\frac{1}{3}\times\frac{1}{2}+\frac{2}{4}\times\frac{1}{3}\times\frac{1}{2}=\frac{1}{3}$，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.某投资公司准备在明年年初将 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、项目二获利的期望相同，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　二项分布的概念

学习任务目标

1.通过具体实例，了解伯努利试验，了解二项分布的概念.(数学抽象)

2.能利用二项分布概率模型解决简单的实际问题.(数学建模、数学运算)

问题式预习

知识清单

知识点一　n 重伯努利试验

(1)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 $\underline{X\sim B(n,p)}$.

由二项式定理，容易得到 $\sum\limits_{k=0}^{n}P(X=k)=\underline{\sum\limits_{k=0}^{n}C_n^kp^k(1-p)^{n-k}=[p+(1-p)]^n=1}$.

概念辨析

1.判断（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，它就有两个结果，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，则 $C_n^kp^k(1-p)^{n-k}$ 就是二项式 $[(1-p)+p]^n$ 的展开式的通项.

任务型课堂

学习任务一　求伯努利试验的概率

例 1　甲、乙两人每次射击，击中目标的概率分别是 $\frac{2}{3}$ 和 $\frac{3}{4}$，假设每次射击是否击中目标相互之间没有影响.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，故 $P(A)=1-P(\overline{A})=1-\left(\frac{2}{3}\right)^3=\frac{19}{27}$.

[一题多思]

思考 1.求甲射击 3 次恰好击中目标 2 次的概率.

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.求两人各射击 2 次，甲恰好击中目标 2 次且乙恰好击中目标 1 次的概率.

解：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 $P(C_1)=C_2^2\times\left(\frac{2}{3}\right)^2=\frac{4}{9}$，$P(C_2)=C_2^1\times\left(\frac{3}{4}\right)^1\times\left(1-\frac{3}{4}\right)=\frac{3}{8}$.由于甲、乙射击相互独立，故 $P(C_1C_2)=\frac{4}{9}\times\frac{3}{8}=\frac{1}{6}$.

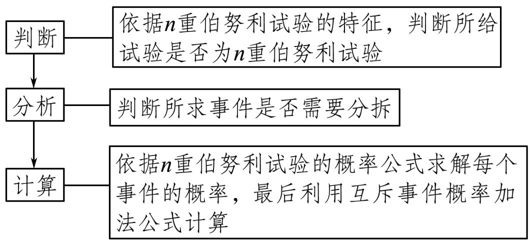
例 2　在平面直角坐标系中，位于原点的一个质点 P 按下列规则移动：质点每次移动一个单位长度，移动的方向为向左或向右，并且向左移动的概率为 $\frac{1}{3}$，向右移动的概率为 $\frac{2}{3}$.质点 P 移动五次后位于点 $(1,0)$ 的概率是　(　　)

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 $C_5^3\times\left(\frac{2}{3}\right)^3\times\left(\frac{1}{3}\right)^2=\frac{80}{243}$.

反思提炼

求 n 重伯努利试验的概率的步骤



探究训练

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"甲试跳成功"为事件 A，"乙试跳成功"为事件 B，"甲、乙两人中至少有一人成功"为事件 C.

由对立事件的概率计算公式得

$P(C)=1-P(\overline{A}\,\overline{B})=1-P(\overline{A})P(\overline{B})=1-0.3\times 0.4=0.88$.

(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，其中 $i=0,1,2$，则所求概率 $p=P(M_1N_0)+P(M_2N_1)=P(M_1)P(N_0)+P(M_2)P(N_1)=C_2^1\times0.7\times0.3\times0.4^2+0.7^2\times C_2^1\times0.6\times0.4=0.302\ 4$.

2. 甲、乙两人进行乒乓球比赛，已知在每局比赛中，甲获胜的概率为 $\frac{2}{3}$，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，则甲获胜的概率 $P_1=\left(\frac{2}{3}\right)^2+C_2^1\times\frac{2}{3}\times\frac{1}{3}\times\frac{2}{3}=\frac{20}{27}$.

(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，则甲获胜的概率 $P_2=\left(\frac{2}{3}\right)^3+C_3^2\times\left(\frac{2}{3}\right)^2\times\frac{1}{3}\times\frac{2}{3}+C_4^2\times\left(\frac{2}{3}\right)^2\times\left(\frac{1}{3}\right)^2\times\frac{2}{3}=\frac{64}{81}$.

学习任务二　二项分布及其应用

例 3　假设某种人寿保险规定：若投保人没活过 65 岁，则保险公司要赔偿 10 万元；若投保人活过 65 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能活过 65 岁的概率都为 0.9，随机抽取其中的 4 个投保人，设其中活过 65 岁的人数为 X.

(1)求 X 的分布列；(参考数据：$0.9^4=0.656\ 1$)

(2)假设保险公司支付给这 4 人的总金额为 Y 万元，写出 Y 与 X 的关系式，并求 $P(Y\geqslant22)$的值.

解：(1)由于每个投保人能活过 65 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 $P(X=k)=C_4^k0.9^k\times(1-0.9)^{4-k}(k=0,1,2,3,4)$，

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，活过 65 岁的人数为 X，则没活过 65 岁的人数为 $4-X$，

因此 $Y=10(4-X)+4X$，

即 $Y=40-6X(X=0,1,2,3,4)$.

由 $Y\geqslant22$，即 $40-6X\geqslant22$，得 $X\leqslant3$.

所以 $P(Y\geqslant22)=P(X\leqslant3)=P(X=0)+P(X=1)+P(X=2)+P(X=3)=1-P(X=4)=1-0.656\ 1=0.343\ 9$.

反思提炼

1. 当 X 服从二项分布时，先确定 $B(n,p)$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 $P(X=k)=C_n^kp^k(1-p)^{n-k}(k=0,1,2,\cdots,n)$，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.

探究训练

1. 口袋里放有大小相同的 2 个红球和 1 个白球，有放回地每次随机摸取一个球，定义数列$\{a_n\}$：$a_n=\begin{cases}-1,\text{第 } n \text{ 次摸到红球},\\1,\text{第 } n \text{ 次摸到白球}.\end{cases}$如果 S_n 为数列$\{a_n\}$的前 n 项和，那么 $S_7=3$ 的概率为　　　　(　　)

A. $C_7^5\times\left(\frac{1}{3}\right)^2\times\left(\frac{2}{3}\right)^5$

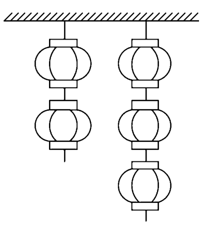
B.$C_7^2\times\left(\frac{2}{3}\right)^2\times\left(\frac{1}{3}\right)^5$

C.$C_7^5\times\left(\frac{1}{3}\right)^2\times\left(\frac{1}{3}\right)^5$

D.$C_7^2\times\left(\frac{1}{3}\right)^2\times\left(\frac{2}{3}\right)^2$

B　**解析**：由 $S_7=3$ 知，在 7 次摸球中有 2 次摸到红球，5 次摸到白球，而每次摸到红球的概率为 $\frac{2}{3}$，摸到白球的概率为 $\frac{1}{3}$，则 $S_7=3$ 的概率为 $C_7^2\times\left(\frac{2}{3}\right)^2\times\left(\frac{1}{3}\right)^5$.故选 B.

2.现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 ξ,η 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 网站购物的概率均为 $\frac{1}{3}$，去 B 网站购物的概率均为 $\frac{2}{3}$.

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，

$P(X=0)=P(A_0)+P(A_4)=C_4^0\times\left(\frac{1}{3}\right)^0\times\left(\frac{2}{3}\right)^4+C_4^4\times\left(\frac{1}{3}\right)^4\times\left(\frac{2}{3}\right)^0=\frac{16}{81}+\frac{1}{81}=\frac{17}{81}$，

$P(X=3)=P(A_1)+P(A_3)$

$=C_4^1\times\left(\frac{1}{3}\right)^1\times\left(\frac{2}{3}\right)^3+C_4^3\times\left(\frac{1}{3}\right)^3\times\left(\frac{2}{3}\right)^1$

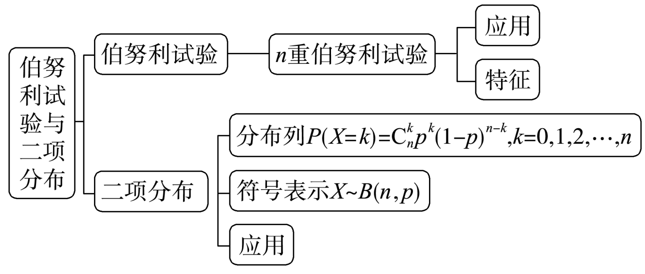
$=\frac{32}{81}+\frac{8}{81}=\frac{40}{81}$，

$P(X=4)=P(A_2)=C_4^2\times\left(\frac{1}{3}\right)^2\times\left(\frac{2}{3}\right)^2=\frac{8}{27}$.

所以随机变量 X 的分布列为

X	0	3	4
P	$\frac{17}{81}$	$\frac{40}{81}$	$\frac{8}{27}$

体系构建



课后素养评价(十四)

基础性·能力运用

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射击 4 次(每次命中目标的概率相同),已知他至少命中一次的概率为$\frac{80}{81}$,则此射击运动员每次射击命中目标的概率为　(　　)

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 ξ,每次射击命中的概率为 p,则 $\xi\sim B(4,p)$.

依题意可知,$P(\xi\geqslant 1)=\frac{80}{81}$,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**:将每位申请人申请房源看作一次试验,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 人申请甲片区的房源的概率 $p=C_4^2\times\left(\frac{1}{3}\right)^2\times\left(\frac{2}{3}\right)^2=\frac{8}{27}$.

5.假设每一架飞机的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.设甲、乙两位同学上学期间,每天 7:30 之前到校的概率均为$\frac{2}{3}$.假定甲、乙两位同学到校情况互不影响,且任一同学每天到校情况相互独立.

(1)用 X 表示甲同学上学期间的三天中 7:30 之前到校的天数,求随机变量 X 的分布列;

(2)设事件 M 为“上学期间的三天中,甲同学在 7:30 之前到校的天数比乙同学在 7:30 之前到校的天数恰好多 2”,求事件 M 发生的概率.

解:(1)因为甲同学上学期间的三天中到校情况相互独立,且每天 7:30 之前到校的概率均为$\frac{2}{3}$,

所以 $X\sim B\left(3,\frac{2}{3}\right)$,

从而 $P(X=k)=C_3^k\left(\frac{2}{3}\right)^k\left(\frac{1}{3}\right)^{3-k}(k=0,1,2,3)$.

所以随机变量 X 的分布列为

X	0	1	2	3
P	$\frac{1}{27}$	$\frac{2}{9}$	$\frac{4}{9}$	$\frac{8}{27}$

(2)设乙同学上学期间的三天中 7:30 之前到校的天数为 Y,则 $Y\sim B\left(3,\frac{2}{3}\right)$,且 $M=\{X=3,Y=1\}\cup\{X=2,Y=0\}$.

由题意知事件$\{X=3,Y=1\}$与$\{X=2,Y=0\}$互斥,且事件$\{X=3\}$与$\{Y=1\}$,事件$\{X=2\}$与$\{Y=0\}$均相互独立,从而由(1)知

$P(M)=P(\{X=3,Y=1\}\cup\{X=2,Y=0\})$

$=P(\{X=3,Y=1\})+P(\{X=2,Y=0\})$

$=P(X=3)P(Y=1)+P(X=2)P(Y=0)$

$=\frac{8}{27}\times\frac{2}{9}+\frac{4}{9}\times\frac{1}{27}=\frac{20}{243}$.

综合性·创新提升

1.假设某射击运动员每次射击命中率相同，且每次射击之间相互没有影响.若在两次射击中至多命中一次的概率是$\frac{16}{25}$，则该射击运动员每次射击的命中率为　（　　）

A.$\frac{9}{25}$　B.$\frac{2}{5}$　C.$\frac{3}{5}$　D.$\frac{3}{4}$

C　**解析**：设该射击运动员每次射击的命中率为 p，两次射击中命中的次数为 X，则 $X \sim B(2,p)$.由题可知 $P(X=0)+P(X=1)=\frac{16}{25}$，即 $C_2^0 p^0(1-p)^2+C_2^1 p(1-p)=\frac{16}{25}$，解得 $p=\frac{3}{5}$.故选 C.

2.某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 ξ 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.（新定义）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，途径 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.如果在路口遇到红灯要停留 1 min.假设他在各路口是否遇到红灯是相互独立的，并且遇到红灯的概率都是$\frac{1}{3}$，则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.设随机变量 $X\sim B(n,p)$，记 $p_k=C_n^k p^k\cdot(1-p)^{n-k}$，$k=0,1,2,\cdots,n$.在研究 p_k 的最大值时，发现：若$(n+1)p$为正整数，则当 $k=(n+1)p$ 时，$p_k=p_{k-1}$，此时 p_k 和 p_{k-1}均为最大值；若$(n+1)p$ 为非整数，则当 k 取$(n+1)p$ 的整数部分时，p_k 是唯一的最大值.以此为理论基础，某同学重复投掷一枚质

地均匀的骰子并实时记录点数 1 出现的次数，当投掷到第 35 次时，记录到点数 1 出现 5 次，若再继续进行 65 次投掷试验，即当投掷到第 100 次时，点数 1 一共出现的次数为______的概率最大.

15 或 16　**解析**：继续再进行 65 次投掷试验，设出现点数 1 的次数为 X，则 $X\sim B\left(65,\frac{1}{6}\right)$.

由 $k=(n+1)p=66\times\frac{1}{6}=11$，结合题中的结论可知，当 $k=11$ 或 $k=10$ 时概率最大.

即后面 65 次中出现 10 次或 11 次点数 1 的概率最大，加上前面 35 次中的 5 次，

所以出现 15 次或 16 次的概率最大.

7. 羽毛球比赛的规则如下：采用 21 分制，3 局 2 胜.每回合中，取胜的一方得 1 分.每局中，先得 21 分的一方获胜；若双方打成 20 平，则领先 2 分的一方该局获胜；若双方打成 29 平，则先得 30 分的一方该局获胜.某次羽毛球比赛中，甲、乙两位选手进行比赛，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 回合甲获胜”为事件 A，

可知甲在第 4 回合取胜，前 3 回合胜 2 个回合，

所以 $P(A)=\frac{3}{4}\times C_3^2\times\left(\frac{3}{4}\right)^2\times\left(\frac{1}{4}\right)=\frac{81}{256}$.

(2) 易知 X 的可能取值为 0，1，2，3，4，且 $X\sim B\left(4,\frac{3}{4}\right)$，

$P(X=0)=C_4^0\times\left(\frac{1}{4}\right)^4=\frac{1}{256}$，

$P(X=1)=C_4^1\times\frac{3}{4}\times\left(\frac{1}{4}\right)^3=\frac{3}{64}$，

$P(X=2)=C_4^2\times\left(\frac{3}{4}\right)^2\times\left(\frac{1}{4}\right)^2=\frac{27}{128}$，

$P(X=3)=C_4^3\times\left(\frac{3}{4}\right)^3\times\left(\frac{1}{4}\right)=\frac{27}{64}$，

$P(X=4)=C_4^4\times\left(\frac{3}{4}\right)^4=\frac{81}{256}$.

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应用问题.

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 $E(X)=\underline{np}$，$D(X)=\underline{np(1-p)}$.特别地，当 $n=1$ 时，X 服从两点分布，此时 $E(X)=\underline{p}$，$D(X)=\underline{p(1-p)}$.

概念辨析

1.判断(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列也一定相同.　(×)

(3)若 X 服从两点分布，则 $E(X)=np$.　(×)

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 $n=1$ 的二项分布；二项分布中的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

学习任务一　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.已知小华投篮10次，每次投篮的命中率为0.7.记10次投篮命中的次数为 X，则 $D(X)=$________.

2.1　**解析**：由题意得 $X\sim B(10,0.7)$，所以 $D(X)=10\times0.7\times0.3=2.1$.

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)求重复投篮5次时，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,重复投篮5次,命中的次数Y服从二项分布,即$Y\sim B(5,0.6)$.

则 $E(Y)=5\times0.6=3$,$D(Y)=5\times0.6\times0.4=1.2$.

反思提炼

若随机变量X服从二项分布,即$X\sim B(n,p)$,则$E(X)=np$,$D(X)=np(1-p)$.

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.

学习任务二　二项分布的均值与方差的实际应用

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经过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ξ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η(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η的分布列.

解:设前k个是绿灯,第$k+1$个是红灯,

则η的分布列为 $P(\eta=k)=\left(\frac{2}{3}\right)^k\times\frac{1}{3}$,$k=0,1,2,3,4,5$,

所以 $P(\eta=0)=\left(\frac{2}{3}\right)^0\times\frac{1}{3}=\frac{1}{3}$,

$P(\eta=1)=\frac{2}{3}\times\frac{1}{3}=\frac{2}{9}$,

$P(\eta=2)=\left(\frac{2}{3}\right)^2\times\frac{1}{3}=\frac{4}{27}$,

$P(\eta=3)=\left(\frac{2}{3}\right)^3\times\frac{1}{3}=\frac{8}{81}$,

$P(\eta=4)=\left(\frac{2}{3}\right)^4\times\frac{1}{3}=\frac{16}{243}$,

$P(\eta=5)=\left(\frac{2}{3}\right)^5\times\frac{1}{3}=\frac{32}{729}$,

若全为绿灯,则 $P(\eta=6)=\left(\frac{2}{3}\right)^6=\frac{64}{729}$.

故η的分布列为

η	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 $P(\xi\geqslant1)=1-P(\xi=0)=1-\left(\frac{2}{3}\right)^6=\frac{665}{729}$.

反思提炼

1. 用二项分布求解实际应用题的步骤

(1)判断随机变量 X 服从二项分布，即 $X\sim B(n,p)$.

(2)根据二项分布概率计算公式 $P(X=k)=C_n^k p^k(1-p)^{n-k}(k=0,1,\cdots,n)$ 求出 X 的分布列.

(3)利用公式 $E(X)=np$，$D(X)=np(1-p)$ 求 X 的均值与方差.

(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.

探究训练

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	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$

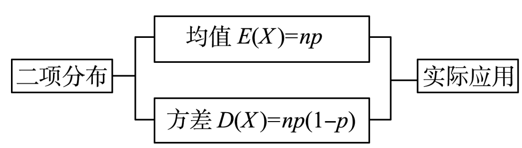
(2)因为 $X\sim B\left(4,\frac{1}{2}\right)$，所以 $E(X)=4\times\frac{1}{2}=2$.

又由题意可知 $Y=2\ 300-100X$，

所以 $E(Y)=E(2\ 300-100X)=2\ 300-100E(X)=2\ 300-100\times 2=2\ 100$.

即所求随机变量 Y 的均值为 2 100.

体系构建



课后素养评价（十五）

基础性·能力运用

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 $E(X)=10\times 0.5=5$.因为 $Y=2X-8$，可得 $E(Y)=2E(X)-8=2\times 5-8=2$.故选 B.

4.某人从家乘车到公司上班，途中经过 3 个路口.假设在各路口遇到红灯的事件是相互独立的，且概率都是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 4 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 $\left(\frac{1}{3}\right)^2\times\left(1-\frac{3}{4}\right)+C_2^1\times\left(\frac{1}{3}\right)\times\left(1-\frac{1}{3}\right)\times\frac{3}{4}=\frac{13}{36}$.

(3)$p_1<p_0$.

7.为防止风沙危害，某地决定建设防护绿化带，种植杨树、沙柳等植物.某人一次种植了 n 株沙柳，各株沙柳成活与否是相互独立的，成活率均为 p.设 ξ 为成活沙柳的株数，数学期望 $E(\xi)=3$，标准差 $\sqrt{D(\xi)}=\frac{\sqrt{6}}{2}$.

(1)求 n,p 的值，并写出 ξ 的分布列；

(2)若有 3 株或 3 株以上的沙柳未成活，则需要补种，求需要补种沙柳的概率.

解：因为每一株沙柳的成活率均为 p，种植了 n 株沙柳，相当于 n 重伯努利试验，因此 ξ 服从二项分布 $B(n,p)$.

(1)由 $E(\xi)=np=3$，$D(\xi)=np(1-p)=\frac{3}{2}$，

得 $1-p=\frac{1}{2}$，从而 $p=\frac{1}{2}$，$n=6$.

所以 ξ 的分布列为

ξ	0	1	2	3	4	5	6
P	$\frac{1}{64}$	$\frac{3}{32}$	$\frac{15}{64}$	$\frac{5}{16}$	$\frac{15}{64}$	$\frac{3}{32}$	$\frac{1}{64}$

(2)记“需要补种沙柳”为事件 A，则 $P(A)=P(\xi\leqslant3)$，

得 $P(A)=\frac{1+6+15+20}{64}=\frac{21}{32}$.

综合性·创新提升

1. 若随机变量 $\xi\sim B(n,0.6)$，且 $E(\xi)=3$，则 $P(\xi=1)$ 的值为（　　）

A. 2×0.4^4　　B. 2×0.4^5

C. 3×0.4^4　　D. 3×0.6^4

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 ξ 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. 一台仪器每启动一次都随机地出现一个 3 位的二进制数 $A=\boxed{a_1}\boxed{a_2}\boxed{a_3}$，其中 A 的各位数字中，$a_k(k=1,2,3)$ 出现 0 的概率为 $\frac{1}{3}$，出现 1 的概率为 $\frac{2}{3}$. 若启动一次出现的数为 100，则称这次试验成功. 若成功一次得 2 分，失败一次得 -1 分，则进行 81 次这样的重复试验的总得分 X 的数学期望和方差分别为（　　）

A. $-63,\frac{50}{9}$　　B. $-63,50$

C. $6,\frac{50}{9}$　　D. $6,50$

B　**解析**：由题可得，启动一次出现的数为 100 的概率 $p=\frac{2}{3}\times\left(\frac{1}{3}\right)^2=\frac{2}{27}$，用 η 表示试验成功的次数，则 $\eta\sim B\left(81,\frac{2}{27}\right)$，

所以 η 的数学期望 $E(\eta)=81\times\frac{2}{27}=6$，

η 的方差 $D(\eta)=81\times\frac{2}{27}\times\frac{25}{27}=\frac{50}{9}$.

因为得分 $X=2\eta+(-1)\times(81-\eta)=3\eta-81$，

所以 $E(X)=E(3\eta-81)=3E(\eta)-81=-63$，

$D(X)=D(3\eta-81)=9D(\eta)=50$.

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 $C_3^2\times\left(\frac{2}{3}\right)^2\times\left(\frac{1}{3}\right)=\frac{4}{9}$；

由题知随机变量 $X\sim B\left(3,\frac{2}{3}\right)$，

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 个选择题，每个选择题有 4 个选项，其中有且只有 1 个正确选项，每一题答对得 5 分，答错或不答得 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. 有一个选手答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 $E(X)=30\times0.8=24$. 由于 $24\times5=120$(分)，所以最可能得到二等奖.

6. 设随机变量 ξ 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 $\Delta=16-4\xi\geqslant0$，即 $\xi\leqslant4$.

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.结合教材实例，了解超几何分布的概念.(数学抽象)
2.会利用公式求服从超几何分布的随机变量的概率、均值.(数学运算)
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实际问题.(数学建模)

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 $P(X=k)=\underline{\dfrac{C_M^kC_{N-M}^{n-k}}{C_N^n}}$，$k=m,m+1,m+2,\cdots,r$.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：$E(X)=\dfrac{nM}{N}=\underline{np}$，其中 $p=\dfrac{M}{N}$，是 N 件产品的次品率.

概念辨析

1.判断(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________.

$\dfrac{3}{7}$　**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 $P(X=2)=\dfrac{C_6^2C_4^2}{C_{10}^4}=\dfrac{3}{7}$.

3.请思考并回答下列问题：

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，而二项分布的模型则是"独立重复试验"，对于抽样，则是有放回抽样.

任务型课堂

学习任务一　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 表示任选 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 表示任选 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 表示任选 5 个球中白球的个数，则 $P(Z=2)=$________.

(1)A　(2)0，1，2，3　(3)$\dfrac{5}{18}$　**解析**：(1)根据超几何分布的定义知，$N=9$，$M=4$，$n=4$.

(2)由于只选取了 3 个球，因此随机变量 Y 的所有可能取值为 0，1，2，3.

(3)$P(Z=2)=\dfrac{C_2^2C_7^3}{C_9^5}=\dfrac{5}{18}$.

反思提炼

对超几何分布的三点说明

(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 M,N,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、学生中的男生和女生等问题，研究对象往往由差异明显的两部分组成.

学习任务二　超几何分布的分布列

例1　袋中有4个红球、3个黑球，从袋中随机取球，设取到一个红球得2分，取到一个黑球得1分，从袋中任取4个球，求得分 X 的分布列.

解：从袋中任取4个球，有1红3黑、2红2黑、3红1黑、4红四种情况，

得分分别为5，6，7，8，

故 X 的可能取值为5，6，7，8.

$P(X=5)=\frac{C_4^1C_3^3}{C_7^4}=\frac{4}{35}$，$P(X=6)=\frac{C_4^2C_3^2}{C_7^4}=\frac{18}{35}$，

$P(X=7)=\frac{C_4^3C_3^1}{C_7^4}=\frac{12}{35}$，$P(X=8)=\frac{C_4^4C_3^0}{C_7^4}=\frac{1}{35}$.

故随机变量 X 的分布列为

X	5	6	7	8
P	$\frac{4}{35}$	$\frac{18}{35}$	$\frac{12}{35}$	$\frac{1}{35}$

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“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 $E(X)=3\times\frac{1}{5}=\frac{3}{5}$，

$D(X)=3\times\frac{1}{5}\times\left(1-\frac{1}{5}\right)=\frac{12}{25}$.

反思提炼

1.超几何分布问题的求解步骤

(1)辨模型：结合实际情景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“男生、女生”“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.

探究训练

1.袋中共有9个球，其中有4个红球、3个黄球和2个绿球，这些球除颜色外完全相同.从袋中一次随机取出4个球，其中红球、黄球、绿球的个数分别为 x_1，x_2，x_3，随机变量 X 表示 x_1，x_2，x_3 中的最大数，求 X 的分布列.

解：X 的可能取值为2，3，4，

$P(X=4)=\frac{C_4^4}{C_9^4}=\frac{1}{126}$，

$P(X=3)=\frac{C_4^3C_5^1+C_3^3C_6^1}{C_9^4}=\frac{13}{63}$，

于是 $P(X=2)=1-P(X=3)-P(X=4)=\frac{11}{14}$.

故 X 的分布列为

X	2	3	4
P	$\frac{11}{14}$	$\frac{13}{63}$	$\frac{1}{126}$

2. 为了解学生的身体素质情况，某校从全校学生中随机抽取 10 人进行体能测试，测试的成绩(单位：分)从低到高依次为 52，65，72，78，86，86，86，87，87，88.规定成绩不低于 79 分的为优秀.从这 10 人中随机选取 3 人，记 X 表示测试成绩为优秀的学生人数，求 X 的分布列.

解： 由题意可得，X 的可能取值为 0，1，2，3，

$P(X=0)=\frac{C_4^3}{C_{10}^3}=\frac{1}{30}$，$P(X=1)=\frac{C_4^2C_6^1}{C_{10}^3}=\frac{3}{10}$，

$P(X=2)=\frac{C_4^1C_6^2}{C_{10}^3}=\frac{1}{2}$，$P(X=3)=\frac{C_6^3}{C_{10}^3}=\frac{1}{6}$.

所以 X 的分布列为

X	0	1	2	3
P	$\frac{1}{30}$	$\frac{3}{10}$	$\frac{1}{2}$	$\frac{1}{6}$

学习任务 三　超几何分布的应用

例 3　端午节吃粽子是我国的传统习俗.设一盘中装有 10 个粽子，其中豆沙粽 2 个、肉粽 3 个、白粽 5 个，这三种粽子的外观完全相同.从中任意选取 3 个.

(1)设 X 表示取到的豆沙粽的个数，求 X 的分布列；

(2)设 Y 表示取到的粽子的种类数，求 Y 的分布列.

解：(1)豆沙粽的个数 X 服从超几何分布，

且 X 的可能取值为 0，1，2，

因此 $P(X=0)=\frac{C_8^3}{C_{10}^3}=\frac{7}{15}$，

$P(X=1)=\frac{C_2^1C_8^2}{C_{10}^3}=\frac{7}{15}$，

$P(X=2)=\frac{C_2^2C_8^1}{C_{10}^3}=\frac{1}{15}$.

所以 X 的分布列为

X	0	1	2
P	$\frac{7}{15}$	$\frac{7}{15}$	$\frac{1}{15}$

(2)由题意知 Y 的所有可能取值为 1，2，3，且

$P(Y=1)=\frac{C_3^3+C_5^3}{C_{10}^3}=\frac{1+10}{120}=\frac{11}{120}$，

$P(Y=3)=\frac{C_2^1C_3^1C_5^1}{C_{10}^3}=\frac{30}{120}=\frac{1}{4}$，

$P(Y=2)=1-P(Y=1)-P(Y=3)=1-\frac{11}{120}-\frac{30}{120}$

$=\frac{79}{120}$.

所以 Y 的分布列为

Y	1	2	3
P	$\frac{11}{120}$	$\frac{79}{120}$	$\frac{1}{4}$

反思提炼

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，通过公式可以求出随机变量的概率、分布列、期望等，利用这些可以解决实际生活中与之相关的问题.

探究训练

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，测试效果既有优秀的，也有良好的”，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 $P(A)=P(X=4)+P(X=5)=\frac{12}{35}+\frac{18}{35}$

$=\frac{6}{7}$.

故事件 A 发生的概率为$\frac{6}{7}$.

2. 三门是“中国青蟹之乡”，其气候温暖、港湾平静、水质优良，以优越的自然环境成为我国优质青蟹的最佳产区之一.三门青蟹具有“金爪、绯钳、青背、黄肚”的特征，以“壳薄、膏黄、肉嫩、味美”而著称，素有“三门青蟹、横行世界”之美誉；且营养丰富，被誉为“海中黄金，蟹中臻品”.养殖户一般把质量超过 350 g 的青蟹标记为 A 类青蟹.

(1)现有一个小型养蟹池，已知蟹池中有 50 只青蟹，其中 A 类青蟹有 7 只，若从池中随机抓 2 只青蟹，用 ξ 表示其中 A 类青蟹的只数，请写出 ξ 的分布列，并求 ξ 的数学期望 $E(\xi)$；

(2)另有一个养蟹池，为估计蟹池中青蟹的数量 N，小王先从蟹池中抓了 50 只青蟹，做好标记后放回池中，过了一段时间后，又从蟹池中抓了 20 只青蟹，发现有记号的有 x 只，若 $x=5$，试给出蟹池中青蟹数量 N 的估计值.(以使 $P(x=5)$ 取得最大值的 N 为估计值)

解：(1)由题意知 ξ 的所有可能取值为 0,1,2,

$P(\xi=0)=\frac{C_{43}^{2}}{C_{50}^{2}}=\frac{129}{175}$，$P(\xi=1)=\frac{C_{7}^{1}C_{43}^{1}}{C_{50}^{2}}=\frac{43}{175}$，

$P(\xi=2)=\frac{C_{7}^{2}}{C_{50}^{2}}=\frac{3}{175}$，

所以 ξ 的分布列为

ξ	0	1	2
P	$\frac{129}{175}$	$\frac{43}{175}$	$\frac{3}{175}$

$E(\xi)=0\times\frac{129}{175}+1\times\frac{43}{175}+2\times\frac{3}{175}=\frac{7}{25}$.

(2)设 $f(N)=P(x=5)=\frac{C_{50}^{5}C_{N-50}^{15}}{C_{N}^{20}}$，

则 $\frac{f(N+1)}{f(N)}=\frac{(N-49)(N-19)}{(N-64)(N+1)}=$

$\frac{N^2-68N+931}{N^2-63N-64}$.

因为 $(N^2-68N+931)-(N^2-63N-64)=-5N+995$，

所以当 $N=199$ 时，$f(N+1)=f(N)$，即 $f(199)=f(200)$；

当 $N>199$ 时，$f(N+1)<f(N)$；

当 $N<199$ 时，$f(N+1)>f(N)$.

所以当 $N=199$ 或 200 时，$P(x=5)$ 最大，估计蟹池中青蟹数量为 199 只或 200 只.

体系构建

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)

基础性·能力运用

1. 在 15 个村庄中，有 7 个村庄交通不方便.若用随机变量 X 表示从这 15 个村庄中任选的 10 个村庄中交通不方便的村庄的个数，则 X 服从超几何分布，其中的参数的值为　(　　)

A. $N=15$，$M=7$，$n=10$

B. $N=15$，$M=10$，$n=7$

C. $N=22$，$M=10$，$n=7$

D. $N=22$，$M=7$，$n=10$

A　**解析**：根据超几何分布概率模型知选 A.

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. $\frac{C_{80}^{4}C_{10}^{6}}{C_{100}^{10}}$　　B. $\frac{C_{80}^{6}C_{10}^{4}}{C_{100}^{10}}$

C. $\frac{C_{80}^{4}C_{20}^{6}}{C_{100}^{10}}$　　D. $\frac{C_{80}^{6}C_{20}^{4}}{C_{100}^{10}}$

D　**解析**：设随机变量 X 表示任取 10 个球中红球的个数，则 X 服从参数 $N=100$，$M=80$，$n=10$ 的超几何分布.取到的 10 个球中恰有 6 个红球，即 $X=6$，$P(X=6)=\frac{C_{80}^{6}C_{20}^{4}}{C_{100}^{10}}$.

3. 12 人的兴趣小组中有 5 名男生，现从中任选 6 人参加竞赛.若随机变量 X 表示参加竞赛的男生的人

数,则$\frac{C_5^3C_7^3}{C_{12}^6}$等于　　(　　)

A.$P(X=6)$　　B.$P(X=5)$

C.$P(X=3)$　　D.$P(X=7)$

C　**解析**:由题意可知随机变量 X 服从参数 $N=12$,$M=5$,$n=6$ 的超几何分布.由公式 $P(X=k)=\frac{C_M^kC_{N-M}^{n-k}}{C_N^n}$,易知$\frac{C_5^3C_7^3}{C_{12}^6}$表示的是 $X=3$ 的概率.

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,则概率是$\frac{3}{10}$的事件为　　(　　)

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,

且$P(X=0)=\frac{C_6^4}{C_{10}^4}=\frac{1}{14}$,

$P(X=1)=\frac{C_4^1C_6^3}{C_{10}^4}=\frac{8}{21}$,

$P(X=2)=\frac{C_4^2C_6^2}{C_{10}^4}=\frac{3}{7}$,

$P(X=3)=\frac{C_4^3C_6^1}{C_{10}^4}=\frac{4}{35}$,

$P(X=4)=\frac{C_4^4}{C_{10}^4}=\frac{1}{210}$,

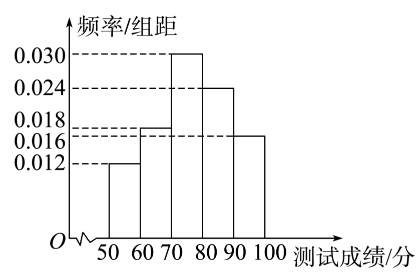
故 $E(X)=0\times\frac{1}{14}+1\times\frac{8}{21}+2\times\frac{3}{7}+3\times\frac{4}{35}+4\times\frac{1}{210}=\frac{8}{5}\left(或 E(X)=4\times\frac{4}{10}=\frac{8}{5}\right)$,故 CD 正确.故选 BCD.

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**:依题意,X 的所有可能取值是 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 ξ,则 ξ 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 ξ 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、颜色都相同的球10个，其中8个标有数字1，2个标有数字5.从中任取2个球，记2个球上的数字之和为 X，求 X 的分布列.

解：由题意知，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	$\frac{28}{45}$	$\frac{16}{45}$	$\frac{1}{45}$

综合性·创新提升

1. 一个小组有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，其中参数为 $N=6,M=1,n=2$，则 $P(X=1)=\frac{C_1^1C_5^1}{C_6^2}=\frac{1}{3}$.

2.（新定义）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. $P(|Z-6|\leqslant 1)=\frac{97}{105}$

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 $P(X=k)=\frac{C_4^kC_6^{4-k}}{C_{10}^4}(k=0,1,2,3,4)$.$P(|Z-6|\leqslant 1)=1-P(Z=4)-P(Z=8)=1-P(X=0)-P(X=4)=\frac{97}{105}$，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瓶已过了保质期.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. 某公司有日生产件数为95的“生产能手”3人，有日生产件数为55的“新手”2人，从这5人中任意抽取2人，则2人的日生产件数之和 X 的标准差为________.

24　**解析**：由题意，可得 X 的所有可能取值为190，150，110，且 $P(X=190)=\frac{C_2^0C_3^2}{C_5^2}=\frac{3}{10}$，$P(X=150)=\frac{C_2^1C_3^1}{C_5^2}=\frac{3}{5}$，$P(X=110)=\frac{C_2^2C_3^0}{C_5^2}=\frac{1}{10}$，则 $E(X)=190\times\frac{3}{10}+150\times\frac{3}{5}+110\times\frac{1}{10}=158$，标准差为 $\sqrt{(190-158)^2\times\frac{3}{10}+(150-158)^2\times\frac{3}{5}+(110-158)^2\times\frac{1}{10}}=24$.

7.5 正态分布

学习任务目标

1.了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.(直观想象)
3.会求正态分布在给定区间的概率,能利用正态分布知识解决实际问题.(数学运算)

问题式预习

知识清单

知识点一 正态分布

函数 $f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$,$x\in\mathbf{R}$,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

(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$\frac{1}{\sigma\sqrt{2\pi}}$.
(3)当$|x|$无限增大时,曲线无限接近 x 轴.
(4)当参数 σ 取固定值时,正态曲线的位置由 μ 确定,且随着 μ 的变化而沿 x 轴平移,如图 1.
(5)当 μ 取定值时,曲线的形状由 σ 确定,当 σ 较小时,正态曲线"瘦高",表示随机变量 X 的分布比较集中;当 σ 较大时,正态曲线"矮胖",表示随机变量 X 的分布比较分散,如图 2.

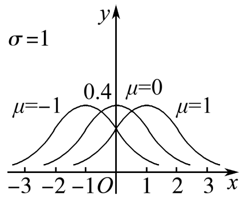


图1

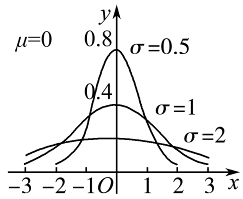


图2

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

(1)$P(\mu-\sigma\leqslant X\leqslant\mu+\sigma)\approx 0.682\ 7$;
(2)$P(\mu-2\sigma\leqslant X\leqslant\mu+2\sigma)\approx 0.954\ 5$;
(3)$P(\mu-3\sigma\leqslant X\leqslant\mu+3\sigma)\approx 0.997\ 3$.
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 3σ 原则.

概念辨析

1.判断(正确的打"√",错误的打"×").
(1)正态密度函数中参数 μ,σ 的意义分别是样本的均值与方差. (×)
(2)若 $X\sim N(\mu,\sigma^2)$,则 $P(X<\mu)=\frac{1}{2}$. (√)
(3)正态分布中,参数 σ 越小,表示随机变量 X 的分布越集中;σ 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 ξ 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 $P(\xi\geqslant 2)=P(\xi\leqslant 0)=P(\xi<1)-P(0<\xi<1)=0.5-0.3=0.2$.
所以 $P(\xi<2)=1-P(\xi\geqslant 2)=1-0.2=0.8$.故选 A.

3.已知随机变量 X 服从正态分布 $N(2,1)$,则 $P(X<1)\approx$________.

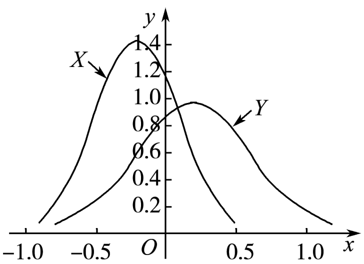
0.158 65 **解析**:因为 $\mu=2$,$\sigma=1$,所以 $P(1\leqslant X\leqslant 3)\approx 0.682\ 7$.
所以 $P(X<1)=P(X>3)\approx\frac{1-0.682\ 7}{2}=0.158\ 65$.

4.请思考并回答下列问题:
一个正态分布由参数 μ 和 σ 完全确定,这两个参数可以由什么近似估计?
提示:参数 μ 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;σ 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

学习任务一　正态曲线及其性质

1. 设 $X\sim N(\mu_1,\sigma_1^2)(\sigma_1>0)$，$Y\sim N(\mu_2,\sigma_2^2)(\sigma_2>0)$，且正态密度曲线如图所示，则有　　(　　)

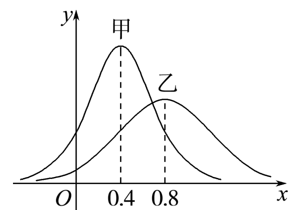


A. $\mu_1<\mu_2,\sigma_1<\sigma_2$　　B. $\mu_1<\mu_2,\sigma_1>\sigma_2$

C. $\mu_1>\mu_2,\sigma_1<\sigma_2$　　D. $\mu_1>\mu_2,\sigma_1>\sigma_2$

A　**解析**：μ 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，由图可知 $\mu_1<\mu_2$；σ 反映的是随机变量的离散程度，σ 越大，越分散，曲线越"矮胖"；σ 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"，故 $\sigma_1<\sigma_2$，
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.

学习任务二　正态分布的概率计算

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 $P(10<X<18)=\dfrac{1-2a}{2}=\dfrac{1}{2}-a$.

反思提炼

正态分布的概率计算的方法

(1)充分利用正态曲线的对称性和曲线与 x 轴之间的区域面积为 1.

(2)熟记 $P(\mu-\sigma\leqslant X\leqslant\mu+\sigma)$，$P(\mu-2\sigma\leqslant X\leqslant\mu+2\sigma)$，$P(\mu-3\sigma\leqslant X\leqslant\mu+3\sigma)$ 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 $P(X<b)=\dfrac{1}{2}-P(b\leqslant X\leqslant\mu)$.

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.

探究训练

1. 已知随机变量 ξ 服从正态分布 $N(0,\sigma^2)$.若 $P(\xi\geqslant 2)=0.023$，则 $P(-2<\xi<2)=$　　(　　)

A. 0.477　　B. 0.625

C. 0.954　　D. 0.977

C　**解析**：$P(-2<\xi<2)=1-2P(\xi\geqslant 2)=1-2\times 0.023=0.954$.

2. 设随机变量 ξ 服从正态分布 $N(2,\sigma^2)$.若 $P(\xi>c)=a$，则 $P(\xi>4-c)=$　　(　　)

A. a　　B. $1-a$

C. $2a$　　D. $1-2a$

B　**解析**：正态曲线的对称轴为直线 $x=2$，

所以 $P(\xi>4-c)=1-P(\xi>c)=1-a$.

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 $P(X<2)=P(X>2)=0.5$，因此 $P(X>2.5)=P(X>2)-P(2<X\leqslant 2.5)=0.5-0.36=0.14$.

学习任务三　正态分布的实际应用

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 $X\sim N(8\ 000,500^2)$.

$P(8\ 000\leqslant X\leqslant 8\ 500)$

$=\dfrac{1}{2}P(8\ 000-500\leqslant X\leqslant 8\ 000+500)$

$\approx 0.341\ 35$.

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：$P(X<7\ 000)=\dfrac{1}{2}(1-P(8\ 000-2\times 500\leqslant X\leqslant 8\ 000+2\times 500))\approx\dfrac{1}{2}\times(1-0.954\ 5)=0.022\ 75$.

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%.

例 3　在某次数学考试中，考生的成绩 ξ 服从正态分布 $N(90,100)$.

(1)试求考试成绩 ξ 位于区间[70,110]内的概率；

(2)若共有 2 000 名考生参加考试，试估计考试成绩在[80,100]内的考生人数.

解：因为 $\xi\sim N(90,100)$，所以 $\mu=90$，$\sigma=10$.

(1)由于正态变量在区间 $[\mu-2\sigma,\mu+2\sigma]$ 内取值的概率是 0.954 5，而该正态分布中，$\mu-2\sigma=90-2\times 10=70$，$\mu+2\sigma=90+2\times 10=110$，于是考试成绩 ξ 位于区间[70,110]内的概率是 0.954 5.

(2)由 $\mu=90$，$\sigma=10$，得 $\mu-\sigma=80$，$\mu+\sigma=100$.由于正态变量在区间 $[\mu-\sigma,\mu+\sigma]$ 内取值的概率是 0.682 7，所以考试成绩 ξ 位于区间[80,100]内的概率是 0.682 7.一共有2 000 名考生，所以考试成绩在[80,100]内的考生大约有$2\ 000\times 0.682\ 7\approx 1\ 365$(人).

反思提炼

正态分布问题的解题策略

解答此类问题的关键在于将待求概率的随机变量的范围向 $[\mu-\sigma,\mu+\sigma]$，$[\mu-2\sigma,\mu+2\sigma]$，$[\mu-3\sigma,\mu+3\sigma]$ 这三个区间进行转化，然后利用随机变量在上述区间内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.

探究训练

1.某厂生产的零件外直径为 X(单位：mm)，且 $X\sim N(8.0,0.022\ 5)$.现从该厂上午与下午生产的零件中各随机取出一个，测得其外直径分别为7.9 mm和7.5 mm，则可认为　　(　　)

A.上午与下午生产情况均为正常

B.上午与下午生产情况均为异常

C.上午生产情况正常，下午生产情况异常

D.上午生产情况异常，下午生产情况正常

C　**解析**：由题可知 $\mu=8.0$，$\sigma=0.15$.根据 3σ 原则，可知 X 在 $[8-3\times 0.15,8+3\times 0.15]$ 即 $[7.55,8.45]$ 之外时为异常.结合已知，可知上午生产情况正常，下午生产情况异常.

2.某面包店声称该店所售出的面包的平均质量是 1 000 g，上下浮动不超过 50 g.用数学语言来表达就是：设该店售出的面包的质量为 X(单位：g)，则 X 服从正态分布$N(1\ 000,50^2)$.假设面包店的说法是真实的，则从该面包店随机购买一个面包，质量不小于 900 g 的概率为　　(　　)

附：若随机变量 X 服从正态分布 $N(\mu,\sigma^2)$，则 $P(\mu-\sigma\leqslant X\leqslant\mu+\sigma)\approx 0.682\ 7$，$P(\mu-2\sigma\leqslant X\leqslant\mu+2\sigma)\approx 0.954\ 5$，$P(\mu-3\sigma\leqslant X\leqslant\mu+3\sigma)\approx 0.997\ 3$.

A.0.841 35　　B.0.972 25

C.0.977 25　　D.0.998 65

C　**解析**：由题意得 $\mu=1\ 000$，$\sigma=50$，则 $900=1\ 000-50\times 2=\mu-2\sigma$，

故面包的质量不小于 900 g 的概率为 $1-\dfrac{1-P(\mu-2\sigma\leqslant X\leqslant\mu+2\sigma)}{2}\approx 1-0.022\ 75=0.977\ 25$.故选 C.

3.随着人们生活水平的不断提高，旅游成了一种生活

时尚，尤其是针对老年人的旅游市场在不断扩大．为了了解老年人每年旅游费用支出 X（单位：元）的情况，相关部门抽取了某地区 1 000 名老年人进行问卷调查，并把所得数据列成如下所示的频数分布表．

组别	频数
[0,1 000)	120
[1 000,2 000)	260
[2 000,3 000)	340
[3 000,4 000)	250
[4 000,5 000)	20
[5 000,6 000]	10

(1)求所得样本的平均数（精确到元）．

(2)根据样本数据，可认为老年人的旅游费用支出 X 服从正态分布 $N(3\ 000,1\ 000^2)$．若该地区共有老年人 95 000 人，试估计有多少位老年人旅游费用支出在 5 000 元以上．

(3)已知旅游费用支出在[5 000,6 000]范围内的 10 名老人中有 7 名女性，3 名男性．现想选其中 3 名老人回访，记选出的男性人数为 ξ，求 ξ 的分布列．

解：(1)设样本平均数为 $\overline{x}$，则有

$$\overline{x}=\frac{1}{1\ 000}\times(500\times120+1\ 500\times260+2\ 500\times340+3\ 500\times250+4\ 500\times20+5\ 500\times10)=2\ 320(\text{元}).$$

(2)因为 $\mu=3\ 000$，$\sigma=1\ 000$，所以 $\mu+2\sigma=5\ 000$．

所以旅游费用支出在 5 000 元以上的概率为

$$P(X>\mu+2\sigma)=\frac{1-P(\mu-2\sigma\leqslant X\leqslant\mu+2\sigma)}{2}\approx0.022\ 75.$$

因为 $95\ 000\times0.022\ 75\approx2\ 161$，

所以估计有 2 161 位老年人旅游费用支出在 5 000 元以上．

(3)由题意可知，ξ 的可能取值为 0,1,2,3，且

$$P(\xi=0)=\frac{C_7^3}{C_{10}^3}=\frac{7}{24},$$

$$P(\xi=1)=\frac{C_7^2C_3^1}{C_{10}^3}=\frac{21}{40},$$

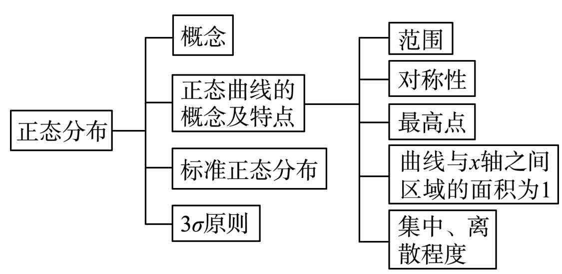
$$P(\xi=2)=\frac{C_7^1C_3^2}{C_{10}^3}=\frac{7}{40},$$

$$P(\xi=3)=\frac{C_3^3}{C_{10}^3}=\frac{1}{120}.$$

所以随机变量 ξ 的分布列为

ξ	0	1	2	3
P	$\frac{7}{24}$	$\frac{21}{40}$	$\frac{7}{40}$	$\frac{1}{120}$

体系构建



课后素养评价（十七）

基础性·能力运用

1.已知随机变量 X 服从正态分布，且 X 的概率分布密度函数为 $f(x)=\frac{1}{\sqrt{8\pi}}e^{-\frac{(x-10)^2}{8}}(x\in\mathbf{R})$，则随机变量 X 的平均数 μ 与标准差 σ 分别是（　　）

A.10 与 8

B.10 与 2

C.8 与 10

D.2 与 10

B　**解析**：因为 $f(x)=\frac{1}{\sqrt{8\pi}}e^{-\frac{(x-10)^2}{8}}=\frac{1}{2\sqrt{2\pi}}e^{-\frac{(x-10)^2}{2\times2^2}}$，所以平均数 $\mu=10$，标准差 $\sigma=2$．故选 B．

2.已知正态密度函数 $\varphi(x)=\frac{1}{\sigma\sqrt{2\pi}}\cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$，$x\in(-\infty,+\infty)$，下列关于正态曲线的说法错误的是（　　）

A.曲线与 x 轴之间区域的面积为 1

B.曲线在 $x=\mu$ 处达到峰值 $\frac{1}{\sigma\sqrt{2\pi}}$

C.当 σ 一定时，曲线的位置由 μ 确定，曲线随着 μ 的变化而沿 x 轴平移

D.当 μ 一定时，曲线的形状由 σ 确定，σ 越小，曲线越“矮胖”

D　**解析**：由随机变量的取值落在该区域的概率为

1 可知 A 正确；因为 $-\frac{(x-\mu)^2}{2\sigma^2}\leqslant 0$，所以 $\varphi(x)\leqslant\frac{1}{\sigma\sqrt{2\pi}}$，当且仅当 $x=\mu$ 时取等号，故 B 正确；当 σ 一定时，曲线的形状是固定的，曲线关于直线 $x=\mu$ 对称，故 C 正确；当 μ 一定时，曲线的对称轴固定，所以 σ 越小，曲线的峰值 $\frac{1}{\sigma\sqrt{2\pi}}$ 越大，故曲线越"瘦高"，故 D 错误.故选 D.

3. 已知随机变量 ξ 服从正态分布 $N(1,\sigma^2)$.若 $P(\xi>3)=0.012$，则 $P(-1\leqslant\xi\leqslant 1)=$　　(　　)

A.0.976　　B.0.024

C.0.488　　D.0.048

C　**解析**：因为随机变量 ξ 服从正态分布 $N(1,\sigma^2)$，故其正态曲线关于直线 $x=1$ 对称.又 $P(\xi>3)=0.012$，故 $P(\xi<-1)=0.012$.因此 $P(-1\leqslant\xi\leqslant 1)=0.5-P(\xi<-1)=0.5-0.012=0.488$.

4. 设随机变量 ξ 服从正态分布 $N(3,\sigma^2)$.若 $P(\xi\geqslant m)=a$，则 $P(\xi\geqslant 6-m)=$　　(　　)

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 $P(\xi\geqslant m)=P(\xi\leqslant 6-m)=a$.则 $P(\xi>6-m)=1-a$.

5. 我国成功举办了 2022 年第 24 届冬季奥林匹克运动会，其中的高山速降运动给我们以速度与激情的完美展现.某选手的速度 ξ(单位：km/h)服从正态分布 $N(100,\sigma^2)(\sigma>0)$.若 ξ 在 $[80,120]$ 内的概率为 0.7，则他的速度超过 120 km/h 的概率为(　　)

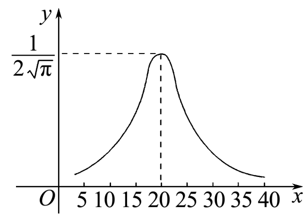
A.0.05　　B.0.1　　C.0.15　　D.0.2

C　**解析**：由题意可得 $\mu=100$，且 $P(80\leqslant\xi\leqslant 120)=0.7$，

所以 $P(\xi>120)=\frac{1-P(80\leqslant\xi\leqslant 120)}{2}=0.15$.

所以他的速度超过 120 km/h 的概率为 0.15.

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，最大值是 $\frac{1}{2\sqrt{\pi}}$，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 ξ(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 $x=4$，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 $0.5\times 0.5=0.25$.

综合性·创新提升

1. 一批电阻的电阻值 X(单位：Ω)服从正态分布 $N(1\ 000,5^2)$.现从甲、乙两箱电阻中各随机抽取一个电阻，测得电阻值分别为 1 011 Ω 和 982 Ω，可以认为　　(　　)

A.甲、乙两箱电阻均合格

B.甲、乙两箱电阻均不合格

C.甲箱电阻合格，乙箱电阻不合格

D.甲箱电阻不合格，乙箱电阻合格

C　**解析**：因为 $X\sim N(1\ 000,5^2)$，所以 $\mu=1\ 000$，$\sigma=5$，所以 $\mu-3\sigma=1\ 000-3\times 5=985$，$\mu+3\sigma=1\ 000+3\times 5=1\ 015$.因为 $1\ 011\in[985,1\ 015]$，$982\notin[985,1\ 015]$，

所以甲箱电阻合格，乙箱电阻不合格.

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 ξ 服从正态分布 $N(70,\sigma^2)$.若 ξ 在

$(50,70)$内的概率为0.4，则ξ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ξ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ξ在$(70,100)$内的概率为0.5.因为ξ在$(50,70)$内的概率为0.4，所以ξ在$(70,90)$内的概率也为0.4，则ξ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$f(x)=\frac{1}{\sqrt{2\pi}\times 10}e^{-\frac{(x-80)^2}{200}}(x\in\mathbf{R})$，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**：因为正态密度函数为

$f(x)=\frac{1}{\sqrt{2\pi}\times 10}\cdot e^{-\frac{(x-80)^2}{200}}(x\in\mathbf{R})$，

所以该市这次考试的数学平均成绩为80，

数学成绩的标准差为10.

易知正态密度曲线关于直线$x=80$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)$P(-4\leqslant X\leqslant 8)=P(2-2\times 3\leqslant X\leqslant 2+2\times 3)\approx 0.954\ 5$.

5.某工厂为了监控某种零件的一条生产线的生产过程，检验员每天从该生产线上随机抽取10个零件，并测量其尺寸（单位：cm）.根据长期生产经验，可以认为这条生产线在正常状态下生产的零件的尺寸服从正态分布$N(10,0.01)$.

(1)假设生产状态正常，记X表示一天内抽取的10个零件中尺寸在$(9.8,10.2)$之外的零件数，求$P(X\geqslant 1)$的值及X的数学期望.

(2)该工厂对生产的零件制定了两种销售方案（假设每种方案对销售量没有影响）：

方案一：每个零件均按85元定价销售.

方案二：若零件的实际尺寸在$(9.8,10.1)$范围内，则该零件为A级零件，每个零件定价100元；否则为B级零件，每个零件定价为30元.

哪种销售方案的利润更大？请根据数据计算说明.

附：若$X\sim N(\mu,\sigma^2)$，则$P(\mu-\sigma\leqslant X\leqslant\mu+\sigma)\approx 0.682\ 7$，$P(\mu-2\sigma\leqslant X\leqslant\mu+2\sigma)\approx 0.954\ 5$，$0.954\ 5^{10}\approx 0.627\ 7$.

解：(1)由题易知，抽取的零件的尺寸在$(9.8,10.2)$之内的概率为0.954 5，

从而零件的尺寸在$(9.8,10.2)$之外的概率为0.045 5.

所以$X\sim B(10,0.045\ 5)$，因此$P(X\geqslant 1)=1-P(X=0)=1-C_{10}^{0}\times 0.045\ 5^{0}\times 0.954\ 5^{10}\approx 0.372\ 3$.

故$E(X)=10\times 0.045\ 5=0.455$.

(2)对方案二，设每个零件的价格为Y元，故Y的可能取值为30，100，

可得$P(Y=100)\approx\frac{0.682\ 7+0.954\ 5}{2}=0.818\ 6$，

$P(Y=30)=1-P(Y=100)=0.181\ 4$.

故$E(Y)=30\times 0.181\ 4+100\times 0.818\ 6=87.302>85$.

又方案一中，每个零件售价均为85，故可得方案二的利润更大.

单元活动研习

单元知识重构

- 随机变量及其分布
  - 条件概率
    - 加法公式
    - 乘法公式
      - 全概率公式、贝叶斯公式*
  - 离散型随机变量
    - 分布列
    - 均值和方差
      - 二项分布
      - 超几何分布
  - 连续型随机变量
    - 正态分布
      - 正态密度曲线
      - 3σ原则

专题任务探究

探究点一　条件概率与全概率公式

例1　采购员要购买10个一包的电器元件，他的采购方法是：从一包中随机抽查3个，如果这3个元件都是好的，他才买下这一包.假定含有4个次品的包数占30%，而其余包中各含1个次品.求采购员拒绝购买的概率.

解：设B_1="取到的一包含4个次品"，

B_2="取到的一包含1个次品"，

A="采购员拒绝购买"，$P(B_1)=\dfrac{3}{10}$，$P(B_2)=\dfrac{7}{10}$.

$P(A|B_1)=1-\dfrac{C_6^3}{C_{10}^3}=\dfrac{5}{6}$，

$P(A|B_2)=1-\dfrac{C_9^3}{C_{10}^3}=\dfrac{3}{10}$.

由全概率公式得到

$P(A)=P(B_1)P(A|B_1)+P(B_2)P(A|B_2)=\dfrac{3}{10}\times\dfrac{5}{6}+\dfrac{7}{10}\times\dfrac{3}{10}=\dfrac{23}{50}$.

[一题多思]

思考. 在采购员拒绝购买的条件下，抽中的一包中含有4个次品的概率.

解：$P(B_1|A)=\dfrac{P(B_1)P(A|B_1)}{P(A)}=\dfrac{\dfrac{3}{10}\times\dfrac{5}{6}}{\dfrac{23}{50}}=\dfrac{25}{46}$.

反思提炼

利用全概率公式求解概率问题的步骤

(1)判断所求问题是否为全概率类型.

(2)若是，正确假设完备事件组$\{A_i\}$及待求概率的事件B.

(3)计算$P(A_i)$，$P(B|A_i)$.

(4)将(3)所得代入全概率公式计算出$P(B)$.

探究训练

1. 将三枚质地均匀的骰子各掷一次，设事件A="向上的点数都不相同"，B="至少出现一个6点向上"，则$P(B|A)=$　　(　　)

A.$\dfrac{60}{91}$　　B.$\dfrac{1}{2}$　　C.$\dfrac{5}{18}$　　D.$\dfrac{91}{216}$

B　**解析**：由题意得$P(A)=\dfrac{A_6^3}{6^3}=\dfrac{5}{9}$，$P(AB)=\dfrac{C_3^1A_5^2}{6^3}=\dfrac{5}{18}$，所以$P(B|A)=\dfrac{P(AB)}{P(A)}=\dfrac{\dfrac{5}{18}}{\dfrac{5}{9}}=\dfrac{1}{2}$.

2. 某商店收进甲厂生产的产品30箱，乙厂生产的同种产品20箱，甲厂每箱装100个，废品率为0.06，乙厂每箱装120个，废品率为0.05.求：

(1)任取一箱，从中任取的一个为废品的概率；

(2)将所有产品开箱混放，从中任取的一个为废品的概率.

解：记事件A="取到甲厂生产的产品"，B="取到乙厂生产的产品"，C="任取的一个为废品".

(1)由题意，得$P(A)=\dfrac{30}{50}=0.6$，$P(B)=\dfrac{20}{50}=0.4$，

$P(C|A)=0.06$，$P(C|B)=0.05$.

由全概率公式，得$P(C)=P(A)P(C|A)+P(B)P(C|B)=0.056$.

(2)$P(A)=\dfrac{30\times100}{30\times100+20\times120}=\dfrac{5}{9}$，

$P(B)=\dfrac{20\times120}{30\times100+20\times120}=\dfrac{4}{9}$，

$P(C|A)=0.06$，$P(C|B)=0.05$.

由全概率公式，得 $P(C)=P(A)P(C|A)+P(B)\cdot P(C|B)=\frac{5}{9}\times\frac{6}{100}+\frac{4}{9}\times\frac{5}{100}=\frac{1}{18}$.

探究点二　离散型随机变量的分布列、均值与方差

例 2　已知随机变量 X 的分布列为 $P(X=i)=\frac{i}{2a}(i=1,2,3,4)$，则 $P(2<X\leqslant4)$ 等于（　　）

A. $\frac{9}{10}$　　B. $\frac{7}{10}$

C. $\frac{3}{5}$　　D. $\frac{1}{2}$

B　**解析**：由分布列的性质知，$\frac{1}{2a}+\frac{2}{2a}+\frac{3}{2a}+\frac{4}{2a}=1$，则 $a=5$，所以 $P(2<X\leqslant4)=P(X=3)+P(X=4)=\frac{3}{10}+\frac{4}{10}=\frac{7}{10}$.

［**一题多思**］

思考 1. 已知随机变量 X 的分布列为 $P(X=i)=\frac{i}{2a}(i=1,2,3,4)$，则 $E(X)=$________.

3　**解析**：由例题可知 $a=5$，所以 $E(X)=1\times\frac{1}{10}+2\times\frac{2}{10}+3\times\frac{3}{10}+4\times\frac{4}{10}=3$.

思考 2. 已知随机变量 X 的分布列为 $P(X=i)=\frac{i}{2a}(i=1,2,3,4)$，则 $D(X)=$________.

1　**解析**：由例题及思考 1 可知，$a=5$，$E(X)=3$，所以 $D(X)=(1-3)^2\times\frac{1}{10}+(2-3)^2\times\frac{2}{10}+(3-3)^2\times\frac{3}{10}+(4-3)^2\times\frac{4}{10}=1$.

例 3　一次同时投掷两枚相同的正方体骰子（骰子质地均匀，且各面分别刻有 1，2，2，3，3，3 六个数字）. 设随机变量 η 表示一次掷得的点数和.

(1)求 η 的分布列；

(2)求 η 的均值和方差.

解：(1)由已知，得随机变量 η 的所有可能取值为 2，3，4，5，6，

$P(\eta=2)=\frac{1}{6\times6}=\frac{1}{36}$，

$P(\eta=3)=\frac{2\times1\times2}{6\times6}=\frac{1}{9}$，

$P(\eta=4)=\frac{2\times1\times3+2\times2}{6\times6}=\frac{5}{18}$，

$P(\eta=5)=\frac{2\times2\times3}{6\times6}=\frac{1}{3}$，

$P(\eta=6)=\frac{3\times3}{6\times6}=\frac{1}{4}$.

故 η 的分布列为

η	2	3	4	5	6
P	$\frac{1}{36}$	$\frac{1}{9}$	$\frac{5}{18}$	$\frac{1}{3}$	$\frac{1}{4}$

(2)由(1)得 $E(\eta)=2\times\frac{1}{36}+3\times\frac{1}{9}+4\times\frac{5}{18}+5\times\frac{1}{3}+6\times\frac{1}{4}=\frac{14}{3}$，$D(\eta)=\left(2-\frac{14}{3}\right)^2\times\frac{1}{36}+\left(3-\frac{14}{3}\right)^2\times\frac{1}{9}+\left(4-\frac{14}{3}\right)^2\times\frac{5}{18}+\left(5-\frac{14}{3}\right)^2\times\frac{1}{3}+\left(6-\frac{14}{3}\right)^2\times\frac{1}{4}=\frac{10}{9}$.

反思提炼

求离散型随机变量 X 的均值与方差的步骤

(1)理解 X 的意义，写出 X 的可能取值.

(2)求 X 取每个值的概率.

(3)写出 X 的分布列.

(4)由均值的定义求 $E(X)$.

(5)由方差的定义求 $D(X)$.

探究训练

1. 设 $a>0$，若随机变量 ξ 的分布列如下：

ξ	-1	0	2
P	a	$2a$	$3a$

则下列方差中最大的是（　　）

A. $D(\xi)$　　B. $D(|\xi|)$

C. $D(2\xi-1)$　　D. $D(2|\xi|+1)$

C　**解析**：由题意 $a+2a+3a=1$，得 $a=\frac{1}{6}$.

$E(\xi)=(-1)\times\frac{1}{6}+0\times\frac{1}{3}+2\times\frac{1}{2}=\frac{5}{6}$，

$E(|\xi|)=1\times\frac{1}{6}+0\times\frac{1}{3}+2\times\frac{1}{2}=\frac{7}{6}$，

$D(\xi)=\frac{1}{6}\times\left(-1-\frac{5}{6}\right)^2+\frac{1}{3}\times\left(0-\frac{5}{6}\right)^2+\frac{1}{2}\times\left(2\right.$

$-\frac{5}{6}\Big)^2=\frac{53}{36}$,

$D(|\xi|)=\frac{1}{6}\times\left(1-\frac{7}{6}\right)^2+\frac{1}{3}\times\left(0-\frac{7}{6}\right)^2+\frac{1}{2}\times\Big(2$

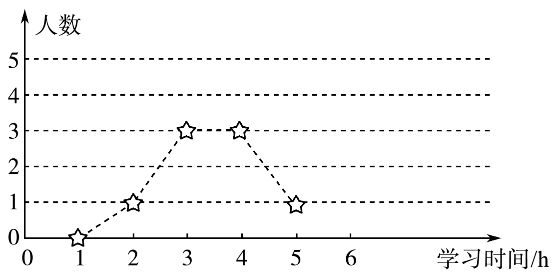
$-\frac{7}{6}\Big)^2=\frac{29}{36}$,

$D(2\xi-1)=4\times\frac{53}{36}=\frac{53}{9}$,

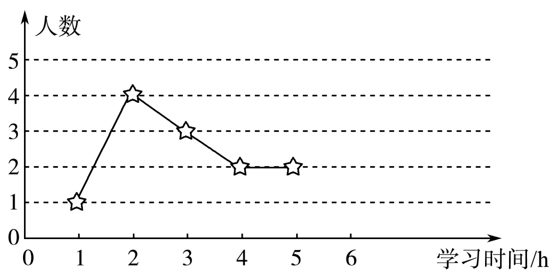
$D(2|\xi|+1)=4\times\frac{29}{36}=\frac{29}{9}$.

综上可知,其中最大的是 $D(2\xi-1)$.故选 C.

2. 为了解学生寒假期间的学习情况,学校对某班男、女学生每天的学习时间进行了调查,学习时间按整小时统计,根据调查结果绘制折线图如下:



女生学习时间统计图



男生学习时间统计图

(1)已知该校有 400 名学生,试估计全校学生中,每天的学习时间不足 4 h 的人数;

(2)若从学习时间不少于 4 h 的学生中任选 4 人,设选的男生人数为 X,求随机变量 X 的分布列及均值 $E(X)$;

(3)试比较男生学习时间的方差 s_1^2 与女生学习时间的方差 s_2^2 的大小.(只需写出结论)

解:(1)由题图可得共抽取了 20 人,其中男生学习时间不足 4 h 的有 8 人,女生学习时间不足 4 h 的有 4 人.

故可估计全校学生中每天学习时间不足 4 h 的人数为 $400\times\frac{12}{20}=240$.

(2)学习时间不少于 4 h 的学生共 8 人,其中男生 4 人,故 X 的所有可能取值为 0,1,2,3,4.

由题意可得

$P(X=0)=\frac{C_4^4}{C_8^4}=\frac{1}{70}$, $P(X=1)=\frac{C_4^1C_4^3}{C_8^4}=\frac{16}{70}=\frac{8}{35}$,

$P(X=2)=\frac{C_4^2C_4^2}{C_8^4}=\frac{36}{70}=\frac{18}{35}$,

$P(X=3)=\frac{C_4^3C_4^1}{C_8^4}=\frac{16}{70}=\frac{8}{35}$,

$P(X=4)=\frac{C_4^4}{C_8^4}=\frac{1}{70}$.

所以随机变量 X 的分布列为

X	0	1	2	3	4
P	$\frac{1}{70}$	$\frac{8}{35}$	$\frac{18}{35}$	$\frac{8}{35}$	$\frac{1}{70}$

$E(X)=0\times\frac{1}{70}+1\times\frac{8}{35}+2\times\frac{18}{35}+3\times\frac{8}{35}+4\times\frac{1}{70}$

$=2$.

(3) $s_1^2>s_2^2$.

探究点三　两点分布与二项分布

例 4　若离散型随机变量 X 的分布列为

X	1	0
P	$2a$	a

则常数 $a=$________,X 的数学期望 $E(X)=$________.

$\frac{1}{3}$　$\frac{2}{3}$　**解析**:由题得 $2a+a=1$,所以 $a=\frac{1}{3}$,所以 $E(X)=1\times\frac{2}{3}+0\times\frac{1}{3}=\frac{2}{3}$.

[一题多思]

思考.若随机变量 X 的分布列为

X	0	1
P	$\frac{2}{3}$	m

则 $D(X)=$________.

$\frac{2}{9}$　**解析**:由题得 $\frac{2}{3}+m=1$,

所以 $m=\frac{1}{3}$,所以 $D(X)=\frac{2}{3}\times\frac{1}{3}=\frac{2}{9}$.

例 5　某市推行“共享汽车”服务,租用该车按行驶里程加用车时间收费,标准是“1 元/km+0.2 元/min”.刚参加工作的小刘拟租用该车上下班,同单位的邻居老李告诉他,上班往返总路程虽然只有 10 km,但偶尔开车上下班总共也需花费大约 1 h”,并将

自己近 50 天往返开车花费的时间（单位：min）情况统计如表：

时间	[15,25)	[25,35)	[35,45)	[45,55)	[55,65)
次数	10	18	12	8	2

将老李统计的各时间段的频率视为相应概率，假定往返的路程不变，而且每次路上开车花费时间视为用车时间.

(1)试估计小刘平均每天支付的租车费用（每个时间段以该时间段中点时间计算）；

(2)小刘认为只要上下班开车总用时不超过 45 min，租用"共享汽车"为他该日的"最优选择"，小刘拟租用该车上下班 2 天，设其中有 ξ 天为"最优选择"，求 ξ 的分布列和数学期望.

解：(1)由题可得如下租车费用（单位：元）与相应频率的统计表：

费用	14	16	18	20	22
频率	0.2	0.36	0.24	0.16	0.04

估计小刘平均每天支付的租车费用为 $14\times0.2+16\times0.36+18\times0.24+20\times0.16+22\times0.04=16.96$.

(2)ξ 可能的取值为 0，1，2，往返开车用时不超过 45 min 的概率为 0.8，由题意知 $\xi\sim B(2,0.8)$，

则 $P(\xi=0)=C_2^0\times0.8^0\times0.2^2=0.04$，

$P(\xi=1)=C_2^1\times0.8^1\times0.2^1=0.32$，

$P(\xi=2)=C_2^2\times0.8^2\times0.2^0=0.64$.

故 ξ 的分布列为

ξ	0	1	2
P	0.04	0.32	0.64

则 $E(\xi)=2\times0.8=1.6$.

反思提炼

随机变量 X 服从二项分布的三个前提条件

(1)每次试验都是在同一条件下进行的.

(2)每次试验的结果相互独立.

(3)每次试验出现的结果只有两个，即事件 A 要么发生，要么不发生.

只有这三个条件均满足时才能说明随机变量 X 服从二项分布，事件 A 在 n 次独立重复试验中恰好发生 k 次的概率可用下面的公式计算.

$P(X=k)=C_n^k p^k(1-p)^{n-k}$，$k=0,1,2,\cdots,n$.

其中 p 为每次试验中事件 A 发生的概率.

探究训练

1.若随机变量 Y 服从二项分布 $B(n,p)$，且 $E(Y)=3.6$，$D(Y)=2.16$，则下列结论正确的是　（　　）

A.$Y\sim B(4,0.9)$

B.$Y\sim B(9,0.4)$

C.$Y\sim B(18,0.2)$

D.$Y\sim B(36,0.1)$

B　**解析**：因为随机变量 Y 服从二项分布 $B(n,p)$，且 $E(Y)=3.6$，$D(Y)=2.16$，

所以 $\begin{cases}np=3.6 ①,\\ np(1-p)=2.16 ②.\end{cases}$

得 $1-p=0.6$，即 $p=0.4$.

代入①解得 $n=9$.

所以 $Y\sim B(9,0.4)$.

2.某公司招聘员工，先由两位专家面试，若这两位专家都同意通过，则通过初审并予以录用；若这两位专家都未同意通过，则未通过初审并不予录用；当这两位专家意见不一致时，再由第三位专家进行复审，若能通过复审则予以录用，否则不予录用.设应聘人员获得每位初审专家通过的概率均为 $\frac{1}{2}$，获得复审专家通过的概率为 $\frac{3}{10}$，各专家评审的结果相互独立.

(1)求某应聘人员被录用的概率；

(2)若 4 人应聘，设 X 为被录用的人数，试求随机变量 X 的分布列及期望.

解：(1)设"某应聘人员被录用"为事件 A，则 $P(A)=\frac{1}{2}\times\frac{1}{2}+C_2^1\times\frac{1}{2}\times\left(1-\frac{1}{2}\right)\times\frac{3}{10}=\frac{2}{5}$.

(2)根据题意，$X\sim B\left(4,\frac{2}{5}\right)$，

则 $P(X=0)=C_4^0\times\left(\frac{3}{5}\right)^4=\frac{81}{625}$，

$P(X=1)=C_4^1\times\frac{2}{5}\times\left(\frac{3}{5}\right)^3=\frac{216}{625}$，

$P(X=2)=C_4^2\times\left(\frac{2}{5}\right)^2\times\left(\frac{3}{5}\right)^2=\frac{216}{625}$，

$P(X=3)=C_4^3\times\left(\frac{2}{5}\right)^3\times\frac{3}{5}=\frac{96}{625}$,

$P(X=4)=C_4^4\times\left(\frac{2}{5}\right)^4=\frac{16}{625}$.

故 X 的分布列为

X	0	1	2	3	4
P	$\frac{81}{625}$	$\frac{216}{625}$	$\frac{216}{625}$	$\frac{96}{625}$	$\frac{16}{625}$

$E(X)=np=4\times\frac{2}{5}=\frac{8}{5}$.

探究点四　超几何分布

例 6　某从事智能教育技术研发的科技公司开发了一个软件，并且在甲、乙两个学校的高二学生中做用户测试.经过一个阶段的试用，为了解该软件对学生学习的促进情况，该公司从参与测试的学生中随机抽200名学生，对他们“向量的数量积”这一知识点的掌握情况进行调查，调查结果如表：

掌握情况	甲校		乙校	
	使用软件	不使用软件	使用软件	不使用软件
基本掌握	32	28	50	30
没有掌握	8	14	12	26

用样本频率估计总体概率，并假设每位学生是否掌握“向量的数量积”这一知识点相互独立.

(1)从两校高二学生中随机抽1人，估计该学生对“向量的数量积”这一知识点基本掌握的概率；

(2)从样本中没有掌握“向量的数量积”这一知识点的学生中随机抽2名学生，以 ξ 表示这2人中使用软件的人数，求 ξ 的分布列和数学期望.

解：(1)在两所学校被调查的200名学生中，

对“向量的数量积”这一知识点基本掌握的学生有140人，

所以估计从两校高二学生中随机抽1人，该学生对“向量的数量积”这一知识点基本掌握的概率为 $\frac{140}{200}=0.7$.

(2)依题意，ξ 的可能取值为0，1，2，且 $P(\xi=0)=\frac{C_{20}^0C_{40}^2}{C_{60}^2}=\frac{26}{59}$,

$P(\xi=1)=\frac{C_{20}^1C_{40}^1}{C_{60}^2}=\frac{80}{177}$，$P(\xi=2)=\frac{C_{20}^2C_{40}^0}{C_{60}^2}=\frac{19}{177}$.

所以 ξ 的分布列为

ξ	0	1	2
P	$\frac{26}{59}$	$\frac{80}{177}$	$\frac{19}{177}$

故 $E(\xi)=\frac{2\times20}{60}=\frac{2}{3}$.

反思提炼

1.超几何分布的概率计算方法

首先明确随机变量是否服从超几何分布，把握等可能、不放回两个特点；其次是明确公式中的参数，即 N,M,n 的值；最后代入公式计算概率.

2.超几何分布与二项分布的联系与区别

	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，总体不改变
	各次抽取不相互独立	各次抽取相互独立
联系	当总体容量很大时，超几何分布可近似看作二项分布	

探究训练

老师要从10篇课文中随机抽3篇让学生背诵，规定至少要背出其中的2篇才能及格.某同学只能背诵其中的6篇，试求：

(1)抽到他能背诵的课文的数量 X 的分布列；

(2)他能及格的概率.

解：(1)依题意，X 的取值可能为0，1，2，3，X 服从超几何分布，

则 $P(X=0)=\frac{C_6^0C_4^3}{C_{10}^3}=\frac{1}{30}$,

$P(X=1)=\frac{C_6^1C_4^2}{C_{10}^3}=\frac{3}{10}$,

$P(X=2)=\frac{C_6^2C_4^1}{C_{10}^3}=\frac{1}{2}$,

$P(X=3)=\frac{C_6^3C_4^0}{C_{10}^3}=\frac{1}{6}$.

故 X 的分布列为

X	0	1	2	3
P	$\frac{1}{30}$	$\frac{3}{10}$	$\frac{1}{2}$	$\frac{1}{6}$

(2)他能及格的概率为 $P(X\geqslant 2)=P(X=2)+P(X=3)=\frac{1}{2}+\frac{1}{6}=\frac{2}{3}$.

探究点五　正态分布

例 7　(1)已知某批零件的长度误差 X(单位:mm)服从正态分布 $N(0,3^2)$,从中随机取一件,其长度误差落在区间(3,6]内的概率为　(　　)

A.0.045 6　　B.0.135 9

C.0.271 8　　D.0.317 4

B　**解析**:因为零件的长度误差服从正态分布 $N(0,3^2)$,所以正态曲线关于直线 $x=0$ 对称.故 $P(3<X\leqslant 6)=\frac{1}{2}(P(-6\leqslant X\leqslant 6)-P(-3\leqslant X\leqslant 3))\approx\frac{1}{2}\times(0.954\ 5-0.682\ 7)=0.135\ 9$.

(2)已知随机变量 X 服从正态分布 $N(3,1)$,且 $P(X>2c-1)=P(X<c+3)$,则 $c=$________.

$\frac{4}{3}$　**解析**:因为 $X\sim N(3,1)$,所以正态曲线关于直线 $x=3$ 对称.又因为 $P(X>2c-1)=P(X<c+3)$,所以 $2c-1+c+3=3\times 2$,所以 $c=\frac{4}{3}$.

反思提炼

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 3σ 原则求概率时,注意把给出的区间或范围与 $[\mu-\sigma,\mu+\sigma]$,$[\mu-2\sigma,\mu+2\sigma]$,$[\mu-3\sigma,\mu+3\sigma]$ 建立联系.

探究训练

现有 1 000 名学生参加数学测试,已知测试成绩 X 近似服从正态分布 $N(110,\sigma^2)(\sigma>0)$,试卷满分 150 分.统计结果显示测试成绩优秀(高于 135 分)的人数占总人数的 $\frac{1}{5}$,则此次测试成绩在 85 分到 110 分之间的人数约为　(　　)

A.200　　B.300　　C.400　　D.500

B　**解析**:因为 $P(X<85)=P(X>135)=0.2$,所以 $P(85\leqslant X\leqslant 135)=1-0.4=0.6$.所以 $P(85\leqslant X\leqslant 110)=\frac{1}{2}P(85\leqslant X\leqslant 135)=0.3$.所以此次数学测试成绩在 85 分到 110 分之间的人数约为 $1\ 000\times 0.3=300$.

第七章综合检测

（时间：120 分钟，分值：150 分）

一、单项选择题（本题共 8 小题，每小题 5 分，共40 分）

1.抛掷一枚骰子两次，下列各量不是随机变量的是　（　　）

A.1 点向上的次数

B.抛掷骰子的次数

C.2 点向上的次数

D.1 点或 2 点向上的次数

B　**解析**：抛掷次数是个确定的值，不是随机变量.

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甲没有准确预报，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 $P(B|A)=\dfrac{P(AB)}{P(A)}=\dfrac{0.4}{0.5}=0.8$.

故选 A.

4.已知随机变量 $X\sim B\left(6,\dfrac{1}{2}\right)$，则 $D(2X+1)$等于　（　　）

A.6　　B.4

C.3　　D.9

A　**解析**：因为 $D(2X+1)=4D(X)$，$D(X)=6\times\dfrac{1}{2}\times\left(1-\dfrac{1}{2}\right)=\dfrac{3}{2}$，所以 $D(2X+1)=4\times\dfrac{3}{2}=6$.

5.箱子中放有 3 个白球、2 个黑球，从中不放回地取球两次，每次任取一球，用 A_1 表示"第一次取得白球"，A_2 表示"第二次取得白球"，则 A_1 和 A_2（　　）

A.是互斥事件

B.是相互独立事件

C.是对立事件

D.不是相互独立的事件

D　**解析**：因为 $P(A_1)=\dfrac{3}{5}$，若 A_1 发生了，$P(A_2)=\dfrac{2}{4}=\dfrac{1}{2}$；若 A_1 不发生，$P(A_2)=\dfrac{3}{4}$，所以 A_1 是否发生对 A_2 发生的概率有影响，所以 A_1 与 A_2 不是相互独立事件.

6.若随机变量 η 的概率分布列如下：

η	-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，

则 $P(A_1)=\frac{5}{10}$，$P(A_2)=\frac{3}{10}$，$P(A_3)=\frac{2}{10}$，

$P(B|A_1)=\frac{1}{10}$，$P(B|A_2)=\frac{1}{15}$，$P(B|A_3)=\frac{1}{20}$.

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（本题共 3 小题，每小题 6 分，共 18 分）

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；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 满足如表所示的分布列，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.（2023·新高考全国Ⅱ卷）在信道内传输 0，1 信号，信号的传输相互独立.发送 0 时，收到 1 的概率为 $\alpha(0<\alpha<1)$，收到 0 的概率为 $1-\alpha$；发送 1 时，收到 0 的概率为 $\beta(0<\beta<1)$，收到 1 的概率为 $1-\beta$.考虑两种传输方案：单次传输和三次传输.单次传输是指每个信号只发送 1 次，三次传输是指每个信号重复发送 3 次.收到的信号需要译码，译码规则如下：单次传输时，收到的信号即为译码；三次传输时，收到的信号中出现次数多的即为译码（例如，若依次收到 1，0，1，则译码为 1）.下列说法正确的是　（　　）

A.采用单次传输方案，若依次发送 1，0，1，则依次收到 1，0，1 的概率为 $(1-\alpha)(1-\beta)^2$

B.采用三次传输方案，若发送 1，则依次收到 1，0，1 的概率为 $\beta(1-\beta)^2$

C.采用三次传输方案，若发送 1，则译码为 1 的概率为 $\beta(1-\beta)^2+(1-\beta)^3$

D.当 $0<\alpha<0.5$ 时，若发送 0，则采用三次传输方案译码为 0 的概率大于采用单次传输方案译码为 0 的概率

ABD　**解析**：对于 A，依次发送 1，0，1，则依次收到 1，0，1 的事件是发送 1 收到 1、发送 0 收到 0、发送 1 收到 1 这 3 个事件的积，它们相互独立，所以所求概率为 $(1-\beta)(1-\alpha)(1-\beta)=(1-\alpha)(1-\beta)^2$，A 正确；

对于 B，三次传输，发送 1，相当于依次发送 1，1，1，则依次收到 1，0，1 的事件是发送 1 收到 1、发送 1 收到 0、发送 1 收到 1 这 3 个事件的积，

它们相互独立，所以所求概率为 $(1-\beta)\cdot\beta\cdot(1-\beta)=\beta(1-\beta)^2$，B 正确；

对于 C，三次传输，发送 1，则译码为 1 的事件是依次收到 1，1，0 或 1，0，1 或 0，1，1 或 1，1，1 这 4 个事件的和，它们互斥，所以所求的概率为 $C_3^2\beta(1-\beta)^2+(1-\beta)^3=(1-\beta)^2(1+2\beta)$，C 错误；

对于 D，由选项 C 知，三次传输，发送 0，则译码为 0 的概率 $P=(1-\alpha)^2(1+2\alpha)$，

单次传输发送 0，则译码为 0 的概率 $P'=1-\alpha$.而 $0<\alpha<0.5$，因此 $P-P'=(1-\alpha)^2(1+2\alpha)-(1-\alpha)=\alpha(1-\alpha)(1-2\alpha)>0$，即 $P>P'$，D 正确.

故选 ABD.

三、填空题（本题共 3 小题，每小题 5 分，共 15 分）

12. 设 $P(A|B)=P(B|A)=\dfrac{1}{2}$，$P(A)=\dfrac{1}{3}$，则 $P(B)$ 等于________.

$\dfrac{1}{3}$　**解析**：因为 $P(B|A)=\dfrac{P(AB)}{P(A)}$，

所以 $P(AB)=P(B|A)\cdot P(A)=\dfrac{1}{2}\times\dfrac{1}{3}=\dfrac{1}{6}$，

所以 $P(B)=\dfrac{P(AB)}{P(A|B)}=\dfrac{\frac{1}{6}}{\frac{1}{2}}=\dfrac{1}{3}$.

13. 袋中有 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 $P(X\leqslant 6)=$________.

$\dfrac{13}{35}$　**解析**：$P(X\leqslant 6)=P(X=4)+P(X=6)$

$=\dfrac{C_4^4+C_4^3C_3^1}{C_7^4}=\dfrac{13}{35}$.

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）

$\dfrac{6}{25}$　**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 $\dfrac{2}{5}$，取得负数的概率为 $\dfrac{1}{2}$，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 $C_3^2\times\left(\dfrac{2}{5}\right)^2\times\left(\dfrac{1}{2}\right)^1=\dfrac{6}{25}$.

四、解答题（本题共 5 小题，共 77 分）

15.（13 分）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 $P(X=k)=\dfrac{C_3^kC_7^{3-k}}{C_{10}^3}(k=0,1,2,3)$.

所以 $P(X=0)=\dfrac{C_3^0C_7^3}{C_{10}^3}=\dfrac{35}{120}=\dfrac{7}{24}$；

$P(X=1)=\dfrac{C_3^1C_7^2}{C_{10}^3}=\dfrac{63}{120}=\dfrac{21}{40}$；

$P(X=2)=\dfrac{C_3^2C_7^1}{C_{10}^3}=\dfrac{21}{120}=\dfrac{7}{40}$；

$P(X=3)=\frac{C_3^3C_7^0}{C_{10}^3}=\frac{1}{120}$.

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.(15分)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 的值可取0，1，2，3，4，5，且服从二项分布 $B\left(5,\frac{1}{3}\right)$，则

$P(X=0)=C_5^0\cdot\left(\frac{1}{3}\right)^0\cdot\left(\frac{2}{3}\right)^5=\frac{32}{243}$，

$P(X=1)=C_5^1\cdot\left(\frac{1}{3}\right)^1\cdot\left(\frac{2}{3}\right)^4=\frac{80}{243}$，

$P(X=2)=C_5^2\cdot\left(\frac{1}{3}\right)^2\cdot\left(\frac{2}{3}\right)^3=\frac{80}{243}$，

$P(X=3)=C_5^3\cdot\left(\frac{1}{3}\right)^3\cdot\left(\frac{2}{3}\right)^2=\frac{40}{243}$，

$P(X=4)=C_5^4\cdot\left(\frac{1}{3}\right)^4\cdot\left(\frac{2}{3}\right)^1=\frac{10}{243}$，

$P(X=5)=C_5^5\cdot\left(\frac{1}{3}\right)^5\cdot\left(\frac{2}{3}\right)^0=\frac{1}{243}$.

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，其可能取值为0，1，2，3，4，5，且

$P(Y=0)=\left(\frac{2}{3}\right)^0\times\frac{1}{3}=\frac{1}{3}$，

$P(Y=1)=\left(\frac{2}{3}\right)^1\times\frac{1}{3}=\frac{2}{9}$，

$P(Y=2)=\left(\frac{2}{3}\right)^2\times\frac{1}{3}=\frac{4}{27}$，

$P(Y=3)=\left(\frac{2}{3}\right)^3\times\frac{1}{3}=\frac{8}{81}$，

$P(Y=4)=\left(\frac{2}{3}\right)^4\times\frac{1}{3}=\frac{16}{243}$，

$P(Y=5)=\left(\frac{2}{3}\right)^5=\frac{32}{243}$.

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.(15分)为研究某种农产品价格变化的规律，收集得到了该农产品连续40天的价格变化数据，如下

表所示.在描述价格变化时,用“+”表示“上涨”,即当天价格比前一天价格高;用“−”表示“下跌”,即当天价格比前一天价格低;用“0”表示“不变”,即当天价格与前一天价格相同.

时段	价格变化																			
第 1 天到第 20 天	−	+	+	0	−	−	−	+	+	0	+	0	−	−	+	−	+	0	0	+
第 21 天到第 40 天	0	+	+	0	−	−	−	+	+	0	+	0	+	−	−	−	+	0	−	+

(1)试估计该农产品价格“上涨”的概率;

(2)假设该农产品每天的价格变化是相互独立的.在未来的日子里任取 4 天,试估计该农产品价格在这 4 天中 2 天“上涨”、1 天“下跌”、1 天“不变”的概率;

(3)假设该农产品每天的价格变化只受前一天价格变化的影响,估计第 41 天该农产品价格“上涨”“下跌”和“不变”的概率哪个最大.

解:(1)根据表格中的数据可以看出,40 天里有 16 个+,即有 16 天价格是“上涨”的,

根据古典概型的概率计算公式,该农产品价格“上涨”的概率为$\frac{16}{40}=0.4$.

(2)在这 40 天里,有 16 天“上涨”,14 天“下跌”,10 天“不变”,则价格“上涨”“下跌”“不变”的概率分别是0.4,0.35,0.25,

所以在未来的日子里任取 4 天,该农产品价格 2 天“上涨”,1 天“下跌”,1 天“不变”的概率是 $C_4^2\times 0.4^2\times C_2^1\times 0.35\times 0.25=0.168$.

(3)由于第 40 天处于“上涨”状态,从前 39 天的 15 次“上涨”进行分析,“上涨”后下一次仍“上涨”的有 4 次,“不变”的有 9 次,“下跌”的有 2 次,

因此估计第 41 天价格“不变”的概率最大.

18.(17 分)经调查统计,某网民在网上光顾某小店,经过一番浏览后,对该店铺中的 A,B,C 三种商品有购买意向.该小店推出每买一种商品送 5 元优惠券的活动.已知该网民购买 A,B,C 三种商品的概率分别为$\frac{2}{3}$,p_1,p_2($p_1<p_2$),至少购买一种商品的概率为$\frac{23}{24}$,最多购买两种商品的概率为$\frac{3}{4}$.假设该网民是否购买这三种商品相互独立.

(1)求该网民分别购买 B,C 两种商品的概率;

(2)用随机变量 X(单位:元)表示该网民购买商品所获得的优惠券金额,求 X 的分布列和数学期望.

解:(1)由题意可知至少购买一种商品的概率为$\frac{23}{24}$,

所以一种商品都不买的概率为 $1-\frac{23}{24}=\frac{1}{24}$,

即$\left(1-\frac{2}{3}\right)(1-p_1)(1-p_2)=\frac{1}{24}$①.

又因为最多购买两种商品的概率为$\frac{3}{4}$,

所以三种都买的概率为 $1-\frac{3}{4}=\frac{1}{4}$,

即$\frac{2}{3}p_1p_2=\frac{1}{4}$②.

联立①②,解得$\begin{cases}p_1=\frac{1}{2},\\ p_2=\frac{3}{4}\end{cases}$或$\begin{cases}p_1=\frac{3}{4},\\ p_2=\frac{1}{2}.\end{cases}$

因为 $p_1<p_2$,所以该网民购买 B,C 两种商品的概率分别为 $p_1=\frac{1}{2}$,$p_2=\frac{3}{4}$.

(2)由题意可得 X 的可能取值为 0,5,10,15.

且 $P(X=0)=\frac{1}{24}$,

$P(X=5)=\frac{2}{3}\times\frac{1}{2}\times\frac{1}{4}+\frac{1}{3}\times\frac{1}{2}\times\frac{1}{4}+\frac{1}{3}\times\frac{1}{2}\times\frac{3}{4}=\frac{1}{4}$,

$P(X=10)=\frac{2}{3}\times\frac{1}{2}\times\frac{1}{4}+\frac{2}{3}\times\frac{1}{2}\times\frac{3}{4}+\frac{1}{3}\times\frac{1}{2}\times\frac{3}{4}=\frac{11}{24}$,

$P(X=15)=\frac{2}{3}\times\frac{1}{2}\times\frac{3}{4}=\frac{1}{4}$.

所以 X 的分布列为

X	0	5	10	15
P	$\frac{1}{24}$	$\frac{1}{4}$	$\frac{11}{24}$	$\frac{1}{4}$

则 $E(X)=0\times\frac{1}{24}+5\times\frac{1}{4}+10\times\frac{11}{24}+15\times\frac{1}{4}$

$=\frac{115}{12}$.

19.(17 分)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 $\frac{1}{2}$,C,D 两枚为非均匀硬币,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.

$P(X=0)=C_2^0\cdot\left(1-\frac{1}{2}\right)^2\cdot C_2^0(1-a)^2=\frac{1}{4}(1-a)^2$,

$P(X=1)=C_2^1\cdot\left(\frac{1}{2}\right)^1\cdot\left(1-\frac{1}{2}\right)^1\cdot C_2^0(1-a)^2+C_2^0\cdot\left(1-\frac{1}{2}\right)^2\cdot C_2^1a(1-a)=\frac{1}{2}(1-a)$,

$P(X=2)=C_2^2\cdot\left(\frac{1}{2}\right)^2\cdot C_2^0(1-a)^2+C_2^1\cdot\left(\frac{1}{2}\right)^1\cdot\left(1-\frac{1}{2}\right)^1\cdot C_2^1a(1-a)+C_2^0\cdot\left(1-\frac{1}{2}\right)^2\cdot C_2^2a^2=\frac{1}{4}(1+2a-2a^2)$,

$P(X=3)=C_2^2\cdot\left(\frac{1}{2}\right)^2\cdot C_2^1a(1-a)+C_2^1\cdot\left(\frac{1}{2}\right)^1\cdot\left(1-\frac{1}{2}\right)^1\cdot C_2^2a^2=\frac{a}{2}$,

$P(X=4)=C_2^2\cdot\left(\frac{1}{2}\right)^2\cdot C_2^2a^2=\frac{1}{4}a^2$.

所以 X 的分布列为

X	0	1	2	3	4
P	$\frac{1}{4}(1-a)^2$	$\frac{1}{2}(1-a)$	$\frac{1}{4}(1+2a-2a^2)$	$\frac{a}{2}$	$\frac{1}{4}a^2$

X 的数学期望为 $E(X)=1\times\frac{1}{2}(1-a)+2\times\frac{1}{4}(1+2a-2a^2)+3\times\frac{a}{2}+4\times\frac{1}{4}a^2=2a+1$.

第八章　成对数据的统计分析

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少的趋势,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 $r<0$. (2)当 $\vert r\vert$ 越接近 1 时,成对样本数据的线性相关程度越强; 当 $\vert r\vert$ 越接近 0 时,成对样本数据的线性相关程度越弱. (3)$\vert r\vert=1$ 的充要条件是成对样本数据确定的点都在一条直线上

概念辨析

1.判断(正确的打"√",错误的打"×").

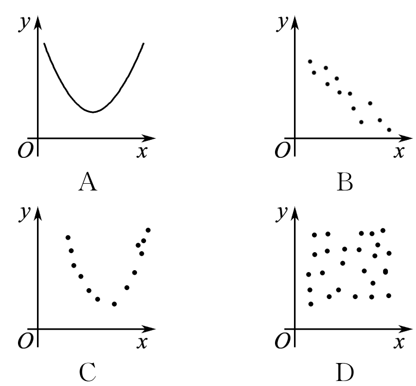
(1)两个变量之间产生相关关系的原因受许多不确定的随机因素的影响.　(√)

(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　(　　)



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的大小可能是　(　　)

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中的关系模型，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判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

学习任务一　变量间相关关系的判断

1. 下列各项中的两个变量是相关关系的是　(　　)

A.出租车费用与行驶的里程

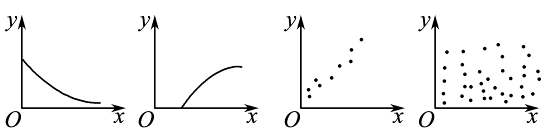
B.正方形的面积与边长

C.身高与体重

D.铁的体积与质量

C　**解析**：A，B，D 都是函数关系，只有 C 是相关关系.

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.

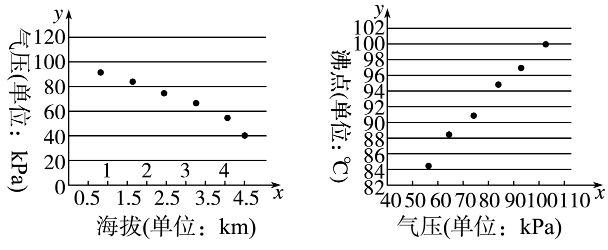
②③　**解析**：利用相关关系的概念进行判断.①④中两个变量之间的关系是一种确定性关系，而②③中的两个变量之间是不确定性关系，它们具有相关关系.

反思提炼

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.

学习任务二　变量正负相关的判断

例 1　某校地理学习兴趣小组根据在某座山测得的海拔、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减小，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减小，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.

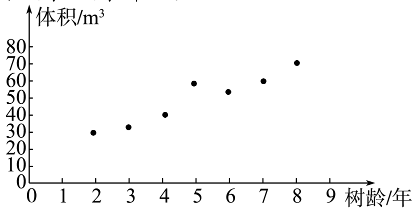
例 2　某种树木树干的体积(单位：m^3)与树龄(单位：年)之间有如下的对应关系：

树龄	2	3	4	5	6	7	8
体积	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.

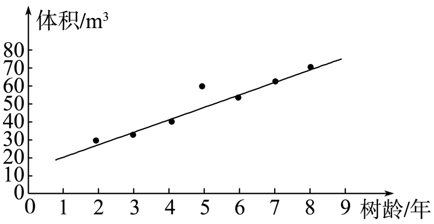


(2)由散点图发现树干体积随着树龄的增加呈现增加的趋势，且散点大致落在一条直线附近，所以树干的体积与树龄呈线性相关关系，且为正相关.

[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.

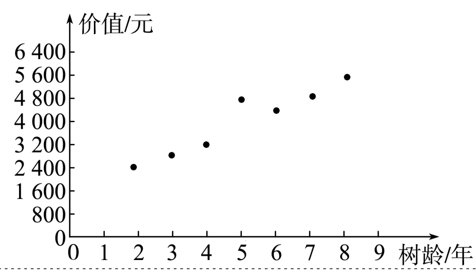


思考 2. 若该种树木的树干每立方米的价值是 80 元，作出树干的价值与树龄之间关系的散点图.

解：树干的价值与树龄之间的关系如表所示.

树龄	2	3	4	5	6	7	8
体积	30	34	40	60	55	62	70
价值	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.



探究训练

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（　　）

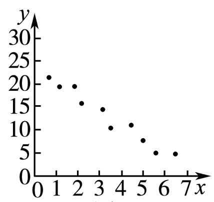


图 1

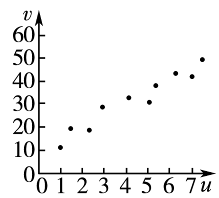


图 2

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.

反思提炼

1. 判断两个变量 x 和 y 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.

学习任务三　样本相关系数

例 3　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 2024 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
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 的线性相关性较强；$|r|<0.75$，则认为 y 与 x 的线性相关性较弱）.

附：样本相关系数 $r=\dfrac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}}$，$\sqrt{1\ 740}\approx41.7$，$\sum\limits_{i=1}^{5}(x_i-\overline{x})(y_i-\overline{y})=41$，$\sqrt{\sum\limits_{i=1}^{5}(x_i-\overline{x})^2}=\sqrt{10}$，$\sqrt{\sum\limits_{i=1}^{5}(y_i-\overline{y})^2}=\sqrt{174}$.

解：(1)由统计表中数据可得

$\overline{y}=\dfrac{7+12+13+19+24}{5}=15$.

(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 $r=\frac{41}{\sqrt{1\ 740}}\approx\frac{41}{41.7}\approx0.98>0.75$，

因此，两个变量具有很强的线性相关关系.

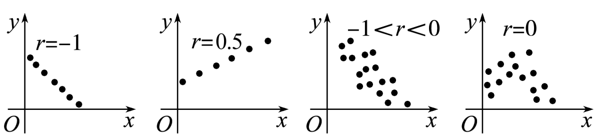
反思提炼

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.需要注意的是样本相关系数 r 的绝对值小，只是说明线性相关程度低，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.

探究训练

1. 下面的各图中，散点图与样本相关系数 r 不符合的是　　　　（　　）



A　　B　　C　　D

B　**解析**：A，B 选项中散点全部集中在一条直线上，且变量之间分别呈负相关、正相关，故样本相关系数 r 的值应分别为 $-1,1$；C 选项变量之间呈负相关，故 $-1<r<0$；D 选项变量之间没有线性相关性，样本相关系数近似为 0.

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 $|r|\geqslant0.75$，则认为 y 与 x 有很强的线性相关关系，否则认为没有很强的线性相关关系）

参考公式及数据：

样本相关系数 $r=\frac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}}$，

$\overline{y}=4.3,\sum\limits_{i=1}^{7}(y_i-\overline{y})^2=7.08$，

$\sum\limits_{i=1}^{7}(x_i-\overline{x})(y_i-\overline{y})=14,\sqrt{198.24}\approx14.08$.

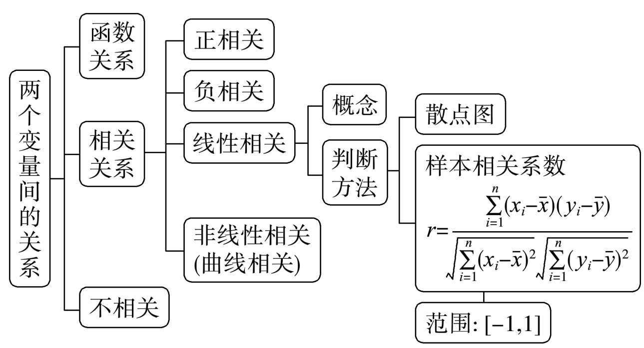
解：由题得 $\overline{x}=\frac{1+2+3+4+5+6+7}{7}=4$，

$\sum\limits_{i=1}^{7}(x_i-\overline{x})^2=9+4+1+0+1+4+9=28$，

所以 $r=\frac{\sum\limits_{i=1}^{7}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{7}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{7}(y_i-\overline{y})^2}}=\frac{14}{\sqrt{28}\times\sqrt{7.08}}\approx0.99>0.75$，

所以 y 与 x 有很强的线性相关关系.

体系构建



课后素养评价(十八)

基础性·能力运用

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具有相关关系;喜鹊发出叫声是它自身的生理反应,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 的几对观测数据统计如表,则这两个变量的关系是　(　　)

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.非线性相关

D.不具有相关性

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 $[0,1]$

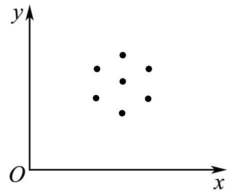
B.r 为正时,两个变量正相关;r 为负时,两个变量负相关

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 $[-1,1]$.

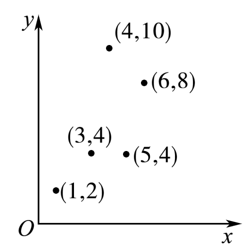
4.变量 x,y 的散点图如图所示,那么 x,y 之间的样本相关系数 r 最接近的值为　(　　)



A.1　　B.-0.5　　C.0　　D.0.5

C　**解析**:根据变量 x,y 的散点图,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,变量 x,y 的五对数据中,去掉________后,剩下的四对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.变量 X 与 Y 的一组样本数据为(10,1),(11.3,2),(11.8,3),(12.5,4),(13,5);变量 U 与 V 的一组样本数据为(10,5),(11.3,4),(11.8,3),(12.5,2),(13,1).若 r_1 表示变量 Y 与 X 之间的样本相关系数,r_2 表示变量 V 与 U 之间的样本相关系数,则 $r_1,r_2,0$ 的大小关系为________.

$r_2<0<r_1$　**解析**:对于变量 X 与 Y 而言,Y 随 X 的增大而增大,故变量 Y 与 X 正相关,即 $r_1>0$;对于变量 U 与 V 而言,V 随 U 的增大而减小,故变量 V 与 U 负相关,即 $r_2<0$.故 $r_2<0<r_1$.

7.已知两个变量 x 和 y 的 7 对数据如表所示.

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 $\overline{x}=\frac{1}{7}\times(21+23+25+27+29+32+35)\approx 27.4$,

$\overline{y}=\frac{1}{7}\times(7+11+21+24+66+115+325)\approx 81.3$,

$\sum\limits_{i=1}^{7}x_i^2=21^2+23^2+25^2+27^2+29^2+32^2+35^2=5\ 414$,

$\sum\limits_{i=1}^{7}x_iy_i=21\times7+23\times11+25\times21+27\times24+29\times66+32\times115+35\times325=18\ 542$,

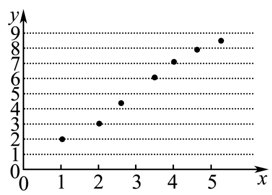
$\sum\limits_{i=1}^{7}y_i^2=7^2+11^2+21^2+24^2+66^2+115^2+325^2=124\ 393$，

所以 $r=\dfrac{\sum\limits_{i=1}^{7}x_iy_i-7\overline{x}\ \overline{y}}{\sqrt{\left(\sum\limits_{i=1}^{7}x_i^2-7\ \overline{x}^2\right)\left(\sum\limits_{i=1}^{7}y_i^2-7\ \overline{y}^2\right)}}\approx\dfrac{18\ 542-7\times27.4\times81.3}{\sqrt{(5\ 414-7\times27.4^2)\times(124\ 393-7\times81.3^2)}}\approx\dfrac{2\ 948.66}{3\ 520.92}\approx0.837\ 5$.

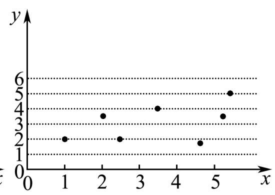
所以 x 与 y 具有较强的线性相关关系.

综合性·创新提升

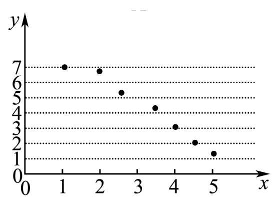
1. 已知 x，y 是两个变量，下列四个散点图中，x，y 呈负相关趋势的是　　(　　)



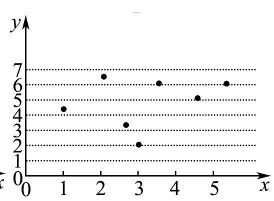
A



B



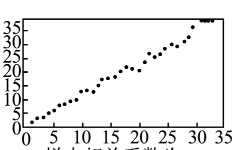
C



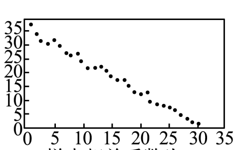
D

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，x 与 y 是正相关关系；对于 B，散点图中的点没有明显的规律，x 与 y 不具有相关关系；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，x 与 y 是负相关关系；对于 D，散点图中的点没有明显的规律，x 与 y 不具有相关关系.

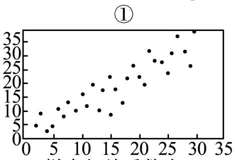
2. 某统计部门对四组数据进行统计分析后，获得如图所示的散点图，关于四组数据的样本相关系数的比较，其中正确的是　　(　　)



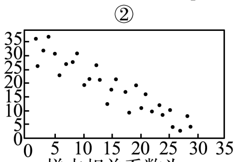
样本相关系数为r_1
①



样本相关系数为r_2
②



样本相关系数为r_3
③



样本相关系数为r_4
④

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. $|r|\leqslant1$，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，表示两个变量之间线性相关性越弱.

4. 下列说法中正确的是________.（填序号）

①样本相关系数 r 的取值范围为$[-1,1]$；

②样本相关系数 r 的绝对值越接近 0，两个变量间的线性相关程度越弱；

③样本相关系数越小，两个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 $|r|\leqslant1$，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.(2022·全国乙卷)某地经过多年的环境治理,已将荒山改造成了绿水青山.为估计一林区某种树木的总材积量,随机选取了10棵这种树木,测量每棵树的根部横截面积(单位:m^2)和材积量(单位:m^3),得到如下数据:

样本号 i	1	2	3	4	5	6	7	8	9	10
根部横截面积 x	0.04	0.06	0.04	0.08	0.08	0.05	0.05	0.07	0.07	0.06
材积量 y	0.25	0.40	0.22	0.54	0.51	0.34	0.36	0.46	0.42	0.40

并计算得$\sum\limits_{i=1}^{10}x_i=0.6$,$\sum\limits_{i=1}^{10}y_i=3.9$,$\sum\limits_{i=1}^{10}x_i^2=0.038$,$\sum\limits_{i=1}^{10}y_i^2=1.615\ 8$,$\sum\limits_{i=1}^{10}x_iy_i=0.247\ 4$.

(1)估计该林区这种树木平均一棵的根部横截面积与平均一棵的材积量.

(2)求该林区这种树木的根部横截面积与材积量的样本相关系数(精确到0.01).

(3)现测量了该林区所有这种树木的根部横截面积,并得到所有这种树木的根部横截面积总和为186 m^2.已知树木的材积量与其根部横截面积近似成正比.利用以上数据求出该林区这种树木的总材积量的估计值.

附:样本相关系数 $r=\dfrac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}}$,$\sqrt{1.896}\approx1.377$.

解:(1)样本中10棵这种树木的根部横截面积的平均数$\overline{x}=\dfrac{0.6}{10}=0.06$,

样本中10棵这种树木的材积量的平均数$\overline{y}=\dfrac{3.9}{10}=0.39$,

据此可估计该林区这种树木平均一棵的根部横截面积为0.06 m^2,平均一棵的材积量为0.39 m^3.

(2)由题中数据计算可得

$$r=\frac{\sum\limits_{i=1}^{10}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{10}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{10}(y_i-\overline{y})^2}}$$

$$=\frac{\sum\limits_{i=1}^{10}x_iy_i-10\overline{x}\,\overline{y}}{\sqrt{(\sum\limits_{i=1}^{10}x_i^2-10\overline{x}^2)(\sum\limits_{i=1}^{10}y_i^2-10\overline{y}^2)}}$$

$$=\frac{0.247\ 4-10\times0.06\times0.39}{\sqrt{(0.038-10\times0.06^2)(1.615\ 8-10\times0.39^2)}}$$

$$=\frac{0.013\ 4}{\sqrt{0.000\ 189\ 6}}\approx\frac{0.013\ 4}{0.013\ 77}\approx0.97.$$

(3)设该林区这种树木的总材积量的估计值为Y m^3,

又已知树木的材积量与其根部横截面积近似成正比,

可得$\dfrac{0.06}{0.39}=\dfrac{186}{Y}$,解得$Y=1\ 209$.

则该林区这种树木的总材积量估计为1 209 m^3.

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

(1)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.

(2)a，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　最小二乘法与经验回归方程

(1)最小二乘法

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.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.

(2)经验回归方程的参数

$$\hat{b}=\frac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}=\frac{\sum\limits_{i=1}^{n}x_iy_i-n\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{n}x_i^2-n\overline{x}^2},$$

$\hat{a}=\overline{y}-\hat{b}\overline{x}$.

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，观测值减去预测值称为残差.

概念辨析

1.判断(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)y关于x的经验回归方程$\hat{y}=\hat{b}x+\hat{a}$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

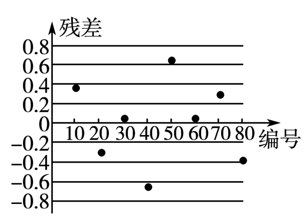
B.因变量Y是由自变量x唯一确定的

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

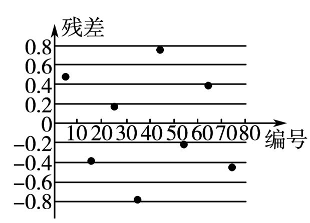
D.随机误差e是由于计算不准确造成的，可通过精确计算避免随机误差e的产生

C　**解析**：在一元线性回归模型中，表达式$Y=bx+a+e$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，因此D错误.

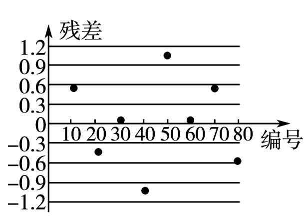
3.对变量x，y进行回归分析时，依据得到的4个不同的线性回归模型画出残差图，则下列模型拟合精度最高的是　(　　)



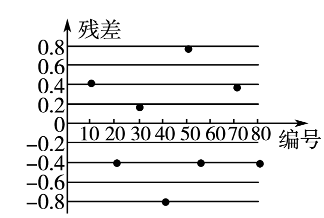
A



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判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

学习任务一　一元线性回归模型的理解

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误差项 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 σ^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 $b=0.7,a=3,|e|\leqslant 0.5$.如果今年该地区财政收入为10亿元,则年支出预计不会超过　(　　)

A.9 亿元

B.9.5 亿元

C.10 亿元

D.10.5 亿元

D　**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.

反思提炼

明确一元线性回归模型的含义是解题的关键,其中 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.

学习任务二　经验回归方程参数的意义

1.(多选)某产品的产量 x(单位:千件)与单位成本 y(单位:元)满足经验回归方程$\hat{y}=77.36-1.82x$,则以下说法中正确的是　(　　)

A.产量与单位成本正相关

B.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),若这组样本数据的样本相关系数为$-1$,则 y 关于 x 的经验回归方程可能是　(　　)

A.$\hat{y}=-\dfrac{1}{2}x+1$

B.$\hat{y}=x-1$

C.$\hat{y}=x+1$

D.$\hat{y}=-x^2$

A　**解析:**因为这组样本数据的样本相关系数为-1,所以 x 与 y 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.

反思提炼

经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$ 中，参数 $\hat{b}$ 的意义可以解释为观测值 x 每增加一个单位，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.

学习任务三　经验回归方程的求法及应用

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：

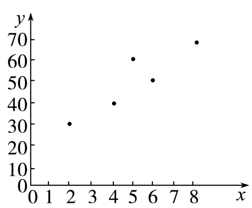
x	2	4	5	6	8
y	30	40	60	50	70

(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

$\overline{x}=\frac{25}{5}=5$，$\overline{y}=\frac{250}{5}=50$，$\sum_{i=1}^{5}x_i^2=145$，$\sum_{i=1}^{5}x_iy_i=1\ 380$.

于是可得 $\hat{b}=\frac{\sum_{i=1}^{5}x_iy_i-5\overline{x}\ \overline{y}}{\sum_{i=1}^{5}x_i^2-5\ \overline{x}^2}=\frac{1\ 380-5\times5\times50}{145-5\times25}=6.5$，$\hat{a}=\overline{y}-\hat{b}\ \overline{x}=50-6.5\times5=17.5$.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？

解：由经验回归方程计算的结果是估计值，所以在作答时结果(数据)前要有“估计”“大约”等限定词语.

反思提炼

利用经验回归方程解决实际问题的步骤

(1)根据变量 x,y 的样本数据画出散点图，从直观上判断 x 与 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 $\overline{x}$，$\overline{y}$，$\sum_{i=1}^{n}x_i^2$，$\sum_{i=1}^{n}x_iy_i$ 的值.

(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.

探究训练

1.观察变量 x,y(存在线性相关关系)的数据：

x	-10	-6.99	-5.01	-2.98	3.98	5	7.99	8.01
y	-9	-7	-5	-3	4.01	4.99	7	8

则 y 关于 x 的经验回归方程为　　(　　)

A.$\hat{y}=\frac{1}{2}x+1$　　B.$\hat{y}=x$

C.$\hat{y}=2x+\frac{1}{3}$　　D.$\hat{y}=x+1$

B　**解析**：由于经验回归直线一定经过样本点的中心$(\overline{x},\overline{y})$，由表中数据可得，$\overline{x}=0$，$\overline{y}=0$，只有 B 项中的方程对应的直线过点(0,0).故选 B.

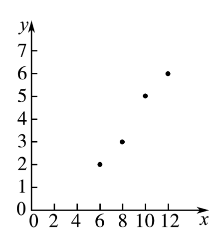
2.某研究机构对高三学生的记忆力 x 和判断力 y 进行统计分析，得下表数据.

x	6	8	10	12
y	2	3	5	6

(1)根据上表数据画出散点图，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.

(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 $\overline{x}=\frac{6+8+10+12}{4}=9$，

$\overline{y}=\frac{2+3+5+6}{4}=4$，

$\sum\limits_{i=1}^{4}x_i^2=6^2+8^2+10^2+12^2=344$，

$\sum\limits_{i=1}^{4}x_iy_i=6\times2+8\times3+10\times5+12\times6=158$，

$\hat{b}=\dfrac{158-4\times9\times4}{344-4\times9^2}=\dfrac{14}{20}=0.7$，

$\hat{a}=\overline{y}-\hat{b}\overline{x}=4-0.7\times9=-2.3$，

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.

学习任务 四　残差计算

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(‰)对海水稻亩产量 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 关于 x 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‰时该品种海水稻的亩产量；

(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，该品种海水稻的亩产量约为 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 $\sum\limits_{i=1}^{5}(y_i-\hat{y}_i)^2=(-0.02)^2+0.02^2+0.01^2+0^2+(-0.01)^2=0.001$.

反思提炼

1. 利用残差分析可以判断回归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.

探究训练

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

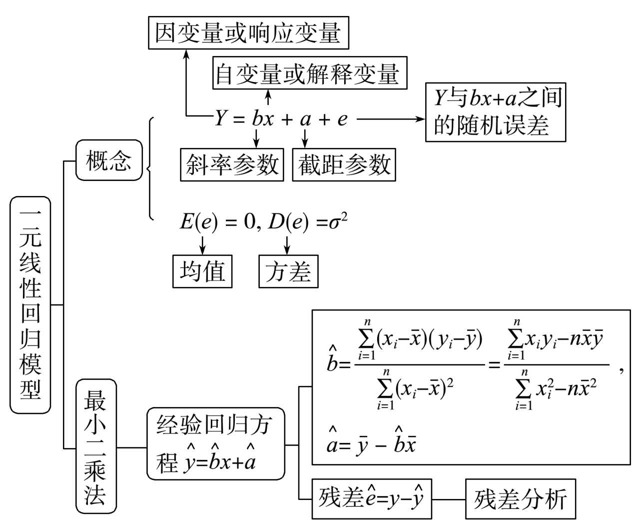
B.残差图的横坐标可以是解释变量或响应变量

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.

体系构建



课后素养评价(十九)

基础性·能力运用

1.已知某产品每吨的生产成本 y(单位:元)关于生产量 x(单位:t)的经验回归方程为 $\hat{y}=256+3x(x\geqslant 1)$，下列说法正确的是　(　　)

A.生产量每增加 1 t,每吨的生产成本增加 259 元

B.生产量每增加 1 t,每吨的生产成本增加 3 元

C.生产量每增加 1 t,每吨的生产成本平均增加 3 元

D.生产量不变,每吨的生产成本为 256 元

C　**解析**:经验回归方程的系数 $\hat{b}$ 表示 x 每增加一个单位,$\hat{y}$ 平均增加 $\hat{b}$.故当生产量增加 1 t 时,每吨的生产成本平均增加 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 $m=0.5$.

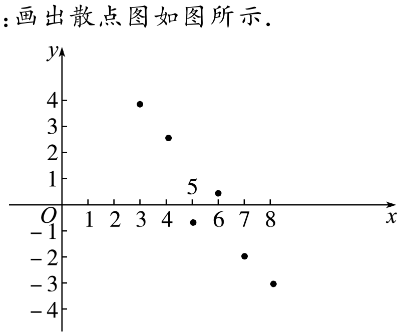
3.根据如下样本数据得到的经验回归方程为 $\hat{y}=\hat{b}x+\hat{a}$,则　(　　)

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.



由图知,$\hat{a}>0,\hat{b}<0$.

4.已知 x 与 y 之间的一组数据如下表:

x	0	1	2	3
y	1	3	5	7

则 y 关于 x 的经验回归直线 $\hat{y}=\hat{b}x+\hat{a}$ 必过点　(　　)

A.(2,2)　　B.(1.5,0)

C.(1,2)　　D.(1.5,4)

D　**解析**:$\overline{x}=\dfrac{0+1+2+3}{4}=1.5$，$\overline{y}=\dfrac{1+3+5+7}{4}=4$,则 y 关于 x 的经验回归直线必过样本点的中心(1.5,4).

5.根据变量 x,y 的 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 $\hat{b}=\dfrac{\sum\limits_{i=1}^{8}x_iy_i-8\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{8}x_i^2-8\overline{x}^2}\approx 2.62$,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 y 关于 x 的经验回归方程为 $\hat{y}=1.5x+0.5$,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直线斜率为 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,$\overline{y}=1.5\times 3+0.5=5$,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数据后重新求得的经验回归直线的斜率为1.2,所以 $5=1.2\times$

$3+\hat{a}$，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）之间有下表关系：

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.

综合性·创新提升

1. 某药厂为了了解某新药的销售情况，将2024年2月至6月的销售额整理如下：

月份	2	3	4	5	6
销售额/万元	19	25	35	37	42

根据表中数据，可求得每月的销售额 y 关于月份 x 的经验回归方程为　　(　　)

参考数据：$\sum\limits_{i=1}^{5}x_iy_i=690$，$\sum\limits_{i=1}^{5}x_i^2=90$.

A. $\hat{y}=5.8x+8.4$

B. $\hat{y}=8.4x+5.8$

C. $\hat{y}=6x-9$

D. $\hat{y}=4x+31.6$

A　**解析**：由题表中的数据得 $\overline{x}=\frac{2+3+4+5+6}{5}=4$，

$\overline{y}=\frac{19+25+35+37+42}{5}=31.6$，

所以 $\hat{b}=\frac{\sum\limits_{i=1}^{5}x_iy_i-5\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{5}x_i^2-5\overline{x}^2}=\frac{690-5\times4\times31.6}{90-5\times4^2}=5.8$，

$\hat{a}=31.6-5.8\times4=8.4$.

因此，y 关于 x 的经验回归方程为 $\hat{y}=5.8x+8.4$. 故选A.

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 $57-54.55=2.45$.

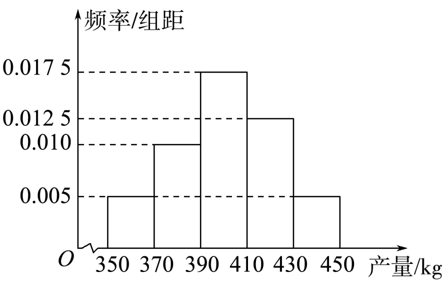
4. 某中学物理兴趣小组通过试验对其中一道竞赛题的两个物理量 u，v 进行测量，得到10对数据：(u_1,v_1)，(u_2,v_2)，…，(u_{10},v_{10})，通过散点图发现 u，v 具有较强的线性相关关系，并且利用最小二乘法求得经验回归方程为 $\hat{v}=1.5u+1$. 由于数据保存失误导致 $\sum\limits_{i=1}^{10}v_i$ 的值丢失，已知 $\sum\limits_{i=1}^{10}u_i=50$，通过所学知识可以求得 $\sum\limits_{i=1}^{10}v_i=$________.

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. 某药企计划种植A，B两种药材，已知药材A的亩产量约为300 kg，其收购价格处于上涨趋势，最近五年的收购价格如下表：

年份	2020	2021	2022	2023	2024
年份编号 x	1	2	3	4	5
单价 y/(元/kg)	18	20	23	25	29

药材 B 的收购价格始终为 20 元/kg,其亩产量的频率分布直方图如下：



(1)已知药材 A 的单价 y(单位:元/kg)与年份编号 x 之间具有线性相关关系,请求出 y 关于 x 的经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$,并估计 2026 年药材 A 的单价.

(2)利用上述频率分布直方图估计药材 B 的平均亩产量.(同一组数据用该组区间中点值为代表)

(3)若不考虑其他因素影响,为使收益最大,试判断 2026 年该药企应当种植药材 A 还是药材 B,并说明理由.

附: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)由题表中数据可得 $\overline{x}=\dfrac{1+2+3+4+5}{5}=3$,

$\overline{y}=\dfrac{18+20+23+25+29}{5}=23$,

$\sum\limits_{i=1}^{5}x_iy_i=1\times18+2\times20+3\times23+4\times25+5\times29=372$,$\sum\limits_{i=1}^{5}x_i^2=1^2+2^2+3^2+4^2+5^2=55$,

则 $\hat{b}=\dfrac{\sum\limits_{i=1}^{5}x_iy_i-5\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{5}x_i^2-5\overline{x}^2}=\dfrac{372-5\times3\times23}{55-5\times3^2}=2.7$,

$\hat{a}=\overline{y}-\hat{b}\,\overline{x}=14.9$,

故经验回归方程为 $\hat{y}=2.7x+14.9$.

当 $x=7$ 时,$\hat{y}=33.8$,

即 2026 年药材 A 的单价预计为 33.8 元/kg.

(2)由题图可得组距为 20,自左向右各组的频率依次为 0.1,0.2,0.35,0.25,0.1,

故药材 B 的平均亩产量为 $360\times0.1+380\times0.2+400\times0.35+420\times0.25+440\times0.1=401$(kg).

(3)预计 2026 年药材 A 每亩产值为 $300\times33.8=10\ 140$(元),

药材 B 每亩产值为 $401\times20=8\ 020$(元).

因为 $10\ 140>8\ 020$,所以药材 A 的每亩产值更高,应该种植药材 A.

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)直接比较两个模型的残差；

(2)直接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)模型 $y=ax^b$ 可以通过 $c=\ln a$,$v=\ln x$,$u=\ln y$ 进行变换,转化为线性回归模型 $\hat{u}=\hat{c}+\hat{b}v$.

(2)模型 $y=ae^{bx}$ 可以通过 $c=\ln a$,$u=\ln y$ 进行变换,转化为线性回归模型 $\hat{u}=\hat{c}+\hat{b}x$.

(3)模型 $y=a+b\ln x$ 可以通过 $v=\ln x$,$u=y$ 进行变换,转化为线性回归模型 $\underline{\hat{u}=\hat{a}+\hat{b}v}$.

概念辨析

1.在对两个变量 y 与 x 进行回归分析时,分别选择了四个不同的模型,且它们的 R^2 的值的大小关系为 $R^2_{模型3}<R^2_{模型4}<R^2_{模型1}<R^2_{模型2}$,则拟合效果最好的是（　　）

A.模型 1　　B.模型 2

C.模型 3　　D.模型 4

B　**解析**:R^2 越大,则残差平方和越小,表示拟合效果越好.所以拟合效果最好的是模型 2.故选 B.

2.以模型 $y=ce^{kx}(c>0)$ 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.$-2,e$　　B.$2,\frac{1}{e}$

C.$-2,\frac{1}{e}$　　D.$2,e$

B　**解析**:由题意得 $\ln y=\ln(ce^{kx})=\ln c+kx$.设 $z=\ln y$,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 $k=2,c=\frac{1}{e}$.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,$\sum_{i=1}^{n}(y_i-\overline{y})^2$ 与经验回归方程无关,残差平方和 $\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$ 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

学习任务一　线性回归分析

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.

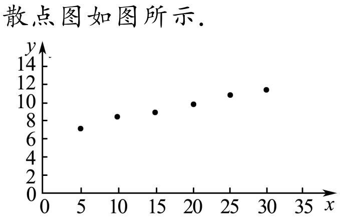
x	5	10	15	20	25	30
y	7.25	8.12	8.95	9.9	10.9	11.8

(1)作出散点图并求 y 关于 x 的经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$;($\hat{a},\hat{b}$ 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:

x	5	10	15	20	25	30
$y-\hat{y}$	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,弹簧长度与悬挂物体的质量具有线性关系.

[**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.

反思提炼

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$\sum\limits_{i=1}^{n}(y_i-\hat{y}_i)^2$越小，模型的拟合效果越好.

(3)决定系数法：$R^2=1-\dfrac{\sum\limits_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}$越大，表示模型的拟合效果越好.

探究训练

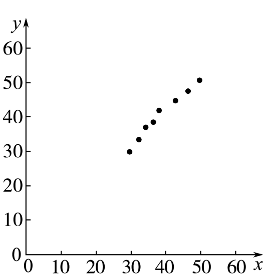
1.某运动员训练次数x与成绩y的数据如表：

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.

(2)作出残差图，计算R^2.

解：(1)作出该运动员训练次数x与成绩y的散点图，如图所示.由散点图可知，它们之间具有线性相关关系.



由题表中数据可得$\overline{x}=39.25$，$\overline{y}=40.875$，$\sum\limits_{i=1}^{8}x_i^2=12\ 656$，$\sum\limits_{i=1}^{8}x_iy_i=13\ 180$，

所以$\hat{b}=\dfrac{\sum\limits_{i=1}^{8}x_iy_i-8\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{8}x_i^2-8\overline{x}^2}\approx1.041\ 5$，

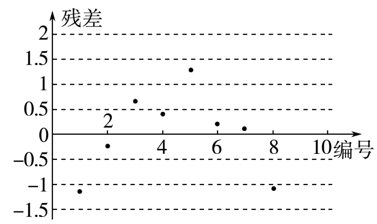
$\hat{a}=\overline{y}-\hat{b}\,\overline{x}\approx-0.003\ 9$.

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训练次数x和成绩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.

计算可得$R^2\approx0.985\ 5$.

2.已知x与y的有关数据如下：

x	2	3	4	5	6
y	2.2	3.8	5.5	6.5	7

(1)求y关于x的经验回归方程$\hat{y}=\hat{b}x+\hat{a}$；

(2)完成下面的残差表并判断(1)中经验回归方程的拟合效果是否良好.(若$R^2>0.9$，则认为拟合效果良好)

x	2	3	4	5	6
$y-\hat{y}$					

解：(1)由题表可知$\overline{x}=4$，$\overline{y}=5$，$\sum\limits_{i=1}^{5}x_i^2=90$，

$\sum\limits_{i=1}^{5}x_iy_i=112.3$，

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$\hat{e}_i=y_i-\hat{y}_i$，

所以$\hat{e}_1=-0.34$，$\hat{e}_2=0.03$，$\hat{e}_3=0.5$，$\hat{e}_4=0.27$，$\hat{e}_5=-0.46$，

则残差表为

x	2	3	4	5	6
$y-\hat{y}$	-0.34	0.03	0.5	0.27	-0.46

因为$\sum\limits_{i=1}^{5}(y_i-\overline{y})^2=(2.2-5)^2+(3.8-5)^2+(5.5-5)^2+(6.5-5)^2+(7-5)^2=15.78$，$\sum\limits_{i=1}^{5}(y_i-\hat{y}_i)^2=0.651$，

所以$R^2=1-\dfrac{0.651}{15.78}\approx0.96>0.9$.

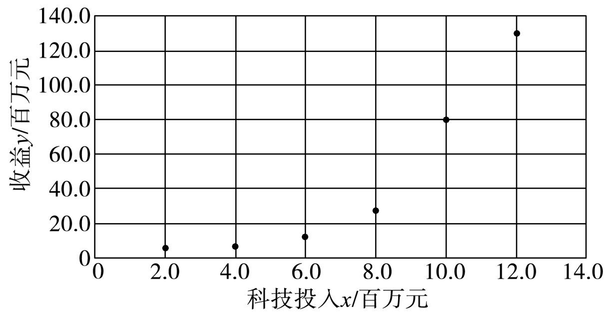
所以该经验回归方程的拟合效果良好.

学习任务二　非线性回归分析

例 2　某企业为了提升行业竞争力，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，如下表：

$\overline{y}$	$\overline{z}$	$\sum\limits_{i=1}^{6}(x_i-\overline{x})\cdot(y_i-\overline{y})$	$\sum\limits_{i=1}^{6}(x_i-\overline{x})\cdot(z_i-\overline{z})$	$\sum\limits_{i=1}^{6}(y_i-\overline{y})^2$	$\sum\limits_{i=1}^{6}(x_i-\overline{x})^2$
43.5	4.5	854	34.7	12 730.4	70

其中 $z_i=\log_2 y_i$，$\overline{z}=\frac{1}{6}\sum\limits_{i=1}^{6}z_i$．

(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？（$\log_2 5\approx 2.3$）

(2)乙认为样本点分布在二次曲线 $y=mx^2+n$ 的附近，并计算得经验回归方程为 $\hat{y}=0.92x^2-12$，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)，其经验回归直线 $\hat{v}=\hat{\alpha}+\hat{\beta}u$ 的斜率和截距的最小二乘估计分别为

$$\hat{\beta}=\frac{\sum\limits_{i=1}^{n}(u_i-\overline{u})(v_i-\overline{v})}{\sum\limits_{i=1}^{n}(u_i-\overline{u})^2},\ \hat{\alpha}=\overline{v}-\hat{\beta}\,\overline{u}.$$

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 $0.5x+1\geqslant\log_2 200$，解得 $x\geqslant 4+4\log_2 5\approx 13.2$．

故科技投入的费用至少为 13.2 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：

x	2	4	6	8	10	12
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．

反思提炼

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，即借助于线性回归模型研究呈非线性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个直线形带状区域，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个曲线形带状区域，要先对变量作适当的变换，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：

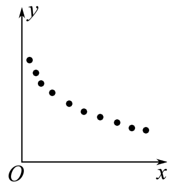
①根据原始数据作出散点图；

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．

探究训练

某公司对某产品做市场调查，获得了该产品的定价 x（单位：万元/t）和每天的销量 y（单位：t）的一组数据，根据这组数据制作了如下散点图和统计表．



$\overline{x}$	$\overline{y}$	$\overline{t}$	$\sum_{i=1}^{10}x_i^2$	$\sum_{i=1}^{10}t_i^2$	$\sum_{i=1}^{10}x_iy_i$	$\sum_{i=1}^{10}t_iy_i$
0.33	10	3	0.164	100	68	350

表中 $t=\frac{1}{x}$.

(1)根据散点图判断 $\hat{y}=\hat{b}x+\hat{a}$ 与 $\hat{y}=\hat{c}x^{-1}+\hat{d}$ 哪一个更适合作为 y 关于 x 的经验回归方程;(给出判断即可,不必说明理由)

(2)根据(1)的判断结果,建立 y 关于 x 的经验回归方程;

(3)若生产 1 t 该产品的成本为 0.25 万元,依据(2)的经验回归方程,定价为多少时,该产品每天的销售利润最大？最大利润是多少？

解:(1)根据散点图可知,$\hat{y}=\hat{c}x^{-1}+\hat{d}$ 更适合作为 y 关于 x 的经验回归方程.

(2)由 $t=\frac{1}{x}$,得 $\hat{y}=\hat{c}t+\hat{d}$,

所以 $\hat{c}=\frac{\sum_{i=1}^{10}t_iy_i-10\,\overline{t}\,\overline{y}}{\sum_{i=1}^{10}t_i^2-10\overline{t}^2}=\frac{350-10\times3\times10}{100-10\times3^2}=5$,

所以 $\hat{d}=\overline{y}-\hat{c}\,\overline{t}=10-5\times3=-5$,

所以 y 关于 t 的经验回归方程为 $\hat{y}=5t-5$,

故 y 关于 x 的经验回归方程为 $\hat{y}=\frac{5}{x}-5$.

(3)设每天的利润为 W 万元,则 $W=y(x-0.25)=\left(\frac{5}{x}-5\right)(x-0.25)=6.25-5\left(x+\frac{0.25}{x}\right)\leqslant6.25-5\times2\sqrt{x\times\frac{0.25}{x}}=6.25-5\times2\times0.5=1.25$,

当且仅当 $x=\frac{0.25}{x}$,即 $x=0.5$ 时等号成立,

所以预计定价为 0.5 万元/t 时,该产品每天的销售利润最大,最大利润是 1.25 万元.

学习任务 三　回归分析的综合应用

例 3　某商场销售一批进价是 30 元/台的小商品,在试销中发现,此商品的销售单价 x(x 取整数,单位:元)与日销售量 y(单位:台)之间有如下关系:

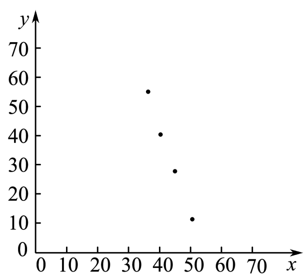
x	35	40	45	50
y	56	41	28	11

(1)画出散点图,并判断 y 与 x 之间是否具有线性相关关系;

(2)求日销售量 y 关于销售单价 x 的经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$;($\hat{b}$ 的值保留一位有效数字)

(3)设经营此商品的日销售利润为 P(单位:元),根据(2)写出 P 关于 x 的函数解析式,并预测当销售单价为多少元时,才能获得最大日销售利润.

解:(1)散点图如图所示,从图中可以看出这些点大致分布在一条直线附近,因此两个变量具有线性相关关系.



(2)因为 $\overline{x}=\frac{1}{4}\times(35+40+45+50)=42.5$,

$\overline{y}=\frac{1}{4}\times(56+41+28+11)=34$,

$\sum_{i=1}^{4}x_iy_i=35\times56+40\times41+45\times28+50\times11=5\ 410$,

$\sum_{i=1}^{4}x_i^2=35^2+40^2+45^2+50^2=7\ 350$,

所以 $\hat{b}=\frac{\sum_{i=1}^{4}x_iy_i-4\,\overline{x}\,\overline{y}}{\sum_{i=1}^{4}x_i^2-4\,\overline{x}^2}=\frac{5\ 410-4\times42.5\times34}{7\ 350-4\times42.5^2}=\frac{-370}{125}\approx-3$.

所以 $\hat{a}=\overline{y}-\hat{b}\,\overline{x}=34-(-3)\times42.5=161.5$.

所以 $\hat{y}=-3x+161.5$.

(3)依题意有 $P=(-3x+161.5)(x-30)=-3x^2+251.5x-4\ 845$,

所以当 $x=\frac{251.5}{6}\approx42$ 时,P 有最大值,约为 426.

所以当销售单价为 42 元时，能获得最大日销售利润.

反思提炼

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.

探究训练

党的二十大报告指出，坚持把发展经济的着力点放在实体经济上，推进新型工业化，加快建设制造强国、质量强国、航天强国、交通强国、网络强国、数字中国.某企业为响应国家号召，汇聚科研力量，加强科技创新，准备增加研发资金的投入.现该企业为了了解年研发资金投入额 x（单位：亿元）对年盈利额 y（单位：亿元）的影响，研究了近 10 年年研发资金投入额和年盈利额的数据，通过对比分析，建立了两个函数模型：①$y=\alpha+\beta x^2$，②$y=e^{\lambda x+t}$，其中 α,β,λ,t 均为常数，e 为自然对数的底数.令 $u_i=x_i^2$，$v_i=\ln y_i$，$i=1,2,\cdots,10$，经计算得到如下数据：

$\overline{x}$	$\overline{y}$	$\sum\limits_{i=1}^{10}(x_i-\overline{x})^2$	$\sum\limits_{i=1}^{10}(y_i-\overline{y})^2$	$\overline{u}$	$\overline{v}$
26	215	65	2	680	5.36

$\sum\limits_{i=1}^{10}(u_i-\overline{u})^2$	$\sum\limits_{i=1}^{10}(u_i-\overline{u})\cdot(y_i-\overline{y})$	$\sum\limits_{i=1}^{10}(v_i-\overline{v})^2$	$\sum\limits_{i=1}^{10}(x_i-\overline{x})\cdot(v_i-\overline{v})$
11 250	130	2.6	12

(1)请从样本相关系数的角度，分析哪一个模型拟合效果更好.

(2)①根据(1)的选择及表中数据，建立 y 关于 x 的经验回归方程(相关参数的值精确到 0.01)；

②若希望明年盈利额为 250 亿元，请预测明年的研发资金投入额为多少亿元(结果精确到0.01).

附：样本相关系数：

$$r=\frac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}}.$$

$\ln 2\approx0.693$，$\ln 5\approx1.609$.

解：(1)设 u 和 y 的样本相关系数为 r_1，x 和 v 的样本相关系数为 r_2，由题意，

$$r_1=\frac{\sum\limits_{i=1}^{10}(u_i-\overline{u})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{10}(u_i-\overline{u})^2}\sqrt{\sum\limits_{i=1}^{10}(y_i-\overline{y})^2}}=\frac{130}{\sqrt{11\ 250\times2}}=\frac{13}{15}\approx0.87,$$

$$r_2=\frac{\sum\limits_{i=1}^{10}(x_i-\overline{x})(v_i-\overline{v})}{\sqrt{\sum\limits_{i=1}^{10}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{10}(v_i-\overline{v})^2}}=\frac{12}{\sqrt{65\times2.6}}=\frac{12}{13}\approx0.92,$$

则 $|r_1|<|r_2|$，因此从样本相关系数的角度，模型 $y=e^{\lambda x+t}$ 的拟合效果更好.

(2)①先建立 v 关于 x 的经验回归方程，

由 $y=e^{\lambda x+t}$，得 $\ln y=t+\lambda x$，即 $v=t+\lambda x$.

$$\hat{\lambda}=\frac{\sum\limits_{i=1}^{10}(x_i-\overline{x})(v_i-\overline{v})}{\sum\limits_{i=1}^{10}(x_i-\overline{x})^2}=\frac{12}{65}\approx0.18,$$

$$\hat{t}=\overline{v}-\hat{\lambda}\,\overline{x}=5.36-\frac{12}{65}\times26=0.56,$$

所以 v 关于 x 的经验回归方程为 $\hat{v}=0.18x+0.56$，

所以 $\ln\hat{y}=0.18x+0.56$，则 $\hat{y}=e^{0.18x+0.56}$.

②若明年盈利额为 250 亿元，

则 $250=e^{0.18x+0.56}$，即 $0.18x+0.56=\ln 250$.

因为 $\ln 250=3\ln 5+\ln 2\approx3\times1.609+0.693=5.52$，

所以 $x\approx\dfrac{5.52-0.56}{0.18}\approx27.56$.

所以预测明年的研发资金投入额约为 27.56 亿元.

体系构建

一元线性回归模型的应用
- 决定系数$R^2=1-\dfrac{\sum\limits_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}$ — 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 — 作适当的变换，转化为线性回归模型

课后素养评价(二十)

基础性·能力运用

1. 根据给定的样本点建立 A 和 B 两个回归模型，若 A，B 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）

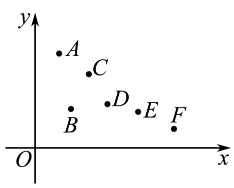
A.若 $a_1<a_2$，则 $b_1<b_2$，A 的拟合效果更好

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越强，

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. 关于 x 与 y，有如下数据：

x	2	4	5	6	8
y	30	40	60	50	70

给出如下的两个回归模型：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的拟合效果更好，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 $Q_1<Q_2$.

4. 在研究两个变量 x，y 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 $\hat{y}=e^{0.25x-2.58}$.

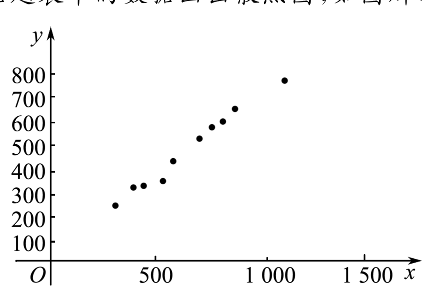
故 y 关于 x 的经验回归方程为 $\hat{y}=e^{0.25x-2.58}$.

5. 某公司为研究员工日均支出与日均收入的相关关系，随机抽取 10 名员工进行调查，其结果如下：

编号	日均收入 x/元	日均支出 y/元
1	300	255
2	390	324
3	420	335
4	520	360
5	570	450
6	700	520
7	760	580
8	800	600
9	850	630
10	1 080	750

试预测该公司日人均收入为 1 100 元和日均收入为 1 200元的两名员工的日均支出.

解：根据题表中的数据画出散点图，如图所示.



由图可知，日均支出与日均收入之间具有线性相关关系.

通过计算可知 $\overline{x}=639$，$\overline{y}=480.4$，

由公式计算得 $\hat{b}\approx0.659\ 9$，$\hat{a}=\overline{y}-\hat{b}\ \overline{x}\approx58.723\ 9$，故 y 关于 x 的经验回归方程为 $\hat{y}=0.659\ 9x+58.723\ 9$.

将 $x=1\ 100$ 代入经验回归方程得 $\hat{y}=784.613\ 9\approx784.61$；将 $x=1\ 200$ 代入经验回归方程得 $\hat{y}=850.603\ 9\approx850.60$.

故预测日均收入分别为 1 100 元和 1 200 元的两名员工的日均支出分别为 784.61 元和 850.60元.

综合性·创新提升

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.可能大也可能小

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 $x=5$，则预测 y 的值可能为　（　　）

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. 已知对于某组数据采用了四个不同的回归模型进行回归分析，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.0.31　　D.0.17

A　**解析**：两个变量的回归模型中，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 是最大的值，所以选项 A 对应的回归模型拟合效果最好.故选 A.

4. 非线性经验回归方程 $\hat{y}=ae^{bx}$进行变换后得到的线性经验回归方程为 $\hat{u}=1-0.6x$，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**：对于 $y=ae^{bx}$，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 $\hat{y}=ae^{bx}$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 $x=\dfrac{3}{10}$，所以函数 $y=x^2+bx+a$ 的单调递增区间为 $\left(\dfrac{3}{10},+\infty\right)$.故选 B.

5. 近年来，云南省保山市龙陵县紧紧围绕打造“中国石斛之乡”的发展定位，大力发展石斛产业.在政府的大力扶持下，龙陵紫皮石斛产量逐年增长，根据 2017 年到 2022 年龙陵县紫皮石斛产量得到如下统计表.

年份	2017	2018	2019	2020	2021	2022
年份代码 x	1	2	3	4	5	6
紫皮石斛产量 y/t	3 200	3 400	3 600	4 200	7 500	9 000

(1)判断 $y=ax+b$ 与 $y=ce^{dx}$(a,b,c,d 均为常数)哪一个更适合作为龙陵县紫皮石斛产量 y 关于年份代码 x 的经验回归方程类型.(给出判断即可，不必说明理由)

(2)经计算得下表中数据，根据(1)中结论，求出 y 关于 x 的经验回归方程，其中 $u=\ln y$，$u_i=\ln y_i$($i=$

1,2,3,4,5,6).

$\overline{x}$	$\overline{y}$	$\overline{u}$	$\sum\limits_{i=1}^{6}(x_i-\overline{x})^2$	$\sum\limits_{i=1}^{6}(x_i-\overline{x})\cdot(y_i-\overline{y})$	$\sum\limits_{i=1}^{6}(x_i-\overline{x})\cdot(u_i-\overline{u})$
3.5	5 150	8.46	17.5	20 950	3.85

(3)龙陵县计划 2025 年实现紫皮石斛年产量达 15 000 t,根据(2)中所求得的经验回归方程,预测该目标是否能完成.

附：$e^{9.45}\approx 12\ 708$，$e^{9.67}\approx 15\ 835$，$\hat{b}=\dfrac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}$,$\hat{a}=\overline{y}-\hat{b}\,\overline{x}$.

解:(1)$y=ce^{dx}$更适合作为龙陵县紫皮石斛产量 y 关于年份代码 x 的经验回归方程类型.

(2)对 $y=ce^{dx}$ 两边取自然对数,得 $\ln y=\ln c+dx$.

令 $u=\ln y$,$v=\ln c$,则 $u=v+dx$.

因为 $\overline{x}=3.5$,$\overline{u}=8.46$,$\sum\limits_{i=1}^{6}(x_i-\overline{x})^2=17.5$,

$\sum\limits_{i=1}^{6}(x_i-\overline{x})(u_i-\overline{u})=3.85$,

所以 $\hat{d}=\dfrac{\sum\limits_{i=1}^{6}(x_i-\overline{x})(u_i-\overline{u})}{\sum\limits_{i=1}^{6}(x_i-\overline{x})^2}=\dfrac{3.85}{17.5}=0.22$,

所以 $\hat{v}=\overline{u}-\hat{d}\,\overline{x}=8.46-0.22\times3.5=7.69$,

所以 $\hat{u}=0.22x+7.69$.

所以龙陵县紫皮石斛产量 y 关于年份代码 x 的经验回归方程为 $\hat{y}=e^{0.22x+7.69}$.

(3)当 $x=9$ 时,$\hat{y}=e^{0.22\times9+7.69}=e^{9.67}\approx 15\ 835>15\ 000$,

故预测该目标可以完成.

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,它们的取值分别为 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.已知一个如下的 2×2 列联表,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.

2.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

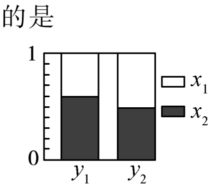
B.残差图和列联表

C.散点图和等高堆积条形图

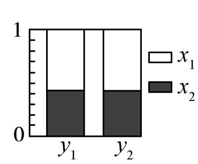
D.等高堆积条形图和列联表

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.

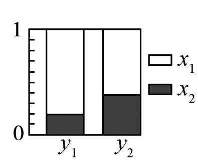
3.观察下列各图,其中两个分类变量 X,Y 之间关系最强的是　　(　　)



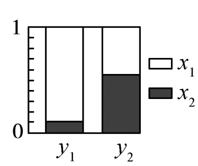
A



B



C



D

D　**解析**:在四幅图中,选项 D 的图中两个深色条的高度差最明显,说明两个分类变量之间的关系最强.

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，就判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

学习任务一　2×2列联表及其应用

例1　某大学一学院甲、乙两个专业的报考和录取情况如表：

性别	甲专业报考人数	乙专业报考人数	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2×2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 $\frac{25+90}{100+300}=28.75\%$，

乙专业的录取率为 $\frac{180+50}{400+100}=46\%$，

所以乙专业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 $\frac{90+50}{300+100}=35\%$.

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 $\frac{25+180}{100+400}=41\%$，

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．

反思提炼

(1)列2×2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2×2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2×2列联表，然后根据频率特征，即将 $\frac{a}{a+b}$ 与 $\frac{c}{c+d}$ 的值相比较，直接判断出两个分类变量间是否相互影响，但方法较粗劣．

探究训练

在下列关于吸烟与肺癌的2×2列联表中，d 的值为　　(　　)

单位：人

吸烟	肺癌		合计
	非肺癌患者	肺癌患者	
非吸烟者	7 775	42	7 817
吸烟者		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 $d=91-42=49$.

学习任务二　利用等高堆积条形图判断两个分类变量是否相关

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 2×2 列联表和等高堆积条形图;

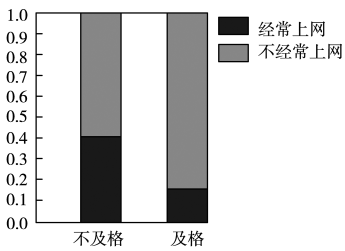
(2)利用(1)中的等高堆积条形图,判断学生经常上网对学习成绩的好坏是否有影响.

解:(1)根据题中所给的数据得到如下 2×2 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

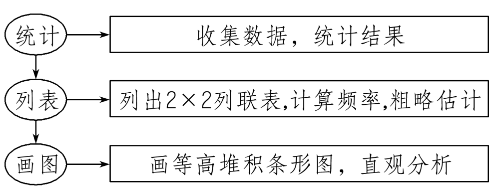
得出等高堆积条形图如图所示.



(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.

反思提炼

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,进而反映它们之间是否有关联.

探究训练

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　　(　　)

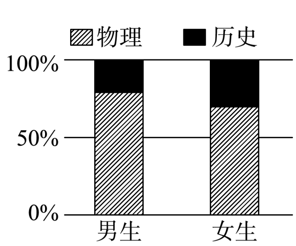


图 1

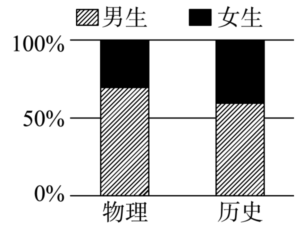


图 2

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;

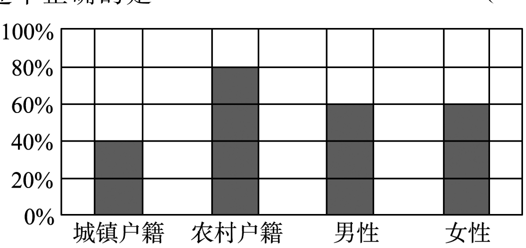
根据题图 2 可知样本中男生人数多于女生人数,故 D 错误;

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)为了了解户籍与性别对生育二胎选择倾向的影响,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二胎与倾向选择不生育二胎的人数比例图(如图所示),其中阴影部分表示倾向选择生育二胎,则下列叙述中正确的是　　(　　)



A.是否倾向选择生育二胎与户籍有关

B.是否倾向选择生育二胎与性别有关

C.倾向选择生育二胎的人员中,男性人数与女性人数相同

D.倾向选择不生育二胎的人员中,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二胎的比例为 40%,农村户籍倾向选择生育二胎的比例为 80%,两者相差较大,所以是否倾向选择生育二胎

与户籍有关，故 A 正确.对于 B，男性倾向选择生育二胎的比例为 60%，女性倾向选择生育二胎的比例为 60%，所以是否倾向选择生育二胎与性别无关，故 B 错误.对于 C，男性倾向选择生育二胎的比例为 60%，人数为 $60\times60\%=36$；女性倾向选择生育二胎的比例为 60%，人数为 $40\times60\%=24$.所以倾向选择生育二胎的人员中，男性人数比女性人数多，故 C 错误.对于 D，倾向选择不生育二胎的人员中，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二胎的人员中，农村户籍人数少于城镇户籍人数，故 D 正确.

体系构建

- 分类变量与列联表
  - 分类变量的意义
  - 列联表
    - 绘制
    - 应用：通过频率大小粗略反映两个分类变量间是否有关联
  - 等高堆积条形图
    - 绘制
    - 应用：通过图形直观反映两个分类变量间是否有关联

课后素养评价(二十一)

基础性·能力运用

1.下面是 X 与 Y 的 2×2 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 2×2 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$\dfrac{c}{c+d}$

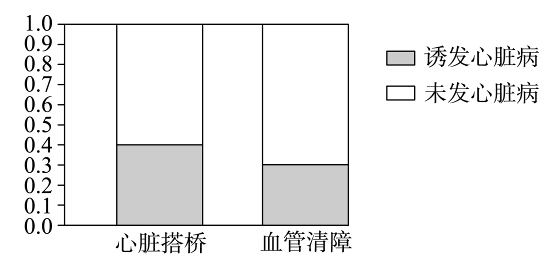
B.$\dfrac{a}{c+d}$与$\dfrac{c}{a+b}$

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，$|ad-bc|$的值越大，两个分类变量有关系的可能性就越大.故选 A.

3.下面的等高堆积条形图可以说明的问题是 ( )



A.“心脏搭桥”手术和“血管清障”手术对“诱发心脏病”的影响是绝对不同的

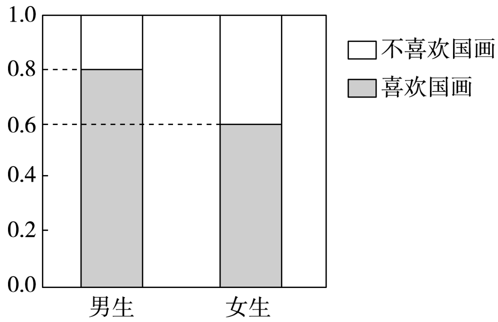
B.“心脏搭桥”手术和“血管清障”手术对“诱发心脏病”的影响没有什么不同

C.无法判断两种手术对“诱发心脏病”的影响有没有不同的地方

D.“心脏搭桥”手术和“血管清障”手术对“诱发心脏病”的影响在某种程度上是不同的，但是没有100%的把握

D **解析**：由题图可知“心脏搭桥”手术和“血管清障”手术对“诱发心脏病”的频率不同，所以“心脏搭桥”手术和“血管清障”手术对“诱发心脏病”的影响在某种程度上是不同的，但是没有 100%的把握，所以选项 D 正确.故选 D.

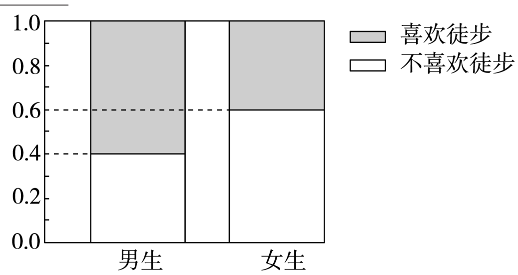
4.某艺术馆为了研究学生性别和是否喜欢国画之间的联系，随机抽取 80 名学生(其中男生 50 名，女生 30 名)进行调查，并绘制等高堆积条形图如图所示，则这 80 名学生中喜欢国画的人数为 ( )



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 $50\times80\%+30\times60\%=58$.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.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,则抽取的男生人数约为________.



15　**解析**:根据题图可知,喜欢徒步的男生人数为 $0.6\times500=300$,喜欢徒步的女生人数为 $0.4\times400=160$,

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,则抽取的男生人数约为 $\frac{300}{460}\times23=15$.

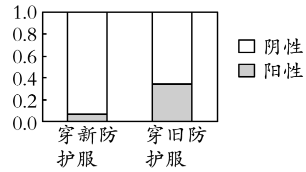
6. 当某矿石粉厂生产一种矿石粉时,在数天内就有部分工人患职业性皮肤炎.在生产期间,随机抽取车间工人抽血化验,75 名穿新防护服的车间工人中有 5 例阳性、70 例阴性,28 名穿旧防护服的车间工人中有 10 例阳性、18 例阴性,请用等高堆积条形图判断这种新防护服对预防工人职业性皮肤炎是否有效.(注:显阴性即未患皮肤炎)

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这种皮肤炎有效.

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,作出考前心情与性格的列联表,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,利用图形判断性格对考前心情是否有影响.

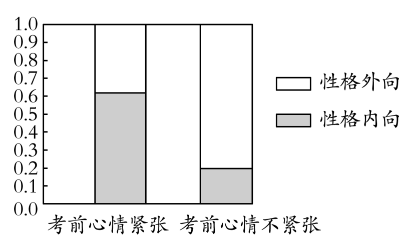
解:(1)作列联表如下:

单位:人

考前心情	性格		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 $\frac{213}{594}=\frac{71}{198}$.

(2)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,考前心情紧张的学生中性格内向的学生占的比例比考前心情不紧张的学生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.

综合性·创新提升

1. 分类变量 X 和 Y 的列联表如下：

X	Y		合计
	$Y=y_1$	$Y=y_2$	
$X=x_1$	200	800	1 000
$X=x_2$	180	a	$180+a$
合计	380	$800+a$	$1\ 180+a$

若两个分类变量之间无关联，则 a 的值可能是（　　）

A.200　B.720　C.100　D.180

B　**解析**：由于 X 和 Y 没有关联，则根据列联表，可知 $\frac{200}{1\ 000}$ 和 $\frac{180}{180+a}$ 基本相等，检验可知，B 满足条件.

2. 下表是一个 2×2 列联表，则表中 $m=$______，$n=$______.

X	Y		合计
	y_1	y_2	
x_1	a	35	45
x_2	7	b	n
合计	m	73	s

17　45　**解析**：由题意，可知 $a+35=45$，解得 $a=10$.

则 $m=a+7=10+7=17$.

又由 $35+b=73$，解得 $b=38$.则 $n=7+38=45$.

3. 某电视台在一次对收看文艺节目和新闻节目观众的抽样调查中，随机抽取了 100 名电视观众，相关的数据如表所示：

单位：名

年龄	电视节目		合计
	文艺节目	新闻节目	
20 至 40 岁	40	18	58
大于 40 岁	15	27	42
合计	55	45	100

由表中数据直观分析，观众收看新闻节目与年龄______关.(填“有”或“无”)

有　**解析**：因为在 20 至 40 岁的 58 名观众中收看新闻节目的频率为 $\frac{18}{58}$，而大于 40 岁的 42 名观众中收看新闻节目的频率为 $\frac{27}{42}$，两者相差较大，所以经直观分析，观众收看新闻节目与年龄是有关的.

4. 在对人们的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 2×2 列联表；

(2)由列联表判断性别对休闲方式是否有影响.

解：(1)2×2 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)中列联表的数据，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，所以性别对休闲方式有影响.

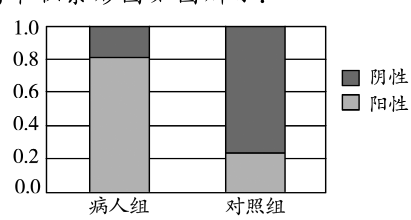
5. 为了解铅的毒素对人的尿棕色素呈阴性或阳性是否有影响，分别对病人组和对照组的尿液作尿棕色素定性检查，结果如下：

单位：人

组别	尿棕色素		合计
	阳性	阴性	
病人组	29	7	36
对照组	9	28	37
合计	38	35	73

试根据列联表画出等高堆积条形图，分析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病人组和对照组中尿棕色素为阳性的频率.由图可以直观地看出，病人组与对照组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 χ^2 的意义.

2. 通过实例,了解 χ^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 χ^2 的取值推断分类变量 X 和 Y 是否独立的方法称为 χ^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 2×2 列联表,计算 χ^2 的值,并与临界值 x_α 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. 下列关于 χ^2 的说法中正确的是　(　　)

A.χ^2 越大,"两个变量有关联"的可信度越小

B.χ^2 越大,"两个变量无关"的可信度越大

C.χ^2 越小,"两个变量有关联"的可信度越小

D.χ^2 越小,"两个变量无关"的可信度越小

C　**解析**:χ^2 越大,"两个变量有关联"的可信度越大,"两个变量无关"的可信度越小;相反,χ^2 越小,"两个变量有关联"的可信度越小,"两个变量无关"的可信度越大.

2. 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 有关系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.$\chi^2=5.014$

3. 请思考并回答下列问题:

如何理解临界值?

提示:(1)小概率值 α 的临界值 x_α 是一个正实数.

(2)$P(\chi^2\geqslant x_\alpha)=\alpha$ 时,x_α 称为 α 的临界值,即对于不同的小概率值 α,有不同的临界值 x_α.

(3)基于小概率值 α 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 α;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.

任务型课堂

学习任务一　独立性检验的基本思想

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)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)$P_1=\frac{200}{350}=\frac{4}{7}$，$P_2=\frac{400}{1\ 000}=\frac{2}{5}$.

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$\alpha=0.001$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：

α	0.1	0.05	0.025	0.01	0.005	0.001
x_α	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)中列联表可知$\chi^2=\frac{1\ 350\times(200\times600-150\times400)^2}{350\times1\ 000\times600\times750}=\frac{216}{7}\approx30.86>10.828=x_{0.001}$，

所以依据小概率值$\alpha=0.001$的独立性检验，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，哪种方法更科学合理些？

解：根据χ^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.

反思提炼

用独立性检验解决实际问题的基本步骤

(1)认真读题，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χ^2的值，并与临界值x_α比较.

(4)根据检验规则得出结论：如果$\chi^2\geqslant x_\alpha$，则“X与Y有关(或不独立)”；如果$\chi^2<x_\alpha$，则“X与Y无关(或独立)”.

探究训练

为考察棉花种子是否经过处理跟是否生病之间的关系，进行试验后得到下表数据：

单位：粒

生病情况	处理情况		合计
	处理	未处理	
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.种子是否经过处理跟是否生病有关

B.种子是否经过处理跟是否生病无关

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$\chi^2=\frac{407\times(32\times213-101\times61)^2}{133\times274\times93\times314}\approx0.164<2.706=x_{0.1}$，即没有充分证据认为种子是否经过处理跟是否生病有关.

学习任务二 独立性检验的综合应用

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 $\hat{y}=-91+100x(1.01\leqslant x\leqslant 1.05)$. 已知合格零件的尺寸为 (1.03 ± 0.01)cm.

(1)根据已知数据完成下面的列联表：

单位：个

加工机床	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 ± 0.01)cm，

故列联表为

单位：个

加工机床	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_0：零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.

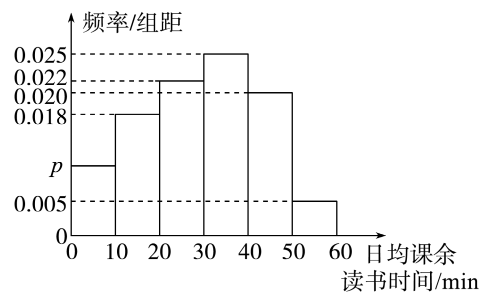
反思提炼

1. 解答独立性检验问题的关键在于正确利用 $\chi^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$ 计算 χ^2 的值，再将它与临界值 x_α 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，要熟悉解题流程，不难理解掌握.

探究训练

1. 书籍是文化的重要载体，读书是传承文化的重要方式.某地区为了解学生课余时间的读书情况，随机抽取了 n 名学生进行调查，根据调查得到的学生日均课余读书时间绘制了如图所示的频率分布直方图，将日均课余读书时间不低于 40 min 的学生称为"读书之星"，日均课余读书时间低于 40 min 的学生称为"非读书之星".已知抽取的样本中日均课余读书时间低于 10 min 的有 10 人.



(1)求 n，p 的值.

(2)根据已知条件完成下面的 2×2 列联表.

单位：人

性别	读书之星		合计
	不是	是	
男			
女		10	55
合计			

依据 $\alpha=0.05$ 的独立性检验，能否认为"读书之星"与性别有关联？

(3)将上述调查所得到的频率视为概率，现从该地区的学生中随机抽取 3 名学生，记其中"读书之星"的人数为随机变量 X，求 X 的分布列和均值 $E(X)$.

附：$\chi^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$，其中 $n=a+b+c+d$.

α	0.1	0.05	0.025	0.01	0.005	0.001
x_α	2.706	3.841	5.024	6.635	7.879	10.828

解：(1)因为$(0.005+p+0.018+0.020+0.022+0.025)\times10=1$，所以$p=0.010$.

所以$n=\dfrac{10}{0.1}=100$.

(2)因为$n=100$，所以"读书之星"有$100\times0.25=25$(人).

从而2×2列联表如下：

单位：人

性别	读书之星		合计
	不是	是	
男	30	15	45
女	45	10	55
合计	75	25	100

零假设为H_0："读书之星"与性别无关联.

将列联表中的数据代入公式计算得

$\chi^2=\dfrac{100\times(30\times10-15\times45)^2}{45\times55\times75\times25}=\dfrac{100}{33}\approx3.030<3.841=x_{0.05}$.

依据$\alpha=0.05$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"读书之星"与性别无关联.

(3)将频率视为概率，即从该地区学生中随机抽取一名学生是"读书之星"的概率为$\dfrac{1}{4}$.

由题意可知$X\sim B\left(3,\dfrac{1}{4}\right)$，

所以$P(X=0)=C_3^0\times\left(\dfrac{1}{4}\right)^0\times\left(1-\dfrac{1}{4}\right)^3=\dfrac{27}{64}$；

$P(X=1)=C_3^1\times\dfrac{1}{4}\times\left(1-\dfrac{1}{4}\right)^2=\dfrac{27}{64}$；

$P(X=2)=C_3^2\times\left(\dfrac{1}{4}\right)^2\times\left(1-\dfrac{1}{4}\right)=\dfrac{9}{64}$；

$P(X=3)=C_3^3\times\left(\dfrac{1}{4}\right)^3=\dfrac{1}{64}$.

所以X的分布列为

X	0	1	2	3
P	$\dfrac{27}{64}$	$\dfrac{27}{64}$	$\dfrac{9}{64}$	$\dfrac{1}{64}$

所以$E(X)=3\times\dfrac{1}{4}=\dfrac{3}{4}$.

2.羽毛球正式比赛的规则是：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.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：$\left\{p_i-\dfrac{1}{3}\right\}$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，并据此判断列联表中的数据是正式比赛数据还是训练数据.

附：$\chi^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$，其中$n=a+b+c+d$.

α	0.15	0.1	0.05	0.01	0.001
x_α	2.072	2.706	3.841	6.635	10.828

(1)**解**：完善2×2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$\chi^2=\dfrac{60\times(18\times24-12\times6)^2}{30\times30\times24\times36}=10>6.635=\chi_{0.01}$，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 $i(i\geqslant 2)$ 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 $p_i-\frac{1}{3}=\frac{2}{5}p_{i-1}-\frac{2}{15}=\frac{2}{5}\left(p_{i-1}-\frac{1}{3}\right)(i\geqslant 2)$.

又 $p_1=1$,所以 $p_1-\frac{1}{3}=\frac{2}{3}\neq 0$,

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 $E(X)=\sum\limits_{i=1}^{60}E(X_i)=\sum\limits_{i=1}^{60}p_{i+1}$

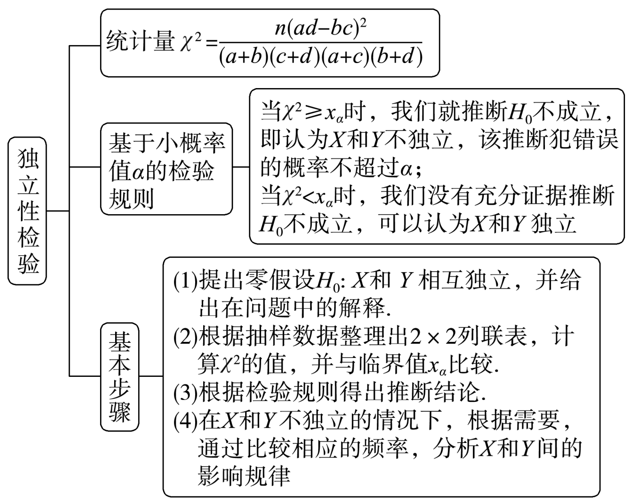
$=\sum\limits_{i=1}^{60}\left[\frac{1}{3}+\frac{2}{3}\times\left(\frac{2}{5}\right)^{i}\right]$

$=\frac{1}{3}\times 60+\dfrac{\frac{2}{3}\times\frac{2}{5}\times\left[1-\left(\frac{2}{5}\right)^{60}\right]}{1-\frac{2}{5}}$

$=20+\frac{4}{9}\cdot\left[1-\left(\frac{2}{5}\right)^{60}\right]=\frac{184}{9}-\frac{4}{9}\times\left(\frac{2}{5}\right)^{60}$.

因为 $E(X)\approx\frac{184}{9}\neq 24$,故列联表中的数据是训练数据.

体系构建



课后素养评价(二十二)

基础性·能力运用

1. 对于分类变量 X 与 Y 的随机变量 χ^2 的值，下列说法正确的是　(　　)

A. χ^2 越大，"X 与 Y 有关系"的可信程度越小

B. χ^2 越小，"X 与 Y 有关系"的可信程度越小

C. χ^2 越接近于 0，"X 与 Y 没有关系"的可信程度越小

D. χ^2 越大，"X 与 Y 没有关系"的可信程度越大

B　**解析**：χ^2 越大，"X 与 Y 没有关系"的可信程度越小，则"X 与 Y 有关系"的可信程度越大；χ^2 越小，"X 与 Y 有关系"的可信程度越小.

2. 为了检验某种血清预防感冒的作用，把 500 名使用血清的人与另外 500 名未使用血清的人一年中的感冒次数作比较，提出零假设 H_0：这种血清不能起到预防感冒的作用.若利用 2×2 列联表计算出 χ^2 的值，根据小概率值 $\alpha=0.01$ 的独立性检验，推断 H_0 不成立，则 χ^2 的一个可能取值为　(　　)

A. 7.879　　B. 5.024

C. 3.841　　D. 2.706

A　**解析**：由题意知 $\chi^2\geqslant6.635$，结合选项知 A 选项符合题意.

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 $\chi^2=\frac{50\times(18\times15-9\times8)^2}{27\times23\times26\times24}\approx5.059>5.024=x_{0.025}$，所以认为"对作业量的评价与学生的性别有关"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 $\frac{1}{2}$

B. 未出现"日落云里走"时，后半夜下雨的概率约为 $\frac{5}{9}$

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；由 $\chi^2\approx19.05>10.828=x_{0.001}$，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,统计各选科人数如表所示,则下列说法中正确的是　(　　)

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后所得的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 $\alpha=0.1$ 的独立性检验,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 $\chi^2=\frac{304\times(63\times82-42\times117)^2}{180\times124\times105\times199}\approx0.041<2.706=x_{0.1}$,所以根据 $\alpha=0.1$ 的独立性检验,认为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	$\frac{3a}{2}$	$\frac{3}{2}a$	$3a$

(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:$\chi^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$.

α	0.05	0.01	0.001
x_α	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 $P(\chi^2\geqslant 6.635)=0.01$，

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 $a\geqslant 9.9525$.

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.

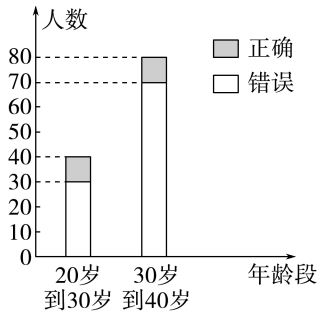
从所有顾客中选出 9 人，所以 X 的所有可能取值是 0，1，2，3，4，5，6，7，8，9，

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：选手从 1～8 号 8 扇大门中选择一扇，按响门上的门铃，门铃会播放一首歌的主旋律，选手需正确回答出这首歌的名称，方可获得该扇门对应的奖金.该电视台经过调查，发现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，能否认为是否猜对歌曲名称与年龄有关系？说明理由.

解：(1)根据题图得到 2×2 列联表如下：

单位：人

年龄段	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：是否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，即认为是否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.

综合性·创新提升

1. 利用独立性检验判断是否喜欢参加体育活动是不是与性别有关，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：是否喜欢参加体育活动与性别有关

D. H_0：是否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. 某校对甲、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 χ^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.某校在两个班进行教学方式对比试验，两个月后进行了一次检测，试验班与对照班成绩统计如下表所示：

单位：人

班级	成绩		合计
	80分及以上	80分以下	
试验班	35	15	50
对照班	20	m	50
合计	55	45	n

(1)$m=$________，$n=$________；

(2)依据小概率值 $\alpha=0.005$ 的独立性检验，可以认为教学方式与成绩________(填"有关系"或"没关系")。

(1)30　100　(2)有关系　**解析**：(1)$m=45-15=30$，$n=50+50=100$.

(2)由题表中的数据得 $\chi^2=\dfrac{100\times(35\times30-15\times20)^2}{50\times50\times55\times45}\approx9.091>7.879=x_{0.005}$，

依据小概率值 $\alpha=0.005$ 的独立性检验，我们推断 H_0 不成立，即认为教学方式与成绩有关系，此推断犯错误的概率不超过0.005.

4.现调查了某市500名居民的工作场所和呼吸系统健康情况，得到2×2列联表如下：

单位：人

呼吸系统疾病	工作场所		合计
	室外	室内	
有呼吸系统疾病	150		
无呼吸系统疾病		100	
合计	200		

(1)补全2×2列联表.

(2)依据小概率值 $\alpha=0.05$ 的独立性检验，能否认为呼吸系统疾病与工作场所有关？

(3)现采用分层随机抽样从室内工作的居民中抽取一个容量为6的样本，将该样本看成一个总体，从中随机地抽取2人，求2人都有呼吸系统疾病的概率.

解：(1)补全2×2列联表如下：

单位：人

呼吸系统疾病	工作场所		合计
	室外	室内	
有呼吸系统疾病	150	200	350
无呼吸系统疾病	50	100	150
合计	200	300	500

(2)零假设为 H_0：呼吸系统疾病与工作场所无关.

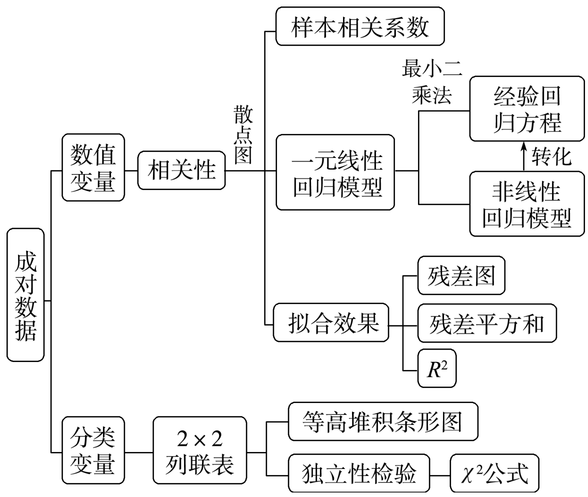
由(1)中列联表的数据计算得 $\chi^2=\dfrac{500\times(150\times100-200\times50)^2}{350\times150\times200\times300}\approx3.968>3.841=x_{0.05}$.

依据小概率值 $\alpha=0.05$ 的独立性检验，推断 H_0 不成立，即认为呼吸系统疾病与工作场所有关，此推断犯错误的概率不超过0.05.

(3)采用分层随机抽样从室内工作的居民中抽取6人，其中有呼吸系统疾病的抽4人，无呼吸系统疾病的抽2人.设事件 A 为"从中随机地抽取2人，2人都有呼吸系统疾病"，则 $P(A)=\dfrac{C_4^2}{C_6^2}=\dfrac{2}{5}$.

单元活动研习

单元知识重构



专题任务探究

探究点一　变量的相关关系

例 1　某市管理部门以周为单位，记录的每周查处的酒驾人数与该周内出现的交通事故数如下：

酒驾人数	80	87	96	100	104	121	147
交通事故数	19	20	25	23	24	30	31

通过表中数据可知，酒驾人数与交通事故数之间是　　（　　）

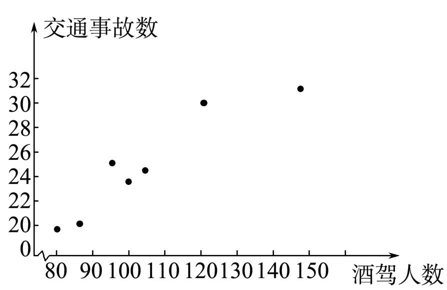
A.正相关

B.负相关

C.不相关

D.函数关系

A　**解析**：由表格中的数据，在直角坐标系中绘出散点图，如图所示．由图可知，散点从左向右呈上升趋势，并集中在一条直线附近，所以酒驾人数与交通事故数具有线性相关关系，且是正相关．

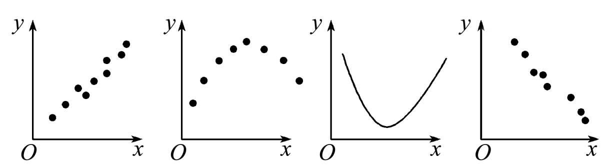


反思提炼

可根据样本数据作出散点图来确定两个变量之间是否具有相关关系，也可通过计算样本相关系数 r 进行判断．

探究训练

1.下列四个图各反映了两个变量的某种关系，其中可以看作具有线性相关关系的是　　（　　）



①　②　③　④

A.①③　　B.①④

C.②③　　D.①②

B　**解析**：对于两个变量的散点图，若样本点成直线形带状分布，则两个变量具有线性相关关系，所以两个变量具有线性相关关系的图是①和④．故选 B.

2.已知变量 x 和 y 满足关系 $y=-2x+1$，变量 y 与 z 正相关，下列结论中正确的是　　（　　）

A.x 与 y 正相关，x 与 z 负相关

B.x 与 y 正相关，x 与 z 正相关

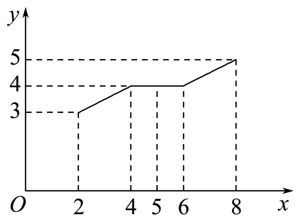
C.x 与 y 负相关，x 与 z 负相关

D.x 与 y 负相关，x 与 z 正相关

C　**解析**：根据题意，变量 x 和 y 满足关系 $y=-2x+1$，其中系数 $-2<0$，所以 x 与 y 负相关．又由变量 y 与 z 正相关，所以 x 与 z 负相关．故选 C.

探究点二　一元线性回归模型

例 2　如图是某采矿厂的污水年排放量 y(单位:t)与矿产品年产量 x(单位:t)的折线图.



(1)依据折线图计算样本相关系数 r(精确到0.01),并据此判断是否可用一元线性回归模型拟合 y 与 x 的关系?(若$|r|>0.75$,则线性相关程度很高,可用一元线性回归模型拟合)

(2)若可用一元线性回归模型拟合 y 与 x 的关系,请建立 y 关于 x 的经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$,并预测矿产品年产量为 10 t 时的污水年排放量.

附:$r=\dfrac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}}$,$\sqrt{0.3}\approx0.55$,$\sqrt{0.9}\approx0.95$.

经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$ 中,$\hat{b}=\dfrac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}$,$\hat{a}=\overline{y}-\hat{b}\,\overline{x}$.

解:(1)由题图中样本数据计算得,

$\overline{x}=5$,$\overline{y}=4$,$\sum\limits_{i=1}^{5}(x_i-\overline{x})(y_i-\overline{y})=6$,

$\sum\limits_{i=1}^{5}(x_i-\overline{x})^2=20$,$\sum\limits_{i=1}^{5}(y_i-\overline{y})^2=2$,

所以样本相关系数 $r=\dfrac{6}{\sqrt{20}\times\sqrt{2}}=\sqrt{0.9}\approx0.95$.

因为 $0.95>0.75$,所以可用一元线性回归模型拟合 y 与 x 的关系.

(2)由(1)中数据可得 $\hat{b}=\dfrac{6}{20}=0.3$,

$\hat{a}=\overline{y}-\hat{b}\,\overline{x}=4-0.3\times5=2.5$,

所以经验回归方程为 $\hat{y}=0.3x+2.5$.

当 $x=10$ 时,$\hat{y}=5.5$.

所以预测年产量为 10 t 时的污水年排放量约为 5.5 t.

[**一题多思**]

根据经验回归方程,预测污水年排放量为 10 t 时的矿产品年产量.

解:将 $\hat{y}=10$ 代入经验回归方程 $\hat{y}=0.3x+2.5$,可得 $x=25$,所以预测污水年排放量为 10 t 时的矿产品年产量为 25 t.

反思提炼

1. 若两个变量具有线性相关关系,则可通过经验回归方程估计和预测变量的值.

2. 正确运用计算 $\hat{b}$,$\hat{a}$ 的公式和准确的计算是求经验回归方程的关键.要充分利用经验回归直线 $\hat{y}=\hat{b}x+\hat{a}$ 过样本点的中心$(\overline{x},\overline{y})$进行求值.

3. 残差平方和越小,决定系数 R^2 越大,所选模型拟合效果就越好.

探究训练

1.(多选)关于变量 x,y 的 n 个样本点(x_1,y_1),(x_2,y_2),…,(x_n,y_n)及其经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$,下列说法正确的有　　(　　)

A.样本相关系数 r 的绝对值越接近 0,表示 x,y 的线性相关程度越强

B.决定系数 R^2 的值越接近 1,表示模型拟合效果越好

C.残差平方和越大,表示模型拟合效果越好

D.若 $\overline{x}=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i$,$\overline{y}=\dfrac{1}{n}\sum\limits_{i=1}^{n}y_i$,则点$(\overline{x},\overline{y})$一定在经验回归直线 $\hat{y}=\hat{b}x+\hat{a}$ 上

BD　**解析**:根据样本相关系数的意义可知,当 r 的绝对值越接近于 0 时,两变量的线性相关程度越弱,则 A 错误;用决定系数 R^2 来刻画模型拟合效果时,R^2 越大,说明模型的拟合效果越好,则 B 正确;模型拟合效果的好坏可由残差平方和来体现,残差平方和越大,模型拟合效果越差,则 C 错误;样本点中心一定在经验回归直线上,则 D 正确.故选 BD.

2. 某高中高三年级对各科考试的评价指标中,有"难度系数"和"区分度"两个指标:

①难度系数$=\dfrac{\text{年级平均分}}{\text{满分}}$,②区分度$=\dfrac{\text{实验班的平均分}-\text{普通班的平均分}}{\text{满分}}$.

(1)某次数学考试(满分为 150 分),学校从实验班和普通班各随机抽取 3 人,实验班 3 人的成绩(单位:分)分别为 147,142,137;普通班 3 人的成绩分别为 97,102,113.通过样本估计本次考试的区分度.(精确到0.01)

(2)如下表格是该校高三年级 6 次数学考试的统计数据.

难度系数 x	0.64	0.71	0.74	0.76	0.77	0.82
区分度 y	0.18	0.23	0.24	0.24	0.22	0.15

①计算样本相关系数 r（精确到 0.01），当 $|r|<0.75$ 时，可认为两个变量线性相关性弱；当 $|r|\geqslant 0.75$ 时，可认为两个变量线性相关性强.通过 r 的值判断，能否利用一元线性回归模型描述 y 与 x 的关系？

②若 $t_i=|x_i-0.74|(i=1,2,\cdots,6)$，求出 y 关于 t 的经验回归方程 $\hat{y}=\hat{b}t+\hat{a}$，并预测当 $x=0.75$ 时 $\hat{y}$ 的值（精确到 0.01）.

附：$\sum\limits_{i=1}^{6}x_iy_i=0.930\ 9$，

$\sqrt{\sum\limits_{i=1}^{6}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{6}(y_i-\overline{y})^2}\approx 0.011\ 2$，

$\sum\limits_{i=1}^{6}t_iy_i=0.048\ 3$，$\sum\limits_{i=1}^{6}(t_i-\overline{t})^2\approx 0.007\ 3$.

样本相关系数

$$r=\frac{\sum\limits_{i=1}^{n}x_iy_i-n\overline{x}\ \overline{y}}{\sqrt{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}}.$$

经验回归直线 $\hat{y}=\hat{b}t+\hat{a}$ 的斜率和截距的最小二乘估计分别为 $\hat{b}=\dfrac{\sum\limits_{i=1}^{n}t_iy_i-n\overline{t}\ \overline{y}}{\sum\limits_{i=1}^{n}(t_i-\overline{t})^2}$，$\hat{a}=\overline{y}-\hat{b}\overline{t}$.

解：(1) 实验班 3 人成绩的均值为 $\dfrac{147+142+137}{3}=142$，

普通班 3 人成绩的均值为 $\dfrac{97+102+113}{3}=104$，

故估计本次考试的区分度为 $\dfrac{142-104}{150}\approx 0.25$.

(2)①由题中的表格可知

$\overline{x}=\dfrac{1}{6}\times(0.64+0.71+0.74+0.76+0.77+0.82)=0.74$，

$\overline{y}=\dfrac{1}{6}\times(0.18+0.23+0.24+0.24+0.22+0.15)=0.21$，

故 $r=\dfrac{\sum\limits_{i=1}^{6}x_iy_i-6\overline{x}\ \overline{y}}{\sqrt{\sum\limits_{i=1}^{6}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{6}(y_i-\overline{y})^2}}$

$\approx\dfrac{0.930\ 9-6\times 0.74\times 0.21}{0.011\ 2}\approx -0.13$.

因为 $|r|<0.75$，所以线性相关性弱.

故不能利用一元线性回归模型描述 y 与 x 的关系.

②y 与 t 的值如下表：

t	0.10	0.03	0	0.02	0.03	0.08
y	0.18	0.23	0.24	0.24	0.22	0.15

因为 $\hat{b}=\dfrac{\sum\limits_{i=1}^{6}t_iy_i-6\overline{t}\ \overline{y}}{\sum\limits_{i=1}^{6}(t_i-\overline{t})^2}\approx -0.86$，

所以 $\hat{a}=\overline{y}-\hat{b}\overline{t}\approx 0.25$.

所以所求经验回归方程为 $\hat{y}=-0.86t+0.25$.

当 $x=0.75$ 时，$t=0.01$，则 $\hat{y}\approx 0.24$.

探究点三　列联表与独立性检验

例 3　某通信公司为了更好地满足消费者对流量的需求，推出了不同定价的流量包，经过一个月的统计，获取了一个容量为 80 万的样本.同时为了进一步了解年龄是否对所选流量包的价格有影响，统计了小于 40 岁和大于等于 40 岁两个年龄段人群购买不同定价的流量包的人数，收集数据整理如表所示.

定价/（元/月）	20	30	50	60
<40 岁/万人	10	15	7	8
$\geqslant 40$ 岁/万人	20	12	6	2

若把 50 元/月以下（不包括 50 元/月）的流量包称为低价流量包，50 元/月以上（包括 50 元/月）的流量包称为高价流量包，根据以上数据完成下面的列联表；依据 $\alpha=0.05$ 的独立性检验，判断年龄段和所选流量包的价格是否有关联.

单位：万人

年龄段	流量包价格		合计
	<50 元/月	$\geqslant 50$ 元/月	
<40 岁			
$\geqslant 40$ 岁			
合计			

附：$\chi^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$，$n=a+b+c+d$.

α	0.1	0.05	0.01	0.005	0.001
x_α	2.706	3.841	6.635	7.879	10.828

解：由题中数据完善 2×2 列联表如下表所示.

单位:万人

年龄段	流量包价格		合计
	<50元/月	$\geqslant 50$元/月	
<40岁	25	15	40
$\geqslant 40$岁	32	8	40
合计	57	23	80

零假设为 H_0:年龄段和所选流量包的价格无关联.

由表中数据计算得 $\chi^2=\dfrac{80\times(25\times8-32\times15)^2}{40\times40\times57\times23}\approx$ $2.990<3.841=x_{0.05}$,

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,

即认为年龄段和所选流量包的价格没有关联.

[一题多思]

试建立购买流量包的总人数 y(单位:万人)关于定价 x(单位:元/月)的经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$,并估计定价为10元/月的流量包的购买人数.

解:由例题表格中的数据可得 $\overline{x}=\dfrac{20+30+50+60}{4}=40$,$\overline{y}=\dfrac{30+27+13+10}{4}=20$,

则 $\hat{b}=\dfrac{\sum\limits_{i=1}^{4}(x_i-\overline{x})(y_i-\overline{y})}{\sum\limits_{i=1}^{4}(x_i-\overline{x})^2}=-0.54$,

$\hat{a}=\overline{y}-\hat{b}\,\overline{x}=20+0.54\times40=41.6$,

所以 y 关于 x 的经验回归方程为 $\hat{y}=-0.54x+41.6$.

当 $x=10$ 时,$\hat{y}=-5.4+41.6=36.2$.

因此,估计定价为10元/月的流量包的购买人数为36.2万人.

反思提炼

独立性检验的基本思想类似于数学中的反证法.先假设“两个分类变量无关”成立,计算随机变量 χ^2 的值.如果 χ^2 的值大于或等于临界值,说明假设不成立.χ^2 越大,两个分类变量有关的可能性越大.

探究训练

1. 一项试验旨在研究臭氧效应,试验方案如下:选40只小白鼠,随机地将其中20只分配到试验组,另外20只分配到对照组,试验组的小白鼠饲养在高浓度臭氧环境中,对照组的小白鼠饲养在正常环境中,一段时间后统计每只小白鼠体重的增加量(单位:g).

试验结果如下:

对照组的小白鼠体重的增加量从小到大排列为

15.2　18.8　20.2　21.3　22.5　23.2　25.8
26.5　27.5　30.1　32.6　34.3　34.8　35.6
35.6　35.8　36.2　37.3　40.5　43.2

试验组的小白鼠体重的增加量从小到大排列为

7.8　9.2　11.4　12.4　13.2　15.5　16.5
18.0　18.8　19.2　19.8　20.2　21.6　22.8
23.6　23.9　25.1　28.2　32.3　36.5

(1)求40只小白鼠体重的增加量的中位数 m,再分别统计两组数据中小于 m 与不小于 m 的数据的个数,完成如下列联表:

单位:只

组别	体重增加量		合计
	$<m$	$\geqslant m$	
对照组			
试验组			
合计			

(2)根据(1)中的列联表,依据小概率值 $\alpha=0.05$ 的独立性检验,能否推断小白鼠在高浓度臭氧环境中与正常环境中体重的增加量有差异?

附:$\chi^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$,$n=a+b+c+d$.

α	0.1	0.05	0.01	0.005	0.001
x_α	2.706	3.841	6.635	7.879	10.828

解:(1)依题意,可知这40只小白鼠体重增加量的中位数是将两组数据合在一起,从小到大排列后第20个数据与第21个数据的平均数,观察数据可得第20个数据为23.2,第21个数据为23.6,

所以 $m=\dfrac{23.2+23.6}{2}=23.4$.

故列联表为

单位:只

组别	体重增加量		合计
	$<m$	$\geqslant m$	
对照组	6	14	20
试验组	14	6	20
合计	20	20	40

(2)零假设为 H_0:小白鼠在高浓度臭氧环境中与正常环境中体重的增加量无差异.由(1)中列联表中的数据可得,$\chi^2=\dfrac{40\times(6\times6-14\times14)^2}{20\times20\times20\times20}=6.4>$

$3.841=x_{0.05}$，

依据小概率值 $\alpha=0.05$ 的独立性检验，我们推断 H_0 不成立，即认为小白鼠在高浓度臭氧环境中与正常环境中体重的增加量有差异，此推断犯错误的概率不大于0.05.

2.党的二十大报告指出：必须牢固树立和践行绿水青山就是金山银山的理念，站在人与自然和谐共生的高度谋划发展.推进垃圾分类处理是落实绿色发展理念的必然选择.为调查居民对垃圾的处理情况，某社区居委会随机抽取400名社区居民参与问卷调查.经统计，有60%的居民对垃圾分类处理，其中女性占$\frac{2}{3}$；有40%的居民不对垃圾分类处理，其中男性和女性各占$\frac{1}{2}$.

(1)请根据以上信息完成2×2列联表，并根据小概率值 $\alpha=0.005$ 的独立性检验，判断能否认为垃圾处理方式与性别有关.

单位：人

性别	垃圾处理方式		合计
	不分类	分类	
男			
女			
合计			

(2)为了提高社区居民对垃圾分类处理的意识，该社区成立了垃圾分类宣传小组，利用周末的时间在社区进行垃圾分类宣传活动，并在每周宣传活动结束后，随机抽取400名社区居民统计其中不对垃圾分类处理的居民人数，统计数据如下：

周次	1	2	3	4	5
不对垃圾分类处理的人数	120	105	100	95	80

请根据所给的数据，建立不对垃圾分类处理的人数 y 关于周次 x 的经验回归方程，并预测第10周不对垃圾分类处理的人数.

解：(1)由题意得，对垃圾分类处理的居民人数为 $400\times0.6=240$，其中女性人数为 $240\times\frac{2}{3}=160$，不对垃圾分类处理的居民人数为 $400\times0.4=160$，其中男性和女性的人数均为 $160\times\frac{1}{2}=80$.

则列联表为

单位：人

性别	垃圾处理方式		合计
	不分类	分类	
男	80	80	160
女	80	160	240
合计	160	240	400

零假设为 H_0：垃圾处理方式与性别无关.

根据列联表中的数据，经计算得到

$\chi^2=\dfrac{400\times(80\times160-80\times80)^2}{160\times240\times160\times240}\approx11.1>7.879=x_{0.005}$.

根据小概率值 $\alpha=0.005$ 的独立性检验，我们推断 H_0 不成立，即认为垃圾处理方式与性别有关，该推断犯错误的概率不大于0.005.

(2)由题可得 $\overline{x}=\dfrac{1+2+3+4+5}{5}=3$，$\overline{y}=\dfrac{120+105+100+95+80}{5}=100$，

$\sum\limits_{i=1}^{5}x_iy_i=1\ 410$，$\sum\limits_{i=1}^{5}x_i^2=55$，

则 $\hat{b}=\dfrac{\sum\limits_{i=1}^{5}x_iy_i-5\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{5}x_i^2-5\overline{x}^2}=\dfrac{1\ 410-1\ 500}{55-45}=-9$，

所以 $\hat{a}=\overline{y}-\hat{b}\,\overline{x}=100-(-9)\times3=127$.

即所求的经验回归方程为 $\hat{y}=-9x+127$.

令 $x=10$，

得 $\hat{y}=-9\times10+127=37$.

所以预测第10周不对垃圾分类处理的人数为37.

第八章综合检测

（时间：120 分钟，分值：150 分）

一、单项选择题（本题共 8 小题，每小题 5 分，共 40 分）

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)，…，(x_n,y_n)，则下列说法中不正确的是（　　）

A.用决定系数 R^2 来刻画回归效果，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 来刻画回归效果，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 之间的线性相关关系越强，正确.故选 A.

2. 已知呈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（　　）

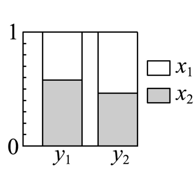
x	0.1	0.2	0.3	0.5
y	2.11	2.85	4.08	10.15

A.(0.1，2.11)　　B.(0.2，2.85)

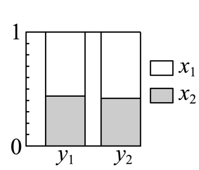
C.(0.3，4.08)　　D.(0.275，4.797 5)

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.

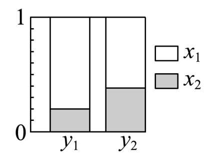
3. 观察下列的等高堆积条形图，其中最有把握认为两个分类变量 x，y 之间没有关系的是（　　）



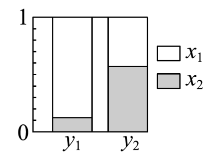
A



B



C



D

B　**解析**：在等高堆积条形图中，当 x_1，x_2 所占比例相差越大时，越有把握认为两个分类变量 x，y 之间有关系.

B 选项中，x_1，x_2 所占比例相差无几，所以最有把握认为两个分类变量 x，y 之间没有关系.故选 B.

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

由散点图可知用水量 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. 某校团委针对学生性别和中学生追星是否有关做了一次抽样调查，样本中女生人数是男生人数的 $\dfrac{1}{2}$，男生喜欢追星的人数占男生人数的 $\dfrac{1}{6}$，女生喜

欢追星的人数占女生人数的$\frac{2}{3}$.若在犯错误的概率不超过 0.05 的前提下认为是否喜欢追星和性别有关，则样本中男生至少有（　　）

附：$\chi^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$.

A.12 人　　B.11 人　　C.10 人　　D.18 人

A　**解析**：设男生人数为 x，依题意可得如下列联表：

单位：人

性别	是否追星		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	$\frac{3x}{2}$

若在犯错误的概率不超过 0.05 的前提下认为是否追星和性别有关，则 $\chi^2\geqslant 3.841$. 由 $\chi^2=\frac{\frac{3x}{2}\left(\frac{x^2}{36}-\frac{5x^2}{18}\right)^2}{x\cdot\frac{x}{2}\cdot\frac{x}{2}\cdot x}=\frac{3}{8}x\geqslant 3.841$ 及$\frac{x}{2}$，$\frac{x}{6}$为整数，可知若在犯错误的概率不超过0.05的前提下认为是否追星和性别有关，则男生至少有 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，在一定条件下可用回归模型 $y=2\lg x$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 与 y，下列说法中正确的是（　　）

①若 $r>0$，则 x 增大时，y 也相应增大；

②若 $r<0$，则 x 增大时，y 也相应增大；

③若 $r=1$ 或 $r=-1$，则 x 与 y 的关系为函数关系，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 $r>0$，表示两个变量正相关，x 增大时，y 也相应增大，故①正确.若 $r<0$，表示两个变量负相关，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强，$|r|=1$ 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（本题共 3 小题，每小题 6 分，共18 分）

9.已知在最小二乘法原理下，具有相关关系的变量 x，y 之间的经验回归方程为 $\hat{y}=-0.7x+10.3$，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 $\hat{b}=-0.7<0$，

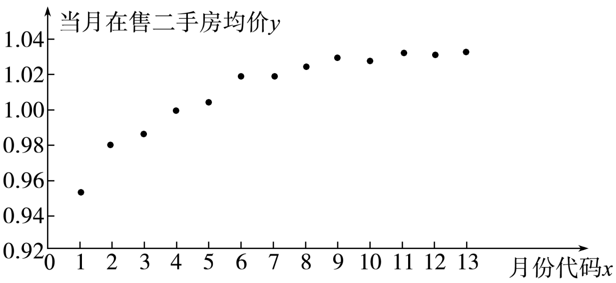
所以变量 x，y 之间呈负相关关系，故 A 错误；

对于 B，当 $x=20$ 时，$\hat{y}=-0.7\times 20+10.3=-3.7$，故 B 错误；

对于 C，由表中数据可知$\overline{x}=9$，$\overline{y}=\frac{6+m+3+2}{4}=\frac{11+m}{4}$，由点$(\overline{x},\overline{y})$必在经验回归直线上，得$\frac{11+m}{4}=-0.7\times 9+10.3$，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.如图是某小区 2023 年 8 月至 2024 年 8 月各月在售二手房均价(单位:万元/m^2)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 $\hat{y}=0.955\ 4+0.030\ 6\ln x$,并得到以下一些统计量的值:

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 $\hat{y}=0.936\ 9+0.028\ 5\sqrt{x}$ 预测该小区 2025 年 8 月在售二手房均价约为 1.079 4 万元/m^2

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 $x=25$,得 $\hat{y}=0.936\ 9+0.028\ 5\times\sqrt{25}=1.079\ 4$,所以可以预测 2025 年 8 月在售二手房均价约为 1.079 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 名居民用餐能否做到"光盘",得到如下的 2×2 列联表,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.进行独立性检验时,零假设为 H_0:能否做到"光盘"与性别有关

C.$\chi^2>3.841$

D.在犯错误的概率不超过 0.1 的前提下,可以认为该市居民能否做到"光盘"与性别有关

AD　**解析**:由 2×2 列联表得到 $a=45,b=10,c=30,d=15$,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"与性别无关.

$$\chi^2=\frac{100\times(675-300)^2}{55\times45\times75\times25}\approx3.030.$$

因为 $2.706<3.030<3.841$,

所以在犯错误的概率不超过 0.1 的前提下,可以认为该市居民能否做到"光盘"与性别有关.

三、**填空题**(本题共 3 小题,每小题 5 分,共 15 分)

12.已知 y 关于 x 的经验回归方程为 $\hat{y}=\hat{b}x+\hat{a}$,当 $x=3$ 时,y 的估计值是 17,当 $x=8$ 时,y 的估计值是 22.y 关于 x 的经验回归方程为________;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. 一般来说，一个人脚越长，他的身高就越高.现对 10 名成年人的脚长 x（单位：cm）与身高 y（单位：cm）进行测量，得到数据作出散点图后，发现各点集中在一条直线附近.经计算得到一些数据：$\overline{x}=24.5$，$\overline{y}=171.5$，$\sum\limits_{i=1}^{10}(x_i-\overline{x})(y_i-\overline{y})=577.5$，$\sum\limits_{i=1}^{10}(x_i-\overline{x})^2=82.5$.某刑侦人员在某案发现场发现一对裸脚印，量得每个脚印长 26.5 cm，则估计嫌疑人的身高为________cm.

185.5　**解析**：由已知得 $\hat{b}=\dfrac{\sum\limits_{i=1}^{10}(x_i-\overline{x})(y_i-\overline{y})}{\sum\limits_{i=1}^{10}(x_i-\overline{x})^2}=\dfrac{577.5}{82.5}=7$，$\hat{a}=\overline{y}-\hat{b}\,\overline{x}=0$，故经验回归方程为 $\hat{y}=7x$.

当 $x=26.5$ 时，$\hat{y}=185.5$.

故答案为 185.5.

四、解答题（本题共 5 小题，共 77 分）

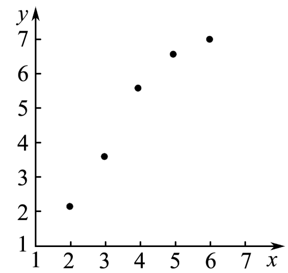
15.（13 分）已知某设备的使用年限 x（单位：年）和所需的维修费用 y（单元：万元）有如下的统计资料：

x	2	3	4	5	6
y	2.2	3.8	5.5	6.5	7.0

（1）y 与 x 之间是否具有线性相关关系？若有，求出经验回归方程.

（2）估计使用年限为 10 年时，维修费用是多少.

解：（1）作出散点图，如图所示.



由散点图可知，y 与 x 呈线性相关关系.

由题可得 $\overline{x}=4$，$\overline{y}=5$，$\sum\limits_{i=1}^{5}x_iy_i=112.3$，$\sum\limits_{i=1}^{5}x_i^2=90$，

所以 $\hat{b}=\dfrac{\sum\limits_{i=1}^{5}x_iy_i-5\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{5}x_i^2-5\,\overline{x}^2}=1.23$，

$\hat{a}=\overline{y}-\hat{b}\overline{x}=5-1.23\times4=0.08$.

所以经验回归方程为 $\hat{y}=1.23x+0.08$.

（2）当 $x=10$ 时，

$\hat{y}=1.23\times10+0.08=12.3+0.08=12.38$.

所以估计使用年限为 10 年时，维修费用是 12.38 万元.

16.（15 分）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	对冬季长跑活动是否感兴趣		合计
	是	否	
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:$\chi^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$,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	对冬季长跑活动是否感兴趣		合计
	是	否	
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 $\chi^2=\frac{100\times(48\times12-8\times32)^2}{56\times44\times80\times20}\approx2.597<2.706=x_{0.1}$,

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 $P(X=0)=\frac{C_5^3}{C_8^3}=\frac{10}{56}=\frac{5}{28}$,

$P(X=1)=\frac{C_5^2C_3^1}{C_8^3}=\frac{30}{56}=\frac{15}{28}$,

$P(X=2)=\frac{C_5^1C_3^2}{C_8^3}=\frac{15}{56}$,

$P(X=3)=\frac{C_3^3}{C_8^3}=\frac{1}{56}$,

所以 X 的分布列为

X	0	1	2	3
P	$\frac{5}{28}$	$\frac{15}{28}$	$\frac{15}{56}$	$\frac{1}{56}$

$E(X)=0\times\frac{5}{28}+1\times\frac{15}{28}+2\times\frac{15}{56}+3\times\frac{1}{56}=\frac{9}{8}$.

17.(15 分)(2022·新高考全国Ⅰ卷)一医疗团队为研究某地的一种地方性疾病与当地居民的卫生习惯(卫生习惯分为良好和不够良好两类)的关系,在已患该疾病的病例中随机调查了 100 例(称为病例组),同时在未患该疾病的人群中随机调查了 100 人(称为对照组),得到如下数据:

单位:人

组别	卫生习惯	
	不够良好	良好
病例组	40	60
对照组	10	90

(1)能否有 99%的把握认为患该疾病群体与未患该疾病群体的卫生习惯有差异?

(2)从该地的人群中任选一人,事件 A 表示"选到的人卫生习惯不够良好",事件 B 表示"选到的人患有该疾病".$\frac{P(B|A)}{P(\bar{B}|A)}$与$\frac{P(B|\bar{A})}{P(\bar{B}|\bar{A})}$的比值是卫生习惯不够良好对患该疾病风险程度的一项度量指标,记该指标为 R.

①证明:$R=\frac{P(A|B)}{P(\bar{A}|B)}\cdot\frac{P(\bar{A}|\bar{B})}{P(A|\bar{B})}$;

②利用调查数据,给出 $P(A|B)$,$P(A|\bar{B})$的估计值,并利用①中的结果给出 R 的估计值.

(1)**解**:由已知得 $\chi^2=\frac{200\times(40\times90-60\times10)^2}{100\times100\times50\times150}=24$.

又因为 $P(\chi^2\geqslant6.635)=0.01$,$24>6.635$,

所以有 99%的把握认为患该疾病群体与未患该疾病群体的卫生习惯有差异.

(2)①**证明**:因为 $R=\frac{P(B|A)}{P(\bar{B}|A)}\cdot\frac{P(\bar{B}|\bar{A})}{P(B|\bar{A})}=\frac{P(AB)}{P(A)}\cdot\frac{P(A)}{P(A\bar{B})}\cdot\frac{P(\bar{A}\bar{B})}{P(\bar{A})}\cdot\frac{P(\bar{A})}{P(\bar{A}B)}$,

所以 $R=\frac{P(AB)}{P(B)}\cdot\frac{P(B)}{P(\bar{A}B)}\cdot\frac{P(\bar{A}\bar{B})}{P(\bar{B})}\cdot\frac{P(\bar{B})}{P(A\bar{B})}$

$=\frac{P(A|B)}{P(\overline{A}|B)}\cdot\frac{P(\overline{A}|\overline{B})}{P(A|\overline{B})}$.

②**解**：由已知得 $P(A|B)=\frac{40}{100}$，$P(A|\overline{B})=\frac{10}{100}$，

$P(\overline{A}|B)=\frac{60}{100}$，$P(\overline{A}|\overline{B})=\frac{90}{100}$，

所以 $R=\frac{P(A|B)}{P(\overline{A}|B)}\cdot\frac{P(\overline{A}|\overline{B})}{P(A|\overline{B})}=6$.

18.(17分)某农科所对冬季昼夜温差与某反季节大豆新品种发芽数之间的关系进行分析研究，他们分别记录了12月1日至12月5日每天的昼夜温差与每100颗种子中的发芽数，得到如下资料：

日期	12月1日	12月2日	12月3日	12月4日	12月5日
温差 x/℃	10	11	13	12	8
发芽数 y/颗	23	25	30	26	16

该农科所确定的研究方案是先从这5组数据中任选2组，用剩下的3组数据求经验回归方程，再对被选取的2组数据进行检验.

(1)求选取的2组数据恰好是不相邻2天数据的概率；

(2)若选取的是12月1日与12月5日的2组数据，请根据12月2日至12月4日的数据，求出 y 关于 x 的经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$；

(3)若由经验回归方程得到的估计数据与所选出的检验数据的误差均不超过2，则认为得到的经验回归方程是可靠的，试问(2)中所得到的经验回归方程是否可靠.

解：(1)设"抽到不相邻的两组数据"为事件 A，因为从5组数据中选取2组数据共有10种情况：(1，2)，(1，3)，(1，4)，(1，5)，(2，3)，(2，4)，(2，5)，(3，4)，(3，5)，(4，5)，其中数据为12月份的日期数.

每种情况都是等可能出现的，事件 A 包含的样本点有6个，所以 $P(A)=\frac{6}{10}=\frac{3}{5}$.

所以选取的2组数据恰好是不相邻2天数据的概率是 $\frac{3}{5}$.

(2)由题表中数据，求得 $\overline{x}=12$，$\overline{y}=27$，$\sum\limits_{i=1}^{3}x_iy_i=977$，$\sum\limits_{i=1}^{3}x_i^2=434$，

所以 $\hat{b}=\frac{977-3\times12\times27}{434-3\times12^2}=2.5$，$\hat{a}=27-2.5\times12=-3$.

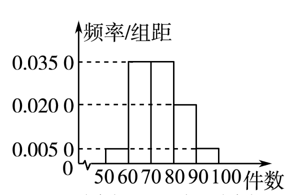
所以 y 关于 x 的经验回归方程为 $\hat{y}=2.5x-3$.

(3)当 $x=10$ 时，$\hat{y}=2.5\times10-3=22$，$|22-23|<2$；

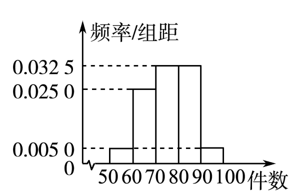
当 $x=8$ 时，$\hat{y}=2.5\times8-3=17$，$|17-16|<2$；

所以该农科所得到的经验回归方程是可靠的.

19.(17分)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小于60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小于80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$p=\frac{7}{10}$.

(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$60\times0.25=15$(名),

25周岁以下组中的生产能手有$40\times0.375=15$(名).

据此可得2×2列联表如下:

单位:名

组别	生产能手		合计
	是	不是	
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.

由数据得$\chi^2=\frac{100\times(15\times25-45\times15)^2}{60\times40\times30\times70}=\frac{25}{14}\approx1.79<2.706=x_{0.1}$.

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.

模块综合检测(一)

（时间：120 分钟，分值：150 分）

一、单项选择题(本题共 8 小题，每小题 5 分，共40 分)

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.在吸烟与患肺病是否有关的研究中，下列属于两个分类变量的是 （　　）

A.吸烟，不吸烟

B.患病，不患病

C.是否吸烟，是否患病

D.以上都不对

C　**解析**：“是否吸烟”是分类变量，它的两个不同取值：吸烟和不吸烟；“是否患病”是分类变量，它的两个不同取值：患病和不患病.可知 A，B 都是一个分类变量所取的两个不同值.故选 C.

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 ξ，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.某商品销售量 y(单位：件)与销售价格 x(单位：元/件)负相关，则其经验回归方程可能是 （　　）

A.$\hat{y}=-10x+200$

B.$\hat{y}=10x+200$

C.$\hat{y}=-10x-200$

D.$\hat{y}=10x-200$

A　**解析**：由 x 与 y 负相关，可排除 B，D 两项，而 C 项中的 $\hat{y}=-10x-200<0$ 恒成立，不符合题意.故选 A.

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

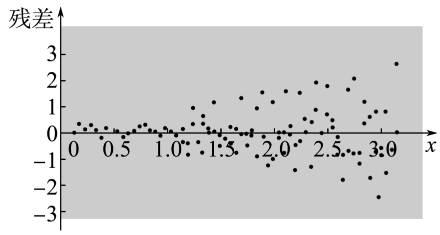
$P(A\cap B)=\frac{3\times2}{4\times3}=\frac{1}{2}$，$P(A)=\frac{3}{4}$，

所以 $P(B|A)=\frac{P(A\cap B)}{P(A)}=\frac{\frac{1}{2}}{\frac{3}{4}}=\frac{2}{3}$.

(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 $P(B|A)=\frac{n(A\cap B)}{n(A)}=\frac{6}{9}=\frac{2}{3}$.

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，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 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(本题共 3 小题，每小题 6 分，共18 分)

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，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 个小球，其中有 3 个红球、7 个白球，现从盒中随机地抽取 4 个小球，则 （　　）

A.恰有 1 个红球的概率为$\dfrac{1}{30}$

B.4 个全是白球的概率为$\dfrac{1}{6}$

C.恰有 2 个白球的概率为$\dfrac{3}{10}$

D.至多有 2 个红球的概率为$\dfrac{1}{3}$

BC　**解析：**设 $X=k$ 表示取出的小球恰有k 个为白球，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 个红球的概率为 $P(X=3)=\dfrac{1}{2}$，

至多有 2 个红球的概率为 $P(X=4)+P(X=3)+P(X=2)=\dfrac{29}{30}$.

11.已知随机变量 ξ 服从正态分布 $N(2,\sigma^2)$，且$P(\xi<4)=0.8$，则 （　　）

A.正态曲线关于直线 $x=2$ 对称

B.$\sigma=2$

C.$P(\xi>4)=P(\xi<0)$

D.$P(0<\xi<2)=0.2$

AC　**解析**：正态曲线关于直线 $x=2$ 对称，根据其对称性求解概率.由 $P(\xi<4)=0.8$，知 $P(\xi>4)=P(\xi<0)=0.2$，故 $P(0<\xi<2)=0.3$.故选 AC.

三、填空题(本题共 3 小题，每小题 5 分，共 15 分)

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.某考察团对全国十个城市进行职工人均工资水平 x(单位：千元)与居民人均消费水平 y(单位：千元)统计调查，得到 y 与 x 具有相关关系，经验回归方程为 $\hat{y}=0.66x+1.562$.当人均工资增加 1 千元时，人均消费水平平均增加________千元.若某城市居民人均消费水平为 7.675 千元，估计该城市人均消费额占人均工资收入的百分比约为________.

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，$x=\dfrac{7.675-1.562}{0.66}\approx9.262$，所以$\dfrac{7.675}{9.262}\approx0.829\approx83\%$.

14.一袋中有大小相同的 4 个红球和 2 个白球，给出下列结论：

①从中任取 3 球，恰有 1 个白球的概率是$\dfrac{3}{5}$；

②从中有放回地取球 6 次，每次任取 1 球，则取到红球次数的方差为$\dfrac{4}{3}$；

③从中不放回地取球 2 次，每次任取 1 球，则在第一次取到红球后，第二次再次取到红球的概率为$\dfrac{2}{5}$；

④从中有放回地取球 3 次，每次任取 1 球，则至少有 1 次取到红球的概率为$\dfrac{26}{27}$.

其中所有正确结论的序号是________.

①②④　**解析**：①恰有 1 个白球的概率

$p=\dfrac{C_2^1C_4^2}{C_6^3}=\dfrac{3}{5}$，故①正确；

②每次任取 1 球，设取到红球的次数为 X，则 $X\sim B\left(6,\dfrac{2}{3}\right)$，其方差为 $6\times\dfrac{2}{3}\times\left(1-\dfrac{2}{3}\right)=\dfrac{4}{3}$，故②正确；

③设 $A=$“第一次取到红球”，$B=$“第二次取到红球”，

则 $P(A)=\dfrac{2}{3}$，$P(AB)=\dfrac{4\times3}{6\times5}=\dfrac{2}{5}$，

所以 $P(B|A)=\dfrac{P(AB)}{P(A)}=\dfrac{3}{5}$，故③错误；

④每次取到红球的概率都为$\dfrac{2}{3}$，

所以至少有 1 次取到红球的概率为 $1-\left(1-\dfrac{2}{3}\right)^3=\dfrac{26}{27}$，故④正确.

四、解答题(本题共 5 小题，共 77 分)

15.(13 分)某校高中部，高一年级有 6 个班，高二年级有 7 个班，高三年级有 8 个班，学校利用星期六组织学生到某厂参加社会实践活动.

(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 $6\times7\times8=336$(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 6×7 种不同的选法;第 2 类,从高一、高三两个年级各选 1 个班,有 6×8 种不同的选法;第 3 类,从高二、高三年级各选 1 个班,有 7×8 种不同的选法.故共有 $6\times7+6\times8+7\times8=146$(种)不同的选法.

16.(15 分)某市开展“寻找身边的好老师”活动,通过市民投票评选“身边的好老师”,并对选出的五位“身边的好老师”的工作年限和得票数进行了统计,得到如下数据:

“身边的好老师”编号	1	2	3	4	5
工作年限 x/年	4	6	8	10	12
得票数 y/百张	10	20	40	60	50

(1)若得票数 y 与工作年限 x 满足线性相关关系,试求经验回归方程 $\hat{y}=\hat{b}x+\hat{a}$,并就此估计身边的“好老师”的工作年限为 15 年时的得票数;

(2)若用 $\dfrac{y_i}{x_i}(i=1,2,3,4,5)$ 表示统计数据时得票数的“即时均值”(四舍五入到整数),从 5 个“即时均值”中任选 2 个,求这 2 个数据之和小于 8 的概率.

解:(1)由题可得 $\overline{x}=8,\overline{y}=36$,

则 $\hat{b}=\dfrac{40+120+320+600+600-5\times8\times36}{16+36+64+100+144-5\times64}=6$,

$\hat{a}=36-48=-12$.所以 $\hat{y}=6x-12$.

当 $x=15$ 时,$\hat{y}=6\times15-12=78$.

(2)5 个“即时均值”分别为 3,3,5,6,4.

从 5 个“即时均值”中任选 2 个,共有 $C_5^2=10$(种)情况,其中 2 个数据之和小于 8 的有(3,3),(3,4),(3,4),共 3 种情况,所以这 2 个数据之和小于 8 的概率为 $\dfrac{3}{10}$.

17.(15 分)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.表中给出一些动物体重与脉搏率对应的数据,图 1 是体重 W 与脉搏率 f 的散点图,图 2 是 $\lg W$ 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

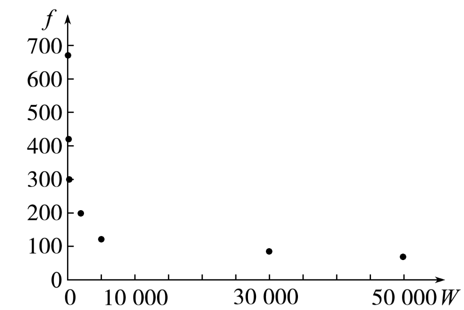


图 1

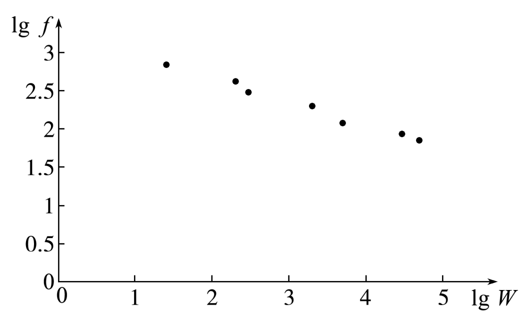


图 2

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$\lg f=k\lg W+b$.

其中,$k,b\in\mathbf{R}$.

(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:根据散点图的特征,题图 2 中的散点基本集中在一条直线附近,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为 W_1，f_1，兔的体重和脉搏率为 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 $f_1=50$.

即马的脉搏率为 50.

18.(17 分)某种可能遭受污染的海产品在进入市场前必须对每件海产品进行两轮检测，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)如果该海产品可以销售，则每件产品可获利 40 元；如果该海产品不能销售，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 ξ 元，求 ξ 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 ξ 的所有可能取值为 -320，-200，-80，40，160.

$P(\xi=-320)=\left(\dfrac{1}{4}\right)^4=\dfrac{1}{256}$，

$P(\xi=-200)=\mathrm{C}_4^1\times\left(\dfrac{1}{4}\right)^3\times\dfrac{3}{4}=\dfrac{3}{64}$，

$P(\xi=-80)=\mathrm{C}_4^2\times\left(\dfrac{1}{4}\right)^2\times\left(\dfrac{3}{4}\right)^2=\dfrac{27}{128}$，

$P(\xi=40)=\mathrm{C}_4^3\times\dfrac{1}{4}\times\left(\dfrac{3}{4}\right)^3=\dfrac{27}{64}$，

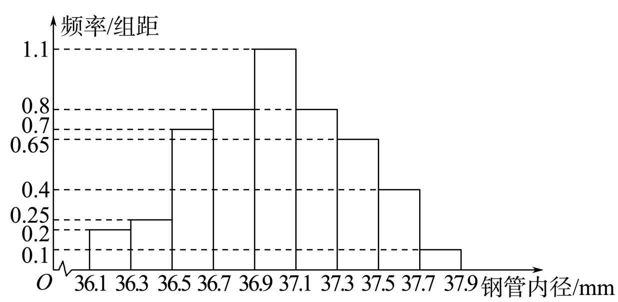
$P(\xi=160)=\left(\dfrac{3}{4}\right)^4=\dfrac{81}{256}$.

所以 ξ 的分布列为

ξ	-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.(17 分)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 μ 的近似值，用样本标准差 s 作为 σ 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 $n(n\geqslant2$，$n\in\mathbf{N})$ 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

$f(p)=C_5^3p^3(1-p)^2=10p^3(1-2p+p^2)=10(p^3-2p^4+p^5)$,

$f'(p)=10(3p^2-8p^3+5p^4)=10p^2(3-8p+5p^2)=10p^2(p-1)(5p-3)$,

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 $\dfrac{216}{625}$.

模块综合检测(二)

（时间：120 分钟，分值：150 分）

一、单项选择题(本题共 8 小题，每小题 5 分，共40 分)

1. 已知 y 关于 x 的经验回归方程 $\hat{y}=5x+1$，则该方程在样本点(1,4)处的残差为　（A）

A.-2　B.1　C.2　D.5

2. 一个礼堂有 4 个门，若从任一门进，从任一门出，则不同的进出方法共有　（C）

A.8 种　B.12 种

C.16 种　D.24 种

3. 已知随机变量 X 服从正态分布 $N(4,3)$，且 $P(X<a-5)=P(X>a+1)$，则实数 a 等于　（　）

A.7　B.6

C.5　D.4

B　**解析**：因为随机变量 X 服从正态分布 $N(4,3)$，$\mu=4$，所以其正态曲线的对称轴是直线 $x=4$.

所以 $\frac{a-5+a+1}{2}=4$，解得 $a=6$.

4. 在一批型号相同的产品中，有 2 件次品、5 件正品，每次随机抽取 1 件测试，直到将 2 件次品全部区分出来为止.假定抽取后不放回，则第 5 次测试后停止的概率是　（　）

A.$\frac{1}{21}$　B.$\frac{5}{21}$

C.$\frac{10}{21}$　D.$\frac{20}{21}$

B　**解析**：由题意知，前 4 次抽取的有 3 件正品、1 件次品，或前 5 次抽取的都是正品，故第 5 次测试后停止的概率为 $\frac{C_2^1C_5^3A_4^4}{A_7^5}+\frac{A_5^5}{A_7^5}=\frac{5}{21}$.

5. 若 $\left(\sqrt{x}-\frac{2}{x^2}\right)^n$ 的展开式中只有第 6 项的二项式系数最大，则展开式中的常数项是　（　）

A.210　B.180

C.160　D.175

B　**解析**：因为 $\left(\sqrt{x}-\frac{2}{x^2}\right)^n$ 的展开式中只有第 6 项的二项式系数最大，所以展开式中共有 11 项，则 $n=10$.

所以展开式的通项为 $T_{k+1}=C_{10}^k\cdot(\sqrt{x})^{10-k}\cdot\left(-\frac{2}{x^2}\right)^k=C_{10}^k(-2)^k\cdot x^{5-\frac{5k}{2}}$.令 $5-\frac{5k}{2}=0$，得 $k=2$.所以常数项是 $T_3=2^2\times C_{10}^2=180$.故选 B.

6. 某商家开展促销活动，促销方案是顾客每消费 1 000 元，便可以获得奖券 1 张，每张奖券中奖的概率为 $\frac{1}{5}$.若中奖，则商家返还中奖的顾客 1 000 元.小王购买一套价格为 2 400 元的西服，只能得到 2 张奖券，于是小王补偿 50 元让一同事购买一件价格为 600 元的便服，这样小王就得到了 3 张奖券.设小王这次消费的实际支出为 ξ，则 $E(\xi)=$　（　）

A.1 850　B.1 720

C.1 560　D.1 480

A　**解析**：根据题意知，ξ 的可能取值为 2 450，1 450，450，-550，且 $P(\xi=2\ 450)=\left(\frac{4}{5}\right)^3=\frac{64}{125}$，$P(\xi=1\ 450)=C_3^1\times\frac{1}{5}\times\left(\frac{4}{5}\right)^2=\frac{48}{125}$，$P(\xi=450)=C_3^2\times\left(\frac{1}{5}\right)^2\times\frac{4}{5}=\frac{12}{125}$，$P(\xi=-550)=\left(\frac{1}{5}\right)^3=\frac{1}{125}$，

故 $E(\xi)=2\ 450\times\frac{64}{125}+1\ 450\times\frac{48}{125}+450\times\frac{12}{125}+(-550)\times\frac{1}{125}=1\ 850$.

7. 先后掷两次质地均匀的骰子，落在水平桌面后，记正面朝上的点数分别为 x，y，设事件 A＝“$x+y$ 为奇数”，事件 B＝“x，y 满足 $x+y<6$”，则 $P(B|A)=$　（　）

A.$\frac{1}{2}$　B.$\frac{1}{3}$

C.$\frac{2}{5}$　D.$\frac{3}{5}$

B　**解析**：掷两次骰子，总的样本点的个数为 $6\times6=36$.

因为事件 A＝“$x+y$ 为奇数”，事件 B＝“x，y 满足

$x+y<6$”，所以事件 A 包含的样本点的个数为 $3\times3\times2=18$，事件 AB 包含的样本点的个数为 6.

由古典概型概率公式，

得 $P(A)=\frac{1}{2}$，$P(AB)=\frac{1}{6}$.

由条件概率公式，得 $P(B|A)=\frac{P(AB)}{P(A)}=\frac{\frac{1}{6}}{\frac{1}{2}}=\frac{1}{3}$.

故选 B.

8. 为了研究男性的年龄与吸烟的关系，抽查了 100 名男性，按年龄超过 40 岁和不超过 40 岁，吸烟量每天多于 20 支和不多于 20 支进行分组，得到如下 2×2 列联表：

吸烟量	年龄		合计
	不超过40岁	超过40岁	
不多于20支/天	50	15	65
多于20支/天	10	25	35
合计	60	40	100

则认为吸烟量与年龄有关犯错误的概率不超过（　　）

A. 0.001　　B. 0.01　　C. 0.05　　D. 0.005

A　**解析**：利用题中列联表，代入公式计算可得

$\chi^2=\frac{100\times(50\times25-15\times10)^2}{65\times35\times60\times40}\approx22.16>10.828=x_{0.001}$.

所以在犯错误的概率不超过0.001的前提下认为吸烟量与年龄有关.

二、多项选择题（本题共 3 小题，每小题 6 分，共 18 分）

9. 若随机变量 X 服从两点分布，$P(X=0)=\frac{1}{3}$，$E(X)$，$D(X)$分别为随机变量 X 的均值与方差，则（　　）

A. $P(X=1)=E(X)$

B. $E(3X+2)=4$

C. $D(3X+2)=4$

D. $D(X)=\frac{4}{9}$

AB　**解析**：因为随机变量 X 服从两点分布，$P(X=0)=\frac{1}{3}$，所以 $P(X=1)=\frac{2}{3}$，$E(X)=0\times\frac{1}{3}+1\times\frac{2}{3}=\frac{2}{3}$，$D(X)=\left(0-\frac{2}{3}\right)^2\times\frac{1}{3}+\left(1-\frac{2}{3}\right)^2\times\frac{2}{3}=\frac{2}{9}$.

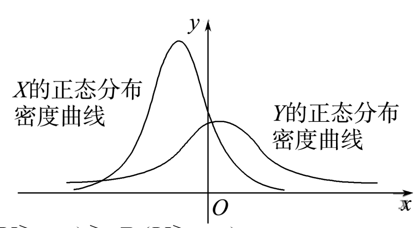
在 A 中，$P(X=1)=E(X)$，故 A 正确；

在 B 中，$E(3X+2)=3E(X)+2=3\times\frac{2}{3}+2=4$，故 B 正确；

在 C 中，$D(3X+2)=9D(X)=9\times\frac{2}{9}=2$，故 C 错误；

D 显然错误.故选 AB.

10. 设 $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，且两个正态分布密度曲线如图所示.下列结论中错误的是（　　）



A. $P(Y\geqslant\mu_2)\geqslant P(Y\geqslant\mu_1)$

B. $P(X\leqslant\sigma_2)\leqslant P(X\leqslant\sigma_1)$

C. 对任意正数 t，$P(X\leqslant t)>P(Y\leqslant t)$

D. 对任意正数 t，$P(X>t)>P(Y>t)$

ABD　**解析**：由题图可知 $\mu_1<0<\mu_2$，$\sigma_1^2<\sigma_2^2$，

所以 $P(Y\geqslant\mu_2)<P(Y\geqslant\mu_1)$，

$P(X\leqslant\sigma_2)>P(X\leqslant\sigma_1)$，故 A，B 错误；

当 t 为任意正数时，由题图可知 $P(X\leqslant t)>P(Y\leqslant t)$，

而 $P(X\leqslant t)=1-P(X>t)$，$P(Y\leqslant t)=1-P(Y>t)$，

所以 $P(X>t)<P(Y>t)$，故 C 正确，D 错误.

故选 ABD.

11. 甲罐中有 5 个红球、2 个白球和 3 个黑球，乙罐中有 4 个红球、3 个白球和 3 个黑球.先从甲罐中随机取出一球放入乙罐，分别以事件 A_1，A_2 和 A_3 表示由甲罐取出的球是红球、白球和黑球；再从乙

罐中随机取出一球，以事件 B 表示由乙罐取出的球是红球.下列结论中正确的是 ( )

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

B.$P(B|A_1)=\frac{5}{11}$

C.事件 B 与事件 A_1 相互独立

D.A_1,A_2,A_3 是两两互斥的事件

BD **解析**：由题意知 $P(A_1)=\frac{5}{10}=\frac{1}{2}$，$P(A_2)=\frac{2}{10}=\frac{1}{5}$，$P(A_3)=\frac{3}{10}$，$P(B|A_1)=\frac{5}{11}$，故 B 正确；$P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+P(A_3)P(B|A_3)=\frac{5}{10}\times\frac{5}{11}+\frac{2}{10}\times\frac{4}{11}+\frac{3}{10}\times\frac{4}{11}=\frac{9}{22}$，故 A，C 不正确；$A_1,A_2,A_3$ 是两两互斥的事件，故 D 正确.故选 BD.

三、填空题（本题共 3 小题，每小题 5 分，共 15 分）

12.甲、乙两名选手进行围棋比赛，每局比赛，甲选手获胜的概率为 $\frac{3}{4}$，乙选手获胜的概率为 $\frac{1}{4}$.有如下两种方案：方案一，三局两胜；方案二，五局三胜.对于乙选手，获胜概率更大的方案是________.

方案一 **解析**：（方案一）设"乙获胜"为事件 A，则事件 A 应包括以下两种情况：

①乙 2∶0 获胜（设为事件 A_1）；②乙 2∶1 获胜（设为事件 A_2）.

这两种情况彼此互斥，根据互斥事件的概率计算公式得

$$P(A)=P(A_1)+P(A_2)=\left(\frac{1}{4}\right)^2+C_2^1\times\frac{1}{4}\times\frac{3}{4}\times\frac{1}{4}=\frac{5}{32}.$$

（方案二）设乙获胜为事件 B，则事件 B 应包括以下三种情况：

①乙 3∶0 获胜（设为事件 B_1）；②乙 3∶1 获胜（设为事件 B_2）；③乙 3∶2 获胜（设为事件 B_3）.

这三种情况两两互斥，根据互斥事件的概率计算公式得

$$P(B)=P(B_1)+P(B_2)+P(B_3)=\left(\frac{1}{4}\right)^3+C_3^2\times\left(\frac{1}{4}\right)^2\times\frac{3}{4}\times\frac{1}{4}+C_4^2\times\left(\frac{1}{4}\right)^2\times\left(\frac{3}{4}\right)^2\times\frac{1}{4}=\frac{53}{512}.$$

因为 $\frac{5}{32}>\frac{53}{512}$，

所以乙获胜概率更大的方案是方案一.

13.在 $(1+3x)(2x-1)^5$ 的展开式中，若按 x 的升幂进行排列，则第 3 项为________.

$-10x^2$ **解析**：由 $(1+3x)(2x-1)^5=(2x-1)^5+3x\cdot(2x-1)^5$，可得展开式中按 x 的升幂进行排列依次为常数项、含 x 的项、含 x^2 的项，…，故第 3 项为含 x^2 的项.

所以第 3 项为 $C_5^3(2x)^2(-1)^3+3x\cdot C_5^4(2x)^1(-1)^4=-40x^2+30x^2=-10x^2$.

14.用 0，1，2，3，4 这五个数字组成无重复数字的自然数.

(1)在组成的三位数中，若十位上的数字比百位上的数字及个位上的数字都大，则称这个数为"凸数"，如 231，243 等都是"凸数"，则"凸数"有________个；

(2)在组成的五位数中，各数位上恰有一个偶数夹在两个奇数之间的有________个.

(1)14 (2)28 **解析**：(1)将这些"凸数"分为三类，且百位不能为 0：

若十位数字为 2，则只有 120 这 1 个"凸数"；

若十位数字为 3，则共有 $C_2^1\cdot C_2^1=4$（个）"凸数"；

若十位数字为 4，则共有 $C_3^1\cdot C_3^1=9$（个）"凸数".

所以共有 14 个符合题意的"凸数".

(2)将符合题意的五位数分为三类：

若两个奇数在万位和百位上，则共有 $A_2^2\cdot A_3^3=12$（个）；

若两个奇数在千位和十位上，则共有 $A_2^2\cdot C_2^1\cdot A_2^2=8$（个）；

若两个奇数在百位和个位上，则共有 $A_2^2\cdot C_2^1\cdot A_2^2$

$=8$(个).

所以共有 28 个符合题意的五位数.

四、解答题(本题共 5 小题,共 77 分)

15.(13 分)已知 $f(x)=(1-x)^{2\,020}=a_0+a_1x+a_2x^2+\cdots+a_{2\,020}x^{2\,020}$.

(1)求 $a_1+a_2+a_3+\cdots+a_{2\,020}$ 的值;

(2)求 $a_1+2a_2+3a_3+\cdots+2\,020a_{2\,020}$ 的值.

解:(1)因为 $f(0)=1=a_0$,

$f(1)=0=a_0+a_1+a_2+\cdots+a_{2\,020}$,

所以 $a_1+a_2+a_3+\cdots+a_{2\,020}=f(1)-f(0)=-1$.

(2)$f'(x)=-2\,020(1-x)^{2\,019}=a_1+2a_2x+3a_3x^2+\cdots+2\,020a_{2\,020}x^{2\,019}$,

所以 $a_1+2a_2+3a_3+\cdots+2\,020a_{2\,020}=f'(1)=0$.

16.(15 分)(2022·全国甲卷(理))甲、乙两个学校进行体育比赛,比赛共设三个项目,每个项目胜方得 10 分,负方得 0 分,没有平局.三个项目比赛结束后,总得分高的学校获得冠军.已知甲学校在三个项目中获胜的概率分别为 0.5,0.4,0.8,各项目的比赛结果相互独立.

(1)求甲学校获得冠军的概率;

(2)用 X 表示乙学校的总得分,求 X 的分布列与期望.

解:(1)设甲学校在三个项目中获胜的事件依次为 A,B,C,

所以甲学校获得冠军的概率为

$P=P(ABC)+P(\bar{A}BC)+P(A\bar{B}C)+P(AB\bar{C})$

$=0.5\times0.4\times0.8+0.5\times0.4\times0.8+0.5\times0.6\times0.8+0.5\times0.4\times0.2=0.16+0.16+0.24+0.04=0.6$.

(2)依题可知,X 的可能取值为 0,10,20,30,

且 $P(X=0)=0.5\times0.4\times0.8=0.16$,

$P(X=10)=0.5\times0.4\times0.8+0.5\times0.6\times0.8+0.5\times0.4\times0.2=0.44$,

$P(X=20)=0.5\times0.6\times0.8+0.5\times0.4\times0.2+0.5\times0.6\times0.2=0.34$,

$P(X=30)=0.5\times0.6\times0.2=0.06$.

所以 X 的分布列为

X	0	10	20	30
P	0.16	0.44	0.34	0.06

期望 $E(X)=0\times0.16+10\times0.44+20\times0.34+30\times0.06=13$.

17.(15 分)一种药用昆虫的产卵数 y(单位:个)在一定范围内与温度 x(单位:℃)有关,现收集了该种药用昆虫的 6 对观测数据如下表:

x	21	23	24	27	29	32
y	6	11	20	27	57	77

(1)若用线性回归模型来拟合 x 与 y 的关系,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 0.1).

(2)若用非线性回归模型来拟合 x 与 y 的关系,得到 y 关于 x 的经验回归方程为 $\hat{y}=0.06e^{0.230\,3x}$,且决定系数 $R^2=0.952\,2$.

①试与(1)中的线性回归模型相比,用 R^2 说明哪种模型的拟合效果更好.

②用拟合效果更好的模型预测温度为 35 ℃ 时该种药用昆虫的产卵数(结果取整数).

附:经验回归直线 $\hat{y}=\hat{b}x+\hat{a}$ 的斜率和截距的最小二乘估计分别为 $\hat{b}=\dfrac{\sum\limits_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum\limits_{i=1}^{n}(x_i-\bar{x})^2}$,$\hat{a}=\bar{y}-\hat{b}\,\bar{x}$,决定系数 $R^2=1-\dfrac{\sum\limits_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum\limits_{i=1}^{n}(y_i-\bar{y})^2}$,$\sum\limits_{i=1}^{6}(x_i-\bar{x})\cdot(y_i-\bar{y})=557$,$\sum\limits_{i=1}^{6}(y_i-\bar{y})^2=3\,930$,$e^{8.060\,5}\approx3\,167$.

解:(1)由题意得 $\bar{x}=\dfrac{21+23+24+27+29+32}{6}=26$,

$\bar{y}=\dfrac{6+11+20+27+57+77}{6}=33$,

$\sum_{i=1}^{6}(x_i-\overline{x})(y_i-\overline{y})=557$，$\sum_{i=1}^{6}(x_i-\overline{x})^2=84$，

所以 $\hat{b}=\dfrac{\sum_{i=1}^{6}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{6}(x_i-\overline{x})^2}=\dfrac{557}{84}\approx 6.6$，

$\hat{a}=33-6.6\times 26=-138.6$.

所以 y 关于 x 的经验回归方程为 $\hat{y}=6.6x-138.6$.

(2)①对于线性回归模型 $\hat{y}=6.6x-138.6$，

易得 $\sum_{i=1}^{6}(y_i-\overline{y})^2=3\ 930$，

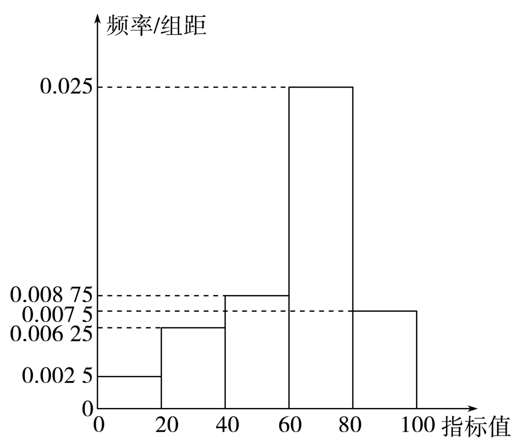
故其决定系数 $R^2=1-\dfrac{\sum_{i=1}^{6}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{6}(y_i-\overline{y})^2}=1-\dfrac{236.64}{3\ 930}\approx 1-0.060\ 2=0.939\ 8$.

因为 $0.939\ 8<0.952\ 2$，所以非线性回归模型 $\hat{y}=0.06e^{0.230\ 3x}$ 比线性回归模型 $\hat{y}=6.6x-138.6$ 拟合效果更好.

②由①得，当 $x=35$ 时，$\hat{y}=0.06e^{0.230\ 3\times 35}=0.06e^{8.060\ 5}\approx 0.06\times 3\ 167\approx 190$.

所以当温度为 35 ℃时，预测该种药用昆虫的产卵数约为 190 个.

18.(17 分)为了检测某种抗病毒疫苗的效果，需要进行动物与人体试验.研究人员将疫苗注射到 200 只小白鼠体内，一段时间后测量小白鼠的某项指标值，按[0,20)，[20,40)，[40,60)，[60,80)，[80,100)分组，绘制频率分布直方图如图所示.试验发现小白鼠体内产生抗体的共有 160 只，其中该项指标值不小于 60 的有 110 只.假设每只小白鼠注射疫苗后是否产生抗体相互独立.



(1)填写下面的 2×2 列联表，并根据 $\alpha=0.05$ 的独立性检验，判断能否认为注射疫苗后小白鼠是否产生抗体与指标值大小有关.

单位：只

抗体	指标值		合计
	小于 60	不小于 60	
有抗体			
没有抗体			
合计			

(2)对第一次注射疫苗后没有产生抗体的 40 只小白鼠进行第二次注射，结果又有 20 只小白鼠产生抗体.

①用频率估计概率，求一只小白鼠注射两次疫苗后产生抗体的概率 p.

②以①中确定的概率 p 作为人体注射两次疫苗后产生抗体的概率，进行人体接种试验，记 n 个人注射两次疫苗后产生抗体的人数为随机变量 X.试验后统计数据显示，当 $k=90$ 时，$P(X=k)(k=0,1,2,\cdots)$取最大值，求参加人体接种试验的人数 n 及 $E(X)$.

附：$\chi^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$，其中 $n=a+b+c+d$.

α	0.1	0.05	0.01	0.005	0.001
x_α	2.706	3.841	6.635	7.879	10.828

解：(1)由题图可知，200 只小白鼠按指标值分布为：

在[0,20)内有 $0.002\ 5\times 20\times 200=10$(只)；

在[20,40)内有 $0.006\ 25\times 20\times 200=25$(只)；

在[40,60)内有 $0.008\ 75\times 20\times 200=35$(只)；

在[60,80)内有 $0.025\times 20\times 200=100$(只)；

在[80,100)内有 $0.007\ 5\times 20\times 200=30$(只).

由题意，有抗体且指标值小于 60 的有 50 只；而指标值小于 60 的小白鼠共有 $10+25+35=70$(只)，所以指标值小于 60 且没有抗体的小白鼠有 20 只.同理，指标值不小于 60 且没有抗体的小白鼠有

20 只，

故 2×2 列联表如下：

单位：只

抗体	指标值		合计
	小于 60	不小于 60	
有抗体	50	110	160
没有抗体	20	20	40
合计	70	130	200

零假设为 H_0：注射疫苗后小白鼠是否产生抗体与指标值大小无关联.

根据列联表中数据，得

$\chi^2=\dfrac{200\times(50\times20-110\times20)^2}{160\times40\times70\times130}\approx4.945>3.841=x_{0.05}$.

根据小概率值 $\alpha=0.05$ 的独立性检验，推断 H_0 不成立，即认为注射疫苗后小白鼠是否产生抗体与指标值大小有关，此推断犯错误的概率不大于 0.05.

(2)①设事件 A＝"小白鼠第一次注射疫苗产生抗体"，事件 B＝"小白鼠第二次注射疫苗产生抗体"，事件 C＝"小白鼠注射两次疫苗后产生抗体".

则 $P(A)=\dfrac{160}{200}=0.8$，$P(B)=\dfrac{20}{40}=0.5$，$P(C)=1-P(\overline{A})P(\overline{B})=1-0.2\times0.5=0.9$.

所以一只小白鼠注射两次疫苗后产生抗体的概率 $p=0.9$.

②由题意知随机变量 $X\sim B(n,0.9)$，

$P(X=k)=C_n^k\times0.9^k\times0.1^{n-k}(k=0,1,2,\cdots,n)$.

因为 $P(X=90)$ 最大，

所以 $\begin{cases}C_n^{90}\times0.9^{90}\times0.1^{n-90}\geqslant C_n^{91}\times0.9^{91}\times0.1^{n-91},\\ C_n^{90}\times0.9^{90}\times0.1^{n-90}\geqslant C_n^{89}\times0.9^{89}\times0.1^{n-89},\end{cases}$

解得 $99\leqslant n\leqslant\dfrac{901}{9}$.

因为 n 是整数，所以 $n=99$ 或 $n=100$，

所以参加接种试验的人数为 99 或 100.

当接种人数为 99 时，$E(X)=np=99\times0.9=89.1$；

当接种人数为 100 时，$E(X)=np=100\times0.9=90$.

19.(17 分)第五届中国国际进口博览会(以下简称进博会)于 2022 年 11 月 5 日至 10 日在国家会展中心(上海)举办.本届进博会共有 284 家世界 500 强和行业龙头企业参展，数量超过上届，其中至少参展过两届及以上进博会的企业占比约为 90%.本届进博会首次运用虚拟现实、三维建模等新技术手段，引入了全新的线上展示技术，为观展者带来了不同以往的观展体验.活动结束后，进博会组委会从观展者中随机抽取 100 人(其中年龄在 50 周岁及以下的有 60 人)了解他们对全新的线上展示活动的满意度，并按年龄分类统计得到如下不完整的 2×2 列联表：

单位：人

年龄	满意度		合计
	不满意	满意	
50 周岁及以下		55	
50 周岁以上	15		
合计			100

(1)根据统计数据完成以上 2×2 列联表，根据小概率值 $\alpha=0.001$ 的独立性检验，能否认为对全新的线上展示活动是否满意与年龄有关联？

(2)从本届参展的 284 家世界 500 强和行业龙头企业中随机抽取 3 家了解他们对组委会的组织工作的满意度，设其中至少参展过两届及以上进博会的企业的个数为 X.以本届参展的世界 500 强和行业龙头企业中至少参展过两届及以上进博会的企业的频率为概率.

①求 X 的分布列和数学期望；

②求 $P(|X-1|\leqslant1)$的值.

附：$\chi^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$，其中 $n=a+b+c+d$.

α	0.1	0.05	0.01	0.001
x_α	2.706	3.841	6.635	10.828

解:(1)由题意知,抽取的100名观展者中年龄在50周岁及以下的有60人,

则年龄在50周岁以上的有40人,所以50周岁及以下不满意的有5人,50周岁以上满意的有25人,补全的2×2列联表如下:

单位:人

年龄	满意度		合计
	不满意	满意	
50周岁及以下	5	55	60
50周岁以上	15	25	40
合计	20	80	100

零假设为H_0:对全新的线上展示活动是否满意与年龄无关.

根据列联表中数据计算得$\chi^2=\frac{100\times(5\times25-55\times15)^2}{60\times40\times20\times80}\approx12.76>10.828=x_{0.001}$.

根据小概率值$\alpha=0.001$的独立性检验,推断H_0不成立,即认为对全新的线上展示活动是否满意与年龄有关联,此推断犯错误的概率不超过0.001.

(2)①由题意可得,一家参展企业至少参展过两届及以上进博会的概率为0.9,

则$X\sim B(3,0.9)$,X的所有可能取值为0,1,2,3,

且$P(X=0)=C_3^0\times0.1^3=0.001$,

$P(X=1)=C_3^1\times0.9\times0.1^2=0.027$,

$P(X=2)=C_3^2\times0.9^2\times0.1=0.243$,

$P(X=3)=C_3^3\times0.9^3=0.729$,

所以X的分布列为

X	0	1	2	3
P	0.001	0.027	0.243	0.729

所以$E(X)=0\times0.001+1\times0.027+2\times0.243+3\times0.729=2.7$.

②$P(|X-1|\leqslant1)=P(X=0)+P(X=1)+P(X=2)=1-P(X=3)=1-0.729=0.271$.