## 第2讲　概率、随机变量及其分布

[考情分析]　1.考查古典概型、互斥事件、相互独立事件、独立重复试验等内容，主要以选择题、填空题的形式出现，中低等难度.2.离散型随机变量的分布列、均值、方差和概率的计算问题常常结合在一起进行考查，中高等难度．

考点一　古典概型

核心提炼



古典概型的概率公式

*P*(*A*)＝＝.

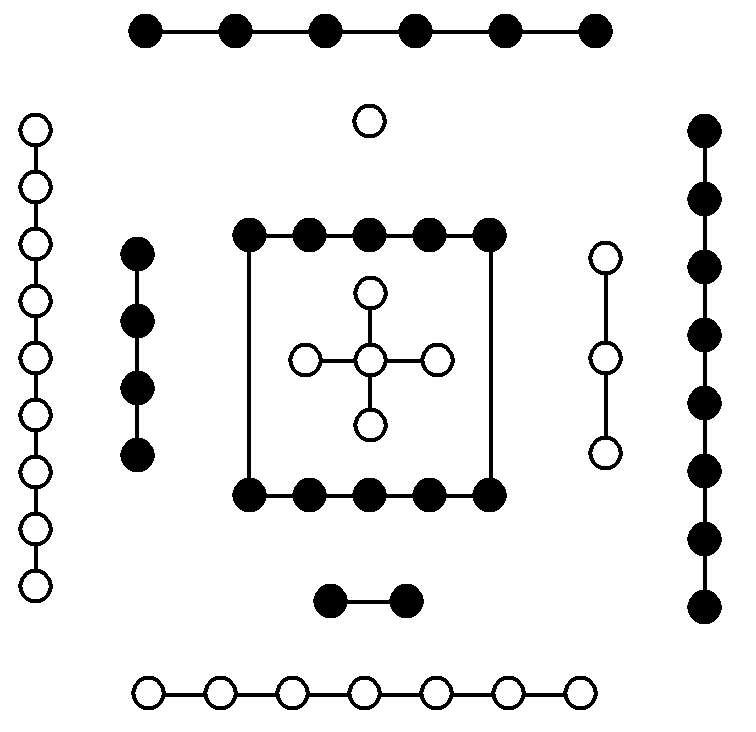
例1　(1)(2020·宁夏六盘山高级中学模拟)2020年春节突如其来的新型冠状病毒肺炎在某省爆发，一方有难八方支援，全国各地的白衣天使走上战场的第一线，某医院抽调甲、乙、丙三名医生，抽调*A*，*B*，*C*三名护士支援某市第一医院与第二医院，参加该市疫情狙击战．其中选一名护士与一名医生去第一医院，其他都在第二医院工作，则医生甲和护士*A*被选去第一医院工作的概率为(　　)

A. B. C. D.

答案　D

解析　根据题意，选一名护士与一名医生去第一医院，有：甲*A*，甲*B*，甲*C*，乙*A*，乙*B*，乙*C*，丙*A*，丙*B*，丙*C*，9种情况，而医生甲和护士*A*被选去第一医院工作有1种情况，所以概率为*P*＝.

(2)河图是上古时代神话传说中伏羲通过黄河中浮出龙马身上的图案，与自己的观察，画出的“八卦”，而龙马身上的图案就叫做“河图”．把一到十分成五组，如图，其口诀：一六共宗，为水居北；二七同道，为火居南；三八为朋，为木居东；四九为友，为金居西；五十同途，为土居中．现从这十个数中随机抽取四个数，则能成为两组的概率是(　　)



A. B. C. D.

答案　C

解析　现从这十个数中随机抽取4个数，基本事件总数*n*＝C，

能成为两组的基本事件个数*m*＝C，则能成为两组的概率是*P*＝＝＝.

规律方法　古典概型求解的关键点

(1)正确求出基本事件总数和所求事件包含的基本事件数，这常常用到排列、组合的有关知识．

(2)对于较复杂的题目计数时要正确分类，分类时应不重不漏．

跟踪演练1　(1)(2018·全国Ⅱ)我国数学家陈景润在哥德巴赫猜想的研究中取得了世界领先的成果．哥德巴赫猜想是“每个大于2的偶数可以表示为两个素数的和”，如30＝7＋23.在不超过30的素数中，随机选取两个不同的数，其和等于30的概率是(　　)

A. B. C. D.

答案　C

解析　不超过30的所有素数为2,3,5,7,11,13,17,19,23,29，共10个，随机选取两个不同的数，共有C＝45(种)情况，而和为30的有7＋23,11＋19,13＋17这3种情况，所以所求概率为＝.

(2)用0与1两个数字随机填入如图所示的5个格子里，每个格子填一个数字，并且从左到右数，不管数到哪个格子，总是1的个数不少于0的个数，则这样填法的概率为(　　)

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

A. B. C. D.

答案　B

解析　由题意可知，填写的可能结果共有如下32种：

00000,00001,00010,00011,00100,00101,00110,00111，

01000,01001,01010,01011,01100,01101,01110,01111，

10000,10001,10010,10011,10100,10101,10110,10111，

11000,11001,11010,11011,11100,11101,11110,11111，

其中满足题意的有10种：10101,10110,10111,11001,11010,11011,11100,11101,11110,11111，由古典概型概率计算公式可得满足题意的概率值*P*＝＝.

考点二　随机变量的分布列

核心提炼



1．超几何分布

在含有*M*件次品的*N*件产品中，任取*n*件，其中恰有*X*件次品，则事件{*X*＝*k*}发生的概率*P*(*X*＝*k*)＝，*k*＝0,1,2，…，*m*，其中*m*＝min{*M*，*n*}，且*n*≤*N*，*M*≤*N*，*n*，*M*，*N*∈**N**\*.

2．二项分布

一般地，在*n*次独立重复试验中，用*X*表示事件*A*发生的次数，设每次试验中事件*A*发生的概率为*p*，则

*P*(*X*＝*k*)＝C*pk*(1－*p*)*n*－*k*，*k*＝0,1,2，…，*n*.

考向一　超几何分布

例2　(2020·天津市滨海新区塘沽第一中学模拟)4月23日是“世界读书日”，某中学开展了一系列的读书教育活动．学校为了解高三学生课外阅读情况，采用分层抽样的方法从高三某班甲、乙、丙、丁四个读书小组(每名学生只能参加一个读书小组)学生中抽取12名学生参加问卷调查．各组人数统计如下：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 小组 | 甲 | 乙 | 丙 | 丁 |
| 人数 | 12 | 9 | 6 | 9 |

(1)从参加问卷调查的12名学生中随机抽取2人，求这2人来自同一个小组的概率；

(2)从已抽取的甲、丙两个小组的学生中随机抽取2个，用*X*表示抽得甲组学生的人数，求随机变量*X*的分布列和均值．

解　(1)由题意得，问卷调查从四个小组中抽取的人数分别为4,3,2,3，从参加问卷调查的12名学生中随机抽取两人的取法共有C＝66(种)，抽取的两名学生来自同一小组的取法共有C＋2C＋C＝13(种)，

所以，抽取的两名学生来自同一个小组的概率为*P*＝.

(2)由(1)知，在参加问卷调查的12名学生中，来自甲、丙两小组的学生人数分别为4,2，所以抽取的两个人中是甲组学生的人数*X*的可能取值为0,1,2，

因为*P*(*X*＝0)＝＝，

*P*(*X*＝1)＝＝，

*P*(*X*＝2)＝＝.

所以随机变量*X*的分布列为

|  |  |  |  |
| --- | --- | --- | --- |
| *X* | 0 | 1 | 2 |
| *P* |  |  |  |

所以随机变量*X*的均值为*E*(*X*)＝0×＋1×＋2×＝.

跟踪演练2　PM2.5是指悬浮在空气中的空气动力学当量直径小于或等于2.5微米的可入肺颗粒物．根据现行国家标准GB3095－2012，PM2.5日均值在35微克/立方米以下空气质量为一级；在35微克/立方米～75微克/立方米之间空气质量为二级；在75微克/立方米以上空气质量为超标．从某自然保护区2018年全年每天的PM2.5监测数据中随机地抽取10天的数据作为样本，监测值频数如下表所示：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| PM2.5日均值(微克/立方米) | [25,35) | [35,45) | [45,55) | [55,65) | [65,75) | [75,85] |
| 频数 | 3 | 1 | 1 | 1 | 1 | 3 |

(1)从这10天的PM2.5日均值监测数据中，随机抽出3天，求恰有一天空气质量达到一级的概率；

(2)从这10天的数据中任取3天数据，记*ξ*表示抽到PM2.5监测数据超标的天数，求*ξ*的分布列．

解　(1)记“从这10天的PM2.5日均值监测数据中，随机抽出3天，恰有一天空气质量达到一级”为事件*A*，

则*P*(*A*)＝＝.

(2)由条件知，*ξ*服从超几何分布，其中*N*＝10，*M*＝3，*n*＝3，且随机变量*ξ*的可能取值为0,1,2,3.

*P*(*ξ*＝*k*)＝(*k*＝0,1,2,3)．

∴*P*(*ξ*＝0)＝＝，*P*(*ξ*＝1)＝＝，

*P*(*ξ*＝2)＝＝，*P*(*ξ*＝3)＝＝.

故*ξ*的分布列为

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *ξ* | 0 | 1 | 2 | 3 |
| *P* |  |  |  |  |

考向二　二项分布

例3　(2020·陕西安康中学模拟)“互联网＋”是“智慧城市”的重要内容，*A*市在智慧城市的建设中，为方便市民使用互联网，在主城区覆盖了免费WiFi，为了解免费WiFi在*A*市的使用情况，调查机构借助网络进行了问卷调查，并从参与调查的网友中抽取了200人进行抽样分析，得到如下列联表(单位：人)：

|  |  |  |  |
| --- | --- | --- | --- |
|  | 经常使用免费WiFi | 偶尔或不用免费WiFi | 总计 |
| 45岁及以下 | 70 | 30 | 100 |
| 45岁以上 | 60 | 40 | 100 |
| 总计 | 130 | 70 | 200 |

(1)根据以上数据，判断是否有90%的把握认为*A*市使用免费WiFi的情况与年龄有关；

(2)将频率视为概率，现从该市45岁以上的市民中用随机抽样的方法每次抽取1人，共抽取3次．记被抽取3人中“偶尔或不用免费WiFi”的人数为*X*，若每次抽取的结果是相互独立的，求*X*的分布列、均值*E*(*X*)和方差*D*(*X*)．

附：*K*2＝，其中*n*＝*a*＋*b*＋*c*＋*d*.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *P*(*K*2≥*k*0) | 0.15 | 0.10 | 0.05 | 0.025 |
| *k*0 | 2.072 | 2.706 | 3.841 | 5.024 |

解　(1)由列联表可知*K*2＝≈2.198，

因为2.198<2.706，所以没有90%的把握认为*A*市使用免费WiFi的情况与年龄有关．

(2)由题意可知*X*～*B*，*X*的所有可能取值为0,1,2,3，

*P*(*X*＝0)＝C×3＝，

*P*(*X*＝1)＝C××2＝，

*P*(*X*＝2)＝C×2×＝，

*P*(*X*＝3)＝C×3＝.

所以*X*的分布列为

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *X* | 0 | 1 | 2 | 3 |
| *P* |  |  |  |  |

*E*(*X*)＝3×＝，*D*(*X*)＝3××＝.

规律方法　随机变量分布列问题的两个关键点

(1)求离散型随机变量分布列的关键是正确理解随机变量取每一个值所表示的具体事件，然后综合应用各类概率公式求概率．

(2)求随机变量均值与方差的关键是正确求出随机变量的分布列，若随机变量服从二项分布，则可直接使用公式求解．

跟踪演练3　某机器生产商对一次性购买2台机器的客户推出2种超过质保期后2年内的延保维修方案：

方案一：交纳延保金6 000元，在延保的2年内一共可免费维修2次，超过2次每次收取维修费1 500元；

方案二：交纳延保金7 740元，在延保的2年内一共可免费维修4次，超过4次每次收取维修费*a*元．

某工厂准备一次性购买2台这种机器，现需决策在购买机器时应购买哪种延保方案，为此搜集并整理了100台这种机器超过质保期后2年内维修的次数，统计得下表：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 维修次数 | 0 | 1 | 2 | 3 |
| 机器台数 | 20 | 10 | 40 | 30 |

以这100台机器维修次数的频率代替1台机器维修次数发生的概率，记*X*表示这2台机器超过质保期后2年内共需维修的次数．

(1)求*X*的分布列；

(2)以所需延保金与维修费用之和的均值为决策依据，该工厂选择哪种延保方案更合算？

解　(1)*X*所有可能的取值为0,1,2,3,4,5,6，

*P*(*X*＝0)＝×＝，

*P*(*X*＝1)＝××2＝，

*P*(*X*＝2)＝×＋××2＝，

*P*(*X*＝3)＝××2＋××2＝，

*P*(*X*＝4)＝×＋××2＝，

*P*(*X*＝5)＝××2＝，

*P*(*X*＝6)＝×＝，

∴*X*的分布列为

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *X* | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| *P* |  |  |  |  |  |  |  |

(2)设选择方案一所需费用为*Y*1元，则*Y*1的分布列为

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Y*1 | 6 000 | 7 500 | 9 000 | 10 500 | 12 000 |
| *P* |  |  |  |  |  |

*E*(*Y*1)＝×6 000＋×7 500＋×9 000＋×10 500＋×12 000＝8 580.

设选择方案二所需费用为*Y*2元，则*Y*2的分布列为

|  |  |  |  |
| --- | --- | --- | --- |
| *Y*2 | 7 740 | 7 740＋*a* | 7 740＋2*a* |
| *P* |  |  |  |

*E*(*Y*2)＝×7 740＋×(7 740＋*a*)＋×(7 740＋2*a*)＝7 740＋.

当*E*(*Y*1)－*E*(*Y*2)＝840－>0，

即0<*a*<2 000时，选择方案二更合算，

当*E*(*Y*1)－*E*(*Y*2)＝840－＝0，

即*a*＝2 000时，选择方案一、方案二均可；

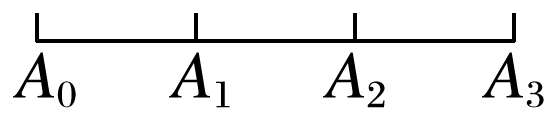
当*E*(*Y*1)－*E*(*Y*2)＝840－<0，

即*a*>2 000时，选择方案一更合算．

## 专题强化练

一、单项选择题

1.某路公交在某段路上有4个站点(如图)，分别记为*A*0，*A*1，*A*2，*A*3，现有甲、乙两人同时从*A*0站点上车，且他们中的每个人在站点*Ai*(*i*＝1,2,3)下车是等可能的，则甲、乙两人不在同一站点下车的概率为(　　)



A. B. C. D.

答案　A

解析　设事件*A*表示“甲、乙两人不在同一站点下车”．甲、乙两人同在站点*A*1下车的概率为×；甲、乙两人同在站点*A*2下车的概率为×；甲、乙两人同在站点*A*3下车的概率为×.所以甲、乙两人在同一站点下车的概率为3××＝，则*P*(*A*)＝1－＝.

2．设离散型随机变量*X*可能的取值为1,2,3,4，*P*(*X*＝*k*)＝*ak*＋*b*，若*X*的均值*E*(*X*)＝3，则*a*－*b*等于(　　)

A. B．0 C．－ D.

答案　A

解析　∵离散型随机变量*X*可能的取值为1,2,3,4，

*P*(*X*＝*k*)＝*ak*＋*b*，

∴(*a*＋*b*)＋(2*a*＋*b*)＋(3*a*＋*b*)＋(4*a*＋*b*)＝1，

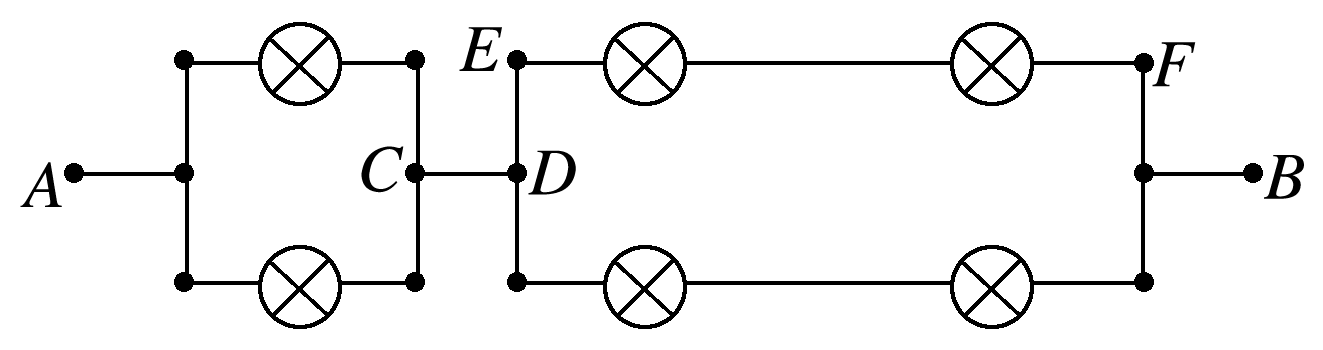
即10*a*＋4*b*＝1，

又*X*的均值*E*(*X*)＝3，

则(*a*＋*b*)＋2(2*a*＋*b*)＋3(3*a*＋*b*)＋4(4*a*＋*b*)＝3，

即30*a*＋10*b*＝3，∴*a*＝，*b*＝0，∴*a*－*b*＝.

3．如图，电路从*A*到*B*上共连接着6个灯泡，每个灯泡断路的概率是，整个电路的连通与否取决于灯泡是否断路，则从*A*到*B*连通的概率是(　　)



A. B. C. D.

答案　B

解析　由题图可知，*A*，*C*之间未连通的概率是2＝，连通的概率是1－＝.*E*，*F*之间连通的概率是2＝，未连通的概率是1－＝，故*D*，*B*之间未连通的概率是2＝，*D*，*B*之间连通的概率是1－＝，故*A*，*B*之间连通的概率是×＝.

4．先后掷两次骰子(骰子的六个面上分别有1,2,3,4,5,6六个点)，落在水平桌面后，记正面朝上的点数分别为*x*，*y*，记事件*A*为“*x*＋*y*为偶数”，事件*B*为“*x*，*y*中有偶数且*x*≠*y*”，则概率*P*(*B*|*A*)等于(　　)

A. B. C. D.

答案　B

解析　正面朝上的点数(*x*，*y*)的不同结果共有C·C＝36(种)，事件*A*：“*x*＋*y*为偶数”包含事件*A*1：“*x*，*y*都为偶数”与事件*A*2：“*x*，*y*都为奇数”两个互斥事件，其中*P*(*A*1)＝＝，*P*(*A*2)＝＝，所以*P*(*A*)＝*P*(*A*1)＋*P*(*A*2)＝＋＝.

事件*B*为“*x*，*y*中有偶数且*x*≠*y*”，所以事件*AB*为“*x*，*y*都为偶数且*x*≠*y*”，所以*P*(*AB*)＝＝，由条件概率的计算公式，得*P*(*B*|*A*)＝＝.

5．(2020·山东枣庄市八中月考)某校有1 000人参加某次模拟考试，其中数学考试成绩近似服从正态分布*N*(105，*σ*2)(*σ*>0)，试卷满分150分，统计结果显示数学成绩优秀(高于120分)的人数占总人数的，则此次数学考试成绩在90分到105分(含90分和105分)之间的人数约为(　　)

A．150 B．200

C．300 D．400

答案　C

解析　因为*P*(*X*<90)＝*P*(*X*>120)＝，

*P*(90≤*X*≤120)＝1－＝，

所以*P*(90≤*X*≤105)＝*P*(90≤*X*≤120)＝，

所以此次数学考试成绩在90分到105分(含90分和105分)之间的人数约为1 000×＝300.

6．某次考试共有12个选择题，每个选择题的分值为5分，每个选择题四个选项有且只有一个选项是正确的，*A*学生对12个选择题中每个题的四个选择项都没有把握，最后选择题的得分为*X*分，*B*学生对12个选择题中每个题的四个选项都能判断其中有一个选项是错误的，对其它三个选项都没有把握，选择题的得分为*Y*分，则*D*(*Y*)－*D*(*X*)等于(　　)

A. B. C. D.

答案　A

解析　设*A*学生答对题的个数为*m*，

则得分*X*＝5*m*，*m*～*B*，

*D*(*m*)＝12××＝，

所以*D*(*X*)＝25×＝；

同理设*B*学生答对题的个数为*n*，

则得分*Y*＝5*n*，*n*～*B*，*D*(*n*)＝12××＝，

所以*D*(*Y*)＝×25＝，

所以*D*(*Y*)－*D*(*X*)＝－＝.

二、多项选择题

7．已知随机变量*ξi*满足*P*(*ξi*＝1)＝*pi*，*P*(*ξi*＝0)＝1－*pi*，*i*＝1,2.若0<*p*1<*p*2<，则下列结论正确的是(　　)

A．*E*(*ξ*1)<*E*(*ξ*2)

B．*E*(*ξ*1)>*E*(*ξ*2)

C．*D*(*ξ*1)<*D*(*ξ*2)

D．*D*(*ξ*1)>*D*(*ξ*2)

答案　AC

解析　∵*E*(*ξ*1)＝*p*1，*E*(*ξ*2)＝*p*2，

∴*E*(*ξ*1)<*E*(*ξ*2)，

∵*D*(*ξ*1)＝*p*1(1－*p*1)，*D*(*ξ*2)＝*p*2(1－*p*2)，

∴*D*(*ξ*1)－*D*(*ξ*2)＝(*p*1－*p*2)(1－*p*1－*p*2)<0，

即*D*(*ξ*1)<*D*(*ξ*2)．

8．某国产杀毒软件的比赛规则为每个软件进行四轮考核，每轮考核中能够准确对病毒进行查杀的进入下一轮考核，否则被淘汰．已知某个软件在四轮考核中能够准确杀毒的概率依次是，，，，且各轮考核能否通过互不影响，则(　　)

A．该软件通过考核的概率为

B．该软件在第三轮考核被淘汰的概率为

C．该软件至少能够通过两轮考核的概率为

D．在此次比赛中该软件平均考核了轮

答案　ABD

解析　设事件*Ai*(*i*＝1,2,3,4)表示“该软件能通过第*i*轮考核”，则*P*(*A*1)＝，*P*(*A*2)＝，*P*(*A*3)＝，*P*(*A*4)＝.该软件通过考核的概率为*P*(*A*1*A*2*A*3*A*4)＝*P*(*A*1)*P*(*A*2)*P*(*A*3)*P*(*A*4)＝×××＝，选项A正确；该软件在第三轮考核被淘汰的概率为*P*(*A*1*A*2)＝*P*(*A*1)*P*(*A*2)*P*()＝××＝，选项B正确；该软件至少能够通过两轮考核的概率为1－*P*()－*P*(*A*1)＝1－－×＝，选项C不正确；设在此次比赛中，该软件考核了*Y*轮，∴*Y*的可能取值为1,2,3,4，*P*(*Y*＝1)＝*P*()＝，*P*(*Y*＝2)＝*P*(*A*1)＝×＝，*P*(*Y*＝3)＝*P*(*A*1*A*2)＝，*P*(*Y*＝4)＝*P*(*A*1*A*2*A*3)＝××＝，∴*E*(*Y*)＝1×＋2×＋3×＋4×＝，故选项D正确．

三、填空题

9．某校高一新生健康检查的统计结果显示：体重超重者占40%，血压异常者占15%，两者都有的占8%，今任选一该校高一新生，已知此人体重超重，则他血压异常的概率为\_\_\_\_\_\_\_\_．

答案　0.2

解析　记事件*A*表示此人体重超重，事件*B*表示此人血压异常，则*P*(*A*)＝0.4，*P*(*AB*)＝0.08，*P*(*B*|*A*)＝＝＝0.2.

10．出租车司机从饭店到火车站途中经过六个交通岗，假设他在各交通岗遇到红灯这一事件是相互独立的，并且概率都是，则这位司机遇到红灯前已经通过了两个交通岗的概率为\_\_\_\_\_\_\_\_．

答案

解析　因为这位司机在第一、二个交通岗未遇到红灯，在第三个交通岗遇到红灯之间是相互独立的，且遇到红灯的概率都是，所以未遇到红灯的概率都是1－＝，所以遇到红灯前已经通过了两个交通岗的概率为××＝.

11．(2020·临沂模拟)《易经》是中国传统文化中的精髓，如图是易经八卦(含乾、坤、巽、震、坎、离、艮、兑八卦)，每一卦由三根线组成(“”表示一根阳线，“”表示一根阴线)，从八卦中任取两卦，这两卦的六根线中恰有两根阳线，四根阴线的概率为\_\_\_\_\_\_\_\_．



答案

解析　观察八卦图可知，含三根阴线的共有一卦，含三根阳线的共有一卦，含两根阳线一根阴线的共有三卦，含一根阳线两根阴线的共有三卦，所以从八卦中任取两卦有C＝28(种)情况．其中抽取的两卦中六根线恰有两根阳线，四根阴线的所有情况是一卦含有三根阴线，另一卦含有两根阳线一根阴线，或者两卦都含有一根阳线两根阴线，即C＋C＝6(种)情况．故所求概率为*P*＝＝.

12．(2020·浙江)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*(*ξ*＝0)＝\_\_\_\_\_\_\_\_，*E*(*ξ*)＝\_\_\_\_\_\_\_\_.

答案　　1

解析　方法一　1个红球，1个绿球，2个黄球，共有A＝12(种)排列．

①红球前面没有黄球，有A＋1＝4(种)，

*P*(*ξ*＝0)＝＝；

②红球前面有1个黄球，有A＋A＝4(种)，

*P*(*ξ*＝1)＝＝；

③红球前面有2个黄球，有1＋A＝4(种)，

*P*(*ξ*＝2)＝＝.

*E*(*ξ*)＝0×＋1×＋2×＝1.

方法二　①第1次就取到红球：*P*(红)＝；

②第2次取到红球：*P*(黄，红)＝×＝，

*P*(绿，红)＝×＝；

③第3次取到红球：*P*(黄，黄，红)＝××＝，

*P*(黄，绿，红)＝××＝，

*P*(绿，黄，红)＝××＝；

④第4次取到红球：

*P*(黄，黄，绿，红)＝××＝，

*P*(黄，绿，黄，红)＝××＝，

*P*(绿，黄，黄，红)＝××＝.

故*P*(*ξ*＝0)＝*P*(红)＋*P*(绿，红)＝＋＝，

*P*(*ξ*＝1)＝*P*(黄，红)＋*P*(黄，绿，红)＋*P*(绿，黄，红)

＝＋＋＝，

*P*(*ξ*＝2)＝*P*(黄，黄，红)＋*P*(黄，黄，绿，红)＋*P*(黄，绿，黄，红)＋*P*(绿，黄，黄，红)

＝＋＋＋＝.

则*E*(*ξ*)＝0×＋1×＋2×＝1.

四、解答题

13．某大学在一次公益活动中聘用了10名志愿者，他们分别来自于*A*，*B*，*C*三个不同的专业，其中*A*专业2人，*B*专业3人，*C*专业5人，现从这10人中任意选取3人参加一个访谈节目．

(1)求3个人来自两个不同专业的概率；

(2)设*X*表示取到*B*专业的人数，求*X*的分布列与均值．

解　(1)令事件*A*表示“3个人来自于两个不同专业”，

事件*A*1表示“3个人来自于同一个专业”，

事件*A*2表示“3个人来自于三个不同专业”，

*P*(*A*1)＝＝，

*P*(*A*2)＝＝＝，

∴3个人来自两个不同专业的概率

*P*(*A*)＝1－*P*(*A*1)－*P*(*A*2)＝1－－＝.

(2)随机变量*X*的取值为0,1,2,3，

*P*(*X*＝0)＝＝＝，

*P*(*X*＝1)＝＝＝，

*P*(*X*＝2)＝＝＝，

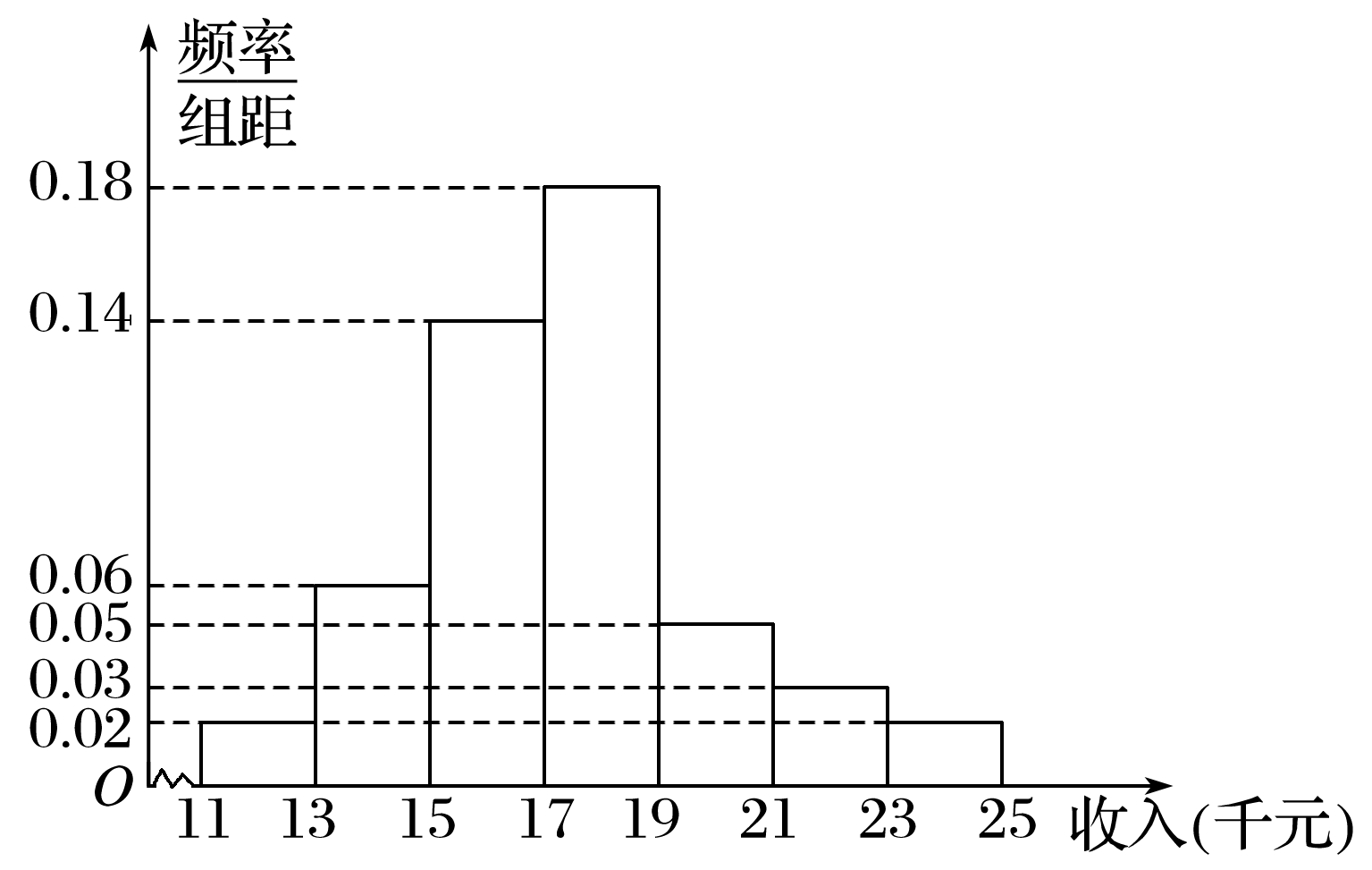
*P*(*X*＝3)＝＝，

∴*X*的分布列为

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *X* | 0 | 1 | 2 | 3 |
| *P* |  |  |  |  |

*E*(*X*)＝0×＋1×＋2×＋3×＝.

14．(2020·寿光市第二中学月考)十九大以来，某贫困地区扶贫办积极贯彻落实国家精准扶贫的政策要求，带领广大农村地区人民群众脱贫奔小康．经过不懈的奋力拼搏，新农村建设取得巨大进步，农民年收入也逐年增加，为了制定提升农民收入力争早日脱贫的工作计划，该地扶贫办统计了2019年50位农民的年收入并制成如下频率分布直方图：



(1)根据频率分布直方图，估计50位农民的年平均收入(单位：千元)(同一组数据用该组数据区间的中点值表示)；

(2)由频率分布直方图，可以认为该贫困地区农民收入*X*服从正态分布*N*(*μ*，*σ*2)，其中*μ*近似为年平均收入，*σ*2近似为样本方差*s*2，经计算得*s*2＝6.92，利用该正态分布，求：

①在扶贫攻坚工作中，若使该地区约有84.14%的农民的年收入不低于扶贫办制定的最低年收入标准，则最低年收入大约为多少千元？

②为了调研“精准扶贫，不落一人”的政策要求落实情况，扶贫办随机走访了1 000位农民．若每位农民的年收入互相独立，这1 000位农民中的年收入不少于12.14千元的人数为*ξ*，求*E*(*ξ*)．

附参考数据：≈2.63，

若随机变量*X*服从正态分布*N*(*μ*，*σ*2)，则

*P*(*μ*－*σ*<*X*≤*μ*＋*σ*)≈0.682 7，*P*(*μ*－2*σ*<*X*≤*μ*＋2*σ*)≈0.954 5，*P*(*μ*－3*σ*<*X*≤*μ*＋3*σ*)≈0.997 3.

解　(1)＝12×0.04＋14×0.12＋16×0.28＋18×0.36＋20×0.10＋22×0.06＋24×0.04＝17.40(千元)，

故估计50位农民的年平均收入为17.40千元．

(2)由题意知*X*～*N*(17.40,6.92)，

①*P*(*X*≥*μ*－*σ*)＝0.5＋≈0.841 4，

所以*μ*－*σ*≈17.40－2.63＝14.77时，满足题意，

即最低年收入大约为14.77千元．

②由*P*(*X*≥12.14)＝*P*(*X*≥*μ*－2*σ*)＝0.5＋≈0.977 3，

每个农民的年收入不少于12.14千元的事件的概率为0.977 3，

则*ξ*～*B*(1 000，*p*)，其中*p*＝0.977 3，

所以*E*(*ξ*)＝1 000×0.977 3＝977.3.