第五节　离散型随机变量及其分布列、数字特征

目录

CONTENTS

PART 1

知识 体系构建

离散型随机变量及其分布列、数字特征

- 离散型随机变量及其分布列
  - 随机变量
    - 定义：对于随机试验样本空间Ω中的每个样本点ω，都有____的实数$X(\omega)$与之对应，称X为随机变量
    - 分类
      - 离散型随机变量：随机变量可能取值为______或可以一一列出，这样的随机变量叫做离散型随机变量
      - 连续型随机变量：随机变量可以取某一区间内的一切值，这样的随机变量叫做连续型随机变量
  - 离散型随机变量的分布列
    - 定义：设离散型随机变量X的可能取值为$x_1,x_2,\cdots,x_n$，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.离散型随机变量的分布列也可以用表格

X	x_1	x_2	$\cdots$	x_n
P	p_1	p_2	$\cdots$	p_n

表示

    - 性质：(1)$p_i\geqslant0,i=1,2,\cdots,n$；(2)$p_1+p_2+\cdots+p_n=$<u>1</u>

离散型随机变量及其分布列、数字特征

离散型随机变量的数字特征

均值

若离散型随机变量X的分布列为

X	x_1	x_2	$\cdots$	x_n
P	p_1	p_2	$\cdots$	p_n

，则称$E(X)=$________$=\sum\limits_{i=1}^{n}x_ip_i$为随机变量$X$的均值或数学期望，它反映了随机变量取值的________

方差（标准差）

若离散型随机变量X的分布列为

X	x_1	x_2	$\cdots$	x_n
P	p_1	p_2	$\cdots$	p_n

，则$(x_i-E(X))^2$描述了$x_i(i=1,2,\cdots,n)$相对于均值$E(X)$的________$D(X)=\sum\limits_{i=1}^{n}(x_i-E(X))^2p_i$为这些偏离程度的加权平均，刻画了随机变量$X$取值与其均值$E(X)$的偏离程度，称$D(X)$为随机变量$X$的方差，并称其算术平方根$\sqrt{D(X)}$为随机变量$X$的标准差，记为$\sigma(X)$

对点自测

1. 袋中有大小相同的6个黑球，5个白球，从袋中每次任意取出1个球且不放回，直到取出的球是白球，记所需要的取球次数为随机变量X，则X的可能取值为（　　）

A. 1，2，3，…，6　　B. 1，2，3，…，7

C. 0，1，2，…，5　　D. 1，2，…，5

解析：　因为取到白球时停止，所以最少取球次数为1，即第一次就取到了白球；最多取球次数是7次，即把所有的黑球取完之后才取到白球.所以取球次数可以是1，2，3，…，7.

2. 某一随机变量ξ的概率分布如表所示，且$m+2n=1.2$，则$m-\frac{n}{2}=$

(　　)

ξ	0	1	2	3
P	0.1	m	n	0.1

A. -0.2　　B. 0.2

C. 0.1　　D. -0.1

解析：　由离散型随机变量分布列的性质可得$m+n+0.2=1$，又$m+2n=1.2$，解得$m=n=0.4$，可得$m-\frac{n}{2}=0.2$.

3. 设随机变量X的概率分布如表所示，且$E(X)=2.5$，则$a-b=$（　　）

X	1	2	3	4
P	$\frac{1}{4}$	$\frac{3}{16}$	a	b

A. $\frac{3}{4}$　　B. $\frac{3}{8}$

C. $\frac{3}{16}$　　D. $\frac{3}{32}$

解析： 由题意得，$\begin{cases}\frac{1}{4}+\frac{3}{16}+a+b=1,\\1\times\frac{1}{4}+2\times\frac{3}{16}+3a+4b=2.5,\end{cases}$解得

$\begin{cases}a=\frac{3}{8},\\b=\frac{3}{16},\end{cases}$ $\therefore a-b=\frac{3}{16}$.故选C.

4. 随机变量X的取值为0，1，2，若$P(X=0)=\frac{1}{4}$，$E(X)=1$，则$D(X)=$（　　）

A. $\frac{1}{4}$　　B. $\frac{1}{2}$

C. $\frac{3}{4}$　　D. 1

解析： 设$P(X=1)=p$，$P(X=2)=q$，由题意得$E(X)=0\times\frac{1}{4}+p+2q=1$，$\frac{1}{4}+p+q=1$，解得$p=\frac{1}{2}$，$q=\frac{1}{4}$，

$\therefore D(X)=\frac{1}{4}\times(0-1)^2+\frac{1}{2}\times(1-1)^2+\frac{1}{4}\times(2-1)^2=\frac{1}{2}$，故选B.

常用结论

1. 若 $Y=aX+b$，其中 a，b 是常数，X 是随机变量，则

(1) $E(k)=k$，$D(k)=0$，其中 k 为常数；

(2) $E(aX+b)=aE(X)+b$，$D(aX+b)=a^2D(X)$；

(3) $E(X_1+X_2)=E(X_1)+E(X_2)$；

(4) $D(X)=E(X^2)-(E(X))^2$；

(5) 若 X_1，X_2 相互独立，则 $E(X_1 \cdot X_2)=E(X_1)\cdot E(X_2)$.

2. 若随机变量 X 服从两点分布，则 $E(X)=p$，$D(X)=p(1-p)$.

应用

1. 已知随机变量ξ和η，其中$\eta=12\xi+7$，且$E(\eta)=34$，若ξ的分布列如下表，则$m=$（　　）

ξ	1	2	3	4
P	$\frac{1}{4}$	m	n	$\frac{1}{12}$

A. $\frac{1}{3}$　　B. $\frac{1}{4}$

C. $\frac{1}{6}$　　D. $\frac{1}{8}$

解析： $\because \eta=12\xi+7$，由结论1得$E(\eta)=12E(\xi)+7$，即$E(\eta)=12\left(1\times\frac{1}{4}+2\times m+3\times n+4\times\frac{1}{12}\right)+7=34$，$\therefore 2m+3n=\frac{5}{3}$，①.又$\frac{1}{4}+m+n+\frac{1}{12}=1$，$\therefore m+n=\frac{2}{3}$，②.由①②，可解得$m=\frac{1}{3}$.

2. 在篮球比赛中，罚球命中1次得1分，不中得0分.如果某运动员罚球命中的概率为0.8，那么他罚球1次的得分 X 的均值为__0.8__.

解析：由结论2易得 $E(X)=0.8$.

课堂演练

考点 分类突破

精选考点 典例研析 技法重悟通

PART 2

考点一 分布列的性质

（基础自学过关）

1. （2024·济宁一模）离散型随机变量X的概率分布规律为$P(X=n)=\dfrac{a}{n(n+1)}(n=1,2,3,4)$，其中$a$是常数，则$P\left(\dfrac{1}{2}<X<\dfrac{5}{2}\right)=$（　　）

A. $\dfrac{2}{3}$　　B. $\dfrac{3}{4}$

C. $\dfrac{4}{5}$　　D. $\dfrac{5}{6}$

解析： 因为$P(X=n)=\dfrac{a}{n(n+1)}$（$n=1$，2，3，4），所以

$\dfrac{a}{2}+\dfrac{a}{6}+\dfrac{a}{12}+\dfrac{a}{20}=1$，即$a=\dfrac{5}{4}$，所以$P\left(\dfrac{1}{2}<X<\dfrac{5}{2}\right)=P(X=1)$

$+P(X=2)=\dfrac{5}{4}\times\dfrac{1}{2}+\dfrac{5}{4}\times\dfrac{1}{6}=\dfrac{5}{6}$.

2. （2024·云南一中检测）设离散型随机变量ξ的分布列如下表所示，则下列各式正确的是（　　）

ξ	-1	0	1	2	3
P	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{2}{5}$

A. $P(\xi<3)=\frac{2}{5}$　　B. $P(\xi>1)=\frac{4}{5}$

C. $P(2<\xi<4)=\frac{2}{5}$ ✓　　D. $P(\xi<0.5)=0$

解析： $P(\xi<3)=\frac{1}{10}+\frac{1}{5}+\frac{1}{10}+\frac{1}{5}=\frac{3}{5}$，A错误；$P(\xi>1)=\frac{1}{5}+\frac{2}{5}=\frac{3}{5}$，B错误；$P(2<\xi<4)=P(\xi=3)=\frac{2}{5}$，C正确；$P(\xi<0.5)=\frac{1}{10}+\frac{1}{5}=\frac{3}{10}$，D错误.

3. 已知随机变量X的分布列为

X	-1	0	1
P	a	b	c

其中a，b，c成等差数列，则$P(|X|=1)$=$\underline{\quad \frac{2}{3} \quad}$，公差$d$的取值范围是$\underline{\quad [-\frac{1}{3}, \frac{1}{3}] \quad}$.

解析：因为a，b，c成等差数列，所以$2b=a+c$.又$a+b+c=1$，所以$b=\frac{1}{3}$，因此$P(|X|=1)=a+c=\frac{2}{3}$.又$a=\frac{1}{3}-d$，$c=\frac{1}{3}+d$，根据分布列的性质，得$0\leqslant\frac{1}{3}-d\leqslant\frac{2}{3}$，$0\leqslant\frac{1}{3}+d\leqslant\frac{2}{3}$，所以$-\frac{1}{3}\leqslant d\leqslant\frac{1}{3}$.

练后悟通

离散型随机变量分布列性质的应用

（1）利用“总概率之和为1”可以求相关参数的取值范围或值；

（2）利用“离散型随机变量在一范围内的概率等于它取这个范围内各个值的概率之和”求某些特定事件的概率；

（3）可以根据性质判断所得分布列结果是否正确.

考点二　离散型随机变量的均值与方差

（定向精析突破）

考向1　均值与方差的性质

【例1】　（多选）设离散型随机变量X的分布列为

X	0	1	2	3	4
P	q	0.4	0.1	0.2	0.2

若离散型随机变量 Y满足 $Y=2X+1$，则下列结果正确的有（　　）

A. $q=0.1$

B. $E(X)=2$，$D(X)=1.4$

C. $E(X)=2$，$D(X)=1.8$

D. $E(Y)=5$，$D(Y)=7.2$

解析： 因为$q+0.4+0.1+0.2+0.2=1$，所以$q=0.1$，故A正确；由已知可得$E(X)=0\times0.1+1\times0.4+2\times0.1+3\times0.2+4\times0.2=2$，$D(X)=(0-2)^2\times0.1+(1-2)^2\times0.4+(2-2)^2\times0.1+(3-2)^2\times0.2+(4-2)^2\times0.2=1.8$，故C正确，B错误；因为$Y=2X+1$，所以$E(Y)=2E(X)+1=5$，$D(Y)=4D(X)=7.2$，故D正确.

解题技法

与均值、方差的性质有关问题的解题思路

若给出的随机变量Y与X的关系为$Y=aX+b$，a，b为常数，一般思路是先求出$E(X)$，$D(X)$，再利用公式$E(aX+b)=aE(X)+b$，$D(aX+b)=a^2D(X)$求$E(Y)$，$D(Y)$；也可以利用X的分布列得到Y的分布列，关键是由X的取值计算Y的取值，对应的概率相等，再由定义法求得$E(Y)$或$D(Y)$.

考向2　离散型随机变量的均值与方差

【例2】　在一场娱乐晚会上，有5位民间歌手（编号为1至5号）登台演唱，由现场数百名观众投票选出最受欢迎歌手.各位观众必须彼此独立地在选票上选3名歌手，其中观众甲是1号歌手的歌迷，他必选1号，不选2号，观众乙和丙对5位歌手的演唱没有偏爱，可以在1至5号中随机选3名歌手.

（1）求观众甲选中3号歌手且观众乙未选中3号歌手的概率；

解：设A表示事件“观众甲选中3号歌手”，B表示事件“观众乙选中3号歌手”，

观众甲必选1号，不选2号，还要从3，4，5号中选2位歌手，因此样本空间共有C_3^2个样本点，事件A有C_2^1个样本点，因此P

$$(A)=\frac{C_2^1}{C_3^2}=\frac{2}{3}.$$

观众乙在1至5号中随机选3名歌手，样本空间有 C_5^3 个样本点，事件 B 有 C_4^2 个样本点，因此 P（B）$=\frac{C_4^2}{C_5^3}=\frac{3}{5}$.

（2）设X表示3号歌手得到观众甲、乙、丙的票数之和，求X的分布列和期望.

解：X的所有可能取值为0，1，2，3，因为三位观众的投票相互独立

所以$P(X=0)=\frac{1}{3}\times\frac{2}{5}\times\frac{2}{5}=\frac{4}{75}$，

$P(X=1)=\frac{2}{3}\times\frac{2}{5}\times\frac{2}{5}+\frac{1}{3}\times\frac{3}{5}\times\frac{2}{5}+\frac{1}{3}\times\frac{2}{5}\times\frac{3}{5}=\frac{20}{75}=\frac{4}{15}$，

$P(X=2)=\frac{2}{3}\times\frac{3}{5}\times\frac{2}{5}+\frac{1}{3}\times\frac{3}{5}\times\frac{3}{5}+\frac{2}{3}\times\frac{2}{5}\times\frac{3}{5}=\frac{33}{75}=\frac{11}{25}$，

$P(X=3)=\frac{2}{3}\times\frac{3}{5}\times\frac{3}{5}=\frac{18}{75}=\frac{6}{25}$，

所以X的分布列为

X	0	1	2	3
P	$\frac{4}{75}$	$\frac{4}{15}$	$\frac{11}{25}$	$\frac{6}{25}$

X的期望为$E(X)=0\times\frac{4}{75}+1\times\frac{4}{15}+2\times\frac{11}{25}+3\times\frac{6}{25}=\frac{28}{15}$.

解题技法

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可能的全部值；

（2）求X取每个值的概率；

（3）写出X的分布列；

（4）由均值、方差的定义求$E(X)$，$D(X)$.

训练

1. 已知ξ的分布列如表所示：

ξ	0	1	2
P	？	！	？

其中，尽管“！”处完全无法看清，且两个“？”处字迹模糊，但能断定这两个“？”处的数值相同.据此计算，下列各式中：①$E(\xi)=1$；②$D(\xi)>1$；③$P(\xi=0)\leqslant\frac{1}{2}$，正确的个数是（　　）

A. 0　　B. 1

C. 2　　D. 3

解析： 设“？”$=a$，“！”$=b$，则a，$b\in[0，1]$，$2a+b=1$.①E（ξ）$=0\times a+1\times b+2\times a=2a+b=1$，因此①正确；②$D$（ξ）=（0-1）$^2\times a+$（1-1）$^2\times b+$（2-1）$^2\times a=2a\leqslant 1$，因此②不正确；③$P$（ξ=0）$=a=\frac{1-b}{2}\leq\frac{1}{2}$，因此③正确.

2. 为推广滑雪运动，某滑雪场开展滑雪促销活动.该滑雪场的收费标准是：滑雪时间不超过1小时免费，超过1小时的部分每小时收费标准为40元（不足1小时的部分按1小时计算）.有甲、乙两人相互独立地来该滑雪场运动，设甲、乙不超过1小时离开的概率分别为$\frac{1}{4}$，$\frac{1}{6}$；1小时以上且不超过2小时离开的概率分别为$\frac{1}{2}$，$\frac{2}{3}$；两人滑雪时间都不会超过3小时.

（1）求甲、乙两人所付滑雪费用相同的概率；

解：两人所付滑雪费用相同，相同的费用可能为0，40，80元，甲、乙两人2小时以上且不超过3小时离开的概率分别为$1-\frac{1}{4}-\frac{1}{2}=\frac{1}{4}$，$1-\frac{1}{6}-\frac{2}{3}=\frac{1}{6}$.

两人都付0元的概率为$P_1=\frac{1}{4}\times\frac{1}{6}=\frac{1}{24}$，

两人都付40元的概率为$P_2=\frac{1}{2}\times\frac{2}{3}=\frac{1}{3}$，

两人都付80元的概率为$P_3=\frac{1}{4}\times\frac{1}{6}=\frac{1}{24}$，

则两人所付滑雪费用相同的概率为$P=P_1+P_2+P_3=\frac{1}{24}+\frac{1}{3}+\frac{1}{24}=\frac{5}{12}$.

（2）设甲、乙两人所付的滑雪费用之和为随机变量ξ，求ξ的分布列与均值E（ξ），方差D（ξ）.

解：ξ的所有可能取值为0，40，80，120，160，则

P（ξ＝0）＝$\frac{1}{4}\times\frac{1}{6}=\frac{1}{24}$,

P（ξ＝40）＝$\frac{1}{4}\times\frac{2}{3}+\frac{1}{2}\times\frac{1}{6}=\frac{1}{4}$,

P（ξ＝80）＝$\frac{1}{4}\times\frac{1}{6}+\frac{1}{2}\times\frac{2}{3}+\frac{1}{4}\times\frac{1}{6}=\frac{5}{12}$,

P（ξ＝120）＝$\frac{1}{2}\times\frac{1}{6}+\frac{1}{4}\times\frac{2}{3}=\frac{1}{4}$,

P（ξ＝160）＝$\frac{1}{4}\times\frac{1}{6}=\frac{1}{24}$.

所以ξ的分布列为

ξ	0	40	80	120	160
P	$\frac{1}{24}$	$\frac{1}{4}$	$\frac{5}{12}$	$\frac{1}{4}$	$\frac{1}{24}$

E（ξ）$=0\times\frac{1}{24}+40\times\frac{1}{4}+80\times\frac{5}{12}+120\times\frac{1}{4}+160\times\frac{1}{24}=80.$

D（ξ）$=$（$0-80$）$^2\times\frac{1}{24}+$（$40-80$）$^2\times\frac{1}{4}+$（$80-80$）$^2\times\frac{5}{12}+$（$120-80$）$^2\times\frac{1}{4}+$（$160-80$）$^2\times\frac{1}{24}=\frac{4\ 000}{3}$.

考点三 均值与方差在决策中的应用

（师生共研过关）

【例3】 （2021·新高考Ⅰ卷18题）某学校组织“一带一路”知识竞赛，有A，B两类问题.每位参加比赛的同学先在两类问题中选择一类并从中随机抽取一个问题回答，若回答错误则该同学比赛结束；若回答正确则从另一类问题中再随机抽取一个问题回答，无论回答正确与否，该同学比赛结束.A类问题中的每个问题回答正确得20分，否则得0分；B类问题中的每个问题回答正确得80分，否则得0分.

已知小明能正确回答A类问题的概率为0.8，能正确回答B类问题的概率为0.6，且能正确回答问题的概率与回答次序无关.

（1）若小明先回答A类问题，记X为小明的累计得分，求X的分布列；

解：由题意得，X的所有可能取值为0，20，100，

$P(X=0)=1-0.8=0.2$，

$P(X=20)=0.8\times(1-0.6)=0.32$，

$P(X=100)=0.8\times0.6=0.48$，

所以X的分布列为

X	0	20	100
P	0.2	0.32	0.48

（2）为使累计得分的期望最大，小明应选择先回答哪类问题？并说明理由.

解：当小明先回答A类问题时，由（1）可得$E(X)=0\times0.2+20\times0.32+100\times0.48=54.4$.

当小明先回答B类问题时，记Y为小明的累计得分，

则Y的所有可能取值为0，80，100，

$P(Y=0)=1-0.6=0.4$，

$P(Y=80)=0.6\times(1-0.8)=0.12$，

$P(Y=100)=0.6\times0.8=0.48$，

所以 Y 的分布列为

Y	0	80	100
P	0.4	0.12	0.48

$E(Y)=0\times0.4+80\times0.12+100\times0.48=57.6.$

因为 $57.6>54.4$，即 $E(Y)>E(X)$，所以为使累计得分的期望最大，小明应选择先回答 B 类问题.

解题技法

利用样本的数字特征解决有关决策问题的关键

（1）建立模型，根据题意准确建立解决问题的概率模型，要注意各种概率模型的差异性，不能混淆；

（2）分析数据，分析题中的相关数据，确定概率模型中的相关参数；

（3）求值，利用概率知识求出概率模型中的数学期望、方差等数字特征；

（4）做出决策，比较概率模型中的数字特征，确定解决问题的最优方案，做出决策.

训练

1. 某投资公司在2024年年初准备将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$\frac{1}{15}$.