概率、统计问题

一、课前热身训练

1. 【答案】C

【解析】不超过 30 的素数有 2，3，5，7，11，13，17，19，23，29，共 10 个，随机选取两个不同的数，共有 $C_{10}^{2}=45$ 种方法，因为 $7+23=11+19=13+17=30$，所以随机选取两个不同的数，其和等于 30 的有 3 种方法，故概率为 $\frac{3}{45}=\frac{1}{15}$，故选 C.

2. 答案：B

解答：由 $X\sim B(10,p)$，$\therefore DX=10p(1-p)=2.4$，$\therefore 10p^2-10p+2.4=0$，解之得 $p_1=0.4,p_2=0.6$，由 $P(X=4)<P(X=6)$，有 $p=0.6$.

3.答案：A

解答：取 $AB=AC=2$,则 $BC=2\sqrt{2}$，

$\therefore$区域Ⅰ的面积为 $S_1=\frac{1}{2}\times2\times2=2$，区域III的面积为 $S_3=\frac{1}{2}\cdot\pi(\sqrt{2})^2-2=\pi-2$，

区域Ⅱ的面积为 $S_2=\pi\cdot1^2-S_3=2$，故 $p_1=p_2$.

4. 【详解】由散点图分布可知，散点图分布在一个对数函数的图象附近，

因此，最适合作为发芽率 y 和温度 x 的回归方程类型的是 $y=a+b\ln x$.

故选：D.

5.答案：A

解答：由于共 9 个评委，将评委所给分数从小到大排列，中位数是第 5 个，假设为 a，去掉一头一尾的最低和最高分后，中位数还是 a，所以不变的是数字特征是中位数。其它的数字特征都会改变。

6.答案：0.98

解答：经停该站的列出共有 40 个车次，所有车次的平均正点率的估计值为 $P=\frac{10\times0.97+20\times0.98+10\times0.99}{40}=0.98$。

典型例题研究

例 1.

解：（1）一轮实验中甲药的得分有三种情况：1、-1、0.

得 1 分时是施以甲药的白鼠治愈且施以乙药的白鼠未治愈，则 $P(X=1)=\alpha(1-\beta)$；

得 -1 分时是施以乙药的白鼠治愈且施以甲药的白鼠未治愈，则 $P(X=-1)=(1-\alpha)\beta$；

得 0 分时是都治愈或都未治愈，则 $P(X=0)=\alpha\beta+(1-\alpha)(1-\beta)$.

则 X 的分布列为：

X	1	-1	0
P	$\alpha(1-\beta)$	$\beta(1-\alpha)$	$\alpha\beta+(1-\alpha)(1-\beta)$

（2）（i）因为$\alpha=0.5$，$\beta=0.8$，

则 $a=P(X=-1)=0.4$， $b=P(X=0)=0.5$， $c=P(X=1)=0.1$.

可得 $p_i=0.4p_{i-1}+0.5p_i+0.1p_{i+1}$，则$0.5p_i=0.4p_{i-1}+0.1p_{i+1}$，

则$0.4(p_i-p_{i-1})=0.1(p_{i+1}-p_i)$，则$\dfrac{p_{i+1}-p_i}{p_i-p_{i-1}}=4$，

所以 $\{p_{i+1}-p_i\}(i=0,1,2,\cdots,7)$为等比数列.

（ii） $\{p_{i+1}-p_i\}(i=0,1,2,\cdots,7)$的首项为 $p_1-p_0=p_1$，那么可得：

$p_8-p_7=p_1\times 4^7$， $p_7-p_6=p_1\times 4^6$，……………… $p_2-p_1=p_1\times 4$，

以上 7 个式子相加，得到 $p_8-p_1=p_1\times(4^7+4^6+\cdots+4)$，

则 $p_8=p_1\times(1+4+\cdots+4^6+4^7)=p_1\times\dfrac{1-4^8}{1-4}=\dfrac{4^8-1}{3}p_1$，则 $p_1=\dfrac{3}{4^8-1}$，

再把后面三个式子相加，得 $p_4-p_1=p_1\times(4+4^2+4^3)$，

则 $p_4=p_1\times(1+4+4^2+4^3)=\dfrac{4^4-1}{3}p_1=\dfrac{4^4-1}{3}\times\dfrac{3}{4^8-1}=\dfrac{1}{4^4+1}=\dfrac{1}{257}$.

p_4表示“甲药治愈的白鼠比乙药治愈的白鼠多 4 只，且甲药的累计得分为 4”，因为$\alpha=0.5$，$\beta=0.8$，$\alpha<\beta$，则实验结果中“甲药治愈的白鼠比乙药治愈的白鼠多 4 只，且甲药的累计得分为 4”这种情况的概率是非常小的，而 $p_4=\dfrac{1}{257}$的确非常小，说明这种实验方案是合理的.

例 2. 解:依题意得$\begin{cases}a+0.2+0.15=0.7\\0.05+b+0.15+0.15+0.2+a=1\end{cases}$，解得$\begin{cases}a=0.35\\b=0.1\end{cases}$.

（1）$0.15\times2+0.2\times3+0.3\times4+0.2\times5+0.1\times6+0.05\times7=4.05$

$0.05\times3+0.1\times4+0.15\times5+0.35\times6+0.2\times7+0.15\times8=6$

得到甲离子残留百分比的平均值为 4.05,，乙离子残留百分比的平均值为 6.

例 3.【思路】:（1）典型的二项分布，利用正态分布的性质计算。（2）考察正态分布，代入运算即可。

【解析】:（1）$P\left(X\geq 1\right)=1-P\left(X=0\right)=1-0.9974^{16}=1-0.9592=0.0408$

由题意可得，X 满足二项分布 $X\sim B\left(16,0.0016\right)$，因此可得 $EX\left(16,0.0016\right)==16\times 0.0016=0.0256$

（2）①由（1）可得 $P\left(X\geq 1\right)=0.0408<5\%$，属于小概率事件，故而如果出现 $(\mu-3\sigma,\mu+3\sigma)$ 的零件，需要进行检查。

②由题意可得 $\mu=9.97,\sigma=0.212\Rightarrow\mu-3\sigma=9.334,\mu+3\sigma=10.606$，故而在 $\left(9.334,10.606\right)$ 范围外存在 9.22 这一个数据，因此需要进行检查。此时：$\mu=\bar{x}=\dfrac{9.97\times 16-9.22}{15}=10.02$，

$\sigma=\sqrt{\dfrac{1}{15}\sum\limits_{i=1}^{15}\left(x-\bar{x}\right)}\approx 0.09$。

例 4.解答:（1）第一种生产方式的平均数为 $\overline{x_1}=84$，第二种生产方式平均数为 $\overline{x_2}=74.7$，

∴ $\overline{x_1}>\overline{x_2}$，所以第一种生产方式完成任务的平均时间大于第二种，

∴第二种生产方式的效率更高.

（2）由茎叶图数据得到 $m=80$，

∴列联表为

	超过m	不超过m	合计
第一种生产方式	15	5	20
第二种生产方式	5	15	20
合计	20	20	40

（3）$K^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}=\dfrac{40(15\times 15-5\times 5)^2}{20\times 20\times 20\times 20}=10>6.635$，

∴有 99% 的把握认为两种生产方式的效率有差异.

二、课堂小结与反思

三、课后巩固训练

1.答案：A **2.**答案：0.18 **3.**答案：C **4.**故选：C.

5.【分析】

（1）根据独立事件的概率乘法公式可求得事件“甲连胜四场”的概率;

（2）计算出四局以内结束比赛的概率，然后利用对立事件的概率公式可求得所求事件的概率;

（3）列举出甲赢的基本事件，结合独立事件的概率乘法公式计算出甲赢的概率，由对称性可知乙赢的概率和甲赢的概率相等，再利用对立事件的概率可求得丙赢的概率.

【详解】（1）记事件 M :甲连胜四场，则 $P\left(M\right)=\left(\dfrac{1}{2}\right)^4=\dfrac{1}{16}$；

（2）记事件 A 为甲输，事件 B 为乙输，事件 C 为丙输，

则四局内结束比赛的概率为

$$P'=P\left(ABAB\right)+P\left(ACAC\right)+P\left(BCBC\right)+P\left(BABA\right)=4\times\left(\frac{1}{2}\right)^4=\frac{1}{4},$$

所以，需要进行第五场比赛的概率为 $P=1-P'=\frac{3}{4}$；

（3）记事件 A 为甲输，事件 B 为乙输，事件 C 为丙输，

记事件 M：甲赢，记事件 N：丙赢，

则甲赢的基本事件包括：$BCBC$、$ABCBC$、$ACBCB$、

$BABCC$、$BACBC$、$BCACB$、$BCABC$、$BCBAC$，

所以，甲赢　概率为 $P\left(M\right)=\left(\frac{1}{2}\right)^4+7\times\left(\frac{1}{2}\right)^5=\frac{9}{32}$.

由对称性可知，乙赢的概率和甲赢的概率相等，

所以丙赢的概率为 $P\left(N\right)=1-2\times\frac{9}{32}=\frac{7}{16}$.

【点睛】本题考查独立事件概率的计算，解答的关键就是列举出符合条件的基本事件，考查计算能力，属于中等题.

6.【分析】（1）利用野生动物数量的估计值等于样区野生动物平均数乘以地块数，代入数据即可；

（2）利用公式 $r=\frac{\sum_{i=1}^{20}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{20}(x_i-\overline{x})^2\sum_{i=1}^{20}(y_i-\overline{y})^2}}$ 计算即可；

（3）各地块间植物覆盖面积差异较大，为提高样本数据的代表性，应采用分层抽样.

【详解】（1）样区野生动物平均数为 $\frac{1}{20}\sum_{i=1}^{20}y_i=\frac{1}{20}\times1200=60$，

地块数为 200，该地区这种野生动物的估计值为 $200\times60=12000$

（2）样本 $\left(x_i,y_i\right)$ (i=1，2，…，20)的相关系数为

$$r=\frac{\sum_{i=1}^{20}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{20}(x_i-\overline{x})^2\sum_{i=1}^{20}(y_i-\overline{y})^2}}=\frac{800}{\sqrt{80\times9000}}=\frac{2\sqrt{2}}{3}\approx0.94$$

（3）由（2）知各样区的这种野生动物的数量与植物覆盖面积有很强的正相关性，

由于各地块间植物覆盖面积差异很大，从俄各地块间这种野生动物的数量差异很大，

采用分层抽样的方法较好地保持了样本结构与总体结构得以执行，提高了样本的代表性，

从而可以获得该地区这种野生动物数量更准确的估计.