概率统计

学校:__________姓名：__________班级：__________考号：__________

一、统计

1．（2023·上海嘉定·统考一模）两位跳水运动员甲和乙，某次比赛中的得分如下表所示，则正确的选项为（　）

	第一跳	第二跳	第三跳	第四跳	第五跳
甲	85.5	96	86.4	75.9	94.4
乙	79.5	80	95.7	94.05	86.4

A．甲和乙的中位数相等，甲的平均分小于乙

B．甲的平均分大于乙，甲的方差大于乙

C．甲的平均分大于乙，甲的方差等于乙

D．甲的平均分大于乙，甲的方差小于乙

2．（2023·上海普陀·统考一模）已知一组数据 3、1、5、3、2，现加入p，q两数对该组数据进行处理，若经过处理后的这组数据的极差为$p-q$，则经过处理后的这组数据与之前的那组数据相比，一定会变大的数字特征是（　）

A．平均数　　B．方差　　C．众数　　D．中位数

3．（2023·上海闵行·统考一模）某校读书节期间，共 120 名同学获奖（分金、银、铜三个等级），从中随机抽取 24 名同学参加交流会，若按高一、高二、高三分层随机抽样，则高一年级需抽取 6 人；若按获奖等级分层随机抽样，则金奖获得者需抽取 4 人．下列说法正确的是（　）

A．高二和高三年级获奖同学共 80 人　　B．获奖同学中金奖所占比例一定最低

C．获奖同学中金奖所占比例可能最高　　D．获金奖的同学可能都在高一年级

4．（2023·上海宝山·统考一模）下列说法中错误的是（　）

A．一组数据的平均数、中位数可能相同

B．一组数据中比中位数大的数和比中位数小的数一样多

C．平均数、众数和中位数都是描述一组数据的集中趋势的统计量

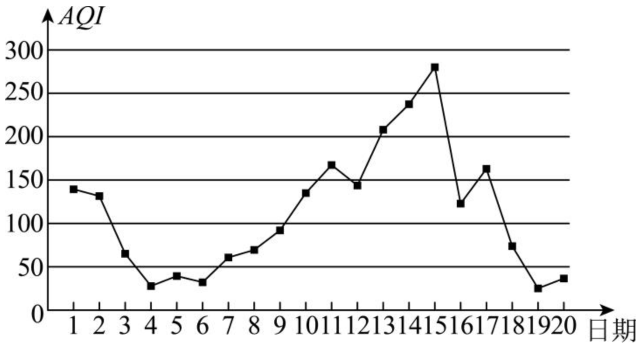
D．极差、方差、标准差都是描述一组数据的离散程度的统计量

5．（2023 上·上海虹口·高三统考期末）空气质量指数 AQI 是反映空气质量状况的指数，其对应关系如下表：

AQI 指数值	0~50	51~100	101~150	151~200	201~300	>300

空气质量	优	良	轻度污染	中度污染	重度污染	严重污染

为监测某化工厂排放废气对周边空气质量指数的影响，某科学兴趣小组在工厂附近某处测得10月1日—20日AQI的数据并绘成折线图如下：



下列叙述正确的是（　　）

A．这20天中AQI的中位数略大于150

B．10月4日到10月11日，空气质量越来越好

C．这20天中的空气质量为优的天数占25%

D．10月上旬AQI的极差大于中旬AQI的极差

6．（2023上·上海松江·高三统考期末）如图所示的茎叶图记录了甲、乙两支篮球队各6名队员某场比赛的得分数据（单位：分）．则下列说法正确的是（　　）

甲队		乙队
7	0	8 9
2 6	1	9 7
0 2	2	7 8
1	3	

A．甲队数据的中位数大于乙队数据的中位数；

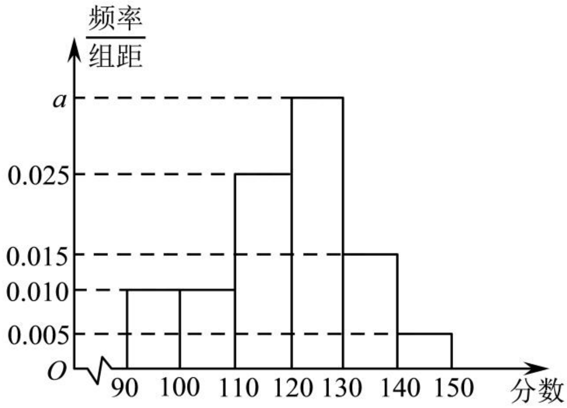
B．甲队数据的平均值小于乙队数据的平均值；

C．甲队数据的标准差大于乙队数据的标准差；

D．乙队数据的第75百分位数为27．

7．（2023上·上海黄浦·高三统考期中）某校共有400名学生参加了趣味知识竞赛（满分：150分），且每位学生的竞赛成绩均不低于90分．将这400名学生的竞赛成绩分组如下：

$[90,100),[100,110),[110,120),[120,130),[130,140),[140,150]$，得到的频率分布直方图如图所示，则这400名学生中竞赛成绩不低于120分的人数为__________.



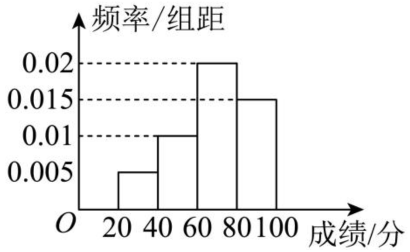
8．（2023·上海崇明·统考一模）如图是小王同学在篮球赛中得分记录的茎叶图，则他平均每场得_______分．

0	3　5　7　8
1	0　1　2　0　0　4

9．（2023上·上海黄浦·高三统考期中）某城市30天的空气质量指数如下：29，26，28，29，38，29，26，26，40，31，35，44，33，28，80，86，65，53，70，34，36，$4y$，31，38，63，60，56，34，74，34.则这组数据的第75百分位数为__________.

10．（2023·上海青浦·统考一模）某家大型超市统计了八次节假日的客流量（单位：百人）分别为29，30，38，25，37，40，42，32，那么这组数据的第75百分位数为_______.

11．（2023·上海徐汇·统考一模）某学校组织全校学生参加网络安全知识竞赛，成绩（单位：分）的频率分布直方图如图所示，数据的分组依次为$[20,40),[40,60),[60,80),[80,100]$，若该校的学生总人数为1000，则成绩低于60分的学生人数为__.



12．（2023·上海奉贤·统考一模）某连锁便利店从2014年到2018年销售商品品种为2000种，从2019年开始，该便利店进行了全面升级，销售商品品种为3000种.下表中列出了从2014年到2023年的利润额.

年份x	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
利润额y/万元	27.6	42.0	38.4	48.0	63.6	63.7	72.8	80.1	60.5	99.3

(1)若某年的利润额超过45.0万元，则该便利店当年会被评选为示范店；若利润额不超过45.0万元，则该便利店当年不会被评选为示范店.试完成2×2列联表，并判断商品品种数量与便利店是否为示范店有关？（显著性水平$\alpha=0.05$，$P(\chi^2\geq3.841)\approx0.05$）

	品种为2000种	品种为3000种	总计
被评为示范店次数			
未被评为示范店次数			
总计			

(2)请根据2014年至2023年（剔除2022年的数据）的数据建立y与x的线性回归模型①；根据2019年至2023年的数据建立y与x的线性回归模型②.分别用这两个模型，预测2024年该便利店的利润额并说明这样的预测值是否可靠？（回归系数精确到0.001，利润精确到0.1万元）

回归系数$\hat{a}$与$\hat{b}$的公式如下：

$$\hat{a}=\frac{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)\left(y_i-\overline{y}\right)}{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2}=\frac{\sum_{i=1}^{n}x_iy_i-n\overline{x}\overline{y}}{\sum_{i=1}^{n}x_i^2-n\overline{x}^2},\ \hat{b}=\overline{y}-\hat{a}\overline{x}=\frac{\sum_{i=1}^{n}y_i-\hat{a}\sum_{i=1}^{n}x_i}{n}$$

二、概率

13.（2023上·上海浦东新·高三统考期末）在100件产品中有90件一等品、10件二等品，从中随机抽取3件产品，则恰好含1件二等品的概率为____________（结果精确到0.01）.

14.（2023·上海金山·统考一模）从1，2，3，4，5这五个数中随机抽取两个不同的数，则所抽到的两个数的和大于6的概率为__________（结果用数值表示）.

15.（2023·上海青浦·统考一模）2023年10月25日至11月12日，青浦曲水园推出以“曲水流觞·花趣水乡”为主题的菊花展．花展结束后，园方挑选数百盆菊花免费赠送给市民．其中有红色、黄色、橙色菊花各1盆，分别赠送给甲、乙、丙三人，每人1盆，则甲没有拿到橙色菊花的概率是___.

16.（2023·上海崇明·统考一模）已知事件A与事件B相互独立，如果$P(A)=0.4$，$P(B)=0.7$，则$P(\overline{A}\cap B)=$___________.

17.（2023·上海杨浦·统考一模）甲和乙两射手射击同一目标，命中的概率分别为0.7和0.8，两人各射击一次，假设事件“甲命中”与“乙命中”是独立的，则至少一人命中目标的概率为________.

18．（2023 上·上海浦东新·高三统考期末）已知事件 A 与事件 B 互斥，且 $P(A)=0.3$，$P(B)=0.4$，则 $P(A\cup B)=$________.

19．（2023 上·上海松江·高三统考期末）有5名同学报名参加暑期区科技馆志愿者活动，共服务两天，每天需要两人参加活动，则恰有1人连续参加两天志愿者活动的概率为________.

20．（2023·上海嘉定·统考一模）已知 11 个大小相同的球，其中 3 个是红球，3 个是黑球，5 个是白球，从中随机取出 4 个形成一组，其中三种颜色都有的概率为____________.

21．（2023·上海闵行·统考一模）2023 年 9 月 23 日至 10 月 8 日，第 19 届亚运会在杭州成功举办，杭州亚运会的志愿者被称为“小青荷”．某运动场馆内共有小青荷 36 名，其中男生 12 名，女生 24 名，这些小青荷中会说日语和会说韩语的人数统计如下：

	男生小青荷	女生小青荷
会说日语	8	12
会说韩语	m	n

其中 m、n 均为正整数，$6\le m\le 8$．

(1)从这 36 名小青荷中随机抽取两名作为某活动主持人，求抽取的两名小青荷中至少有一名会说日语的概率；

(2)从这些小青荷中随机抽取一名去接待外宾，用 A 表示事件“抽到的小青荷是男生”，用 B 表示事件“抽到的小青荷会说韩语”．试给出一组 m、n 的值，使得事件 A 与 B 相互独立，并说明理由．

22．（2023·上海宝山·统考一模）一个盒子中装有4张卡片，卡片上分别写有数字1、2、3、4．现从盒子中随机抽取卡片．

(1)若一次抽取3张卡片，事件 A 表示“3张卡片上数字之和大于7”，求 $P(A)$；

(2)若第一次抽取1张卡片，放回后再抽取1张卡片，事件 B 表示“两次抽取的卡片上数字之和大于6”，求 $P(B)$；

(3)若一次抽取2张卡片，事件 C 表示“2张卡片上数字之和是3的倍数”，事件 D 表示“2张卡片上数字之积是4的倍数”.验证 C、D 是独立的．

23．（2023·上海长宁·统考一模）已知等差数列 $\{a_n\}$ 的前 n 项和为 S_n，公差 $d=2$．

(1)若 $S_{10}=100$，求 $\{a_n\}$ 的通项公式；

(2)从集合 $\{a_1,a_2,a_3,a_4,a_5,a_6\}$ 中任取 3 个元素，记这 3 个元素能成等差数列为事件 A，求事件 A 发生的概率 $P(A)$．

三、随机变量及其分布

24．（2023·上海长宁·统考一模）“$P(A\cap B)=P(A)P(B)$”是“事件A与事件$\overline{B}$互相独立”（　）

A．充分不必要条件　　B．必要不充分条件

C．充要条件　　D．既不充分也不必要条件

25．（2023·上海嘉定·统考一模）已知事件A和B独立，$P(A)=\dfrac{1}{4},P(B)=\dfrac{1}{13}$，则$P(A\cap B)=$____________．

26．（2023·上海奉贤·统考一模）某公司生产的糖果每包标识质量是$500\,\mathrm{g}$，但公司承认实际质量存在误差．已知糖果的实际质量X服从$\mu=500$的正态分布．若随意买一包糖果，假设质量误差超过5克的可能性为p，则$P(495\le X\le 500)$的值为____________．（用含p的代数式表达）

概率统计

学校:___________姓名：___________班级：___________考号：___________

一、统计

1．（2023·上海嘉定·统考一模）两位跳水运动员甲和乙，某次比赛中的得分如下表所示，则正确的选项为（　）

A．甲和乙的中位数相等，甲的平均分小于乙

B．甲的平均分大于乙，甲的方差大于乙

C．甲的平均分大于乙，甲的方差等于乙

D．甲的平均分大于乙，甲的方差小于乙

【答案】B

【分析】计算出两者的中位数，平均分和方差，比较后得到结论.

【详解】甲的比赛得分从小到大排序为$75.9,85.5,86.4,94.4,96$，

选择第三个数86.4作为中位数，

甲的平均分为$\frac{75.9+85.5+86.4+94.4+96}{5}=87.64$，

甲的方差为$\frac{(75.9-87.64)^2+(85.5-87.64)^2+(86.4-87.64)^2+(94.4-87.64)^2+(96-87.64)^2}{5}=50.8568$，

乙的比赛得分从小到大排序为$79.5,80,86.4,94.05,95.7$，

选择第三个数86.4作为中位数，

乙的平均分为$\frac{79.5+80+86.4+94.05+95.7}{5}=87.13$，

乙的方差为$\frac{(79.5-87.13)^2+(80-87.13)^2+(86.4-87.13)^2+(94.05-87.13)^2+(95.7-87.13)^2}{5}=46.1836$，

甲和乙的中位数相等，因为$87.64>87.13$，故甲的平均分大于乙的平均数，

因为$50.8568>46.1836$，所以甲的方差大于乙的方差.

故选：B

2．（2023·上海普陀·统考一模）已知一组数据 3、1、5、3、2，现加入p，q两数对该组数据进行处理，若经过处理后的这组数据的极差为$p-q$，则经过处理后的这组数据与之前的那组数据相比，一定会变大的数字特征是（　）

A．平均数　　B．方差　　C．众数　　D．中位数

【答案】B

【分析】根据平均数、方差、众数和中位数的概念，并通过举反例即可判断.

【详解】对 A，将原数据从小到大进行排序得 1,2,3,3,5；其平均数为 $\bar{x}=\frac{1+2+3+3+5}{5}=\frac{14}{5}$，众数为 3，中位数为 3，

若加入的数据为 $p=-\frac{2}{5},q=6$，则平均数 $\bar{x}'=\frac{-\frac{2}{5}+1+2+3+3+5+6}{7}=\frac{14}{5}$，众数为 3，中位数为 3，平均数、众数和中位数均不变，故 ACD 错误；

对 B，因为加入 p，q 两数后，极差变为 $p-q$，则数据波动程度变大，则方差一定变大，故 B 正确.

故选：B.

3．（2023·上海闵行·统考一模）某校读书节期间，共 120 名同学获奖（分金、银、铜三个等级），从中随机抽取 24 名同学参加交流会，若按高一、高二、高三分层随机抽样，则高一年级需抽取 6 人；若按获奖等级分层随机抽样，则金奖获得者需抽取 4 人．下列说法正确的是（　）

A．高二和高三年级获奖同学共 80 人　　B．获奖同学中金奖所占比例一定最低

C．获奖同学中金奖所占比例可能最高　　D．获金奖的同学可能都在高一年级

【答案】D

【分析】直接根据分层抽样的比例关系计算得到答案.

【详解】对选项 A：高二和高三年级获奖同学共 $120-120\times\frac{6}{24}=90$，错误；

对选项 B：不能确定银奖和铜奖的人数，错误；

对选项 C：金奖人数为 $120\times\frac{4}{24}=20$，银奖和铜奖的人数和为 100 人，

故获奖同学中金奖所占比例不可能最高，错误；

对选项 D：高一年级人数为 30，金奖人数为 20，故获金奖的同学可能都在高一年级，

正确；

故选：D

4．（2023·上海宝山·统考一模）下列说法中错误的是（　）

A．一组数据的平均数、中位数可能相同

B．一组数据中比中位数大的数和比中位数小的数一样多

C．平均数、众数和中位数都是描述一组数据的集中趋势的统计量

D．极差、方差、标准差都是描述一组数据的离散程度的统计量

【答案】B

【分析】A 选项，可举出实例；B 选项，可举出反例；CD 选项，根据平均数、众数和中位数，极差、方差、标准差的定义进行判断.

【详解】A 选项，例如$1,2,3$，这组数据的平均数、中位数相同，均为 2，A 正确；

B 选项，例如$1,1,2,2,5$，中位数为 2，这组数据中比中位数大的数只有 1 个，比中位数小的数有 2 个，两者不一样多，B 错误；

C 选项，平均数、众数和中位数都是描述一组数据的集中趋势的统计量，C 正确；

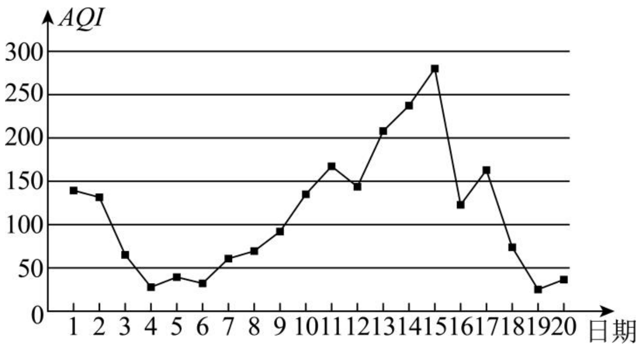
D 选项，极差、方差、标准差都是描述一组数据的离散程度的统计量，D 正确.

故选：B

5.（2023 上·上海虹口·高三统考期末）空气质量指数 AQI 是反映空气质量状况的指数，其对应关系如下表：

AQI 指数值	0~50	51~100	101~150	151~200	201~300	>300
空气质量	优	良	轻度污染	中度污染	重度污染	严重污染

为监测某化工厂排放废气对周边空气质量指数的影响，某科学兴趣小组在工厂附近某处测得 10 月 1 日—20 日 AQI 的数据并绘成折线图如下：



下列叙述正确的是（　　）

A．这 20 天中 AQI 的中位数略大于 150

B．10 月 4 日到 10 月 11 日，空气质量越来越好

C．这 20 天中的空气质量为优的天数占 25%

D．10 月上旬 AQI 的极差大于中旬 AQI 的极差

【答案】C

【分析】利用折线图中数据信息以及变换趋势，对选项一一分析判断即可.

【详解】对于 A，由折线图知 100 以上有 10 个，100 以下有 10 个，中位数是 100 两边最近的两个数的均值，观察这两个数，比 100 大的数离 100 远点，因此两者均值大于 100 但小于 150，故 A 错误；

对于 B，由折线图知 10 月 4 日到 10 月 11 日，AQI 越来越大，则空气质量越来越差，故 B 错误；

对于 C，由折线图知 AQI 小于 50 的有 5 天，则 20 天中的空气质量为优的天数占 25%，故 C 正确；

对于 D，由折线图知 10 月上旬 AQI 的最小值与中旬 AQI 的最小值差不多，但 10 月上旬 AQI 的最大值比中旬 AQI 的最大值小的多，则 10 月上旬 AQI 的极差小于中旬 AQI 的极差，故 D 错误；

故选：C.

6．（2023 上·上海松江·高三统考期末）如图所示的茎叶图记录了甲、乙两支篮球队各 6 名队员某场比赛的得分数据（单位：分）．则下列说法正确的是（　）

甲队		乙队
7	0	8　9
2　6	1	9　7
0　2	2	7　8
1	3	

A．甲队数据的中位数大于乙队数据的中位数；

B．甲队数据的平均值小于乙队数据的平均值；

C．甲队数据的标准差大于乙队数据的标准差；

D．乙队数据的第 75 百分位数为 27．

【答案】D

【分析】根据中位数、平均数、方程、百分位数等知识对选项进行分析，从而确定正确答案．

【详解】A 选项，甲队的中位数是 $\frac{16+20}{2}=18$，乙队的中位数是 $\frac{17+19}{2}=18$，

两者相等，所以 A 选项错误．

B 选项，甲队的平均数为 $\frac{7+12+16+20+22+31}{6}=\frac{108}{6}=18$，

乙队的平均数为 $\frac{8+9+17+19+27+28}{6}=\frac{108}{6}=18$，

两者相等，所以 B 选项错误．

C 选项，甲队的标准差为：

$$\sqrt{\frac{(7-18)^2+(12-18)^2+(16-18)^2+(20-18)^2+(22-18)^2+(31-18)^2}{6}}=\sqrt{\frac{175}{3}},$$

乙队的标准差为：

$$\sqrt{\frac{(8-18)^2+(9-18)^2+(17-18)^2+(19-18)^2+(27-18)^2+(28-18)^2}{6}}=\sqrt{\frac{182}{3}},$$

所以甲队数据的标准差小于乙队数据的标准差，所以 C 选项错误．

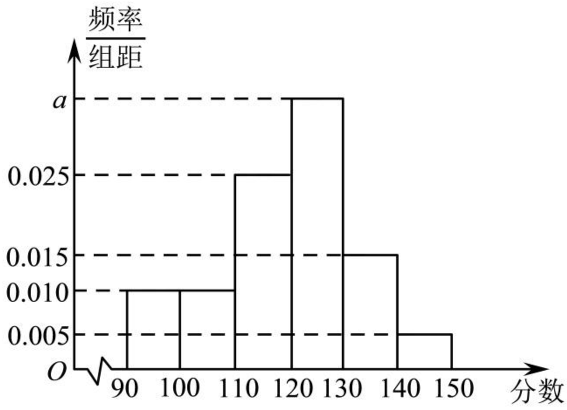
D 选项，乙队的数据为 $8,9,17,19,27,28$，$6\times0.75=4.5$，

所以乙队数据的第 75 百分位数为 27，D 选项正确．

故选：D

7．（2023 上·上海黄浦·高三统考期中）某校共有 400 名学生参加了趣味知识竞赛（满分：150 分），且每位

学生的竞赛成绩均不低于 90 分．将这 400 名学生的竞赛成绩分组如下：$[90,100),[100,110),[110,120),[120,130),[130,140),[140,150]$，得到的频率分布直方图如图所示，则这 400 名学生中竞赛成绩不低于 120 分的人数为__________．



【答案】220

【分析】由频率分布直方图的面积和为1求出a，再计算出结果即可.

【详解】由频率分布直方图可知$(0.010+0.010+0.025+a+0.015+0.005)\times 10=1$，解得$a=0.035$，

这 400 名学生中竞赛成绩不低于 120 分的人数为$400\times(0.035+0.015+0.005)\times 10=220$，

故答案为：220

8．（2023·上海崇明·统考一模）如图是小王同学在篮球赛中得分记录的茎叶图，则他平均每场得_______分．

0	3　5　7　8
1	0　1　2　0　0　4

【答案】9

【分析】根据平均数的求法求得平均数.

【详解】平均数为$\frac{3+5+7+8+10+11+12+10+10+14}{10}=9$.

故答案为：9

9．（2023 上·上海黄浦·高三统考期中）某城市 30 天的空气质量指数如下：29，26，28，29，38，29，26，26，40，31，35，44，33，28，80，86，65，53，70，34，36，$4y$，31，38，63，60，56，34，74，34.则这组数据的第 75 百分位数为__________．

【答案】56

【分析】把给定数据按由小到大的顺序排列，再根据第p百分位数的定义求解即得.

【详解】显然$40\le 4y<50$，30 个数据由小到大排列为：

26，26，26，28，28，29，29，29，31，31，33，34，34，34，35，36，38，38，40，

44，$4y$，53，56，60，63，65，70，74，80，86，

或者 26，26，26，28，28，29，29，29，31，31，33，34，34，34，35，36，38，38，

40，$4y$，44，53，56，60，63，65，70，74，80，86，

由$30\times75\%=22.5$，得这组数据的第 75 百分位数为上述排列后的从小到大的第 23 个数 56.

故答案为：56

10.（2023·上海青浦·统考一模）某家大型超市统计了八次节假日的客流量（单位：百人）分别为 29，30，38，25，37，40，42，32，那么这组数据的第75百分位数为_______.

【答案】39

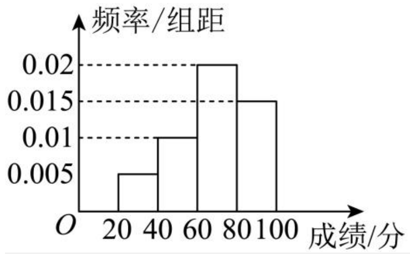
【分析】根据第 75 百分位数的定义计算可得答案.

【详解】将这 8 个数据按从小到大排列为：25，29，30，32，37，38，40，42，

因为$8\times75\%=6$，所以第 75 百分位数为$\dfrac{38+40}{2}=39$.

故答案为：39.

11.（2023·上海徐汇·统考一模）某学校组织全校学生参加网络安全知识竞赛，成绩（单位：分）的频率分布直方图如图所示，数据的分组依次为$[20,40),[40,60),[60,80),[80,100]$，若该校的学生总人数为 1000，则成绩低于 60 分的学生人数为__.



【答案】300

【分析】先利用频率分布直方图求得成绩低于 60 分的频率，进而求得该校成绩低于 60 分的学生人数.

【详解】图中成绩低于 60 分的频率为$20(0.01+0.005)=0.3$，

则该校成绩低于 60 分的学生人数为$1000\times0.3=300$（人）

故答案为： 300

12.（2023·上海奉贤·统考一模）某连锁便利店从2014年到2018年销售商品品种为2000种，从2019年开始，该便利店进行了全面升级，销售商品品种为3000种.下表中列出了从2014年到2023年的利润额.

年份x	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
利润额y/万元	27.6	42.0	38.4	48.0	63.6	63.7	72.8	80.1	60.5	99.3

(1)若某年的利润额超过45.0万元，则该便利店当年会被评选为示范店；若利润额不超过45.0万元，则该便利店当年不会被评选为示范店.试完成2×2列联表，并判断商品品种数量与便利店是否为示范店有关？（显

著性水平 $\alpha=0.05$，$P(\chi^2\geq 3.841)\approx 0.05$）

	品种为2000种	品种为3000种	总计
被评为示范店次数			
未被评为示范店次数			
总计			

(2)请根据2014年至2023年（剔除2022年的数据）的数据建立y与x的线性回归模型①；根据2019年至2023年的数据建立y与x的线性回归模型②.分别用这两个模型，预测2024年该便利店的利润额并说明这样的预测值是否可靠？（回归系数精确到0.001，利润精确到0.1万元）

回归系数$\hat{a}$与$\hat{b}$的公式如下：

$$\hat{a}=\frac{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)\left(y_i-\overline{y}\right)}{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2}=\frac{\sum_{i=1}^{n}x_iy_i-n\overline{x}\overline{y}}{\sum_{i=1}^{n}x_i^2-n\overline{x}^2},\ \hat{b}=\overline{y}-\hat{a}\overline{x}=\frac{\sum_{i=1}^{n}y_i-\hat{a}\sum_{i=1}^{n}x_i}{n}$$

【答案】(1)列联表见解析，商品品种的提升与该便利店是否是示范店有关.

(2)答案见解析

【分析】（1）列出列联表后计算出χ^2后，与3.841比较大小即可得；

（2）分别计算出线性回归模型后，结合所得数据进行判断即可得.

【详解】（1）列联表为

	品种为2000种	品种为3000种	总计
被评为示范店次数	2	5	7
未被评为示范店次数	3	0	3
总计	5	5	10

$\chi^2=\frac{10(0-15)^2}{5\times 5\times 7\times 3}\approx 4.29>3.841$，

可以判断商品品种的提升与该便利店是否是示范店有关.

（2）线性回归模型①：

$\overline{x}=\frac{2014+2015+2016+2017+2018+2019+2020+2021+2023}{9}\approx 2018.1111$，

$\overline{y}=\frac{27.6+42+38.4+48+63.6+63.7+72.8+80.1+99.3}{9}=59.5$，

则 $\hat{a}=\dfrac{\sum_{i=1}^{n}x_iy_i-n\overline{x}\overline{y}}{\sum_{i=1}^{n}x_i^2-n\overline{x}^2}\approx 7.627$，

则 $\hat{b}=\overline{y}-\hat{a}\overline{x}=59.5-7.627\times 2018.1111\approx -15332.20$，

故 $\hat{y}=7.627x-15332.20$，

当 $x=2024$ 时，预测值为 $\hat{y}=7.627\times 2024-15332.20\approx 104.9$；

线性回归模型②：

$\overline{x}'=\dfrac{2019+2020+2021+2022+2023}{5}=2021$，

$\overline{y}'=\dfrac{63.7+72.8+80.1+60.5+99.3}{5}=75.28$，

则 $\hat{a}'=\dfrac{\sum_{i=1}^{n}x_iy_i-n\overline{x}'\cdot\overline{y}'}{\sum_{i=1}^{n}x_i^2-n\overline{x}'^2}\approx 5.89$，

$\hat{b}'=\overline{y}'-\hat{a}'\overline{x}'=75.28-5.89\times 2021\approx -11828.41$，

故 $\hat{y}=5.89x-11828.41$，

当 $x=2024$ 时，预测值为 $\hat{y}=5.89\times 2024-11828.41=93.0$.

模型①的预测不可靠，根据（1）可以知道商品品种与便利店的品质有关，影响了利润额，

因此按照经济发展规律，应该用比较新的数据即品种为 3000 种的数据进行预测；

模型②的预测不可靠，2022 年可能因为受疫情影响或者其它不可因素，

其利润额 60.5 为异常数据，应该剔除.

二、概率

13．（2023 上·上海浦东新·高三统考期末）在 100 件产品中有 90 件一等品、10 件二等品，从中随机抽取 3 件产品，则恰好含 1 件二等品的概率为____________（结果精确到 0.01）.

【答案】0.25

【分析】由题意先求出事件总数，再求出恰好有一件二等品的事件，结合古典概型的概率公式计算即可求解.

【详解】从这批产品中抽取 3 件，则事件总数为 C_{100}^{3}，

其中恰好有一件二等品的事件有 $C_{90}^{2}C_{10}^{1}$，

所以恰好有一件二等品的概率为 $P=\dfrac{C_{90}^{2}C_{10}^{1}}{C_{100}^{3}}=\dfrac{267}{1078}\approx 0.25$.

故答案为：0.25

14.（2023·上海金山·统考一模）从 1，2，3，4，5 这五个数中随机抽取两个不同的数，则所抽到的两个数的和大于 6 的概率为__________（结果用数值表示）.

【答案】$\frac{2}{5}$/0.4

【分析】求出所有的基本事件个数以及符合题意的基本事件个数，利用古典概型求概率即可.

【详解】根据题意，从 1，2，3，4，5 这五个数中随机抽取两个不同的数共有 $C_5^2=10$，

所抽到两个数的和大于 6 共有 $(2,5)$，$(3,5)$，$(4,5)$，$(3,4)$共 4 种，

所以所抽到的两个数的和大于 6 的概率为 $P=\frac{4}{10}=\frac{2}{5}$.

故答案为：$\frac{2}{5}$

15.（2023·上海青浦·统考一模）2023 年 10 月 25 日至 11 月 12 日，青浦曲水园推出以“曲水流觞·花趣水乡”为主题的菊花展．花展结束后，园方挑选数百盆菊花免费赠送给市民．其中有红色、黄色、橙色菊花各1盆，分别赠送给甲、乙、丙三人，每人1盆，则甲没有拿到橙色菊花的概率是___.

【答案】$\frac{2}{3}$

【分析】根据题意甲、乙、丙三人拿到橙色菊花概率相等，都为$\frac{1}{3}$，进而求出甲没有拿到橙色菊花的概率.

【详解】设事件 $A=$ 甲拿到橙色菊花，

根据题意有红色、黄色、橙色菊花各1盆，分别赠送给甲、乙、丙三人，每人1盆，

甲、乙、丙三人拿到橙色菊花概率相等，都为$\frac{1}{3}$，

所以 $P(A)=\frac{1}{3}$，则甲没有拿到橙色菊花的概率 $P(\overline{A})=1-P(A)=1-\frac{1}{3}=\frac{2}{3}$.

故答案为：$\frac{2}{3}$

16.（2023·上海崇明·统考一模）已知事件 A 与事件 B 相互独立，如果 $P(A)=0.4$，$P(B)=0.7$，则

$P(\overline{A}\cap B)=$___________.

【答案】$0.42/\frac{21}{50}$

【分析】根据独立事件和对立事件的概率公式计算可得答案

【详解】由事件 A 与事件 B 相互独立，则事件 $\overline{A}$ 与事件 B 相互独立，

又 $P(A)=0.4$，$P(B)=0.7$，

则 $P(\overline{A}\cap B)=P(\overline{A})P(B)=(1-P(A))P(B)=(1-0.4)\times 0.7=0.42$

故答案为：0.42.

17.（2023·上海杨浦·统考一模）甲和乙两射手射击同一目标，命中的概率分别为 0.7 和 0.8，两人各射击

一次，假设事件“甲命中”与“乙命中”是独立的，则至少一人命中目标的概率为________.

【答案】$0.94/\frac{47}{50}$

【分析】利用独立事件的乘法公式分别求出“仅有一人命中目标”的概率和“两人同时命中目标”的概率，即可得出结果.

【详解】根据题意可知“至少一人命中目标”包括“仅有一人命中目标”和“两人同时命中目标”两个基本事件;

可得“仅有一人命中目标”的概率为$P_1=(1-0.7)\times0.8+0.7\times(1-0.8)=0.38$;

“两人同时命中目标”的概率为$P_2=0.7\times0.8=0.56$;

所以至少一人命中目标的概率为$P=P_1+P_2=0.94$.

故答案为：0.94

18.（2023上·上海浦东新·高三统考期末）已知事件A与事件B互斥，且$P(A)=0.3$，$P(B)=0.4$，则$P(A\cup B)=$________.

【答案】$0.7/\frac{7}{10}$

【分析】根据互斥事件的概率加法公式，即可求解.

【详解】因为随机事件A与B互斥，且$P(A)=0.3$，$P(B)=0.4$，

所以$P(A\cup B)=P(A)+P(B)=0.3+0.4=0.7$.

故答案：0.7.

19.（2023上·上海松江·高三统考期末）有5名同学报名参加暑期区科技馆志愿者活动，共服务两天，每天需要两人参加活动，则恰有1人连续参加两天志愿者活动的概率为________.

【答案】$\frac{3}{5}$

【分析】由分布乘法计数原理的知识结合古典概型的概率公式可解.

【详解】每天从5名同学中抽取2名参加志愿者活动，一共有$C_5^2C_5^2=100$种方式，

恰有一人连续参加两天志愿者活动有$C_5^1C_4^1C_3^1=60$种方式，

由古典概型的概率公式可得恰有1人连续参加两天志愿者活动的概率为$\frac{60}{100}=\frac{3}{5}$，

故答案为：$\frac{3}{5}$.

20.（2023·上海嘉定·统考一模）已知11个大小相同的球，其中3个是红球，3个是黑球，5个是白球，从中随机取出4个形成一组，其中三种颜色都有的概率为____________.

【答案】$\frac{6}{11}$

【分析】4 个球有三个颜色，肯定有两个球同色，按同色的球的颜色分情况讨论，再结合古典概型概率的计算公式可求答案.

【详解】从 11 个球中随机取出 4 个球的取法有：$C_{11}^{4}=\frac{11\times10\times9\times8}{4\times3\times2\times1}=330$.

又 4 个球有三种颜色，所以必定有且只有两个球同色.

若同色的两个球为红色，满足条件的取法有：$C_3^2C_3^1C_5^1=45$；

若同色的两个球为黑色，满足条件的取法有：$C_3^1C_3^2C_5^1=45$；

若同色的两个球为白色，满足条件的取法有：$C_3^1C_3^1C_5^2=90$.

∴取出的 4 个球中三种颜色都有的概率为：$P=\frac{45+45+90}{330}=\frac{6}{11}$

故答案为：$\frac{6}{11}$

21．（2023·上海闵行·统考一模）2023 年 9 月 23 日至 10 月 8 日，第 19 届亚运会在杭州成功举办，杭州亚运会的志愿者被称为“小青荷”．某运动场馆内共有小青荷 36 名，其中男生 12 名，女生 24 名，这些小青荷中会说日语和会说韩语的人数统计如下：

	男生小青荷	女生小青荷
会说日语	8	12
会说韩语	m	n

其中 m、n 均为正整数，$6\le m\le 8$.

(1)从这 36 名小青荷中随机抽取两名作为某活动主持人，求抽取的两名小青荷中至少有一名会说日语的概率；

(2)从这些小青荷中随机抽取一名去接待外宾，用 A 表示事件“抽到的小青荷是男生”，用 B 表示事件“抽到的小青荷会说韩语”．试给出一组 m、n 的值，使得事件 A 与 B 相互独立，并说明理由．

【答案】(1)$\frac{17}{21}$

(2)$m=6$，$n=12$ 或 $m=7$，$n=14$ 或 $m=8$，$n=16$，均符合题意.理由见解析

【分析】(1）求出从这 36 名小青荷中随机抽取两名作为某活动主持人，共有几种抽法，再求出抽取的两名小青荷中至少有一名会说日语的抽法，根据古典概型的概率公式即可求得答案；

（2）分别求出事件 A,B,AB 的概率的值或表达式，根据独立事件的乘法公式列式计算，即可求得答案.

【详解】（1）从这 36 名小青荷中随机抽取两名作为某活动主持人，共有 C_{36}^{2} 种抽法，

抽取的两名小青荷中至少有一名会说日语的抽法有 $(C_{20}^{1}C_{16}^{1}+C_{20}^{2}C_{16}^{0})$ 种，

故抽取的两名小青荷中至少有一名会说日语的概率为 $P=\dfrac{C_{20}^{1}C_{16}^{1}+C_{20}^{2}C_{16}^{0}}{C_{36}^{2}}=\dfrac{20\times16+10\times19}{18\times35}=\dfrac{17}{21}$；

（2）由题意得 $P(A)=\dfrac{12}{36}=\dfrac{1}{3}$，$P(B)=\dfrac{m+n}{36}$，$P(AB)=\dfrac{m}{36}$，

要使得事件 A 与 B 相互独立，则需满足 $P(A)P(B)=P(AB)$，

即 $\dfrac{1}{3}\times\dfrac{m+n}{36}=\dfrac{m}{36}$，即 $n=2m$，

由于 $6\le m\le 8$，故 $m=6$ 时，$n=12$；$m=7$ 时，$n=14$；

$m=8$ 时，$n=16$，均符合题意，取其中一组即可.

22．（2023·上海宝山·统考一模）一个盒子中装有 4 张卡片，卡片上分别写有数字 1、2、3、4．现从盒子中随机抽取卡片.

(1)若一次抽取 3 张卡片，事件 A 表示“3 张卡片上数字之和大于 7”，求 $P(A)$；

(2)若第一次抽取 1 张卡片，放回后再抽取 1 张卡片，事件 B 表示“两次抽取的卡片上数字之和大于 6”，求 $P(B)$；

(3)若一次抽取 2 张卡片，事件 C 表示“2 张卡片上数字之和是 3 的倍数”，事件 D 表示“2 张卡片上数字之积是 4 的倍数”.验证 C、D 是独立的.

【答案】(1) $\dfrac{1}{2}$

(2) $\dfrac{3}{16}$

(3)事件 C 与事件 D 是独立

【分析】（1）利用古典概型的概率求解；

（2）利用古典概型的概率求解；

（3）利用古典概型的概率分别求得 $P(C)$，$P(D)$，$P(C\cap D)$ 判断.

【详解】（1）解：若一次抽取 3 张卡片，共包含 $(1,2,3)$、$(1,2,4)$、$(1,3,4)$、$(2,3,4)$ 共 4 个基本事件.

其中事件 $A=\{(1,3,4),(2,3,4)\}$ 包含 2 个基本事件

所以 $P(A)=\dfrac{2}{4}=\dfrac{1}{2}$；

（2）若第一次抽取 1 张卡片，放回后再抽取 1 张卡片，共包含 $4\times4=16$ 个基本事件，

其中事件 $B=\{(3,4)、(4,4)、(4,3)\}$ 包含 3 个基本事件

所以 $P(B)=\frac{3}{16}$

（3）一次抽取 2 张卡片,共包含 $C_4^2=6$ 个基本事件，

事件 $C=\{(1,2),(2,4)\}$，

所以 $P(C)=\frac{2}{6}=\frac{1}{3}$

事件 $D=\{(1,4)、(2,4)、(3,4)\}$，所以 $P(D)=\frac{3}{6}=\frac{1}{2}$

当 C、D 同时发生，即 2 张卡片上数字之和是 3 的倍数同时积是 4 的倍数，只有一种取法 $(2,4)$，

所以 $P(C\cap D)=\frac{1}{6}$

因为 $P(C\cap D)=P(C)P(D)$，

所以事件 C 与事件 D 是独立的.

23.（2023·上海长宁·统考一模）已知等差数列 $\{a_n\}$ 的前 n 项和为 S_n，公差 $d=2$.

(1)若 $S_{10}=100$，求 $\{a_n\}$ 的通项公式；

(2)从集合 $\{a_1,a_2,a_3,a_4,a_5,a_6\}$ 中任取 3 个元素，记这 3 个元素能成等差数列为事件 A，求事件 A 发生的概率 $P(A)$.

【答案】(1) $a_n=2n-1$

(2) $\frac{3}{10}$

【分析】（1）根据题意，利用等差数列的求和公式，列出方程，求得 $a_1=1$，进而求得数列的通项公式；

（2）根据题意，得到所有的不同取法有 20 种，再利用列举法求得事件 A 中所包含的基本事件的个数，结合古典概型的概率计算公式，即可求解.

【详解】（1）解：由等差数列 $\{a_n\}$ 的前 n 项和为 S_n，公差 $d=2$，

因为 $S_{10}=100$，可得 $10a_1+\frac{10\times 9}{2}\times 2=100$，解得 $a_1=1$，

所以 $a_n=1+(n-1)\times 2=2n-1$，即数列 $\{a_n\}$ 的通项公式为 $a_n=2n-1$.

（2）解：由题意，从集合 $\{a_1,a_2,a_3,a_4,a_5,a_6\}$ 中任取 3 个元素，共有 20 种不同的取法，

其中这 3 个元素能成等差数列有 $\{(a_1,a_2,a_3),(a_1,a_3,a_5),(a_2,a_3,a_4),(a_2,a_4,a_6),(a_3,a_4,a_5),$

$(a_4,a_5,a_6)\}$，有 6 种不同的取法，

所以事件 A 的概率为 $P(A)=\frac{6}{20}=\frac{3}{10}$.

三、随机变量及其分布

24.（2023·上海长宁·统考一模）“$P(A\cap B)=P(A)P(B)$”是“事件 A 与事件 $\overline{B}$ 互相独立”（　）

A．充分不必要条件　　B．必要不充分条件

C．充要条件　　D．既不充分也不必要条件

【答案】C

【分析】根据事件互斥，对立，独立的关系得出即可.

【详解】因为对于任意两个事件 A,B，如果 $P(A\cap B)=P(A)P(B)$，则事件 A 与事件 B 相互独立，若事件 A 与事件 B 相互独立，则事件 A 与事件 $\overline{B}$ 也互相独立，所以充分性成立；

若事件 A 与事件 $\overline{B}$ 互相独立，则事件 A 与事件 B 也相互独立，则 $P(A\cap B)=P(A)P(B)$ 成立，所以必要性成立；

故选：C

25.（2023·上海嘉定·统考一模）已知事件 A 和 B 独立，$P(A)=\frac{1}{4},P(B)=\frac{1}{13}$，则 $P(A\cap B)=$____________.

【答案】$\frac{1}{52}$

【分析】根据独立事件的概率计算公式直接求解出结果.

【详解】因为事件 A,B 互相独立，

所以 $P(A\cap B)=P(A)P(B)=\frac{1}{4}\times\frac{1}{13}=\frac{1}{52}$，

故答案为：$\frac{1}{52}$.

26.（2023·上海奉贤·统考一模）某公司生产的糖果每包标识质量是 $500\,\mathrm{g}$，但公司承认实际质量存在误差. 已知糖果的实际质量 X 服从 $\mu=500$ 的正态分布．若随意买一包糖果，假设质量误差超过 5 克的可能性为 p，则 $P(495\le X\le 500)$ 的值为____________.　（用含 p 的代数式表达）

【答案】$\frac{1-p}{2}$

【分析】根据正态分布的性质直接求解即可.

【详解】由题知，$\sigma=5$，

则 $P(495\le X\le 500)=P(\mu-\sigma\le X\le\mu)$

$=\frac{1}{2}P(\mu-\sigma\leq X\leq\mu+\sigma)$

$=\frac{1}{2}[1-P(X>\mu+\sigma)-P(X<\mu-\sigma)]$

$=\frac{1}{2}[1-P(X>505)-P(X<495)]$

$=\frac{1}{2}(1-p)$.

故答案为：$\frac{1-p}{2}$