第九章　统计

9.1　随机抽样

9.1.1　简单随机抽样

对点上分

1. C　【解析】了解一批节能灯泡的使用寿命，调查过程带有破坏性，只能采用抽样调查的方式，而不能将整批节能灯泡全部用于试验，**故 A 错误**；要调查某个班级同学的身高，采用全面调查的方式，**故 B 错误**；调查海河某段水域的水质情况，采用抽样调查的方式，**故 C 正确**；调查全市高中生每天的睡眠时间，采用抽样调查的方式，**故 D 错误. 故选 C.**

2. C　【解析】根据总体、个体、样本和样本容量的定义进行判断. 100 名同学的视力情况（数据）是从总体中抽取的一部分个体所组成的集合，所以是总体的一个样本，**故 C 正确.**

3. B　【解析】根据样本容量的定义，12 个班，每班抽取 10 人的志愿，共抽取 120 人的志愿，所以样本容量是 120，**故 B 正确.**

4. ABC　【解析】对于 A，1 000 名运动员的年龄是总体，故正确；对于 B，所抽取的 10 名运动员的年龄是一个样本，故正确；对于 C，样本容量为 10，故正确；对于 D，每名运动员被抽到的机会相等，故错误.

5. ABD　【解析】对于 A，从 20 名同学中随机抽取 5 名同学参加义务劳动，是简单随机抽样，**故 A 正确**；

对于 B，是简单随机抽样，一次性抽取 3 个零件，等价于逐个抽取零件 3 次，**故 B 正确**；

对于 C，不是简单随机抽样，不符合“等可能性”，因为 5 名同学是指定的，而不是随机抽取的，**故 C 错误**；

对于 D，中国福利彩票 30 选 7，得到 7 个

彩票号码,是简单随机抽样,**故 D 正确.**

故选 ABD.

方法总结 **简单随机抽样的特点**

(1)总体中的个体数有限;

(2)等可能抽样.

要判断所给的抽样方法是否是简单随机抽样,关键看它本身的特点与简单随机抽样的特点是否完全符合,其中等可能性是出题的热点.

6. B 【解析】在简单随机抽样过程中,考虑的最主要原则是保证样本能够很好地代表总体. 随机抽样的出发点是使每个个体都有相同的机会被抽中,这是基于对样本数据代表性的考虑,与第几次抽样无关,**故 B 正确.**

易错警示 **认为个体被抽中的概率与抽样顺序有关而致误**

简单随机抽样遵循随机原则,即简单随机抽样不受任何主观因素及其他系统性因素的影响,所以某个个体被抽中的概率是相等的,并且是确定的.

7. B 【解析】第二次抽取时,余下的每个个体被抽到的概率为$\frac{1}{3}$,则$\frac{13-1}{n-1}=\frac{1}{3}$,解得$n=37$,故在整个抽样过程中,每个个体被抽到的可能性为$\frac{13}{37}$. **故 B 正确.**

规律点拨 简单随机抽样中,每个个体被抽到的可能性都是相等的,与抽取的顺序无关. 本题已知的是第二次抽取时余下的每个个体被抽到的概率,此时需要考虑的是还剩下多少个个体.

8. B 【解析】从含有 N 个个体的总体中抽取样本容量为 $n(1\leqslant n<N)$ 的样本,总体和样本容量都不大时,采用抽签法的方式. 由题知 B 选项的总体和样本容量都不大,故可采用抽签法的方式. **故 B 正确.**

规律点拨 应用抽签法的两个关键

一个抽样试验能否用抽签法，关键看两点：一是制签是否方便；二是号签是否容易被搅匀. 一般地，当总体容量和样本容量都较小时可用抽签法的方式. 若总体容量非常大，采用抽签法的方式就比较费时、费力，也不方便，搅拌不均匀有失公平性，从而使产生代表性差的样本的可能性增加.

9. D 【解析】找到第 8 行第 4 列的数开始向右读，第一个数 258 符合条件，第二个数 392 符合条件，第三个数 120 符合条件，第四个数 676 大于 499 要舍去，第五个数 630 大于 499 要舍去，第六个数 164 符合条件，第七个数 785 大于 499 要舍去，第八个数 916 大于 499 要舍去，第九个数 955 大于 499 要舍去，第十个数 567 大于 499 要舍去，第十一个数 199 符合条件，即 258，392，120，164，199 符合条件. **故 D 正确**.

10.【解】第一步：列出只能由 2，3，5，7 中任意两个数（数字可重复）组成的区域代码，共 16 个，用抽签法随机抽取 3 个.

第二步：用抽签法或计算机生成法产生若干个 0~9 之间的随机整数，5 个一组，构成 00000~99999 之间的随机数.

第三步：用随机数表产生随机数的方法选出 15 个不重复的五位数即为所选号码，分成 3 组.

第四步：第 1，2，3 组前分别加上选出的第 1，2，3 个区域代码.

11. D 【解析】根据样本估计总体知，样本平均数具有随机性，只能估计总体平均数，**故 D 正确**.

12. B 【解析】这 3 000 个数据的平均数为

$$\frac{78.1\times800+85\times1\ 300+91.9\times900}{3\ 000}=85.23.$$

用样本平均数估计总体平均数，可估计这 40 000 个数据的平均数为 85. 23.

13.【解】(1) 用现金的顾客的平均消费额为 $\frac{1}{10}\times(19.50+39.50+78.60+35.70+$

31.80+98.30+23.50+108.00+29.00+32.20)=49.61(元),

用电子钱包的顾客的平均消费额为$\frac{1}{12}\times$(45.90+89.70+123.60+98.75+59.30+45.35+109.45+95.50+103.00+24.00+65.50+67.20)≈77.27(元).

(2)由题意得使用现金的顾客的比例为0.32,所以使用电子钱包的顾客的比例为0.68,

所以估计该超市的顾客的平均消费额为49.61×0.32+77.27×0.68≈68.42(元).

9.1.2　分层随机抽样

对点上分

1. C　【解析】①的总体中的个体数较少,宜采用简单随机抽样;

②中300户家庭中收入存在较大差异,层次比较明显,宜采用分层随机抽样. **故选 C.**

2. D　【解析】因为总体是由差异明显的三部分组成,所以考虑用分层随机抽样.

因为总人数为28+54+81=163,样本容量为36,按$\frac{36}{163}$的比例抽样,无法得到整数解,

因此考虑先剔除1人,将抽样比变为$\frac{36}{162}=\frac{2}{9}$.

若从老年人中随机地剔除1人,则老年人应抽取$27\times\frac{2}{9}=6$(人),中年人应抽取$54\times\frac{2}{9}=12$(人),青年人应抽取$81\times\frac{2}{9}=18$(人),从而组成容量为36的样本. **故选 D.**

3. A　【解析】由题可得,被抽到的研发人员有$600\times\frac{120}{1\ 200}=60$(人),销售人员有$400\times\frac{120}{1\ 200}=40$(人),所以被抽到的研发人员人数比销售人员人数多60−40=20. **故 A 正确.**

4. D　【解析】该社区21岁至65岁的居民共有840+700+560=2 100(人),设抽样

调查抽取的总人数为 n，则 $\frac{n}{2\ 100}=\frac{100}{700}$，解得 $n=300$. **故 D 正确.**

5. A 【解析】不妨设总人数为 k，则 $\frac{3}{8}k-\frac{1}{8}k=8$，所以 $k=32$，所以计分组人数为 $32\times\frac{1}{2}=16$.

6. C 【解析】由题意，估计总样本平均数为 $\frac{600}{2\ 000}\times93+\frac{800}{2\ 000}\times81+\frac{600}{2\ 000}\times99=90$.

故 C 正确.

7. C 【解析】设高三年级抽取的学生人数为 x，平均睡眠时间为 t 小时，由题意可得 $\frac{30}{1\ 200+1\ 000+800}=\frac{x}{800}$，解得 $x=8$，同理可得高二年级抽取 10 人，高一年级抽取 12 人，由 $8t+10\times7.8+12\times8.5=30\times8$，解得 $t=7.5$. **故 C 正确.**

一题多解 设高三年级抽取的学生人数为 x，高二年级抽取的学生人数为 y，高一年级抽取的学生人数为 z，高三年级抽取的学生的平均睡眠时间为 t 小时，高三、高二、高一抽取人数之比为 $800:1\ 000:1\ 200=4:5:6=x:y:z$，由题意得 $x+y+z=30$，解得 $x=8$，$y=10$，$z=12$，由 $8t+10\times7.8+12\times8.5=30\times8$，解得 $t=7.5$. **故 C 正确.**

能力上分

1. BD 【解析】因为各地区的饮食习惯各不相同，差异较大，所以用按比例分配的分层随机抽样更合适，

又各地区人数比为 $5:3:2$，若抽取人数为 100，

则抽取东部地区学生 50 人、中部地区学生 30 人、西部地区学生 20 人，**故 A 错误，B 正确；**

采取按比例分配的分层随机抽样，每个学生被抽取的可能性均为 $\frac{100}{5\ 000}=\frac{1}{50}$，各地区学生被选中的可能性大小是相等的，**故 C 错误；**

事件的总体为该高校全体大一新生，**故 D 正确. 故选 BD.**

2. A 【解析】由题意知抽样比为 $\frac{10}{100+300}=\frac{1}{40}$,

则 $\frac{50}{100+300+150+450+z+600}=\frac{1}{40}$，解得 $z=400$. **故选 A.**

3. ACD 【解析】设中年市民、青年市民抽到的人数分别为 x,y,

从而 $\frac{80}{4}=\frac{x}{3}=\frac{y}{3}$，解得 $x=y=60$, $n=\frac{80}{\frac{4}{4+3+3}}=200$,

抽取的中年与青年市民人数之和减去老年市民人数为 $60+60-80=40$，**故 ACD 正确，B 错误. 故选 ACD.**

4. (1) 60%　(2) 36　【解析】(1) 设登山看日出组的人数为 x，在海边看日落组中，设高一学生、高二学生、高三学生所占的比例分别为 a,b,c,

则 $\frac{x\times50\%+3ax}{4x}=20\%$，$\frac{x\times30\%+3bx}{4x}=30\%$，解得 $a=10\%$，$b=30\%$，

所以 $c=1-a-b=1-10\%-30\%=60\%$.

故在海边看日落组中，高三学生所占的比例为 60%.

(2) 由 (1) 可得在海边看日落组中高三学生所占的比例为 60%，

所以在海边看日落组中，高三年级应抽取的人数为 $80\times\frac{3}{4}\times60\%=36$.

5. 【解】(1) 总样本的均值约为 $\frac{400}{600}\times174+\frac{200}{600}\times162=170$ (cm).

(2) 总样本的均值约为 $\frac{100}{200}\times174+\frac{100}{200}\times162=168$ (cm)，

该均值不能作为总体均值的估计值，因为分层随机抽样中未按比例抽样，

总体中每个个体被抽到的可能性不完全相同，所以样本的代表性差.

9.1.3 获取数据的途径

对点上分

1. C 【解析】因为“中国天眼”为 500 米口径球面射电望远镜，易知其获取数据的方式是观察. 故 C 正确.

2. C 【解析】对于 A，调查一个地区糖尿病的发病率，调查数量较多，不适合全面调查，故错误；对于 B，了解一批水稻种子的发芽率，调查数目较多，且具有破坏性，不适合全面调查，故错误；对于 C，了解一个班级学生的身高情况，适合全面调查，故正确；对于 D，了解某城市居民的生活水平，调查数目较多，不适合全面调查，故错误.

3. AB 【解析】由直接获取数据的概念可知，试验、统计调查是直接获取数据的方法，通过收集企业的经营报表数据、某科研机构发布的研究数据获得数据是间接获取数据的方法. 故 A，B 正确.

4. B 【解析】①了解全校同学喜欢的课程的情况，应在各班进行抽样，同时不能仅限于男同学，不合理；②“神舟二十号”飞船发射前，应采用全面调查检查其各零部件的合格情况，不合理；③了解国内外观众对刚上映的某部电影的观影感受，采用抽样调查，合理；④调查某市居民对“垃圾分类”有关内容的了解程度，将问题放到某网站上，受调查人群有局限性，不合理. 故 B 正确.

5. C 【解析】某市中小学生每天的运动时间以及某地 2024 年的交通事故情况，均可通过调查研究获取，某地的降水量可通过观察获取，而某种针对流感病毒的新药的疗效，需通过试验获取. 故 C 正确.

6. C 【解析】对于小凉，邯郸市今年“五一”前后的气温通过观察获取数据；对于小爽，判断某种新药是否有效通过试验获取数据；对于小夏，“山西新闻联播”的收视率通过调查获取数据；对于小天，近年来我市普通高中入学人数通过查询获取数据. 所以通过调查获取数据的是小夏. 故 C 正确.

7.【解】(1)用调查法;(2)用观察法;(3)用试验法;(4)用查询法.

9.1 节测上分

1. C 【解析】样本数据7天代收快递的件数的平均数为$\frac{1}{7}\times(285+367+463+290+335+719+698)=451$,则每月代收快递约为$451\times30=13\ 530$(件),所以该驿站每月收取的服务费约为$13\ 530\times0.8=10\ 824$(元),**故C正确.**

2. A 【解析】对于A,理学比工学抽取的人数多,但张三和李四作为一个个体被抽到的概率相等,故张三与李四被抽到的可能性一样大,**故A错误**;对于B,理学专业应抽取的人数为$100\times30\%=30$,工学专业应抽取的人数为$100\times20\%=20$,**故B正确**;对于C,因为各专业差异比较大,所以采用分层随机抽样比简单随机抽样更合理,**故C正确**;对于D,该问题中的样本容量为100,**故D正确.**

3. C 【解析】用样本平均数估计该小区业主对户型结构满意度的平均分为$8\times\frac{2}{5}+9\times\frac{3}{5}=\frac{43}{5}=8.6$(分).**故选C.**

4. C 【解析】设北面共有x人,由题意得$\frac{x}{400+200+x}=\frac{10}{60}$,解得$x=120$,**故C正确.**

5. B 【解析】由题意可知样本中高中生的人数为$\frac{3}{5+4+3}n=\frac{1}{4}n$,小学生的人数为$\frac{5}{5+4+3}n=\frac{5}{12}n$,则$\frac{1}{4}n+20=\frac{5}{12}n$,解得$n=120$.**故B正确.**

6.【解】设自选快餐窗口a个,商务套餐窗口b个,现炒现做窗口c个,自动售货机d台.自选快餐高峰期就餐总人数为$400\times\frac{50}{100}=200$,每个窗口等待人数为$\frac{200}{a}$,最长等待时长为$1\cdot\frac{200}{a}=\frac{200}{a}$(分钟);商务套餐高峰期就餐总人数为$400\times\frac{30}{100}=120$,每个窗口等待人数为$\frac{120}{b}$,最长等待时长为$0.5\cdot\frac{120}{b}=\frac{60}{b}$(分钟);现炒现做

高峰期就餐总人数为 $400\times\frac{15}{100}=60$，每个窗口等待人数为 $\frac{60}{c}$，最长等待时长为 $5\cdot\frac{60}{c}=\frac{300}{c}$（分钟）；自动售货机高峰期就餐总人数为 $400\times\frac{5}{100}=20$，每台自动售货机等待人数为 $\frac{20}{d}$，最长等待时长为 $1\cdot\frac{20}{d}=\frac{20}{d}$（分钟）.

依题意，从等待时长和公平的角度上考虑，要求每个队伍的最长等待时长大致相等，则可得 $\frac{200}{a}=\frac{60}{b}=\frac{300}{c}=\frac{20}{d}$，即有 $a:b:c:d=10:3:15:1$，则 $18\times\frac{10}{29}\approx 6$（个），$18\times\frac{3}{29}\approx 2$（个），$18\times\frac{15}{29}\approx 9$（个），$18\times\frac{1}{29}\approx 1$（台），故应设置自选快餐、商务套餐、现炒现做的取餐窗口分别为 6 个、2 个、9 个，自动售货机 1 台.

9.2　用样本估计总体

9.2.1　总体取值规律的估计

对点上分

1. C　【解析】依据定义，每个小组的频数与样本容量之比是这个小组的频率，**故 C 正确.**

2. D　【解析】样本数据共有 20 个. 根据选项，可将样本数据分为 4 组，各组的频数和频率如表所示：

分组	频数	频率
[5.5,7.5)	2	0.1
[7.5,9.5)	6	0.3
[9.5,11.5)	8	0.4
[11.5,13.5]	4	0.2
合计	20	1

从表中可以看出频率为 0.2 的是[11.5,13.5]，**故 D 正确.**

3. D　【解析】对于 A，参加 3 场公益讲座的学生约为 $1\ 000\times26\%=260$（人），**故 A 错**

误;对于 B,参加 2 场或 4 场公益讲座的学生约为 1 000 × (18% + 20%) = 380(人),**故 B 错误**;对于 C,参加不高于 2 场公益讲座的学生约为 1 000×(8%+10%+20%)= 380(人),**故 C 错误**;对于 D,参加不低于 4 场公益讲座的学生约为 1 000 × (18% + 12% + 4% + 2%) = 360(人),**故 D 正确**.

4. C 【解析】对于 A,显然(2)班学生的数学成绩的优秀率最高,**故 A 正确**;对于 B,只根据优秀率的大小,无法比较每班学生的数学成绩优秀人数的多少,**故 B 正确**;对于 C,该年级全体学生数学成绩的优秀率为全年级数学成绩优秀的学生人数与全年级学生总人数之比,由于各班的学生人数未知,所以不能计算该年级全体学生数学成绩的优秀率,**故 C 错误**;对于 D,设(1)班、(2)班数学成绩优秀的人数分别为 $x,y,x>0,y>0$,(1)班、(2)班人数分别为 $a,b,a>0,b>0$,则$\frac{x}{a}=80\%$,$\frac{y}{b}=85\%$,得 $x=80\%a$,$y=85\%b$,又(1)班和(2)班的数学成绩放在一起统计的优秀率为 83%,所以$\frac{x+y}{a+b}=83\%$,即$\frac{80\%a+85\%b}{a+b}=83\%$,可得 $2b=3a$,则 $a<b$,**故 D 正确**.

易错警示 忽略表格所给数据意义而致错

表格中所给的是优秀率,即数学成绩在 120 分以上的学生人数与该班学生总人数之比,而题中未给出人数等相关数据,故无法进行全年级数学成绩优秀率的计算,在做关于频率分布表的问题时,切记要看清表格所给数据的意义.

5. B 【解析】由题图可得 $10(m+2m+0.015+0.020\times 2+0.030)=1$,解得 $m=0.005$,故物理成绩大于等于 60 分的人数为 $300\times[1-10\times(0.005+0.015)]=$

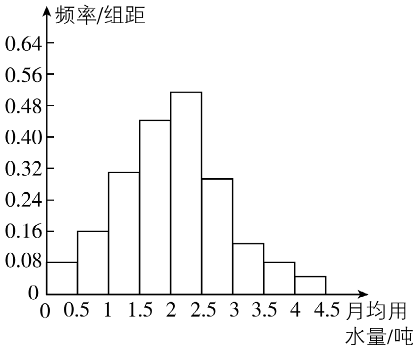
240. **故选 B.**

6. C 【解析】由题得 $10\times(0.005+0.035+a+0.020+0.010)=1$，所以 $a=0.030$. 成绩在[120,130)内的学生有 $100\times10\times0.030=30$(人)，所以应从[120,130)内抽取$\frac{20}{100}\times30=6$(人)，即 $b=6$. **故 C 正确.**

7. B 【解析】由题图可知，后两个小组的频率为$(0.013+0.037)\times5=0.25$，从左到右的前三个小组的频率之比为 $1:2:3$，所以第一小组的频率为$(1-0.25)\times\frac{1}{1+2+3}=0.125$. 又第一小组的频数为 4，所以报考飞行员的学生人数是 $4\div0.125=32$. **故 B 正确.**

8.【解】(1)在区间[0.5,1)内的频率为 0.08，则频数为 8，在区间[4,4.5]内的频率为 0.02，则频数为 2，则 $x=100-(4+8+15+22+14+6+4+2)=25$，$y=\frac{6}{100}=0.06$.

(2)绘制频率分布直方图如下，



因为从左往右数第 4 个矩形对应的频率为 0.22，且表中数据的组距为 0.5，所以所求高度为$0.22\div0.5=0.44$.

9. C 【解析】对于 A，2025 年 1~2 月，商品零售总额同比增长 2.9%，**故 A 错误**；对于 B，2024 年 8 月，餐饮收入总额同比增加，**故 B 错误**；对于 C，2024 年 6~10 月，商品零售总额同比都增加，**故 C 正确**；对于 D，2024 年 12 月，餐饮收入总额环比增速并未告知，**故 D 错误.**

10. C 【解析】对于 A，由题图可知，2017 年至 2024 年，知识付费用户数量逐年增加，**故 A 正确.** 对于 B，C，知识付费用户数量的逐年增加量分别为 2018 年：

0. 96－0. 48＝0. 48（亿人次）. 2019 年：1. 88－0. 96＝0. 92（亿人次）. 2020 年：2. 95－1. 88＝1. 07（亿人次）. 2021 年：3. 56－2. 95＝0. 61（亿人次）. 2022 年：4. 15－3. 56＝0. 59（亿人次）. 2023 年：4. 77－4. 15＝0. 62（亿人次）. 2024 年：5. 27－4. 77＝0. 5（亿人次）. 可知知识付费用户数量的逐年增加量 2020 年最多，不是逐年递增的，**故 B 正确，C 错误**. 对于 D，由于 5. 27>10×0. 48，则 2024 年知识付费用户数量超过 2017 年知识付费用户数量的 10 倍，**故 D 正确.**

11. ACD 【解析】由参保险种比例图可知，丁险种参保人数比例为 1－0. 02－0. 04－0. 1－0. 3＝0. 54，**故 A 正确**；由参保人数比例图可知，41 周岁以上参保人数占总参保人数的 45%，不到五成，**故 B 错误**；由不同年龄段人均参保费用图和参保人数比例图可知，18～29 周岁人群人均参保费用最少，在（3 000，4 000）之间，且这类人所占比例为 15%，54 周岁及以上参保人数最少，比例为 10%，但是 54 周岁及以上人群人均参保费用为 6 000 元，所以 18～29 周岁人群参保的总费用最少，**故 C 正确**；由不同年龄段人均参保费用图和参保人数比例图可知，总体的人均参保费用不超过 5 000 元，**故 D 正确.**

12. D 【解析】由扇形图可知，成绩在前 200 名的学生中，高一人数比高二人数多 200×（45%－30%）＝30，**故 A 正确**；成绩在第 1～50 名的学生中，高一人数为 200×45%×0. 2＝18，因此高三最多有 32 人，**故 B 正确**；由条形图知高一学生的成绩在第 101～150 名的人数为 200×45%×0. 4＝36，而高三学生的成绩在第 1～50 名的人数最多为 32，故高一学生的成绩在第 101～150 名的人数一定比高三学生的成绩在第 1～50 名的人数多，**故 C 正确**；成绩在第 51～100 名的学生中，高一人数为 200×45%×0. 3＝27，高二成绩在第 51～100 名的人数最多为 23，即成绩在第

51~100 名的学生中，高一的人数一定比高二的人数多，**故 D 错误**.

能力上分

1. D　【解析】生鲜区的净利润占比为 65.8%，65.8%>50%，**故 A 正确**.

生鲜区的营业利润率为 $\frac{65.8\%}{48.6\%}\times$ 32.5%≈44%>40%，**故 C 正确**.

熟食区的营业利润率为 $\frac{-4.3\%}{15.8\%}\times$ 32.5%<0；

乳制品区的营业利润率为 $\frac{16.5\%}{20.1\%}\times$ 32.5%≈26.68%；

其他区的营业利润率为 $\frac{1.8\%}{4.7\%}\times32.5\%\approx$ 12.45%；

日用品区的营业利润率为 $\frac{20.2\%}{10.8\%}\times$ 32.5%≈60.79%，可以看出，本季度日用品区的营业利润率最高，**故 B 正确**. 由题中数据知，其他区营业收入占比最低，为 4.7%，**故 D 错误**.

故选 D.

2. B　【解析】对于①，速度在 80 千米/时以下时，相同条件下每消耗 1 升汽油，丙车行驶的里程比乙车多，所以该市用丙车比用乙车更省油，**故①正确**；对于②，从题图中可以看出乙车的最高燃油效率大于 5 千米/升，**故②错误**；对于③，同样速度甲车消耗 1 升汽油行驶的里程比乙车、丙车的多，所以以相同速度行驶相同里程，甲车油耗最少，**故③正确**；对于④，甲车以 80 千米/时的速度行驶，1 升汽油行驶 10 千米，所以行驶 1 小时，即行驶 80 千米，消耗 8 升汽油，**故④错误**.

3. BCD　【解析】对于 A，由 $(0.005+0.025+a+0.025+0.010)\times10=1$，得 $a=0.035$，**故 A 错误**；对于 B，1 200 名参赛的学生中不及格的人数约为 $1\,200\times0.005\times10=60$，**故 B 正确**；对于 C，选取 100 人的成绩中，成绩落在[80，90)的人

数为 $100\times0.025\times10=25$,成绩落在[70,80)的人数为 $100\times0.035\times10=35$,则成绩落在[80,90)的人数是落在[70,80)的人数的$\frac{5}{7}$,**故 C 正确**;对于 D,以频率估计概率,从 1 200 名参赛学生中随机抽取 1 人,则该学生成绩不低于 90 分的概率为 $0.010\times10=0.1$,**故 D 正确**.

4. AD 【解析】由题图可知,在调查中 C——乘坐公交的人数为 30,所占比例为 25%,所以调查的总人数为 $30\div25\%=120$. 对于 A,$1\ 300\times\frac{42+30}{120}=780$(人),**故 A 正确**;对于 B,$\frac{120-42-30-18}{120}=\frac{1}{4}$,**故 B 错误**;对于 C,$\frac{42}{120}=0.35=35\%$,**故 C 错误**;对于 D,$\frac{42+18}{120}=0.5=50\%$,**故 D 正确**.

5. ①③④ 【解析】当 $m>127$ 时,数学成绩为"A 等级"的 8 人的学号为 1,2,3,4,5,6,7,10 号;

当 $m=127$ 时,数学成绩为"A 等级"的人数不为 8 人,不合题意;

当 $m<127$ 时,数学成绩为"A 等级"的 8 人的学号为 1,2,3,4,5,6,7,8 号.

当 $n>98$ 时,语文成绩为"A 等级"的 7 人的学号为 1,2,3,4,5,6,10 号;

当 $n=98$ 时,语文成绩为"A 等级"的人数不为 7 人,不合题意;

当 $n<98$ 时,语文成绩为"A 等级"的 7 人的学号为 1,2,3,4,5,6,9 号.

故当 $m>127$,$n>98$ 时,数学与语文两科成绩全是"A 等级"的有 1,2,3,4,5,6,10 号,共 7 人,不合题意;

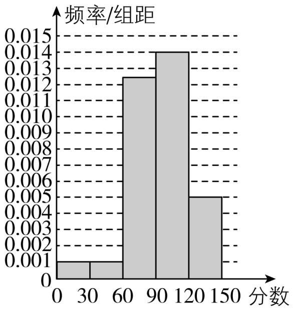
当 $m>127$,$n<98$ 时,数学与语文两科成绩全是"A 等级"的有 1,2,3,4,5,6 号,共 6 人,符合题意;

当 $m<127$,$n>98$ 时,数学与语文两科成绩全是"A 等级"的有 1,2,3,4,5,6 号,共 6 人,符合题意;

当 $m<127$，$n<98$ 时，数学与语文两科成绩全是"A 等级"的有 1，2，3，4，5，6 号，共 6 人，符合题意.

综上可知，当 $m>127$ 时，$n<98$，①**正确**；当 $m<127$ 时，$n\neq 98$，②**错误**；当 $m>127$，$n<98$ 或 $m<127$，$n>98$ 或 $m<127$，$n<98$ 时，两科均不是"A 等级"的学生依次为 8，9，10 号，均恰有 1 名，③**正确**；1～6 号的学生两科成绩全是"A 等级"，④**正确**.

6.【解】（1）由频率分布表得 $M=\dfrac{3}{0.03}=100$，所以 $m=100-(3+3+37+15)=42$，$n=\dfrac{42}{100}=0.42$，$N=0.03+0.03+0.37+0.42+0.15=1$. 频率分布直方图如图所示.



（2）由题意，知全校成绩在 90 分以上（包含 90 分）的学生人数约为 $\dfrac{42+15}{100}\times 600=342$.

7.【解】（1）由题图可得 $(0.002+0.009\ 5+0.011+0.012\ 5+x+0.005+0.002\ 5)\times 20=1$，解得 $x=0.007\ 5$，所以频率分布直方图中 x 的值是 0.007 5.

（2）月平均用电量为 $[220,240)$ 的居民有 $0.012\ 5\times 20\times 100=25$（户），月平均用电量为 $[240,260)$ 的居民有 $0.007\ 5\times 20\times 100=15$（户），月平均用电量为 $[260,280)$ 的居民有 $0.005\times 20\times 100=10$（户），月平均用电量为 $[280,300]$ 的居民有 $0.002\ 5\times 20\times 100=5$（户），所以月平均用电量不低于 220 千瓦时的有 $25+15+10+5=55$（户）.

（3）由（2）可知，抽取比例为 $\dfrac{11}{55}=\dfrac{1}{5}$，所

以月平均用电量在$[220,240)$的分组中应抽取$25\times\frac{1}{5}=5$(户).

9.2.2 总体百分位数的估计

对点上分

1. C 【解析】将数据按从小到大的顺序排列为3,4,5,6,7,8,9,10,共8个,则$8\times25\%=2$,则这组数据的25%分位数是$\frac{4+5}{2}=4.5$,**故C正确.**

2. C 【解析】由题可知数据的总个数为6,$6\times75\%=4.5$.

所以该组数据的上四分位数,即该组数据的75%分位数,即为这组数据从小到大排列后的第5个数据,则$m\leqslant8$,即实数m的取值范围是$(-\infty,8]$. **故选C.**

3. 6 【解析】由题易知共有13个数字,这13个数字从小到大排列为3,4,4,4,6,6,6,7,7,8,8,9,9. 由于$13\times0.4=5.2$,则这组数据的40%分位数是第6项,所以40%分位数为6.

4. 90 【解析】由题意得,成绩大于或等于75分的学生至少有$300-300\times70\%=90$(名).

5. AB 【解析】对于A,由A班的频率分布直方图可知,$[10,20)$与$[60,70]$的纵坐标相同,组距均为10,所以两组的频率相同,**故A正确;**

对于B,$10\times(0.005+0.020+m+0.020+0.015+0.015)=1$,解得$m=0.025$,前4个矩形面积之和为0.7,前5个矩形面积之和为0.85,故x_2位于$[50,60)$中,所以$0.7+(x_2-50)\times0.015=0.75$,解得$x_2=\frac{160}{3}$,**故B正确;**对于C,因为A班、B班产生塑料饮料瓶数在$[40,50)$之间的频率都是0.2,且该校有学生1 000人,则5月份产生塑料饮料瓶数在$[40,50)$之间的人数约为$1\,000\times0.2=200$,**故C错误;**对于D,由B知$m=0.025$,**故D错误.**

易错警示 不能正确理解百分位数在频率分布直方图中的含义而致错

根据频率分布直方图计算样本数据的百分位数的易错之处是不能正确理解百分位数的含义，由百分位数的含义可知一组数据的百分位数即为累积频率，因此首先要根据频率分布直方图中各组数据的频率估计百分位数所在的区间，再应用方程的思想方法设出百分位数进行求解.

6.【解】(1)根据按比例分配的分层随机抽样知，应抽取小吃类商家 $80\times(1-30\%-15\%-10\%-5\%-5\%)=28$(家)，生鲜类商家 $80\times15\%=12$(家)，所以应抽取小吃类商家 28 家，生鲜类商家 12 家.

(2)(i)根据题意可得 $(0.002\times3+2a+0.006)\times50=1$，解得 $a=0.004$，设 75% 分位数为 x，因为 $(0.002+0.004+0.006)\times50=0.6$，$0.6+0.004\times50=0.8$，所以 $(x-450)\times0.004+0.6=0.75$，解得 $x=487.5$，所以该直播平台商家平均日利润的 75% 分位数为 487.5 元.

(ii) $\left(\dfrac{500-480}{50}\times0.004+0.002+0.002\right)\times50\times1\,000=280$，所以估计该直播平台"优质商家"的个数为 280.

9.2.3 总体集中趋势的估计

对点上分

1. BCD 【解析】对于 A，这 5 位同学成绩的众数是 72，**故 A 正确**；对于 B，这 5 位同学成绩的平均数为 $\dfrac{84+72+76+80+72}{5}=76.8$，**故 B 错误**；对于 C，将这 5 位同学的成绩从小到大排列为 72，72，76，80，84，则这 5 位同学成绩的中位数为 76，**故 C 错误**；对于 D，$5\times75\%=3.75$，则将这 5 位同学的成绩按从小到大排列，第 4 个数为 75% 分位数，即 80，**故 D 错误**.

易错警示 错误理解表格数据含义而致错

读懂表格显示的是每位同学对应的成绩,计算相关数据时应予以注意.

2. D 【解析】众数为频率分布直方图中最高的矩形的底边中点横坐标值,因此题图①中的众数位于第二列矩形底边的中点处,数据第二、三列较多,且右侧拖尾,所以平均数大于中位数. 即在题图①中,众数<中位数<平均数.

同理在题图②中,平均数<中位数<众数.

故选 D.

3. BD 【解析】由题可得,原样本数据的平均数 $\bar{x}=\dfrac{a_1+a_2+a_3+\cdots+a_9}{9}$,新样本数据的平均数为$\dfrac{9\bar{x}+t}{10}$,因为 $t<\bar{x}$,所以$\dfrac{9\bar{x}+t}{10}<\bar{x}$,**故 A 错误;**

若原样本数据中存在两个众数,不妨令 $a_1=a_2=a_3$,$a_4=a_5=a_6$,设 $t=a_1$,则新样本数据有一个众数,**故 B 正确;**

原样本数据的极差为 a_9-a_1,当 $t<a_1$ 时,新样本数据的极差变大,当 $a_1\leqslant t<\bar{x}\leqslant a_9$ 时,极差不变,**故 C 错误;**

原样本数据的中位数为 a_5,当 $t<a_5$ 时,新样本数据的中位数为新样本数据由小到大排列的第 5 个数据和第 6 个数据的平均数,可能小于 a_5,如原样本数据为 1,2,3,4,5,6,7,8,9,加入 $t=3$,则新样本数据的中位数为$\dfrac{4+5}{2}=4.5<5$,**故 D 正确. 故选 BD.**

4. B 【解析】将实际数据从小到大排序为 4.0,5.9,7.7,9.6,9.6,14.9,其平均数为$\dfrac{4.0+5.9+7.7+9.6+9.6+14.9}{6}\approx 8.62$,中位数为$\dfrac{7.7+9.6}{2}=8.65$,极差为 $14.9-4.0=10.9$. 将错误数据从小到大排序为 4.0,5.9,7.7,9.6,14.9,96,其平均数为$\dfrac{4.0+5.9+7.7+9.6+14.9+96}{6}\approx 23.02$,中

位数为$\frac{7.7+9.6}{2}=8.65$,极差为$96-4.0=92$.根据众数的含义可知,实际数据中的众数是9.6,错误数据没有众数,所以没有发生变化的量是中位数,**故B正确**.

归纳总结 **平均数的有关性质**

(1)平均数与样本的每一个数据都有关,它可以反映出更多关于样本数据总体的信息,其中任何一个数据的改变都会引起平均数的变化.

(2)平均数受数据中极端值的影响较大,使平均数在估计总体时的可靠性降低.

5.【解】(1)在题目所给的80个数据中,2 000出现了22次,出现的次数最多,因此这组数据的众数是2 000.

把这80个数据按从小到大的顺序排列后,位于中间的数是2 000,2 500,因此这组数据的中位数是$\frac{2\,000+2\,500}{2}=2\,250$.

这组数据的平均数为$\bar{x}=\frac{18\,000+12\,000\times2+8\,000\times3+\cdots+1\,200\times6}{80}=\frac{249\,200}{80}=3\,115$.

(2)由于大多数员工的月奖金达不到平均数3 115,显然用平均数作为该公司员工月奖金的代表值并不合适;众数2 000及中位数2 250在一定程度上代表了大多数人的月奖金水平,较能反映月奖金水平的实际情况.

(3)由于全国各地的公司月奖金差异性较大,因而不能用一个公司的数据估计全国该类公司的月奖金水平.(言之有理即可)

6.【解】(1)由七年级学生成绩统计图可得,众数是8,所以$a=8$,

合格率为$1-20\%=80\%$,所以$b=80\%$.

由八年级学生成绩统计图可得,成绩为5分的有3人,成绩为6分的有2人,成绩为7分的有5人,此时总共有10人,八年级共抽取20人,所以中位数为$\frac{7+8}{2}=$

7.5，即 $c=7.5$.

(2)利用合格率估计合格的总人数，该校八年级有 600 名学生，合格人数约为 $85\%\times600=510$.

7. AB　【解析】依题意，$(0.015+0.025+0.035+0.005+2a)\times10=1$，解得 $a=0.010$. 对于 A，∵ 最高小矩形的中点横坐标为 75，∴ 众数是 75，**故 A 正确**；对于 B，平均数为 $45\times0.1+55\times0.15+65\times0.25+75\times0.35+85\times0.1+95\times0.05=68.5$，**故 B 正确**；对于 C，∵ $10\times(0.010+0.015+0.025)=0.5$，∴ 样本的中位数为 70，**故 C 错误**；对于 D，估计该校学生中得分超过 80 分的约占 $(0.010+0.005)\times10\times100\%=15\%$，**故 D 错误**.

易错警示　**不能正确根据频率分布直方图求数字特征**

解决有关频率分布直方图的问题时，一定要注意纵轴是$\dfrac{频率}{组距}$.

8. C　【解析】由众数的定义可知，100 户居民用户的月均用水量在区间$[4.2,7.2)$内的居民用户最多，再由样本估计总体可知，**A 正确**；

由图可知，随着月均用水量的增加，居民用户数量先增后减，整体呈现降低的趋势，**B 正确**；

平均数为$(2.7\times0.077+5.7\times0.107+8.7\times0.043+11.7\times0.030+14.7\times0.030+17.7\times0.017+20.7\times0.010+23.7\times0.013+26.7\times0.007)\times3=8.9604$，

因为 $0.077\times3=0.231$，$(0.077+0.107)\times3=0.552$，则中位数在第二组内，

设中位数为 x，则 $0.231+(x-4.2)\times0.107=0.5$，得 $x\approx6.7$，

故可以估计该市月均用水量的平均数大于中位数，**C 错误**；

因为$(0.077+0.107+0.043+0.030)\times3=0.771$，$(0.077+0.107+0.043+0.030+0.030)\times3=0.861$，

则 80% 分位数在第五组内，

设该市居民用户的月均用水量的 80% 分位数约为 y，则 $0.771+(y-13.2)\times 0.030=0.8$，

得 $y\approx 14.2$，**故 D 正确. 故选 C.**

能力上分

1. B 【解析】这组数据一共有五个数，中位数为 8，则数据从小到大排列，8 的前面有两个数，后面也有两个数，又因为唯一的众数为 9，所以有两个 9，其余的数均只出现一次，则最大数字为 9. 又因为极差为 3，所以最小的数为 6，所以这组数据为 6，7，8，9，9，所以平均数为 $\frac{6+7+8+9+9}{5}=7.8$，**故 B 正确.**

2. D 【解析】将除 x 外的成绩按从小到大的顺序排列为 91，92，92，95，100，由题知 x 的值必定在 92，93，94 三个数值之中，若 x 为 92，则众数是 92，中位数也是 92，符合题意；若 x 是 93，则中位数是 92.5，与众数 92 不相等，不符合题意；若 x 是 94，则中位数为 93，与众数 92 不相等，不符合题意. 故 x 为 92，所以这 6 名同学的数学成绩的平均数是 $\frac{91+92+92+92+95+100}{6}\approx 93.7$，**故 D 正确.**

3. AB 【解析】这 100 名学生中“综合体能测试”成绩高于 80 的学生人数为 $18+30+24+10=82$，所占比例为 $82\div 100=0.82>0.8$，**所以 A 正确；**

成绩不超过 85 的学生人数为 $6+12+18=36<50$，**所以 B 正确；**

成绩的众数不一定为 85，**所以 C 不正确；**

若同一组数据都用右端点值来估计，则这 100 名学生的“综合体能测试”成绩平均值最大为 $\frac{1}{100}\times(75\times 6+80\times 12+85\times 18+90\times 30+95\times 24+100\times 10)=89.2<90$，**所以 D 不正确. 故选 AB.**

4. $\{-9,5,19\}$ 【解析】设丢失的数据为 x，则这七个数的平均数为 $\frac{x+37}{7}$，众数为 4，

因为这组数据的平均数与众数的和是中位数的2倍，将除x以外的六个数据由小到大的顺序排列后，分以下几种情况讨论：

若$x\leqslant4$，则中位数为4，此时，$\frac{x+37}{7}+4=2\times4$，解得$x=-9$；

若$4<x<6$，则中位数为x，此时，$\frac{x+37}{7}+4=2x$，解得$x=5$；

若$x\geqslant6$，则中位数为6，此时，$\frac{x+37}{7}+4=2\times6$，解得$x=19$.

综上可知，丢失的数据的所有可能取值构成的集合为$\{-9,5,19\}$.

5.【解】(1)依题意得平均评分$\bar{x}=1\times0.08+3\times0.22+5\times0.36+7\times0.3+9\times0.04=5$.

(2)由(1)可得评分位于区间$[0,5)$，$[5,7)$，$[7,9)$，$[9,10]$内的大黄梨分别定为普通果、良果、优果、特优果，则评分位于区间$[0,5)$内的频率为$0.08+0.22+0.18=0.48$；评分位于区间$[5,7)$内的频率为$0.18+0.15=0.33$；评分位于区间$[7,9)$内的频率为$0.02+0.15=0.17$；评分位于区间$[9,10]$内的频率为0.02.

又果农现有5 000千克大黄梨，所以估计普通果的质量为$0.48\times5\ 000=2\ 400$(千克)；良果的质量为$0.33\times5\ 000=1\ 650$(千克)；优果的质量为$0.17\times5\ 000=850$(千克)；特优果的质量为$0.02\times5\ 000=100$(千克).

(3)采用方案2较好，理由如下：

若按方案1：不分等级卖出，单价为8元/千克. 若按方案2：分等级卖出，则单价的平均值为$7\times0.48+8\times0.33+12\times0.17+15\times0.02=8.34$(元/千克). 因为$8.34>8$，所以从果农的角度考虑，采用方案2较好.

6.【解】(1)根据平均数的计算公式可知，甲厂数据的平均数是$\frac{8+9+9+9+9+11+13+16+17+19}{10}=12$，

乙厂数据的平均数是 $\frac{10+10+12+12+12+13+14+16+18+19}{10}=13.6$，

丙厂数据的平均数是 $\frac{8+8+8+10+11+13+17+19+20+20}{10}=13.4$.

甲厂、乙厂、丙厂数据的众数分别是 9，12，8.

甲厂数据的中位数为 $\frac{9+11}{2}=10$，乙厂数据的中位数为 $\frac{12+13}{2}=12.5$，丙厂数据的中位数为 $\frac{11+13}{2}=12$.

(2)甲厂用平均数作为该电子产品推销广告中的待机时间，乙厂用众数作为该电子产品推销广告中的待机时间，丙厂用中位数作为该电子产品推销广告中的待机时间.

(3)我会选乙厂的产品. 因为乙厂产品的平均数最大，众数最大，中位数最大，所以待机时间更长些，稳定性也较好.

9.2.4 总体离散程度的估计

对点上分

1. C 【解析】由题意，数据的中位数为 16，可得 $\frac{m+n}{2}=16$，所以 $m+n=32$，所以这 6 周的慢走里程的平均数为 $\frac{11+12+m+n+20+27}{6}=17$. 要使这 6 周的慢走里程的标准差最小，需要 $(m-17)^2+(n-17)^2$ 最小，又由 $(m-17)^2+(n-17)^2=(m-17)^2+(32-m-17)^2=2m^2-64m+514$，故当标准差最小时，$m=-\frac{-64}{2\times2}=16$，**故 C 正确.**

2. ACD 【解析】由题图可知样本 A 的最高点与最低点的高度差小于样本 B 的最高点与最低点的高度差，所以样本 A 的极差小于样本 B 的极差，**A 正确.**

因为中位数是按大小顺序排列后，位于中间位置的数值. 由题可知样本数据为偶数个，所以中位数为中间两个数值的

平均数,由题图可知样本 A 的样本数据的中间两位数的平均数小于样本 B 的样本数据的中间两位数的平均数,B **不正确**.

由题图可知样本 A 的每个样本数据均小于样本 B 的对应数据,所以样本 A 的平均数小于样本 B 的平均数,C **正确**.

样本 A 的离散程度小于样本 B 的离散程度,所以样本 A 的方差小于样本 B 的方差,D **正确**. **故选** ACD.

3. BD 【**解析**】因为样本数据 x_1,x_2,x_3 的平均数为 2,方差为 1,所以数据 $3x_1-1$, $3x_2-1,3x_3-1$ 的平均数为 $3\times2-1=5$,**故 A 错误**;数据 $3x_1-1,3x_2-1,3x_3-1$ 的方差为 $3^2\times1=9$,**故 B 正确**;$x_1+x_2+x_3=3\times2=6$,$(x_1-2)^2+(x_2-2)^2+(x_3-2)^2=1\times3=3$,数据 x_1,x_2,x_3,2 的平均数为 $\frac{x_1+x_2+x_3+2}{4}=2$,所以方差为 $\frac{1}{4}[(x_1-2)^2+(x_2-2)^2+(x_3-2)^2+(2-2)^2]=\frac{3}{4}$,**故 C 错误**;由 $x_1+x_2+x_3=6$,$(x_1-2)^2+(x_2-2)^2+(x_3-2)^2=3$,得 $x_1^2+x_2^2+x_3^2-4(x_1+x_2+x_3)+12=3$,所以 $x_1^2+x_2^2+x_3^2=15$,所以数据 x_1^2,x_2^2,x_3^2 的平均数为 $\frac{x_1^2+x_2^2+x_3^2}{3}=5$,**故 D 正确**.

易错警示 平均数与方差性质混淆

(1)如果数据 $x_1,x_2,\cdots,x_n$ 的平均数为 $\bar{x}$,方差为 s^2,那么一组新数据 $x_1+b,x_2+b,\cdots,x_n+b$ 的平均数为 $\bar{x}+b$,方差为 s^2.

(2)如果一组数据 $x_1,x_2,\cdots,x_n$ 的平均数为 $\bar{x}$,方差为 s^2,那么一组新数据 $ax_1,ax_2,\cdots,ax_n$ 的平均数为 $a\bar{x}$,方差为 a^2s^2.

(3)如果一组数据 $x_1,x_2,\cdots,x_n$ 的平均数为 $\bar{x}$,方差为 s^2,那么一组新数据 $ax_1+b,ax_2+b,\cdots,ax_n+b$ 的平均数为 $a\bar{x}+b$,方差为 a^2s^2.

4. A 【解析】对于甲，由甲的统计数据可知，甲至少有 3 天的合格品件数不低于 24，最低合格品件数不低于 22，所以甲一定能通过. 对于乙，设乙每天的合格品件数为 $a_i(i=1,2,3,4,5)$，$a_i\in\mathbf{Z}$，则 $\frac{\sum\limits_{i=1}^{5}(a_i-23)^2}{5}\leqslant 1$，即 $\sum\limits_{i=1}^{5}(a_i-23)^2\leqslant 5$. 若乙有不止 1 天的合格品件数低于 22，则 $\sum\limits_{i=1}^{5}(a_i-23)^2>5$，不符合题意；若乙只有 1 天的合格品件数低于 22，不妨取 $a_1=21$，$(a_1-23)^2=4$，因为平均数为 23，则至少有 1 天的合格品件数为 25 或至少有 2 天的合格品件数为 24，无论哪种情况，都可以得到 $\sum\limits_{i=1}^{5}(a_i-23)^2>5$，不符合题意，所以乙的这 5 天的合格品件数都不低于 22，故乙一定能通过. 对于丙，若丙的合格品件数为 21，22，23，23，23，则丙的众数为 23，方差为 0. 64，符合丙的统计数据，但此时丙不能通过. 所以甲、乙一定能通过，**故 A 正确.**

5. A 【解析】抽取的 100 人中，男生有 60 人，女生有 40 人，

这 100 位学生的平均身高为 $\frac{60\times160+40\times155}{100}=158$，

则这 100 位学生身高的方差为 $\frac{3}{5}\times[4+(160-158)^2]+\frac{2}{5}\times[3+(155-158)^2]=9.6$，

所以估计该校学生身高的总体方差是 9. 6. **故选 A.**

6. 7　2 【解析】因为该市学生的男、女生人数比为 2∶3，所以设男、女生人数分别为 $2a,3a(a>0)$，所以该市学生每天睡眠时长的平均数为 $\frac{7.3\cdot 2a+6.8\cdot 3a}{2a+3a}=7$（小时）. 由题中方差公式可得该市学生每天睡眠时长的方差 $b^2=\frac{1}{2a+3a}\cdot\left[(2a\cdot 2+3a\cdot 1.9)+\frac{2a\cdot 3a}{2a+3a}\cdot\right.$

$(7.3-6.8)^2\Big]=2.$

能力上分

1. A 【解析】删去一个数之后，平均值没有改变，所以删除的数为 5.

由题意可得 $s^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-5)^2=32$，即 $\sum_{i=1}^{n}(x_i-5)^2=32n$，删除一个数后的方差为 $\frac{1}{n-1}\left[\sum_{i=1}^{n}(x_i-5)^2-(5-5)^2\right]=40$，即 $\frac{32n}{n-1}=40$，解得 $n=5$. **故选** A.

2. ABD 【解析】由题中的频率分布直方图可知，$(0.005+0.010+0.020+a+0.025+0.010)\times10=1$，解得 $a=0.03$.

由题图可以看出众数在 $[70,80)$ 区间内，所以众数为 $\frac{70+80}{2}=75$，**所以 A 正确**；

上四分位数即为第 75 百分位数，因为前四组数据的频率和为 $(0.005+0.010+0.020+0.03)\times10=0.65$，而前五组数据的频率和为 $(0.005+0.010+0.020+0.03+0.025)\times10=0.90$，所以第 75 百分位数位于 $[80,90)$ 区间内，设第 75 百分位数为 x，则 $x=80+\frac{0.75-0.65}{0.025}=84$，**所以 B 正确**；

成绩的极差通过频率分布直方图只能估计，由题图可知，成绩介于 $[40,100]$ 之间，极差小于或等于 60，**所以 C 错误**；

由频率分布直方图可求得在区间 $[50,60)$ 的样本数为 $0.010\times10\times100=10$，在 $[60,70)$ 的样本数为 $0.020\times10\times100=20$. 所以两组成绩的平均数 $\bar{x}=\frac{54\times10+66\times20}{10+20}=62$. 两组成绩的方差 $s^2=\frac{10\times[2+(54-62)^2]+20\times[5+(66-62)^2]}{10+20}=36$，所以两组成绩的总标准差为 6. **所以 D 正确. 故选** ABD.

3. C 【解析】若随机删去其中任意一轮的成绩，恰好不是五轮成绩中的最高成绩和最低成绩，此时新数据的极差等于原数据的极差，A **错误**；

不妨设 $x_1<x_2<x_3<x_4<x_5$，当 $\frac{1}{2}(x_2+x_4)=x_3$ 时，若随机删去的成绩是 x_3，此时新数据的中位数等于原数据的中位数，B **错误**；

若 $\bar{x}=\bar{y}$，即删去的数据恰为平均数，根据方差的计算公式，分子不变，分母变小，此时方差会变大，C **正确**；

不妨设 $x_1<x_2<x_3<x_4<x_5$，在按从小到大的顺序排列的五个数据中，可得 $5\times60\%=3$，此时原数据的 60% 分位数为第三个数和第四个数的平均数，即 $\frac{x_3+x_4}{2}$，删去一个数据后的四个数据，按从小到大的顺序排列，可得 $4\times60\%=2.4$，此时新数据的 60% 分位数为第三个数，即 x_3 或 x_4，则 $x_3<\frac{x_3+x_4}{2}<x_4$，显然新数据的 60% 分位数不等于原数据的 60% 分位数，D **错误**. **故选** C.

4.【解】(1)由题知，高一年级学生竞赛成绩的平均数 $\bar{x}_1=\frac{1}{50}\times(65\times15+75\times5+85\times15+95\times15)=81$，

方差 $s_1^2=\frac{1}{50}\times[(65-81)^2\times15+(75-81)^2\times5+(85-81)^2\times15+(95-81)^2\times15]=144$，

高二年级学生竞赛成绩的平均数 $\bar{x}_2=\frac{1}{50}\times(65\times10+75\times10+85\times20+95\times10)=81$，

方差 $s_2^2=\frac{1}{50}\times[(65-81)^2\times10+(75-81)^2\times10+(85-81)^2\times20+(95-81)^2\times10]=104$.

$\because \bar{x}_1=\bar{x}_2, s_1^2>s_2^2$，$\therefore$ 样本中平均成绩一样，但高二学生的成绩更稳定，用样本估计总体可以认为高二学生这次竞赛成绩更好.

(2)设选择方案一时一位高一学生获得的食堂代金券金额为 X 元，则 X 的可能取值为 10，25，35，对应的频率分别为 0.3，0.4，0.3，$\therefore$ 其获得食堂代金券金额的平均数 $\bar{X}=10\times0.3+25\times0.4+35\times0.3=23.5$(元)；

设选择方案二时一位高一学生获得的食

堂代金券金额为 Y 元，则其获得食堂代金券金额的平均数 $\overline{Y}=10\times0.5+30\times0.5=20$(元). $\because\ \overline{X}>\overline{Y}$, $\therefore$ 从统计的角度看，高一年级组长应该选择方案一.

9.2 节测上分

1. C 【解析】因为数据在$[0.2,0.8)$内的频率为 0.75，所以数据在$[0.2,0.8)$内的频数为 $600\times0.75=450$，故样本中数据在$[0.4,0.8)$内的个数为 $450-95-120=235$. **故选 C.**

2. B 【解析】根据题中数据 3,3,4,5,7,8,9,9 得，极差为 6，众数为 3,9，**所以 A 正确，B 错误.**

数据的平均数 $\bar{x}=\dfrac{3+3+4+5+7+8+9+9}{8}=6$，**所以 C 正确.**

数据的方差 $s^2=\dfrac{1}{8}\times[(3-6)^2+(3-6)^2+(4-6)^2+(5-6)^2+(7-6)^2+(8-6)^2+(9-6)^2+(9-6)^2]=\dfrac{23}{4}$，**所以 D 正确. 故选 B.**

3. ABC 【解析】由题图可知甲的得分从小到大排列为 81,81,82,83,84,87，乙的得分从小到大排列为 78,79,80,81,82,86.

甲得分的中位数为$\dfrac{82+83}{2}=82.5$，乙得分的中位数为$\dfrac{80+81}{2}=80.5$，所以甲得分的中位数大于乙得分的中位数，**A 正确；**

易知甲得分的极差为 6，乙得分的极差为 8，所以甲得分的极差小于乙得分的极差，**B 正确；**

计算可得甲得分的平均数为$\dfrac{1}{6}\times(81+81+82+83+84+87)=83$，乙得分的平均数为$\dfrac{1}{6}\times(78+79+80+81+82+86)=81$，所以甲得分的平均数大于乙得分的平均数，**C 正确；**

甲得分的方差为$\dfrac{1}{6}\times(4+4+1+0+1+16)=\dfrac{13}{3}$，乙得分的方差为$\dfrac{1}{6}\times(9+4+1+0+1+25)=\dfrac{20}{3}$，所以甲得分的方差小于

乙得分的方差,**D 错误. 故选 ABC.**

4. ABD 【解析】对于 A,由频率分布直方图中矩形面积之和为 1 得 $50\times(0.0024+0.0036+0.0060+x+0.0024+0.0012)=1$,解得 $x=0.0044$,**故 A 正确**;对于 B,月用电量不超过 200 kW·h 的频率为 $50\times(0.0024+0.0036+0.0060)=0.6$,所以户数为 $1500\times0.6=900$,**故 B 正确**;对于 C,平均数为 $(75\times0.0024+125\times0.0036+175\times0.0060+225\times0.0044+275\times0.0024+325\times0.0012)\times50=186$,设中位数为 y,$(0.0024+0.0036)\times50=0.3$,$0.3+0.0060\times50=0.6$,则 y 在第三组$(150,200]$中,即 $0.3+(y-150)\times0.0060=0.5$,解得 $y\approx183.33$,故平均数大于中位数,**故 C 错误**;对于 D,设 45% 分位数为 z,则 z 在第三组$(150,200]$中,$0.3+(z-150)\times0.0060=0.45$,解得 $z=175$,**故 D 正确.**

5. ABC 【解析】对于 A,从题图中可看出,电动汽车销量逐年递增,**故 A 正确**;

对于 B,将汽车销量数据从小到大排序,因为 $0.6\times6=3.6$,所以销量数据的 60% 分位数为第 4 个数,即 536.5,**故 B 正确**;

对于 C,2018—2019 年的增长率为 $\frac{111.5-97.2}{97.2}\times100\%\approx14.7\%$,2019—2020 年的增长率为 $\frac{291.6-111.5}{111.5}\times100\%\approx161.5\%$,2020—2021 年的增长率为 $\frac{536.5-291.6}{291.6}\times100\%\approx84.0\%$,2021—2022 年的增长率为 $\frac{668.5-536.5}{536.5}\times100\%\approx24.6\%$,2022—2023 年的增长率为 $\frac{756.8-668.5}{668.5}\times100\%\approx13.2\%$,2019—2020 年的增长率超过其他年份的增长率,**故 C 正确**;

对于 D,这六年销量的平均值为 $(97.2+111.5+291.6+536.5+668.5+756.8)\div6=410.35>291.6$,**故 D 错误.**

6. C 【解析】对于 A,假设甲组存在选手失分超过 7 分,最高失 8 分,根据极差为 5,

得最低失分为 3 分，此时中位数为 3，故假设可以成立，**故 A 错误**；对于 B，假设乙组的失分情况为 0，0，1，1，2，2，2，2，2，8，满足平均数为 2，众数为 2，但该组不为“优秀小组”，**故 B 错误**；对于 C，假设丙组的失分情况从小到大排列依次为 $x_1,x_2,\cdots,x_{10}$，丙组平均数为 2，方差为 3，即 $(x_1-2)^2+(x_2-2)^2+\cdots+(x_{10}-2)^2=30$，若 $x_{10}=8$，则 $(x_{10}-2)^2=36>30$，不符合要求，故 $x_{10}\leqslant 7$，所以该组每位选手失分都不超过 7 分，则该组为“优秀小组”，**故 C 正确**；对于 D，$85\%\times 10=8.5$，故从小到大排序后，第 9 个数即为 85% 分位数，即从小到大排序后的第 9 个数为 7，假设丁组失分情况为 0，0，0，0，0，0，0，5，7，8，满足平均数为 2，85% 分位数为 7，但不是“优秀小组”，**故 D 错误**.

7. AC　【解析】样本数据 S_1 的平均数为 $\frac{x_1+x_2+x_3}{3}$，

样本数据 S_2 的平均数为 $\frac{x_1+x_2+x_3}{3}$，

所以 S_1 的平均数等于 S_2 的平均数，**所以 A 正确**；

由 $x_1<x_2<x_3$，可得数据 S_1 的中位数为 x_2，

数据 S_2 的中位数为 $\frac{x_1+x_3}{2}$，此时 x_2 不一定小于 $\frac{x_1+x_3}{2}$，**所以 B 不正确**；

样本数据 S_1 的极差为 x_3-x_1，

样本数据 S_2 的极差为 $\frac{x_2+x_3}{2}-\frac{x_1+x_2}{2}=\frac{x_3-x_1}{2}$，

因为 $x_1<x_2<x_3$，所以 $x_3-x_1>0$，可得 $x_3-x_1>\frac{x_3-x_1}{2}$，**所以 C 正确**；

由于 $x_1<\frac{x_1+x_2}{2}<\frac{x_1+x_3}{2}<\frac{x_2+x_3}{2}<x_3$，且数据 S_1 和 S_2 的平均数相同，所以 $s_1^2>s_2^2$，**所以 D 不正确. 故选 AC.**

8. D　【解析】原数据的平均数 $\bar{x}=\frac{1}{5}\times(1+2+3+4+5)=3$，原数据的方差 $s_x^2=\frac{1}{5}\times[(1-3)^2+(2-3)^2+(3-3)^2+$

$(4-3)^2+(5-3)^2]=2.$

若 $m=n=3$,则满足 $m+n=6$,此时所得新数据的平均数 $\bar{y}=\frac{1}{7}\times(3\times5+6)=3$,方差 $s_y^2=\frac{1}{7}\times[(1-3)^2+(2-3)^2+3\times(3-3)^2+(4-3)^2+(5-3)^2]=\frac{10}{7}$,方差变小,**故 A 错误;**

若极差不变,则 m,n 的取值可能是 $\begin{cases}m=1,\\n=1,\end{cases}\begin{cases}m=2,\\n=1,\end{cases}\cdots$,不一定满足 $m+n=6$,**故 B 错误;**

若 $m+n=6$,则当 $m=3,n=3$ 时,新数据 1,2,3,3,3,4,5 的中位数是 3,因为原数据 1,2,3,4,5 的中位数也是 3,所以中位数不变,**故 C 错误;**

新数据的平均数 $\bar{y}=\frac{1}{7}\times(3\times5+m+n)$,由 $\frac{1}{7}\times(3\times5+m+n)=3$,得 $m+n=6$,**故 D 正确. 故选 D.**

9. 21.6　15.5　【解析】根据题意,样本中有 90 名男员工、50 名女员工,该公司全体人员的 BMI 值的平均数 $\bar{x}=\frac{90\times22.1+50\times20.7}{140}=21.6$;方差 $s^2=\frac{9}{14}\times[14.3+(22.1-21.6)^2]+\frac{5}{14}[16.4+(20.7-21.6)^2]=15.5.$

10.【解】(1) A 公司员工月均工资的平均数为 $0.3\times0.21+0.5\times0.29+0.7\times0.27+0.9\times0.21+29\times0.02=1.166$(万元). 由题图①可知 A 公司员工月均工资在 0.6 万元以下的比例为 $0.21+0.29=0.5$,所以估计 A 公司员工月均工资的中位数为 0.6 万元. 用中位数更能反映该公司普通员工的工资水平,理由如下:因为平均数受每一个数据的影响,越离群的数据对平均数的影响越大,该公司极少数员工的月均工资很高,在这种情况下平均数并不能较好地反映普通员工的工资水平,而中位数不受少数极端数据的影响,可以较好地反映普通员工的工资水平.

(2) B 公司员工月均工资的平均数为

$(0.3\times0.375+0.5\times0.750+0.7\times2.750+0.9\times1.000+1.1\times0.125)\times0.2=0.69$(万元).

由题图②知,B公司员工月均工资在0.6万元以下的频率为$(0.375+0.750)\times0.2=0.225$,在0.8万元以下的频率为$(0.375+0.750+2.750)\times0.2=0.775$.

设B公司员工月均工资的中位数为x万元,则$(x-0.6)\times2.750=0.5-0.225$,解得$x=0.7$.

小明应选择B公司应聘,理由如下:B公司员工工资数据较为集中,月均工资的平均数和中位数均能反映该公司普通员工的平均工资水平,B公司员工月均工资平均数为0.69万元,中位数为0.7万元,均大于A公司员工月均工资的中位数0.6万元,所以以公司普通员工的工资水平作为决策依据,小明应该选B公司应聘.

9.3 统计案例 公司员工的肥胖情况调查分析(略)

真题上分

1. B 【解析】依题意,全球年平均气温在区间[14.35,14.75]内的频率是$0.2\times0.5+0.2\times0.65=0.23$,故全球年平均气温在区间[14.35,14.75]内的有$100\times0.23=23$(年).**故选B.**

2. B 【解析】由题意知,讲座前问卷答题的正确率分别为65%,60%,70%,60%,65%,75%,90%,85%,80%,95%,将其按照从小到大的顺序排列为60%,60%,65%,65%,70%,75%,80%,85%,90%,95%,所以其中位数为$\frac{70\%+75\%}{2}=72.5\%$,**故A不正确;**

讲座后问卷答题的正确率分别为90%,85%,80%,90%,85%,85%,95%,100%,85%,100%,所以其平均数为$\frac{1}{10}\times(90\%+85\%+80\%+90\%+85\%+85\%+95\%+100\%+85\%+100\%)=89.5\%$,**故B正确;**

由题图可知，讲座前问卷答题的正确率比讲座后正确率的波动更大，所以讲座前问卷答题的正确率的标准差大于讲座后正确率的标准差，**故 C 不正确**；

由题图可知，讲座前问卷答题的正确率的极差为 $95\%-60\%=35\%$，讲座后问卷答题的正确率的极差为 $100\%-80\%=20\%$，**故 D 不正确. 故选** B.

3.【解】(1) 设 X 为患病者指标，Y 为未患病者指标，由患病者指标的频率分布直方图，知 $p(c)=P(X\leqslant c)=(c-95)\times 0.002=0.5\%$，解得 $c=97.5$.

则 $q(c)=P(Y>c)=(100-97.5)\times 0.010+5\times 0.002=0.035=3.5\%$.

(2) 当 $95\leqslant c\leqslant 100$ 时，

$p(c)=(c-95)\times 0.002$，$q(c)=(100-c)\times 0.010+5\times 0.002$，

所以 $f(c)=p(c)+q(c)=-0.008c+0.82$；

当 $100<c\leqslant 105$ 时，

$p(c)=5\times 0.002+(c-100)\times 0.012$，$q(c)=(105-c)\times 0.002$，

所以 $f(c)=p(c)+q(c)=0.01c-0.98$.

综上所述，$f(c)=\begin{cases}-0.008c+0.82, & 95\leqslant c\leqslant 100,\\ 0.01c-0.98, & 100<c\leqslant 105.\end{cases}$

由一次函数的单调性知，函数 $f(c)$ 在 $[95,100]$ 上单调递减，在 $(100,105]$ 上单调递增，

所以 $f(c)_{\min}=f(100)=-0.008\times 100+0.82=0.02$.

4. C 【解析】平均数为 $\frac{1}{5}\times(2+8+14+16+20)=12$. **故选** C.

5. C 【解析】A 选项，因为 $6+12+18=36<50$，$36+30=66>50$，所以 100 块稻田亩产量的中位数不小于 1 050 kg，**A 错误**；B 选项，因为 100 块稻田中亩产量低于 1 100 kg 的稻田有 66 块，所占比例为 $66\%<80\%$，**所以 B 错误**；C 选项，100 块稻田亩产量的极差的最大值小于 $1\ 200-900=300$，最小值大于 $1\ 150-950=200$，所以极差介于 200 kg 至 300 kg 之间，**C 正确**；D 选项，同一组中的数据都用左端

点值来估计，则这 100 块稻田亩产量的平均值的最小值为$\frac{1}{100}\times(6\times900+12\times950+18\times1\ 000+30\times1\ 050+24\times1\ 100+10\times1\ 150)=1\ 042>1\ 000$，所以平均值不介于 900 kg 至 1 000 kg 之间，D 错误. 故选 C.

6. BD 【解析】对于选项 A：$\because x_1, x_6$ 不确定，$\therefore x_1, x_2, \cdots, x_6$ 的平均数不确定，如 1,2,2,2,2,4 的平均数不等于 2,2,2,2 的平均数，故 A 错误；

对于选项 B：不妨设 $x_2\leqslant x_3\leqslant x_4\leqslant x_5$，则 x_2, x_3, x_4, x_5 的中位数为$\frac{x_3+x_4}{2}$，$x_1, x_2, x_3, x_4, x_5, x_6$ 的中位数为$\frac{x_3+x_4}{2}$，故 B 正确；

对于选项 C：$\because x_1, x_2, x_3, x_4, x_5, x_6$ 的波动性不小于 x_2, x_3, x_4, x_5 的波动性，$\therefore x_2, x_3, x_4, x_5$ 的标准差不大于 $x_1, x_2, x_3, x_4, x_5, x_6$ 的标准差，故 C 错误；

提示：标准差反映数据的离散程度，数据越离散，标准差越大

对于选项 D：不妨设 $x_2\leqslant x_3\leqslant x_4\leqslant x_5$，则 $x_1\leqslant x_2\leqslant x_3\leqslant x_4\leqslant x_5\leqslant x_6$，$\therefore x_5-x_2\leqslant x_6-x_1$，即 x_2, x_3, x_4, x_5 的极差不大于 $x_1, x_2, x_3, x_4, x_5, x_6$ 的极差，故 D 正确.

定义：极差为样本数据的最大值减去最小值

故选 BD.

7. 【解】(1) 因为 $z_i=x_i-y_i(i=1,2,\cdots,10)$，

所以 $z_1=9, z_2=6, z_3=8, z_4=-8, z_5=15, z_6=11, z_7=19, z_8=18, z_9=20, z_{10}=12$，

所以 $\bar{z}=\frac{1}{10}\times(9+6+8-8+15+11+19+18+20+12)=11$，

所以 $s^2=\frac{1}{10}\times(4+25+9+361+16+0+64+49+81+1)=61$.

(2) 因为 $2\sqrt{\frac{s^2}{10}}=2\times\sqrt{\frac{61}{10}}=\sqrt{24.4}<11$，即 $2\sqrt{\frac{s^2}{10}}<\bar{z}$，

所以甲工艺处理后的橡胶产品的伸缩率较乙工艺处理后的橡胶产品的伸缩率有显著提高.

素养上分

1. ABD 【解析】由题图可知，每一年的数据点对应的圆的圆心所在的高度呈现上升趋势，故自 1990 年至 2023 年，我国人口总数大致呈增长趋势，A 正确；

由于这些圆呈现变大的趋势，故半径呈现变大的趋势，因此城镇化率大至呈增长趋势，B 正确；

由于我国人口总数大致呈增长趋势且城镇化率也呈增长趋势，因此自 1990 年至 2023 年，我国城镇化率与人口总数呈正相关，D 正确；

根据题图，无法得知人口增长率的变化情况，故 C 错误. 故选 ABD.

2. AD 【解析】由雷达图可知，该校秋季运动会有 100 米、200 米、铅球、1 000 米、跳远、跳高这六个项目，A 班的得分分别为 4，4，3，5，4，3，B 班的得分分别为 5，3，4，5，3，4.

对于 A，在 200 米项目中，A 班的得分比 B 班的得分高，A 正确.

对于 B，$6\times0.80=4.8$，故第 80 百分位数为第五个数，A 班的得分从小到大排序为 3，3，4，4，4，5；

B 班的得分从小到大排序为 3，3，4，4，5，5，所以 A 班的第 80 百分位数为 4，B 班的第 80 百分位数为 5，故 A 班的第 80 百分位数比 B 班的第 80 百分位数低，B 错误.

对于 C，A 班的众数为 4，B 班的数据中 3，4，5 出现的次数一样多，众数为 3，4，5，C 错误.

对于 D，A 班的总分为 $4+4+3+5+4+3=23$，B 班的总分为 $5+3+4+5+3+4=24$，所以 B 班的总分比 A 班的总分高，D 正确. 故选 AD.

3. 【解】（1）由题可得，$10+a+30+30+10=100$，

解得 $a=20$，则 $b=\dfrac{20}{100}=0.2$，

样本成绩的平均数的估计值为 $55\times0.1+65\times0.2+75\times0.3+85\times0.3+95\times0.1=76$.

（2）由题表可知，分数在区间［50，60）内

的频数为 10，在区间[60,70)内的频数为 20，

故两组成绩的总平均数 $\bar{z}=\frac{10}{10+20}\times 56+\frac{20}{10+20}\times 65=62$，

两组成绩的总方差 $s^2=\frac{10}{10+20}\times[7+(56-62)^2]+\frac{20}{10+20}\times[4+(65-62)^2]=23$.

所以两组成绩的总平均数是 62，总方差是 23.

4.【解】(1)甲品种产量的样本均值 $\overline{x_甲}=\frac{47+51+49+50+53}{5}=50(\text{kg})$，极差为 $53-47=6$；

乙品种产量的样本均值 $\overline{x_乙}=\frac{44+51+60+58+52}{5}=53(\text{kg})$，极差为 $60-44=16$.

所以从均值与极差来看，甲品种的产量略低于乙品种，但比较稳定；乙品种的产量较高，但波动较大.

(2)甲品种产量的样本方差 $s_甲^2=\frac{1}{5}\times[(47-50)^2+(51-50)^2+(49-50)^2+(50-50)^2+(53-50)^2]=4$，

所以甲品种产量的变异系数 $CV_1=\frac{\sqrt{4}}{50}\times 100\%=4\%$；

乙品种产量的样本方差 $s_乙^2=\frac{1}{5}\times[(44-53)^2+(51-53)^2+(60-53)^2+(58-53)^2+(52-53)^2]=32$，

所以乙品种产量的变异系数 $CV_2=\frac{\sqrt{32}}{53}\times 100\%=\frac{4\sqrt{2}}{53}\times 100\%\approx 10.7\%$.

因为 $CV_1<CV_2$，所以甲品种的产量更稳定，生产的风险也更小，更适合推广.

5.【解】(1)根据频率分布直方图可得，

第一组[20,25)的频率为 $0.01\times 5=0.05$，

第二组[25,30)的频率为 $0.07\times 5=0.35$，

第三组[30,35)的频率为 $0.06\times 5=0.3$，

第四组[35,40)的频率为 $b\times 5=5b$，

第五组[40,45)的频率为 $0.02\times 5=0.1$，

因为频率之和为 1，所以 $0.05+0.35+$

$0.3+5b+0.1=1$，即 $5b=0.2$，解得 $b=0.04$.

因为第一组[20,25)的频率为 0.05，且第一组有 10 人，所以 $\frac{10}{m}=0.05$，解得 $m=200$.

(2)设平均年龄为 $\bar{x}$，则 $\bar{x}=0.05\times22.5+0.35\times27.5+0.3\times32.5+0.2\times37.5+0.1\times42.5=32.25$(岁).

设第 80 百分位数为 a，根据每一组频率得前三组的频率为 $0.05+0.35+0.3=0.7<0.8$，前四组的频率为 $0.05+0.35+0.3+0.2=0.9>0.8$，因此第 80 百分位数在第四组，

所以 $0.7+(a-35)\times0.04=0.8$，解得 $a=37.5$.

综上，平均年龄为 32.25 岁，第 80 百分位数为 37.5.

6. BC 【解析】根据题意，同时选 A,B 的人数在 10% 到 20% 之间，换算成人数为 $2\,017\times10\%\approx202$，$2\,017\times20\%\approx403$，即同时选择 A,B 两门课的人数在 202 到 403 之间，因此符合题意的选项有 B，C.

故选 BC.

第九章 全章上分

1. B 【解析】该市场监管局的调查方法是随机抽样,**故 A 错误**;个体是每种冷藏饮品的质量,**故 B 正确**;调查的总体是超市在售的 40 种冷藏饮品的质量,**故 C 错误**;样本容量是 20,**故 D 错误**.

2. D 【解析】在简单随机抽样的过程中,每个个体被抽到的概率都相等,从该年级文科生中用简单随机抽样的方法抽出 20 人,所有班的学生被抽到的概率都一样,男生、女生被抽到的概率也都一样,即其中任意两个人被同时抽到的概率一样,**故 D 正确**.

3. D 【解析】对于 A,2021 年至 2022 年,两项访问量的增长量分别为 54 880,346 913,显然增长幅度是较大的,**故 A 正确**;对于 B,由题表可知 2022 年至 2023 年,两项访问量都有回落,**故 B 正确**;对于 C,2023 年至 2024 年,两项访问量的增长量分别为 16 989,24 505,**故 C 正确**;对于 D,由 B 分析知,该市政府部门网站的两项访问量在 2022 年至 2023 年都有回落,而不是逐年增长态势,**故 D 错误**.

4. B 【解析】因为感染人群中 O 型血、A 型血、B 型血、AB 型血的人数比为 4∶3∶3∶2,所以抽取样本量为 n 的样本中,O 型血的人数为 $\frac{4}{4+3+3+2}n$,AB 型血的人数为 $\frac{2}{4+3+3+2}n$,所以 $\frac{4}{4+3+3+2}n-\frac{2}{4+3+3+2}n=20$,解得 $n=120$. **故 B 正确**.

一题多解 由题知,$\frac{4-2}{4+3+3+2}=\frac{20}{n}$,解得 $n=120$. **故 B 正确**.

5. D 【解析】由随机数法可知,以 3 个数字为单位抽取数字,且数字不能大于 240,还要去掉重复数字,据此第一个数字为 114,第二个数字为 165,第三个数字为 100,第四个数字为 210. **故 D 正确**.

6. B 【解析】选择物理的学生人数为 40−

$30+10=20$,即该校选择物理的学生人数与该校学生总人数比值的估计值为$\frac{20}{100}=0.2$. **故 B 正确.**

7. C 【解析】由于$\frac{12}{1-0.4-0.1-0.26}=50$,所以该考场总共有 50 人,所以化学考试获得一等奖的有 $50\times(1-0.16-0.38-0.38)=4$(人),**所以①正确**;估计全校物理考试获得二等奖的有 $1\ 000\times0.24=240$(人),**所以②正确**;如果采用按比例分配的分层随机抽样从全校抽取 200 人,则化学考试被淘汰的人数为 $200\times0.38=76$,**所以③错误. 故 C 正确.**

8. A 【解析】由题意进行数据分析,可得
$0.020\times(10-0)+0.010\times(20-10)+0.030\times(30-20)+0.015\times(x_1-30)=0.75$,解得 $x_1=40$;
$0.010\times(10-0)+0.020\times(20-10)+0.030\times(30-20)+0.025\times(x_2-30)=0.75$,解得 $x_2=36$,
所以 $x_1>x_2$. 比较两个频率分布直方图可以看出,雪上项目的数据更分散,冰上项目的数据更集中,由方差的意义可以得到 $s_1^2>s_2^2$. **故 A 正确.**

9. AD 【解析】对于 A,由扇形图可知,30~41 周岁的理财人数最多,**故 A 正确**;对于 B,由折线图可知,随着年龄的增长,人均理财费用越来越多,**故 B 错误**;对于 C,由扇形图可知,30 周岁及以上的理财人数约占总理财人数的 80%,**故 C 错误**;对于 D,由柱状图可知,采用丁理财方式的比例最高,**故 D 正确.**

10. BCD 【解析】对于 A,甲的数据介于[1.5,7.5]之间,极差小于或等于6;乙的数据介于[2.5,8.5]之间,极差小于或等于 6,从而甲和乙的极差可能相等,**故 A 错误.** 对于 B,根据频率分布直方图可知,甲的众数的估计值为 3,4,5,乙的众数的估计值为 6,乙的众数大于甲的众数,**故 B 正确.** 对于 C,甲的数据较分散,乙的数据集中分布在中间,因此甲的方差大于乙的方差,**故 C 正确.** 对于 D,对

于甲，各组频率依次为 0.15，0.20，0.20，0.20，0.15，0.10，因为前两组频率之和为 $0.15+0.20=0.35<0.5$，前三组频率之和为 $0.15+0.20+0.20=0.55>0.5$，故中位数位于[3.5，4.5)之间；同理，对于乙，各组频率依次为 0.05，0.10，0.15，0.35，0.20，0.15，前三组频率之和为 $0.05+0.10+0.15=0.3<0.5$，前四组频率之和为 $0.05+0.10+0.15+0.35=0.65>0.5$，故中位数位于[5.5，6.5)之间，所以乙的中位数大于甲的中位数，**故 D 正确**.

11. CD 【解析】对于 A，$\because (0.015+0.033+a+0.011+0.011)\times10=1$，$\therefore a=0.03$，**故 A 错误**；对于 B，由频率分布直方图知短视频观众年龄在 10～20 岁的对应频率为 0.15，$\therefore$ 短视频观众年龄在 10～20 岁的有 $4\ 000\times0.15=600$(人)，**故 B 错误**；对于 C，估计平均年龄为 $(0.015\times15+0.033\times25+0.03\times35+0.011\times45+0.011\times55)\times10=32$(岁)，**故 C 正确**；对于 D，设 75% 分位数为 x，由题图知年龄在 10～20 岁和 20～30 岁两组的频率是 $(0.015+0.033)\times10=0.48$，年龄在 10～20 岁，20～30 岁和 30～40 岁三组的频率是 $(0.015+0.033+0.03)\times10=0.78$，$\therefore$ 75% 分位数位于年龄在 30～40 岁这一组内，则 $0.48+(x-30)\times0.03=0.75$，解得 $x=39$，**故 D 正确**.

12. 6 【解析】由题意得男性、女性人数比为 5∶3，则女性参会人员应抽取的人数为 $16\times\dfrac{3}{5+3}=6$.

13. 甲、丙 【解析】①甲地：5 个数据的中位数为 24，众数为 22，根据数据得出甲地连续 5 天的日平均温度的记录数据由小到大为 22，22，24，x，y($x\neq y$)，其连续 5 天的日平均气温均不低于 22 ℃；②乙地：5 个数据的中位数为 27，总体均值为 24，当 5 个数据为 19，20，27，27，27 时，可知其连续 5 天的日平均温度有低于 22 ℃的，故不确定乙地是否进入夏季；③丙地：5 个数据中有 1 个数据是 32，总体均值为 26，若有低于 22 ℃的，假设取

21,此时方差就超过了 10.8,可知其连续 5 天的日平均温度均不低于 22 ℃,则肯定进入夏季的地区有甲、丙.

14. B 【解析】设 A 小区全部居民每月用电量的平均数和方差分别为 $\overline{w_A}$ 和 s_A^2,B 小区全部居民每月用电量的平均数和方差分别为 $\overline{w_B}$ 和 s_B^2,用样本估计总体,则估计 $\overline{w_A}=\frac{1}{12+15+13}\times(12\times198.5+15\times210+13\times168)=192.9$,所以 $s_A^2=\frac{1}{40}\times\{12\times[10.22+(198.5-192.9)^2]+15\times[20.22+(210-192.9)^2]+13\times[28.56+(168-192.9)^2]\}=340.4955$,$\overline{w_B}=\frac{1}{10+20+20}\times(10\times170.6+20\times200.2+20\times180.5)=186.4$,$s_B^2=\frac{1}{50}\times\{10\times[27.15+(170.6-186.4)^2]+20\times[14.55+(200.2-186.4)^2]+20\times[19.56+(180.5-186.4)^2]\}=159.102$,因为 $s_A^2>s_B^2$,所以推测 B 小区每月用电量更稳定.

15. 【解】(1)根据题意可知,方式 1 采用的是简单随机抽样法,方式 2 采用的是分层随机抽样法.

(2)方式 1 抽样的步骤如下:

在全年级 10 个班中用抽签法任意抽取 1 个班级,考查他们的成绩.

方式 2 抽样的步骤如下:

第一步:分层,把该校高三年级的学生按成绩分成优秀、良好、普通三个级别;

第二步:确定各个层抽取的人数,由于样本容量与总体个数比值为 $\frac{40}{400}=\frac{1}{10}$,所以每层抽取的个体数依次为 $60\times\frac{1}{10}=6$,$180\times\frac{1}{10}=18$,$160\times\frac{1}{10}=16$;

第三步:按层分别抽取样本人数,在优秀学生中用简单随机抽样法抽取 6 人,在良好学生中用简单随机抽样法抽取 18 人,在普通学生中用简单随机抽样法抽取 16 人,并考查他们的成绩.

16. 【解】(1)由频率分布直方图可知,$(0.010+0.020+a+0.050+0.065+a+$

$0.015+0.010+0.005)\times4=1$，解得 $a=0.0375$.

(2)居民用水量为 20 m^3 时，收费为 60 元，所以用水费用不超过 60 元，则用水量小于等于 20 m^3，由频率分布直方图可知，用水量小于等于 20 m^3 的频率为 $(0.010+0.020+0.0375+0.050+0.065)\times4=0.73$，$20\times0.73=14.6$（万户），所以估计全市居民中月均用水费用不超过 60 元的用户有 14.6 万户.

(3)抽取的 100 户居民月均用水量不超过 28 m^3 的频率为 $(0.010+0.020+0.0375+0.050+0.065+0.0375+0.015)\times4=0.94$，$0.94<0.95$，所以现行收费标准不符合要求，抽取的 100 户居民月均用水量不超过 32 m^3 的频率为 $(0.010+0.020+0.0375+0.050+0.065+0.0375+0.015+0.010)\times4=0.98$，$\frac{0.95-0.94}{0.98-0.94}\times(32-28)=1$，所以需将第二阶梯用水量的上限至少上调到 29 m^3.

17.【解】(1)60 岁以上人口比例是 $(0.01+0.003+0.0025+0.0005)\times10=0.16$，所以由①可分析出该地区人口已经老龄化.

(2)由折线统计图可知，该地区年龄在 71～80 岁且已签约家庭医生的居民有 $0.03\times0.7\times1000=21$（万人）.

(3)由图①和图②可知该地区年龄段为 18～30 岁的人口在 180～230 万之间，签约率为 30.3%；

年龄段为 31～50 岁的人口为 $(0.02+0.016)\times10\times1000=360$（万），签约率为 37.1%；

年龄段为 51～60 岁的人口为 $0.015\times10\times1000=150$（万），签约率为 55.7%；

年龄段为 61～70 岁的人口为 $0.01\times10\times1000=100$（万），签约率为 61.7%；

年龄段为 71～80 岁的人口为 $0.003\times10\times1000=30$（万），签约率为 70%；

年龄段为 81 岁及以上的人口为 $(0.0025+0.0005)\times10\times1000=30$（万），签约率为 75.8%.

由以上数据可知，这个地区在 31～50 岁

这个年龄段人数为 360 万，基数较其他年龄段是最大的，且签约率仅为 37.1%，比较低，所以应着重提高此年龄段的签约率.

18.【解】(1)由题可知 $8+8+10+24=50$，所以这 100 人在 5 年使用期内更换刷头的个数的中位数为$\frac{17+18}{2}=17.5$.

(2)若购买该品牌电动牙刷的同时购买刷头，则每个刷头 20 元；若单独购买刷头，则每个刷头 30 元，则当 $n\leqslant 18$ 时，$y=20n$，当 $n>18$ 时，$y=360+30(n-18)=30n-180$，所以 y 关于 n 的函数解析式为

$$y=\begin{cases}20n, n\leqslant 18,\\ 30n-180, n>18,\end{cases} n\in\mathbf{N}.$$

(3)若这 100 人购买 1 个该品牌电动牙刷的同时都购买了 17 个刷头，则其中有 50 人购买刷头的费用为 340 元，28 人购买刷头的费用为 370 元，12 人购买刷头的费用为 400 元，10 人购买刷头的费用为 430 元，因此这 100 人购买刷头的费用的平均数为$\frac{1}{100}\times(50\times 340+28\times 370+12\times 400+10\times 430)=364.6$(元)；

若这 100 人购买 1 个该品牌电动牙刷的同时都购买了 18 个刷头，则其中有 78 人购买刷头的费用为 360 元，12 人购买刷头的费用为 390 元，10 人购买刷头的费用为 420 元，因此这 100 人购买刷头的费用的平均数为$\frac{1}{100}\times(78\times 360+12\times 390+10\times 420)=369.6$(元).

因为 $364.6<369.6$，所以购买 1 个该品牌电动牙刷的同时应购买 17 个刷头.

19.【解】(1)因为频率之和为 1，所以 Y 之和为 0.1 得 $0.1\times(k+2k+4k+4k+3k+2k)=0.1$，解得 $k=\frac{1}{16}$.

(2)根据 B 校学生成绩的频率分布直方图，设其中位数为 x，则 x 在使频率分布直方图中矩形左右两边面积相等的位置，易知 $x=50$，设众数为 y，由众数的估计值为最高矩形底边的中点值，易知 $y=45$.

(3)设总误判率为$f(t)$,又$t\in[50,70)$,

则当$t\in[50,60)$时,$f(t)=0.1\times 10\times\frac{1}{16}+0.1\times\frac{1}{8}\times(t-50)+0.0375\times(60-t)+0.125=-\frac{1}{40}t+\frac{29}{16}$;

当$t\in[60,70)$时,$f(t)=0.1\times 10\times\frac{1}{16}+0.1\times 10\times\frac{1}{8}+0.1\times(t-60)\times\frac{1}{4}+0.0625+0.00625\times(70-t)=\frac{3t}{160}-\frac{13}{16}$,

由$f(t)$的单调性知,当$t=60$时,$f(t)$最小,此时$f(t)=\frac{5}{16}$,所以总误判率最小为$\frac{5}{16}$,此时$t=60$.