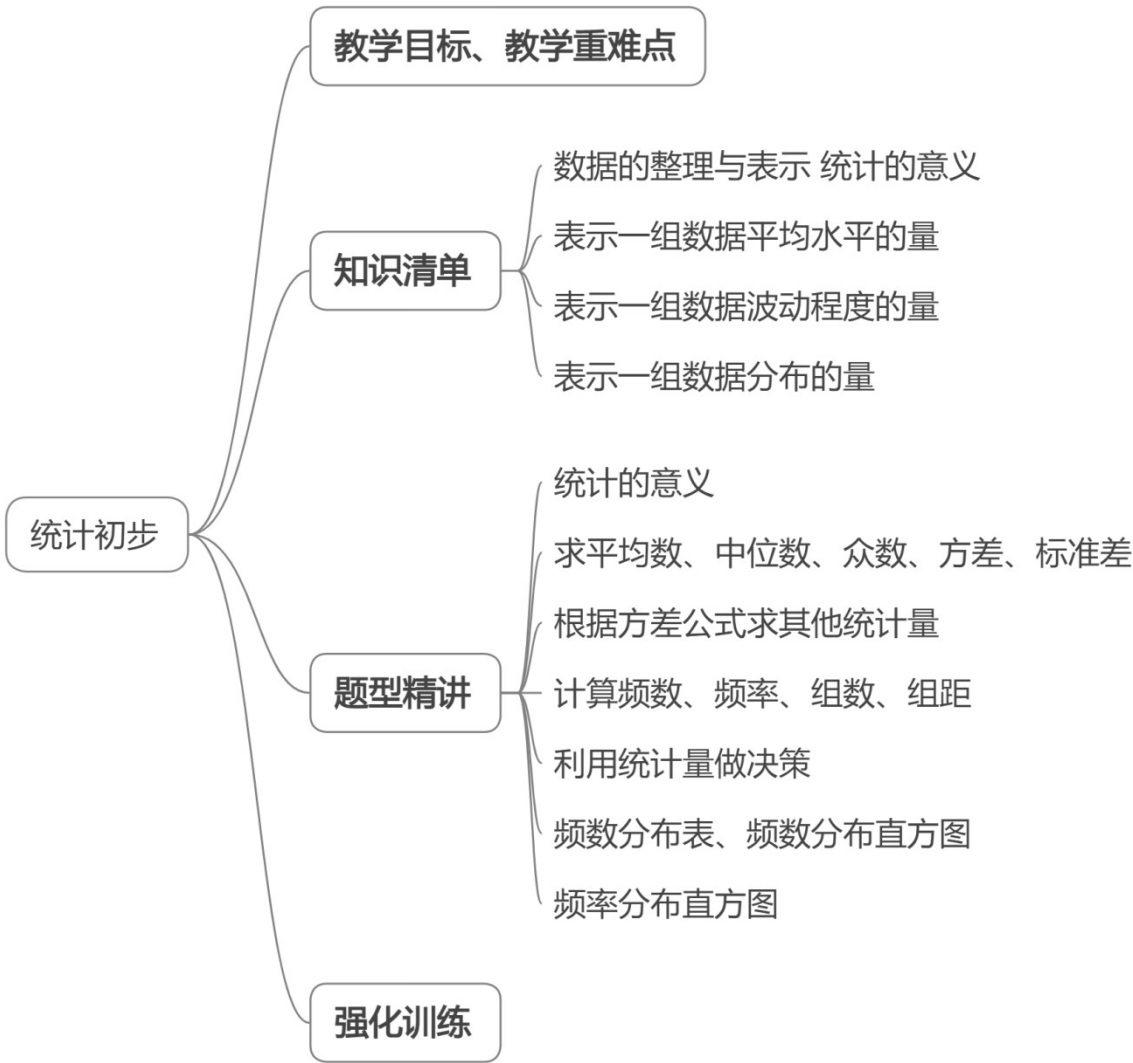


第 28 章 统计初步

内容概览



教学目标、教学重难点

教学目标	1. 学习数据描述的方法；统计图的选择； 2. 知道统计的意义；学会用样本估计总体； 3. 会求平均数，加权平均数，截尾平均数； 4. 了解中位数与众数的统计意义；会求方差与标准差；学会用它们做决策； 5. 掌握频数分布直方图；频率分布直方图及其应用。
------	---

教学重难点	1.重点 (1) 统计图； (2) 表示一组数据平均水平的量；表示一组数据波动程度平的量； (3) 表示一组数据分布量。 2.难点 (1) 平均数、中位数、众数等含参数问题； (2) 统计初步的综合应用。
--------------	---

知识清单

知识点 1 数据整理与表示 统计的意义

一、数据的描述

描述数据的方法有两种：统计表和统计图。

统计表：利用表格将要统计的数据填入相应的表格内，表格统计法可以很好地整理数据

统计图：利用“条形图”、“扇形图”、“折线图”描述数据，这样做的最大优点是将表格中的数据所呈现出来的信息直观化。

要点：

(1) **条形统计图**：用线段长度表示数据，根据数据的多少画成长短不同的长方形直条，然后按顺序把这些直条排列起来，条形统计图很容易看出数据的大小，便于比较，但不能清楚地反映各部分占总体的百分比。

(2) **扇形统计图**：用整个圆表示总体，用圆内各个扇形的大小表示各部分数量，从扇形上可清楚地看出各部分量和总数量之间的关系，但不能直接表示出各个项目的具体数据。

(3) **折线统计图**：用一个单位长度表示一定的数量，根据数量的多少描出各点，然后把各点用线段依次连接起来，折线图不但可以表示出数量的多少，而且能够清楚地表示出数量的增减变化情况，但不能清楚地反映数据的分布情况。

二、统计图的选择

统计图：利用“条形图”、“扇形图”、“折线图”描述数据，这样做的最大优点是将表格中的数据所呈现出来的信息直观化。

要点：

(1) **条形统计图**：用线段长度表示数据，根据数据的多少画成长短不同的长方形直条，然后按顺序把这些直条排列起来，条形统计图很容易看出数据的大小，便于比较，但不能清楚地反映各部分占总体的百分比。

(2) **扇形统计图**：用整个圆表示总体，用圆内各个扇形的大小表示各部分数量，从扇形上可清楚地看出各部分量和总数量之间的关系，但不能直接表示出各个项目的具体数据。

(3) **折线统计图**：用一个单位长度表示一定的数量，根据数量的多少描出各点，然后把各点用线段依次连接起来，折线图不但可以表示出数量的多少，而且能够清楚地表示出数量的增减变化情况，但不能清

楚地反映数据的分布情况。

三、统计的意义

1. 统计学相关概念

总体：调查时，调查对象的全体叫做总体。

个体：组成总体的每一个调查对象叫做个体。

样本：从总体中取出的一部分个体叫做总体的一个样本。

样本容量：样本中个体的数量叫做样本容量（不带单位）。

要点：

(1)“调查对象的全体”一般是指调查对象的某种数量指标的全体，如对于一个班级，如果考察的是这个班学生的身高，那么总体是指这个班学生身高的全体，不能错误地理解为学生的全体是总体。

(2)样本是总体的一部分，一个总体中可以有多样本，样本在一定程度上能够反映总体，为了使样本能较好地反映总体情况，在选取样本时要注意使其具有一定的代表性。

(3) 样本容量是一个数字，不能有单位。一般地，样本容量越大，通过样本对总体的估计越精确，在实际研究中，要根据具体情况确定样本容量的大小。例如：“从 5 万名考生的数学成绩中抽取 2000 名考生的数学成绩进行分析”，样本是“2000 名考生的数学成绩”，而样本容量是“2000”，不能将其误解为“2000 名考生”或“2000 名”。

2. 调查的方法：普查和抽样调查

(1) **普查**：考察全体对象的调查叫做普查。

要点：

(1)普查又叫“普查”，它是指在统计的过程中，为了某种特定的目的而对所有考察的对象一一作出的调查，在记录数据时，通常用划记法进行记录数据。

(2)一般来说，普查能够得到全体被调查对象的全面、准确的信息，但有时总体中的个体的数目非常大，普查的工作量太大；有时受条件的限制，无法进行普查；有时调查具有破坏性(例如：测试一批灯泡的使用寿命或炮弹的杀伤半径等)，不能进行普查。

(2) **抽样调查**：从调查对象中抽取部分对象进行调查，然后根据调查的数据推断全体对象的情况，这种调查方式称为抽样调查。

要点：

(1)从总体中抽取部分个体进行调查的方式，我们称抽样调查，在抽取的过程中，总体中的每一个个体都有相等的机会被抽到，像这样的抽样方式是一种简单随机抽样。

(2)抽样调查方便、快捷，能够减少调查统计的工作量但调查的结果不如“普查”得到的结果准确。

(3) **调查方法的选择：**

①普查是对考查对象的全体调查，它要求对考查范围内所有个体进行一个不漏的逐个准确统计；而抽样调查则只是对总体中的部分个体进行调查，以样本来估计总体的情况。

②在调查实际生活中的相关问题时，要灵活处理，既要考虑问题本身的需要，又要考虑实现的可能性和所付出代价的大小。

3. 随机样本

具有代表性的样本叫作随机样本。

4. 用样本估计总体

比如在保护区内不同的地方，将 20 只黑叶猴背上涂一个色块做标记，再放归野外，一个月后如果在保护区

内不同的地方观察到 60 只黑叶猴，发现其中 2 只黑叶猴有记号，那么我们就能粗略估计该自然保护区里黑叶猴的数量，这里假定有记号的黑叶猴在自然保护区里是均匀分布的，观察到的黑叶猴又是随机的，因此 60 只黑叶猴中出现有记号的黑叶猴数与 60 的比，应该与保护区所有黑叶猴中存在 20 只有记号的黑叶猴数与黑叶猴总数的比大致相同。

经计算，自然保护区里大约有 600 只黑叶猴。

【即学即练】

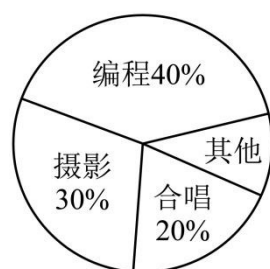
1. 下列调查中，适宜采用抽样调查的是 ()

- A. 调查冬奥会高山滑雪运动员兴奋剂的使用情况
- B. 调查某批次汽车的抗撞击能力
- C. 神舟二十一号载人飞船发射前对零部件的检查
- D. 调查全班观看电影《哪吒 2》的情况

2. 要制作空气中各成分所占百分比，选取的统计图是 ()

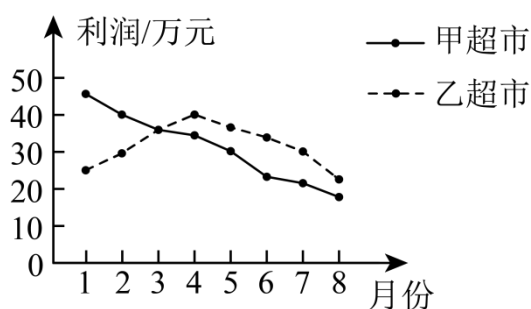
- A. 条形统计图
- B. 折线统计图
- C. 扇形统计图
- D. 三种都可以

3. 为了解某校七年级 a 名学生参加社团的情况，小郑随机抽取部分学生进行调查统计，并绘制如图所示的扇形统计图，那么下列说法不正确的是 ()



- A. 参加编程的学生有 $0.4a$ 人
- B. 参加摄影所在扇形的圆心角度数为 120°
- C. 参加编程的人数是参加合唱人数的 2 倍
- D. 参加其他社团的人数占总人数的 10%

4. 甲、乙两家超市 1~8 月的月利润情况如图所示，下列说法中，不正确的是 ()



- A. 甲超市的月利润逐月减少 B. 4~8月乙超市的月利润逐月减少
C. 3月甲、乙两家超市的月利润相等 D. 6月甲、乙两家超市的月利润相差最大

知识点2 表示一组数据平均水平的量

一、平均数和加权平均数

一般地, 对于 n 个数 $x_1, x_2, x_3, \dots, x_n$, 我们把 $\frac{1}{n}(x_1 + x_2 + x_3 + \dots + x_n)$ 叫做这 n 个数的平均数, 记作 $\bar{x}$. 计算公式为 $\bar{x} = \frac{1}{n}(x_1 + x_2 + x_3 + \dots + x_n)$.

要点:

平均数表示一组数据的“平均水平”, 反映了一组数据的集中趋势.

(1) 当一组数据较大时, 并且这些数据都在某一常数 a 附近上、下波动时, 一般选用简化计算公式 $\bar{x} = \bar{x}' + a$. 其中 $\bar{x}'$ 为新数据的平均数, a 为取定的接近这组数据的平均数的较“整”的数.

(2) 平均数的大小与一组数据里的每个数据均有关系, 其中任一数据的变动都会相应引起平均数的变动. 所以平均数容易受到个别特殊值的影响.

若 n 个数 $x_1, x_2, \dots, x_n$ 的权分别是 $w_1, w_2, \dots, w_n$, 则 $\frac{x_1 w_1 + x_2 w_2 + \dots + x_n w_n}{w_1 + w_2 + \dots + w_n}$ 叫做这 n 个数的加权平均数.

要点:

(1) 相同数据 x_i 的个数 w_i 叫做权, w_i 越大, 表示 x_i 的个数越多, “权”就越重. 数据的权能够反映数据的相对“重要程度”.

(2) 加权平均数实际上是算术平均数的另一种表现形式, 是平均数的简便运算.

二、样本平均数与总体平均数

我们把样本中所有个体的平均数称为样本平均数; 把总体中所有个体的平均数称为总体平均数. 随机样本的容量越大, 样本平均数就越接近于总体平均数. 必要时, 可以用样本平均数来估计总体平均数.

三、截尾平均数

截尾平均数: 去掉一个最大值, 去掉一个最小值, 剩下的数求平均数.

四、中位数和众数

1. 中位数的概念: 将一组数据按照由小到大 (或由大到小) 的顺序排列, 如果数据的个数是奇数, 则处于中间位置的数称为这组数据的中位数; 如果数据的个数是偶数, 则中间两个数据的平均数称为这组数据的中位数.

要点: (1) 一组数据的中位数是唯一的; 一组数据的中位数不一定出现在这组数据中.

(2) 由一组数据的中位数可以知道中位数以上和以下数据各占一半.

2. 众数的概念：一组数据中出现次数最多的数据称为这组数据的众数.

要点：(1) 一组数据的众数一定出现在这组数据中；一组数据的众数可能不止一个；如果所有数据出现的次数都一样，那么这组数据就没有众数.

(2) 众数是一组数据中出现次数最多的数据而不是数据出现的次数.

五、平均数、中位数与众数的联系与区别

联系：平均数、众数、中位数都是用来描述数据集中趋势的量，其中以平均数最为重要.

区别：平均数的大小与每一个数据都有关，任何一个数的波动都会引起平均数的波动，当一组数据中有个别数据太高或太低，用平均数来描述整体趋势则不合适，用中位数或众数则较合适. 中位数与数据排列位置有关，个别数据的波动对中位数没影响；众数主要研究各数据出现的频数，当一组数据中不少数据多次重复出现时，可用众数来描述.

【即学即练】

1. 有一组数据：6, 4, 6, 5, 3，则这组数据的平均数、众数、中位数分别是 ()

- A. 5, 6, 5 B. 5, 5, 5 C. 4.8, 6, 6 D. 4.8, 6, 5

2. 一名射击运动员连续射靶 8 次，命中的环数如下：8, 9, 10, 9, 7, 8, 10, 8. 这名运动员射击环数的众数与中位数分别是 ()

- A. 9 环与 8 环 B. 8 环与 8.5 环 C. 8.5 环与 9 环 D. 8 环与 8 环

3. 为建设生态城市，某中学在植树节那天，组织九年级八个班的学生到西城新区植树，各班植树情况如下表：

班 级	一	二	三	四	五	六	七	八	合计
棵 数	15	18	22	25	29	14	18	19	160

这组数据的中位数、众数分别是 ()

- A. 18, 18 B. 18.5, 18 C. 19, 19 D. 19.5, 19

知识点 3 表示一组数据波动程度的量

一、*极差、方差和标准差

用一组数据中的最大值减去最小值所得的差来反映这组数据的变化范围，用这种方法得到的差称为***极差**，***极差** = 最大值 - 最小值.

要点：极差是最简单的一种度量数据波动情况的量，它受极端值的影响较大. 一组数据极差越小，这组数据就越稳定.

方差是反映一组数据的整体波动大小的特征的量.方差 s^2 的计算公式是：

$$S^2 = \frac{1}{n} \left[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2 \right]$$

要点：(1) 方差反映的是一组数据偏离平均值的情况.方差越大，数据的波动越大；方差越小，数据的波动越小.

(2) 一组数据的每一个数都加上（或减去）同一个常数，所得的一组新数据的方差不变.

(3) 一组数据的每一个数据都变为原来的 k 倍，则所得的一组新数据的方差变为原来的 k^2 倍.

方差的算术平方根叫做这组数据的**标准差**，用符号 s 表示，即：

$$S = \sqrt{\frac{1}{n} \left[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2 \right]}$$
；标准差的数量单位与原数据一致.

二、*极差、方差和标准差的联系与区别

联系：*极差与方差、标准差都是表示一组数据离散程度的特征数.

区别：*极差表示一组数据波动范围的大小，它受极端数据的影响较大；方差反映了一组数据与其平均值的离散程度的大小.方差越大，稳定性也越小；反之，则稳定性越好.所以一般情况下只求一组数据的波动范围时用极差，在考虑到这组数据的稳定性时用方差.

三、用样本估计总体

在考察总体的平均水平或方差时，往往都是通过抽取样本，用样本的平均水平或方差近似估计得到总体的平均水平或方差.

要点：(1) 如果总体数量太多，或者从总体中抽取个体的试验带有破坏性，都应该抽取样本.取样必须具有尽可能大的代表性.

(2) 用样本估计总体时，样本容量越大，样本对总体的估计也越精确.样本容量的确定既要考虑问题本身的需要，又要考虑实现的可能性所付出的代价.

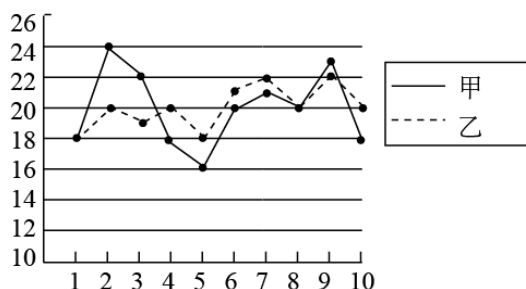
【即学即练】

1. 某小组8位学生一次数学测试的分数为121.5，123.5，123.5，124.5，126.5，127.5，128.5，128.5，那么这个小组测试分数的标准差是_____.

2. 在一场物理实验探究中，甲、乙、丙、丁四名同学分别对同一测量仪器进行10次测量操作，他们测量的平均值相同，测量数据的方差分别是 $s_{\text{甲}}^2 = 2.1$ ， $s_{\text{乙}}^2 = 0.3$ ， $s_{\text{丙}}^2 = 1.7$ ， $s_{\text{丁}}^2 = 0.8$ ，则这四名同学中测量成绩最稳定的是（ ）

A. 甲 B. 乙 C. 丙 D. 丁

3. 甲、乙两地4月下旬的日平均气温统计图如图所示，那么由图中信息可知甲、乙两地这10天日平均气温比较稳定的是_____。（填“甲”或“乙”）



4. 小莹在计算一组数据的方差时，列出没有化简的算式： $s^2 = \frac{(5-\bar{x})^2 + (2-\bar{x})^2 + (5-\bar{x})^2 + (4-\bar{x})^2}{4}$ 关于这

组数据，下列说法正确的是 ()

①平均数是4；②众数是5；③中位数是4；④样本容量是3.

A. ①②

B. ③④

C. ①③

D. ②④

5. 已知一组数据 x_1, x_2, x_3, x_4, x_5 的平均数是2，方差是 $\frac{1}{3}$ ，那么另一组数据

$3x_1 - 2, 3x_2 - 2, 3x_3 - 2, 3x_4 - 2, 3x_5 - 2$ 的平均数是____，方差是_____.

知识点4 表示一组数据分布的量

一、组距、频数、频率与频数分布表

1. **组距**：把所有数据分成若干组，每个小组的两个端点之间的距离（组内数据的取值范围）.

2. **频数**：在统计数据时，某个对象出现的次数或落在某个组别中的数据的个数称为频数.

3. **频率**：频数与总次数的比值称为频率.

4. **频数分布表**：把各个组别中相应的频数分布用表格的形式表示出来，所得表格就是频数分布表.

频数分布表能清楚地反映一组数据的大小分布情况.将一批数据分组，一般数据越多，分的组也越多.当数据在100个以内时，按照数据的多少，常分成5~12组.在分组时，要灵活确定组距，使所分组数合适，

一般组数为 $\frac{\text{最大值}-\text{最小值}}{\text{组距}}$ 的整数部分+1.

要点：

(1) 频数之和等于样本容量，各频率之和等于1；

(2) 制作频数分布表的一般步骤：①计算最大值与最小值的差；②决定组距和组数；③确定分点；

④列频数分布表.

二、频数分布直方图

1. 频数分布直方图

根据频数分布表，用横轴表示各分组数据、纵轴表示各组数据的频数，绘制条形统计图.这样的条形统计图，直观地呈现了频数的分布特征和变化规律，称为频数分布直方图.

2. 画频数分布直方图的步骤

(1) 计算最大值与最小值的差；

(2) 决定组距与组数；

(3)列频数分布表；

(4)画频数分布直方图。

3. 频数分布直方图与条形图的联系与区别

(1)联系：它们都是用矩形来表示数据分布情况的；当矩形的宽度相等时，都是用矩形的高来表示数据分布情况的；频数分布直方图是特殊的条形统计图。

(2)区别：①由于分组数据具有连续性，频数分布直方图中各“条形”之间通常是连续排列，中间没有间隙，而条形图中各“条形”是分开排列的，中间有一定的间隙；②条形统计图用横向指标表示考察对象的类别，用纵向指标表示不同对象的数量。频数分布直方图横向指标表示考察对象数据的变化范围，用纵向指标表示相应范围内数据的频数。

要点：

(1)频数分布直方图简称直方图，它是条形统计图的一种。

(2)注意直方图与条形图、扇形图、折线图在表示数据方面的优缺点。

三、**频率分布直方图**：通常在频率分布直方图中，用每小组对应的小矩形的面积表示该小组的组频率。因此

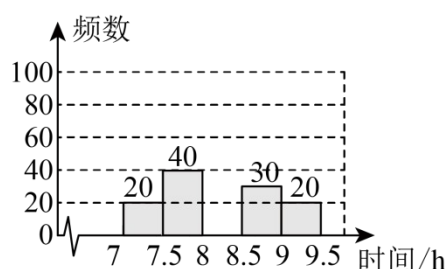
在频率分布直方图中，纵轴表示频率与组距的商，即 $\frac{\text{频率}}{\text{组距}}$ ，横轴的意义与频数分布直方图相同。

【即学即练】

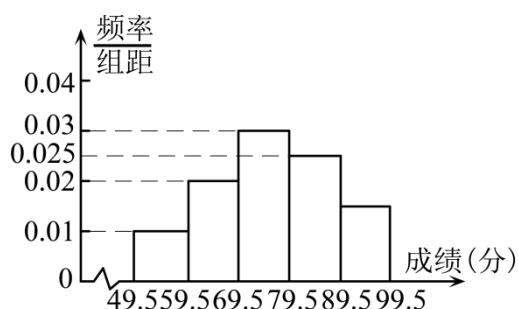
1. 已知一个 50 个数据的样本，把它分成 6 组，第一组到第四组的频数分别是 8、6、11、7，第五组的频率是 0.2，那么第六组的频数是_____。

2. 一组数共有 80 个，最大值是 136，最小值是 52，用频数分布直方图描述这一数据，取组距为 10，则可以分成_____组。

3. 某学校对部分学生的睡眠时间进行调查统计，得到的频数分布直方图（每组含前一个边界值，不含后一个边界值）如图所示，其中睡眠时间在 7.5~8 小时的学生所占百分比为 20%，则睡眠时间在 8~9 小时的学生有_____人。



4. 某区有 1200 名学生参加了“垃圾分类”知识竞赛，为了解本次竞赛成绩分布情况，竞赛组委会从中随机抽取部分学生的成绩(得分都是整数)作为样本，绘制成频率分布直方图(如图)。请根据提供的信息估计该区本次竞赛成绩在 89.5 分~99.5 分的学生有_____名。



题型精讲

题型 01 统计的意义

【典例 1】. 下列调查中，适宜采用普查方式的是 ()

- A. 调查一批新型节能灯泡的使用寿命
- B. 调查上海市中小学生的课外阅读时间
- C. 全市中学生对《流浪地球》影评
- D. 对我国六代战斗机“歼-36”试飞前整机零部件质量的调查

【变式 1】. 为了解某校学生的睡眠时间，下列抽样调查中样本具有代表性的是 ()

- A. 选择九年级一个班的学生进行调查；
- B. 选择全校的男生进行调查；
- C. 对全校成绩排名前 20% 的学生进行调查；
- D. 每个班级随机抽取 10% 的学生进行调查。

【变式 2】. 某校为了了解学生对“二十四节气”的知晓情况，从全校 6000 名学生中，随机抽取了 120 名学生进行调查，在这次调查中 ()

- A. 6000 名学生是总体
- B. 所抽取的 120 名学生是总体的一个样本
- C. 6000 名是样本容量
- D. 所抽取的 120 名学生对“二十四节气”的知晓情况是总体的一个样本

题型 02 求平均数、中位数、众数、方差、标准差

【典例 1】. 一组数据：3，4，4，6，8. 这组数据的平均数是 ()

A . 3 B . 4 C . 5 D . 6

【变式 1】 . 一组数据：5，7，10，5，7，5，6，这组数据的众数和中位数分别是（ ）.

A . 10，7 B . 5，7 C . 6，7 D . 5，6

【变式 2】 . 数据 5,6,7,8,9 的方差等于_____.

【变式 3】 . 某小组 8 位学生一次数学测试的分数为 121.5，123.5，123.5，124.5，126.5，127.5，128.5，128.5，那么这个小组测试分数的标准差是_____.

【变式 4】 . 某小组 5 名同学在一周内参加家务劳动的时间如下表所示，关于“劳动时间”的这组数据，以下说法正确的是（ ）

劳动时间（小时）	3	3.5	4	4.5
人数	1	1	2	1

A . 中位数是 4，平均数是 3.8 B . 众数是 4，平均数是 3.75

C . 众数是 2，平均数是 3.75 D . 众数是 2，中位数是 4

题型 03 根据方差公式求其他统计量

【典例 1】 . 如果样本的方差为 $S^2 = \frac{1}{3}[(x_1 - 8)^2 + (x_2 - 8)^2 + (x_3 - 8)^2]$ ，那么它的样本容量为_____，平均数为_____.

【变式 1】 . 一组数据的方差计算公式如下： $s^2 = \frac{1}{6}[2(7 - \bar{x})^2 + (8 - \bar{x})^2 + 3(9 - \bar{x})^2]$ ，则这组数据的平均数是_____.

【变式 2】 . 已知一组数据的方差计算公式为 $s^2 = \frac{1}{4}[(2 - \bar{x})^2 + (0 - \bar{x})^2 + (2 - \bar{x})^2 + (4 - \bar{x})^2]$ ，则这组数据的中位数是_____，标准差是_____.

题型 04 计算频数、频率、组数、组距

【典例 1】 . “woaiwuxijingkai”中，字母“i”出现的频率为_____.

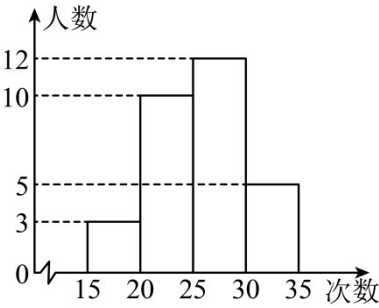
【变式 1】 . 已知在一个样本中，50 个数据分别落在 5 个组内，第一、二、三、四、五组数据的个数分别是 1，9，16，20，4，则第四组的频数是_____.

【变式 2】 . 一个样本容量为 60 的样本，最大值是 135，最小值为 60，取组距为 10，则可以分成_____组.

【变式3】 . 某校为了了解学生在校午餐所需的时间，抽查了 20 名同学在校午餐所需的时间，获得如下数据（单位：分）：10，12，15，10，16，18，19，18，20，34，22，25，20，18，18，20，15，16，21，16．若将这些数据分为 5 组，则组距是（ ）

- A . 4 分 B . 5 分 C . 6 分 D . 7 分

【变式4】 . 某校九年级随机抽查一部分学生进行了 1 分钟仰卧起坐次数的测试，并将其绘制成如图所示的频数分布直方图．那么仰卧起坐次数在 25~30 次的人数占抽查总人数的频率是_____．



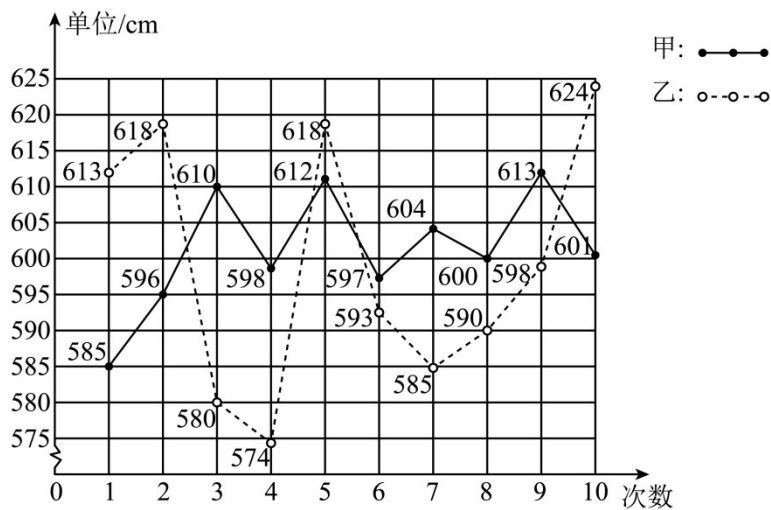
题型 05 利用统计量做决策

【典例1】 . 一家鞋店在一段时间内销售了某种女鞋 30 双，销售量如表：根据表中的数据，可建议鞋店进货时，多进尺码为_____ cm 的女鞋．

尺码/cm	22	22.5	23	23.5	24	24.5	25
销售量/双	1	5	12	6	3	2	1

【变式1】 . 学校为选拔数学竞赛选手，对甲、乙两名同学进行了 4 次模拟测试．已知两人成绩的方差分别为： $S^2_{甲} = 2.5$ ， $S^2_{乙} = 0.5$ ，且两人 4 次测试成绩如下：甲：78，82，79，81，乙：80，81，79，80，根据平均数和方差，应选_____同学参赛．（填“甲”或“乙”）

【变式2】 . 某校要从甲、乙两个跳远运动员中挑选一人参加一项比赛．在最近的 10 次选拔赛中，他们的成绩（单位：cm）折线统计图如图所示：



历届比赛成绩表明，成绩达到 5.96m 就很可能夺冠．若为了稳妥夺冠，则应选择参赛的运动员是 _____
(填“甲”或“乙”) ．

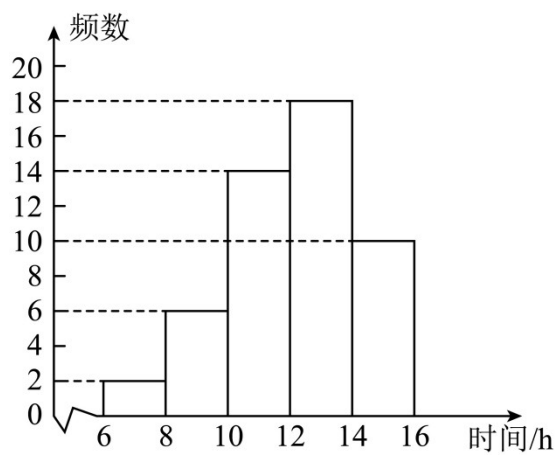
【变式 3】 ．在某次体育测试中，甲、乙两班成绩的平均数、中位数、方差如下表：

班 级	人数	平均数/分	中位数/分	方差
甲 班	45	82	91	19.3
乙 班	45	87	89	5.8

规定学生个人成绩大于 90 分为优秀，则甲、乙两班中优秀人数更多的是_____班 (填“甲”或“乙”) ．

题型 06 频数分布表、频数分布直方图

【典例 1】 ．某校为了解七年级 700 名学生上学期参加社会实践活动的时间，随机对该年级部分学生进行了调查．根据收集的数据绘制了下面的频数分布直方图，则以下说法正确的是 ()



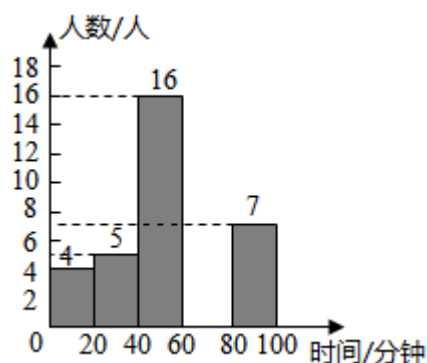
- A．一共调查了 40 名学生
- B．图中五个小长方形的面积比是1:9:49:81:25
- C．估计七年级 700 名学生参加社会实践活动时间少于10h 的有 112 名学生
- D．随机抽取的学生中参加社会实践活动时间不少于10h 的有 32 名的学生

【变式 1】．为了了解全区近 4800 名初三学生数学学习状况，从中随机抽取 500 名学生的测试成绩作为样本，将他们的成绩整理后分组情况如下：（ 每组数据可含最低值，不含最高值 ）

分组 (分)	40 ~ 50	50 ~ 60	60 ~ 70	70 ~ 80	80 ~ 90	90 ~ 100
频数	12	18	160			
频率					0.18	0.04

根据上表信息，由此样本请你估计全区此次成绩在 70~80 分的人数大约是_____．

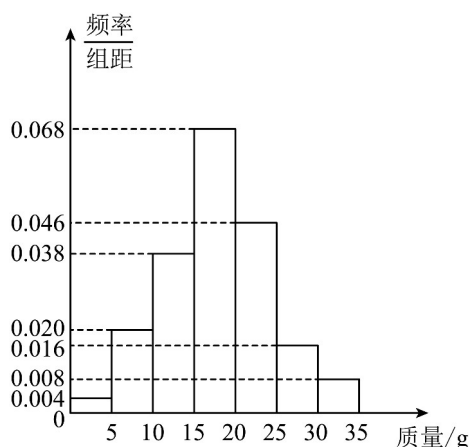
【变式 2】．为了解学生假期每天帮忙家长做家务活动情况，学校团委随机抽取了部分学生进行线上调查，并将调查结果绘制成频数直方图（ 不完整， 每组含最小值， 不含最大值 ），并且知道 80 ~ 100 分钟占所抽查学生的 17.5%，根据提供信息，以下说法不正确的是（ ）



- A . 本次共随机抽取了 40 名学生 ;
- B . 抽取学生中每天做家务时间的中位数落在 40 ~ 60 分钟这一组 ;
- C . 如果全校有 800 名学生 , 那么每天做家务时间超过 1 小时的大约有 300 人 ;
- D . 扇形统计图中 0 ~ 20 分钟这一组的扇形圆心角的度数是 30° ;

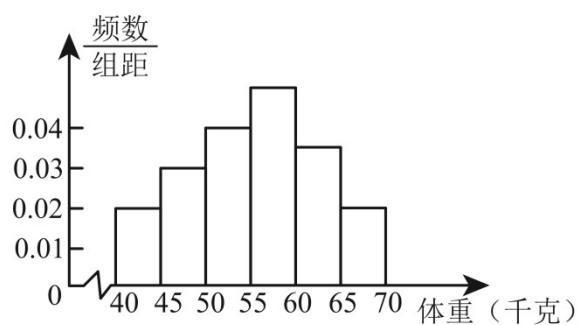
题型 07 频率分布直方图

【典例 1】 . 如图 , 上海某有机草莓农场为了解今年草莓的收成情况 , 随机选择了一个大棚摘取草莓并逐一称重 (精确到 1g) , 绘制出频率分布直方图 (每组数据含最低值 , 不含最高值) . 如果质量不小于 20g 的草莓为“大果” , 则可估计 500kg 草莓中“大果”的总质量是 ()

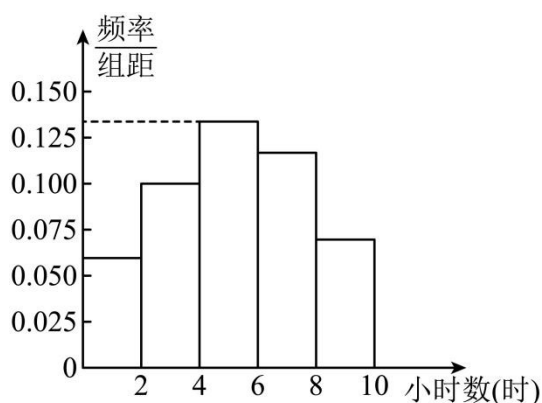


- A . 35kg B . 170kg C . 175kg D . 380kg

【变式 1】 . 为了解全区 5000 名初中毕业生的体重情况 , 随机抽测了 200 名学生的体重 , 频率分布如图所示 (每小组数据可含最小值 , 不含最大值) , 其中从左至右前六个小长方形的高依次为 0.02、0.03、0.04、0.05、0.035、0.025 , 由此可估计全区初中毕业生的体重在 50 到 55 千克的学生人数约为_____人 .



【变式2】 . 某校为了解本校学生每周阅读课外书籍的时间，对本校全体学生进行了调查，并绘制如图所示的频率分布直方图，那么图中这五个小矩形的面积之和为_____ .



强化训练

一、单选题

- 下列调查中，哪一项适合用普查()
 - 夏季冷饮市场上的冰淇淋的质量
 - 对学校设立读报角的看法
 - 人们环境保护的意识
 - 调查青年人对音乐的喜爱情况
- 为了了解某地区初一年级 4500 名学生的体重情况，从中抽取了 500 名学生的体重，就这个问题来说，下面说法中正确的是()
 - 样本容量是 500
 - 每个学生是个体
 - 500 名学生是所抽取的一个样本
 - 4500 名学生是总体
- 下列调查中，选取的样本最具有代表性的是()
 - 调查某校 3000 名学生的体检情况，选取该校初二年级的学生进行调查
 - 调查某校学生每周课余体育锻炼时间，选取该校体育社团中的 50 名同学进行调查
 - 为了解某社区老年人的健康状况，在该社区随机对 10 名正在健身的老人进行调查

D. 为了解某公司1200名员工的每日睡眠时长，随机选取该公司120位员工进行调查

4. 一组数据：3，4，4，6，8. 这组数据的平均数是 ()

- A. 3 B. 4 C. 5 D. 6

5. 某校为了了解学生在校午餐所需的时间，抽取了20名学生在校午餐所需时间，获得如下的数据（单位：分）：12、12、15、11、16、18、19、18、20、18、18、20、28、22、30、20、15、16、21、16. 若将这些数据以4分为组距进行分组，则组数是 ()

- A. 4组 B. 5组 C. 6组 D. 7组

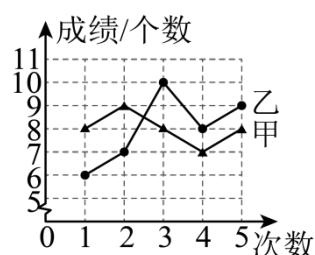
6. 某农民在池塘里养了许多鱼，有草鱼、鲢鱼、鲤鱼、鲫鱼等，为了能更清楚地表示出各种鱼的条数，最适合使用的统计图是 ()

- A. 扇形统计图 B. 折线统计图 C. 条形统计图 D. 以上都可以

7. 有一组数据分别为：4，1，7， x ，5；已知这组数的平均数为4，那么这组数据的中位数是 () .

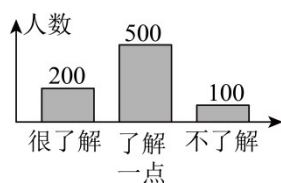
- A. 7 B. 5 C. 4 D. x

8. 如图是甲、乙两名同学的5次篮球训练中练习投篮成绩的折线统计图，下列判断正确的是 ()



- A. 甲的成绩的中位数比乙的成绩的中位数大 B. 甲的成绩的众数是9个
C. 甲的成绩的平均数比乙的成绩的平均数大 D. 甲的成绩比乙的成绩稳定

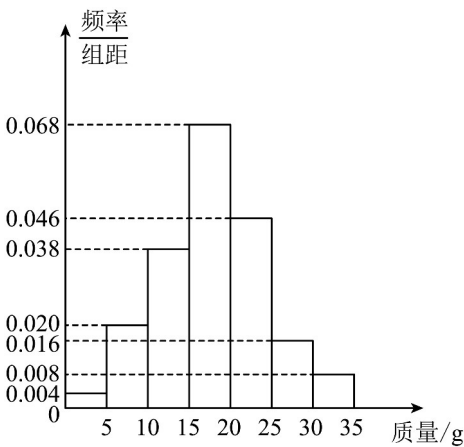
9. 为了调查全校师生对人工智能Deepseek的熟悉程度，某数学小组对全校2000名师生发放了问卷，随机回收了800份，将回收问卷的调查结果绘制成统计图如图，由此估计全校师生对人工智能Deepseek“不了解”的约有 ()



- A. 500人 B. 750人 C. 250人 D. 1200人

10. 如图，上海某有机草莓农场为了了解今年草莓的收成情况，随机选择了一个大棚摘取草莓并逐一称重

(精确到 1g)，绘制出频率分布直方图(每组数据含最低值，不含最高值)。如果质量不小于 20g 的草莓为“大果”，则可估计 500kg 草莓中“大果”的总质量是 ()



- A . 35kg B . 170kg C . 175kg D . 380kg

二、填空题

11. “神舟十八号”载人飞船于 2024 年 4 月 25 日在酒泉卫星发射中心发射，要想调查飞船零件的质量，适合采用_____ (填“普查”或“抽样调查”) .

12. 一组数据 1，3，5， x 的平均数与中位数相同，则 x 的值是_____ .

13. 一个射箭运动员连续射靶 5 次，所得环数分别是：8，6，10，7，9，则这个运动员所得环数的标准差为_____ .

14. 某校为迎接“五四”青年节，举办了校园歌曲比赛. 每名选手最后得分为去掉一个最高分和一个最低分后的平均分. 已知七位评委给小萌的分数分别为：94，98，97，92，95，96，93，则小萌同学最后的得分为_____分 .

15. 小明在计算一组数据的方差时，列出的算式如下： $S^2 = \frac{1}{6} \left[2(7-\bar{x})^2 + 3(8-\bar{x})^2 + (9-\bar{x})^2 \right]$ ，根据算式信息，这组数据的平均数是_____ .

16. 已知某班学生理化实验操作测试成绩的统计结果如下表：

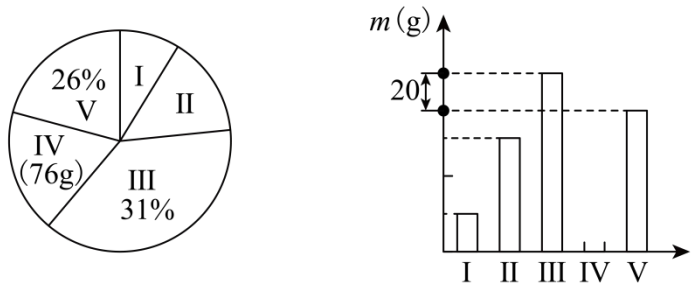
成绩 (分)	4	5	6	7	8	9	10
人数	1	2	2	6	9	11	9

则这些学生成绩的众数是_____分 .

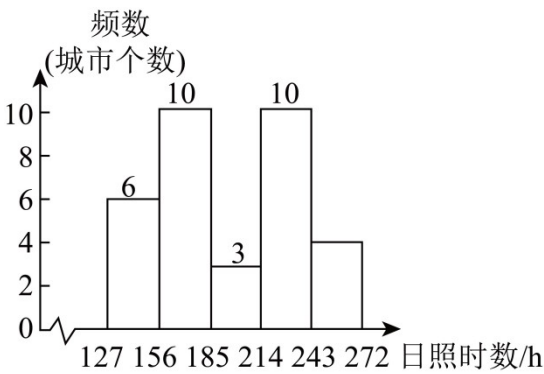
17．五种不发生反应的化合物I、II、III、IV、V 在一个密封的容器中，经过物质检验，得到如下两张图．

如果条形图中每个横线刻度间的距离相等，那么化合物II的质量是_____g．

化学物质的质量和质量占全体比重 化学物质质量的关系



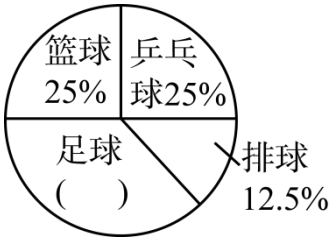
18．根据 33 个全国主要城市 2023 年 7 月的日照时数（单位：h），绘制了不完整的频数分布直方图如图所示（数据分成 5 组： $127 \leq x < 156$ ， $156 \leq x < 185$ ， $185 \leq x < 214$ ， $214 \leq x < 243$ ， $243 \leq x < 272$ ）．下面三个结论：①日照时数在 $185 \leq x < 214$ 范围的城市数量最少；②有 4 个城市日照时数在 243h 至 272h（不含 272h）的范围；③ 2023 年 7 月，北京的日照时数是 232.8h，比这 33 个全国主要城市中一半以上城市的日照时数都长．所有正确的结论的序号是_____．



三、解答题

19．某校开展阳光体育运动，调查了七年级学生喜欢的球类活动（每人只选一项自己最喜欢的球类项目），并将调查情况制成如下统计表和统计图（不完整）．请将统计表和统计图补充完整．

球类项目	乒乓球	篮球	足球	排球
人数	30 人	_____人	_____人	_____人



20．七年级(5)班 20 名女生的身高如下(单位: cm):

153 156 152 158 156 160 163 145 152 153
162 153 165 150 157 153 158 157 158 158

(1)请你在下表中填出身高在以下各个范围的频数，百分比(每个范围包含下限，但不包含上限)：

身高(cm)	140 ~ 150	150 ~ 160	160 ~ 170
频数			
百分比			

(2)上表把身高分成__组，组距是__；

(3)身高在__范围的人数最多．

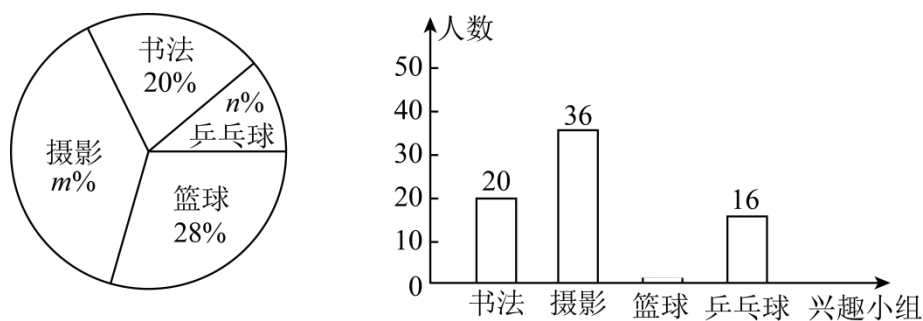
21．甲、乙两人在 5 次打靶测试中命中的环数如下：甲：8，8，7，8，9；乙：5，9，7，10，9．

选 手	平均数	众 数	中位数	方差
甲	8	8	8	c
乙	a	9	b	3.2

(1)表格中 $a =$ _____， $b =$ _____， $c =$ _____；

(2)教练根据这 5 次成绩，选择甲参加射击比赛，教练的理由是什么？

22．某校计划组织学生参加学校书法、摄影、篮球、乒乓球四个课外兴趣小组，要求每人必须参加并且只能选择其中的一个小组，为了了解学生对四个课外小组的选择情况，学校从全体学生中随机抽取部分学生进行问卷调查，并把调查结果制成如下所示的两幅不完整的统计图



请根据给出的信息解答下列问题：

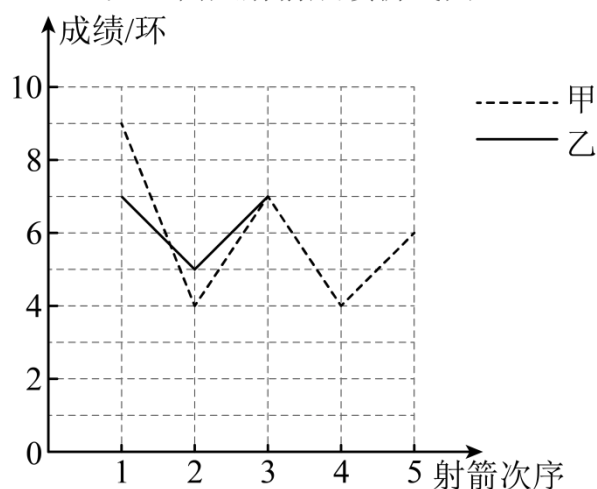
(1)求该校参加这次问卷调查的学生人数，并补全条形统计图；

(2) $m = \underline{\quad}$ ， $n = \underline{\quad}$ ；

(3)若该校共有学生 4000 人，试估计该校选择“乒乓球”课外兴趣小组的学生有多少人？

23．某社区准备在甲乙两位射箭爱好者中选出一人参加集训，两人各射了 5 箭，他们的总成绩（单位：环）相同，小宇根据他们的成绩绘制了尚不完整的统计图表，并计算了甲成绩的平均数和方差（见小宇的作业）。

甲、乙两人射箭成绩折线图



小宇的作业：

$$\text{解：}\overline{X_{\text{甲}}} = \frac{1}{5}(9+4+7+4+6) = 6$$

$$\begin{aligned} S_{\text{甲}}^2 &= \frac{1}{5}[(9-6)^2 + (4-6)^2 + (7-6)^2 + (4-6)^2 + (6-6)^2] \\ &= \frac{1}{5}(9+4+1+4+0) \\ &= 3.6 \end{aligned}$$

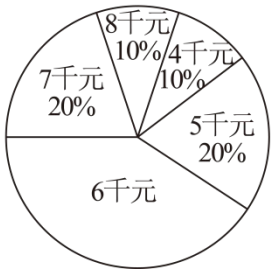
甲、乙两人射箭成绩统计表

	第 1 次	第 2 次	第 3 次	第 4 次	第 5 次
甲成 绩	9	4	7	4	6
乙成 绩	7	5	7	a	7

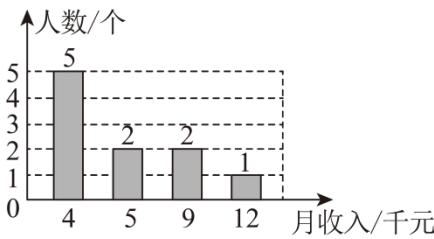
- (1) $a = \underline{\hspace{2cm}}$, $\overline{x_Z} = \underline{\hspace{2cm}}$;
- (2) 请完成图中表示乙成绩变化情况的折线 ;
- (3) ① 观察图 , 可看出 的成绩比较稳定 (填“甲”或“乙”) .
- ② 请你从平均数和方差的角度分析 , 谁将被选中 .

24 . 近年来网约车给人们的出行带来了便利 , 小明和数学兴趣小组的同学对甲、乙两家网约车公司司机月收入进行了抽样调查 , 两家公司分别抽取的 10 名司机月收入 (单位 : 千元) 如图所示 :

甲公司司机月收入人数分布
扇形统计图



乙公司司机月收入人数分布
条形统计图



根据以上信息 , 整理分析数据如下 :

	平均月收入/千元	中位数	众数	方差
甲公司	6	6	c	1.2
乙公司	a	b	4	

- (1) 填空 : $b = \underline{\hspace{2cm}}$, $c = \underline{\hspace{2cm}}$;
- (2) 求出乙公司的平均月收入 a 以及方差 d ; ($s^2 = \frac{1}{n} \left[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2 \right]$)
- (3) 小明的叔叔计划从两家公司中选择一家做网约车司机 , 如果你是小明 , 你建议他选择哪家公司 ? 请说明

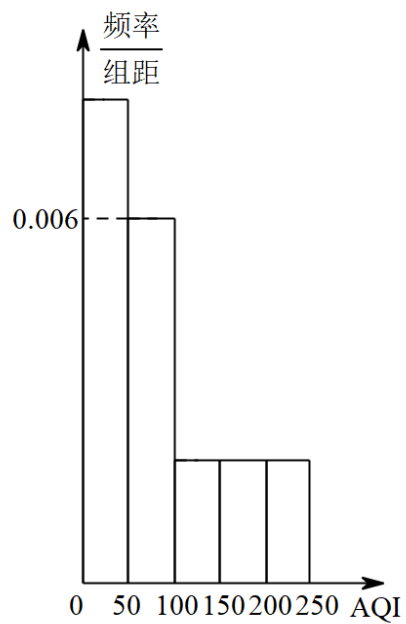
理由 .

25 . 空气质量指数 (*Air Quality Index* , 缩写 *AQI*) 是定量描述空气状况的非线性无量纲指数 . 其数值越大、级别和类别越高 , 说明空气污染状况越严重 , 对人体的健康危害也就越大 , 适用于表示某地区的短期空气质量状况和变化趋势 . (空气污染指数为 0~50 是优 ; 空气污染指数为 50~100 是良好 ; 空气污染指数为 100~150 是轻度污染 ; 空气污染指数为 150~200 是中度污染 ; 空气污染指数为 200~250 是重度污染 .)

如图表示的是某地区 2022 年 11 月份 30 天日均 *AQI* 指数的频率分布直方图 .

空气质量指数 (<i>AQI</i>)	0~50	50~100	100~150	150~200	200~250
天数	a	b	3	3	3
频率	c	d	0.1	0.1	0.1

(注 : 每组数据可含最高值 , 不含最低值)



(1) 请你根据上述频率分布直方图及表格完成下面的填空 :

这个地区 11 月份空气为轻度污染的天数是_____天 . $a =$ _____ ; $b =$ _____ ; $c =$ _____ ; $d =$ _____ .

(2) 为了进一步改善生活环境和空气质量 , 提高人民的生活质量 , 当地政府计划从 2023 年开始增加绿化面积 . 已知 2022 年底该地区的绿化面积为 20 万亩 , 如果到 2024 年底 , 该地区的绿化面积比为 2022 年的绿化面积增加了 50% , 假设这两年绿化面积的年增长率相同 , 求这两年中绿化面积每年的增长率 (精确到

0.01) (参考数据： $\sqrt{2} \approx 1.414$, $\sqrt{3} \approx 1.732$, $\sqrt{5} \approx 2.236$, $\sqrt{6} \approx 2.449$)