

统计-习题

一、选择题

1 下面的抽样方法是简单随机抽样的是 () .

- A. 在某年明信片销售活动中, 规定每100万张为一个开奖组, 通过随机抽取的方式确定号码的后四位为2709的为三等奖
- B. 某车间包装一种产品, 在自动包装的传送带上, 每隔30分钟抽一包产品, 称其质量是否合格
- C. 某学校分别从行政人员、教师、后勤人员中抽取2人、14人、4人了解学校机构改革的意义
- D. 用抽签法从10件产品中选取3件进行质量检验

答案 D

解析 A、B不是简单随机抽样, 因为抽取的个体间的间隔是固定的;
C不是简单随机抽样, 因为总体的个体有明显的层次; D是简单随机抽样.

2 从1002个编号中抽取50个号码, 若采用系统抽样法抽样, 则抽样间隔为 () .

- A. 19
- B. 20
- C. 21
- D. 以上都不对

答案 B

解析 剔除两个元素后, 组距为 $\frac{1000}{50} = 20$.

3 某校做了一次关于“感恩父母”的问卷调查, 从8~10岁, 11~12岁, 13~14岁, 15~16岁四个年龄段回收的问卷依次为: 120份, 180份, 240份, x 份因调查需要, 从回收的问卷中按年龄段分层抽取容量为300的样本, 其中在11~12岁学生问卷中抽取60份, 则在15~16岁学生中抽取的问卷份数为 () .

- A. 60
- B. 80
- C. 120
- D. 180

答案 C

解析 11 ~ 12岁回收的问卷为180份，其中在11 ~ 12岁学生问卷中抽取60份，则抽样比为 $\frac{1}{3}$ 。
 $\therefore$ 从回收的问卷中按年龄段分层抽取容量为300的样本，
 $\therefore$ 从8 ~ 10岁，11 ~ 12岁，13 ~ 14岁，15 ~ 16岁四个年龄段回收的问卷总数为 $300 \div \frac{1}{3} = 900$ ，
 $\therefore$ 15 ~ 16岁回收问卷份数为 $x = 900 - 120 - 180 - 240 = 360$ ，
 $\therefore$ 在15 ~ 16岁学生中抽取的问卷份数为 $360 \times \frac{1}{3} = 120$ 。
 故选C。

- 4 为了检查某超市货架上的饮料是否含有塑化剂，要从编号依次为1到50的塑料瓶装饮料中抽取5瓶进行检验，用每部分选取的号码间隔一样的系统抽样法确定所选取的5瓶饮料的编号可能是（ ）。
- A. 5, 10, 15, 20, 25 B. 2, 4, 6, 8, 10
 C. 1, 2, 3, 4, 5 D. 7, 17, 27, 37, 47

答案 D

解析 由系统抽样法性质可知，需将50瓶饮料分为5组，
 $\therefore$ 每组有 $\frac{50}{5} = 10$ 瓶，
 故选取出的5瓶饮料编号间隔为10，故D符合。
 故选D。

- 5 将参加夏令营的600名学生编号为001, 002, ..., 600。现采用系统抽样方法抽取一个容量为50的样本，且随机抽得的号码为003。这600名学生分住在三个营区，从001到300为第Ⅰ营区，从301到495为第Ⅱ营区，从496到600为第Ⅲ营区。三个营区被抽中的人数依次为（ ）。
- A. 25, 17, 8 B. 25, 16, 9 C. 26, 16, 8 D. 24, 17, 9

答案 A

解析 因为总体容量为600，样本容量是50，所以 $600 \div 50 = 12$ 。

因此，每隔12个号能抽到一名。由于随机抽得的第一个号码为003，

因此按照系统抽样的操作步骤，在第Ⅰ营区应抽取25人，

第Ⅱ营区应抽取17人，第Ⅲ营区应抽取8人，故选A。

6 某校有住宿的男生400人，住宿的女生600人，为了解住宿生每天运动时间，通过分层随机抽样的方法抽到100名学生，其中男生、女生每天运动时间的平均值分别为100分钟、80分钟。结合此数据，请你估计该校全体住宿学生每天运动时间的平均值为（ ）。

A. 98分钟

B. 90分钟

C. 88分钟

D. 85分钟

答案 C

解析 $\frac{400}{400 + 600} = 0.4$ ，

$\frac{600}{400 + 600} = 0.6$ ，

$\therefore \bar{x} = 0.4 \times 100 + 0.6 \times 80$

$= 40 + 48 = 88$ 。

故选C。

7 以下茎叶图记录了甲、乙两组各五名学生在一次英语听力测试中的成绩（单位：分）。已知甲组数据的中位数为15，乙组数据的平均数为16.8，则 x, y 的值分别为（ ）。

甲组			乙组	
9		0	9	
x	2	1	5	y 8
7	4	2	4	

A. 2, 5

B. 5, 5

C. 5, 8

D. 8, 8

答案 C

解析 乙组数据平均数 $= (9 + 15 + 18 + 24 + 10 + y) \div 5 = 16.8$;

$$\therefore y = 8 ;$$

甲组数据可排列成 : $9, 12, 10 + x, 24, 27$. 所以中位数为 : $10 + x = 15$,

$$\therefore x = 5 .$$

故选C .

8 从某中学抽取10名同学，他们的数学成绩如下：82，85，88，90，92，92，92，96，96，98（单位：分），则这10名同学数学成绩的众数、第25百分位数分别为（ ） .

A. 92，85

B. 92，88

C. 95，88

D. 96，85

答案 B

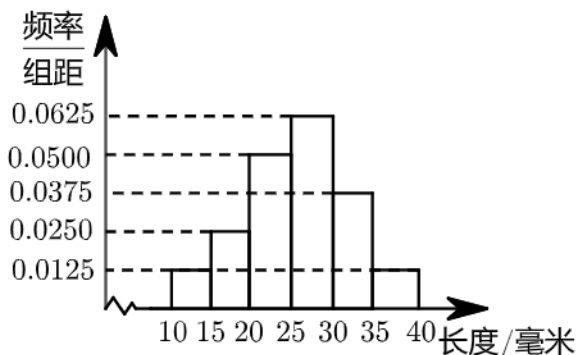
解析 将数据从小到大排列82，85，88，90，92，92，92，96，96，98，

众数为92， $10 \times 0.25 = 2.5$ ，

$\therefore$ 第25百分位数为88

故选B .

9 对一批产品的长度（单位：毫米）进行抽样检测，样本容量为200，如图为检测结果的频率分布直方图，根据产品标准，单件产品长度在区间[25, 30)的为 一等品，在区间[20, 25)和[30, 35)的为二等品，其余均为三等品，则样本中三等品的件数为（ ） .



A. 30

B. 40

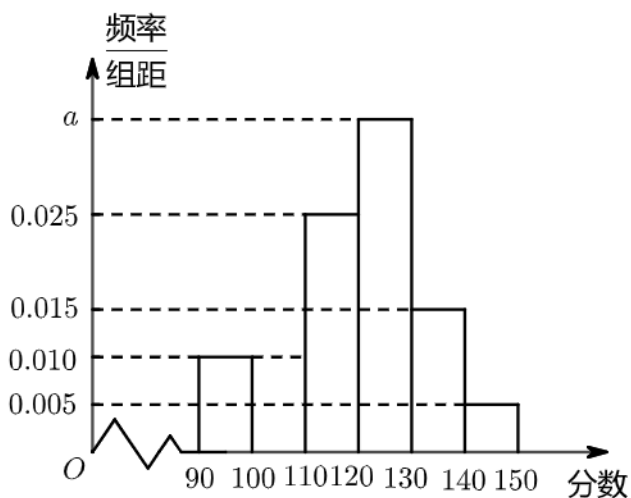
C. 50

D. 60

答案 C

解析 $(0.0125 + 0.025 + 0.0125) \times 5 \times 200$
 $= 0.05 \times 5 \times 200$
 $= 50 .$

- 10 某校高三年级共有800名学生参加了数学测验，将所有学生的数学成绩分组如下： $[90, 100)$ ， $[100, 110)$ ， $[110, 120)$ ， $[120, 130)$ ， $[130, 140)$ ， $[140, 150]$ ，得到的频率分布直方图如图所示，则下列说法中正确的是（ ）。



- ①成绩不低于120分的学生人数为440；
 ②这800名学生中数学成绩的众数为125；
 ③这800名学生数学成绩的中位数的近似值为121.4；
 ④这800名学生数学成绩的平均数为120。

A. 1个 B. 2个 C. 3个 D. 4个

答案 D

解析 $(0.005 + 0.01 + 0.01 + 0.015 + 0.025 + a) \times 10 = 1$,
 $\therefore 0.065 + a = 0.1$,
 $\therefore a = 0.035$,
 ① $(0.035 + 0.015 + 0.005) \times 10 = 0.055 \times 10 = 0.55$,
 $800 \times 0.55 = 440$, 故①正确；

②众数为 $\frac{120 + 130}{2} = 125$ ，故②正确；

③ $(0.01 + 0.01 + 0.025) \times 10 = 0.45$ ，

$0.035 \times 10 = 0.35$ ，

$\therefore$ 中位数为 $120 + \frac{0.5 - 0.45}{0.35} \times 10 \approx 121.4$ ，故③正确；

④ $0.01 \times 10 \times 95 + 0.01 \times 10 \times 105 + 0.025 \times 10 \times 115 +$

$0.035 \times 10 \times 125 + 0.015 \times 10 \times 135 + 0.005 \times 10 \times 145 = 120$ ，故④正确。

故选D。

11 对于数据3, 3, 2, 3, 6, 3, 10, 3, 6, 3, 2, 有以下结论：

- ①这组数据的众数是3，
- ②这组数据的众数与中位数的数值不等，
- ③这组数据的中位数与平均数的数值相等，
- ④这组数据的平均数与众数的数值相等

其中正确的结论有（ ）。

- A. 1个 B. 2个 C. 3个 D. 4个

答案 A

解析 将题中数据由小到大排列得：

2, 2, 3, 3, 3, 3, 3, 3, 6, 6, 10,

易知众数为3，

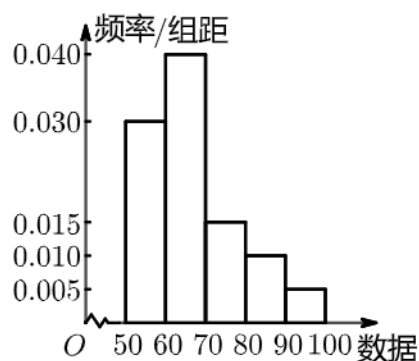
中位数为3，

平均数为 $\frac{2 \times 2 + 3 \times 6 + 6 \times 2 + 10}{11} = \frac{4 + 18 + 12 + 10}{11} = 4$ ，

$\therefore$ 只有①正确。

故选A。

12 已知一组数据的频率分布直方图如图所示，则众数、中位数、平均数分别为（ ）。



A. 63, 64, 66

B. 65, 65, 67

C. 55, 64, 66

D. 64, 65, 64

答案 B

解析 众数： $\frac{60 + 70}{2} = 65$ ，

平均数： $55 \times 0.3 + 65 \times 0.4 + 75 \times 0.15 + 85 \times 0.1 + 95 \times 0.05 = 67$ ，

设中位数为 x ，

$$(x - 60) \times 0.04 + 0.03 \times 10 = 0.5$$

$$x = 65.$$

故选B.

13 某工厂技术人员对三台智能机床生产数据统计后发现，甲车床每天生产次品数的平均数为1.5，标准差为1.28；乙车床每天生产次品数的平均数为1.2，标准差为0.87；丙车床每天生产次品数的平均数为1.2，标准差为1.28. 由此数据可以判断生产性能最好且较稳定的为（ ）.

A. 无法判断

B. 甲车床

C. 乙车床

D. 丙车床

答案 C

解析 $\frac{\overline{X}_{\text{甲}}}{S_{\text{甲}}} = \frac{1.5}{1.28} = 1.17$ ，

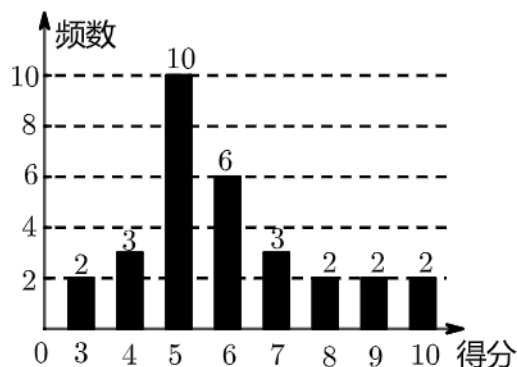
$$\frac{\overline{X}_{\text{乙}}}{S_{\text{乙}}} = \frac{1.2}{0.87} = 1.38$$
，

$$\frac{\overline{X}_{\text{丙}}}{S_{\text{丙}}} = \frac{1.2}{1.28} = 0.94$$
，

$$\frac{\overline{X}_{\text{乙}}}{S_{\text{乙}}} \text{ 最大.}$$

故选C.

- 14 为了普及环保知识，增强环保意识，某大学随机抽取30名学生参加环保知识测试，得分（十分制）如图所示，假设得分值的中位数为 m_e ，众数为 m_0 ，平均值为 $\bar{x}$ ，则（ ）。



- A. $m_e = m_0 = \bar{x}$ B. $m_e = m_0 < \bar{x}$ C. $m_e < m_0 < \bar{x}$ D. $m_0 < m_e < \bar{x}$

答案 D

解析 由题图可知，名学生的得分情况如下：2个人得3分，
 个人得4分，10个人得5分，6个人得6分，3个人得7分，2个人得8分，
 个人得9分，2个人得10分。中位数为第15，
 16个数（分别是5，6）的平均数，即 $m_e = 5.5$ ，5出现的次数最多，所以 $m_0 = 5$ 。

$$\bar{x} = \frac{2 \times 3 + 3 \times 4 + 10 \times 5 + 6 \times 6 + 3 \times 7 + 2 \times 8 + 2 \times 9 + 2 \times 10}{30} \approx 5.97$$
。于是得
 $m_0 < m_e < \bar{x}$ 。

- 15 在“世界杯”足球赛闭幕后，某中学学生会对本校高一年级1000名学生收看比赛的情况用随机抽样方式进行调查，样本容量为50，将数据分组整理后，列表如下：

现场观看	0	1	2	3	4	5	6	7
观看人数占调查人数的百分比	8%	10%	20%	26%	$m\%$	12%	6%	2%

从表中可以得出正确的结论为（ ）。

- A. 表中 m 的数值为8 B. 估计观看比赛不低于4场的学生约为360人
 C. 估计观看比赛不低于4场的学生约为720人 D. 估计观看比赛场数的众数为2

答案 B

解析 $8\% + 10\% + 20\% + 26\% + m\% + 12\% + 6\% + 2\% = 1$,

$$\therefore m = 16,$$

$\therefore$ A错,

不低于4场学生占比为 $16\% + 12\% + 6\% + 2\% = 36\%$, $1000 \times 36\% = 360$,

$\therefore$ B对, C错,

众数为3, $\therefore$ D错,

故选B.

二、填空题

- 16 我国古代数学名著《数书九章》有“米谷粒分”题：粮仓开仓收粮，有人送来米1534石，验得米内夹谷，抽样取米一把，数得254粒内夹谷28粒，则这批米内夹谷约为 _____（石）.

答案 169

解析 设这批米内夹谷量为 x 石，

$$\frac{x}{1534} = \frac{28}{254},$$

$$\therefore x \approx 169.$$

- 17 采用随机数表法从编号为01, 02, 03, …, 30的30个个体中选取7个个体，指定从下面随机数表的第一行第5列开始，由左向右选取两个数字作为应取个体的号码，则选取的第6个个体号码是 _____；

03 47 43 86 36 16 47 80 45 69 11 14 16 95 36 61 46 98 63 71 62 33 26 36 77

97 74 24 67 62 42 81 14 57 20 42 53 32 37 32 27 07 36 07 52 24 52 79 89 73

答案 20

解析 原则是数字应在01 – 30中且与之前数字无重复，
 $\therefore$ 第1个个体为16，
 第2个个体为11，
 第3个个体为14，
 第4个个体为26，
 第5个个体为24，
 第6个个体为20。
 故为20。

18 一次游戏中，获奖者可得到5件不同的奖品，这些奖品要从已编号为1 ~ 50的50种不同商品中随机抽取确定，那么分段间隔为 _____，可分为 _____ 段，各段编号依次为 _____。

答案 1:10
 2:5
 3:1 ~ 10，11 ~ 20，21 ~ 30，31 ~ 40，41 ~ 50

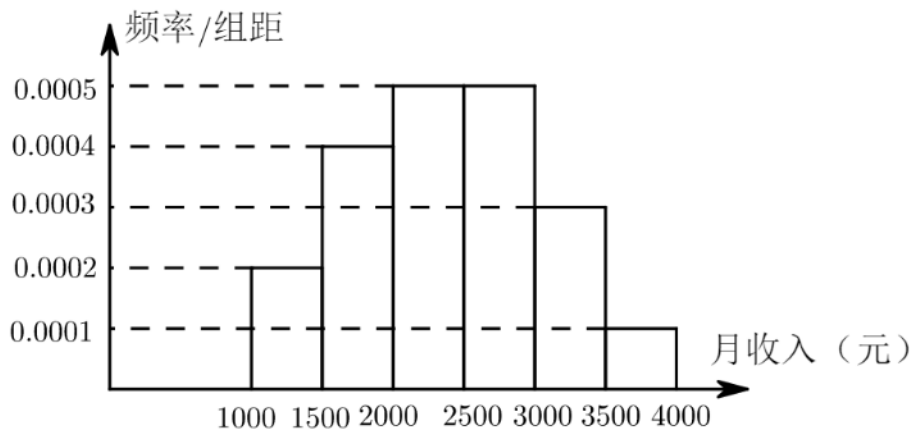
解析 结合系统抽样的结论易知。

19 某单位有职工750人，其中青年职工350人，中年职工250人，老年职工150人，为了了解该单位职工的健康情况，用分层抽样的方法从中抽取样本。若样本中的青年职工为21人，则样本容量为 _____。

答案 45

解析 根据分层抽样的定义和方法，每个个体被抽到的概率等于 $\frac{21}{350} = \frac{3}{50}$ 。
 设样本容量等于 n ，则有 $\frac{n}{750} = \frac{3}{50}$ ，解得 $n = 45$ 。

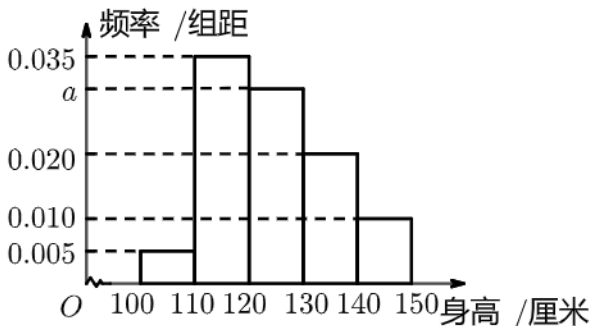
一个社会调查机构就某地居民的月收入调查了10000人，并根据所得数据画了样本的频率分布直方图（如图，每组数据含最小值，不含最大值）. 为了分析居民的收入与年龄、学历、职业等方面的关系，要从这10000人中再用分层抽样方法抽出100人进一步调查，则在 $[2500, 3000)$ （单位：元/月）收入段应抽出 _____ 人 .



答案 25

解析 收入在 $[2500, 3000)$ （单位：元/月）收入段的人数占总人数的频率为
 $(3000 - 2500) \times 0.0005 = 0.25$ ，故应抽出 $100 \times 0.25 = 25$ （人）.

21 从某小学随机抽取100名同学，将他们的身高（单位：厘米）数据绘制成频率分布直方图（如图），由图中数据可知 $a = \underline{\hspace{2cm}}$. 若要从身高在 $[120, 130)$ ， $[130, 140)$ ， $[140, 150]$ 三组内的学生中，用分层抽样的方法选取18人参加一项活动，则从身高在 $[140, 150]$ 内的学生中选取的人数应为 _____ .



答案 1:0.030

2:3

解析

由于频率分布直方图中各个小矩形的面积即为相应区间的频率，

且其频率之和为1，故有 $10a = 1 - (0.005 + 0.010 + 0.020 + 0.035) \times 10 = 0.300$ 故 $a = 0.030$

又由于身高为 $[120, 130)$, $[130, 140)$, $[140, 150]$ 的频率分别为0.3, 0.2, 0.1，频率之比为3 : 2 : 1，

故若样本容量为18，故身高在 $[140, 150]$ 内的频数为 $18 \times \frac{1}{1+2+3} = 3$ 人。

三、解答题

22 从某珍珠公司生产的产品中，任意抽取12颗珍珠，得到它们的质量（单位：g）如下：

7.9, 9.0, 8.9, 8.6, 8.4, 8.5, 8.5, 8.5, 9.9, 7.8, 8.3, 8.0.

(1) 分别求出这组数据的第25, 75, 95百分位数。

(2) 若用第25, 50, 95百分位数把公司生产的珍珠划分为次品、合格品、优等品和特优品，依照这个样本的数据，给出该公司珍珠等级的划分标准。

答案

(1) 8.15 ; 8.75 ; 9.9 .

(2) $[0, 8.15)$ 为次品，
 $[8.15, 8.75)$ 为合格品，
 $[8.75, 9.9)$ 为优等品，
 $[9.9, +\infty)$ 为特优品。

解析

(1) 将12个数据从小到大排列7.8、7.9、8.0、8.3、8.4、8.5、8.5、8.5、8.6、8.9、9.0、9.9

$$12 \times 0.25 = 3, \frac{8.0 + 8.3}{2} = 8.15,$$

$$12 \times 0.75 = 9, \frac{8.6 + 8.9}{2} = 8.75,$$

$$12 \times 0.95 = 11.4, \therefore 9.9,$$

$\therefore$ 第25百分位数为8.15，

第75百分位数为8.75，

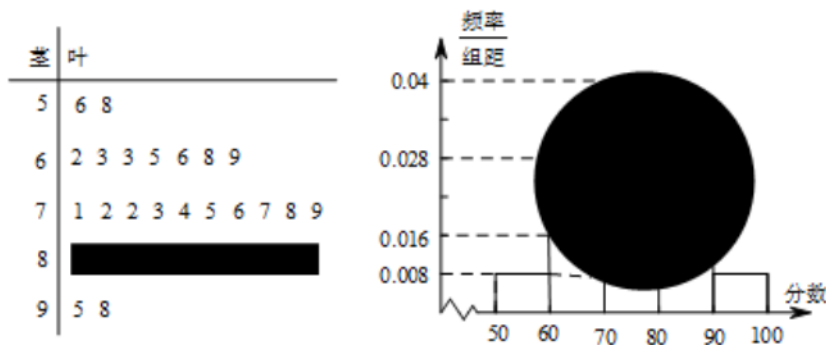
第95百分位数为9.9。

(2) $[0, 8.15)$ 为次品，
 $[8.15, 8.75)$ 为合格品，

$[8.75, 9.9)$ 为优等品，

$[9.9, +\infty)$ 为特优品。

- 23 某校高三(1)班的一次数学测试成绩的茎叶图和频率分布直方图都受到不同程度的破坏，但可见部分如下，据此解答如下问题。



- 求全班人数及分数在 $[80, 90)$ 之间的频数；
- 估计该班的平均分数，并计算频率分布直方图中 $[80, 90)$ 间的矩形的高。

答案

(1) 4.

(2) 0.016.

解析

- 由茎叶图知，分数在 $[50, 60)$ 之间的频数为2，频率为 $0.008 \times 10 = 0.08$ ，
全班人数为 $\frac{2}{0.08} = 25$ 。所以分数在 $[80, 90)$ 之间的频数为 $25 - 2 - 7 - 10 - 2 = 4$ 。
- 分数在 $[50, 60)$ 之间的总分为 $56 + 58 = 114$ ；
分数在 $[60, 70)$ 之间的总分为 $60 \times 7 + 2 + 3 + 3 + 5 + 6 + 8 + 9 = 456$ ；
分数在 $[70, 80)$ 之间的总分数为 $70 \times 10 + 1 + 2 + 3 + 3 + 4 + 5 + 6 + 7 + 8 + 9 = 747$ ；
分数在 $[80, 90)$ 之间的总分数约为 $85 \times 4 = 340$ ；分数在 $[90, 100]$ 之间的总分数为
 $95 + 98 = 193$ ；
所以，该班的平均分数为 $\frac{114 + 456 + 747 + 340 + 193}{25} = 74$ 。
估计平均分时，以下解法也算对
分数在 $[50, 60)$ 之间的频率为 $\frac{2}{25} = 0.08$ ；
分数在 $[60, 70)$ 之间的频率为 $\frac{7}{25} = 0.28$ ；
分数在 $[70, 80)$ 之间的频率为 $\frac{10}{25} = 0.40$ ；
分数在 $[80, 90)$ 之间的频率为 $\frac{4}{25} = 0.16$ ；

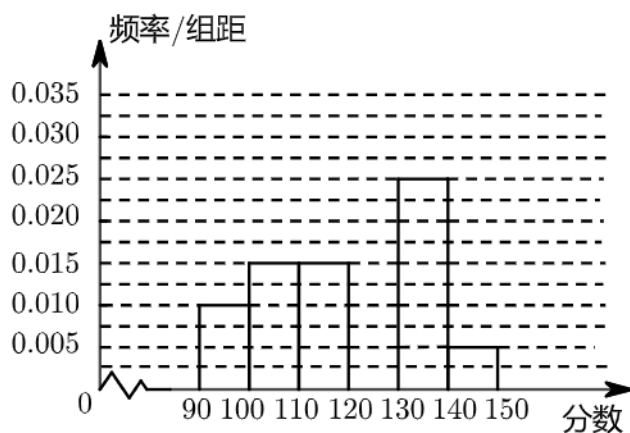
分数在 $[90, 100]$ 之间的频率为 $\frac{2}{25} = 0.08$;

所以, 该班的平均分约为

$$55 \times 0.08 + 65 \times 0.28 + 75 \times 0.40 + 85 \times 0.16 + 95 \times 0.08 = 73.8$$

频率分布直方图中 $[80, 90)$ 间的矩形的高为 $\frac{4}{25} \div 10 = 0.016$.

- 24 某校从参加高二期中考试的学生中随机抽取60名学生, 将其数学成绩(均为整数)分成六组 $[90, 100)$, $[100, 110)$, $\dots$, $[140, 150]$ 后得到如下部分频率分布直方图, 观察图形的信息, 回答下列问题:



- (1) 求这60名学生中分数在 $[120, 130)$ 内的人数.
- (2) 估计本次考试的中位数和平均分.

答案

- (1) 18.
- (2) 123.34; 121.

解析

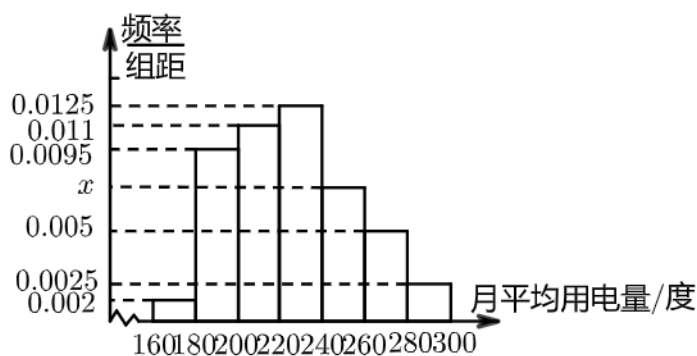
(1) 分数在 $[120, 130)$ 内的数率为 $1 - (0.1 + 0.15 + 0.15 + 0.25 + 0.05) = 0.3$, 则人数
 $= 60 \times 0.3 = 18$ 人.

(2) (2) 由(1)易知中位数在 $[120, 130)$ 内设中位数为 a , 则

$$0.1 + 0.15 + 0.15 + 0.03(a - 120) = 0.5, \text{ 则 } a = 123.34,$$

$$\bar{x} = 95 \times 0.1 + 105 \times 0.15 + 115 \times 0.15 + 125 \times 0.3 + 135 \times 0.25 + 145 \times 0.05 = 121.$$

- 25 某城市100户居民的月平均用电量(单位:度), 以 $[160, 180)$, $[180, 200)$, $[200, 220)$, $[220, 240)$, $[240, 260)$, $[260, 280)$, $[280, 300)$ 分组的频率分布直方图如图.



- (1) 求直方图中 x 的值 .
- (2) 求月平均用电量的众数和中位数 .
- (3) 在月平均用电量为 $[220, 240)$, $[240, 260)$, $[260, 280)$, $[280, 300)$ 的四组用户中 , 用分层抽样的方法抽取11户居民 , 则月平均用电量在 $[220, 240)$ 的用户中应抽取多少户 ?

答案

- (1) 0.0075 .
- (2) 月平均用电量的众数是230 ; 月平均用电量的中位数为224 .
- (3) 月平均用电量在 $[220, 240)$ 的用户中应抽取5户 .

解析

- (1) 由直方图的性质可得 $(0.002 + 0.0095 + 0.011 + 0.0125 + x + 0.005 + 0.0025) \times 20 = 1$,
解方程可得 $x = 0.0075$,
 $\therefore$ 直方图中 x 的值为0.0075 .
- (2) 月平均用电量的众数是 $\frac{220 + 240}{2} = 230$,
 $\therefore (0.002 + 0.0095 + 0.011) \times 20 = 0.45 < 0.5$,
 $\therefore$ 月平均用电量的中位数在 $[220, 240)$ 内 ,
设中位数为 a , 由 $(0.002 + 0.0095 + 0.011) \times 20 + 0.0125 \times (a - 220) = 0.5$ 可得 $a = 224$,
 $\therefore$ 月平均用电量的中位数为224 .
- (3) 月平均用电量为 $[220, 240)$ 的用户有 $0.0125 \times 20 \times 100 = 25$,
月平均用电量为 $[240, 260)$ 的用户有 $0.0075 \times 20 \times 100 = 15$,
月平均用电量为 $[260, 280)$ 的用户有 $0.005 \times 20 \times 100 = 10$,
月平均用电量为 $[280, 300)$ 的用户有 $0.0025 \times 20 \times 100 = 5$,
 $\therefore$ 抽取比例为 $\frac{11}{25 + 15 + 10 + 5} = \frac{1}{5}$,
 $\therefore$ 月平均用电量在 $[220, 240)$ 的用户中应抽取 $25 \times \frac{1}{5} = 5$ 户 .

