

国庆作业--- 第九章统计

一、单项选择题(本大题共 8 小题, 每小题 5 分, 共 40 分)

1. 答案 A

解析 以上调查属于抽样调查, 故 A 错误.

2. 答案 A

解析 ①中, 考试成绩在不同分数段之间的同学有明显的差异, 用分层随机抽样比较恰当; ②中, 总体包含的个体较少, 用简单随机抽样比较恰当.

3. 答案 B

解析 找到第 8 行第 7 列的数开始向右读, 前五个符合条件的是 785, 667, 199, 507, 175, 故抽取检测的第 5 袋牛奶的编号是 175.

4. 答案 C

解析 估计该校学生数学竞赛成绩的平均分 $\bar{x} = 65 \times 0.2 + 75 \times 0.4 + 85 \times 0.3 + 95 \times 0.1 = 78$.

5. 答案 D

解析 对于 A, 由甲班体能测试等级条形图可知甲班人数为 $8 + 22 + 14 + 6 = 50$, 故甲、乙两个班共有学生 100 人, 故 A 正确;

对于 B, 由乙班体能测试等级扇形图可知, 乙班体能测试等级不合格的人数为 $50 \times 10\% = 5$, 故 B 正确;

对于 C, 甲班体能测试等级为良好以上(包含良好)的人数为 $8 + 22 = 30$, 乙班体能测试等级为良好以上(包含良好)的人数为 $48\% \times 50 + 12\% \times 50 = 30$, 即甲班体能测试等级为良好以上(包含良好)的人数与乙班一样, 故 C 正确;

对于 D, 甲班体能测试等级为合格的人数为 14, 乙班等级为合格的人数为 $(1 - 48\% - 12\% - 10\%) \times 50 = 15$, 故 D 不正确.

6. 答案 D

解析 设样本 $x_1, x_2, \dots, x_n$ 的平均数为 $\bar{x}$,

$$\text{由已知可得 } \frac{\sum_{i=1}^n (x_i + 2)}{n} = \frac{\sum_{i=1}^n x_i + 2n}{n} = \bar{x} + 2 = 10, \text{ 解得 } \bar{x} = 8,$$

$$\frac{\sum_{i=1}^n [(x_i + 2) - 10]^2}{n} = \frac{\sum_{i=1}^n [(x_i + 2) - (\bar{x} + 2)]^2}{n} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = 2,$$

所以样本 $2x_1 + 3, 2x_2 + 3, \dots, 2x_n + 3$ 的平均数为

$$\frac{\sum_{i=1}^n (2x_i + 3)}{n} = \frac{2 \sum_{i=1}^n x_i + 3n}{n} = 2\bar{x} + 3 = 19,$$

方差为

$$\frac{\sum_{i=1}^n [(2x_i + 3) - (2\bar{x} + 3)]^2}{n} = \frac{\sum_{i=1}^n (2x_i - 2\bar{x})^2}{n} = \frac{4 \sum_{i=1}^n (x_i - \bar{x})^2}{n} = 4 \times 2 = 8.$$

7. 答案 A

解析 由中位数的定义, 得

$$M = \frac{5+6}{2} = 5.5, \text{ 众数 } N = 5, \text{ 平均数 } \bar{x} = \frac{2 \times 3 + 3 \times 4 + 10 \times 5 + 6 \times 6 + 3 \times 7 + 2 \times 8 + 2 \times 9 + 2 \times 10}{30} \approx 5.97, \text{ 所以 } N < M < \bar{x}.$$

8. 答案 C

解析 设这个班有 n 个同学, 得分分别为 $a_1, a_2, \dots, a_i, \dots, a_n$, 第 i 个同学的得分没登记 ($a_i = \bar{x}$),

第一次计算时总分为 $(n-1)\bar{x}$,

$$\text{方差为 } s^2 = \frac{1}{n-1} [(a_1 - \bar{x})^2 + \dots + (a_{i-1} - \bar{x})^2 + (a_{i+1} - \bar{x})^2 + \dots + (a_n - \bar{x})^2].$$

$$\text{第二次计算时, 平均数 } \bar{x}_1 = \frac{(n-1)\bar{x} + \bar{x}}{n} = \bar{x},$$

$$\text{方差 } s_1^2 = \frac{1}{n} [(a_1 - \bar{x})^2 + \dots + (a_{i-1} - \bar{x})^2 + (a_{i+1} - \bar{x})^2 + \dots + (a_n - \bar{x})^2 + (\bar{x} - \bar{x})^2] = \frac{n-1}{n} s^2, \text{ 故 } s^2 > s_1^2.$$

二、多项选择题(本大题共 4 小题, 每小题 5 分, 共 20 分. 全部选对的得 5 分, 部分选对的得 2 分, 有选错的得 0 分)

9. 答案 ABC

解析 将数据从小到大排列为 46, 48, 51, 53, 53, 56, 56, 56, 58, 64, 66, 71, 共 12 个数据, 因为 $12 \times 20\% = 2.4$,

所以 20% 分位数为第三个数据, 即为 51, 故 A 正确;

因为 $12 \times 75\% = 9$, 所以 75% 分位数为 $\frac{58+64}{2} = 61$, 故 B 正确;

该组数据的中位数为 $\frac{56+56}{2} = 56$, 故 C 正确;

该组数据的平均数为

$$\frac{46+48+51+53+53+56+56+56+58+64+66+71}{12} = 56.5, \text{ 故 D 错误.}$$

10. 答案 BCD

解析 因为甲队平均每场进球数为 3.2, 乙队平均每场进球数为 1.8, 甲队平均数大于乙队平均数较多, 所以甲队技术比乙队好, 所以 A 不正确;

因为甲队全年进球数的标准差为 3, 乙队全年进球数的标准差为 0.3, 乙队的标准差小于甲队, 所以乙队发挥比甲队发挥稳定, 所以 B 正确;

因为乙队的标准差为 0.3, 说明每次进球数接近平均值, 乙队几乎每场都进球, 甲队标准差为 3, 说明甲队表现时好时坏, 所以 C, D 正确.

11. 答案 BCD

解析 对于 A, 若将一组数据中的每个数都加上一个相同的正数 x , 则平均数增加 x , 方差不变, 故 A 不正确;

对于 B, 抽样比为 $\frac{160}{500} = \frac{8}{25}$, 所以 $\frac{k}{8+5+4+k} = \frac{8}{25}$, 解得 $k=8$, 故 B 正确;

对于 C, 若这组数据出现 6, 设其他 4 个数分别为 x_1, x_2, x_3, x_4 , 则方差

$$s^2 = \frac{(x_1-2)^2 + (x_2-2)^2 + (x_3-2)^2 + (x_4-2)^2 + (6-2)^2}{5} \geq \frac{(6-2)^2}{5} = \frac{16}{5} > 2.4, \text{ 所以这组数据一定没有出现 } 6, \text{ 故 C 正确};$$

确;

对于 D, 因为数据 1, 4, 4, x , 7, 8 (其中 $x \neq 7$) 的中位数为 $\frac{4+x}{2}$, 众数为 4, 所以 $\frac{4+x}{2} = \frac{5}{4} \times 4$, 解得 $x=6$, 所以该组

数据的平均数为 $\frac{1+4+4+6+7+8}{6} = 5$, 故 D 正确.

12. 答案 CD

解析 对于 A, $\because (0.015 + 0.033 + a + 0.011 + 0.011) \times 10 = 1$,

$\therefore a = 0.03$, 故 A 错误;

对于 B, 由频率分布直方图, 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 $4000 \times 0.15 = 600$ (人), 故 B 错误;

对于 C, 平均年龄为

$$\bar{x} = (0.015 \times 15 + 0.033 \times 25 + 0.03 \times 35 + 0.011 \times 45 + 0.011 \times 55) \times 10 = 32 \text{ (岁)}, \text{ 故 C 正确};$$

对于 D, 设 75% 分位数为 x ,

则 $0.015 \times 10 + 0.033 \times 10 + (x - 30) \times 0.03 = 0.75$, 解得 $x = 39$, 故 D 正确.

三、填空题(本大题共 4 小题, 每小题 5 分, 共 20 分)

13. 答案 90, 70 84.375

解析 由题意可得高一年级抽取的样本量为 $\frac{450}{450+350} \times 160 = 90$, 高二年级抽取的样本量为 $\frac{350}{450+350} \times 160 =$

70, 此次数学竞赛的平均分约为 $\bar{x} = \frac{90}{90+70} \times 80 + \frac{70}{90+70} \times 90 = 84.375(\text{分})$.

14. 答案 4.7 4.75

解析 由图可知, 众数的估计值为 4.7.

第五小组的频率为 $0.5 \times 0.3 = 0.15$,

由从左至右五个小组的频率之比是 5 : 7 : 12 : 10 : 6,

可得第一小组的频率为 $0.15 \times \frac{5}{6} = 0.125$,

第二小组的频率为 $0.15 \times \frac{7}{6} = 0.175$.

第三小组的频率为 $0.15 \times \frac{12}{6} = 0.3$,

所以中位数在第三小组, 第三小组的矩形面积为 0.3, 则第三小组的高为 $\frac{0.3}{0.3} = 1$.

设中位数的估计值为 x , 则 $0.125 + 0.175 + (x - 4.55) \times 1 = 0.5$, 解得 $x = 4.75$.

15. 答案 1.6

解析 依题意得 $x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 = 20$,

设 x_1, x_2, x_3, x_4, x_5 的平均数为 $\bar{x}$,

根据方差的计算公式有

$$\frac{1}{5}[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + (x_4 - \bar{x})^2 + (x_5 - \bar{x})^2] = 1.44,$$

$$\therefore (x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2) - 2\bar{x}(x_1 + x_2 + x_3 + x_4 + x_5) + 5\bar{x}^2 = 7.2,$$

$$\text{即 } 20 - 10\bar{x}^2 + 5\bar{x}^2 = 7.2, \text{ 又 } \bar{x} > 0, \therefore \bar{x} = 1.6.$$

16. 答案 146

解析 由题意知, 总样本平均数为

$$\bar{x} = \frac{100}{300} \times 167 + \frac{100}{300} \times 170 + \frac{100}{300} \times 173 = 170,$$

$\therefore$ 总样本方差为

$$s^2 = \frac{100}{300} \times [120 + (167 - 170)^2] + \frac{100}{300} \times [150 + (170 - 170)^2] + \frac{100}{300} \times [150 + (173 - 170)^2]$$

$$= 146.$$

四、解答题(本大题共 6 小题, 共 70 分)

17. 解 (1)数据整理如表所示:

健康情况	健康	基本健康	不健康尚能自理	不能自理
80 岁及以上	20	45	20	15
80 岁以下	200	225	50	25

从图表中知,

$$80 \text{ 岁及以上老人应抽取 } 8 \times \frac{15}{15+25} = 3(\text{人}).$$

$$80 \text{ 岁以下老人应抽取 } 8 \times \frac{25}{15+25} = 5(\text{人}).$$

$$(2)(20+45+20+15) \div 600 = \frac{1}{6},$$

$$\frac{66 \times \frac{1}{6}}{400} \times 100\% = 2.75\%,$$

估算该市 80 岁及以上老人占全市户籍人口的 2.75%.

$$18. \text{ 解 } (1) \text{ 由样本频率分布直方图得样本在 } [15, 18) \text{ 内的频率等于 } \frac{4}{75} \times 3 = \frac{4}{25}.$$

$$(2) \because \text{样本在 } [15, 18) \text{ 内的频数为 } 8, \text{ 由(1)可知, 样本容量为 } 8 \div \frac{4}{25} = 50.$$

$$(3) \because \text{在 } [12, 15) \text{ 内的小矩形面积为 } 0.06, \text{ 故样本在 } [12, 15) \text{ 内的频率为 } 0.06, \text{ 故样本在 } [15, 33) \text{ 内的频数为 } 50 \times (1 - 0.06) = 47, \text{ 又在 } [15, 18) \text{ 内的频数为 } 8, \text{ 故在 } [18, 33) \text{ 内的频数为 } 47 - 8 = 39.$$

19 解 (1)根据题意, 读取的编号依次是 512,916(超界), 935(超界), 805,770,951(超界), 512(重复), 687,858,554,876,647,547,332.

将有效的编号由小到大排序, 得 332,512,547,554,647,687,770,805,858,876,

$$\text{故样本编号的中位数为 } \frac{647+687}{2} = 667.$$

(2)设样本中选择 A 题目的成绩的平均数为 $\bar{x}$, 方差为 s^2 ;

样本中选择 B 题目的成绩的平均数为 $\bar{y}$, 方差为 t^2 ,

$$\text{则 } \bar{x} = 7, s^2 = 4, \bar{y} = 8, t^2 = 1,$$

$$\text{所以样本的平均数为 } \frac{8}{8+2} \bar{x} + \frac{2}{8+2} \bar{y} = \frac{4}{5} \times 7 + \frac{1}{5} \times 8 = 7.2,$$

$$\text{方差} = \frac{8}{8+2} \times [s^2 + (\bar{x} - 7.2)^2] + \frac{2}{8+2} \times [t^2 + (\bar{y} - 7.2)^2]$$

$$= \frac{4}{5} \times [4 + (7 - 7.2)^2] + \frac{1}{5} \times [1 + (8 - 7.2)^2] = 3.56.$$

故该校 900 名学生的选做题得分的平均数约为 7.2，方差约为 3.56.

20 解 (1)A 公司员工月均工资的平均数为

$$0.3 \times 0.21 + 0.5 \times 0.29 + 0.7 \times 0.27 + 0.9 \times 0.21 + 2.9 \times 0.02 = 1.166(\text{万元}).$$

由图 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 A 公司员工月均工资的中位数约为 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,该公司少数员工的月收入很高,在这种情况下平均数并不能较好地反映普通员工的收入水平,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收入水平.

(2)B 公司员工月均工资的平均数为

$$(0.3 \times 0.375 + 0.5 \times 0.75 + 0.7 \times 2.75 + 0.9 \times 1 + 1.1 \times 0.125) \times 0.2 = 0.69(\text{万元}),$$

由图 2 知, B 公司员工月均工资在 0.6 万元以下的频率为 $(0.375 + 0.75) \times 0.2 = 0.225$, 在 0.8 万元以下的频率为

$$(0.375 + 0.75 + 2.75) \times 0.2 = 0.775.$$

设 B 公司员工月均工资的中位数为 x 万元,

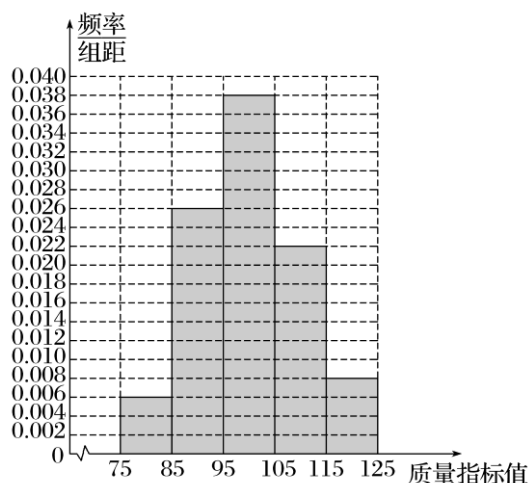
$$\text{则 } (x - 0.6) \times 2.75 = 0.5 - 0.225, \text{ 解得 } x = 0.7.$$

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收入水平, B 公司员工月均工资平均数为 0.69, 中位数为 0.7, 大于 A 公司员工月均工资的中位数 0.6, 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.

21. 解 (1)由题意可知, 分组 $[75,85)$, $[85,95)$, $[95,105)$, $[105,115)$, $[115,125]$ 对应的频率分别为 0.06, 0.26, 0.38, 0.22, 0.08.

则频率分布直方图如图所示.



(2)质量指标值的样本平均数为

$$\bar{x} = 80 \times 0.06 + 90 \times 0.26 + 100 \times 0.38 + 110 \times 0.22 + 120 \times 0.08 = 100.$$

质量指标值的样本方差为

$$s^2 = (100 - 80)^2 \times 0.06 + (100 - 90)^2 \times 0.26 + (100 - 100)^2 \times 0.38 + (100 - 110)^2 \times 0.22 + (100 - 120)^2 \times 0.08 = 104.$$

(3)由题意知,质量指标值落在区间[75,105)内的产品有 70 件,设质量指标值为 $x_1, x_2, \dots, x_{70}$, 平均数为 $\bar{x} = 94$, 方差为 $s_x^2 = 40$,

则质量指标值落在区间[105,125]内的产品有 30 件,设质量指标值分别为 $y_1, y_2, \dots, y_{30}$, 平均数为 $\bar{y}$, 方差为 s_y^2 ,

则这 100 件产品的质量指标值的平均数为 $\bar{z} = 100$, 方差为 $s_z^2 = 202$, 则 $100\bar{z} = 70\bar{x} + 30\bar{y}$,

所以 $\bar{y} = 114$, 又因为 $s_x^2 = \frac{1}{70} \sum_{i=1}^{70} x_i^2 - \bar{x}^2$, 则

$$\sum_{i=1}^{70} x_i^2 = 621\,320, \text{ 又因为 } s_z^2 = \frac{1}{100} \left(\sum_{i=1}^{70} x_i^2 + \sum_{j=1}^{30} y_j^2 \right) - \bar{z}^2,$$

$$\text{则 } \sum_{j=1}^{30} y_j^2 = 398\,880, \text{ 所以 } s_y^2 = \frac{1}{30} \sum_{j=1}^{30} y_j^2 - \bar{y}^2 = 300.$$

22. 解 (1)由频率分布直方图得,一刀宣纸有正牌 $100 \times 0.1 \times 4 = 40$ (张),有副牌 $100 \times 0.05 \times 4 \times 2 = 40$ (张),有废品 $100 \times 0.025 \times 4 \times 2 = 20$ (张),

所以该公司一刀宣纸的利润的估计值为 $40 \times 15 + 40 \times 8 + (-20) \times 20 = 520$ (元),

可以估计该公司的年利润为 520 万元.

(2)由频率分布直方图得,

$$\bar{x} = 42 \times 0.025 \times 4 + 46 \times 0.05 \times 4 + 50 \times 0.1 \times 4 + 54 \times 0.05 \times 4 + 58 \times 0.025 \times 4 = 50.$$

由表格数据可得一刀宣纸中正牌的张数估计为

$$100 \times 0.6827 = 68.27,$$

$$\text{废品的张数估计为 } 100 \times (1 - 0.9545) = 4.55,$$

$$\text{副牌的张数估计为 } 100 \times (0.9545 - 0.6827) = 27.18,$$

$$\text{所以一刀宣纸的利润为 } 68.27 \times 12 + 27.18 \times 5 - 4.55 \times 20 = 864.14(\text{元}),$$

$$10\,000 \text{ 刀宣纸的利润为 } 10\,000 \times 864.14 = 8\,641\,400(\text{元}) = 864.14(\text{万元}),$$

$$\text{所以购买这种机器后该公司的利润为 } 864.14 - 100 = 764.14(\text{万元}),$$

因为 $764.14 > 520$ ，所以建议该公司购买这种机器。

仅供会东县中学使用