

文科概统第三节抽样与频率分布

1、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 7 人，则样本容量为_____

解： $\frac{n}{750} = \frac{7}{350}, n = 15$ ，注： $a:b:c = A:B:C, \frac{a}{A} = \frac{b}{B} = \frac{c}{C} = \frac{n}{N}$

2、将参加夏令营的 600 名学生编号为：001, 002, ……600，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 在第 I 营区，从 301 到 495 住在第 II 营区，从 496 到 600 在第 III 营区，三个营区被抽中的人数依次为_____

解： $n = 50, d = \frac{600}{50} = 12$ ，

设在 001 到 300 中有 n_1 个，则 $3 + 12(n_1 - 1) \leq 300, 1 + 4(n_1 - 1) \leq 100, n_1 \leq \frac{99}{4} + 1 = 25\frac{3}{4}$ ，

设在 001 到 300 中，首项 3，末项 291，于是 $n_1 = \frac{291 - 3}{12} + 1 = 25$ ，

从 301 到 495 中，首项 304，末项 495，于是 $n_2 = \frac{495 - 303}{12} + 1 = 17, n_3 = 50 - 25 - 17 = 8$

3、8 个班参加合唱比赛的得分茎叶图，

则这组数据的中位数和平均数分别是(A)

A、91.5 和 91.5 B、91.5 和 92 C、91 和 91.5 D、92 和 92

茎	叶
8	9 7
9	3 1 6 4 0 2

解：中位数 $= (91 + 92) / 2 = 91.5$ $\bar{x} = 90 + \frac{1}{8}(-1 - 3 + 3 + 1 + 6 + 4 + 0 + 2) = 91.5$

4. 一样本频率分布直方图(如图)

(1) 众数 \approx _____ (2) 中位数 \approx _____ (3) 平均体重数 \approx _____

解：(1) 众数 ≈ 55

(2) 中位数 $\approx 60 + 60 + 10 \times \frac{1}{3} = 63.33$

(3) $\bar{x} = 45 \times 0.05 + 55 \times 0.35 + 65 \times 0.3 + 75 \times 0.2 + 85 \times 0.1 = 64.5$

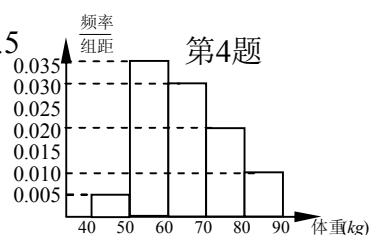
补 100 名体重在 [60,70), [70,80), [80,90] 三组内用分层抽样的方法取 12 人则选取的 12 人中体重在 [60,80) 内的男生要抽取的人数为_____

解 1:

	频率	频数
[60,70)	0.3	30
[70,80)	0.2	20
[80,90]	0.1	10

$$\frac{n}{50} = \frac{12}{60}, n = 10$$

解 2: [60,70), [70,80), [80,90] 的比例 3:2:1, [60,80) 有 5 份, 于是 $12 \times \frac{5}{6} = 10$



	频率
[40,50)	0.05
[50,60)	0.35
[60,70)	0.30
[70,80)	0.2
[80,90]	0.1

5、为了了解某工厂开展群众体育活动的情况，拟采用分层抽样的方法从 A, B, C 三个区中抽取 7 个工厂进行调查，已知 A, B, C 区中分别有 18, 27, 18 个工厂 (I) 求从 A, B, C 区中分别抽取的工厂个数；(II) 若从抽取的 7 个工厂中随机抽取 2 个进行调查结果的对比，用列举法计算这 2 个工厂中至少有 1 个来自 A 区的概率。

解：(1) $18:27:18 = 2:3:2, 7 \times \frac{2}{7} = 2, 7 \times \frac{3}{7} = 3, 7 \times \frac{2}{7} = 2$

从 A, B, C 区中分别抽取的工厂个数分别是 2, 3, 2

(2) 设 $M = \{\text{这2个工厂中至少有1个来自A区}\}$

设从 A, B, C 抽得工厂分别记为 $a_1, a_2; b_1, b_2, b_3; c_1, c_2$

试验的基本事件有:

$\{a_1, a_2\}, \{a_1, b_1\}, \{a_1, b_2\}, \{a_1, b_3\}, \{a_1, c_1\}, \{a_1, c_2\},$
 $\{a_2, b_1\}, \{a_2, b_2\}, \{a_2, b_3\}, \{a_2, c_1\}, \{a_2, c_2\},$
 $\{b_1, b_2\}, \{b_1, b_3\}, \{b_1, c_1\}, \{b_1, c_2\},$
 $\{b_2, b_3\}, \{b_2, c_1\}, \{b_2, c_2\},$
 $\{b_3, c_1\}, \{b_3, c_2\},$
 $\{c_1, c_2\}$, 共 21 种

事件的基本事件有:

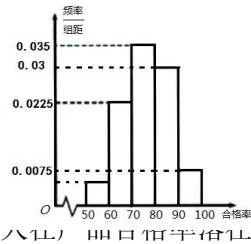
$\{a_1, a_2\}, \{a_1, b_1\}, \{a_1, b_2\}, \{a_1, b_3\}, \{a_1, c_1\}, \{a_1, c_2\},$

$\{a_2, b_1\}, \{a_2, b_2\}, \{a_2, b_3\}, \{a_2, c_1\}, \{a_2, c_2\}$ 共 11 种, $P(A) = \frac{11}{21}$

6、某工厂有工人 40 人, 在一次产品大检查中, 将每人加工的产品合格率绘制成如图所示的频率分布直方图

(I) 求本次产品大检查中, 产品合格率 $[50, 60)$ 内的工人人数

(II) 从产品合格率落在 $[50, 70)$ 内的工人中随机选取 2 人, 求至少一人在 $[50, 60)$ 内的概率



解: (I) 产品合格率为 $[50, 60)$ 内的频率为:

$$1 - (0.035 + 0.03 + 0.0225 + 0.0075) \times 10 = 0.05,$$

所以产品合格率为 $[50, 60)$ 内的人数共有 $40 \times 0.05 = 2$ 人

(II) 同 (I) 可得产品合格率为 $[60, 70)$ 内的人数有

$$40 \times 0.0075 \times 10 = 3,$$

所以产品合格率为 $[50, 70)$ 内的人数共有 5 人。

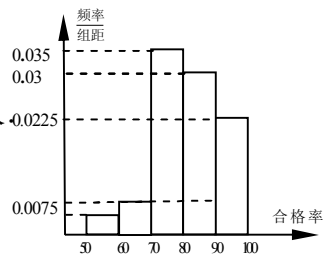
记为 a, b, c, d, e, 则所有的基本事件有

$\{a, b\}, \{a, c\}, \{a, d\}, \{a, e\}, \{b, c\}, \{b, d\}, \{b, e\}, \{c, d\}, \{c, e\}, \{d, e\}$

设至少一人在产品合格率为 $[50, 60)$ 内为事件 A

则事件 A 所含的基本事件有: $\{a, b\}, \{a, c\}, \{a, d\}, \{a, e\}, \{b, c\}, \{b, d\}, \{b, e\}$

于是 $P(A) = \frac{7}{10}$ 答



7、小明订了一份日报, 暑假期间他收集了每天报纸送达的时间的数据, 并绘制成频率分布直方图如图所示

(I) 请你根据图中的数据信息, 写出众数 $x_0 =$ _____

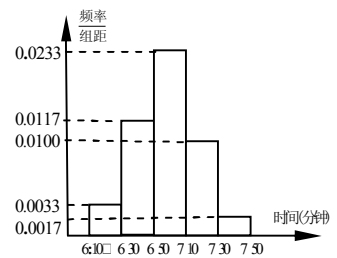
(II) 小明的父亲离家去上班的时间 y 在上午 7:00-7:30 之间,

为此小明要求送报人每在 x_0 时间前后半小时内把报纸送达

(每个时间点送达的可能性相等)

(1) 求小明的父亲在去上班前能取到取到报纸(称为事件 A)的概率

(2) 求小明的父亲一周 5 天(假日除外)能取到取到报纸的的天数 X 的数学期望



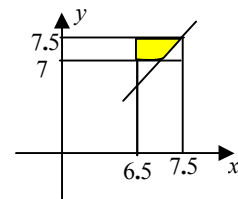
解: (I) $x_0 = 7:00$

(II) (1) 设报纸送达时间为 $x, x \in [6.5, 7.5]$

试验的基本事件为区域 $\Omega = \{(x, y) | 6.5 \leq x \leq 7.5, 7 \leq y \leq 7.5\}$

事件 A 的基本事件为区域 $\{(x, y) | 6.5 \leq x \leq 7.5, 7 \leq y \leq 7.5, x \leq y\}$

作出上述区域如图



$$S_{\Omega} = 1 \times \frac{1}{2} = \frac{1}{2}, \quad S_A = \frac{1}{2} - \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{3}{8}$$

$$P(A) = \frac{S_A}{S_{\Omega}} = \frac{\frac{3}{8}}{\frac{1}{2}} = \frac{3}{4}$$

8. 某班有甲、乙两个研究性学习小组，甲组有 10 个学生，乙组有 5 个学生，他们的小组个人综合成绩如茎叶图所示。

- (1) 问甲、乙哪个小组的平均成绩更高？
 (2) 从乙两组中共抽取 3 个学生的成绩，求至少一人成绩优秀的概率；
 (成绩在 80 分以上为优秀)

甲组			乙组	
3	7	8	6	6
		2	7	6
2	3	5	8	9
1	4	5	9	5

解：(1) $\bar{x}_甲 = \frac{1}{10}(63+67+68+72+82+83+85+91+92+95) = 80$

$\bar{x}_乙 = \frac{1}{10}(66+76+79+89+95) = 81$

因 $\bar{x}_甲 < \bar{x}_乙$ ，故乙组平均成绩更高

(2)

9、2012 年 3 月 2 日，国家环保部门发布《环境空气质量标准》，其中规定：居民区的 PM2.5 的年平均浓度不得超过 35 微克/立方米，PM2.5 的 24 小时平均浓度不得超过 75 微克/立方米。某城市环保部门随机抽取了一居民区去年 20 天 PM2.5 的 24 小时平均浓度的监测数据，数据统计如下

组别	PM2.5 微克/立方米	频 数 (天)	频率
第一组	(0,25]	5	0.25
第二组	(25, 50]	10	0.5
第三组	(50,75]	3	0.15
第四组	(75,100]	2	0.1

(I) 从样本 PM2.5 的 24 小时平均浓度超过 50 微克/立方米的 5 天中随机抽取 2 天，求恰好有一天 PM2.5 的 24 小时平均浓度超过 75 微克/立方米的概率；

(II) 求样本平均数，并根据样本估计总体的思想，从 PM2.5 的年平均浓度考虑，判断该居民区的环境是否需要改进？说明理由。

解：(I) 设这五天分 PM2.5 在 (50,75] 的 3 天为 a_1, a_2, a_3 ，在 (75,100] 的 2 天为 b_1, b_2

随机抽取 2 天所有可能的结果为：

$\{a_1, a_2\}, \{a_1, a_3\}, \{a_1, b_1\}, \{a_1, b_2\}, \{a_2, a_3\}, \{a_2, b_1\}, \{a_2, b_2\}, \{a_3, b_1\}, \{a_3, b_2\}, \{b_1, b_2\}$

共 10 个， 设所求事件为 A，则 A 包含的基本事件为：

$\{a_1, b_1\}, \{a_1, b_2\}, \{a_2, b_1\}, \{a_2, b_2\}, \{a_3, b_1\}, \{a_3, b_2\}$ ，共 6 个，故 $P(A) = \frac{6}{10} = 0.6$

(II) $\bar{x} = \frac{25}{2} \times 0.25 + \frac{75}{2} \times 0.5 + \frac{125}{2} \times 0.15 + \frac{175}{2} \times 0.1 = 40$ (微克/立方米)

因为 $\bar{x} = 40 > 35$ ，所抽的 20 天中有 2 天也超标
 因此该居民区的环境需要改进