

专题 16 概率与统计（解答题）（文科专用）

1. 【2022 年全国甲卷】甲、乙两城之间的长途客车均由 A 和 B 两家公司运营，为了解这两家公司长途客车的运行情况，随机调查了甲、乙两城之间的 500 个班次，得到下面列联表：

	准点班次数	未准点班次数
A	240	20
B	210	30

(1) 根据上表，分别估计这两家公司甲、乙两城之间的长途客车准点的概率；

(2) 能否有 90% 的把握认为甲、乙两城之间的长途客车是否准点与客车所属公司有关？

$$\text{附： } K^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)},$$

$P(K^2 \geq k)$	0.100	0.050	0.010
k	2.706	3.841	6.635

【答案】(1) A , B 两家公司长途客车准点的概率分别为 $\frac{12}{13}$, $\frac{7}{8}$

(2) 有

【解析】

【分析】

(1) 根据表格中数据以及古典概型的概率公式可求得结果；

(2) 根据表格中数据及公式计算 K^2 ，再利用临界值表比较即可得结论.

(1)

根据表中数据， A 共有班次 260 次，准点班次有 240 次，

设 A 家公司长途客车准点事件为 M ，

$$\text{则 } P(M) = \frac{240}{260} = \frac{12}{13};$$

B 共有班次 240 次，准点班次有 210 次，

设 B 家公司长途客车准点事件为 N ，

$$\text{则 } P(N) = \frac{210}{240} = \frac{7}{8}.$$

A 家公司长途客车准点的概率为 $\frac{12}{13}$;

B 家公司长途客车准点的概率为 $\frac{7}{8}$.

(2)

列联表

	准点班次数	未准点班次数	合计
A	240	20	260
B	210	30	240
合计	450	50	500

$$K^2 = \frac{n(ad - bc)^2}{(a + b)(c + d)(a + c)(b + d)}$$

$$= \frac{500 \times (240 \times 30 - 210 \times 20)^2}{260 \times 240 \times 450 \times 50} \approx 3.205 > 2.706,$$

根据临界值表可知，有 90% 的把握认为甲、乙两城之间的长途客车是否准点与客车所属公司有关.

2. 【2022 年全国乙卷】某地经过多年的环境治理，已将荒山改造成了绿水青山. 为估计一林区某种树木的总材积量，随机选取了 10 棵这种树木，测量每棵树的根部横截面积（单位： m^2 ）和材积量（单位： m^3 ），得到如下数据：

样本号 i	1	2	3	4	5	6	7	8	9	10	总和
根部横截面积 x_i	0.04	0.06	0.04	0.08	0.08	0.05	0.05	0.07	0.07	0.06	0.6
材积量 y_i	0.25	0.40	0.22	0.54	0.51	0.34	0.36	0.46	0.42	0.40	3.9

并计算得 $\sum_{i=1}^{10} x_i^2 = 0.038$, $\sum_{i=1}^{10} y_i^2 = 1.6158$, $\sum_{i=1}^{10} x_i y_i = 0.2474$.

(1) 估计该林区这种树木平均一棵的根部横截面积与平均一棵的材积量；

(2) 求该林区这种树木的根部横截面积与材积量的样本相关系数（精确到 0.01）；

(3) 现测量了该林区所有这种树木的根部横截面积，并得到所有这种树木的根部横截面积总和为 186m^2 . 已知树木的材积量与其根部横截面积近似成正比. 利用以上数据给出该林区这种树木的总材积量的估计值.

附：相关系数 $r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}, \sqrt{1.896} \approx 1.377$.

【答案】(1) 0.06m^2 ; 0.39m^3

(2) 0.97

(3)1209m³

【解析】

【分析】

(1) 计算出样本的一棵根部横截面积的平均值及一棵材积量平均值，即可估计该林区这种树木平均一棵的根部横截面积与平均一棵的材积量；

(2) 代入题给相关系数公式去计算即可求得样本的相关系数值；

(3) 依据树木的材积量与其根部横截面积近似成正比，列方程即可求得该林区这种树木的总材积量的估计值．

(1)

样本中 10 棵这种树木的根部横截面积的平均值 $\bar{x} = \frac{0.6}{10} = 0.06$

样本中 10 棵这种树木的材积量的平均值 $\bar{y} = \frac{3.9}{10} = 0.39$

据此可估计该林区这种树木平均一棵的根部横截面积为 0.06m²，

平均一棵的材积量为 0.39m³

(2)

$$\begin{aligned} r &= \frac{\sum_{i=1}^{10} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{10} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{10} (y_i - \bar{y})^2}} = \frac{\sum_{i=1}^{10} x_i y_i - 10\bar{x}\bar{y}}{\sqrt{(\sum_{i=1}^{10} x_i^2 - 10\bar{x}^2)(\sum_{i=1}^{10} y_i^2 - 10\bar{y}^2)}} \\ &= \frac{0.2474 - 10 \times 0.06 \times 0.39}{\sqrt{(0.038 - 10 \times 0.06^2)(1.6158 - 10 \times 0.39^2)}} = \frac{0.0134}{\sqrt{0.0001896}} \approx \frac{0.0134}{0.01377} \approx 0.97 \end{aligned}$$

则 $r \approx 0.97$

(3)

设该林区这种树木的总材积量的估计值为 Y m³，

又已知树木的材积量与其根部横截面积近似成正比，

可得 $\frac{0.06}{0.39} = \frac{186}{Y}$ ，解之得 $Y=1209$ m³．

则该林区这种树木的总材积量估计为 1209m³

3. 【2021 年甲卷文科】甲、乙两台机床生产同种产品，产品按质量分为一级品和二级品，为了比较两台机床产品的质量，分别用两台机床各生产了 200 件产品，产品的质量情况统计如下表：

	一级品	二级品	合计
--	-----	-----	----

甲机床	150	50	200
乙机床	120	80	200
合计	270	130	400

(1) 甲机床、乙机床生产的产品中一级品的频率分别是多少?

(2) 能否有 99% 的把握认为甲机床的产品质量与乙机床的产品质量有差异?

$$\text{附: } K^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$$

$P(K^2 \geq k)$	0.050	0.010	0.001
k	3.841	6.635	10.828

【答案】(1) 75%; 60%;

(2) 能.

【解析】

【分析】

根据给出公式计算即可

【详解】

(1) 甲机床生产的产品中的一级品的频率为 $\frac{150}{200} = 75\%$,

乙机床生产的产品中的一级品的频率为 $\frac{120}{200} = 60\%$.

$$(2) K^2 = \frac{400(150 \times 80 - 120 \times 50)^2}{270 \times 130 \times 200 \times 200} = \frac{400}{39} > 10 > 6.635,$$

故能有 99% 的把握认为甲机床的产品与乙机床的产品质量有差异.

4. 【2021 年乙卷文科】某厂研制了一种生产高精产品的设备，为检验新设备生产产品的某项指标有无提高，用一台旧设备和一台新设备各生产了 10 件产品，得到各件产品该项指标数据如下：

旧设备	9.8	10.3	10.0	10.2	9.9	9.8	10.0	10.1	10.2	9.7
-----	-----	------	------	------	-----	-----	------	------	------	-----

新设备	10.1	10.4	10.1	10.0	10.1	10.3	10.6	10.5	10.4	10.5
-----	------	------	------	------	------	------	------	------	------	------

旧设备和新设备生产产品的该项指标的样本平均数分别记为 $\bar{x}$ 和 $\bar{y}$ ，样本方差分别记为 s_1^2 和 s_2^2 。

(1) 求 $\bar{x}$ ， $\bar{y}$ ， s_1^2 ， s_2^2 ；

(2) 判断新设备生产产品的该项指标的均值较旧设备是否有显著提高（如果

$\bar{y} - \bar{x} \geq 2\sqrt{\frac{s_1^2 + s_2^2}{10}}$ ，则认为新设备生产产品的该项指标的均值较旧设备有显著提高，否则不认为有显著提高）。

【答案】(1) $\bar{x} = 10, \bar{y} = 10.3, s_1^2 = 0.036, s_2^2 = 0.04$ ；(2) 新设备生产产品的该项指标的均值较旧设备有显著提高。

【解析】

【分析】

(1) 根据平均数和方差的计算方法，计算出平均数和方差。

(2) 根据题目所给判断依据，结合 (1) 的结论进行判断。

【详解】

$$(1) \bar{x} = \frac{9.8 + 10.3 + 10 + 10.2 + 9.9 + 9.8 + 10 + 10.1 + 10.2 + 9.7}{10} = 10,$$

$$\bar{y} = \frac{10.1 + 10.4 + 10.1 + 10 + 10.1 + 10.3 + 10.6 + 10.5 + 10.4 + 10.5}{10} = 10.3,$$

$$s_1^2 = \frac{0.2^2 + 0.3^2 + 0 + 0.2^2 + 0.1^2 + 0.2^2 + 0 + 0.1^2 + 0.2^2 + 0.3^2}{10} = 0.036,$$

$$s_2^2 = \frac{0.2^2 + 0.1^2 + 0.2^2 + 0.3^2 + 0.2^2 + 0 + 0.3^2 + 0.2^2 + 0.1^2 + 0.2^2}{10} = 0.04.$$

$$(2) \text{依题意, } \bar{y} - \bar{x} = 0.3 = 2 \times 0.15 = 2\sqrt{0.15^2} = 2\sqrt{0.0225}, \quad 2\sqrt{\frac{0.036 + 0.04}{10}} = 2\sqrt{0.0076},$$

$$\bar{y} - \bar{x} \geq 2\sqrt{\frac{s_1^2 + s_2^2}{10}}, \text{ 所以新设备生产产品的该项指标的均值较旧设备有显著提高.}$$

5. 【2020 年新课标 1 卷文科】某厂接受了一项加工业务，加工出来的产品(单位：件)按标准分为 A，B，C，D 四个等级.加工业务约定：对于 A 级品、B 级品、C 级品，厂家每件分别收取加工费 90 元，50 元，20 元；对于 D 级品，厂家每件要赔偿原料损失费 50 元.该厂有甲、

乙两个分厂可承接加工业务.甲分厂加工成本费为 25 元/件，乙分厂加工成本费为 20 元/件. 厂家为决定由哪个分厂承接加工业务，在两个分厂各试加工了 100 件这种产品，并统计了这些产品的等级，整理如下：

甲分厂产品等级的频数分布表

等级	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
频数	40	20	20	20

乙分厂产品等级的频数分布表

等级	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
频数	28	17	34	21

- (1) 分别估计甲、乙两分厂加工出来的一件产品为 A 级品的概率；
- (2) 分别求甲、乙两分厂加工出来的 100 件产品的平均利润，以平均利润为依据，厂家应选哪个分厂承接加工业务？

【答案】(1) 甲分厂加工出来的 A 级品的概率为 0.4，乙分厂加工出来的 A 级品的概率为 0.28；

(2) 选甲分厂，理由见解析.

【解析】

【分析】

- (1) 根据两个频数分布表即可求出；
- (2) 根据题意分别求出甲乙两厂加工 100 件产品的总利润，即可求出平均利润，由此作出选择.

【详解】

(1) 由表可知，甲厂加工出来的一件产品为 A 级品的概率为 $\frac{40}{100} = 0.4$ ，乙厂加工出来的一件产品为 A 级品的概率为 $\frac{28}{100} = 0.28$ ；

(2) 甲分厂加工 100 件产品的总利润为

$$40 \times (90 - 25) + 20 \times (50 - 25) + 20 \times (20 - 25) - 20 \times (50 + 25) = 1500 \text{ 元},$$

所以甲分厂加工 100 件产品的平均利润为 15 元每件；

乙分厂加工100件产品的总利润为

$$28 \times (90 - 20) + 17 \times (50 - 20) + 34 \times (20 - 20) - 21 \times (50 + 20) = 1000 \text{ 元},$$

所以乙分厂加工100件产品的平均利润为10元每件.

故厂家选择甲分厂承接加工任务.

【点睛】

本题主要考查古典概型的概率公式的应用, 以及平均数的求法, 并根据平均值作出决策, 属于基础题.

6. 【2019年新课标1卷文科】某商场为提高服务质量, 随机调查了50名男顾客和50名女顾客, 每位顾客对该商场的服务给出满意或不满意的评价, 得到下面列联表:

	满意	不满意
男顾客	40	10
女顾客	30	20

(1) 分别估计男、女顾客对该商场服务满意的概率;

(2) 能否有95%的把握认为男、女顾客对该商场服务的评价有差异?

$$\text{附: } K^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}.$$

$P(K^2 \geq k)$	0.050	0.010	0.001
k	3.841	6.635	10.828

【答案】(1) $\frac{4}{5}, \frac{3}{5}$;

(2) 能有95%的把握认为男、女顾客对该商场服务的评价有差异.

【解析】

【分析】

(1) 从题中所给的 2×2 列联表中读出相关的数据, 利用满意的人数除以总的人数, 分别算出相应的频率, 即估计得出的概率值;

(2) 利用公式求得观测值与临界值比较, 得到能有95%的把握认为男、女顾客对该商场服

务的评价有差异.

【详解】

(1) 由题中表格可知, 50 名男顾客对商场服务满意的有 40 人,

所以男顾客对商场服务满意率估计为 $P_1 = \frac{40}{50} = \frac{4}{5}$,

50 名女顾客对商场满意的有 30 人,

所以女顾客对商场服务满意率估计为 $P_2 = \frac{30}{50} = \frac{3}{5}$,

(2) 由列联表可知 $K^2 = \frac{100(40 \times 20 - 30 \times 10)^2}{70 \times 30 \times 50 \times 50} = \frac{100}{21} \approx 4.762 > 3.841$,

所以能有 95% 的把握认为男、女顾客对该商场服务的评价有差异.

【点睛】

该题考查的是有关概率与统计的知识, 涉及到的知识点有利用频率来估计概率, 利用列联表计算 K^2 的值, 独立性检验, 属于简单题目.

7. 【2019 年新课标 2 卷文科】某行业主管部门为了解本行业中小企业的生产情况, 随机调查了 100 个企业, 得到这些企业第一季度相对于前一年第一季度产值增长率 y 的频数分布表.

y 的分组	$[-0.20, 0)$	$[0, 0.20)$	$[0.20, 0.40)$	$[0.40, 0.60)$	$[0.60, 0.80)$
企业数	2	24	53	14	7

(1) 分别估计这类企业中产值增长率不低于 40% 的企业比例、产值负增长的企业比例;

(2) 求这类企业产值增长率的平均数与标准差的估计值 (同一组中的数据用该组区间的中点值为代表). (精确到 0.01)

附: $\sqrt{74} \approx 8.602$.

【答案】(1) 增长率超过 40% 的企业比例为 $\frac{21}{100}$, 产值负增长的企业比例为 $\frac{2}{100} = \frac{1}{50}$; (2) 平均数 0.3; 标准差 0.17.

【解析】

【分析】

(1) 本题首先可以通过题意确定 100 个企业中增长率超过 40% 的企业以及产值负增长的企业个数, 然后通过增长率超过 40% 的企业以及产值负增长的企业个数除随机调查的企业

总数即可得出结果；

(2)可通过平均值以及标准差的计算公式得出结果.

【详解】

(1)由题意可知，随机调查的100个企业中增长率超过 40% 的企业有 $14+7=21$ 个，

产值负增长的企业有2个，

所以增长率超过 40% 的企业比例为 $\frac{21}{100}$ ，产值负增长的企业比例为 $\frac{2}{100}=\frac{1}{50}$.

(2)由题意可知，平均值 $\bar{y}=\frac{2'(-0.1)+24'0.1+53'0.3+14'0.5+7'0.7}{100}=0.3$ ，

标准差的平方：

$$\begin{aligned}s^2 &= \frac{1}{100} [2'(-0.1-0.3)^2 + 24'(0.1-0.3)^2 + 53'(0.3-0.3)^2 + 14'(0.5-0.3)^2 + 7'(0.7-0.3)^2] \\&= \frac{1}{100} [0.32 + 0.96 + 0.56 + 1.12] = 0.0296,\end{aligned}$$

所以标准差 $s=\sqrt{0.0296}=\sqrt{0.0004}\times 0.02\sqrt{4}\times 0.02\sqrt{4}\times 0.17=0.17$.

【点睛】

本题考查平均值以及标准差的计算，主要考查平均值以及标准差的计算公式，考查学生从信息题中获取所需信息的能力，考查学生的计算能力，是简单题.

8.【2018年新课标1卷文科】某家庭记录了未使用节水龙头50天的日用水量数据(单位： m^3)

和使用了节水龙头50天的日用水量数据，得到频数分布表如下：

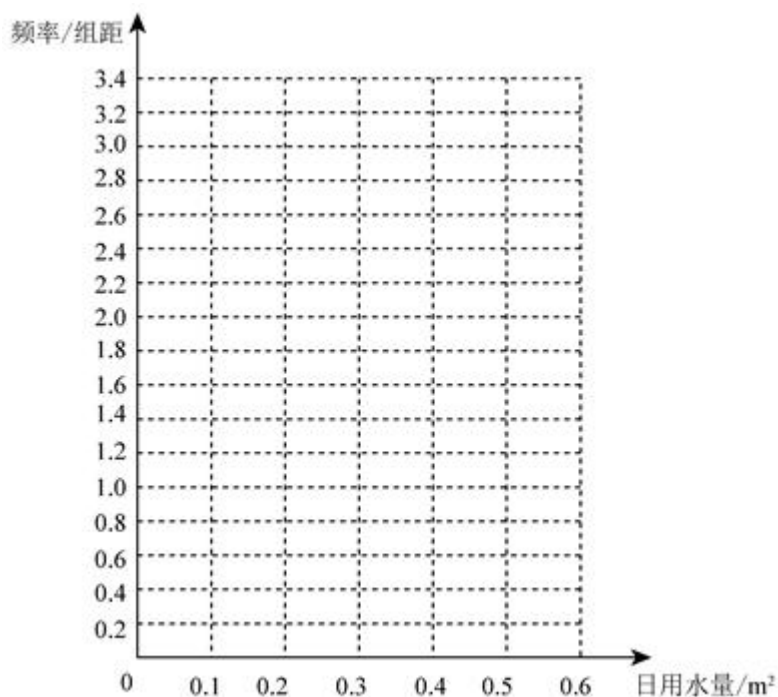
未使用节水龙头50天的日用水量频数分布表

日用水量	[0,0.1)	[0.1,0.2)	[0.2,0.3)	[0.3,0.4)	[0.4,0.5)	[0.5,0.6)	[0.6,0.7)
频数	1	3	2	4	9	26	5

使用了节水龙头50天的日用水量频数分布表

日用水量	[0,0.1)	[0.1,0.2)	[0.2,0.3)	[0.3,0.4)	[0.4,0.5)	[0.5,0.6)
频数	1	5	13	10	16	5

(1)在答题卡上作出使用了节水龙头50天的日用水量数据的频率分布直方图：



(2) 估计该家庭使用节水龙头后，日用水量小于 $0.35m^3$ 的概率；

(3) 估计该家庭使用节水龙头后，一年能节省多少水？（一年按 365 天计算，同一组中的数据以这组数据所在区间中点的值作代表.）

【答案】 (1) 直方图见解析；(2) 0.48；(3) $47.45m^3$.

【解析】

【分析】

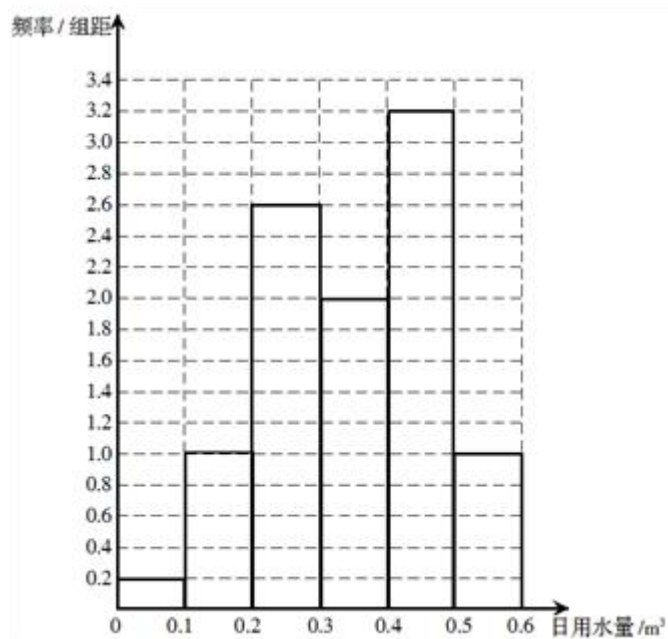
(1) 根据题中所给的使用了节水龙头 50 天的日用水量频数分布表，算出落在相应区间上的频率，借助于直方图中长方形的面积表示的就是落在相应区间上的频率，从而确定出对应矩形的高，从而得到直方图；

(2) 结合直方图，算出日用水量小于 0.35 的矩形的面积总和，即为所求的频率；

(3) 根据组中值乘以相应的频率作和求得 50 天日用水量的平均值，作差乘以 365 天得到一年能节约用水多少 m^3 ，从而求得结果.

【详解】

(1) 频率分布直方图如下图所示：



(2) 根据以上数据, 该家庭使用节水龙头后 50 天日用水量小于 $0.35m^3$ 的频率为

$$0.2 \times 0.1 + 1 \times 0.1 + 2.6 \times 0.1 + 2 \times 0.05 = 0.48;$$

因此该家庭使用节水龙头后日用水量小于 $0.35m^3$ 的概率的估计值为 0.48;

(3) 该家庭未使用节水龙头 50 天日用水量的平均数为

$$\bar{x}_1 = \frac{1}{50} (0.05 \times 1 + 0.15 \times 3 + 0.25 \times 2 + 0.35 \times 4 + 0.45 \times 9 + 0.55 \times 26 + 0.65 \times 5) = 0.48.$$

该家庭使用了节水龙头后 50 天日用水量的平均数为

$$\bar{x}_2 = \frac{1}{50} (0.05 \times 1 + 0.15 \times 5 + 0.25 \times 13 + 0.35 \times 10 + 0.45 \times 16 + 0.55 \times 5) = 0.35.$$

估计使用节水龙头后, 一年可节省水 $(0.48 - 0.35) \times 365 = 47.45 (m^3)$.

【点睛】

该题考查的是有关统计的问题, 涉及到的知识点有频率分布直方图的绘制、利用频率分布直方图计算变量落在相应区间上的概率、利用频率分布直方图求平均数, 在解题的过程中, 需要认真审题, 细心运算, 仔细求解, 就可以得出正确结果.

