

## 第21期

第3-4版同步周测参考答案

## 一、单项选择题

1.B 提示:由题意知,线性回归方

程必过点 $(\bar{X}, \bar{Y})$ ,  $\bar{X} = \frac{2+4+6+8}{4} = 5$ ,  $\bar{Y} =$  $\frac{1+3+5+7}{4} = 4$ , 故选 B.2.B 提示:依题意,  $\bar{x} = 2 + \frac{m}{5}$ ,  $\bar{y} = 9$ , 则这个样本点的中心为 $(2 + \frac{m}{5}, 9)$ , 因此,  $9 = 5(2 + \frac{m}{5}) - 7$ , 解得  $m = 6$ , 故

选 B.

3.C 提示:由题意知,  $\bar{x} = \frac{1}{5} \times (1+2+3+4+5) = 3$ ,  $\bar{y} =$  $\frac{1}{5} \times (9+a+17+b+27) = \frac{53+a+b}{5}$ ,所以样本点的中心的坐标为 $(3, \frac{53+a+b}{5})$ , 代入线性回归方程  $y = 4.5x + 3.7$ ,得  $\frac{53+a+b}{5} = 4.5 \times 3 + 3.7$ , 解得  $a+b = 33$ . 故选 C.4.C 提示:由条件可知,  $\bar{x} = \frac{-1+0+1+2}{4} = \frac{1}{2}$ ,  $\bar{y} =$  $\frac{-1+1+2+4}{4} = \frac{3}{2}$ ,所以回归直线过点 $(\frac{1}{2}, \frac{3}{2})$ , 将其代入直线  $y =$  $1.6x + \hat{a}$ , 得  $\frac{3}{2} = 1.6 \times \frac{1}{2} + \hat{a}$ , 得  $\hat{a} = 0.7$ .所以经验回归方程为  $y = 1.6x + 0.7$ , 当  $x = 5$  时,  $y = 1.6 \times$  $5 + 0.7 = 8.7$ . 故选 C.5.B 提示:由题意可得,  $\bar{x} = \frac{12}{6} = 2$ ,  $\bar{y} = \frac{18}{6} = 3$ , 因为回归直线方程为  $y = 2x + \hat{a}$ ,所以  $3 = 2 \times 2 + \hat{a}$ , 解得  $\hat{a} = -1$ . 故选 B.6.D 提示:对于 A, 由  $\hat{b} = 0.8$ , 知父亲身高每增加

1cm, 则其女儿身高大约增加 0.8cm. 故 A 正确;

对于 B, 若女儿身高为 166cm, 则其父亲身高可能

为  $(166 - 26) \div 0.8 = 175$ cm, 故 B 正确;

对于 C, D, 若父亲身高为 170cm, 则其女儿身高可

能为  $0.8 \times 170 + 26 = 162$ cm, 而不是一定, 故 C 正确, D 错误.

故选 D.

7.D 提示:因为  $\bar{x} = \frac{1}{5} \times (9+9.5+10+10.5+11) = 10$ ,  $\bar{y} =$  $\frac{1}{5} \times (11+10+8+6+5) = 8$ ,所以  $y$  与  $x$  的回归直线方程  $y = -3.2x + \hat{a}$  恒过定点 $(10, 8)$ . 所以  $8 = -3.2 \times 10 + \hat{a}$ , 解得  $\hat{a} = 40$ . 故 A 正确; 回归直线方程为  $y = -3.2x + 40$ , 即售价变量  $x$  每增加 1 个单位时, 销售变量  $y$  大约减少 3.2 个单位. 故 B 正确; 当 $x = 8.5$  时,  $y = -3.2 \times 8.5 + 40 = 12.8$ . 即当  $x = 8.5$  时,  $y$  的估计

值为 12.8, 故 C 正确. D 错误. 故选 D.

8.C 提示:由表中数据, 计算得  $\bar{x} = \frac{1}{5} \times (1+2+3+4+5) =$  $3$ ,  $\bar{y} = \frac{1}{5} \times (75+84+93+98+100) = 90$ , 所以线性回归方程 $y = 6.4x + \hat{a}$  过样本点的中心为  $(3, 90)$ , 即  $\hat{a} = 90 - 6.4 \times 3 = 70.8$ .所以  $y = 6.4x + 70.8$ , 当  $x = 38$  时,  $y = 6.4 \times 38 + 70.8 = 314$ .

即预测从 2023 年 2 月 1 日起的第 38 天到该专营店

购物的人数(单位:万人)为 314. 故选 C.

二、多项选择题

9.ABC 提示:由  $54.54 = 0.85x - 85.71$ , 得  $x = 165$ . 故A 正确; 线性回归直线过点 $(\bar{x}, \bar{y})$ , 故 B 正确; 根据线性回归方程  $y = 0.85x - 85.71$  的斜率为 0.85, 可知该高中

的女生身高增加 1cm, 其体重约增加 0.85kg, 所以 C 正

确; 线性回归方程确定之后只能用于预测, 所以 D 错

误. 故选 ABC.

10.ABD 提示:对于 A,  $\bar{x} = \frac{1}{5} \times (1+2+3+4+5) = 3$ ,  $\bar{y} =$  $\frac{1}{5} \times (4.9+5.1+5.5+5.7+5.8) = 5.4$ .因为  $y$  关于  $x$  的线性回归方程为  $y = 0.24x + \hat{a}$ , 所以 $5.4 = 0.24 \times 3 + \hat{a}$ , 解得  $\hat{a} = 4.68$ . 故 A 正确;

对于 B, 借阅量 4.9, 5.1, 5.5, 5.7, 5.8 的 75% 分位数

为第 4 个数, 即 5.7. 故 B 正确;

对于 C, 线性回归方程为  $y = 0.24x + 4.68$ , 当  $x = 6$  时, $y = 6.12$ , 故 2023 年的借阅量约为 6.12 万册, 故 C 错误;对于 D, 当  $x = 9$  时,  $y = 6.84$ . 故 2026 年的借阅量大

约是 6.84 万册, 故 D 正确.

故选 ABD.

模型拟合  $y$  与  $x$  的关系,所以  $\hat{b} = \frac{20}{10} = 2$ ,  $\hat{a} = \bar{y} - \hat{b}\bar{x} = 13 - 10 = 3$ , 则  $y$  关于  $x$  的线性回归方程为  $y = 2x + 3$ .

(2) 2x2 列联表如下所示:

	喜欢	不喜欢	总计
男	70	30	100
女	40	60	100
总计	110	90	200

计算得  $\chi^2 = \frac{200 \times (70 \times 60 - 30 \times 40)^2}{110 \times 90 \times 100 \times 100} \approx 18.182 > 6.635$ ,

所以有 99% 的把握认为“游客是否喜欢该网红景点与性别有关”.

20. 解: (1) 由 5 月至 9 月的数据可知  $\bar{t} = \frac{1}{5} \times (1+2+3+$  $4+5) = 3$ ,  $\bar{y} = \frac{1}{5} \times (3+2.4+2.2+2+1.8) = 2.28$ , $\sum_{i=1}^5 t_i y_i = 1 \times 3 + 2 \times 2.4 + 3 \times 2.2 + 4 \times 2 + 5 \times 1.8 = 31.4$ , $\sum_{i=1}^5 (t_i - \bar{t})^2 = 4 + 1 + 0 + 1 + 4 = 10$ , $\sum_{i=1}^5 (y_i - \bar{y})^2 = 0.72^2 + 0.12^2 + 0.08^2 + 0.28^2 + 0.48^2 = 0.848$ ,所以  $r = \frac{\sum_{i=1}^5 t_i y_i - 5\bar{t}\bar{y}}{\sqrt{\sum_{i=1}^5 (t_i - \bar{t})^2} \sqrt{\sum_{i=1}^5 (y_i - \bar{y})^2}} = \frac{31.4 - 5 \times 3 \times 2.28}{\sqrt{10} \times \sqrt{0.848}} =$  $\frac{-2.8}{\sqrt{8.48}} \approx -0.962$ ,因为样本相关系数的绝对值  $|r| = |-0.962| = 0.962 >$ 0.75, 所以认为  $y$  与  $t$  具有很强的线性相关关系.(2) 由题意知,  $\sum_{i=1}^5 t_i^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 = 55$ , $\hat{b} = \frac{\sum_{i=1}^5 t_i y_i - 5\bar{t}\bar{y}}{\sum_{i=1}^5 t_i^2 - 5\bar{t}^2} = \frac{31.4 - 5 \times 3 \times 2.28}{55 - 5 \times 3^2} = \frac{-2.8}{10} = -0.28$ ,所以  $\hat{a} = \bar{y} - \hat{b}\bar{t} = 2.28 - (-0.28) \times 3 = 3.12$ , 所以  $y$  关于  $t$  的经验回归方程为  $y = -0.28t + 3.12$ .当  $t = 6$  时,  $y = -0.28 \times 6 + 3.12 = 1.44$ . 因为  $1.44 < 1.5$ , 所

以 10 月收入从预测看不能突破 1.5 万元.

21. 解: (1) 由题意知, 补全 2x2 列联表如下:

	城镇居民	农村居民	合计
经常阅读	100	30	130
不经常阅读	40	30	70
合计	140	60	200

则  $\chi^2 = \frac{200 \times (100 \times 30 - 40 \times 30)^2}{140 \times 60 \times 130 \times 70} \approx 8.477 > 6.635$ ,

所以有 99% 的把握认为经常阅读与居民居住地有关.

(2) 在城镇居民 140 人中, 经常阅读的有 100 人, 不

经常阅读的有 40 人,

采取分层随机抽样抽取 7 人, 则其中经常阅读的有

 $\frac{100}{140} \times 7 = 5$  人, 不经常阅读的有 2 人, 从这 7 人中随机选取2 人作交流发言, 所有可能的情况为  $C_7^2 = 21$  (种),

其中被选中的 2 位居民都是经常阅读的居民的情

况为  $C_5^2 = 10$  (种), 所以所求概率为  $P = \frac{10}{21}$ .

22. 解: (1) 根据频率分布直方图可求得甲车间的

样本中此次参加“反诈”知识竞赛的合格率为  $0.02 \times 10 +$  $0.03 \times 10 + 0.02 \times 10 + 0.01 \times 10 = 0.8$ , 则甲车间职工此次“反

诈”知识竞赛的合格率约为 80%.

(2) 根据题中统计数据可填写 2x2 列联表如下,

	甲车间	乙车间	合计
合格人数	80	60	140
不合格人数	20	40	60
合计	100	100	200

则  $\chi^2 = \frac{200 \times (80 \times 40 - 20 \times 60)^2}{100 \times 100 \times 140 \times 60} \approx 9.524 > 6.635$ ,

所以有 99% 的把握认为此次职工“反诈”知识竞

赛的成绩与职工所在车间有关系.

(3) 由题意知, 从甲车间职工中, 抽取 1 人, 竞赛成

绩合格的概率为  $\frac{80}{100} = \frac{4}{5}$ .设这 20 名职工中, 竞赛成绩合格的人数  $X$ , 则  $X \sim B$  $(20, \frac{4}{5})$ ,所以  $P(X=k) = C_{20}^k (\frac{4}{5})^k (\frac{1}{5})^{20-k}$ , 且  $0 \leq k \leq 20$ ,  $k \in \mathbb{N}$ ,由题意, 设  $\begin{cases} P(X=k) \geq P(X=k+1), \\ P(X=k) \geq P(X=k-1), \end{cases}$ 则  $\begin{cases} C_{20}^k (\frac{4}{5})^k (\frac{1}{5})^{20-k} \geq C_{20}^{k+1} (\frac{4}{5})^{k+1} (\frac{1}{5})^{19-k}, \\ C_{20}^k (\frac{4}{5})^k (\frac{1}{5})^{20-k} \geq C_{20}^{k-1} (\frac{4}{5})^{k-1} (\frac{1}{5})^{21-k}, \end{cases}$ 化简得  $\begin{cases} k+1 \geq 4(20-k), \\ 4(21-k) \geq k, \end{cases}$  解得  $\frac{79}{5} \leq k \leq \frac{84}{5}$ , 所以  $k = 16$ .

所以竞赛成绩合格的人数最有可能是 16.

所以当施肥量为 1.8 吨时, 预计姜的产量约为 9.48 吨, 故 D 错误. 故选 ABC.

11. BD 提示: 对于 A,  $P(\text{走读生学习效率}) = \frac{10}{50} =$ 20%, 故 A 错误; 对于 B,  $P(\text{寄宿生学习效率}) = \frac{30}{50} =$ 

60%, 故 B 正确;

因为  $\chi^2 = \frac{100 \times (30 \times 40 - 10 \times 20)^2}{50 \times 50 \times 40 \times 60} \approx 16.667 > 6.635$ , 所

以有 99% 的把握认为“学生学习效率高低与晚上睡眠

是否充足有关”. 故 C 错误, D 正确. 故选 BD.

12. BC 提示: 因为在 105 人中随机抽取 1 人, 成

绩优秀的概率为  $\frac{3}{7}$ , 所以成绩优秀的人数为  $105 \times \frac{3}{7} = 45$ ,非优秀人数为  $105 - 45 = 60$ , 所以  $c = 45 - 10 = 35$ ,  $b =$  $60 - 30 = 30$ . 故 A 错误, B 正确;因为  $\chi^2 \approx \frac{105 \times (10 \times 30 - 35 \times 30)^2}{45 \times 60 \times 40 \times 65} \approx 8.413 > 6.635$ , 所以依据  $\alpha = 0.01$  的独立性检验, 能认为“成绩与班级有

关系”, 故 C 正确, D 错误. 故选 BC.

三、填空题

13. 6 提示: 因为  $\sum_{i=1}^{10} x_i = 20$ , 所以  $\bar{x} = \frac{1}{10} \sum_{i=1}^{10} x_i = 2$ , 因为回归方程为  $y = -2.2x + 5$ ,所以  $\bar{y} = -2.2 \times 2 + 5 = 0.6$ , 所以  $\sum_{i=1}^{10} y_i = 10 \times 0.6 = 6$ .14. 6.061; 95% 提示:  $\chi^2 = \frac{200 \times (90 \times 30 - 60 \times 20)^2}{150 \times 50 \times 110 \times 90} \approx$ 

6.061.

因为  $6.061 > 3.841$ , 所以有 95% 的把握认为性别与

喜欢体育课有关.

15. 30 提示: 设男生人数为  $x$ , 依题意可得列联表

如下:

	追星	不追星	合计
男生	$\frac{1}{3}x$	$\frac{2}{3}x$	$x$
女生	$\frac{1}{3}x$	$\frac{1}{6}x$	$\frac{1}{2}x$
合计	$\frac{2}{3}x$	$\frac{5}{6}x$	$\frac{3}{2}x$

计算  $\chi^2 = \frac{\frac{3}{2}x \times (\frac{1}{3}x \times \frac{1}{6}x - \frac{2}{3}x \times \frac{1}{3}x)^2}{\frac{2}{3}x \times \frac{1}{2}x \times \frac{3}{2}x \times \frac{5}{6}x} = \frac{3}{20}x > 3.841$ ,解得  $x > \frac{20 \times 3.841}{3} \approx 25.6$ , 又  $x = 6k$ ,  $k \in \mathbb{N}$ , 所以  $x_{\min} = 30$ .即根据小概率值  $\alpha = 0.05$  的独立性检验, 判断中学生追

星与性别有关, 则男生至少有 30 人.

16. 1.6; 3.65 提示: 由题意得,  $\bar{x} = \frac{1.8+2.2+2.6+3.0}{4} =$  $2.4$ ,  $\bar{y} = \frac{2.0+2.8+3.2+4.0}{4} = 3$ .因为样本中心满足回归直线方程, 可得  $3 = 2.4\hat{b} -$  $0.84$ , 解得  $\hat{b} = 1.6$ . 所以  $y = 1.6x - 0.84$ . 设 2023 年出口总额达到 5 千亿元时, 预计该年进口总额为  $x$  千亿元, 则  $5 = 1.6x - 0.84$ , 解得  $x = 3.65$ .

四、解答题

17. 解: 因为  $\sum_{i=1}^7 (t_i - \bar{t})(y_i - \bar{y}) \approx 2.89$ , $\sqrt{\sum_{i=1}^7 (y_i - \bar{y})^2} \approx 0.55$ , 计算  $\bar{t} = \frac{1}{7} \times (1+2+3+4+5+$  $6+7) = 4$ , $\sum_{i=1}^7 (t_i - \bar{t})^2 = (1-4)^2 + (2-4)^2 + (3-4)^2 + (4-4)^2 + (5-$  $4)^2 + (6-4)^2 + (7-4)^2 = 28$ ,所以  $\sqrt{\sum_{i=1}^7 (t_i - \bar{t})^2} = \sqrt{28} = 2\sqrt{7} \approx 2 \times 2.646 =$ 

5.292,

所以  $y$  与  $t$  的相关系数为  $r = \frac{2.89}{5.292 \times 0.55} \approx 0.99$ , 所以  $y$  与  $t$  的线性相关程度很强.18. 解: (1) 由  $\bar{x} = \frac{1}{5} \times (3+4+5+6+7) = 5$ ,  $\bar{y} = \frac{1}{5} \times (1+$  $1.1+1.5+1.9+2.2) = 1.54$ ,得  $\hat{b} = \frac{41.7-5 \times 5 \times 1.54}{135-5 \times 5^2} = -0.32$ ,  $\hat{a} = 1.54 - 0.32 \times 5 = -0.06$ .故  $y$  关于  $x$  的线性回归方程为  $y = 0.32x - 0.06$ .(2) 由 (1) 知, 线性回归方程为  $y = 0.32x - 0.06$ ,当  $x = 10$  时,  $y = 0.32 \times 10 - 0.06 = 3.14$ , 所以预测该月

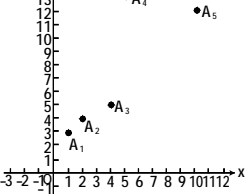
的用户数量为 3.14 万人.

19. 解: (1) 由题意知,  $\bar{x} = \frac{1}{5} \times (3+4+5+6+7) = 5$ ,  $\bar{y} = \frac{1}{5} \times$  $(10+12+11+12+20) = 13$ ,  $\sum_{i=1}^5 (x_i - \bar{x})^2 = 10$ ,  $\sum_{i=1}^5 (y_i - \bar{y})^2 = 64$ , $\sum_{i=1}^5 (x_i - \bar{x})(y_i - \bar{y}) = 20$ ,所以  $r = \frac{20}{\sqrt{10 \times 64}} = \frac{5}{\sqrt{10}} \approx 0.791$ ,因为  $r \approx 0.791$  接近于 1,所以  $y$  与  $x$  的线性相关关系很强, 可用线性回归第 24 期  
第 2-3 版节测试参考答案

## 一、单项选择题

1. A 提示: 因为相关系数  $r$  为正, 表示正相关, 回归直线方程上升,  $r$  为负, 表示负相关, 回归直线方程下降, 所以  $\hat{b}$  与  $r$  的符号相同. 故选 A.

2. D 提示: 根据题意, 散点图如图所示,



(第 2 题图)

在 5 个数据中去掉  $A_5(5, 1$

## 一、单项选择题

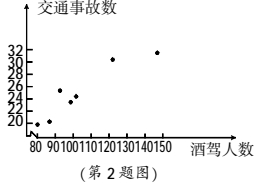
1.B 提示:对于 A,水稻产量与施肥量之间没有明确的等量关系,是相关关系,故 A 不符合题意;

对于 B,正方形的面积与边长之间有着明确的等量关系,不是相关关系,故 B 符合题意;

对于 C,商品销售收入与其广告费支出之间没有明确的等量关系,是相关关系,故 C 不符合题意;

对于 D,人体内的脂肪含量与年龄之间没有明确的等量关系,是相关关系,故 D 不符合题意.

2.A 提示:由表格中的数据,在直角坐标系中描出数据的散点图,如图所示,直观判断散点从左向右成带状分布,在一条直线附近,所以具有线性相关关系,且是正相关.故选 A.



(第 2 题图)

3.C 提示:①中,汽车的重量和汽车每消耗 1 升汽油所行驶的距离是负相关的关系;②中,平均日学习时间和平均学习成绩的关系是正相关的关系;③中,某人每日吸烟量和其身体健康情况是负相关的关系;④中,圆的半径与面积是函数关系;⑤中,汽车的重量和每千米耗油量关系是正相关的关系.所以②⑤中的两个变量成正相关.故选 C.

4.A 提示:由题意可知,这组样本数据呈完全负相关,所以其相关系数  $r=-1$ .故选 A.

5.B 提示:对于 A,数据组①的相关系数  $r_1=0$ ,无相关性,故 A 错误;

对于 B,C,4 个数据组中,数据组②的相关系数  $r_2=-0.95$ ,绝对值最大,其两变量线性相关性最强,故 B 正确.C 错误;

对于 D,由  $r_4>r_1$ ,得 D 错误.故选 B.

6.C 提示:因为相关系数  $r=0.8245>0$ ,且散点图呈左下角到右上角的带状分布,所以花瓣长度和花萼长度呈正相关.若从样本中抽取一部分,则这部分的相关系数不一定是 0.8245.故选 C.

7.B 提示:由  $T=2\pi\sqrt{\frac{l}{g}}$ ,得  $T^2=\frac{4\pi^2}{g}l$ ,故  $T^2$  与  $l$  的相关系数最大.故选 B.

8.A 提示:由给出的四组数据的散点图可以看出,图 3 和图 5 是正相关,相关系数大于 0,

图 4 和图 6 是负相关,相关系数小于 0,图 3 和图 4 的点相对更加集中,所以相关性更强,所以  $r_1$  接近于 1, $r_2$  接近于 -1,由此可得  $r_2<r_1<0<r_3<r_4$ ,故选 A.

## 二、多项选择题

9.AC 提示:通过对诗句的理解可判断选项 A 和 C 为相关关系,选项 B 不是相关关系,选项 D 中理想状态下自由下落的距离为  $s=\frac{1}{2}gt^2$ ,与速度无关,故选 D 不是相关关系,故选 AC.

10.ACD 提示:对于 A,C,变量  $x,y$  的散点图从左向右是下降的,所以  $r<0$ ,故 A,C 错误;

对于 B,变量  $x,y$  的散点图从左向右是上升的且各点不在一条直线上,所以  $0<r<1$ ,故 B 可能符合;

对于 D,变量  $x,y$  的散点图从左向右呈上升的带状分布,所以  $0<r<1$ ,故 D 错误.故选 ACD.

11.AC 提示:对于 A,样本相关系数即为其标准化数据向量夹角的余弦值,故 A 正确;

对于 B,C,样本相关系数的取值范围是  $[-1,1]$ ,故 B 错误,C 正确;

对于 D,变量  $x$  与  $y$  的线性回归方程为  $y=1.5x-2$ , $1.5>0$ ,则  $x$  与  $y$  正相关,故 D 错误.故选 AC.

12.ABD 提示:对于 A,由表中数据,最低气温大致随着最高气温的升高而升高,故该城市的各月最低温与最高温具有相关关系,且最低温度与最高温度为正相关,故 A 正确;

对于 B,由表中数据,前 8 个月每月最高温度与最低温度的平均值依次为 -3.5,3.5,6.12,20.5,23,26.5,故前 8 个月逐月增加,故 B 正确;

对于 C,由表中数据,月温差依次为 17,12,8,10,

10,7,8,7,6,11,月温差的最大值出现在 1 月,故 C 错误;

对于 D,由 C 的结论可得 1 月至 4 月的月温差相

对于 7 月至 10 月波动性更大,故 D 正确.故选 ABD.

## 三、填空题

13.有 提示:物品大小的值由小变大时,销售价格也由小变大,因此,两个变量间有相关关系.

14.甲 提示:因为线性相关系数的绝对值越大,线性相关性越强,甲、乙、丙、丁四组数据的线性相关系数分别为 -0.98,0.87,0.58,0.92,所以甲组数据的线性相关性最强.

15. $r_1>r_2$  提示:由变量 X 与 Y 相对应的一组数据为 (10,1),(11,3,2),(11,8,3),(12,5,4),(13,5),可得变量 Y 与 X 之间为正相关,所以  $r_1>0$ .而由变量 U 与 V 相对应的一组数据为 (10,5),(11,3,4),(11,8,3),(12,5,2),(13,1),可得变量 V 与 U 之间为负相关,所以  $r_2<0$ .所以  $r_1>r_2$ .

16. $\frac{8}{9}$  提示:因为  $\sum_{i=1}^{20} x_i=60$ ,  $\sum_{i=1}^{20} y_i=1200$ ,所以  $\bar{x}=3$ ,

$\bar{y}=60$ ,根据夹角公式的定义  $r=\cos\langle a,b\rangle=\frac{a\cdot b}{|a||b|}$ ,因为

$a\cdot b=\sum_{i=1}^{20} (x_i-\bar{x})(y_i-\bar{y})=\sum_{i=1}^{20} x_i y_i-20\bar{x}\bar{y}=4400-20\times 3\times 60=$

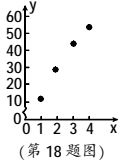
800.

所以  $r=\cos\langle a,b\rangle=\frac{a\cdot b}{|a||b|}=\frac{800}{9\times 100}=\frac{8}{9}$ .

## 四、解答题

17.解:因为“Secchi 深度”越大说明湖水越清洁,湖水被水藻污染的程度越小,而叶绿素的总浓度越大湖水被水藻污染的程度越大,所以“Secchi 深度”和叶绿素的总浓度,这两个变量间是负相关.

18.解:作出散点图如图:



(第 18 题图)

由散点图可知,各点大致分布在一条直线附近,由此推断  $x$  与  $y$  线性相关.

由题中所给表格及参考数据得:

$\bar{x}=\frac{5}{2}$ ,  $\bar{y}=\frac{69}{2}$ ,  $\sum_{i=1}^4 x_i y_i=418$ ,  $\sqrt{\sum_{i=1}^4 (y_i-\bar{y})^2} \approx 32.6$ ,

$\sum_{i=1}^4 x_i^2=30$ ,  $\sum_{i=1}^4 (x_i-\bar{x})(y_i-\bar{y})=\sum_{i=1}^4 x_i y_i-4\bar{x}\bar{y}=418-4\times \frac{5}{2}\times \frac{69}{2}=$

$\frac{69}{2}=73$ ,  $\sqrt{\sum_{i=1}^4 (x_i-\bar{x})^2}=\sqrt{\sum_{i=1}^4 x_i^2-4\bar{x}^2}=\sqrt{30-4\times (\frac{5}{2})^2}=\sqrt{5} \approx 2.24$ ,所以  $r=\frac{\sum_{i=1}^4 (x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^4 (x_i-\bar{x})^2} \sqrt{\sum_{i=1}^4 (y_i-\bar{y})^2}} \approx$

$\frac{73}{2.24\times 32.6} \approx 0.9997$ .

因为  $y$  与  $x$  的相关系数近似为 0.9997,可以推断该公司的年销量  $y$  与第  $x$  年呈正线性相关,且线性相关程度很强.

19.解:(1)由表格中的数据,可得  $\bar{x}=\frac{5+2+9+8+11}{5}=7$ ,

$\bar{y}=\frac{2+3+10+5+15}{5}=7$ ,

外卖甲的日接单量的方差

$s_1^2=\frac{(5-7)^2+(2-7)^2+(9-7)^2+(8-7)^2+(11-7)^2}{5}=10$ ,

外卖乙的日接单量的方差

$s_2^2=\frac{(2-7)^2+(3-7)^2+(10-7)^2+(5-7)^2+(15-7)^2}{5}=23.6$ ,

因为  $\bar{x}=y$ ,  $s_1^2<s_2^2$ ,即外卖甲平均日接单量与外卖乙平均日接单量相同,但外卖甲日接单量更集中一些,所以外卖甲比外卖乙经营状况更好.

(2)因为  $r=\frac{\sum_{i=1}^n (x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^n (x_i-\bar{x})^2} \sqrt{\sum_{i=1}^n (y_i-\bar{y})^2}}$ ,

又  $\sum_{i=1}^5 (x_i-\bar{x})(y_i-\bar{y})=66$ ,  $\sqrt{\sum_{i=1}^5 (x_i-\bar{x})^2} \sqrt{\sum_{i=1}^5 (y_i-\bar{y})^2} \approx 77$ ,

所以代入计算可得,相关系数  $r \approx \frac{66}{77} \approx 0.857>0.75$ ,

所以可认为  $y$  与  $x$  之间有较强的线性相关关系.

20. 解:(1)样本区域水生动物平均数为  $\frac{1}{20} \sum_{i=1}^{20} y_i=$

$\frac{1}{20} \times 1200=60$ ,小水域数为 100,所以该地区 A 种水生

动物数量的估计值为  $100 \times 60=6000$ .

(2)样本  $(x_i, y_i)$  ( $i=1,2, \cdots, 20$ ) 的相关系数为  $r=$

$\frac{\sum_{i=1}^{20} (x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{20} (x_i-\bar{x})^2} \sqrt{\sum_{i=1}^{20} (y_i-\bar{y})^2}}=\frac{1000}{\sqrt{120 \times 9000}}=\frac{5\sqrt{3}}{9} \approx$

0.96.

(3)由(2)知各样本区域的这种水生动物的数量与水草覆盖面积有很强的正相关性,由于各小水域间水草覆盖面积差异很大,从而各小水域间这种野生动物的数量差异很大,所以可以采用分层随机抽样的方法,这样能够较好地保持样本结构与总体结构的一致性,提高了样本的代表性,从而可以获得该地区这种水生动物数量更准确的估计.

21.解:(1)该食品加工厂这六个月内这种袋装食品的平均每袋出厂价格为

$\bar{x}=\frac{1}{6} \times (10.5+10.9+11+11.5+12+12.5)=11.4$ (元),

平均月销售量为  $\bar{y}=\frac{1}{6} \times (2.2+2+1.9+1.8+1.5+1.4)=$

1.8(万袋),

平均月销售收入为  $\frac{1}{6} \sum_{i=1}^6 x_i y_i=\frac{1}{6} \times 122=\frac{61}{3}$ (万元).

(2)由已知,每袋出厂价格与月销售量的样本相关

系数为  $r=\frac{\sum_{i=1}^6 x_i y_i-6\bar{x}\bar{y}}{\sqrt{\sum_{i=1}^6 x_i^2-6\bar{x}^2} \sqrt{\sum_{i=1}^6 y_i^2-6\bar{y}^2}}=\frac{122-6\times 11.4\times 1.8}{\sqrt{782.56-6\times 11.4^2} \times \sqrt{19.9-6\times 1.8^2}}=\frac{-1.12}{\sqrt{2.8\times 0.46}}=-\frac{1.12}{2\sqrt{0.7\times 0.46}}=-\frac{1.12}{2\sqrt{0.322}} \approx -0.98$ .

(3)由于每袋出厂价格与月销售量的样本相关系数  $|r| \approx 0.98>0.75$ ,

所以该食品加工厂制定的每袋食品的出厂价格与月销售量之间有较强的相关性.

22.解:(1)由表格易得,水库的平均水位  $\bar{x}=\frac{1}{10} \times$

758.01=75.801,

HN1 号渗压计管内平均水位  $\bar{y}=\frac{1}{10} \times 729.32=72.932$ ,

所以  $r=\frac{\sum_{i=1}^{10} x_i y_i-10\bar{x}\bar{y}}{\sqrt{\sum_{i=1}^{10} x_i^2-10\bar{x}^2} \sqrt{\sum_{i=1}^{10} y_i^2-10\bar{y}^2}}=\frac{55\ 283.2-10\times 75.801\times 72.932}{\sqrt{5\ 745\ 798-10\times 75.801^2} \times \sqrt{53\ 190.77-10\times 72.932^2}} \approx \frac{0.01468}{10^{-3}\sqrt{240.6}} \approx \frac{14.68}{15.51} \approx 0.95$ .

(2)因为  $\hat{b}=\frac{\sum_{i=1}^{10} x_i y_i-10\bar{x}\bar{y}}{\sum_{i=1}^{10} x_i^2-10\bar{x}^2} \approx \frac{55\ 283.2-10\times 75.801\times 72.932}{57\ 457.98-10\times 75.801^2} \approx 0.23$ ,

$\hat{a}=\bar{y}-\hat{b}\bar{x} \approx 72.932-0.23\times 75.801 \approx 55.50$ .

所以 HN1 号渗压计管内水位关于水库水位的经验回归方程为  $y=0.23x+55.50$ ,

当  $x=76$  时,预测值  $y=0.23\times 76+55.50=72.98$ ,即水库的水位为 76m 时,HN1 号渗压计管内水位的估计值为 72.98m.

## 第 23 期

## 一、单项选择题

1.C

2.C 提示:“是否吸烟”是分类变量,它的两个不同取值:吸烟和不吸烟.“是否患病”是分类变量,它的两个不同取值:患病和不患病.可知 A,B 都是一个分类变量所取的两个不同值.故选 C.

3.C 提示:由题意,得  $c=120-73-25=22$ ,  $a=74-22=52$ ,  $b=73-52=21$ ,所以  $a-b-c=52-21-22=9$ .故选 C.

4.D 提示:根据分类变量 X 与 Y 的观察数据知,  $\chi^2=3.174>2.706$ .且  $3.174<3.841$ ,

所以变量 X 与 Y 不独立,这个结论犯错误的概率不超过 10%.故选 D.

5.C 提示:因为有 95%的把握但没有 99%的把握认为偏爱蔬菜还是肉类与性别有关,

所以  $\chi^2$  的取值范围为 (3.841, 6.635),结合选项知,  $\chi^2$  的值可能为 6.625.故选 C.

6.D 提示:根据题意,填写列联表如下:

	专业 A	专业 B	合计
女生	12	4	16
男生	38	46	84
合计	50	50	100

则  $\chi^2=\frac{100\times (12\times 46-4\times 38)^2}{16\times 84\times 50\times 50}=\frac{100}{21} \approx 4.762$ , 又  $4.762>$

3.841,所以认为工科院校中“性别”与“专业”有关的把握有 95%.故选 D.

7.B 提示:因为有 90%以上的把握认为是否喜爱足球运动与学生性别有关,

所以  $\frac{160\times [(70-m)(30-m)-(10+m)(50+m)]^2}{80\times 80\times 120\times 40}$

2.706,即  $(m-10)^2>20.295$ ,因为  $y=-(m-10)^2$  在  $10\leq m\leq 20$ ,  $m\in\mathbb{N}$  时单调递增,且  $(14-10)^2<20.295$ ,  $(15-10)^2>20.295$ ,所以  $m$  的最小值为 15.故选 B.

8.C 提示:对于 A,选考物理类的学生中选择政治的比例为  $\frac{80}{220}=\frac{4}{11}$ ,选考历史类的学生中选择政治

的比例为  $\frac{50}{80}=\frac{5}{8}$ ,显然  $\frac{4}{11}<\frac{5}{8}$ ,故 A 错误;

对于 B,选考物理类的学生中选择地理的比例为  $\frac{100}{220}=\frac{5}{11}$ ,选考历史类的学生中选择地理的比例为  $\frac{45}{80}=\frac{9}{16}$ ,  $\frac{5}{11}<\frac{9}{16}$ ,故 B 错误;

对于 C,D,根据已知,可列出 2×2 列联表:

	选择生物	不选择生物	合计
物理类	115	105	220
历史类	35	45	80
合计	150	150	300

$\chi^2=\frac{300\times (115\times 45-105\times 35)^2}{150\times 150\times 220\times 80} \approx 1.705<2.706$ ,所以

根据小概率值  $\alpha=0.1$  的独立性检验,我们认为选择生物与选考类别无关,故 C 正确,D 错误.故选 C.

## 二、多项选择题

9.BD 提示:由题意知,  $\chi^2=\frac{100\times (35\times 25-25\times 15)^2}{60\times 40\times 50\times 50} \approx$

4.17,

因为  $\chi^2>3.841$ ,所以有 95%的把握认为“喜欢足球与性别相关”,故 A 错误,B 正确;

因为  $\chi^2<6.635$ ,所以没有 99%的把握认为“喜欢足球与性别有关”,故 C 错误,D 正确.故选 BD.

10.BC 提示:因为在这 105 人中随机抽取 1 人,成绩优秀的概率为  $\frac{2}{7}$ ,

所以“优秀生”的人数为  $105\times \frac{2}{7}=30$ ,“潜力生”的人数为  $105-30=75$ .

所以  $c=30-10=20$ ,  $b=75-30=45$ ,故 A 错误,B 正确;

因为  $\chi^2=\frac{105\times (10\times 30-20\times 45)^2}{30\times 75\times 50\times 55} \approx 6.109>3.841$ ,所以有 95%的把握认为成绩与班级有关,故 C 正确,D 错误.故选 BC.

11.CD 提示:由题意知,被调查的男女生人数相同,设男生的人数为  $5m$  ( $m\in\mathbb{N}_+$ ),则可列出 2×2 列联表:

	男生	女生	合计
喜欢躺着看书	2m	4m	6m
不喜欢躺着看书	3m	m	4m
合计	5m	5m	10m

$\chi^2=\frac{10m(2m\cdot m-4m\cdot 3m)^2}{6m\cdot 4m\cdot 5m\cdot 5m}=\frac{5m}{3}$ .因为有 95%的把握

认为是否喜欢躺着看书和性别有关,所以  $3.841<\frac{5m}{3}<$

6.635,解得  $11.523<5m<19.905$ ,因为  $m\in\mathbb{N}_+$ ,所以 A,B 错误,C,D 正确.故选 CD.

12.AC 提示:根据题意,喜欢该电影的男性观众有 160 人,可得  $\frac{160}{400}=\frac{2}{5}$ ,故 A 正确;

由男性不喜欢该电影的比例为  $\frac{40}{200}$ ,女性不喜欢

该电影的比例为  $\frac{60}{200}$ ,可得  $\frac{60}{200}>\frac{40}{200}$ ,故 B 错误;由  $\chi^2=\frac{400\times (160\times 60-140\times 40)^2}{300\times 100\times 200\times 200}=\frac{16}{3} \approx 5.333$ ,因为  $3.841<$

5.333<6.635,

所以根据小概率值  $\alpha=0.05$  的独立性检验,可以认为被调查观众的性别与对电影的喜爱程度有差异,故 C 正确,D 错误.故选 AC.

## 三、填空题

13.女教授人数,男教授人数,女副教授人数,男副教授人数

14.88 提示:因为志愿者的总人数为 100,所以  $a+b+12+d=100$ ,解得  $a+b+d=88$ .

15.4.844;能 提示:由表得  $\chi^2=\frac{50\times (13\times 20-7\times 10)^2}{23\times 27\times 20\times 30} \approx$

4.844>3.841,所以能在犯错误的概率不超过 0.05 的前提下认为主修统计专业与性别有关系.

16.6 提示:根据独立性检验思想可得,

$\chi^2=\frac{20\times (30\times 2-20\times 2)^2}{10\times 10\times 9\times 11\times}=\frac{20\times}{99}>2.706$ ,

解得  $x>13.3947$ ,因为  $x\in\mathbb{N}$ ,且  $x<20$ ,所以  $x=14$  或  $x=15$  或  $x=16$  或  $x=17$  或  $x=18$  或  $x=19$ .

所以  $x$  的所有可能取值个数是 6.

## 四、解答题

17.解:(1)依题意,补全列联表如下:

	休闲方式	看手机	看书	合计
性别				
男		10	50	60
女		10	10	20
合计		20	60	80

用频率估计概率,可得以看书为休闲方式的男性的概率为  $P=\frac{50}{60}=\frac{5}{6}$ ,

则随机调查 3 名在该社区的男性,记“这 3 人中至少有 1 人是以看书为休闲方式的”为事件 A,

则事件 A 的对立事件为“这 3 人中没有 1 人是以看书为休闲方式”,其概率为  $(1-P)^3=(1-\frac{5}{6})^3=\frac{1}{216}$ .

所以这 3 人中至少有 1 人是以看书为休闲方式的概率为  $P(A)=1-\frac{1}{216}=\frac{215}{216}$ .

(2)由(1)中的列联表可得

$\chi^2=\frac{80\times (10\times 10-50\times 10)^2}{60\times 20\times 20\times 60} \approx 8.889>6.635$ ,

所以我们有 99%的把握认为“在晚上 8 点至 10 点时间段的休闲方式与性别有关系”.

18.解:(1)众数约为  $\frac{30+40}{2}=35$ ,

平均数约为  $5\times 0.1+15\times 0.18+25\times 0.22+35\times 0.25$