



过马路”的路人的概率是  $\frac{3}{5}$ , 即  $\frac{10+b}{30} = \frac{3}{5}$ , 解得  $b=8$ .

又  $10+b+c=30$ , 解得  $c=4$ .

由列联表填写列联表得表1:

表1: 列联表，包含反感、不反感、合计，以及男性、女性、合计。

(2)由列联表得  $K^2 = \frac{30 \times (10 \times 8 - 4 \times 8)^2}{14 \times 16 \times 18 \times 12} \approx 1.429 < 3.841$ .

所以没有95%的把握认为反感“中国式过马路”与性别有关.

10.解:(1)由已知,得2x2列联表如表2:

表2: 列联表，包含优分、非优分、合计，以及男生、女生、合计。

(2)  $K^2 = \frac{100 \times (15 \times 25 - 15 \times 45)^2}{60 \times 40 \times 30 \times 70} \approx 1.786$ .

因为  $1.786 < 2.706$ , 所以没有90%以上的把握认为“数学成绩与性别有关”.

同步分层能力测试题(二) 基础练

- 一、选择题 1~6.CCBBDC. 提示: 1.正确的为①③④.故选C. 2.等高条形图展示的是列联表数据的频率特征,条形图的高表示分类变量的频率.故选C. 3.∵P(K^2 ≥ 6.635) = 0.010, P(K^2 ≥ 7.879) = 0.005, 又把握性超过99%但不超过99.5%, ∴6.635 ≤ K^2 ≤ 7.879.故选B. 4.统计的结果只是说明事件发生可能性的大小,具体到一个个不一定发生.故选D. 5.根据独立性检验知|ad-bc|越大,X与Y之间的关系越强,即(ad-bc)^2越大,X与Y的关系越强.故选C. 6.由表格中的数据可得K^2 = 110 \* (10 \* 30 - 20 \* 50)^2 / (60 \* 50 \* 30 \* 80) ≈ 7.486, 所以P(K^2 ≥ 6.635) = 0.01, 因此有99%的把握认为“成绩与班级有关系”.故选C. 二、填空题 7.①; 8.47, 92, 88, 82, 53. 提示: 7.根据已知能判断有95%的把握认为“这种血清能起到预防感冒的作用”,所以选项①是正确的;如果某人未使用该血清,不能说明他在一年中有95%的可能性得感冒,所以选项②是错误的;这种血清预防感冒的有效率为95%,不是5%,所以选项③是错误的.故答案为:①. 三、解答题 9.解:因为K^2的观测值k = 90 \* (10 \* 38 - 35 \* 7)^2 / (45 \* 45 \* 17 \* 73) ≈ 0.653 < 2.706, 所以在犯错误的概率不超过0.1的前提下不能判定学生的成绩与班级有关. 10.解:由K^2 = n(ad-bc)^2 / ((a+b)(c+d)(a+c)(b+d)),

得k = 160 \* (19 \* 20 - 7 \* 6 \* 45)^2 / (95 \* 65 \* 64 \* 96) ≈ 38.974.

因为38.974 > 10.828, 所以在犯错误的概率不超过0.001的前提下认为运动员受伤与不做热身运动有关.

11.解:(1)K^2的观测值k = 830 \* (52 \* 218 - 466 \* 94)^2 / (146 \* 684 \* 518 \* 312) ≈ 54.21, ∴54.21 > 10.828,

因此在出错的概率不超过0.001的前提下,我们认为该地区这种传染病与饮用不干净水有关.

(2)依题意得2x2列联表(如表1):

表1: 列联表，包含得病、不得病、总计，以及干净水、不干净水、总计。

此时, K^2的观测值k = 86 \* (5 \* 22 - 50 \* 9)^2 / (14 \* 72 \* 55 \* 31) ≈ 5.785.

由于5.785 > 2.706, 所以在出错的概率不超过0.1的前提下,我们认为该种疾病与饮用不干净水有关.

两个样本都能统计得到传染病与饮用不干净水有关这一相同结论,但(1)中我们在出错的概率不超过0.001的前提下肯定结论的正确性,(2)中我们在出错的概率不超过0.1的前提下肯定.

12.解:由条件可得2x2的列联表如表2:

表2: 列联表，包含年轻人、非年轻人、总计，以及经常使用微信支付、不常使用微信支付、合计。

∴K^2 = 300 \* (165 \* 30 - 45 \* 60)^2 / (210 \* 90 \* 75 \* 225) = 100 / 21 ≈ 4.762 > 3.841,

∴有95%的把握认为“使用微信支付与年龄有关”.

提升练

- 一、选择题 1~6.AAABBC. 提示: 1.充分利用独立性检验的意义来分析. ①符合事件独立的意义,故正确; ②中的关系密切不等于相关,所以不对; ③不是唯一根据,还有独立事件的判断、假设检验等方法; ④独立性检验推断的是相关的把握,有犯错误的可能.故选A. 2.由题意可得,当a/a+10与c/c+30相差越大,X与Y有关系的可能性最大,分析四组选项,A中的a,c的值最符合题意,故选A. 3.代入公式计算可得k ≈ 16.373.故选A. 4.由题意,根据调查数据计算得到k^2 ≈ 15.968, 因为P(K^2 ≥ 10.828) = 0.001, 所以这种判断犯错误的概率不超过0.001.故选B. 5.由X,Y独立,得ad=bc,由此可得A,C,D正确.故选B. 6.由等高条形图可知,女生中喜欢理科的百分比约为1-0.8=0.2=20%,男生中喜欢理科的百分比约为1-0.4=0.6=60%,因此男生比女生喜欢理科的可能性大些.故选C. 二、填空题 7.大于等于6.635; 8.28.

提示: 7.查表可知观测值k ≥ 6.635时满足题意. 8.由2x2列联表得a+6=18,所以a=12. 因为a+b=20,所以b=8.因为6+d=30,所以d=24,所以a-b+d=12-8+24=28.

故答案为:28.

三、解答题

9.解:由题意,可得到表3.

表3: 列联表，包含药物处理、未经过药物处理、总计，以及青花病、无青花病、总计。

根据公式得k = 470 \* (25 \* 200 - 185 \* 60)^2 / (210 \* 260 \* 85 \* 385) ≈ 9.788.

由于9.788 > 7.879, 所以在犯错误的概率不超过0.005的前提下认为经过药物处理跟发生青花病是有关系的.

10.解:(1)由题意,样本中喜欢的有6/60 \* 40 = 4(名),不喜欢的有6/60 \* 20 = 2(名).

(2)∴k = 140 \* (40 \* 20 - 20 \* 60)^2 / (60 \* 80 \* 100 \* 40) ≈ 1.167 < 5.024, 所以不能在犯错误的概率不超过0.025的前提下认为观众性别与喜爱乐嘉有关.

(3)设男性观众中喜欢乐嘉的为1,2,3,4,不喜欢的为a,b,则从6名观众中选2名共有:12,13,14,1a,1b,23,24,2a,2b,3a,3b,4a,4b,4c,都喜欢的有6种,所以选到的2名观众都喜欢乐嘉的概率是6/15 = 2/5.

第一章能力测试题

- 一、选择题 1~6.AACBBB; 7~12.ABBCBA. 提示: 1.选项B,C,D中都是一种不确定的关系,是相关关系,而函数关系是一种确定的关系,A中判别式Δ和变量b是一种确定的表达式,所以是一种函数关系.故选A. 2.由样本点的中心(x̄,ȳ)落在回归直线上,A正确;x和y的相关系数表示x与y之间的线性相关程度,不表示直线l的斜率,故B错;x和y的相关系数应在-1到1之间,故C错;分布在回归直线两侧的样本点的个数并不绝对平均,且无论样本点个数是奇数还是偶数,故D错.故选A. 3.四个模型的相关系数R^2分别为0.25,0.50,0.98,0.80,相关系数的绝对值越靠近1,拟合效果越好.故选C. 4.根据表中数据,得x̄ = 1/8 \* (-10 - 6.99 - 5.01 - 2.98 + 3.98 + 5 + 7.99 + 8.01) = 0, ȳ = 1/8 \* (-9 - 7 - 5 - 3 + 4.01 + 4.99 + 7 + 8) = 0, 由两变量x,y的线性回归方程过样本点中心(0,0),可以排除A,C,D选项.故选B. 5.根据K^2的意义可知选B. 6.散点图里变量的对应点分布在一条直线附近,且比较密集,故可判断语文成绩和英语成绩之间具有较强的线性相关关系,且直线斜率小于1.故选B. 7.由已知数据得如表1所示的2x2列联表: 表1: 列联表，包含杂质高、杂质低、总计，以及旧设备、新设备、总计。

由公式,得k = 382 \* (37 \* 202 - 121 \* 22)^2 / (158 \* 224 \* 59 \* 323) ≈ 13.11.

由于13.11 > 10.828, 故在犯错误的概率不超过0.001的前提下含杂质的低与设备是否改造是有关的. 故选A.

8. x̄ = (2.2+2.6+4.0+5.3+5.9)/5 = 4, ȳ = (0.2+1.5+2.0+2.5+3.8)/5 = 2.2

2.代入回归直线方程,得2=0.8x+â,解得â=-1.2,所以回归直线方程为ŷ=0.8x-1.2.当x=7时,支出为4.4亿元.故选B.

9. K^2 = n(ad-bc)^2 / ((a+b)(c+d)(a+c)(b+d)) = 90 \* (380 - 245)^2 / (17 \* 73 \* 45 \* 45) ≈ 0.653.

∴0.455 < 0.653 < 0.708, ∴P ∈ (0.4, 0.5), 则有错误的概率介于0.4~0.5之间.故选B.

10.由题意可得关于商品和服务评价的2x2列联表如表2:

表2: 列联表，包含对服务好、对服务不满意、总计，以及对商品好评、对商品不满意、总计。

根据公式得k = 200 \* (80 \* 10 - 40 \* 70)^2 / (150 \* 50 \* 120 \* 80) ≈ 11.111 > 10.828, 可以在犯错误的概率不超过0.1的前提下,认为商品好评与服务好评有关.故选C.

11. x̄\_0, y\_0为这10组数据的平均值,又因为回归直线ŷ = b̂x + â必过样本中心点(x̄,ȳ),因此(x\_0, y\_0)一定满足线性回归方程,但坐标准足线性回归方程的点不一定是(x̄,ȳ),故选B.

12.由2x2列联表知c+d=35,即d=35-c.故K^2的观测值k = 66 \* [10(35-c) - 21c]^2 / (31 \* 35 \* (10+c) \* (56-c)) ≥ 5.024.

将选项A,B,C,D代入验证可知应选A.

二、填空题

13.体重预测值=0.72 \* 身高 - 58.5; 14.甲;

15.2.45; 16.②③.

提示: 13.4磅/英寸=4 \* (0.45kg/2.5cm) = 0.72(kg/cm), 130磅=130 \* 0.45kg = 58.5kg.

14.可以根据表中数据分析两个回归方程对数据预测的准确率进行判断,甲回归方程的数据准确率为32/40 = 4/5, 而乙回归方程的数据准确率为40/60 = 2/3, 显然甲的准确率更高些,因此甲回归方程好些.故答案:甲.

15.当x=165时,ȳ=0.85 \* 165 - 85.7 = 54.55, 所以方程在样本(165, 57)处的残差是57 - 54.55 = 2.45.

16.①在线性回归模型中,相关指数R^2表示解释变量x对于预报变量y的贡献率,R^2越接近于1,表示回归效果越好,所以①错误;

②在回归直线方程ŷ=0.8x-12中,当解释变量x每增加一个单位时,预报变量平均增加0.8个单位.正确;

③两个变量相关性越强,则相关系数的绝对值就越接近于1.正确;

④对分类变量X与Y,对它们的随机变量K^2的观测值k来说,k越小,则“X与Y有关系”的把握程度越小,所以④错误,故正确命题的序号是②③.

三、解答题

17.解:等高条形图如图1所示,由图形观察可以看出父母吸烟者中子女吸烟的比例要比父母不吸烟者中子女

吸烟的比例高,因此可以在某种程度上认为“子女吸烟与父母吸烟有关”.

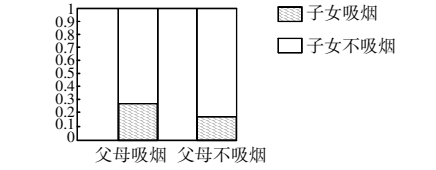


图1

18.解:(1)散点图如图2所示,从图中可以看出这些点大致分布在一条直线附近,因此两个变量具有线性相关关系.

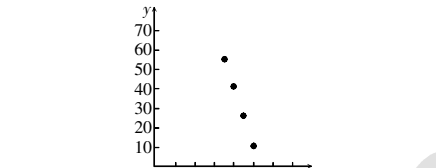


图2

设线性回归方程为ŷ = b̂x + â, 由题知x̄ = 42.5, ȳ = 34, 则求得b̂ = (∑(x\_i - x̄)(y\_i - ȳ)) / (∑(x\_i - x̄)^2) = -370 / 125 ≈ -3.

∴â = ȳ - b̂x̄ = 34 - (-3) \* 42.5 = 161.5.

∴ŷ = -3x + 161.5.

(2)依题意有P = (-3x + 161.5)(x - 30) = -3x^2 + 251.5x - 4845.

= -3(x - 251.5/6)^2 + 251.5^2/12 - 4845,

∴当x = 251.5/6 ≈ 42时,P有最大值,

即预测销售单价为42元时,能获得最大日销售利润.

19.解:(1)作出散点图(如图3),从图中可看出x与y具有相关关系.

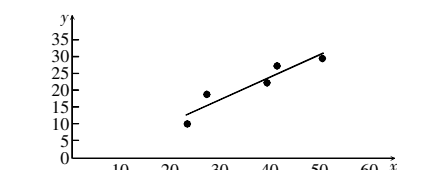


图3

(2)由已知,可得y对x的回归直线方程为ŷ = 0.6512x - 2.72.

当x=23时,ŷ = 12.2576, y - ŷ = 9.5 - 12.2576 = -2.7576;

当x=50时,ŷ = 29.84, y - ŷ = 28.2 - 29.84 = -1.64. 所以23岁和50岁的残差分别为-2.7576和-1.64.

20.解:(1)由折线图中数据和参考数据得t̄ = 4, ∑(t\_i - t̄)^2 = 28, √(∑(y\_i - ȳ)^2) = 0.55, √7 ≈ 2.646, ∑(t\_i - t̄)(y\_i - ȳ) = 2.89, r = 2.89 / (0.55 \* 2 \* 2.646) ≈ 0.99.

因为y与t的相关系数近似为0.99,说明y与t的线性相关程度相当高,从而可以用线性回归模型拟合y与t的关系.

(2)由ȳ = 9.32/7 ≈ 1.331及(1)得b̂ = 2.89/28 ≈ 0.103, â = ȳ - b̂t̄ ≈ 1.331 - 0.103 \* 4 ≈ 0.92.

所以y关于t的回归方程为y = 0.92 + 0.10t. 将2020年对应的t=13代入回归方程得ŷ = 0.92 + 0.10 \* 13 = 2.22.

所以预测2020年我国生活垃圾无害化处理量将约为2.22亿吨.

21.解:(1)列联表如表3. 表3: 列联表，包含关注、不关注、合计，以及年轻人、中老年人、合计。

其中a=10, b=30, c=40, d=20代入公式得K^2 ≈ 16.67 > 6.635, 故有99%的把握认为关注“中国湖北(潜江)龙虾节”和年龄段有关.

(2)抽取的6位中老年人中有4人关注,2人不关注,设事件“选取的3人中恰有2人关注‘中国湖北(潜江)龙虾节’”为事件A,记关注的四人为A\_1, A\_2, A\_3, A\_4,记不关注的两人分别为B\_1, B\_2, 从这6人中选3人的选法有(A\_1, A\_2, A\_3), (A\_1, A\_2, A\_4), (A\_1, A\_2, B\_1), (A\_1, A\_2, B\_2), (A\_1, A\_3, A\_4), (A\_1, A\_3, B\_1), (A\_1, A\_3, B\_2), (A\_1, A\_4, B\_1), (A\_1, A\_4, B\_2), (A\_2, A\_3, A\_4), (A\_2, A\_3, B\_1), (A\_2, A\_3, B\_2), (A\_2, A\_4, B\_1), (A\_2, A\_4, B\_2), (A\_3, A\_4, B\_1), (A\_3, A\_4, B\_2), (A\_1, B\_1, B\_2), (A\_2, B\_1, B\_2), (A\_3, B\_1, B\_2), (A\_1, B\_1, B\_2), 共20种,

其中有12种情况满足题意,故P(A) = 12/20 = 3/5.

22.解:(1)第一组的人数为120/0.6 = 200, 频率为0.04 \* 5 = 0.2, 所以n = 200/0.2 = 1000. 由题可知,第二组的频率为1 - (0.04 \* 2 + 0.03 + 0.02 + 0.01) \* 5 = 1 - 0.7 = 0.3, 所以第二组的人数为1000 \* 0.3 = 300(人), 所以p = 195/300 = 0.65. 又第四组的频率为0.03 \* 5 = 0.15, 所以第四组的人数为1000 \* 0.15 = 150(人), 所以a = 150 \* 0.4 = 60. 频率分布直方图如图4.



图4

(2)由已知数据,完成表格如表4. 表4: 列联表，包含是否低碳族、是否非低碳族、总计，以及青年、老年、总计。

代入公式得k = 1000 \* (415 \* 195 - 105 \* 285)^2 / (520 \* 700 \* 300 \* 480) ≈ 49.622 > 10.828. 所以在犯错误的概率不超过0.001的前提下能判断该地区的生活习惯是否符合低碳观念与人的年龄有关.

第28期《堂堂清》参考答案 §2.1.1 合情推理 1.B. 2.B.