

回归直线独立性检验综合

一、单选题

1. 下列说法中，错误的有（ ）

- ①回归直线 $\hat{y} = \hat{b}x + \hat{a}$ 恒过点 $(\bar{x}, \bar{y})$ ，且至少过一个样本点；
- ②根据 2×2 列联表中的数据计算得出 $\chi^2 \geq 6.635$ ，而 $P(\chi^2 \geq 6.635) \approx 0.01$ ，则有 99% 的把握认为两个分类变量有关系，即有 1% 的可能性使得“两个分类变量有关系”的推断出现错误；
- ③在做回归分析时，可以用决定系数 R^2 刻画模型的回归效果，若 R^2 越大，则说明模型拟合的效果越好；
- ④某项测量结果 ξ 服从正态分布 $N(1, \sigma^2)$ ，若 $P(\xi \leq 5) = 0.81$ ，则 $P(\xi \leq -3) = 0.19$

A. 1 个 B. 2 个 C. 3 个 D. 4 个

2. 以下四个命题中错误的是（ ）

- A. 在独立性检验中，由计算得 χ^2 的值，若 χ^2 的值越大，则两个变量相关的可能性就越大
- B. 在做回归分析时，可以用决定系数 R^2 刻画模型的回归效果，若 R^2 越大，则说明模型拟合的效果越好
- C. 在回归直线方程 $\hat{y} = 2 - 3x$ 中，变量 x 每增加 1 个单位时， y 平均增加 2 个单位
- D. 若变量 y 和 x 之间的相关系数为 $r = -0.99$ ，则变量 y 和 x 之间具有很强的线性相关性，而且是负相关

3. 以下结论错误的是（ ）

- A. 根据 2×2 列联表中的数据计算得出 $\chi^2 \geq 6.635$ ，而 $P(\chi^2 \geq 6.635) \approx 0.01$ ，则根据小概率值 $\alpha = 0.01$ 的独立性检验，认为两个分类变量有关系
- B. 在回归直线 $\hat{y} = 0.5x - 85$ 中，变量 $x = 200$ 时，变量 y 的值一定是 15
- C. χ^2 的值越大，两个事件的相关性的可能性就越大
- D. 在回归分析中，相关指数 R^2 越大，说明残差平方和越小，回归效果越好

4. 下列命题错误的是（ ）

- A. 设 $\xi \sim B(n, p)$ ，若 $E(\xi) = 30$ ， $D(\xi) = 20$ ，则 $n = 90$
- B. 独立性检验是对两个变量是否具有线性相关关系的一种检验

C. 两个随机变量的线性相关性越强，相关系数的绝对值越接近于 1

D. 一个袋子中有 100 个大小相同的球，其中有 40 个黄球、60 个白球，从中不放回地随机摸出 20 个球作为样本，用随机变量 X 表示样本中的黄球个数，则 X 服从超几何分布，且 $E(X)=8$

5. 在性别与吃零食这两个分类变量的计算中，下列说法正确的是（ ）

①若 χ^2 的观测值为 $k=6.635$ ，我们有 99% 的把握认为吃零食与性别有关系，那么在 100 个吃零食的人中必有 99 人是女性；

②从独立性检验可知有 99% 的把握认为吃零食与性别有关系时，我们说某人吃零食，那么此人是女性的可能性为 99%；

③若从统计量中求出有 99% 的把握认为吃零食与性别有关系，是指有 1% 的可能性使得出的判断出现错误.

A. ①②

B. ①③

C. ②③

D. ③

6. 为了验证牛的毛色（黑色、红色）和角（有角、无角）这两对相对性状是否相关，某学院进行了一次数据统计，根据形成的 2×2 列联表，计算得到 $\chi^2 \approx 2.786$ ，根据小概率值

$\alpha=0.05$ 的独立性检验（已知 χ^2 独立性检验中 $x_{0.05}=3.841$ ），下列结论正确的是（ ）

A. 牛的毛色与角无关

B. 牛的毛色与角无关，此推断犯错误的概率不超过 0.05

C. 牛的毛色与角有关

D. 牛的毛色与角有关，此推断犯错误的概率不超过 0.05

7. 下列说法正确的个数为（ ）

①根据具有线性相关关系的两个变量的统计数据所得的回归直线方程为 $\hat{y}=\hat{b}x+\hat{a}$ ，若 $\hat{b}=2$ ， $\bar{x}=1$ ， $\bar{y}=3$ ，则 $\hat{a}=1$ ；

②分类变量 A 与 B 的统计量 χ^2 越大，说明“ A 与 B 有关系”的可信度越小；

③一元线性回归模型中，如果相关系数 $r=0.98$ ，表明两个变量的相关程度很强.

A. 0

B. 1

C. 2

D. 3

8. 下列结论中，正确的选项个数是（ ）

(1)对具有线性相关关系的变量 x, y ，且回归方程为 $y=0.3x-m$ ，若样本点的中心为 $(m, 2.8)$ ，则实数 m 的值是 -4

(2) 若随机变量 $\xi \sim N(1, \sigma^2)$, $P(\xi \leq -2) = 0.21$, 则 $P(\xi \leq 4) = 0.79$

(3) 若随机变量 ξ, η , 满足 $\eta = 3\xi - 2$, 则 $E(\eta) = 3E(\xi) - 2$, $D(\eta) = 9D(\xi) - 4$

(4) 根据分类变量 X 与 Y 成对样本数据, 计算得到 $\chi^2 = 9.632$, 依据小概率值 $\alpha = 0.001$ 的 χ^2 独立性检验 ($x_{0.001} = 10.828$), 可判断 X 与 Y 有关联, 此推断犯错误的概率不大于 0.001

A. 1 个 B. 2 个 C. 3 个 D. 4 个

9. 下列说法中, 正确的个数是 ()

(1) 两个随机变量的线性相关程度越强, 相关系数的绝对值越接近于 1

(2) 两个变量的 2×2 列联表中, 对角线上数据的乘积相差越大, 说明两个变量有关系成立的可能性就越大

(3) 从独立性检验可知: 有 99% 的把握认为吸烟与患肺癌有关, 是指在犯错误的概率不超过 1% 的前提下认为吸烟与患肺癌有关

A. 0 个 B. 1 个 C. 2 个 D. 3 个

10. 下列说法中, 不正确的是 ()

A. 在 1, 3, 6, 7, 9, 10, 12, 15 这组数据中, 第 50 百分位数为 8

B. 分类变量 A 与 B 的统计量 χ^2 越大, 说明“ A 与 B 有关系”的可信度越大

C. 已知经验回归方程为 $y = \hat{b}x + 1.8$ 且 $\bar{x} = 2, \bar{y} = 20$, 则 $\hat{b} = 9.1$

D. 两个模型中, 残差平方和越大的模型拟合的效果越好

11. 下列命题中错误的是 ()

A. 在回归分析中, 相关系数 r 的绝对值越大, 两个变量的线性相关性越强

B. 若变量 y 与 x 之间存在线性相关关系, 且根据最小二乘法得到的经验回归方程为

$\hat{y} = 2x + \hat{a}$, 样本点中心为 $(3, 6.5)$, 则样本点 $(2.5, 7)$ 的残差为 1.5

C. 在回归分析中, 残差平方和越小, 模型的拟合效果越好

D. 对分类变量 X 与 Y , 它们的随机变量 K^2 的观测值 k 越小, 说明“ X 与 Y 有关系”的把握越大

12. 根据吸烟与患肺癌这两个分类变量的样本数据, 计算得出 $\chi^2 \approx 6.816$, 经查阅 χ^2 独立性检验的小概率值和相应的临界值 $x_{0.01} = 6.635$, 则下列说法正确的是 ()

A. 在 100 个吸烟的人中就会有 99 人患肺癌

- B. 若某人吸烟，那么他有 99%的可能患肺癌
- C. 若某人患肺癌，那么他有 99%的可能为吸烟者
- D. 吸烟与患肺癌有关联，此推断犯错误的概率不大于 1%

13. 下列实际问题不适合用独立性检验解决的是 ()

- A. 不良的饮食习惯是否会导致肠胃疾病
- B. 某公司的营业额在过去 5 年逐年变化的情况
- C. 参加课外辅导能否提高学习成绩
- D. 男性和女性在职业选择偏好上是否有差异

14. 下列说法中，错误的命题是 ()

- A. 在刻画回归模型的拟合效果时， R^2 的值越大，说明拟合的效果越好
- B. 线性相关系数 r 越大，两个变量的线性相关性越强，反之，线性相关性越弱
- C. 设随机变量 X 服从正态分布 $N(4, 2^2)$ ，则 $P(X > 4) = \frac{1}{2}$
- D. 对分类变量 X 与 Y ，若计算出的 χ^2 越大，则判断“ X 与 Y 有关系”的犯错误的概率越小

15. 下列说法中，正确的有 ()

- ①回归直线 $\hat{y} = \hat{b}x + \hat{a}$ 恒过点 $(\bar{x}, \bar{y})$ ，且至少过一个样本点；
- ②根据 2×2 列联表中的数据计算得出 $\kappa^2 \geq 6.635$ ，而 $P(\kappa^2 \geq 6.635) \approx 0.01$ ，则有 99% 的把握认为两个分类变量有关系，即有 1% 的可能性使得“两个分类变量有关系”的推断出现错误；
- ③ k^2 是用来判断两个分类变量是否相关的随机变量，当 k^2 的值很小时可以推断两类变量不相关；
- ④某项测量结果 ξ 服从正态分布 $N(1, a^2)$ ，若 $P(\xi \leq 5) = 0.81$ ，则 $P(\xi \leq -3) = 0.19$.

- A. 1 个
- B. 2 个
- C. 3 个
- D. 4 个

16. 为了研究高中学生对乡村音乐的态度 (喜欢和不喜欢两种态度) 与性别的关系，运用 2×2 列联表进行独立性检验，经计算得 $\chi^2 = 7.01$ ，则认为“喜欢乡村音乐与性别有关系”的把握约为 ()

- A. 0.1%
- B. 1%
- C. 99%
- D. 99.9%

17. 在对吸烟与患肺病这两个分类变量的独立性检验中，下列说法正确的是 (参考数据：

$$P(K^2 \geq 6.635) = 0.01) ()$$

- ①若 K^2 的观测值满足 $K^2 \geq 6.635$ ，我们有 99% 的把握认为吸烟与患肺病有关系；
- ②若 K^2 的观测值满足 $K^2 \geq 6.635$ ，那么在 100 个吸烟的人中约有 99 人患有肺病；
- ③从独立性检验可知，如果有 99% 的把握认为吸烟与患肺病有关系时，那么我们就认为：每个吸烟的人有 99% 的可能性会患肺病；
- ④从统计量中得知有 99% 的把握认为吸烟与患肺病有关系时，是指有 1% 的可能性使推断出现错误。

- A. ②③ B. ②③④ C. ①②④ D. ①④

18. 下列说法错误的是 ()

- A. 线性相关系数 $|r|$ 越接近 1，两个变量的线性相关程度越强；
- B. 独立性检验可以 100% 确定两个变量之间是否具有某种关系；
- C. 在回归分析中，可用残差图判断模型的拟合效果，残差点比较均匀地落在水平的带状区域中，说明这样的模型比较合适，带状区域的宽度越窄，说明模型的拟合精度越高；
- D. 甲、乙两个模型的决定系数 R^2 分别约为 0.88 和 0.80，则模型甲的拟合效果更好。

19. 下列结论错误的是 ()

- A. 若随机变量 ξ, η 满足 $\eta = 2\xi + 1$ ，则 $D(\eta) = 4D(\xi)$
- B. 数据 1, 3, 5, 7, 9, 11, 13 的第 60 百分位数为 9
- C. 用简单随机抽样的方法从 51 个个体中抽取 2 个个体，则每个个体被抽到的概率都是 $\frac{1}{51}$
- D. 根据分类变量 X 与 Y 的成对样本数据，计算得到 $\chi^2 = 4.712$. 依据 $\alpha = 0.05$ 的独立性检验 ($\chi_{0.05}^2 = 3.841$)，可判断 X 与 Y 有关

20. 下列说法正确的个数是 ()

- ①线性相关系数 $|r|$ 越接近 1，两个变量的线性相关程度越强；
- ②独立性检验可以 100% 确定两个变量之间是否具有某种关系；
- ③在回归分析中，可用残差图判断模型的拟合效果，残差点比较均匀地落在水平的带状区域中，说明这样的模型比较合适，带状区域的宽度越窄，说明模型的拟合精度越高；
- ④甲、乙两个模型的决定系数 R^2 分别约为 0.88 和 0.80，则模型甲的拟合效果更好。

- A. 1 B. 2 C. 3 D. 4