

山东大学

二〇一七年招收攻读硕士学位研究生入学考试试题

科目代码 432

科目名称 统计学

(答案必须写在答卷纸上, 写在试题上无效)

一、单项选择题 (本题包括 1—30 题共 30 个小题, 每小题 2 分, 共 60 分)

1. 以下那一种情形涉及定型数据的收集 ()?
- A. 质量控制工程师测量灯泡的寿命
B. 社会学家通过抽样调查来估计广州市市民的平均年收入
C. 运动器材厂家在区分各大俱乐部棒球选手是左撇子还是右撇子时作的调查
D. 婚礼策划公司通过抽样调查来估计上海市市民举办婚礼的平均开销
2. 为了调查某校学生的购书费用支出, 将全校学生的名单按拼音顺序排列后, 每隔 50 名学生抽取一名学生进行调查, 这种调查方法是 ()。
- A. 简单随机抽样 B. 整群抽样 C. 系统抽样 D. 分层抽样
3. 一家公司的人力资源部主管需要研究公司雇员的饮食习惯, 改善公司餐厅的现状。他将问卷发给就餐者, 填写后再收上来。他的数据收集方法是 ()。
- A. 自填式问卷调查 B. 面访式问卷调查 C. 实验调查 D. 观察式调查
4. 下面的哪一个图形适合于比较研究两个或多个样本或总体的结构性问题 ()。
- A. 环形图 B. 饼图 C. 直方图 D. 茎叶图
5. 经验法则表明, 当一组数据对称分布时, 在平均数加减 2 个标准差的范围之内大约有 ()。
- A. 68% 的数据 B. 95% 的数据 C. 99% 的数据 D. 100% 的数据
6. 样本 X_1, \dots, X_n 取自标准正态分布总体 $N(0,1)$, \bar{X} 与 S 分别是样本均值和样本标准差, 则 ()。
- A. $\bar{X} \sim N(0,1)$ B. $n\bar{X} \sim N(0,1)$ C. $\frac{1}{n} \sum_{i=1}^n X_i^2 \sim \chi^2(n)$ D. $\frac{\bar{X}}{S} \sim t(n-1)$
7. 设总体 X 服从参数为 $\lambda (\lambda > 0)$ 的泊松分布, X_1, \dots, X_n 是来自总体 X 的简单随机样本, 则对应的统计量 $T_1 = \frac{1}{n} \sum_{i=1}^n X_i$, $T_2 = \frac{1}{n-1} \sum_{i=1}^{n-1} X_i + \frac{1}{n} X_n$ 满足 ()。
- A. $ET_1 > ET_2, DT_1 > DT_2$ B. $ET_1 > ET_2, DT_1 < DT_2$

C. $ET_1 < ET_2, DT_1 > DT_2$ D. $ET_1 < ET_2, DT_1 < DT_2$

8. 某厂家生产的灯泡寿命的均值为 60 小时, 标准差为 4 小时。如果从中随机抽取 30 只灯泡进行检测, 则样本均值 ()。

- A. 抽样分布的标准差为 4 小时
B. 抽样分布近似等同于总体分布
C. 抽样分布的中位数为 60 小时
D. 抽样分布近似等同于正态分布, 均值为 60 小时

9. 设随机变量 X_1, \dots, X_n 相互独立且同分布, $DX_i = \sigma^2$, $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$,

$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$, 则 S ()。

- A. 是 σ 的无偏估计 B. 是 σ 的极大似然估计
C. 是 σ 的一致估计 D. 与 \bar{X} 相互独立

10. 根据一个具体的样本, 计算总体均值的置信水平为 90% 的置信区间, 则该区间 ()。

- A. 以 90% 的概率包含总体均值 B. 有 10% 的可能性包含总体均值
C. 绝对包含总体均值 D. 绝对包含总体均值或绝对不包含总体均值

11. 设 X_1, X_2, X_3, X_4 是来自总体 X 的样本, $EX = \mu$, 则 () 是 μ 的最有效估计。

- A. $\hat{\mu} = \frac{1}{4} X_1 + \frac{1}{4} X_2 + \frac{1}{4} X_3 + \frac{1}{4} X_4$
B. $\hat{\mu} = \frac{1}{5} X_1 + \frac{2}{5} X_2 + \frac{1}{5} X_3 + \frac{1}{5} X_4$
C. $\hat{\mu} = \frac{1}{9} X_1 + \frac{2}{9} X_2 + \frac{1}{9} X_3 + \frac{1}{9} X_4$
D. $\hat{\mu} = \frac{1}{3} X_1 + \frac{1}{3} X_2 + \frac{1}{6} X_3 + \frac{1}{6} X_4$

12. 设总体 X 的分布中未知参数 θ 的置信度为 $1-\alpha$ 的置信区间为 $(\hat{\theta}_1, \hat{\theta}_2)$, 即 $P(\hat{\theta}_1 \leq \hat{\theta}_2) = 1-\alpha$ 。则下面说法正确的是 ()。

- A. 对 $\hat{\theta}_1, \hat{\theta}_2$ 的观测值 a, b , 恒有 $\theta \in (a, b)$
B. θ 的数学期望 $E\theta$ 必属于 $(\hat{\theta}_1, \hat{\theta}_2)$
C. θ 以 $1-\alpha$ 的概率落入区间 $(\hat{\theta}_1, \hat{\theta}_2)$
D. 区间 $(\hat{\theta}_1, \hat{\theta}_2)$ 以 $1-\alpha$ 的概率包含 θ

13. 从一个正态总体中随机抽取一个容量为 n 的样本, 其均值和标准差分别为 33 和 4。当 $n=25$ 时, 构造总体均值的 95% 的置信区间为 ()。

A. 33 ± 4.97 B. 33 ± 2.22 C. 33 ± 1.65 D. 33 ± 1.96

14. 指出下列假设检验形式的写法哪一个是错误的 ()?

A. $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$ B. $H_0: \mu \geq \mu_0, H_1: \mu < \mu_0$

C. $H_0: \mu \leq \mu_0, H_1: \mu > \mu_0$ D. $H_0: \mu > \mu_0, H_1: \mu \leq \mu_0$

15. 1998 年的一次网络民意调查中, 共 7553 人接受调查, 其中 4381 人认为“白水事件”调查执行官 Kenneth Start 应该询问所有可能找到的证人, 包括时任总统克林顿的助手。由此, 你能得出以下哪个结论 ()?

A. 以上数据提供了充分证据, 说明多数人认为应该询问所有可能找到的证人, 包括时任总统克林顿的助手

B. 以上数据未提供充分证据, 说明多数人认为应该询问所有可能找到的证人, 包括时任总统克林顿的助手

C. 以上数据提供了充分证据, 说明多数人认为不应询问所有可能找到的证人, 包括时任总统克林顿的助手

D. 以上数据未提供充分证据, 说明多数人认为不应询问所有可能找到的证人, 包括时任总统克林顿的助手

16. 在一项涉及 1602 名儿童的流感疫苗试验中, 接受疫苗的 1070 人中只有 14 人患了流感, 而接受安慰剂的 532 名儿童中有 98 人患了流感。在 $\alpha = 0.05$ 的显著性水平下, 检验“疫苗减少了儿童患流感的可能性”, 即检验假设 $H_0: \pi_1 - \pi_2 \geq 0, H_1: \pi_1 - \pi_2 < 0$, 检验的结论是 ()。

A. 拒绝 H_0 B. 不拒绝 H_0

C. 可以拒绝也可以不拒绝 H_0 D. 可能拒绝也可能不拒绝 H_0

17. 对线性回归方程的显著性检验, 通常采用 3 种方法, 即相关系数检验法、 F 检验法和 t 检验法, 下面说法正确的是 ()。

A. F 检验法最有效

B. t 检验法最有效

C. 3 种方法是相通的, 检验效果是相同的

D. F 检验法和 t 检验法, 可以代替相关系数检验法

18. 在 DW 检验中, 无序列相关的区间为 ()。

A. $0 \leq DW \leq d_u$ B. $d_u < DW < 4 - d_u$

C. $4 - d_u \leq DW \leq 4 - d_l$ D. $4 - d_u < DW < 4$

19. 对模型 $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$ 的最小二乘回归结果显示, 多重判定系数 $R^2 = 0.92$, 样本容量为 30, 总离差平方和为 500, 则估计的标准误差为 ()。

A. 1.217 B. 1.482 C. 4.152 D. 5.214

20. 在多元线性回归模型中, 若自变量 x_j 对因变量 y 的影响不显著, 那么它的回归系数 β_j 的取值 ()。

A. 可能为 0 B. 可能为 1 C. 可能小于 0 D. 可能大于 1

21. 在多元线性回归分析中, 多重共线性是指 ()。

A. 两个或两个以上的自变量彼此相关

B. 两个或两个以上的自变量彼此无关

C. 因变量与一个自变量相关

D. 因变量与一个或一个以上的自变量相关

22. 关于方差分析, 以下说法正确的是 ()。

A. 方差分析的目的是分析各组总体方差是否相同

B. 方差分析的组间平方和仅仅衡量了随机误差的变异大小

C. 各组数据呈严重偏态时, 也可以做方差分析

D. 方差分析的目的是分析各组总体均值是否相同

23. 在单因素方差分析中, 从 4 个总体种各选取了 5 个观察值, 得到组间平方和 $SSA = 375$, 组内平方和 $SSE = 480$, 则 F 统计量的值为 ()。

A. 0.78 B. 1.30 C. 2.50 D. 4.17

24. 移动平均法适合于预测 ()。

A. 平稳序列 B. 非平稳序列 C. 有趋势成分的序列 D. 有季节成分的序列

25. 某商场 2012 年 8 月份的商品销售额为 100 万元, 该月的季节指数为 125%, 在消除季节因素后该月的销售额为 ()。

A. 80 万元 B. 100 万元 C. 125 万元 D. 以上都不对

26. 对于任意二事件 A 和 B , 则 ()。

A. 若 $AB \neq \phi$, 则 A, B 一定独立 B. 若 $AB \neq \phi$, 则 A, B 有可能独立

C. 若 $AB = \phi$, 则 A, B 一定独立 D. 若 $AB = \phi$, 则 A, B 一定不独立

27. 设 A 和 B 是任意两个概率不为零的互不相容事件, 则下列结论肯定正确的是 ()。

A. \bar{A} 和 \bar{B} 不相容

B. \bar{A} 和 \bar{B} 相容

C. $P(AB) = P(A)P(B)$

D. $P(A-B) = P(A)$

28. 10 件产品中有 3 件次品, 随机从中抽取两件, 至少抽到一件次品的概率 ()。

A. $\frac{1}{3}$ B. $\frac{2}{5}$ C. $\frac{7}{15}$ D. $\frac{8}{15}$

29. 设 X 的概率密度函数为 $f(x)$, 则 $f(x)$ 一定满足 ()。

A. $0 \leq f(x) \leq 1$

B. $\int_{-\infty}^{+\infty} f(x) dx = 1$

C. $p\{X > x\} = \int_x^{\infty} f(t)dt$ D. $f(+\infty) = 1$

30. 设 X 服从参数为 λ 的泊松分布, 已知 $P(X=1) = \frac{1}{2}P(X=2)$, 则 $\lambda =$ ()。

- A. 2 B. 1 C. 4 D. 0.25

二、简要回答下列问题 (本题包括 1—4 题共 4 个小题, 每小题 10 分, 共 40 分)。

1. 一组数据的分布特征包括哪几个方面? 分别用哪些指标进行测度?
2. 在多元线性回归中, 对整个方程进行的 F-检验和对每个变量进行的 t-检验等价吗? 为什么?
3. 时间序列的影响因素有哪些? 各自的含义如何?
4. 试写出概率的公理化定义。

三、计算与分析题 (本题包括 1—3 题共 3 个小题, 第 1 小题和第 2 小题每题 20 分, 第 3 小题 10 分, 共 50 分)。

1. 设两位化验员 A, B 独立地对某种聚合物含氯量用相同的方法各做 10 次测定, 其测定值的样本均值分别为, $\bar{X}_A = 1.2034$, $\bar{X}_B = 2.3268$, 样本方差依次为 $S_A^2 = 0.5419$, $S_B^2 = 0.6065$. 设 σ_A^2, σ_B^2 分别为 A, B 所测定的测定值总体的方差, 且总体均服从正态分布, 求

- 1) 方差比 σ_A^2 / σ_B^2 的置信度为 0.90 的置信区间;
 - 2) 均值差 $\mu_1 - \mu_2$ 的置信度为 0.95 的置信区间 (这里假设 $\sigma_A^2 = \sigma_B^2$).
- ($F_{0.05}(9,9) = 3.18$, $F_{0.05}(10,10) = 0.3358$, $t_{0.025}(20) = 2.086$, $t_{0.025}(18) = 2.101$)

2. 在铜线含碳量对于电阻的效应的研究中, 得到数据如下,

含碳量 x_i (%)	0.10	0.30	0.40	0.55	0.70	0.80	0.95
电阻 y_i (微欧)	15	18	19	21	22.6	23.8	26

- (1) 作出散点图;
- (2) 求线性回归方程 $y = \hat{a} + \hat{b}x$;
- (3) 假设线性关系显著, 试对 $x = 0.98$ 时, 预测 y 的值。

(已算得 $\sum_{i=1}^n x_i = 3.8$, $\sum_{i=1}^n y_i = 145.4$, $\sum_{i=1}^n x_i^2 = 2.595$, $\sum_{i=1}^n x_i y_i = 85.61$)

3. 设随机变量 X 的密度函数为 $f(x) = \begin{cases} Ax(1-x), & 0 < x < 1 \\ 0, & \text{其它} \end{cases}$ 。

求: (1) 常数 A ;

(2) X 的分布函数;

(3) X 的数学期望 EX 和方差 DX ;