

安徽师范大学

2019 年硕士研究生招生考试初试试题

科目代码: 432

科目名称: 统计学

注意: 可使用不带有存储功能的计算器。

一、单项选择题 (本题包括 1-25 题共 25 个小题, 每小题 2 分, 共 50 分。在每小题给出的四个选项中, 只有一个符合题目要求, 把所选项前的字母填在答题纸相应的序号内)。

- 为了调查某高校学生的购书费用支出, 从男生中抽取 60 名学生调查, 从女生中抽取 40 名学生调查, 这种抽样方法属于【 】。
A. 简单随机抽样 B. 等距抽样 C. 分层抽样 D. 整群抽样
- 要反映我国上市公司业绩的整体水平, 总体是【 】。
A. 我国所有上市公司
B. 我国每一家上市公司
C. 我国上市公司总数
D. 我国上市公司的利润总额
- 以下数据不属于顺序数据的是【 】。
A. 性别 (如: 男、女) 数据
B. 评级 (如: 优、良、差) 数据
C. 产品质量 (如: 一等品、二等品、三等品等) 数据
D. 受教育程度 (如: 小学、初中、高中等) 数据
- 如果你的业务是销售运动衫, 哪一种运动衫号码的度量对你更为有用?【 】。
A. 均值 B. 中位数
C. 众数 D. 四分位数
- 一组数据的离散系数为 0.4, 平均数为 20, 则标准差为【 】。
A. 8 B. 4 C. 80 D. 0.02
- 与直方图相比, 茎叶图【 】。
A. 没保留原始数据的信息
B. 不能有效展示数据的分布
C. 保留了原始数据的信息
D. 更适合描述分类数据
- 对数据进行标准化变换是一种重要的数据预处理方法, 其计算公式是用某一原始数据减去这组数据的均值, 再除以这组数据的【 】。
A. 均值 B. 中位数 C. 众数 D. 标准差
- 一个 95% 的置信区间是指【 】。
A. 总体参数有 95% 的概率落在这一区间内
B. 在用同样方法构造的总体参数的多个区间中, 有 95% 的区间包含该总体参数

- 总体参数有 5% 的概率未落在这一区间内
D. 在用同样方法构造的总体参数的多个区间中, 有 5% 的区间不包含该总体参数
- 在一次假设检验中, 当显著性水平 $\alpha = 0.01$ 时原假设被拒绝, 则用 $\alpha = 0.05$ 时【 】。
A. 一定会被拒绝 B. 需要重新检验
C. 一定不会被拒绝 D. 有可能拒绝原假设
- 在方差分析中, 进行多重比较的前提是【 】。
A. 拒绝原假设 B. 不拒绝原假设
C. 可以拒绝原假设也可以不拒绝原假设 D. 各样本均值相等
- 一个估计量的一致性是指【 】。
A. 该估计量的数学期望等于被估计的总体参数
B. 该估计量的方差比其他估计量小
C. 随着样本量的增大, 该估计量的值越来越接近于被估计的总体参数
D. 该估计量的方差比其他估计量大
- 某班学生年龄分布是左偏的, 均值为 25, 标准差是 4.45。如果采取重复抽样的方法从该班抽取容量为 100 的样本, 则样本均值的抽样分布是【 】。
A. 正态分布, 均值为 25, 标准差为 0.445
B. 分布形状未知, 均值为 25, 标准差为 4.45
C. 正态分布, 均值为 25, 标准差为 4.45
D. 分布形状未知, 均值为 25, 标准差为 0.445
- 某饮料生产企业研制了一种新型饮料, 饮料有五种颜色。如果要考察颜色是否会影响销售, 则水平为【 】。
A. 2 B. 3 C. 4 D. 5
- 设 A_1, A_2, A_3 是三个事件, 则事件 $A_1 \cup A_2 \cup A_3$ 表示【 】。
A. A_1, A_2, A_3 都发生 B. A_1, A_2, A_3 都不发生
C. A_1, A_2, A_3 至少有一个发生 D. A_1, A_2, A_3 不多于一个发生
- 设 $X \sim N(\mu, 2^2), Y \sim N(\mu, 3^2)$, 记 $a_1 = P(X \leq \mu - 2), a_2 = P(Y \geq \mu + 3)$, 则【 】。
A. $a_1 < a_2$ B. $a_1 = a_2$ C. $a_1 > a_2$ D. 无法判断
- 设两个随机变量 X 和 Y 相互独立且同分布: $P\{X = -1\} = P\{Y = -1\} = 1/2$,
 $P\{X = 1\} = P\{Y = 1\} = 1/2$, 则下列各式中成立的是【 】。
A. $P\{X = Y\} = \frac{1}{2}$ B. $P\{X = Y\} = 1$
C. $P\{X + Y = 0\} = \frac{1}{4}$ D. $P\{XY = 1\} = \frac{1}{4}$
- 设事件 A 与 B 互不相容, 且 $P(A) > 0, P(B) > 0$, 则有【 】。
A. $P(\overline{AB}) = 1$ B. $P(A) = 1 - P(B)$ C. $P(AB) = P(A)P(B)$ D. $P(A \cup B) = 1$
- 设随机变量 $X \sim N(1, 4), Y \sim N(-1, 6)$, 且 X 和 Y 独立, 则 $D(2Y - X) =$ 【 】。
A. 8 B. 16 C. 28 D. 20
- 设随机变量 X 和 Y 不相关, 则下列式子中错误的是【 】。
A. $D(XY) = D(X) \cdot D(Y)$ B. $E(XY) = E(X) \cdot E(Y)$
C. $D(X + Y) = D(X) + D(Y)$ D. $Cov(X, Y) = 0$
- 设总体 $X \sim N(\mu, \sigma^2)$, 其中 $\sigma^2 > 0$ 未知, μ 已知, (X_1, X_2, \dots, X_n) 是来自 X 的样本, 则下

列不是统计量的是【 】。

- A. $\frac{1}{n} \sum_{i=1}^n X_i$ B. $\sum_{i=1}^n X_i / \sigma^2$ C. $\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$ D. $\max\{X_1, X_2, \dots, X_n\}$

21. 设 X_1, X_2, \dots, X_6 是来自总体 $X \sim N(0,1)$ 的样本, 则统计量 $\frac{(X_1^2 + X_5^2)/2}{(X_2^2 + X_4^2 + X_6^2)/3}$ 服从【 】。

- A. $\chi^2(5)$ B. $t(3)$ C. $F(3,2)$ D. $F(2,3)$

22. 设 R 为列联表的行数, C 为列联表的列数; 则 χ^2 分布的自由度为【 】。

- A. $R \times (C-1)$ B. $(R-1) \times C$ C. $(R-1) \times (C-1)$ D. $R \times C$

23. 按指数的对比指标性质不同, 统计指数可分为【 】。

- A. 分类指数和总体指数 B. 个体指数和总指数
C. 动态指数和静态指数 D. 数量指数和质量指数

24. 从商学院、经济学院、财经学院和环境学院分别抽取 30 名学生, 得到每个学生的数学考试分数。要分析各学院的考试分数是否有显著差异, 适合的分析方法是【 】。

- A. 独立样本 t 检验 B. χ^2 拟合优度检验
C. 回归分析 D. 方差分析

25. 在多元线性回归分析中, 如果 t 检验表明各回归系数不显著, 则意味着【 】。

- A. 整个回归方程的线性关系不显著
B. 整个回归方程的线性关系显著
C. 自变量与因变量之间的线性关系不显著
D. 自变量与因变量之间的线性关系显著

二、判断题 (本题包括 26-30 题共 5 个小题, 每小题 2 分, 共 10 分。若论断正确, 在题后的括号内打 \checkmark ; 若论断不正确, 在题后的括号内打 \times)。

26. 在总体方差未知的情况下进行均值检验, 一定要用 t 统计量。【 】
27. 进行假设检验时, 检验统计量的值落在拒绝域之外, 并不能说明原假设一定为真。【 】
28. 当置信水平一定时, 置信区间的宽度随着样本量的增大而减少。【 】
29. 对季度数据序列, 用移动平均法测定其长期趋势值时, 可采用四项移动平均。【 】
30. 相关系数为 0 表示两个变量之间不存在任何关系。【 】

三、简要回答下列问题 (本题包括 31-34 题共 4 个小题, 每小题 10 分, 共 40 分)。

31. 一项对成年人和幼儿的身高 (厘米) 进行抽样调查, 要比较成年组和幼儿组的身高差异, 你会采用什么样的指标测度值? 为什么?
32. 什么是相关关系? 请给出相关关系的分类。
33. 方差分析是用来分析什么问题的? 方差分析需要哪些基本假定?
34. 试述假设检验的一般步骤。

四、计算与分析题 (本题包括 35-39 题共 5 个小题, 每小题 10 分, 共 50 分)

参考数据: $t_{0.025}(14) = 2.14$, $Z_{0.11} \approx 1.25$, $Z_{0.025} = 1.96$, $Z_{0.05} = 1.65$,

$t_{0.025}(12) = 2.18$, $t_{0.05}(12) = 1.782$, $F_{0.05}(7,5) = 4.88$, $F_{0.05}(5,7) = 3.97$

35. 有人做过记录, 在晚上 8:30 以后某一娱乐频道两次广告的间隔时间为: (单位: 分钟)

6.0	6.6	5.8	7.0	6.3	6.2	7.2	5.7	6.4	7.0
6.5	6.2	6.0	6.5	7.2	7.3	7.6	6.8	6.0	6.2

问题:

- (1) 计算均值和中位数; (5 分)
(2) 计算上四分位数和下四分位数。 (5 分)

36. 从一个标准差为 5 的总体中抽出一个容量为 100 的样本, 样本均值为 25。

- (1) 样本均值的抽样标准差 $\sigma_{\bar{x}}$ 等于多少? (4 分)
(2) 在 95% 的置信水平下, 允许误差是多少? (4 分)
(3) 求总体均值在 95% 的置信水平下的置信区间。 (2 分)

37. 设 X_1, \dots, X_n 是来自总体 $P(\lambda)$ 的样本, $\lambda > 0$ 未知, 求 λ 的最大似然估计量 $\hat{\lambda}$, 并验证它是否是 λ 的无偏估计?

38. 有甲、乙两台机床加工同样产品, 从它们的产品中分别随机抽取 8 件和 6 件, 测得产品直径 (单位: mm) 数据为:

$$\bar{x}_1 = 201, \bar{x}_2 = 198, s_1^2 = 0.17, s_2^2 = 0.14$$

假定两个总体都服从正态分布。

- (1) 在显著性水平 $\alpha = 0.1$ 下, 检验这两个总体的方差是否相等。 (5 分)
(2) 在显著性水平 $\alpha = 0.05$ 下, 检验这两台机床加工的产品平均直径有无显著差异。 (5 分)

39. 某企业用三种方法组装一种新产品, 为确定哪种方法每小时生产的产品数量最多, 随机抽取 30 名工人, 并指定每个人使用其中的一种方法。方差分析表数值如下表:

差异源	SS	df	MS	F	P-value	F crit
组间	A	2	210	E	0.245946	3.354131
组内	3836	C	D	—	—	—
总计	B	29	—	—	—	—

- (1) 求表中 A、B、C、D、E 的值 (结果保留 3 位小数); (5 分)
(2) 在显著性水平 $\alpha = 0.05$ 下, 检验三种方法组装的产品数量之间是否有显著差异? (5 分)