

安徽师范大学

2017 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—30 题共 30 个小题, 每小题 2 分, 共 60 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卡相应的序号内)。

1. 对一组样本数据: 5, 11, 11, 23, 24, 25, 35, 35, 47, 其众数和中位数分别是 ()。

- A. 众数是 11, 中位数是 23.5
- B. 众数是 35, 中位数是 24
- C. 众数是 11 和 35, 中位数是 23.5
- D. 众数是 11 和 35, 中位数是 24

2. 抽签属于 ()。

- A. 简单随机抽样
- B. 分层抽样
- C. 整群抽样
- D. 非概率抽样

3. 以下不适用于展示定性数据的是 ()。

- A. 频数频率分布表
- B. 柱状图
- C. 环形图
- D. 直方图

4. 箱线图(box plot)中不包括以下哪个统计量 ()。

- A. 中位数
- B. 均值
- C. 上四分位数
- D. 下四分位数

5. 莎士比亚在戏剧里的用字长度的分布是右偏的, 说明莎翁 ()。

- A. 更习惯用长的字
- B. 更习惯用短的字
- C. 用字的长度无规律
- D. 不能说明任何问题

6. 一名统计学专业的学生为了完成其统计论文, 其在《统计年鉴》中找到 2014 年各省的城镇家庭的人均收入数据。这一数据属于 ()。

- A. 分类数据
- B. 顺序数据
- C. 截面数据
- D. 时间/序列数据

7. 两组数据相比较 ()。

- A. 标准差大的离散程度就大
- B. 标准差大的离散程度小
- C. 离散系数大的离散程度就大
- D. 离散系数大的离散程度小

8. 统计学中的许多“系数”是不带量纲、可直接进行比较的, 下列哪个“系数”不具有这种特征 ()。

- A. 离散系数
- B. 偏态系数
- C. 相关系数
- D. 回归系数

9. 关于回归分析中决定系数(R Square)的描述, 错误的是 ()。

- A. 决定系数的功能是评价模型拟合效果, 是一个绝对的拟合效果的度量
- B. 决定系数表示由自变量的线性回归模型解释的因变量的变化所占总变化的比例
- C. 决定系数表示引入自变量后模型的残差平方和减小的比例
- D. 如果有两上及两个以上的自变量, 应采用调整的决定系数(Adjusted R Square)来评价模型拟合效果

10. 95%的置信水平是指 ()。

- A. 总体参数落在一个特定的样本所构造的区间内的概率为 95%
- B. 总体参数落在一个特定的样本所构造的区间内的概率为 5%

考生请注意: 答案必须写在答题纸上, 写在本试题纸上的无效!

- C. 在用同样方法构造的总体参数的多个区间中, 包含总体参数的区间比例为 95%
 D. 在用同样方法构造的总体参数的多个区间中, 包含总体参数的区间比例为 5%
11. 关于原假设和备择假设的说法中, 错误的是 ()。
- A. 带等号的假设一般作为原假设
 B. 需要保护的假设一般作为备择假设
 C. 想要证明的假设一般作为备择假设
 D. 选定原假设后, 原假设的对立假设可以作为备择假设
12. 以下关于抽样分布的说法中, 错误的是 ()。
- A. 抽样分布主要用于评价估计量的效果以及构造置信区间和拒绝域
 B. χ^2 分布、t 分布以及 F 分布是三个常见的抽样分布
 C. 抽样分布一般与样本容量无关
 D. 当精确的抽样分布不易推导出来时, 可以通过随机模拟等方法来获得近似的抽样分布
13. 在单因子方差分析中, 若在显著性水平 0.05 下拒绝了原假设, 则表明 ()。
- A. 在显著性水平 0.05 下, 因子各个水平的均值全不相等
 B. 在显著性水平 0.05 下, 因子各个水平的均值不全相等
 C. 在显著性水平 0.05 下, 因子中至少有两个水平的均值相等
 D. 以上说明均不正确
14. 对于一个方差已知的正态总体均值进行区间估计, 采用常用的区间估计公式, 下列说法中正确的是 ()。
- A. 若给定置信度, 则样本容量越大, 区间会越短
 B. 若给定置信度, 则样本容量越大, 区间会越长
 C. 若给定置信度, 则样本容量越大, 区间长度保持不变
 D. 若给定置信度, 则样本容量的变化与区间长度的变化之间没有任何关系
15. 下列关于假设检验问题中 P 值的说法中正确的是 ()。
- A. P 值就是第一类错误的概率
 B. P 值就是第二类错误的概率
 C. 当 P 值大于显著性水平时, 应拒绝原假设
 D. P 值表示原假设为真的条件下, 所得样本结果会像实际观测结果那么极端或更极端的概率
16. 若原假设为 H_0 : 某生产过程的不合格品率不大于 p_0 , 则第二类错误是指 ()。
- A. 认为该过程生产的不合格品过多, 但实际并不多
 B. 认为该过程生产的不合格品不过多, 但实际过多
 C. 认为该过程生产的不合格品不过多, 实际也不过多
 D. 认为该过程生产的不合格品过多, 实际也过多
17. 下列关于参数估计的说法中, 错误的是 ()。
- A. 常用参数估计方法有矩估计法和最大似然估计法
 B. 矩估计的原理是建立总体与被估参数的关系, 然后用样本矩去代替总体矩, 从而得到估计量
 C. 无偏是指估计量的期望等于被估参数
 D. 均方误差越大, 说明估计量越好
18. 下列关于相关系数的说法中, 错误的是 ()。
- A. 相关系数的取值范围是 $[-1, 1]$
 B. 相关系数的绝对值越大, 表示两变量之间的线性相关程度就越大
 C. 相关系数为负数, 表示两变量之间存在负线性相关关系

考生请注意: 答案必须写在答题纸上, 写在本试题纸上的无效!

- D. 相关系数为零, 表示两变量之间不存在任何关系
19. 方差分析中的多重比较是通过配对比较来进一步检验 ()。
 A. 哪两个总体均值之间有差异
 B. 哪两个总体方差之间有差异
 C. 哪两个样本均值之间有差异
 D. 哪两个样本方差之间有差异
20. 设总体服从正态分布 $N(\mu, \sigma^2)$, 其中 σ^2 已知, 从中抽取一个样本容量为 10 的简单随机样本, 则样本均值的抽样分布为 ()。
 A. $N(\mu, \sigma^2/10)$
 B. $N(\mu/10, \sigma^2)$
 C. $N(\mu/10, \sigma^2/100)$
 D. $N(\mu, \sigma^2/100)$
21. 设总体服从正态分布 $N(\mu, \sigma^2)$, 如果在显著性水平 $\alpha = 0.05$ 下拒绝了 $H_0: \mu = \mu_0$, 则在显著性水平 $\alpha = 0.01$ 下 ()。
 A. 接受 H_0
 B. 拒绝 H_0
 C. 可能接受 H_0 , 也可能拒绝 H_0
 D. 不接受 H_0 , 也不拒绝 H_0
22. 设 $X \sim N(0, 1)$, 记 u_α 为其上侧分位数, 即 $P(X > u_\alpha) = \alpha$, 则下列结论中不成立的是 ()。
 A. $u_\alpha + u_{1-\alpha} = 0$
 B. $u_{0.5} = 0$
 C. 若 $\alpha > \beta$, 则 $u_\alpha > u_\beta$
 D. $P(X > u_{\alpha/2}) + P(X < u_{1-\alpha/2}) = \alpha$
23. 设总体 $X \sim N(\mu, \sigma^2)$, 其中 $\sigma^2 > 0$ 未知, μ 已知. 设 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. $\mu + \sum_{i=1}^n X_i / \sigma^2$
 C. $\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$
 D. $X_n - \mu$
24. 事件 A 与任何事件均独立, 则 $P(A)$ 等于 π ()。
 A. 0
 B. 0.5
 C. 1
 D. 0 或 1
25. 在独立重复试验中, 每次试验成功的概率为 $p(0 < p < 1)$, 则直到第 $n+1$ 次试验才取得第 $r(1 \leq r \leq n+1)$ 次成功的概率为 ()。
 A. $C_{n-1}^{r-1} p^r (1-p)^{n-r}$
 B. $C_{n-1}^{r-1} p^{r-1} (1-p)^{n-r}$
 C. $C_n^{r-1} p^r (1-p)^{n-r+1}$
 D. $C_n^{r-1} p^{r-1} (1-p)^{n-r+1}$
26. 设 X 的概率密度为 $f(x) = \frac{1}{2\sqrt{2\pi}} e^{-\frac{(x-2)^2}{8}}$, $-\infty < x < \infty$, 则 $E(X^2)$ 等于 ()。
 A. 8
 B. 16
 C. 4
 D. 2
27. 设 X, Y 是随机变量, 且 X, Y 均不为常数, 则下列与 $E(XY) = E(X)E(Y)$ 不等价的是 ()。

考生请注意: 答案必须写在答题纸上, 写在本试题纸上的无效!

- A. $\text{Cov}(X, Y) = 0$ B. $D(X + Y) = D(X) + D(Y)$
 C. X, Y 不相关 D. $D(XY) = D(X)D(Y)$

28. 设随机变量 X 服从参数为 2 的泊松(Poisson)分布, 则 $2X-1$ 的方差等于 ()。

- A. 5 B. 8 C. 9 D. 16

29. 同时掷两枚骰子, 则出现的点数之和为 3 的概率是 ()。

- A. 1/36 B. 1/18 C. 1/12 D. 1/6

30. 设日光灯的使用寿命 X (单位: 月)服从指数分布, 已知 $P(X > 10) = p$, 现有一只日光灯, 已经使用了 10 个月仍未坏, 则它还能再使用 10 个月以上的概率为 ()。

- A. p B. $1-p$ C. $2p$ D. $p(1-p)$

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. 谈谈 t 分布与正态分布有什么区别和联系。
2. 先简述最大似然估计法的思想; 设总体服从区间 $[\alpha, \beta]$ 上的均匀分布, 再求参数 α 和 β 的最大似然估计。
3. 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?
4. 假设检验中为什么可能会犯错误? 会犯什么样的错误?

三、计算与分析题 (本题包括 1—4 题共 4 个小题, 第 1 小题和第 2 小题每题 14 分, 第 3 小题和第 4 小题每题 11 分, 共 50 分)。

1. 甲袋中有 3 个白球 2 个黑球, 乙袋中有 4 个白球 4 个黑球。现从甲袋中任取两球放入乙袋, 再从乙袋中任取一球,

- (1) 求从乙袋中取出的一球是白球的概率;
 - (2) 若已知从乙袋中取出的是白球, 求从甲袋中取出的两球是一白一黑的概率。
2. 设总体 X 的密度函数为

$$f(x) = \frac{1}{2\theta} e^{-|x|/\theta}, \quad -\infty < x < \infty.$$

其中 $\theta > 0$ 未知, (X_1, \dots, X_n) 是来自 X 的样本,

(1) 试求 θ 的最大似然估计 $\hat{\theta}$;

(2) $\hat{\theta}$ 是 θ 的无偏估计吗? 为什么?

3. 化肥厂用自动包装机打包, 每包的标准重量是 100 斤, 每天开工后需检验一次包装机工作是否正常。已知包重服从正态分布, 某日开工后测 9 包产品, 计算后得平均重量是 100.4 斤, 样本标准差为 1.2 斤, 问该天包装机工作是否正常? (α 取 0.05)

(已知 $t_{0.025}(8) = 2.31$, $t_{0.025}(9) = 2.26$, $t_{0.05}(8) = 1.86$, $t_{0.05}(9) = 1.833$)

考生请注意: 答案必须写在答题纸上, 写在本试题纸上的无效!

4. 有四个品牌（因素 1）的彩电在五个地区（因素 2）销售，对每个品牌在各地区的销售量随机抽取了一些样本，利用统计软件中的无重复双因素方差分析功能，运行后得到以下数据：

SUMMARY	观测数	求和	平均	方差
行 1	5	1721	344.20	233.70
行 2	5	1739	347.80	295.70
行 3	5	1685	337.00	442.50
行 4	5	1424	284.80	249.20
列 1	4	1356	339.00	1224.67
列 2	4	1321	330.25	1464.25
列 3	4	1357	339.25	822.92
列 4	4	1273	318.25	1538.92
列 5	4	1262	315.50	241.67

方差分析

差异源	SS	df	MS	F	P-value	F crit
行	13004.55		4334.85	18.11	0.00	3.49
列	2011.7		502.93	<u>B</u>	0.14	3.26
误差	2872.7	<u>A</u>	239.39			
总计	17888.95	19				

- (1) 将表格中 A, B 处缺少的数据填写完整；
- (2) 从输出结果来看，你可以得出什么统计结论？