

山 大 学

二〇一六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. 下面哪种调查方式的样本不是随机选取的 ()。
 - A. 分层抽样
 - B. 系统抽样
 - C. 整群抽样
 - D. 判断抽样

2. 为了解小区居民对物业服务的意见和看法, 管理人员随机抽取了 50 户居民, 上门通过问卷进行调查。这种数据收集方法是 ()。
 - A. 面访式问卷调查
 - B. 实验调查
 - C. 观察式调查
 - D. 自填式问卷调查

3. 为了研究多个不同变量在不同样本间的相似性, 适合采用的图形是 ()。
 - A. 环形图
 - B. 茎叶图
 - C. 雷达图
 - D. 箱线图

4. 经验法则表明, 当一组数据对称分布时, 在平均数加减 1 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

5. 如果一个样本因人故意操纵而出现偏差, 这种误差属于 ()。
 - A. 抽样误差
 - B. 实验误差
 - C. 设计误差
 - D. 非抽样误差

6. 设 $N(0, \sigma^2)$, 则服从 $t(n-1)$ 的随机变量是 ()。
 - A. $\frac{\sqrt{n}\bar{X}}{S}$
 - B. $\frac{\sqrt{n-1}\bar{X}}{S}$
 - C. $\frac{\sqrt{n}\bar{X}}{S^2}$
 - D. $\frac{\sqrt{n-1}\bar{X}}{S^2}$

7. 设总体 $X \sim U\left(\frac{1}{2}-\theta, \frac{1}{2}+\theta\right)$, 其中 θ 为未知参数。设 X_1, \dots, X_n 是来自 X 的样本, 则当 $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$ 中的 $\mu =$ () 时, S^2 不是统计量。
 - A. $\frac{1}{n} \sum_{i=1}^n X_i$
 - B. $\max_{1 \leq i \leq n} X_i$
 - C. EX
 - D. DX

8. 某厂家生产的灯泡寿命的均值为 1000 小时, 标准差为 4 小时。如果从中随机抽取 16 只灯泡进行检测, 则样本均值 ()。
 - A. 抽样分布的标准差为 1 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 1000 小时 D. 抽样分布服从正态分布, 均值为 1000 小时
9. 样本 X_1, \dots, X_n 来自正态分布总体 $N(\mu, \sigma^2)$, 则使得 $c \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2$ 为 σ^2 的无偏估计的常数 $c =$ ()。
 - A. $\frac{1}{n-1}$
 - B. $\frac{1}{2(n-1)}$
 - C. $\frac{1}{2n}$
 - D. $\frac{1}{n-1}$
10. 在其他条件相同的情况下, 95% 的置信区间比 90% 的置信区间 ()。
 - A. 要宽
 - B. 要窄
 - C. 相同
 - D. 可能宽也能窄
11. 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被估计的总体参数
 - D. 该估计量的方差比其他估计量大
12. 随机抽取一个有 290 名教师组成的样本, 让每个人对一些说法表明自己的态度。第一种说法是“年龄偏大的学生对班上的讨论比年龄偏小的学生更积极”。态度按 5 分制来衡量: 1=非常同意; 2=同意; 3=没有意见; 4=不同意; 5=很不同意。对这一看法, 样本的平均态度得分为 1.94, 标准差为 0.92。用 98% 的置信水平估计教师对这一看法的平均态度得分的置信区间为 ()。
 - A. 1.94 ± 0.13
 - B. 1.94 ± 1.13
 - C. 1.94 ± 1.96
 - D. 1.94 ± 2.58
13. 某地区的写字楼月租金的标准差为 80 元, 要估计总体均值的 95% 的置信区间, 希望的边际误差为 15 元, 则应抽取的样本量为 ()。
 - A. 100
 - B. 110
 - C. 120
 - D. 30
14. 如果原假设 H_0 为真, 所得到的样本结果会像实际观测结果那么极端或更极端的概率称为 ()。
 - A. 临界值
 - B. 统计量
 - C. P 值
 - D. 事先给定的显著性水平 α
15. 在假设检验中, 当样本容量一定时, 若缩小犯第一类错误的概率, 则犯第二类错误的概率会相应 ()。
 - A. 增大
 - B. 减少
 - C. 不变
 - D. 不确定
16. 抽自两个总体的独立随机样本提供的信息如下表所示:

样本1	样本2
$n_1 = 80$	$n_2 = 70$
$\bar{x}_1 = 104$	$\bar{x}_2 = 106$
$s_1 = 8.4$	$s_2 = 7.6$

则在显著性水平 $\alpha = 0.05$ 时, 检验假设 $H_0: \mu_1 - \mu_2 = 0, H_1: \mu_1 - \mu_2 \neq 0$, 得到的结论是()。

- 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. 条件不足, 无法判断

17. 如果 y 关于 x 的回归方程为 $\hat{y} = 2 - x$, 而且这个回归方程的 $R^2 = 0.81$, 则 x 与 y 之间的相关系数()。

- A. $r = 1$
- B. $r = -1$
- C. $r = 0.9$
- D. $r = -0.9$

18. 以回归方程 $Y = a + bX$ 作相关分析和回归分析, 关于样本相关系数 r 与回归系数 b , 下列各论断中哪项是正确的()?

- A. $r > 0$ 时 $b < 0$
- B. $r > 0$ 时 $b > 0$
- C. $|r| = 1$ 时 $|b| = 1$
- D. $|r| = 1$ 时 $b = 0$

19. 某研究人员发现, 举重运动员的体重与他能举起的重量之间的相关系数为 0.6, 则()。

- A. 举重能力的 60% 归因于体重
- B. 平均来说, 运动员能举起其体重 60% 的重量
- C. 如果运动员的体重增加 10 公斤, 则可多举起 6 公斤的重量
- D. 运动员体重和能举起的重量之间有正相关关系

20. 在 k 元线性回归中, n 为样本容量, SSE 为残差平方和, SSR 为回归平方和, 则对回归方程线性关系的显著性进行检验时, 构造的 F 统计量为()。

- A. $\frac{SSR}{SST}$
- B. $\frac{SSE}{SST}$
- C. $\frac{SSR/k}{SSE/(n-k-1)}$
- D. $1 - \frac{SSR/k}{SSE/(n-k-1)}$

21. 在多元线性回归分析中, 当 F 检验表明线性关系显著时, 部分回归系数的 t 检验却并不显著, 这意味着()。

- A. 不显著的回归系数所对应的自变量对因变量的影响不显著
- B. 所有的自变量对因变量的影响都不显著
- C. 模型中可能存在多重共线性
- D. 整个回归模型的线性关系不显著

22. 关于单因素方差分析中的 F 检验, 下面说法正确的是()。

- A. 拒绝域在 F 分布曲线的右侧
- B. F 统计量的样本观测值可能为负值
- C. 拒绝域在 F 分布曲线的左侧和右侧
- D. 以上表述都不对

23. 从两个总体中分别抽取 $n_1 = 7$ 和 $n_2 = 6$ 的两个独立随机样本, 经计算得到下面的方差分析表:

差异源	SS	df	MS	F	P-值
组间	A	1	7.50	3.15	0.10
组内	26.19	11	2.38		
总计	33.69	12			

表中 “A” 单元格内的结果是()。

- A. 4.50
- B. 5.50
- C. 6.50
- D. 7.50

24. 移动平均法是通过计算逐项移动的时序平均数, 来形成派生序列, 从而达到()对序列的影响。

- A. 消除偶然因素引起的不规则变动
- B. 消除非偶然因素引起的不规则变动
- C. 消除人为误差
- D. 消除计算误差

25. 要通过移动平均法消除季节变动, 则移动平均项数()。

- A. 应选择奇数
- B. 应选择偶数
- C. 应和季节周期长度一致
- D. 可任意取值

26. 设两事件 A 与 B 同时发生时, 事件 C 必发生, 则()发生。

- A. $P(C) \leq P(A) + P(B) - 1$
- B. $P(C) \geq P(A) + P(B) - 1$
- C. $P(C) = P(AB)$
- D. $P(C) = P(A) + P(B)$

27. 如果 $P(A) + P(B) > 1$, 则事件 A 与 B 必定()。

- A. 独立
- B. 不独立
- C. 相容
- D. 不相容

28. 已知随机变量 X 服从区间 $[a, b]$ 上的均匀分布, 且期望 $EX = 2$, 方差 $DX = 3$,

$P\{X \leq c\} = \frac{1}{6}$, 则参数 a, b, c 的值为()。

- A. $a = 5, b = -1, c = 1$
- B. $a = -1, b = 5, c = 1$
- C. $a = 5, b = -1, c = 0$
- D. $a = -1, b = 5, c = 0$

29. 设随机变量 X 的密度函数为 $\phi(x)$, 且满足 $\phi(x) = \phi(-x)$, X 的分布函数为 $F(x)$, 则对任意实数 a , $F(-a) =$ ()。

- A. $1 - \int_0^a \phi(x)dx$
- B. $\frac{1}{2}(1 - \int_{-a}^a \phi(x)dx)$
- C. $F(a)$
- D. $2F(a) - 1$

30. 设 X 的密度函数为 $f(x) = \frac{1}{\pi(1+x^2)}$, 则 $Y=3X$ 的密度函数为 ()。

- A. $\frac{1}{\pi(1+9y^2)}$ B. $\frac{3}{\pi(9+y^2)}$ C. $\frac{1}{\pi(9+y^2)}$ D. $\frac{1}{\pi(1+y^2)}$

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. 影响抽样误差大小的因素主要有哪些?
2. 回归分析和相关分析的联系和区别?
3. 什么是区间估计? 如何理解置信区间与参数真值的关系?
4. 试简述事件的独立性与随机变量的独立性的异同?

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. 设总体 X 的概率分布为:

X	0	1	2	3
P	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中: $\theta(0 < \theta < \frac{1}{2})$ 是未知参数, 利用样本观测值: 3, 1, 3, 0, 3, 1, 2, 3,

求: θ 的矩估计值和极大似然估计值。

2. 已知某企业 10 年间的研究费用 x 与利润 y 的统计资料如下:

年份	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
研究费用 x (万元)	10	10	8	8	8	12	12	12	11	11
利润 y (万元)	100	150	200	180	250	300	280	310	320	300

(其中: $\sum x_i = 102, \sum y_i = 2390, \sum x_i^2 = 1066, \sum y_i^2 = 624300, \sum x_i y_i = 25040$)

- (1) 建立研究费用 x 与利润 y 的直线回归方程, 并说明回归系数的经济意义;
- (2) 在 0.05 的显著性水平下, 检验所建立的方程线性回归效果是否显著?

已知: $t_{0.05}(9)=1.8331, t_{0.05}(8)=1.8595, t_{0.025}(9)=2.2622, t_{0.025}(8)=2.3060$

3. 玻璃杯成箱出售, 每箱 20 只, 假设各箱含 0, 1, 2 只残次品的概率相应为 0.8, 0.1, 0.1。顾客欲购一箱玻璃杯, 在购买时, 售货员随意取一箱, 而顾客开箱随机查看四只: 若无残次品则买下该箱玻璃杯, 否则退回。试求: 顾客买下该箱玻璃杯的概率。

