

# 苏州大学

## 2011 年硕士研究生入学考试初试试题 (A 卷)

科目代码: 432 科目名称: 统计学 满分: 150 分

注意: ①认真阅读答题纸上的注意事项; ②所有答案必须写在答题纸上, 写在本试题纸或草稿纸上均无效; ③本试题纸须随答题纸一起装入试题袋中交回!

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)。

- 关于随机抽样, 下列说法中正确的是 ( )。
  - 抽样时应使得总体内的每个个体都有相同的机会被抽取
  - 随机抽样即随意抽取个体作为样本
  - 研究者在抽样时应精心挑选个体, 以使样本更具代表性
  - 应选取符合研究者意愿的样本
- 2010 年 11 月 1 日零点的全国人口普查是 ( )。
  - 一次性调查和非全面调查
  - 经常性调查和非全面调查
  - 一次性调查和全面调查
  - 经常性调查和全面调查
- 下列指标中不属于离散趋势指标的是 ( )。
  - 算术平均数
  - 极差
  - 四分位数间距
  - 方差
- 观察得 10 只养犬发现症状到死亡的时间分别为 7、8、8、9、11、12、12、13、14、14 天, 则这组数据的中位数为 ( )。
  - 11
  - 12
  - 11.5
  - 10.8
- 已知总体的均值为 50, 标准差为 8, 从该总体中随机抽取样本量为 16 的样本, 则样本均值的数学期望和标准差分别为 ( )。
  - 50 和 8
  - 50 和 2
  - 50 和 0.5
  - 8 和 2

- 设  $\theta$  为总体  $X$  的未知参数, 随机区间  $[\theta_1(X_1, \dots, X_n), \theta_2(X_1, \dots, X_n)]$  是  $\theta$  的置信水平为  $1-\alpha$  ( $0 < \alpha < 1$ ) 的置信区间, 则有 ( )。
  - $P(\theta \leq \theta_2(X_1, \dots, X_n)) = 1-\alpha$
  - $P(\theta_1(X_1, \dots, X_n) \leq \theta \leq \theta_2(X_1, \dots, X_n)) = \alpha$
  - $P(\theta \leq \theta_1(X_1, \dots, X_n)) = \alpha$
  - $P(\theta_1(X_1, \dots, X_n) \leq \theta \leq \theta_2(X_1, \dots, X_n)) = 1-\alpha$
- 在假设检验中, 记  $H_1$  为备择假设, 则称 ( ) 为犯第 I 类错误。
  - 若  $H_1$  为真, 接受  $H_1$
  - 若  $H_1$  不真, 接受  $H_1$
  - 若  $H_1$  为真, 拒绝  $H_1$
  - 若  $H_1$  不真, 拒绝  $H_1$
- 利用 P 值进行显著性检验 (水平为  $\alpha$ ) 时, 接受原假设的条件是 ( )。
  - P 值  $< \alpha$
  - P 值  $\neq \alpha$
  - P 值  $> \alpha$
  - P 值  $> \beta$
- 设总体分布形式和总体方差都未知, 对总体均值进行假设检验时, 若抽取一个容量为 100 的样本, 则可采用 ( )。
  - Z 检验法
  - t 检验法
  - $\chi^2$  检验法
  - F 检验法
- 调查 50 个房地产公司, 房屋销售面积与广告费用之间的相关系数为 0.76, 这说明 ( )。
  - 二者之间有较强的正相关关系
  - 平均看来, 销售面积的 76% 归因于其广告费用
  - 如要多销售 1 万平方米的房屋, 则要增加广告费用 7600 元
  - 如果广告费用增加 1 万元, 可以多销售 7600 平方米的房屋
- 在回归分析中, 因变量的预测区间估计是指 ( )。

- A. 对于自变量  $x$  的一个给定值  $x_0$ , 求出因变量  $y$  的平均值的区间
- B. 对于自变量  $x$  的一个给定值  $x_0$ , 求出因变量  $y$  的个别值的区间
- C. 对于因变量  $y$  的一个给定值  $y_0$ , 求出自变量  $x$  的平均值的区间
- D. 对于因变量  $y$  的一个给定值  $y_0$ , 求出自变量  $x$  的平均值的区间
12. 基本统计推断的两大主要内容是 ( )。
- A. 统计估计与假设检验
- B. 总体均值与方差
- C. 矩估计与最大似然估计
- D. 参数检验与非参数检验
13. 在下面的假定中, 哪一个不属于方差分析中的假定 ( )。
- A. 每个总体都服从正态分布
- B. 各总体的方差相等
- C. 观测值是独立的
- D. 各总体的均值等于 1
14. 单因素方差分析中, 各平方和之间一定有关系式 ( )。
- A.  $SS_{\text{组内}} < SS_{\text{组间}}$
- B.  $MS_{\text{组内}} < MS_{\text{组间}}$
- C.  $MS_{\text{总}} = MS_{\text{组内}} + MS_{\text{组间}}$
- D.  $SS_{\text{总}} = SS_{\text{组内}} + SS_{\text{组间}}$
15. 某批产品共计有 4000 件, 为了了解这批产品的质量, 从中随机抽取 200 件进行质量检验, 发现其中有 30 件不合格。根据抽样结果进行推断, 下列说法不正确的是 ( )。
- A. 样本量为 30
- B. 总体合格率的点估计值是 85%
- C. 总体合格率是一个未知参数
- D. 样本合格率是一个统计量
16. 某企业产值 2000 年比 1991 年增长 200%, 计算其年平均发展速度的算式是 ( )。
- A.  $\bar{x} = \sqrt[10]{3}$
- B.  $\bar{x} = \sqrt[10]{2}$
- C.  $\bar{x} = \sqrt[3]{2}$
- D.  $\bar{x} = \sqrt[3]{3}$
17. 如果时间序列的环比增长量大致相等, 则应采用的趋势模型为 ( )。
- A. 直线趋势模型
- B. 指数曲线趋势模型

- C. 二次曲线趋势模型
- D. 修正指数曲线趋势模型
18. 如果变量  $x$  和变量  $y$  之间的相关系数为 0, 这说明两个变量之间一定是 ( )。( )。
- A. 没有关系
- B. 完全线性相关关系
- C. 高度线性相关关系
- D. 以上都不对
19. 对于  $p$  阶时间序列自回归模型来说, 可用于估计其模型参数的有效样本容量为 ( )。
- A.  $n$
- B.  $p$
- C.  $n-p$
- D.  $n+p$
20. 设总体  $X$  的方差  $\sigma^2$  存在,  $X_1, X_2, \dots, X_n$  是  $X$  的样本, 若总体  $X$  的期望  $EX$  未知, 则总体方差  $\sigma^2$  的无偏估计量为 ( )。
- A.  $\hat{\sigma}_1^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - EX)^2$
- B.  $\hat{\sigma}_2^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$
- C.  $\hat{\sigma}_3^2 = \frac{1}{n} \sum_{i=1}^n (X_i - EX)^2$
- D.  $\hat{\sigma}_4^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$
21. 设  $X_1, X_2, X_3, X_4$  是来自正态总体  $N(0, 9)$  的一个简单随机样本, 则统计量  $\xi = \frac{(X_2 + X_3 + X_4)^2}{3X_1^2}$  服从的分布为 ( )。
- A.  $F(2, 1)$
- B.  $F(1, 1)$
- C.  $\chi^2(3)$
- D.  $\chi^2(2)$
22. 在不重置抽样中, 抽样单位数从 5% 增加到 25%, 抽样平均误差 ( )。
- A. 增加 39.7%
- B. 增加约  $\frac{3}{5}$
- C. 减少约  $\frac{3}{5}$
- D. 没有什么变化
23. 设总体  $X$  服从区间  $(0, \theta)$  上的均匀分布, 即  $X \sim U(0, \theta)$ , 其中  $\theta > 0$  为未知参数,  $X_1, X_2, \dots, X_n$  是

来自总体  $X$  的样本。现有  $\theta$  的四个估计量:  $\hat{\theta}_1 = X_{(n)}$ ,  $\hat{\theta}_2 = 2\bar{X}$ ,  $\hat{\theta}_3 = \frac{n+1}{n}X_{(n)}$  和  $\hat{\theta}_4 = \frac{n+2}{n+1}X_{(n)}$ 。则此四个估计量中是  $\theta$  的无偏估计的是 ( )。

- A.  $\hat{\theta}_1$
- B.  $\hat{\theta}_1$  和  $\hat{\theta}_2$
- C.  $\hat{\theta}_2$  和  $\hat{\theta}_3$
- D.  $\hat{\theta}_2$  和  $\hat{\theta}_4$

24. 设总体  $X$  服从正态分布  $N(\mu, \sigma^2)$ , 其中参数  $\mu$  和  $\sigma^2$  都未知,  $X_1, X_2, \dots, X_n$  是总体  $X$  的一个样本, 样本均值为  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ , 样本方差为  $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ , 记  $t(n-1)$  分布的上侧  $\alpha$  分位数为  $t_{\alpha}(n-1)$  (即  $P(T \geq t_{\alpha}(n-1)) = \alpha$ ), 则参数  $\mu$  的置信水平为  $1-\alpha$  的置信区间为 ( )。

- A.  $[\bar{X} \pm \frac{S}{\sqrt{n}} t_{\alpha}(n-1)]$
- B.  $[\bar{X} \pm \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1)]$
- C.  $[\bar{X} \pm \frac{S}{n} t_{\alpha}(n-1)]$
- D.  $[\bar{X} \pm \frac{S}{n} t_{\alpha/2}(n-1)]$

25. 设总体  $X \sim N(\mu, 4)$ ,  $X_1, X_2, \dots, X_{25}$  是来自总体  $X$  的一个容量为 25 的样本, 现要检验假设:  $H_0: \mu=2 \leftrightarrow H_1: \mu=4$ 。若检验法则由拒绝域  $W = \{(x_1, x_2, \dots, x_n) | \bar{x} \geq 3.2\}$  确定, 则该检验犯第一类错误的概率  $\alpha$  和犯第二类错误的概率  $\beta$  分别为 ( )。

- A.  $\alpha = \Phi(3)$ ,  $\beta = \Phi(2)$
- B.  $\alpha = 1 - \Phi(3)$ ,  $\beta = \Phi(2)$
- C.  $\alpha = 1 - \Phi(3)$ ,  $\beta = 1 - \Phi(2)$
- D.  $\alpha = \Phi(3)$ ,  $\beta = 1 - \Phi(2)$

26. 设正态随机变量  $X \sim N(2, 4)$ , 则随机变量  $Y = 2X + 2$  所服从的分布为正态分布 ( )。

- A.  $N(6, 16)$
- B.  $N(6, 8)$
- C.  $N(8, 6)$
- D.  $N(16, 6)$

27. 设  $X$  和  $Y$  是独立同分布随机变量, 共同分布是参数为  $\alpha$  的指数随机变量, 则随机变量  $X+Y$  是 ( )。

- A. 参数为  $2\alpha$  的指数随机变量
- B. 参数为  $\alpha^2$  的指数随机变量
- C. 参数为  $(\alpha, 2)$  的  $\Gamma$ -随机变量

D. 参数为  $(2, \alpha)$  的  $\Gamma$ -随机变量

28. 若事件  $A$  和  $B$  的概率满足  $P(A) = P(B) = 1$ , 则事件  $A \cap B$  的概率是 ( )。

- A. 1
- B. 0
- C. 0.5
- D. 0.8

29. 已知事件  $A, B, C$  且  $P(A \cap B) = 0.5$ ,  $P(A \cap B \cap C) = 0.3$ , 概率  $P_A(\bullet)$  定义为  $P_A(B) = \frac{P(A \cap B)}{P(A)}$ 。则

条件概率  $P_A(C|B)$  的值为 ( )。

- A. 0.2
- B. 0.4
- C. 0.6
- D. 0.8

30. 离散型随机变量  $\xi$  的分布列为  $\begin{pmatrix} 1 & 2 & 3 \\ 0.4 & a & b \end{pmatrix}$ , 其中  $a, b$  是未知数, 如果已知  $\xi$  取 2 的概率和取 3 的概率相等, 则  $a = ( )$ 。

- A. 0.2
- B. 0.3
- C. 0.4
- D. 0.5

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. 简述假设检验的基本步骤。
2. 简述无偏估计与一致最小方差无偏估计的定义。
3. 简述回归分析与相关分析的区别。
4. 一般两个随机变量的协方差为零并不意味着这两个随机变量独立, 但若两个正态随机变量的协方差为零, 则二者独立, 给出这一事实的解释。

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  $X$  的密度函数为  $f(x) = \begin{cases} \theta x^{\theta-1} & 0 < x < 1 \\ 0 & \text{其它} \end{cases}$ , 其中未知参数  $\theta > 0$

若  $x_1, x_2, \dots, x_n$  是  $X$  的一组样本观察值, 求  $\theta$  的矩法估计与最大似然估计。

2. 为了调查某广告对销售收入的影响, 某商店记录了 5 个月的销售收入  $y$  (万元) 和广告费用  $x$  (万元), 数据见表 2.6, 要求用手工计算:

月份	1	2	3	4	5
$X$	1	2	3	4	5
$Y$	10	10	20	20	40

- (1) 用最小二乘估计求出回归方程;
- (2) 求回归标准误差  $\hat{\sigma}$ ;
- (3) 计算  $x$  与  $y$  的决定系数;
- (4) 对回归方程作方差分析 (显著性水平  $\alpha = 0.05$ );
- (5) 给出  $\beta_1$  的置信度为 95% 的区间估计。

注: F 分布的上侧分位数  $F_{0.05}(1, 3) = 10.13$

t 分布的上侧分位数  $t_{0.025}(3) = 3.182$

3. 某工厂有甲、乙、丙三条生产线生产螺丝钉, 它们的产量各占 25%, 35%, 40%, 在各自的产品里, 不合格品各占 5%, 4%, 3%。现在从产品中任取一支, (1). 该产品是不合格品的概率是多少? (2). 若该产品是不合格, 则它是由机器甲生产的概率是多少?