

# 2011 年华东师范大学金融与统计学院 432 统计学[专业硕士]考研真题



启航教育

2011 年华东师范大学金融与统计学院 432 统计学[专业硕士]考研真题

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，先在答题纸上写上序号，再把所选项前的字母填在相应的序号后）。

1. 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
2. 下列说法中错误的是（ ）。
  - A. 简单随机抽样是一种概率抽样
  - B. 采用概率抽样方法，不可计算和控制抽样误差
  - C. 采用概率抽样方法，则每个样本都有事先确定的被抽到的概率（ $\lambda$ 样概率）
  - D. 采用概率抽样方法，则由 $\lambda$ 样概率确定随机抽的样本
3. 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（ ），再除以这组数据的标准差。
  - A. 均值
  - B. 中位数
  - C. 众数
  - D. 标准差
4. 为探索一组定量数据的分布形态，下列图形中最适合的是（ ）。
  - A. 饼图
  - B. 柱状图
  - C. 直方图
  - D. 雷达图
5. 以下关于回归模型假定的检验的说法中，错误的是（ ）。
  - A. 对正态性的假定进行检验，应通过检验残差的正态性来进行
  - B. 对独立性的假定进行检验，应通过检验残差的独立性来进行
  - C. 对方差齐性的假定进行检验，应通过检验残差的方差齐性来进行
  - D. 无法通过残差来检验回归函数线性的假定
6. 在单因子方差分析中，若在显著性水平  $\alpha=0.05$  下拒绝了原假设，则表明（ ）。
  - A. 在显著性水平  $\alpha=0.05$  下，因子各个水平的均值全不相等
  - B. 在显著性水平  $\alpha=0.05$  下，因子各个水平的均值不全相等
  - C. 在显著性水平  $\alpha=0.05$  下，因子中至少有两个水平的均值相等
  - D. 在显著性水平  $\alpha=0.05$  下，因子不显著
7. 设有 5 人获得了某项劳务收入，分别为 300 元、350 元、400 元、500 元、600 元，他们收入的中位数等于（ ）。
  - A. 375 元
  - B. 400 元
  - C. 430 元
  - D. 450 元
8. 下列统计量中，不能用于描述数据差异情况的统计是（ ）。
  - A. 极差
  - B. 标准差
  - C. 方差
  - D. 众数

9. 设准妈妈的怀孕期（单位：天）服从正态分布  $N(\mu, 16^2)$ 。调查了 100 个准妈妈的怀孕期，得到的样本均值为 266 天。设  $\mu_r$  表示标准正态分布的概率为  $r$  的下分位数。那么，参数  $\mu$  的置信度为 95% 的区间估计为（ ）

- A.  $[266 - 16^2 u_{0.975} / \sqrt{100}, 266 + 16^2 u_{0.975} / \sqrt{100}]$
- B.  $[266 - 16 u_{0.975} / \sqrt{100}, 266 + 16 u_{0.975} / \sqrt{100}]$
- C.  $[266 - 16 u_{0.975}, 266 + 16 u_{0.975}]$
- D.  $[266 - 16 u_{0.95} / \sqrt{100}, 266 + 16 u_{0.95} / \sqrt{100}]$

10. 对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，下列说法正确的是（ ）

- A. 若给定置信度，样本量越大区间会越短
- B. 若给定置信度，样本量越大区间会越长
- C. 若给定置信度，样本量变大，区间长度保持不变
- D. 若给定置信度，样本量的变化与区间长度的变化没有规律

11. 以下关于显著性检验的说法，错误的是（ ）。

- A. 若做出拒绝原假设的判断，这是有充分证据的
- B. 若做出接受原假设的判断，这是有充分证据的
- C. 显著性检验有保护原假设的作用
- D. 不能轻易拒绝的假设一般应设置为原假设

12. 设某一批零件的直径服从正态分布  $N(20, 2^2)$ ，若直径在 20 加减 4 之间为合格品，则该批零件的合格频率约为（ ）

- A. 68%
- B. 90%
- C. 95%
- D. 99%

13. 设有两个正态总体  $N(\mu_1, 1)$  和  $N(\mu_2, 1)$ ，为对假设检验问题  $H_0: \mu_1 = \mu_2, H_1: \mu_1 > \mu_2$  进行检验，从两个总体中各抽取了样本容量为  $n$  的样本，得样本均值分别为  $\bar{x}_1$  和  $\bar{x}_2$ ，则检验的拒绝域形式应为（ ）

- A.  $\{\bar{x}_1 - \bar{x}_2 > c\}$ （ $c$  为常数）
- B.  $\{\bar{x}_1 - \bar{x}_2 > c / \sqrt{n}\}$ （ $c$  为常数）
- C.  $\{\bar{x}_1 - \bar{x}_2 < c\}$ （ $c$  为常数）

D.  $\{\bar{x}_1 - \bar{x}_2 < c / \sqrt{n}\}$  (c 为常数)

14. 在单因子方差分析中，组内平方和反映的是 ( )

- A. 全部观测值的误差
- B. 因子各个水平的样本均值之间的误差
- C. 因子各个水平内部的样本观测值之间的误差
- D. 因子各个水平之间和各个水平内部的误差和

15. 某研究小组分析身高与体重的相关关系，计算了一个相关系数。但后来发现因测体重的秤不准，每位测量对象的体重数据偏大了 1 公斤，因此需要进行校正从而使数据准确。那么，相对于之前的相关系数，采用校正后的体重数据来计算的相关系数会 ( )。

- A. 减小
- B. 增大
- C. 不变
- D. 无法判断

16. 下列关于假设检验中 P 值的说法正确的是 ( )

- A. P 值等于第一类错误概率
- B. P 值等于第二类错误概率
- C. 当 P 值大于显著性水平  $\alpha$ ，应拒绝原假设
- D. P 值表示原假设为真的假设下，所得到样本结果会像实际观测结果那么极端或更极端的概率

17. 以下关于多重共线性的描述，错误的是 ( )。

- A. 若存在多重共线性，回归模型的效果不会受到影响
- B. 若存在多重共线性，则表明某些自变量之间存在一定的线性相关关系
- C. 若存在多重共线性，则可能导致增减自变量时回归系数的估计量变化很大
- D. 若存在多重共线性，则可能导致某些回归系数不显著

18. 已知某种商品的销售量存在周期性，其时间序列预测模型为  $\hat{X}_t = 0.8X_{t-2} + 0.3X_{t-4}$ 。设  $X_1=30$ ,  $X_2=10$ ,  $X_3=40$ ,  $X_4=20$ ,  $X_5=50$ ，则  $\hat{X}_6 = ( )$ 。

- A. 19
- B. 41
- C. 46
- D. 52

19. 以下关于时间序列的说法，错误的是 ( )。

- A. 对存在周期变化的数据，采用时间序列模型进行描述是比较合适的
- B. 对随时间推移而逐渐增长或衰减的数据，采用时间序列模型进行描述是比较合适的
- C. 某些情况下，可采用时间序列模型来解决回归模型的多重共线性问题
- D. 无法利用时间序列模型进行预测

20. 下列哪些统计量不易受到极端值的影响 ( )

- A. 均值
- B. 方差
- C. 中位数
- D. 标准差

21. 常采用五数概括 (即最小值、下四分位数、中位数、四分位数、最大值) 描述这一组数据的情况，并由此做出箱线图。下列说法错误的是 ( )

- A. 可通过箱线图大致了解分布的中心位置  
 B. 可通过箱线图大致了解分布的离散程度  
 C. 可通过箱线图大致了解分布的形态  
 D. 可通过箱线图, 无法大致比较不同总体分布的善异
22. 设原假设为  $H_0$ : 某产生过程的不合格品率不大于  $P_0$ 。则第二错误指的是 ( )。
- A. 认为该过程生产不合格品过多, 但实际并不多  
 B. 认为该过程生产不合格品不过多, 但实际过多  
 C. 认为该过程生产不合格品不过多, 实际也不过多  
 D. 认为该过程生产不合格品过多, 实际也过多
23. 设某因子有 3 个水平, 每个水平下各重复 2 次试验。在单因子方差分析中, 得到  $F=17$ , 已知  $F_{0.95}(2,3)=9.55, F_{0.99}(2,3)=30.80$ , 则 ( )。
- A. 因子显著  
 B. 因子不显著  
 C. 在显著性水平  $\alpha=0.05$  下, 因子显著  
 D. 在显著性水平  $\alpha=0.01$  下, 因子显著
24. 参数的估计量是一个 ( ), 衡量其效果最重要的指标是其 ( )。
- A. 随机变量, 均方误差  
 B. 随机变量, 期望  
 C. 常数, 均方误差  
 D. 常数, 期望
25. 以下关于回归分析的说法中, 错误的是 ( )。
- A. 解读回归分析结果时, 应首先看回归模型是否显著, 在回归模型显著地情况下再看各回归系数是否显著。  
 B. 对回归系数  $\beta$  进行区间估计, 需要正态分布的假定  
 C. 对  $X$  和  $Y$  建立一元线性回归方程, 则  $X$  和  $Y$  的相关系数与  $\beta$  成正比  
 D. 若对同样的数据建立两个回归方程, 则  $R^2$  越小的回归方程越好
26. 同时掷两个均匀骰子 (为六面的正方体, 分别标有 1, 2, 3, 4, 5, 6 个点), 则两骰子点数之和为 3 的概率为 ( )。
- A. 1/36  
 B. 2/36  
 C. 3/36  
 D. 6/36
27. 已知  $P(B) > 0$ , 则  $P(A|B)$  与  $P(A)$  的大小关系是 ( )。
- A.  $P(A|B) \geq P(A)$   
 B.  $P(A|B) \leq P(A)$   
 C.  $P(A|B) = P(A)$   
 D. 上述三种情况都有可能
28. 设随机变量  $X$  服从二项分布  $b(2, 0.3)$ , 则  $P(X=2)$  为 ( )。
- A. 0.09  
 B. 0.18  
 C. 0.30  
 D. 0.49

29 设随机变量  $X$  服从正态分布  $N(\mu, \sigma^2)$ , 则  $X$  的期望和方差分别为 ( )

- A.  $\mu, \sigma$       B.  $\mu, \sigma^2$       C.  $\sigma, \mu$       D.  $\sigma^2, \mu$

30. 已知随机变量  $X$  的期望和方差分别为 -1 和 1, 则  $2X+3$  的方差为 ( )

- A. 1      B. 4      C. 5      D. 7

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. 在假设检验中, 为什么说无论是作出接受原假设还是拒绝原假设的判断, 都有可能犯错误? 它们是什么样的错误?

2. 重复测量某物体的重量, 得到数据  $X_1, X_2, \dots, X_{50}$ , 应当采用  $\frac{1}{25} \sum_{i=1}^{25} X_i$  和  $\frac{1}{50} \sum_{i=1}^{50} X_i$  中的哪一个作为物体重量的估计为好? 为什么?

3. 简述时间序列的乘法模型。

4. 先说说矩估计的统计思想, 再求出区间  $[\alpha, \beta]$  上均匀分布的总体参数的矩估计。

三、计算与分析题(本题包括 1~3 题共 3 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. 某生产车间 50 名工人日加工零件数如下(单位: 个) 117、122、124、129、139、107、117、130、122、125、108、131、125、117、122、133、126、122、118、108、110、118、123、126、133、134、127、123、118、112、112、134、127、123、119、113、120、123、127、135、137、114、120、128、124、115、139、128、124、121。请先适当分组, 在作出频率直方图, 最后根据图形, 判断这组数据的分布形状。

2. 某商业银行 2007 年各分行贷款情况见下表 3 (单位: 亿元)

表 1 某商业银行 2007 年各分行贷款情况

不良贷款 $y$	各项贷款余额 $x_1$	累计应收贷款 $x_2$
0.58	4.21	4.25
0.71	6.96	12.40
...	...	...
1.99	6.39	7.51

运用 Excel 中的“回归分析”, 得到下列结果:

SUMMARY OUTPUT	
Multiple R	0.8541
R Square	0.74667
Adjusted R Square	A
标准误差	1.21959
观测值	25

方差分析

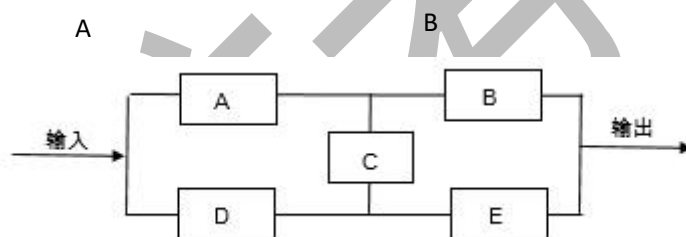
	df	SS	MS	F	Significance F
回归分析		96.4488	48.2244	C	2.7575E-07

残差	B	32.7228	1.4874		
总计		129.172			

	Coefficients	标准误差	t Stat	P-value	Lower95%	Upper95%
Intercept	-0.921986908	0.4783984	-1.9272	0.00669	-1.9141244	0.07015063
X Variable 1	0.27840674	0.0674517	D	0.00044	0.1385239	0.41829309
X Variable 2	E	0.085462	2.28313	0.03244	0.01788363	0.3723581

根据以上结果，回答下列问题：

- (1) 将上述表中的数据 A, B, C, D, E 求出：
  - (2) 你认为用线性回归模型的拟合原始数据是否合适？为什么？
  - (3) 写出不良贷款  $y$  与各项贷款余额  $x_1$  及累计应收贷款  $x_2$  之间线性回归方程，并解释各个回归系数的意义：
  - (4) 您认为各个回归系数是不是显著不等于 0？为什么？（取显著性水平为 0.05）
  - (5) 找出各个回归参数的 95% 区间估计。
3. 图中的五个字母代表五个同类元件，假设它们正常工作的概率都是  $P$ ，且某个元件是否正常工作与其它元件无关，求系统正常工作的概率  $P$ 。



哎呀这里只有部分真题

加群 779335571

可获取全部真题答案资料及相应答疑

你还在等什么？

启航 2020 应用统计考研交流群 779335571