

机密★启封前

# 湖北汽车工业学院

## 2023 年攻读硕士学位研究生入学考试试题

科目名称：统计学原理

(☐A 卷☒B 卷) 科目代码：810

考试时间：3 小时 满分 150 分

注意：本试题共四大题，共 6 页；所有答题内容必须写在答题纸上，写在试题或草稿纸上的一律无效；考完后试题和答题纸一同装入试卷袋密封交回。

### 一、名词解释（共 5 小题，每小题 4 分，共 20 分）

- 数值型数据
- 方差分析
- 中心极限定理
- 置信区间
- 区间估计

### 二、单项选择题（共 20 小题，每小题 2 分，共 40 分）

- 抽样分布是指（ ）  
A. 一个样本各观测值的分布  
B. 总体中各观测值的分布  
C. 样本统计量的分布  
D. 样本数量的分布
- 一家餐厅记录了自开业 5 年以来每天的营业额，每天营业额的均值为 2500 元，标准差为 400 元。由于在特殊节日的营业额偏高，所以每天营业额的分布是右偏的，假设从这 5 年中随机抽取 100 天，计算这 100 天的平均营业额，则样本的抽样分布是（ ）  
A. 正态分布，均值为 2500 元，标准差为 40 元  
B. 正态分布，均值为 250 元，标准差为 40 元  
C. 右偏，均值为 2500 元，标准差为 400 元  
D. 正态分布，均值为 2500 元，标准差为 400 元

3. 从均值为 200、标准差为 50 的总体中抽取容量为 100 的简单随机样本，样本均值的标准差是（ ）

- A. 10      B. 15      C. 50      D. 5

4. 在参数估计中，要求通过样本的统计量来估计总体参数，评价统计量的标准之一是使它与总体参数的离差越小越好。这种评价标准称为（ ）

- A. 无偏性  
B. 有效性  
C. 一致性  
D. 充分性

5. 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. 总体参数有 5%的概率未落在这一区间内  
C. 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  
D. 在用同样方法构造的总体参数的多个区间中，有 95%的区间不包含该总体参数

6. 一个估计量的有效性是指（ ）

- A. 该估计量的数学期望等于被估计的总体参数  
B. 该估计量的一个具体数值等于被估计的总体参数  
C. 该估计量的方差比其他估计量大  
D. 该估计量的方差比其他估计量小

7. 在置信水平不变的条件下，要缩小置信区间，则（ ）

- A. 需要增加样本量  
B. 需要减少样本量  
C. 需要保持样本量不变  
D. 需要改变统计量的抽样标准差

8. 使用统计量  $z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$  估计总体均值的条件是（ ）

- A. 总体为正态分布
- B. 总体为正态分布且方差已知
- C. 大样本
- D. 总体为正态分布但方差未知

9. 某地区估计营养不良人数高达 20%，但有人认为这个比例实际上还要高，如果想要检验该说法是否正确，则假设形式为（ ）

- A.  $H_0: \pi \leq 0.2; H_1: \pi > 0.2$
- B.  $H_0: \pi = 0.2; H_1: \pi \neq 0.2$
- C.  $H_0: \pi \geq 0.3; H_1: \pi < 0.3$
- D.  $H_0: \pi \leq 0.3; H_1: \pi < 0.3$

10. 一种零件的标准长度 5cm，检验某天生产的零件是否符合标准要求，建立的原假设和备择假设应为（ ）

- A.  $H_0: \mu = 5; H_1: \mu \neq 5$
- B.  $H_0: \mu \neq 5; H_1: \mu = 5$
- C.  $H_0: \mu \leq 5; H_1: \mu > 5$
- D.  $H_0: \mu \geq 5; H_1: \mu < 5$

11. 在假设检验中，第一类错误是指（ ）

- A. 当原假设错误时拒绝原假设
- B. 当原假设正确时拒绝原假设
- C. 当备择假设正确时拒绝备择假设
- D. 当备择假设不正确时未拒绝备择假设

12. 在方差分析中，检验统计量  $F$  是（ ）

- A. 组间平方和除以组内平方和

- B. 组间均方除以组内均方
- C. 组间平方除以总平方和
- D. 组间均方除以总均方

13. 在方差分析中，不同水平下样本数据之间的误差称为（ ）

- A. 组内误差
- B. 组间误差
- C. 组内平方
- D. 组间平方

14. 下面是一个方差分析表，表中 A, B, C, D, E 五个单元格内的数据分别是（ ）

差异源	SS	df	MS	F
组间	24.7	4	C	E
组内	A	B	D	
总计	62.7	34		

- A. 38, 30, 6.175, 1.27, 4.86
- B. 38, 29, 6.175, 1.27, 4.86
- C. 38, 30, 6.175, 1.27, 5.86
- D. 27.7, 29, 6.175, 1.27, 4.86

15. 下面的各问题中，哪个不是相关分析要解决的问题（ ）

- A. 判断变量之间是否存在关系
- B. 判断一个变量数值的变化对另一个变量的影响
- C. 描述变量之间的关系强度
- D. 判断样本所反映的变量之间的关系能否代表总体变量之间的关系

16. 如果变量之间的关系近似地表现为一条直线，则称两个变量之间为（ ）

- A. 正线性相关关系
- B. 负线性相关关系
- C. 线性相关关系

D. 非线性相关关系

17. 如果相关系数  $r=0$ ，则表明两个变量之间（ ）

- A. 相关程度很低
- B. 不存在任何关系
- C. 不存在线性相关关系
- D. 存在非线性相关关系

18. 在回归分析中，F 检验主要是用来检验（ ）

- A. 相关系数的显著性
- B. 回归系数的显著性
- C. 线性关系的显著性
- D. 估计标准误差的显著性

19. 由最小二乘法得到的回归直线，要求满足因变量的（ ）

- A. 平均值与其估计值的离差平方和最小
- B. 实际值与其平均值的离差平方和最小
- C. 实际值与其估计值的离差和为 0
- D. 实际值与其估计值的离差平方和最小

20. 残差平方和  $SSE$  反映了  $y$  的总变差中（ ）

- A. 由于  $x$  与  $y$  之间的线性关系引起的  $y$  的变化部分
- B. 除了  $x$  对  $y$  的线性影响之外的其他因素对  $y$  变差的影响
- C. 由于  $x$  与  $y$  之间的非线性关系引起的  $y$  的变化部分
- D. 由于  $x$  与  $y$  之间的函数关系引起的  $y$  的变化部分

### 三、简答题（共 4 小题，每小题 10 分，共 40 分）

1. 什么是估计量？评价估计量的标准有哪些？
2. 假设检验和参数估计有什么相同点和不同点？
3. 简述回归分析的一般过程。

4. 单因素方差的实质是什么？并说明单因素方差分析的步骤。

四、计算与分析题（共 50 分）

1. 为分析影响消费的因素，随机抽取 2000 个家庭，得到家庭人均消费水平、家庭人均收入、家庭储蓄总额、投资总额等 4 个变量的有关数据。假定要建立一个线性回归模型预测人均消费水平，请你写出建模的基本步骤，并做简要说明。（共 18 分）

2. 在运动会比赛的 10 米气手枪比赛中，每个运动员首先进行预赛，根据预赛总成绩确定进入决赛的 8 名运动员。决赛时 8 名运动员再进行 10 枪射击，然后用预赛成绩加上决赛成绩确定最后的名次。在某一次运动会中，男子 10 米气手枪决赛中，最后获得金牌和银牌的两名运动员 10 枪的决赛成绩如下表所示：（共 15 分）

运动员	决赛成绩（环）									
甲	9.3	10.3	10.5	10.3	10.4	10.3	10.7	10.4	10.7	9.3
乙	9.5	9.9	10.6	10.3	9.4	10.2	10.1	10.8	9.9	9.8

根据该表计算的运动员乙的平均环数是 10.05 环，标准差是 0.445 环。那么，请问：

- (1) 运动员甲、乙两人决赛成绩的中位数是多少？（4 分）
- (2) 请计算运动员甲决赛成绩的平均数和样本标准差？（5 分）
- (3) 请比较分析两位运动员中，哪一位发挥的更稳定？（6 分）

3. 某种酒类饮料的标签上显示：酒精含量 $\leq 10\%$ 。为了检验标签上的说法是否属实，质监部门随机抽取 20 个批次的样本进行检测，得到的结果如下（单位：%）：（共 17 分）

10.1	10.3	10.0	10.0	10.2	9.9	9.8	10.1	10.1	9.9
9.8	9.9	10.3	9.9	10.2	10.1	10.3	10.0	9.8	9.8

- (1) 写出检验的原假设和备择假设。（4 分）
- (2) 检验所使用的分布是什么？使用这一分布的假定条件是什么？（4 分）
- (3) 检验得到的  $P\text{-value}=0.2645$ 。解释  $P$  值的含义。（4 分）
- (4) 根据样本数据得到的 95% 的置信区间为 (9.94, 10.11)，该酒类饮料的酒精含量总体均值是否在这一区间？说明理由。（5 分）