

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）

- 1.统计量
- 2.点估计
- 3.判定系数
- 4.时间序列数据
- 5.非概率抽样

二、单项选择题（共 20 小题，每小题 2 分，共 40 分）

- 1.总体均值为 50，标准差为 8，从此总体中随机抽取容量为 64 的样本，则样本均值的抽样分布的均值和标准误差分别为（ ）。
A.50,1 B.50,8 C.50,4 D.8,8
- 2.某班学生的年龄分布是右偏的，均值为 22，标准差为 4.45。如果采取重复抽样的方法从该班抽取容量为 100 的样本，则样本均值的抽样分布是（ ）。
A.正态分布，均值为 22，标准差为 0.445
B.分布形状未知，均值为 22，标准差为 4.45
C.正态分布，均值为 22，标准差为 4.45
D.分布形状未知，均值为 22，标准差为 0.445
- 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，样本均值的期望值是（ ）。

A.150 B.200 C.100 D.250

4.估计量的含义是指 ()。

- A.总体参数的名称
- B.用来估计总体参数的统计量的具体数值
- C.用来估计总体参数的统计量的名称
- D.总体参数的具体数值

5.95%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.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%
- C.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%
- 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%

6.在总体均值和总体比例的区间估计中, 边际误差由 ()。

- A.置信水平确定
- B.统计量的抽样标准差确定
- C.置信水平和统计量的抽样标准差确定
- D.统计量的抽样方差确定

7.根据 $n = 250$, $p = 0.38$ 的样本计算的样本比例的抽样标准差为 ()。

A.0.031 B.0.016 C.0.043 D.0.052

8.若边际误差 $E = 5$, $\sigma = 40$, 要估计总体均值 μ 的 95%的置信区间所需要的样本量为 ()。

A.146 B.246 C.446 D.346

9.指出下列假设检验哪一个属于左侧检验 ()。

A. $H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$

B. $H_0: \mu \geq \mu_0, H_1: \mu < \mu_0$

C. $H_0: \mu > \mu_0, H_1: \mu \leq \mu_0$

D. $H_0: \mu \leq \mu_0, H_1: \mu > \mu_0$

10. 在一项科学研究中发现, 大学生中吸烟的比例高达 30%, 为检验这一说法是否属实, 建立的原假设和备择假设应为 ()。

A. $H_0: \pi \neq 30\%, H_1: \pi = 30\%$

B. $H_0: \pi = 30\%, H_1: \pi \neq 30\%$

C. $H_0: \pi \geq 30\%, H_1: \pi < 30\%$

D. $H_0: \pi \leq 30\%, H_1: \pi > 30\%$

11. 检验一个正态总体的方差时所使用的分布为 ()。

A. 正态分布 B. t 分布 C. χ^2 分布 D. F 分布

12. 方差分析的主要目的是判断 ()。

A. 各总体是否存在方差

B. 各样本数据之间是否有显著差异

C. 分类型自变量对数值型因变量的影响是否显著

D. 分类型因变量对数值型自变量的影响是否显著

13. 在方差分析中, 某一水平下样本数据之间的误差称为 ()。

A. 随机误差 B. 非随机误差 C. 系统误差 D. 非系统误差

14. 在方差分析中, 多重比较的目的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

15.在方差分析中,用于度量自变量与因变量之间关系强度的统计量是 R^2 , 其计算方法为 ()。

$$A. R^2 = \frac{\text{组间平方和}}{\text{组内平方和}} \quad B. R^2 = \frac{\text{组间平方和}}{\text{总平方和}}$$

$$C. R^2 = \frac{\text{组间方差}}{\text{组内方差}} \quad D. R^2 = \frac{\text{组内平方和}}{\text{总平方和}}$$

16.如果一个变量的取值完全依赖于另一个变量,各观测点落在一条直线上,称为两个变量之间为 ()。

- A.完全相关关系
- B.非线性相关关系
- C.正线性相关关系
- D.负线性相关关系

17.变量 x 与 y 之间的负相关是指 ()。

- A. x 值增大时 y 值也随之增大
- B. x 值减小时 y 值也随之减小
- C. y 的取值几乎不受 x 取值的影响
- D. x 值增大时 y 值随之减少,或 x 值减小时 y 值随之增大

18.在回归分析中,利用估计的回归方程,对于 x 的一个特定值 x_0 , 求出 y 的平均值的一个估计值 $E(y_0)$, 称为 ()。

- A.平均值的置信区间估计
- B.个别值的点估计
- C.平均值的点估计
- D.个别值的预测区间估计

19.已知回归平方和 SSR=4854，残差平方和 SSE=146，则判定系数 $R^2=$ （ ）。
A.3.01% B.2.92% C.97.08% D.33.25%

20.标准化残差图主要用于直观地判断（ ）。

A.回归模型的线性关系是否显著

B.回归系数是否显著

C.误差项 ε 等方差的假定是否成立

D.误差项 ε 服从正态分布的假定是否成立

三、简答题（共 4 小题，每小题 10 分，共 40 分）

- 1.请简述方差分析的目的以及基本假定。
- 2.请简述一元线性回归模型的基本假定。
- 3.请简述假设检验中的两类错误。
- 4.请简述中心极限定理的意义。

四、计算与分析题（共 50 分）

1.在一项调查中获得容量为 20 的样本数据，如下表。

2	9	5	8	10	12	12	21	4	6
4	8	5	7	17	13	11	18	8	3

试求以下内容。（本小题 20 分）

- （1）该数据的中位数位置、中位数、均值是多少？（6 分）
- （2）以 0-4、5-9、10-14、15-19 等为组限，绘制等距式频数分布表及累积频数分布表，
请将表格中字母处填写为相应的数字（如 A=1）。（14 分）

分组	频数	频率（%）	向上累积		向下累积	
			频数	频率（%）	频数	频率（%）
0-4	4	A	B	20	20	100
5-9	C	40	12	D	16	E
10-14	5	F	G	85	H	40
15-19	2	I	19	95	3	J
20-24	K	5	L	M	1	5
合计	N	100	—	—	—	—

2.利用下面的信息，构建总体均值的置信区间。 $z_{0.1/2}=1.645$ ， $z_{0.05/2}=1.96$ ， $z_{0.01/2}=2.58$ 。（本小题 16 分）

（1）总体服从正态分布，已知 $\sigma=500$ ， $n=15$ ， $\bar{x}=8900$ ，置信水平为 95%。（4 分）

（2）总体不服从正态分布，已知 $\sigma=500$ ， $n=35$ ， $\bar{x}=8900$ ，置信水平为 95%。（4 分）

（3）总体不服从正态分布， σ 未知， $n=35$ ， $\bar{x}=8900$ ， $s=500$ ，置信水平为 90%。（4 分）

（4）总体不服从正态分布， σ 未知， $n=35$ ， $\bar{x}=8900$ ， $s=500$ ，置信水平为 99%。（4 分）

3.为分析不同行业的工资水平是否存在显著差异，从制造业、旅游业、农业随机抽取 1000 人，得到的月工资收入数据如下（单位：元）：（本小题 14 分）

序号	制造业	旅游业	农业
1	9905	8335	4698
2	8798	7837	4350
3	11023	5900	3879
4	8506	6803	4755
...
1000	10003	8485	5200

（1）要分析不同行业的工资水平是否显著，采用什么统计方法？（2 分）

（2）该分析中的假定有哪些？简要说明这些假定。（6 分）

（3）简要说明分析过程。（6 分）