

中国海洋大学 2019 年硕士研究生招生考试试题

科目代码: 432

科目名称: 统计学

注:计算结果请保留两位小数,可能用到的临界值如下:

$$t_{0.05}(8) = 1.86, t_{0.025}(8) = 2.31, t_{0.05}(9) = 1.83, t_{0.025}(9) = 2.26,$$
$$\chi_{0.1}^2(8) = 13.36, \chi_{0.05}^2(8) = 15.51, \chi_{0.1}^2(3) = 6.25, \chi_{0.05}^2(3) = 7.82.$$

一、填空题 (共 14 分, 每题 2 分)

1. 在相同或近似相同的地点上收集的数据称为_____.
2. 一组数据的最大值与最小值之差称为_____.
3. 峰态通常是与标准正态分布相比较而言的。如果一组数据服从标准正态分布, 则峰态系数的值等于_____.
4. 在假设检验中, 第一类错误是指_____.
5. 若变量 x 与 y 之间的相关系数 $r = 0.8$, 则线性回归方程的判定系数 $R^2 =$ _____.
6. 设自变量的个数为 5, 样本容量为 20。在多元回归分析中, 估计标准误差的自由度为_____.
7. 如果某月份的商品销售额为 91 万元, 该月的季度指数等于 1.3, 在消除季节因素后该月的销售额为_____万元.

二、单项选择题 (共 18 分, 每题 2 分)

1. 到商场购物停车, 管理人员希望掌握顾客找到停车位的平均时间, 某管理人员跟踪了 50 名顾客, 并记录下他们找到车位的时间, 这管理人员感兴趣的总体是()
A. 管理人员跟踪过的 50 名顾客 B. 在商场停车的所有顾客
C. 上午在商场停车的顾客 D. 到商场购物的所有顾客
2. 与概率抽样相比, 非概率抽样的缺点是 ()
A. 样本统计量的分布是确定的
B. 调查的成本比较高
C. 不适合搜索性研究
D. 无法使用样本的结果对总体相应的参数进行推断
3. 设多元线性回归方程为 $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_k x_k$, 若自变量 x_i 的回归系数 $\hat{\beta}_i$ 的取值接近 0, 这表明 ()

特别提醒: 答案必须写在答题纸上, 若写在试卷或草稿纸上无效。

- A. 因变量 y 对自变量 x_i 的影响不显著 B. 因变量 y 对自变量 x_i 的影响显著
 C. 自变量 x_i 对因变量 y 的影响不显著 D. 自变量 x_i 对因变量 y 的影响显著
4. 在方差分析中, 所提出的原假设是 $H_0: \mu_1 = \mu_2 = \dots = \mu_k$, 则备择假设是 ()
- A. $H_1: \mu_1 \neq \mu_2 \neq \dots \neq \mu_k$ B. $H_1: \mu_1 > \mu_2 > \dots > \mu_k$
 C. $H_1: \mu_1 < \mu_2 < \dots < \mu_k$ D. $H_1: \mu_1, \mu_2, \dots, \mu_k$ 不全相等
5. 在假设检验中, 不拒绝原假设意味着 ()
- A. 原假设肯定是正确的 B. 原假设肯定是错误的
 C. 没有证据证明原假设是正确的 D. 没有证据证明原假设是错误的
6. 假设总体比例为 0.4, 采取重复抽样的方法从此总体中抽取一个容量为 100 的简单随机样本, 则样本比例的期望值是 ()
- A. 0.3 B. 0.4 C. 0.5 D. 0.45
7. 如果两个变量之间完全线性相关, 则以下结论中不正确的是 ()
- A. 相关系数 $r=1$ B. 判定系数 $R^2=1$
 C. 回归系数 $\beta=0$ D. 估计标准差等于 0
8. 方差分析中, 检验行因素对因变量的影响是否显著构造的统计量是 ()
- A. $F_R = \frac{MSR}{MSE}$ B. $F_C = \frac{MSC}{MSE}$ C. $F_R = \frac{SSR}{SSE}$ D. $F_C = \frac{SSA}{SST}$
9. 若两个变量存在正线性相关关系, 则建立的一元线性回归方程的判定系数 R^2 的取值范围是 ()
- A. $[-1, 1]$ B. $[-1, 0]$ C. $[0, 1]$ D. 小于 0 的任意数

三、多项选择题 (共 20 分, 每题 4 分)

1. 指出下面的陈述中正确的是 ()
- A. 抽样误差只存在于概率抽样中
 B. 非抽样误差只存在于非概率抽样中
 C. 无论是概率抽样还是非概率抽样都存在非抽样误差
 D. 在全面调查中也存在非抽样误差
 E. 以上全正确
2. 指出下列是回归分析能解决的问题 ()
- A. 从一组样本数据出发, 确定变量之间的数学关系式
 B. 对数学关系式的可信程度进行各种统计检验, 并从影响某一特定变量的诸多变量

特别提醒: 答案必须写在答题纸上, 若写在试卷或草稿纸上无效。

中找出哪些变量的影响是显著的，哪些是不显著的

- C. 利用所求的关系式，根据一个或几个变量的取值来估计或预测另一个特定变量的取值
- D. 度量两个变量之间的关系程度
- E. 度量各个变量之间的关系程度
3. 下面的陈述不正确的是 ()
- A. 若 F 检验表明回归方程的线性关系显著，则每个自变量同因变量的关系都显著
- B. 若 F 检验表明回归方程的线性关系显著，则每个自变量同因变量的关系都不显著
- C. 若 F 检验表明回归方程的线性关系显著，则至少有一个自变量同因变量的关系都显著
- D. 若 F 检验表明回归方程的线性关系显著，则至少有一个自变量同因变量的关系不显著
- E. 以上均不正确
4. 对于一元线性回归方程的检验，可以 ()
- A. t 检验
- B. F 检验
- C. t 检验与 F 检验的结论是一致的
- D. t 检验与 F 检验的结论是不同的
- E. 用判定系数
5. 进行方差分析需要满足的假设条件是 ()
- A. 各总体服从正态分布
- B. 各总体方差随因素水平增加而增加
- C. 各总体方差相等
- D. 各总体方差随因素水平减少而减少
- E. 样本间相互独立

四、简答题 (共 18 分，每题 6 分)

1. 简述相关分析与回归分析的关系。
2. 简述平稳时间序列的预测方法。
3. 参数估计和假设检验的区别和联系。

五、计算分析题 (7 小题，共 80 分)

1. (15 分) 某银行现有两种办理业务的流程，为了比较哪种流程更为节省时间，从两种流程中各随机抽取 9 个服务对象，得到第一种流程的平均花费时间为 6.2 分钟，标准差为 1.07 分钟。第二种流程中，9 个服务对象的花费时间 (单位：分钟) 如下：

5.2 6.3 5.7 8.2 4.6 5.8 7.4 7.2 6.9

- (1) 计算第二种流程中花费时间的平均值和标准差。

特别提醒：答案必须写在答题纸上，若写在试卷或草稿纸上无效。

- (2) 比较两种流程中花费时间的离散程度。
- (3) 如果让你选择一种服务流程，你会选择哪一种？试说明理由。
- (4) 若假设服务对象所花费的时间 X 是正态分布，利用抽样的 9 个样本建立第二种流程下花费的时间 X 的平均值的置信区间（置信水平为 95%）。
2. (10 分) 假定有 10 种投资项目，其中 3 种项目投资后可以获利，其余 7 种将会亏损。如果从这 10 种项目中随机选出 4 种进行投资，求所有 3 种可以获利的项目都被选中的概率是多少？至少有 2 种获利的项目被选中的概率是多少？
3. (10 分) 一种电子元件由四家工厂提供，产品的质量被分成两个等级。从这批电子元件中随机抽取 420 个进行检验，结果如下表。请检验元件的质量是否与生产厂家有关（显著性水平为 0.1）。

	一厂	二厂	三厂	四厂	合计
一等品	68	75	57	79	279
二等品	32	45	33	31	141
合计	100	120	90	110	420

4. (10 分) 从某行业随机抽取 10 家企业，对其产量 y 和生产费用 x 进行调查，并对 10 组数据进行分析，得到部分结果如下：

	df	SS	MS	F	Significance F
回归		1160.26			0.016
残差				—	—
总计	9	2162.40	—	—	—

- (1) 请将方差分析表中缺少的数据补全。
- (2) 根据表格数据计算产量 y 和生产费用 x 之间的样本相关系数。
5. (10 分) 表 1 是对变量 Y 、 X_1 和 X_2 随机抽样得到的数据，利用两个变量 X_1 和 X_2 建立 Y 的回归方程，输出的参数估计表（表 2）如下：

特别提醒：答案必须写在答题纸上，若写在试卷或草稿纸上无效。

Y	96	93	85	98	84	92	91	84	82	96
X_1	5	5.2	4	6	3.3	4	4.2	4.1	4.1	5
X_2	3.2	3.3	4.1	3.8	2.3	4.1	5.3	2.3	3.9	2.7

	Coefficients	标准误差	t Stat	P-value
Intercept	61.03	8.48	7.19	0.000179
X Variable 1	6.06	1.60	3.77	0.006913
X Variable 2	0.52	1.35	0.38	0.711954

(1) 计算 Y 和 X_1 的样本相关系数, 并在显著性水平为 0.05 下, 检验 Y 和 X_1 的线性关系是否显著。

(2) 根据参数估计表对多元线性回归的结果进行分析和解释。

6. (10 分) 1991 年至 2000 年某市粮食产量如下表,

年份	粮食产量 (万吨)	年份	粮食产量 (万吨)
1991	6.22	1996	8.94
1992	6.66	1997	9.28
1993	7.15	1998	9.80
1994	7.89	1999	10.54
1995	8.72	2000	10.80

- 移动间隔 $k=3$, 计算各年的粮食产量预测值, 并预测 2001 年的产量。
- 选择平滑系数为 0.3, 计算各年的粮食产量预测值, 并预测 2001 年的产量。
- 根据最小二乘法确定直线趋势方程, 并计算趋势预测的误差。

7. (15 分) 某地区三种农作物的生产情况如下表, 分析该地区农作物总的平均产量变动情况及其原因。

农作物	播种面积 (公顷)		平均产量 (千克/公顷)	
	2004 年	2005 年	2004 年	2005 年
A	2000	1000	200	200
B	2200	2000	300	315
C	4000	4600	400	440

特别提醒: 答案必须写在答题纸上, 若写在试卷或草稿纸上无效。