

安徽工业大学 2008 年招收攻读硕士学位研究生专业课试卷 (A)

课程名称 统计学原理

代码 882

一、单项选择题 (从下列每小题的四个选项中, 选出一个正确的, 每小题 3 分, 共 30 分)

1、某校经济管理类的学生学习《统计学》的时间 x 与考试成绩 y 之间建立线性回归方程 $y = \beta_0 + \beta_1 x$ 。经计算, 方程为 $y = 20 - 0.8x$, 该方程参数的计算()。

- A. β_0 值是明显不对的 B. β_1 值是明显不对的
C. β_0 值和 β_1 值都是不对的 D. β_0 值和 β_1 值都是正确的

2、在简单随机重复抽样条件下, 当抽样平均误差缩小为原来的 $1/2$ 时, 则样本单位数为原来的()。

- A. 2 倍 B. 3 倍
C. 4 倍 D. $1/4$ 倍

3、在统计学的形成和发展过程中, 首先将古典概率论引入社会经济现象研究的学者是()。

- A. 阿道夫·凯特勒 B. 威廉·配第
C. 约翰·格朗特 D. 海尔曼·康令

4、社会经济统计学是一门()。

- A. 自然科学 B. 新兴科学
C. 方法论科学 D. 实质性科学

5、在其它条件不变的情况下, 提高估计的概率保证程度, 其估计的精确程度()。

- A. 随之扩大 B. 随之缩小
C. 保持不变 D. 无法确定

6、已知各期的环比增长速度为 10%、15%、18%, 则定基增长速度为()。

- A. $10\% \times 15\% \times 18\%$ B. $10\% \times 15\% \times 18\% - 100\%$
C. $110\% \times 115\% \times 118\%$ D. $110\% \times 115\% \times 118\% - 100\%$

7、如果现在进行左侧检验, 检验概率 p 值表示当原假设成立时, 样本可能的结果() 实际观测结果的概率。

- A. 不高于 B. 不低于
C. 等于 D. 不等于

8、某企业的职工工资水平比上年提高 5%, 职工人数增加 2%, 则企业工资总额增长()。

- A. 10% B. 7.1%
C. 7% D. 11%

9、几何平均数主要适用于计算()。

- A. 具有等差关系的数列
B. 变量值的连乘积等于总比率或总速度的数列
C. 变量值为偶数项的数列
D. 变量值的连乘积等于变量值之和的数列

10、某企业生产某种产品, 其产量年年增加 5 万吨, 则该产品产量的环比增长速度()。

- A. 年年下降 B. 年年增长
C. 年年保持不变 D. 无法做结论

二、简答题 (每小题 10 分, 共 40 分)

- 1、什么是统计分组？统计分组的作用有哪些？
- 2、统计指数的作用主要表现在哪些方面？
- 3、什么是抽样推断？抽样推断都有哪几方面的特点？
- 4、相关分析与回归分析有何区别？

三、计算分析题（第 2、3、4 题要写出计算公式、计算过程，结果保留 2 位小数。每小题 12 分，共 60 分）

1、某车间 120 人日生产产品 478 件，具体情况如下表所示：

日产量（件） x	1	2	3	4	5	6	7	8
人数（个） f	5	12	20	38	25	10	8	2

- (1) 日产量的中位数是（ ）。
A. 4 B. 4.5 C. 5 D. 8
- (2) 日产量的众数是（ ）。
A. 4 B. 4.5 C. 5 D. 8
- (3) 该车间工人日产量全距是（ ）。
A. 4.5 B. 4.5 C. 8 D. 7
- (4) 该车间工人日产量的平均差是（ ）。
A. 4 B. 478/120 C. 2
D. $\frac{|1-4.15|\times 5+|2-4.15|\times 12+\dots+|8-4.15|\times 2}{120} = \frac{140.5}{120} = 1.17$
- (5) 该车间工人日产量的方差与标准差是（ ）。
A. 4, 2 B. 4.5, 2.1 C. 2, 1.14
D. $\sigma = \sqrt{\frac{\sum (x-\bar{x})^2 f}{\sum f}}$, $\sigma^2 = \frac{\sum (x-\bar{x})^2 f}{\sum f}$
- (6) 该车间工人日产量的平均差系数是（ ）。
A. $\frac{1.17}{4.15}$ B. $\frac{2}{4.15}$ C. $\frac{4}{4.15}$ D. $\frac{4.5}{4.5}$

2、某工厂生产一种新型灯泡 5000 只，随机抽取 100 只作耐用时间试验。测试结果，平均寿命为 4500 小时，标准差 300 小时。

- (1) 试在 95.45% 概率保证下 ($t=2$)，估计该新式灯泡平均寿命区间；
- (2) 假定概率保证程度提高到 99.73% ($t=3$)，允许误差缩小一半，试问应抽取多少只灯泡进行测试？
(均用重置抽样公式计算)

3、某公司三种商品销售额及价格变动资料如下：

名称	商品销售额（万元）		价格变动率(%)
	基期	报告期	
甲	500	650	2
乙	200	200	-5
丙	1000	1200	10

计算三种商品价格总指数和销售量总指数。并分析销售量和价格因素变动对零售总额变动的绝对影响值。

4、某部门所属 20 个企业全员劳动生产率 (x) 与销售利润 (y) 的调查资料经初步加工

整理如下： $n = 20$ $\Sigma x = 30.8$ $\Sigma y = 961.3$ $\Sigma xy = 1652.02$ $\Sigma x^2 = 52.44$

$$\Sigma y^2 = 65754.65$$

要求：(1) 计算全员劳动生产率与销售利润之间的相关系数，并分析相关的密切程度和方向。

(2) 建立销售利润为因变量、全员劳动生产率为自变量的直线回归方程。

5、根据动态指标的相互关系，计算下表中所缺指标数值，填入相应的空格内，并计算平均增长速度。

年份	棉布产量 (百万米)	与上年比较		
		增长量	发展速度%	增长速度%
2003	95.2	——	——	——
2004		4.8		
2005			104	
2006				5.8

四、抽样方案设计分析及分析（每小题 5 分，共 20 分）

某钢铁公司有 50 个企业共 50000 名职工，现欲抽取 500 名职工进行职工文化技术素质检查，为此调查人员设计了以下两个调查方案：

方案一：采用简单随机抽样，在 50000 名职工中随机抽取 500 名进行调查；

方案二：采用分层随机抽样，根据每个企业所具有的人数按比例随机抽取 500 名学生进行调查。

问题：(1) 方案一和方案二的抽样框各是什么；

(2) 分别给出两方案下，全体职工中具有高等学历（或高级职称）的抽样误差（用符号表示）；

(3) 你认为上述两个方案哪个较好，说明理由；

(4) 抽样设计的原则是什么。

安徽工业大学 2008 年招收攻读硕士学位研究生专业课

《统计学原理》试卷(A)

答案及评分标准(供参考)

一、单项选择题(从下列每小题的四个选项中,选出一个正确的,每小题 3 分,共 30 分)

1. B 2. C 3. A 4. C 5. B
6. D 7. A 8. B 9. B 10. A

二、简答题(每小题 10 分,共 40 分)

1、统计分组,是指根据统计研究的目的,按照一定的标志将总体划分为若干个性不同的部分的一种统计方法。(2 分)统计分组同时具有两个方面的含义:对总体而言是“分”,即将整体划分为性质相异的若干部分;对总体单位而言是“合”,即将性质相同的总体单位合并到同一组。(2 分)

统计分组的作用主要有:(1)区分社会经济现象的类型。(2 分)(2)研究总体的内部结构及其变化。(2 分)(3)探讨现象之间的依存关系。(2 分)

2、(1)综合反映复杂现象总体数量上的变动状态;(3 分)

(2)分析现象总体变动中受各个因素变动的的影响程度;(3 分)

(3)利用连续编制的指数数列,对复杂现象总体长时间发展变化趋势进行分析。(2 分)

3、抽样推断是在抽样调查的基础上,利用样本的实际资料计算样本指标,并据以推算总体相应数量特征的统计分析方法。(2 分)特点:(1)是由部分推算整体的一种认识方法论;(2 分)(2)建立在随机取样的基础上;(2 分)(3)运用概率估计的方法;(2 分)(4)抽样推断的误差可以事先计算并加以控制。(2 分)

4、回归和相关都是研究两个变量相互关系的分析方法。相关分析研究两个变量之间相关的方向和相关的密切程度。(2 分)回归方程则是通过一定的数学方程来反映变量之间相互关系的具体形式,以便从一个已知量来推测另一个未知量。为估算预测提供一个重要的方法。(2 分)

相关分析既可以研究因果关系的现象也可以研究共变的现象,不必确定两变量中谁是自变量,谁是因变量。而回归分析是研究两变量具有因果关系的数学形式,因此必须事先确定变量中自变量与因变量的地位。(2 分)

计算相关系数的两变量是对等的,可以都是随机变量,各自接受随机因素的影响,改变两变量的地位并不影响相关系数的数值。(2 分)在回归分析中因变量是随机的,自变量是可控制的解释变量,不是随机变量。因此回归分析只能用自变量来估计因变量,而不允许由因变量来推测自变量。(2 分)

三、计算分析题(第 2、3、4 题要写出计算公式、计算过程,结果保留 2 位小数。每小题 12 分,共 60 分)

1、(1) A (2) A (3) D (4) D (5) D (6) A (各 2 分)

2、解: $n=100$ $\bar{x}=4500$ $c=300$ $t=2$ (2 分)

$$(1) \mu_x = \frac{c}{\sqrt{n}} = \frac{300}{\sqrt{100}} = 30 \quad (2 \text{ 分})$$

$$\Delta x = t\mu_x = 2 \times 30 = 60 \quad (2 \text{ 分})$$

该新式灯泡的平均寿命的区间范围是：

$$\bar{x} - \Delta x \leq \bar{X} \leq \bar{x} + \Delta x \quad (2 \text{ 分})$$

$$4500 - 60 \leq \bar{X} \leq 4500 + 60$$

$$4440 \leq \bar{X} \leq 4560 \quad (2 \text{ 分})$$

$$(2) \quad n = \frac{t^2 \sigma^2}{\Delta_x^2} = \frac{3^2 \times 300^2}{\left(\frac{60}{2}\right)^2} = 900 \quad (4 \text{ 分})$$

应抽取 900 只灯泡进行测试。

3、解：三种商品物价总指数：

$$\frac{\sum q_1 p_1}{\sum \frac{1}{k} q_1 p_1} = \frac{650 + 200 + 1200}{\frac{650}{1+2\%} + \frac{200}{1-5\%} + \frac{1200}{1+10\%}} = \frac{2050}{1938.69} = 105.74\% \quad (4 \text{ 分})$$

由于销售量变动对零售总额变动的绝对影响值 $2050 - 1938.69 = 111.31$ (万元) (2 分)

销售量总指数 = 销售额指数 ÷ 价格指数

$$= \frac{\sum q_1 p_1}{\sum q_0 p_0} \div \frac{\sum q_1 p_1}{\sum \frac{1}{k} q_1 p_1} = \frac{650 + 200 + 1200}{500 + 200 + 1000} \div 105.74\% = 114.04\% \quad (4 \text{ 分})$$

由于价格因素变动对零售总额变动的绝对影响值 $1938.69 - 1700 = 238.69$ (万元) (2 分)

4、(1) 全员劳动生产率与销售利润的相关系数：

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{20 \times 52.44 - 30.8^2} \sqrt{20 \times 65754.65 - 961.3^2}} = 0.55 \quad (3 \text{ 分})$$

可以看出，全员劳动生产率与销售利润之间存在着的正相关关系。 (2 分)

(2) 设销售利润倚全员劳动生产率的直线回归方程为 $y_c = a + bx$

$$b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2} \quad (2 \text{ 分})$$

$$= \frac{20 \times 1652.02 - 30.8 \times 961.3}{20 \times 52.44 - 30.8^2} = 34.27 \quad (2 \text{ 分})$$

$$a = \frac{1}{n} \sum y - b \frac{1}{n} \sum x$$

$$= \frac{1}{20} \times 961.3 - \frac{1}{20} \times 30.8 \times 34.27 = -4.71 \quad (2 \text{ 分})$$

故销售利润倚全员劳动生产率的直线回归方程为 $y_c = -4.71 + 34.27x$ (1 分)

5、

年份	棉布产量 (百万米)	与上年比较		
		增长量	发展速度%	增长速度%
2003	95.2	—	—	—
2004	100	4.8	105.04	5.04
2005	104	4	104	4.0
2006	110.032	6.032	105.8	5.8

(6分)

$$\text{平均增长速度} = \sqrt[3]{1.0504 \times 1.04 \times 1.058} - 1 = 0.0495 \quad (6 \text{分})$$

四、抽样方案设计及分析 (每小题 5 分, 共 20 分)

答案要点: (1) 10000 名学生名单; 20 个学院名单与 10000 名学生名单。

(2) 简单随机抽样: 第一步, 根据样本资料计算样本出勤率和标准差, 并推算抽样平均误差。公式略。

第二步, 根据给定的置信度 $F(t)$ 。

第三步, 根据概率度和抽样平均误差计算抽样极限误差, 并估计总体出勤率的上下限。

$$\Delta_p = t\mu_p$$

$$\text{下限} = p - \Delta_p$$

$$\text{上限} = p + \Delta_p$$

分层随机抽样: 第一步, 在各层分别取样, 可以计算各层抽样平均数。

$$\bar{P}_i = \frac{\sum_{j=1}^{n_i} P_{ij}}{n_i} \quad (i=1,2,\dots,20)$$

第二步, 将各层样本平均数以各层样本单位数或总体单位数为权数加权平均, 即为所求的样本平均数。

分层抽样的抽样平均误差的计算步骤为:

第一步, 计算各层内方差。

$$\sigma_i^2 = \frac{\sum (P_i - \bar{P}_i)^2}{N_i} \quad (i=1,2,\dots,20)$$

第二步, 以各层样本单位数为权数, 计算各层内方差的平均数。

$$\bar{\sigma}_i^2 = \frac{\sum \sigma_i^2 n_i}{n}$$

第三步, 计算抽样平均误差。

$$\text{重复抽样 } \mu_p = \sqrt{\frac{P(1-P)}{n}}$$

$$\text{不重复抽样 } \mu_p = \sqrt{\frac{P(1-P)}{n} \left(1 - \frac{n}{N}\right)}$$

(3) 分层抽样。分析略

(4) 在抽样设计中，首先要保证随机原则的实现。

其次，要考虑样本容量和结构问题。

再次，要认识到不同的抽样组织形式，会有不同的抽样误差，因而抽样的效果也是不同的。

最后，在抽样设计中必须重视调查费用这个基本因素。