

南京审计大学  
2018-2019 学年第 1 学期《统计学》试题

一、单项选择题（每题 2 分，满分 10 分）

1. 为了解 2018 年首届中国国际进口博览会（CIIE）40 多万中外采购商对该届博览会的基本看法，某研究机构于展会期间编制了一份调查问卷，分别从国籍、年龄、性别、职业、文化程度、年收入，以及对一些特定问题（注：仅看作一个问题）的看法（如“非常满意”“满意”“非常不满意”或“非常赞同”“赞同”“非常不赞同”等）等方面，对随机抽取的 500 位客商进行了调查。在上述调查变量中，用于搜集分类数据的有

- A. 1 项      B. 2 项      C. 3 项      D. 4 项

2. 若想用一张统计图展示 2017 年我国第一产业、第二产业和第三产业增加值及其所占国内生产总值的比重情况，则下列最适宜的图形是

- A. 直方图    B. 散点图    C. 圆饼图    D. 条（柱）型图

3. 下列说法正确的有

- (1) 小概率事件原理是指小概率事件在试验中几乎不可能发生  $\times$   
 (2) 假设检验中，显著性水平设置得越大，表明研究者越想拒绝原假设  $\alpha$   
 (3) 在其他条件相同的情况下，置信水平越高，所求总体均值的置信区间越短  $\times$

- A. 0 项    B. 1 项    C. 2 项    D. 3 项    1-2

4. 下列说法正确的有

- (1) 利用某高校学生阅读手机时间和其已获学分绩点，可以建立一元回归模型  $\checkmark$   
 (2) 利用某地历年第一、第二、第三产业增加值和地区生产总值，可以建立三元回归模型  $\times$   
 (3) 利用某地各高校大一、大二、大三、大四、研究生人数和学生总人数，可以建立五元回归模型  $\times$

- A. 0 项    B. 1 项    C. 2 项    D. 3 项

5. 下列说法正确的有

- (1) 统计图的标题通常应置于图的正下方  $\checkmark$   
 (2) 一段统计文字中的数据通常应保留相同的小数位数  $\times$   
 (3) 若统计表中所有数据的计量单位相同，则通常应将该单位置于表的右上角  $\checkmark$

- A. 0 项    B. 1 项    C. 2 项    D. 3 项

二、判断题（每题 2 分，满分 10 分）

1. 调查标准时点为 2018 年 12 月 31 日的我国第四次经济普查，调查时不会产生非抽样误差。  $\times$

2. 在单因素方差分析中，因素 A 有  $A_1$ 、 $A_2$ 、 $A_3$ 、 $A_4$ 、 $A_5$  5 个水平，且在这 5 个水平下分别进行了 7

$$\frac{SSA}{5-1}$$

$\times$  次、8 次、11 次、12 次和 15 次试验。对试验所得 53 个数据的计算表明：水平（组）间的数据误差是 30，总误差是 150，则检验统计量  $F$  的值是 3.5。

$$F = \frac{MSA}{MSE} \quad SSA = 30 \quad SSE = 120$$

3. 只要对总体中的所有单位一一进行调查，即使仅仅只调查一个事项也是全面调查。

$\times$  4. 在建立回归模型时，通常要求作为自变量的必须是随机变量，作为因变量的必须不是随机变量。

$\times$  5. 根据一个时间序列计算的环比增长速度之连乘积等于相应的定基增长速度，相邻的两个定基增长速度之商等于相应的环比增长速度。

$$y = b_0 + b_1 x + \varepsilon \quad \varepsilon \sim N(0, \sigma^2) \quad y \sim N(\Delta, \Delta)$$

三、填空题（每题 2 分，满分 10 分）

1. 要研究南京审计大学学生参与课外体育锻炼的有关问题，调查总体应为\_\_\_\_\_。

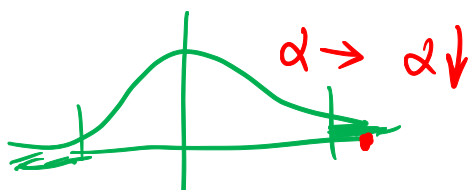
2. 2013-2017 年某省经济增长速度分别为 7.8%、8.5%、8.3%、8.1%和 7.6%，则 2013-2017 年该省经济年平均增长速度是 8.0534 %。（保留 4 位小数）

$$\sqrt[5]{1.078 \times 1.085 \times 1.083 \times 1.081 \times 1.076}$$

3. 若采用移动平均法，测定 2013 年 1 月至 2018 年 11 月某大型啤酒生产企业啤酒产量的长期趋势，则移动平均的项数（步长）至少应为 12 项。

$$-1 = 8.0534\%$$

4. 在对一个总体的均值进行检验时，若在取较大的显著性水平时拒绝了原假设，那么在其他条件保



持不变的情况下，仅将显著性水平变得小一些，则 可能... 也可能不拒绝...

5. 2018 年 10 月某地猪肉价格同比上涨 2.5%，则 2018 年 10 月该地一个居民家庭用 450 元购买的猪肉，若在上年同期购买最有可能少支出 10.99 元（保留 2 位小数）。

四、简答题（每题 5 分，满分 10 分）

$$x(1+2.5\%) = 450$$

$$x = 439.024$$

1. 请简述分层随机抽样组织方式组织实施的基本步骤。

- ① 按某种特征或规则划分不同的层
- ② 根据样本容量确定每层... 抽样... 数目
- ③ 从不同的层随机、独立抽取样本
- ④ ... 结合...

2. 若已掌握 2018 年 11 月 11 日，南京审计大学所有学生当日午饭消费支出的数据及其当日阅读手机花费时间的数据，则请简述比较这两组数据差异程度大小的基本步骤。

- ① ...  $\bar{x}$
  - ② ...  $s^2$  ...  $s$
  - ③  $\frac{s}{\bar{x}}$
  - ④ 变异系数越小，... 小。
- 离散系数 / 变异系数  $\frac{s}{\bar{x}}$

五、计算分析题（本大题有 3 题，第 1 题 30 分、第 2 题 8 分、第 3 题 12 分，共 50 分）

1. （本题满分 30 分）2018 年 10 月，某研究人员在研究大学生在校消费的有关问题时，采用简单随机重复抽样组织方式，从南京 S 高校 18500 名学生中随机抽取 1400 人并设法调查其 2018 年 10 月某日在校消费支出及其性别、年级、专业和生源地等数据。其中，抽选出的这 1400 人已按其 2018 年 10 月该日的在校消费支出进行了分组（见表 1）。

表 1 2018 年 10 月某日 S 高校学生在校消费支出情况

按该日校消费支出（元）分组	人数（人）	占比（%）
15 以下 <u>15-2.5-12.5</u>	98 人	7%
15-20 <u>17.5</u>	182 人	13%
20-25 <u>22.5</u>	728 人	52%
25-30 <u>27.5</u>	308 人	22%
30-40 <u>35</u>	56 人	4%
40 以上 <u>45</u>	28 人	2%
合计	1400 人	100%

要求：（1）作为一张统计表，表 1 中主要不规范之处有哪些？（3 分） 上线、下线加粗

（2）该研究人员调查大学生“生源地”的主要目的可能是什么？（2 分）

（3）指出该次调查的样本和样本（容）量：（4 分） 1400

（4）计算表中学生该日在校消费支出的样本中位数、样本均值和样本标准差；（保留 1 位小数）（9 分）

（5）测算 2018 年 10 月该日 S 高校学生在校平均消费支出 95% 的置信区间。（5 分）

（6）在  $\alpha = 0.05$  时，2018 年 10 月该日 S 高校学生在校消费支出超过 30 元的人数占比会大于 5% 吗？

（4）中位数组 20-25 （7 分）

$$B + \frac{\frac{n}{2} - S_{m-1}}{f_m} \times C = 20 + \frac{700 - 280}{728} \times 5 = 22.88 \approx 22.9$$

$$\bar{x} = \frac{\sum m_i f_m}{\sum f_m} = \frac{12.5 \times 98 + \dots + 45 \times 28}{1400} = 23.2$$

第 2 页 共 4 页

$$s^2 = \frac{(12.5 - 23.2)^2 \cdot 98 + \dots + (45 - 23.2)^2 \cdot 28}{1399} = 31.6576$$

$$s = 5.6$$

$$(5) \bar{x} \pm t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \quad \times$$

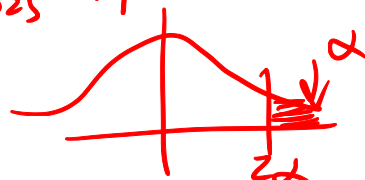
$$\bar{x} \pm Z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \quad 1-\alpha=0.95 \quad Z_{0.025}=1.96$$

$$(22.9053, 23.4947)$$

$$(6) H_0: \pi \leq 5\%$$

$$Z = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \sim N(0,1)$$

$$p \sim N(\pi, \frac{\pi(1-\pi)}{n})$$



$$Z_{0.05} = 1.64$$

$$Z_0 = \frac{6\% - 5\%}{\sqrt{\frac{0.05(1-0.05)}{n}}} = 1.7168 \quad \text{拒绝 } H_0$$

2. (本题满分8分) 2018年10月下旬、中旬以及2017年10月下旬, 南京某超市三种商品销售数据见表2. (注: 百分号前数值保留1位小数, 绝对数保留整数!)

表2 南京某超市三种商品的销售情况

商品名称	计量单位	销售价格(元)			销售量		
		2018年 10月下旬	2018年 10月中旬	2017年 10月下旬	2018年 10月下旬	2018年 10月中旬	2017年 10月下旬
粳米	千克	6.0	5.5	5.7	2500	2510	2550
花生油	升	16.5	15.5	15.2	1500	1100	1000
五花肉	千克	27.8	24.5	23.5	4800	4600	4500

试问: (1) 2018年10月下旬, 南京该超市上述三种商品销售价格同比上涨了多少? 价格变动对销售额的影响是多少? (4分)

(2) 2018年10月下旬, 南京该超市上述三种商品销售量环比增长了多少? 销售量的变动对销售额的影响是多少? (4分)

$$(1) I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{6 \times 2500 + 16.5 \times 1500 + 27.8 \times 4800}{5.7 \times 2500 + \dots + 23.5 \times 4800} = 1.1357$$

$$\text{上涨 } 13.6\%$$

$$\sum p_1 q_1 - \sum p_0 q_1 = \dots = 23340 (\text{元})$$

$$(2) I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0} = \frac{5.5 \times 2500 + 15.5 \times 1500 + 24.5 \times 4800}{5.5 \times 2510 + 15.5 \times 1100 + 24.5 \times 4600} = 1.0769 \quad \text{增长 } 7.7\%$$

$$\sum q_1 p_0 - \sum q_0 p_0 = \dots = 11045 (\text{元})$$

3. (本题满分 12 分) 2018 年 10 月, 某研究人员为研究江苏规模以上工业企业生产经营的有关问题, 收集了 1998—2017 年江苏规模以上工业企业的利润总额及其主营业务收入等方面的数据 (见表 3); 并准备利用它们之间的关系建立一元线性回归模型  $y = \alpha + \beta x + \varepsilon$ ; 该模型参数的估计值及其他相关检验值的 Excel 软件输出结果如下。

表 3 1998—2017 年江苏规模以上工业企业主要经营数据

单位: 亿元

年份	利润总额	主营业务收入	年份	利润总额	主营业务收入
1998	152	7375	2008	3973	66482
1999	234	8256	2009	4100	71725
2000	370	9971	2010	5971	91077
2001	420	11248	2011	7074	107030
2002	554	13535	2012	7250	119287
2003	794	18020	2013	8380	133606
2004	1111	24492	2014	9057	141956
2005	1385	32098	2015	9687	147074
2006	1907	41015	2016	10574	156591
2007	2766	52594	2017	10360	155000

注: 表中数据见《江苏统计年鉴-2017》。

Excel 软件输出结果:

回归统计

Multiple R 0.99788

R Square ①

Adjusted R Square 0.99554

标准误差 ②

观测值 20

方差分析

	Df	SS	MS	F	Significance F
回归分析	1	279845529	279845529	4240.727	7.999E-23
残差	③ 18	④ 1187820	③	-	-
总计	19	281033349	-	-	-

	Coefficients	标准误差	t-Stat	P-value	Lower 95%	Upper 95%
Intercept	-524.7261 $\beta_0$	93.8194	-5.593	2.621E-05	-721.8334	-327.6188
Variable 1	0.0686 $\beta_1$	④	⑤ 65.121	⑥ 7.999E-23	0.0664	0.0708

请根据以上所给信息, 解答以下问题:

(1) 研究人员建立该模型所选择的因变量是\_\_\_\_\_ (2 分)

(2) 请写出 Excel 软件输出结果中在划线处应填写的数值: (4 分)

(3) 可决 (判定) 系数是 0.99577, 它表明 99.577%; (2 分)

(4) 该模型对应的回归方程为  $\hat{y} = -524.7261 + 0.0686x$ ; (2 分)

(5) 在显著性水平为 0.01 时, 自变量对因变量有高度线性影响吗? 为什么? (2 分)

$$p = 7.999E-23 < 0.05$$

## 六、案例分析题（每题 2 分，满分 10 分）

2018 年前三季度，我国规模以上工业原煤产量同比增长 5.1%，增速比上年加快 1.2 个百分点；原油产量下降 1.9%，收窄 0.1 个百分点；天然气产量增长 6.2%，加快 1.6 个百分点；发电量增长 7.4%。其中，水电、核电、风电、太阳能发电等清洁能源发电量增长 8.7%，占全部发电的比重为 26.7%。2016 年 3 季度至 2018 年 3 季度我国规模以上工业发电量情况见表 4。

表 4 2016 年 3 季度至 2018 年 3 季度我国规模以上工业发电量情况 单位：亿千瓦时

时期	火力	水力	核能	风力	太阳能
2016 年 3 季度	11710	3297	580	433	117
4 季度	11669	2410	583	615	102
2017 年 1 季度	11357	1938	548	621	123
2 季度	10858	2675	606	706	166
3 季度	12311	3534	680	587	181
4 季度	11589	2672	646	781	178
2018 年 1 季度	12148	1933	611	872	199
2 季度	11739	2685	689	841	228
3 季度	13031	3708	789	655	234

注：（1）文字来源于 <http://www.chinanews.com/sh/2018/10-30/8663268.shtml>；（2）表中数据据国家统计局网站-数据查询-“月度数据”整理。（3）在下列设问中出现的“清洁能源发电”仅指水电、核电、风电、太阳能发电；“全部发电量”仅指火力、水电、核电、风电、太阳能发电的总和。

1. 2018 年上半年，我国规模以上工业原油产量同比

- A. 下降了 1.9 个百分点      B. 下降了 2.0 个百分点  
C. 增长了 1.8 个百分点      D. 增长了 2.0 个百分点

2. 2018 年前三季度，我国规模以上工业清洁能源发电量同比增加了

- A. 404 亿千瓦时      B. 943 亿千瓦时      C. 1079 亿千瓦时      D. 1124 亿千瓦时

3. 2018 年 3 季度，我国规模以上工业清洁能源发电量同比增速最快的是

- A. 水力      B. 核能      C. 风力      D. 太阳能

4. 2016 年 4 季度至 2018 年 3 季度，我国规模以上工业清洁能源发电量季平均增速最快的是

- A. 水力      B. 核能      C. 风力      D. 太阳能

5. 下列判断正确的有

- （1）2016 年 3 季度至 2018 年 3 季度，我国规模以上工业核能发电量均超过火力发电量的 5%  
（2）2018 年前三季度，我国规模以上工业清洁能源发电量占全部发电量的比重同比有所提高  
（3）2016 年 4 季度至 2018 年 3 季度，我国规模以上工业清洁能源发电量均呈逐季增加的态势

- A. 0 项      B. 1 项      C. 2 项      D. 3 项