

南京审计大学

2020 年硕士研究生招生考试初试（笔试）试题（ A 卷 ）

科目代码： 432

科目名称： 统计学

满分： 150 分

注意：①认真阅读答题纸上的注意事项；②所有答案必须写在答题纸上，写在本试题纸或草稿纸上均无效；③本试题纸须随答题纸一起装入试题袋中交回！

一、简答题（本题共 6 小题，每小题 10 分，计 60 分）

1. 简述统计数据几种常用的分类依据及其分类情况。
2. 简述统计调查中非抽样误差的来源。
3. 简述中位数、众数、算术平均数的特点以及它们之间的关系和各自的应用场合。
4. 在估计总体均值置信区间时，需要确定样本容量。简述在确定样本容量时应考虑的因素及其影响情况。
5. 简述我国消费者价格指数（CPI）编制的基本步骤。
6. 简述采用移动平均趋势剔除法测定月度时间序列季节指数的基本步骤，并解释月度季节指数大小的基本含义。

二、计算分析题（本题共 5 小题，第 1 小题 30 分，其余各小题 15 分，计 90 分）

1. 研究人员采用简单随机重复抽样组织方式，从某地 18—35 岁人群中抽选出 50 人，调查其 2019 年“世界睡眠日”当天的睡眠时间（单位：小时）如下：5.8、4.2、6.2、6.5、6.8、7.4、7.5、7.8、8.1、8.5、4.7、5.6、6.3、5.4、7.1、7.4、7.6、7.8、8.2、8.4、8.2、8.3、5.2、4.3、5.3、8.7、8.4、8.3、5.9、7.8、7.4、6.3、5.7、4.9、8.3、7.6、7.8、6.6、7.2、7.2、7.4、6.9、7.3、7.5、6.5、5.2、6.3、6.4、7.7、7.9。

要求：（计算保留 1 位小数）

- (1) 以 4.0 为最小下限、1.0 为组距，将上述 50 个数据分为 5 组；
 - (2) 绘制规范的统计表和适当的统计图，展示 (1) 的结果；
 - (3) 若需要对这 50 人的睡眠时间进行概括性度量，应主要从哪几个方面进行？
 - (4) 以 (1) 的分组结果为依据，估算这组数据的离散系数；
 - (5) 估计 2019 年“世界睡眠日”当天，该地 18—35 岁人群平均睡眠时间 95% 的置信区间。
2. 海关检验人员对来自 A 国的进口水产品实施入关检验时，发现该水产品说明书标明某成分“每包重量 ≤ 100 克”。根据以往经验，该成分每包重量的标准差为 20 克。如果海关检验人员准备采用简单随机重复抽样组织方式，对来自 A 国的水产品进行检验，要求达到 90% 的置信水平且估计误差不能超过 6 克。
- 要求：
- (1) 若要满足上述要求，海关检验人员至少应抽取多少包 A 国水产品进行检验？
 - (2) 如果海关人员随机抽取了 144 包来自 A 国的水产品进行检验，结果显示该成分的平均重量为 103 克、标准差为 18 克，那么在显著性水平为 0.05 时，来自 A 国的该批水产品能否入关？
- （注：假定该成分达标即可入关）

3. 2018 年 11 月和 2019 年 10—11 月，南京市某农贸市场三种产品的销售情况见表 1。

表 1 南京市某农贸市场三种产品的销售情况

产品名称	计量单位	2018 年 11 月		2019 年 10 月		2019 年 11 月	
		价格(元)	销售量	价格(元)	销售量	价格(元)	销售量
甲	斤	0.5	2000	0.6	2500	0.5	3000
乙	袋	5.0	300	6.0	350	7.0	340
丙	只	10.0	90	10.5	100	11.0	110

要求：

- (1) 2019 年 11 月，南京市该农贸市场上这三种产品的销售价格同比平均上涨了多少？对销售额产生了多大的影响？

(2) 2019年11月,南京市该农贸市场这三种产品的销售量环比平均增长了多少?对销售额产生了多大的影响?

4.某家用净水器销售公司,仅在全国所有地市级城市的电台、电视台和报纸投放产品促销广告,产品销量逐年稳步增加。2019年10月该公司随机抽取了36个城市,得到2019年9月这36个城市分别在电台、电视台和报纸上实际广告费支出(单位:万元)和产品销售量(单位:千台)数据。在假定这三类广告费总支出(X)与产品销售总量(Y)之间存在线性关系的前提下,构建了一个线性回归模型。Excel软件输出的部分结果如下:

回归统计	
Multiple R	0.488
R Square	0.239
Adjusted R Square	0.217
标准误差	13.623
观测值	36

方差分析					
	df	SS	MS	F	Significance F
回归分析	1	2035.529	2035.529	①	0.00216
残差	34	③	②		
总计	35	④			

	Coefficients	标准误差	t Stat	P—Value
Intercept	20.105	14.497	1.387	0.17426
广告费总支出	0.617	0.186	⑤	0.00216

要求:

- 根据所缺数值与其它数值之间的关系,填上所缺的数值或计算式(如⑥= $211+2.36\times 50$);
- 该公司能以在36个城市的电台广告费支出、电视台广告费支出、报纸广告费支出为自变量,以三类广告费总支出为因变量,构建三元回归模型吗?为什么?
- 根据上述Excel软件输出的部分结果,写出相应的回归方程;
- 在显著性水平为0.01时,该公司构建的线性回归模型能通过统计检验吗?为什么?
- 设2019年10月该公司在上述36个城市的三类广告费总支出为100万元,预测在95%的置信水平下,2019年10月该公司在这36个城市的产品销售总量至少可以达到多少?(注:可以按大样本近似)

5.研究人员随机抽取了800名正在购买手机的消费者,调查其购买手机的偏好。调查结果显示:有20%的消费者最看重“品牌”、30%的最看重“功能”、50%的最看重“价格”;若上述调查结果与被调查者“受教育程度”结合再进行交叉分组,则分组结果见表2。

表2 不同受教育程度的消费者购买手机的偏好情况

受教育程度	初中及以下	高中	大专及以上	合计
品牌	30	69	61	160
功能	30	81	129	240
价格	60	150	190	400
合计	120	300	380	800

要求:(计算保留1位小数)

- 在显著性水平为0.05时,消费者购买手机的偏好与其受教育程度有关吗?
- 根据上表计算的Cramer's V的值会大于0.3吗?它能说明什么?

(注: $\chi_{0.025}^2(4) = 11.14$, $\chi_{0.05}^2(4) = 9.49$, $\chi_{0.025}^2(9) = 16.02$, $\chi_{0.05}^2(9) = 16.92$)