

南京审计大学

2019 年硕士研究生入学考试初试（笔试）试题（ B 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（每小题 10 分，共 6 小题，总计 60 分）

- 1、分层抽样与整群抽样有何异同？它们分别适用于什么场合？
- 2、简述概率的古典定义、频率定义和主观定义，并说明其适用场合。
- 3、简述假设检验的基本原理和一般流程。
- 4、什么是判定系数（可决系数）？多元回归分析中为什么需要使用修正的判定系数（可决系数）来比较方程的拟合效果？是如何计算的？
- 5、简述指数平滑法的基本含义及平滑系数的选择方法。
- 6、简述平均数变动因素分解的基本思想。

二、计算分析题（每小题 15 分，共 6 小题，总计 90 分，结果保留小数点后两位）

- 1、由 10 名学生组成一个随机样本，让他们分别采用 A 和 B 两套试卷进行测试所得成绩如下表：

编号	1	2	3	4	5	6	7	8	9	10
A 卷	78	63	72	89	91	49	68	76	85	55
B 卷	71	44	61	84	74	51	55	60	77	39

要求：

- (1) 分别对两套试卷成绩计算常用汇总统计量，说明两套试卷成绩的整体水平和分散程度？
- (2) 用汇总统计量描述数据的优缺点是什么？如果考试成绩服从正态分布，根据经验法则，1 个标准差相当于告诉我们什么信息？
- (3) 计算 B 卷成绩标准误（差），标准误（差）给我们提供什么的信息？

- 2、尼尔森的调查报告显示，约 86% 的 90 后受访者表示他们每天至少上网一到两次，20% 左右的受访者表示，他们一般全天都会通过个人电脑或者是智能手机和平板电脑保持在线的状态。其中，在对 1250 名中国 90 后受访者调查时，当问及“想象一下，没有手机的生活”时，有 850 名的受访者表示“没法好好生活了”。

要求：

- (1) 请估计中国 90 后中，认为“没有手机没法好好生活了”人群占比的置信水平为 95% 的置信区间；
- (2) 是否可以认为总体比率有 95% 的可能落在该置信区间内？为什么？
- (3) 如果保持相同的置信水平，要使得置信区间缩短为原来的一半，需要抽取多少受访者作为样本？

你认为这样做是否很有必要，为什么？

- 3、有一项针对上海市大学生的情感现状调查，所得有效样本有 1105 名学生，其中：学生干部占 350 人，非学生干部占 755 人。在回答“成功感”题项时（最高为 4 分，最低为 1 分），学生干部所得均分 2.97 分，标准差 0.75，非学生干部均分 2.78 分，标准差 0.64 分。现要求检验学生干部是否比非学生干部更具有成功感。

- (1) 检验的零假设和备择假设分别是什么？
- (2) 如果显著性水平设为 0.05，是否拒绝零假设。你在做出该项判断时可能会犯何种类型的错误？

(3) 假设检验中的 P 值是什么？是零假设错误的概率吗？在该项检验中 P 值会小于 0.05 吗？

4、有一项有关大学生收看和认知“诗词大会”电视节目的调查，其中某大学回收有效问卷 446 份，分别按性别、学历和专业对是否看过“诗词大会”得到频率汇总表如下，该表还分别给出了 chi-square 统计量及相应 p-value。

对于观众与非观众按性别、学历和专业的分类统计表

	性别		学历		专业	
	男生	女生	本科生	研究生	文科生	非文科生
没看过	41.7%	23.3%	33.0%	29.9%	30.1%	37.2%
看过	58.3%	76.7%	67.0%	70.1%	69.9%	62.8%
	chi-square	p-value	chi-square	p-value	chi-square	p-value
	18.183	0.000	0.334	0.563	2.295	0.130

要求：

- (1) 该问题的自变量和因变量分别是什么？
- (2) 根据表中数据求各 Cram \acute{e} 's V 相关系数。
- (3) 性别、学历和专业与是否收看节目之间相关程度如何？

5、某咨询公司最近进行了一项关于居民价值取向与心理状况的调查。根据 20 个国家相关数据，以“对成功和赚钱有很大压力”作为因变量，“根据自己拥有的东西衡量个人成功”作为自变量，估计线性回归方程，所得结果如下：

	df	SS	MS	F	Significance F
回归分析	1	1393.997	1393.997	13.88634	0.001545
残差	18	1806.953	100.3863		
总计	19	3200.95			

	Coefficients	标准误差	t Stat	P-value
Intercept	27.92183	5.08922	5.486465	3.28E-05
根据拥有的物质 衡量成功 (%)	0.525561	0.141036	3.726437	0.001545

要求：

- (1) 根据表中数据写出回归方程；
- (2) 回归系数 0.525561 表明了什么？
- (3) 两个变量之间的相关关系如何？回归方程有多大的解释效应？
- (4) 为提高方程的解释效应，你认为还可以添加哪些变量来解释各国居民压力的差异？

6、某公司三种产品的有关资料如下表

产品名称	个体产量指数	基期产值 (万元)	报告期产值 (万元)
甲	1.25	100	120
乙	1.10	100	115
丙	1.50	60	85

要求：(1) 三种产品产量平均增长了多少，产量增长对产值有什么影响？

- (2) 三种产品价格平均变动了多少，价格变动对产量有什么影响？