

2021年10月高等教育自学考试全国统一考试

计算机统计分析方法

(课程代码 08229)

注意事项:

1. 本试卷分为两部分, 第一部分为选择题, 第二部分为非选择题。
2. 应考者必须按试题顺序在答题卡(纸)指定位置上作答, 答在试卷上无效。
3. 涂写部分、画图部分必须使用2B铅笔, 书写部分必须使用黑色字迹签字笔。

第一部分 选择题

一、单项选择题: 本大题共15小题, 每小题1分, 共15分。在每小题列出的备选项中只有一项是最符合题目要求的, 请将其选出。

1. 下列选项中不是SPSS的内部文件的是
A. 数据文件(*.sav) B. 输出文件(*.spv)
C. 语法文件(*.sps) D. 视频文件(*.avi)
2. 相关系数的取值范围是
A. $[0, +\infty)$ B. $(-1, 1)$
C. $[-1, 1]$ D. $[0, 1]$
3. SPSS默认的数值型变量的对齐方式是
A. 右对齐 B. 中间对齐
C. 左对齐 D. 两端对齐
4. 聚类分析中不属于个体与小类、小类与小类间“亲疏程度”的度量方法的是
A. 组间平均连锁距离(Between-groups linkage)
B. 欧氏距离(Euclidean distance)
C. 最近邻居距离(Nearest Neighbor)
D. 重心距离(Centroid clustering)
5. 以下统计分析方法中要求变量之间必须具有较高相关性的是
A. 线性回归分析 B. 偏相关分析
C. 聚类分析 D. 因子分析

6. 相关分析是一种
A. 以定量分析为前提的定性分析 B. 定量分析
C. 以定性分析为前提的定量分析 D. 定性分析
7. 在SPSS的单因素方差分析时, Post Hoc键的作用是
A. 弹出均值分布图 B. 弹出检验结果
C. 弹出显著性 D. 弹出各组均值的两两比较窗口
8. 方差分析用来检验不同数组的
A. 数据差异 B. 标准差差异
C. 均值差异 D. 误差差异
9. SPSS中的Measuer栏目的作用是
A. 选择显示宽度 B. 缺失选择
C. 选择测度级别 D. 变量选择
10. 通过在被关注特性与潜在影响因素之间建立模型来研究其相互之间因果关系的统计技术称为
A. 实验设计 B. 假设检验
C. 测量分析 D. 回归分析
11. One-Sample T Test 用于
A. 检验某个变量的总体方差和指定值之间是否存在显著差异
B. 检验某个变量的总体均值和指定值之间是否存在显著差异
C. 检验两个变量的总体均值是否存在显著差异
D. 检验两个变量的总体方差是否存在显著差异
12. 改变直方图形状可用的方法是
A. 精确制图 B. 改变组距
C. 数据变换 D. 组限变换
13. 下列测定, 既可进行加减运算, 也可进行乘除运算的是
A. 定名测定 B. 序列测定
C. 定距测定 D. 定比测定
14. 在线性回归模型中, 随机误差被假定服从
A. 正态分布 B. 二项分布
C. 指数分布 D. t分布
15. 二项分布的数学期望为
A. $n(1-n)p$ B. $np(1-p)$
C. np D. $n(1-p)$

二、多项选择题：本大题共 5 小题，每小题 2 分，共 10 分。在每小题列出的备选项中至少有两项是符合题目要求的，请将其选出，错选、多选或少选均无分。

16. SPSS 是多窗口软件，其界面提供的窗口包括
- A. 数据编辑窗口
 - B. 输出窗口
 - C. 语法窗口
 - D. 脚本窗口
 - E. 输入窗口
17. 多因素方差分析中下列选项属于非饱和模型的有
- A. $SST=SSA+SSB+SSE$
 - B. $SST=SSA+SSB+SSAB+SSE$
 - C. $SST=SSA+SSB+SSC+SSAB+SSAC+SSBC+SSE$
 - D. $SST=SSA+SSB+SSC+SSABC+SSE$
 - E. $SST=SSA+SSB+SSE+SSBC$
18. SPSS 的对数线性模型检验方式包括
- A. 拟合模型检验
 - B. 交互效应检验
 - C. 主效应检验
 - D. 拟合优度检验
 - E. 单项效应检验
19. 利用 SPSS 进行数据分析的基本步骤有
- A. 建立 SPSS 数据文件
 - B. 建立数据图形
 - C. 数据的加工处理
 - D. 统计分析
 - E. 解释分析结果
20. SPSS 数据预处理的基本功能包括
- A. 数据的排序
 - B. 变量计算
 - C. 数据选取
 - D. 分类汇总
 - E. 数据分组

三、判断题：本大题共 10 小题，每小题 1 分，共 10 分。

判断下列各题正误，正确的在答题卡相应位置涂“A”，错误的涂“B”。

21. 启动 SPSS 时，数据编辑器、语法编辑器、输出查看器、脚本编辑器都不可以同时打开多个。
22. 参数的点估计量所选统计量的样本量越大，估计值离真实值的差异就越小。
23. 对于分类变量，原始类别频数和构成百分比是最常用的描述指标。
24. 定类尺度代表了事物的类别和属性，不仅能测定类别差，也能比较各类之间的大小。
25. 频率过程基本上按照数据的集中趋势、离散趋势、百分位数和分布指标四大模块将各描述指标进行归类。
26. 常见的连续型分布中最重要和常用的分布是均匀分布。
27. 正态分布曲线是一条对称曲线，而该曲线的高矮形状则与标准差有关。
28. 对于分布不对称的数据，均值比中位数更适合描述数据的集中趋势。
29. SPSS 的标识重复个案功能可通过简单的菜单操作，仅能发现个别变量值的重复。
30. P-P 图和 Q-Q 图最常应用于判断变量是否服从正态分布。

第二部分 非选择题

四、填空题：本大题共 10 小空，每小空 1 分，共 10 分。

31. SPSS 数据文件管理功能中的“转换”主要是实现变量级别的数据管理；而“数据”主要是实现_____的管理。
32. 多选题问题的分析常用两种方法，即多选题二分法和多选题_____。
33. 数据报表中变量的测量尺度主要分为两大类即分类变量和_____。
34. SPSS 中频率功能基本上是按照数据的集中趋势、离散趋势、百分位数和分布指标四大块将各_____进行了归类。
35. 单因素方差分析中，除了基本的分析外，一般还应做方差齐性检验和_____检验。
36. 常见的刻化离散程度的描述统计量是样本标准差、样本方差和全距；刻化分布形态的描述统计量是_____。
37. 相关分析中，相关系数 $|r|$ 越接近于 1，说明相关性越好；相关系数 $|r|$ 越接近于 0，说明_____。
38. 多配对样本的非参数检验方法包括：Friedman 检验、Cochran Q 检验和_____。
39. 设总体 X 的方差为 1，从总体中随机取容量为 100 的样本，得样本均值 $\bar{x}=5$ ，则总体均值的置信水平为 99% 的置信区间_____。($Z_{0.005}=2.58$)
40. 卡方检验中的 X^2 值表示观察值与理论值之间的_____。

五、名词解释题：本大题共 3 小题，每小题 3 分，共 9 分。

41. 协方差分析
42. 变量赋值
43. 单一样本 T 检验

六、简答题：本大题共 4 小题，每小题 6 分，共 24 分。

44. 单样本非参数检验的推断方法主要有哪几种？卡方检验的基本理论依据是什么？
45. 简述条形图、饼图、直方图的含义及适用的情况。
46. 简述对应分析与因子分析的相同和不同处。
47. 在 SPSS 中可以使用哪些方法输入数据？

七、分析、计算题：本大题共 2 小题，每小题 11 分，共 22 分。

48. 调查不同职业的男性和女性的收入情况，分析作为财务管理、计算机程序员、药剂师等不同职业的男女之间月薪是否有明显差异，为此，从每个职业中分别选取 5 名男子和 5 名女子组成样本，样本中每个人的月薪（单位：元）数据如下表所示。用双因素方差分析法分析职业、性别及其交互作用对月薪的影响。

不同职业的男性和女性的收入表：（部分数据略）

序号	月薪/元	职业	性别	序号	月薪/元	职业	性别
1	3488	财务管理	男	8	3536	程序员	女
2	3436	财务管理	男	9	3060	程序员	女
3	4112	财务管理	男	10	2740	程序员	女
4	4468	财务管理	男	11	2800	程序员	女
5	4076	财务管理	男	12	2684	程序员	女
6	2076	财务管理	女	13	4420	药剂师	男
7	2808	财务管理	女	14	4576	药剂师	男

分析结果如下表所示

表 1

Levene's Test of Equality of Error Variances^a

Dependent Variable: 月薪

F	df1	df2	Sig.
.383	5	24	.856

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+职业+性别+职业*性别

表 2

Tests of Between-Subjects Effects

Dependent Variable: 月薪

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	9822080.000 ^a	5	1964416.000	11.761	.000
Intercept	350892000	1	350892000.0	2100.714	.000
职业	4424960.000	2	2212480.000	13.246	.000
性别	3550080.000	1	3550080.000	21.254	.000
职业*性别	1847040.000	2	923520.000	5.529	.011
Error	4008832.000	24	167034.667		
Total	364722912	30			
Corrected Total	13830912.0	29			

a. R Squared = .710 (Adjusted R Squared = .650)

请对此结果进行分析和解释。（注：要求对每个表都要分别分析。）

49. 从某一行业中随机抽取 5 家企业，所得产品产量与生产费用的数据如下：

产品产量（台） x_i	40	50	50	70	80
生产费用（万元） y_i	130	140	145	150	156

要求：①利用最小二乘法求出估计的回归方程；②计算判定系数 R^2 。

附： $\sum_{i=1}^5 (x_i - \bar{x})^2 = 1080$ $\sum_{i=1}^5 (y_i - \bar{y})^2 = 392.8$ $\bar{x} = 58$ $\bar{y} = 144.2$

$\sum_{i=1}^5 x_i^2 = 17900$ $\sum_{i=1}^5 y_i^2 = 104361$ $\sum_{i=1}^5 x_i y_i = 42430$