

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. SPSS 数据文件的扩展名是
A. .sav
B. .exe
C. .dat
D. .frm
2. 快速查找异常值、极端值可以利用
A. 快速定位
B. 冻结行列
C. 排序功能
D. 数据导入
3. 逻辑运算中, 优先级最高的是
A. AND
B. OR
C. NOT
D. NE
4. 用来显示、管理 SPSS 统计分析结果、报表及图形的是
A. 数据窗口
B. 输出窗口
C. 语法窗口
D. 脚本窗口
5. 工资、年龄、成绩等变量一般定义成哪种数据类型
A. 数值型
B. 字符型
C. 数据型
D. 日期型
6. 表示数据分布形态是尖峰分布的峰度系数值是
A. -1.47
B. -1/2
C. 0
D. 3.14

7. 频数分析中通常用于描述分类变量的单变量图是
A. 直方图
B. 饼图
C. 箱图
D. 散点图
8. SPSS 中进行参数检验应选择的主窗口菜单是
A. 视图
B. 编辑
C. 文件
D. 分析
9. 为研究某种减肥茶减肥效果是否显著, 可以采用的分析方法是
A. 单样本 t 检验
B. 两独立样本 t 检验
C. 两配对样本 t 检验
D. 方差分析
10. 两个独立样本的非参数检验方法中, 应用最广的方法是
A. Mann-Whitney U 检验
B. Wilcoxon 符号秩检验
C. Kruskal-Wallis H 检验
D. Friedman 检验
11. 某公司生产的一批 10000 件产品质量的众数为 498g, 表示
A. 10000 件产品的平均质量为 498g
B. 10000 件产品中, 质量为 498g 的件数最多
C. 10000 件产品中质量最大的为 498g
D. 10000 件产品中有 5000 件的质量小于等于 498g
12. 表示 x 和 y 之间是线性回归模型的是
A. $E(Y_i) = \beta_0 + \beta_1 X_i$
B. $Y_i = \beta_0 + \beta_1 X_i + \mu$
C. $Y_i = \beta_0 + \beta_1 X_i^2$
D. $Y_i = \beta_0 + \beta_1$
13. 解释变量不断剔除出回归方程的方法是
A. 进入法
B. 向前法
C. 向后法
D. 逐步回归法
14. SPSS 中无效的变量名是
A. cd_1
B. abc1
C. 王 1
D. ab_
15. 现有一批数据: 2.3, 1.2, 3.8, 6, 9, 则 6 的秩是
A. 2
B. 3
C. 4
D. 6

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 中变量的基本类型包括
A. 数值型
B. 货币型
C. 逻辑型
D. 字符型
E. 日期型

17. SPSS 的函数包括

- A. 算术函数
- B. 统计函数
- C. 分布函数
- D. 逻辑函数
- E. 字符串函数

18. 可以通过频率过程获得的值有

- A. 分位数
- B. 均值
- C. 标准差
- D. 方差
- E. 偏度

19. 可以表示多对变量间统计关系的有

- A. 简单散点图
- B. 重叠散点图
- C. 矩阵散点图
- D. 三维散点图
- E. 分组散点图

20. 以下属于均数间的多重比较法的有

- A. LSD 法
- B. Sidak 法
- C. Bonferroni 法
- D. Scheffe 法
- E. Dunnett 法

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. 在多重二分法中如果有多个变量，其变量值标签的定义可以有多重设置。
- 22. 在同一算术表达式中的常量及变量，数据类型不一致时也可以计算。
- 23. 当样本量 n 足够大时，其抽样均值都会近似服从正态分布。
- 24. SPSS 中绘制散点图应选择数据主窗口菜单。
- 25. 假设检验推断过程依据的原理是小概率事件。
- 26. SPSS 中得到的直方图上面可以附加正态分布曲线，便于与正态分布做比较。
- 27. 非参数检验和总体分布无关。
- 28. 卡方检验可以用来考察某无序分类变量各水平在两组或多组间的分布是否一致。
- 29. 当数据为有序变量时，通常考察其直线相关的问题。
- 30. 在合并文件中自动生成的来源变量中，0 表示来自当前文件，1 表示来自第二份数据文件。

第二部分 非选择题

四、填空题：本大题共 10 小题，每小题 1 分，共 10 分。

- 31. 在统计软件中用于多选题的常见方法有：_____和多重分类法。
- 32. 函数中如果有多个参数，各参数之间需要用_____隔开。
- 33. 多变量排序时需要使用“_____”对话框来进行操作。
- 34. 参数点估计可以采用矩法和_____法。
- 35. 序列 00110111000100100010 总的游程数=_____。

36. 针对连续变量的统计推断方法中，最常用的是_____和方差分析。

37. 纵向合并数据文件是依据两份数据文件的_____进行数据对接。

38. 从红、橙、黄、绿、蓝、黑、紫中选出三种你最喜爱的颜色，如果采用多重二分法设置变量，应该设置的变量数目是_____个。

39. 变量名最长不能超过_____个字符。

40. 某次调查发放了 900 份问卷，回收得到的有效问卷有 800 份，则 SPSS 所建立的相关数据文件中的行数为_____行。

五、名词解释题：本大题共 3 小题，每小题 3 分，共 9 分。

- 41. 数据字典
- 42. 正态分布
- 43. 曲线相关

六、简答题：本大题共 4 小题，每小题 6 分，共 24 分。

- 44. 请简述对数据文件进行分类汇总的步骤。
- 45. 一组数据的分布特征可以从哪几个方面进行测度？
- 46. 请简述方差分析的应用条件。
- 47. 相关系数的特点是什么？

七、分析、计算题：本大题共 2 小题，每小题 11 分，共 22 分。

48. 在一项关于城市租房的研究中，调查的情况资料如下表所示：

小区的租房情况资料

小区编号	月平均家庭收入/元	离市中心的距离/km	有三居室的住宅月租金/元
1	3400	7.5	1500
2	3700	6.3	1600
3	3900	5.0	1800
4	4200	3.6	2200
5	4100	4.5	2000
6	4500	4.8	2050

分析结果如下表所示：

Correlations

		月平均收入	离市中心的距离	三居室
月平均收入	Pearson Correlation	1	-.642*	.911**
	Sig. (2-tailed)		.033	.000
	N	11	11	11
离市中心的距离	Pearson Correlation	-.642*	1	-.851**
	Sig. (2-tailed)	.033		.001
	N	11	11	11
三居室	Pearson Correlation	.911**	-.851**	1
	Sig. (2-tailed)	.000	.001	
	N	11	11	11

*. Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

试对月平均家庭收入、离市中心的距离和有三居室的住宅月租金三个变量进行相关分析。

49. 一种新型减肥食品正在作投入市场前的检验。一个随机样本有 8 人组成，在他们服用这种减肥食品前记录下每个人的体重。然后让他们连续服用这种食品一个疗程，再次记录他们的体重，试验结果如下表所示：

受试者	服用减肥食品前的体重/kg	服用减肥食品后的体重/kg
1号	75	71.5
2号	96	92.5
3号	90	85
4号	100	96
5号	78	73
6号	83	79
7号	67	66
8号	71	71

分析结果如下表所示：

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair1 减肥前-减肥后	3.25000	1.81265	.64087	1.73458	4.76542	5.071	7	.001

请对此结果进行分析和解释。