**浙江工商大学研究生**

**2020/2021学年第一 学期考试试卷**

**课程名称**  应用多元统计分析 **考试方式 上机操作 完成时限：120分钟**

**学号 姓名 得分**

**根据给定的数据，上机操作进行分析。**

**数据：** SAS数据集innovation（或见Excel表：innovation）

反映我国大陆30个地区（由于西藏部分数据缺失，暂不考虑）创新发展能力的12项指标，请运用因子分析方法、聚类分析、典型相关分析等方法对各地区的创新能力进行分析和评价。要求以表格的形式列出相关结果，加以文字说明和分析。

1.运用合适的旋转方法进行因子分析，保留4个因子，根据旋转后的因子载荷矩阵解释各公共因子的含义；

2.根据前4个因子得分，计算综合因子得分F，并根据综合因子得分F降序排序，输出结果。

3.根据前4个因子得分，进行聚类分析。要求运用Ward法进行聚类，画出聚类谱系图。分类后，对每一类进行命名和简要分析。

4.以反映创新投入的变量X3、X4、X5、X6作为第一组变量，反映创新产出的变量X9、X10、X11、X12为第二组变量，进行典型相关分析。请问保留几对典型相关系数比较合适，并说明理由。

6.请解释典型变量的含义，以及典型相关系数的含义。

7.请进行冗余分析。

**表1地区创新发展能力评价指标体系**

|  |  |  |  |
| --- | --- | --- | --- |
| **一级指标** | **二级指标** | **单位** | **符号** |
| 创新环境 | 人均地区生产总值 | 元 | X1 |
| 本科以上学历人数占就业人数的比重 | % | X2 |
| 创新投入 | R&D经费支出 | 亿元 | X3 |
| R&D人员全时当量 | 亿元 | X4 |
| R&D经费支出占GDP比重 | % | X5 |
| 每万人研究与开发人员数 | （人年/万人） | X6 |
| 创新效率 | 每千R&D人员发明专利申请数 | 件/千人 | X7 |
| 每千R&D人员有效发明专利数 | 件/千人 | X8 |
| 创新成果 | 每万人发明专利申请数 | 件/万人员 | X9 |
| 每万人有效发明专利数 | 件/万人员 | X10 |
| 创新效益 | 新产品销售收入占产品销售收入的比重 | % | X11 |
| 单位能源创造的GDP | 万元/吨标准煤 | X12 |