《工程应用统计学》

实验报告

**实验学期： 2024~2025春季**

**实验班级： 工业（1）班**

**学生姓名： 余泓为**

**学生学号： 3123000627**

**广东工业大学 机电工程学院**

**二０二四年四月印刷**

广东工业大学实验报告

评分标准与得分：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **撰写规范性**  **(20分)** | **实验过程描述**  **(30分)** | **结果与分析**  **(35分)** | **实验讨论**  **(15分)** | **成绩**  **评定** | |
| **实验一：基于Excel的数据描述统计分析**  第\_9周星期 二 第 5 节 | |  |  |  |  |  | |
| **实验二：基于SPSS的数据推断统计分析**  第10周星期 二 第 5 节 | |  |  |  |  |  | |
|  | **平均成绩：** | | | | | |  |
|  | **教师签名：** | | | | | |  |

实验一 基于Excel的数据描述统计分析

1. **实验目的**
   1. 掌握统计学有关描述统计的基本概念与原理
   2. 学会使用Excel进行数据整理和图示
2. **简述常用的概率抽样方法，分析其中哪些适用于本实验的学生成绩抽样？**

常用的概率抽样方法：

* 1. 简单随机抽样
  2. 系统抽样
  3. 分层抽样
  4. 整群抽样
  5. 多阶段抽样

本次适用：

1. 系统抽样
2. 简单随机抽样
3. **简述分类数据和数值数据的整理与图示方法**
   1. 利用数据透视表生成频数分布表
   2. 利用条形图、帕累托图、饼图等展示频数分析
4. **借助Excel进行简单随机非重复抽样，获得一个容量为30的学生成绩样本，描述具体实现步骤并列出样本数据。**（注意，要求每位同学单独抽样，理论上所得样本数据不太可能完全相同！另外，若抽样数据中没有包含全部5种成绩等级，则可以直接按实际样本数据分析，也可以重新抽样直到包含5种成绩等级。由于Excel在每次修改和保存时会重新生成随机数，导致样本数据发生变化，因此建议获得样本数据后及时拷贝数值保存到Excel其它位置，然后再进行分析计算）
   1. 在数据开头插入两列，分别命名为“随机数”（A列）与“排名”（B列）；
   2. 在A2中输入“=RAND()”，B2中输入“=RANK(A2,A:A)”并将两公式拖至91行
   3. 在I2中输入“=VLOOKUP(ROW()-1,$B:$E,2,0)”并将公式拖至31行

数据表格如下：

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 性别 | 男 | 男 | 男 | 女 | 男 | 男 | 男 | 女 | 男 | 男 |
| FZ成绩 | 66 | 90 | 98 | 83 | 75 | 87 | 77 | 84 | 76 | 89 |
| 成绩等级 | 及格 | 优秀 | 优秀 | 良好 | 中等 | 良好 | 中等 | 良好 | 中等 | 良好 |
| 性别 | 男 | 男 | 男 | 男 | 男 | 男 | 男 | 男 | 男 | 男 |
| FZ成绩 | 82 | 85 | 67 | 94 | 80 | 76 | 81 | 76 | 79 | 82 |
| 成绩等级 | 良好 | 良好 | 及格 | 优秀 | 良好 | 中等 | 良好 | 中等 | 中等 | 良好 |
| 性别 | 男 | 男 | 男 | 女 | 女 | 男 | 男 | 男 | 男 | 男 |
| FZ成绩 | 70 | 96 | 64 | 73 | 72 | 89 | 95 | 90 | 94 | 46 |
| 成绩等级 | 中等 | 优秀 | 及格 | 中等 | 中等 | 良好 | 优秀 | 优秀 | 优秀 | 不及格 |

1. **样本数据分析**
2. 对样本中的分类数据（性别、成绩等级）进行整理和图示，生成单变量简单频数分布表，以及两变量二维列联表，并分别至少选用一种图示方法展示结果。

1. 对样本中的数值数据（成绩）进行分组，并采用直方图展示结果；
2. 对样本中的数值数据（成绩）进行描述统计，计算相应的描述统计量，并据此分析样本数据中成绩的分布特征。

|  |  |
| --- | --- |
| 描述统计 | |
|  |  |
| 平均 | 79.33333333 |
| 标准误差 | 2.330868978 |
| 中位数 | 81 |
| 众数 | 81 |
| 标准差 | 12.76669518 |
| 方差 | 162.9885057 |
| 峰度 | 3.818417836 |
| 偏度 | -1.613994814 |
| 区域 | 60 |
| 最小值 | 38 |
| 最大值 | 98 |
| 求和 | 2380 |
| 观测数 | 30 |

注：如需原文件，请访问此网站：

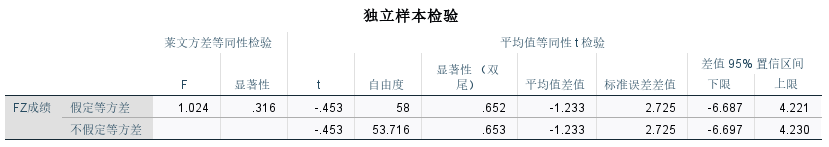
<https://github.com/YuusiHoshizora/YuusiHoshizora.github.io/tree/main/fileStorage>

实验二 基于SPSS的数据推断统计分析

1. **实验目的**
   1. 掌握统计学有关推断统计的基本概念与原理
   2. 学会使用SPSS进行数据统计分析
2. **简述推断统计分析包含哪些内容**
   1. 单个总体的参数估计
   2. 假设检验
   3. 独立性检验
   4. 方差分析
   5. 相关分析
3. **借助SPSS进行简单随机非重复抽样，获得每一届学生成绩样本（样本容量为30），描述具体实现步骤并列出样本数据。**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 序号 | 1-性别 | 1-FZ | 序号 | 2-性别 | 2-FZ | 序号 | 3-性别 | 3-FZ | 3-TJ |
| 1 | 男 | 64 | 3 | 男 | 77 | 1 | 男 | 88 | 67 |
| 4 | 男 | 68 | 7 | 男 | 90 | 2 | 男 | 73 | 60 |
| 5 | 男 | 75 | 11 | 男 | 81 | 3 | 男 | 71 | 84 |
| 9 | 男 | 77 | 12 | 男 | 83 | 8 | 男 | 57 | 64 |
| 10 | 男 | 90 | 13 | 男 | 85 | 9 | 男 | 61 | 87 |
| 11 | 男 | 80 | 15 | 男 | 84 | 16 | 男 | 85 | 65 |
| 12 | 男 | 81 | 27 | 男 | 86 | 18 | 男 | 60 | 60 |
| 14 | 男 | 69 | 28 | 男 | 69 | 19 | 男 | 73 | 71 |
| 18 | 男 | 80 | 33 | 男 | 95 | 20 | 男 | 72 | 92 |
| 20 | 男 | 76 | 36 | 男 | 75 | 22 | 男 | 70 | 75 |
| 21 | 男 | 82 | 41 | 女 | 89 | 23 | 男 | 65 | 60 |
| 23 | 男 | 74 | 42 | 女 | 87 | 27 | 男 | 79 | 86 |
| 24 | 男 | 38 | 48 | 男 | 83 | 29 | 男 | 80 | 77 |
| 27 | 男 | 90 | 50 | 男 | 93 | 37 | 女 | 84 | 82 |
| 29 | 男 | 81 | 55 | 男 | 77 | 38 | 女 | 70 | 84 |
| 31 | 男 | 70 | 56 | 男 | 81 | 39 | 女 | 87 | 80 |
| 32 | 男 | 96 | 60 | 男 | 79 | 43 | 男 | 81 | 92 |
| 33 | 男 | 86 | 62 | 男 | 77 | 50 | 男 | 75 | 76 |
| 34 | 男 | 73 | 63 | 男 | 71 | 51 | 男 | 79 | 83 |
| 41 | 女 | 74 | 65 | 男 | 99 | 63 | 男 | 83 | 96 |
| 43 | 女 | 100 | 66 | 男 | 73 | 65 | 男 | 84 | 89 |
| 50 | 男 | 81 | 69 | 男 | 77 | 66 | 男 | 72 | 66 |
| 60 | 男 | 70 | 71 | 男 | 64 | 68 | 男 | 73 | 80 |
| 65 | 男 | 90 | 75 | 男 | 74 | 71 | 男 | 89 | 70 |
| 66 | 男 | 94 | 78 | 男 | 74 | 73 | 男 | 83 | 79 |
| 70 | 男 | 70 | 81 | 男 | 77 | 74 | 男 | 81 | 78 |
| 72 | 男 | 89 | 83 | 女 | 67 | 76 | 男 | 80 | 74 |
| 79 | 男 | 75 | 84 | 女 | 83 | 80 | 女 | 85 | 90 |
| 84 | 男 | 90 | 85 | 女 | 60 | 84 | 女 | 84 | 94 |

1. **基于样本数据进行以下推断统计分析，简洁描述主要步骤。**
   1. **估计第一届学生的FZ成绩均值的95%置信区间和优良率的95%置信区间。**



第一届学生的FZ成绩均值的95%置信区间：

解得其对应区间为(74.25,82.81)

第一届学生的FZ优良率的95%置信区间：

解得其对应区间为(0.313,0.678)

* 1. **分析第一届与第二届学生的FZ成绩均值有无显著差异？**

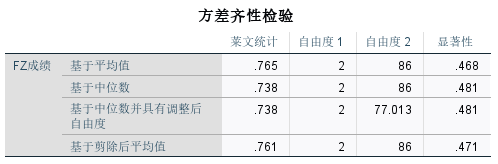
表格

AI 生成的内容可能不正确。

由于P值为0.652，表明两届FZ成绩均值差异并不显著。

* 1. 表格

     AI 生成的内容可能不正确。**分析三届学生的FZ成绩均值有无显著差异？**



表格

AI 生成的内容可能不正确。

方差齐性结果表明：莱文统计量为0.765，P值为0.468，因此不能拒绝原假设

ANOVA结果表明：F统计量为0.223，P值为0.801，因此不能拒绝原假设

LSD多重比较结果表明：两两对比结果的P值均大于0.05，从而表明三届学生的FZ成绩均值无显著差异

* 1. **图表, 散点图

     AI 生成的内容可能不正确。表格

     AI 生成的内容可能不正确。分析第三届学生的FZ成绩与TJ成绩是否相关？若相关，两者的相关程度如何？**

**两门课程之间的皮尔逊相关系数为0.337，表明两者之间存在一定的正线性相关关系，但是相关性不是特别强。**

**显著性（双尾）为0.073，表明检验P值小于0.05，从统计学的角度来说两者之间存在显著的线性相关性。**

注：如需原文件，请访问此网站：

<https://github.com/YuusiHoshizora/YuusiHoshizora.github.io/tree/main/fileStorage>