**实 验 报 告**

**学 院：数学与计算机科学学院**

**课程名称：统计分析**

**专业班级： 智能科学与技术**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 实验项目 | 实验一 | 实验二 | 实验三 | 实验四 | 实验五 | 实验六 | 实验七 | 实验八 | 实验九 | 实验十 | **总评**  **成绩** |
| 评分 |  |  |  |  |  |  |  |  |  |  |  |

**姓 名：刘糠杰**

**学 号：202305130222**

**学生实验报告**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 实验项目 | python软件并掌握其应用，利用其对数据进行方差分析 | | | |
| ▇必修 □选修 | | □演示性实验▇验证性实验 □操作性实验 ▇综合性实验 | | |
| 实验地点 | 10-315 | | 实验仪器台号 |  |
| 指导教师 | 刘洪久 | | 实验日期及节次 |  |

**一、实验目的及要求：**

1、**目的**

熟悉python软件并掌握其应用，利用其对数据进行方差分析。

2、**内容及要求**

用python软件进行方差分析，具体见第七章第一节例7-1。

（1）绘制单因素方差分析表。

（2）分析是否显著

**二、仪器用具：**

|  |  |  |  |
| --- | --- | --- | --- |
| **仪器名称** | **规格/型号** | **数量** | **备注** |
| 计算机 |  | 1 |  |
| Python 软件 |  | 1 |  |
| Jupyter软件 |  | 1 |  |

**三、实验资料**

《Python统计分析》

**四、实验注意事项**

1．提前复习和实验内容相关的理论。

2．准确输入实验数据。

3．熟悉代码语句，准确输入代码。

**五、实验步骤（代码）**

import numpy as np  
import scipy.stats as stats  
import pandas as pd  
from statsmodels.stats.anova import anova\_lm  
from statsmodels.formula.api import ols  
  
*# 数据：每个小麦品种的产量数据*data = {  
 'A': [5.2, 5.5, 5.6, 5.3, 5.4],  
 'B': [6.1, 6.3, 6.2, 6.5, 6.4],  
 'C': [4.8, 5.0, 4.9, 5.1, 4.7],  
 'D': [5.9, 6.0, 5.8, 5.7, 6.1],  
 'E': [5.3, 5.5, 5.2, 5.4, 5.3]  
}  
  
*# 将数据转换为长格式*df = pd.DataFrame(data)  
df = df.melt(var\_name='品种', value\_name='产量')  
  
*# 打印数据*print("数据框：")  
print(df)  
  
*# 使用ols模型进行单因素方差分析*model = ols('产量 ~ C(品种)', data=df).fit()  
  
*# 进行方差分析*anova\_table = anova\_lm(model)  
  
*# 输出方差分析表*print("\n方差分析表：")  
print(anova\_table)  
  
*# 提取p值*p\_value = anova\_table['PR(>F)'][0]  
  
*# 显示结论*alpha = 0.05 *# 显著性水平*if p\_value < alpha:  
 print(f"\np值 = {p\_value:.4f}，小麦品种对产量的影响是显著的，拒绝原假设。")  
else:  
 print(f"\np值 = {p\_value:.4f}，小麦品种对产量的影响不显著，无法拒绝原假设。")

**数据框：**

**品种 产量**

**0 A 5.2**

**1 A 5.5**

**2 A 5.6**

**3 A 5.3**

**4 A 5.4**

**5 B 6.1**

**6 B 6.3**

**7 B 6.2**

**8 B 6.5**

**9 B 6.4**

**10 C 4.8**

**11 C 5.0**

**12 C 4.9**

**13 C 5.1**

**14 C 4.7**

**15 D 5.9**

**16 D 6.0**

**17 D 5.8**

**18 D 5.7**

**19 D 6.1**

**20 E 5.3**

**21 E 5.5**

**22 E 5.2**

**23 E 5.4**

**24 E 5.3**

**方差分析表：**

**df sum\_sq mean\_sq F PR(>F)**

**C(品种) 4.0 5.8624 1.4656 64.849558 3.632729e-11**

**Residual 20.0 0.4520 0.0226 NaN NaN**

**六、实验结果及分析**