

地理建模实验2 实验报告

42109232 吕文博 地信2101班

2024-05-13

Pearson相关系数计算和散点图绘制

```
df_pearson = readxl::read_xls('../data/exp2/2.xls', sheet = 'Pearson')
summary(df_pearson)
```

```
##      人名              数学              化学
## Length:18          Min.   :50.00      Min.   :60.00
## Class :character    1st Qu.:79.25      1st Qu.:81.00
## Mode  :character    Median :87.00      Median :88.00
##                               Mean   :83.56      Mean   :86.61
##                               3rd Qu.:89.75      3rd Qu.:96.00
##                               Max.   :99.00      Max.   :99.00
```

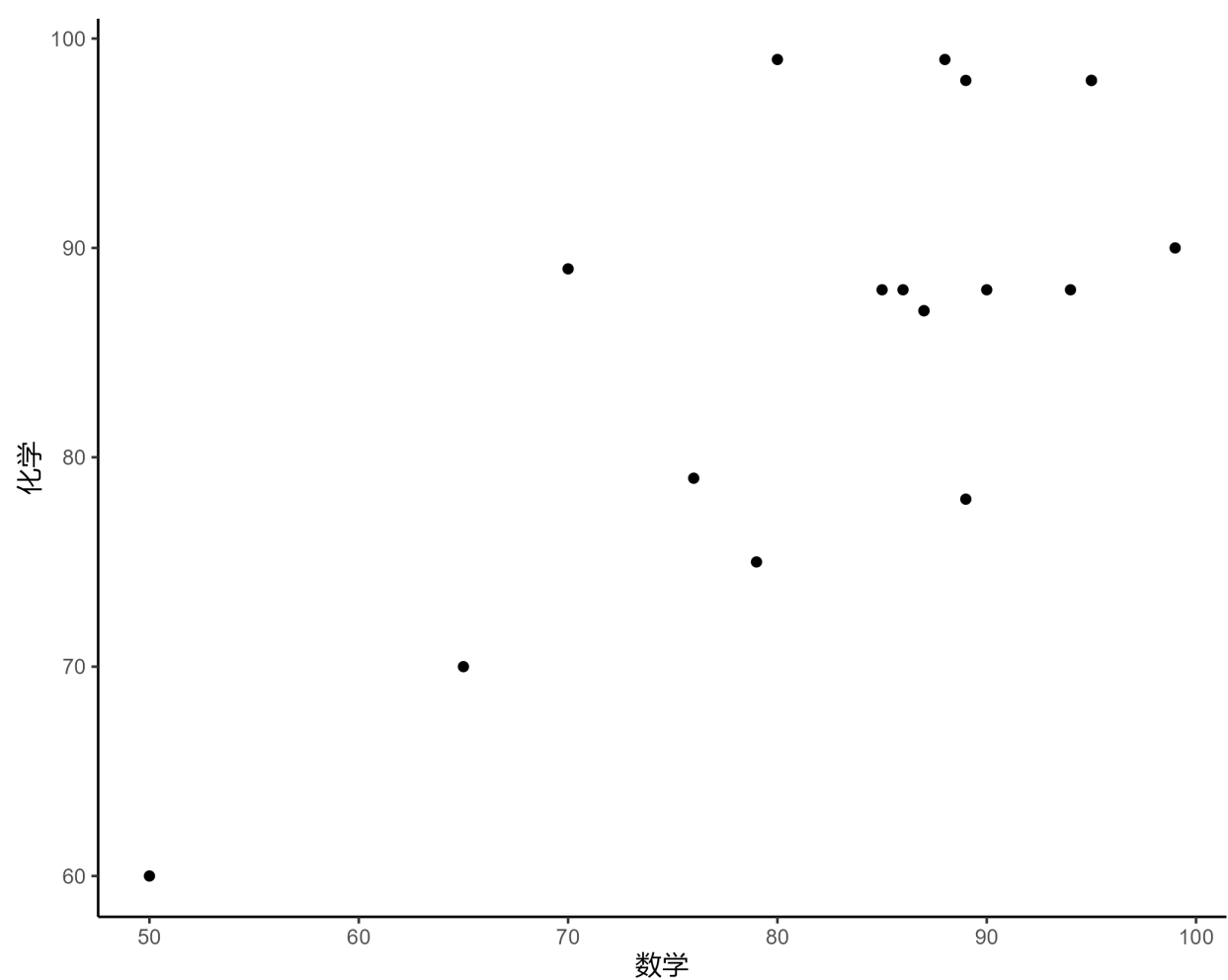
```
cor.test(df_pearson$`数学`,
          df_pearson$`化学`,
          method = "pearson")
```

```
##
## Pearson's product-moment correlation
##
## data:  df_pearson$数学 and df_pearson$化学
## t = 4.4281, df = 16, p-value = 0.0004219
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.4210747 0.8978697
## sample estimates:
##      cor
## 0.7420644
```

数学和化学的pearson相关系数为0.7420644, P值为0.0004219, 两者在0.01显著性水平下相关性显著

```
library(ggplot2)

ggplot() +
  geom_point(data = df_pearson,
            aes(x = 数学,
                y = 化学)) +
  coord_fixed() +
  theme_classic()
```



Spearman等级相关系数、Kendall's 等级相关系数计算

```
df_rank = readxl::read_xls('../data/exp2/2.xls', sheet = 'Spearman')
# 计算Spearman等级相关系数
cor.test(df_rank$`作文1`,
         df_rank$`作文2`,
         method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  df_rank$作文1 and df_rank$作文2
## S = 122.56, p-value = 2.199e-06
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.8735153
```

作文1和作文2的spearman相关系数为0.8735153, P值为2.199e-06, 两者在0.01显著性水平下相关性显著

```
# 计算Kendall's等级相关系数
cor.test(df_rank$`作文1`,
         df_rank$`作文2`,
         method = "kendall")
```

```
##
##  Kendall's rank correlation tau
##
## data:  df_rank$作文1 and df_rank$作文2
## z = 4.2307, p-value = 2.33e-05
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##      tau
## 0.7451175
```

作文1和作文2的kendall相关系数为0.7451175, P值为2.33e-05, 两者在0.01显著性水平下相关性显著

偏相关系数计算

```
df_p = readxl::read_xls('../data/exp2/2.xls', sheet = '偏相关')
# 控制变量为温度
```

```
ppcor::pcor.test(x = df_p$产量,
                 y = df_p$降雨量,
                 z = df_p$温度)
```

```
##      estimate    p.value statistic  n gp Method
## 1 0.7802799 0.01310611  3.300809 10  1 pearson
```

当控制变量为温度时,产量和降雨量的偏相关系数为**0.7802799**,且P值**0.01310611**小于**0.05**,说明产量和降雨量相关性显著.

控制变量为降雨量

```
ppcor::pcor.test(x = df_p$产量,
                 y = df_p$温度,
                 z = df_p$降雨量)
```

```
##      estimate    p.value statistic  n gp Method
## 1 0.8462227 0.00402605  4.201899 10  1 pearson
```

当控制变量为降雨量时,产量和温度的偏相关系数为**0.8462227**,且P值**0.00402605**小于**0.05**,说明产量和温度相关性显著.