



数据结构与画图

作者 李伟

目录

多元统计数据格式

变量分类

常用数字特征

R 语言中的数据结构

读取 EXCEL 文件

常用绘图函数

示例代码与课堂小练习

参考文献

# 数据结构与画图

作者 李伟

太原工业学院 理学系

2018 年 3 月 8 日



# 目录

数据结构与画图

作者 李伟

目录

多元统计数据格式

变量分类

常用数字特征

R 语言中的数据结构

读取 EXCEL 文件

常用绘图函数

示例代码与课堂小练习

参考文献

- 1 多元统计数据格式
- 2 变量分类
- 3 常用数字特征
- 4 R 语言中的数据结构
- 5 读取 EXCEL 文件
- 6 常用绘图函数
- 7 示例代码与课堂小练习
- 8 参考文献



# 多元数据格式

数据结构与画图

作者 李伟

目录

多元统计数据格式

变量分类

常用数字特征

R 语言中的数据结构

读取 EXCEL 文件

常用绘图函数

示例代码与课堂小练习

参考文献

$$\begin{array}{l} \text{sample1} \\ \text{sample2} \\ \vdots \\ \text{sampleN} \end{array} \begin{pmatrix} \text{variable1} & \text{variable2} & \cdots & \text{variableP} \\ x_{11} & x_{12} & \cdots & x_{1P} \\ x_{21} & x_{22} & \cdots & x_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{NP} \end{pmatrix}$$



# 多元数据格式

数据结构与画  
图

作者 李伟

目录

多元统计数据  
格式

变量分类

常用数字特征

R 语言中的数  
据结构

读取 EXCEL  
文件

常用绘图函数

示例代码与课  
堂小练习

参考文献

- 比如我们的变量为 AQI 的六个指标:  
 $CO, SO_2, NO_2, O_3, PM_{2.5}, PM_{10}$ .
- 样本分别为某一时刻太原, 北京, 上海, 深圳, 广州, 南京等地的指标值.



# 变量的分类

数据结构与画  
图

作者 李伟

目录

多元统计数据  
格式

变量分类

常用数字特征

R 语言中的数  
据结构

读取 EXCEL  
文件

常用绘图函数

示例代码与课  
堂小练习

参考文献

- 定性变量, 比如性别, 天气的阴晴, 物种的分类等;
- 定量变量, 比如年龄, 身高, 血压等.
- 我们的数据是没有缺失的.



# 常用数字特征

数据结构与画图

作者 李伟

目录

多元统计数据格式

变量分类

常用数字特征

R 语言中的数据结构

读取 EXCEL 文件

常用绘图函数

示例代码与课堂小练习

参考文献

- 多维随机变量的期望与协方差阵;
- 多维抽样中的样本均值与样本协方差阵.



# 期望

数据结构与画  
图

作者 李伟

目录

多元统计数据  
格式

变量分类

常用数字特征

R 语言中的数  
据结构

读取 EXCEL  
文件

常用绘图函数

示例代码与课  
堂小练习

参考文献

设  $X = (X_1, X_2, \dots, X_p)$ ,  $Y = (Y_1, Y_2, \dots, Y_q)$  是两个随机向量,  $X$  的期望定义如下:

$$EX = \begin{pmatrix} EX_1 \\ EX_2 \\ \vdots \\ EX_p \end{pmatrix}$$



# 协方差阵

数据结构与画  
图

作者 李伟

目录

多元统计数据  
格式

变量分类

常用数字特征

R 语言中的数  
据结构

读取 EXCEL  
文件

常用绘图函数

示例代码与课  
堂小练习

参考文献

$X, Y$  的协方差阵定义如下:

$$\text{cov}(X, Y) = \begin{pmatrix} \text{cov}(X_1, Y_1) & \text{cov}(X_1, Y_2) & \cdots & \text{cov}(X_1, Y_q) \\ \text{cov}(X_2, Y_1) & \text{cov}(X_2, Y_2) & \cdots & \text{cov}(X_2, Y_q) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(X_p, Y_1) & \text{cov}(X_p, Y_2) & \cdots & \text{cov}(X_p, Y_q) \end{pmatrix}$$

即  $\text{cov}(X, Y) = E[(X - EX)(Y - EY)']$





# 协方差阵

数据结构与画图

作者 李伟

目录

多元统计数据格式

变量分类

常用数字特征

R 语言中的数据结构

读取 EXCEL 文件

常用绘图函数

示例代码与课堂小练习

参考文献

- 样本的期望即用均值代替相应的数学期望;
- 样本的协方差阵就是用样本的协方差代替相应的方差.



# 向量

数据结构与画图

作者 李伟

目录

多元统计数据格式

变量分类

常用数字特征

R 语言中的数据结构

读取 EXCEL 文件

常用绘图函数

示例代码与课堂小练习

参考文献

- `x <- c(...)`
- `x <- seq(from = , to = , by = )`
- `x <- rep(x = , times = )`



# 矩阵

数据结构与图

作者 李伟

目录

多元统计数据格式

变量分类

常用数字特征

R 语言中的数据结构

读取 EXCEL 文件

常用绘图函数

示例代码与课堂小练习

参考文献

- `x <- matrix(data = , nrow = , ncol = , byrow = )`



# 数据框

数据结构与画图

作者 李伟

目录

多元统计数据格式

变量分类

常用数字特征

R 语言中的数据结构

读取 EXCEL 文件

常用绘图函数

示例代码与课堂小练习

参考文献

- `x <- data.frame(...)`



# 读取 EXCEL 文件

数据结构与画图

作者 李伟

目录

多元统计数据格式

变量分类

常用数字特征

R 语言中的数据结构

读取 EXCEL 文件

常用绘图函数

示例代码与课堂小练习

参考文献

- 复制 EXCEL 中的相应数据;  
在命令行敲入以下命令:  

```
data1 <-  
read.table(file = "clipboard", header = TRUE)
```
- 推荐使用 `read.xlsx()`



# plot()

数据结构与画  
图

作者 李伟

目录

多元统计数据  
格式

变量分类

常用数字特征

R 语言中的数  
据结构

读取 EXCEL  
文件

常用绘图函数

示例代码与课  
堂小练习

参考文献

常用的参数:

- xlim;ylim
- xlab;ylab
- lty;pch



# 其它作图函数

数据结构与画图

作者 李伟

目录

多元统计数据类型

变量分类

常用数字特征

R 语言中的数据结构

读取 EXCEL 文件

常用绘图函数

示例代码与课堂小练习

参考文献

- hist(): 直方图;
- barplot(): 条状图;
- boxplot(): 箱线图.
- lines()
- points()



# 示例代码

数据结构与画  
图

作者 李伟

目录

多元统计数据  
格式

变量分类

常用数字特征

R 语言中的数  
据结构

读取 EXCEL  
文件

常用绘图函数

示例代码与课  
堂小练习

参考文献

- 详见 chapter2.R





# 基于 table() 与 barplot() 的小练习

数据结构与画图

作者 李伟

目录

多元统计数据格式

变量分类

常用数字特征

R 语言中的数据结构

读取 EXCEL 文件

常用绘图函数

示例代码与课堂小练习

参考文献

- 读数据：读取 chapter2.xlsx 文件 Sheet2;
- 分别给出” 风险”, ” 教育” 两列的列联表, 并且画出相应的条状图.
- 给出” 教育” 与” 风险” 两者之间的列联表, 并且画出相应的条状图。



# 基于抽烟与学习时间长短的分析 (P37)

数据结构与画图

作者 李伟

目录

多元统计数据格式

变量分类

常用数字特征

R 语言中的数据结构

读取 EXCEL 文件

常用绘图函数

示例代码与课堂小练习

参考文献

- 分别画出”抽烟情况”与”学习时间”的列联表
- 画出两者之间的二维列联表
- 利用 `chisq.test()` 检验两者是否存在相关性 (卡方检验)
- 数据来源: chapter2.xlsx 中的 Exercise3



# References

数据结构与画图

作者 李伟

目录

多元统计数据格式

变量分类

常用数字特征

R 语言中的数据结构

读取 EXCEL 文件

常用绘图函数

示例代码与课堂小练习

参考文献



王斌会

多元统计分析及 *R* 语言建模, 第四版.  
暨南大学出版社, 2016.



黄正华

*beamer* 幻灯片模板.  
黄正华



数据结构与画  
图

作者 李伟

目录

多元统计数据  
格式

变量分类

常用数字特征

R 语言中的数  
据结构

读取 EXCEL  
文件

常用绘图函数

示例代码与课  
堂小练习

参考文献

谢谢!

AUTHOR: LI Wei  
EMAIL: [lww1993@163.com](mailto:lww1993@163.com)