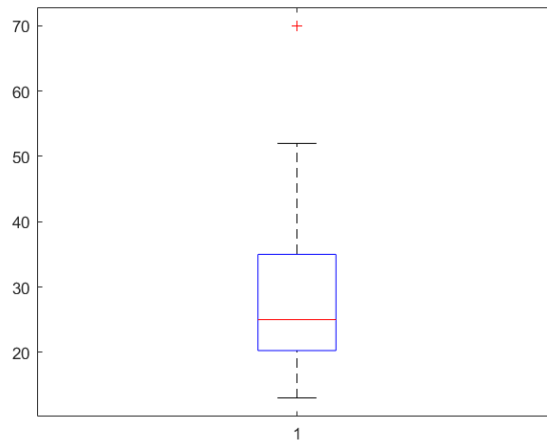


## 第一次作业答案

Answer1:

- (1) 均值: 29.963 中位数: 25
- (2) 众数: 25 和 35
- (3) 41.5
- (4) 第一个四分位: 20 第三个四分位: 35
- (5) 一个分布的五个数汇总由最小值、第一个四分位数、中间值、第三个四分位数和最大值组成。这个数据是: 13, 20, 25, 35, 70



Answer2:

$$E(X) = \frac{(153 + 181 + \dots + 189)}{7} \approx 172.43$$

$$E(Y) = \frac{(44 + 64 + \dots + 84)}{7} \approx 63.86$$

$$\sigma(X) = \frac{\sum_{i=1}^7 (X_i - E(X))^2}{7} \approx 107.67$$

$$\sigma(Y) = \frac{\sum_{i=1}^7 (Y_i - E(Y))^2}{7} \approx 128.98$$

$$\text{cov}(X, Y) = \frac{\sum_{i=1}^7 (X_i - E(X))(Y_i - E(Y))}{7 - 1} \approx 115.07$$

Answer3:

$$(1) D_p = \sqrt{\sum_{i=1}^3 (x_i - y_i)^2} = 2\sqrt{2}$$

$$(2) D_M = \sum_{i=1}^3 |x_i - y_i| = 4$$

$$(3) D_C = \max_{i=1,2,3} (|x_i - y_i|) = 2$$

(4) 由于没有给出数据的协方差，无法计算马氏距离

Answer4:

$$(1) D_R = \frac{p-m}{p}, \text{ 其中 } m \text{ 为对象相同的属性数, } p \text{ 为对象的属性总数}$$

$$(2) D_J = \frac{q}{q+r+s}, \text{ 其中 } q \text{ 是样本 } x \text{ 和 } y \text{ 都取 } 1 \text{ 的属性数; } r \text{ 是样本 } x \text{ 中取 } 1, \text{ 在}$$

样本  $y$  中取 0 的属性数;  $s$  是在样本  $x$  中取 0, 在样本  $y$  中取 1 的属性数

(3) 可用欧式距离、曼哈顿距离、切比雪夫距离、闵可夫斯基距离等距离来描述

Answer5:

定义不同、取值不同、对单位敏感程度不同

## 作业 1.1

### 1. MATLAB 代码:

```
clc
clear all

% 给定数据
data = [13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70];

% 计算均值和中位数
mean_value = mean(data);
median_value = median(data);

% 计算众数
[counts, values] = hist(data, unique(data));
mode_value = values(counts == max(counts));

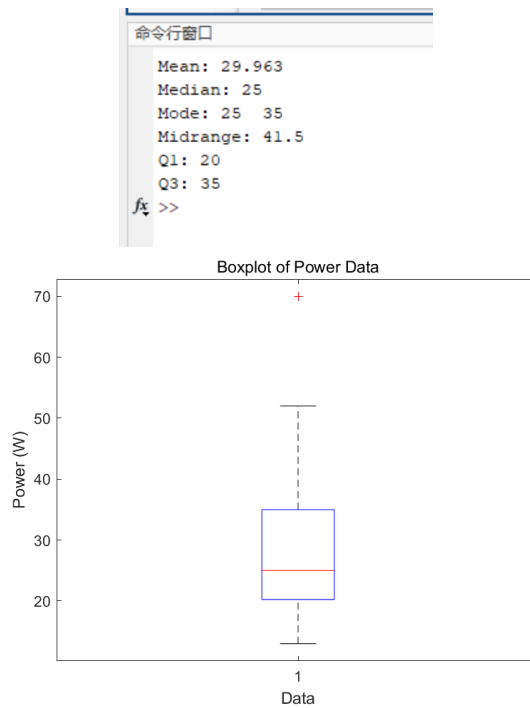
% 计算中列数
midrange_value = (max(data) + min(data)) / 2;

% 计算四分位数
sorted_data = sort(data);
n = length(sorted_data);
q1 = sorted_data(ceil(0.25 * n));
q3 = sorted_data(ceil(0.75 * n));

% 绘制箱线图
figure;
boxplot(data);
title('Boxplot of Power Data');
xlabel('Data');
ylabel('Power (W)');

% 显示计算结果
disp(['Mean: ', num2str(mean_value)]);
disp(['Median: ', num2str(median_value)]);
disp(['Mode: ', num2str(mode_value)]);
disp(['Midrange: ', num2str(midrange_value)]);
disp(['Q1: ', num2str(q1)]);
disp(['Q3: ', num2str(q3)]);
```

### 2. 运行结果:



## 作业 1.2

### 1. MATLAB 代码:

```
clc;
```

```
clear all;
```

```
% 给定样本数据
```

```
X = [153, 181, 170, 172, 174, 168, 189];
```

```
Y = [44, 64, 70, 57, 61, 67, 84];
```

```
% 计算期望
```

```
mean_X = mean(X);
```

```
mean_Y = mean(Y);
```

```
% 计算方差
```

```
var_X = sum((X - mean_X).^2) / length(X);
```

```
var_Y = sum((Y - mean_Y).^2) / length(Y);
```

```
% 计算协方差
```

```
covariance = sum((X - mean_X) .* (Y - mean_Y)) / (length(X) - 1);
```

```
% 显示计算结果
```

```
disp(['期望 E(X): ', num2str(mean_X)]);
```

```
disp(['期望 E(Y): ', num2str(mean_Y)]);
```

### 2. 运行结果:

```
命令窗口
期望 E(X): 172.4286
期望 E(Y): 63.8571
方差 Var(X): 107.6735
方差 Var(Y): 128.9796
协方差 Cov(X, Y): 115.0714
fx >>
```

## 作业 1.3

### 1.MATLAB 代码:

```
clc
clear all;

X = [1, 7, 7];
Y = [3, 9, 7];

% 欧氏距离
euclidean_dist = norm(X - Y);

% 曼哈顿距离
manhattan_dist = sum(abs(X - Y));

% 切比雪夫距离
chebyshev_dist = max(abs(X - Y));

% 显示计算结果
disp(['欧氏距离: ', num2str(euclidean_dist)]);
disp(['曼哈顿距离: ', num2str(manhattan_dist)]);
disp(['切比雪夫距离: ', num2str(chebyshev_dist)]);

try
    % 将 X 和 Y 组成一个集合
    data = [X; Y];

    % 计算协方差矩阵
    cov_matrix = cov(data);

    % 检查协方差矩阵是否对称正定
    [~, p] = cholcov(cov_matrix);
    if p == 0
        % 计算马氏距离
        mahalanobis_dist = mahal(X, Y, cov_matrix);
        disp(['马氏距离: ', num2str(mahalanobis_dist)]);
    else
        disp('无法计算马氏距离，协方差矩阵不满足要求');
```

```
end  
catch ME  
    disp('无法计算马氏距离');  
end
```

2.运行结果:

