回归分析（非线性回归）

多项式回归，分类型变量回归，分类型和数值型混合回归

虽然是非线性回归，但X Y是线性关系

- 简单线性回归：

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

- 多项式回归：

$$y_i = \beta_0 + \beta_1 {x_i}^1 + \beta_2 {x_i}^2 + \cdots + \varepsilon_i$$

- 非线性关系：

$$y_i = \beta_0 + \beta_1 x_1 + {\beta_1}^2 x_2 + \varepsilon_i$$

虽然是非线性回归，但X Y是线性关系

- 简单线性回归：

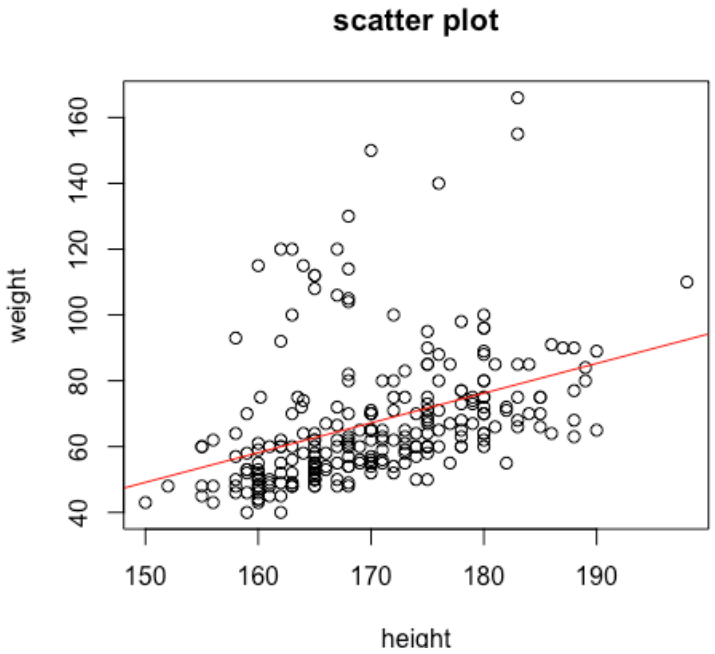
$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

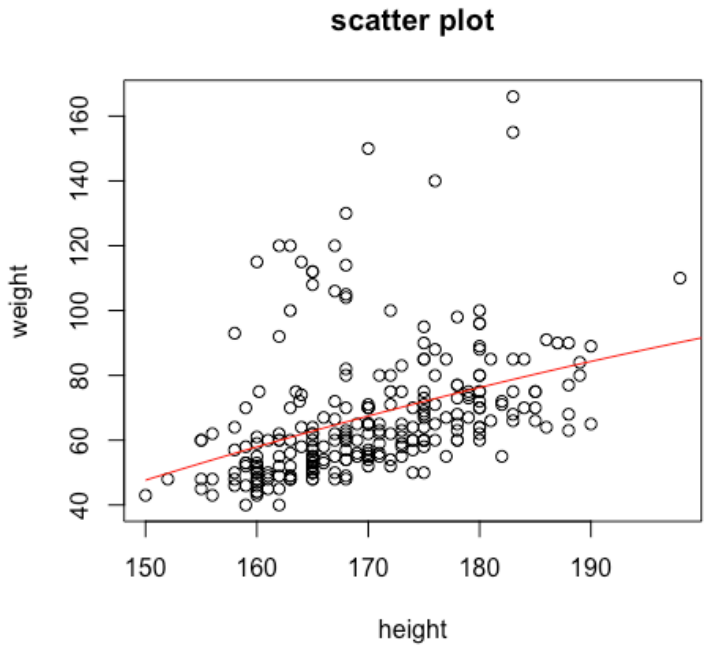
- 二次多项式回归：

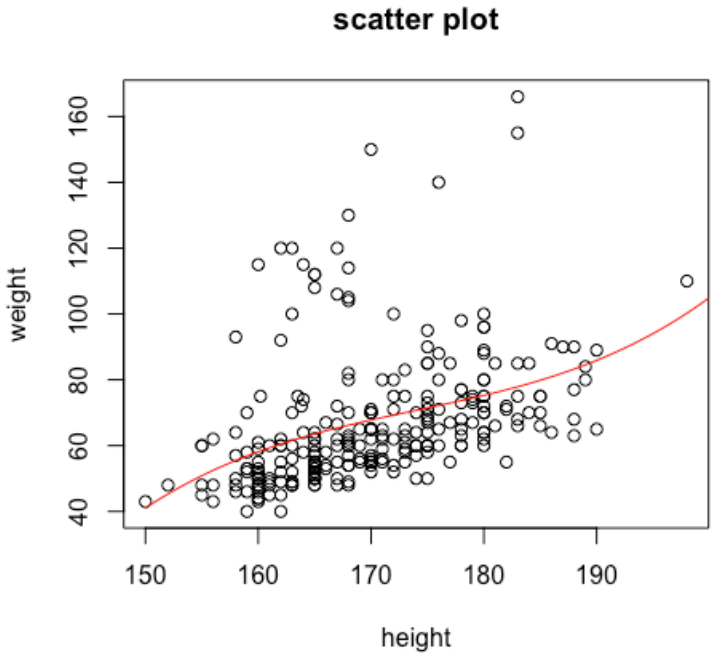
$$y_i = \beta_0 + \beta_1 {x_i}^1 + \beta_2 {x_i}^2 + \varepsilon_i$$

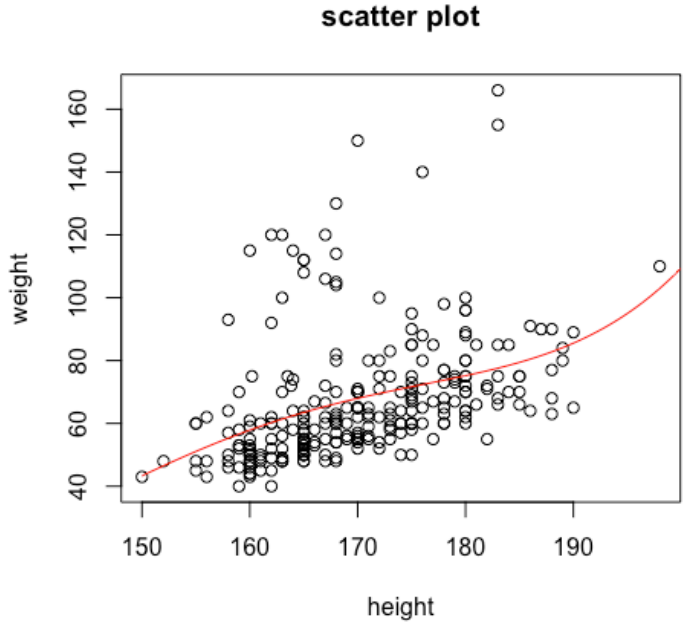
- 三次多项式回归：

$$y_i = \beta_0 + \beta_1 {x_i}^1 + \beta_2 {x_i}^2 + \beta_3 {x_i}^3 + \varepsilon_i$$

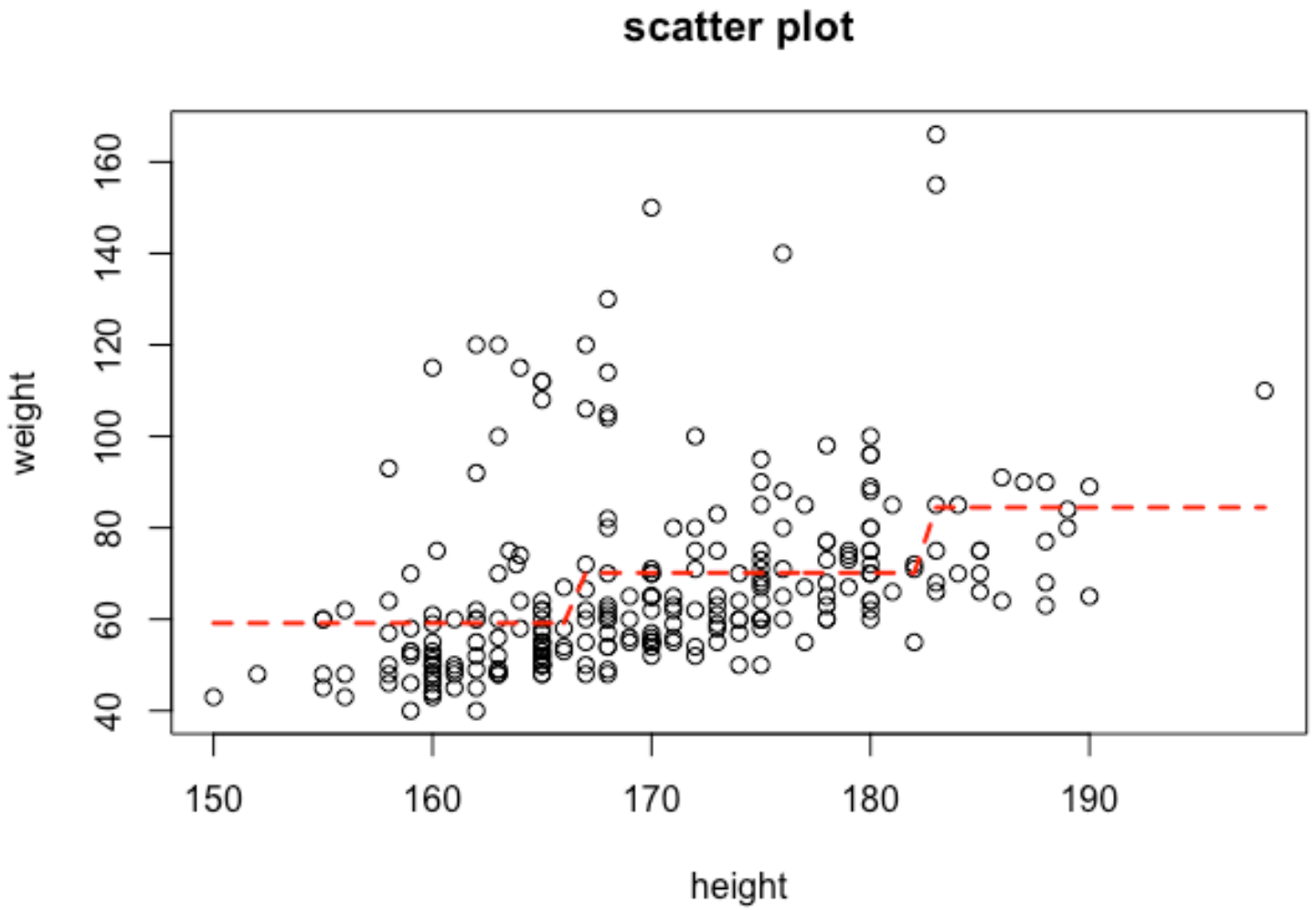

scatter plot
weight
height
40
60
80
100
120
140
160
150
160
170
180
190


scatter plot
weight
height
40
60
80
100
120
140
160
150
160
170
180
190


scatter plot
weight
height
40
60
80
100
120
140
160
150
160
170
180
190

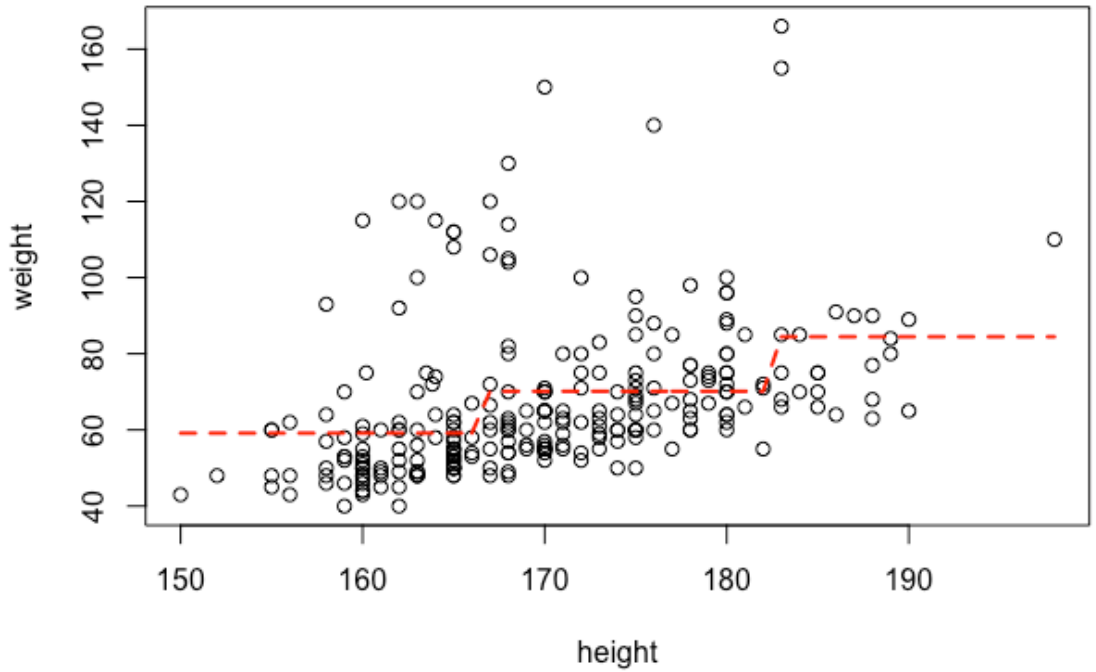

scatter plot
weight
height
40
60
80
100
120
140
160
150
160
170
180
190

分类型变量回归（跃阶函数）



```
Call:
lm(formula = weight ~ cut(height, 3))

Coefficients:
            (Intercept)  cut(height, 3)(166,182]  cut(height, 3)(182,198]
                  59.17                    10.98                    25.29
```



$$y = 59.17 + 10.98 \cdot I(\text{身高属于}(166,182]) + 25.29 \cdot I(\text{身高属于}(182,198]) = \begin{cases} 59.17\ \text{kg} & (\text{身高属于}150,166) \\ 70.15\text{kg} & (\text{身高属于}166,182) \\ 84.46\ kg & (\text{身高属于}182,198) \end{cases}$$

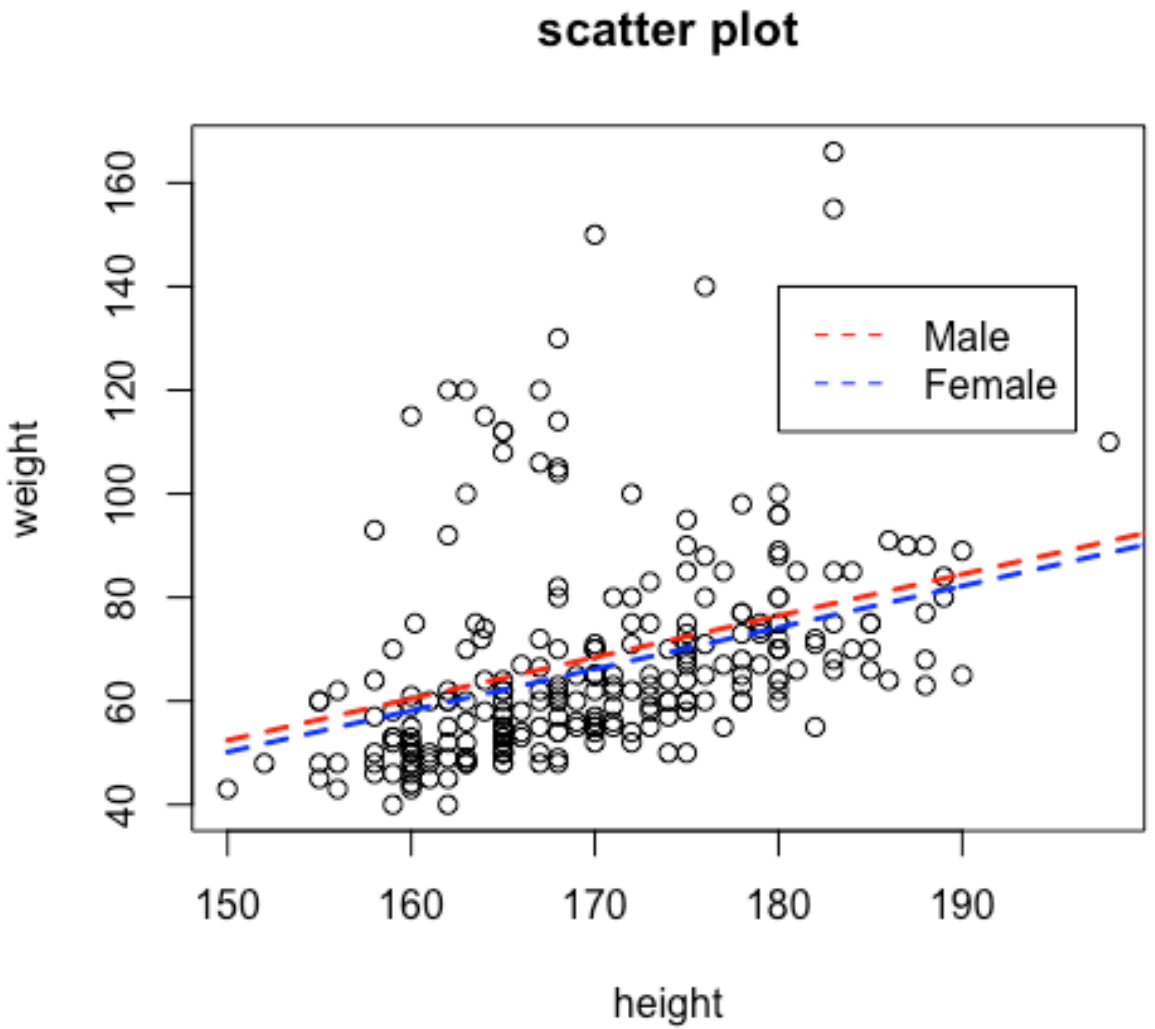
多分类型变量回归

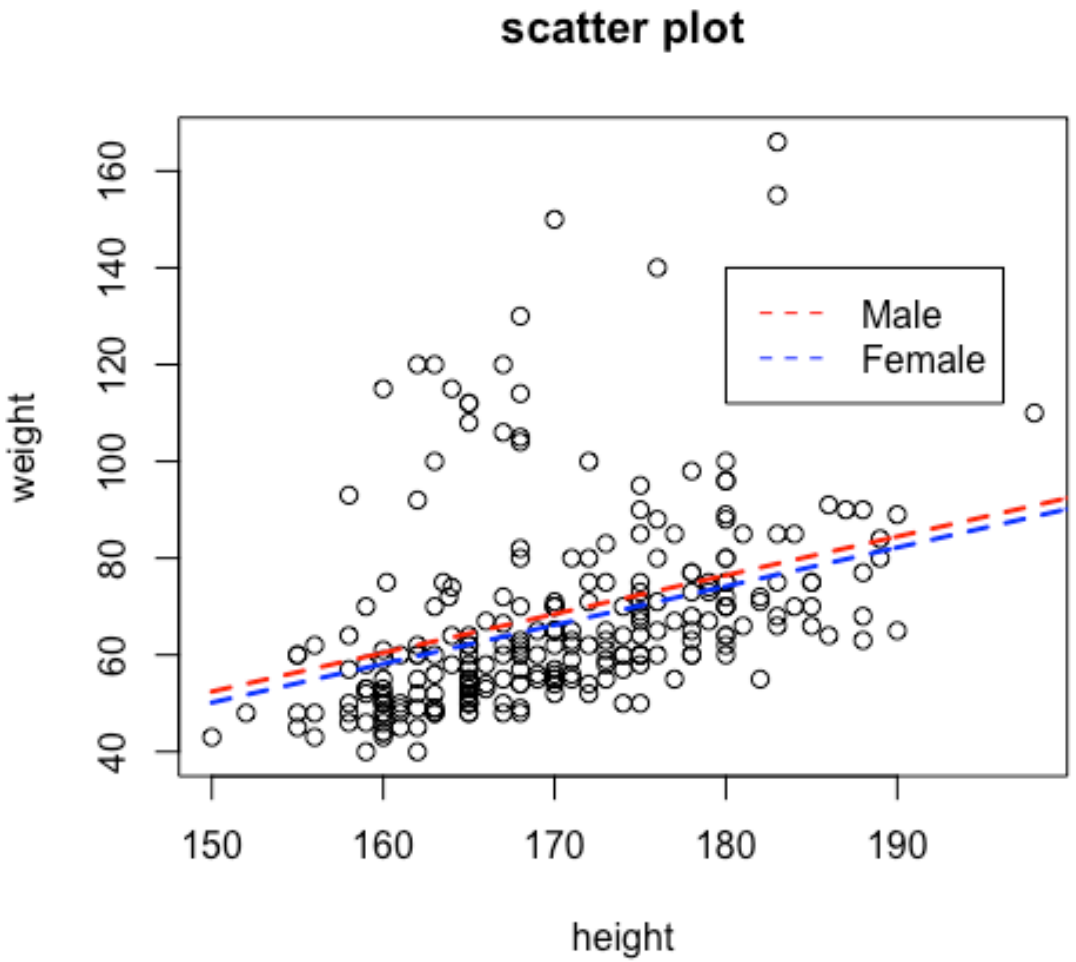
```
Call:
lm(formula = weight ~ cut(height, 3) + gender)

Coefficients:
            (Intercept)  cut(height, 3)(166,182]  cut(height, 3)(182,198]              genderMale
                 59.069                    7.923                   20.395                   4.994
```

$$weight$$
$$= 59.069 + 7.923 \cdot I(\text{身高属于}(166,182]) + 20.395 \cdot I(\text{身高属于}(182,198]) + 4.994 \cdot I(\text{男性})$$

分类型数值型混合回归


scatter plot
weight
height
40
60
80
100
120
140
160
150
160
170
180
190
Male
Female



$$weight = -69.9777 + 0.8009 * height + 2.2707 \cdot I(\text{男性})$$

$$= \begin{cases} \text{如果为男性：}(-69.9777 + 2.2707) + 0.8009 * height \\ \text{如果为女性：} -69.9777 + 0.8009 * height \end{cases}$$

```
Call:
lm(formula = weight ~ height + gender)

Coefficients:
(Intercept)       height   genderMale
   -69.9777       0.8009       2.2707
```

分类型数值型混合回归（含交互作用）

```
Call:
lm(formula = weight ~ height * gender)

Coefficients:
      (Intercept)             height         genderMale  height:genderMale
         -57.3487             0.7241            -24.1839             0.1547
```

$$weight = -57.3487 + 0.7241 * height - 24.1839 \cdot I(\text{男性}) + 0.1547 * I(\text{男性}) * height$$

$$= \begin{cases} \text{如果为男性：} (-57.3487 - 24.1839) + (0.7241 + 0.1547) * height \\ \text{如果为女性：} -57.3487 + 0.7241 * height \end{cases}$$