# 对于剩下三个数据的回归

## 基本分析

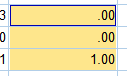
对给定数据进行基本分析：我们使用二元逻辑回归：

|  |  |  |  |
| --- | --- | --- | --- |
| 回归结果 | | | |
| 未加权个案数a | | 个案数 | 百分比 |
| 选定的个案 | 包括在分析中的个案数 | 38 | 90.5 |
| 缺失个案数 | 4 | 9.5 |
| 总计 | 42 | 100.0 |
| 未选定的个案 | | 0 | .0 |
| 总计 | | 42 | 100.0 |
| a. 如果权重为生效状态，请参阅分类表以了解个案总数。 | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **分类表a** | | | | | |
|  | 实测 | | 预测 | | |
|  | fruit\_name=apple | | 正确百分比 |
|  | .00 | 1.00 |
| 步骤 1 | fruit\_name=apple | .00 | 15 | 4 | 78.9 |
| 1.00 | 5 | 14 | 73.7 |
| 总体百分比 | |  |  | 76.3 |
| a. 分界值为 .500 | | | | | |

可以发现一共38个数据, 结果并不是很好, 仅有76%的正确率.

预测结果:



显著性水平：

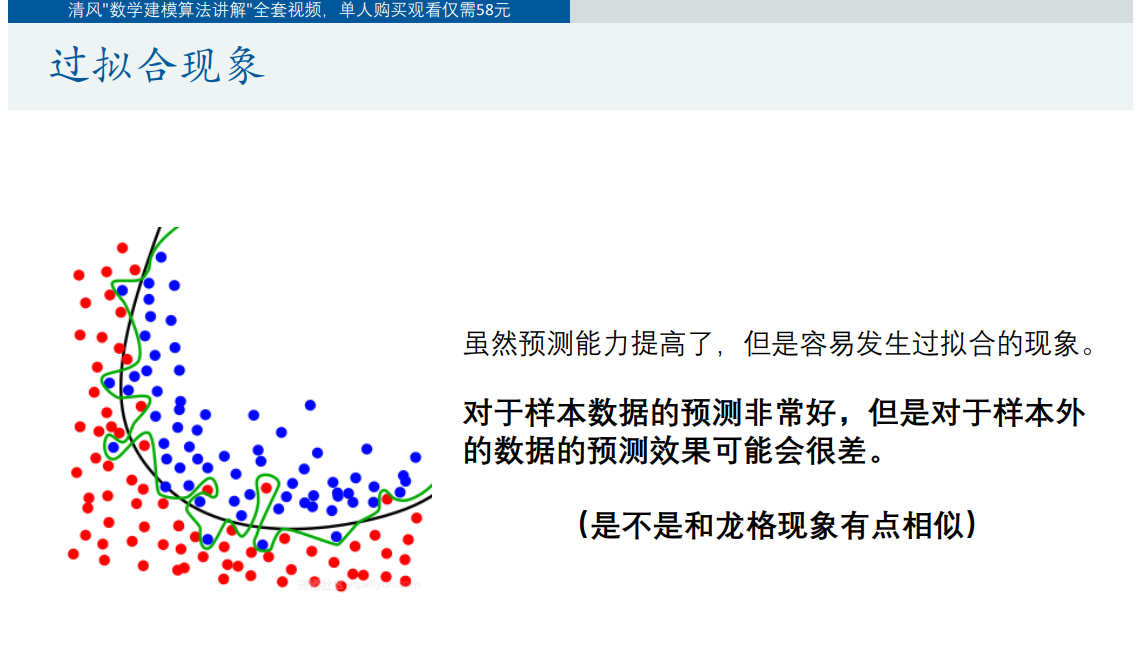
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | |
|  | | B | 标准误差 | 瓦尔德 | 自由度 | 显著性 | Exp(B) |
| 步骤 1a | mass | -.024 | .024 | .965 | 1 | .326 | .977 |
| width | 4.307 | 1.844 | 5.452 | 1 | .020 | 74.199 |
| height | -3.750 | 1.641 | 5.224 | 1 | .022 | .024 |
| color\_score | 9.891 | 5.746 | 2.964 | 1 | .085 | 19758.273 |
| 常量 | -7.202 | 14.503 | .247 | 1 | .620 | .001 |
| a. 在步骤 1 输入的变量：mass, width, height, color\_score。 | | | | | | | |

## 加入平方项后的模型：

|  |  |  |  |
| --- | --- | --- | --- |
| **个案处理摘要** | | | |
| 未加权个案数a | | 个案数 | 百分比 |
| 选定的个案 | 包括在分析中的个案数 | 42 | 100.0 |
| 缺失个案数 | 0 | .0 |
| 总计 | 42 | 100.0 |
| 未选定的个案 | | 0 | .0 |
| 总计 | | 42 | 100.0 |
| a. 如果权重为生效状态，请参阅分类表以了解个案总数。 | | | |

回归结果：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **分类表a** | | | | | |
|  | 实测 | | 预测 | | |
|  | fruit\_name=apple | | 正确百分比 |
|  | .00 | 1.00 |
| 步骤 1 | fruit\_name=apple | .00 | 23 | 0 | 100.0 |
| 1.00 | 0 | 19 | 100.0 |
| 总体百分比 | |  |  | 100.0 |
| 1. 分界值为 .500   可见，模型运行的结果正确率达到了百分百正确，我们分析，由于数据的过多，回归可能出现了过拟合现象。 | | | | | |

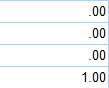


# Fisher线性判别：

通过Fisher线性判别，我们得到分析结果，有76%的正确率得到正确的结果。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **分类结果a** | | | | | |
|  |  | fruit\_name=apple | 预测组成员信息 | | 总计 |
|  |  | .00 | 1.00 |
| 原始 | 计数 | .00 | 15 | 4 | 19 |
| 1.00 | 5 | 14 | 19 |
| 未分组个案 | 3 | 1 | 4 |
| % | .00 | 78.9 | 21.1 | 100.0 |
| 1.00 | 26.3 | 73.7 | 100.0 |
| 未分组个案 | 75.0 | 25.0 | 100.0 |
| a. 正确地对 76.3% 个原始已分组个案进行了分类。 | | | | | |

预测的四个结果



# 多分类判别:

## 使用Fisher线性判别对多水果类别回归

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **分类结果a** | | | | | | | |
|  |  | kind | 预测组成员信息 | | | | 总计 |
|  |  | 1 | 2 | 3 | 4 |
| 原始 | 计数 | 1 | 14 | 0 | 5 | 0 | 19 |
| 2 | 0 | 16 | 0 | 0 | 16 |
| 3 | 4 | 1 | 14 | 0 | 19 |
| 4 | 0 | 0 | 0 | 5 | 5 |
| 未分组个案 | 2 | 1 | 4 | 1 | 8 |
| % | 1 | 73.7 | .0 | 26.3 | .0 | 100.0 |
| 2 | .0 | 100.0 | .0 | .0 | 100.0 |
| 3 | 21.1 | 5.3 | 73.7 | .0 | 100.0 |
| 4 | .0 | .0 | .0 | 100.0 | 100.0 |
| 未分组个案 | 25.0 | 12.5 | 50.0 | 12.5 | 100.0 |
| a. 正确地对 83.1% 个原始已分组个案进行了分类。 | | | | | | | |

显然：此次模型有83.1%的概率得到正确的结果。

所以本次回归的结果：

