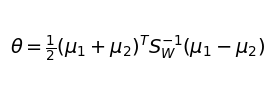
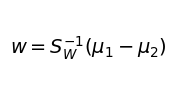
# 判别分析例题

## 题目描述

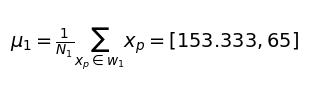
小明新入职的一家互联网公司通过员工的工作时长和客户满意度来评估员工的表现，并将员工分为高效员工和低效员工两类。以下是部分员工的数据。请你根据员工的工作时长和客户满意度数据，计算出判别函数，并解释判别分析中间变量的实际意义。小明的工作时长为160小时/月，客户满意度为75，请问他是高效员工吗？你认为这种评估方式对公司有何意义？

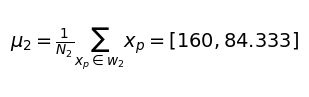
|  |  |  |  |
| --- | --- | --- | --- |
| 员工编号 | 工作时长（小时/月） | 客户满意度（满分100） | 类别 |
| 1 | 160 | 85 | 高效 |
| 2 | 170 | 90 | 高效 |
| 3 | 150 | 78 | 高效 |
| 4 | 140 | 65 | 低效 |
| 5 | 155 | 60 | 低效 |
| 6 | 165 | 70 | 低效 |

## Step 1: 判别向量和阈值

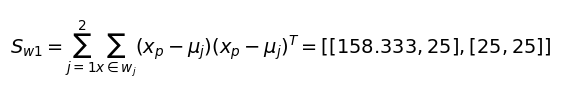


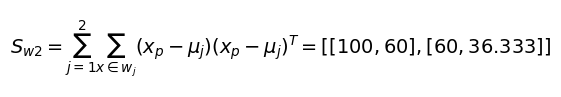
## Step 2: 两个类的均值向量

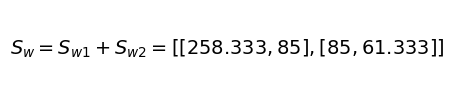




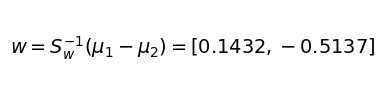
## Step 3: 散度矩阵

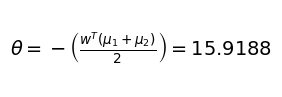




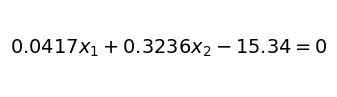


## Step 4: 计算 w 和 θ

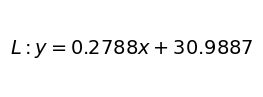




## Step 5: 判别函数方程



## Step 6: 直线方程

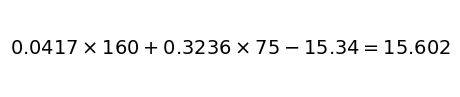


## Step 7: 解释判别分析中间变量的实际意义

判别向量 w 指示了哪些特征（工作时长和客户满意度）对区分高效和低效员工更重要。阈值 θ 用于确定分类的边界，帮助将员工分类为高效或低效。

## Step 4: 小明预测

小明的工作时长为160小时/月，客户满意度为75，代入判别函数：



由于判别函数值为正数，说明该员工属于高效员工。

## Step 5: 意义

通过这种判别分析，公司可以：  
- 识别和培养高效员工：重点关注和培养在工作时长和客户满意度上表现突出的员工。  
- 优化培训和激励措施：针对低效员工，提供更有针对性的培训和激励，帮助他们提升表现。  
- 决策支持：使用判别分析的结果作为决策依据，优化员工管理和绩效评估体系。

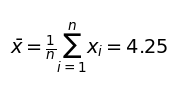
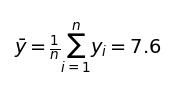
# 线性回归分析题目

## 题目描述

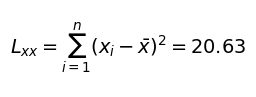
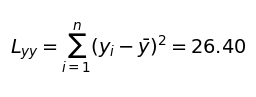
已知某公司广告费用（单位：万元）与销售额（单位：万元）的数据如下表所示，请你根据线性回归分析模型预测，当广告费用为8万元时，销售额为多少？

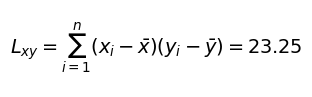
|  |  |
| --- | --- |
| 广告费用 x（万元） | 销售额 y（万元） |
| 2.0 | 5.0 |
| 2.5 | 5.5 |
| 3.0 | 6.0 |
| 3.5 | 7.0 |
| 4.0 | 7.5 |
| 4.5 | 8.0 |
| 5.0 | 8.5 |
| 5.5 | 9.0 |
| 6.0 | 9.5 |
| 6.5 | 10.0 |

## Step 1: 计算平均值

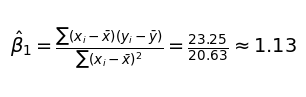
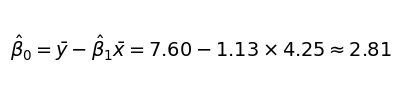
 

## Step 2: 计算相关系数

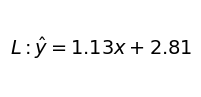
 



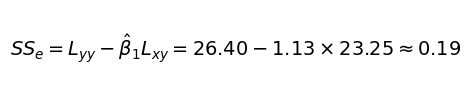
## Step 3: 计算估计参数

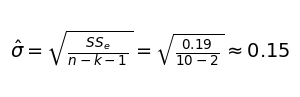
 

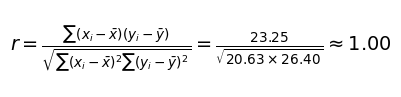
## Step 4: 直线方程

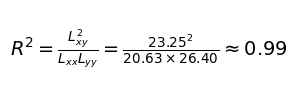


## Step 6: 误差分解



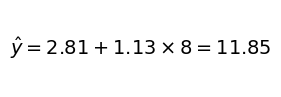


接近完美线性相关

表明模型解释了约 99% 的变异性

## 模型预测

根据模型预测，当广告费用为8万元时，销售额为：



# 同一题目的支持向量机分析

已知一个银行的用户数据如下：[存款金额（单位：万=美元），信誉积分]。

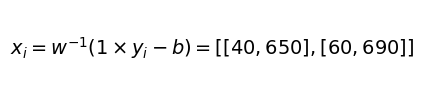
标记为1是办卡成功，标记为0则办卡失败。

请你分析出支持向量机分析算出用户办卡的决策边界方程

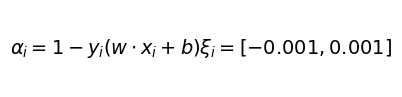
## Step 1: 数据和标签

|  |  |
| --- | --- |
| X | y |
| [45, 720] | 1 |
| [40, 650] | 0 |
| [85, 800] | 1 |
| [30, 580] | 0 |
| [60, 690] | 1 |

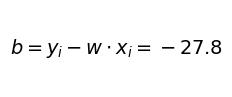
## Step 2: 支持向量



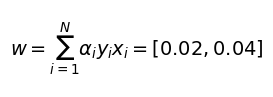
## Step 3: 拉格朗日乘子



## Step 4: 偏置项



## Step 5: 超平面的法向量



## Step 6: 决策边界方程

