

题目一（15 分）：

假设有一组数据  $X = \{(x_1, y_1), \dots, (x_m, y_m)\}$ ，我们希望求出对应的一元线性模型来拟合这一组数据：

$$y = \beta_0 + \beta_1 x$$

请证明，按照最小二乘法的计算方式，最终的结果如下：

$$\beta_0 = \bar{y} - \beta_1 \bar{x}$$

$$\beta_1 = \frac{\sum_{i=1}^m x_i y_i - \bar{y} \sum_{i=1}^m x_i}{\sum_{i=1}^m x_i^2 - \bar{x} \sum_{i=1}^m x_i} = \frac{\sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^m (x_i - \bar{x})^2}$$

其中  $\bar{x}$  和  $\bar{y}$  是  $x$  和  $y$  的平均值

题目二（10 分）：

请说明，当求解  $n$  组数据的标准差时，为什么分母除以的是  $(n-1)$  而不是  $n$ 。