

第 15 次作业（提交截止时间：6 月 10 日上午 9:50）

1. 假定变量 Y （响应变量）与 X （预测变量）之间的关系可用如下的线性模型刻画： $Y = \beta_0 + \beta_1 X + \varepsilon$ ，其中 β_0 是常数项， β_1 称为模型的回归系数，皆为

常数， ε 为随机误差项，均值为 0，方差为 σ^2 。假设有 (X, Y) 的独立观测：

(x_i, y_i) ($i = 1, \dots, n$)。参数 β_0, β_1 的最小二乘估计分别表示为 $\hat{\beta}_0, \hat{\beta}_1$ 。

(1) 验证：最小二乘法拟合的直线经过点 (\bar{x}, \bar{y}) 。

(2) 计算 $\text{Cov}(\hat{\beta}_0, \hat{\beta}_1)$ 。什么时候 $\hat{\beta}_0, \hat{\beta}_1$ 不相关？

(3) 假设 $x_i \in [-1, 1]$ ($i = 1, \dots, n$)，为最小化 $\text{Var}(\hat{\beta}_1)$ ，应该如何选择 x_i ($i = 1, \dots, n$)？

(4) 如果模型可以事先假设 $\beta_0 = 0$ ，则请在此情况下给出 β_1 的最小二乘估计。

2. 进一步假设随机误差 $\varepsilon \sim N(0, \sigma^2)$ 。

(1) 给出参数 σ^2 的极大似然估计。

(2) 证明： $\frac{\text{RSS}}{n-2}$ 是 σ^2 的无偏估计。

(3) 已知模型可以事先假设 $\beta_0 = 0$ ，请给出 σ^2 的无偏估计，并给出

$X = x_0$ 时响应变量 Y 取值的 $(1-\alpha)$ 置信的区间估计。