

Пусть $\{X_i\}_{i=1}^n$ - таргеты обучающей выборки, попавшей в лист, а y - таргет исследуемого объекта. Тогда в первом случае математическое ожидание ошибки равно

$$\begin{aligned} E(MSE_1) &= E \left(y - \frac{\sum_{i=1}^n X_i}{n} \right)^2 = Ey^2 + nE \frac{X_1^2}{n^2} - 2nEyE \frac{X_1}{n} + n(n-1) \left(E \frac{X_1}{n} \right)^2 = \\ &= Ey^2 - 2EyEX_1 + \frac{EX_1^2}{n} + \left(1 - \frac{1}{n}\right)(EX_1)^2, \end{aligned}$$

а во втором оно равно

$$E(MSE_2) = E(y - X_k)^2 = Ey^2 - 2EyEX_1 + EX_1^2.$$

$$E(MSE_2 - MSE_1) = EX_1^2 \left(1 - \frac{1}{n}\right) - \left(1 - \frac{1}{n}\right)(EX_1)^2 = EX_1 \left(1 - \frac{1}{n}\right) > 0$$

Таким образом, первая стратегия приводит к меньшему математическому ожиданию.