

EXAM 2022

Problem 2

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2)$$

$$C(\beta_0, \beta_1) = \sum_{i=1}^N (y_i - \beta_0 - \beta_1 x_i)^2 + \lambda(\beta_1 - 1.5)^2$$

$$a - \tilde{x}_i = x_i - \bar{x} \rightarrow x_i = \tilde{x}_i + \bar{x}$$

$$\rightarrow C(\beta_0, \beta_1) = \sum_{i=1}^N (y_i - \beta_0 - \beta_1(\tilde{x}_i + \bar{x}))^2 + \lambda(\beta_1 - 1.5)^2$$

$$= \sum_{i=1}^N (y_i - \underbrace{\beta_0 + \beta_1 \bar{x}}_{-\tilde{\beta}_0} - \beta_1 \tilde{x}_i)^2 + \lambda(\beta_1 - 1.5)^2$$

$$= \sum_{i=1}^N (y_i - \tilde{\beta}_0 - \beta_1 \tilde{x}_i)^2 + \lambda(\beta_1 - 1.5)^2$$

$$= C(\tilde{\beta}_0, \beta_1)$$

$$\text{with } \tilde{\beta}_0 = \beta_0 + \beta_1 \bar{x}$$

Further, we have

$$\begin{aligned} \sum_{i=1}^N \tilde{x}_i &= \sum_{i=1}^N (x_i - \bar{x}) = \sum_{i=1}^N x_i - N \cdot \bar{x} \\ &= \sum_{i=1}^N x_i - \cancel{N} \cdot \frac{1}{\cancel{N}} \underbrace{\sum_{i=1}^N x_i}_{N \cdot \bar{x}} \\ &= \underline{\underline{0}} \end{aligned}$$

b. We want to minimise $C(\tilde{\beta}_0, \beta_1)$ w.r.t $\tilde{\beta}_0$ and β_1 :

$$\frac{\partial C(\tilde{\beta}_0, \beta_1)}{\partial \tilde{\beta}_0} = \frac{\partial}{\partial \tilde{\beta}_0} \left(\sum_{i=1}^N (y_i - \tilde{\beta}_0 - \beta_1 \tilde{x}_i)^2 + \lambda(\beta_1 - 1.5)^2 \right)$$

$$= 2 \sum_{i=1}^N (y_i - \tilde{\beta}_0 - \beta_1 \tilde{x}_i) \cdot (-1)$$

$$= -2 \sum_{i=1}^N (y_i - \tilde{\beta}_0) + 2 \underbrace{\sum_{i=1}^N \beta_1 \tilde{x}_i}_{\beta_1 \underbrace{\sum_{i=1}^N \tilde{x}_i}_{=0}}$$

$$= -2 \sum_{i=1}^N (y_i - \tilde{\beta}_0)$$

$$\longrightarrow -2 \sum_{i=1}^N (y_i - \hat{\tilde{\beta}}_0) = 0$$

$$\longrightarrow \sum_{i=1}^N y_i - N \cdot \hat{\tilde{\beta}}_0 = 0 \longrightarrow \hat{\tilde{\beta}}_0 = \frac{1}{N} \sum_{i=1}^N y_i = \bar{y}$$

$$\frac{\partial C(\tilde{\beta}_0, \beta_1)}{\partial \beta_1} = \frac{\partial}{\partial \beta_1} \left(\sum_{i=1}^N (y_i - \tilde{\beta}_0 - \beta_1 \tilde{x}_i)^2 + \lambda(\beta_1 - 1.5)^2 \right)$$

$$= 2 \sum_{i=1}^N (y_i - \tilde{\beta}_0 - \beta_1 \tilde{x}_i) (-\tilde{x}_i) + 2\lambda(\beta_1 - 1.5)$$

$$= -2 \sum_{i=1}^N (y_i - \tilde{\beta}_0 - \beta_1 \tilde{x}_i) \tilde{x}_i + 2\tilde{\beta}_0 \underbrace{\sum_{i=1}^N \tilde{x}_i}_{=0} + 2\lambda\beta_1 - 2 \cdot 1.5\lambda$$

$$= -2 \sum_{i=1}^N y_i \tilde{x}_i + 2\beta_1 \sum_{i=1}^N \tilde{x}_i^2 + 2\lambda\beta_1 - 2 \cdot 1.5\lambda$$

$$= 2\beta_1 \left(\sum_{i=1}^N \tilde{x}_i^2 + \lambda \right) - 2 \left(\sum_{i=1}^N y_i \tilde{x}_i + 1.5\lambda \right)$$

$$\longrightarrow 2\hat{\beta}_1 \left(\sum_{i=1}^N \tilde{x}_i^2 + \lambda \right) - 2 \left(\sum_{i=1}^N y_i \tilde{x}_i + 1.5\lambda \right) = 0$$

$$\longrightarrow \hat{\beta}_1 = \frac{\sum_{i=1}^N (x_i - \bar{x}) y_i + 1.5\lambda}{\sum_{i=1}^N (x_i - \bar{x})^2 + \lambda}$$

$$\hat{\beta}_0 = \hat{\tilde{\beta}}_0 - \hat{\beta}_1 \bar{x}$$

$$c. \quad \hat{\beta}_1^{OLS} = \frac{\sum_{i=1}^N (x_i - \bar{x}) y_i}{\sum_{i=1}^N (x_i - \bar{x})^2}$$

In b., we found

$$\hat{\beta}_1 = \frac{\sum_{i=1}^N (x_i - \bar{x}) y_i + 1,5 \lambda}{\sum_{i=1}^N (x_i - \bar{x})^2 + \lambda}$$

$$= \frac{\sum_{i=1}^N (x_i - \bar{x}) y_i}{\sum_{i=1}^N (x_i - \bar{x})^2 + \lambda} + \frac{1,5 \lambda}{\sum_{i=1}^N (x_i - \bar{x})^2 + \lambda}$$

$$= \underbrace{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{\sum_{i=1}^N (x_i - \bar{x})^2 + \lambda}}_{\alpha} \underbrace{\frac{\sum_{i=1}^N (x_i - \bar{x}) y_i}{\sum_{i=1}^N (x_i - \bar{x})^2}}_{\hat{\beta}_1^{OLS}} + \frac{\lambda}{\sum_{i=1}^N (x_i - \bar{x})^2 + \lambda} \cdot 1,5$$

$$\alpha = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{\sum_{i=1}^N (x_i - \bar{x})^2 + \lambda} \rightarrow 1 - \alpha = 1 - \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{\sum_{i=1}^N (x_i - \bar{x})^2 + \lambda}$$

$$= \frac{\cancel{\sum_{i=1}^N (x_i - \bar{x})^2} + \lambda - \cancel{\sum_{i=1}^N (x_i - \bar{x})^2}}{\sum_{i=1}^N (x_i - \bar{x})^2 + \lambda}$$

$$= \frac{\lambda}{\sum_{i=1}^N (x_i - \bar{x})^2 + \lambda}$$

$$\rightarrow \hat{\beta}_1 = \alpha \cdot \hat{\beta}_1^{OLS} + (1 - \alpha) \cdot 1,5, \text{ with } \alpha = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{\sum_{i=1}^N (x_i - \bar{x})^2 + \lambda}$$

Problem 3

a. In a logistic regression model:

$$\text{odds}(x) = \frac{P(Y=1|x)}{P(Y=0|x)} = e^{x^T \beta}$$

For two covariate vectors x and x^* , we have

$$\frac{\text{odds}(x^*)}{\text{odds}(x)} = \frac{e^{x^{*T} \beta}}{e^{x^T \beta}} = e^{(x^* - x)^T \beta}$$

Let $x_j^* = x_j + 1$ and $x_h^* = x_h$, $h \neq j$, then

$$\frac{\text{odds}(x^*)}{\text{odds}(x)} = e^{\beta_j}$$

Use the odds ratios to interpret the effects.

- b. The log-likelihood "always" becomes lower when we reduce the number of parameters because we are optimizing over a more constrained space.

$$AIC = -2 \log\text{-like} + 2 \cdot (\underbrace{\text{number of free parameters}}_{\text{degrees of freedom}})$$

$$BIC = -2 \log\text{-like} + \log(N) \left(\frac{\text{degrees of freedom}}{n} \right)$$

Full model (Model 1):

$$AIC = -2 \cdot (-4164.612) + 2 \cdot 11 = \underline{\underline{8351.224}}$$

$$BIC = \text{—————} + \log(9020) \cdot 11 = \underline{\underline{8429.403}}$$

Model 2:

$$AIC = -2 \cdot (-4337.698) + 2 \cdot 3 = \underline{\underline{8681.396}}$$

$$BIC = -2 \cdot (\text{—————}) + \log(9020) \cdot 3 = \underline{\underline{8702.718}}$$

The full model has the lowest AIC and BIC, and therefore preferred.