Modelo de regressão linear: Parte I

Prof. Caio Azevedo

Modelo de regressão normal linear

- Modelo

$$
\begin{aligned}
Y_i &= \beta_0 + \sum_{j=1}^{p-1} X_{ij}\beta_j + \xi_i, i = 1, 2, ...n \\
\xi &\overset{iid}{\sim} N(0, \sigma^2)
\end{aligned}
$$

- Assim, temos que $Y_i \overset{iid}{\sim} N(\beta_0 + \sum_{j=1}^{p-1} X_{ij}\beta_j, \sigma^2)$.
- Consideremos, por enquanto σ^2 conhecido.

Aplicação no modelo de regressão normal linear

- Modelo

$$\mathbf{Y} = \mathbf{X}\beta + \boldsymbol{\xi}$$

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & X_{11} & \ldots & X_{1p} \\ 1 & X_{21} & \ldots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \ldots & X_{np} \end{bmatrix}, \boldsymbol{\xi} = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}$$

- Suposição $\boldsymbol{\xi} \sim N_n(\mathbf{0}, \sigma^2 \mathbf{I}_n)$.
- Estimador de mínimos quadrados de $\boldsymbol{\beta}$, minimizar $(\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)$, equivale a resolver $(\mathbf{X}'\mathbf{X})\widehat{\beta} = \mathbf{X}'\mathbf{Y}$.

Estimação por mínimos quadrados

- Solução $\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$.
- Sob normalidade $\widehat{\boldsymbol{\beta}} \sim N_p(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1})$.
- Consequentemente, $\beta_j \sim N_1(\beta_j, \boldsymbol{\Psi}_{jj}), j = 1, .., p$ em que $\boldsymbol{\Psi}_{jj} = (\sigma^2(\mathbf{X}'\mathbf{X})^{-1})_{jj}$.
- IC exato: $\widehat{\boldsymbol{\beta}}_j \pm z_{\frac{1-\gamma}{2}}\sqrt{\boldsymbol{\Psi}_{jj}}$. $P(Z > z_{\frac{1-\gamma}{2}}) = \frac{1-\gamma}{2}$, $Z \sim N(0,1)$.
- Testes de hipótese: $H_0 : \beta_j = 0$ vs $H_1 : \beta_j \neq 0$. Teste exato, rejeitar H_0 se $|Z_c| > z_{\frac{1-\gamma}{2}}$, em que

$$Z_c = \frac{\widehat{\boldsymbol{\beta}}_j}{\sqrt{\boldsymbol{\Psi}_{jj}}}$$

.

Inferência Baysiana

- Considere que $\boldsymbol{\beta} \sim N_p(\boldsymbol{\mu}_{\boldsymbol{\beta}}, \boldsymbol{\Psi}_{\boldsymbol{\beta}})$.
- Pode-se provar que $\beta|\mathbf{y} \sim N_p(\boldsymbol{\Psi}^*_{\boldsymbol{\beta}}\boldsymbol{\mu}^*_{\boldsymbol{\beta}}, \boldsymbol{\Psi}^*_{\boldsymbol{\beta}})$, em que

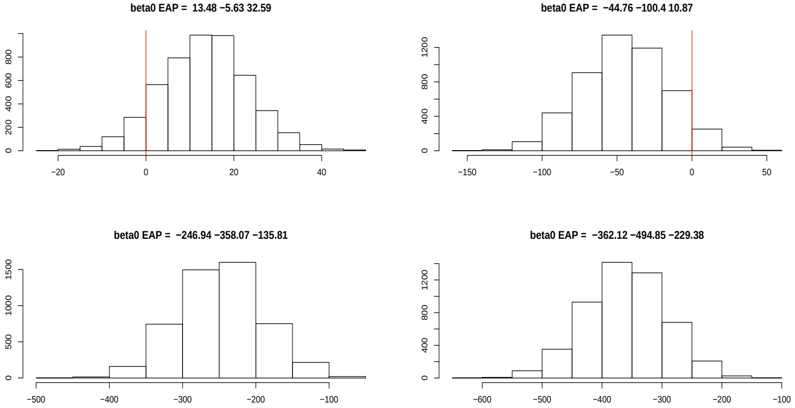
$$
\begin{aligned}
\boldsymbol{\Psi}^*_{\boldsymbol{\beta}} &= \left((\sigma^2)^{-1}\left(\mathbf{X}'\mathbf{X}\right) + \boldsymbol{\Psi}^{-1}_{\boldsymbol{\beta}}\right)^{-1} \\
\boldsymbol{\mu}^*_{\boldsymbol{\beta}} &= \boldsymbol{\Psi}^{-1}_{\boldsymbol{\beta}}\boldsymbol{\mu}_{\boldsymbol{\beta}} + (\sigma^2)^{-1}\mathbf{X}'\mathbf{Y}
\end{aligned}
$$

- Assim, $\beta_j|\mathbf{y} \sim N_1(\boldsymbol{\mu}^*_{\boldsymbol{\beta}_j}, \boldsymbol{\Psi}^*_{\boldsymbol{\beta}_{jj}})$, em que $\boldsymbol{\mu}^*_{\boldsymbol{\beta}_j} = \left(\boldsymbol{\Psi}^*_{\boldsymbol{\beta}}\boldsymbol{\mu}^*_{\boldsymbol{\beta}}\right)_j$ e $\boldsymbol{\Psi}^*_{\boldsymbol{\beta}_{jj}} = \left(\boldsymbol{\Psi}^*_{\boldsymbol{\beta}}\right)_{jj}$.

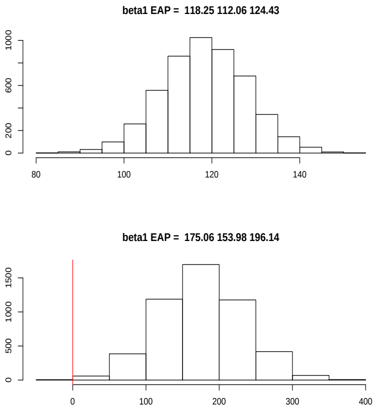
Cont.

- $IC_B = \boldsymbol{\mu}^*_{\boldsymbol{\beta}_j} \pm z_{\frac{1-\gamma}{2}} \sqrt{\boldsymbol{\Psi}^*_{\boldsymbol{\beta}_{jj}}}$.
- Teste de hipótese $\dfrac{P(\beta_j \in \Theta_1|\mathbf{y})}{P(\beta_j \in \Theta_0|\mathbf{y})}$
- Estimação (empírica) dos hiperparâmetros $(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma})$
- Dist. preditivia a priori: $Y_{i+1}|\mathbf{y} \sim N(\mathbf{x}_i'\boldsymbol{\mu}_{\boldsymbol{\beta}}, \mathbf{x}_i'\boldsymbol{\Sigma}_{\boldsymbol{\beta}}\mathbf{x}_i)$.
- Dist. preditivia a posteriori: $Y_{i+1}|\mathbf{y} \sim N(\mathbf{x}_i'\boldsymbol{\Sigma}^*_{\boldsymbol{\beta}}\boldsymbol{\mu}^*_{\boldsymbol{\beta}}, \mathbf{x}_i'\boldsymbol{\Sigma}^*_{\boldsymbol{\beta}}\mathbf{x}_i)$.
- Estimativa de mínimos quadrados: $-381,28(69,04), 199,83(13,03)$.

Posteriores



Posteriores



Cont.

- Resultados fortemente influenciados pelas prioris (n pequeno).
- Neste caso, a menos que haja alguma informação à priori confiável, melhor escolher

  $p(\boldsymbol{\beta}) \propto \mathbb{1}_{\mathcal{R}^p}(\boldsymbol{\beta})$