Bayesian Linear Models

Felipe José Bravo Márquez

July 5, 2021

Bayesian Linear Models

- In this class we are going to revisit the linear regression model from a Bayesian point of view.
- The idea is the same as in the frequentist approach, to model the relationship of a numerical dependent variable $\mathbf{y}$ with n independent variables $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$.
- We again use a Gaussian distribution to describe our uncertainty about the response variable: $y_i \sim N(\mu_i, \sigma^2)$.
- And we also assume that each attribute has a linear relationship to the mean of the outcome.

$$\mu_i = \beta_0 + \beta_1 x_i + \ldots \beta_n x_n$$

- Instead of using least squares or maximum likelihood estimation we are going to estimate the joint posterior distribution of all the parameters of the model:

$$f(\theta|d) = f(\beta_0, \beta_1, \ldots, \beta_n, \sigma|d)$$

- This approach is more flexible as it allows incorporating prior information.
- It also allows to interpret the uncertainty of the model in a clearer way.

Bayesian Linear Models

- Notice the the parameters of the model are $\beta_0, \beta_1, \ldots, \beta_b$ and σ.
- The mean of the outcome μ_i is not treated as parameter because it is determined deterministically from the linear model's coefficients.
- To complete the model, we need a joint prior density:

$$f(\theta) = f(\beta_0, \beta_1, \ldots, \beta_n, \sigma)$$

- And the posterior gets specified as follows:

$$f(\theta|d) = \frac{\prod_{i=1}^{m} f(d_i|\beta_0, \beta_1, \cdots, \beta_n, \sigma) * f(\beta_0, \beta_1, \ldots, \beta_n, \sigma)}{f(d)}$$

here d_i corresponds to each data point containing values for y and $x_1, \ldots, x_n$.

Bayesian Linear Models

- The evidence is expressed by a multiple integral:

$$f(d) = \int \int \cdots \int \prod_{i=1}^{m} f(d_i|\beta_0, \beta_1, \cdots, \beta_n, \sigma) * f(\beta_0, \beta_1, \ldots, \beta_n, \sigma) d\beta_0 d\beta_1 \cdots d\beta_n d\sigma$$

- In most cases, priors are specified independently for each parameter, which amounts to assuming:

$$f(\beta_0, \beta_1, \cdots, \beta_b, \sigma) = f(\beta_0) * f(\beta_1) * \cdots * f(\beta_n) * f(\sigma).$$

- This class is based on chapters 4 and 5 of [McElreath, 2020]

A model of height revisited

- Let's fit again a linear model relating height and weight for the !Kung San people.
- But his time we will use a Bayesian approach.
- We need to define a probabilistic model specifying all the components of the Bayesian model:

$$
\begin{aligned}
h_i &\sim \text{Normal}(\mu_i, \sigma) && \text{[likelihood]} \\
\mu_i &= \alpha + \beta x_i && \text{[linear model]} \\
\alpha &\sim \text{Normal}(178, 100) && [\alpha \text{ prior}] \\
\beta &\sim \text{Normal}(0, 10) && [\beta \text{ prior}] \\
\sigma &\sim \text{Uniform}(0, 50) && [\sigma \text{ prior}]
\end{aligned}
$$

Conclusions

- Blabla

References I

McElreath, R. (2020).
Statistical rethinking: A Bayesian course with examples in R and Stan.
CRC press.