Networks in Context Lab, 2024

Bayesian Statistics and Multilevel Models

Barum Park
Department of Sociology
Cornell University

March 6, 2024

Let us start by fitting a simple linear regression in `Stan`:

$$y_i = \alpha + \beta x_i + \epsilon, \quad \sigma_\epsilon \text{ Normal}(0, \sigma_\epsilon)$$

or equivalently

$$y_i \sim \text{Normal}(\alpha + \beta x_i, \sigma_\epsilon)$$

Let us start by fitting a simple linear regression in `Stan`:

$$y_i = \alpha + \beta x_i + \epsilon, \quad \sigma_\epsilon \text{ Normal}(0, \sigma_\epsilon)$$

or equivalently

$$y_i \sim \text{Normal}(\alpha + \beta x_i, \sigma_\epsilon)$$

The data consists of two vectors:

1. the outcome $y = [y_1, ..., y_i, ..., y_N]^\top$; and
2. the predictor $x = [x_1, ..., x_i, ..., x_N]^\top$

Let us start by fitting a simple linear regression in `Stan`:

$$y_i = \alpha + \beta x_i + \epsilon, \quad \sigma_\epsilon \text{ Normal}(0, \sigma_\epsilon)$$

or equivalently

$$y_i \sim \text{Normal}(\alpha + \beta x_i, \sigma_\epsilon)$$

The data consists of two vectors:

1. the outcome $y = [y_1, ..., y_i, ..., y_N]^\top$; and
2. the predictor $x = [x_1, ..., x_i, ..., x_N]^\top$

We specify this in the data `data{}` block.

Let us start by fitting a simple linear regression in `Stan`:

$$y_i = \alpha + \beta x_i + \epsilon, \quad \sigma_\epsilon \ \mathrm{Normal}(0, \sigma_\epsilon)$$

or equivalently

$$y_i \sim \mathrm{Normal}(\alpha + \beta x_i, \sigma_\epsilon)$$

The data consists of two vectors:

1. the outcome $y = [y_1, ..., y_i, ..., y_N]^\top$; and
2. the predictor $x = [x_1, ..., x_i, ..., x_N]^\top$

We specify this in the data `data{}` block.

```
data {
    int N;            // no. of obs.
    vector[N] x;      // predictor
    vector[N] y;      // outcome
}
```

The only parameters in the model are

The only parameters in the model are

1. $\alpha \in \mathbb{R}$: the intercept

The only parameters in the model are

1. $\alpha \in \mathbb{R}$: the intercept
2. $\beta \in \mathbb{R}$: the slope coefficient

The only parameters in the model are

1. $\alpha \in \mathbb{R}$: the intercept
2. $\beta \in \mathbb{R}$: the slope coefficient
3. $\sigma_\epsilon \in \mathbb{R}_+$, the residual standard deviation.

The only parameters in the model are

1. $\alpha \in \mathbb{R}$: the intercept
2. $\beta \in \mathbb{R}$: the slope coefficient
3. $\sigma_\epsilon \in \mathbb{R}_+$, the residual standard deviation.

We can declare them in the `parameters{}` block:

```
parameters {
    real alpha;
    real beta;
    real<lower = 0> sigma_epsilon;
}
```

Notice that we specified `real<lower = 0>` to indicate that σ_ϵ has to be positive

Lastly, we specify the (log) posterior distribution up to a constant.

Lastly, we specify the (log) posterior distribution up to a constant.This is equivalent to specifying the likelihood and prior

Lastly, we specify the (log) posterior distribution up to a constant.This is equivalent to specifying the likelihood and prior

Let us use the following weakly informative priors

$$\alpha \sim \text{Normal}(0, 3)$$
$$\beta \sim \text{Normal}(0, 1)$$
$$\sigma_\epsilon \sim \text{Exponential}(3)$$

As for the likelihood, notice that the model postulates

$$y_i = \alpha + x_i\beta + \epsilon_i$$

where $\epsilon_i \overset{\text{iid}}{\sim} \text{Normal}(0, \sigma_\epsilon)$.

As for the likelihood, notice that the model postulates

$$y_i = \alpha + x_i\beta + \epsilon_i$$

where $\epsilon_i \overset{\text{iid}}{\sim} \text{Normal}(0, \sigma_\epsilon)$. Hence, the likelihood contribution of a single observation is

$$p(y_i \,|\, \alpha, \beta, \sigma_\epsilon) = \text{Normal}(\alpha + x_i\beta, \sigma_\epsilon)$$

As for the likelihood, notice that the model postulates

$$y_i = \alpha + x_i\beta + \epsilon_i$$

where $\epsilon_i \overset{\text{iid}}{\sim} \text{Normal}(0, \sigma_\epsilon)$. Hence, the likelihood contribution of a single observation is

$$p(y_i \,|\, \alpha, \beta, \sigma_\epsilon) = \text{Normal}(\alpha + x_i\beta, \sigma_\epsilon)$$

and the likelihood of the entire dataset is

$$p(y \,|\, \alpha, \beta, \sigma_\epsilon) = \prod_{i=1}^{N} \text{Normal}(\alpha + \beta x_i, \sigma_\epsilon)$$

We code this up in the `model` block as follows:

```
model {
    // linear predictor (local variable)
    vector[N] xb;

    for (n in 1:N)
        xb[n] = alpha + beta * x[n];

    // priors
    alpha ~ normal(0, 3);
    beta ~ normal(0, 1);
    sigma_epsilon ~ exponential(3);

    // vectorized log-likelihood
    y ~ normal(xb, sigma_epsilon);

}
```