Bayesian Regression

Now we move to bayesian regression which is nothing but fitting a model on a data to determine the outcome of new test data.

First of all few points about bayesian regression -

1. In Bayesian regression The response, y, is not estimated as a single value, but is assumed to be drawn from a probability distribution. So if we see Bayesian linear regression the output variable y is not a single value(as in frequentist approach) but a probability distribution.

2. The aim of Bayesian Linear Regression is not to find the single “best” value of the model parameters, but rather to determine the posterior distribution for the model parameters.

3. Not only for linear but also for logistic or poisson bayesian draws a probabilistic approach of estimating model parameters(implemented in previous parts).

Poisson regression

Since our model is assumed to have poisson distribution , therefore we need to consider bayesian poisson regression.

Linear regression reminder

When we have a response y that is continuous from $-\infty$ to ∞, we can consider using a linear regression represented by:

$$y = N(\mu, \sigma)$$

$$\mu = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ \ldots$$

We read this as: our response is normally distributed around μ with a standard deviation of σ. The value of μ is described by a linear function of explanatory variables $X\beta$ with a baseline intercept β_0.

Now since our distribution is not normally distributed but it is poisson distribution therefore we require to have some Link function to scale our result for poisson distribution.

Link Function

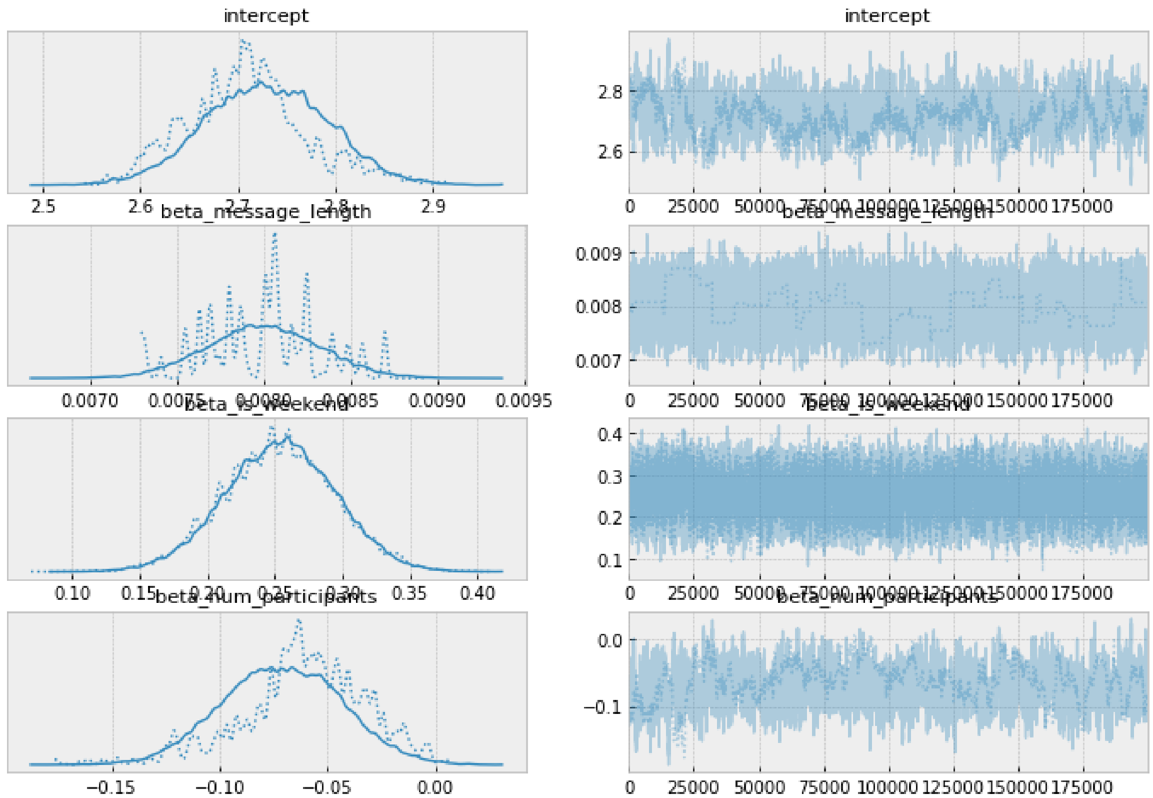
For a Poisson distribution, the canonical link function used is the log link. This can be formally described as:

$$y = Poi(\mu)$$

$$log(\mu) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ \dots$$

$$\mu = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ \dots}$$

This is considered to be a fixed effects model.



As you can see from the above results, the baseline intercept β_0 has an estimated value of between 2.5 and 2.9.

In the linear regression, we could say for every unit increase in x, $\hat{y}$ increases by β. However, in the Poisson regression we need to consider the link function.

For a Poisson model, given a unit change in x, the fitted $\hat{y}$ changes by $\hat{y}\left(e^{\beta} - 1\right)$

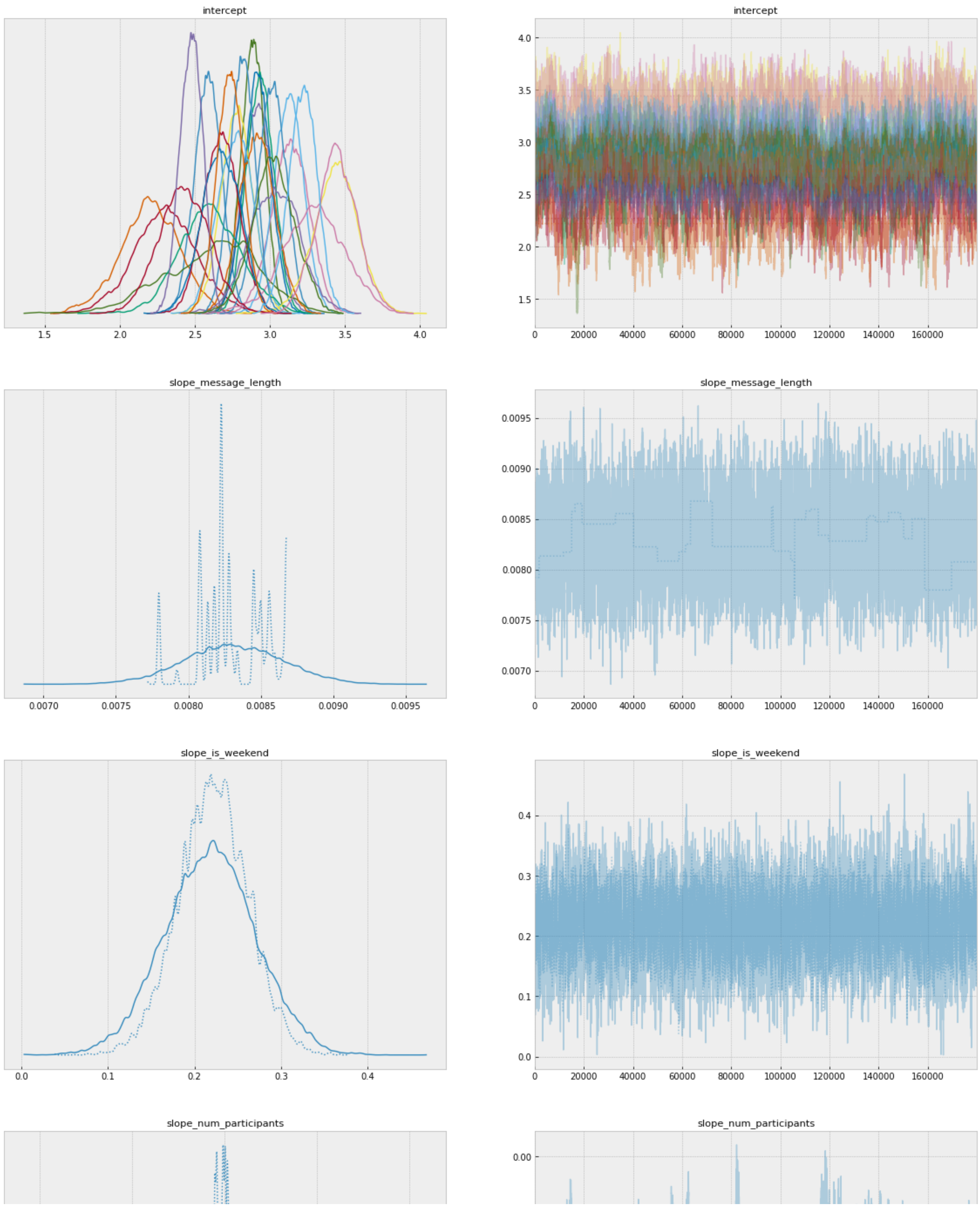
So an important observation in this is that the effect of changing x depends on the current value of y. Unlike the simple linear regression, a unit change in x does not cause a consistent change in y.

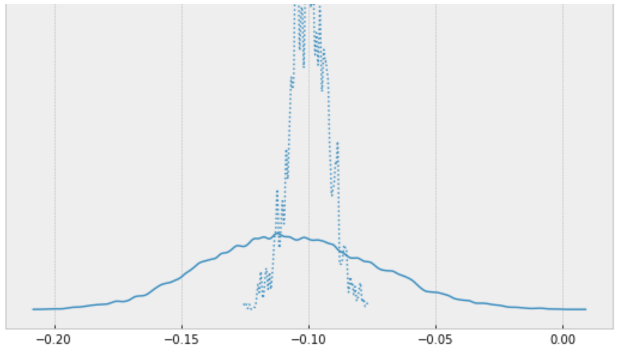
We can extend the above model by including a random intercept parameter. This will allow to estimate a baseline parameter value β_0 for everybody the person communicate with. For all the other parameters, a parameter over all the entire population will be estimated.

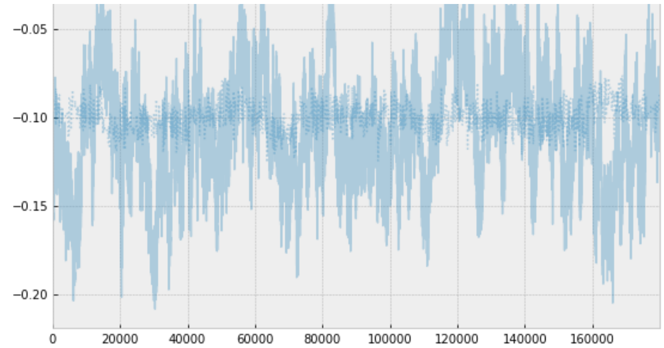
For each person i and each message j, this is represented as

$$y_{ji} \sim \text{Poisson}(\mu)$$
$$\mu = \beta_{0_i} + \beta_1 x_1 \cdot \cdots \cdot \beta_n x_n$$

By introducing this random effects parameter β_0 for each person i, it allows the model to establish a different baseline for each person responded to - whilst estimating the effects of the covariates on the response for the entire population.



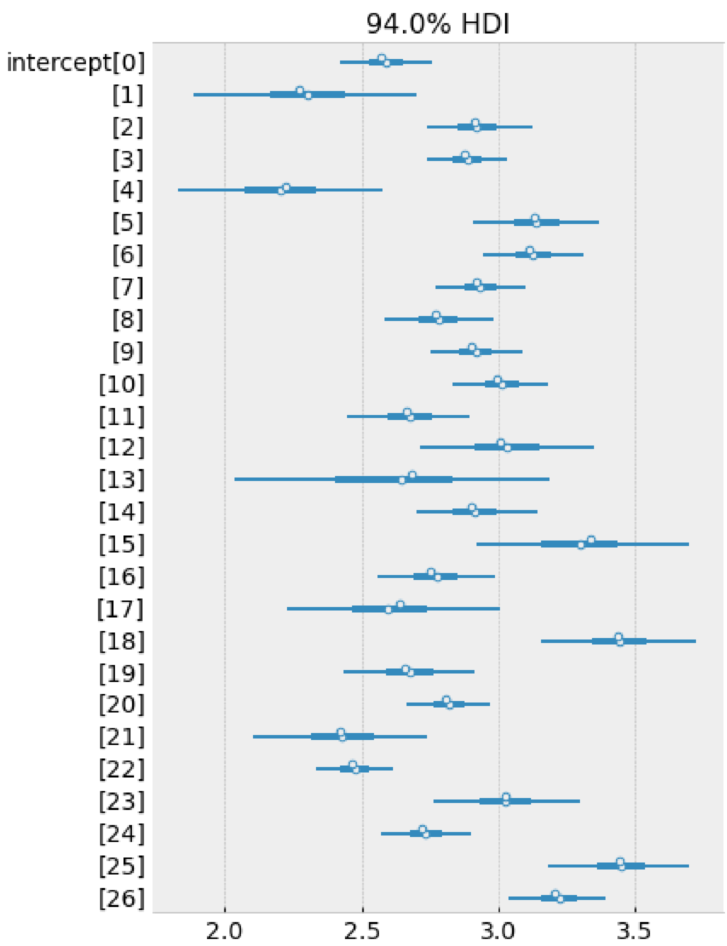




The interpretation of the above results are interesting:

- Each person has a different baseline response rate
- Longer messages take marginally longer to respond to
- A person is more likely to get a slow response if they message the person on the weekend
- The person tend to reply more quickly to conversations that have multiple people added to it (group chat)

And after accounting for the effect of each covariate on the response, the model estimates the below β_0 parameters.



The above image has no names as they were removed from the dataset for privacy concerns.