

separate build

Consider the simple two-part model where we model our response as

$$\begin{aligned} p_i &= \frac{1}{1 + e^{-(x_i\gamma_1 + \gamma_0)}} \\ y_i^* &= x_i\beta_1 + \beta_0 + \epsilon_i \quad \text{where} \quad \epsilon_i \sim \mathcal{N}(0, \sigma^2) \\ y_i &= p_i \cdot y_i^* \end{aligned}$$

Under a Bayesian frame we could define a full posterior for all model parameters as

$$\begin{aligned} f(\beta_1, \beta_0, \gamma_1, \gamma_0, \sigma^2 \mid y) &\propto f(y \mid \beta_1, \beta_0, \gamma_1, \gamma_0, \sigma^2) f(\beta_1, \beta_0, \gamma_1, \gamma_0, \sigma^2) \\ &= L(\beta_1, \beta_0, \gamma_1, \gamma_0, \sigma^2 \mid y) f(\beta_1, \beta_0, \gamma_1, \gamma_0, \sigma^2) \\ &= L(\beta_1, \beta_0, \gamma_1, \gamma_0, \sigma^2 \mid y) f(\beta_1) f(\beta_0) f(\gamma_1) f(\gamma_0) f(\sigma^2) \end{aligned}$$

We can expand this by writing out the likelihood more fully based on whether y is zero or not:

$$f(\beta_1, \beta_0, \gamma_1, \gamma_0, \sigma^2 \mid y) \propto \left[\prod_{i:y_i=0} (1 - p_i) \prod_{i:y_i>0} p_i f(y_i \mid \beta_1, \beta_0, \sigma^2) \right] f(\beta_1) f(\beta_0) f(\gamma_1) f(\gamma_0) f(\sigma^2)$$

We then can group these terms based on the parameters that they use

$$= \left[\left(\prod_{i:y_i=0} (1 - p_i) \prod_{i:y_i>0} p_i \right) f(\gamma_1) f(\gamma_0) \right] \left[\left(\prod_{i:y_i>0} f(y_i \mid \beta_1, \beta_0, \sigma^2) \right) f(\beta_1) f(\beta_0) f(\sigma^2) \right]$$

But now, if we look at this closely we can see that what we really have here is a full separation of the likelihoods for the individual models for p_i and y_i^* .

$$\begin{aligned}
f(\beta_1, \beta_0, \gamma_1, \gamma_0, \sigma^2 \mid y) &\propto \left[L(\gamma_1, \gamma_0 \mid y) f(\gamma_1) f(\gamma_0) \right] \cdot \left[L(\beta_1, \beta_0, \sigma \mid y) f(\beta_1) f(\beta_0) f(\sigma^2) \right] \\
&= \left[f(y \mid \gamma_1, \gamma_0) f(\gamma_1, \gamma_0) \right] \cdot \left[f(y \mid \beta_1, \beta_0, \sigma^2) f(\beta_1, \beta_0, \sigma^2) \right] \\
&= \left[f(\gamma_1, \gamma_0 \mid y) f(y) \right] \left[f(\beta_1, \beta_0, \sigma^2 \mid y) f(y) \right] \\
&\propto f(\gamma_1, \gamma_0 \mid y) f(\beta_1, \beta_0, \sigma^2 \mid y)
\end{aligned}$$

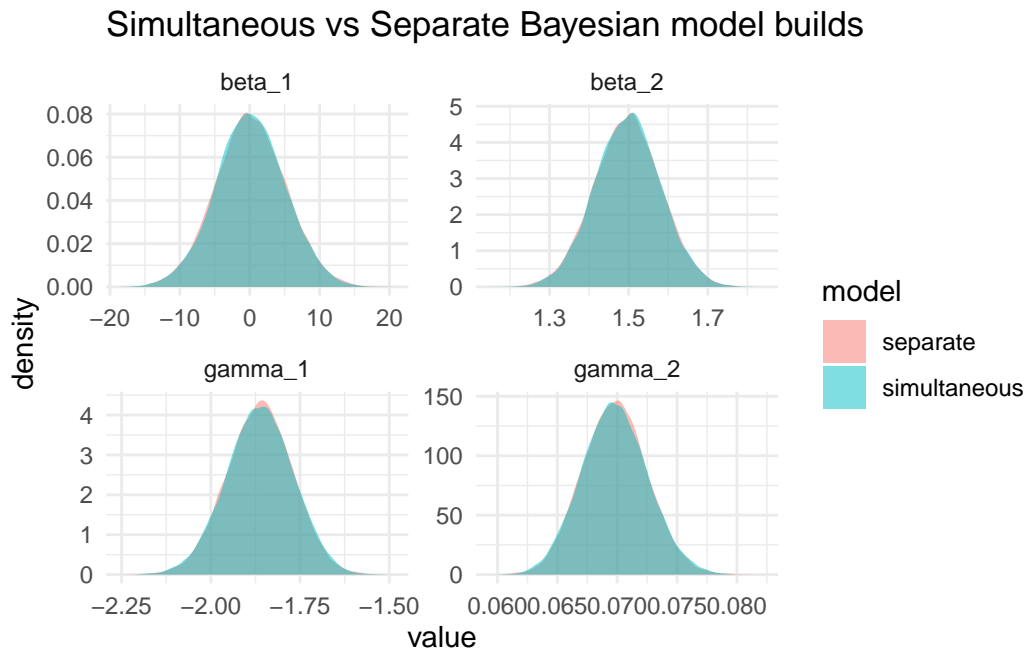
So we've just shown that the full posterior for the model built simultaneously is proportional to the posteriors for each model built separately. What this means for us is that we can fit the models separately, and then combine the results at the end to get our posterior predictive distributions.

And indeed, when we fit the models simultaneously and also fit them separately making use of MCMC to generate samples from their posteriors, we find that they are nearly identical

```

full_comp %>%
  ggplot(aes(x = value, fill = model)) +
  geom_density(alpha = 0.5, color = NA) +
  facet_wrap(~param, scales = "free") +
  labs(
    title = "Simultaneous vs Separate Bayesian model builds"
  ) +
  theme_minimal()

```



The major upshot here is that as long as we don't build in any correlations between the parameters in the two models, then we can build each model as complex as we might desire without having to worry about how we will eventually build the two models together. As we learned above, we can simply build them separately and combine the results at the end.