separate build

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

$$p_i = \frac{1}{1 + e^{-(x_i\gamma_1 + \gamma_0)}}$$

$$y_i^* = x_i\beta_1 + \beta_0 + \epsilon_i \qquad \text{where} \qquad \epsilon_i \sim \mathcal{N}(0, \sigma^2)$$

$$y_i = p_i \cdot y_i^*$$

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

$$f(\beta_1, \beta_0\gamma_1, \gamma_0, \sigma^2 \mid y) \propto f(y \mid \beta_1, \beta_0, \gamma_1, \gamma_0, \sigma^2) f(\beta_1, \beta_0, \gamma_1, \gamma_0, \sigma^2)$$

$$= L(\beta_1, \beta_0, \gamma_1, \gamma_0, \sigma^2 \mid y) f(\beta_1, \beta_0, \gamma_1, \gamma_0, \sigma^2)$$

$$= L(\beta_1, \beta_0, \gamma_1, \gamma_0, \sigma^2 \mid y) f(\beta_1) f(\beta_0) f(\gamma_1) f(\gamma_0) f(\sigma^2)$$

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

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

$$= \left[\prod_{i:y_i=0} \left(1 - \frac{1}{1 + e^{-\gamma x_i}}\right) \right] \left[\prod_{i:y_i>0} \left(\frac{1}{1 + e^{-\gamma x_i}}\right) \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_i - x_i\beta)^2}{2\sigma^2}} \right] f(\beta_1) f(\beta_0) f(\gamma_1) f(\gamma_0) f(\sigma$$

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

$$= \left[\prod_{i:y_i=0} \left(1 - \frac{1}{1 + e^{-\gamma x_i}}\right) \prod_{i:y_i>0} \left(\frac{1}{1 + e^{-\gamma x_i}}\right) \right] f(\gamma) \left[\prod_{i:y_i>0} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_i - x_i\beta)^2}{2\sigma^2}} \right] f(\beta) f(\sigma^2)$$

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

$$f(\beta, \gamma, \sigma^2 \mid y) \propto \left[L(\gamma \mid y)f(\gamma)\right] \cdot \left[L(\beta, \sigma \mid y)f(\beta)f(\sigma^2)\right]$$
$$= \left[f(y \mid \gamma)f(\gamma)\right] \cdot \left[f(y \mid \beta, \sigma^2)f(\beta, \sigma^2)\right]$$
$$= \left[f(\gamma \mid y)f(y)\right]\left[f(\beta, \sigma^2 \mid y)f(y)\right]$$
$$\propto f(\gamma \mid y)f(\beta, \sigma^2 \mid y)$$

So we've just shown that the full posterior for the model built simultaneously is proportionaly to the posteriors for each separate model. What this means for us is that we can fit the models separately.