Online Supplement B: Default Priors Rstan

Software packages have default prior settings that ensure estimation of the results even if the users do not choose priors themselves. It is important to realize that many of these default priors do not take the complexity of the hierarchical models into account (i.e., the hierarchical structure and the resulting number of parameters), therefore, it is *not* recommended to use the default priors for Bayesian hierarchical modeling. Here, we will provide details on the default priors for `rstan`, show predictions for data from these priors, and highlight resulting issues.

Rstan

When no prior is specified in `rstan` (Stan Development Team, 2019), a uniform distribution is applied using the defined range of the variable (Carpenter et al., 2017). The uniform distribution is shown in the figure below.
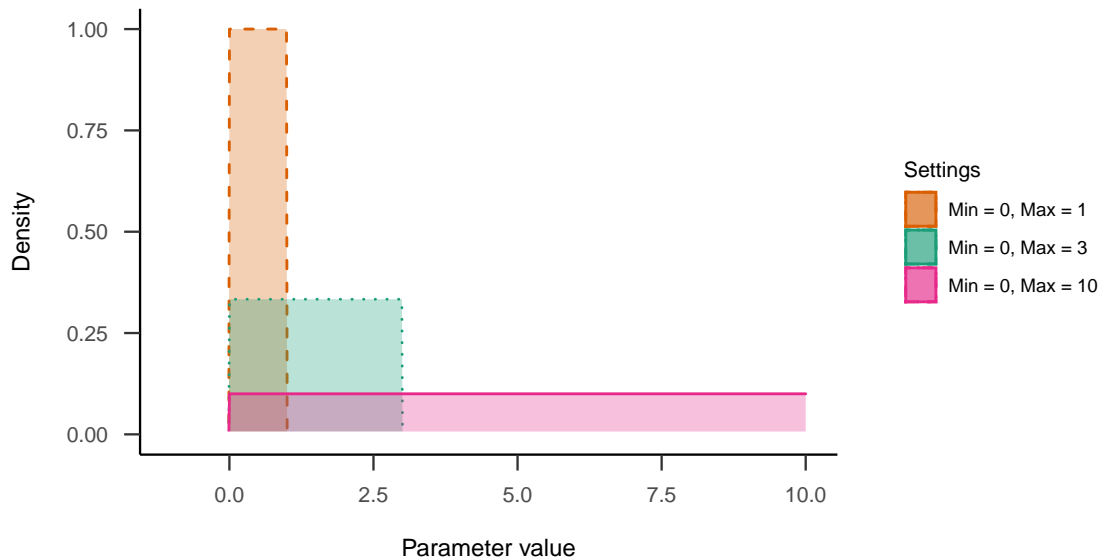

Figure B1: Different uniform distributions that vary in range (i.e., minimum and maximum).

The figure shows the uniform distribution for different settings of the minimum and maximum. The minimum and maximum define the range of the distribution where the minimum represents the lower bound and the maximum the upper bound. The figure illustrates that increasing the range, would result in wider distribution (i.e., the pink distribution is much wider than the green and orange distribution). In contrast, decreasing the range would result in a very specific prior with the probability spread out over a narrow range of values (i.e.,

the orange distribution is very narrow compared to the green and pink distribution).

However, if the range of the variable is not defined in `Stan`, the bounds will revert to $-\infty$ and ∞, resulting in an improper prior. This means that the probability distribution is not proper: integrating over a range of $-\infty$ to ∞ does not result in 1.

What would happen if these default priors were applied to the symbolic distance effect? In our scenario, we do not have a specified range for the parameters. Only the lower bound of the variances is specified as 0 (i.e., only positive variances are possible). Therefore, σ^2 and the hyperpriors can be defined in the following way:

$$
\begin{aligned}
\mu_\gamma, \mu_\beta, \mu_\delta, \sigma^2, \sigma_\gamma^2, \sigma_\beta^2, \sigma_\delta^2, &\sim \mathrm{U}(-\infty, \infty), \\
\sigma^2, \sigma_\gamma^2, \sigma_\beta^2, \sigma_\delta^2, &\sim \mathrm{U}(0, \infty).
\end{aligned}
\tag{1}
$$

All mean parameters have a flat prior, ranging from minus $-\infty$ to ∞, while the prior on the variances ranges from zero to ∞. Since at least one of the bounds (i.e., minimum or maximum) is not defined, the priors are improper. We cannot sample from $-\infty$ to ∞. As a result, prior visualization and prior prediction is impossible. However, `rstan` would still provide results as it does not identify improper priors (Stan Development Team, 2018, sec. 24.3).

References

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1).

Stan Development Team. (2018). Stan modeling language users guide, version 2.18.0. Retrieved from https://mc-stan.org/docs/2_18/stan-users-guide

Stan Development Team. (2019). RStan: The R interface to Stan. Retrieved from http://mc-stan.org/