Use truncated SBP with probit to do model selection for covariates and number of groups

* Do we select the optimal number of groups?
* Are the response curves still meaningful?
* Are spatial and temporal predictions good?

We adopt the following priors:

In relation to , we adopt a probit regression formulation of a truncated stick breaking prior:

We want to pick such that:

Which is a moment matching strategy, derived under the assumption that

We allow intercepts to vary for each l and k because:

1. ideally this model would revert back to the TSBP if all the slope coefficients were zero. This is sort of what happens because, when slopes are zero, then we have the identity:
2. If we don’t allow these intercepts to vary, then we run into the problem of assuming that the ’s are completely explained by the covariates.

This probit regression formulation enables us to automatically determine the optimal number of groups and create response curves for the individual species groups.

#----------------------------------------

How do we make this operational? Say we have four groups

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| 1 | >0 | NA | NA | NA |
| 2 | <0 | >0 | NA | NA |
| 3 | <0 | <0 | >0 | NA |
| 4 | <0 | <0 | <0 | >0 |

#------------------------------------------

FCD’s

* : I sample these guys from a truncated normal distribution

where is a matrix that assigns the corresponding means to their respective . Notice that only contains those elements that are not missing and contains the corresponding deltas

Therefore:

Notice that just counts the number of non-missing in each location l and group c and just sums the non-missing for each location. I need to check this