

# Chapter 1

## Other variable importance metrics in the Bayesian framework

Different variable importance metrics exist. The conditional variance formula allowed us to calculate the  $R^2$  of the submodels from the posterior sample of the full model. The focus of this master thesis was on the LMG formula. For each posterior sample the LMG formula can be applied for the submodels of each posterior sample. A lot of the variable importance metrics are based on the  $R^2$  of the full model compared to the submodels. Instead of the LMG formula we could as well have used another variable importance metric after we have calculated the  $R^2$  of all the submodels. Commonality and dominance analysis seem to be interesting extensions of the LMG framework. Both provide besides the LMG information some other information about the variance decomposition of the predictors. The relaimpo package provides some more bootstrap options like pariwise differences that could also be easily transfered to the Bayesian framework.

### 1.1 Conclusion

The Bayesian frameworks provides us with a nice option to include the uncertainty about parameters. Posterior distributions of the parameters allow us to calculate a distribution of  $R^2$  values for each model. Using the conditional variance formula allowd us to calculate the  $R^2$  of all the submodels from the posterior parameter distributions of the full model. Instead of fitting  $2^{p-1}$  models, only the full model needs to be fitted. Doing it this way has some nice properties. The interdependence of the submodels to each other is respected. The  $R^2$  of the submodels does not decrease when adding predictors. The non-negativity of the shares is therefore respected in the LMG formula. This property of using the conditional variance formula is also interesting when the LMG formula is applied to random intercept models fitted by maximum likelihood.

A disadvantage about calculating the  $R^2$  of all the submodels with the conditional variance formula seems to be the restriction to the linear model. Although this may be a topic of further research.

Assuming non-stochastic or stochastic predictors can have a big impact on the the uncertainty of the explained variance and the LMG values. Although the posterior regression parameter distribtuions is the same in both cases (under some assumptions described in chapter ...) the explained variance of a model is directly dependent on the covarinace matrix. Inference about

the covariance of the predictors  $\mathbf{X}$  is therefore an important part when stochastic predictors are assumed. However, this does in general not seem to be an easy problem. Non-parametric bootstrap provides a practical solution.