Non-informative Prior Distribution Does not Exist

a_statistician

April 7, 2017

For clarity, let the first order parameter be the parameter estimated by the Markov Chain Monte Carlo (MCMC) algorithm, and the second order parameter be the parameter used to specify the prior distribution of the first order parameter.

When we use the MCMC algorithm to estimate the first order parameters, the prior distributions of the first order parameters must be provided. If we have no information on the first order parameters (which is very common at present), we provide the so called non-informative prior distribution. But does non-informative prior distribution exist?

To specify the prior distribution, we need to not only point out the name of the distribution, such as uniform distribution, log-normal distribution, but also give the exact value(s) of the second order parameters of the distribution. When the distribution and its parameters are fixed, there is no ambiguity about the distribution of the first order parameters. It means when we specify the prior distribution, we admit that we have more or less the information about the first order parameters, instead of no information on the first order parameters.

For example, a stranger will practice shooting a target. I could not say anything about his rate of shooting on target, because I have no idea about this person. But it implies that his mean rate is 0.5 if I assume his rate follows standard uniform distribution. Therefore, having no information on his rate and his rate following standard uniform distribution are not the same.

It is common to specify the standard uniform distribution as prior for event rate in binomial distribution. If you really have no information on event rate, it would not be possible to give out the mean of the event rate. But the standard uniform distribution as prior implies that you have an idea that mean event rate is 0.5. Therefore, the commonly used standard uniform distribution for the rate of binomial distribution is not the non-informative distribution. This reasoning applies to any kind of so called non-informative distribution. Having no information means you cannot say anything about the first order parameters, but from any prior distribution of the first order parameter it is easy to get the mean of the first order parameter or say that the mean does not exist. Therefore, the real non-informative prior distribution does not exist.

The MCMC algorithm developed from Bayesian statistics which requires the prior distribution of the parameters. The Bayesian statistics combines the information from old prior distribution and the new data to get the posterior distribution of the parameters, and the estimate of the parameters is generated from the posterior distribution. For the MCMC algorithm, the prior distribution of all parameters is needed to start the algorithm.

The name "non-informative prior" misleads us to think that non-informative prior plus no informative data will give us no answer. But the fake "non-informative prior" makes the MCMC give us the fake answer. When the non-informative prior is provided to MCMC, we expect that answer is totally based on the data. But in fact, this is not true. When the data contain numerous information on first order parameters, the influence of non-informative prior is ignorable. When the information in the data is tiny, the posterior distribution of the first order parameters may be dominated by fake non-informative prior. Meta-analysis is an example. In meta-analysis, the study effect is treated as random. It is common very few studies can be used in the one meta-analysis. Therefore, the information for estimating parameters related to this random effect is very limited and the estimate is dominated by the prior. When the prior is changed, the estimate is also changed.

In conclusion, any prior distribution brings some information into MCMC. When information in the data is limited, we need to pay special attention to the influence of prior distribution on the estimates of first order parameters from MCMC.

Previous Topics:

Estimating Variance Using MCMC When Sample Size Is One
Interaction of Covariates in Statistics
Scope of the Appliaction for Survival Analysis
Log Transformation in Simple Linear Regression (3/3)
Log Transformation in Simple Linear Regression (2/3)
Log Transformation in Simple Linear Regression (1/3)