

MC3：多核马尔科夫链蒙特卡洛

|  |  |
| --- | --- |
| X.data | 要输入的数据集。在输入数据时，第一列默认是因变量，后面的是协变量。 |
| burn | 在使用MC3取样时，从第1轮开始，要放弃的迭代次数。 |
| iter | 在使用MC3取样时，总的迭代次数。Iter-burn就是最后留下的迭代。 |
| nmodel | 存储信息的最佳模型的数量(默认为500)。最佳模型用于似然和MCMC频率之间的收敛分析，以及基于似然的推理。 |
| mcmc | 表示要使用的模型采样器的类型。就是马尔科夫链蒙特卡洛模拟的不同逻辑。 |
| g | 参数的先验设定。用超参数 来衡量， 代表对于参数为0的确定程度，当 的值越大则表示对于系数值非0的把握也就越大。 |
| mprior | 模型先验分布的设定。一致先验分布（Uniform），固定先验分布（Fixed）和随机先验分布（Random）等。 |
| mprior.size | 模型的先验期望值。一般貌似不太用特别设定。 |
| user.int | 在进行完本句计算后，是否输出模型计算结果。（有关各个变量的后验概率等一系列因素的那个大表格） |
| start.value | 指定迭代链的起始模型。start. model=numeric(K)从空模型开始，只包括一个常量项。 start.value=c(3,6)对于一个只包括协变量3和6的起始模型。 |
| g.stats | 是否要统计g。设置g.stats=FALSE以加快迭代。 |
| logfile | 在执行工作过程中，生成一个名为“test.log”的日志文件，以 |
| logstep | 指定将后验图信息写入日志文件的次数 |
| force.full.ols | =TRUE会减慢采样速度，但可以更好地处理高度共线的数据 |
| fixed.reg | x数据的索引或变量名，这些数据是固定的回归变量，总是包含在每个抽样模型中。备注:参数mprior.size规模指的是包括这些固定回归量在内的先验模型规模。 |

>data(dataset)：加载数据集dataset。

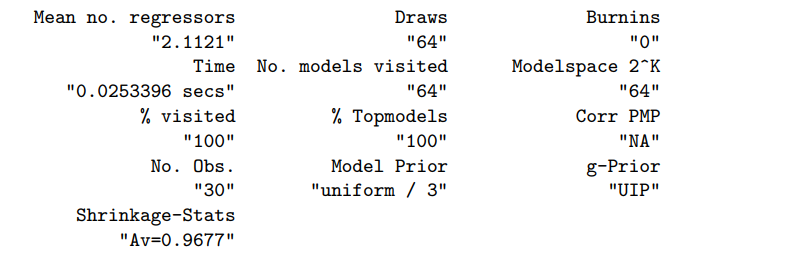
> library(BMS)：运行第三方数据包

> test = bms(dataset, mprior = "uniform", g="UIP", user.int=F)：文章的第一种计算。没有输出。

> coef(test)：输出计算的模型集test的各个协变量的参数相应计算结果，得出原文中的回归结果。

\*\* coef(test, std.coefs=T, order.by.pip=F, include.constant=T)也可以对coef输出结果做一系列的改动。

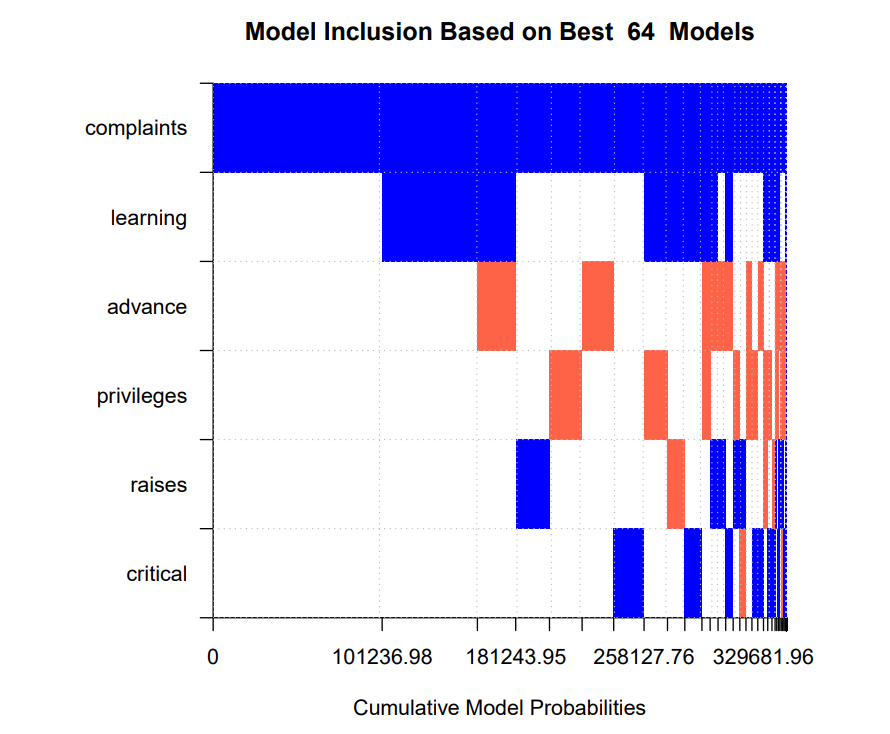
> summary(test)



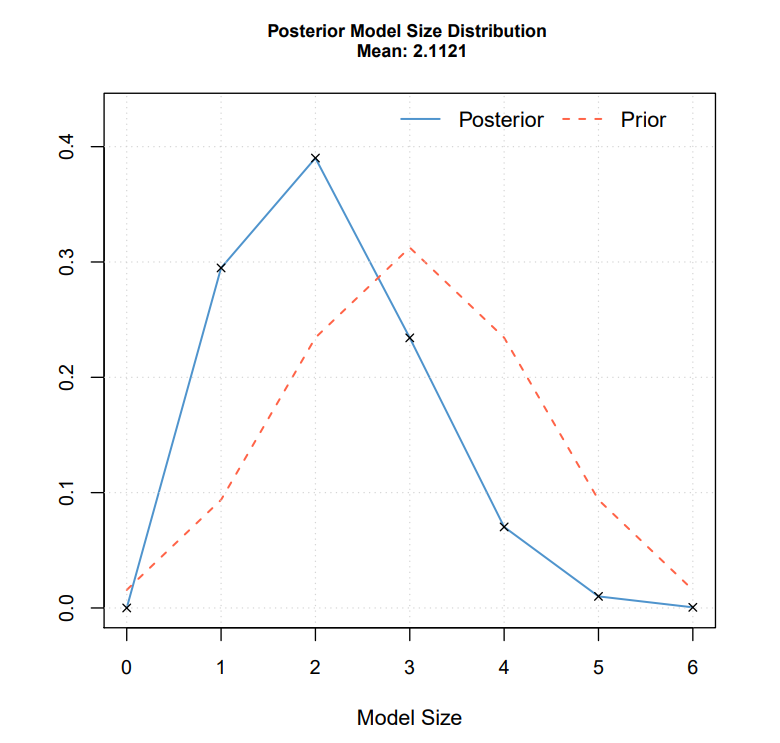
可得抽样程序的其他一些信息。

> topmodels.bma(test)[,1:3]：输出性能最好的三个模型的相关信息，包括模型包含哪些协变量等。

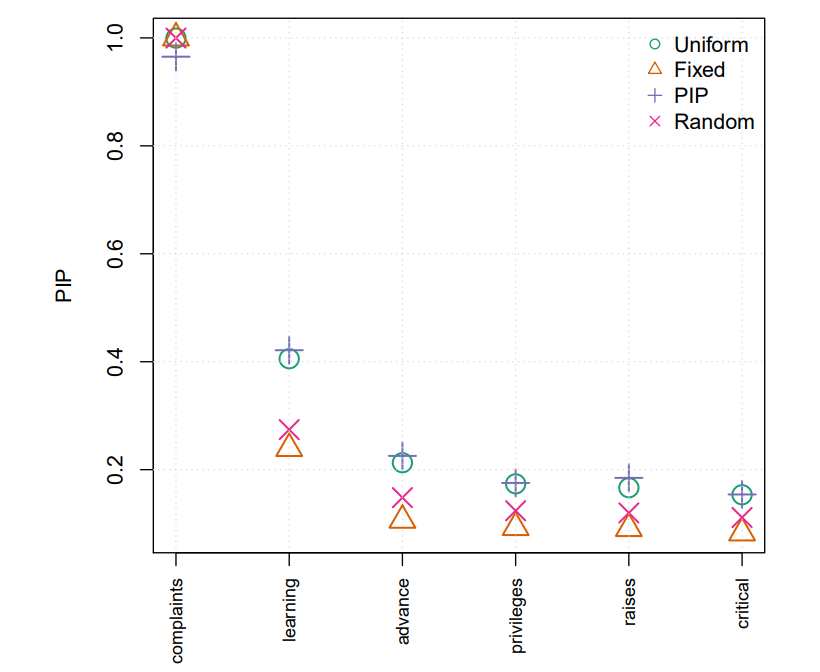
> image(test)：所有的模型后验概率由大到小排列，不同模型包含的变量，以及其变量的符号正负。得出来的是下面这种的图。



> plotModelsize(test)：比较先验和后验模型的模型大小（包含几个协变量）



> plotComp(Uniform=test, Fixed=test\_fixed, PIP=test\_pip, Random=test\_random)：更改mprior，看模型是否稳定的协变量比较，和文章里的图片一样。



补充材料：

1. [什么是马尔可夫链蒙特卡洛（MCMC）？ - 知乎 (zhihu.com)](https://zhuanlan.zhihu.com/p/116725922)
2. [R package BMS - Bayesian Model Averaging (zeugner.eu)](http://bms.zeugner.eu/)
3. [towardsdatascience.com/a-zero-math-introduction-to-markov-chain-monte-carlo-methods-dcba889e0c50](https://link.zhihu.com/?target=https%3A//towardsdatascience.com/a-zero-math-introduction-to-markov-chain-monte-carlo-methods-dcba889e0c50)