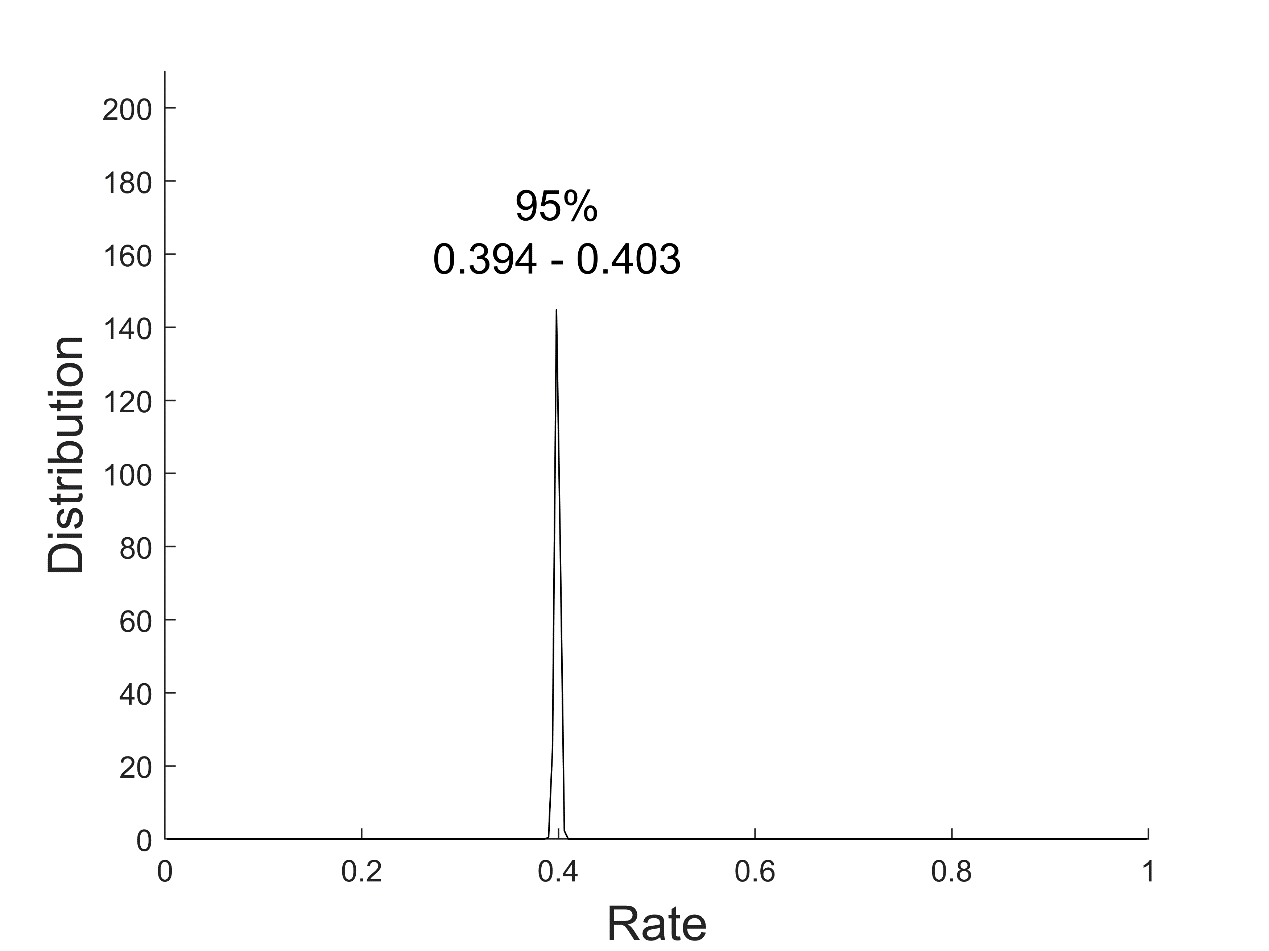
题1

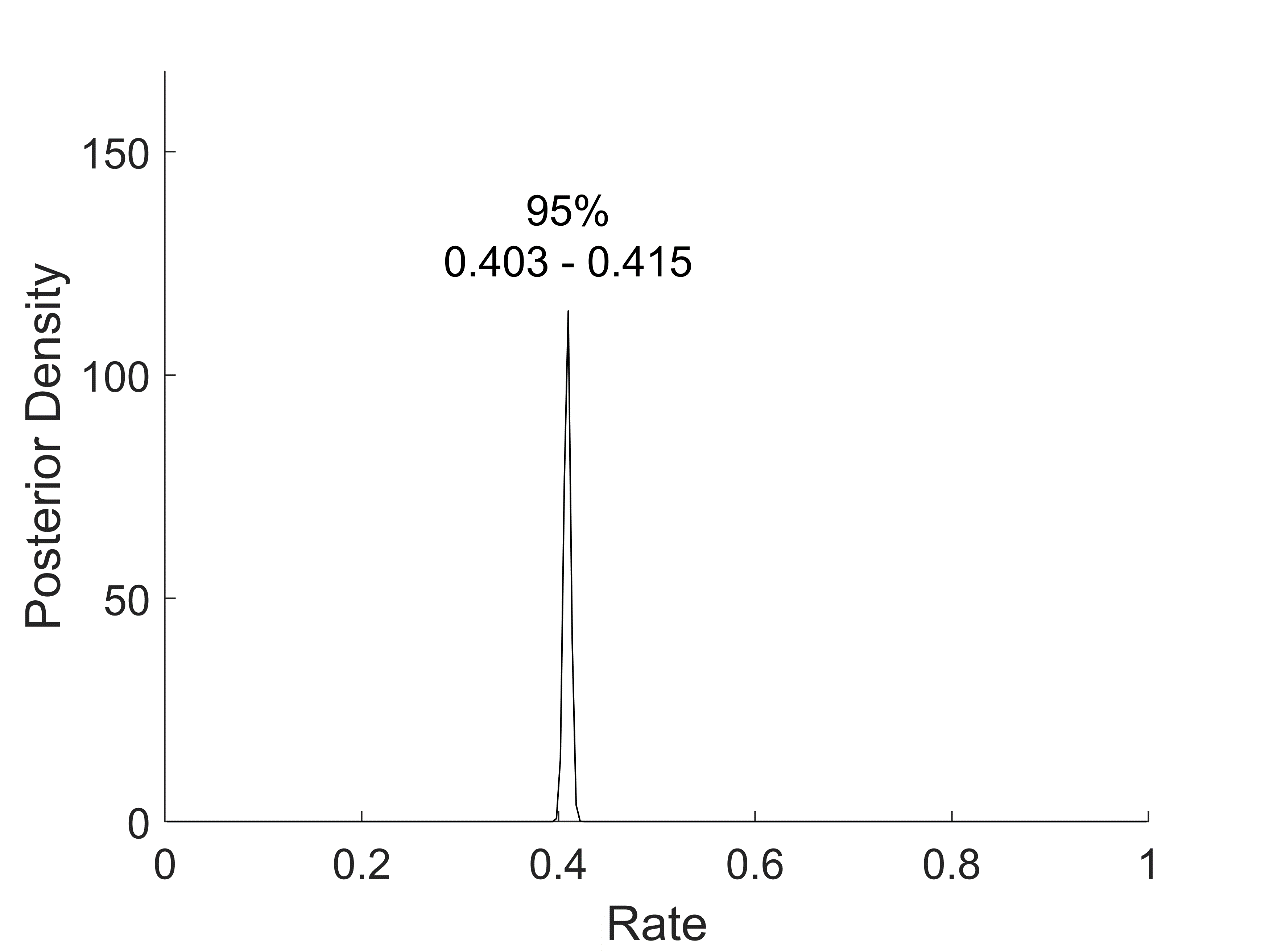
1)

θ的MLE估计为0.399，由parametric bootstrap(重采样10000次)得到的95% Confidence Interval为[0.394, 0.403]，见图1-1.



**图1-1 由MLE及parametric bootstrap得到的ChaSaSoon的的分布**

θ的MAP估计为0.410，95% Credible Interval为[0.403,0.415]，见图1-2.



**图1-2 由Bayesian Inference得到的ChaSaSoon的θ的后验分布**

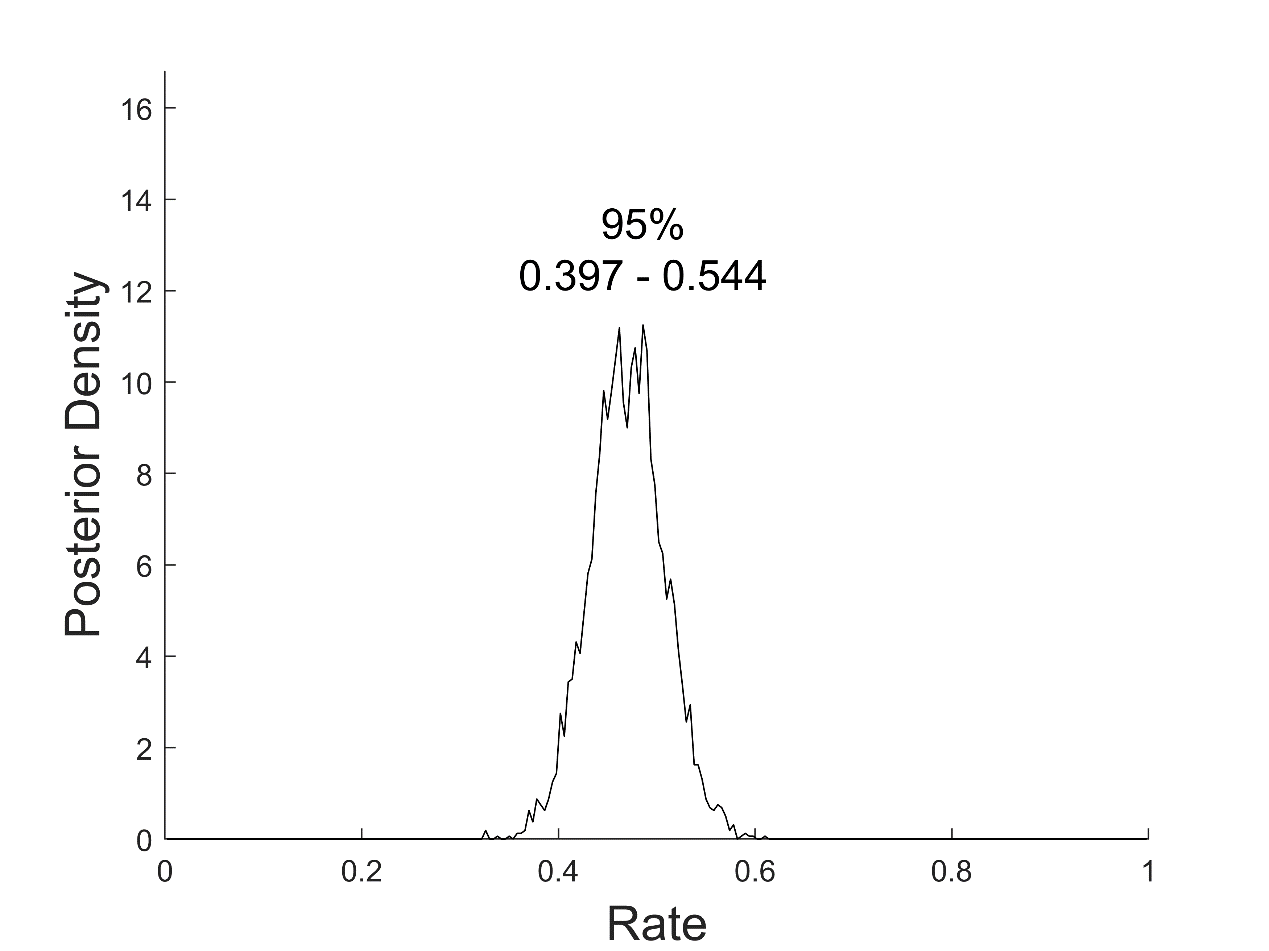
贝叶斯推理和最大似然估计的结果基本类似。事实上，由于前949次几乎都位于15至25之间，显然结果在0.4附近。但事实上，贝叶斯推理和最大似然估计还是有些微的不同。

由于4选1的选择题中我们先验地知道θ位于0.25至1，且假设为均匀分布，故在贝叶斯推理中加入了这一因素，而最大似然估计中θ没此影响，可能造成了分布上的差异。

然而由于Prior是0.25到1上的均匀分布，本不应该造成最大后验估计和最大似然估计的区别。由于原结果是基于JAGS实现的，重新用BUGS而非JAGS运行了教材中的程序后发现，可能由于随机取样的原因，BUGS下θ的后验分布整体比JAGS下θ的后验分布向左平移，和最大似然估计无二致。

2)

Grandpa’s θ的MAP估计为0.486，95% Credible Interval为[0.397,0.544]，见图1-3



**图1-3 由Bayesian Inference得到的Grandpa的θ的后验分布**

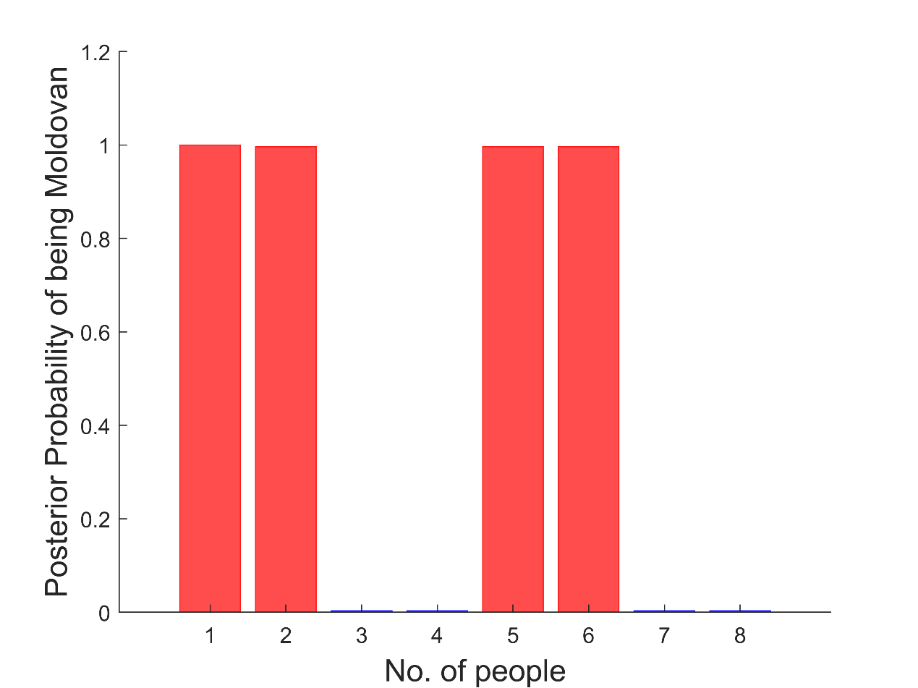
比较ChaSaSoon的θ的后验分布和Grandpa的θ的后验分布。两者的主要差别在于ChaSaSoon的θ的后验分布的众数为0.410，而Grandpa的θ的后验分布的众数为0.486。这是由于尽管两者的先验分布和模型相同，但两者的数据不同。两者都只有1次非删失的30的数据点，但前者比后者有更多的在15-25的删失数据点。故而后者更加强烈地受到了非删失数据的影响。此外，后者分布的方差要比前者分布的方差大很多，这同样是因为前者的数据尽管删失但大量集中于15-25的区间中。

题2

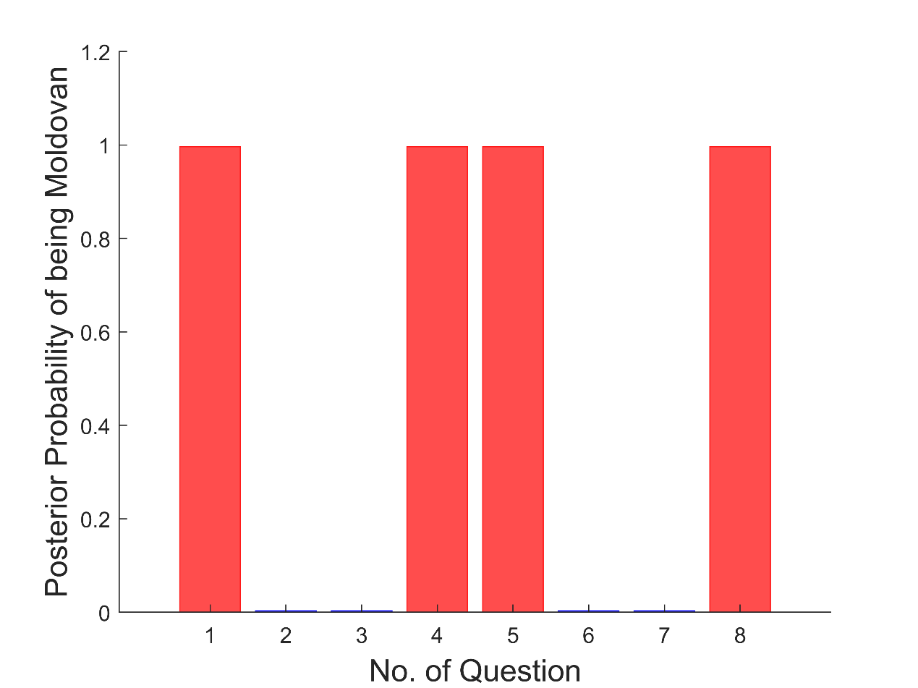
1)

1,2,5,6号人为Moldovan人的概率为1.00，3,4,7,8号人为Moldovan人的概率为0.00，见图2-1.

1,4,5,8号题为关于Moldovan的历史题的概率为1.00，2,3,6,7号题为关于Moldovan的历史题的概率为0.00，见图2-2.



**图2-1 由Bayesian Inference得到的是Moldovan人的后验分布**



**图2-2 由Bayesian Inference得到的是关于Moldovan的历史题的后验分布**

题3

1)

**xi**

**zj**

**kij**

**i people**

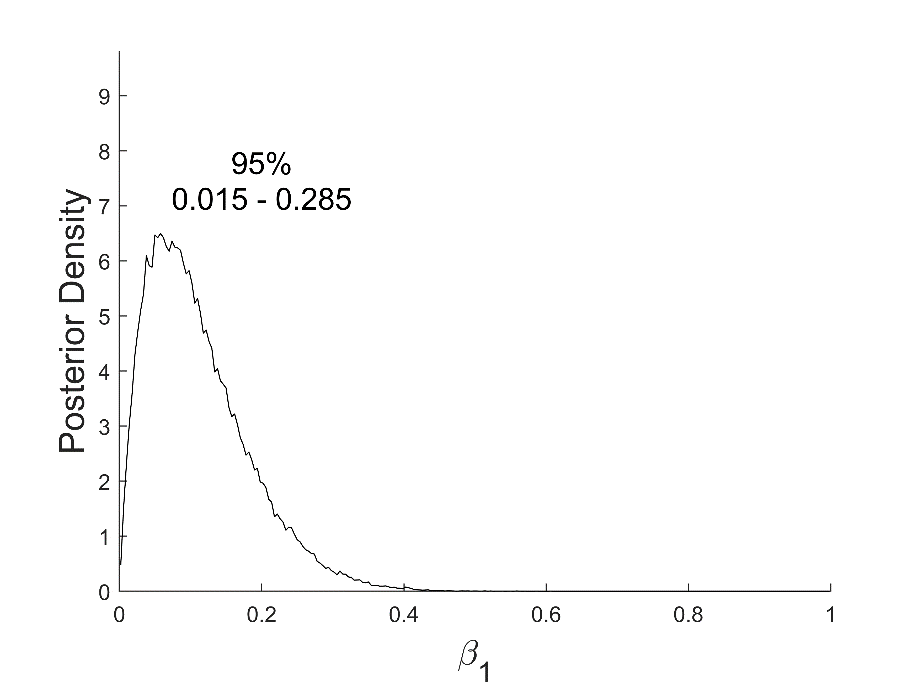
**j questions**

由于本模型重点在于考察参数β1，β2。为了解决模型不确定性的问题，不失一般性，设x1=0，从而上述xi仅对2至8成立。值得注意的是，这种讨论下0或1所代表的国家是不确定的，我们也并不关心。

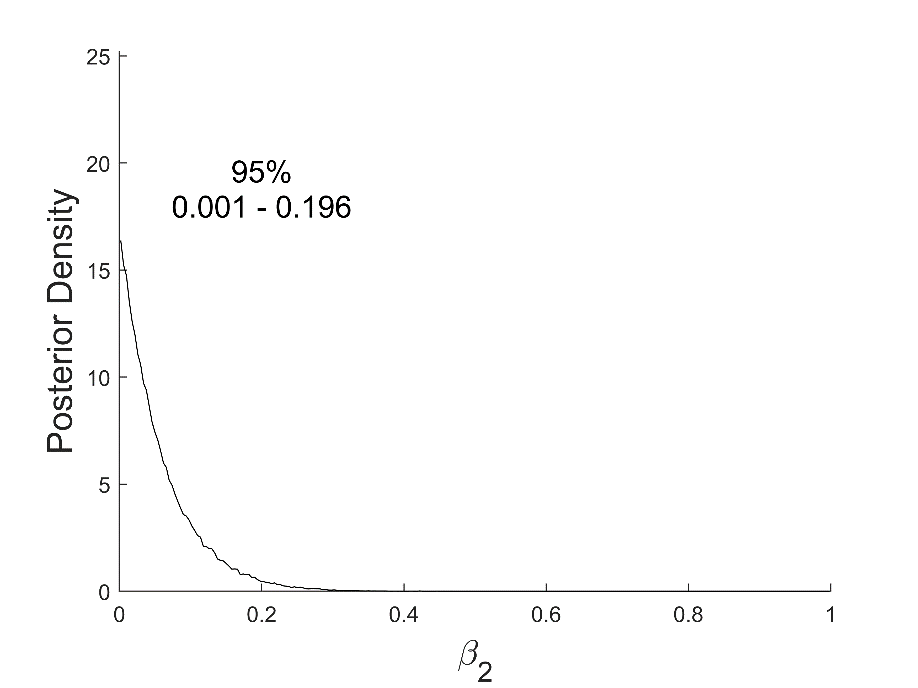
2)3)

β1的95% Credible Interval为[0.015,0.285]，见图3-1

β2的95% Credible Interval为[0.001,0.196]，见图3-2



**图3-1 由Bayesian Inference得到的β1的后验分布**



**图3-2 由Bayesian Inference得到的β2的后验分布**

4)

model 1(β1=β2)’s DIC=33.44

model 2(β1≠β2)’s DIC=35.52

因此，相比于model 2，model 1可能更好。

原因在于加入了β1(β2)后并没有十分明显地提高模型的预测力，缺徒增了模型的复杂性