

## 题 1

1)

$\theta$  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.

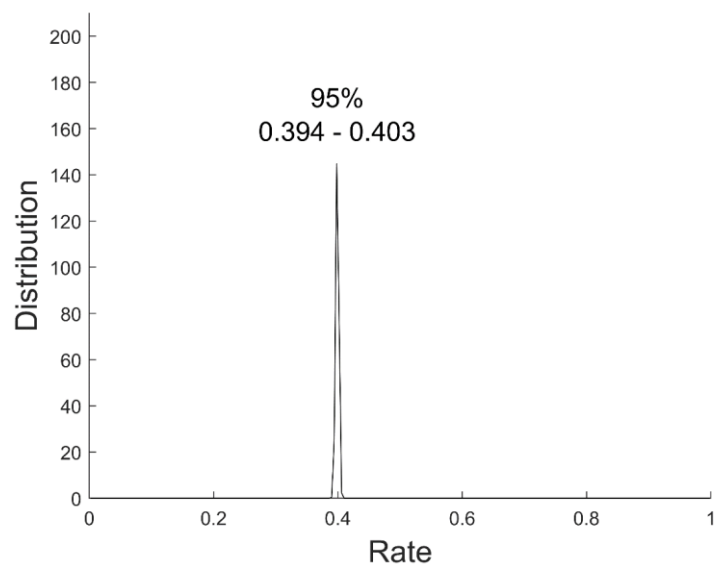


图 1-1 由 MLE 及 parametric bootstrap 得到的 ChaSaSoon 的  $\hat{\theta}$  的分布

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

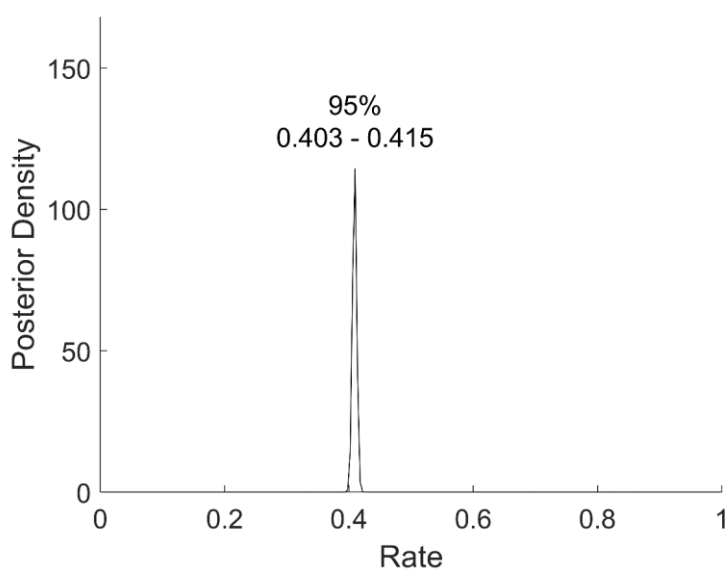


图 1-2 由 Bayesian Inference 得到的 ChaSaSoon 的  $\theta$  的后验分布

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。

$$\text{Posterior} \propto \text{Prior} * \text{Likelihood}$$

由于 4 选 1 的选择题中我们先验地知道  $\theta$  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  $\theta$  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  $\theta$  的后验分布整体比 JAGS 下  $\theta$  的后验分布向左平移，和最大似然估计无二致。

2)

Grandpa's  $\theta$  的 MAP 估计为 0.486，95% Credible Interval 为 [0.397, 0.544]，见图 1-3

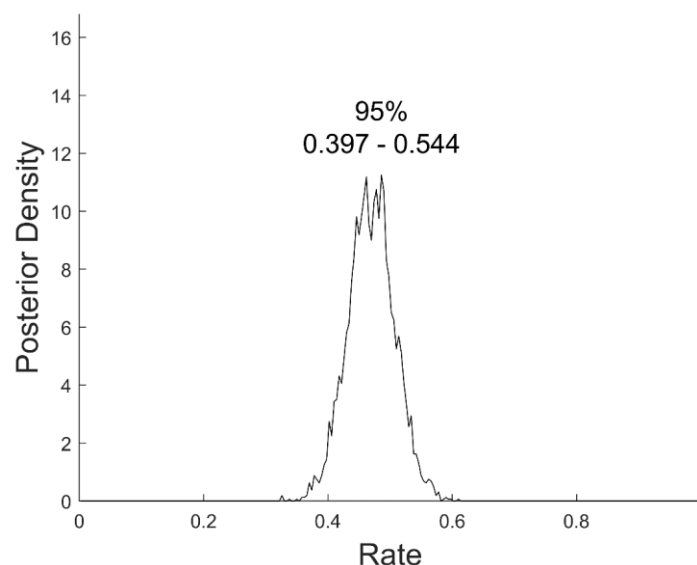


图 1-3 由 Bayesian Inference 得到的 Grandpa 的  $\theta$  的后验分布

比较 ChaSaSoon 的  $\theta$  的后验分布和 Grandpa 的  $\theta$  的后验分布。两者的主要差别在于 ChaSaSoon 的  $\theta$  的后验分布的众数为 0.410，而 Grandpa 的  $\theta$  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.

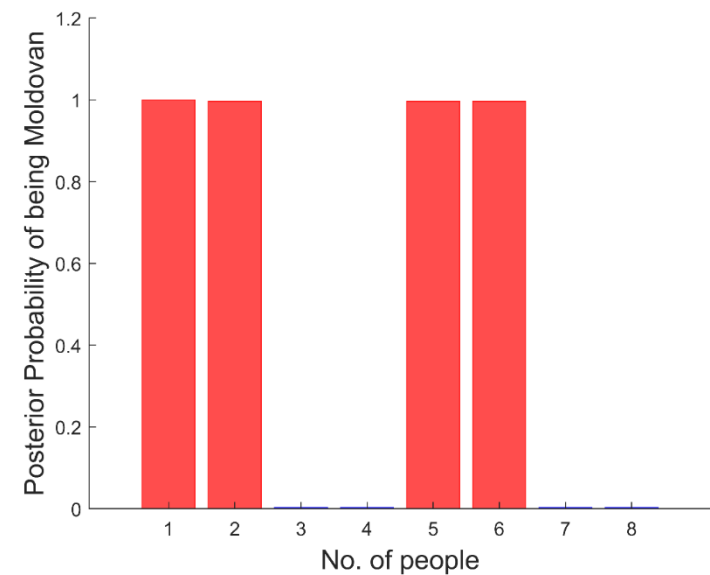


图 2-1 由 Bayesian Inference 得到的是 Moldovan 人的后验分布

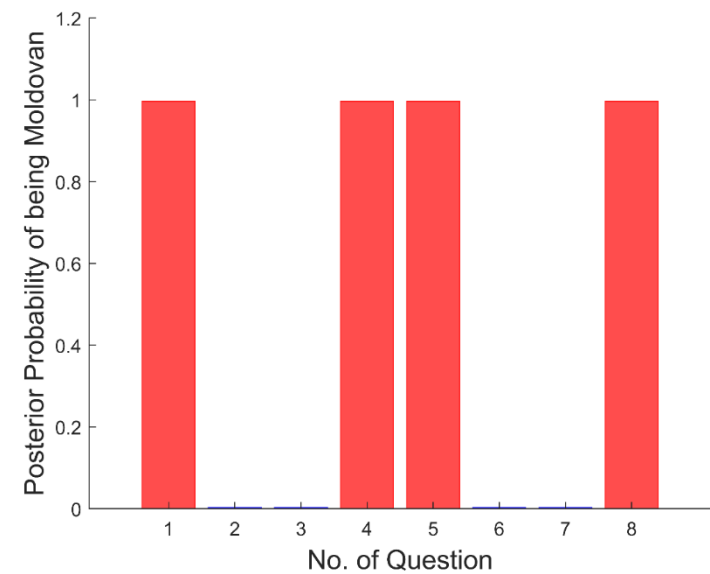
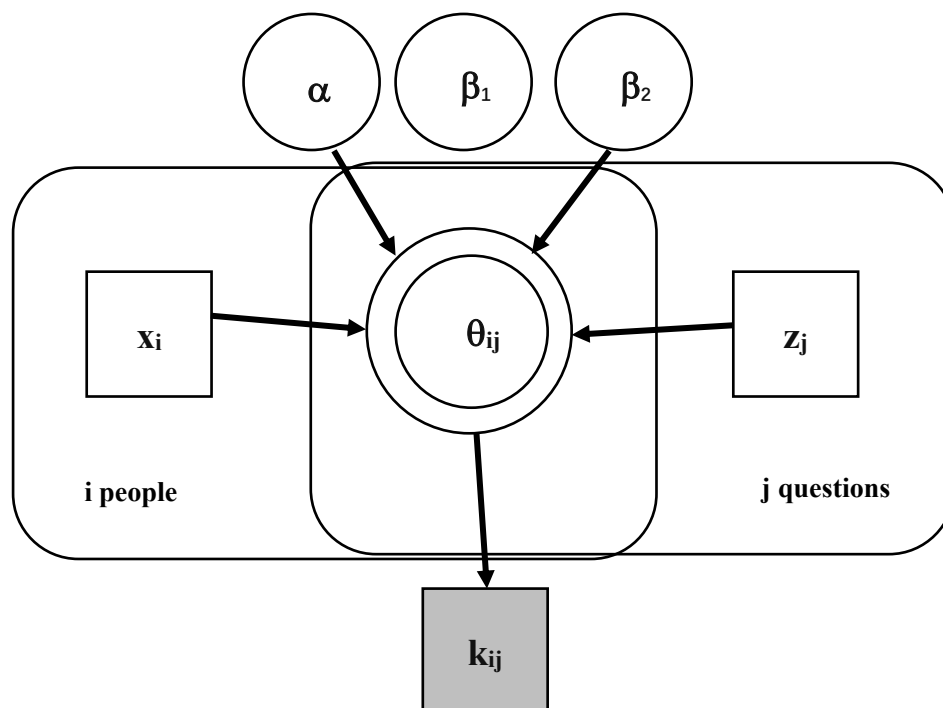


图 2-2 由 Bayesian Inference 得到的是关于 Moldovan 的历史题的后验分布

题 3

1)



$$\begin{aligned}
 \alpha &\sim \text{Uniform}(0,1) \\
 \beta_1 &\sim \text{Uniform}(0,\alpha) \\
 \beta_2 &\sim \text{Uniform}(0,\alpha) \\
 x_i &\sim \text{Bernoulli}(0.5), \quad i = 1:8 \\
 z_j &\sim \text{Bernoulli}(0.5), \quad j = 1:8 \\
 \theta_{ij} &= \begin{cases} \alpha & \text{if } x_i = z_j \\ \beta_1 & \text{if } x_i = 0 \text{ and } z_j = 1 \\ \beta_2 & \text{if } x_i = 1 \text{ and } z_j = 0 \end{cases} \\
 k_{ij} &\sim \text{Bernoulli}(\theta_{ij})
 \end{aligned}$$

由于本模型重点在于考察参数 $\beta_1, \beta_2$ 。为了解决模型不确定性的问题，不失一般性，设 $x_1=0$ ，从而上述 $x_i$ 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)

$\beta_1$  的 95% Credible Interval 为[0.015,0.285], 见图 3-1

$\beta_2$  的 95% Credible Interval 为[0.001,0.196], 见图 3-2

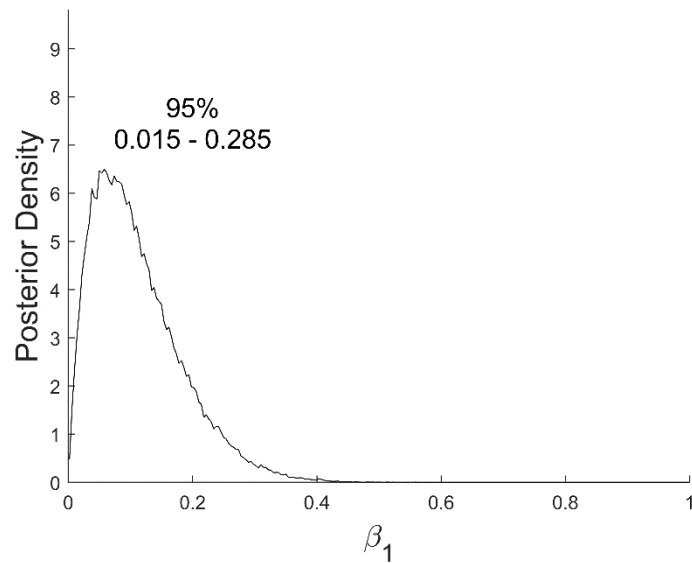


图 3-1 由 Bayesian Inference 得到的 $\beta_1$ 的后验分布

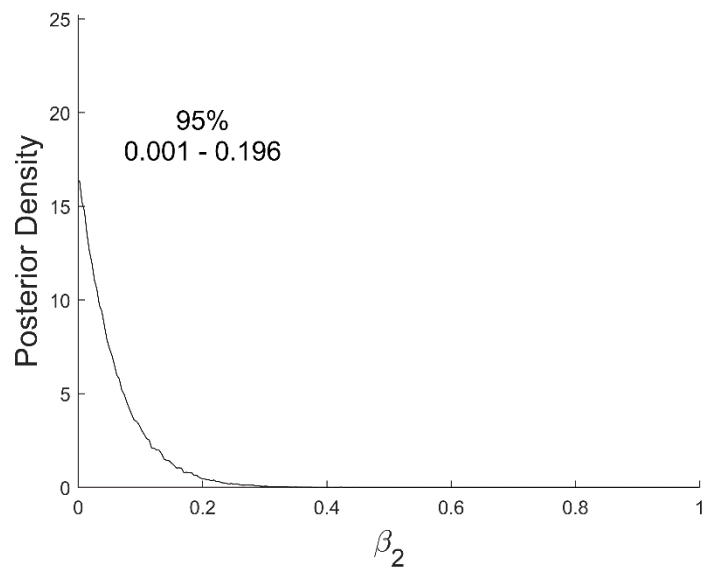


图 3-2 由 Bayesian Inference 得到的 $\beta_2$ 的后验分布

4)

model 1( $\beta_1=\beta_2$ )'s DIC=33.44

model 2( $\beta_1\neq\beta_2$ )'s DIC=35.52

因此, 相比于 model 2, model 1 可能更好。

原因在于加入了 $\beta_1(\beta_2)$ 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