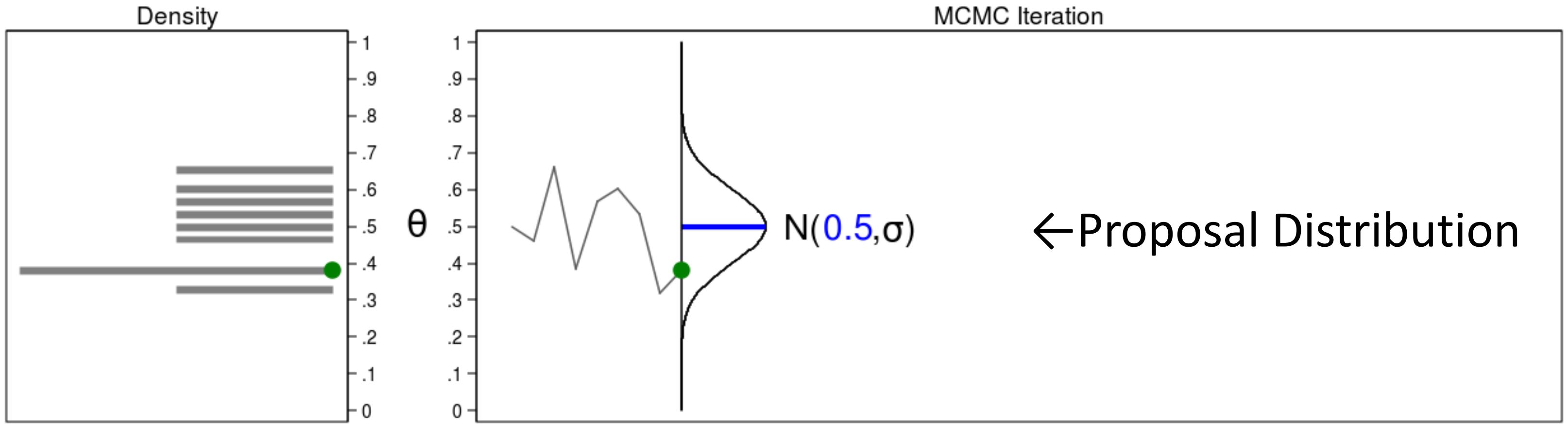
Markov Chain Monte Carlo

Often the posterior distribution does not have a simple form. We can use Markov Chain Monte Carlo (MCMC) with the Metropolis-Hastings algorithm to generate a sample from the posterior distribution.

MCMC and Metropolis-Hastings

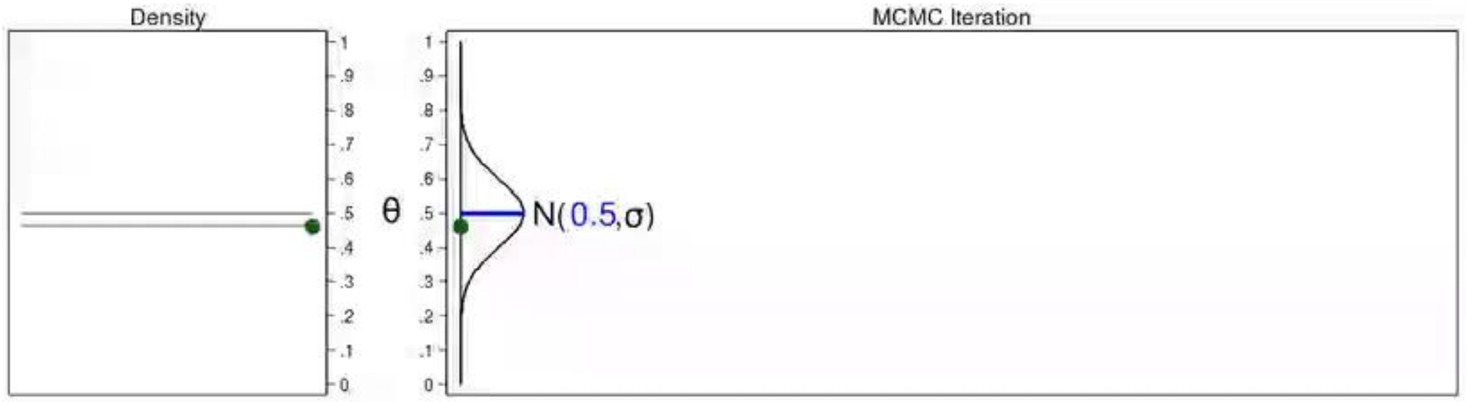
1. Monte Carlo
2. Markov Chains
3. Metropolis-Hastings

Monte Carlo



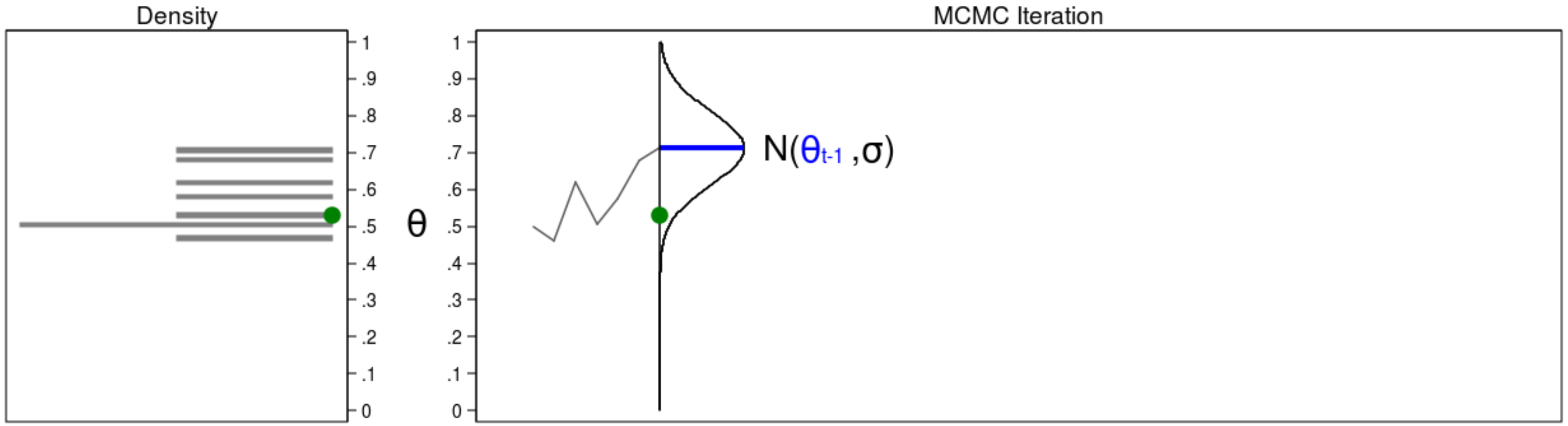
Draw $\theta_t \sim \text{Normal}(0.5, \sigma) = 0.381$

Monte Carlo



Draw $\theta_t \sim \text{Normal}(0.5, \sigma) = 0.460$

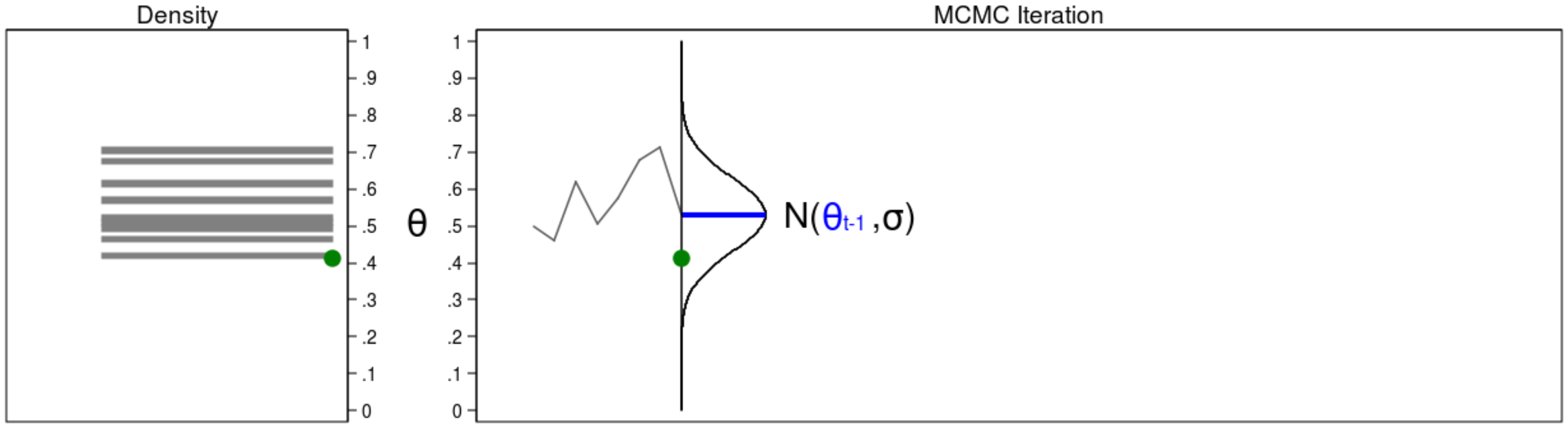
Markov Chain Monte Carlo



Draw $\theta_t \sim \text{Normal}(\theta_{t-1},\sigma)$

$\text{Normal}(0.712,\sigma) = 0.530$

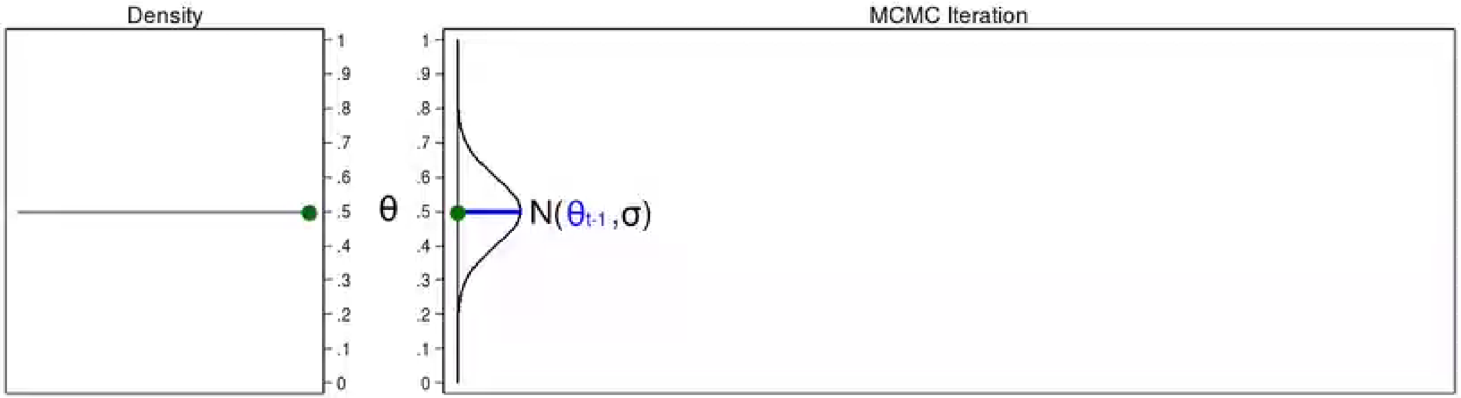
Markov Chain Monte Carlo



Draw $\theta_t \sim \text{Normal}(\theta_{t-1},\sigma)$

$\text{Normal}(0.530,\sigma) = 0.411$

Markov Chain Monte Carlo



Draw $\theta_t \sim \text{Normal}(\theta_{t-1}, \sigma)$

$\text{Normal}(0.500, \sigma) = 0.497$

MCMC with Metropolis-Hastings

$$r(\theta_{new}\,, \theta_{t-1}) = \frac{Posterior\ probability\ of\ \theta_{new}}{Posterior\ probability\ of\ \theta_{t-1}}$$

$$= \frac{Beta(1,1,\theta_{new})\ \times\ Binomial(10,4,\theta_{new})}{Beta(1,1,\theta_{t-1})\ \times\ Binomial(10,4,\theta_{t-1})}$$

MCMC with Metropolis-Hastings

$$acceptance\ probability = \alpha(\theta_{new}\ , \theta_{t-1})$$

$$= min[\ r(\theta_{new}, \theta_{t-1})\ , 1\ ]$$

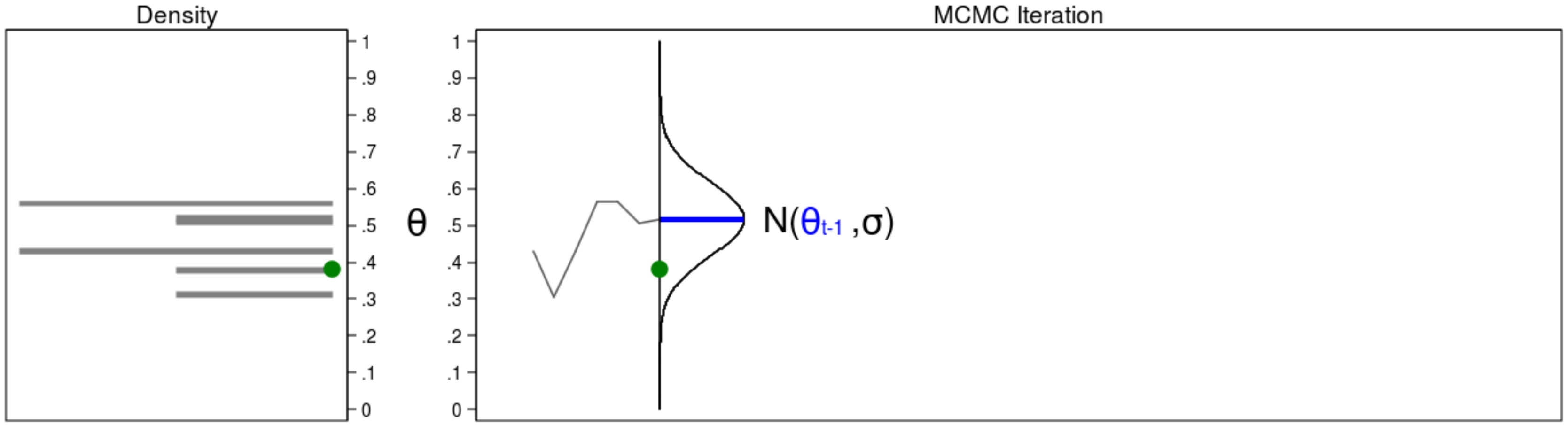
MCMC with Metropolis-Hastings

$$Draw\ u \sim Uniform(0,1)$$

$$If \quad u < \alpha(\theta_{new}, \theta_{t-1}) \qquad Then\ \theta_t = \theta_{new}$$

$$Otherwise\ \theta_t = \theta_{t-1}$$

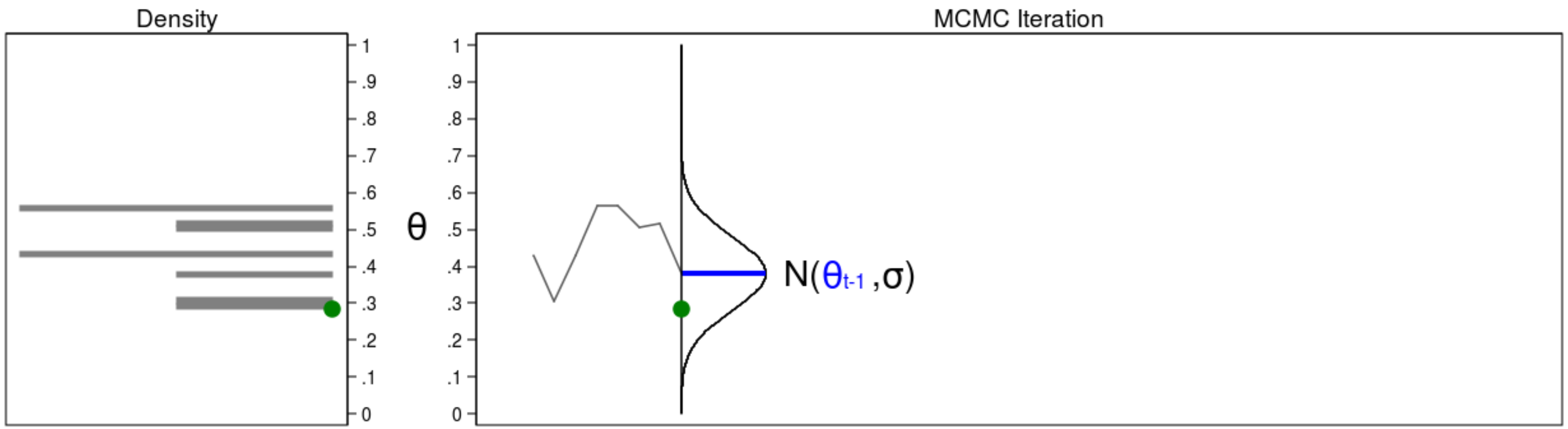
MCMC with Metropolis-Hastings



Step 1: $r(\theta_{new}, \theta_{t-1}) = \dfrac{\text{Posterior}(\theta_{new})}{\text{Posterior}(\theta_{t-1})} = \dfrac{\text{Beta}(1,1,0.380) \times \text{Binomial}(10,4,0.380)}{\text{Beta}(1,1,0.517) \times \text{Binomial}(10,4,0.517)} = 1.307$

Step 2: Acceptance probability $\alpha(\theta_{new}, \theta_{t-1}) = \min\{r(\theta_{new}, \theta_{t-1}), 1\} = \min\{1.000, 1\} = 1.000$

MCMC with Metropolis-Hastings



Step 1: $r(\theta_{new}, \theta_{t-1}) = \dfrac{\text{Posterior}(\theta_{new})}{\text{Posterior}(\theta_{t-1})} = \dfrac{\text{Beta}(1,1,0.286) \times \text{Binomial}(10,4,\ 0.286)}{\text{Beta}(1,1,0.380) \times \text{Binomial}(10,4,\ 0.380)} = 0.747$

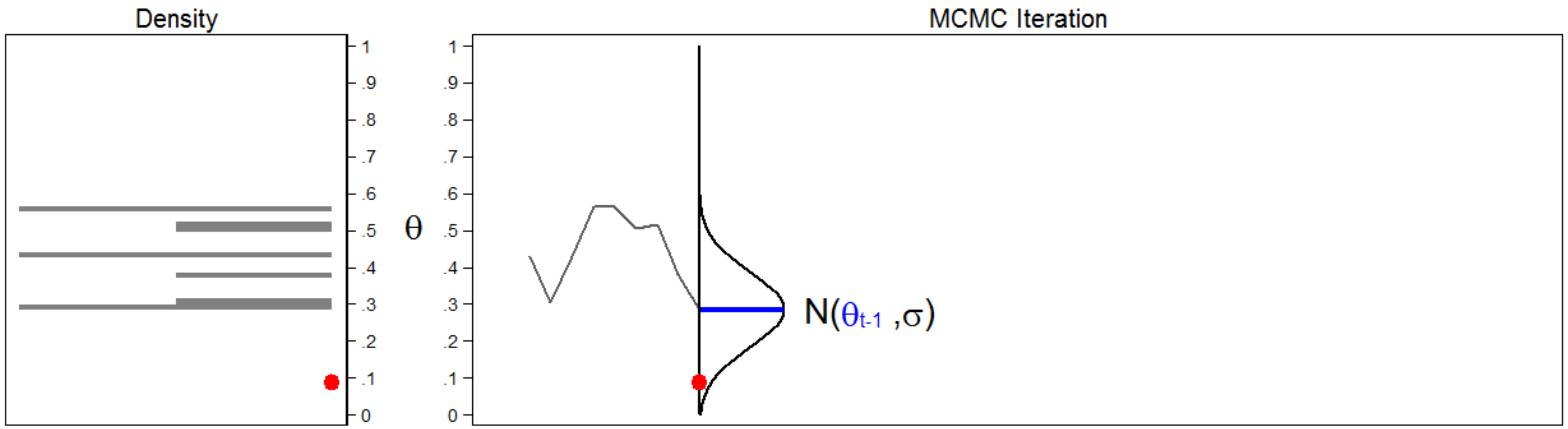
Step 2: Acceptance probability $\alpha(\theta_{new}, \theta_{t-1}) = \min\{r(\theta_{new}, \theta_{t-1}), 1\} = \min\{0.747, 1\} = 0.747$

Step 3: Draw $u \sim \text{Uniform}(0,1) = 0.094$

Step 4: If $u < \alpha(\theta_{new}, \theta_{t-1}) \rightarrow$ If $0.094 < 0.747$ Then $\theta_t = \theta_{new} = 0.286$

Otherwise $\theta_t = \theta_{t-1} = 0.380$

MCMC with Metropolis-Hastings



Step 1: $r(\theta_{new}, \theta_{t-1}) = \dfrac{\text{Posterior}(\theta_{new})}{\text{Posterior}(\theta_{t-1})} = \dfrac{\text{Beta}(1,1,0.088) \times \text{Binomial}(10,4,0.088)}{\text{Beta}(1,1,0.286) \times \text{Binomial}(10,4,0.286)} = 0.039$

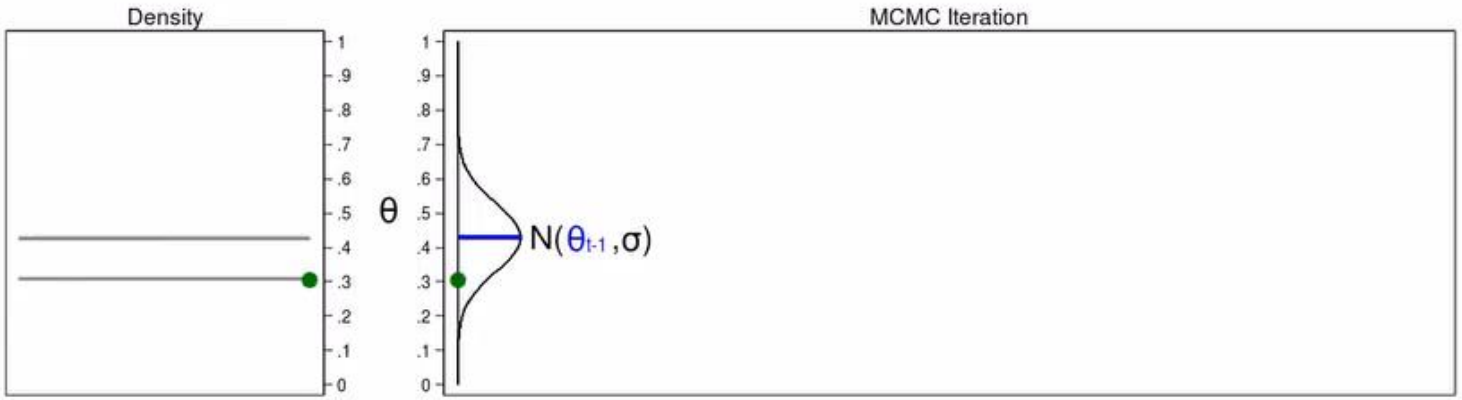
Step 2: Acceptance probability $\alpha(\theta_{new}, \theta_{t-1}) = \min\{r(\theta_{new}, \theta_{t-1}), 1\} = \min\{0.039, 1\} = 0.039$

Step 3: Draw $u \sim \text{Uniform}(0,1) = 0.247$

Step 4: If $u < \alpha(\theta_{new}, \theta_{t-1}) \rightarrow$ If $0.247 < 0.039$ Then $\theta_t = \theta_{new} = 0.088$

Otherwise $\theta_t = \theta_{t-1} = 0.286$

MCMC with Metropolis-Hastings



Step 1: $r(\theta_{new}, \theta_{t\text{-}1}) = \dfrac{\text{Posterior}(\theta_{new})}{\text{Posterior}(\theta_{t\text{-}1})} = \dfrac{\text{Beta}(1,1,0.306) \times \text{Binomial}(10,4,0.306)}{\text{Beta}(1,1,0.429) \times \text{Binomial}(10,4,0.429)} = 0.834$

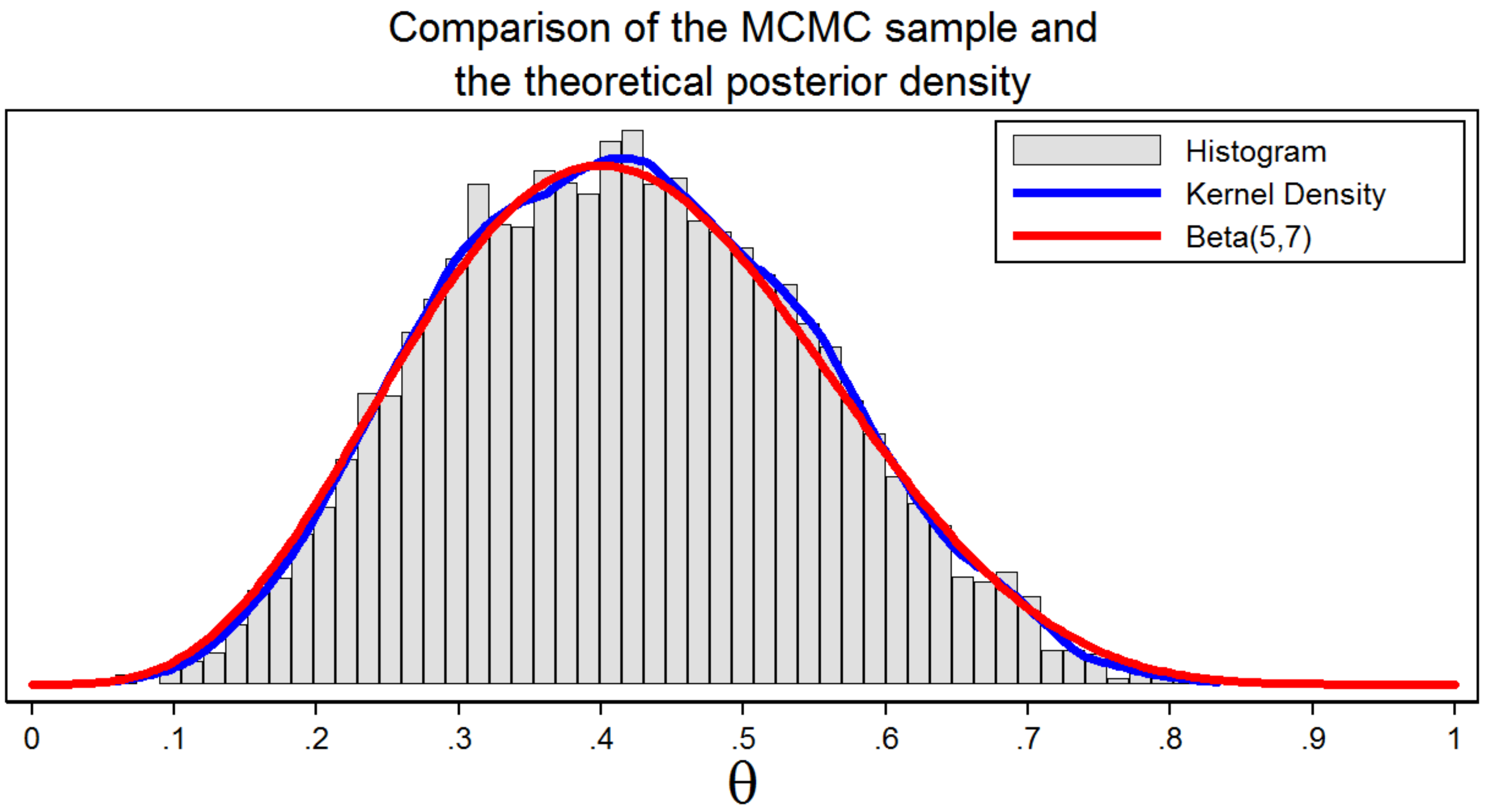
Step 2: Acceptance probability $\alpha(\theta_{new}, \theta_{t\text{-}1}) = \min\{r(\theta_{new}, \theta_{t\text{-}1}), 1\} = \min\{0.834, 1\} = 0.834$

Step 3: Draw $u \sim \text{Uniform}(0,1) = 0.617$

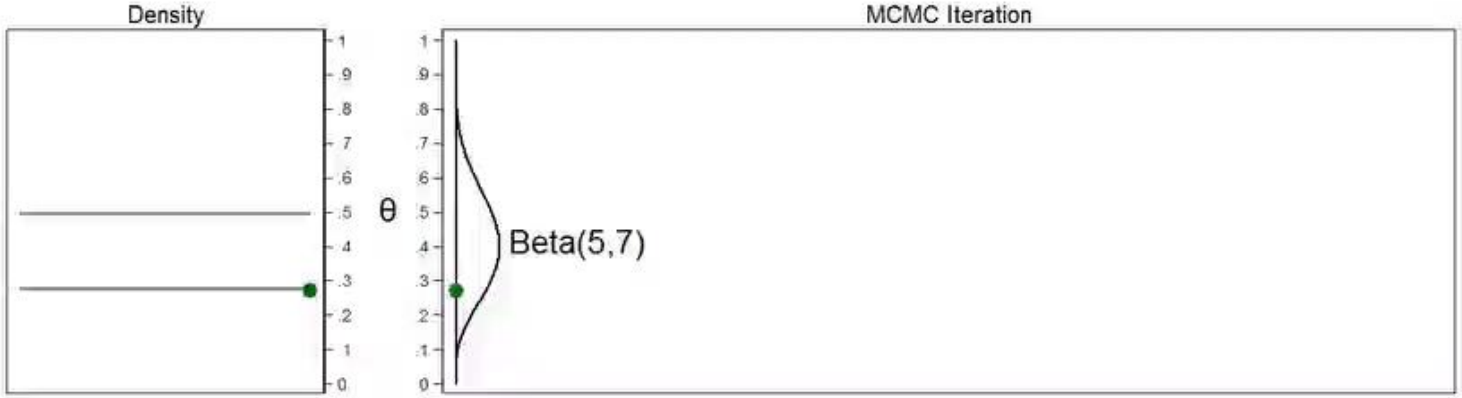
Step 4: If $u < \alpha(\theta_{new}, \theta_{t\text{-}1}) \rightarrow$ If $0.617 < 0.834$ Then $\theta_t = \theta_{new} = 0.306$

Otherwise $\theta_t = \theta_{t\text{-}1} = 0.429$

MCMC with Metropolis-Hastings



MCMC with Gibbs Sampling



Draw $\theta_t \sim \text{Beta}(1,1) \times \text{Binomial}(4,10)$

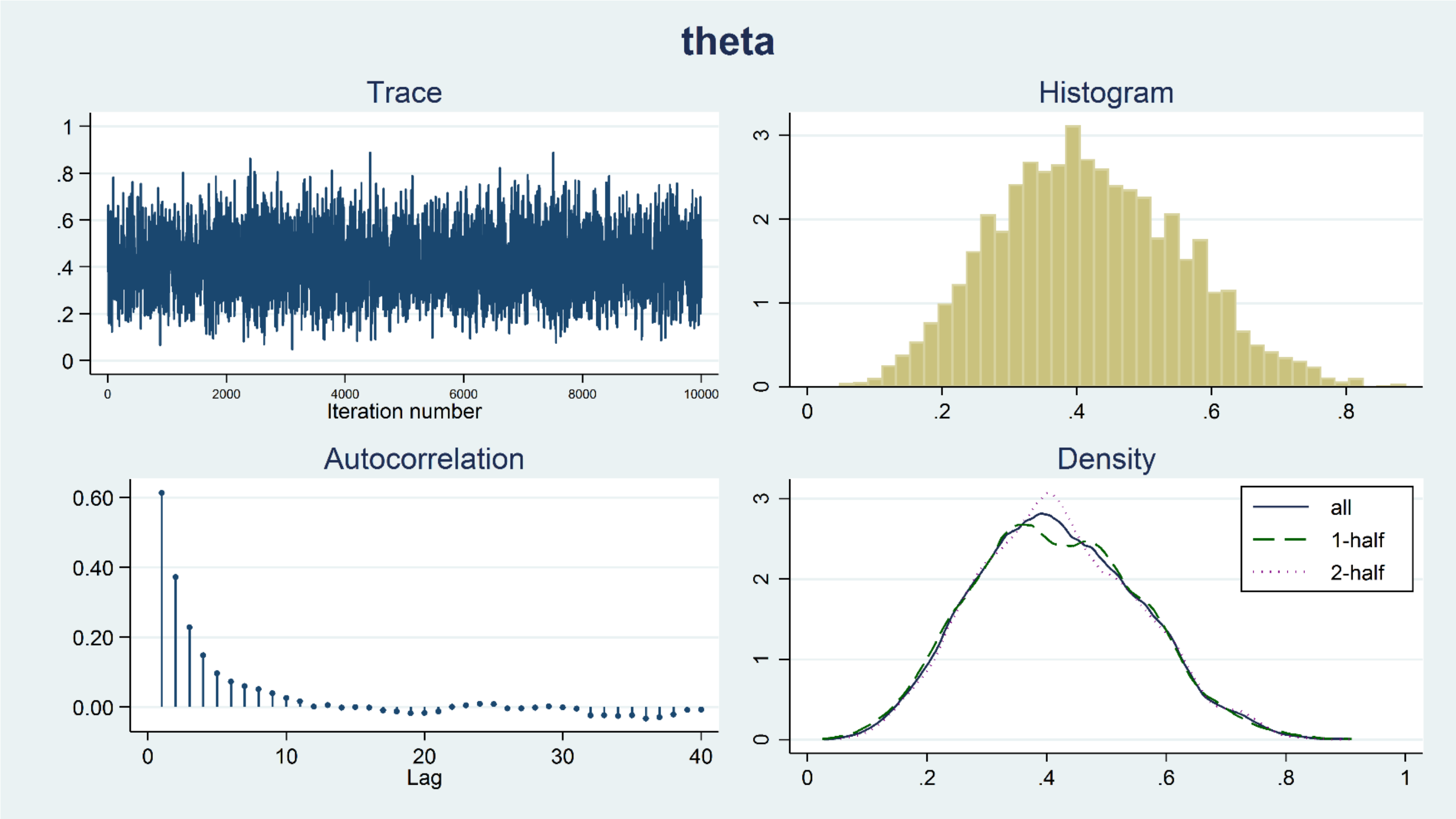
$\text{Beta}(5,7) = 0.271$

The **bayesmh** command

```
bayesmh heads,                                ///
     likelihood(dbernoulli({theta}))    ///
     prior({theta}, beta(1,1))
```

```
. bayesmh heads, likelihood(dbernoulli({theta})) prior({theta}, beta(1,1))

Burn-in ...
Simulation ...

Model summary
------------------------------------------------------------------------------
Likelihood:
  heads ~ bernoulli({theta})

Prior:
  {theta} ~ beta(1,1)
------------------------------------------------------------------------------

Bayesian Bernoulli model                         MCMC iterations  =     12,500
Random-walk Metropolis-Hastings sampling         Burn-in          =      2,500
                                                 MCMC sample size =     10,000
                                                 Number of obs    =         10
                                                 Acceptance rate  =      .4823
Log marginal likelihood = -7.8194591             Efficiency       =      .2291

------------------------------------------------------------------------------
             |                                                Equal-tailed
             |      Mean   Std. Dev.     MCSE     Median  [95% Cred. Interval]
-------------+----------------------------------------------------------------
       theta |  .4187117   .1342192   .002804   .4152274   .1746616   .6876875
------------------------------------------------------------------------------
```

Diagnostic Plots




```
bayesgraph diagnostics {theta}
```

Outline

- Introduction to Bayesian Analysis
- Coin Toss Example
- Priors, Likelihoods, and Posteriors
- Markov Chain Monte Carlo (MCMC)
- **Bayesian Linear Regression**
- Advantages and Disadvantages of Bayes

Bayesian Linear Regression

```
. webuse nhanes2, clear

. generate id = _n

. rename bpsystol sbp

. describe id sbp highbp healthstatus age sex race

              storage   display    value
variable name   type    format     label      variable label
-------------------------------------------------------------------------
id              float   %9.0g                 Identification Number
sbp             int     %9.0g                 Systolic Blood Pressure (mm/Hg)
highbp          byte    %8.0g      highbp     sbp >= 140 or dbp >= 90
healthstatus    byte    %9.0g      hs         Health Status
age             byte    %9.0g                 Age (years)
sex             byte    %9.0g      sex        Sex
race            byte    %9.0g      race       Race
```

We will ignore the sample weights to keep things simple.

$$sbp_i = \beta_0 + \beta_1 age_i + \beta_2 sex_i + e_i$$

```
. regress sbp age sex
```

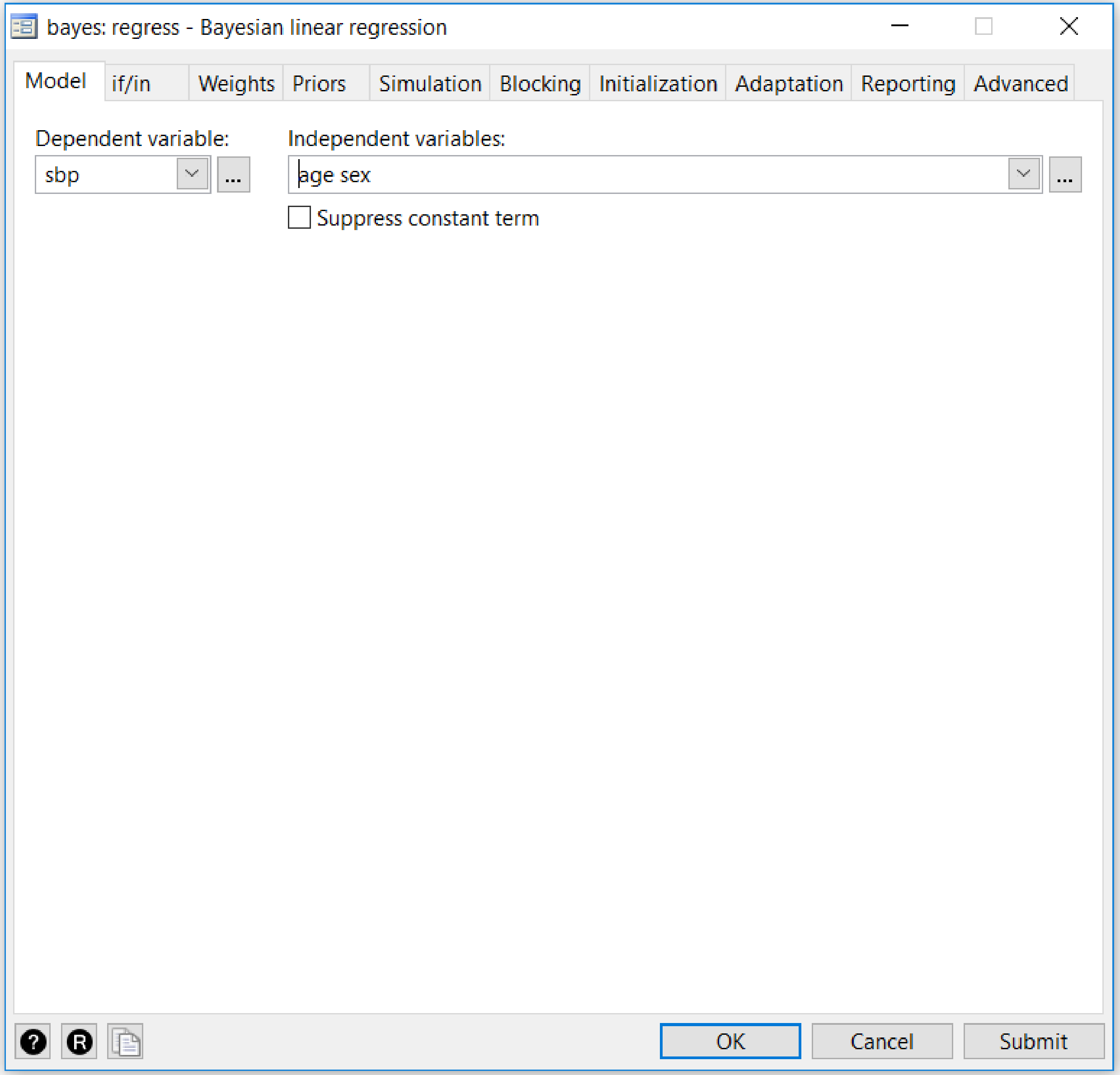
Source	SS	df	MS
Model	1345777.75	2	672888.876
Residual	4288892.27	10,348	414.465817
Total	5634670.03	10,350	544.412563

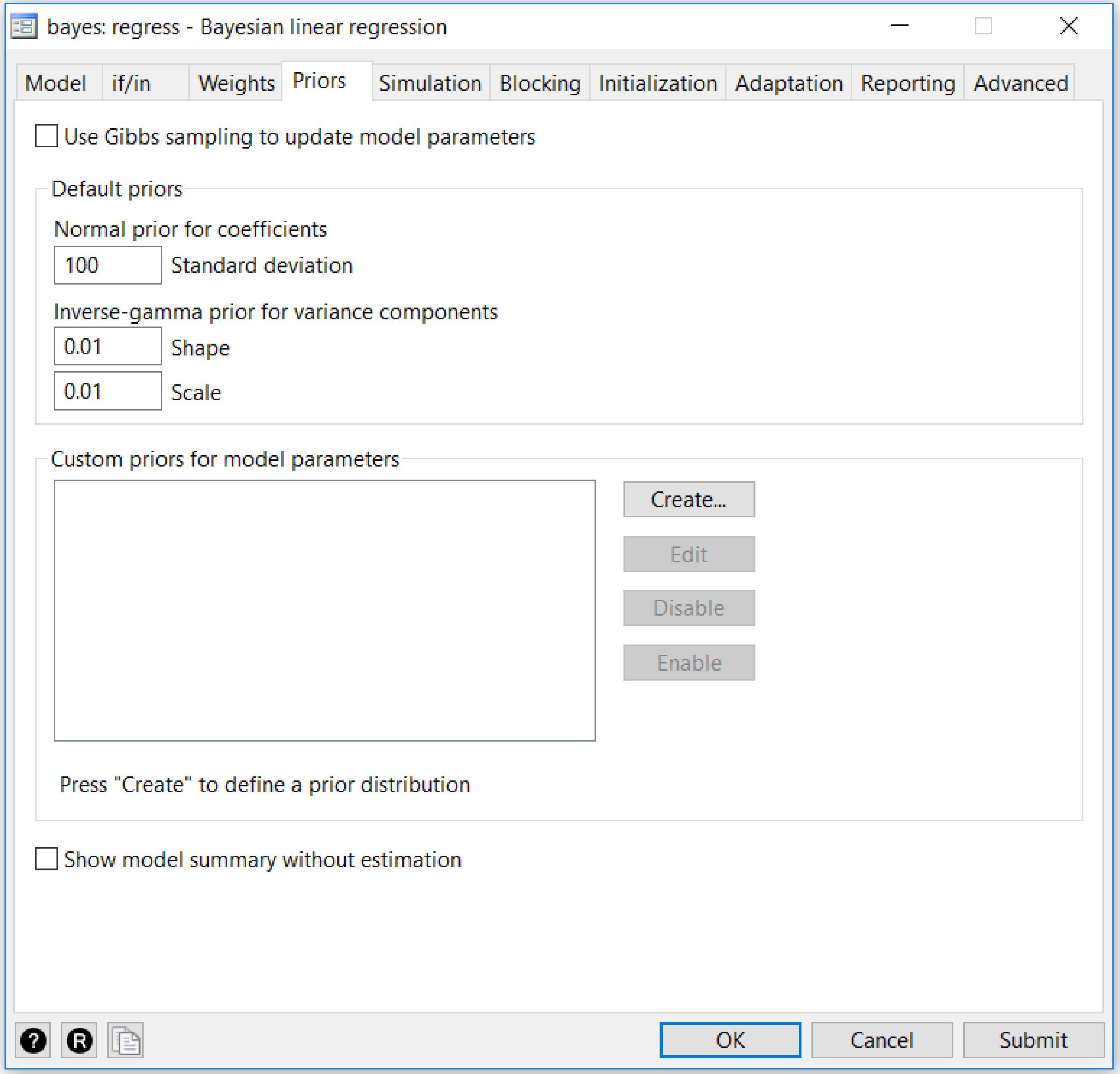
Number of obs	=	10,351
F(2, 10348)	=	1623.51
Prob > F	=	0.0000
R-squared	=	0.2388
Adj R-squared	=	0.2387
Root MSE	=	20.358

sbp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.6530798	.0116249	56.18	0.000	.6302929	.6758668
sex	-4.013619	.4007284	-10.02	0.000	-4.799124	-3.228114
_cons	105.9298	.8447992	125.39	0.000	104.2738	107.5857

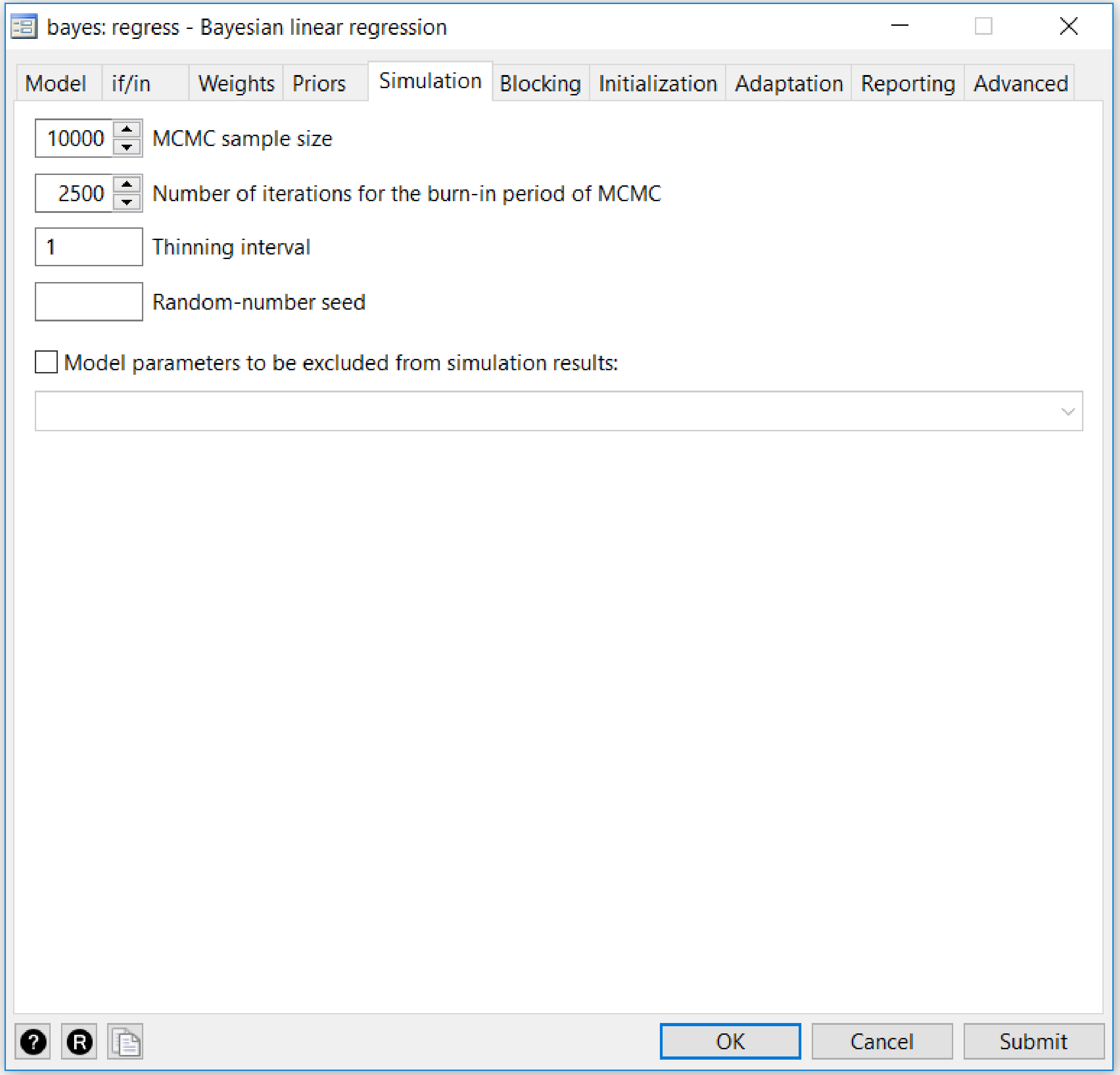
```
. bayes: regress sbp age sex

Burn-in ...
Simulation ...

Model summary
------------------------------------------------------------------------------
Likelihood:
  sbp ~ regress(xb_sbp,{sigma2})

Priors:
  {sbp:age sex _cons} ~ normal(0,10000)                                    (1)
             {sigma2} ~ igamma(.01,.01)
------------------------------------------------------------------------------
(1) Parameters are elements of the linear form xb_sbp.

Bayesian linear regression                       MCMC iterations  =     12,500
Random-walk Metropolis-Hastings sampling         Burn-in          =      2,500
                                                 MCMC sample size =     10,000
                                                 Number of obs    =     10,351
                                                 Acceptance rate  =      .3484
                                                 Efficiency:  min =     .07126
                                                              avg =      .1153
Log marginal likelihood = -45908.106                          max =      .2286

------------------------------------------------------------------------------
             |                                                Equal-tailed
             |      Mean   Std. Dev.     MCSE     Median  [95% Cred. Interval]
-------------+----------------------------------------------------------------
sbp          |
         age |  .6529088   .0113266   .000398   .6530755    .630286     .67464
         sex | -3.998945   .3923714   .013853  -3.991372  -4.774898  -3.224494
       _cons |  105.9215   .8241964   .030876    105.917    104.343    107.542
-------------+----------------------------------------------------------------
      sigma2 |  414.5504    5.84057   .122166   414.3584   403.5949   426.2577
------------------------------------------------------------------------------
Note: Default priors are used for model parameters.
```

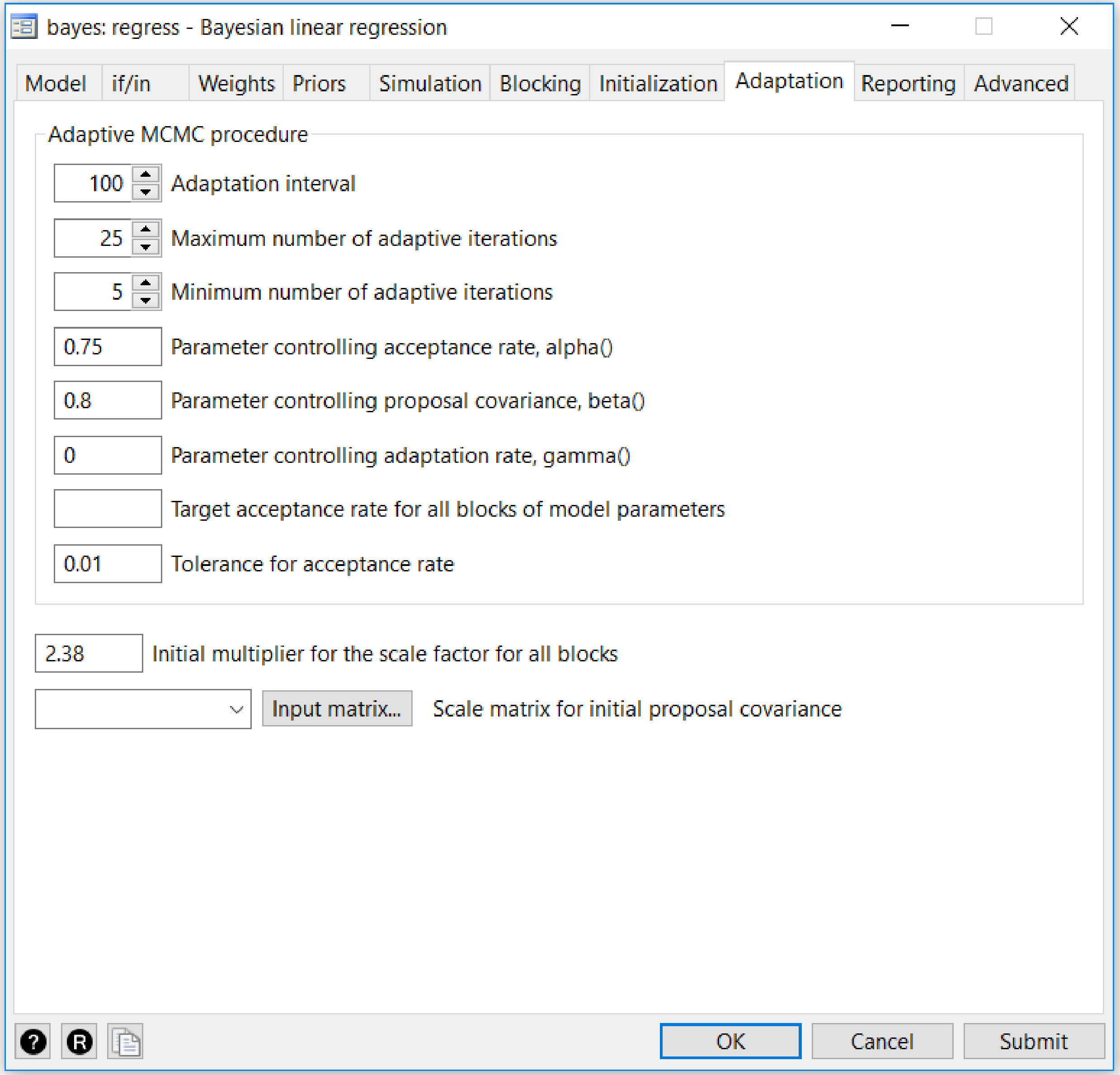
Data Analysis and Statistical Software




bayes: regress - Bayesian linear regression
Model
if/in
Weights
Priors
Simulation
Blocking
Initialization
Adaptation
Reporting
Advanced
Use Gibbs sampling to update model parameters
Default priors
Normal prior for coefficients
100
Standard deviation
Inverse-gamma prior for variance components
0.01
Shape
0.01
Scale
Custom priors for model parameters
Create...
Edit
Disable
Enable
Press "Create" to define a prior distribution
Show model summary without estimation
OK
Cancel
Submit

Data Analysis and Statistical Software




bayes: regress - Bayesian linear regression
Model
if/in
Weights
Priors
Simulation
Blocking
Initialization
Adaptation
Reporting
Advanced
Adaptive MCMC procedure
100 Adaptation interval
25 Maximum number of adaptive iterations
5 Minimum number of adaptive iterations
0.75 Parameter controlling acceptance rate, alpha()
0.8 Parameter controlling proposal covariance, beta()
0 Parameter controlling adaptation rate, gamma()
Target acceptance rate for all blocks of model parameters
0.01 Tolerance for acceptance rate
2.38 Initial multiplier for the scale factor for all blocks
Input matrix...
Scale matrix for initial proposal covariance
OK
Cancel
Submit

bayes options

```
. bayes: regress sbp age sex, noheader

Burn-in ...
Simulation ...
```

	Mean	Std. Dev.	MCSE	Median	Equal-tailed [95% Cred.	Interval]
sbp						
age	.6530982	.0110625	.000405	.6533669	.6318849	.6737548
sex	-4.010928	.3865944	.013339	-4.015244	-4.734603	-3.241869
_cons	105.9341	.8156657	.030091	105.9519	104.2668	107.5468
sigma2	414.5182	5.636057	.11081	414.3449	403.8492	425.8631

bayes options

```
. bayes, rseed(15): regress sbp age sex, noheader

Burn-in ...
Simulation ...
```

	Mean	Std. Dev.	MCSE	Median	Equal-tailed [95% Cred.	Interval]
sbp						
age	.6529088	.0113266	.000398	.6530755	.630286	.67464
sex	-3.998945	.3923714	.013853	-3.991372	-4.774898	-3.224494
_cons	105.9215	.8241964	.030876	105.917	104.343	107.542
sigma2	414.5504	5.84057	.122166	414.3584	403.5949	426.2577