

## Introduction to BOA

- **BOA** stands for **Bayesian Output Analysis**.
- The package and related documentation are available at <http://www.public-health.uiowa.edu>
- BOA is a set of functions written for R and S-Plus to perform MCMC convergence diagnostics.
- BOA is compatible with BUGS output, R objects and text files.
- BOA is a menu-driven package.
- Available for both UNIX and Windows platforms.
- We will show the capabilities of BOA using the genetic linkage example.
- To install BOA, we need to download and unzip a .tgz file.
- First, we start R or S-plus and type  

```
source('boa.r') or source('<program directory>/boa.r')  
boa.menu()
```

- After starting BOA the following menu window appears:

BOA MAIN MENU

\*\*\*\*\*

1:File >>  
2:Data >>  
3:Analysis >>  
4:Plot >>  
5:Options >>  
6:Window >>

- To read in data generated from an MCMC algorithm, you need to type "2". This will allow to read data from different formats including ASCII files or BUGS output files.

IMPORT DATA MENU

-----

1:BUGS Output File  
2:Flat ASCII File  
3:Data Matrix Object  
4:View Format Specifications  
5:Options...  
6:Back  
7:Return to Main Menu

Selection: 2

- By selecting “2”, we are asking BOA to read in an ASCII file.
- The ASCII file must be a column file where each column is separated by a blank space, a tab delimiter or a comma.
- The first row of this file must have the names for the parameters.
- The remaining rows are the values of the parameters from the MCMC.
- In this case below, we are reading the file link1.txt that corresponds to the genetic linkage example.

```
Enter filename prefix without the .txt extension
```

```
[Working Directory: ""]
```

```
1: link1
```

```
Read 1 items
```

```
+++ Data successfully imported +++
```

- link1.txt is a one column file of the parameter "theta" for the genetic linkage model.
- Since some of the BOA diagnostics use multiple chains, I also created files link2.txt, link3.txt, link4.txt.

Enter filename prefix without the .txt extension

[Working Directory: ""]

1: link2

Read 1 items

+++ Data successfully imported +++

- Each file considers 1000 draws from the Metropolis-Hastings algorithm for the genetic linkage model from different starting points.
- The option 3 of the main BOA menu brings you to the following menu:

ANALYSIS MENU

=====

1:Descriptive Statistics >>

2:Convergence Diagnostics >>

3:Options...

4:Return to Main Menu

Selection: 1

- By selecting 1, BOA will take you to the descriptive statistics menu.

DESCRIPTIVE STATISTICS MENU

- 
- 1: Autocorrelations
  - 2: Correlation Matrix
  - 3: Highest Probability Density Intervals
  - 4: Summary Statistics
  - 5: Back
  - 6: Return to Main Menu

Selection: 3

- After typing 3, BOA will produce interval estimates for the parameter  $\theta$  of the genetic linkage model based on all four files.

HIGHEST PROBABILITY DENSITY INTERVALS:

=====

Alpha level = 0.05

Chain: link1

-----

	Lower Bound	Upper Bound
theta	0.5266789	0.7255235

Chain: link2

-----

	Lower Bound	Upper Bound
theta	0.5234056	0.7175498

- The first option for the descriptive statistics menu gives autocorrelation values for the parameter of our model.

LAGS AND AUTOCORRELATIONS:

=====

Chain: link1

-----

	Lag 1	Lag 5	Lag 10	Lag 50
theta	0.643131	0.1479772	0.03049042	-0.01126595

Chain: link2

-----

	Lag 1	Lag 5	Lag 10	Lag 50
theta	0.6218542	0.0760426	0.001186962	0.01503285

Chain: link3

-----

	Lag 1	Lag 5	Lag 10	Lag 50
theta	0.633444	0.1003947	0.01048863	-0.002252984

Chain: link4

-----

	Lag 1	Lag 5	Lag 10	Lag 50
theta	0.632891	0.1097898	0.03234036	0.01445291

- The option 2 for in the analysis leads to the following menu.

CONVERGENCE DIAGNOSTICS MENU

-----

1:Brooks, Gelman & Rubin

2:Geweke

3:Heidelberger & Welch

4:Raftery & Lewis

5:Back

6:Return to Main Menu

Selection: 1

BROOKS, GELMAN AND RUBIN CONVERGENCE DIAGNOSTICS:

=====

Iterations used = 5002:10001

Potential Scale Reduction Factors

-----

theta

1.000180

GEWEKE CONVERGENCE DIAGNOSTIC:

=====

Fraction in first window = 0.1

Fraction in last window = 0.5

Chain: link1

-----

theta

Z-Score 0.1528488

p-value 0.8785175

Press <ENTER> to continue

Chain: link2

-----

theta

Z-Score 1.1465992

p-value 0.2515473

Chain: link3

-----

theta

Z-Score -0.0744460

p-value 0.9406555

Press <ENTER> to continue

Chain: link4

-----

theta

Z-Score -0.1002423

p-value 0.9201520



- We can also obtain results for the Raftery and Lewis diagnostic.

RAFTERY AND LEWIS CONVERGENCE DIAGNOSTIC:

=====

Quantile = 0.025

Accuracy = +/- 0.005

Probability = 0.95

Chain: link1

-----

	Thin	Burn-in	Total	Lower Bound	Dependence	Factor
theta	5	20	26340	3746		7.0315

Chain: link2

-----

	Thin	Burn-in	Total	Lower Bound	Dependence	Factor
theta	5	15	20465	3746		5.463161

Chain: link3

-----

	Thin	Burn-in	Total	Lower Bound	Dependence	Factor
theta	5	15	21215	3746		5.663374

Chain: link4

-----

	Thin	Burn-in	Total	Lower Bound	Dependence	Factor
theta	5	15	21935	3746		5.855579

- BOA also has available a plot menu for descriptive statistics and graphical displays

#### PLOT MENU

=====

- 1:Descriptive
- 2:Convergence Diagnostics
- 3:Options...
- 4:Return to Main Menu

- For descriptive statistics we may obtain ACF plots, density estimators, ergodic means and trace plots

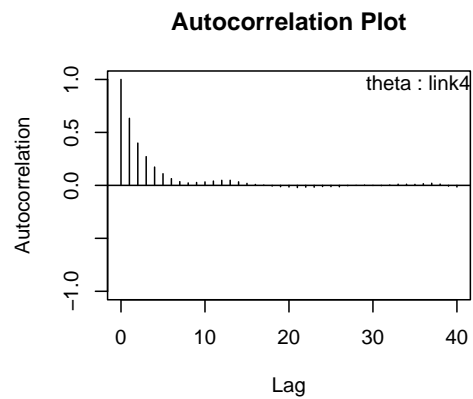
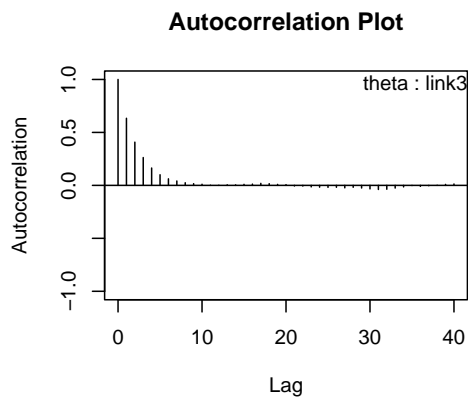
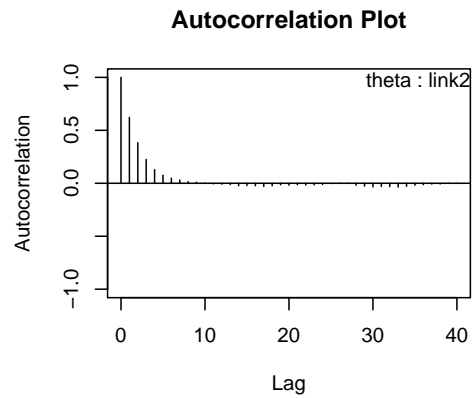
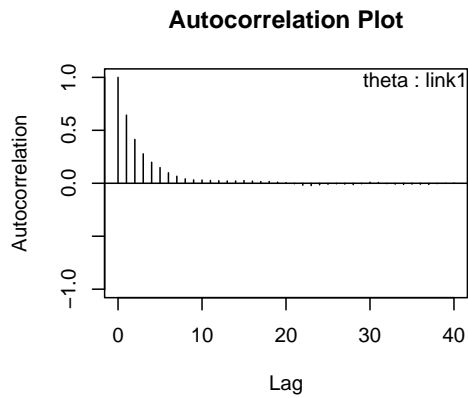
#### DESCRIPTIVE PLOT MENU

-----

- 1:Autocorrelations
- 2:Density
- 3:Running Mean
- 4:Trace
- 5:Back
- 6:Return to Main Menu

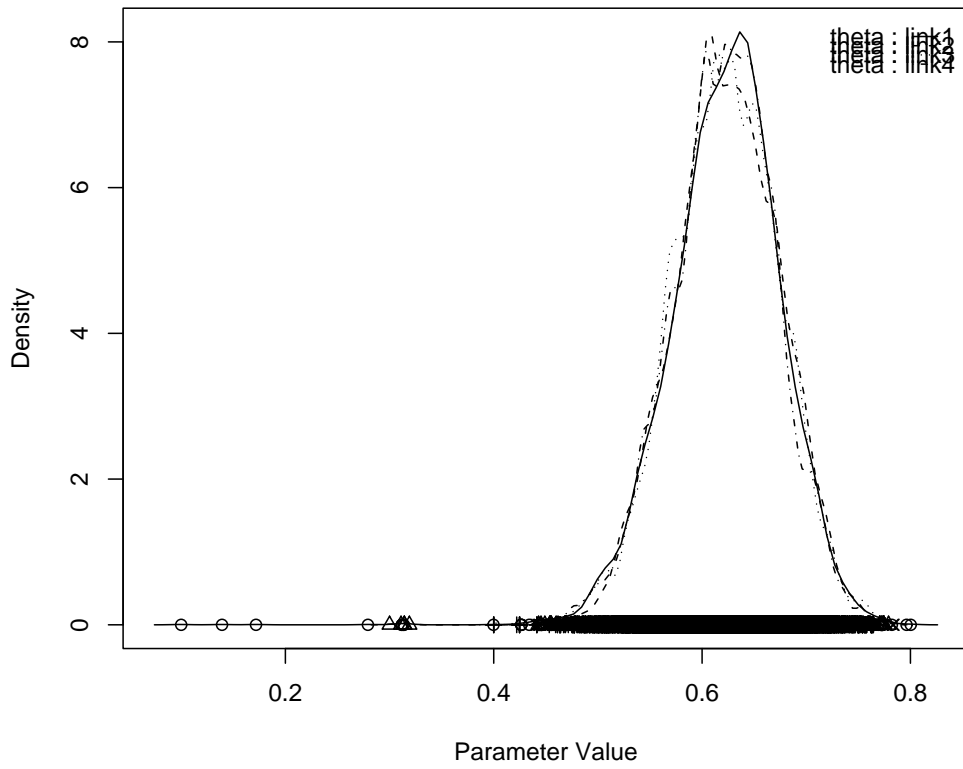
- Here, I am showing some of the plots obtained with this menu.

## Bayesian Output Analysis



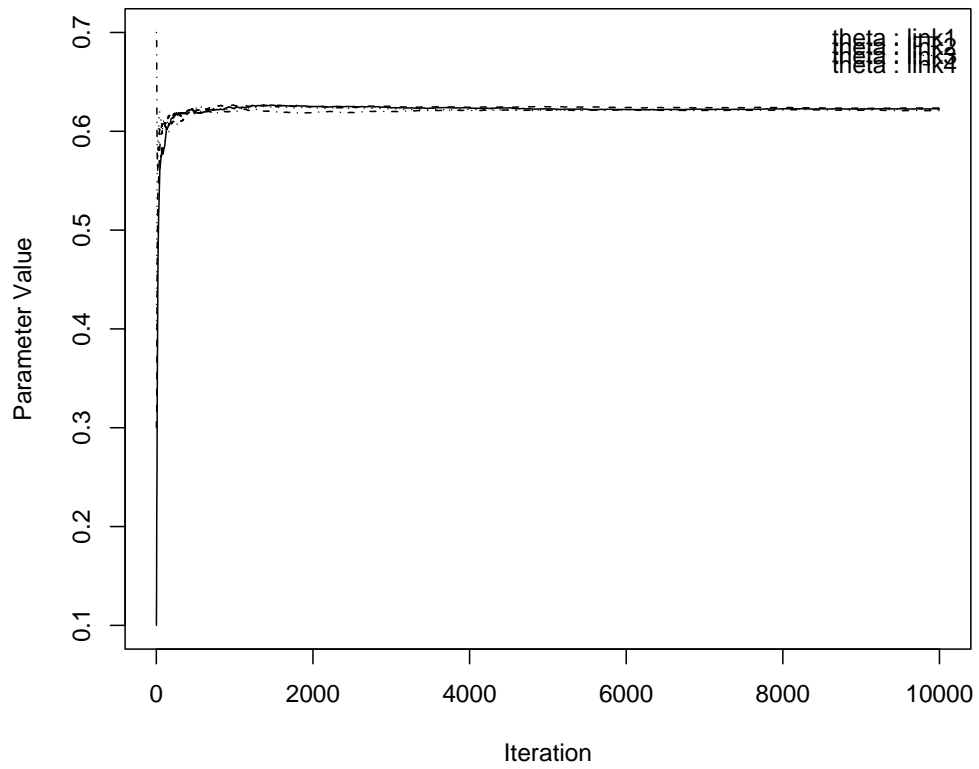
# Bayesian Output Analysis

## Density Plot



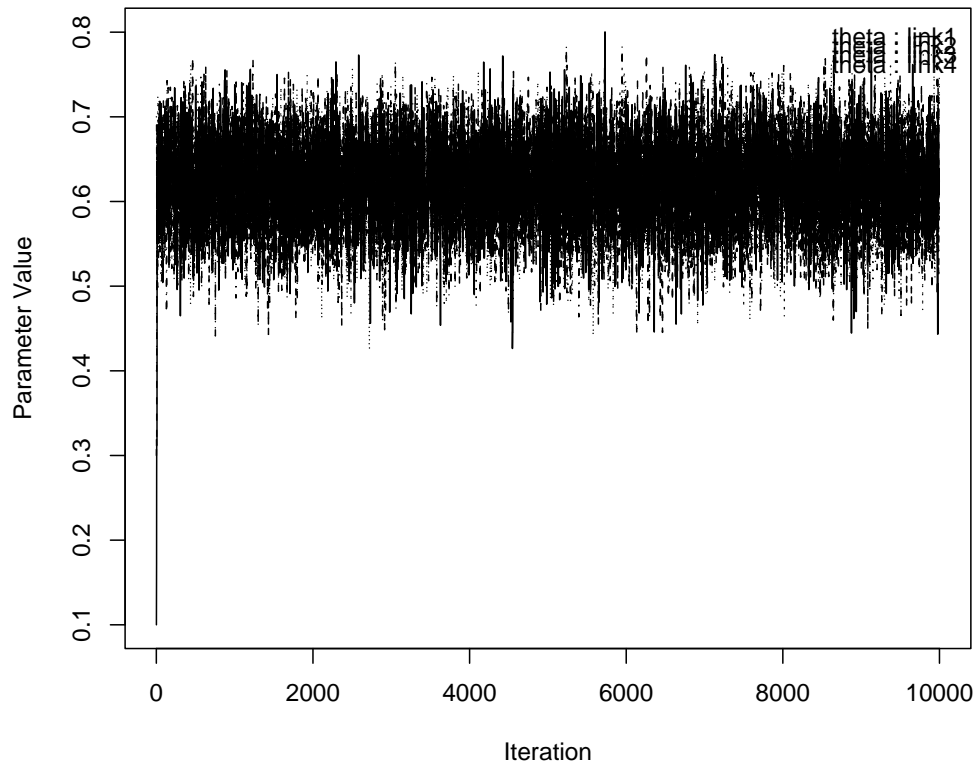
# Bayesian Output Analysis

## Running Mean Plot



# Bayesian Output Analysis

Trace Plot



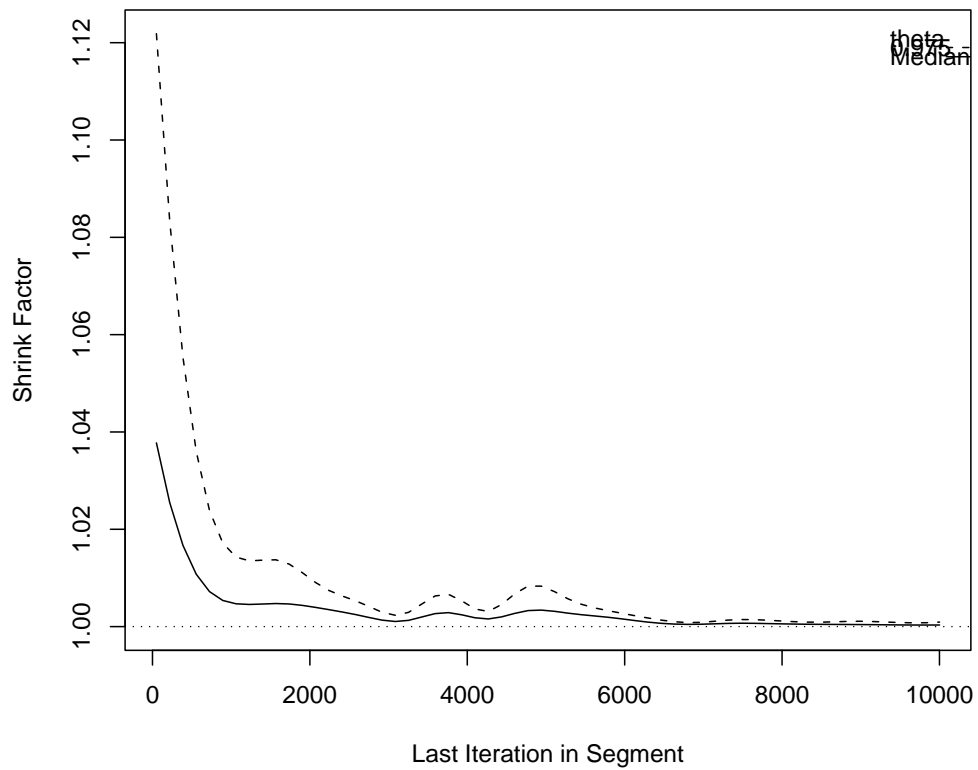
- For convergence diagnostics you have several options as well.

### CONVERGENCE DIAGNOSTICS PLOT MENU

- 
- 1: Brooks & Gelman
  - 2: Gelman & Rubin
  - 3: Geweke
  - 4: Back
  - 5: Return to Main Menu

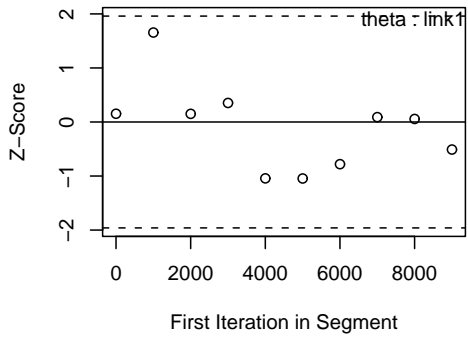
## Bayesian Output Analysis

### Gelman & Rubin Shrink Factors

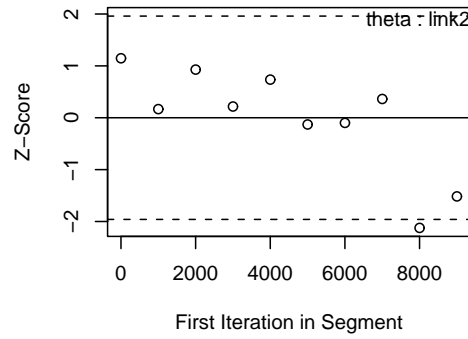


# Bayesian Output Analysis

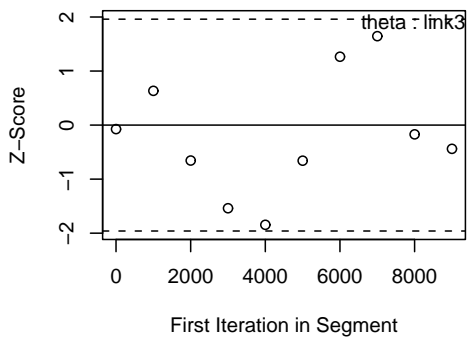
**Geweke Convergence Diagnostic**



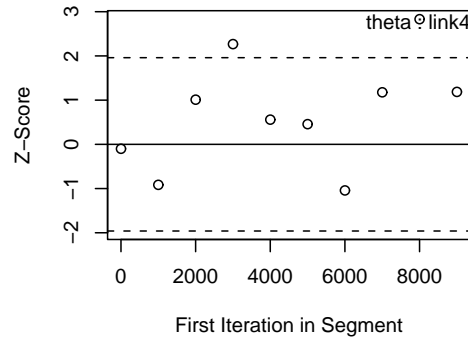
**Geweke Convergence Diagnostic**



**Geweke Convergence Diagnostic**



**Geweke Convergence Diagnostic**



- To save a plot in a postscript is a bit inconvenient. After creating the plot, you need to go back to the main menu and then to the window menu.

## WINDOW 1 MENU

=====

- 1: Previous
- 2: Next
- 3: Save to Postscript File
- 4: Close
- 5: Close All
- 6: Return to Main Menu



One last thing....

- **HW exercise Due April 8** For the HW you turned in today where the  $N(0,1)$  is the target distribution and with the  $U(\theta - \delta, \theta + \delta)$  perform convergence diagnostics using BOA. You will need to run the Metropolis-Hastings from different starting points to produce multiple chains and apply diagnostics such as Gelman and Rubin's.