

# Performing Bayesian analysis in Stata using WinBUGS

Tom Palmer, John Thompson & Santiago Moreno

Department of Health Sciences,  
University of Leicester, UK

13<sup>th</sup> UK Stata Users Group Meeting,  
10 September 2007



- 1 The Bayesian approach & WinBUGS
- 2 The `winbugsfromstata` package
- 3 How to run an analysis
- 4 Summary & developments

# The Bayesian approach

## Bayes Theorem

$$\textit{Posterior} \propto \textit{Likelihood} \times \textit{prior}$$

## Bayes Theorem

$$\textit{Posterior} \propto \textit{Likelihood} \times \textit{prior}$$

- Direct probability statements - not frequentist - subjective

## Bayes Theorem

$$\textit{Posterior} \propto \textit{Likelihood} \times \textit{prior}$$

- Direct probability statements - not frequentist - subjective
- Complex posterior marginal distributions - estimation via simulation

## Bayes Theorem

$$\textit{Posterior} \propto \textit{Likelihood} \times \textit{prior}$$

- Direct probability statements - not frequentist - subjective
- Complex posterior marginal distributions - estimation via simulation
- Markov chain Monte Carlo (MCMC) methods

- Bayesian statistics using Gibbs sampling

- Bayesian statistics using Gibbs sampling
- MRC Biostatistics unit  
<http://www.mrc-bsu.cam.ac.uk/bugs>

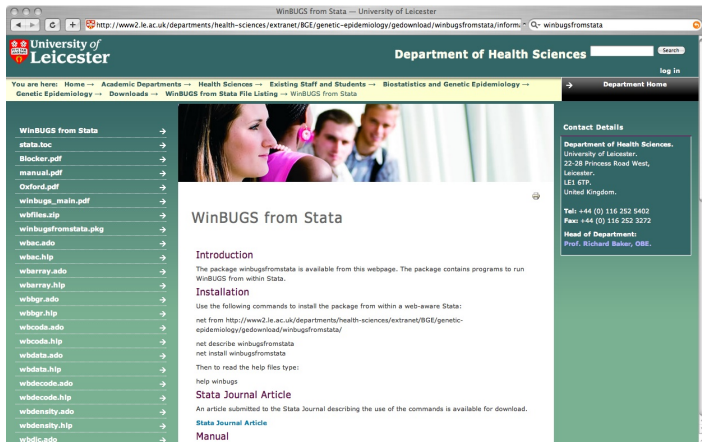


- Bayesian statistics using Gibbs sampling
- MRC Biostatistics unit  
<http://www.mrc-bsu.cam.ac.uk/bugs>
- Health Economics, Medical Statistics

- Bayesian statistics using Gibbs sampling
- MRC Biostatistics unit  
<http://www.mrc-bsu.cam.ac.uk/bugs>
- Health Economics, Medical Statistics
- Disadvantages: data management, post-processing of results, graphics

# The winbugsfromstata package

- Stata interface to WinBUGS [Thompson et al., 2006]  
`http://www2.le.ac.uk/departments/health-sciences/extranet/BGE/genetic-epidemiology/gedownload/information`



The screenshot shows a web browser window displaying the University of Leicester website. The page title is "WinBUGS from Stata - University of Leicester". The URL in the address bar is `http://www2.le.ac.uk/departments/health-sciences/extranet/BGE/genetic-epidemiology/gedownload/winbugsfromstata/inform...`. The page content includes a navigation menu on the left with links to various files like `stata.toc`, `Blocker.pdf`, `manual.pdf`, `Oxford.pdf`, `winbugs_main.pdf`, `wbfiles.zip`, `winbugsfromstata.pkg`, `wbec.ado`, `wbec.hlp`, `wbarray.ado`, `wbarray.hlp`, `wbbgr.ado`, `wbbgr.hlp`, `wbcodes.ado`, `wbcodes.hlp`, `wbdata.ado`, `wbdata.hlp`, `wbdecode.ado`, `wbdecode.hlp`, `wbdensity.ado`, `wbdensity.hlp`, and `wbolic.ado`. The main content area features a header image of three people, followed by the title "WinBUGS from Stata" and an "Introduction" section. The introduction states: "The package winbugsfromstata is available from this webpage. The package contains programs to run WinBUGS from within Stata." Below this is an "Installation" section with instructions on how to install the package using Stata commands: `net from http://www2.le.ac.uk/departments/health-sciences/extranet/BGE/genetic-epidemiology/gedownload/winbugsfromstata/`, `net describe winbugsfromstata`, and `net install winbugsfromstata`. It also mentions that help files are available and provides a link to a "Stata Journal Article" and a "Manual". On the right side of the page, there is a "Contact Details" section for the Department of Health Sciences, University of Leicester, including the address, telephone number (+44 (0) 116 252 5402), fax number (+44 (0) 116 252 3272), and the head of department, Prof. Richard Baker, OBE.

# The winbugsfromstata package

WinBUGS from Stata — University of Leicester

http://www2.le.ac.uk/departments/health-sciences/extranet/BCE/genetic-epidemiology/gedownload/winbugsfromstata/inform... winbugsfromstata

wbdic.ado →  
wbdic.hlp →  
wbgeweke.ado →  
**wbgeweke.hlp**  
wbhull.ado →  
wbhull.hlp →  
wbinervals.ado →  
wbinervals.hlp →  
wbrun.ado →  
wbrun.hlp →  
wbscalar.ado →  
wbscalar.hlp →  
wbscript.ado →  
wbscript.hlp →  
wbsection.ado →  
wbsection.hlp →  
wbstats.ado →  
wbstats.hlp →  
wbstructure.ado →  
wbstructure.hlp →  
wbttrace.ado →  
wbttrace.hlp →  
wbvector.ado →  
wbvector.hlp →  
winbugs.hlp →

Useful Links  
Seminars

## Manual

A guide to the programs presented in the style of the Stata Reference Manuals is available for download.

### Manual

### Examples

A number of worked examples demonstrating the use of the programs are provided below.

**Oxford example**  
**Blocker example**

### Ado-files and help files for manual installation

The ado and help files are available for download. Save the files on the adopath.

#### wbfiles

### Contact details

If you have any questions about the programs please read the help files, manual and worked examples first.

If you can't find an answer to your question or if you find any errors in the programs please contact Professor John Thompson ([john.thompson@le.ac.uk](mailto:john.thompson@le.ac.uk)). Please begin the subject line of an email with *winbugsfromstata*.

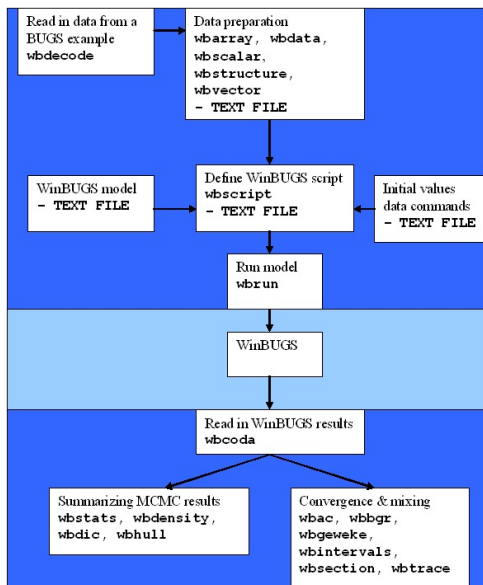
### Links

Darren Greenwood's bugsutils package is available from <http://www.personal.leeds.ac.uk/~hssdg/Stata/>.

### Notes

Use the following Stata command to check you have the latest version:  
adoupdate winbugsfromstata

# How to run an analysis



# help winbugs

Viewer (#1) [view "Z:\winbugsfromstata\wb.commands.6feb07\winbugs.hlp"]

view "Z:\winbugsfromstata\wb.commands.6feb07\winbugs.hlp"

Advice Contents What's New News

---

**help winbugs**

---

**Title**

**winbugs** — help on available routines for running WinBUGS

**Description**

This help file describes the commands available for running WinBUGS from within Stata. There is no executable command **winbugs**.

The files are

<b>wbarray</b>	writes data from Stata as a WinBUGS array
<b>wbdata</b>	writes mixed data (scalars, vectors & structures) from Stata as a WinBUGS list
<b>wbscalar</b>	writes scalars from Stata as a WinBUGS list
<b>wbstructure</b>	writes data from Stata as a WinBUGS structure
<b>wbvector</b>	writes data from Stata as a WinBUGS vector
<b>wbcoda</b>	reads data from a WinBUGS coda file into Stata
<b>wbdecode</b>	reads data from a WinBUGS list into Stata
<b>wbrun</b>	runs a pre-prepared WinBUGS script file from within Stata
<b>wbscript</b>	writes & runs a WinBUGS script file from within Stata
<b>wbac</b>	autocorrelation plots
<b>wbbgr</b>	Brooks-Gelman-Rubin plot
<b>wbgeweke</b>	test of means for two sections of a chain
<b>wbintervals</b>	interval plots for sections of a chain
<b>wbsection</b>	density plots of subsections of a chain
<b>wbtrace</b>	trace (history) plot(s) of an MCMC run
<b>wbdensity</b>	smoothed posterior density estimates
<b>wbdic</b>	read Deviance Information Criterion (DIC) statistics in a winBUGS log-file into Stata
<b>wbhull</b>	contours for pairs of parameters
<b>wbstats</b>	summary statistics from an MCMC chain

**Reference**

The WinBUGS Manual is available from [www.mrc-bsu.cam.ac.uk/bugs](http://www.mrc-bsu.cam.ac.uk/bugs).

**Author**

John Thompson, Department of Health Sciences, University of Leicester. Please report any errors to [john.thompson@le.ac.uk](mailto:john.thompson@le.ac.uk).

# Example analysis: Schools

- Schools example [Goldstein et al., 1993],[Spiegelhalter et al., 2004]

# Example analysis: Schools

- Schools example [Goldstein et al., 1993],[Spiegelhalter et al., 2004]
- Between-school variation in exam results from inner London schools



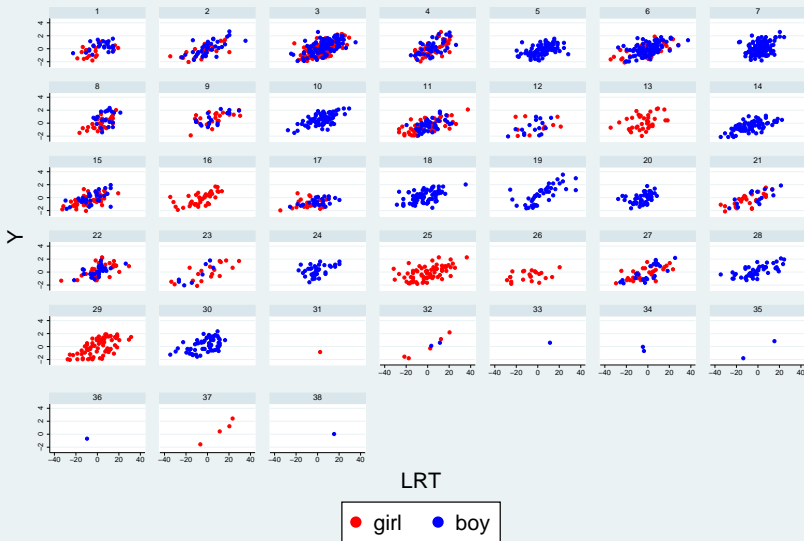
# Example analysis: Schools

- Schools example [Goldstein et al., 1993],[Spiegelhalter et al., 2004]
- Between-school variation in exam results from inner London schools
- Standardized mean scores ( $Y$ ) 1,978 pupils, 38 schools

# Example analysis: Schools

- Schools example [Goldstein et al., 1993],[Spiegelhalter et al., 2004]
- Between-school variation in exam results from inner London schools
- Standardized mean scores ( $Y$ ) 1,978 pupils, 38 schools
- LRT: London Reading Test, VR: verbal reasoning, Gender intake of school, denomination of school

# Data for the Schools example



Graphs by school

# The model

- Hierarchical model; specified the mean and variance

# The model

- Hierarchical model; specified the mean and variance
- Model:

$$Y_{ij} \sim N(\mu_{ij}, \tau_{ij})$$

$$\mu_{ij} = \gamma_{1j} + \gamma_{2j}LRT_{ij} + \gamma_{3j}VR1_{ij} + \beta_1LRT_{ij}^2 + \beta_2VR2_{ij}$$

$$+ \beta_3Gir1_{ij} + \beta_4Gsch_j + \beta_5Bsch_j + \beta_6CEsch_j + \beta_7RCsch_j + \beta_8Osch_j$$

$$\log \tau_{ij} = \theta + \phi LRT_{ij}$$

# WinBUGS model statement

```
model{
for(p in 1 : N){
Y[p] ~ dnorm(mu[p], tau[p])
mu[p] <- alpha[school[p], 1] + alpha[school[p], 2] * LRT[p]
+ alpha[school[p], 3] * VR[p, 1] + beta[1] * LRT2[p]
+ beta[2] * VR[p, 2] + beta[3] * Gender[p]
+ beta[4] * School.gender[p, 1] + beta[5] * School.gender[p, 2]
+ beta[6] * School.denom[p, 1] + beta[7] * School.denom[p, 2]
+ beta[8] * School.denom[p, 3]
log(tau[p]) <- theta + phi * LRT[p]
sigma2[p] <- 1 / tau[p]
LRT2[p] <- LRT[p] * LRT[p]
}

min.var <- exp(-(theta + phi * (-34.6193))) # lowest LRT score = -34.6193
max.var <- exp(-(theta + phi * (37.3807))) # highest LRT score = 37.3807

# Priors for fixed effects:
for (k in 1 : 8){
beta[k] ~ dnorm(0.0, 0.0001)
}
theta ~ dnorm(0.0, 0.0001)
phi ~ dnorm(0.0, 0.0001)

# Priors for random coefficients:
for (j in 1 : M) {
alpha[j, 1 : 3] ~ dnmnorm(gamma[1:3 ], T[1:3 ,1:3 ])
alpha1[j] <- alpha[j,1]
}

# Hyper-priors:
gamma[1 : 3] ~ dnmnorm(mn[1:3 ], prec[1:3 ,1:3 ])
T[1 : 3, 1 : 3 ] ~ dwish(R[1:3 ,1:3 ], 3)
}
```

## Do-file for the example

```
// winbugsfromstata demo, 16august2007
cd "Z:/conferences/stata.users.uk.2007/schools"
wbdecode, file(Schoolsdata.txt) clear

wbscript, sav('c(pwd)'/script.txt, replace) ///
model('c(pwd)'/Schoolsmodel.txt) ///
data('c(pwd)'/Schoolsdata.txt) ///
inits('c(pwd)'/Schoolsinit.txt) ///
coda('c(pwd)'/out) ///
burn(500) update(1000) ///
set(beta gamma phi theta) dic ///
log('c(pwd)'/winbugslog.txt) ///
quit

wbrun , sc('c(pwd)'/script.txt) ///
win(Z:/winbugs/WinBUGS14/WinBUGS14.exe)

clear
set memory 500m
wbcoda, root(out) clear

wbstats gamma* beta* phi theta

wbtrace beta_1 gamma_1 phi theta
wbdensity beta_1 gamma_1 phi theta
wbac beta_1 gamma_1 phi theta
wbhull beta_1 beta_2 gamma_2, peels(1 5 10 25)

wbgeweke beta_1 gamma_1 phi theta

wbdic using winbugslog.txt
```

# wbrun screenshot 1

The screenshot displays the Stata/IC 10.0 interface. The main window shows the Stata logo and version information. A 'Review' pane on the left lists three commands: 'Browse', 'do "C:\temp\ST0000000000.dmp"', and 'wbrun, sc "qwe\font140\W...'. A 'Variables' pane is empty. A 'Command' window at the bottom shows the execution of the 'wbrun' command. A separate window titled 'C:\WINDOWS\system32\cmd.exe' is open over the main window, showing a blue background and the text 'C:\WINDOWS\system32\cmd.exe'.

Stata/IC 10.0 - [Results]

File Edit Data Graphics Statistics User Window Help

Review

Command JC

```
1 Browse
2 do "C:\temp\ST0000000000.dmp"
3 wbrun, sc "qwe\font140\W...
```

STATISTICS 10.0  
Statistics/Data Analysis

copyright 1984-2007  
StataCorp  
4901 Lakeside Drive  
College Station, Texas 77801 USA  
800-STATA-PC http://www.stata.com  
512-692-4600 stata@stata.com  
978-699-4601 (Fax)

C:\WINDOWS\system32\cmd.exe

Variables

Name

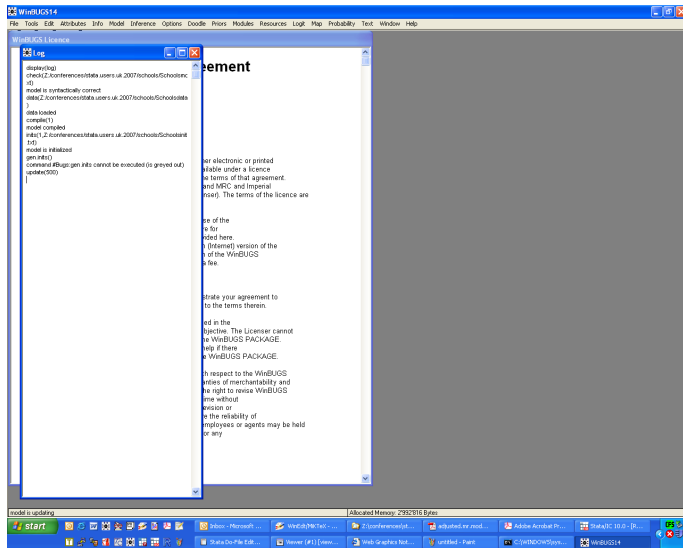
Command

start | zbox - Microsoft Out... | Web of Me Test: [2]... | z:\preferences\stata... | adkubad.nv.mocdel.w... | webappforstata.pr... | 05

Address Acrobat Profile... | Web Graphics Notes... | Stata/IC 10.0 - [Data... | Stata Do-File Editor... | C:\WINDOWS\sys...



# wbrun screenshot 2



## wbstats output

```
. wbstats gamma* beta* phi theta
Parameter      n      mean      sd      sem      median      95% CrI
gamma_1        500    -0.715    0.103    0.0179   -0.715 (  -0.951,  -0.523 )
gamma_2        500     0.031    0.010    0.0005    0.031 (   0.010,   0.052 )
gamma_3        500     0.967    0.105    0.0225    0.972 (   0.750,   1.168 )
beta_1         500     0.000    0.000    0.0000    0.000 (   0.000,   0.000 )
beta_2         500     0.433    0.072    0.0099    0.435 (   0.284,   0.576 )
beta_3         500     0.173    0.048    0.0031    0.172 (   0.085,   0.271 )
beta_4         500     0.151    0.141    0.0230    0.164 (  -0.156,   0.392 )
beta_5         500     0.091    0.105    0.0150    0.087 (  -0.094,   0.318 )
beta_6         500    -0.279    0.183    0.0279   -0.290 (  -0.618,   0.108 )
beta_7         500     0.170    0.105    0.0158    0.169 (  -0.029,   0.380 )
beta_8         500    -0.109    0.209    0.0376   -0.124 (  -0.485,   0.357 )
phi            500    -0.003    0.003    0.0002   -0.003 (  -0.009,   0.003 )
theta          500     0.579    0.032    0.0016    0.579 (   0.513,   0.649 )
```

## wbstats output

```
. wbstats gamma* beta* phi theta
Parameter      n      mean      sd      sem      median      95% CrI
gamma_1        500    -0.715    0.103    0.0179   -0.715 (  -0.951,  -0.523 )
gamma_2        500     0.031    0.010    0.0005    0.031 (   0.010,   0.052 )
gamma_3        500     0.967    0.105    0.0225    0.972 (   0.750,   1.168 )
beta_1         500     0.000    0.000    0.0000    0.000 (   0.000,   0.000 )
beta_2         500     0.433    0.072    0.0099    0.435 (   0.284,   0.576 )
beta_3         500     0.173    0.048    0.0031    0.172 (   0.085,   0.271 )
beta_4         500     0.151    0.141    0.0230    0.164 (  -0.156,   0.392 )
beta_5         500     0.091    0.105    0.0150    0.087 (  -0.094,   0.318 )
beta_6         500    -0.279    0.183    0.0279   -0.290 (  -0.618,   0.108 )
beta_7         500     0.170    0.105    0.0158    0.169 (  -0.029,   0.380 )
beta_8         500    -0.109    0.209    0.0376   -0.124 (  -0.485,   0.357 )
phi            500    -0.003    0.003    0.0002   -0.003 (  -0.009,   0.003 )
theta          500     0.579    0.032    0.0016    0.579 (   0.513,   0.649 )
```

- regress  $\gamma_2$ : 0.030, 95% C.I. (0.026, 0.034)

## wbgeweke output

```
. wbgeweke beta_1
Parameter: beta_1 first 10.0% (n=50) vs last 50.0% (n=250)
Means (se)      0.0003 ( 0.0000)  0.0003 ( 0.0000)
Autocorrelations 0.3736  0.4114
Mean Difference (se) 0.0000 ( 0.0000) z = 1.030 p = 0.3031
```

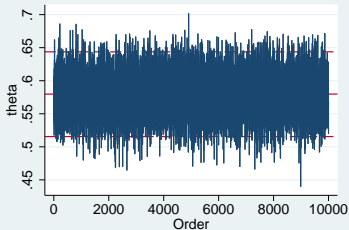
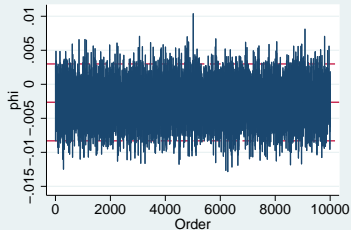
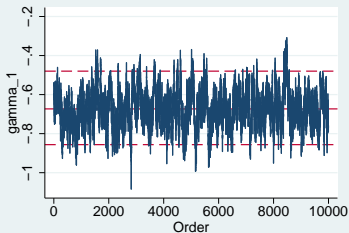
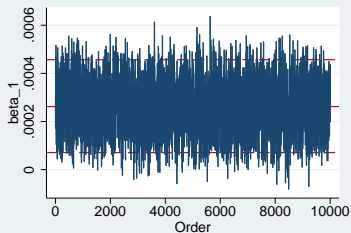
## wbgeweke output

```
. wbgeweke beta_1
Parameter: beta_1 first 10.0% (n=50) vs last 50.0% (n=250)
Means (se)      0.0003 ( 0.0000)    0.0003 ( 0.0000)
Autocorrelations 0.3736  0.4114
Mean Difference (se) 0.0000 ( 0.0000) z = 1.030 p = 0.3031
```

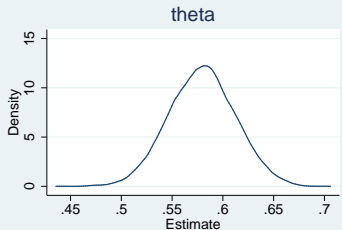
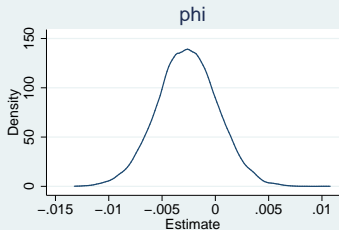
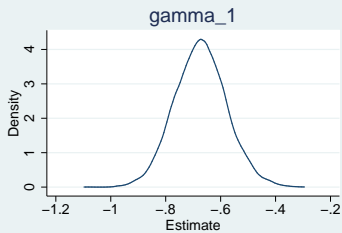
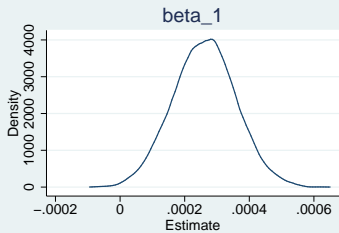
## wbdic output

```
. wbdic using winbugslog.txt
DIC statistics 1
DIC
Dbar = post.mean of -2logL; Dhat = -2LogL at post.mean of stochastic nodes
      Dbar    Dhat    pD    DIC
Y      4466.330    4393.470    72.861  4539.190
total  4466.330    4393.470    72.861  4539.190
```

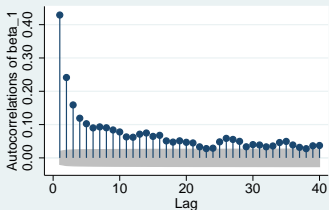
# wbtrace output



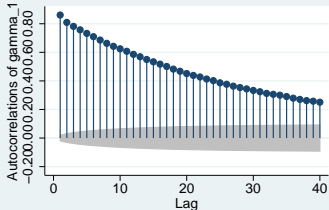
# wbdensity output



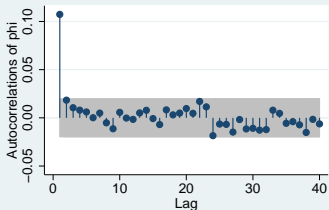
# wbac output



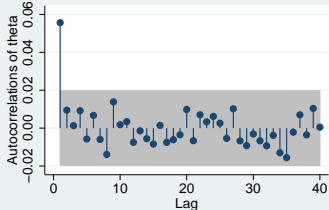
Bartlett's formula for MA(q) 95% confidence bands



Bartlett's formula for MA(q) 95% confidence bands



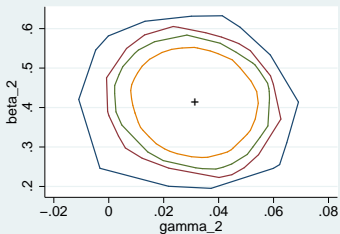
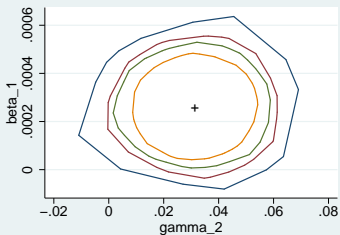
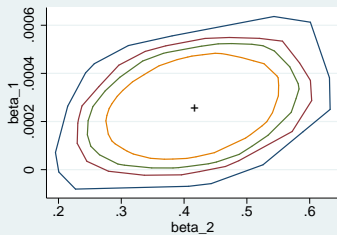
Bartlett's formula for MA(q) 95% confidence bands



Bartlett's formula for MA(q) 95% confidence bands



# wbhull output



# Summary

- WinBUGS - easy & flexible

# Summary

- WinBUGS - easy & flexible
- `winbugsfromstata` - data preparation, analysis of MCMC output, graphics

# Summary

- WinBUGS - easy & flexible
- `winbugsfromstata` - data preparation, analysis of MCMC output, graphics
- Prior distributions - *controversial*

# Summary

- WinBUGS - easy & flexible
- `winbugsfromstata` - data preparation, analysis of MCMC output, graphics
- Prior distributions - *controversial*
- Check complex Stata models - *vague* prior distributions

# Summary

- WinBUGS - easy & flexible
- `winbugsfromstata` - data preparation, analysis of MCMC output, graphics
- Prior distributions - *controversial*
- Check complex Stata models - *vague* prior distributions
- Fit complex models not possible in Stata

- Bayesian residuals and model checking [Lu et al., 2007]

- Bayesian residuals and model checking [Lu et al., 2007]
- Automate WinBUGS model statement



- Bayesian residuals and model checking [Lu et al., 2007]
- Automate WinBUGS model statement
- Mac users: WinBUGS runs under Darwin

- Bayesian residuals and model checking [Lu et al., 2007]
- Automate WinBUGS model statement
- Mac users: WinBUGS runs under Darwine
- OpenBUGS (version 3.0.1), WinBUGS (version 1.4.2)  
<http://mathstat.helsinki.fi/openbugs/>

# References



Goldstein, H., Rasbash, J., Yang, M., Woodhouse, G., Pan, H., Nuttall, D., and Thomas, S. (1993).  
A multilevel analysis of school examination results.  
*Oxford Review of Education*, 19(4):425–433.



Lu, G., Ades, A. E., Sutton, A. J., Cooper, N. J., Briggs, A. H., and Caldwell, D. M. (2007).  
Meta-analysis of mixed treatment comparisons at multiple follow-up times.  
*Statistics in Medicine*.  
in press.



Spiegelhalter, D. J., Thomas, A., Best, N., and Lunn, D. (2004).  
*WinBUGS User Manual, version 1.4.1*.  
MRC Biostatistics Unit, Cambridge, UK.



Thompson, J., Palmer, T., and Moreno, S. (2006).  
Bayesian Analysis in Stata using WinBUGS.  
*The Stata Journal*, 6(4):530–549.

## Acknowledgements



MRC Capacity Building PhD Studentship in Genetic Epidemiology