Bayesian Analysis with Stata: application to neonatal mortality in the UK

John Thompson
john.thompson@le.ac.uk

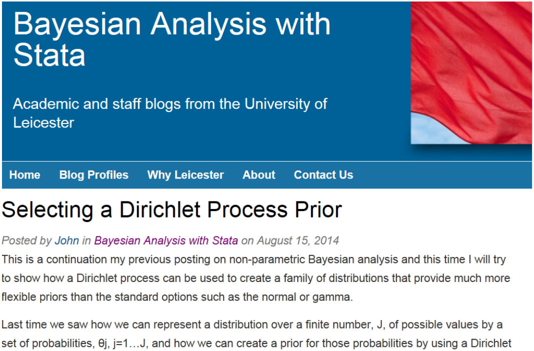
University of Leicester

12th September 2014

Book

Bayesian Analysis
With Stata
$p(\theta_1|y) = \int p(\theta_1, \theta_2|y)d\theta_2$

Blog



Bayesian Analysis with Stata

Academic and staff blogs from the University of Leicester

Selecting a Dirichlet Process Prior

Posted by John in Bayesian Analysis with Stata on August 15, 2014

This is a continuation my previous posting on non-parametric Bayesian analysis and this time I will try to show how a Dirichlet process can be used to create a family of distributions that provide much more flexible priors than the standard options such as the normal or gamma.

Last time we saw how we can represent a distribution over a finite number, J, of possible values by a set of probabilities, θj, j=1…J, and how we can create a prior for those probabilities by using a Dirichlet

http://staffblogs.le.ac.uk/bayeswithstata/

Pros and Cons of Bayes

Advantages

- Direct answers to research questions
- Purely model based: no ad hoc modifications
- Computation can be done by simulation

Pros and Cons of Bayes

Advantages

- Direct answers to research questions
- Purely model based: no ad hoc modifications
- Computation can be done by simulation

Disadvantages

- Bayes is unfamiliar to some people
- Users must specify their priors
- Computation by simulation can be slow
- Bayesian analysis is not yet a core part of Stata

Outline of the talk

- MCMC: Computation by simulation in Stata
- Neonatal mortality data
- Model the data for 2005
    - Stata
    - Mata
    - WinBUGS
- Model time trends 1999-2009; predict 2010

MCMC: Computation by Simulation

Researcher specifies:

Model: $p(y|\theta)$
Prior: $p(\theta)$

Posterior:

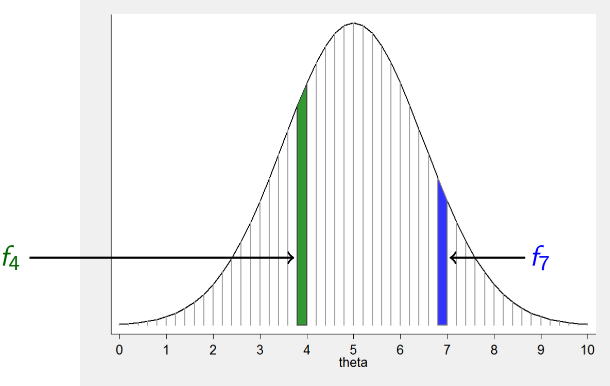
$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{\int p(y|\theta)p(\theta)d\theta} \propto p(y|\theta)p(\theta)$$

Approximate the posterior by simulating many values of θ

For example:
mean of simulations approximates mean of posterior

Simulating θ

Posterior distribution $p(\theta|y)$



$\theta =$ ....,4.0,7.8,5.2,...,2.1,7.0,5.5,4.0,..

4.0 and 7.0 must occur in the ratio f_4:f_7

Markov chain Monte Carlo

$\theta = \ldots.\ 4.0 \curvearrowright 7.8 \curvearrowright 5.2\ \ldots\ 2.1 \curvearrowright 7.0 \curvearrowright 7.0 \curvearrowright 4.0\ ..$

Either move to a new value or repeat the old value with transition probabilities t(old,new)

Markov chain Monte Carlo

$\theta = ...\ 4.0 \curvearrowright 7.8 \curvearrowright 5.2\ ...\ 2.1 \curvearrowright 7.0 \curvearrowright 7.0 \curvearrowright 4.0\ ..$

Either move to a new value or repeat the old value with transition probabilities t(old,new)

$$\begin{bmatrix} f_4 & f_7 \end{bmatrix} \begin{bmatrix} 1 - t(4,7) & t(4,7) \\ t(7,4) & 1 - t(7,4) \end{bmatrix} = \begin{bmatrix} f_4 & f_7 \end{bmatrix}$$

Markov chain Monte Carlo

$\theta = \ldots.\ 4.0 \curvearrowright 7.8 \curvearrowright 5.2\ \ldots\ 2.1 \curvearrowright 7.0 \curvearrowright 7.0 \curvearrowright 4.0\ ..$

Either move to a new value or repeat the old value with transition probabilities t(old,new)

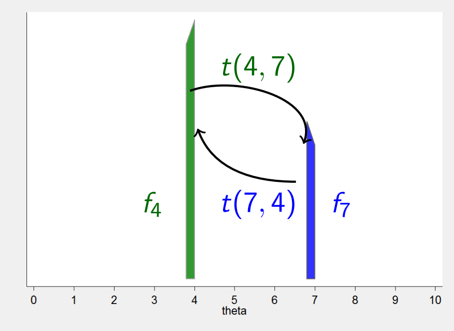
$$\begin{bmatrix} f_4 & f_7 \end{bmatrix} \begin{bmatrix} 1 - t(4,7) & t(4,7) \\ t(7,4) & 1 - t(7,4) \end{bmatrix} = \begin{bmatrix} f_4 & f_7 \end{bmatrix}$$

$$f_4[1 - t(4,7)] + f_7 t(7,4) = f_4$$

Detailed Balance: $f_4 t(4,7) = f_7 t(7,4)$

for all pairs of values

Metropolis-Hastings



P(propose 4 to 7) $= \pi(4,7)$ P(accept 4 to 7) $= a(4,7)$

Detailed Balance becomes,

$$f_4\pi(4,7)a(4,7) = f_7\pi(7,4)a(7,4)$$

e.g. set a(7,4)=1 and make,

$$a(4,7) = \frac{f_7\pi(7,4)}{f_4\pi(4,7)}$$

Multi-parameter models

Block updating

propose new values for all of the parameters and update in a single MH step

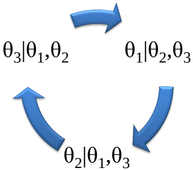
Multi-parameter models

Block updating

propose new values for all of the parameters and update in a single MH step

Gibbs sampling

update each parameter in turn keeping the others fixed at their current value


$\theta_3|\theta_1,\theta_2$
$\theta_1|\theta_2,\theta_3$
$\theta_2|\theta_1,\theta_3$

Bayesian Computing

Metropolis-Hastings: rejected moves lead to repeated values

$$\theta = \text{....,4.0,4.0,5.2,...,2.1,7.0,5.5,5.5,4.0,..}$$

Many repeats ⇒ slow convergence ⇒ a long chain

The trick is to choose proposals that,

- Move freely across the posterior
- Have a 'reasonable' chance of being accepted

Bayesian Computing

Metropolis-Hastings: rejected moves lead to repeated values

$$\theta = \text{....,4.0,4.0,5.2,...,2.1,7.0,5.5,5.5,4.0,..}$$

Many repeats ⇒ slow convergence ⇒ a long chain

The trick is to choose proposals that,

- Move freely across the posterior
- Have a 'reasonable' chance of being accepted

Bayesian computing = Designing efficient algorithms

- Inefficient algorithms can take days to run
- *Bayesian Analysis with Stata* presents guidelines and some programs
- WinBUGS, OpenBUGS, JAGS, Stan provide black-box solutions

A Simple Example

Data: twenty values from N(5,sd=2)

Researcher specifies:

Model: $p(y|\theta) = N(\theta, 2)$

Prior: $p(\theta) = N(4, 1)$

Posterior:

$$p(\theta|y) \propto \left[\prod exp(-0.125(y_i - \theta)^2\right] exp(-0.5(\theta - 4)^2)$$

log Posterior:

$$log\left[p(\theta|y)\right] = \texttt{constant} + \sum -0.125(y_i - \theta)^2 - 0.5(\theta - 4)^2$$

A Stata Program

```
set obs 20
gen y = rnormal(5,2)
local theta = 4
forvalues iter=1/100 {
    gen LogL = -0.125*(y-‘theta’)^2
    qui su LogL
    local logpost = r(sum) -0.5*(‘theta’-4)^2
    local newtheta = ‘theta’ + rnormal(0,0.25)
    qui replace LogL = -0.125*(y-‘newtheta’)^2
    qui su LogL
    local newlogpost = r(sum) -0.5*(‘newtheta’-4)^2
    if log(runiform()) < (‘newlogpost’ - ‘logpost’) local theta = ‘newtheta’
    di %6.2f ‘theta’
    drop LogL
}
```

A Stata Program

Simulate the data

```
set obs 20
gen y = rnormal(5,2)
local theta = 4
forvalues iter=1/100 {
    gen LogL = -0.125*(y-‘theta’)^2
    qui su LogL
    local logpost = r(sum) -0.5*(‘theta’-4)^2
    local newtheta = ‘theta’ + rnormal(0,0.25)
    qui replace LogL = -0.125*(y-‘newtheta’)^2
    qui su LogL
    local newlogpost = r(sum) -0.5*(‘newtheta’-4)^2
    if log(runiform()) < (‘newlogpost’ - ‘logpost’) local theta = ‘newtheta’
    di %6.2f ‘theta’
    drop LogL

}
```

A Stata Program

Initial value

```
set obs 20
gen y = rnormal(5,2)
local theta = 4
forvalues iter=1/100 {
    gen LogL = -0.125*(y-‘theta’)^2
    qui su LogL
    local logpost = r(sum) -0.5*(‘theta’-4)^2
    local newtheta = ‘theta’ + rnormal(0,0.25)
    qui replace LogL = -0.125*(y-‘newtheta’)^2
    qui su LogL
    local newlogpost = r(sum) -0.5*(‘newtheta’-4)^2
    if log(runiform()) < (‘newlogpost’ - ‘logpost’) local theta = ‘newtheta’
    di %6.2f ‘theta’
    drop LogL
}
```

A Stata Program

Evaluate log-posterior

```
set obs 20
gen y = rnormal(5,2)
local theta = 4
forvalues iter=1/100 {
    gen LogL = -0.125*(y-‘theta’)^2
    qui su LogL
    local logpost = r(sum) -0.5*(‘theta’-4)^2
    local newtheta = ‘theta’ + rnormal(0,0.25)
    qui replace LogL = -0.125*(y-‘newtheta’)^2
    qui su LogL
    local newlogpost = r(sum) -0.5*(‘newtheta’-4)^2
    if log(runiform()) < (‘newlogpost’ - ‘logpost’) local theta = ‘newtheta’
    di %6.2f ‘theta’
    drop LogL

}
```

A Stata Program

```
set obs 20
gen y = rnormal(5,2)
local theta = 4
forvalues iter=1/100 {
    gen LogL = -0.125*(y-‘theta’)^2
    qui su LogL
    local logpost = r(sum) -0.5*(‘theta’-4)^2
    local newtheta = ‘theta’ + rnormal(0,0.25)
    qui replace LogL = -0.125*(y-‘newtheta’)^2
    qui su LogL
    local newlogpost = r(sum) -0.5*(‘newtheta’-4)^2
    if log(runiform()) < (‘newlogpost’ - ‘logpost’) local theta = ‘newtheta’
    di %6.2f ‘theta’
    drop LogL
}
```

Make a proposal

A Stata Program

```
set obs 20
gen y = rnormal(5,2)
local theta = 4
forvalues iter=1/100 {
    gen LogL = -0.125*(y-‘theta’)^2
    qui su LogL
    local logpost = r(sum) -0.5*(‘theta’-4)^2
    local newtheta = ‘theta’ + rnormal(0,0.25)
    qui replace LogL = -0.125*(y-‘newtheta’)^2
    qui su LogL
    local newlogpost = r(sum) -0.5*(‘newtheta’-4)^2
    if log(runiform()) < (‘newlogpost’ - ‘logpost’) local theta = ‘newtheta’
    di %6.2f ‘theta’
    drop LogL
}
```

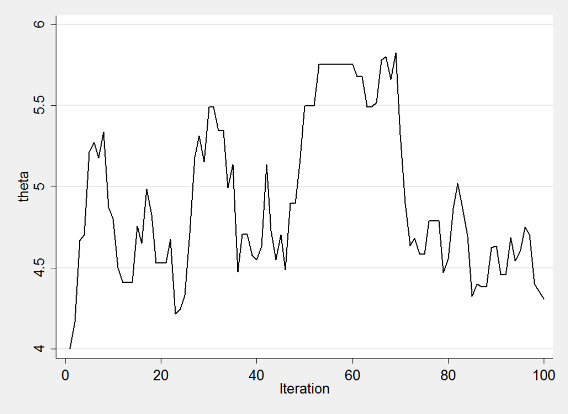
Evaluate new log-posterior

A Stata Program

```
set obs 20
gen y = rnormal(5,2)
local theta = 4
forvalues iter=1/100 {
    gen LogL = -0.125*(y-‘theta’)^2
    qui su LogL
    local logpost = r(sum) -0.5*(‘theta’-4)^2
    local newtheta = ‘theta’ + rnormal(0,0.25)
    qui replace LogL = -0.125*(y-‘newtheta’)^2
    qui su LogL
    local newlogpost = r(sum) -0.5*(‘newtheta’-4)^2
    if log(runiform()) < (‘newlogpost’ - ‘logpost’) local theta = ‘newtheta’
    di %6.2f ‘theta’
    drop LogL
}
```

Accept or Reject

Results



Simpler Code

```
set obs 20
gen y = rnormal(5,2)

program logpost
    args logp b
    local theta = ‘b’[1,1]
    scalar ‘logp’ = 0
    logdensity normal ‘logp’ y ‘theta’ 2
    logdensity normal ‘logp’ ‘theta’ 4 1
end

matrix b = 4
mcmcrun logpost b using temp.csv, replace ///
    sampler(mhsnorm , sd(0.25)) par(theta) burn(100) update(5000)
insheet using temp.csv, comma clear
mcmctrace theta
```

Simpler Code

Parameters in row matrix b

Log-posterior returned in scalar logp

```
set obs 20
gen y = rnormal(5,2)

program logpost
    args logp b
    local theta = ‘b’[1,1]
    scalar ‘logp’ = 0
    logdensity normal ‘logp’ y ‘theta’ 2
    logdensity normal ‘logp’ ‘theta’ 4 1
end

matrix b = 4
mcmcrun logpost b using temp.csv, replace ///
    sampler(mhsnorm , sd(0.25)) par(theta) burn(100) update(5000)
insheet using temp.csv, comma clear
mcmctrace theta
```

Simpler Code

```
set obs 20
gen y = rnormal(5,2)

program logpost
    args logp b
    local theta = ‘b’[1,1]
    scalar ‘logp’ = 0
    logdensity normal ‘logp’ y ‘theta’ 2
    logdensity normal ‘logp’ ‘theta’ 4 1
end

matrix b = 4
mcmcrun logpost b using temp.csv, replace ///
    sampler(mhsnorm , sd(0.25)) par(theta) burn(100) update(5000)
insheet using temp.csv, comma clear
mcmctrace theta
```

logdensity knows the formulae

for standard distributions

Simpler Code

```
set obs 20
gen y = rnormal(5,2)

program logpost
    args logp b
    local theta = ‘b’[1,1]
    scalar ‘logp’ = 0
    logdensity normal ‘logp’ y ‘theta’ 2
    logdensity normal ‘logp’ ‘theta’ 4 1
end

matrix b = 4
mcmcrun logpost b using temp.csv, replace ///
    sampler(mhsnorm , sd(0.25)) par(theta) burn(100) update(5000)
insheet using temp.csv, comma clear
mcmctrace theta
```

mcmcrun creates the chain

and saves the values

Simpler Code

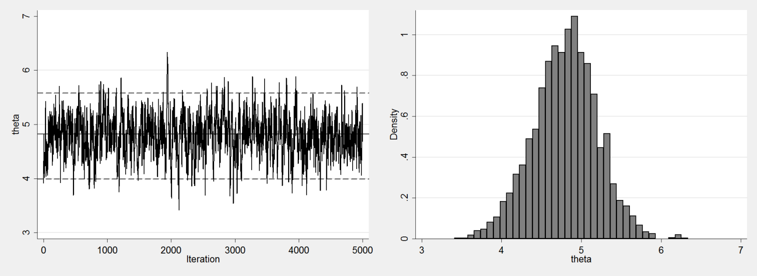
```
set obs 20
gen y = rnormal(5,2)

program logpost
    args logp b
    local theta = ‘b’[1,1]
    scalar ‘logp’ = 0
    logdensity normal ‘logp’ y ‘theta’ 2
    logdensity normal ‘logp’ ‘theta’ 4 1
end

matrix b = 4
mcmcrun logpost b using temp.csv, replace ///
    sampler(mhsnorm , sd(0.25)) par(theta) burn(100) update(5000)
insheet using temp.csv, comma clear
mcmctrace theta
```

mcmctrace plots the chain

Results


theta
3
4
5
6
7
0
1000
2000
3000
4000
5000
Iteration
Density
0
.2
.4
.6
.8
1
3
4
5
6
7
theta

Robust Model

```
set obs 20
gen y = rnormal(5,2)
qui replace y = 12 in 10
program logpost
    args logp b
    local theta = ‘b’[1,1]
    scalar ‘logp’ = 0
    logdensity t ‘logp’ y ‘theta’ 2 4
    logdensity normal ‘logp’ ‘theta’ 4 1
end
matrix b = 4
mcmcrun logpost b using temp.csv, replace ///
sampler(mhsnorm , sd(0.25)) par(theta) burn(100) update(5000)
insheet using temp.csv, comma clear
mcmctrace theta
```

Programs introduced in the book

Family	Example	Purpose
	logdensity	Calculate log-likelihoods & log-priors
mhs	mhsnorm	Various Metropolis-Hastings samplers
mcmc	mcmctrace	Run or inspect an MCMC analysis
gbs	gbsslice	Gibbs samplers
wbs	wbsrun	Communication with WinBUGS

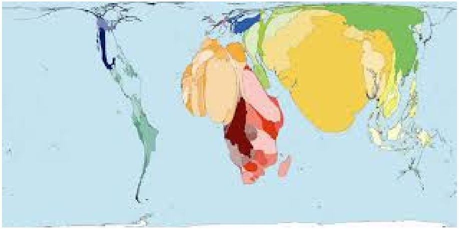
Neonatal Mortality in the UK

Mortality during the first 28 days of life following a live birth

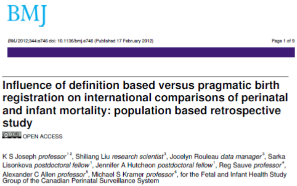
Often divided into

Early Neonatal (0-7 days): pregnancy

Late Neonatal (8-27days): environmental



Why do UK & USA do poorly?



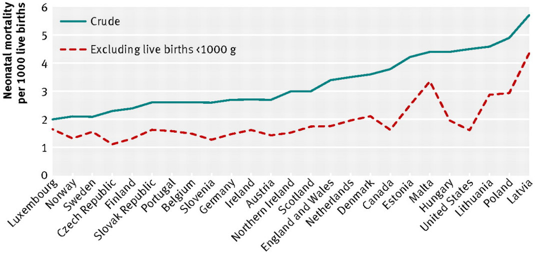
BMJ

BMJ 2012;344:e746 doi: 10.1136/bmj.e746 (Published 17 February 2012) Page 1 of 9

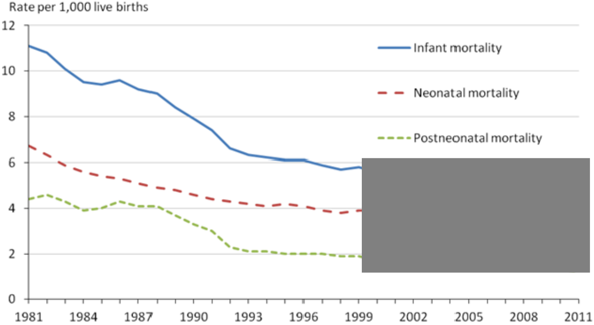
Influence of definition based versus pragmatic birth registration on international comparisons of perinatal and infant mortality: population based retrospective study

OPEN ACCESS

K S Joseph *professor*[1 2], Shiliang Liu *research scientist*[3], Jocelyn Rouleau *data manager*[3], Sarka Lisonkova *postdoctoral fellow*[1], Jennifer A Hutcheon *postdoctoral fellow*[1], Reg Sauve *professor*[4], Alexander C Allen *professor*[5], Michael S Kramer *professor*[6], for the Fetal and Infant Health Study Group of the Canadian Perinatal Surveillance System

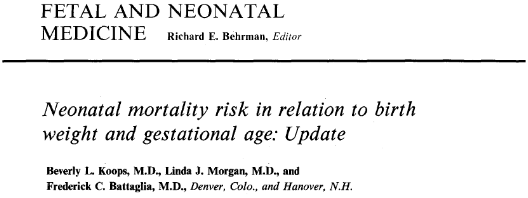


UK Neonatal Mortality 1999-2010



http://www.ons.gov.uk/ons/rel/vsob1/child-mortality-statistics
–childhood–infant-and-perinatal/2011/sty-infant-mortality.html

Prior knowledge

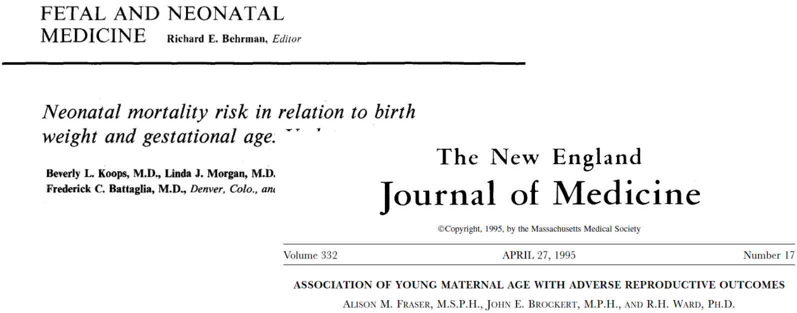


FETAL AND NEONATAL MEDICINE **Richard E. Behrman,** *Editor*

Neonatal mortality risk in relation to birth weight and gestational age: Update

Beverly L. Koops, M.D., Linda J. Morgan, M.D., and Frederick C. Battaglia, M.D., *Denver, Colo., and Hanover, N.H.*

Prior knowledge



FETAL AND NEONATAL MEDICINE Richard E. Behrman, *Editor*

Neonatal mortality risk in relation to birth weight and gestational age.

Beverly L. Koops, M.D., Linda J. Morgan, M.D.
Frederick C. Battaglia, M.D., *Denver, Colo., an*

The New England Journal of Medicine

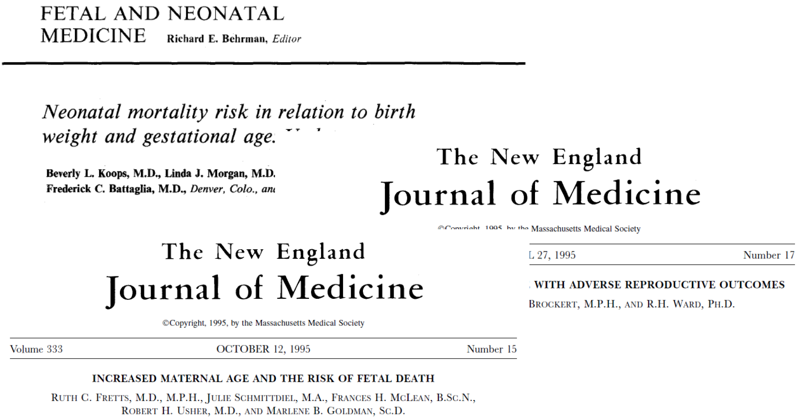
©Copyright, 1995, by the Massachusetts Medical Society

Volume 332 APRIL 27, 1995 Number 17

ASSOCIATION OF YOUNG MATERNAL AGE WITH ADVERSE REPRODUCTIVE OUTCOMES

Alison M. Fraser, M.S.P.H., John E. Brockert, M.P.H., and R.H. Ward, Ph.D.

Prior knowledge



FETAL AND NEONATAL MEDICINE Richard E. Behrman, *Editor*

Neonatal mortality risk in relation to birth weight and gestational age.

Beverly L. Koops, M.D., Linda J. Morgan, M.D. Frederick C. Battaglia, M.D., *Denver, Colo., an*

The New England Journal of Medicine

©Copyright, 1995, by the Massachusetts Medical Society

L 27, 1995 Number 17

WITH ADVERSE REPRODUCTIVE OUTCOMES

BROCKERT, M.P.H., AND R.H. WARD, PH.D.

The New England Journal of Medicine

©Copyright, 1995, by the Massachusetts Medical Society

Volume 333 OCTOBER 12, 1995 Number 15

INCREASED MATERNAL AGE AND THE RISK OF FETAL DEATH

RUTH C. FRETTS, M.D., M.P.H., JULIE SCHMITTDIEL, M.A., FRANCES H. MCLEAN, B.SC.N., ROBERT H. USHER, M.D., AND MARLENE B. GOLDMAN, SC.D.

Model

deaths $\sim$ Poisson(Rate per 1000 x Births/1000)

log(Rate) $= \mu + \alpha[\mathit{age}] + \beta[\mathit{bwt}]$
Constraint: α[20-24] $= \beta$[3000-3499] $= 0$

μ represents the log rate in the baseline group
α represents the log relative rate (baseline:20-24 years)
β represents the log relative rate (baseline: 3000-3499 grams)

Model

deaths $\sim$ Poisson(Rate per 1000 x Births/1000)

log(Rate) $= \mu + \alpha[age] + \beta[bwt]$
Constraint: α[20-24] $= \beta$[3000-3499] $=$ 0

μ represents the log rate in the baseline group
α represents the log relative rate (baseline:20-24 years)
β represents the log relative rate (baseline: 3000-3499 grams)

ML analysis in Stata

```
. glm deaths ib2.age ib6.bwt , fam(poi) off(lnBirths)
```

My Priors

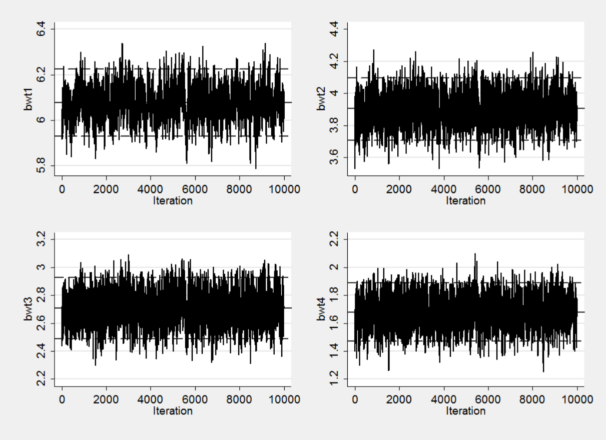
Maternal Age	Relative Rate	α	Birth Weight	Relative Rate	β
<20	2 (1,4)	0.69 (0.34)	<1000	100(10,1000)	4.61 (1.15)
20-24	1	0.0	1000-1499	10(2,50)	2.30 (0.80)
25-29	1(0.5,2)	0.00 (0.34)	1500-1999	3(1,9)	1.10 (0.55)
30-34	1(0.5,2)	0.00 (0.34)	2000-2499	1(0.5,2)	0.00 (0.34)
35-39	1(0.5,2)	0.00 (0.34)	2500-2999	1(0.5,2)	0.00 (0.34)
40+	2(1,4)	0.69 (0.34)	3000-3499	1	0.0
			3500-3999	1(0.5,2)	0.00 (0.34)
			4000+	2(1,4)	0.69 (0.34)

Mortality in baseline category
0.25 per 1,000 births (0.08,0.75) μ -1.39(0.55)

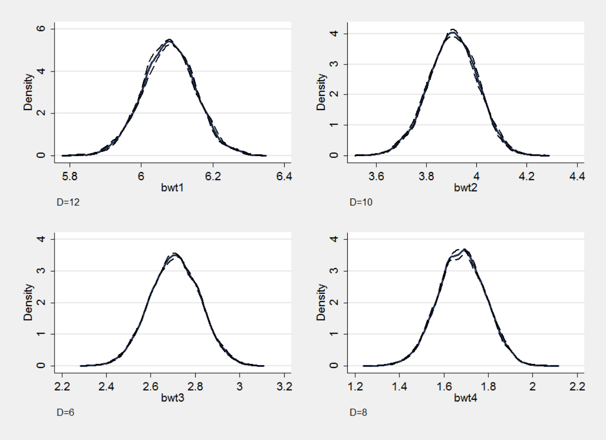
log-posterior(slow and inefficient)

```
program logpost
    args logp b
    tempvar lnMU j MU
    gen ‘lnMU’ = lnBirths + ‘b’[1,1]
    replace ‘lnMU’ = ‘lnMU’ + ‘b’[1,b] if bwt != 6
    replace ‘lnMU’ = ‘lnMU’ + ‘b’[1,a] if age != 2
    gen ‘MU’ = exp(‘lnMU’)
    scalar ‘logp’ = 0
    logdensity poisson ‘logp’ deaths ‘MU’
    logdensity normal ‘logp’ ‘b’[1,1] -1.39 0.55
    logdensity normal ‘logp’ ‘b’[1,2] 4.61 1.15
    logdensity normal ‘logp’ ‘b’[1,3] 2.30 0.80
    logdensity normal ‘logp’ ‘b’[1,4] 1.10 0.55
    logdensity normal ‘logp’ ‘b’[1,5] 0 0.34
    logdensity normal ‘logp’ ‘b’[1,6] 0 0.34
    logdensity normal ‘logp’ ‘b’[1,7] 0 0.34
    logdensity normal ‘logp’ ‘b’[1,8] 0.69 0.34
    logdensity normal ‘logp’ ‘b’[1,9] 0.69 0.34
    logdensity normal ‘logp’ ‘b’[1,10] 0 0.34
    logdensity normal ‘logp’ ‘b’[1,11] 0 0.34
    logdensity normal ‘logp’ ‘b’[1,12] 0 0.34
    logdensity normal ‘logp’ ‘b’[1,13] 0.69 0.34
end
```

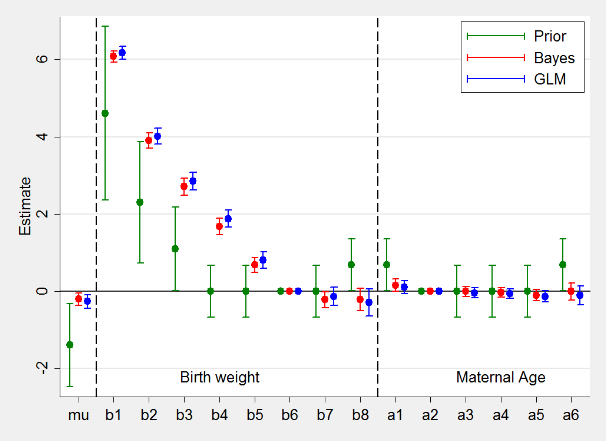
Convergence: Trace Plots



Convergence: Section Plots



Parameter Estimates



Prediction

Predictive distribution of new values y* given previous values y

$$p(y^*|y) = \int p(y^*|\theta)p(\theta|y)d\theta$$

approximate the integral in an MCMC algorithm

simulate new data y* from $p(y^*|\theta)$ using the current θ

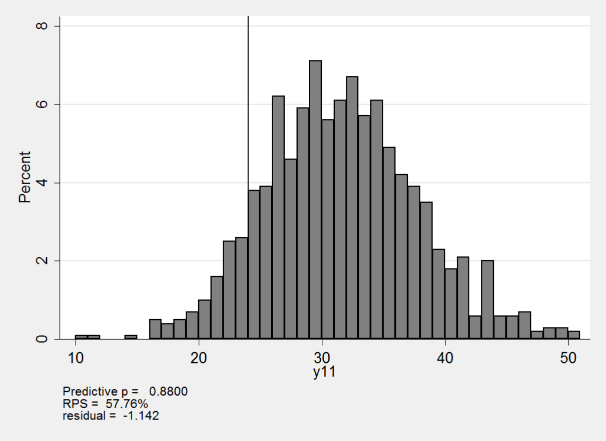
Predictions

```
program y , rclass
    args b

    tempvar lnMU j MU y
    tempname pd
    gen ‘lnMU’ = lnBirths + ‘b’[1,1]
    qui replace ‘lnMU’ = ‘lnMU’ + ‘b’[1,b] if bwt != 6
    qui replace ‘lnMU’ = ‘lnMU’ + ‘b’[1,a] if age != 2
    gen ‘MU’ = exp(‘lnMU’)
    gen ‘y’ = rpoisson(‘MU’)
    mkmat ‘y’ , matrix(‘pd’)
    matrix ‘pd’ = ‘pd”
    return matrix pred = ‘pd’
end

mcmcrun logpost b using temp.csv, replace ///
  samplers( 13(mhsnorm , sd(0.1)) ) burn(1000) updates(20000) ///
   thin(20) par(mu bwt1-bwt5 bwt7 bwt8 age1 age3-age6) pred(y)
```

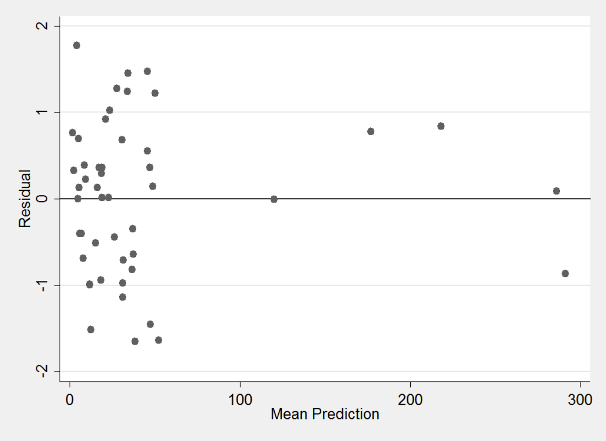
Predictive Distribution of Observation 11



Predictive p = 0.8800
RPS = 57.76%
residual = -1.142

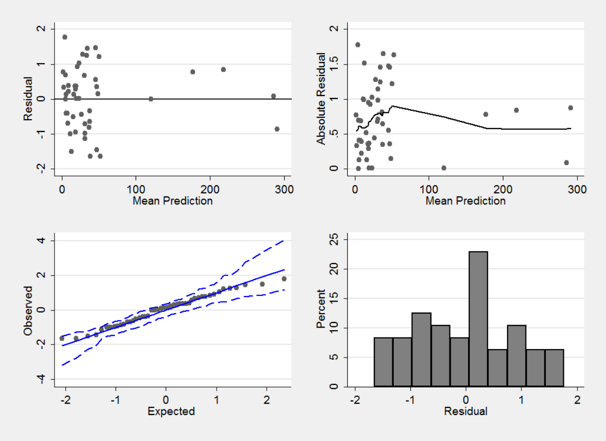
```
mcmccheck , d(deaths[11]) df(paed2005.dta) p(y11) pf(mcmc2005.dta)
gopt(width(1))
```

Residuals vs Fit



mcmccheck , d(deaths) df(paed2005.dta) p(y) pf(mcmc2005.dta)

Summary Plot



```
mcmccheck , d(deaths) df(paed2005.dta) p(y) pf(mcmc2005.dta) plot(summary)
```

Mata Analysis

```
set matastrict on
mata:
mata clear
real scalar logpost(real matrix X,real rowvector b,real scalar ipar)
{
  real colvector lnM
  real scalar i

  lnM = X[,2] : + b[1]
  for(i=1;i<=48;i++) {
    if( X[i,5] ! = 2 ) lnM[i] = lnM[i] + b[X[i,3]]
    if( X[i,6] ! = 6 ) lnM[i] = lnM[i] + b[X[i,4]]
  }
  return( sum(X[,1]:*lnM :- exp(lnM)) -1.652893*(b[1]-1.39)*(b[1]-1.39)
  -0.378072*(b[2]-4.61)*(b[2]-4.61)-0.78125*(b[3]-2.3)*(b[3]-2.3)
  -1.652893*(b[4]-1.1)*(b[4]-1.1)-4.32526*(b[5]*b[5]+b[6]*b[6]+b[7]*b[7]
  +(b[8]-0.69)*(b[8]-0.69)+(b[9]-0.69)*(b[9]-0.69)+b[10]*b[10]+b[11]*b[11]
  +b[12]*b[12]+(b[13]-0.69)*(b[13]-0.69)) )
}
end
```

Mata Analysis

```
use paediatric.dta, clear
keep if year == 2005
drop if bwt == 9
gen Rate = 1000*deaths/births
gen lnBirths = log(births/1000)
gen b = bwt + (bwt<6)
gen a = 7 + age + (age<2)
matrix theta = (-1.5,4.5,2.5,1,0,0,0,1,1,0,0,0,1)
matrix s = J(1,13,0.1)
mcmcrun logpost X theta using temp1.csv, samp( 13(mhsnorm , sd(s)) ) ///
  burn(1000) adapt update(50000) thin(5) jpost dots(0) ///
  par(mu bwt1-bwt5 bwt7 bwt8 age1 age3-age6) replace ///
  data(X=(deaths lnBirths a b age bwt) theta s) mata
```

WinBUGS Analysis

- Write text files containing:
  - The model file (c.f. logpost)
  - The data in WinBUGS format
  - The initial values
  - The script (batch commands to control the fit)
- Run the script file
  - Results stored in a text file
- Read results into Stata
- Process the results

WinBUGS model file

```
model {
  for( i in 1:6 ) {
    for( j in 1:8 ) {
      log(m[i,j]) <- mu + alpha[i] + beta[j] + LB[i,j]
      D[i,j] ~ dpois(m[i,j])
    }
  }
  mu ~ dnorm(-1.39,3.306)
  alpha[1] ~ dnorm(0.69,8.651)
  alpha[2] <- 0
  alpha[3] ~ dnorm(0.0,8.651)
  alpha[4] ~ dnorm(0.0,8.651)
  alpha[5] ~ dnorm(0.0,8.651)
  alpha[6] ~ dnorm(0.69,8.651)
  beta[1] ~ dnorm(4.61,0.756)
  beta[2] ~ dnorm(2.30,1.563)
  beta[3] ~ dnorm(1.10,3.306)
  beta[4] ~ dnorm(0.0,8.651)
  beta[5] ~ dnorm(0.0,8.651)
  beta[6] <- 0
  beta[7] ~ dnorm(0.0,8.651)
  beta[8] ~ dnorm(0.69,8.651)
}
```

WinBUGS data file

Store data logBirths and deaths in Stata matrices LB and D

```
. wbslist (matrix LB D) using data.txt , replace
```

```
list( LB=structure(.Data=c(
-1.248, -0.931, -0.224, 0.901, 2.218, 2.840, 2.420, 1.152,
-0.496, -0.076, 0.646, 1.840, 3.173, 3.812, 3.468, 2.370,
-0.197, 0.153, 0.914, 2.043, 3.343, 4.081, 3.840, 2.866,
-0.200, 0.240, 0.967, 2.090, 3.366, 4.186, 4.037, 3.151,
-0.601, -0.172, 0.538, 1.530, 2.750, 3.565, 3.449, 2.610,
-2.017, -1.483, -0.863, 0.143, 1.273, 2.016, 1.839, 1.055),.Dim = c(6,8)),
D=structure(.Data=c(120,20,8,7,19,17,10,3,
234,33,29,35,27,35,26,6,
272,50,35,42,57,50,25,8,
288,60,27,44,50,39,31,13,
190,24,21,14,23,24,19,6,
35,10,7,5,6,5,8,3),.Dim = c(6,8)))
```

WinBUGS initial values file

```
. wbslist (mu=-1.4,alpha=c(0.7,NA,0,0,0,0.7), ///
    beta=c(4.6,2.3,1.1,0,0,NA,0,0.7)) ///
    using init.txt , replace
```

```
list( mu=-1.4,alpha=c(0.7,NA,0,0,0,0.7),
 beta=c(4.6,2.3,1.1,0,0,NA,0,0.7))
```

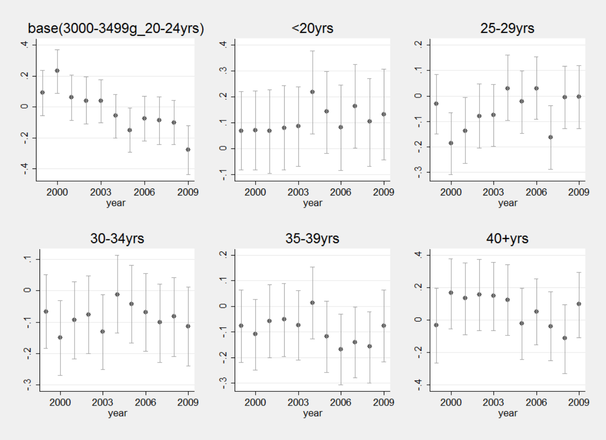
WinBUGS script file

```
. wbsscript using script.txt, model(model.txt) ///
    data(data.txt) init(init.txt) burn(1000) update(10000) ///
    set(mu alpha beta) coda(NMR) log(NMR.log) replace
```

```
display(’log’)
check(’E:/StataUsers/Data/model.txt’)
data(’E:/StataUsers/Data/data.txt’)
compile(1)
inits(1,’E:/StataUsers/Data/init.txt’)
gen.inits()
refresh(100000)
update(1000)
set(’mu’)
set(’alpha’)
set(’beta’)
update(10000)
coda(*,’E:/StataUsers/Data/NMR’)
save(’E:/StataUsers/Data/NMR.log’)
```

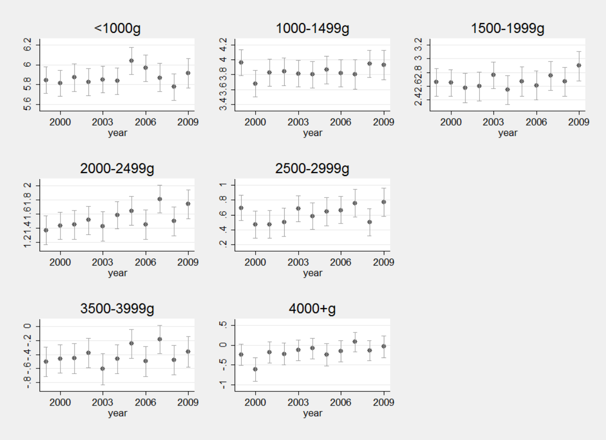
Analyses for 1999-2009

Baseline and age effects



Analyses for 1999-2009

Birth weight effects

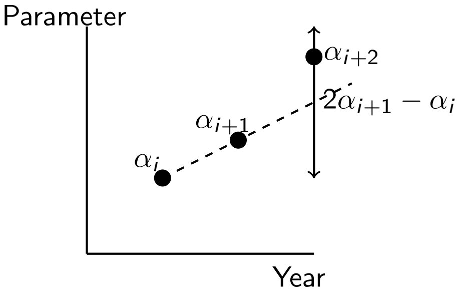


Smoothing the parameters

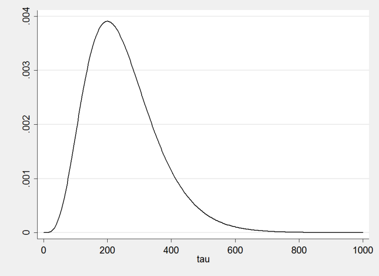
The parameter estimates show improbable jumps between years and would be more credible if the trends were smoother.

$$\alpha_{i+2} \sim N(2\alpha_{i+1} - \alpha_i, \tau)$$

τ acts as a smoothing parameter



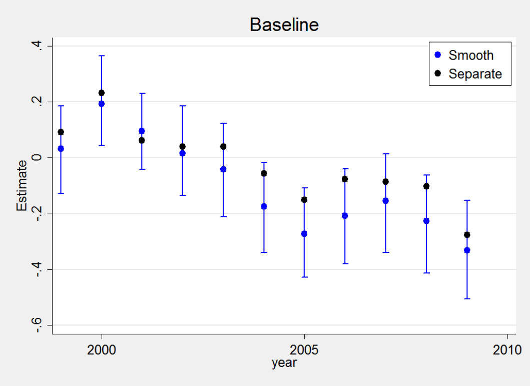
Prior for tau



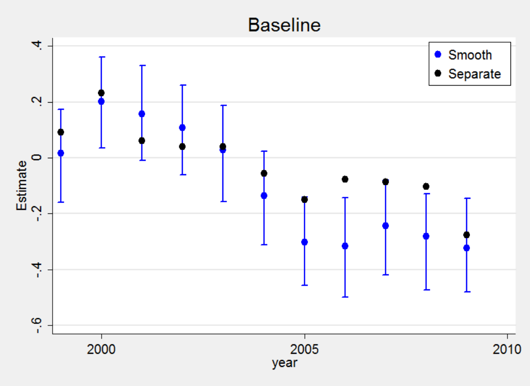
Prior G(5,50), prior mean = 250, prior sd = 112

τ=250 => sd=0.063

Smoothed Estimates



Stronger smoothing



Prior G(50,5), prior mean 250, prior sd=35

Average predictions for 2010

mean prediction (actual number)

	AGE(yrs)					
BWT(g)	< 20	20−	25−	30−	35−	40+
< 1000	90(72)	217(205)	317(257)	283(278)	210(173)	84(49)
1000−	13(17)	32(21)	35(46)	43(31)	33(26)	12(11)
1500−	11(4)	30(25)	42(28)	40(34)	29(17)	12(8)
2000−	12(11)	33(24)	45(37)	37(34)	27(22)	10(8)
2500−	17(10)	49(38)	65(51)	53(44)	34(27)	11(8)
3000−	12(10)	35(37)	50(48)	42(49)	26(17)	8(5)
3500−	6(3)	19(21)	30(13)	27(19)	17(15)	5(5)
4000+	2(0)	8(6)	15(13)	14(12)	9(10)	3(2)

Slightly better predictions if τ is allowed to differ with the parameter

Summary

- The Bayesian approach has many advantages
- MCMC is straightforward but can be slow
- MCMC allows flexibility in modelling (cf ml)
- MCMC is practical in Stata for small problems
- Larger problems require Mata or WinBUGS
- Other issues covered in the book include:
  - Convergence checking
  - Gibbs sampling
  - Model comparison & tests
  - Validation of software
  - Writing Bayesian programs for general use

Recommendations

- Stata should provide facilities for communication with other software, e.g.
  - Stata $\longleftrightarrow$ WinBUGS
  - Stata $\longleftrightarrow$ R
- Stata needs to be able to handle datasets that are too large for the Editor
- Bayesian analysis should be fully integrated into Stata
- It would be possible to re-write WinBUGS/JAGS in Mata
- This is a job for StataCorp