

Review of WINBUGS

by

Harvey Goldstein

Institute of Education

University of London

h.goldstein@ioe.ac.uk

1. Introduction

This review is based upon the beta release of WINBUGS 1.4 which has several enhancements over version 1.3.

WINBUGS, a Bayesian MCMC package, is distributed freely and is the result of many years of development by a team of statisticians and programmers at the Medical research Council Biostatistics Research Unit in Cambridge (<http://www.mrc-bsu.cam.ac.uk/bugs/>). Models are represented by a flexible language, and there is also a graphical feature, DOODLEBUGS, that allows users to specify their model in terms of a directed graph. For complex models the latter can be extremely useful, but we will not use it in this review since all the models dealt with are relatively straightforwardly expressed using the command language.

The topic of Bayesian estimation using Markov Chain Monte Carlo (MCMC) methods has an extensive literature and users are well advised to introduce themselves to this before using these methods. The WINBUGS manual provides a good starting place. In this review we shall not be exploring the full range of possibilities provided by WINBUGS, although, where appropriate these will be referred to. We shall give examples of WINBUGS code where this is illuminative, but since this review is not a guide to using WINBUGS, we will not be exhaustive.

Users of simulation based methods such as MCMC should remember that, unlike maximum likelihood procedures, there is a stochastic element to the estimates obtained and that procedures for judging convergence and accuracy are more complex and a certain amount of judgement and experience is involved. The WINBUGS manual is a useful source here.

The advantage of MCMC estimation, apart from the obvious one for those who wish to adopt a Bayesian approach, is that, like bootstrapping, it provides accurate estimates for parameter quantiles. One disadvantage is the time taken to obtain results and often the use of likelihood based methods for initial model exploration prior to using MCMC will be efficient. Also, there are still many issues whose solution is less than clearcut. For multilevel models, one of these is the issue of choice of priors for variance and covariance parameters; we have adopted the WINBUGS gamma and Wishart defaults, but see Browne (1998) for a further discussion. Another issue is that of model comparison and we have used the Bayesian Deviance Information Criterion for this (Spiegelhalter et al. 2002).

1.1 Data interface

Data is input to WINBUGS as three components that can be in the same or different files. The first is a series of model specification instructions, the second contains the observed data and the third contains a set of starting values. In fact WINBUGS will itself obtain starting values if required, but generally it is better for the user to specify these. We shall illustrate the use of these components in the first example below. WINBUGS runs under various versions of windows. Older versions will also run under other operating systems such as LINUX – see the WINBUGS website for more information. While WINBUGS 1.4 has many useful graphical tools, there is also a suite of routines, CODA, that can provide additional plotting features for WINBUGS output in SPLUS or R. The principal features will be illustrated in our first example.

1.2 Model types

The range of model types that can be fitted in WINBUGS is very large. The language allows a wide variety of linear and nonlinear model forms, including the standard set of generalised linear models. For the random effects a variety of distributional forms can be specified, including the multivariate Normal and T-distributions. Spatial models and latent variable models can be fitted. In addition a range of prior distributions is available with standard defaults. All of the models in this set of reviews can be fitted in WINBUGS.

2. Examples

2.1 Two and three level Normal models

The first model we fit is a variance components 2-level model using the [EXAM](#) data set and we shall use it to illustrate the use of the WINBUGS modelling language.

The model is

$$\begin{aligned} y_{ij} &= (X\beta)_{ij} + u_{0j} + e_{0ij} \\ u_{0j} &\sim N(0, \sigma_{u0}^2) \quad e_{0ij} \sim N(0, \sigma_{e0}^2) \end{aligned} \tag{1}$$

where i, j respectively index level 1 and level 2 and $(X\beta)_{ij}$ represents the fixed part contribution for the i, j - th record. The fixed part here includes the intercept (β_0) and four covariates, $\beta_1, \beta_2, \beta_3, \beta_4$, being LRT score, gender, school gender composition (2 dummy variables; boys school and girls school contrasted with base category, mixed school). We discuss priors below.

The WINBUGS model specification code for fitting this model is as follows:

```

model
{
# Level 1 definition
for(i in 1:N) {
normexam[i] ~ dnorm(mu[i],tau)
mu[i]<- beta[1] * cons[i]
+ beta[2] * standlrt[i]
+ beta[3] * gender[i]
+ beta[4] * boysch[i]
+ beta[5] * girlsch[i]
+ u2[school[i]] * cons[i]
}
# Higher level definitions
for (j in 1:n2) {
u2[j] ~ dnorm(0,tau.u2)
}
# Priors for fixed effects
for (k in 1:5) { beta[k] ~ dflat() }
# Priors for random terms
tau ~ dgamma(0.001,0.001)
sigma2 <- 1/tau
tau.u2 ~ dgamma(0.001,0.001)
sigma2.u2 <- 1/tau.u2
}

```

Notes:

- Comments (starting with #) are inserted for ease of reading
- Line 5 specify that the response is Normal, and has a mean $\mu(i)$ and *precision* τ . The precision is simply the inverse of the variance and this relationship as specified in the penultimate line by the *logical or deterministic* (i.e. non-stochastic) relationship $\sigma^2.u2 <- 1/\tau.u2$.
- Lines 6 – 10 specify $\mu(i)$ as a linear additive function of the intercept and four covariates and line 11 adds the set of random effects where $school(i)$ is the school ID (1...65) for record i and is multiplied by $u2$ which is a random variable (see below). This product is also multiplied by $cons(i)$ which here is strictly unnecessary but is retained for consistency with general random coefficient models where the explanatory variable associated with a random coefficient is required.
- Note that constants N , $n2$ used here have values that are input with the data (see later), but we could have specified the values, 4059, 65, instead.
- Under priors for the fixed effects we use a ‘flat’ i.e. uniform prior across the whole real line, for each regression (fixed effect) coefficient. An alternative, ‘proper prior’ which is functionally equivalent is a Normal with a very large variance, e.g. $N(0,1*10^6)$ which is used in WINBUGS tutorials.
- Under priors for random terms we give the precisions Gamma (0.001,0.001) priors – the WINBUGS default. Other choices are available – see Browne (1998) for more discussion.
- Under ‘higher level definitions’ the distribution of $u2(j)$ ($j=1\dots n2$) is specified as Normal mean 0 and precision $\tau.u2$. The final line relates the level 2 variance to this precision parameter.
- We can also specify further logically derived parameters. For example the variance partition coefficient (VPC – Goldstein et al, 2002) which is here

plots for the parameter posterior distributions.

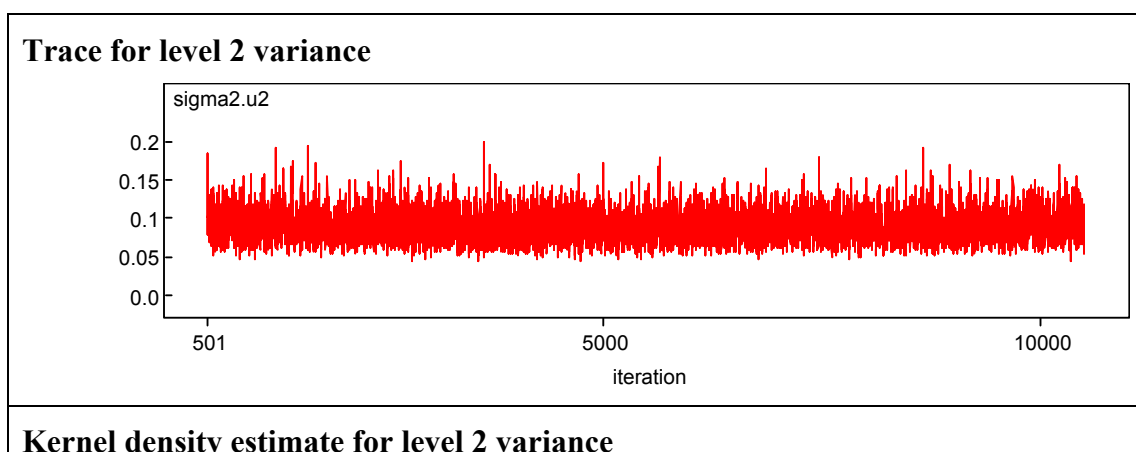
Table 1 compares two WINBUGS runs with a 500 burn in and 5000 and 10000 following iterations, together with the REML estimates.

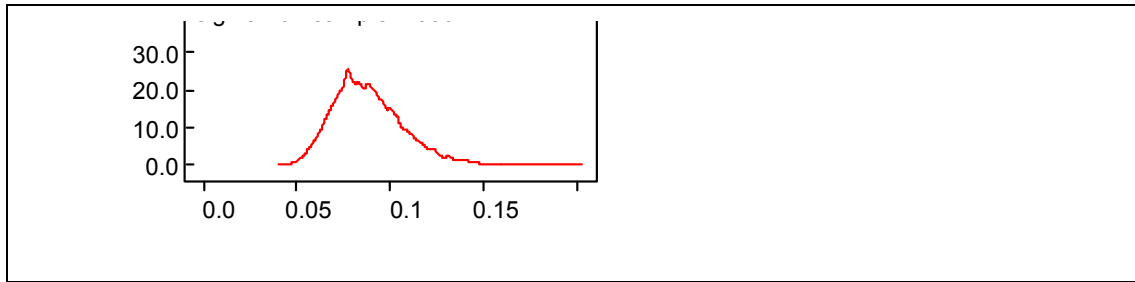
Table 1. Two level Normal response model for EXAM data			
Parameter	REML estimator (s.e.)	WINBUGS (5000) (95% interval)	WINBUGS (10000) (95% interval)
β_0	-0.168	-0.171	-0.171
β_1	0.560 (0.013)	0.560 (0.535, 0.584)	0.560(0.535, 0.584)
β_2	0.167 (0.034)	0.167 (0.099, 0.234)	0.167 (0.098, 0.233)
β_3	0.178 (0.114)	0.176 (-0.036, 0.383)	0.180 (-0.033, 0.394)
β_4	0.159 (0.089)	0.162 (-0.006, 0.329)	0.162 (-0.012, 0.338)
σ_{u0}^2	0.086 (0.017)	0.088 (0.057, 0.131)	0.088 (0.058, 0.131)
σ_{e0}^2	0.563 (0.013)	0.563 (0.539, 0.588)	0.563 (0.539, 0.588)
DIC		9245.7	9245.6
Time	0.4 secs (SAS)	7 mins. 12 secs*	11 mins. 45 secs*

*includes 500 iterations to convergence (Burn in). MCMC estimates are the chain means. Converted to equivalent time on 433Mhz Pentium II under Windows.

As can be seen there is very little change between 5000 and 10000 iterations. The point estimates are similar to the REML estimates and 95% confidence intervals using a Normal approximation are also close, except for the level 2 variance parameter. If the modal or median rather than mean are used from MCMC we would expect closer agreement with REML; in the present case the median for the level 2 variance is 0.086.

The following are the trace and kernel density plot for the level 2 variance for the analysis using 10000 iterations. The kernel density is unsmoothed; this may be changed in later versions.





In the remaining examples we shall give results from running 10000 iterations with a 500 burn in.

We now fit a model where the level 1 variance is a function of gender and this is followed by a model where the log(level 1 variance) is a function of STANDLRT. The model specifications are given in the following boxes and the results in Table 2.

Finally we show below the model specification for a model where STANDLRT has a coefficient random at the school level.

```

model
{
  # Level 1 definition
  for(i in 1:N) {
    normexam[i] ~ dnorm(mu[i],tau)
    mu[i] <- beta[1] * cons[i]
    + beta[2] * standlrt[i]
    + beta[3] * gender[i]
    + beta[4] * boysch[i]
    + beta[5] * girlsch[i]
    + u2[school[i],1] * cons[i]
    + u2[school[i],2] * standlrt[i]
  }
  # Higher level definitions
  for (j in 1:n2) {
    u2[j,1:2] ~ dmnorm(zero2[1:2],tau.u2[1:2,1:2])
  }
  # Priors for fixed effects
  for (k in 1:5) { beta[k] ~ dflat() }
  # Priors for random terms
  tau ~ dgamma(0.001000,0.001000)
  sigma2 <- 1/tau
  for (i in 1:2) {zero2[i] <- 0}
  tau.u2[1:2,1:2] ~ dwish(R2[1:2, 1:2],2)
  sigma2.u2[1:2,1:2] <- inverse(tau.u2[,i])
}

```

Note that we have used a ‘minimally informative’ Wishart distribution as prior for the inverse covariance matrix by giving it degrees of freedom equal to the order of the matrix and we can choose the Wishart matrix parameters as equal to the (inverse) covariance matrix based upon ML. These values, R2, will be declared with the data. Other priors are possible including a uniform.

2.2 Complex level 1 variation

We fit a separate variance for boys and girls as follows:

$$\begin{aligned}
y_{ij} &= (X\beta)_{ij} + u_{0j} + e_{ij} \\
e_{ij} &= e_{1ij}(1 - x_{2ij}) + e_{2ij}x_{2ij} \\
u_{0j} &\sim N(0, \sigma_{u0}^2) \quad \begin{pmatrix} e_{1ij} \\ e_{2ij} \end{pmatrix} \sim N \begin{pmatrix} \sigma_{e1}^2 & \\ 0 & \sigma_{e2}^2 \end{pmatrix}
\end{aligned} \tag{2}$$

$x_{2ij} = 1$ if girl

Note that WINBUGS can fit other complex level 1 variance functions. A useful one is the logarithm of the level 1 variance as a linear function of the STANDLRT score which guarantees a positive variance. This would be written

$$\begin{aligned}
y_{ij} &= (X\beta)_{ij} + u_{0j} + e_{ij} \\
u_{0j} &\sim N(0, \sigma_{u0}^2) \quad e_{ij} \sim N(0, \sigma_{eij}^2) \\
\log_e(\sigma_{eij}^2) &= \alpha_0 + \alpha_1 x_{1ij}
\end{aligned} \tag{3}$$

Note that in WINBUGS we need to express the precision as a linear function, in which case the coefficients simply reverse their sign.

For the first model the WINBUGS specification would be

```

model
{
# Level 1 definition
for(i in 1:N) {
normexam[i] ~ dnorm(mu[i],tau[i])
mu[i] <- beta[1] * cons[i]
+ beta[2] * standlrt[i]
+ beta[3] * boysch[i]
+ beta[4] * girlsch[i]
+ beta[5] * girl[i]
+ u2[school[i]] * cons[i]
}
# Higher level definitions
for (j in 1:n2) {
u2[j] ~ dnorm(0,tau.u2)
}
# Priors for fixed effects
for (k in 1:5) { beta[k] ~ dflat() }
# Priors for random terms
for(i in 1:N) {
tau[i] <- 1/sigma2[i]
sigma2[i] <- va0[1] * girl[i] * girl[i]
+ va0[2] * boy[i] * boy[i]
}
for(i in 1:2) {
va0[i] ~ dflat()
}
tau.u2 ~ dgamma(0.001000,0.001000)
sigma2.u2 <- 1/tau.u2
}

```

Note that we use ‘girl’ in the fixed part and both the dummy variables ‘girl’ and ‘boy’ for level 1. This requires that the data for ‘boy’ is also entered. Note also that we have used flat priors for the level 1 variance parameters; when we have a general function of explanatory variables - we cannot use gamma priors for variances. In fact in the gender case we could have specified gamma priors since we are really fitting two separate variances.

Table 2. Two level Normal response model for EXAM data with complex level 1 variance

	Separate gender variances
Parameter	Estimate (95% interval)
β_0	-0.196
β_1	0.548 (0.493, 0.603)
β_2	0.169 (0.102, 0.232)
β_3	0.193 (-0.003, 0.430)
β_4	0.189 (0.022, 0.340)
β_5	0.008 (-0.051, 0.064)
σ_{u0}^2	0.088 (0.058, 0.128)
σ_{u01}	0.022 (0.010, 0.038)
σ_{u1}^2	0.005 (0.009, 0.027)
σ_{e1}^2	0.589 (0.551, 0.631)
σ_{e2}^2	0.524 (0.499, 0.556)
α_0	
α_1	
DIC	9187.2
Time	2 mins 31 secs.

2.3 Three level Normal response model

For this example we use the [chem97](#) data set consisting of 31022 students nested within 2279 schools within 130 Education authorities.

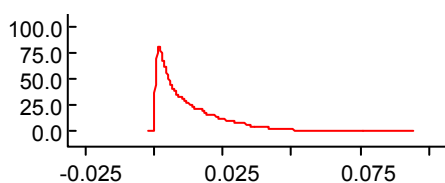
$$\begin{aligned}
 y_{ij} &= (X\beta)_{ij} + v_{0k} + u_{0jk} + e_{0ijk} \\
 v_{0k} &\sim N(0, \sigma_{v0}^2), \quad u_{0j} \sim N(0, \sigma_{u0}^2) \quad e_{0ij} \sim N(0, \sigma_{e0}^2)
 \end{aligned}
 \tag{4}$$

where we have the intercept and GCSE average score as predictors.

Table 3 presents the results.

Table 3. Three level Normal response model for CHEM97 data	
Parameter	WINBUGS (10000) (95% interval)
β_0	5.635
β_1	2.473 (2.440, 2.505)
σ_{v0}^2	0.009 (0.001, 0.045)
σ_{u0}^2	1.170 (1.066, 1.281)
σ_{e0}^2	5.155 (5.069, 5.241)
DIC	140415.3
Time	82 mins.

The level 3 variance is very poorly estimated, and in fact the mode of the distribution is at zero as is shown by the following kernel density estimate.



2.4 A binary response model

The dataset used is the fertility survey data BANG. In the fixed part the intercept and four covariates are fitted; urban, age, and four dummy variables for number of children in family (1, 2, 3+) with no children as base category. The logit link function model is

$$\begin{aligned}
 \text{logit}(\pi_{ij}) &= (X\beta)_{ij} + u_{0j} \\
 y_{ij} &\sim \text{Bin}(1, \pi_{ij}) \\
 u_{0j} &\sim N(0, \sigma_{u0}^2)
 \end{aligned} \tag{5}$$

For the probit link function we can write the model as:

$$\begin{aligned}
 \text{probit}(\pi_{ij}) &= (X\beta)_{ij} + u_{0j} \\
 y_{ij} &\sim \text{Bin}(1, \pi_{ij}) \\
 u_{0j} &\sim N(0, \sigma_{u0}^2)
 \end{aligned} \tag{6}$$

but we can also estimate the parameters of the underlying Normal.

Consider the 2-level model

$$\begin{aligned}
 y_{ij} &= (XB)_{ij} + u_{0j} + e_{0ij} \\
 e_{0ij} &\sim N(0, 1) \quad u_{0j} \sim N(0, \sigma_{u0}^2)
 \end{aligned} \tag{7}$$

where y_{ij} is the probit that determines the probability of a correct response, namely

$$\int_{-\infty}^{y_{ij}} \phi(t) dt, \quad \phi(t) \text{ is pdf of } N(0,1) \quad (8)$$

By inserting another step into the MCMC algorithm we can generate a chain from the underlying Normal distribution at level 1. Threshold models such as (7) & (8) have a number of uses where the binary responses can be thought of as generated from truncation of an underlying continuous distribution. Furthermore, estimation using this procedure is often faster since Gibbs sampling rather than, say, Metropolis Hastings or adaptive rejection methods can be used. A similar device can be used with the logistic link where we define an underlying logistic distribution. See Browne (2003) for more details.

In the present case we use the standard WINBUGS probit function and the model specification is as follows:

```
model
{
# Level 1 definition
for(i in 1:N) {
use[i] ~ dbin(p[i],denom[i])
probit(p[i]) <- beta[1] * cons[i]
+ beta[2] * urban[i]
+ beta[3] * agecentered[i]
+ beta[4] * one[i]
+ beta[5] * two[i]
+ beta[6] * three_[i]
+ u2[district[i]] * cons[i]
}
# Higher level definitions
for (j in 1:n2) {
u2[j] ~ dnorm(0,tau.u2)
}
# Priors for fixed effects
for (k in 1:6) { beta[k] ~ dflat() }
# Priors for random terms
tau.u2 ~ dgamma(0.001000,0.001000)
sigma2.u2 <- 1/tau.u2
}
```

Table 4. Two level binary response model for BANG data with logistic and probit link functions.

	Logit	Probit
Parameter	Estimate (95% interval)	Estimate (95% interval)
β_0	-1.732	-1.039
β_1	0.735 (0.501, 0.963)	0.450 (0.308, 0.588)
β_2	-0.028 (-0.043, -0.012)	-0.017 (-0.026, -0.007)
β_3	1.132 (0.834, 1.448)	0.676 (0.486, 0.855)
β_4	1.403 (1.063, 1.750)	0.842 (0.632, 1.048)
β_5	1.382 (1.022, 1.728)	0.825 (0.600, 1.053)
σ_{u0}^2	0.241 (0.117, 0.437)	0.086 (0.041, 0.156)
DIC	2407.1	2406.8
Time	41mins. 20 secs.	60 mins.

Note that the DIC values are almost identical suggesting that the models fit equally well.

2.5 A multivariate model with missing responses

The data set is GCSEMV, with two responses from pupils to a written and coursework assessment.

The model used is as follows:

$$y_{1ij} = \beta_{10} + \beta_{11}x_{1ij} + u_{1j} + e_{1ij}$$

$$y_{2ij} = \beta_{20} + \beta_{21}x_{1ij} + u_{2j} + e_{2ij}$$

$$\begin{pmatrix} u_{1j} \\ u_{2j} \end{pmatrix} \sim N \begin{pmatrix} 0 & \sigma_{u1}^2 \\ 0 & \sigma_{u12} & \sigma_{u2}^2 \end{pmatrix} \quad (9)$$

$$\begin{pmatrix} e_{1ij} \\ e_{2ij} \end{pmatrix} \sim N \begin{pmatrix} 0 & \sigma_{e1}^2 \\ 0 & \sigma_{e12} & \sigma_{e2}^2 \end{pmatrix}$$

The model specification is as follows, using a similar formulation for the Wishart priors as with the random coefficient model. We have called the response vector MV_RESP. The covariate is gender (girls=1).

```

model
{
# Level 1 definition
for(i in 1:N) {
MV_RESP[i,1:2] ~ dnorm(mu[i,1:2],tau.u2[1:2,1:2])
mu[i,1] <- beta[1] * CONS[i]
+ beta[3] * FEMALE[i]
+ u3[MV_SCHOO[i],1] * CONS[i]
mu[i,2] <- beta[2] * CONS[i]
+ beta[4] * FEMALE[i]
+ u3[MV_SCHOO[i],2] * CONS[i]
}
# Higher level definitions
for (j in 1:n3) {
u3[j,1:2] ~ dnorm(zero3[1:2],tau.u3[1:2,1:2])
}
# Priors for fixed effects
for (k in 1:4) { beta[k] ~ dflat() }
# Priors for random terms
tau.u2[1:2,1:2] ~ dwish(R2[1:2,1:2],2)
sigma2.u2[1:2,1:2] <- inverse(tau.u2[,])
for (i in 1:2) {zero3[i] <- 0}
tau.u3[1:2,1:2] ~ dwish(R3[1:2,1:2],2)
sigma2.u3[1:2,1:2] <- inverse(tau.u3[,])
}

```

In WINBUGS missing values are denoted by NA and are handled using imputation for a multivariate Normal response, although this is not available for a multivariate t-distribution response.

Table 5 gives the results

Table 5. Bivariate Normal response for written and coursework assessment.		
	Written	Coursework
Parameter	Estimate (95% interval)	Estimate (95% interval)
β_0	49.6 (47.8, 51.4)	69.7 (67.5, 72.1)
β_1	-2.53 (-3.60, -1.38)	6.73 (5.37, 8.04)
σ_{u1}^2	49.7 (33.2, 72.7)	
σ_{u12}	26.4 (10.2, 47.7)	
σ_{u2}^2	79.6 (54.4, 115.5)	
σ_{e1}^2	124.9 (116.8, 133.7)	
σ_{e12}	73.2 (65.2, 81.6)	
σ_{e2}^2	180.6 (169.9, 193.4)	
DIC	29900.3	
Time	2 mins. 30 secs.	

2.6 Cross classified data

The data set is 2LEV-XC consisting of examination results for students cross classified by primary and secondary school. The model is

$$\begin{aligned}
 Y_i &= \beta_0 + \beta_1 x_i + u_{0,sec(i)}^{(3)} + u_{0,prim(i)}^{(2)} + e_{0i} \\
 u_{0,sec(i)}^{(3)} &\sim N(0, \sigma_{u0(3)}^2) \\
 u_{0,prim(i)}^{(2)} &\sim N(0, \sigma_{u0(2)}^2) \\
 e_{0i} &\sim N(0, \sigma_{e0}^2)
 \end{aligned}
 \tag{10}$$

This uses the ‘classification notation’ introduced by Browne et al. (2001). The superscript denotes the classification number (not the level) where the pupil level term is at level 1, and the superscript is omitted for convenience. The term $sec(i)$ denotes the secondary school to which pupil i belongs and similarly $prim(i)$ the primary school. The covariate x_1 is a dummy variable for gender (girl=1).

The WINBUGS model specification is as follows

```

model
{
# Level 1 definition
for(i in 1:N) {
ATTAIN[i] ~ dnorm(mu[i],tau)
mu[i] <- beta[1] * CONS[i]
+ beta[2] * SEX[i]
+ u2[PID[i]] * CONS[i]
+ u3[SID[i]] * CONS[i]
}
# Higher level definitions
for (j in 1:n2) {
u2[j] ~ dnorm(0,tau.u2)
}
for (j in 1:n3) {
u3[j] ~ dnorm(0,tau.u3)
}
# Priors for fixed effects
for (k in 1:2) { beta[k] ~ dflat() }
# Priors for random terms
tau ~ dgamma(0.001000,0.001000)
sigma2 <- 1/tau
tau.u2 ~ dgamma(0.001000,0.001000)
sigma2.u2 <- 1/tau.u2
tau.u3 ~ dgamma(0.001000,0.001000)
sigma2.u3 <- 1/tau.u3
}

```

Where PID is the pupil identification. The data are sorted on secondary school within primary school so that the first few pupil records for the secondary school data and for the primary school data are respectively

```

9,9,9,9,9,1,1,9,9,
9,9,9,9,9,1,9,9,9,
9,9,9,9,9,9,9,9,9,
9,9,9,9,9,1,18,1,9,
1,9,9,9,9,9,1,9,1,
9,9,9,9,7,7,7,7,7,7,
7,5,5,5,9,9,9,9,9,9,
6,1,1,1,1,1,1,1,1,1,

```

```

1,1,1,1,1,1,1,1,1,1,

```

1,1,1,1,1,1,1,1,1,1,
1,1,1,1,1,1,1,1,1,1,
1,1,1,1,1,1,1,1,1,1,
1,1,1,1,1,1,1,1,1,1,
1,1,1,1,2,2,2,2,2,2,
2,3,3,3,4,4,4,4,4,4,
4,5,5,5,5,5,5,5,5,5,

The results are given in Table 6.

Table 6. Two level cross classified model using 2LEV-XC attainment data.	
Parameter	Estimate (95% interval)
β_0	5.253
β_1	0.498 (0.304, 0.690)
$\sigma_{u0(3)}^2$	0.411 (0.138, 0.940)
$\sigma_{u0(2)}^2$	1.130 (0.770, 1.600)
σ_{e0}^2	8.062 (7.687, 8.457)
DIC	17024.8
Time	8mins. 2 secs.

The wide interval estimate for the secondary school variance reflects the fact that there are only 19 such schools in the data set.

2.7 Multicategorical response models

The dataset used is the social attitudes data with the response being the number of positive answers to seven questions with seven codes being 0 or 1, 2, 3, 4, 5, 6, 7. This scale is ordered but for illustration we shall fit both ordered and unordered models. Each respondent has a response in each of 4 years. For the purpose of the present analysis we treat the years as independent replications so that there are just 4 ‘trials’ for each respondent.

For the unordered case we fit the multivariate logit model for t categories

$$E(y_{ij}^{(s)}) = \pi_{ij}^{(s)}$$

$$\log \left(\frac{\pi_{ij}^{(s)}}{\pi_{ij}^{(t)}} \right) = \beta_0^{(s)} + u_j^{(s)}, \quad s=1, \dots, t-1 \quad (11)$$

$$u_j \sim MVN(0, \Omega_u)$$

where we fit only intercept terms. The covariance matrix Ω_u is of order $(t-1) \times (t-1)$ in general but for present purposes we assume a common

single variance at level 2.

For the ordered case we fit the model for the cumulative probabilities defined as

$$E(y_{ij}^{(s)}) = \gamma_{ij}^{(s)} = \sum_{h=1}^s \pi_{ij}^{(h)}, \quad s = 1, \dots, t-1 \quad (12)$$

with

$$\gamma_{ij}^{(s)} = \{1 + \exp[-\beta_0^{(s)} + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \beta_3 x_{3ij} + u_{0j}]\}^{-1}$$
$$u_{0j} \sim N(0, \sigma_{u0}^2) \quad (13)$$

where we have included three covariates which are dummy variables for the religious categories Roman Catholic, Protestant and others and where only the intercept (threshold) term varies across categories.

The WINBUGS model specification code for the unordered and ordered models respectively is

and

The results are given in Table 7.

Table 7. Two level ordered and unordered multinomial response model for SOCATT attitude data with logistic link function. Category 7 is the base category.

	Unordered	Ordered (20,000 iterations)
Parameter	Estimate (95% interval)	Estimate (95% interval)
$\beta_0^{(1)}$	-1.923 (-2.456, -1.410)	-3.952 (-5.273, -2.695)
$\beta_0^{(2)}$	-1.258 (-1.697, -0.833)	-2.161 (-3.450, -0.952)
$\beta_0^{(3)}$	0.312 (-0.072, 0.687)	0.669 (-0.607, 1.871)
$\beta_0^{(4)}$	-0.360 (-0.760, 0.047)	1.617 (0.332, 2.831)
$\beta_0^{(5)}$	-0.382 (-0.782, 0.014)	2.518 (1.229, 3.746)
$\beta_0^{(6)}$	-0.264 (-0.653, 0.133)	3.571 (2.284, 4.803)
β_1		-2.248 (-3.519, -0.862)
β_2		-0.976 (-2.332, 0.554)
β_3		-3.086 (-4.396, -1.669)
σ_{u0}^2	5.945 (3.962, 8.609)	5.438 (4.164, 7.002)
DIC	3292.2	2929.5
Time		

3. Documentation, user support and general conclusions

There is an on line manual supplied with the software which is a fairly basic description of the packages features. There is no HELP system, and this would be a very welcome future addition. There is a useful tutorial example in the manual which, together with a series of worked examples with real datasets should enable new users to get started.

There is a discussion list for users of WINBUGS where issues are discussed. To access this go to <http://www.jiscmail.ac.uk/lists/BUGS.html>. This is an active list where users share ideas and problems. In addition the developers of WINBUGS are happy to receive feedback; go to <http://www.mrc-bsu.cam.ac.uk/bugs/overview/list.shtml>.

WINBUGS is a very impressive package, especially given its free distribution policy. It is perhaps the best known Bayesian modelling package and is flexible enough to handle a very wide range of models. In this review we have only touched upon a few of its facilities. Because of its generality it is not necessarily as computationally efficient as other packages which can fit MCMC models, such as MlwiN (see the comparative tables) but its greater model flexibility does allow new models to be developed. One drawback of WINBUGS is that is rather cumbersome to use if one wishes to carry out extensive model exploration trying different model formulations. For this purpose one of the likelihood based packages could be used with final

estimates obtained from WINBUGS. Another issue is that there is a great deal of embedded statistical theory within WINBUGS which will make considerable demands on less statistically experienced users. There are other issues associated with Bayesian modelling, such as choice of priors, but it is not the purpose of this review to enter into these.

If you want a very flexible MCMC Bayesian package for multilevel modelling, WINBUGS is an excellent choice.

References

Browne, W. (1998). *Applying MCMC methods to multilevel models. Statistics*. Bath, University of Bath.

Browne, W., Goldstein, H. and Rasbash, J. (2001). Multiple membership multiple classification (MMMC) models. *Statistical Modelling* **1**: 103-124.

Spiegelhalter, D., Best, N., Carlin, B. P. and Van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, B* **64**: 583-640.