Workshop on Bayesian Modeling Using WinBUGS

Sessions 10–11 (B) : Advanced GLMs and Extensions

Presentation is available at: http://stat-athens.aueb.gr/~jbn/current.htm

Workshop on Bayesian Modeling Using WinBUGS

Sessions 10–11 (B) : Advanced GLMs and Extensions

Ioannis Ntzoufras

Department of Statistics,
Athens University of Economics and Business,
Athens, Greece;

e-mail: ntzoufras@aueb.gr.

Presentation is available at: http://stat-athens.aueb.gr/~jbn/current.htm

This presentation is based on Chapter 8 of Ntzoufras (2009): *Bayesian Modeling Using WinBUGS*, Wiley.

Synopsis

8.1 Models with nonstandard distributions

- WinBUGS allows for modeling of nonstandard distributions (i.e., for a distribution that is not listed in WinBUGS' prespecified distributions) using the zero–ones trick.
- Here we provide general guidelines and then illustrate this approach using an example by fitting the inverse Gaussian model.

8.1.1 Specification of arbitrary likelihood using the zeros–ones trick

- We can use either the Bernoulli or the Poisson distribution to indirectly specify any arbitrary model likelihood.
- Denote the log-likelihood by $l_i = \log f(y_i|\boldsymbol{\theta})$.
- The model likelihood can be written as

$$f(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_{i=1}^{n} e^{l_i} = \prod_{i=1}^{n} \frac{e^{-(-l_i)}(-l_i)^0}{0!} = \prod_{i=1}^{n} f_P(0; -l_i)$$

- $f_P(x;\lambda)$: Poisson probability function evaluated at x with mean λ.
- $f_P(0;-l_i)$: Poisson probability function evaluated at $x=0$ with mean $-l_i$.
- Hence, the model likelihood can be written as
  - The product of the densities of new pseudorandom variables Ξ_i $(i = 1, \ldots, n)$,
  - with $\Xi_i \sim \text{Poisson}(-l_i)$ and
  - all observed values of Ξ_i are zero

The zeros trick
Ensuring that the mean of the Poisson is positive

- To ensure the positivity of the mean $(-kl_i)$ of each Ξ_i, we add a positive constant term C to the mean.
- This is equivalent to multiplying each likelihood term by e^{-C}.
- This action does not affect the likelihood and the posterior inference since it is equivalent to multiplying the resulting (unnormalized) posterior distribution by a constant term equal to e^{-nC}.
- With this approach, the likelihood becomes equal to

$$f(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_{i=1}^{n} \frac{e^{-(-l_i+C)}(-l_i+C)^0}{0!} = \prod_{i=1}^{n} f_P(0; -l_i + C)\ ;$$

C must be selected in such way that $-l_i + C > 0$ for all $i = 1, 2, \ldots, n$.

The zeros trick - WinBUGS Code

$C = 10000$
for $i = 1, \ldots, n$
$\xi_i = 0$
$\xi_i \sim \text{Poisson}(\mu_{\xi_i})$
$\mu_{\xi_i} = -l_i + C$

$l_i = \log f(y_i|\boldsymbol{\theta})$

```
C <- 10000
for (i in 1:n) {
    zeros[i] <- 0
    zeros[i] ~ dpois( zeros.mean[i])
    zeros.mean[i] <- -l[i] + C
    # write here the expression of the log-
        likelihood for i observation
    l[i] <- ...
    }
```

`l[i]` must be specified accordingly for each model.

e.g. for Normal model:

$$l_i = -0.5\log(2\pi) - 0.5\log(\sigma^2) - \frac{(y_i - \mu_i)^2}{2\sigma^2}$$

```
l[i] <- -0.5*log(2*3.14) -0.5*log(s2) -0.5*pow( y[i]-mu[i], 2 )/s2
```


The Ones trick

- Instead of the Poisson–zeros strategy, the Bernoulli distribution can be also used for the same purpose.
- With this approach

$$f(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_{i=1}^{n} e^{l_i} = \prod_{i=1}^{n} \left(e^{l_i}\right)^1 \left(1 - e^{l_i}\right)^0 = \prod_{i=1}^{n} f_B(1; e^{l_i}, 1),$$

  where $f_B(1; e^{l_i}, 1)$ is the binomial probability function with success probability e^{l_i} and $N = 1$.
- Hence, the model likelihood can be expressed as
  - the product of the densities of new pseudorandom variables Ξ_i,
  - with $\Xi_i \sim \text{Bernoulli}(e^{l_i})$
  - with all observed values of Ξ_i equal to one.

The Ones trick (2)

- Success probability should be lower than one $\Rightarrow$ multiply each likelihood term by e^{-C}.
- Now the likelihood is given by

$$f(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_{i=1}^{n} \left(e^{l_i - C}\right)^1 \left(1 - e^{l_i - C}\right)^0 = \prod_{i=1}^{n} f_B(1; e^{l_i - C}, 1) \ .$$

$C = 100$
for $i = 1, \ldots, n$
$\xi_i = 1$
$\xi_i \sim \text{Bernoulli}(\pi_{\xi_i})$
$\pi_{\xi_i} = e^{-l_i + C}$

$l_i = \log f(y_i|\boldsymbol{\theta})$

```
C <- 100
for (i in 1:n) {
    ones[i] <- 1
    ones[i] ~ dbern( ones.p[i])
    ones.p[i] <- exp( l[i] - C )
    # log-likelihood for i
        observation
    l[i] <- ...
}
```

`l[i]` must be specified accordingly for each model (as in the Poisson–zeros trick).

Some Comments

- Both approaches have the same effect.
- Use the Poisson–zeros approach because it avoids overflow problems due to the simpler likelihood expression.
- The same approach can be followed to specify a prior distribution of nonstandard form (Spiegelhalter et al., 2003).

8.1.2 The inverse Gaussian model

The inverse Gaussian distribution

The density of the inverse Gaussian distribution $Y \sim \text{IGaussian}(\mu, \lambda)$ is given by

$$f(y|\mu,\lambda) \quad = \quad \left(\frac{\lambda}{2\pi y^3}\right)^{1/2} \exp\left(-\frac{\lambda(y-\mu)^2}{2\mu^2 y}\right) \quad \text{for} \quad y > 0.$$

- $E(Y) = \mu$ and $V(Y) = \mu^3/\lambda$.
- In GLMs, the parametrization $\mu, \sigma^2 = \lambda^{-1}$ is frequently encountered.
- For $\lambda \rightarrow \infty$ (or $\sigma^2 \rightarrow 0$) $\Rightarrow$ the inverse Gaussian distribution becomes similar to a normal (Gaussian) distribution.
- $Y > 0 \Rightarrow$ the inverse Gaussian does not result in the inverse of a normal (Gaussian) distribution (Seshadri, 1993).

Model formulation

$$Y_i \sim \text{IGaussian}(\mu_i, \lambda),$$

Canonical (squared reciprocal) link:	$\mu_i^{-2} = \eta_i \Leftrightarrow \mu_i = \sqrt{1/\lvert\eta_i\rvert}$
Log-link:	$\log \mu_i = \eta_i \Leftrightarrow \mu_i = e^{\eta_i}$.

The log-link is preferred because of its the easier interpretation (similar to the Poisson log-linear models).

WinBUGS Syntax

The log-likelihood for the original parametrization is given by

```
l[i] <-   0.5*( log( lambda ) - log(2*3.14) - 3*log(y[i]) )
        - 0.5* lambda * pow( (y[i]-mu[i])/mu[i], 2 )/y[i]
```

Canonical link: $\mu_i = \sqrt{1/|\eta_i|}$

```
mu[i] <- sqrt( 1/abs(eta[i]) )
```

or by

Log link: $\mu_i = e^{\eta_i}$

```
log(mu[i]) <- eta[i]
```

$\sigma^2 = 1/\lambda$

```
s2 <- 1/lambda
```

- $\sigma^2 \Rightarrow$ logical node (i.e. deterministic function)
- Normal priors for β_j
- Gamma prior (similar to the precision of the normal regression model) for λ

Example 8.1. An inverse Gaussian simulated dataset.

- *Simulated dataset with $n = 100$ random values from*

$$\text{IGaussian}(\log \mu = 3 + 2X_1 - 1X_2, \lambda = 10).$$

- *Four standardized normal variables X_j, $j = 1,2,3,4$ were generated as possible covariates.*
- *Data are available at the Website of Ntzoufras (2009).*

Table 8.1: Posterior summaries of inverse Gaussian model parameters for simulated data of Example 8.1

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
beta[1]	3.220	0.156	0.014	2.936	3.215	3.539	4001	5000
beta[2]	2.151	0.083	0.007	1.998	2.151	2.324	4001	5000
beta[3]	-0.950	0.111	0.006	-1.183	-0.945	-0.745	4001	5000
beta[4]	-0.081	0.083	0.004	-0.242	-0.083	0.096	4001	5000
beta[5]	0.013	0.122	0.004	-0.242	0.020	0.247	4001	5000
lambda	12.140	1.779	0.033	8.913	12.070	15.82	4001	5000
s2	0.084	0.013	0.0002	0.063	0.083	0.112	4001	5000

Results

- Both the Poisson–zeros and the Bernoulli–ones approaches worked satisfactorily with similar results.
- 4000 burnin and an additional 5000 iterations finally kept are presented in Table 8.1.
- The constant C was set equal to $10,000$ (Poisson) and 100 (Bernoulli).
- The log-link was used in both cases.

Estimated model (based on the posterior means):

$$Y_i \sim \text{IGaussian}(\mu = e^{3.22+2.15X_1-0.95X_2+0.08X_3+0.01X_4}, \lambda = 12.14),$$

Actual model:

$$Y_i \sim \text{IGaussian}(\mu = e^{3.00+2.00X_1-1.00X_2+0.00X_3+0.00X_3}, \lambda = 10.00)$$

Table 8.2: WinBUGS code for inverse Gaussian model used for simulated data of Example 8.1

```
model{
   C <- 10000
   for (i in 1:n) {
       zeros[i] <- 0
       zeros[i] ~ dpois( zeros.mean[i])
       zeros.mean[i] <- -l[i] + C
       l[i] <-   0.5*( log( lambda ) - log(2*3.14) - 3*log(y[i]) )
               - 0.5* lambda * pow( (y[i]-mu[i])/mu[i], 2 )/y[i]
       log(mu[i]) <- beta[1] + beta[2]*x1[i] + beta[3]*x2[i]
                             + beta[4]*x3[i] + beta[5]*x4[i]
    }
    # priors
    for (j in 1:5){ beta[j] ~ dnorm( 0.0 ,  0.001 ) }
    lambda ~ dgamma( 0.01, 0.01)
    s2 <- 1/lambda
}
```

Poisson–zeros trick; code for the ones trick is available at the Website of Ntzoufras (2009)

References

Ntzoufras, I. (2009), *Bayesian Modeling Using WinBUGS*, Wiley Series in Computational Statistics, Hoboken, NJ.

Seshadri, V. (1993), *The Inverse Gaussian Distribution: A Case Study in Exponential Families*, Oxford Science Publications, UK.

Spiegelhalter, D., Thomas, A., Best, N. and Lunn, D. (2003), *WinBUGS User Manual,* Version 1.4, MRC Biostatistics Unit, Institute of Public Health and Department of Epidemiology and Public Health, Imperial College School of Medicine, UK, available at `http://www.mrc-bsu.cam.ac.uk/bugs`.