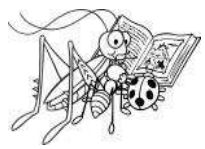
BUGS

BUGS is freely available software for constructing Bayesian statistical models and evaluating them using MCMC methodology.



BUGS and WINBUGS are distributed freely and are the result of many years of development by a team of statisticians and programmers at the Medical research Council Biostatistics Research Unit in Cambridge, see the Bugs Project: `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 DOODLEBUGS can be extremely useful. I trust you have visited `www.mrc-bsu.cam.ac.uk/bugs/` and down-loaded the latest version of BUGS software. As of mid-October 2004, the latest version is 1.4 with patch 1.4.1 and beta 1.5. You should also request the password that will alow you unlimited free use of the software.

So let's start with a simple example. Consider the regression model

$$\begin{aligned}
[y_i|\mu_i,\tau] &\sim \mathcal{N}(\mu_i,\tau),\ i=1,\ldots,n\\
\mu_i &= \alpha+\beta(x_i-\bar{x}),\\
\alpha &\sim \mathcal{N}(0,10^{-4})\\
\beta &\sim \mathcal{N}(0,10^{-4})\\
\tau &\sim \mathcal{G}amma(0.001,0.001).
\end{aligned}$$

The normal distributions here are parameterized in terms of *precision* parameters, $\tau = 1/\sigma^2$ and 10^{-4}. Natural distributions for the precision parameters are Gamma and small values of the precision reflect the flatness (noninformativeness) of the priors. Assume that (x,y)-pairs $(1,1),(2,3),(3,3),(4,3)$, and $(5,5)$ are observed and that initial values of the parameters are $\alpha_0 = 0.1$, $\beta_0 = 0.6$, and $\tau = 1$.

Estimators in classical, Least Square regression of y on $x-\bar{x}$, are given in the following table.

```
Coef      LSEstimate  SE Coef     t      p
ALPHA         3.0000   0.3266  9.19  0.003
BETA          0.8000   0.2309  3.46  0.041
S = 0.730297   R-Sq = 80.0%   R-Sq(adj) = 73.3%
```

How about Bayesian estimators?

Start BUGS and read or cut-and-paste the following code in. [**File** ⟩ **New**]. You can save yourself some typing if you download the code from the course page: `simpleexample.vi` at Codes/Software menu.

The code contains three parts: model, data, and initial values.

```
# A simple example
model{
  for (i in 1:N) {
  Y[i] ~ dnorm(mu[i],tau);
```

```
  mu[i] <- alpha + beta * (x[i] - x.bar);
  }
x.bar <- mean(x[]);
alpha ~ dnorm(0, 0.0001);
beta ~ dnorm(0, 0.0001);
tau ~ dgamma(0.001, 0.001);
sigma <- 1.0/sqrt(tau);
}

#----------------------------
#this window is data
list( x=c(1,2,3,4,5), Y=c(1,3,3,3,5), N=5);

#----------------------------
#initials
list(alpha = 0.1, beta = 0.6, tau = 1);
```

Next, highlight the line in the code that contains word "model." Go tho the **Model** menu and open **Specification. Specification Tool** window will pop-out. If your model is highlighted, you may **check model** in the specification tool window. If the model is correct, the response on the lower bar of the BUGS window should be: **model is syntactically correct.** Next, highlight the "list" statement in the data-part of your code. In the Specification Tool window select **load data.** If the data are in correct format, you should receive response on the bottom bar of BUGS window: **data loaded.** You need to compile your model on order to activate **inits**-buttons. Select **compile** in the Specification Tool window. The response should be: **model compiled**, and the buttons **load inits** and **gen inits** are active. Finally, highlight the "list" statement in the initials-part of your code and in the Specification Tool window select **load inits.** The response should be: **model is initialized**, and this finishes reading in the model. In the **Model** menu, you can choose **Update...** and open **Update Tool** to check if your model updates and also to burn-in some simulations.

From the **Inference** menu, open **Samples...**. Window titled **Sample Monitor Tool** will pop out. In the **node** sub-window input the names of variables you want to monitor. In this case these are `alpha, beta,` and `tau.` If you correctly input the variable the **set** button will become active and you can set the variable. Do this for all 3 variables of interest. In fact, `sigma` is available, as well.

Choose, now `alpha` from the subwindow in **Sample Monitor Tool**. All buttons (**clear, set, trace, history, density, stats, coda, quantiles, bgr diag, auto cor**) are now active.

Return to **Update Tool** and select 10000 in **updates** subwindow. Press **update** button.

Return to **Sample Monitor Tool** and check **trace** for the part of MC trace for α, **history** for the complete trace, **density** for the density estimator of α, etc. For example, pressing **stats** button will produce something like the following table

	mean	sd	MCerror	val2.5pc	median	val97.5pc	start	sample
alpha	3.003	0.549	0.003614	1.977	3.004	4.057	10000	20001

For all parameters a comparative table is

	mean	sd	MCerror	val2.5pc	median	val97.5pc	start	sample
alpha	3.003	0.549	0.003614	1.977	3.004	4.057	10000	20001
beta	0.7994	0.3768	0.002897	0.07088	0.7988	1.534	10000	20001
tau	1.875	1.521	0.01574	0.1399	1.471	5.851	10000	20001
sigma	1.006	0.7153	0.009742	0.4134	0.8244	2.674	10000	20001

If you want to save the trace for α in a file and process it in MATLAB, say, select **coda** and data window will open with an information window as well. Keep data window active and select **Save As** from the **File** menu. Save the α's in `alphas.txt.` ready to be imported to MATLAB.

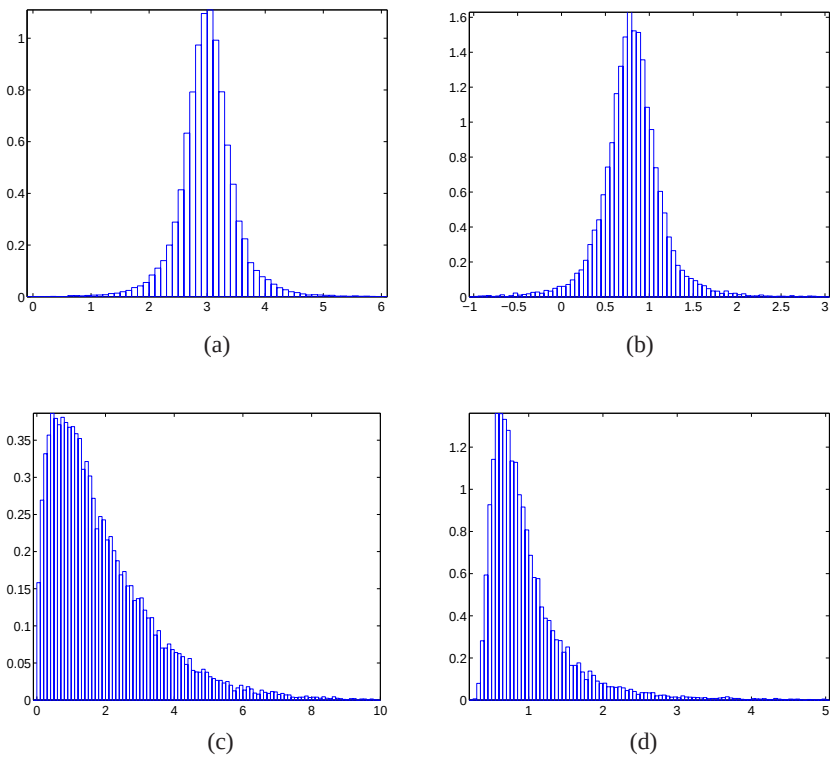


Figure 1: Simple example: Traces of α, β, τ, and σ from BUGS are imported to MATLAB and histograms plotted.

Some BUGS Examples

Anticlons: Is the cloning of humans moral? .

Recent Gallup Poll estimates that about 88% Americans opposed cloning humans (Let's name them "anti-clons"). Results are based on telephone interviews with a randomly selected national sample of $n = 1000$ adults, aged 18 and older, conducted May 2-4, 2004. In these 1000 interviews, 882 adults opposed cloning humans.

Figure 2: CLONAID:"world's first cloning of babies - Eve" - Fact or Fraud?

(i) Using BUGS estimate the proportion anticlons. Use non-informative prior for the proportion.
(ii) Test the hypothesis that the true proportion is less than 0.87.

(iii) Pretend that the original poll had $n = 1062$ adults, i.e., results for 62 adults are missing. Estimate the number of anticlons among the 62 missing in the poll.

```
Anticlons
model {     anticlons ~  dbin(prob,npolled) ;
       lessthan87 <- step(prob-0.87)
       anticlons.missing ~ dbin(prob,nmissing)
       prob ~ dbeta(1,1)}

Data
list(anticlons=882,npolled= 1000, nmissing=62)
```

Exercises

1. [From David Madigan's Bayesian Course Site] Bayesian binary regression with a probit model using BUGS.

(i) Finney (1947) describes a binary regression problem with two continuous valued predictors and a binary response. Here are the data in BUGS-ready format:

```
list(n=39,x1=c(3.7,3.5,1.25,0.75,0.8,0.7,0.6,1.1,0.9,0.9,0.8,0.55,0.6,1.4,0.75,2.3,3.2,
0.85,1.7,1.8,0.4,0.95,1.35,1.5,1.6,0.6,1.8,0.95,1.9,1.6,2.7,2.35,1.1,1.1,1.2,0.8,
0.95,0.75,1.3),x2=c(0.825,1.09,2.5,1.5,3.2,3.5,0.75,1.7,0.75,0.45,0.57,2.75,3.0,
2.33,3.75,1.64,1.6,1.415,1.06,1.8,2.0,1.36,1.35,1.36,1.78,1.5,1.5,1.9,0.95,0.4,
0.75,0.03,1.83,2.2,2.0,3.33,1.9,1.9,1.625),y=c(1,1,1,1,1,1,0,0,0,0,0,0,0,1,1,
1,1,1,0,1,0,0,0,0,1,0,1,0,1,0,1,0,0,1,1,1,0,0,1))
```

The objective here is to build a predictive model that predicts y using x_1 and x_2. One approach is the probit model: $P(y = 1|x_1, x_2) = \Phi(\beta_0 + \beta_1\ x_1 + \beta_2\ x_2)$ where Φ is the standard normal cumulative distribution function. Use BUGS to compute posterior distributions for β_0, β_1 and β_2 using diffuse normal priors for each. Please provide your BUGS code as well as the posterior distributions.

(ii) Suppose instead of the diffuse normal prior for $\beta_i,\ i = 0, 1, 2,$ you use a normal prior with mean zero and variance v_i, and assume the v_i's are independently exponentially distributed with some hyperparameter gamma (i.e., a hierarchical model). Fit this model using BUGS. How different are the posterior distributions from this model? How sensitive are they to the choice of gamma?

2. Flipping a fair coin in BUGS.

```
#coin.bug:
model coin;
{
flip12 ~ dcat(p.coin[])
coin <- flip12 - 1
}
#coin.dat:
list(p.coin=c(0.5, 0.5))
# just generate initials
```

3. IVF Success Rate and Age. The highly publicized (recent TV reports) *in vitro fertilization* success cases for women in their late fifties all involve donor's egg. If the egg is woman's own, the story is quite different.

In vitro fertilization (IVF), one of the assisted reproductive technology (ART) procedures, involves extracting a woman's eggs, fertilizing the eggs in the laboratory, and then transferring the resulting embryos into the womans uterus through the cervix. Fertilization involves a specialized technique known as intracytoplasmic sperm injection (ICSI).

The table shows the live-birth success rate per transfer rate from own eggs, by age of recipient. The data are for year 1999, published by CDC:

(`http://www.cdc.gov/reproductivehealth/ART99/index99.htm`)

Age (x)	24	25	26	27	28	29	30	31	32	33	34	35
Percentage (y)	38.7	38.6	38.9	41.4	39.7	41.1	38.7	37.6	36.3	36.9	35.7	33.8
Age (x)	36	37	38	39	40	41	42	43	44	45	46	
Percentage(y)	33.2	30.1	27.8	22.7	21.3	15.4	11.2	9.2	5.4	3.0	1.6	

Assume the change-point regression model

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_i + \epsilon_i,\ i = 1, \ldots, \tau \\ y_i &= \gamma_0 + \gamma_1 x_i + \epsilon_i,\ i = \tau + 1, \ldots, n \\ \epsilon_i &\sim \mathcal{N}(0, \sigma^2). \end{aligned}$$

Propose priors (with possibly hyperpriors) on σ^2,β_0, β_1, γ_0, and γ_1.

Take discrete uniform prior on τ. Program this exercise in BUGS.

4. BUGS **at NIST.**

Check: Bayesian Version of Consensus Means

`http://www.itl.nist.gov/div898/bayesian/bugscons.htm`

and Bayesian Analysis of the SRM 1946 chlorinated data

`http://www.itl.nist.gov/div898/bayesian/srm1946.htm`

for some very applied uses of BUGS at NIST (Blaza Toman).

References

[1] Spiegelhalter, D. J., Thomas, A., Best, N. G., and Gilks, W. R. (1996). BUGS Examples Volume 1, Version 0.5, (version ii). Cambridge: Medical Research Council Biostatistics Unit. (PDF)

[2] Finney, D. J. (1947). The estimation from individual records of the relationship between dose and quantal response. Biometrika 34, 320–334.