

ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΔΙΠΛΩΜΑ ΒΙΟΣΤΑΤΙΣΤΙΚΗΣ

ΜΑΘΗΜΑ: ΜΠΕΨΖΙΑΝΗ ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑ

ΔΙΔΑΣΚΟΝΤΕΣ: Ι. Ντζούφρας, Π. Τσιαμυρτζής

ΔΙΔΑΚΤΙΚΟ ΥΛΙΚΟ ΓΙΑ ΤΟ ΜΑΘΗΜΑ

ΜΠΕΨΖΙΑΝΗ ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑ:

Βιοστατιστική κατά Bayes με τη χρήση

του Λογισμικού BUGS

ΙΩΑΝΝΗΣ ΝΤΖΟΥΦΡΑΣ



*Χίος, Μάιος 2004
Αθήνα, Ιανουάριος 2006*

ΜΕΤΑΠΤΥΧΙΑΚΟ ΔΙΠΛΩΜΑ ΒΙΟΣΤΑΤΙΣΤΙΚΗΣ

ΜΑΘΗΜΑ: ΜΠΕΥΖΙΑΝΗ ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑ

ΔΙΔΑΣΚΟΝΤΕΣ: Ι. Ντζούφρας, Π. Τσιαμυρτζής

ΜΕΡΟΣ II: Βιοστατιστική κατά Bayes με τη χρήση του Λογισμικού BUGS

Λιδάσκων: Ιωάννης Ντζούφρας,

Διεύθυνση Ηλεκτρονικού Ταχυδρομείου: ntzoufras@aueb.gr

Ιστοσελίδα Μαθήματος (Μέρος II):

<http://stat-athens.aueb.gr/~jbn/courses/bugs2/home.html>

Προγράμματα/MCMC Software

1. *Classic BUGS 0.6,*

available at <http://www.mrc-bsu.cam.ac.uk/bugs/classic/bugs06/prog06.exe>

2. *WINBUGS 1.4*

available at <http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/WinBUGS14.exe>

Registration page <http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/register.shtml>

3. *CODA*

available at <http://www.mrc-bsu.cam.ac.uk/bugs/classic/coda04/cdaprg04.exe>

Εγχειρίδια/Manuals

1. Spiegelhalter, D., Thomas, A., Best, N. and Gilks, W. (1996). *BUGS 0.5: Bayesian Inference Using Gibbs Sampling Manual*. MRC Biostatistics Unit, Institute of Public health, Cambridge, UK.

available at <http://www.mrc-bsu.cam.ac.uk/bugs/documentation/Download/manual05.pdf>

2. Spiegelhalter, D., Thomas, A., Best, N. and Gilks, W. (1996). *BUGS 0.5: Examples Volume 1*. MRC Biostatistics Unit, Institute of Public health, Cambridge, UK.

available at <http://www.mrc-bsu.cam.ac.uk/bugs/documentation/Download/eg05vol1.pdf>

3. Spiegelhalter, D., Thomas, A., Best, N. and Gilks, W. (1996). *BUGS 0.5: Examples Volume 2*. MRC Biostatistics Unit, Institute of Public health, Cambridge, UK.

available at <http://www.mrc-bsu.cam.ac.uk/bugs/documentation/Download/eg05vol2.pdf>

4. 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 & Public Health, Imperial College School of Medicine.

available at <http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/manual114.pdf> .

5. Additional documentation and manuals for BUGS/CODA available at <http://www.mrc-bsu.cam.ac.uk/bugs/documentation/contents.shtml> .
6. New examples for WinBUGS available at <http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/examples.shtml>
7. Additional electronic material (tutorial, courses papers) for WINBUGS available at <http://www.mrc-bsu.cam.ac.uk/bugs/weblinks/webresource.shtml> .

Βιβλιογραφία/Bibliography

1. Scollnik D.P.M. (2000). Actuarial modeling with MCMC and BUGS: Additional Worked Examples. *Actuarial Research Clearing House*, **2000.2**, 433-585, available at <http://www.math.ucalgary.ca/~scollnik/abcd> .
2. Scollnik, D.P.M (2001). Actuarial modeling with MCMC and BUGS. *North American Actuarial Journal*, **5(2)**, 96-125, available at <http://www.math.ucalgary.ca/~scollnik/abcd> .
3. Congdon, P. (2001). *Bayesian Statistical Modelling*. Willey and Sons, ISBN: 0-471-49600-6 .
4. Jackman, S. (2004). *Estimation and inference via MCMC: a Resource for Social Scientists*; available at <http://tamarama.stanford.edu/mcmc/> .
5. Ibrahim, J.G., Chen, M.-H. and Sinha, D. (2001). *Bayesian Survival Analysis*, examples are available at <http://merlot.stat.uconn.edu/~mhchen/survbook/>.

Επιπλέον Πηγές/ Additional Resources

1. **Social science:** [Simon Jackman's MCMC Resource for Social Scientists](#) features a wide range of models concerned with ordered outcomes, missing data, random coefficients, generalized link functions, latent autoregressive structure and so on. WinBUGS code and Splus data files are provided, as well as tutorial papers on MCMC for social scientists.
 - Web-page: <http://tamarama.stanford.edu/mcmc/> .
2. **Pharmacokinetics:** [David Lunn's PKBugs Page](#) contains details of an 'add-on' to WinBUGS for pharmacokinetic modelling, developed by David Lunn at Imperial College. This can be run using WinBUGS 1.3.
 - Web-page: http://www.med.ic.ac.uk/divisions/60/pkbugs_web/home.html .
3. **Actuarial science:** [Actuarial Modelling with MCMC and BUGS](#) has been provided by David Scollnik in Calgary, and has a range of worked examples designed for an actuarial context but using models of much wider applicability.

An excellent tutorial paper on WinBUGS can also be downloaded - better than the WinBUGS documentation!

- Web-page: <http://www.math.ucalgary.ca/~scollnik/abcd/> .
- 4. **Population genetics:** [Kent Holsinger's Population Genetics course](http://darwin.eeb.uconn.edu/eeb348/) (<http://darwin.eeb.uconn.edu/eeb348/>) has a whole set of examples using WinBUGS for estimating inbreeding coefficients, selfing rates, analysing variability selection and so on. Kent also has a set of [notes and WinBUGS code](http://darwin.eeb.uconn.edu/summer-institute/summer-institute.html) (<http://darwin.eeb.uconn.edu/summer-institute/summer-institute.html>) from the Summer Institute for Statistical Genetics at NC State, which form an introduction to using WinBUGS in population genetics.
- 5. **Cost-effectiveness analysis:** [Tony O'Hagan's Research Page](#) contains draft papers and WinBUGS code for running Bayesian cost-effectiveness analysis.
 - Web-page: <http://www.shef.ac.uk/~st1ao/res.html> .
- 6. **Programs for analysing imperfect diagnostic tests:** [The Epidemiologic Diagnostics group at UC Davis](#) provide WinBUGS code and examples for analyzing data derived from imperfect diagnostic test.
 - Web-page: <http://www.epi.ucdavis.edu/diagnostictests/downloads.html>.
- 7. **Complex epidemiological modelling:** [Tom Smith at the Swiss Tropical Institute](http://www.sti.ch/biomet.htm) (<http://www.sti.ch/biomet.htm>) has models and documentation for 1) A latent class model for non-parametric resolution of a two component mixture, with a training set available for one component: 2) Two-state Hidden Markov Model with covariates 3) A non-linear regression model with Poisson errors in both x and y. [Brad Carlin's software page](#) also has a variety of examples for longitudinal and spatial models (<http://www.biostat.umn.edu/~brad/software.html>).
- 8. **Educational testing:** [Dan Segall's site](#) concerns an item-response model for test compromise, and includes a manuscript and related WinBUGS programs.
 - Web-page: <http://segalld.home.netcom.com/compromise.html> .
- 9. **Archeology:** [Andrew Millard's 'WinBUGS and Bayesian tools for Archaeology' site](#) shows how to use WinBUGS to analyse many Bayesian examples from the archeological literature.
 - Web-page: <http://www.dur.ac.uk/a.r.millard/BUGS4Arch.html> .

Bayesian BioStatistics Using BUGS



**Βιοστατιστική κατά Bayes
με τη χρήση του Λογισμικού
BUGS**

Ioannis Ntzoufras

E-mail: ntzoufras@aueb.gr



Department of Statistics,
Athens University of
Economics & Business

ΠΕΡΙΕΧΟΜΕΝΑ

⌘ **ΜΑΘΗΜΑ 1: ΕΙΣΑΓΩΓΗ ΣΤΟ BUGS**

⌘ **ΜΑΘΗΜΑ 2: ΕΙΣΑΓΩΓΗ ΣΤΟ WINBUGS**

⌘ **ΜΑΘΗΜΑ 3: ΣΥΝΘΕΤΑ ΠΑΡΑΔΕΙΓΜΑΤΑ**

⌘ **ΜΑΘΗΜΑ 4: ΣΥΓΚΡΙΣΗ ΜΟΝΤΕΛΩΝ**

Bayesian BioStatistics Using BUGS



ΜΑΘΗΜΑ 1: ΕΙΣΑΓΩΓΗ ΣΤΟ BUGS

Ioannis Ntzoufras

E-mail: ntzoufras@aueb.gr



Department of Statistics,
Athens University of
Economics & Business

ΠΕΡΙΕΧΟΜΕΝΑ 1ου ΜΑΘΗΜΑΤΟΣ

- ⌘ **1...ΕΙΣΑΓΩΓΗ ΣΤΑ ΓΕΝΙΚΕΥΜΕΝΑ
ΓΡΑΜΜΙΚΑ ΜΟΝΤΕΛΑ**
- ⌘ **2... ΕΙΣΑΓΩΓΗ ΣΤΟΥΣ ΑΛΓΟΡΙΘΜΟΥΣ
MCMC**
- ⌘ **3... ΜΠΕΥΖΙΑΝΗ
ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑ ΜΕ ΤΗ ΧΡΗΣΗ
ΤΟΥ BUGS**

1... ΕΙΣΑΓΩΓΗ ΣΤΑ ΓΕΝΙΚΕΥΜΕΝΑ ΓΡΑΜΜΙΚΑ ΜΟΝΤΕΛΑ

⌘ ΓΕΝΙΚΕΥΣΗ ΤΩΝ ΠΑΛΙΝΔΡΟΜΙΚΩΝ
ΜΟΝΤΕΛΩΝ

⌘ ΞΕΚΙΝΗΣΑΝ ΑΠΟ ΤΟΝ LEGENDRE
(1805) ΚΑΙ ΤΟΝ GAUSS (1809)

1.1. ΔΕΔΟΜΕΝΑ

⌘ **RESPONSE VARIABLE (Y)**: εξαρτημένη ή
ενδογενής μεταβλητή ή μεταβλητή
απόκρισης/αντίδρασης

☒ Y είναι τυχαία μεταβλητή

⌘ **EXPLANATORY VARIABLES (X_j)**:

Ανεξάρτητες ή εξογενείς ή επεξηγηματικές
μεταβλητές

☒ X_j θεωρούνται συνήθως σταθερές από το
σχεδιασμό του πειράματος

1.2. 3 ΚΥΡΙΕΣ ΣΥΝΙΣΤΩΣΕΣ

⌘ ΤΥΧΑΙΑ ΣΥΝΙΣΤΩΣΑ (random component)

☒ $Y_i \sim \text{ΚΑΤΑΝΟΜΗ}(\underline{\theta})$

☒ $\underline{\theta}$: διάνυσμα παραμέτρων

⌘ ΣΥΣΤΗΜΑΤΙΚΗ ΣΥΝΙΣΤΩΣΑ (systematic component)

☒ $\eta_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi}$

☒ η_i λέγεται γραμμικός προσδιορισμός του μοντέλου (linear predictor)

⌘ ΣΥΝΔΕΤΙΚΗ ΣΥΝΑΡΤΗΣΗ (link function)

☒ Ο τρόπος συνδεσης των παραμέτρων της τυχαίας συνιστώσας με το γραμμικό προσδιορισμό

☒ $g(\underline{\theta}) = \eta_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi}$

☒ συνήθως $\underline{\theta}$ είναι ο μέσος της Y

1.3. ΚΑΝΟΝΙΚΟ ΜΟΝΤΕΛΟ

⌘ ΤΥΧΑΙΑ ΣΥΝΙΣΤΩΣΑ:

☒ Y ποσοτική μεταβλητή

☒ $Y_i \sim \text{NORMAL}(\mu_i, \sigma^2)$

⌘ ΣΥΣΤΗΜΑΤΙΚΗ ΣΥΝΙΣΤΩΣΑ:

☒ X_j ποσοτικές ή κατηγορικές

⌘ ΣΥΝΔΕΤΙΚΗ ΣΥΝΑΡΤΗΣΗ

☒ $\mu_i = \eta_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi}$

☒ αρα συνδετική συνάρτηση: $g(\mu) = \mu$

1.4. ΜΟΝΤΕΛΑ BERNOULLI

⌘ ΤΥΧΑΙΑ ΣΥΝΙΣΤΩΣΑ:

☒ Y δίτιμη μεταβλητή (0/1)

☒ $Y_i \sim \text{Bernoulli}(p_i)$

⌘ ΣΥΣΤΗΜΑΤΙΚΗ ΣΥΝΙΣΤΩΣΑ:

☒ X_j ποσοτικές ή κατηγορικές

⌘ ΣΥΝΔΕΤΙΚΗ ΣΥΝΑΡΤΗΣΗ

☒ $\log(p_i / (1 - p_i)) = \eta_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi}$

☒ αρα συνδετική συνάρτηση: $g(p) = \text{logit}(p)$

1.5. ΔΙΩΝΥΜΙΚΑ ΜΟΝΤΕΛΑ

⌘ ΤΥΧΑΙΑ ΣΥΝΙΣΤΩΣΑ:

☒ Y : # επιτυχιών σε σύνολο n επαναλήψεων

☒ $Y_i \sim \text{Binomial} (p_i, n_i)$

⌘ ΣΥΣΤΗΜΑΤΙΚΗ ΣΥΝΙΣΤΩΣΑ:

☒ X_j ποσοτικές ή κατηγορικές

⌘ ΣΥΝΔΕΤΙΚΗ ΣΥΝΑΡΤΗΣΗ

☒ $\log (p_i / (1 - p_i)) = \eta_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi}$

☒ αρα συνδετική συνάρτηση: $g(p) = \text{logit}(p)$

1.6. ΜΟΝΤΕΛΑ POISSON

⌘ ΤΥΧΑΙΑ ΣΥΝΙΣΤΩΣΑ:

☒ Y : # επιτυχιών σε ένα χρονικό διάστημα

☒ $Y_i \sim \text{Poisson} (\lambda_i)$

⌘ ΣΥΣΤΗΜΑΤΙΚΗ ΣΥΝΙΣΤΩΣΑ:

☒ X_j ποσοτικές ή κατηγορικές

⌘ ΣΥΝΔΕΤΙΚΗ ΣΥΝΑΡΤΗΣΗ

☒ $\log (\lambda_i) = \eta_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi}$

☒ αρα συνδετική συνάρτηση: $g(\lambda) = \log(\lambda)$

1.7. ΓΕΝΙΚΕΣ ΑΡΧΕΣ ΜΟΝΤΕΛΟΠΟΙΗΣΗΣ

- ⌘ ΕΙΝΑΙ ΤΕΧΝΗ
- ⌘ ΟΛΑ ΤΑ ΜΟΝΤΕΛΑ ΕΙΝΑΙ ΛΑΘΟΣ
 - ☒ ΜΕΡΙΚΑ ΠΙΟ ΧΡΗΣΙΜΑ
 - ☒ ΨΑΧΝΟΥΜΕ ΕΝΑ ΜΟΝΤΕΛΟ ΠΟΥ ΠΕΡΙΓΡΑΦΕΙ ΠΕΡΙΛΗΠΤΙΚΑ ΤΗΝ ΠΡΑΓΜΑΤΙΚΟΤΗΤΑ
- ⌘ ΕΛΕΓΧΟΥΜΕ ΠΟΛΛΑ ΔΙΑΦΟΡΕΤΙΚΑ ΜΟΝΤΕΛΑ
- ⌘ ΠΡΕΠΕΙ ΝΑ ΓΙΝΕΤΑΙ ΕΛΕΓΧΟΣ ΚΑΛΗΣ ΠΡΟΣΑΡΜΟΓΗΣ

1.8. ΜΕΡΙΚΑ ΙΣΤΟΡΙΚΑ ΣΤΟΙΧΕΙΑ

- ⌘ **1805**: LEGENDRE:
 - ☒ ορίζει τα κατάλοιπα
 - ☒ Εκτιμάει τις παραμέτρους παλινδρόμησης μέσω των ελαχίστων τετραγώνων
- ⌘ **1809**: GAUSS: Εισάγει την Κανονική κατανομή

1.8. ΜΕΡΙΚΑ ΙΣΤΟΡΙΚΑ ΣΤΟΙΧΕΙΑ

⌘ 1805: LEGENDRE:

- ☒ ορίζει τα κατάλοιπα
- ☒ Εκτιμάει τις παραμέτρους παλινδρόμησης μέσω των ελαχίστων τετραγώνων

⌘ 1823: GAUSS (Theoria Combinationis)

- ☒ Εγκαταλείπει την κανονικότητα
- ☒ Η ομοσκεδαστικότητα διατηρεί κάποιες καλές ιδιότητες
- ☒ Wedderburn (1974, biometrika) απέδειξε το ίδιο για τα GLM

1.8. ΜΕΡΙΚΑ ΙΣΤΟΡΙΚΑ ΣΤΟΙΧΕΙΑ

⌘ 1919-1929: Fisher διατυπώνει τη θεωρία "Πειραματικού σχεδιασμού" (Experimental Design).

⌘ 1922: Ο Έλεγχος Περιεκτικότητας του Fisher και το 1ο γενικευμένο γραμμικό μοντέλο (RSS)

⌘ 1935: Η Μελέτη βιολογική περιεκτικότητας του Bliss. Τα μοντέλα Probit (Ann.Appl.Biol.).

⌘ 1944 & 1951: Ο Berkston εισάγει τα μοντέλα logit (JASA & Biometrics)

⌘ 1952 (biometrics) Dyke & Patterson εφαρμογή λογιστικής παλινδρόμησης σε κατηγορικά δεδομένα με 4 μεταβλητές

Bayesian Biostatistics Using BUGS

2...ΕΙΣΑΓΩΓΗ ΣΤΟΥΣ ΑΛΓΟΡΙΘΜΟΥΣ MCMC



Department of Statistics,
Athens University of
Economics & Business

2...ΕΙΣΑΓΩΓΗ ΣΤΟΥΣ ΑΛΓΟΡΙΘΜΟΥΣ MCMC

- ⌘ Η ΜΠΕΥΖΙΑΝΗ ΠΡΟΣΕΓΓΙΣΗ
- ⌘ ΠΡΟΣΟΜΟΙΩΣΗ ΑΠΟ ΤΗΝ
ΕΚ-ΤΩΝ-ΥΣΤΕΡΩΝ (posterior) ΚΑΤΑΝΟΜΗ
 - ☒ Ο Αλγόριθμος Metropolis-Hastings
 - ☒ Ο Δειγματολήπτης Gibbs
- ⌘ ΠΑΡΑΔΕΙΓΜΑ ΑΠΛΗΣ ΠΑΛΙΝΔΡΟΜΙΣΗΣ

2.1. Η Μπευζιανή προσέγγιση

⌘ Κλασσική στατιστική:

- ☒ βασίζεται στην πιθανοφάνεια $f(\mathbf{y}|\underline{\theta})$
- ☒ $\underline{\theta}$ διάνυσμα παραμέτρων => άγνωστες προς εκτίμηση ποσότητες
- ☒ Εκτίμηση γίνεται μέσω εκτιμητριών συναρτήσεων με κάποιες καλές ιδιότητες (π.χ. μεροληψία)
- ☒ Εκτιμήτριες βρίσκονται μεγιστοποιώντας την πιθανοφάνεια
- ☒ ΠΑΡΑΔΕΙΓΜΑ: για $Y_i \sim N(\mu, \sigma^2)$
 - ☒ μ εκτιμάται από το δειγματικό μέσο $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$

⌘ Στατιστική κατά Μπέυες:

- ☒ Θεωρεί τις παραμέτρους τυχαίες μεταβλητές
- ☒ Ορίζει εκ-των-προτέρων (prior) κατανομών $f(\underline{\theta})$
- ☒ βασίζεται στην εκ-των-υστερων (posterior) κατανομή $f(\underline{\theta}|\mathbf{y})$

⌘ Πλεονεκτήματα

- ☒ Πιο καθάρη και πιθανοθεωρητική προσέγγιση
- ☒ Μπορεί να συμπεριλάβει πληροφορία ειδικών ή αποτελέσματα άλλων μελετών (meta-analysis)

⌘ Μειονεκτήματα

- ☒ Υποκειμενικότητα αποτελεσμάτων λόγω της prior
- ☒ Δυσκολίες υπολογισμού των posterior Κατανομών

⌘ Η Εκ-των-υστέρων κατανομή υπολογίζεται από το Θεώρημα του Bayes

$$\text{⌘ } f(\underline{\theta}|\underline{y}) = f(\underline{\theta}, \underline{y}) / f(\underline{y})$$

$$\text{⌘ } = f(\underline{y}|\underline{\theta})f(\underline{\theta}) / f(\underline{y})$$

$$\text{⌘ } \propto f(\underline{y}|\underline{\theta})f(\underline{\theta})$$

$$\text{⌘ Posterior} \propto \text{Likelihood} \times \text{Prior}$$

ΠΑΡΑΔΕΙΓΜΑ: ΕΚΤΙΜΗΣΗ ΜΕΣΟΥ

⌘ **ΠΙΘΑΝΟΦΑΝΕΙΑ:** $Y_i \sim N(\mu, \sigma^2)$

☒ σ^2 γνωστή σταθερή ποσότητα

⌘ **PRIOR:** $\mu \sim N(\mu_0, \tau^2)$

⌘ **POSTERIOR:**

$$f(\mu|\mathbf{y}) = N(w \bar{y} + (1-w) \mu_0, w \sigma^2/n)$$

$$w = \tau^2 / (\tau^2 + \sigma^2/n)$$

2.2. ΠΡΟΣΟΜΟΙΩΣΗ ΑΠΟ ΤΗΝ ΕΚ-ΤΩΝ-ΥΣΤΕΡΩΝ ΚΑΤΑΝΟΜΗ

⌘ Υπολογισμός της posterior Κατανομής είναι δύσκολος.

☒ Συζυγείς εκ-των-προτέρων κατανομές (conjugate priors, 70s)

☒ Ασυμπτωτικές Προσεγγίσεις (80s)

☒ Προσομοίωση μέσω MCMC (90s)

2.2. ΠΡΟΣΟΜΟΙΩΣΗ ΑΠΟ ΤΗΝ ΕΚ-ΤΩΝ-ΥΣΤΕΡΩΝ ΚΑΤΑΝΟΜΗ

⌘ MCMC

- ☒ Προϋπήρχαν σε άλλες επιστήμες
- ☒ 1954 Metropolis et al. (Metropolis Algorithm)
- ☒ 1970 Hastings (Metropolis-Hastings Algorithm)
- ☒ 1984 Geman and Geman (Gibbs Sampling)
- ☒ 1990 Smith et al (Εφαρμογή των MCMC σε Μπευζιανά Προβλήματα)
- ☒ 1995 Green (Reversible Jump MCMC)

2.2. ΠΡΟΣΟΜΟΙΩΣΗ ΑΠΟ ΤΗΝ ΕΚ-ΤΩΝ-ΥΣΤΕΡΩΝ ΚΑΤΑΝΟΜΗ

Η ΙΔΕΑ:

ΑΦΟΥ ΔΕΝ ΜΠΟΡΟΥΜΕ ΝΑ ΥΠΟΛΟΓΙΣΟΥΜΕ
ΤΗΝ ΕΚ-ΤΩΝ-ΥΣΤΕΡΩΝ ΚΑΤΑΝΟΜΗ ΤΟΤΕ
ΝΑ ΒΡΟΥΜΕ ΕΝΑ ΑΛΓΟΡΙΘΜΟ ΚΑΙ ΝΑ
ΠΡΟΣΟΜΟΙΩΣΟΥΜΕ ΤΥΧΑΙΕΣ ΤΙΜΕΣ ΑΠΟ
ΑΥΤΗ

2.2. ΠΡΟΣΟΜΟΙΩΣΗ ΑΠΟ ΤΗΝ ΕΚ-ΤΩΝ-ΥΣΤΕΡΩΝ ΚΑΤΑΝΟΜΗ

Η ΛΟΓΙΚΗ:

ΦΤΙΑΧΝΟΥΜΕ ΜΙΑ ΜΑΡΚΟΒΙΑΝΗ ΑΛΥΣΙΔΑ ΤΗΣ
ΟΠΟΙΑΣ Η ΣΤΑΣΙΜΗ ΚΑΤΑΝΟΜΗ ΕΙΝΑΙ ΙΣΗ ΜΕ
ΤΗ ΖΗΤΟΥΜΕΝΗ (ΕΚ-ΤΩΝ-ΥΣΤΕΡΩΝ)
ΚΑΤΑΝΟΜΗ.

ΚΑΘΕ ΒΗΜΑ ΕΞΑΡΤΑΤΑΙ ΜΟΝΟ ΑΠΟ ΤΟ
ΠΡΟΗΓΟΥΜΕΝΟ

ΧΡΗΣΙΜΟΠΟΙΟΥΜΕ ΑΥΤΗ ΤΗΝ ΑΛΥΣΙΔΑ ΓΙΑ ΝΑ
“ΚΑΤΑΣΚΕΥΑΣΟΥΜΕ” ΕΝΑ ΔΕΙΓΜΑ ΑΠΟ ΤΗ
ΣΤΑΣΙΜΗ ΚΑΤΑΝΟΜΗ

2.2. ΠΡΟΣΟΜΟΙΩΣΗ ΑΠΟ ΤΗΝ ΕΚ-ΤΩΝ-ΥΣΤΕΡΩΝ ΚΑΤΑΝΟΜΗ

⌘ Η ΔΙΑΔΙΚΑΣΙΑ

- ☒ ΟΡΙΖΟΥΜΕ ΑΥΘΑΙΡΕΤΑ ΚΑΠΟΙΕΣ ΑΡΧΙΚΕΣ
ΤΙΜΕΣ ΓΙΑ ΤΙΣ ΠΑΡΑΜΕΤΡΟΥΣ ΜΑΣ $\theta^{(0)}$
- ☒ ΓΙΑ $t=1, \dots, T$ ΓΕΝΝΑΜΕ ΤΙΜΕΣ ΣΥΜΦΩΝΑ ΜΕ
ΤΟΝ ΑΛΓΟΡΙΘΜΟ ΜΑΣ
- ☒ ΟΤΑΝ ΒΕΒΑΙΩΘΟΥΜΕ ΟΤΙ Η ΑΛΥΣΙΔΑ ΓΕΝΝΑΕΙ
ΤΙΜΕΣ ΑΠΟ ΤΗ ΣΤΑΣΙΜΗ ΔΙΑΔΙΚΑΣΙΑ (ΔΗΛΑΔΗ
ΕΧΕΙ **ΣΥΓΚΛΙΝΕΙ**) ΣΤΑΜΑΤΑΜΕ
- ☒ ΠΕΤΑΜΕ ΤΙΣ ΠΡΩΤΕΣ K ΤΙΜΕΣ ΓΙΑ ΝΑ
ΑΠΟΦΥΓΟΥΜΕ ΤΥΧΟΝ ΕΠΙΡΡΟΗ ΤΟΥ ΑΡΧΙΚΟΥ
ΣΗΜΕΙΟΥ.

2.2. ΠΡΟΣΟΜΟΙΩΣΗ ΑΠΟ ΤΗΝ ΕΚ-ΤΩΝ-ΥΣΤΕΡΩΝ ΚΑΤΑΝΟΜΗ

⌘ ΟΡΟΛΟΓΙΑ

- ☒ **INITIAL VALUES (ΑΡΧΙΚΕΣ ΤΙΜΕΣ)** ΤΗΣ ΑΛΥΣΙΔΑΣ ΛΕΓΟΝΤΑΙ ΟΙ ΤΙΜΕΣ $\theta^{(0)}$ ΑΠΟ ΟΠΟΥ ΞΕΚΙΝΑΜΕ ΤΟΝ ΑΛΓΟΡΙΘΜΟ ΜΑΣ
- ☒ **ITERATION (ΕΠΑΝΑΛΗΨΗ)**: ΚΑΘΕ ΠΡΟΣΟΜΟΙΩΜΕΝΗ ΠΑΡΑΤΗΡΗΣΗ ΑΝΤΙΣΤΟΙΧΕΙ ΣΕ ΜΙΑ ΕΠΑΝΑΛΗΨΗ ΤΟΥ ΑΛΓΟΡΙΘΜΟΥ
- ☒ **BURN-IN PERIOD**: ΟΙ ΠΕΡΙΟΔΟΣ ΜΕΧΡΙ ΝΑ ΑΡΧΙΣΟΥΜΕ ΝΑ ΠΕΡΝΟΥΜΕ ΤΙΜΕΣ ΑΠΟ ΤΗ ΣΤΑΣΙΜΗ ΚΑΤΑΝΟΜΗ

2.2. ΠΡΟΣΟΜΟΙΩΣΗ ΑΠΟ ΤΗΝ ΕΚ-ΤΩΝ-ΥΣΤΕΡΩΝ ΚΑΤΑΝΟΜΗ

⌘ ΟΡΟΛΟΓΙΑ

- ☒ **CONVERGENCE (ΣΥΓΚΛΙΣΗ)**: ΟΤΑΝ Η ΔΙΑΔΙΚΑΣΙΑ ΜΑΣ ΕΧΕΙ ΔΩΣΕΙ ΤΙΜΕΣ ΑΠΟ ΤΗ ΣΤΑΣΙΜΗ ΚΑΤΑΝΟΜΗ
- ☒ **CONVERGENCE DIAGNOSTICS**: ΔΙΑΓΝΩΣΤΙΚΟΙ ΕΛΕΓΧΟΙ ΣΥΓΓΛΙΣΗΣ
- ☒ **EQUILIBRIUM**: ΣΤΑΣΙΜΗ ΚΑΤΑΝΟΜΗ
- ☒ **MCMC OUTPUT**: ΤΟ ΠΡΟΣΟΜΟΙΩΜΕΝΟ ΔΕΙΓΜΑ

2.2. ΠΡΟΣΟΜΟΙΩΣΗ ΑΠΟ ΤΗΝ ΕΚ-ΤΩΝ-ΥΣΤΕΡΩΝ ΚΑΤΑΝΟΜΗ

⌘ ΟΡΟΛΟΓΙΑ

- ☒ ΕΠΕΙΔΗ ΟΙ ΑΛΓΟΡΙΘΜΟΙ MCMC ΒΑΣΙΖΟΝΤΑΙ ΣΕ ΜΑΡΚΟΒΙΑΝΕΣ ΑΛΥΣΙΔΕΣ, ΤΟ ΠΡΟΣΟΜΟΙΩΜΕΝΟ ΔΕΙΓΜΑ ΔΕΝ ΕΙΝΑΙ I.I.D.
- ☒ ΕΞΑΛΕΙΨΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗΣ
 - ☒ ΒΛΕΠΟΥΜΕ ΤΙΣ ΑΥΤΟΣΥΣΧΕΤΙΣΕΙΣ (AUTOCORRELATIONS)
 - ☒ ΚΡΑΤΑΜΕ 1 ΠΑΡΑΤΗΡΗΣΗ ΑΝΑ L ΕΠΑΝΑΛΗΨΕΙΣ.
- ☒ Η ΠΟΣΟΤΗΤΑ L ΛΕΓΕΤΑΙ **ΤΗΙΝ (ΛΕΠΤΥΝΣΗ)** ΤΗΣ ΑΛΥΣΙΔΑΣ
- ☒ ΛΕΠΤΥΝΣΗ ΕΝΟΣ ΠΡΟΣΟΜΟΙΩΜΕΝΟΥ ΔΕΙΓΜΑΤΟΣ ΜΠΟΡΕΙ ΝΑ ΓΙΝΕΙ ΚΑΙ ΛΟΓΩ ΑΝΑΛΥΣΗΣ Η ΑΠΟΘΗΚΕΥΤΙΚΟΥ ΧΩΡΟΥ

2.2. ΠΡΟΣΟΜΟΙΩΣΗ ΑΠΟ ΤΗΝ ΕΚ-ΤΩΝ-ΥΣΤΕΡΩΝ ΚΑΤΑΝΟΜΗ

⌘ ΑΛΓΟΡΙΘΜΟΙ

- ☒ ΑΛΓΟΡΙΘΜΟΣ METROPOLIS-HASTINGS
- ☒ ΔΕΙΓΜΑΤΟΛΗΠΤΗΣ GIBBS

2.2.1. ΑΛΓΟΡΙΘΜΟΣ METROPOLIS-HASTINGS

Έστω $\underline{\theta}^{\text{old}}$ η τρέχουσα τιμή των παραμέτρων

⌘ Δειγματοληπτούμε $\underline{\theta}^{\text{can}}$ από μία *κατανομή εισήγησης (proposal distribution)* $q(\underline{\theta}^{\text{can}} | \underline{\theta}^{\text{old}})$.

⌘ Υπολογίζουμε $a = \min \left\{ 1, \frac{f(\underline{\theta}^{\text{can}} | \underline{y}) q(\underline{\theta}^{\text{old}} | \underline{\theta}^{\text{can}})}{f(\underline{\theta}^{\text{old}} | \underline{y}) q(\underline{\theta}^{\text{can}} | \underline{\theta}^{\text{old}})} \right\}$

⌘ Θέτουμε $\underline{\theta}^{\text{new}} = \underline{\theta}^{\text{can}}$ με πιθανότητα a και $\underline{\theta}^{\text{new}} = \underline{\theta}^{\text{old}}$ με πιθανότητα $(1-a)$

2.2.1. ΑΛΓΟΡΙΘΜΟΣ METROPOLIS-HASTINGS

Συνήθως

⌘ $q(\underline{\theta}^{\text{can}} | \underline{\theta}^{\text{old}}) = N(\underline{\theta}^{\text{old}}, c^2)$.

⌘ c^2 λέγεται και tuning parameter και επηρεάζει τη σύγκλιση. Επιλέγεται έτσι ώστε να δεχόμαστε 30-40%

⌘ Η πιθανότητα αποδοχής a απλοποιείται

$$a = \min \left\{ 1, \frac{f(\underline{\theta}^{\text{can}} | \underline{y})}{f(\underline{\theta}^{\text{old}} | \underline{y})} \right\}$$

2.2.2. ΔΕΙΓΜΑΤΟΛΗΠΤΗΣ GIBBS

Έστω $\underline{\theta}^{\text{old}}$ η τρέχουσα τιμή των παραμέτρων
και $\underline{\theta}^{\text{old}} = (\theta_1^{\text{old}}, \dots, \theta_p^{\text{old}})^T$

$$\text{⌘ } \theta_1^{\text{new}} \sim f(\theta_1 | \theta_2^{\text{old}}, \dots, \theta_p^{\text{old}}, \mathbf{y})$$

⌘ ...

$$\text{⌘ } \theta_j^{\text{new}} \sim f(\theta_j | \theta_1^{\text{new}}, \dots, \theta_{j-1}^{\text{new}}, \theta_{j+1}^{\text{old}}, \dots, \theta_p^{\text{old}}, \mathbf{y})$$

⌘ ...

$$\text{⌘ } \theta_p^{\text{new}} \sim f(\theta_p | \theta_1^{\text{new}}, \dots, \theta_{p-1}^{\text{new}}, \mathbf{y})$$

2.2.2. ΔΕΙΓΜΑΤΟΛΗΠΤΗΣ GIBBS

$$\text{⌘ } f(\theta_j | \theta_1^{\text{new}}, \dots, \theta_{j-1}^{\text{new}}, \theta_{j+1}^{\text{old}}, \dots, \theta_p^{\text{old}}, \mathbf{y})$$

ονομάζεται full conditional posterior
distribution και συμβολίζεται $f(\theta_j | \bullet)$

2.2.2. ΔΕΙΓΜΑΤΟΛΗΠΤΗΣ GIBBS

⌘ Διαφορές με M-H

- ☒ $\underline{\theta}^{\text{can}} \neq \underline{\theta}^{\text{old}}$ σε κάθε επανάληψη
- ☒ Gibbs υπο-περίπτωση M-H για $q() = f(\theta_j | \bullet)$
- ☒ Κάθε φορά ανανεώνουμε μία-μία τις τιμές
- ☒ $f(\theta_j | \bullet)$ μπορεί να είναι άγνωστη
 - ☒ => Χρήση adaptive rejection sampling για log-concave κατανομές (Gilks & Wild, 1992)
 - ☒ Στα GLM οι posterior είναι log-concave (Dellaportas & Smith, 1993)
 - ☒ Χρησιμοποιείται στο BUGS

2.3 ΠΑΡΑΔΕΙΓΜΑ: GIBBS SAMPLING ΣΤΗΝ ΑΠΛΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

- ⌘ $Y_i \sim N(\mu_i, \sigma^2)$ για $i=1,2,\dots,n$
- ⌘ $\mu_i = \alpha + \beta X_i$
- ⌘ $\underline{\theta} = (\alpha, \beta, \sigma^2)^T$
- ⌘ PRIORS: $f(\underline{\theta}) = f(\alpha, \beta, \sigma^2) = f(\alpha)f(\beta)f(\sigma^2)$
 - ☒ $f(\alpha) = \text{Normal}(\mu_\alpha, \tau_\alpha^2)$
 - ☒ $f(\beta) = \text{Normal}(\mu_\beta, \tau_\beta^2)$
 - ☒ $f(\sigma^2) = \text{Inverse Gamma}(\gamma, \delta)$
 - ☒ αν $\tau = \sigma^{-2}$ τότε $f(\tau) = \text{Gamma}(\gamma, \delta)$

2.3 ΠΑΡΑΔΕΙΓΜΑ: GIBBS SAMPLING ΣΤΗΝ ΑΠΛΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

⌘ Full Conditional Posteriors

$$\boxtimes f(\alpha|\beta, \sigma^2, \mathbf{y}) = N\left(w_1(\bar{y} - b\bar{x}) + (1 - w_1)\mu_\alpha, w_1 \frac{\sigma^2}{n}\right)$$

$$\boxtimes w_1 = \tau^2 / (\tau^2 + \sigma^2/n)$$

$$\boxtimes f(\beta|\alpha, \sigma^2, \mathbf{y}) = N\left(w_2 \frac{\sum_{i=1}^n x_i y_i - an\bar{x}}{\sum_{i=1}^n x_i^2} + (1 - w_2)\mu_\beta, w_2 \frac{\sigma^2}{\sum_{i=1}^n x_i^2}\right)$$

$$\boxtimes w_2 = \tau^2 / (\tau^2 + \sigma^2 / \sum x_i^2)$$

2.3 ΠΑΡΑΔΕΙΓΜΑ: GIBBS SAMPLING ΣΤΗΝ ΑΠΛΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

⌘ Full Conditional Posteriors

$$\boxtimes f(\sigma^2|\alpha, \beta, \mathbf{y}) = \text{IG}\left(\gamma + n/2, \delta + \sum (y_i - \alpha - \beta x_i)^2/2\right)$$

Bayesian Biostatistics Using BUGS

3... Μπεϋζιανή Συμπερασματολογία με τη Χρήση του BUGS



Department of Statistics,
Athens University of
Economics & Business

3.1. Εισαγωγή: Τι είναι το BUGS

- ⌘ BUGS: Bayesian inference Using Gibbs Sampling
- ⌘ Γλώσσα προγραμματισμού που ορίζουμε μοντέλο (πιθανοφάνεια, prior)
- ⌘ Υπολογίζει τις full conditional και προσομοιώνει από log-concnave posterior κατανομές

3.1. Εισαγωγή: Τι είναι το BUGS

- ⌘ Ξεκίνησε γύρω στο 1995
- ⌘ Γύρω στο 1998 βγήκε η έκδοση για Windows (Winbugs)
- ⌘ Οφείλεται σε μια ομάδα του MRC στο Cambridge (Spiegelhalter, Gilks, Best, Thomas)

3.1. Εισαγωγή: Τι είναι το BUGS

ΠΡΟΣΟΧΗ:

**Ο ΔΕΙΓΜΑΤΟΛΗΠΤΗΣ GIBBS ΜΠΟΡΕΙ
ΝΑ ΕΙΝΑΙ ΕΠΙΚΙΝΔΥΝΟΣ
ΓΙΑΤΙ;**

3.1. Εισαγωγή: Τι είναι το BUGS

ΓΙΑΤΙ;

- ⌘ ΛΑΘΟΣ ΑΠΟΤΕΛΕΣΜΑΤΑ ΛΟΓΩ ΜΗ ΣΥΓΚΛΙΣΗΣ
- ⌘ ΠΟΛΥ ΑΡΓΗ ΣΥΓΚΛΙΣΗ
- ⌘ ΚΑΚΕΣ ΑΡΧΙΚΕΣ ΤΙΜΕΣ
- ⌘ ΚΑΚΗ ΕΠΙΛΟΓΗ ΜΟΝΤΕΛΟΥ
- ⌘ ΥΠΕΡ-ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ (over-parametrization)
- ⌘ ΣΠΑΣΙΜΟ ΝΕΥΡΩΝ

3.1. Εισαγωγή: Τι είναι το BUGS

ΓΙΑΤΙ;

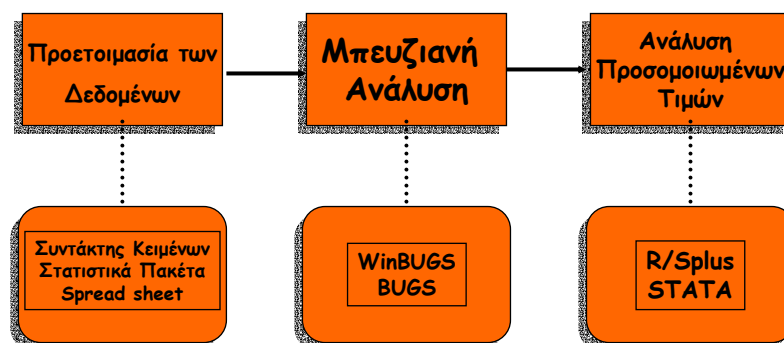
- ⌘ ΛΑΘΟΣ ΑΠΟΤΕΛΕΣΜΑΤΑ ΛΟΓΩ ΜΗ ΣΥΓΚΛΙΣΗΣ
- ⌘ ΠΟΛΥ ΑΡΓΗ ΣΥΓΚΛΙΣΗ
- ⌘ ΚΑΚΕΣ ΑΡΧΙΚΕΣ ΤΙΜΕΣ
- ⌘ ΚΑΚΗ ΕΠΙΛΟΓΗ ΜΟΝΤΕΛΟΥ
- ⌘ ΥΠΕΡ-ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ (over-parametrization)
- ⌘ ΣΠΑΣΙΜΟ ΝΕΥΡΩΝ

3.1. Εισαγωγή: Τι είναι το BUGS

ΑΡΧΕΙΑ BUGS

- ⌘ ***.DAT**: Αρχείο Δεδομένων
- ⌘ ***.INI**: Αρχείο Αρχικών τιμών
- ⌘ ***.BUG**: Αρχείο Μοντέλου
- ⌘ ***.CMD**: Αρχείο Εντολών προσομοίωσης
- ⌘ ***.LOG**: Αρχείο Αποτελεσμάτων προσομοίωσης
- ⌘ ***.OUT**: Αρχείο Προσομοιωμένων τιμών
- ⌘ ***.IND**: Αρχείο με Περιεχόμενα του ***.OUT**

3.1. Εισαγωγή: Τι είναι το BUGS



3.2. Ένα Απλό Παράδειγμα στο BUGS

⌘ Green & Touchston (1963, *Am. Jour. Of Obstetrics & Gynecology*)

⌘ Μελέτη σχέσης

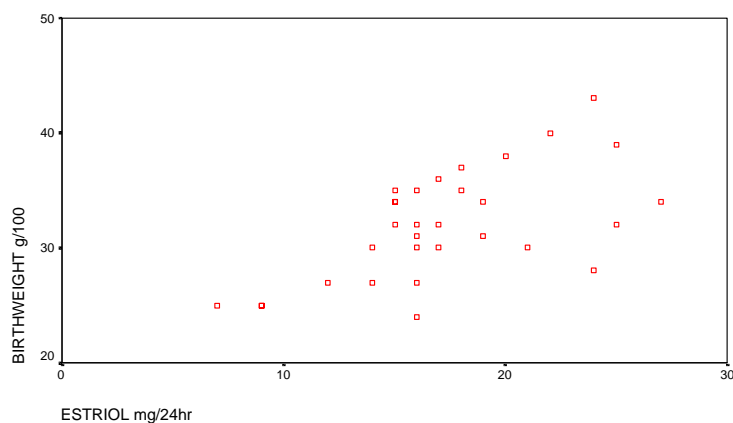
☒ Y : Βάρος γέννησης (birthweight) ενός παιδιού

☒ X : Επίπεδο οιστριόλης (estriol) των εγκύων γυναικών

☒ n=31

⌘ Το γράφημα που ακολουθεί δείχνει ότι υπάρχει συσχέτιση

3.2. Ένα Απλό Παράδειγμα στο BUGS



3.2.1. Χτίζοντας το Μοντέλο

- ⌘ ΤΥΧΑΙΑ ΣΥΝΙΣΤΩΣΑ: $\text{Birth}_i \sim \text{Normal}(\mu_i, \sigma^2)$
- ⌘ ΣΥΣΤΗΜΑΤΙΚΗ ΣΥΝΙΣΤΩΣΑ: $\eta_i = \alpha + \beta \times \text{Estriol}_i$
- ⌘ ΣΥΝΔΕΤΙΚΗ ΣΥΝΑΡΤΗΣΗ: $\mu_i = \eta_i = \alpha + \beta \times \text{Estriol}_i$
- ⌘ για $i=1, \dots, 31$
- ⌘ PRIORS (Non-informative)
 - ☒ $f(\alpha) = \text{Normal}(0, 10^4)$
 - ☒ $f(\beta) = \text{Normal}(0, 10^4)$
 - ☒ $f(\sigma^2) = \text{Inverse Gamma}(10^{-4}, 10^{-4})$ ή
 - ☒ $f(\tau = \sigma^{-2}) = \text{Gamma}(10^{-4}, 10^{-4})$

3.2.2. Δημιουργία Μοντέλων στο BUGS

- ⌘ ΣΕ *.BUG Ορίζουμε το μοντέλο μας
- ⌘ Εντολές => BUGS MANUAL σελ 17-18
- ⌘ ΔΟΜΗ:
 - ☒ **Προκαταρκτικό κομμάτι:** δηλώνουμε μεταβλητές, σταθερές, δεδομένα και αρχικές τιμές
 - ☒ **Κυρίως μοντέλο:** εδώ ορίζουμε πιθανοφάνεια και priors

3.2.2. Δημιουργία Μοντέλων στο BUGS

⌘ ΠΡΟΚΑΤΑΡΚΤΙΚΟ ΚΟΜΜΑΤΙ

- ☒ `<;>`: τελειώνει κάθε εντολή του BUG file
- ☒ `model <όνομα μοντέλου>;`
- ☒ `const <σταθερά 1=#>, ..., <σταθερά κ=#>;` Ορισμός σταθερών παραμέτρων
- ☒ `var <μεταβλητή 1>, ... , <μεταβλητή κ>;` Ορισμός τυχαίων μεταβλητών
- ☒ `data <μεταβλητή 1>, ... , <μεταβλητή κ> in '<ονομα αρχείου>'`, `<μεταβλητή κ+1>, ... , <μεταβλητή κ+λ> in '<ονομα αρχείου2>'`: ορισμός αρχείων δεδομένων
- ☒ `inits in '<ονομα αρχείου>'`: ορισμός αρχείων με αρχικές τιμές

3.2.2. Δημιουργία Μοντέλων στο BUGS

⌘ ΚΥΡΙΩΣ ΜΟΝΤΕΛΟ

- ☒ Το κυρίως μοντέλο αρχίζει και τελειώνει με `{}`
- ☒ `~` : ορίζουμε τις τυχαίες μεταβλητές
- ☒ `<-` : ισότητα/ ανάθεση
- ☒ Οι κατανομές ορίζονται στις σελίδες 17-18
- ☒ Αν θέλουμε να περιορίσουμε μία κατανομή σε ένα διάστημα (α, β) τότε η κατανομή ακολουθείται από **$I(\alpha, \beta)$**

3.2.2. Δημιουργία Μοντέλων στο BUGS

⌘ ΚΥΡΙΩΣ ΜΟΝΤΕΛΟ

☒ **Κανονική Κατανομή:** $y \sim \text{dnorm}(\mu, \tau)$

☒ **μ** = μέσος

☒ **τ** = ακρίβεια (precision) = $1/\sigma^2$

☒ **Γάμμα Κατανομή:** $y \sim \text{dgamma}(a, b)$

☒ **μέσος** = a/b

☒ **$x[i]$** : i στοιχείο του διανύσματος x

☒ **$d[i, j]$** : στοιχείο της i γραμμής και j στήλης του πίνακα d

3.2.3. Εισαγωγή του Μοντέλου του Παραδείγματος στο BUGS

(1) $\text{Birth}_i \sim \text{Normal}(\mu_i, \sigma^2)$ ⌘ for (i in 1:n) {
 (2) $\eta_i = a + \beta \times \text{Estriol}_i$ ⌘ Birth[i] ~ dnorm(mu[i], tau)
 (3) $\mu_i = \eta_i = a + \beta \times \text{Estriol}_i$ }
 για $i = 1, \dots, 31$ ⌘ mu[i] <- a + b * estriol[i]
 ⌘ }

PRIORS

☒ $f(a) = \text{Normal}(0, 10^4)$ ⌘ a ~ dnorm(0.0, 1.0E-04)
 ☒ $f(\beta) = \text{Normal}(0, 10^4)$ ⌘ b ~ dnorm(0.0, 1.0E-04)
 ☒ $f(\tau = \sigma^{-2}) = \text{Gamma}(10^{-4}, 10^{-4})$ ⌘ tau ~ dgamma(1.0E-04, 1.0E-04)
 ☒ $\sigma^2 = 1/\tau$ ⌘ s2 <- 1/tau

3.2.3. Εισαγωγή του Μοντέλου του Παραδείγματος στο BUGS

ΠΡΟΚΑΤΑΡΚΤΙΚΟ ΚΟΜΜΑΤΙ

```
% model example1;
% const n=31; # n=sample size
% var   estriol[n], # estriol level of pregnant woman
%       birth[n],   # birthweight
%       mu[n],      # regression expected value
%       a,b,tau,s2; # model parameters,
%                   # tau = precision, s2=1/tau variance
% data estriol,birth in 'estriol.dat';
% inits in 'estriol.in';
```

3.2.4. Προσομοιώνοντας από την posterior

ΕΝΤΟΛΕΣ ΠΡΟΣΟΜΟΙΩΣΗΣ (σελ 31, BUGS MANUAL)

% bugs [enter]	ΞΕΚΙΝΗΜΑ ΤΟΥ BUGS
% compile('estriol.bug')	ΤΣΕΚΑΡΕΙ ΤΗ ΣΥΝΤΑΞΗ ΤΟΥ ΜΟΝΤΕΛΟΥ, ΒΡΙΣΚΕΙ ΤΙΣ CONDITIONALS ΚΑΙ ΤΙΣ ΑΡΧΙΚΕΣ ΤΙΜΕΣ
% update(1000)	ΠΡΟΣΟΜΟΙΩΝΕΙ 1000 ΤΙΜΕΣ
% monitor(a)	} ΑΠΟ ΤΗ ΣΤΙΓΜΗ ΑΥΤΗ ΑΠΟΘΗΚΕΥΕΙ ΤΙΣ ΠΡΟΣΟΜΟΙΩΜΕΝΕΣ ΤΙΜΕΣ ΑΠΟΘΗΚΕΥΣΗ ΑΝΑ 10 (THIN=10)
% monitor(b)	
% monitor(s2)	
% monitor(s2,10)	

3.2.4. Προσομοιώνοντας από την posterior

ΕΝΤΟΛΕΣ ΠΡΟΣΟΜΟΙΩΣΗΣ (σελ 31, BUGS MANUAL)

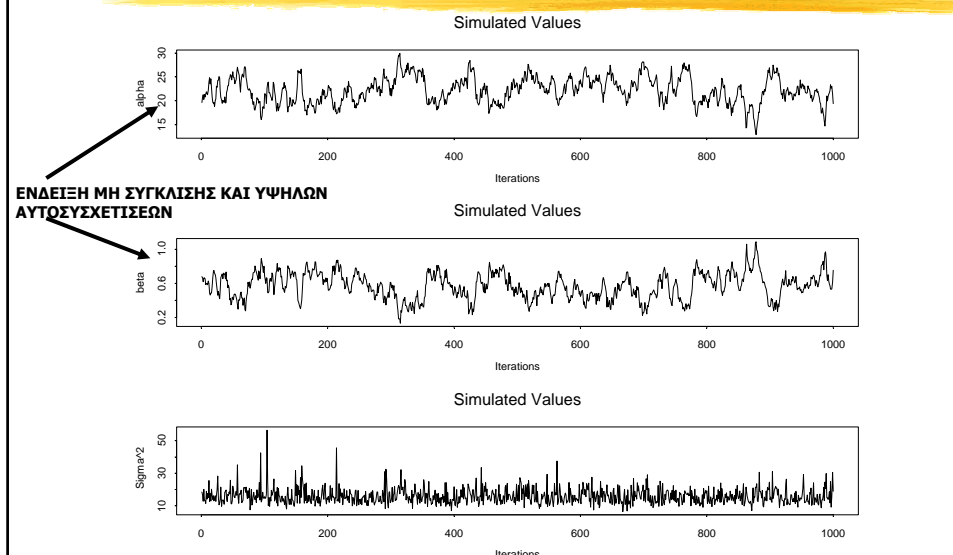
⌘ update(1000)	ΠΡΟΣΟΜΟΙΩΝΟΥΜΕ 1000 ΤΙΜΕΣ
⌘ stats(a)	} ΣΤΑΤΙΣΤΙΚΟΙ ΔΕΙΚΤΕΣ ΤΩΝ ΠΡΟΣΟΜΟΙΩΜΕΝΩΝ ΤΙΜΩΝ (ΤΗΣ posterior ΚΑΤΑΝΟΜΗΣ)
⌘ stats(b)	
⌘ stats(s2)	
⌘ diag(a)	} ΔΙΑΓΝΩΣΤΙΚΟ ΤΕΣΤ ΣΥΓΚΛΙΣΗΣ
⌘ diag(b)	
⌘ diag(s2)	
⌘ q()	ΕΞΟΔΟΣ ΑΠΟ ΤΟ BUGS
⌘ EDIT BUGS.LOG	ΒΛΕΠΟΥΜΕ ΟΛΑ ΤΑ ΑΠΟΤΕΛΕΣΜΑΤΑ
⌘ EDIT BUGS.IND	ΠΕΡΙΕΧΟΜΕΝΑ BUGS.OUT
⌘ ΕΙΣΑΓΟΥΜΕ ΤΟ BUGS.OUT ΣΕ ΣΤΑΤΙΣΤΙΚΟ ΠΑΚΕΤΟ ΓΙΑ ΑΝΑΛΥΣΗ	

3.2.4. Προσομοιώνοντας από την posterior

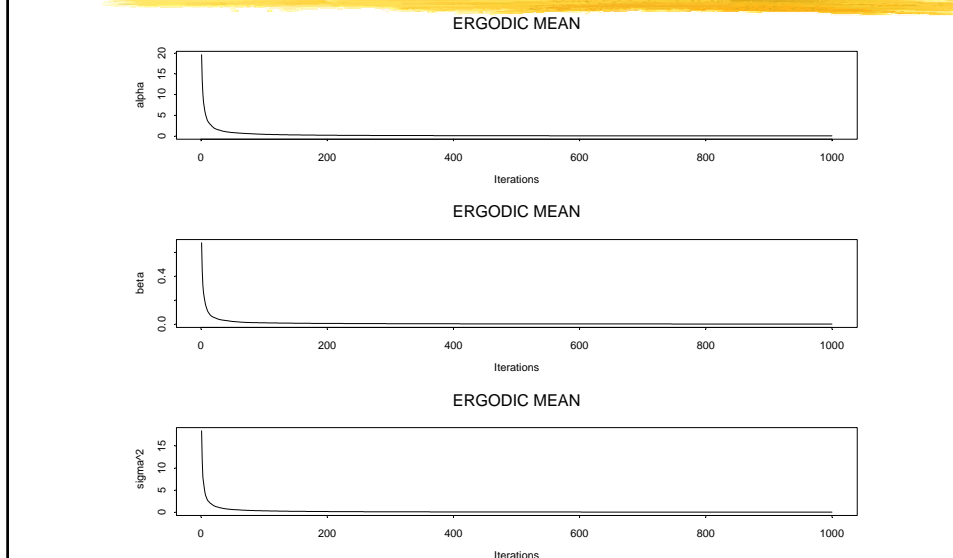
ΠΡΟΣΟΜΟΙΩΣΗ ΣΤΟ ΠΑΡΑΣΚΗΝΙΟ

- ⌘ Γράφουμε τις εντολές ένα αρχείο με κατάληξη CMD πχ.
Example1.cmd
- ⌘ Τρέχουμε το μοντέλο στο παρασκήνιο με την εντολή
backbugs <όνομα αρχείου CMD>
π.χ. **backbugs example.cmd**
- ⌘ Τα αποτελέσματα τα βλέπουμε με τον ίδιο τρόπο όπως παραπάνω δηλ.
- ⌘ EDIT BUGS.LOG ΒΛΕΠΟΥΜΕ ΟΛΑ ΤΑ ΑΠΟΤΕΛΕΣΜΑΤΑ
- ⌘ EDIT BUGS.IND ΠΕΡΙΕΧΟΜΕΝΑ BUGS.OUT
- ⌘ ΕΙΣΑΓΟΥΜΕ ΤΟ BUGS.OUT ΣΕ ΣΤΑΤΙΣΤΙΚΟ ΠΑΚΕΤΟ ΓΙΑ

ΔΙΑΓΡΑΜΜΑ 1: Trace of Posterior Values

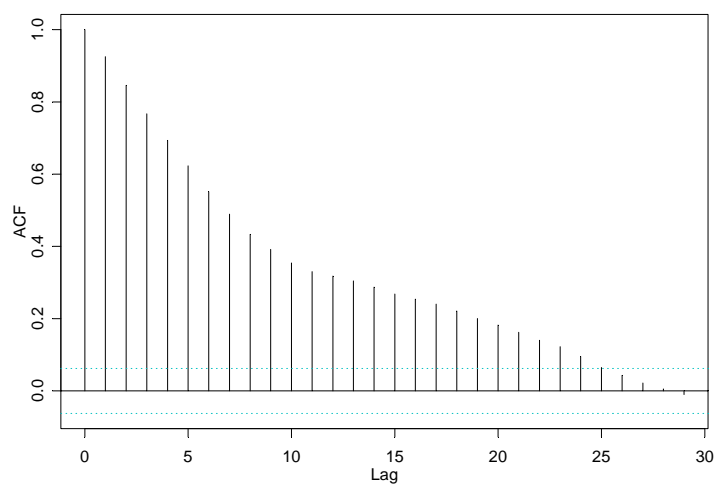


ΔΙΑΓΡΑΜΜΑ 2: ΕΡΓΟΔΙΚΟΙ ΜΕΣΟΙ



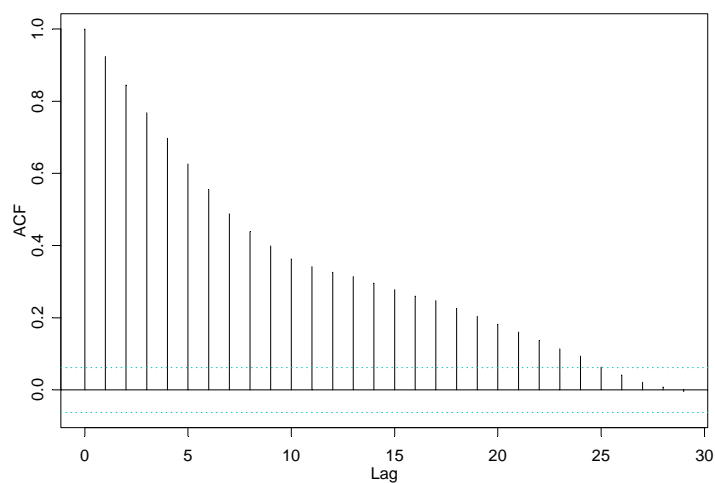
ΔΙΑΓΡΑΜΜΑ 3:
ΑΥΤΟΣΥΣΧΕΤΙΣΕΙΣ ΓΙΑ ΤΗΝ ΠΑΡΑΜΕΤΡΟ α

Series : a



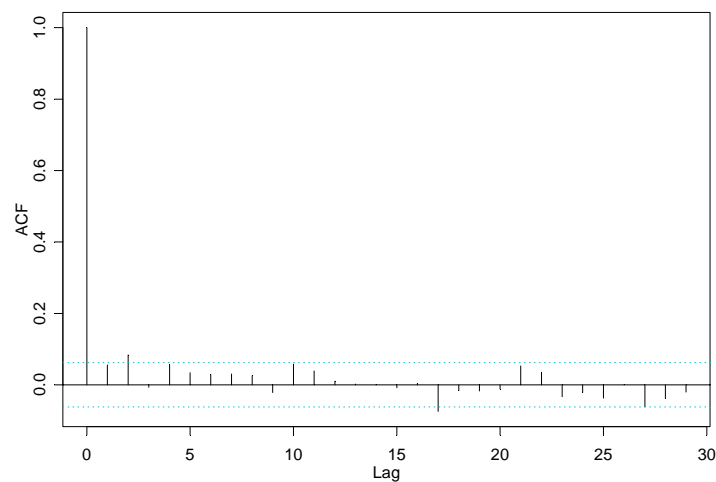
ΔΙΑΓΡΑΜΜΑ 4:
ΑΥΤΟΣΥΣΧΕΤΙΣΕΙΣ ΓΙΑ ΤΗΝ ΠΑΡΑΜΕΤΡΟ β

Series : b



ΔΙΑΓΡΑΜΜΑ 5: ΑΥΤΟΣΥΣΧΕΤΙΣΕΙΣ ΓΙΑ ΤΗΝ ΠΑΡΑΜΕΤΡΟ σ^2

Series : s2



3.2.4. Προσομοιώνοντας από την posterior

ΔΙΟΡΘΩΣΗ ΤΗΣ ΑΥΤΟΣΥΣΧΕΤΙΣΗΣ

(1) ΑΥΞΑΝΟΥΜΕ ΤΟ ΔΙΑΣΤΗΜΑ ΛΕΠΤΥΝΣΗΣ

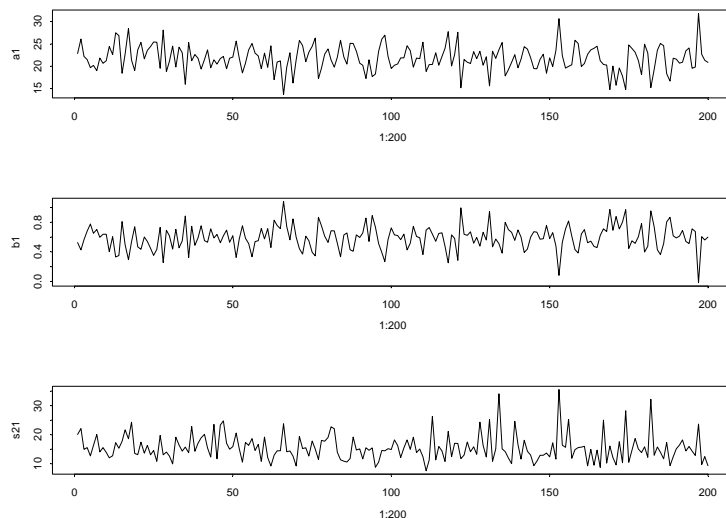
(Thin=25)

⌘ ΣΤΟ CMD ΑΡΧΕΙΟ ΤΟΥ BUGS

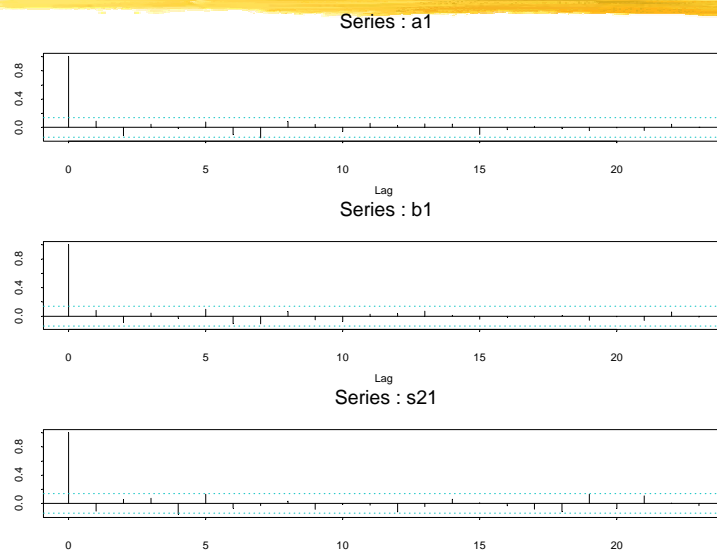
⌘ `monitor(a,25)`

⌘ `monitor(b,25)`

ΔΙΑΓΡΑΜΜΑ 6:
Trace of Posterior Values
(5000 Iterations, Thin=25)



ΔΙΑΓΡΑΜΜΑ 7:
ΑΥΤΟΣΥΣΧΕΤΙΣΕΙΣ
(5000 Iterations, Thin=25)



3.2.4. Προσομοιώνοντας από την posterior

ΔΙΟΡΘΩΣΗ ΤΗΣ ΑΥΤΟΣΥΣΧΕΤΙΣΗΣ

(2) ΑΛΛΑΖΟΥΜΕ ΤΗΝ ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

⌘ ΑΦΑΙΡΕΣΗ ΤΟΥ ΜΕΣΟΥ ΑΠΟ ΚΑΘΕ ΕΠΕΞΗΓΗΜΑΤΙΚΗ ΜΕΤΑΒΛΗΤΗ ΒΟΗΘΑΕΙ ΣΤΗ ΣΥΓΚΛΙΣΗ ΑΡΑ

$$\eta_i = \alpha^* + \beta(x_i - \bar{x})$$

$$\alpha = \alpha^* - \beta \bar{x}$$

3.2.4. Προσομοιώνοντας από την posterior

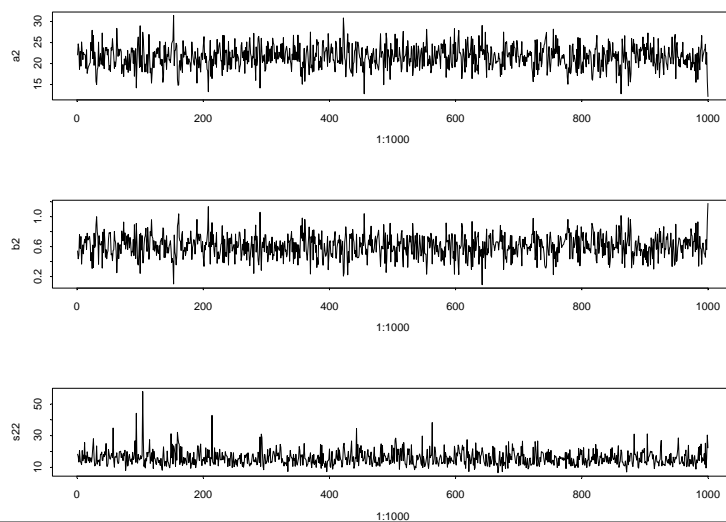
ΔΙΟΡΘΩΣΗ ΤΗΣ ΑΥΤΟΣΥΣΧΕΤΙΣΗΣ

(2) ΑΛΛΑΖΟΥΜΕ ΤΗΝ ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

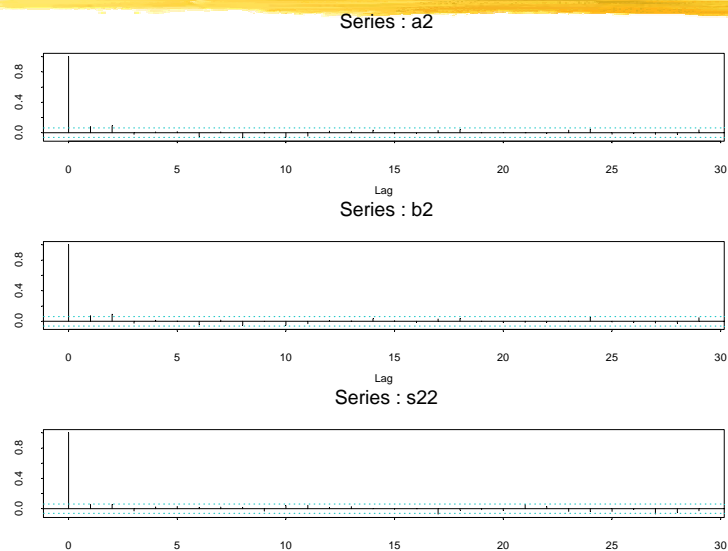
⌘ ΑΦΑΙΡΕΣΗ ΤΟΥ ΜΕΣΟΥ ΑΠΟ ΚΑΘΕ ΕΠΕΞΗΓΗΜΑΤΙΚΗ ΜΕΤΑΒΛΗΤΗ ΒΟΗΘΑΕΙ ΣΤΗ ΣΥΓΚΛΙΣΗ ΑΡΑ ΣΤΟ BUGS

```
mu[i] <-
  a.star+b*(estriol[i]-mean(estriol[]));
a<-a.star-b*mean(estriol[]);
a.star~dnorm(0.0, 1.0E-04);
```

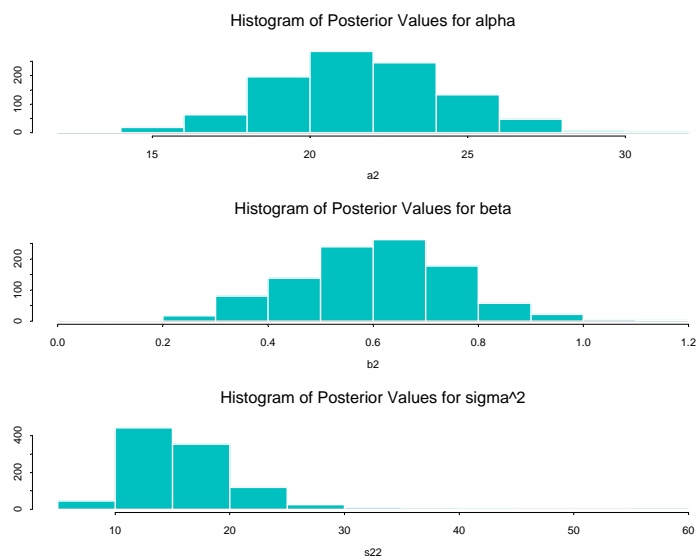
ΔΙΑΓΡΑΜΜΑ 8: Trace of Posterior Values



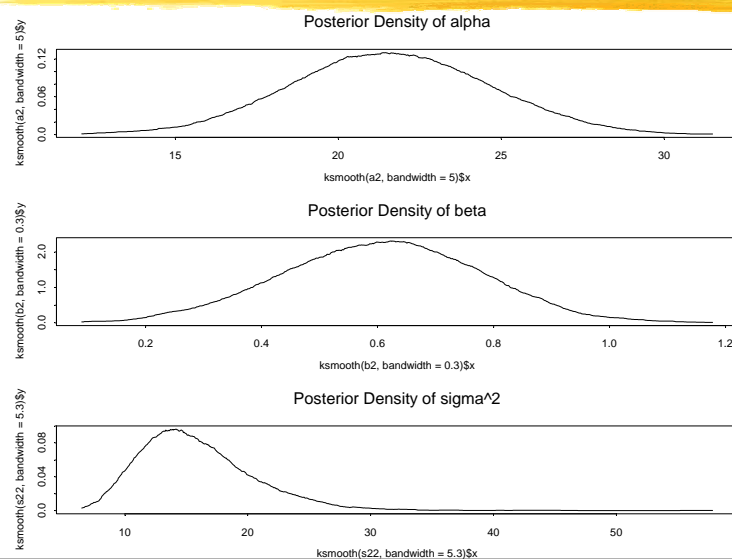
ΔΙΑΓΡΑΜΜΑ 9: ΑΥΤΟΣΥΣΧΕΤΙΣΕΙΣ



ΔΙΑΓΡΑΜΜΑ 10: ΙΣΤΟΓΡΑΜΜΑΤΑ POSTERIOR ΤΙΜΩΝ



ΔΙΑΓΡΑΜΜΑ 11: ΕΚΤΙΜΩΜΕΝΕΣ POSTERIOR ΚΑΤΑΝΟΜΕΣ



3.2.5. Προσομοίωση στο Παρασκήνιο

- ⌘ ΔΗΜΙΟΥΡΓΟΥΜΕ ΕΝΑ ΑΡΧΕΙΟ ΜΕ ΤΙΣ ΠΡΟΗΓΟΥΜΕΝΕΣ ΕΝΤΟΛΕΣ (ESTRIOL.CMD)
- ⌘ ΠΡΟΣΟΜΟΙΩΝΟΥΜΕ ΤΙΜΕΣ ΜΕ BACKBUGS ESTRIOL.CMD ΣΤΟΝ ΚΑΤΑΛΟΓΟ ΤΟΥ BUGS
- ⌘ ΒΛΕΠΟΥΜΕ ΤΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΣΤΟ ΑΡΧΕΙΟ BUGS.LOG

3.3. Παράδειγμα 2: Μοντέλα για Διωνυμικά Δεδομένα

- ⌘ ΠΑΡΑΔΕΙΓΜΑ 12, BUGS EXAMPLES vol 2 σελ. 43
 - ⌘ Bliss (1935)
- Εκθέτουμε 8 ομάδες εντόμων σε διαφορετικά επίπεδα Carbon disphide και καταγράφουμε
- ⌘ Συγκέντρωση (X_i)
 - ⌘ Συνολικός Αριθμός εντόμων στην ομάδα (n_i)
 - ⌘ Αριθμός εντόμων που απεβίωσαν (r_i)

3.3. Παράδειγμα 2: Μοντέλα για Διωνυμικά Δεδομένα

```

⌘ (1)  $r_i \sim \text{Binomial}(p_i, n_i)$ 
⌘ (2)  $\eta_i = \alpha + \beta(x_i - \bar{x})$ 
⌘ (3)  $\text{logit}(p_i) = \eta_i$ 
⌘ για  $i=1, \dots, 8$ 
⌘ PRIORS
  ⌘  $f(\alpha) = \text{Normal}(0, 10^4)$ 
  ⌘  $f(\beta) = \text{Normal}(0, 10^4)$ 
⌘  $\alpha = \alpha - \beta \bar{x}$ 
⌘ LINK FUNCTIONS
  ⌘  $\text{logit}(p) = \log\{p/(1-p)\}$ 
  ⌘  $\text{probit}(p) = \Phi^{-1}(p)$ 
  ⌘  $\text{cloglog}(p) = \log(-\log(1-p))$ 

```

for (i in 1:n) {
 r[i]~dbinom(p[i],n[i])
 logit(p[i])<-a.star+b*x[i]
}

a~dnorm(0.0,1.0E-04)
b~dnorm(0.0,1.0E-04)
a<-a.star-b*mean(x[])

3.3. Παράδειγμα 2: Μοντέλα για Διωνυμικά Δεδομένα

ΜΕΡΙΚΑ ΣΧΟΛΙΑ

⌘ ANAMENOMENES TIMES ΥΠΟΛΟΓΙΖΟΝΤΑΙ ΩΣ

⌘ $r.\text{hat}[i] <- n[i] * p[i];$

⌘ ODDS RATIO ΓΙΑ LOGIT MODELS

⌘ $\text{odds.ratio} <- \exp(b);$

3.4. Παράδειγμα 3: Μοντέλα για Δίτιμες Μεταβλητές

Ένα δείγμα ηλικιωμένων ατόμων εξετάστηκαν ψυχιατρικά αν έχουν κάποια συμπτώματα γηρατειών (senility symptoms).

Μια επεξηγηματική μεταβλητή είναι είναι το σκορ σε ένα υπο-τεστ της κλίμακας ενήλικης ευφυΐας Wechsler (Wechsler Adult Intelligence Scale - WAIS).

Να βρεθεί ποιο x αντιστοιχεί σε $p=1/2$ και να βρεθεί και η posterior κατανομή της πιθανότητας για κάποιον με WAIS ίσο με το μέσο όρο

3.4. Παράδειγμα 3: Μοντέλα για Δίτιμες Μεταβλητές

ΜΕΡΙΚΑ ΣΧΟΛΙΑ

⌘ Το μοντέλο

```
for (i in 1:n) {
  symptom[i]~dbern( p[i] );
  logit( p[i] ) <- a+b*wais[i]; }
```

⌘ Το x για $p=1/2$ βρίσκεται ως

```
x.half<- -a/b;
```

⌘ Το ποσοστό για κάποιον με μέσο σκορ WAIS

```
p.mean<-
  exp(a+b*mean(x[ ]))/(1+exp(a+b*mean(x[ ])));
```

3.5. Παράδειγμα 4: Μοντέλα Poisson για 2x2 Πίνακες Συνάφειας & Η Posterior Κατανομή του Odds Ratio

Mahon *et.al.* (1970) bulletin of the world health organization

Μελέτη για την πιθανή θετική σχέση μεταξύ ηλικίας στην 1η γέννα και καρκίνου του μαστού.

Οι περιπτώσεις (cases) προέρχονται από επιλεγμένα νοσοκομεία στις ΗΠΑ, Ελλάδα, Γιουγκοσλαβία, Βραζιλία, Ταϊβάν & Ιαπωνία

Οι Μάρτυρες (controls) επιλέχθηκαν από γυναίκες με συγκρίσιμη ηλικία από τα ίδια νοσοκομεία

3.5. Παράδειγμα 4: Μοντέλα Poisson για 2x2 Πίνακες Συνάφειας & Η Posterior Κατανομή του Odds Ratio

STATUS	AGE AT FIRST BIRTH	
	Age>29 (1)	Age<30 (0)
Case (1)	683	2537
Control (0)	1498	8747

3.5. Παράδειγμα 4: Μοντέλα Poisson για 2x2 Πίνακες Συνάφειας & Η Posterior Κατανομή του Odds Ratio

Status	Age	Counts
1	1	683
1	0	2537
0	1	1498
0	0	8747

3.5. Παράδειγμα 4: Μοντέλα Poisson για 2x2 Πίνακες Συνάφειας & Η Posterior Κατανομή του Odds Ratio

TO MONTELO

```
for (i in 1:4) {
  counts[i]~dpois(lambda[i]);
  log (lambda[i])<-mu+
    a*status[i]+b*age[i]+ab*status[i]*age[i];
}
```

ODDS RATIO

```
odds.ratio<-exp(ab);
```

3.6. Παράδειγμα 5: Εκτίμηση κοινού Odds Ratio σε 2x2xJ Πίνακες Συνάφειας

Σε πίνακες 2x2xJ η εκτίμηση ενός κοινού OR γίνεται από το
Maentel-Haenzel $OR_{MH} = (\sum a_i d_i / n_i) / (\sum b_i c_i / n_i)$

Sandler, Everson & Wilcox (1985) *Amer. Journal of Epidemiology*

Μελέτη με 518 καρκινοπαθείς με ηλικίες 15-59 και 518 μάρτυρες
(ομάδα ελέγχου) ταιριασμένοι (matched) ως προς φύλο και
ηλικία

Σκοπός: εκτίμηση της επίδρασης του παθητικού καπνίσματος στον
κίνδυνο εμφάνισης καρκίνου. Το παθητικό κάπνισμα ορίστηκε
θετικά αν η σύζυγος κάπνιζε τουλάχιστον 1 τσιγάρο ημερησίως
τους τελευταίους 6 μήνες.

Συγχυτικός παράγοντας (confounder) αν το ίδιο άτομο καπνίζει

3.6. Παράδειγμα 5: Εκτίμηση κοινού Odds Ratio σε 2x2xJ Πίνακες Συνάφειας

	Non Smokers (0)		Smokers (1)	
	Passive Smoker (1)	Non Passive Smoker(0)	Passive Smoker (1)	Non Passive Smoker (0)
Case(1)	120	111	161	117
Control (0)	80	155	130	124

3.6. Παράδειγμα 5: Εκτίμηση κοινού Odds Ratio σε 2x2xJ Πίνακες Συνάφειας

⌘ ΚΑΝΟΥΜΕ ΔΥΟ ΑΝΑΛΥΣΕΙΣ

⌘ ΑΝΑΛΥΣΗ 1:

- ⌘ ΧΕΙΡΙΖΟΜΑΣΤΕ ΞΕΧΩΡΙΣΤΑ ΤΟΝ ΚΑΘΕ ΠΙΝΑΚΑ ΚΑΙ ΕΚΤΙΜΟΥΜΕ ΕΝΑ ODDS RATIO ΓΙΑ ΚΑΘΕ ΠΙΝΑΚΑ (ΟΠΩΣ ΠΑΡΑΔΕΙΓΜΑ 4)
- ⌘ ΣΥΓΚΡΙΝΟΥΜΕ ΤΙΣ POSTERIOR ΚΑΤΑΝΟΜΕΣ ΤΟΥΣ

⌘ ΑΝΑΛΥΣΗ 2:

- ⌘ ΧΕΙΡΙΖΟΜΑΣΤΕ ΞΕΧΩΡΙΣΤΑ ΤΟΝ ΚΑΘΕ ΠΙΝΑΚΑ
- ⌘ ΕΚΤΙΜΟΥΜΕ ΕΝΑ ΚΟΙΝΟ ODDS RATIO ΓΙΑ ΚΑΘΕ ΠΙΝΑΚΑ

3.6. Παράδειγμα 5: Εκτίμηση κοινού Odds Ratio σε 2x2xJ Πίνακες Συνάφειας

⌘ ΑΝΑΛΥΣΗ 1:

⌘ ΠΡΟΚΑΤΑΡΤΙΚΟ ΚΟΜΜΑΤΙ

⌘ `var ..., b[2,4], or[2];`

⌘ ΜΟΝΤΕΛΟ

```
⌘ #model for 1st table (nonsmokers)
⌘ for (i in 1:4) {
⌘   counts[i]~dpois( lambda[i] );
⌘   log(lambda[i])<-b[1,1]+b[1,2]*status[i]
⌘                                     +b[1,3]*passive[i]
⌘                                     +b[1,4]*status[i]*passive[i];
⌘ }
```


3.6. Παράδειγμα 5: Εκτίμηση κοινού Odds Ratio σε 2x2xJ Πίνακες Συνάφειας

⌘ ΑΝΑΛΥΣΗ 1:

⌘ ΠΡΟΚΑΤΑΡΤΙΚΟ ΚΟΜΜΑΤΙ

```
⌘ var ..., b[2,4], or[2];
```

⌘ ΜΟΝΤΕΛΟ

```
⌘ #model for 2nd table (smokers)
```

```
⌘ for (i in 5:8) {
```

```
⌘ counts[i]~dpois( lambda[i] );
```

```
⌘ log(lambda[i])<-b[2,1]+b[2,2]*status[i]
```

```
⌘                               +b[2,3]*passive[i]
```

```
⌘                               +b[2,4]*status[i]*passive[i];
```

```
⌘ }
```

3.6. Παράδειγμα 5: Εκτίμηση κοινού Odds Ratio σε 2x2xJ Πίνακες Συνάφειας

⌘ ΑΝΑΛΥΣΗ 1:

⌘ ΠΡΟΚΑΤΑΡΤΙΚΟ ΚΟΜΜΑΤΙ

```
⌘ var ..., b[2,4], or[2];
```

⌘ ΜΟΝΤΕΛΟ

```
⌘ #priors
```

```
⌘ for (i in 1:2){
```

```
⌘   for (j in 1:p){
```

```
⌘     b[i,j]~dnorm(0.0, 1.0E-04)
```

```
⌘   }
```

```
⌘ }
```

3.6. Παράδειγμα 5: Εκτίμηση κοινού Odds Ratio σε 2x2xJ Πίνακες Συνάφειας

⌘ ΑΝΑΛΥΣΗ 2:

☒ ΜΟΝΤΕΛΟ

```

☒ #model for 1st table (nonsmokers)
☒ for (i in 1:4) {
☒   counts[i]~dpois( lambda[i] );
☒   log(lambda[i])<-b[1,1]+b[1,2]*status[i]
☒                                     +b[1,3]*passive[i]
☒                                     +b[1,4]*status[i]*passive[i];}
☒ #model for 2nd table (smokers)
☒ for (i in 5:8) {
☒   counts[i]~dpois( lambda[i] );
☒   log(lambda[i])<-b[2,1]+b[2,2]*status[i]
☒                                     +b[2,3]*passive[i]
☒                                     +b[2,4]*status[i]*passive[i]; }

```

3.6. Παράδειγμα 5: Εκτίμηση κοινού Odds Ratio σε 2x2xJ Πίνακες Συνάφειας

⌘ ΑΝΑΛΥΣΗ 2:

☒ ΜΟΝΤΕΛΟ

```

☒ #model for 1st table (nonsmokers)
☒ for (i in 1:4) {
☒   counts[i]~dpois( lambda[i] );
☒   log(lambda[i])<-b[1,1]+b[1,2]*status[i]
☒                                     +b[1,3]*passive[i]
☒                                     + ab *status[i]*passive[i];}
☒ #model for 2nd table (smokers)
☒ for (i in 5:8) {
☒   counts[i]~dpois( lambda[i] );
☒   log(lambda[i])<-b[2,1]+b[2,2]*status[i]
☒                                     +b[2,3]*passive[i]
☒                                     + ab *status[i]*passive[i]; }

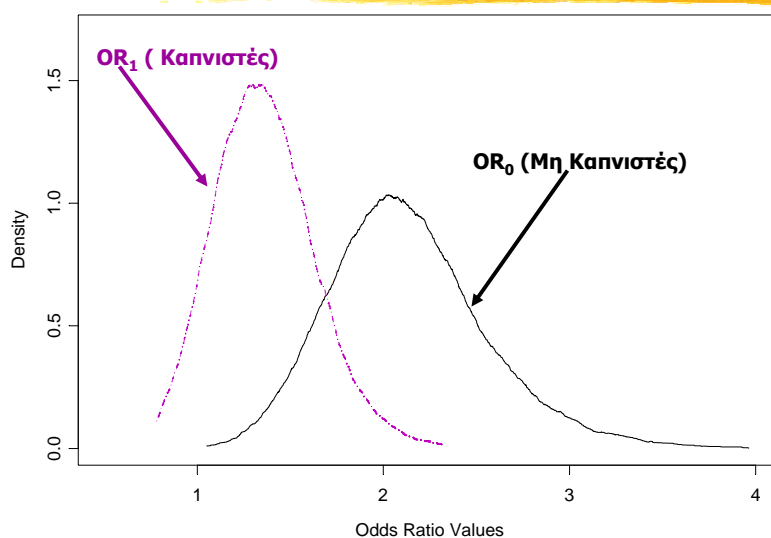
```

3.6. Παράδειγμα 5: Εκτίμηση κοινού Odds Ratio σε 2x2xJ Πίνακες Συνάφειας

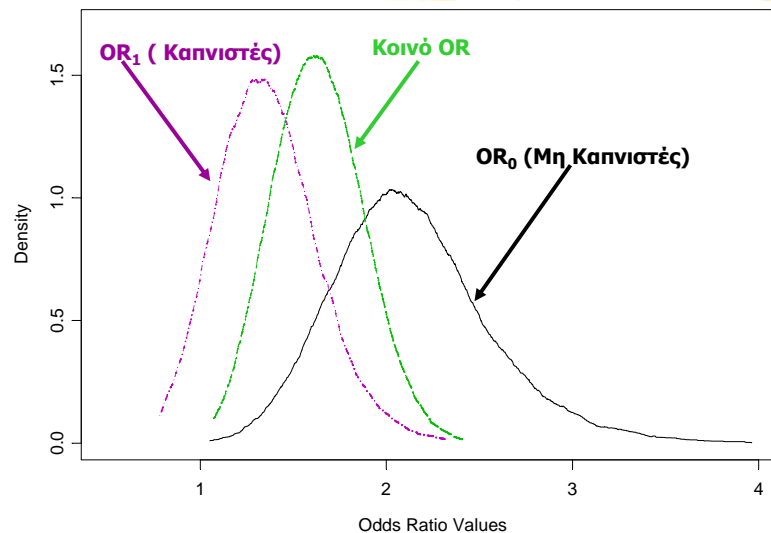
ΑΠΟΤΕΛΕΣΜΑΤΑ

	<u>MLE</u>	<u>Posterior Mean</u>	<u>95% posterior credible interval</u>
<u>ΑΝΑΛΥΣΗ 1</u>			
OR₀	2.09	2.07±0.036	1.47 - 3.09
OR₁	1.31	1.33±0.022	0.97 - 1.88
<u>ΑΝΑΛΥΣΗ 2</u>			
Common			
OR_{MH}	1.63	1.61 ± 0.0087	1.27 - 2.06

ΔΙΑΓΡΑΜΜΑ 12: ΕΚΤΙΜΩΜΕΝΕΣ POSTERIOR ΚΑΤΑΝΟΜΕΣ ΤΩΝ ODDS RATIOS



**ΔΙΑΓΡΑΜΜΑ 13:
ΕΚΤΙΜΩΜΕΝΕΣ POSTERIOR ΚΑΤΑΝΟΜΕΣ
ΤΩΝ ODDS RATIOS**



**Bayesian Biostatistics
Using BUGS**

**ΤΕΛΟΣ ΕΙΣΑΓΩΓΙΚΟΥ
ΜΑΘΗΜΑΤΟΣ**

ΕΠΟΜΕΝΟ: WINBUGS

E-mail: ntzoufras@aueb.gr



Department of Statistics,
Athens University of
Economics & Business

ΠΑΡΑΡΤΗΜΑ Α (1^ο ΜΑΘΗΜΑΤΟΣ): ΠΑΡΑΔΕΙΓΜΑΤΑ BUGS

1 EXAMPLE 1: BIRTHWEIGHT & ESTRIOIOL LEVEL

1.1 Model (Bug file)

```
model example1;
const n=31; # n=sample size
var   estriol[n], # estriol level of pregnant woman
      birth[n],   # birthweight
      mu[n],       # regression expected value
      a.star,a,b,tau,s2; # model parameters,
                        # tau = precision, s2=1/tau error variance
data estriol,birth in 'estriol.dat';
inits in 'estriol2.ini';
{
#       definition of likelihood function
#
  for (i in 1:n) {
    birth[i]~dnorm( mu[i], tau ); # random component
#       systematic component & link function
    mu[i]<-a.star+b*(estriol[i]-mean(estriol[]));
  }
#       prior distributions
#
  a.star~dnorm( 0, 1.0E-04 ); # normal prior for a
  b~dnorm( 0, 1.0E-04 ); # normal prior for b
  tau~dgamma( 1.0E-04 , 1.0E-04 ); # gamma prior for precision
  s2<-1/tau;
  a<-a.star-b*mean(estriol[]);
}
```

1.2 Initial Values (estriol2.ini)

```
list(a.star=0.0, b=0.0, tau=1.0)
```

1.3 Data (estriol.dat)

```
7      25
9      25
9      25
12     27
14     27
16     27
16     24
14     30
16     30
16     31
17     30
19     31
21     30
24     28
15     32
16     32
17     32
25     32
27     34
15     34
15     34
15     35
16     35
19     34
18     35
17     36
18     37
20     38
22     40
25     39
```

2 EXAMPLE 2: BEETLES DATASET

2.1 Model (Bug file)

```

model beetles;
const
  N = 8;      # number of doses
var
  r[N],p[N],x[N],n[N],alpha,alpha.star,beta,r.hat[N], odds.ratio;
data x, n, r in "beetles.dat";
inits in "beetles.in";
{
  for (i in 1:N) {
    r[i] ~ dbin(p[i], n[i]);
    logit(p[i]) <- alpha.star + beta*(x[i]-mean(x[]));
    r.hat[i] <- p[i]*n[i]; # fitted values
  }
  alpha.star ~ dnorm(0.0, 1.0E-3);
  beta ~ dnorm(0.0, 1.0E-3);
  alpha <- alpha.star - beta*mean(x[]);
  odds.ratio <- exp( beta )
}

```

2.2 Initial Values (beetles.ini)

```
list(alpha.star=0, beta=0)
```

2.3 Data (beetles.dat)

```

1.6907  59   6
1.7242  60  13
1.7552  62  18
1.7842  56  28
1.8113  63  52
1.8369  59  53
1.8610  62  61
1.8839  60  60

```

3 EXAMPLE 3: WAIS & SENILITY SYMPTOMS

3.1 Model (Bug file)

```
model example3;
const n=54;          # number of observations
var   wais[n],       # wais measurement
      symptom[n],    # symptom binary indicator
      p[n],          # probability of symptom appearance
      alpha,         # alpha parameter
      beta,          # beta parameter
      odds.ratio,    # odds ratio
      x.fifty,       # value of x for p=1/2
      p.meanx;       # fitted probability for x=mean(x)
data wais, symptom in 'wais.dat';
inits in 'wais.ini';
{
  #      Likelihood Definition
  for (i in 1:n) {
    #          random component
    symptom[i]~dbern( p[i] );
    #          systemantic component & link function
    logit( p[i] ) <- alpha + beta * wais[i];
  }
  #      calculation of odds ratio
  odds.ratio<-exp(beta);
  #      priors for alpha and beta
  alpha~dnorm(0.0, 0.00001);
  beta~dnorm(0.0, 0.00001);
  #      x.fifty
  x.fifty<- -alpha/beta
  #      symptom probability for mean of observed wais
  p.meanx<-exp( alpha+beta*mean(wais[]) ) /
  (1+exp(alpha+beta*mean(wais[]) ) )
}
```

3.2 Initial Values (wais.ini)

```
list( alpha=0.0, beta=0.0 )
```

3.3 Data (wais.dat)

```
9      1
13     1
6      1
8      1
10     1
4      1
14     1
8      1
11     1
7      1
9      1
7      1
5      1
14     1
13     0
16     0
10     0
12     0
11     0
14     0
15     0
18     0
7      0
16     0
```

9	0
9	0
11	0
13	0
15	0
13	0
10	0
11	0
6	0
17	0
14	0
19	0
9	0
11	0
14	0
10	0
16	0
10	0
16	0
14	0
13	0
13	0
9	0
15	0
10	0
11	0
12	0
4	0
14	0
20	0

4 EXAMPLE 4: BREAST CANCER & AGE AT 1st BIRTH

4.1 Model (Bug file)

```
model example4;
const n=4;          # number of cells
var  status[n],     # study group (1=case, 0=control)
     age[n],        # age at 1st birth (0=age<30, 1=age>29)
     counts[n],     # cell counts
     lambda[n],     # expected number of cells
     mu,            # constant parameter
     a,             # status effect
     b,             # age effect
     ab,            # interaction between status and age
     odds.ratio;    # odds ratio
data status, age, counts in 'breast.dat';
inits in 'breast.ini';
{
  #      Likelihood Definition
    for (i in 1:n) {
  #          random component
        counts[i]~dpois( lambda[i] );
  #      systematic component & link function
        log( lambda[i] ) <- mu + a*status[i] +b*age[i]+
ab*status[i]*age[i];
    }
  #      calculation of odds ratio
    odds.ratio<-exp(ab);
  #      priors for model parameters
    mu~dnorm(0.0, 0.00001);
    a~dnorm(0.0, 0.00001);
    b~dnorm(0.0, 0.00001);
    ab~dnorm(0.0, 0.00001);
}
```

4.2 Initial Values (breast.ini)

```
list(mu=0,a=0,b=0,ab=0)
```

4.3 Data (breast.dat)

```
1 1 683
1 0 2537
0 1 1498
0 0 8747
```

5 EXAMPLE 5: 2x2x2 COMMON ODDS RATIO ESTIMATION: PASSIVE SMOKING & CANCER ADJUSTING FOR SMOKING STATUS

5.1 ANALYSIS 1: DIFFERENT RISK PER SMOKING STATUS

5.1.1 Model (Bug File)

```
model ex5notcommonOR;
const n=8,
      p=4;
var
    smoking[n],
    status[n],
    passive[n],
    counts[n],
    b[2,p],
    lambda[n],
    or[2];
data smoking, status, passive, counts in 'smoke.dat';
inits in 'smoke.ini';
{
#
#      model for 1st table (nonsmokers)
#      for (i in 1:4) {
#          counts[i]~dpois( lambda[i] );
#          log(lambda[i])<- b[1,1] + b[1,2]*status[i] +
b[1,3]*passive[i] + b[1,4]*status[i]*passive[i];
#      }
#
#      model for 2nd table (smokers)
#      for (i in 5:8) {
#          counts[i]~dpois( lambda[i] );
#          log(lambda[i])<- b[2,1] + b[2,2]*status[i] +
b[2,3]*passive[i] + b[2,4]*status[i]*passive[i];
#      }
#      priors
#      for (i in 1:2){
#          for (j in 1:p){
#              b[i,j]~dnorm(0.0, 1.0E-04)
#          }
#      }
#      odds ratios
#      or[1]<-exp(b[1,4])
#      or[2]<-exp(b[2,4])
}
```

5.1.2 Initial Values (smoke.ini)

```
list(b=c(0,0,0,0,0,0,0,0,0))
```

5.1.3 Data (smoke.dat)

```
0 1 1 120
0 1 0 111
0 0 1 80
0 0 0 155
1 1 1 161
1 1 0 117
1 0 1 130
1 0 0 124
```

5.2 ANALYSIS 2: COMMON RISK PER SMOKING STATUS

5.2.1 Model (Bug File)

```
model ex5commonOR;
const n=8,
      p=4;
var
  smoking[n],
  status[n],
  passive[n],
  counts[n],
  b[2,p-1],
  ab,
  lambda[n],
  or;
data smoking, status, passive, counts in 'smoke.dat';
inits in 'smoke2.ini';
{
#
#      model for 1st table (nonsmokers)
#      for (i in 1:4) {
#          counts[i]~dpois( lambda[i] );
#          log(lambda[i])<- b[1,1] + b[1,2]*status[i] +
b[1,3]*passive[i] + ab*status[i]*passive[i];
#      }
#
#      model for 2nd table (smokers)
#      for (i in 5:8) {
#          counts[i]~dpois( lambda[i] );
#          log(lambda[i])<- b[2,1] + b[2,2]*status[i] +
b[2,3]*passive[i] + ab*status[i]*passive[i];
#      }
#      priors
#      for (i in 1:2){
#          for (j in 1:p-1){
#              b[i,j]~dnorm(0.0, 1.0E-04);
#          }
#      }
#      ab~dnorm(0.0, 1.0E-04);
#      odds ratios
#      or<-exp(ab)
#
}
```

5.2.2 Initial Values (smoke2.ini)

```
list(b=c(0,0,0,0,0,0),ab=0)
```

Bayesian Biostatistics Using BUGS



Βιο-Στατιστική κατά Bayes με τη χρήση του Λογισμικού BUGS

ΜΑΘΗΜΑ 2: ΕΙΣΑΓΩΓΗ ΣΤΟ WINBUGS

I. Ntzoufras

E-mail: ntzoufras@aueb.gr



Department of Statistics,
Athens University of
Economics & Business

4... WINBUGS version 1.4

- ⌘ **4.1. Εισαγωγή: Διαφορές από το BUGS.**
- ⌘ **4.2. Ένα Απλό Παράδειγμα.**
- ⌘ **4.3. Καθορισμός Νέας Πιθανοφάνειας και Prior [μόνο στο WINBUGS].**
- ⌘ **4.4. Προσομοίωση στο Παρασκήνιο [μόνο στο 1.4].**

4... WINBUGS

4.1. ΕΙΣΑΓΩΓΗ: ΔΙΑΦΟΡΕΣ ΜΕ ΤΟ ΚΛΑΣΣΙΚΟ BUGS

⌘ Γενικά η λογική είναι ίδια με το κλασσικό BUGS

- ☒ Επιπλέον νέες μεθοδολογίες και αλγόριθμοι (Metropolis-Hastings και Slice Gibbs)
- ☒ Γράφουμε κατευθείαν το κώδικα για το «κυρίως μοντέλο»
- ☒ Δεν χρειαζόμαστε το προπαρασκευαστικό κώδικα (preamble).

4... WINBUGS

4.1. ΕΙΣΑΓΩΓΗ: ΔΙΑΦΟΡΕΣ ΜΕ ΤΟ ΚΛΑΣΣΙΚΟ BUGS (2)

- ⌘ Γραφική αναπαράσταση του μοντέλου μέσω του DoodleBUGS
- ⌘ Φιλικά μενού για το χειρισμό της προσομοίωσης
- ⌘ Μπορούμε να έχουμε κατευθείαν περιγραφική ανάλυση και διαγράμματα της posterior κατανομής
- ⌘ Υπάρχει δυνατότητα «Αποκοπής και επικόλλησης» (Cut/paste) σε άλλα προγράμματα

4... WINBUGS

4.1. ΕΙΣΑΓΩΓΗ: ΔΙΑΦΟΡΕΣ ΜΕ ΤΟ ΚΛΑΣΣΙΚΟ BUGS (3)

- ⌘ Μπορούμε να ορίσουμε Πιθανοφάνεια και Prior που δε συμπεριλαμβάνονται στην προκαθορισμένη λίστα του WINBUGS
- ⌘ Στην έκδοση 1.4 μπορούμε να προσομοιώσουμε στο παρασκήνιο κάτι που γινόταν στο BUGS αλλά όχι στον WINBUGS 1.3

4... WINBUGS

4.2. Ένα Απλό Παράδειγμα

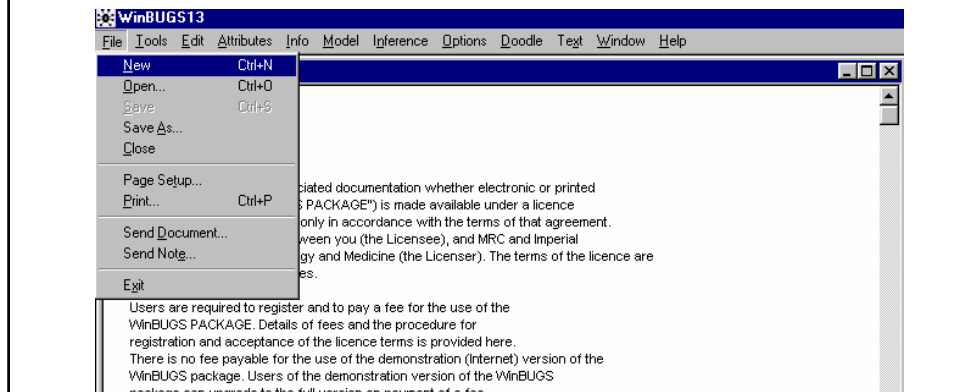
- ⌘ Green & Touchston (1963, *Am.Jour. Of Obstetrics & Gynecology*)
- ⌘ Μελετάμε τη σχέση μεταξύ
 - ☒ Y : Βάρος γέννησης (Birthweight)
 - ☒ X : Επίπεδο Εστριόλης της μητέρας (Estriol level)
 - ☒ Μέγεθος Δείγματος $n=31$

4... WINBUGS

4.2.1. Ορισμός του Μοντέλου στο WINBUGS

⌘ Ξεκινάμε το WinBUGS

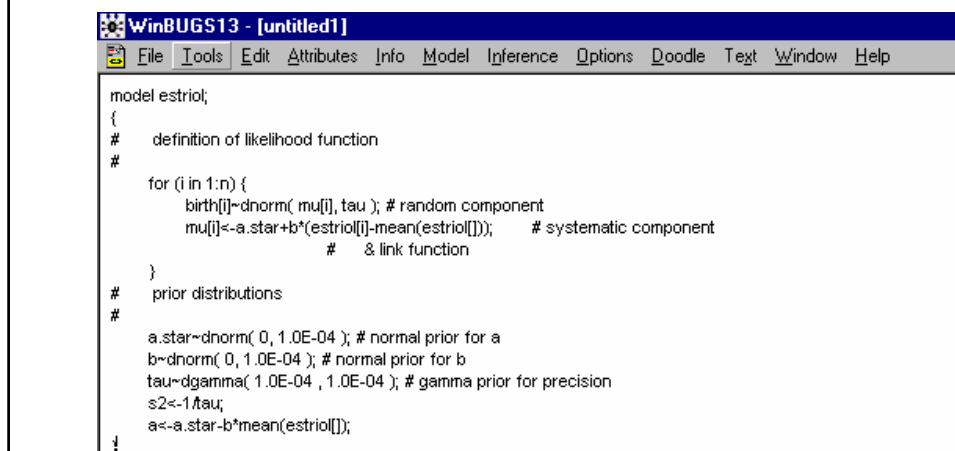
⌘ Επιλέγουμε "New" στο μενού *file*



4... WINBUGS

4.2.1. Ορισμός του Μοντέλου στο WINBUGS

⌘ Γράφουμε τον κώδικα του μοντέλου



4... WINBUGS

4.2.1. Ορισμός του Μοντέλου στο WINBUGS

⌘ Γράφουμε τον κώδικα του μοντέλου

⌘ Ακολουθούμε από τις αρχικές τιμές

```
list(a.star=0.0, b=0.0, tau=1.0)
```

⌘ και τα δεδομένα

```
list(n=31)
```

```
estriol[] birth[]
```

```
7 25
```

```
9 25
```

```
... ..
```

```
24 43
```

```
END
```

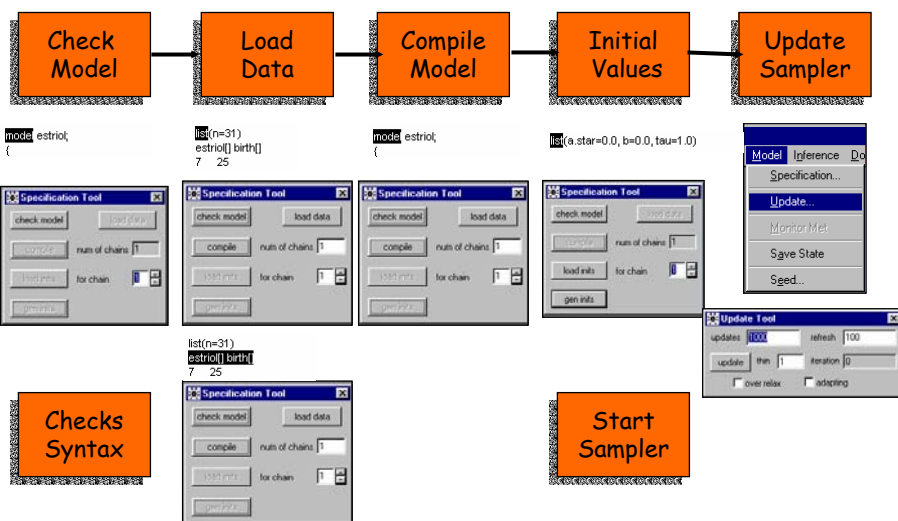
↵ [ENTER]

↵ [ENTER]

ΠΡΟΣΟΧΗ: ΑΥΤΟ ΔΕΝ ΧΡΕΙΑΖΕΤΑΙ
ΣΤΟ WINBUGS 1.3

4... WINBUGS

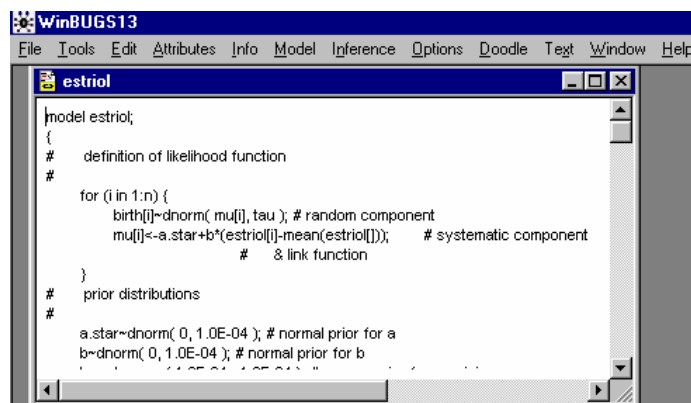
4.2.2. Προσομοίωση στο WINBUGS



4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘ 1... Έλεγχος του Μοντέλου (Check Model)



```

WinBUGS13
File Tools Edit Attributes Info Model Inference Options Doodle Text Window Help
estriol
{
  # definition of likelihood function
  #
  for (i in 1:n) {
    birth[i]~dnorm( mu[i], tau ); # random component
    mu[i]~a.star+b*(estriol[i]-mean(estriol[])); # systematic component
    # & link function
  }
  # prior distributions
  #
  a.star~dnorm( 0, 1.0E-04 ); # normal prior for a
  b~dnorm( 0, 1.0E-04 ); # normal prior for b
}

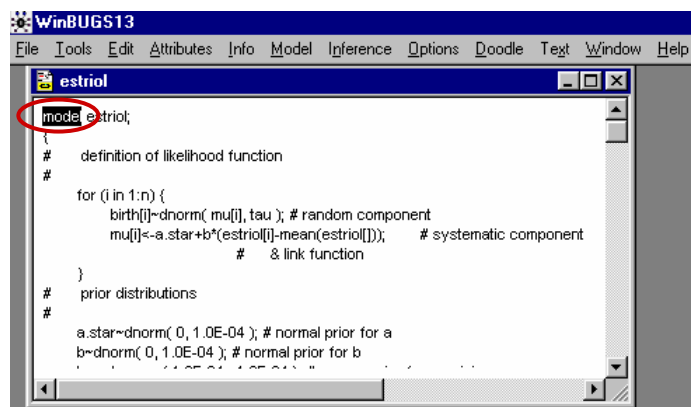
```

4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘ 1... Έλεγχος του Μοντέλου (Check Model)

MAYRIZOYME (με *double click*) ΤΗΝ ΕΝΤΟΛΗ **model**



```

WinBUGS13
File Tools Edit Attributes Info Model Inference Options Doodle Text Window Help
estriol
{
  # definition of likelihood function
  #
  for (i in 1:n) {
    birth[i]~dnorm( mu[i], tau ); # random component
    mu[i]~a.star+b*(estriol[i]-mean(estriol[])); # systematic component
    # & link function
  }
  # prior distributions
  #
  a.star~dnorm( 0, 1.0E-04 ); # normal prior for a
  b~dnorm( 0, 1.0E-04 ); # normal prior for b
}

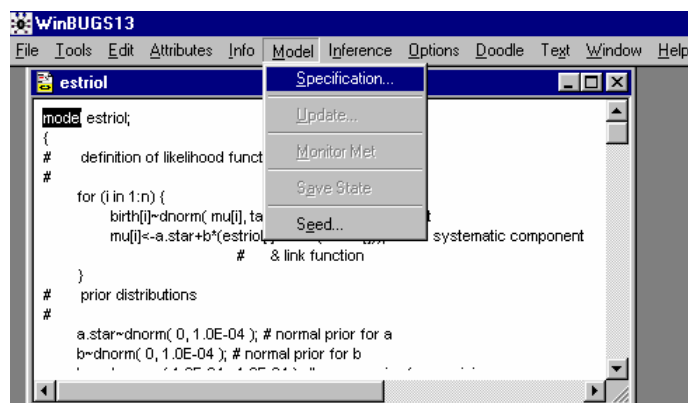
```

4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘ 1... Έλεγχος του Μοντέλου (Check Model)

ΕΠΙΛΕΓΟΥΜΕ "SPECIFICATION" ΑΠΟ ΤΟ ΜΕΝΟΥ "MODEL"

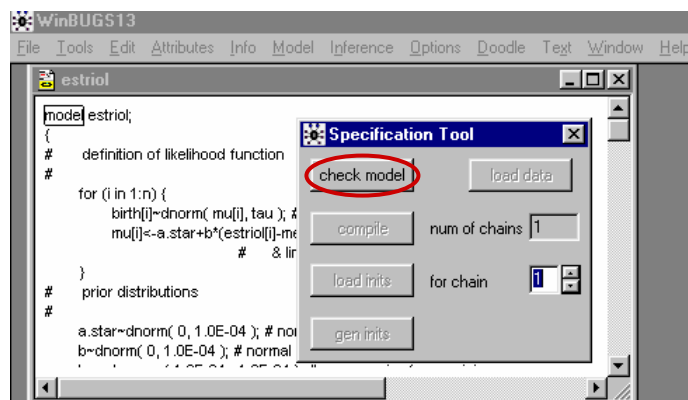


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘ 1... Έλεγχος του Μοντέλου (Check Model)

ΕΠΙΛΕΓΟΥΜΕ ΤΟ ΚΟΥΤΙ "CHECK MODEL"

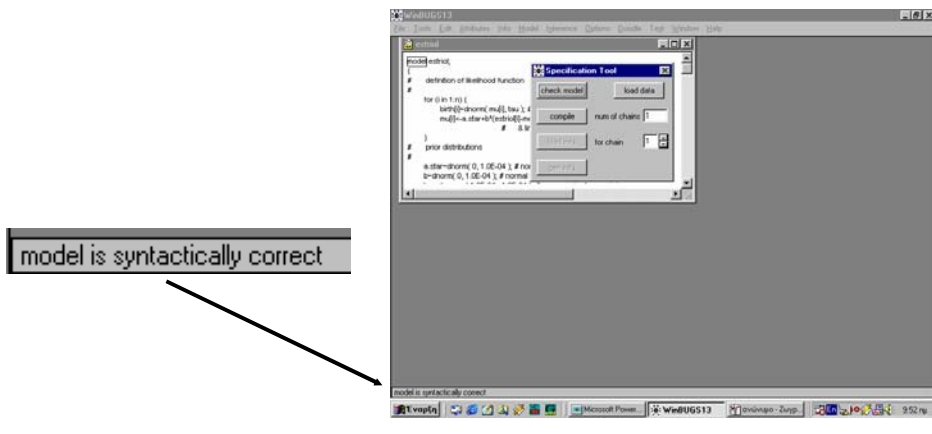


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘ 1... Έλεγχος του Μοντέλου (Check Model)

ΕΑΝ Η ΣΥΝΤΑΞΗ ΕΙΝΑΙ ΣΩΣΤΗ ΤΟΤΕ

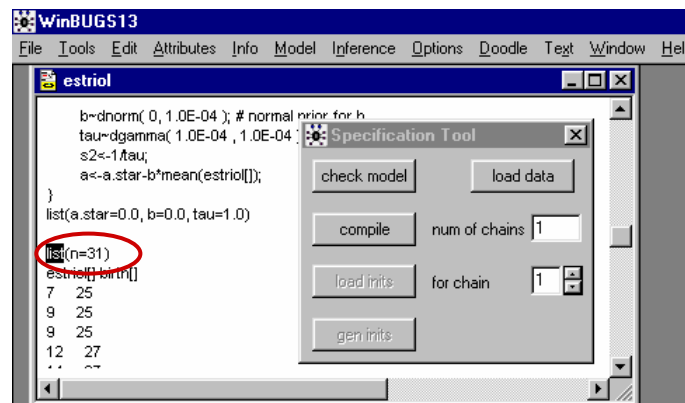


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘ 2... Φόρτωση των Δεδομένων (Load Data)

ΜΑΥΡΙΖΟΥΜΕ ΤΟ **list** Ή ΤΗΝ 1η ΓΡΑΜΜΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ

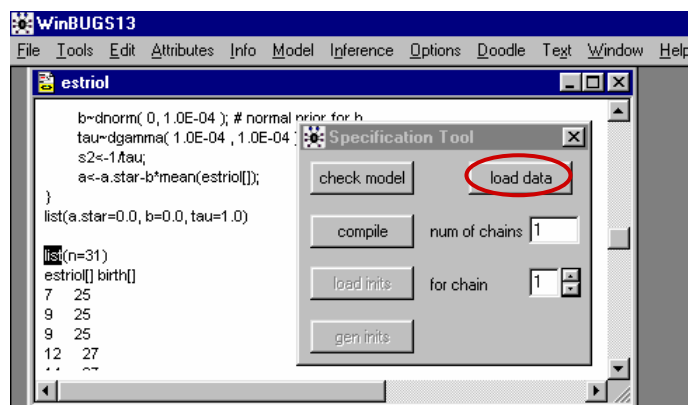


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘ 2... Φόρτωση των Δεδομένων (Load Data)

ΕΠΙΛΕΓΟΥΜΕ ΤΟ ΚΟΥΤΙ "LOAD DATA"

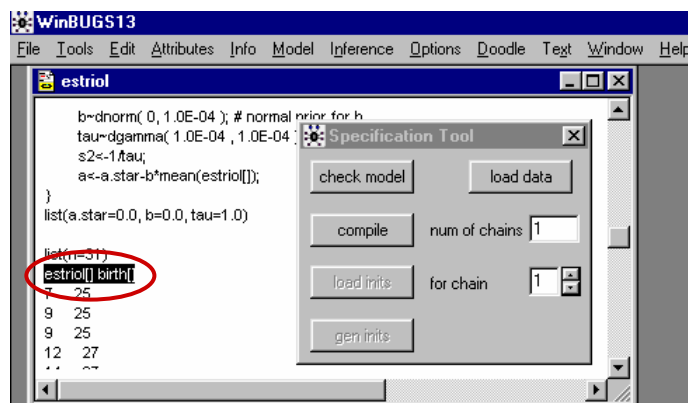


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘ 2... Φόρτωση των Δεδομένων (Load Data)

ΜΑΥΡΙΖΟΥΜΕ ΕΠΙΠΛΕΟΝ ΔΕΔΟΜΕΝΑ ΓΙΑ ΦΟΡΤΩΜΑ

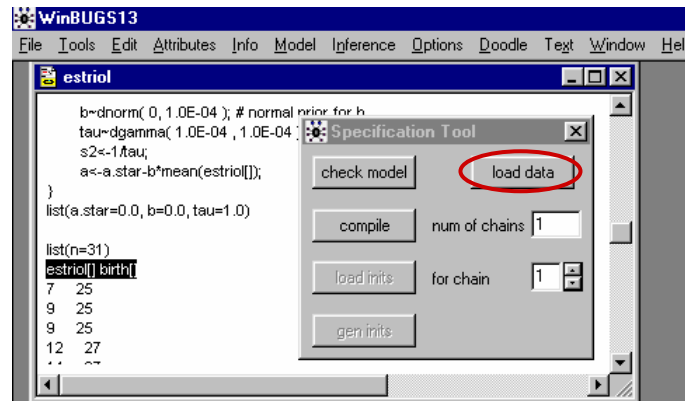


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘ 2... Φόρτωση των Δεδομένων (Load Data)

ΞΑΝΑΕΠΙΛΕΓΟΥΜΕ ΤΟ ΚΟΥΤΙ "LOAD DATA"



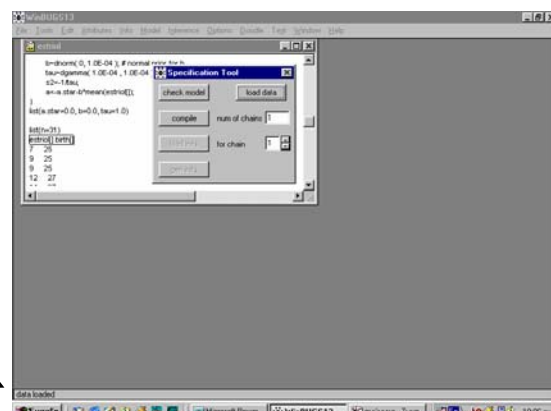
4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘ 2... Φόρτωση των Δεδομένων (Load Data)

ΕΑΝ ΤΑ ΔΕΔΟΜΕΝΑ ΦΟΡΤΩΘΟΥΝ ΣΩΣΤΑ ΤΟΤΕ

data loaded

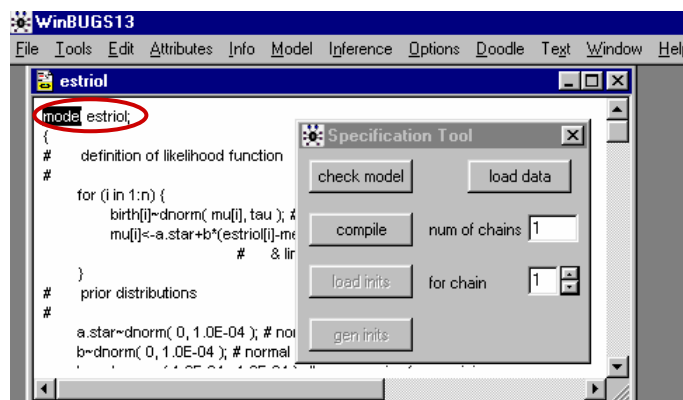


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘3... Εκκίνηση του Μοντέλου (Compile Model)

ΜΑΥΡΙΖΟΥΜΕ ΤΗΝ ΕΝΤΟΛΗ **model** (ΞΑΝΑ)

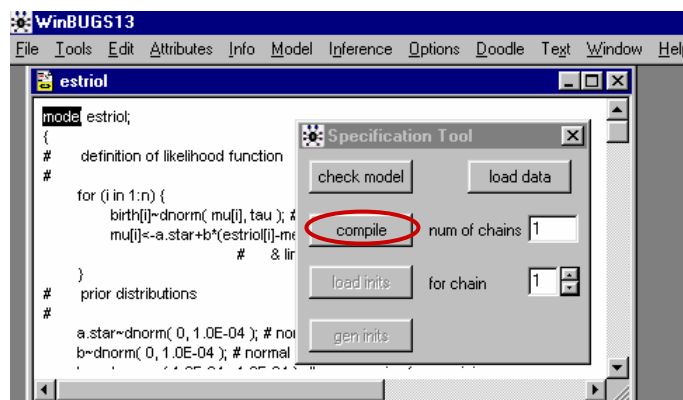


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘3... Εκκίνηση του Μοντέλου (Compile Model)

ΕΠΙΛΕΓΟΥΜΕ ΤΟ ΚΟΥΤΙ "**COMPILE**"



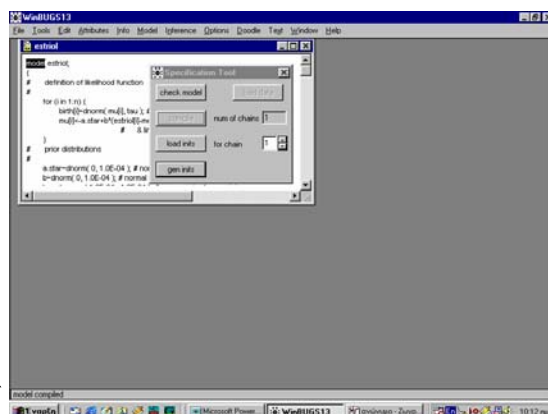
4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘3... Εκκίνηση του Μοντέλου (Compile Model)

ΕΑΝ ΤΟ ΜΟΝΤΕΛΟ ΕΚΚΙΝΗΘΕΙ ΣΩΣΤΑ ΤΟΤΕ

model compiled

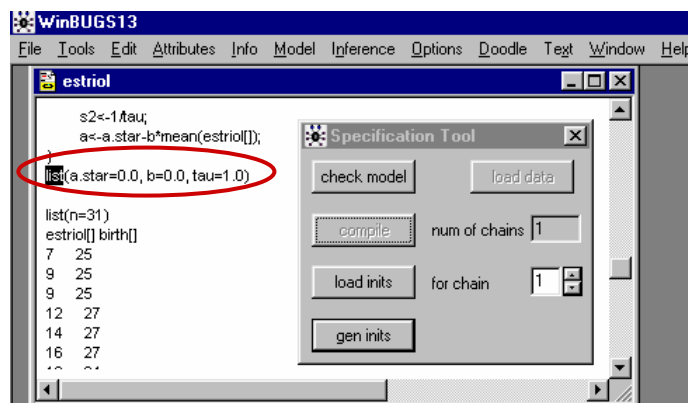


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘4... Φόρτωση ή Προσομοίωση Αρχικών Τιμών

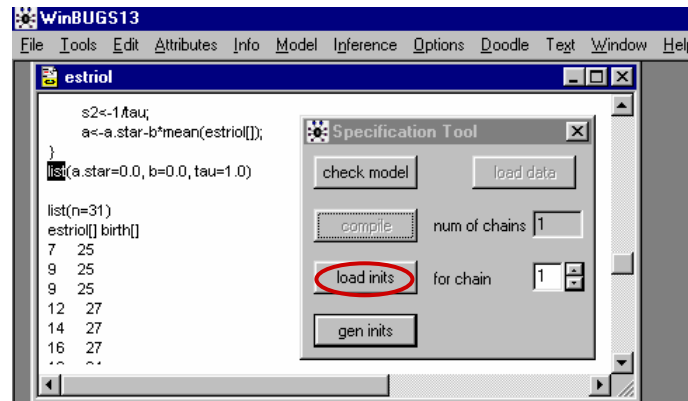
ΜΑΥΡΙΖΟΥΜΕ ΤΗΝ ΕΝΤΟΛΗ **list** ΤΩΝ ΑΡΧΙΚΩΝ ΤΙΜΩΝ



4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

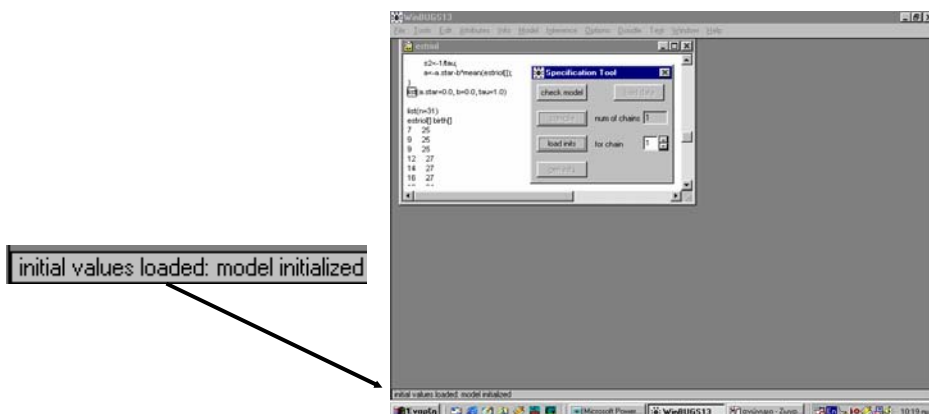
⌘ 4... Φόρτωση ή Προσομοίωση Αρχικών Τιμών
ΕΠΙΛΕΓΟΥΜΕ ΤΟ ΚΟΥΤΙ "LOAD INITS"



4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘ 4... Φόρτωση ή Προσομοίωση Αρχικών Τιμών
ΕΑΝ ΟΙ ΑΡΧΙΚΕΣ ΤΙΜΕΣ ΦΟΡΤΩΘΟΥΝ ΣΩΣΤΑ ΤΟΤΕ



4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘ 4... Φόρτωση ή Προσομοίωση Αρχικών Τιμών

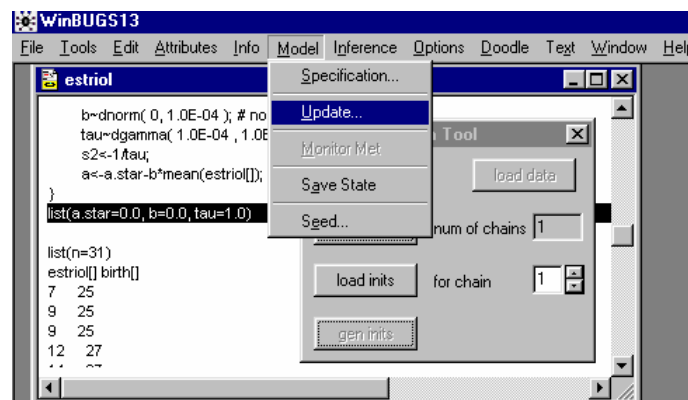
⌘ ΤΟ WINBUGS ΕΙΝΑΙ ΕΤΟΙΜΟ ΓΙΑ ΝΑ ΠΡΟΣΟΜΟΙΩΣΕΙ ΕΝΑ ΔΕΙΓΜΑ ΑΠΟ ΤΗΝ *ΕΚ-ΤΩΝ-ΥΣΤΕΡΩΝ ΚΑΤΑΝΟΜΗ* ΜΕ ΤΗ ΧΡΗΣΗ ΤΟΥ ΔΕΙΓΜΑΤΟΛΗΠΤΗ GIBBS .

4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘ 5... Προσομοίωση Τιμών Burn-in

ΕΠΙΛΕΓΟΥΜΕ "UPDATE" ΣΤΟ ΜΕΝΟΥ "MODEL"

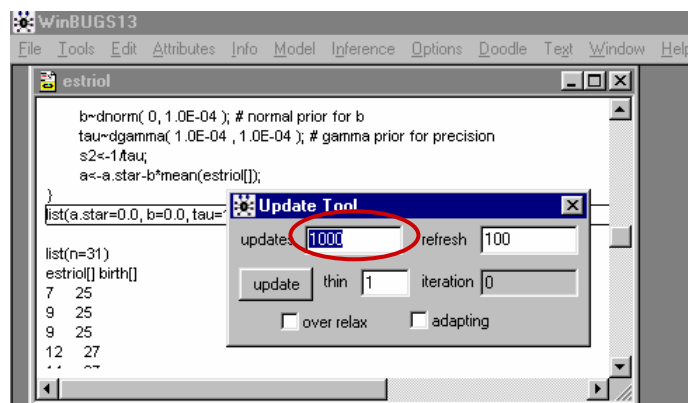


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘5... Προσομοίωση Τιμών Burn-in

ΓΡΑΦΟΥΜΕ ΣΤΟ **updates** ΤΟΝ ## ΤΩΝ BURN-IN ΕΠΑΝΑΛΗΨΕΩΝ

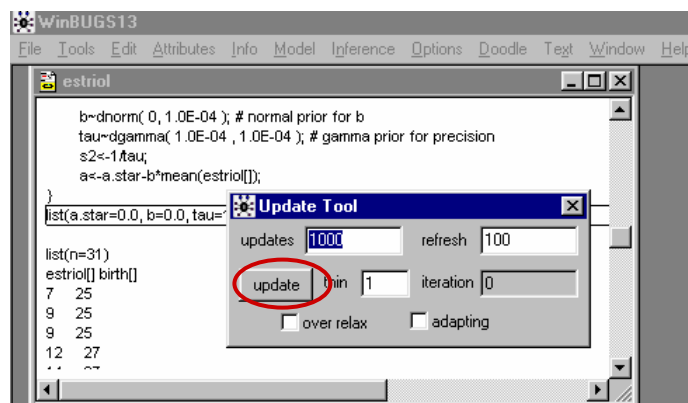


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘5... Προσομοίωση Τιμών Burn-in

ΕΠΙΛΕΓΟΥΜΕ ΤΟ **UPDATE** ΓΙΑ ΝΑ ΠΡΟΣΟΜΟΙΩΣΟΥΜΕ

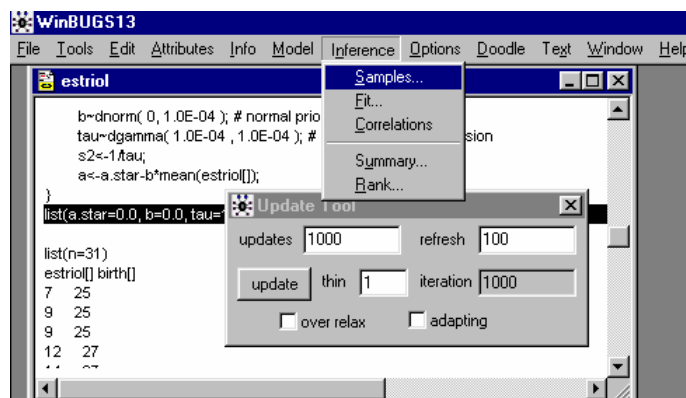


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘6... Παρακολούθηση Παραμέτρων

ΕΠΙΛΕΓΟΥΜΕ ΤΟ "**SAMPLES**" ΣΤΟ ΜΕΝΟΥ "**INFERENCE**"

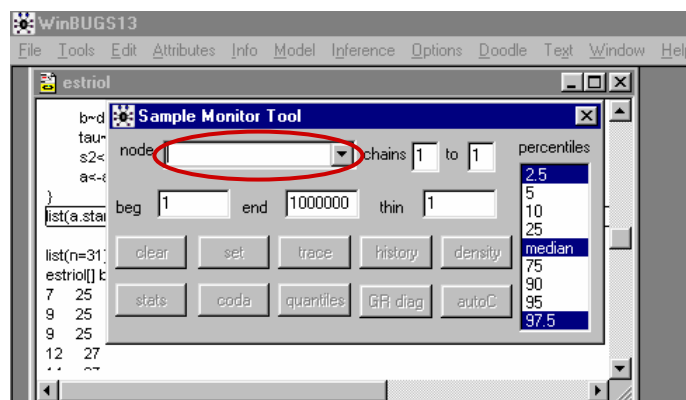


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘6... Παρακολούθηση Παραμέτρων

ΓΡΑΦΟΥΜΕ ΣΤΟ **node** ΤΟ ΟΝΟΜΑ ΤΗΣ ΠΑΡΑΜΕΤΡΟΥ

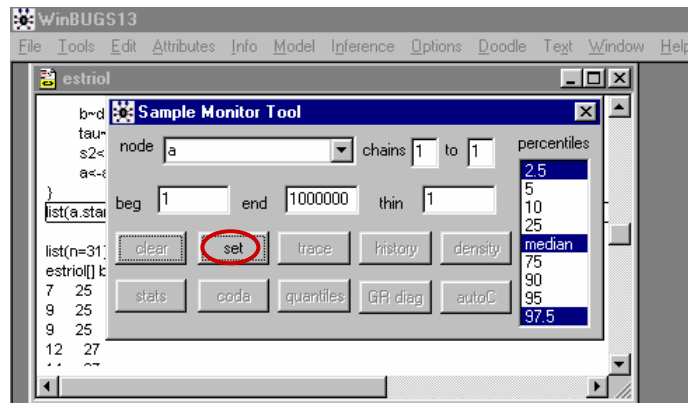


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘6... Παρακολούθηση Παραμέτρων

ΚΑΙ ΜΕΤΑ ΕΠΙΛΕΓΟΥΜΕ ΤΟ ΚΟΥΤΙ "SET"



4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘6... Παρακολούθηση Παραμέτρων

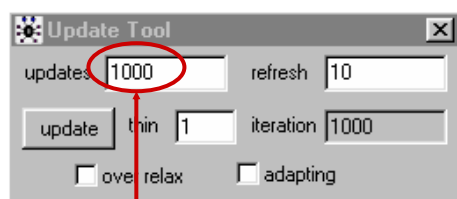
- ☒ ΓΡΑΦΟΥΜΕ "a" ΚΑΙ ΕΠΙΛΕΓΟΥΜΕ ΤΟ ΚΟΥΤΙ "SET"
- ☒ ΓΡΑΦΟΥΜΕ "b" ΚΑΙ ΕΠΙΛΕΓΟΥΜΕ ΤΟ ΚΟΥΤΙ "SET"
- ☒ ΓΡΑΦΟΥΜΕ "s2" ΚΑΙ ΕΠΙΛΕΓΟΥΜΕ ΤΟ ΚΟΥΤΙ "SET"

4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘7... Προσομοίωση Εκ-των-υστέρων Τιμών

ΕΠΙΛΕΓΟΥΜΕ ΤΟ "UPDATE TOOL"



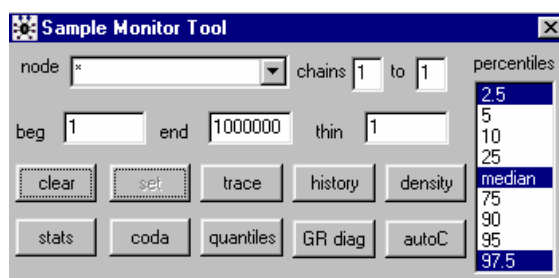
ΓΡΑΦΟΥΜΕ ΣΤΟ *updates* ΤΟΝ ΑΡΙΘΜΟ ΤΩΝ ΕΠΑΝΑΛΗΨΕΩΝ ΠΟΥ ΕΠΙΘΥΜΟΥΜΕ (ΕΔΩ 1000)

4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘8... Περιγραφή της Εκ-των-Υστέρων Κατανομής

ΕΠΙΛΕΓΟΥΜΕ ΤΟ "SAMPLE MONITOR TOOL".



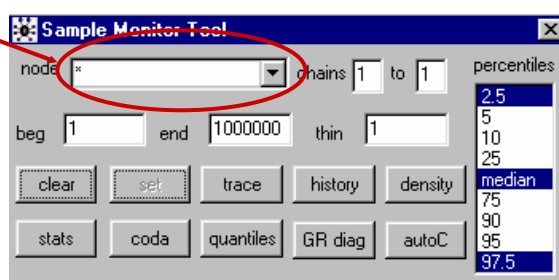
4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘8... Περιγραφή της Εκ-των-Υστέρων Κατανομής

ΓΡΑΦΟΥΜΕ ΣΤΟ **node** ΤΟ ΟΝΟΜΑ ΤΗΣ ΠΑΡΑΜΕΤΡΟΥ ΠΟΥ ΘΕΛΟΥΜΕ ΝΑ ΠΕΡΙΓΡΑΦΟΥΜΕ

[*=ΟΛΕΣ ΟΙ ΠΑΡΑΜΕΤΡΟΙ ΠΟΥ ΔΗΛΩΣΑΜΕ ΣΤΟ MONITOR]

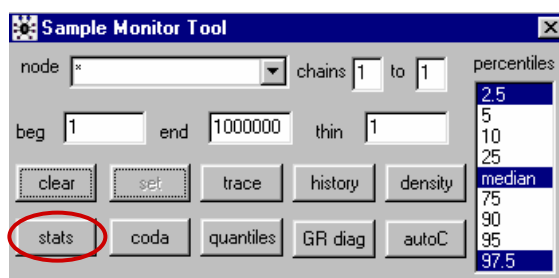


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘8... Περιγραφή της Εκ-των-Υστέρων Κατανομής

ΕΠΙΛΕΓΟΥΜΕ ΤΟ ΚΟΥΤΙ "**STATS**".

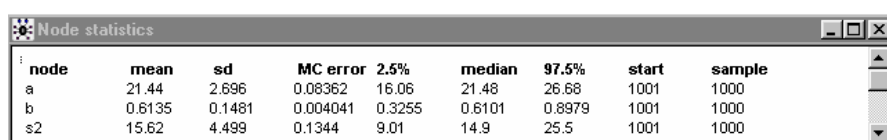


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

8... Περιγραφή της Εκ-των-Υστέρων Κατανομής

ΒΛΕΠΟΥΜΕ ΤΟΥΣ ΠΕΡΙΓΡΑΦΙΚΟΥΣ ΣΤΑΤΙΣΤΙΚΟΥΣ ΔΕΙΚΤΕΣ ΤΗΝ ΕΚ-ΤΩΝ-ΥΣΤΕΡΩΝ ΚΑΤΑΝΟΜΗΣ.



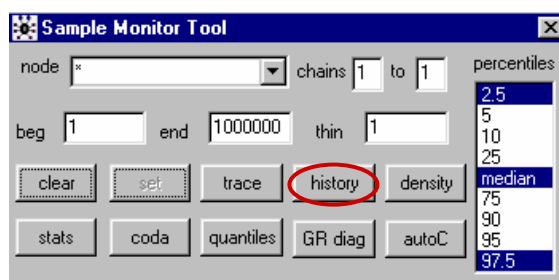
node	mean	sd	MC error	2.5%	median	97.5%	start	sample
a	21.44	2.696	0.08362	16.06	21.48	26.68	1001	1000
b	0.6135	0.1481	0.004041	0.3255	0.6101	0.8979	1001	1000
s2	15.62	4.499	0.1344	9.01	14.9	25.5	1001	1000

4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

8... Περιγραφή της Εκ-των-Υστέρων Κατανομής

ΕΠΙΛΕΓΟΝΤΑΣ ΤΟ ΚΟΥΤΙ "HISTORY" ΣΤΟ "SAMPLE MONITOR TOOL"...

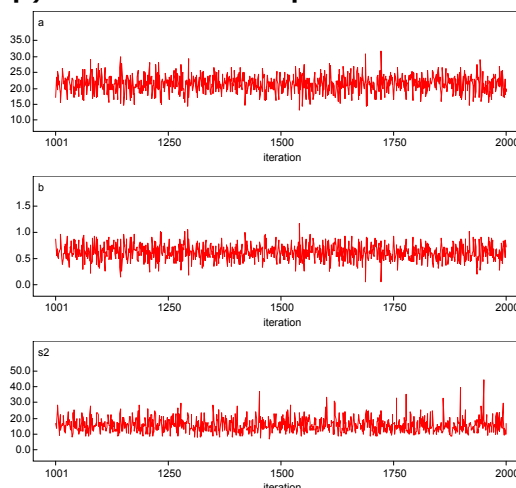


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘8... Περιγραφή της Εκ-των-Υστέρων Κατανομής

ΠΑΙΡΝΟΥΜΕ ΤΟ
ΔΙΑΓΡΑΜΜΑ ΤΩΝ
ΠΡΟΣΟΜΟΙΩΜΕΝΩΝ
ΔΙΑΓΡΑΜΜΑΤΩΝ
(TRACE PLOTS) .

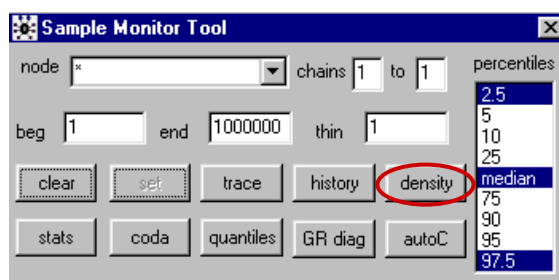


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘8... Περιγραφή της Εκ-των-Υστέρων Κατανομής

ΕΠΙΛΕΓΟΝΤΑΣ ΤΟ ΚΟΥΤΙ "DENSITY" ΣΤΟ "SAMPLE MONITOR TOOL" . . .

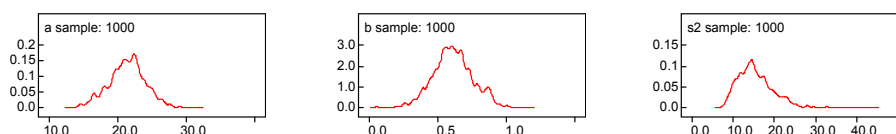


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘8... Περιγραφή της Εκ-των-Υστέρων Κατανομής

ΠΑΙΡΝΟΥΜΕ ΤΑ ΔΙΑΓΡΑΜΜΑΤΑ ΤΗΣ ΕΚΤΙΜΩΜΕΝΗΣ ΣΥΝΑΡΤΗΣΗΣ ΠΥΚΝΟΤΗΤΑΣ ΠΙΘΑΝΟΤΗΤΑΣ ΤΗΣ ΕΚ-ΤΩΝ-ΥΣΤΕΡΩΝ ΚΑΤΑΝΟΜΗΣ

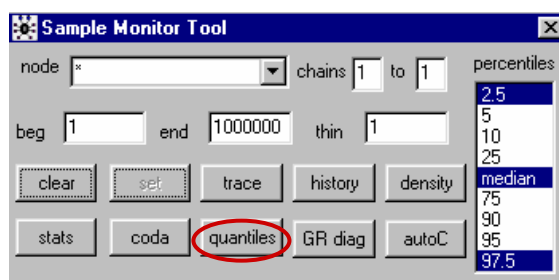


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘8... Περιγραφή της Εκ-των-Υστέρων Κατανομής

ΕΠΙΛΕΓΟΝΤΑΣ ΤΟ ΚΟΥΤΙ "QUANTILES" ΣΤΟ "SAMPLE MONITOR TOOL"...

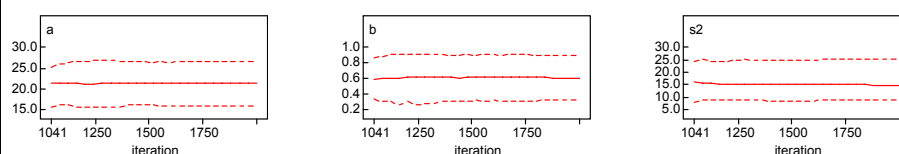


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘8... Περιγραφή της Εκ-των-Υστέρων Κατανομής

ΠΑΙΡΝΟΥΜΕ ΤΑ ΔΙΑΓΡΑΜΜΑΤΑ ΤΩΝ ΠΟΣΟΣΤΗΜΟΡΙΩΝ ΤΗΣ ΕΚ-ΤΩΝ-ΥΣΤΕΡΩΝ ΚΑΤΑΝΟΜΗΣ

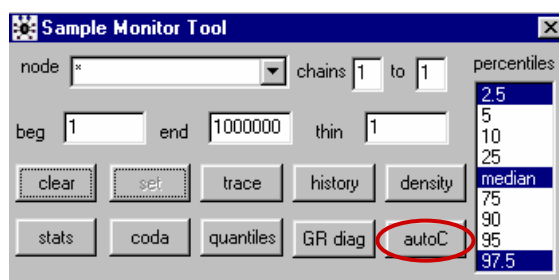


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘8... Περιγραφή της Εκ-των-Υστέρων Κατανομής

ΕΠΙΛΕΓΟΝΤΑΣ ΤΟ ΚΟΥΤΙ "AUTOC" ΣΤΟ "SAMPLE MONITOR TOOL"...

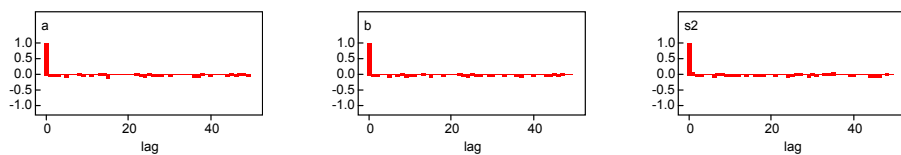


4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘8... Περιγραφή της Εκ-των-Υστέρων Κατανομής

ΠΑΙΡΝΟΥΜΕ ΤΑ ΔΙΑΓΡΑΜΜΑΤΑ ΤΩΝ ΑΥΤΟ-ΣΥΣΧΕΤΙΣΕΩΝ ΤΩΝ ΠΡΟΣΟΜΟΙΩΜΕΝΩΝ ΤΙΜΩΝ



4... WINBUGS

4.2.2. Προσομοίωση στο WINBUGS

⌘8... Περιγραφή της Εκ-των-Υστέρων Κατανομής

- ☒ ΤΟ ΚΟΥΤΙ "**TRACE**" ΠΑΡΑΓΕΙ ΕΝΑ ΔΥΝΑΜΙΚΟ ΔΙΑΓΡΑΜΜΑ ΤΩΝ ΠΡΟΣΟΜΟΙΩΜΕΝΩΝ ΤΙΜΩΝ ON LINE
- ☒ ΤΟ ΚΟΥΤΙ "**CODA**" ΠΑΡΑΓΕΙ ΜΙΑ ΛΙΣΤΑ ΜΕ ΤΙΣ ΠΡΟΣΟΜΟΙΩΜΕΝΕΣ ΤΙΜΕΣ ΠΟΥ ΜΠΟΡΕΙ ΝΑ ΔΙΑΒΑΣΤΕ ΑΠΟ ΤΟ ΠΡΟΓΡΑΜΜΑ CODA ΠΟΥ ΔΟΥΛΕΥΕΙ ΣΤΟ R Ή ΣΤΟ SPLUS.
- ☒ ΤΟ ΚΟΥΤΙ "**GR DIAG**" ΠΑΡΑΓΕΙ ΕΝΑ ΔΙΑΓΝΩΤΙΚΟ ΕΛΕΓΧΟ ΓΙΑ ΤΗ ΣΥΓΚΛΙΣΗ ΤΗΣ ΑΛΥΣΙΔΑΣ (ΓΙΑ ΠΟΛΛΑΠΛΕΣ ΑΛΥΣΙΔΕΣ)

4... WINBUGS

4.3. Καθορισμός Νεας Πιθανοφάνειας και Prior

⌘ [MONO STO WINBUGS]

⌘ Τι γίνεται αν θέλουμε να χρησιμοποιήσουμε μια κατανομή που δεν υπάρχει στο WINBUGS;

⌘ **ΤΟ ΚΟΛΠΟ ΜΕ ΤΑ ΜΗΔΕΝΙΚΑ!**

4... WINBUGS

4.3.1 Καθορισμός Νεας Prior

⌘ [MONO STO WINBUGS]

⌘ **ΤΟ ΚΟΛΠΟ ΜΕ ΤΑ ΜΗΔΕΝΙΚΑ!**

Έστω ότι θέλουμε να ορίσουμε μια prior $f(\theta)$ για την παράμετρο θ

⌘ 1... Θέτουμε ότι η prior του θ είναι επίπεδη στο πεδίο τιμών της (`dflat()` ή `dunif()`)

⌘ 2... Θέτουμε μια μεταβλητή (π.χ. `zero`) ίση με μηδέν

⌘ 3... Ορίζουμε ότι ακολουθεί την κατανομή **Poisson** μέ μέσο **λ**

⌘ 4... Θέτουμε **$\lambda = -\log f(\theta)$**

4... WINBUGS

4.3.1 Καθορισμός Νεας Prior

⌘ [ΜΟΝΟ ΣΤΟ WINBUGS]

⌘ ΤΟ ΚΟΛΠΟ ΜΕ ΤΑ ΜΗΔΕΝΙΚΑ!

Π.χ. $f(\theta) = z(z + \omega\theta)^{\theta-1} e^{-(z + \omega\theta)}/(\theta!)$.

Γενικευμένη/Lagrangian Poisson με μέσο $= z/(1-\omega)$,

Διακύμανση $= z/(1-\omega)^3$, $DI = 1/(1-\omega)^2$

⌘ `theta ~ dflat()`

⌘ `zero <- 0`

⌘ `zero ~ dpois(lambda)`

⌘ `lambda <- -(log(zeta)+(theta-1)*
log(zeta+omega*theta)-(zeta+omega*theta)-
logfact(theta))`

4... WINBUGS

4.3.1 Καθορισμός Νεας Prior

Γιατί δουλεύει αυτό το κόλπο;

$f(\text{zero}|\theta)$ είναι η Poisson με μέση τιμή ίση $-\log\text{-likelihood}$

$\varphi(\theta)$ είναι η επίπεδη ψεύτο-prior που χρησιμοποιούμε για να ορίσουμε έμμεσα την πραγματική prior.

$f(\theta|\text{zero})$ είναι η prior που τελικά ορίζουμε.

$$f(\theta|\text{zero}) \propto f(\text{zero}=0|\theta) \times \varphi(\theta) = e^{-\lambda} \times 1 = e^{-(-\log f(\theta))} \times 1 = f(\theta)$$

4... WINBUGS

4.3.1 Καθορισμός Νεας Prior

⌘ [ΜΟΝΟ ΣΤΟ WINBUGS]

⌘ ΤΟ ΚΟΛΠΟ ΜΕ ΤΑ ΜΗΔΕΝΙΚΑ!

- ⌘ Δείτε στο manual του WINBUGS στην ενότητα [new-prior](#) για παράδειγμα με την κανονική κατανομή
- ⌘ ΠΡΟΣΟΧΗ: Αυτή η μέθοδος παράγει δείγματα με
 - ☑ Μεγάλη αυτο-συσχέτιση
 - ☑ Αργή Σύγκλιση
 - ☑ Υψηλά Μόντε Κάρλο σφάλματα
- ⌘ είναι δηλ. αργή υπολογιστικά και χρειάζεται να αφήσουμε το WINBUGS να τρέξει για αρκετές επαναλήψεις.

4... WINBUGS

4.3.2 Καθορισμός Νεας Πιθανοφάνειας

Έστω ότι θέλουμε να ορίσουμε με $y_i \sim f(y_i | \theta)$ με παράμετρο θ

- ⌘ 1... Θέτουμε μια σταθερά C ίση με ένα μεγάλο νούμερο για να εξασφαλίσουμε ότι το
- ⌘ 2... Θέτουμε ένα διάνυσμα ψευτο-δεδομένων **zero** (με μήκος ίσο με το μέγεθος των πραγματικών μας δεδομένων) ίσο με μηδέν
- ⌘ 3... Ορίζουμε ότι ακολουθεί την κατανομή **Poisson** μέ μέσο λ_i
- ⌘ 4... Θέτουμε $\lambda_i = -\log f(y_i | \theta)$

4... WINBUGS

4.3.2 Καθορισμός Νεας Πιθανοφάνειας

Π.χ. $f(y_i | z, \omega) = \frac{z (z + \omega y_i)^{y_i-1} e^{-(z + \omega y_i)}}{(y_i!)}$.

Γενικευμένη/Lagrangian Poisson με

παραμέτρους $\theta = (z, \omega)$

μέσος $= z/(1-\omega)$ και Διακύμανση $= z/(1-\omega)^3$,

```
C <- 10000
```

```
for (i in 1:N) {
  zeros[i] <- 0
  zeros[i] ~ dpois(lambda[i])
  lambda[i] <- -L[i] + C
  L[i] <- -( log(zeta)+(y[i]-1)*
log(zeta+omega*y[i])-(zeta+omega*y[i])-
logfact(y[i])) )
}
```

4... WINBUGS

4.3.2 Καθορισμός Νεας Πιθανοφάνειας

Γιατί δουλεύει αυτό το κόλπο;

$$\begin{aligned} \prod_{i=1}^n f(zero_i | y_i, \theta) &= \prod_{i=1}^n e^{-\lambda_i} = \prod_{i=1}^n e^{-(-L_i + C)} \\ &= \prod_{i=1}^n e^{\log f(y_i | \theta) - C} \propto \prod_{i=1}^n f(y_i | \theta) \end{aligned}$$

4... WINBUGS

4.3.2 Καθορισμός Νεας Πιθανοφάνειας

ΠΑΡΑΔΕΙΓΜΑ ΓΕΝΙΚΕΥΜΕΝΗΣ POISSON

(Rosner 1994, page 94)

Αριθμός Θανάτων από πολιομυελίτιδα την περίοδο
1968-76

<u>1968</u>	<u>1969</u>	<u>1970</u>	<u>1971</u>	<u>1972</u>	<u>1973</u>	<u>1974</u>	<u>1975</u>	<u>1976</u>
24	13	7	18	2	10	3	9	16

4... WINBUGS

4.3.2 Καθορισμός Νεας Πιθανοφάνειας

ΠΑΡΑΔΕΙΓΜΑ ΓΕΝΙΚΕΥΜΕΝΗΣ POISSON

```

model {
  C<-10000
  for (i in 1:9) {
    zeros[i]<-0
    zeros[i]~dpois( lambda[i] )
    lambda[i]<- C - loglike[i]
    loglike[i] <- log(zeta)+(y[i]-1)* log(zeta+omega*y[i]) -
      (zeta+omega*y[i]) - logfact(y[i])
  }
  zeta~dgamma(0.001, 0.001)
  omega~dbeta(1,1)
  mean<-zeta/(1-omega)
  var<-zeta/pow(1-omega,3)
  DI<-1/((1-omega)*(1-omega))
}
DATA : list( y=c(24, 13, 7, 18, 2, 10, 3, 9, 16) )
INITS: list( zeta=1, omega=0.5 )

```


4... WINBUGS

4.4. Προσομοίωση στο Παρασκήνιο

4.4.1 ΕΙΣΑΓΩΓΗ

⌘ BATCH MODE METHOD: SCRIPTING

⌘ [μόνο στο WINBUGS 1.4].

⌘ Εναλλακτικός τρόπος προσομοίωσης χωρίς να περιμένουμε για αποτελέσματα

⌘ Χρειάζονται τουλάχιστον 4 αρχεία σε WINBUGS (*.odc) ή text (*.txt) format

☒ Εντολές Προσομοίωσης (script)

☒ Μοντέλο

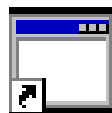
☒ Δεδομένων (μπορεί να είναι περισσότερα από 1)

4... WINBUGS

4.4. Προσομοίωση στο Παρασκήνιο

Παράδειγμα script.odc

Επιλέξτε



BackBugs14.Ink

Στον κατάλογο του WINBUGS14

συνήθως **c:\Program Files\Winbugs14**

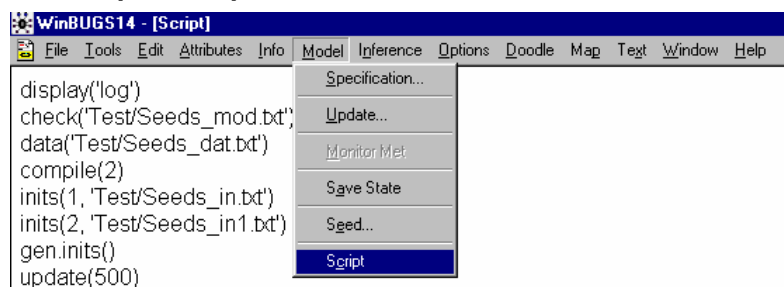
4... WINBUGS

4.4. Προσομοίωση στο Παρασκήνιο

Ανοίγουμε το script file

c:\Program Files\Winbugs14\script.odc

και τρέχουμε το μοντέλο στο παρασκήνιο
επιλέγοντας **MODEL>SCRIPT**



4... WINBUGS

4.4. Προσομοίωση στο Παρασκήνιο

4.4.2. Μερικές Σημαντικές Εντολές

- ⌘ **display('log')** : Ανοίγει ένα αρχείο log που αποθηκεύει τα αποτελέσματα
- ⌘ **check('Test/Seeds_mod.txt')** : Έλεγχος μοντέλου που βρίσκεται στο αρχείο **Seeds_mod.txt** και στον υποκατάλογο του WINBUGS, **Test**.
- ⌘ **data('Test/Seeds_dat.txt')** : Φόρτωση Δεδομένων από το αρχείο **Seeds_dat.txt** και στον υποκατάλογο του WINBUGS, **Test**.
- ⌘ **compile(2)** : Εκκίνηση 2 προσομοιωμένων αλυσίδων.
- ⌘ **inits(1, 'Test/Seeds_in.txt')** : Φόρτωση Αρχικών Τιμών της 1ης αλυσίδας το αρχείο **Seeds_in.txt** και στον υποκατάλογο του WINBUGS, **Test**.

4... WINBUGS

4.4. Προσομοίωση στο Παρασκήνιο

4.4.2. Μερικές Σημαντικές Εντολές

- ⌘ **gen.inits()** : Προσομοίωση Αρχικών Τιμών .
- ⌘ **update(500)** : Προσομοίωση 500 Τιμών (Burn-in) .
- ⌘ **set(alpha0)** : Αρχίζουμε και αποθηκεύουμε τις προσομοιωμένες τιμές για την παράμετρο **alpha0** .
- ⌘ **update(1000)** : Προσομοίωση 1000 Τιμών .
- ⌘ **stats(*)** : Στατιστικοί δείκτες για το προσομοιωμένο δείγμα όλων των παραμέτρων έχουν οριστεί μέσω του της εντολής **set** .
- ⌘ **history(*)** : Διάγραμμα προσομοιωμένων τιμών ανά επανάληψη για τις παραμέτρους που έχουν οριστεί μέσω του της εντολής **set** .
- ⌘ **trace(*)** : Δυναμικό (on-line) Διάγραμμα προσομοιωμένων τιμών ανά επανάληψη για τις παραμέτρους που έχουν οριστεί μέσω του της εντολής **set** .

4... WINBUGS

4.4. Προσομοίωση στο Παρασκήνιο

4.4.2. Μερικές Σημαντικές Εντολές

- ⌘ **density(*)** : Διάγραμμα της εκτιμώμενης εκ-των-υστέρων συνάρτησης πιθανότητας ή πυκνότητας πιθανότητας για τις παραμέτρους που έχουν οριστεί μέσω του της εντολής **set** .
- ⌘ **autoC(*)** : Διάγραμμα αυτοσυσχετίσεων των προσομοιωμένων τιμών για τις παραμέτρους που έχουν οριστεί μέσω του της εντολής **set** .
- ⌘ **quantiles(*)** : Διάγραμμα ποσοστημορίων των προσομοιωμένων τιμών ανά επανάληψη για τις παραμέτρους που έχουν οριστεί μέσω του της εντολής **set** .

4... WINBUGS

4.4. Προσομοίωση στο Παρασκήνιο

4.4.2. Μερικές Σημαντικές Εντολές

- ⌘ **coda(*,output)** : Αποθήκευση όλων των monitored παραμέτρων στο αρχείο **output** το οποίο είναι σε format CODA. Αν το όνομα του αρχείου μείνει κενό ανοίγει παράθυρο στο WINBUGS.
- ⌘ **save('seedsLog')** : Αποθήκευση όλων των αποτελεσμάτων του log παραθύρου στο αρχείο **seedLog.odc** (format WINBUGS μαζί με διαγράμματα). Αν το αρχείο έχει κατάληξη **txt** τότε αποθηκεύονται σε αρχείο text μόνο το κείμενο (όχι διαγράμματα και άλλα γραφικά).
- ⌘ **quit()** : Έξοδος από το WINBUGS.

4... WINBUGS

4.4. Προσομοίωση στο Παρασκήνιο

4.4.2. Μερικές Σημαντικές Εντολές

- ⌘ **coda(*,output)** : Αποθήκευση όλων των monitored παραμέτρων στο αρχείο **output** το οποίο είναι σε format CODA. Αν το όνομα του αρχείου μείνει κενό ανοίγει παράθυρο στο WINBUGS.
- ⌘ **save('seedsLog')** : Αποθήκευση όλων των αποτελεσμάτων του log παραθύρου στο αρχείο **seedLog.odc** (format WINBUGS μαζί με διαγράμματα). Αν το αρχείο έχει κατάληξη **txt** τότε αποθηκεύονται σε αρχείο text μόνο το κείμενο (όχι διαγράμματα και άλλα γραφικά).
- ⌘ **quit()** : Έξοδος από το WINBUGS.

4... WINBUGS

4.4. Προσομοίωση στο Παρασκήνιο

4.4.3. Προσομοιώνοντας το Παράδειγμα Εστριόλης στο Παρασκήνιο

1... ΦΤΙΑΧΝΟΥΜΕ 5 ΑΡΧΕΙΑ

⌘ script.odc

⌘ model.odc

⌘ data.odc

⌘ data2.odc

⌘ inits.odc

2... ΑΝΟΙΓΟΥΜΕ ΤΟ **SCRIPT.ODC** ΚΑΙ ΤΡΕΧΟΥΜΕ ΤΟ ΜΟΝΤΕΛΟ
ΣΤΟ ΠΑΡΑΣΚΗΝΙΟ ΑΠΟ ΤΟ ΜΕΝΟΥ **MODEL>SCRIPT**.

Bayesian Biostatistics Using BUGS

ΤΕΛΟΣ ΔΕΥΤΕΡΟΥ ΜΑΘΗΜΑΤΟΣ

E-mail: ntzoufras@aueb.gr



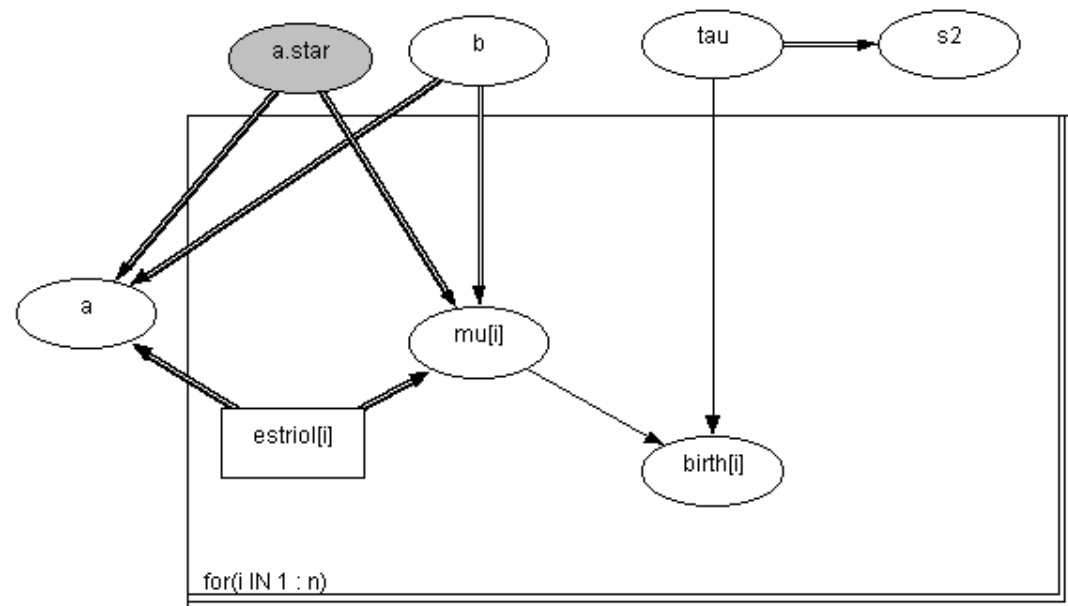
Department of Statistics,
Athens University of
Economics & Business

ΠΑΡΑΡΤΗΜΑ Β (2ου ΜΑΘΗΜΑΤΟΣ): ΠΑΡΑΔΕΙΓΜΑΤΑ WINBUGS

1 EXAMPLE 1: BIRTHWEIGHT & ESTRIOI LEVEL

```
{
#      definition of likelihood function
#
  for (i in 1:n) {
    birth[i]~dnorm( mu[i], tau ); # random component
    mu[i]<-a.star+b*(estriol[i]-mean(estriol[]));           # systematic component
                                                            #      & link function
  }
#      prior distributions
#
  a.star~dnorm( 0, 1.0E-04 ); # normal prior for a
  b~dnorm( 0, 1.0E-04 ); # normal prior for b
  tau~dgamma( 1.0E-04 , 1.0E-04 ); # gamma prior for precision
  s2<-1/tau;
  a<-a.star-b*mean(estriol[]);
}
list(a.star=0.0, b=0.0, tau=1.0) # initial values

# data
list(n=31)
estriol[] birth[]
7      25
9      25
9      25
12     27
14     27
16     27
16     24
14     30
16     30
16     31
17     30
19     31
21     30
24     28
15     32
16     32
17     32
25     32
27     34
15     34
15     34
15     35
16     35
19     34
18     35
17     36
18     37
20     38
22     40
25     39
24     43
END
```



2 Example 2: BEETLES DATASET

```

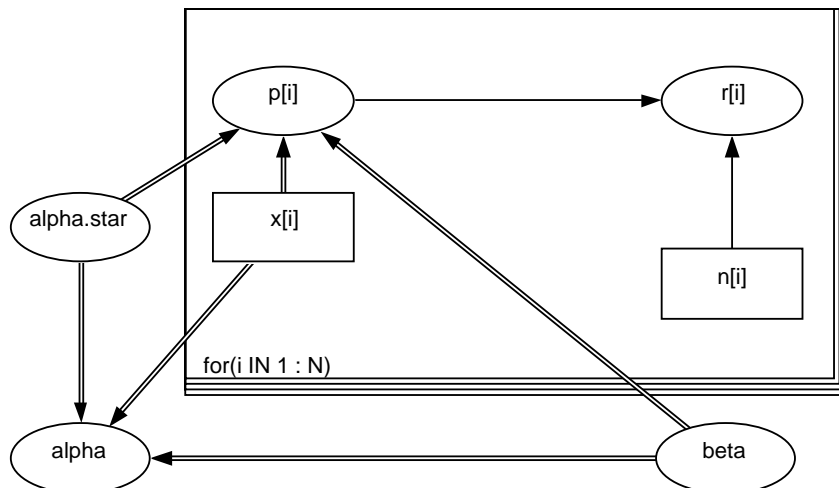
model beetles;
{
  for (i in 1:N) {
    r[i] ~ dbin(p[i], n[i]);
    logit(p[i]) <- alpha.star + beta*(x[i]-mean(x[]));
    r.hat[i] <- p[i]*n[i]; # fitted values
  }
  alpha.star ~ dnorm(0.0, 1.0E-3);
  beta ~ dnorm(0.0, 1.0E-3);
  alpha <- alpha.star - beta*mean(x[]);
  odds.ratio <- exp( beta )
}

```

```

list(x = c(1.6907, 1.7242, 1.7552, 1.7842, 1.8113, 1.8369, 1.8610, 1.8839)   n = c(59, 60, 62, 56, 63, 59, 62, 60)   r = c(6, 13, 18, 28,
52, 53, 61, 60), N=8) #data
list(alpha.star=0, beta=0) #initial values

```



3 EXAMPLE 3: ΧΡΗΣΗ ΝΕΑΣ ΠΙΘΑΝΟΦΑΝΕΙΑΣ (ΓΕΝΙΚΕΥΜΕΝΗ POISSON)

```
model
{
  C<-10000
  for (i in 1:9) {
    zeros[i]<-0
    zeros[i]~dpois( lambda[i] )
    lambda[i]<- C - loglike[i]
    loglike[i] <- log(zeta)+(y[i]-1)* log(zeta+omega*y[i])-(zeta+omega*y[i])-logfact(y[i])
  }
  zeta~dgamma(0.001, 0.001)
  omega~dbeta(1,1)

  mean<-zeta/(1-omega)
  var<-zeta/pow(1-omega,3)
  DI<-1/((1-omega)*(1-omega))
}
```

DATA

```
list( y=c(24, 13, 7, 18, 2, 10, 3, 9, 16) )
```

INITS

```
list( zeta=1, omega=0.5 )
```

4 EXAMPLE 4: ΠΡΟΣΟΜΟΙΩΣΗ ΣΤΟ ΠΑΡΑΣΚΗΝΙΟ ΤΟΥ ΠΑΡΑΔΕΙΓΜΑΤΟΣ ESTRIOL

4.1 APXEIO script.odc

```
display('log')
check('c:/myfiles/courses/BUGS_course/01_presentations/Bugs_files/lec2/ex1_estriol_back/model.odc')
data('c:/myfiles/courses/BUGS_course/01_presentations/Bugs_files/lec2/ex1_estriol_back/data.odc')
data('c:/myfiles/courses/BUGS_course/01_presentations/Bugs_files/lec2/ex1_estriol_back/data2.odc')
compile(1)
inits(1,'c:/myfiles/courses/BUGS_course/01_presentations/Bugs_files/lec2/ex1_estriol_back/inits.odc')

update(1000)
set(a)
set(b)
update(1000)
trace(*)
update(1000)
stats(*)
history(*)
density(*)
autoC(*)
quantiles(*)
dic.stats()
coda(*,output)
save('seedsLog')
```

4.2 APXEIO model.odc

```
{
#      definition of likelihood function
#
      for (i in 1:n) {
        birth[i]~dnorm( mu[i], tau ); # random component
        mu[i]<-a.star+b*(estriol[i]-mean(estriol[])); # systematic component
                                                    #      & link function
      }
#      prior distributions
#
      a.star~dnorm( 0, 1.0E-04 ); # normal prior for a
      b~dnorm( 0, 1.0E-04 ); # normal prior for b
      tau~dgamma( 1.0E-04 , 1.0E-04 ); # gamma prior for precision
      s2<-1/tau;
      a<-a.star-b*mean(estriol[]);
}
```

4.3 APXEIO data.odc

```
estriol[] birth[]
7      25
9      25
9      25
12     27
14     27
16     27
16     24
14     30
16     30
16     31
17     30
19     31
21     30
24     28
15     32
16     32
17     32
25     32
27     34
15     34
15     34
15     35
16     35
19     34
18     35
17     36
18     37
20     38
22     40
25     39
24     43
END
```

4.4 APXEIO data2.odc

```
list(n=31)
```

4.5 APXEIO inits.odc

```
list(a.star=0.0, b=0.0, tau=1.0) # initial values
```

Bayesian Biostatistics Using BUGS



**Βιο-Στατιστική κατά Bayes
με τη χρήση του Λογισμικού
BUGS**

ΜΑΘΗΜΑ 3: ΣΥΝΘΕΤΑ ΜΟΝΤΕΛΑ

Ioannis Ntzoufras

E-mail: ntzoufras@aueb.gr



Department of Statistics,
Athens University of
Economics & Business

5. ΣΥΝΘΕΤΑ ΜΟΝΤΕΛΑ ΣΤΟ BUGS

ΜΑΘΗΜΑ 3 (ΠΕΡΙΕΧΟΜΕΝΑ):

- ⌘ 5.1 Prior Distributions
- ⌘ 5.2 Parameterization
- ⌘ 5.3 Random Effects
- ⌘ 5.4 Examples
 - ☒ 5.4.1 Rats (Repeated Measures Regression)
 - ☒ 5.4.2 Seeds (Random Effects Logistic Regression)
 - ☒ 5.4.3 Surgical (Institutional Ranking)
 - ☒ 5.4.4 Equiv (Bioequivalence in Cross-over Trials)

5. ΣΥΝΘΕΤΑ ΜΟΝΤΕΛΑ ΣΤΟ BUGS

ΜΑΘΗΜΑ 3 (ΠΕΡΙΕΧΟΜΕΝΑ):

- ⌘ 5.1 Prior Distributions
- ⌘ 5.2 Parameterization
- ⌘ 5.3 Random Effects
- ⌘ 5.4 Examples
 - ☒ 5.4.5 Dyes (Variance Component Models)
 - ☒ 5.4.6 Blocker (Meta analysis & random Effects)
 - ☒ 5.4.7 Mice (Weibull Survival Analysis)
 - ☒ 5.4.8 Alli(Multinomial-Logistic Models)

5.1. ΚΑΘΟΡΙΣΜΟΣ PRIOR ΚΑΤΑΝΟΜΩΝ

- ⌘ Εάν έχουμε εκ-των-προτέρων πληροφορία από άλλες μελέτες τότε μπορούμε να χρησιμοποιήσουμε:
 - ☒ Κανονική για ποσοτικές (μέσος=prior belief, διακύμανση = αβεβαιότητα για την prior γνώση)
 - ☒ Γάμμα ή Λογαριθμο-κανονική για ποσοτικές θετικές
 - ☒ Βήτα για ποσοστά ή παραμέτρους στο $[0,1]$

5.1. ΚΑΘΟΡΙΣΜΟΣ PRIOR ΚΑΤΑΝΟΜΩΝ

Μη πληροφοριακές εκ-των-προτέρων κατανομές (non-informative prior distributions) :

- ⌘ δεν επηρεάζουν την συμπερασματολογία μας
- ⌘ βασιζόμαστε μόνο στα δεδομένα
- ⌘ εκ-των-υστερων αποτελέσματα είναι παρόμοια με της μέγιστης πιθανοφάνειας

5.1. ΚΑΘΟΡΙΣΜΟΣ PRIOR ΚΑΤΑΝΟΜΩΝ

Μη πληροφοριακές εκ-των-προτέρων κατανομές :

- ⌘ δίνουν ίδια πιθανότητα εμφάνισης μίας τιμής ή ενός ισομηκούς διαστήματος (ομοιόμορφη κατανομή).
- ⌘ Η ομοιόμορφη ορίζεται σε ένα διάστημα $[a, \beta]$.
- ⌘ Για μεταβλητές στο $R \Rightarrow U(-1000, 1000)$
- ⌘ Για μεταβλητές στο $[0, +\infty) \Rightarrow U(0, 1000)$
- ⌘ Για μεταβλητές στο $[0, 1] \Rightarrow U(0, 1)$
- ⌘ **ΚΙΝΔΥΝΟΣ:** Να χασουμε τιμές λόγω των ορίων a, β

5.1. ΚΑΘΟΡΙΣΜΟΣ PRIOR ΚΑΤΑΝΟΜΩΝ

Μη πληροφοριακές εκ-των-προτέρων κατανομές:

- ⌘ Εναλλακτικά μπορούμε να χρησιμοποιήσουμε improper prior κατανομές $f(x) \propto 1/c$ (σταθερά)
- ⌘ Στο BUGS προτείνουμε
 - ☒ Για μεταβλητές στο $R \Rightarrow$ Κανονική με μεγάλη διακύμανση (μικρή ακρίβεια)
 - ☒ Για μεταβλητές στο $[0, +\infty) \Rightarrow$ Γάμμα με $\alpha=\beta$ (μικρές τιμές)
[μέση τιμή 1, διακύμανση μεγάλη = $1/\beta$].

5.2. ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ:

- ⌘ Σημαντική για τον υπολογισμό των παραμέτρων
- ⌘ Μια παράμετρος δεν είναι προσδιορίσιμη (identifiable) όταν τα δεδομένα δεν δίνουν αρκετή πληροφορία για τον υπολογισμό της.
- ⌘ Για το λόγο αυτό θέτουμε περιορισμούς:
 - ☒ Περιορισμός Αθροίσματος στο μηδέν (Sum-to-zero constraint)
 - ☒ Γωνιακός Περιορισμός (corner constraint) [treatment effect]

5.2. ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ:

- ⌘ Αν έχουμε το ίδιο μοντέλο με διαφορετική παραμετροποίηση τότε
 - ☒ Η ερμηνεία των παραμέτρων είναι διαφορετική
 - ☒ Οι προσαρμοσμένες ή προβλεπόμενες τιμές είναι ίδιες
 - ☒ Οι παράμετροι του ίδιου μοντέλου μπορούν να μετασχηματιστούν σε παραμέτρους με άλλους περιορισμούς με απλές συναρτήσεις
- ⌘ Η παραμετροποίηση αλλάζει και χρησιμοποιώντας διαφορετικές ψευδομεταβλητές (dummy variables).

5.2. ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

ΠΑΡΑΔΕΙΓΜΑ 1: One-way Anova

Έχουμε μία ποσοτική Y και μία κατηγορική A με k επίπεδα

$$Y_i \sim \text{Normal}(\mu_i, \sigma^2)$$

- ⌘ **Παραμετροποίηση 1:** Εκτιμούμε κατευθείαν τους μέσους ανά ομάδα

5.2. ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

ΠΑΡΑΔΕΙΓΜΑ 1: One-way Anova

Έχουμε μία ποσοτική Y και μία κατηγορική A με k επίπεδα

$$Y_i \sim \text{Normal}(\mu_i, \sigma^2)$$

⌘ **Παραμετροποίηση 2:** $\mu_i = \mu + \alpha_i$

⌘ $\alpha_1 + \alpha_2 + \dots + \alpha_k = 0$ (STZ)

⌘ Ερμηνεία:

⌘ μ = συνολικό μέσο

⌘ α_i = διαφορά της ομάδας i από το συνολικό μέσο (group effect)

5.2. ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

ΠΑΡΑΔΕΙΓΜΑ 1: One-way Anova

Έχουμε μία ποσοτική Y και μία κατηγορική A με k επίπεδα

$$Y_i \sim \text{Normal}(\mu_i, \sigma^2)$$

⌘ **Παραμετροποίηση 3:** $\mu_i = \mu + \alpha_i$

⌘ $\alpha_1 = 0$ (CR)

⌘ Ερμηνεία:

⌘ μ = μέσος μίας ομάδας βάσης (baseline group)
συνήθως η ομάδα ελέγχου (control group)

α_i = διαφορά της ομάδας i από τη ομάδα βάσης

5.2. ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

ΨΕΥΔΟΜΕΤΑΒΛΗΤΕΣ: One-way Anova

⌘ **Παραμετροποίηση 3:** $\mu_i = \mu + \alpha_i$

⌘ $\alpha_1 = 0$ (CR)

⌘ $\mu = \mathbf{X} \beta$, με $\beta^T = (\mu, \alpha_2, \dots, \alpha_k)$

⌘ $k-1$ dummies (D_2, D_3, \dots, D_k)

☒ $D_j = 1$ αν $A=j$

☒ $D_j = 0$ διαφορετικά

5.2. ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

ΨΕΥΔΟΜΕΤΑΒΛΗΤΕΣ: One-way Anova

⌘ **Παραμετροποίηση 2:** $\mu_i = \mu + \alpha_i$

⌘ $\alpha_1 + \alpha_2 + \dots + \alpha_k = 0$ (STZ)

⌘ $\mu = \mathbf{X} \beta$, με $\beta^T = (\mu, \alpha_2, \dots, \alpha_k)$

⌘ $k-1$ dummies (D_2, D_3, \dots, D_k)

☒ $D_j = -1$ αν $A=1$

☒ $D_j = 1$ αν $A=j$

☒ $D_j = 0$ διαφορετικά

5.2. ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

ΨΕΥΔΟΜΕΤΑΒΛΗΤΕΣ: One-way Anova

για $k=3$

j	D_2	D_3	S_2	S_3
1	0	0	-1	-1
2	1	0	1	0
3	0	1	0	1

5.2. ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

ΣΧΕΣΕΙΣ ΠΑΡΑΜΕΤΡΩΝ: One-way Anova

$$\mu_i = \mu^{CR} + \alpha_i^{CR} \Rightarrow$$

$$\mu_i = \mu^{CR} + \bar{a} + \alpha_i^{CR} - \bar{a} \Rightarrow$$

$$\mu_i = \mu^{STZ} + \alpha_i^{STZ} \text{ με}$$

$$\mu^{STZ} = \mu^{CR} + \bar{a}$$

$$\alpha_i^{STZ} = \alpha_i^{CR} - \bar{a}$$

$$\text{Γενικά θέτουμε } \mathbf{X}^{STZ} \boldsymbol{\beta}^{STZ} = \mathbf{X}^{CR} \boldsymbol{\beta}^{CR} \Rightarrow$$

$$\boldsymbol{\beta}^{STZ} = (\mathbf{X}^{STZ})^{-1} \mathbf{X}^{CR} \boldsymbol{\beta}^{CR}$$

$$\boldsymbol{\beta}^{CR} = (\mathbf{X}^{CR})^{-1} \mathbf{X}^{STZ} \boldsymbol{\beta}^{STZ}$$

5.2. ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

ΠΑΡΑΔΕΙΓΜΑ 2: Απλή Παλινδρόμηση

Έχουμε μία ποσοτική Y και μία επεξηγηματική ποσοτική $X \Rightarrow Y_i \sim \text{Normal}(\mu_i, \sigma^2)$

⌘ **Παραμετροποίηση 1:** $\mu_i = \alpha + \beta X_i$

⌘ **Παραμετροποίηση 2:** $\mu_i = \alpha^* + \beta^*(X_i - \mu_x)$

⌘ **Παραμετροποίηση 3:** $\mu_i = \alpha^{**} + \beta^{**}(X_i - \mu_x)/\sigma_x$

5.2. ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

ΠΑΡΑΔΕΙΓΜΑ 2: Απλή Παλινδρόμηση

ΕΡΜΗΝΕΙΑ ΠΑΡΑΜΕΤΡΩΝ

⌘ **Παραμετροποίηση 1:**

⌘ α = Αναμενόμενος μέσος της Y όταν $X=0$

⌘ β = Αναμενόμενη αύξηση της Y όταν X αυξηθεί κατά μία μονάδα

5.2. ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

ΠΑΡΑΔΕΙΓΜΑ 2: Απλή Παλινδρόμηση
ΕΡΜΗΝΕΙΑ ΠΑΡΑΜΕΤΡΩΝ

⌘ **Παραμετροποίηση 2:**

⌘ α^* = Αναμενόμενος μέσος της Y όταν $X = \mu_x$

⌘ β^* = Αναμενόμενη αύξηση της Y όταν X αυξηθεί κατά μία μονάδα

5.2. ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

ΠΑΡΑΔΕΙΓΜΑ 2: Απλή Παλινδρόμηση
ΕΡΜΗΝΕΙΑ ΠΑΡΑΜΕΤΡΩΝ

⌘ **Παραμετροποίηση 3:**

⌘ α^{**} = Αναμενόμενος μέσος του Y όταν $X = \mu_x$

⌘ β^{**} = Αναμενόμενη αύξηση της Y όταν X αυξηθεί κατά ποσότητα ίση με την τυπική απόκλιση του X [είναι ίσος με το συντελεστή συσχέτισης, λέγεται και beta parameter]

5.2. ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

ΠΑΡΑΔΕΙΓΜΑ 2: Απλή Παλινδρόμηση
ΣΧΕΣΕΙΣ ΠΑΡΑΜΕΤΡΩΝ

Παραμετροποίηση 2:

$$\text{⌘} \mu_i = \alpha^* + \beta^*(X_i - \mu_x) = \alpha^* - \beta^* \mu_x + \beta^* X_i$$

$$\text{☒} \alpha = \alpha^* - \beta^* \mu_x$$

$$\text{☒} \beta = \beta^*$$

Παραμετροποίηση 3:

$$\text{⌘} \mu_i = \alpha^{**} + \beta^{**}(X_i - \mu_x)/\sigma_x = \alpha^{**} - \beta^{**} \mu_x/\sigma_x + \beta^{**} X_i/\sigma_x$$

$$\text{☒} \alpha = \alpha^{**} - \beta^{**} \mu_x/\sigma_x$$

$$\text{☒} \beta = \beta^{**}/\sigma_x$$

5.2. ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

Διαφορετική προσδιορίσιμη παραμετροποίηση **δεν επιδρά** στις posterior

Διαφορετική προσδιορίσιμη παραμετροποίηση **επιδρά** στη σύγκλιση του MCMC [πρέπει οι παράμετροι να είναι όσο το δυνατόν λιγότερο συσχετισμένοι]

Για τις μη-προσδιορίσιμες παράμετρους ενός μοντέλου

1... MCMC -> δε συγκλίνει

2... Η posterior θα είναι improper

3... Συναρτήσεις αυτών μπορεί να συγκλίνουν!?

5.3. RANDOM EFFECTS

Τι είναι τα random effects

- ⌘ Τυχαίο σφάλμα στον γραμμικό προσδιορισμό (η)
- ⌘ Οι παράμετροι (στην κλασσική στατιστική) δεν είναι σταθερές αλλά τυχαίες μεταβλητές

5.3. RANDOM EFFECTS

Πότε χρησιμοποιούμε random effects

- ⌘ Επαναλαμβανόμενες μετρήσεις στα ίδια άτομα την ίδια χρονική περίοδο (διασπορά λόγω σφάλματος μέτρησης ή τυχαιότητας του ίδιου παράγοντα)
- ⌘ Επαναλαμβανόμενες μετρήσεις σε διαφορετικές χρονικές περιόδους
- ⌘ Χρήση πιο σύνθετων κατανομών (Mixtures)
- ⌘ Διόρθωση Υπερ-διασποράς (overdispersion)
- ⌘ Εκτίμηση της within-subject variability

Bayesian Biostatistics Using BUGS



5.4. ΠΑΡΑΔΕΙΓΜΑΤΑ ΣΤΟ BUGS



Department of Statistics,
Athens University of
Economics & Business

5.4.1. Normal Hierarchical Model with Missing Data (Rats)

BUGS Examples Vol.1, σελ.4, Example 1

- ⌘ Section 6 of Gelfand *et al* (1990, JASA),
- ⌘ 30 young rats
- ⌘ weights for 5 weeks.

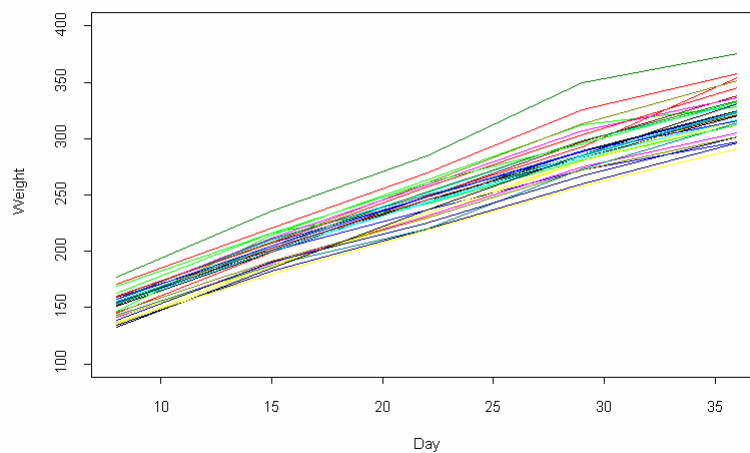
5.4.1. Normal Hierarchical Model with Missing Data (Rats)

Part of the data:

	Weights Y_{ij} of rat i on day x_j				
	$x_j = 8$	15	22	29	36
Rat 1	151	199	246	283	320
Rat 2	145	199	249	293	354
...
Rat 30	153	200	244	286	324

Y_{ij} is the weight of the i th rat measured at age x_j .

5.4.1. Normal Hierarchical Model with Missing Data (Rats)



5.4.1. Normal Hierarchical Model with Missing Data (Rats)

The model is essentially a random effects linear growth curve

$$Y_{ij} \sim \text{Normal}(\alpha_i + \beta_i(x_i - \text{xbar}), \tau_c)$$

$$\alpha_i \sim \text{Normal}(\alpha_c, \tau_\alpha)$$

$$\beta_i \sim \text{Normal}(\beta_c, \tau_\beta)$$

where $\text{xbar} = 22$, and τ represents the *precision* (1/variance) of a normal distribution.

see the `Birats` example in Volume 2 which models the covariance between α_i and β_i .

standardise the x_i 's around their mean to reduce dependence (for the full balanced data, complete independence is achieved)

5.4.1. Normal Hierarchical Model with Missing Data (Rats)

a_c, t_a, b_c, t_b, t_c are given independent "noninformative" priors.

Interest particularly focuses on the intercept at zero time (birth), denoted $a_0 = a_c - b_c \text{xbar}$.

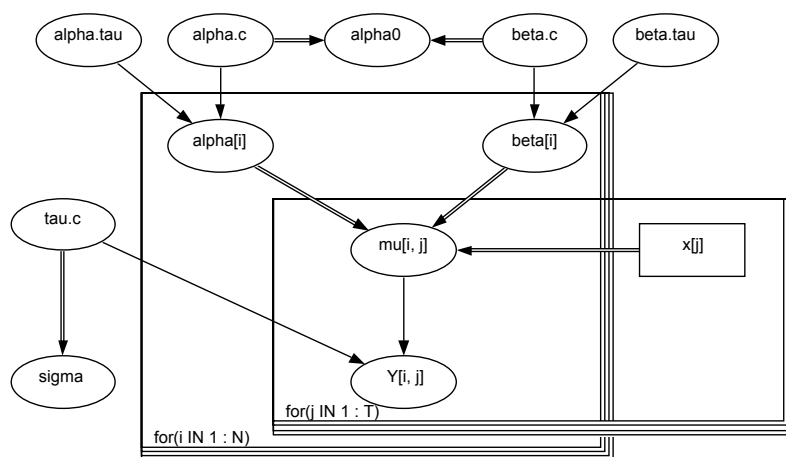
RANDOM EFFECTS => two sources of variability

1... within each subject (rat)

2... across subjects

5.4.1. Normal Hierarchical Model with Missing Data (Rats)

Graphical model for rats example:

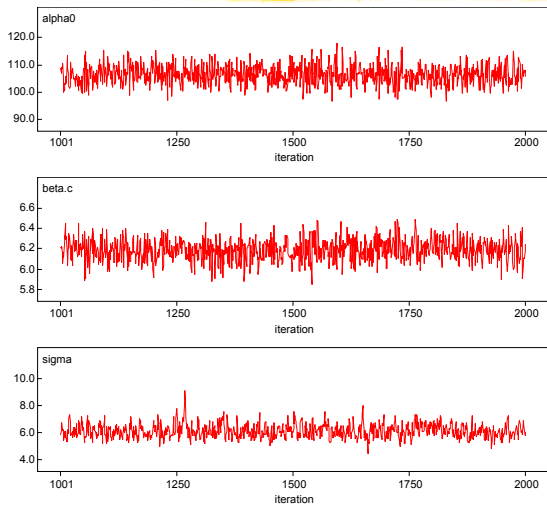


5.4.1. Normal Hierarchical Model with Missing Data (Rats)

BUGS language for rats example:

```
model
{
  for(i in 1 : N){
    for(j in 1 : T){
      Y[i, j] ~ dnorm(mu[i, j], tau.c)
      mu[i, j] <- alpha[i] + beta[i] * (x[j] - xbar)
    }
    alpha[i] ~ dnorm(alpha.c, alpha.tau)
    beta[i] ~ dnorm(beta.c, beta.tau)
  }
  tau.c ~ dgamma(0.001, 0.001)
  sigma <- 1 / sqrt(tau.c)
  alpha.c ~ dnorm(0.0, 1.0E-6)
  alpha.tau ~ dgamma(0.001, 0.001)
  beta.c ~ dnorm(0.0, 1.0E-6)
  beta.tau ~ dgamma(0.001, 0.001)
  alpha0 <- alpha.c - xbar * beta.c
}
```

5.4.1. Normal Hierarchical Model with Missing Data (Rats)



5.4.1. Normal Hierarchical Model with Missing Data (Rats)

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
alpha0	106.4	3.518	0.1071	99.87	106.4	113.7	1001	1000
beta.c	6.187	0.1052	0.003881	5.978	6.189	6.396	1001	1000
sigma	6.123	0.5053	0.02685	5.288	6.065	7.231	1001	1000

5.4.1. Normal Hierarchical Model with Missing Data (Rats)

MISSING VALUES ANALYSIS(ratmiss)

Missing Value -> NA

Rat 26 -> 153 NA NA NA NA

Monitor (Y[26,]) -> predicted values of Missing

node	mean	sd	MC error	2.5%	median	97.5%
Y[26,2]	205.3	8.711	0.3794	187.6	205.2	222.3
Y[26,3]	251.5	10.39	0.5242	230.7	251.4	272.7
Y[26,4]	297.1	12.54	0.6756	271.1	297.4	321.0
Y[26,5]	343.3	15.55	0.8251	313.7	343.6	373.6

5.4.1. Normal Hierarchical Model with Missing Data (Rats)

FULL DATA ANALYSIS

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
alpha0	106.4	3.518	0.1071	99.87	106.4	113.7	1001	1000
beta.c	6.187	0.1052	0.003881	5.978	6.189	6.396	1001	1000
sigma	6.123	0.5053	0.02685	5.288	6.065	7.231	1001	1000

MISSING DATA ANALYSIS

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
alpha0	100.9	3.75	0.1106	93.63	101.0	108.3	1001	2000
beta.c	6.581	0.1449	0.00655	6.298	6.583	6.872	1001	2000
sigma	6.129	0.7062	0.03879	4.93	6.065	7.728	1001	2000

5.4.2. Random Effects Logistic Regression (Seeds)

BUGS Examples Vol.1, σελ.10, Example 3

- ⌘ Table 3 of Crowder (1978, *Ap.Stat.*), also analysed by Breslow & Cleyton (1993, *Ap.Stat.*)
- ⌘ % of germinated seeds on 21 plates
- ⌘ 2 by 2 factorial layout by seed and type of root extract
- ⌘ r_i = # germinated seeds in i plate
- ⌘ n_i = total # seeds in i plate
- ⌘ $i=1, \dots, N=21$ plates

5.4.2. Random Effects Logistic Regression (Seeds)

seed <i>O. aegyptiaco</i> 75						seed <i>O. aegyptiaco</i> 73					
Bean			Cucumber			Bean			Cucumber		
r	n	r/n	r	n	r/n	r	n	r/n	r	n	r/n
10	39	0.26	5	6	0.83	8	16	0.50	3	12	0.25
23	62	0.37	53	74	0.72	10	30	0.33	22	41	0.54
23	81	0.28	55	72	0.76	8	28	0.29	15	30	0.50
26	51	0.51	32	51	0.63	23	45	0.51	32	51	0.63
17	39	0.44	46	79	0.58	0	4	0.00	3	7	0.43
			10	13	0.77						

5.4.2. Random Effects Logistic Regression (Seeds)

DATA IN BUGS

r[]	n[]	x1[]	x2[]
10	39	0	0

...

5	6	0	1
---	---	---	---

...

8	16	1	0
---	----	---	---

...

3	12	1	1
---	----	---	---

x1 = dummy for seed type

x2 = dummy for root extract

5.4.2. Random Effects Logistic Regression (Seeds)

MODEL

1.. RANDOM COMPONENT:

⌘ $r_i \sim \text{Binomial}(p_i, n_i)$

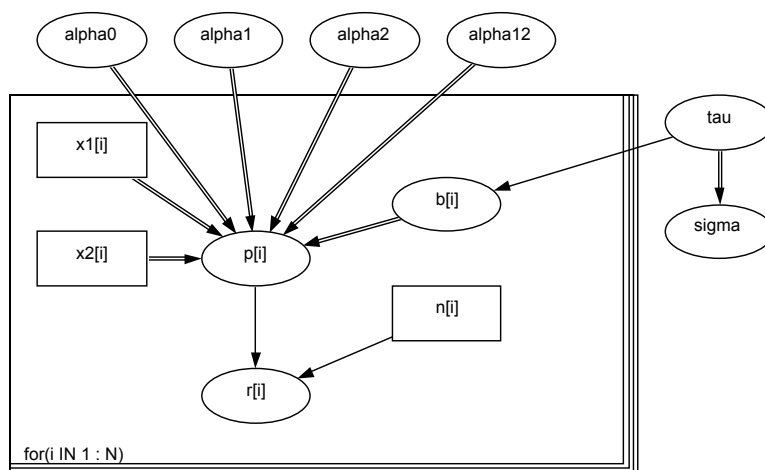
2.. Systematic Component & link:

⌘ $\text{logit}(p_i) = \alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2i} + \alpha_{12} x_{1i} x_{2i} + b_i$

3.. Random Effect:

⌘ $b_i \sim \text{Normal}(0, \tau)$

5.4.2. Random Effects Logistic Regression (Seeds)



5.4.2. Random Effects Logistic Regression (Seeds)

```

model
{
  # Model Likelihood
  for( i in 1 : N ) {
    r[i] ~ dbin(p[i],n[i]) # random component
    b[i] ~ dnorm(0.0,tau) # random effect
    #systematic component
    logit(p[i]) <- alpha0 + alpha1 * x1[i] +
    alpha2 * x2[i] + alpha12 * x1[i] * x2[i] +
    b[i]
  }
}

```

5.4.2. Random Effects Logistic Regression (Seeds)

...

```
# Prior distributions
alpha0 ~ dnorm(0.0,1.0E-6)
alpha1 ~ dnorm(0.0,1.0E-6)
alpha2 ~ dnorm(0.0,1.0E-6)
alpha12 ~ dnorm(0.0,1.0E-6)
tau ~ dgamma(0.001,0.001)
sigma <- 1 / sqrt(tau)
}
```

5.4.2. Random Effects Logistic Regression (Seeds)

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
alpha0	-0.5546	0.1941	0.007696	-0.9353	-0.5577	-0.1597	1001	10000
alpha1	0.08497	0.3127	0.01283	-0.5814	0.09742	0.6679	1001	10000
alpha12	-0.8229	0.4321	0.01785	-1.697	-0.8218	0.01641	1001	10000
alpha2	1.356	0.2743	0.01236	0.8257	1.347	1.909	1001	10000
sigma	0.2731	0.1437	0.007956	0.04133	0.2654	0.5862	1001	10000

We may compare simple logistic, maximum likelihood (from EGRET), penalized quasi-likelihood (PQL) Breslow and Clayton (1993) with the *BUGS* results

variable	Logistic regression		maximum likelihood		PQL	
	β	SE	β	SE	β	SE
α_0	-0.558	0.126	-0.546	0.167	-0.542	0.190
α_1	0.146	0.223	0.097	0.278	0.77	0.308
α_2	1.318	0.177	1.337	0.237	1.339	0.270
α_{12}	-0.778	0.306	-0.811	0.385	-0.825	0.430
σ	---	---	0.236	0.110	0.313	0.121

5.4.2. Random Effects Logistic Regression (Seeds)

CONSTRAINING RANDOM EFFECTS

- 1... Sum-to-zero constraints
- 2... Hierarchical Centering

ALTERNATIVE PARAMETRAZATION FOR Sigma

5.4.2. Random Effects Logistic Regression (Seeds)

CONSTRAINING RANDOM EFFECTS

- 1... Sum-to-zero constraints

Why:

- (a) Unconstrained b_i may not converge
- (b) Estimate fixed effects for each subject

Use $c[i]$:

```
c[i] ~ dnorm(0, tau)
b[i] <- c[i] - mean(c[])
```

5.4.2. Random Effects Logistic Regression (Seeds)

CONSTRAINING RANDOM EFFECTS

2... Hierarchical Centering

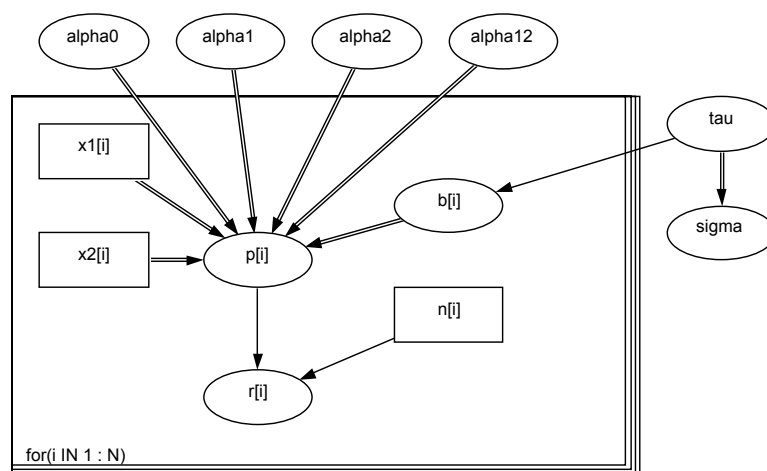
$$\mu_i = \alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2i} + \alpha_{12} x_{1i} x_{2i}$$

$$\beta_i = \mu_i + b_i$$

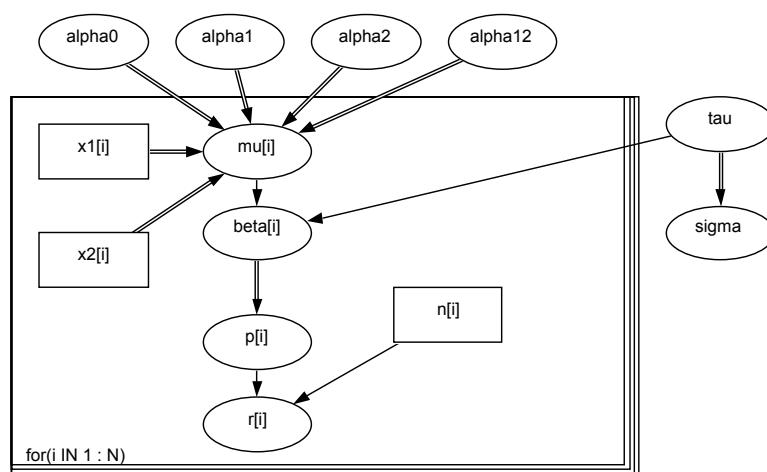
$$\text{logit}(p_i) = \beta_i$$

$$\beta_i \sim \text{Normal}(\mu_i, \tau)$$

5.4.2. Random Effects Logistic Regression (Seeds)



5.4.2. Random Effects Logistic Regression (Seeds)



5.4.2. Random Effects Logistic Regression (Seeds)

CONSTRAINING RANDOM EFFECTS

2... Hierarchical Centering

advantages:

- (a) Gibbs sampler with better correlation properties
- (b) Conditionals for the α parameters are conjugate \Rightarrow reduced generation time

5.4.2. Random Effects Logistic Regression (Seeds)

ALTERNATIVE PARAMETRAZATION FOR Sigma

```
b[i] ~ dnorm(0.0,1.0)  # random effect
#systematic component
logit(p[i]) <- alpha0 + alpha1 * x1[i] +
alpha2 * x2[i] + alpha12 * x1[i] * x2[i] +
sigma*b[i]
# exponential prior on sigma
sigma~dexp(1.0)
```

5.4.3. Institutional Ranking (Surgical)

BUGS Examples Vol.1, σελ.15, Example 4

Mortality rates in 12 hospitals performing cardiac surgery in babies.

5.4.3. Institutional Ranking (Surgical)

Hospital	No of ops	No of deaths
A	47	0
B	148	18
C	119	8
D	810	46
E	211	8
F	196	13
G	148	9
H	215	31
I	207	14
J	97	8
K	256	29
L	360	24

5.4.3. Institutional Ranking (Surgical)

MODEL

$r_i \sim \text{Binomial}(p_i, n_i)$

(1) **Fixed Effects Model:** $p_i \sim \text{Beta}(1.0, 1.0)$

[Uniform Prior on each p_i]

(2) **Random Effects Model:**

$\text{logit}(p_i) = b_i$

$b_i \sim \text{Normal}(\mu, \tau)$

$\mu \sim \text{Normal}(0.0, 10^{-6})$ [Prior]

$\tau \sim \text{Gamma}(10^{-6}, 10^{-6})$ [Prior]

5.4.3. Institutional Ranking (Surgical)

BUGS CODE: FIXED EFFECTS MODEL

```
model{
  for( i in 1 : N ) {

    r[i] ~ dbin(p[i], n[i])

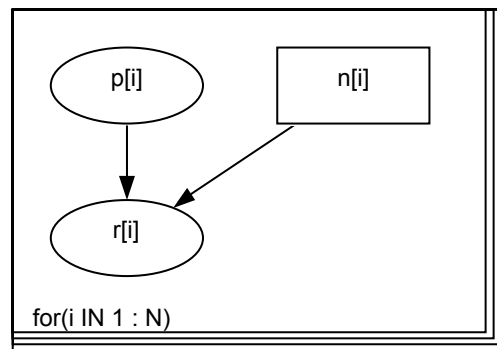
    p[i] ~ dbeta(1.0, 1.0)
  }
}
```

5.4.3. Institutional Ranking (Surgical)

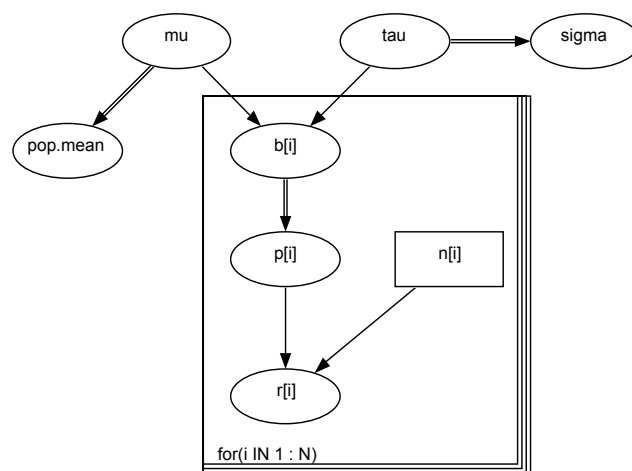
BUGS CODE: RANDOM EFFECTS MODEL

```
model{
  for( i in 1 : N ) {
    b[i] ~ dnorm(mu, tau)
    r[i] ~ dbin(p[i], n[i])
    logit(p[i]) <- b[i]
  }
  pop.mean <- exp(mu) / (1 + exp(mu))
  mu ~ dnorm(0.0, 1.0E-6)
  sigma <- 1 / sqrt(tau)
  tau ~ dgamma(0.001, 0.001)
}
```

5.4.3. Institutional Ranking (Surgical)



5.4.3. Institutional Ranking (Surgical)



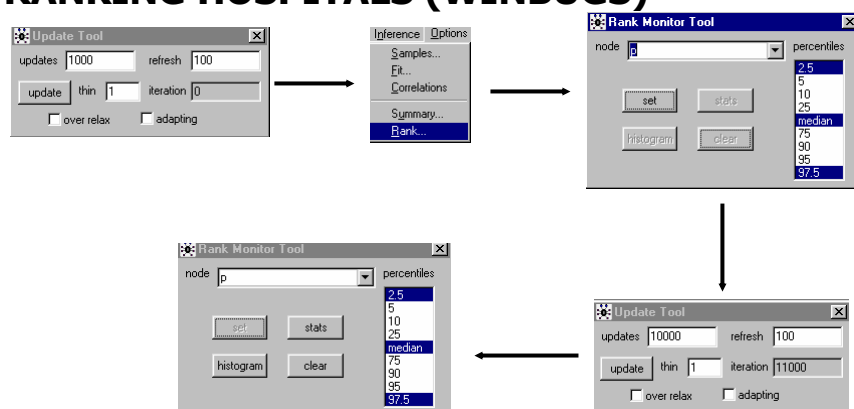
5.4.3. Institutional Ranking (Surgical)

RANKING HOSPITALS (WINBUGS)

- 1... Update 1000 [burnin]
- 2... Select ***Inference>Rank***
- 3... Set the rank monitor for p
- 4... Update 10,000 iterations
- 5... Select "histogram" or "Stats" option.

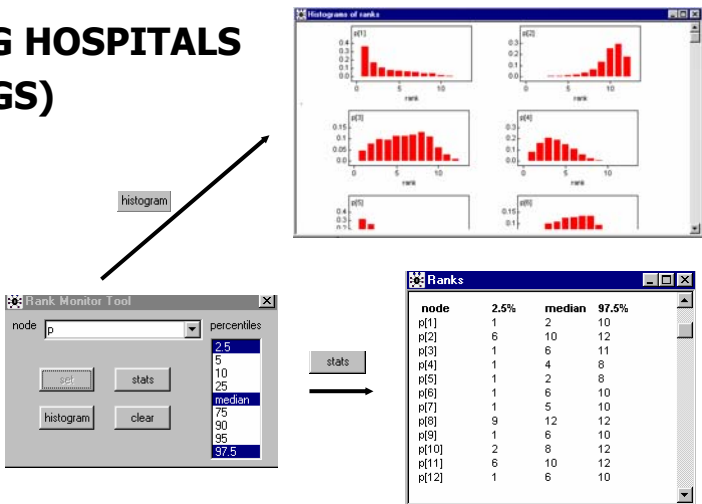
5.4.3. Institutional Ranking (Surgical)

RANKING HOSPITALS (WINBUGS)

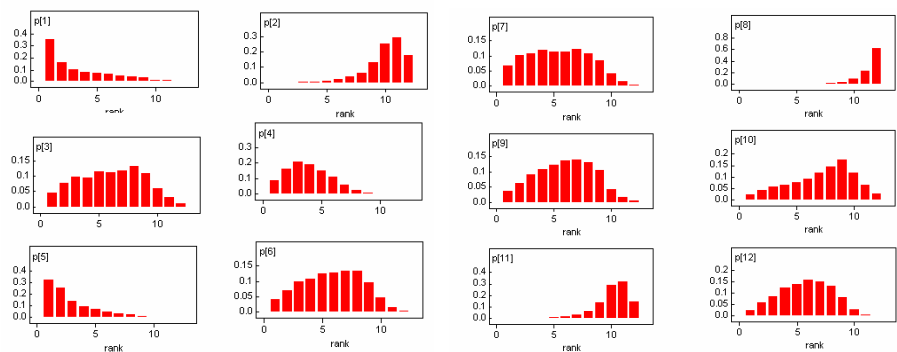


5.4.3. Institutional Ranking (Surgical)

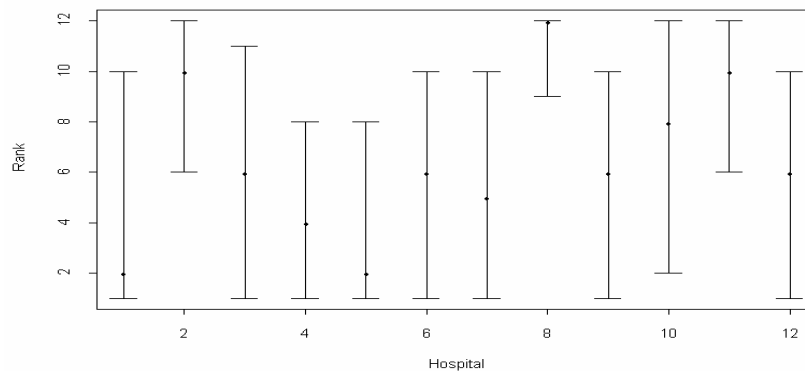
RANKING HOSPITALS (WINBUGS)



5.4.3. Institutional Ranking (Surgical)



5.4.3. Institutional Ranking (Surgical)



5.4.3. Institutional Ranking (Surgical)

RESULTS

Considerable uncertainty associated with 'league tables':

- (a) only 2 hospitals (H and K) exclude the median rank
- (b) none of intervals fall completely within the lower or upper quartiles

5.4.3. Institutional Ranking (Surgical)

RANKING HOSPITALS (Classic BUGS)

Compute Ranks in Bugs:

```
for (i in 1:N) {  
  for (j in 1:N) {  
    not.less.than[i,j]<-step(p[i]-p[j]);  
  }  
  rank[i]<-sum(not.less.than[i,]);  
}  
step(x)=1 if x≥0 and step(x)=0 if x<0  
monitor(rank) # monitor ranks in BUGS  
stats(rank)   # view stats of ranks in BUGS
```

5.4.4. Cross Over Trial (equiv)

BUGS Examples Vol.1, σελ. 21, Example 6

Data from a two-treatment, two-period crossover trial to compare 2 tablets A and B, as reported by Gelfand *et al* (1990, JASA).

5.4.4. Cross Over Trial (equiv)

Data

Subject i	Sequence	seq	Period 1	T_{i1}	Period 2	T_{i2}
1	AB	1	1.40	1	1.65	2
2	AB	1	1.64	1	1.57	2
3	BA	-1	1.44	2	1.58	1
....						
8	AB	1	1.25	1	1.44	2
9	BA	-1	1.25	2	1.39	1
10	BA	-1	1.30	2	1.52	1

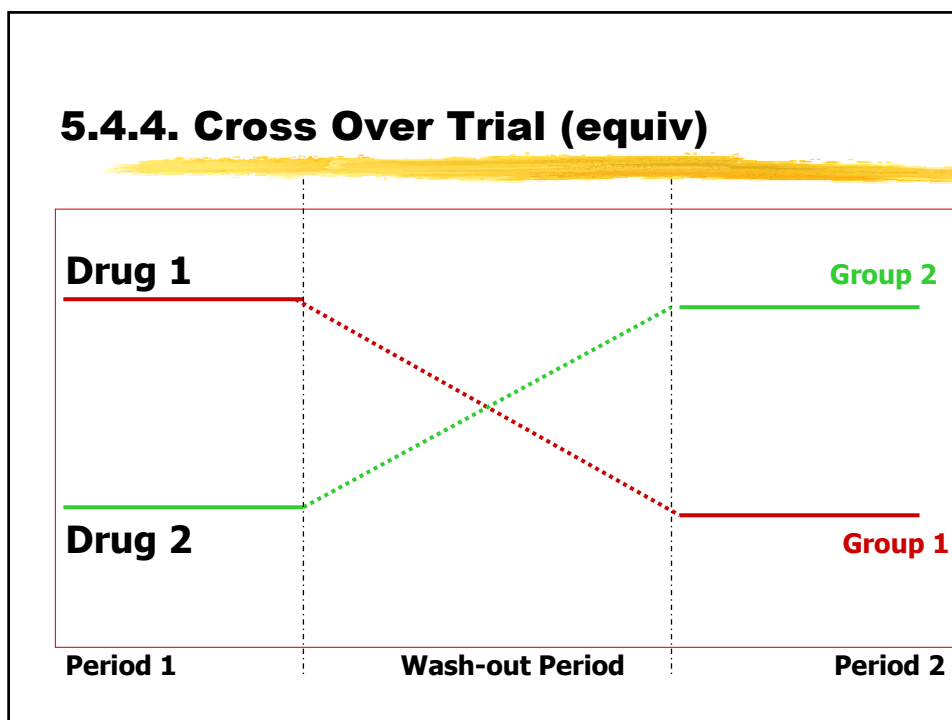
5.4.4. Cross Over Trial (equiv)

Definitions

Two-treatment, Two-period Cross-over trial

- ⌘ Type of randomized clinical trial
- ⌘ Each subject is randomly allocated in two groups
- ⌘ Group 1: receive drug A in period 1 and drug B in period 2
- ⌘ Group 2: receive drug B in period 1 and drug A in period 2
- ⌘ *Wash-out Period*: period of no-treatment between two active drug periods

5.4.4. Cross Over Trial (equiv)



5.4.4. Cross Over Trial (equiv)

Definitions

Treatment or Drug Effect: The effect of different treatment or medication

Carry Over Effect: residual biological effect of a 1st period study treatment
Wash-out period often does not eliminate carry-over effects

5.4.4. Cross Over Trial (equiv)

Classical Inference for Treatment Effects

Y_{ik} measurement of i subject, k period

Find the differences [Drug 1 -Drug 2]

$$\Delta_{i1} = D_{i1} = Y_{i1} - Y_{i2} \text{ for group 1 } [T_1 - T_2 + P_1 - P_2]$$

$$\Delta_{i2} = -D_{i2} = Y_{i2} - Y_{i1} \text{ for group 2 } [T_1 - T_2 + P_2 - P_1]$$

Estimate overall effect by

$$\bar{\Delta} = (\bar{\Delta}_1 + \bar{\Delta}_2) / 2 = (\bar{D}_1 - \bar{D}_2) / 2$$

Test H_0 : Drug-effect $[E(\Delta)] = 0$ using one-sample t-test

5.4.4. Cross Over Trial (equiv)

Classical Inference for Carry-over Effects

Previous analysis assumed no carry-over effects

Find the sum [Drug 1 +Drug 2]: $S_{ij} = Y_{i1} + Y_{i2}$

Estimate overall carry-over effect by $\bar{S} = \bar{S}_1 - \bar{S}_2$

Test H_0 : Carry-over effect = 0 using two-sample t-test

5.4.4. Cross Over Trial (equiv)

MODEL BUILDING

Find the differences [Drug 1 -Drug 2]

$$Y_{i1} - Y_{i2} \sim \text{Normal}(\mu_1, \tau^2) \text{ via Group 1 [AB]}$$

$$Y_{i1} - Y_{i2} \sim \text{Normal}(\mu_2, \tau^2) \text{ via Group 2 [BA]}$$

$$\mu_1 = \pi_1 - \pi_2 + d_1 - d_2 \text{ and } \mu_2 = \pi_1 - \pi_2 - d_1 + d_2$$

$$\text{Drug effect } \phi = (\mu_2 - \mu_1)/2 = (d_2 - d_1)/2$$

$$\text{Period Effect } \pi = (\mu_1 + \mu_2)/2 = (\pi_1 - \pi_2)/2$$

[if $\pi_1 = \pi_2$ no period/carry over effect]

$$\text{Reparametrize } \mu_1 = (\pi - \phi), \mu_2 = (\pi + \phi)/2$$

$$\text{therefore } D_i = Y_{i1} - Y_{i2} \sim \text{Normal}((\pi + \text{seq } \phi)/2, \tau^2)$$

seq=-1 if group AB, 1 if group BA

5.4.4. Cross Over Trial (equiv)

MODEL BUILDING

$$Y_{i1} + Y_{i2} \sim \text{Normal}(m_1, \tau^{*2}) \text{ via Group 1 [AB]}$$

$$Y_{i1} + Y_{i2} \sim \text{Normal}(m_2, \tau^{*2}) \text{ via Group 2 [BA]}$$

$$\text{Carry Over effect cof} = m_1 - m_2$$

If cof=0 ($m_1 = m_2$) no carry-over effect

$$\text{Reparametrize } m_1 = \mu - \text{cof}, m_2 = \mu + \text{cof}$$

5.4.4. Cross Over Trial (equiv)

Random Effects model (No carry-over effect)

$$\begin{aligned}
 Y_{ik} &\sim \text{Normal}(m_{ik}, t1) \\
 m_{ik} &= m + (-1)^{Tik-1} \varphi/2 + (-1)^{k-1} \pi/2 + d_i \\
 d_i &\sim \text{Normal}(0, t2)
 \end{aligned}$$

5.4.4. Cross Over Trial (equiv)

```

Random Effects model (BUGS CODE)
model{
  for( k in 1 : P ) {
    for( i in 1 : N ) {
      Y[i,k]~dnorm(m[i,k], tau1)
      m[i,k]<-mu+sign[T[i,k]]*phi/2+
      sign[k]*pi/2 +delta[i]
      T[i,k]<-group[i]*(k-1.5)+1.5}}
  for( i in 1 : N ) {delta[i]~dnorm(0.0, tau2)}

```


5.4.4. Cross Over Trial (equiv)

Random Effects model (BUGS CODE)

```

tau1 ~ dgamma(0.001, 0.001)
sigma1 <- 1 / sqrt(tau1)
tau2 ~ dgamma(0.001, 0.001)
sigma2 <- 1 / sqrt(tau2)
mu ~ dnorm(0.0, 1.0E-6)
phi ~ dnorm(0.0, 1.0E-6)
pi ~ dnorm(0.0, 1.0E-6)
theta <- exp(phi)
equiv <- step(theta - 0.8) - step(theta - 1.2)
}

```

5.4.4. Cross Over Trial (equiv)

	Period 1	Period 2
AB	$\mu + \phi/2 + \pi/2$	$\mu - \phi/2 - \pi/2$
BA	$\mu - \phi/2 + \pi/2$	$\mu + \phi/2 - \pi/2$
$\text{cof} = [(\mu + \phi/2 + \pi/2) + (\mu - \phi/2 - \pi/2)]$ $- [(\mu - \phi/2 + \pi/2) + (\mu + \phi/2 - \pi/2)] = 0$		

5.4.4. Cross Over Trial (equiv)

Random Effects model (carry-over effect)

$$Y_{ik} \sim \text{Normal}(m_{ik}, t1)$$

$$m_{ik} = \mu + (-1)^{Tik-1} \phi/2 + (-1)^{k-1} \pi/2 + d_i \\ + (k-1) (-1)^{Tik} \text{cof}/2$$

$$d_i \sim \text{Normal}(0, t2)$$

5.4.4. Cross Over Trial (equiv)

	Period 1	Period 2
AB	$\mu + \phi/2 + \pi/2$	$\mu - \phi/2 - \pi/2 + \text{cof}/2$
BA	$\mu - \phi/2 + \pi/2$	$\mu + \phi/2 - \pi/2 - \text{cof}/2$

$$\text{cof} = [(\mu + \phi/2 + \pi/2) + (\mu - \phi/2 - \pi/2 + \text{cof}/2)] \\ - [(\mu - \phi/2 + \pi/2) + (\mu + \phi/2 - \pi/2 - \text{cof}/2)]$$

5.4.5. Variance Components Models (Dyes)

BUGS Examples Vol.1, σελ. 24, Example 7

Box and Tiao (1973) analyse data first presented by Davies (1967) concerning batch to batch variation in yields of dyestuff.

The data (shown below) arise from a balanced experiment whereby the total product yield was determined for 5 samples from each of 6 randomly chosen batches of raw material.

5.4.5. Variance Components Models (Dyes)

Batch	Yield (in grams)				
1	1545	1440	1440	1520	1580
2	1540	1555	1490	1560	1495
3	1595	1550	1605	1510	1560
4	1445	1440	1595	1465	1545
5	1595	1630	1515	1635	1625
6	1520	1455	1450	1480	1445

5.4.5. Variance Components Models (Dyes)

Object of the study:

determine the relative importance of ***between batch variation*** versus ***variation due to sampling and analytic errors***.

Assumption:

batches and samples vary independently, and contribute additively to the total error variance

5.4.5. Variance Components Models (Dyes)

MODEL

$$y_{ij} \sim \text{Normal}(m_i, t_{\text{within}})$$

$$m_i \sim \text{Normal}(q, t_{\text{between}})$$

y_{ij} = yield for sample j of batch i ,

m_i = true yield for batch i ,

t_{within} is the inverse of the within-batch variance

s_{within}^2 variation due to sampling and analytic error

q true average yield for all batches

t_{between} inverse of the between-batch variance

s_{between}^2 .

5.4.5. Variance Components Models (Dyes)

MODEL

Total variation $\sigma_{\text{total}}^2 = \sigma_{\text{within}}^2 + \sigma_{\text{between}}^2$

Relative contributions of each component to the total variance are

$$f_{\text{within}} = \sigma_{\text{within}}^2 / \sigma_{\text{total}}^2 \text{ and}$$

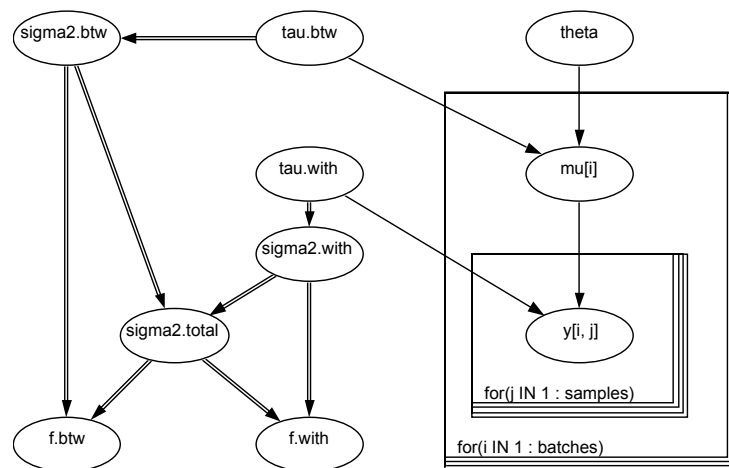
$$f_{\text{between}} = \sigma_{\text{between}}^2 / \sigma_{\text{total}}^2.$$

5.4.5. Variance Components Models (Dyes)

BUGS CODE

```
model{
  for( i in 1 : batches ) {
    for( j in 1 : samples ) {
      y[i , j] ~ dnorm(m[i], tau.with) }
    m[i] ~ dnorm(theta, tau.btw) }
  sigma2.with <- 1 / tau.with
  sigma2.btw <- 1 / tau.btw
  sigma2.total<- sigma2.btw+ sigma2.with
  f.with<- sigma2.with/ sigma2.total
  f.btw<- sigma2.btw/ sigma2.total
  tau.with ~ dgamma(0.001, 0.001)
  tau.btw ~ dgamma(0.001, 0.001)
  theta ~ dnorm(0.0, 1.0E-10)
}
```

5.4.5. Variance Components Models (Dyes)



5.4.5. Variance Components Models (Dyes)

node	mean	sd	MC error	2.5%	median	97.5%
f.btw	0.3311	0.2463	0.006816	1.503E-6	0.3261	0.8137
f.with	0.6689	0.2463	0.006816	0.1863	0.6739	1.0
sigma2.btw	2194.0	3832.0	49.2	0.006255	1277.0	10230.0
sigma2.with	3059.0	1138.0	25.65	1562.0	2815.0	5880.0
theta	1527.0	21.84	0.2532	1485.0	1527.0	1571.0

5.4.6. Random Effects Meta-analysis of Clinical Trials(Blocker)

BUGS Examples Vol.1, σελ. 33, Example 10

Carlin (1992, Stats in Medicine)

22 trials of beta-blockers to prevent mortality after myocardial infraction

Study	Mortality: deaths / total	
	<i>Treated</i>	<i>Control</i>
1	3/38	3/39
2	7/114	14/116
.....		
22	22/680	39/674

5.4.6. Random Effects Meta-analysis of Clinical Trials(Blocker)

Random effects meta-analysis:

Assumption:

the true effect (on a log-odds scale) δ_i in a trial i is drawn from some population distribution.

r_i^C number of events in the control group in trial i ,

r_i^T number of events under active treatment in trial i .

MODEL

$$r_i^C \sim \text{Binomial}(p_i^C, n_i^C)$$

$$r_i^T \sim \text{Binomial}(p_i^T, n_i^T)$$

$$\text{logit}(p_i^C) = \mu_i$$

$$\text{logit}(p_i^T) = \mu_i + \delta_i$$

$$\delta_i \sim \text{Normal}(d, \tau)$$

5.4.6. Random Effects Meta-analysis of Clinical Trials(Blocker)

BUGS CODE

```
model{
  for( i in 1 : Num ) {
    rc[i] ~ dbin(pc[i], nc[i])
    rt[i] ~ dbin(pt[i], nt[i])
    logit(pc[i]) <- mu[i]
    logit(pt[i]) <- mu[i] + delta[i]
    mu[i] ~ dnorm(0.0,1.0E-5)
    delta[i] ~ dnorm(d, tau)    }
  d ~ dnorm(0.0,1.0E-6)
  tau ~ dgamma(0.001,0.001)
  delta.new ~ dnorm(d, tau)
  sigma <- 1 / sqrt(tau) }
```

5.4.7. Weibull Regression in Censored Survival Analysis

BUGS Examples Vol.1, σελ. 54, Example 16

Dellaportas and Smith (1993, Ap.Stats) analyse data from Grieve (1987, Statistician)

Photocarcinogenicity in four groups, each containing 20 mice.

Recorded:

survival time

whether they died or were censored at that time.

A portion of the data, giving survival times in weeks, are shown below. A * indicates censoring.

5.4.7. Weibull Regression in Censored Survival Analysis

DATA				
Mouse	Irradiated	Vehicle	Test	Positive
	control	control	substance	control
1	12	32	22	27
.....				
18	*40	30	24	12
19	31	37	37	17
20	36	27	29	26

* indicates censoring.

5.4.7. Weibull Regression in Censored Survival Analysis

The survival distribution is assumed to be Weibull.

$$t_i \sim \text{Weibull}(\tau, \mu_i)$$

$$\mu_i = \exp(\beta \mathbf{z}_i)$$

$$f(t_i, \mathbf{z}_i) = r \mu_i t_i^{r-1} \exp(-\mu_i \exp(t_i^r))$$

t_i failure time of an individual with covariate vector \mathbf{z}_i and β is a vector of unknown regression coefficients.

Baseline hazard function: $\lambda_0(t_i) = r t_i^{r-1}$

For censored observations, the survival distribution is a truncated Weibull, with lower bound corresponding to the censoring time.

5.4.7. Weibull Regression in Censored Survival Analysis

Prior distributions

$f(\beta) = \text{Normal}(0, \tau = 0.0001)$

$f(r) = \text{Gamma}(1, 0.0001)$

Median survival for individuals with covariate vector \mathbf{z}_i

$$: m_i = (\exp(-\beta \mathbf{z}_i) \log 2)^{1/r}$$

5.4.7. Weibull Regression in Censored Survival Analysis

BUGS CODE

```
model{
  for(i in 1 : M) {    for(j in 1 : N) {
    t[i, j] ~ dweib(r, mu[i]) I(t.cen[i, j],)
  }
  mu[i] <- exp(beta[i])
  beta[i] ~ dnorm(0.0, 0.001)
  median[i] <- pow(log(2) * exp(-beta[i]), 1/r)
}
r ~ dexp(0.001)
veh.control <- beta[2] - beta[1]
test.sub <- beta[3] - beta[1]
pos.control <- beta[4] - beta[1] }
```

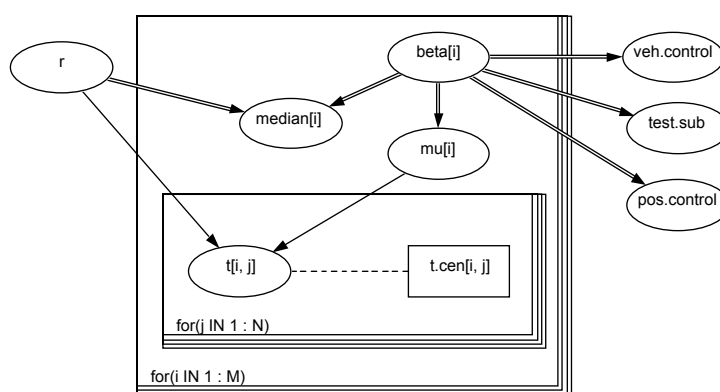
5.4.7. Weibull Regression in Censored Survival Analysis

```

y[i, j] is observed time
ind[i, j] is censoring indicator (0=survival,
1=censored)
if i, j observation is survival time then
  t[i, j]=y[i, j]
  t.cen[i, j]=0
  dweib(r, mu[i]) I(0,)= dweib(r, mu[i])
if i, j observation is censored observation
then
  t[i, j]=NA
  t.cen[i, j]=censored time = ind[i, j]*y[i, j]
  dweib(r, mu[i]) I(t.cen[i, j],)

```

5.4.7. Weibull Regression in Censored Survival Analysis



5.4.7. Weibull Regression in Censored Survival Analysis

Other Survival Examples:

BUGS Examples Vol.1, σελ. 57, Example 17
(Weibull Regression with Random Effects, kidney)

BUGS Examples Vol.1, σελ. 60, Example 18
(Survival Analysis Using Cox Regression, Leuk)

5.4.8. Multinomial Logistic Regression (Alli/Alligator)

BUGS Examples Vol.2, σελ. 51, Example 14

Agresti(1990, *Categorical Data Analysis*)

⌘ 4x2x5 Contingency Table

⌘ 221 alligators cross classified by

⌘ food type (fish, invertebrate, reptile, bird, other)
[Y]

⌘ length of alligator (≤ 2.3 and > 2.3)

⌘ lake (Hancock, Oklawaha, Trafford, George)

5.4.8. Multinomial Logistic Regression (Alli/Alligator)

Lake	Size	Primary Food Choice				
		Fish	Inv. Reptile	Bird	Other	
Hancock	≤ 2.3	23	4	2	2	8
	> 2.3	7	0	1	3	5
Oklawaha	≤ 2.3	5	11	1	0	3
	> 2.3	13	8	6	1	0
Trafford	≤ 2.3	5	11	2	1	5
	> 2.3	8	7	6	3	5
George	≤ 2.3	16	19	1	2	3
	> 2.3	17	1	0	1	3

5.4.8. Multinomial Logistic Regression (Alli/Alligator)

Assume multinomial response with a logit link.

$$X_{ij.} \sim \text{Multinomial}(p_{ij.}, n_{ij.})$$

$$p_{ijk} = \phi_{ijk} / \sum_k \phi_{ijk}$$

$$\phi_{ijk} = \exp(\alpha_k + \beta_{ik} + \gamma_{jk})$$

$$X_{ij.} = X_{ij1}, \dots, X_{ij5}$$

$$n_{ij} = \sum_k X_{ijk},$$

$$\alpha_1 = \beta_{11} = \beta_{1k} = \gamma_{11} = \gamma_{1k} = 0 \text{ (corner constraints)}$$

5.4.8. Multinomial Logistic Regression (Alli/Alligator)

BUGS CODE

...

LIKELIHOOD

```
for (i in 1 : I) { for (j in 1 : J) {
  X[i,j,1:K] ~ dmulti( p[i,j,1:K] , n[i,j])
  n[i,j] <- sum(X[i,j,])
  for (k in 1:K) {
    p[i,j,k]<- phi[i,j,k]/sum(phi[i,j,])
    log(phi[i,j,k]) <- alpha[k] + beta[i,k]
    + gamma[j,k]
  }
}
```

5.4.8. Multinomial Logistic Regression (Alli/Alligator)

Alternatively assume Poisson response

$$X_{ijk} \sim \text{Poisson}(\mu_{ijk})$$

$$\log(\mu_{ijk}) = \lambda_{ij} + \alpha_k + \beta_{ik} + \gamma_{jk}$$

POISSON LIKELIHOOD

```
for (i in 1 : I) { for (j in 1 : J) {
  for (k in 1:K){
    X[i,j,k] ~ dpois( mu[i,j,k])
    log(mu[i,j,k]) <- lambda[i,j] + alpha[k] +
                      beta[i,k] + gamma[j,k]
  }
}}
```

5.4.9. Other Interesting Examples

BUGS Examples Vol.1, σελ. 30, Example 9
(Epil: repeated measures on Poisson Counts)

BUGS Examples Vol.2, σελ. 18, Example 7
(Jaw: repeated measures ANOVA)

BUGS Examples Vol.2, σελ. 55, Example 15
(Endo: conditional inference for matched samples in case-control studies)

Bayesian Biostatistics Using BUGS



ΤΕΛΟΣ 3ου ΜΑΘΗΜΑΤΟΣ

**ΕΠΟΜΕΝΟ:
MODEL DIAGNOSTICS AND
VARIABLE SELECTION**

E-mail: ntzoufras@aueb.gr



Department of Statistics,
Athens University of
Economics & Business