

MCMC DIAGNOSTICS

- Πόσο πρέπει να περιμένουμε για να επιτευχθεί η στασιμότητα;
- Πόσο μεγάλο πρέπει να είναι το m (μετά την στασιμότητα για πόσο πρέπει να τρέξεις την αλυσίδα σου);
- Από που να ξεκινήσεις;

Για να βεβαιωθούμε για όλα τα παραπάνω ερωτήματα χρησιμοποιούμε **διαγνωστικούς ελέγχους (MCMC diagnostics)**. Τέσσερις από τους πιο γνωστούς διαγνωστικούς ελέγχους μπορούν να εφαρμοστούν στα MCMC δεδομένα με την βοήθεια της R βιβλιοθήκης **CODA**. Ας τους εφαρμόσουμε πάνω στο τελευταίο παράδειγμα με τα NB10 δεδομένα.

CODA

- Αφού τρέξουμε το WinBUGS, πηγαίνουμε στο "Sample Monitor Tool" (Επιλέγοντας το "Samples" από το μενού "Inference"), βάζουμε * στο "node", για να επιλέξουμε όλες τις παραμέτρους, και επιλέγουμε το κουτί "CODA". Δύο παράθυρα εμφανίζονται, ένα με όλες τις προσομοιωμένες τιμές για κάθε παράμετρο, και ένα που επεξηγεί ποιες τιμές αναφέρονται σε ποιες παραμέτρους. Επιλέγουμε το κάθε ένα παράθυρο χωριστά και το αποθηκεύουμε σαν .txt αρχείο, από το "Save As" στο μενού "File". Εν συνέχεια αλλάζουμε τις καταλήξεις στα δύο αρχεία, το μεν με τις προσομοιωμένες τιμές σε .out και το επεξηγηματικό σε .ind. Ας υποθέσουμε λοιπόν ότι έχουμε τα αρχεία extendmore.out και extendmore.ind
- Ανοίγουμε την R και μεταφερόμαστε στον φάκελο που σώσαμε τα δύο αρχεία. Εγκαθιστούμε τις βιβλιοθήκες CODA και Lattice (**install packages**) και εν συνέχεια τις φορτώνουμε (**load packages**).

CODA

> codamenu()
CODA startup menu

- 1: Read BUGS output files
- 2: Use an mcmc object
- 3: Quit

Selection: 1

Enter CODA index file name
(or a blank line to exit)

1: extendmore.ind

Enter CODA output file names, separated by return key
(leave a blank line when you have finished)

1: extendmore.out

2:

Abstracting mu ... 40000 valid values

Abstracting nu ... 40000 valid values

Abstracting sigma ... 40000 valid values

Abstracting y.new ... 40000 valid values

Checking effective sample size ...OK

CODA Main Menu

- 1: Output Analysis
- 2: Diagnostics
- 3: List/Change Options
- 4: Quit

Selection: 2

CODA

CODA Diagnostics Menu

- 1: Geweke
- 2: Gelman and Rubin
- 3: Raftery and Lewis
- 4: Heidelberger and Welch
- 5: Autocorrelations
- 6: Cross-Correlations
- 7: List/Change Options
- 8: Return to Main Menu

Geweke

- Ο Geweke (1992) πρότεινε έναν απλό διαγνωστικό έλεγχο, στηριζόμενος σε απλές ιδέες χρονοσειρών. Αν σκεφτούμε κάθε στήλη από τα MCMC δεδομένα ως μία χρονοσειρά, τότε είναι λογικό πως αν η αλυσίδα έχει συγκλίνει στην στάσιμη κατανομή, οι μέσοι των πρώτων 10% (π.χ.) και τελευταίων 50% (π.χ.) των παρατηρήσεων πρέπει να είναι σχεδόν ίσοι. Άρα μπορούμε να προβούμε σε ένα z-test για να ελέγξουμε αυτήν την ισότητα. Αν για κάποια παράμετρο η απόλυτη τιμή του z-score είναι πολύ μεγαλύτερη από το 2, τότε αυτό είναι μια ένδειξη ότι η αλυσίδα δεν έχει συγκλίνει και θα πρέπει να αυξηθεί η περίοδος burn-in.

Geweke

GEWEKE CONVERGENCE DIAGNOSTIC (Z-score)

=====

Iterations used = 1001:41000
Thinning interval = 1
Sample size per chain = 40000

\$extendmore.out

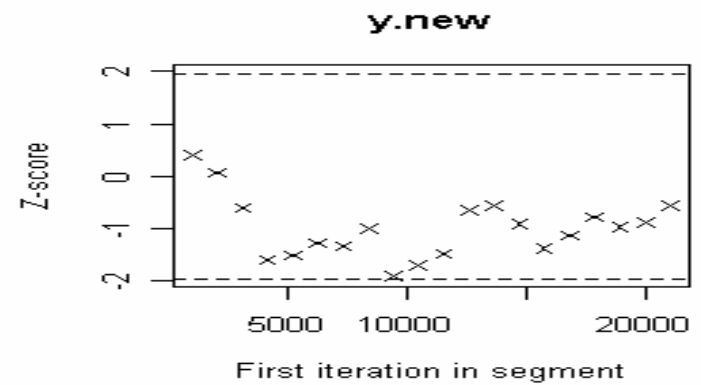
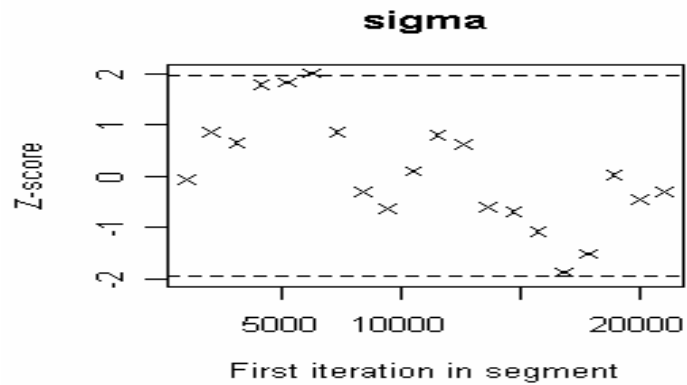
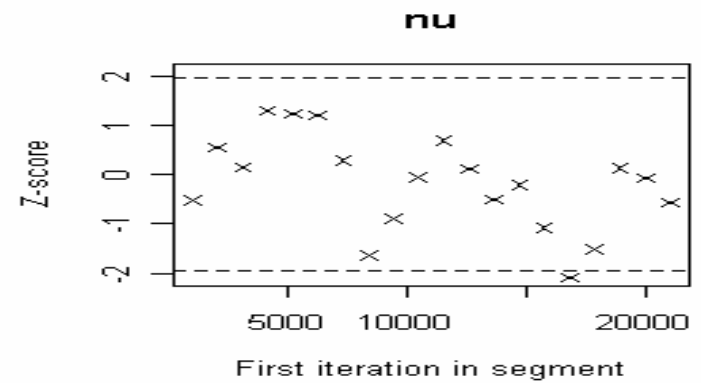
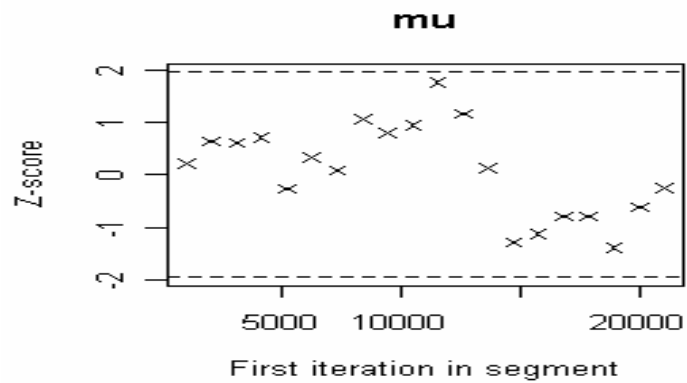
Fraction in 1st window = 0.1
Fraction in 2nd window = 0.5

mu	nu	sigma	y.new
0.22074	-0.51724	-0.05966	0.39525

Geweke plots menu

- 1: Change window size
- 2: Plot Z-scores
- 3: Change number of bins for plot
- 4: Return to Diagnostics Menu

Geweke



Gelman and Rubin

- Οι Gelman and Rubin (1992), πρότειναν έναν έλεγχο που κοιτά για πολυκόρυφη εκ των υστέρων κατανομή. Αν η εκ των υστέρων κατανομή έχει 2 κορυφές, την μία μακριά από την άλλη και ξεκινήσεις την αλυσίδα από τιμές κοντά στην μία μόνο κορυφή μπορεί να μην επισκεφτείς ποτέ την άλλη. Η ιδέα λοιπόν είναι να τρέξεις την αλυσίδα 2 ή περισσότερες φορές από διαφορετικές αρχικές τιμές και να ελέγξεις αν οι αλυσίδες συγκλίνουν στο ίδιο σημείο. Η μέθοδος βασίζεται στην σύγκριση των ανάμεσα των αλυσίδων (within) και μεταξύ των αλυσίδες (between) διασπορών (ANOVA) για κάθε παράμετρο. Η σύγκριση αυτή γίνεται για να εκτιμήσουμε τον παράγοντα κατά τον οποίο η παράμετρος μεταβλητότητας της περιθώριας εκ των υστέρων κατανομής για την κάθε παράμετρο μπορεί να μειωθεί (**shrink factor – παράγων συρρίκνωσης**) αν η αλυσίδα έτρεχε για άπειρο χρόνο. Ως αποτέλεσμα παίρνουμε τα 50% και 97.5% ποσοστιαία σημεία της κατανομής των παραγόντων συρρίκνωσης. Τιμές κοντά στο 1 δηλώνουν μικρή ένδειξη διαφοροποίησης μεταξύ των κατανομών στις οποίες οι αλυσίδες συγκλίνουν.

Gelman and Rubin

GELMAN AND RUBIN DIAGNOSTIC

=====

Iterations used = 1001:41000
Thinning interval = 1
Sample size per chain = 40000

Potential scale reduction factors:

	Point est.	97.5% quantile
mu	1.00	1.00
nu	1.00	1.00
sigma	1.00	1.00
y.new	1.01	1.01

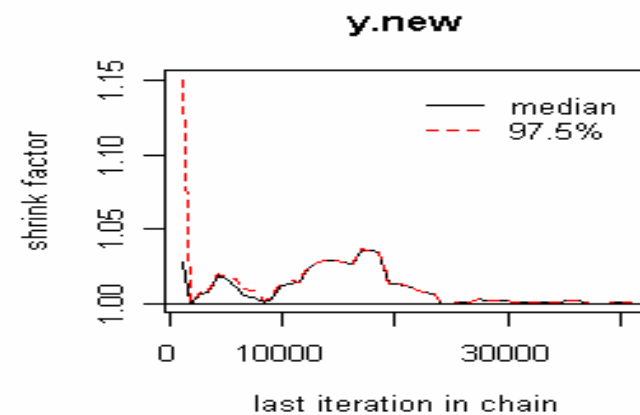
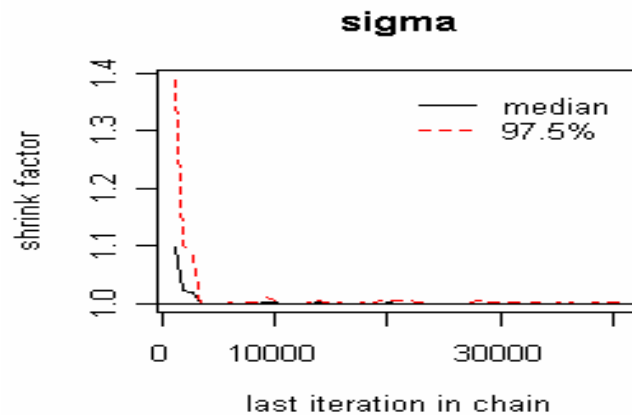
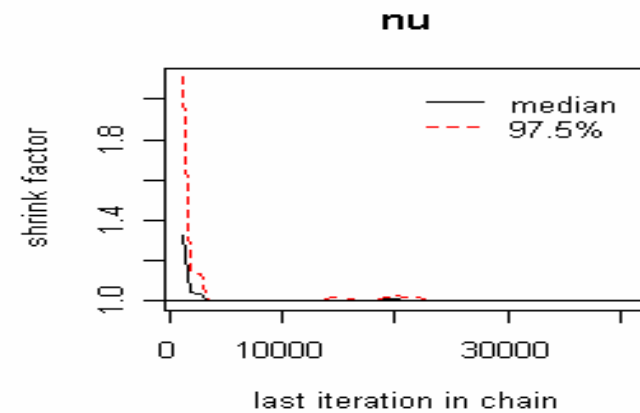
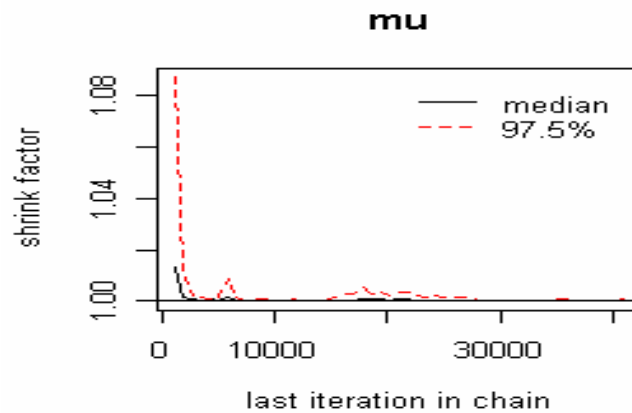
Multivariate psrf

1

Gelman & Rubin menu

- 1: Shrink Factor Plots
- 2: Change bin size for shrink plot
- 3: Return to Diagnostics Menu

Gelman and Rubin



Raftery and Lewis

- Οι Raftery and Lewis (1992) , πρότειναν έναν έλεγχο που απαντάει στο ερώτημα πόσο μεγάλα πρέπει να είναι τα b και m . Η απάντηση του ερωτήματος εξαρτάται άμεσα από το πόσο ακριβείς θέλουμε οι περιγραφικοί στατιστικοί δείκτες της εκ των υστέρων κατανομής να είναι.
 1. Ποια ποσοστιαία σημεία των περιθωρίων εκ των υστέρων μας ενδιαφέρουν; (συνήθως 2.5% και 97.5%)
 2. Τι MCMC ακρίβεια; (συνήθως 0.005)
 3. Με ποια ελάχιστη πιθανότητα θέλουμε να επιτύχουμε την παραπάνω ακρίβεια; (συνήθως 0.95)

Raftery and Lewis

- Ως αποτελέσματα του ελέγχου έχουμε τα ακόλουθα:
 1. Ένα προβλεπόμενο διάστημα αραίωσης (**thinning interval**). Όταν η αλυσίδα δεν αναμιγνύεται σωστά, συνήθως έχουμε μεγάλες θετικές αυτοσυσχετίσεις, π.χ. το μ_{1000} έχει υψηλή εξάρτηση από το μ_{999} , μ_{998} , κλπ. Με άλλα λόγια οι τυχαίες τιμές κατά την διαδικασία προσομοίωσης δεν κινούνται γύρω από τον παραμετρικό χώρο γρήγορα. Ένας τρόπος επίλυσης αυτού του προβλήματος είναι να τρέξουμε την αλυσίδα για μεγαλύτερο διάστημα και να κατοχυρώνουμε κάθε k επανάληψη (**thinning interval**). (**Μόνο στο παλιό CODA**)
 2. Προτεινόμενο επιπλέον burn-in.
 3. Προτεινόμενος συνολικός αριθμός επαναλήψεων N (συμπεριλαμβανομένου του burn-in).
 4. Ένα κάτω όριο του συνολικού αριθμού επαναλήψεων N_{\min} (αν οι τιμές ήταν IID).
 5. Ο λόγος $I=N/N_{\min}$ (**dependent factor**), τιμές κοντά στο 1 δηλώνουν καλή μίξη, ενώ τιμές μεγαλύτερες του 5 συνήθως υποδηλώνουν πρόβλημα.

Raftery and Lewis

RAFTERY AND LEWIS CONVERGENCE DIAGNOSTIC

=====

Iterations used = 1001:41000
Thinning interval = 1
Sample size per chain = 40000

\$extendmore.out

Quantile (q) = 0.025
Accuracy (r) = +/- 0.005
Probability (s) = 0.95

	Burn-in (M)	Total (N)	Lower bound (Nmin)	Dependence factor (I)
mu	3	6218	3746	1.66
nu	9	12579	3746	3.36
sigma	9	13440	3746	3.59
y.new	2	3850	3746	1.03

- 1: Change parameters
- 2: Return to diagnostics menu

Heidelberger and Welch

- Οι Heidelberger and Welch (1983), πρότειναν έναν διαγνωστικό έλεγχο που χρησιμοποιεί το **Cramer-von-Mises τεστ** για στασιμότητα. Το CODA επίσης εφαρμόζει το **half-width τεστ**, το οποίο προσπαθεί να αποφανθεί αν με συγκεκριμένη ακρίβεια μπορούμε να εκτιμήσουμε τους εκ των υστέρων μέσους.

Heidelberger and Welch

HEIDELBERGER AND WELCH STATIONARITY AND INTERVAL HALFWIDTH TESTS

=====

Iterations used = 1001:41000
Thinning interval = 1
Sample size per chain = 40000

Precision of halfwidth test = 0.1

\$extendmore.out

	Stationarity start test	iteration	p-value
mu	passed	1	0.704
nu	passed	1	0.783
sigma	passed	1	0.443
y.new	passed	1	0.241

	Halfwidth Mean test	Halfwidth
mu	passed 404.29	0.00627
nu	passed 3.63	0.04512
sigma	passed 3.86	0.01261
y.new	passed 404.30	0.10553

- 1: Change precision
- 2: Return to diagnostics menu

Autocorrelations

AUTOCORRELATIONS WITHIN EACH CHAIN:

=====

Iterations used = 1001:41000

Thinning interval = 1

Sample size per chain = 40000

\$extendmore.out

, , mu

	mu	nu	sigma	y.new
Lag 0	1.00000	0.06661	0.06035	0.03869
Lag 1	0.32200	0.06608	0.06023	0.01277
Lag 5	0.00159	0.04092	0.02698	-0.00186
Lag 10	0.00402	0.02491	0.02840	-0.00429
Lag 50	0.00377	-0.00564	0.00483	-0.00718

Autocorrelations

, , nu

	mu	nu	sigma	y.new
Lag 0	0.0666	1.00000	0.51619	0.00234
Lag 1	0.0618	0.83617	0.51917	0.00298
Lag 5	0.0296	0.44756	0.30549	-0.00168
Lag 10	0.0189	0.21468	0.14458	-0.00799
Lag 50	-0.0104	0.00158	0.00186	-0.00049

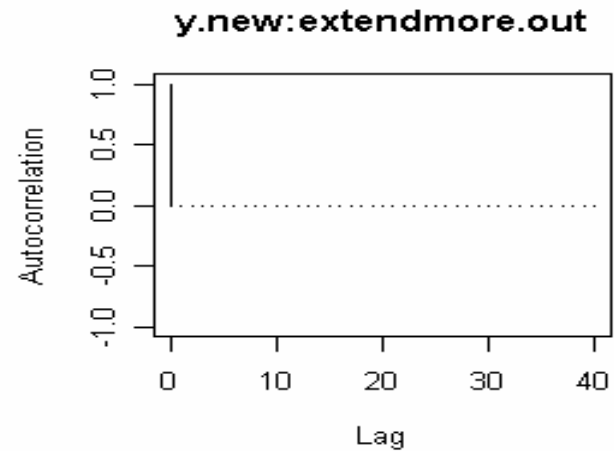
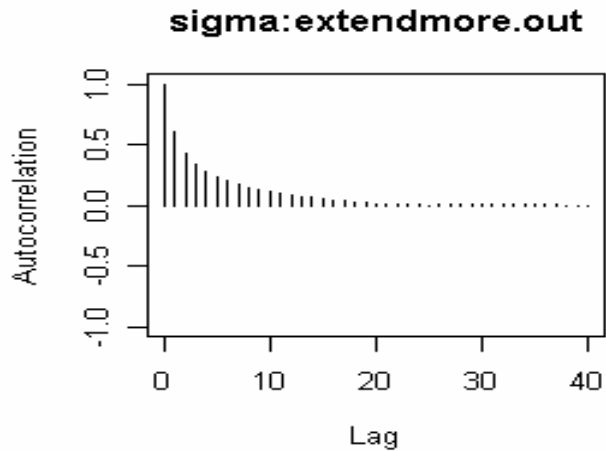
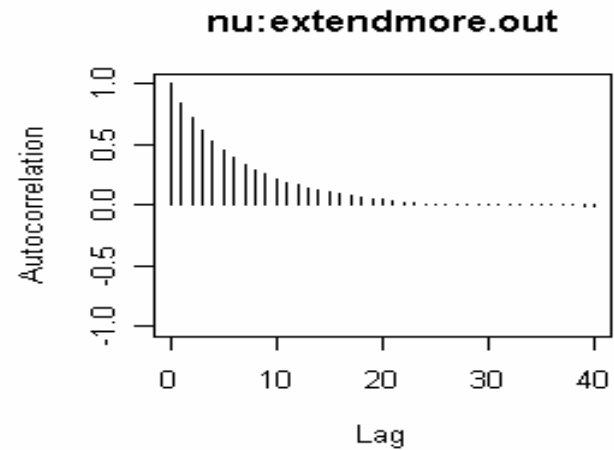
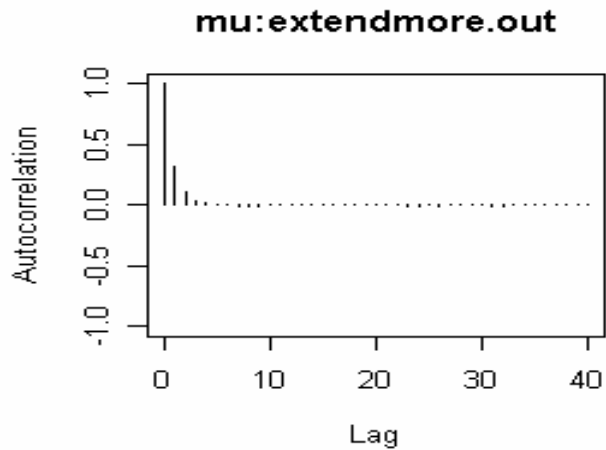
, , sigma

	mu	nu	sigma	y.new
Lag 0	0.060349	0.51619	1.00000	-0.002128
Lag 1	0.058392	0.51838	0.60252	0.004562
Lag 5	0.024566	0.31298	0.23634	-0.005980
Lag 10	0.016039	0.15780	0.11083	-0.002247
Lag 50	-0.000554	0.00862	-0.00481	-0.000639

, , y.new

	mu	nu	sigma	y.new
Lag 0	0.038686	0.002343	-2.13e-03	1.00000
Lag 1	0.021490	0.002859	4.30e-04	0.00116
Lag 5	0.000316	0.001636	-7.48e-03	0.00396
Lag 10	0.005483	0.002913	-5.80e-03	0.00166
Lag 50	0.007340	0.000822	6.24e-05	-0.00432

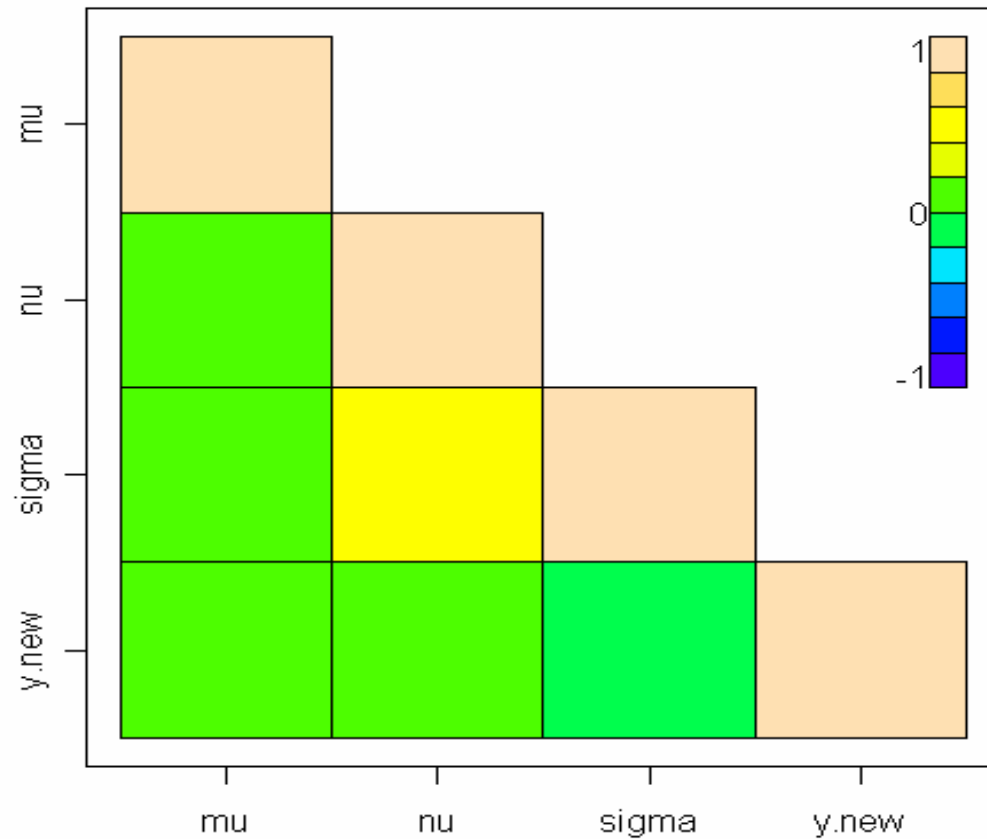
Autocorrelations



Cross-Correlations

	mu	nu	sigma	y.new
mu	1.0000	0.06661	0.06035	0.03869
nu	0.0666	1.00000	0.51619	0.00234
sigma	0.0603	0.51619	1.00000	-0.00213
y.new	0.0387	0.00234	-0.00213	1.00000

Cross-Correlations



Output Analysis

- Το CODA τέλος μπορεί να μας δώσει και διάφορα γραφήματα ή περιγραφικά μέτρα των προσομοιωμένων τιμών τα οποία είναι ανάλογα με αυτά που μας δίνει και το WinBUGS.

Back to Autocorrelations

- Η χρονοσειρά θ_t^* καλείτε αυτοπαλινδρομική ανέλιξη τάξεως p (**autoregressive process of order p - AR_p**) αν:

$$\theta_t^* = \alpha_1 \theta_{t-1}^* + \dots + \alpha_p \theta_{t-p}^* + e_t$$

με e_t μία IID ανέλιξη μέσου 0 και διασποράς σ_e^2 . Η μερική συνάρτηση αυτοσυσχέτισης (**partial autocorrelation function - PACF**) υπολογίζει την συσχέτιση μεταξύ θ_t^* και θ_{t+k}^* και μας βοηθάει να προσδιορίσουμε την τάξη p . Για μία AR_p η PACF στις υστερήσεις $1, \dots, p$ θα είναι σημαντικά διάφορη του 0, και κοντά στο 0 για υστερήσεις $> p$. Για το v στο παράδειγμα μας μπορούμε να θεωρήσουμε ότι πρόκειται για AR_1 ανέλιξη με πρώτης τάξης αυτοσυσχέτιση περίπου $r_1 = 0.836$.

MCMC Accuracy

- Ας υποθέσουμε ότι θ_t^* είναι μια στάσιμη χρονοσειρά με μέσο μ_θ και διασπορά σ_θ^2 . Αν η παραπάνω χρονοσειρά είναι AR_1 με πρώτης τάξης αυτοσυσχέτιση ρ_1 , μπορεί να αποδειχτεί ότι

$$V(\bar{\theta}^*) = \frac{\sigma_\theta^2}{m} \left(\frac{1 + \rho_1}{1 - \rho_1} \right).$$

- Άρα άμα θέλουμε να εκτιμήσουμε το μέσο μιας ποσότητας θ , με συγκεκριμένη ακρίβεια T , πρέπει ο αριθμός των επαναλήψεων m να ικανοποιεί την σχέση:

$$SE(\bar{\theta}^*) = \frac{\hat{\sigma}_\theta}{\sqrt{m}} \sqrt{\frac{1 + \hat{\rho}_1}{1 - \hat{\rho}_1}} \leq T \Rightarrow m \geq \left(\frac{\hat{\sigma}_\theta}{T} \right)^2 \left(\frac{1 + \hat{\rho}_1}{1 - \hat{\rho}_1} \right).$$

MCMC Accuracy

- Παρατηρούμε ότι για $\rho_1 \rightarrow 1$ έχουμε $m \rightarrow \infty$.
- Για το n για παράδειγμα έχουμε $\hat{\rho}_1 = r_1$ και η δειγματική διασπορά είναι 1.161. Αν θέλουμε το $T = 0.01$ τότε:

$$m \geq \left(\frac{1.161}{0.01} \right)^2 \left(\frac{1+0.836}{1-0.836} \right) \approx 150000.$$