

User Manual for the Linux/Unix Console Version of
BACC (Bayesian Analysis, Computation, and
Communication)

William McCausland John J. Stevens

October 29, 2003

Important Information

About this Manual

This manual describes the software developed in connection with the project Bayesian Communication in the Social Sciences, Siddhartha Chib and John Geweke principle investigators. Acknowledgement in any resulting published work would be appreciated. This project was supported, in part, by Grants SBR-9600040 and SBR-9731037 from the National Science Foundation.

Help Keep This Software Free

The National Science Foundation supports this software and its continued development. It is important that we document the use of BACC. We respectfully request that all publications and working papers reporting the results of research using BACC software, include the following acknowledgement and reference:

Computations reported in this paper were undertaken [in part] using the Bayesian Analysis, Computation and Communication software (<http://www.econ.umn.edu/~bacc>) described in:

Geweke, J. (1999) "Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication" (with discussion and rejoinder), *Econometric Reviews* 18: 1-126.

BACC Software and Documentation

BACC software and documentation is available on the web at

<http://www.econ.umn.edu/~bacc>

Please send any comments or questions to

bacc@econ.umn.edu

Contents

Important Information	3
1 Getting Started with BACC	11
1.1 Introduction	11
1.2 Requirements	12
1.3 Installation and Configuration	12
2 Models	15
2.1 Introduction	15
2.1.1 Dimension parameters	15
2.1.2 Unknown Quantities	15
2.1.3 Known Quantities	15
2.1.4 Data Generating Process	16
2.1.5 Prior Distribution	16
2.1.6 Creating a Model Instance	16
2.1.7 Sampling Algorithms	16
2.1.8 Marginal Likelihood	16
2.2 The Normal Linear Model	17
2.2.1 Dimension parameters	17
2.2.2 Unknown Quantities	17
2.2.3 Known Quantities	17
2.2.4 Data Generating Process	17
2.2.5 Priors	17
2.2.6 Creating a Model Instance	18
2.2.7 Sampling Algorithms	18
2.3 The Seemingly Unrelated Regressions Model	19
2.4 The I.I.D. Finite State Model	20
2.4.1 Dimension parameters	20
2.4.2 Unknown Quantities	20
2.4.3 Known Quantities	20
2.4.4 Data Generating Process	20
2.4.5 Priors	20
2.4.6 Creating a Model Instance	20
2.4.7 Sampling Algorithms	20

2.4.8	Marginal Likelihood	21
2.5	The Non-Stationary First Order Markov Finite State Model	22
2.5.1	Dimension parameters	22
2.5.2	Unknown Quantities	22
2.5.3	Known Quantities	22
2.5.4	Data Generating Process	22
2.5.5	Priors	22
2.5.6	Creating a Model Instance	23
2.5.7	Sampling Algorithms	23
2.5.8	Marginal Likelihood	23
2.6	The Stationary First Order Markov Finite State Model	24
2.6.1	Dimension parameters	24
2.6.2	Unknown Quantities	24
2.6.3	Known Quantities	24
2.6.4	Data Generating Process	24
2.6.5	Priors	24
2.6.6	Creating a Model Instance	25
2.6.7	Sampling Algorithms	25
2.7	The Poisson Model	26
2.7.1	Dimension parameters	26
2.7.2	Unknown Quantities	26
2.7.3	Known Quantities	26
2.7.4	Data Generating Process	26
2.7.5	Priors	26
2.7.6	Creating a Model Instance	26
2.7.7	Sampling Algorithms	26
2.7.8	Marginal Likelihood	27
2.8	The Uniform Model	28
2.8.1	Dimension parameters	28
2.8.2	Unknown Quantities	28
2.8.3	Known Quantities	28
2.8.4	Data Generating Process	28
2.8.5	Priors	28
2.8.6	Creating a Model Instance	28
2.8.7	Sampling Algorithms	29
2.8.8	Marginal Likelihood	29
2.9	A Univariate Linear Model with Normal Disturbances	30
2.10	The Dichotomous Choice Model (with normally distributed disturbances)	32
2.10.1	Dimension parameters	32
2.10.2	Unknown Quantities	32
2.10.3	Known Quantities	32
2.10.4	Data Generating Process	32
2.10.5	Priors	32
2.10.6	Creating a Model Instance	32
2.10.7	Sampling Algorithms	33

2.11	The Censored Linear Model (with normally distributed disturbances)	34
2.11.1	Dimension parameters	34
2.11.2	Unknown Quantities	34
2.11.3	Known Quantities	34
2.11.4	Data Generating Process	34
2.11.5	Priors	34
2.11.6	Creating a Model Instance	34
2.11.7	Sampling Algorithms	35
2.12	The Univariate Latent Linear Model (with normally distributed disturbances)	36
2.12.1	Dimension parameters	36
2.12.2	Unknown Quantities	36
2.12.3	Known Quantities	36
2.12.4	Data Generating Process	36
2.12.5	Priors	36
2.12.6	Creating a Model Instance	36
2.12.7	Sampling Algorithms	37
2.13	A Univariate Linear Model with Student t Disturbances	38
2.14	The Dichotomous Choice Model (with Student t distributed disturbances)	40
2.14.1	Dimension parameters	40
2.14.2	Unknown Quantities	40
2.14.3	Known Quantities	40
2.14.4	Data Generating Process	40
2.14.5	Priors	40
2.14.6	Creating a Model Instance	40
2.14.7	Sampling Algorithms	41
2.15	The Censored Linear Model (with Student t distributed disturbances)	42
2.15.1	Dimension parameters	42
2.15.2	Unknown Quantities	42
2.15.3	Known Quantities	42
2.15.4	Data Generating Process	42
2.15.5	Priors	42
2.15.6	Creating a Model Instance	42
2.15.7	Sampling Algorithms	43
2.16	The Univariate Latent Linear Model (with Student t distributed disturbances)	44
2.16.1	Dimension parameters	44
2.16.2	Unknown Quantities	44
2.16.3	Known Quantities	44
2.16.4	Data Generating Process	44
2.16.5	Priors	44
2.16.6	Creating a Model Instance	44
2.16.7	Sampling Algorithms	45

2.17	A Univariate Linear Model with Finite Mixtures of Normals Disturbances . . .	46
2.18	The Dichotomous Choice Model (with a scale mixture of normals distribution for the disturbances)	48
2.18.1	Dimension parameters	48
2.18.2	Unknown Quantities	48
2.18.3	Known Quantities	48
2.18.4	Data Generating Process	48
2.18.5	Priors	48
2.18.6	Creating a Model Instance	49
2.18.7	Sampling Algorithms	49
2.19	The Censored Linear Model (with a scale mixture of normals distribution for the disturbances)	50
2.19.1	Dimension parameters	50
2.19.2	Unknown Quantities	50
2.19.3	Known Quantities	50
2.19.4	Data Generating Process	50
2.19.5	Priors	50
2.19.6	Creating a Model Instance	51
2.19.7	Sampling Algorithms	51
2.20	The Univariate Latent Linear Model (with a scale mixture of normals distribution for the disturbances)	52
2.20.1	Dimension parameters	52
2.20.2	Unknown Quantities	52
2.20.3	Known Quantities	52
2.20.4	Data Generating Process	52
2.20.5	Priors	52
2.20.6	Creating a Model Instance	53
2.20.7	Sampling Algorithms	53
2.21	An Autoregression Model	54
2.21.1	Dimension Parameters	54
2.21.2	Unknown Quantities	54
2.21.3	Known Quantities	54
2.21.4	Data Generating Process	54
2.21.5	Prior Distribution	55
2.21.6	Creating a Model Instance	55
2.21.7	Sampling Algorithm	55
2.22	An Autoregression Model with State Dependant Means	56
2.22.1	Dimension Parameters	56
2.22.2	Unknown Quantities	56
2.22.3	Known Quantities	56
2.22.4	Data Generating Process	57
2.22.5	Prior Distribution	57
2.22.6	Creating a Model Instance	58
2.22.7	Sampling Algorithm	58

3	BACC Commands	59
3.1	Overview of BACC Commands	59
3.2	Command Line Issues	60
3.2.1	Matrix Files	60
3.2.2	Filter Files	60
3.2.3	Batch Files	61
3.3	Command Line Issues	61
3.3.1	Matrix Files	61
3.3.2	Filter Files	61
3.3.3	Batch Files	61
3.4	Detailed Description of Commands	61
3.4.1	The <code>dirichletSim</code> Command	62
3.4.2	The <code>expect1</code> Command	63
3.4.3	The <code>expectN</code> Command	66
3.4.4	The <code>extract</code> Command	68
3.4.5	The <code>gammaSim</code> Command	69
3.4.6	The <code>gaussianSim</code> Command	70
3.4.7	The <code>listModelSpecs</code> Command	71
3.4.8	The <code>listModels</code> Command	72
3.4.9	The <code>miDelete</code> Command	73
3.4.10	The <code>miLoad</code> Command	74
3.4.11	The <code>miLoadAscii</code> Command	75
3.4.12	The <code>miSave</code> Command	76
3.4.13	The <code>miSaveAscii</code> Command	77
3.4.14	The <code>minst</code> Command	78
3.4.15	The <code>mlike</code> Command	79
3.4.16	The <code>paretoSim</code> Command	81
3.4.17	The <code>postfilter</code> Command	82
3.4.18	The <code>postsim</code> Command	83
3.4.19	The <code>postsimHM</code> Command	84
3.4.20	The <code>priorRobust</code> Command	85
3.4.21	The <code>priorfilter</code> Command	87
3.4.22	The <code>priorsim</code> Command	88
3.4.23	The <code>setseedconstant</code> Command	89
3.4.24	The <code>setseedtime</code> Command	90
3.4.25	The <code>weightedSmooth</code> Command	91
3.4.26	The <code>wishartSim</code> Command	93
4	A BACC Tutorial	95
4.1	Creating a Model Instance	96
4.2	Simulating the Model	96
4.3	Computing Marginal likelihoods	97
4.4	Computing Moments of Functions of Interest	97
4.5	Sensitivity to the Prior	98
4.6	Kernel Smoothing of Simulations	98
4.7	Saving and Loading Model Instances	99

4.7.1	To/From Text Files	99
4.8	Drawing from Various Distributions	99
4.8.1	Dirichlet	99
4.8.2	Gamma	99
4.8.3	Gaussian	99
4.8.4	Pareto	100
4.8.5	Wishart	100
A	Distributions	101
A.1	The Dirichlet Distribution	101
A.2	The Gamma Distribution	101
A.3	The Normal Distribution	102
A.4	The Pareto Distribution	102
A.5	The Poisson Distribution	102
A.6	The Wishart Distribution	103
	Bibliography	104

Chapter 1

Getting Started with BACC

1.1 Introduction

The BACC software provides the user several commands for doing Bayesian analysis and communications. This document describes the function of these commands and their inputs and outputs. It also outlines some of the theory behind the commands, and provides references to the relevant literature.

The following versions of the BACC software and documentation are available:

- Windows Matlab
- Linux/Unix Matlab
- Windows Splus
- Linux/Unix Splus
- Windows R
- Linux/Unix R
- Windows Gauss
- Linux/Unix Gauss
- Windows Console
- Linux/Unix Console

This particular manual is for the Linux/Unix Console version of BACC.

1.2 Requirements

The Linux/Unix Console version of the BACC software requires no additional software to run. It serves two purposes. First, it allows model developers to debug new model specifications using standard debugging software, such as `dbx` or `gdb`. Second, it is the version of last resort for those who do not have installed any of the applications that other versions run under. Note, however, that there are now Linux/Unix R and Windows R versions of BACC, and that R can be downloaded for free from www.r-project.org. There are both Windows and Linux/Unix versions.

The Windows SPlus version of the BACC software requires Splus Professional Edition for Microsoft Windows 95 or higher. The Splus standard Edition does not have commands windows to work with dynamic linked libraries (DLLs). If you do not have the Professional Edition, you can use the Windows R version of the BACC software.

The Windows R Version of the BACC software requires R 1.1 or higher for Microsoft Windows 95 or higher.

1.3 Installation and Configuration

Follow these steps to install the Linux/Unix Console version of BACC.

1. Download the tarball file `bacc.tar.gz` from the software page of the BACC website <http://www.econ.umn.edu/~bacc>.

2. Create a directory for BACC.

```
% cd ~
% mkdir bacc
```

Variations: The directory does not need to be named `bacc`.

3. Put the file `bacc.tar.gz` into this directory.
4. Create the BACC directory structure from the file `bacc.tar.gz`.

```
% cd ~/bacc
% gunzip bacc.tar.gz
% tar -xpf bacc.tar
```

5. If you are building BACC under Unix, you need to make two minor modifications to the file `~/bacc/Makefile`. Search for the string 'Unix' and follow the instructions in the file. If you are building BACC under Linux, do not make these changes.
6. Make the Linux/Unix console version of BACC.

```
% cd ~/bacc
% make console
```

It will take some time to build BACC, and there will be a long stream of messages, which can be ignored.

7. If necessary, set execute privileges. If you are not sure whether it is necessary or not, it will not hurt to do it anyway.

```
chmod u+x /bacc/Console/*
```

8. Copy all files from `~/bacc/Console` to some directory in your path. In the example below, it is assumed that you have a local `bin` directory, and that it is part of your path. See a Unix manual if you are unsure how to add a directory to your path.

```
% cp ~/bacc/Console/* ~/bin
```

The Linux/Unix Console version of BACC is now installed. Follow these steps to run the sample program `testBACC`.

1. Change directory to the Test directory.

```
% cd /bacc/Console/Test
```

2. Execute the test program.

```
% testBACC
```


Chapter 2

Models

2.1 Introduction

This document specifies the models currently supported by the BACC system. Each section following this one describes one of the supported models. Each model description is organized into subsections, following the pattern of this section. Appendix A gives the probability density and mass functions of the distributions used throughout the document.

2.1.1 Dimension parameters

All the quantities relevant to a model are treated as matrix valued. All matrix sizes are specified in terms of these dimension parameters. Examples of dimension parameters include the number of times a variable is observed, the number of individuals in a cross section, and the number of equations in a linear model. This subsection lists and describes the dimension parameters for a particular model.

2.1.2 Unknown Quantities

Unknown quantities are all the unobserved elements in a model. They include unknown parameters of the model, latent variables, and missing data. Separate sub-sub-sections discuss unknown quantities in each of these categories. Posterior simulation involves drawing these quantities from their posterior distribution; that is, their conditional distribution given known quantities.

2.1.3 Known Quantities

Known quantities are all the observed or user-specified values in a model. They include prior parameters, which index distributions within a family of prior distributions, and observed data. Separate sub-sub-sections discuss known quantities in both of these categories. The user of the BACC software must specify all the known quantities of a model in order to create an instance of the model.

2.1.4 Data Generating Process

This section specifies the conditional distribution of the endogenous observed data, given the unknown quantities and any observed data ancillary with respect to the unknown quantities.

2.1.5 Prior Distribution

This section specifies the marginal distribution of the unknown quantities, reflecting the user's prior beliefs about these quantities. These unknown quantities may or may not be independent. An example where they are not is a hierarchical prior, in which the prior density is expressed as the product of marginal densities of the "lowest level" unknowns and conditional densities of "higher level" unknowns given "lower level" unknowns.

2.1.6 Creating a Model Instance

This section gives all the model specific information a user requires to create a model instance. It specifies a short mnemonic label that identifies the model, the order in which the user gives the names to assign the unknown quantities, and the order in which to supply all the known quantities. To create a model instance, the user issues the `minst` command, with appropriate arguments (see section 3.4.14).

2.1.7 Sampling Algorithms

This section has brief descriptions of the algorithm used to generate samples of unknown quantities from their prior and posterior distributions. One subsection each concerns the prior distribution and the posterior distribution. For further details on the algorithms, the user should consult the internal (source code) documentation for the BACC system.

2.1.8 Marginal Likelihood

Where there is an analytical expression for the marginal likelihood in a model, this subsection provides that expression.

2.2 The Normal Linear Model

2.2.1 Dimension parameters

There are m equations, k covariate coefficients, and T observations of each variable.

2.2.2 Unknown Quantities

Unknown Parameters

There is a $k \times 1$ coefficient parameter β and a $m \times m$ precision parameter H .

2.2.3 Known Quantities

Prior Parameters

There is a $k \times 1$ coefficient mean vector $\underline{\beta}$, a $k \times k$ positive definite coefficient precision matrix \underline{H} , a precision degrees of freedom parameter $\underline{\nu} > \frac{m-1}{2}$, and a positive definite precision inverse scale parameter \underline{S} .

Data

There are m vectors of observations of dependent variables: y_1, \dots, y_m . Each vector is $T \times 1$.

There are m matrices of observations of ancillary (with respect to unknown quantities) variables: Z_1, \dots, Z_m . Each matrix is $T \times k$.

2.2.4 Data Generating Process

$$y \equiv \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} Z_1 \\ \vdots \\ Z_m \end{bmatrix} \beta + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_m \end{bmatrix} \equiv Z\beta + \epsilon$$

$$\begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_m \end{bmatrix} \mid \begin{bmatrix} Z_1 \\ \vdots \\ Z_m \end{bmatrix} \sim N(0, H^{-1} \otimes I_T)$$

2.2.5 Priors

The unknown parameters β and H are a-priori independent, and have the following marginal distributions.

$$\beta \sim N(\underline{\beta}, \underline{H}_\beta^{-1})$$

$$H \sim Wi(\underline{S}^{-1}, \underline{\nu})$$

When $m = 1$, the distribution of $\underline{S}H$ is chi-squared with $\underline{\nu}$ degrees of freedom.

2.2.6 Creating a Model Instance

The mnemonic label identifying the model is `nlm`.

Supply the names you wish to give the unknown quantities in the following order: first the name of β (“beta” for example) and then the name of H .

Supply the known quantities in the following order: $\underline{\beta}$, \underline{H}_β , $\underline{\nu}$, \underline{S} , Z , y .

2.2.7 Sampling Algorithms

Generating Prior Samples

Samples from the prior distribution of β and H are generated independently.

Generating Posterior Samples

The algorithm to generate samples from the posterior distribution $\beta, H|Z, y$ is a Gibbs sampling algorithm with two blocks, based on the following conditional posterior distributions.

$$\beta|H, y, Z \sim \text{N}(\bar{\beta}, \bar{H}_\beta^{-1})$$

$$H|\beta, y, Z \sim \text{W}(\bar{S}^{-1}, \bar{\nu})$$

where

$$\bar{H}_\beta = \underline{H}_\beta + Z'(H \otimes I_T)Z$$

$$\bar{\beta} = \bar{H}_\beta^{-1}[\underline{H}_\beta \underline{\beta} + Z'(H \otimes I_T)y]$$

$$\bar{S} = \underline{S} + [s_{ij}], s_{ij} = (y_i - Z_i \beta)'(y_i - Z_i \beta)$$

$$\bar{\nu} = \underline{\nu} + T$$

2.3 The Seemingly Unrelated Regressions Model

This is a special case of the Normal Linear Model with $m > 1$. Please see section 2.2.

2.4 The I.I.D. Finite State Model

2.4.1 Dimension parameters

There are m states, N individuals and T observation times.

2.4.2 Unknown Quantities

Unknown Parameters

There is a $1 \times m$ state probability vector π .

2.4.3 Known Quantities

Prior Parameters

There is a $1 \times m$ parameter $\underline{\alpha}$ indexing the prior distribution of π .

Data

There are state observations $s_{ti} \in \{1, \dots, m\}$ for each individual i of N individuals and each observation period t of T periods.

$$S = \begin{bmatrix} s_{1,1} & \cdots & s_{N,1} \\ \vdots & \ddots & \vdots \\ s_{T,1} & \cdots & s_{N,T} \end{bmatrix}$$

2.4.4 Data Generating Process

Each observation s_{ti} is independently and identically distributed as follows.

$$\Pr(s_{ti} = s) = \pi_s \quad s = 1, \dots, m$$

2.4.5 Priors

$$\pi \sim \text{Di}(\underline{\alpha})$$

2.4.6 Creating a Model Instance

The mnemonic label identifying the model is `iidfs`.

Supply the name you wish to give the unknown quantity π (“pi” for example).

Supply the known quantities in the following order: $\underline{\alpha}$, S .

2.4.7 Sampling Algorithms

Generating Prior Samples

Samples from the prior distribution of π are generated independently.

Generating Posterior Samples

In this model, the posterior distribution for π is the following familiar distribution.

$$\pi|S \sim \text{Di}(\bar{\alpha})$$

where

$$\bar{\alpha} = \underline{\alpha} + n$$

$$n = [n_1 \cdots n_m]$$

and n_s is the number of observations for which $s_{ti} = s$. Posterior samples are drawn independently from this distribution.

2.4.8 Marginal Likelihood

The marginal likelihood is given by

$$p(S) = \frac{\Gamma(\sum_{i=1}^m \alpha_i) \prod_{i=1}^m \Gamma(\bar{\alpha}_i)}{\prod_{i=1}^m \Gamma(\alpha_i) \Gamma(\sum_{i=1}^m \bar{\alpha}_i)}.$$

2.5 The Non-Stationary First Order Markov Finite State Model

2.5.1 Dimension parameters

There are m states, N individuals and T observation times.

2.5.2 Unknown Quantities

Unknown Parameters

There is a $1 \times m$ initial state probability vector π and an $m \times m$ Markov transition probability matrix P .

2.5.3 Known Quantities

Prior Parameters

The prior parameters are a $1 \times m$ vector $\underline{\alpha}_0$ indexing the prior distribution of π , and an $m \times m$ matrix $\underline{\alpha}$ indexing the prior distribution of P .

Data

There are state observations $s_{ti} \in \{1, \dots, m\}$ for each individual i and each observation time t .

$$S = \begin{bmatrix} s_{11} & \cdots & s_{1N} \\ \vdots & \ddots & \vdots \\ s_{T1} & \cdots & s_{TN} \end{bmatrix}$$

2.5.4 Data Generating Process

The N observation sequences $\{s_{ti}\}_{t=1}^T$ are i.i.d., with each sequence being first order Markov. The initial distribution is π and the Markov transition matrix is P .

$$\Pr(s_{1i} = s) = \pi_s \quad s = 1, \dots, m$$

$$\Pr(s_{ti} = s' | s_{t-1,i} = s) = P_{ss'}$$

2.5.5 Priors

The m rows P_s of P and π are mutually independent, and have the following marginal distributions.

$$\begin{aligned} \pi &\sim \text{Di}(\underline{\alpha}_0) \\ P_s &\equiv [P_{s1}, \dots, P_{sm}] \sim \text{Di}(\underline{\alpha}_s) \quad s = 1, \dots, m \\ \underline{\alpha}_0 &\equiv [\alpha_{01} \quad \cdots \quad \alpha_{0m}] \end{aligned}$$

$$\underline{\alpha}_s \equiv [\underline{\alpha}_{s1} \quad \cdots \quad \underline{\alpha}_{sm}] \quad s = 1, \dots, m$$

2.5.6 Creating a Model Instance

The mnemonic label identifying the model is `nsfomfs`.

Supply the names you wish to give the unknown quantities in the following order: first the name of π (“pi” for example), and then the name of P .

Supply the known quantities in the following order: $\underline{\alpha}_0$, $\underline{\alpha}$, S .

2.5.7 Sampling Algorithms

Generating Prior Samples

Samples from the prior distributions of π and P are generated independently.

Generating Posterior Samples

In the posterior distribution $\pi, P|S$, the parameters π and P are conditionally independent, and their marginal posterior distributions are the following familiar distributions.

$$\pi|S \sim \text{Di}(\bar{\alpha}_0)$$

$$P_s|S \sim \text{Di}(\bar{\alpha}_s) \quad s = 1, \dots, m$$

where

$$\begin{aligned} \bar{\alpha}_0 &\equiv \underline{\alpha}_0 + n_0 \\ \bar{\alpha} &\equiv \underline{\alpha} + n \\ \bar{\alpha}_s &\equiv [\bar{\alpha}_{s1} \quad \cdots \quad \bar{\alpha}_{sm}] \\ n_0 &\equiv [n_{01} \quad \cdots \quad n_{0m}] \\ n &\equiv \begin{bmatrix} n_{11} & \cdots & n_{1m} \\ \vdots & \ddots & \vdots \\ n_{m1} & \cdots & n_{mm} \end{bmatrix} \end{aligned}$$

where n_{0s} is the number of individuals starting in state s , and $n_{ss'}$ is the number of transitions from state s to state s' in the data.

Posterior samples are drawn independently from this distribution.

2.5.8 Marginal Likelihood

The marginal likelihood is available in closed form:

$$p(S) = \frac{\Gamma(\sum_{s=1}^m \underline{\alpha}_{0s}) \prod_{s=1}^m \Gamma(\bar{\alpha}_{0s})}{\prod_{s=1}^m \Gamma(\underline{\alpha}_{0s}) \Gamma(\sum_{s=1}^m \bar{\alpha}_{0s})} \cdot \prod_{s=1}^m \left[\frac{\Gamma(\sum_{s'=1}^m \underline{\alpha}_{ss'}) \prod_{s'=1}^m \Gamma(\bar{\alpha}_{ss'})}{\prod_{s'=1}^m \Gamma(\underline{\alpha}_{ss'}) \Gamma(\sum_{s'=1}^m \bar{\alpha}_{ss'})} \right]$$

2.6 The Stationary First Order Markov Finite State Model

2.6.1 Dimension parameters

There are m states, N individuals and T observation times.

2.6.2 Unknown Quantities

Unknown Parameters

There is an $m \times m$ Markov transition probability matrix P .

2.6.3 Known Quantities

Prior Parameters

The prior parameter is an $m \times m$ matrix $\underline{\alpha}$ indexing the prior distribution of P .

Data

There are state observations $s_{ti} \in \{1, \dots, m\}$ for each individual i and each observation time t .

$$S = \begin{bmatrix} s_{11} & \cdots & s_{1N} \\ \vdots & \ddots & \vdots \\ s_{T1} & \cdots & s_{TN} \end{bmatrix}$$

2.6.4 Data Generating Process

The N observation sequences $\{s_{ti}\}_{t=1}^T$ are i.i.d., with each sequence being first order Markov with transition matrix P . The initial distribution vector is assumed to be the invariant distribution π for P .

$$\Pr(s_{1i} = s) = \pi_s \quad s = 1, \dots, m$$

$$\Pr(s_{ti} = s' | s_{t-1,i} = s) = P_{ss'}$$

where π is the left eigenvector of P corresponding to the eigenvalue $\lambda = 1$.

2.6.5 Priors

The m rows P_s of P are mutually independent, and have the following marginal distributions.

$$P_s \equiv [P_{s1}, \dots, P_{sm}] \sim \text{Di}(\underline{\alpha}_s) \quad s = 1, \dots, m$$

$$\underline{\alpha}_s \equiv [\underline{\alpha}_{s1} \quad \cdots \quad \underline{\alpha}_{sm}] \quad s = 1, \dots, m$$

2.6.6 Creating a Model Instance

The mnemonic label identifying the model is `sfomfs`.

Supply the name you wish to give the unknown quantity P .

Supply the known quantities in the following order: $\underline{\alpha}$, S .

2.6.7 Sampling Algorithms

Generating Prior Draws

Samples from the prior distribution of P are generated independently.

Generating Posterior Draws

In this model, an independence Metropolis-Hastings chain is used to draw from the posterior distribution for P . The distribution $P^*|S$ of candidate draws is

$$P_s^*|S \sim \text{Di}(\bar{\alpha}_s) \quad s = 1, \dots, m$$

where

$$\begin{aligned} \bar{\alpha} &\equiv \underline{\alpha} + n \\ \bar{\alpha}_s &\equiv [\bar{\alpha}_{s1} \quad \cdots \quad \bar{\alpha}_{sm}] \\ n &\equiv \begin{bmatrix} n_{11} & \cdots & n_{1m} \\ \vdots & \ddots & \vdots \\ n_{m1} & \cdots & n_{mm} \end{bmatrix} \end{aligned}$$

where $n_{ss'}$ is the number of transitions from state s to state s' in the data.

The Hastings ratio for this block is given by

$$\prod_{i=1}^N \frac{\pi_{S_{1i}}^*}{\pi_{S_{1i}}}$$

2.7 The Poisson Model

2.7.1 Dimension parameters

There are N observations.

2.7.2 Unknown Quantities

Unknown Parameters

There is a scalar mean parameter λ .

2.7.3 Known Quantities

Prior Parameters

There is a scalar shape parameter $\underline{\alpha} > 0$ and a scalar scale parameter $\underline{\beta} > 0$ indexing the prior distribution of λ .

Data

Each observation x_i is a non-negative integer.

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$$

2.7.4 Data Generating Process

The observations x_i are independently and identically Poisson distributed.

$$x_i \sim \text{Po}(\lambda)$$

2.7.5 Priors

$$\lambda \sim \text{Ga}(\underline{\alpha}, \underline{\beta})$$

2.7.6 Creating a Model Instance

The mnemonic label identifying the model is `poisson`

Supply the names you wish to give the unknown quantity λ (“lambda” for example).

Supply the known quantities in the following order: $\underline{\alpha}$, $\underline{\beta}$, X .

2.7.7 Sampling Algorithms

Generating Prior Samples

Samples from the prior distribution of λ are generated independently.

Generating Posterior Samples

In this model, the posterior distribution for λ is the following familiar distribution.

$$\lambda|X \sim \text{Ga}(\bar{\alpha}, \bar{\beta})$$

$$\bar{\alpha} = \underline{\alpha} + \sum_{i=1}^N x_i$$

$$\bar{\beta} = \underline{\beta} + N$$

Posterior samples are drawn independently from this distribution.

2.7.8 Marginal Likelihood

The marginal likelihood is given by the following expression.

$$p(X) = \frac{\underline{\beta}^N}{(\underline{\beta} + N)^{\underline{\alpha} + r}} \frac{\Gamma(\underline{\alpha} + r)}{\Gamma(\underline{\alpha})} \frac{1}{\prod_{i=1}^N x_i!},$$

where

$$r = \sum_{i=1}^N x_i$$

2.8 The Uniform Model

2.8.1 Dimension parameters

There are N observations.

2.8.2 Unknown Quantities

Unknown Parameters

There is a scalar support parameter θ .

2.8.3 Known Quantities

Prior Parameters

There is a scalar notional count parameter $\underline{\alpha} > 0$ and a scalar notional maximal element parameter $\underline{\beta} \geq 0$ indexing the prior distribution of θ .

Data

Each observation x_i is a non-negative real-valued scalar.

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$$

2.8.4 Data Generating Process

Each observation x is independently and identically distributed with a uniform distribution on $[0, \theta]$.

$$x_i \sim \text{i.i.d.} U(0, \theta)$$

2.8.5 Priors

$$\theta \sim \text{Pa}(\underline{\alpha}, \underline{\beta})$$

2.8.6 Creating a Model Instance

The mnemonic label identifying the model is **uniform**. Supply the name you wish to give the unknown quantity θ (“theta” for example).

Supply the known quantities in the following order: $\underline{\alpha}$, $\underline{\beta}$, X .

2.8.7 Sampling Algorithms

Generating Prior Samples

Samples from the prior distribution of θ are generated independently.

Generating Posterior Samples

In this model, the posterior distribution for θ is the following familiar distribution.

$$\theta|X \sim Pa(\bar{\alpha}, \bar{\beta})$$

where

$$\begin{aligned}\bar{\alpha} &= \underline{\alpha} + N \\ \bar{\beta} &= \max \left\{ \underline{\beta}, \max_i x_i \right\}\end{aligned}$$

Posterior samples are drawn independently from this distribution.

2.8.8 Marginal Likelihood

The marginal likelihood is given by the following expression.

$$p(X) = (\underline{\alpha}/\bar{\alpha}) \prod_{i=1}^N (p_i/\bar{\beta}_i)$$

where

$$\bar{\beta}_i = \max \{ \underline{\beta}, \max_{j \leq i} x_j \}$$

and

$$p_i = \begin{cases} (\frac{\bar{\beta}_i}{x_i})^{(1+\underline{\alpha}+i)} & \text{if } x_i > \bar{\beta}_i \\ 1 & \text{otherwise.} \end{cases}$$

2.9 A Univariate Linear Model with Normal Disturbances

The mnemonic label identifying the model is `n_u1m`.

Dimension Parameters

T number of observations

K number of covariates

Unknown Quantities

β ($K \times 1$) vector of covariate coefficients

h (1×1) precision of disturbance

Known Quantities

$\underline{\beta}$ ($K \times 1$) prior mean of β

\underline{H}_β ($K \times K$) prior precision of β

\underline{s}^2 (1×1) prior inverse scale of h

$\underline{\nu}$ (1×1) prior degrees of freedom of h

X ($T \times K$) covariates

y ($T \times 1$) dependant variable

Data Generating Process

The observables y are given by

$$y = X\beta + u,$$

where u is a $T \times 1$ vector of i.i.d. normal disturbances, with $u_t|h \sim N(0, h^{-1})$:

$$p(u|X, \beta, h) = (2\pi)^{-T/2} h^{T/2} \exp(-hu'u/2).$$

Prior Distribution

The vectors β and h , together with X , are mutually independent. The covariate coefficient vector β has distribution $N(\underline{\beta}, \underline{H}_\beta)$:

$$p(\beta) = (2\pi)^{-K/2} |\underline{H}_\beta|^{1/2} \exp[-(\beta - \underline{\beta})' \underline{H}_\beta (\beta - \underline{\beta})/2].$$

The precision parameter h has a scaled chi-squared distribution, with $\underline{s}^2 h \sim \underline{\nu}$:

$$p(h) = 2^{-\underline{\nu}/2} \Gamma(\underline{\nu}/2)^{-1} (\underline{s}^2)^{\underline{\nu}/2} h^{(\underline{\nu}-2)/2} \exp(-\underline{s}^2 h/2).$$

Sampling Algorithm

See Example 3.4.1 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.10 The Dichotomous Choice Model (with normally distributed disturbances)

2.10.1 Dimension parameters

There are k covariate coefficients and T observations of each variable.

2.10.2 Unknown Quantities

Unknown Parameters

There is a $k \times 1$ coefficient parameter β , a scalar precision parameter h , and a $T \times 1$ vector \tilde{y} of latent outcomes.

2.10.3 Known Quantities

Prior Parameters

There is a $k \times 1$ coefficient mean vector $\underline{\beta}$, a $k \times k$ positive definite coefficient matrix \underline{H}_β , a precision degrees of freedom parameter $\underline{\nu}$, and a positive definite precision inverse scale parameter \underline{S} .

Data

There is a $T \times 1$ vector of observations of dependent variables y taking values in $\{0, 1\}$.

There is a $T \times k$ matrix X of observations of ancillary (with respect to unknown quantities) variables.

2.10.4 Data Generating Process

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.10.5 Priors

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.10.6 Creating a Model Instance

The mnemonic label identifying the model is `udcht`.

Supply the names you wish to give the unknown quantities in the following order: first the name of β (“beta” for example), then the name of h (“hHomo” for example), and finally the name of the latent variable \tilde{y} (“yTilde” for example).

Supply the known quantities in the following order: $\underline{\beta}$, \underline{H}_β , \underline{S} , $\underline{\nu}$, X , y .

2.10.7 Sampling Algorithms

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html> .

2.11 The Censored Linear Model (with normally distributed disturbances)

2.11.1 Dimension parameters

There are k covariate coefficients and T observations of each variable.

2.11.2 Unknown Quantities

Unknown Parameters

There is a $k \times 1$ coefficient parameter β , a scalar precision parameter h , and a $T \times 1$ vector \tilde{y} of (possibly) latent outcomes.

2.11.3 Known Quantities

Prior Parameters

There is a $k \times 1$ coefficient mean vector $\underline{\beta}$, a $k \times k$ positive definite coefficient matrix \underline{H}_β , a precision degrees of freedom parameter $\underline{\nu}$, a positive definite precision inverse scale parameter \underline{S} , and a censoring parameter \underline{c} .

Data

There is a $T \times 1$ vector of observations of dependent variables y .

There is a $T \times k$ matrix X of observations of ancillary (with respect to unknown quantities) variables.

2.11.4 Data Generating Process

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.11.5 Priors

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.11.6 Creating a Model Instance

The mnemonic label identifying the model is **ucensor**.

Supply the names you wish to give the unknown quantities in the following order: first the name of β (“beta” for example), then the name of h (“hHomo” for example), and finally the name of the latent variable \tilde{y} (“yTilde” for example).

Supply the known quantities in the following order: $\underline{\beta}$, \underline{H}_β , \underline{S} , $\underline{\nu}$, \underline{c} , X , y .

2.11.7 Sampling Algorithms

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.12 The Univariate Latent Linear Model (with normally distributed disturbances)

2.12.1 Dimension parameters

There are k covariate coefficients and T observations of each variable.

2.12.2 Unknown Quantities

Unknown Parameters

There is a $k \times 1$ coefficient parameter β , a scalar precision parameter h , and a $T \times 1$ vector \tilde{y} of (possibly) latent outcomes.

2.12.3 Known Quantities

Prior Parameters

There is a $k \times 1$ coefficient mean vector $\underline{\beta}$, a $k \times k$ positive definite coefficient matrix \underline{H}_β , a precision degrees of freedom parameter $\underline{\nu}$, and a positive definite precision inverse scale parameter \underline{S} .

Data

Corresponding to the (possibly) latent outcome \tilde{y} , there are two $T \times 1$ vectors c and d , $c \geq d$, which describe the observed, set-valued outcome.

There is a $T \times k$ matrix X of observations of ancillary (with respect to unknown quantities) variables.

2.12.4 Data Generating Process

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html> .

2.12.5 Priors

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html> .

2.12.6 Creating a Model Instance

The mnemonic label identifying the model is `ullm`.

Supply the names you wish to give the unknown quantities in the following order: first the name of β (“beta” for example), then the name of h (“hHomo” for example), and finally the name of the latent variable \tilde{y} (“yTilde” for example).

Supply the known quantities in the following order: $\underline{\beta}$, \underline{H}_β , \underline{S} , $\underline{\nu}$, X , c , d .

2.12.7 Sampling Algorithms

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html> .

2.13 A Univariate Linear Model with Student t Disturbances

The mnemonic label identifying the model is `t_ultm`.

Dimension Parameters

T number of observations

K number of covariates

Unknown Quantities

β ($K \times 1$) vector of covariate coefficients

h (1×1) precision of Student t distribution

\tilde{h} ($T \times 1$) time varying latent precision variable

λ (1×1) degrees of freedom of Student t distribution

Known Quantities

$\underline{\beta}$ ($K \times 1$) prior mean of β

\underline{H}_β ($K \times K$) prior precision of β

\underline{s}^2 (1×1) prior inverse scale of h

$\underline{\nu}$ (1×1) prior degrees of freedom of h

$\underline{\lambda}$ (1×1) prior mean of λ

X ($T \times K$) covariates

y ($T \times 1$) dependant variable

Data Generating Process

The observables y are given by

$$y = X\beta + u,$$

where u is a $T \times 1$ vector of independant Student t disturbances, with $u_t|h, X \sim t(0, h^{-1}, \lambda)$. Conditioning on the latent \tilde{h} gives $u_t|h, \tilde{h} \sim N(0, (h\tilde{h}_t)^{-1})$:

$$p(u|X, \beta, \tilde{h}, h) = (2\pi)^{-T/2} h^{T/2} \prod_{t=1}^T \tilde{h}_t^{1/2} \exp(-h\tilde{h}_t u_t^2/2).$$

See Section 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html> for details.

Prior Distribution

The vectors β , h , and (\tilde{h}, λ) , together with X , are mutually independent. The covariate coefficient vector β has distribution $N(\underline{\beta}, \underline{H}_\beta)$:

$$p(\beta) = (2\pi)^{-K/2} |\underline{H}_\beta|^{1/2} \exp[-(\beta - \underline{\beta})' \underline{H}_\beta (\beta - \underline{\beta})/2].$$

The precision parameter h has a scaled chi-squared distribution, with $\underline{s}^2 h \sim \underline{\nu}$:

$$p(h) = 2^{-\underline{\nu}/2} \Gamma(\underline{\nu}/2)^{-1} (\underline{s}^2)^{\underline{\nu}/2} h^{(\underline{\nu}-2)/2} \exp(-\underline{s}^2 h/2).$$

The time varying latent precision parameters \tilde{h} are i.i.d. scaled chi-squared variates, with $\lambda \tilde{h}_t \sim \chi^2(\lambda)$:

$$p(\tilde{h}|\lambda) = [2^{\lambda/2} \Gamma(\lambda/2)]^{-T} \lambda^{T\lambda/2} \prod_{t=1}^T \tilde{h}_t^{(\lambda-2)/2} \exp(-\lambda \tilde{h}_t/2)$$

The degrees of freedom parameter λ is distributed $\exp(\underline{\lambda})$:

$$p(\lambda) = \underline{\lambda}^{-1} \exp(-\lambda/\underline{\lambda}).$$

Sampling Algorithm

See Section 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.14 The Dichotomous Choice Model (with Student t distributed disturbances)

2.14.1 Dimension parameters

There are k covariate coefficients and T observations of each variable.

2.14.2 Unknown Quantities

Unknown Parameters

There is a $k \times 1$ coefficient parameter β , a scalar precision parameter h , a $T \times 1$ vector of precision parameters h_t , a $T \times 1$ vector \tilde{y} of latent outcomes, and a scalar degrees of freedom parameter λ .

2.14.3 Known Quantities

Prior Parameters

There is a $k \times 1$ coefficient mean vector $\underline{\beta}$, a $k \times k$ positive definite coefficient matrix \underline{H}_β , a precision degrees of freedom parameter $\underline{\nu}$, a positive definite precision inverse scale parameter \underline{S} , and a degrees of freedom parameter $\underline{\lambda}$.

Data

There is a $T \times 1$ vector of observations of dependent variables y taking values in $\{0, 1\}$.

There is a $T \times k$ matrix X of observations of ancillary (with respect to unknown quantities) variables.

2.14.4 Data Generating Process

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.14.5 Priors

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.14.6 Creating a Model Instance

The mnemonic label identifying the model is `t_udcht`.

Supply the names you wish to give the unknown quantities in the following order: first the name of β (“beta” for example), then the name of h (“hHomo” for example), then the name of h_t (“hHetero” for example), then the name of the latent variable \tilde{y} (“yTilde” for example), and finally the name of the degrees of freedom parameter λ (“lambda” for example).

Supply the known quantities in the following order: $\underline{\beta}$, \underline{H}_β , \underline{S} , $\underline{\nu}$, $\underline{\lambda}$, X , y .

2.14.7 Sampling Algorithms

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.15 The Censored Linear Model (with Student t distributed disturbances)

2.15.1 Dimension parameters

There are k covariate coefficients and T observations of each variable.

2.15.2 Unknown Quantities

Unknown Parameters

There is a $k \times 1$ coefficient parameter β , a scalar precision parameter h , a $T \times 1$ vector of precision parameters h_t , a $T \times 1$ vector \tilde{y} of (possibly) latent outcomes, and a scalar degrees of freedom parameter λ .

2.15.3 Known Quantities

Prior Parameters

There is a $k \times 1$ coefficient mean vector $\underline{\beta}$, a $k \times k$ positive definite coefficient matrix \underline{H}_β , a precision degrees of freedom parameter $\underline{\nu}$, a positive definite precision inverse scale parameter \underline{S} , a degrees of freedom parameter $\underline{\lambda}$, and a censoring parameter \underline{c} .

Data

There is a $T \times 1$ vector of observations of dependent variables y .

There is a $T \times k$ matrix X of observations of ancillary (with respect to unknown quantities) variables.

2.15.4 Data Generating Process

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.15.5 Priors

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.15.6 Creating a Model Instance

The mnemonic label identifying the model is `t_ucensor`.

Supply the names you wish to give the unknown quantities in the following order: first the name of β (“beta” for example), then the name of h (“hHomo” for example), then the name of h_t (“hHetero” for example), then the name of the latent variable \tilde{y} (“yTilde” for example), and finally the name of the degrees of freedom parameter λ (“lambda” for example).

Supply the known quantities in the following order: $\underline{\beta}$, \underline{H}_β , \underline{S} , $\underline{\nu}$, $\underline{\lambda}$, \underline{c} , X , y .

2.15.7 Sampling Algorithms

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.16 The Univariate Latent Linear Model (with Student t distributed disturbances)

2.16.1 Dimension parameters

There are k covariate coefficients and T observations of each variable.

2.16.2 Unknown Quantities

Unknown Parameters

There is a $k \times 1$ coefficient parameter β , a scalar precision parameter h , a $T \times 1$ vector of precision parameters h_t , a $T \times 1$ vector \tilde{y} of (possibly) latent outcomes, and a scalar degrees of freedom parameter λ .

2.16.3 Known Quantities

Prior Parameters

There is a $k \times 1$ coefficient mean vector $\underline{\beta}$, a $k \times k$ positive definite coefficient matrix \underline{H}_β , a precision degrees of freedom parameter $\underline{\nu}$, a positive definite precision inverse scale parameter \underline{S} , and a degrees of freedom parameter $\underline{\lambda}$.

Data

Corresponding to the (possibly) latent outcome \tilde{y} , there are two $T \times 1$ vectors c and d , $c \geq d$, which describe the observed, set-valued outcome.

There is a $T \times k$ matrix X of observations of ancillary (with respect to unknown quantities) variables.

2.16.4 Data Generating Process

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.16.5 Priors

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.16.6 Creating a Model Instance

The mnemonic label identifying the model is `t_ullm`.

Supply the names you wish to give the unknown quantities in the following order: first the name of β (“beta” for example), then the name of h (“hHomo” for example), then the name of h_t (“hHetero” for example), then the name of the latent variable \tilde{y} (“yTilde” for example), and finally the name of the degrees of freedom parameter λ (“lambda” for example).

Supply the known quantities in the following order: $\underline{\beta}$, \underline{H}_β , \underline{S} , $\underline{\nu}$, $\underline{\lambda}$, X , c , d .

2.16.7 Sampling Algorithms

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.17 A Univariate Linear Model with Finite Mixtures of Normals Disturbances

The mnemonic label identifying the model is `fmm_ulm`.

Dimension Parameters

T	number of observations
K	number of covariates
m	number of mixture components (or states)

Unknown Quantities

γ	$((m + K) \times 1)$	vector of state means and covariate coefficients
h	(1×1)	constant multiplicative precision component
\mathbf{h}	$(m \times 1)$	state dependant multiplicative precision component
\tilde{s}	$(T \times 1)$	time varying latent discrete state
π	$(1 \times m)$	state probabilities

Known Quantities

\underline{h}_α	(1×1)	prior precision parameter for state means
$\underline{\beta}$	$(K \times 1)$	prior mean of covariate coefficients
\underline{H}_β	$(K \times K)$	prior precision of covariate coefficients
\underline{s}^2	(1×1)	prior inverse scale of h
$\underline{\nu}$	(1×1)	prior degrees of freedom of h
m	(1×1)	number of states
$\underline{\nu}$	(1×1)	degrees of freedom parameter for state precisions
r	(1×1)	Dirichlet parameter for state probabilities
X	$(T \times K)$	covariates
y	$(T \times 1)$	dependant variable

Data Generating Process

The observables y are given by

$$y = X\beta + u,$$

where u is a $T \times 1$ vector of independant discrete normal mixture disturbances, with $u_t|h, \pi, \alpha, \mathbf{h}, X$ given by:

$$p(u_t|h, \pi, \alpha, \mathbf{h}, X) = (2\pi)^{-1/2} h^{1/2} \sum_{j=1}^m \pi_j h_j^{1/2} \exp[-h \cdot h_j (u_t - \alpha_j)^2 / 2]$$

Conditioning on the latent states gives

$$p(u_t|h, \alpha, \mathbf{h}, X) = (2\pi)^{-1/2} h^{1/2} h_{\bar{s}_t}^{1/2} \exp[-h \cdot h_{\bar{s}_t} (u_t - \alpha_{\bar{s}_t})^2 / 2].$$

See Section 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html> for details.

Prior Distribution

The vectors (γ, h) , \mathbf{h} , and (\bar{s}, π) , together with X , are mutually independant. The γ parameter vertically stacks the parameters α and β , where α is the $m \times 1$ vector of state dependant means and β is the $K \times 1$ vector of covariate coefficients. They are independant, with $\alpha|h \sim N(0, (\underline{h}_\alpha h)^{-1})$ and $\beta \sim N(\underline{\beta}, \underline{H}_\beta)$:

$$\begin{aligned} p(\gamma|h) = p(\alpha|h) \cdot p(\beta) &= (2\pi)^{-m/2} (\underline{h}_\alpha h)^{m/2} \exp(-\underline{h}_\alpha h \alpha' \alpha / 2) \\ &\cdot (2\pi)^{-K/2} |\underline{H}_\beta|^{1/2} \exp[-(\beta - \underline{\beta})' \underline{H}_\beta (\beta - \underline{\beta}) / 2]. \end{aligned}$$

The precision parameter h has a scaled chi-squared distribution, with $\underline{s}^2 h \sim \underline{\nu}$:

$$p(h) = 2^{-\underline{\nu}/2} \Gamma(\underline{\nu}/2)^{-1} (\underline{s}^2)^{\underline{\nu}/2} h^{(\underline{\nu}-2)/2} \exp(-\underline{s}^2 h / 2).$$

The state dependant precisions h_j are i.i.d., with $\underline{\nu} \cdot h_j \sim \chi^2(\underline{\nu})$:

$$p(\mathbf{h}) = 2^{m\underline{\nu}} \Gamma(\underline{\nu}/2) (\underline{\nu}^2)^{m\underline{\nu}/2} \prod_{j=1}^m h_j^{(-\underline{\nu}-2)/2} \exp(-\underline{\nu} \cdot h_j / 2).$$

The latent states are i.i.d., with the probability $\Pr[s_t = j]$ given by π_j , for $j = 1, \dots, m$:

$$p(\bar{s}|\pi) = \prod_{t=1}^T \pi_{\bar{s}_t}$$

The vector π of probabilities is distributed Dirichlet(r, \dots, r):

$$p(\pi) = \Gamma(mr) \Gamma(r)^{-m} \prod_{j=1}^m \pi_j^{r-1}.$$

Sampling Algorithm

See section 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.18 The Dichotomous Choice Model (with a scale mixture of normals distribution for the disturbances)

2.18.1 Dimension parameters

There are k covariate coefficients, T observations of each variable, and m components for the mixture of normals (i.e., m states).

2.18.2 Unknown Quantities

Unknown Parameters

There is a $k \times 1$ coefficient parameter γ , a scalar precision parameter h , a $T \times 1$ vector of state indices, a $1 \times m$ vector of probabilities, an $m \times 1$ vector of precision parameters h_j , and a $T \times 1$ vector \tilde{y} of latent outcomes.

2.18.3 Known Quantities

Prior Parameters

There is a $k \times 1$ coefficient mean vector $\underline{\beta}$, a positive scalar precision parameter \underline{H}_α , a $k \times k$ positive definite coefficient matrix \underline{H}_β , a precision degrees of freedom parameter $\underline{\nu}$, a positive definite precision inverse scale parameter \underline{S} , an $m \times 1$ vector of precision degrees of freedom parameters $\underline{\nu}_j$, an $m \times 1$ vector of positive definite precision inverse scale parameter \underline{S}_j , and a $1 \times m$ vector of hyperparameters $\underline{\rho}$.

Data

There is a $T \times 1$ vector of observations of dependent variables y taking values in $\{0, 1\}$.

There is a $T \times k$ matrix X of observations of ancillary (with respect to unknown quantities) variables.

2.18.4 Data Generating Process

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.18.5 Priors

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.18.6 Creating a Model Instance

The mnemonic label identifying the model is `fmn.udcht`.

Supply the names you wish to give the unknown quantities in the following order: first the name of γ (“gamma” for example), then the name of s , then the name of p , then the name of h_j (“hState” for example), and finally the name of the latent outcome variable \tilde{y} (“yTilde” for example).

Supply the known quantities in the following order: $\underline{\beta}$, \underline{H}_α , \underline{H}_β , \underline{S} , $\underline{\nu}$, \underline{S}_j , $\underline{\nu}_j$, \underline{r} , X , y .

2.18.7 Sampling Algorithms

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.19 The Censored Linear Model (with a scale mixture of normals distribution for the disturbances)

2.19.1 Dimension parameters

There are k covariate coefficients, T observations of each variable, and m components for the mixture of normals (i.e., m states).

2.19.2 Unknown Quantities

Unknown Parameters

There is a $k \times 1$ coefficient parameter γ , a scalar precision parameter h , a $T \times 1$ vector of state indices, a $1 \times m$ vector of probabilities, an $m \times 1$ vector of precision parameters h_j , and a $T \times 1$ vector \tilde{y} of (possibly) latent outcomes.

2.19.3 Known Quantities

Prior Parameters

There is a $k \times 1$ coefficient mean vector $\underline{\beta}$, a positive scalar precision parameter \underline{H}_α , a $k \times k$ positive definite coefficient matrix \underline{H}_β , a precision degrees of freedom parameter $\underline{\nu}$, a positive definite precision inverse scale parameter \underline{S} , an $m \times 1$ vector of precision degrees of freedom parameters $\underline{\nu}_j$, an $m \times 1$ vector of positive definite precision inverse scale parameter \underline{S}_j , a $1 \times m$ vector of hyperparameters $\underline{\tau}$, and a censoring parameter \underline{c} .

Data

There is a $T \times 1$ vector of observations of dependent variables y .

There is a $T \times k$ matrix X of observations of ancillary (with respect to unknown quantities) variables.

2.19.4 Data Generating Process

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.19.5 Priors

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.19.6 Creating a Model Instance

The mnemonic label identifying the model is `fmn.ucensor`.

Supply the names you wish to give the unknown quantities in the following order: first the name of γ (“gamma” for example), then the name of s , then the name of p , then the name of h_j (“hState” for example), and finally the name of the outcome variable \tilde{y} (“yTilde” for example).

Supply the known quantities in the following order: $\underline{\beta}$, \underline{H}_α , \underline{H}_β , \underline{S} , $\underline{\nu}$, \underline{S}_j , $\underline{\nu}_j$, \underline{r} , \underline{c} , X , y .

2.19.7 Sampling Algorithms

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.20 The Univariate Latent Linear Model (with a scale mixture of normals distribution for the disturbances)

2.20.1 Dimension parameters

There are k covariate coefficients, T observations of each variable, and m components for the mixture of normals (i.e., m states).

2.20.2 Unknown Quantities

Unknown Parameters

There is a $(m+k) \times 1$ coefficient parameter γ , a scalar precision parameter h , a $T \times 1$ vector of state indices, a $1 \times m$ vector of probabilities, an $m \times 1$ vector of precision parameters h_j , and a $T \times 1$ vector \tilde{y} of (possibly) latent outcomes.

2.20.3 Known Quantities

Prior Parameters

There is a $k \times 1$ coefficient mean vector $\underline{\beta}$, a positive scalar precision parameter \underline{H}_α , a $k \times k$ positive definite coefficient matrix \underline{H}_β , a precision degrees of freedom parameter $\underline{\nu}$, a positive definite precision inverse scale parameter \underline{S} , an $m \times 1$ vector of precision degrees of freedom parameters $\underline{\nu}_j$, an $m \times 1$ vector of positive definite precision inverse scale parameter \underline{S}_j , and a $1 \times m$ vector of hyperparameters \underline{x} .

Data

Corresponding to the (possibly) latent outcome \tilde{y} , there are two $T \times 1$ vectors c and d , $c \geq d$, which describe the observed, set-valued outcome.

There is a $T \times k$ matrix X of observations of ancillary (with respect to unknown quantities) variables.

2.20.4 Data Generating Process

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.20.5 Priors

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.20.6 Creating a Model Instance

The mnemonic label identifying the model is `fmn.ullm`.

Supply the names you wish to give the unknown quantities in the following order: first the name of γ (“gamma” for example), then the name of s , then the name of p , then the name of h_j (“hState” for example), and finally the name of the outcome variable \tilde{y} (“yTilde” for example).

Supply the known quantities in the following order: $\underline{\gamma}$, \underline{H}_α , \underline{H}_β , \underline{S} , $\underline{\nu}$, \underline{S}_j , $\underline{\nu}_j$, $\underline{\tau}$, X , c , d .

2.20.7 Sampling Algorithms

See Sections 4.5 and 4.8 in “Contemporary Bayesian Econometrics and Statistics,” by John Geweke, at <http://www.cirano.qc.ca/~bacc/bacc2003/resources.html>

2.21 An Autoregression Model

2.21.1 Dimension Parameters

The model features the following dimension parameters.

Dimension Parameter	Description
T	number of observations
K	number of covariates
p	autoregressive order

2.21.2 Unknown Quantities

The model features the following unknown quantities.

Unknown Quantity	Dimensions	Description
β	$K \times 1$	covariate coefficient vector
h	1×1	residual precision
ϕ	$p \times 1$	vector of autoregression coefficients

2.21.3 Known Quantities

The model features the following known quantities.

Known Quantity	Dimensions	Description
$\bar{\beta}$	$K \times 1$	prior mean of β
\bar{H}_β	$K \times K$	prior precision of β
$\bar{\nu}$	1×1	prior degrees of freedom of h
\bar{s}^2	1×1	prior inverse scale of h
$\bar{\phi}$	$p \times 1$	prior mean of ϕ before truncation
\bar{H}_ϕ	$p \times p$	prior precision of ϕ before truncation
X	$T \times K$	covariates
y	$T \times 1$	dependant variable

2.21.4 Data Generating Process

The data generating process is given by

$$y_t = \beta' x_t + \epsilon_t$$

where x_t is the t 'th row of X , as a column vector,

$$\epsilon_t = \sum_{i=1}^p \phi_i \epsilon_{t-i} + u_t,$$

and

$$u_t \sim \text{i.i.d. } N(0, h^{-1})$$

2.21.5 Prior Distribution

The unknowns are a-priori independent and have the following distributions.

$$\beta \sim N(\bar{\beta}, \bar{H}_\beta^{-1})$$

$$\bar{s}^2 h \sim \chi^2(\bar{\nu})$$

The prior for ϕ is obtained by truncating the following density to the region for which y is stationary.

$$\phi \sim N(\bar{\phi}, \bar{H}_\phi^{-1})$$

2.21.6 Creating a Model Instance

The mnemonic label identifying the model is **AR**.

Supply the names you wish to give the unknown quantities in the same order as they appear in the table of unknown quantities. Supply the known quantities in the same order as they appear in the table of known quantities.

2.21.7 Sampling Algorithm

The sampling algorithm for prior simulation features three blocks, each making independent draws from the prior distribution of one of the unknown quantities. The sampling algorithm for posterior simulation features three blocks, each making draws from the conditional posterior distribution of one of the unknown quantities.

2.22 An Autoregression Model with State Dependant Means

2.22.1 Dimension Parameters

The model features the following dimension parameters.

Dimension Parameter	Description
T	number of observations
K	number of covariates
p	autoregressive order
m	number of states

2.22.2 Unknown Quantities

The model features the following unknown quantities.

Unknown Quantity	Dimensions	Description
γ	$(m + K) \times 1$	vertical stack of alpha and beta
h	1×1	residual precision
ϕ	$p \times 1$	vector of autoregression coefficients
P	$m \times m$	state transition probability matrix
s	$T \times 1$	latent states
f	$T \times m$	filter probabilities

2.22.3 Known Quantities

The model features the following known quantities.

Known Quantity	Dimensions	Description
$\bar{\gamma}$	$(m + K) \times 1$	prior mean of γ
\bar{H}_γ	$(m + K) \times (m + K)$	prior precision of γ
$\bar{\nu}$	1×1	prior degrees of freedom of h
\bar{s}^2	1×1	prior inverse scale of h
$\bar{\phi}$	$p \times 1$	prior mean of ϕ before truncation
\bar{H}_ϕ	$p \times p$	prior precision of ϕ before truncation
\bar{A}	$m \times m$	parameters of prior for P
X	$T \times K$	covariates
y	$T \times 1$	dependant variable

2.22.4 Data Generating Process

The data generating process is given by

$$y_t = \alpha_{s_t} + \beta' x_t + \epsilon_t$$

where x_t is the t 'th row of X , as a column vector, and α ($m \times 1$) and β ($K \times 1$) are obtained by partitioning γ ,

$$\epsilon_t = \sum_{i=1}^p \phi_i \epsilon_{t-i} + u_t,$$

and

$$u_t \sim \text{i.i.d. } N(0, h^{-1})$$

2.22.5 Prior Distribution

The unknowns are a-priori independent and have the following distributions.

$$\begin{aligned} \gamma &\sim N(\bar{\gamma}, \bar{H}_\gamma^{-1}) \\ \bar{s}^2 h &\sim \chi^2(\bar{\nu}) \end{aligned}$$

The prior for ϕ is obtained by truncating the following density to the region for which y is stationary.

$$\begin{aligned} \phi &\sim N(\bar{\phi}, \bar{H}_\phi^{-1}) \\ (P_{i1}, \dots, P_{im}) &\sim \text{i.i.d. Di}(\bar{A}_{i1}, \dots, \bar{A}_{im}) \\ \Pr[s_t = j | s_{t-1} = i] &= P_{ij} \end{aligned}$$

The unknown quantity f gives, for each observation time t , the state probabilities at t given previous states, previous values of the observed variables, and the other unknown quantities. It is not a primitive unknown quantity, and it is included to give the user access to filtered probabilities.

2.22.6 Creating a Model Instance

The mnemonic label identifying the model is `Hamilton`.

Supply the names you wish to give the unknown quantities in the same order as they appear in the table of unknown quantities. Supply the known quantities in the same order as they appear in the table of known quantities.

2.22.7 Sampling Algorithm

The sampling algorithm for prior simulation features five blocks. Four blocks make independent draws from the prior distributions of γ , h , ϕ and P . The fifth makes draws from the distribution $s|P$. The sampling algorithm for posterior simulation features five blocks, each making draws from the conditional posterior distribution of one of the unknown quantities.

Chapter 3

BACC Commands

3.1 Overview of BACC Commands

The following is a list of BACC commands with brief descriptions.

<code>dirichletSim</code>	Generates a sample from a multiple Dirichlet distribution.
<code>expect1</code>	Calculates, for a weighted random sample, the sample mean and standard deviation, estimates of the numerical standard error for the mean, and estimates of the relative numerical efficiency.
<code>expectN</code>	Calculates combined sample means, with numerical standard errors, for a set of different weighted random samples, and tests for the equality of their individual population means.
<code>extract</code>	Returns simulation matrices for a model instance.
<code>gammaSim</code>	Generates a sample from a gamma distribution.
<code>gaussianSim</code>	Generates a sample from a Gaussian distribution.
<code>listModelSpecs</code>	Lists all available model specifications (e.g. <code>nlm</code> , <code>poisson</code>).
<code>listModels</code>	Lists all open model instances.
<code>miDelete</code>	Closes without saving a (or all) model instances.
<code>miLoad</code>	Loads a model instance stored in a binary file.
<code>miLoadAscii</code>	Loads a model instance stored in a text file.
<code>miSave</code>	Saves a model instance in a binary file.
<code>miSaveAscii</code>	Saves a model instance in a text file.
<code>minst</code>	Creates an instance of a particular model specification.

<code>mlike</code>	Computes various estimates of the marginal likelihood for a model instance, with numerical standard errors.
<code>paretoSim</code>	Generates a sample from a Pareto distribution.
<code>postfilter</code>	Filters out previously generated draws from the posterior simulation matrix of a given model instance.
<code>postsim</code>	Generates or appends to the posterior simulation matrix of a given model instance.
<code>postsimHM</code>	Generates or appends to the posterior HM simulation matrix of a given model instance.
<code>priorRobust</code>	Calculates upper and lower bounds on the mean of a posterior function of interest, as the prior distribution is varied from its original specification.
<code>priorfilter</code>	Filters out previously generated draws from the prior simulation matrix of a given model instance.
<code>priorsim</code>	Generates or appends to the prior simulation matrix of a given model instance.
<code>setseedconstant</code>	Sets the seeds of the random number generators to a constant value.
<code>setseedtime</code>	Sets the seeds of the random number generators to the number of seconds since the beginning of 1970.
<code>weightedSmooth</code>	Estimates a univariate density function for a weighted random sample, using a kernel smoothing algorithm adapted to weighted samples.
<code>wishartSim</code>	Generates a sample from a Wishart distribution.

3.2 Command Line Issues

3.2.1 Matrix Files

Each input and output matrix is stored in its own file, even if it is a scalar. A file containing a matrix must be an ASCII text file, have a `.mtx` filename extension, and contain only floating point numbers. Elements in a row must be separated by tabs and/or spaces. Rows are separated by carriage returns.

Output matrices are furthermore guaranteed to have row elements separated only by single tab characters. This allows easy loading into spreadsheets and other software supporting the tab-separated column standard.

3.2.2 Filter Files

A filter file must be an ASCII text file, have a `.ftr` filename extension, and contain only positive integers. Filter elements are separated by carriage returns.

3.2.3 Batch Files

BACC commands entered at the command line without arguments will print a statement about how to use the command. Commands can be entered by hand at the command line, or a batch file can be created which contains multiple commands.

3.3 Command Line Issues

3.3.1 Matrix Files

Each input and output matrix is stored in its own file, even if it is a scalar. A file containing a matrix must be an ASCII text file, have a `.mtx` filename extension, and contain only floating point numbers. Elements in a row must be separated by tabs and/or spaces. Rows are separated by carriage returns.

Output matrices are furthermore guaranteed to have row elements separated only by single tab characters. This allows easy loading into spreadsheets and other software supporting the tab-separated column standard.

3.3.2 Filter Files

A filter file must be an ASCII text file, have a `.ftr` filename extension, and contain only positive integers. Filter elements are separated by carriage returns.

3.3.3 Batch Files

BACC commands entered at the command line without arguments will print a statement about how to use the command. Commands can be entered by hand at the command line, or a batch file can be created which contains multiple commands.

3.4 Detailed Description of Commands

Each BACC command is described in detail in one of the following sections.

3.4.1 The `dirichletSim` Command

Description

Generates a sample from a multiple Dirichlet distribution.

Usage

```
dirichletSim A.mtx n.mtx
```

Inputs

<code>A</code>	<code>m</code> by <code>K</code> matrix: Dirichlet parameters
<code>n</code>	Integer: number of draws to generate

Outputs

<code>sample</code>	<code>n</code> by <code>nK</code> matrix: sample generated from multiple Dirichlet distribution
---------------------	---

See Also

`paretoSim`, `gaussianSim`, `gammaSim`, `wishartSim`.

Example

```
echo -e "1.0 2.0 3.0  
n4.0 5.0 6.0" > A.mtx  
  
dirichletSim A.mtx 1000.mtx 1000.mtx
```

Details

The sample consists of n draws. Each of the n draws of the sample is an m by K matrix with independent rows. Each row has a Dirichlet distribution with parameters given by the corresponding row of A .

The result is given as a n by mK matrix, and each column gives a draw in column major order. See Appendix A for the parameterization of the Dirichlet distribution.

3.4.2 The `expect1` Command

Description

Calculates, for a weighted random sample, the sample mean and standard deviation, estimates of the numerical standard error for the mean, and estimates of the relative numerical efficiency.

Usage

```
expect1 logWeight.mtx sample.mtx
expect1 logWeight.mtx sample.mtx taper.mtx
```

Inputs

<code>logWeight</code>	Vector of length M : log sample weights
<code>sample</code>	Vector of length M : sample of scalar draws
<code>taper</code>	Vector of length K : taper half-widths (optional)

Outputs

<code>mean</code>	Real scalar: weighted sample mean
<code>std</code>	Real scalar: weighted sample standard deviation
<code>nse</code>	Vector of length $K + 1$: estimated numerical standard errors
<code>rne</code>	Vector of length $K + 1$: estimated relative numerical efficiency

See Also

`expectN`, `priorRobust`.

Example

```
# Use default taper values
expect1 lw.mtx z.mtx z.mtx z.mtx z.mtx z.mtx
# Use alternate taper values
echo -e "4.0 8.0" > taper.mtx
expectN lw.mtx z.mtx taper.mtx taper.mtx taper.mtx taper.mtx taper.mtx
```

Details

Let $z = (z_1, \dots, z_M)$ be the sample and $(\log w_1, \dots, \log w_M)$ be the vector of log weights. Let $\lambda = (\lambda_1, \dots, \lambda_K)$ be the vector of half-widths. The sample is broken into T groups of size $J \equiv M \text{ div } T$ and the last $M \bmod T$ elements are ignored. Thus $M_{\text{use}} \equiv JT$ elements are used.

The sample mean and standard deviation are calculated as follows:

$$\bar{z} = \frac{\sum_{m=1}^{M_{use}} w_m z_m}{\sum_{m=1}^{M_{use}} w_m}$$

$$\sigma_z = \left[\frac{\sum_{m=1}^{M_{use}} w_m (z_m - \bar{z})^2}{\sum_{m=1}^{M_{use}} w_m} \right]^{\frac{1}{2}}$$

For the calculation of the first numerical standard error τ_0 , we assume no serial correlation in (z_1, \dots, z_M) . This is appropriate for independence or importance sampling. Following Geweke (1989) [3], this leads to

$$\tau \approx \tau_0 \equiv \left[\frac{\sum_{m=1}^{M_{use}} w_m^2 (z_m - \bar{z})^2}{\left(\sum_{m=1}^{M_{use}} w_m \right)^2} \right]^{\frac{1}{2}}$$

For the calculation of τ_1 through τ_K , the remaining K estimates of the numerical standard error, the following method is used. First, `expect1` calculates group and sample means of the numerator quantity $w_m z_m$ and the denominator quantity w_m :

$$n(t) = \frac{1}{J} \sum_{m=(t-1)J+1}^{tJ} w_m z_m \quad d(t) = \frac{1}{J} \sum_{m=(t-1)J+1}^{tJ} w_m \quad t = 1, \dots, T$$

$$\bar{n} = \frac{1}{M_{use}} \sum_{m=1}^{M_{use}} w_m z_m \quad \bar{d} = \frac{1}{M_{use}} \sum_{m=1}^{M_{use}} w_m$$

Then it calculates the following sample autocorrelation and autocovariance functions:

$$\gamma_{nn}(t) = \frac{1}{T} \sum_{s=t+1}^T (n(s) - \bar{n})(n(s-t) - \bar{n}) \quad t = 0, \dots, T-1$$

$$\gamma_{dd}(t) = \frac{1}{T} \sum_{s=t+1}^T (d(s) - \bar{d})(d(s-t) - \bar{d}) \quad t = 0, \dots, T-1$$

$$\gamma_{nd}(t) = \frac{1}{T} \sum_{s=t+1}^T (n(s) - \bar{n})(d(s-t) - \bar{d}) \quad t = 0, \dots, T-1$$

$$\gamma_{dn}(t) = \frac{1}{T} \sum_{s=t+1}^T (d(s) - \bar{d})(n(s-t) - \bar{n}) \quad t = 0, \dots, T-1$$

Then it calculates, for each $k \in \{1, \dots, K\}$, estimates $\sigma_{n(k)}^2$, $\sigma_{d(k)}^2$, and $\sigma_{nd(k)}$ of $\sigma_n^2 \equiv \text{Var}[n(t)]$, $\sigma_d^2 \equiv \text{Var}[d(t)]$ and $\sigma_{nd} \equiv \text{Cov}[n(t), d(t)]$, based on the taper half-width λ_k :

$$\sigma_{nn(k)}^2 = \gamma_{nn}(0) + 2 \sum_{s=1}^{\lambda_k-1} \frac{\lambda_k - s}{\lambda_k} \gamma_{nn}(s)$$

$$\sigma_{dd(k)}^2 = \gamma_{dd}(0) + 2 \sum_{s=1}^{\lambda_k-1} \frac{\lambda_k - s}{\lambda_k} \gamma_{dd}(s)$$

$$\sigma_{nd(k)}^2 = \gamma_{nd}(0) + \sum_{s=1}^{\lambda_k-1} \frac{\lambda_k - s}{\lambda_k} [\gamma_{nd}(s) + \gamma_{dn}(s)]$$

These calculations are based on conventional time series methods for a wide sense stationary process, described in Geweke (1992) [4].

By the conventional asymptotic expansion, the square of the numerical standard error is approximated by

$$\tau^2 = \text{Var}\left(\frac{n}{d}\right) \approx \begin{bmatrix} \frac{1}{d} & -\frac{1}{d^2} \end{bmatrix} \begin{bmatrix} \sigma_n^2 & \sigma_{nd} \\ \sigma_{nd} & \sigma_d^2 \end{bmatrix} \begin{bmatrix} \frac{1}{d} \\ -\frac{1}{d^2} \end{bmatrix}$$

For each $k \in \{1, \dots, K\}$ it calculates the approximation τ_k using $\sigma_{nn(k)}^2$, $\sigma_{dd(k)}^2$ and $\sigma_{nd(k)}^2$ defined above.

Relative numerical efficiencies (ν_0, \dots, ν_K) are calculated using

$$\nu_k \equiv \left(\frac{\tau_0}{\tau_k}\right)^2 \quad k = 0, \dots, K$$

3.4.3 The expectN Command

Description

Calculates combined sample means, with numerical standard errors, for a set of different weighted random samples, and tests for the equality of their individual population means.

Usage

```
expectN logWeight1.mtx sample1.mtx logWeight2.mtx sample2.mtx
expectN logWeight1.mtx sample1.mtx logWeight2.mtx sample2.mtx taper.mtx
```

Inputs

logWeight1	Vector of length M_1 : log sample weights for first sample
sample1	Vector of length M_1 : first sample of scalar draws
logWeight2	Vector of length M_2 : log sample weights for second sample
sample2	Vector of length M_2 : second sample of scalar draws
taper	Vector of length K : taper half-widths (optional)

Outputs

mean	Vector of length $K + 1$: estimated combined weighted sample means
nse	Vector of length $K + 1$: estimated numerical standard errors
equal	Vector of length $K + 1$: marginal significance levels for a chi-squared test of the equality of the population means

See Also

expect1.

Example

```
# Use default taper values
expectN lw1.mtx z1.mtx lw2.mtx z2.mtx z2.mtx z2.mtx z2.mtx
# Use alternate taper values
echo -e "4.0 8.0" > taper.mtx
expectN lw1.mtx z1.mtx lw2.mtx z2.mtx taper.mtx taper.mtx taper.mtx
taper.mtx
```

Details

In general, there are N pairs of weighted samples, not just two. For each sample $z^{(i)}$, **expectN** calculates individual sample moments $\bar{z}^{(i)}$ and estimates of numerical standard errors $(\tau_0^{(i)}, \dots, \tau_K^{(i)})$ from the samples $(z_1^{(i)}, \dots, z_{M_i}^{(i)})$, the log weights $(\log z_1^{(i)}, \dots, \log z_{M_i}^{(i)})$, and the half-taper values $(\lambda_1, \dots, \lambda_K)$, in the same way that **expect1** calculates \bar{z} and (τ_0, \dots, τ_K) from (z_1, \dots, z_M) , $(\log z_1, \dots, \log z_M)$, and $(\lambda_1, \dots, \lambda_K)$.

The estimated sample means \bar{z}_k are given by

$$\bar{z}_k = \sum_{i=1}^N \frac{\bar{z}^{(i)}}{\tau_k^{2(i)}} \bigg/ \sum_{i=1}^N \frac{1}{\tau_k^{2(i)}} \quad k = 1, \dots, K$$

The estimated numerical standard errors τ_k are given by

$$\frac{1}{\tau_k^2} = \sum_{i=1}^N \frac{1}{\tau_k^{2(i)}}$$

For each k , the marginal significance level is the value of p_k such that

$$\begin{bmatrix} \bar{z}_k^{(2)} - \bar{z}_k^{(1)} & \dots & \bar{z}_k^{(N)} - \bar{z}_k^{(N-1)} \end{bmatrix} \cdot \Sigma^{-1} \cdot \begin{bmatrix} \bar{z}_k^{(2)} - \bar{z}_k^{(1)} \\ \bar{z}_k^{(3)} - \bar{z}_k^{(2)} \\ \vdots \\ \bar{z}_k^{(N)} - \bar{z}_k^{(N-1)} \end{bmatrix} = \chi_{1-p_k}^2(N-1)$$

where Σ is the following matrix

$$\begin{bmatrix} \tau_k^{2(1)} + \tau_k^{2(2)} & -\tau_k^{2(2)} & 0 & \dots & 0 & 0 \\ -\tau_k^{2(2)} & \tau_k^{2(2)} + \tau_k^{2(3)} & -\tau_k^{2(3)} & \dots & 0 & 0 \\ 0 & -\tau_k^{2(3)} & \tau_k^{2(3)} + \tau_k^{2(4)} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \tau_k^{2(N-2)} + \tau_k^{2(N-1)} & -\tau_k^{2(N-1)} \\ 0 & 0 & 0 & \dots & -\tau_k^{2(N-1)} & \tau_k^{2(N-2)} + \tau_k^{2(N-1)} \end{bmatrix}$$

3.4.4 The extract Command

Description

Returns simulation matrices for a model instance.

Usage

```
extract modelInst.mtx
```

Inputs

`modelInst` Integer: model instance identifier.

Outputs

`sim` Structure: simulation matrices

Example

```
extract mi.mtx mi.mtx
```

Details

The return value is a structure (named list in S-PLUS and R) with the following fields (components in S-PLUS and R):

id Model instance identifier.

logWeightPost Log weights for posterior draws.

logPrior Value of log prior for posterior draws.

logXPrior Value of transformed log prior values for posterior draws

logLike Value of log likelihood for posterior draws.

logPriorHM Value of log prior for posterior HM draws.

logLikeHM Value of log likelihood for posterior HM draws.

logWeightPrior Log weights for prior draws.

logPriorPrior Value of log prior for prior draws.

logLikePrior Value of log likelihood for prior draws.

* Posterior simulation matrix of unknown quantity named *.

***Prior** Prior simulation matrix of unknown quantity named *.

***HM** Posterior HM simulation matrix of unknown quantity named *.

All simulation matrices have three dimensions. The first two dimensions give the row and column of the unknown quantity. The third dimension is the simulation dimension. Each value of the third index gives a different draw of the unknown quantity.

3.4.5 The gammaSim Command

Description

Generates a sample from a gamma distribution.

Usage

```
gammaSim alpha.mtx beta.mtx n.mtx
```

Inputs

alpha	Real scalar: shape parameter of gamma distribution
beta	Real scalar: inverse scale parameter of gamma distribution
n	Integer: number of draws to generate

Outputs

sample	n by 1 matrix: sample generated from gamma distribution
--------	---

See Also

paretoSim, gaussianSim, dirichletSim, wishartSim.

Example

```
echo 3.0 > alpha.mtx
echo 5.0 > beta.mtx
gammaSim alpha.mtx beta.mtx 1000.mtx 1000.mtx
```

Details

Each of the n draws of the sample is a scalar with a gamma distribution. The result is given as a n by 1 matrix.

See Appendix A for the parametrization of the gamma distribution.

3.4.6 The gaussianSim Command

Description

Generates a sample from a Gaussian distribution.

Usage

```
gaussianSim mean.mtx precision.mtx n.mtx
```

Inputs

mean	Vector of length K : mean of Gaussian distribution
precision	K by K matrix: precision of Gaussian distribution
n	Integer: number of draws to generate

Outputs

sample	n by K matrix: sample generated from Gaussian distribution
--------	--

See Also

paretoSim, dirichletSim, gammaSim, wishartSim.

Example

```
echo -e "1.0  
n2.0" > mean.mtx  
  
echo -e "1.0 0.0  
n0.0 1.0" > precision.mtx  
  
gaussianSim mean.mtx precision.mtx 1000.mtx 1000.mtx
```

Details

Each of the n draws of the sample is a vector of length K with a Gaussian distribution. The result is given as a n by K matrix.

See Appendix A for the parametrization of the Gaussian distribution.

3.4.7 The `listModelSpecs` Command

Description

Lists all available model specifications (e.g. `nlm`, `poisson`).

Usage

```
listModelSpecs
```

Inputs

None.

Outputs

None.

See Also

`minst`, `listModels`.

Example

```
listModelSpecs
```

Details

A printed message gives a list of model specifications.

3.4.8 The `listModels` Command

Description

Lists all open model instances.

Usage

```
listModels
```

Inputs

None.

Outputs

None.

See Also

`minst`, `miDelete`, `listModelSpecs`.

Example

```
listModels
```

Details

A printed message gives the model instance identification number, the name of the model specification (e.g. `nlm`, `poisson`), and the number of prior, posterior, and posterior HM draws.

3.4.9 The miDelete Command

Description

Closes without saving a (or all) model instances.

Usage

```
miDelete modelInst.mtx
```

Inputs

`modelInst` Integer: model instance identifier.

Outputs

None.

See Also

`minst`, `listModels`.

Example

```
miDelete mi.mtx
```

3.4.10 The miLoad Command

Description

Loads a model instance stored in a binary file.

Usage

```
miLoad filename.mtx
```

Inputs

<code>filename</code>	String: name of binary file storing the model instance
-----------------------	--

Outputs

<code>modelInst</code>	Integer: model instance identifier.
------------------------	-------------------------------------

See Also

`miSave`, `minst`, `miLoadAscii`.

Example

```
miLoad "miFile".mtx "miFile".mtx
```

3.4.11 The miLoadAscii Command

Description

Loads a model instance stored in a text file.

Usage

```
miLoadAscii filename.mtx
```

Inputs

filename String: name of text file storing the model instance

Outputs

modelInst Integer: model instance identifier.

See Also

miSaveAscii, minst, miLoad.

Example

```
miLoadAscii "miFile.txt".mtx "miFile.txt".mtx
```

3.4.12 The miSave Command

Description

Saves a model instance in a binary file.

Usage

```
miSave modelInst.mtx filename.mtx
```

Inputs

<code>modelInst</code>	Integer: model instance identifier.
<code>filename</code>	String: name of binary file in which to store the model instance

Outputs

None.

See Also

`miLoad`, `minst`, `miSaveAscii`.

Example

```
miSave mi.mtx "miFile".mtx
```

Details

If the file already exists, it is written over.

3.4.13 The `miSaveAscii` Command

Description

Saves a model instance in a text file.

Usage

```
miSaveAscii modelInst.mtx filename.mtx
```

Inputs

<code>modelInst</code>	Integer: model instance identifier.
<code>filename</code>	String: name of text file in which to store the model instance

Outputs

None.

See Also

`miLoadAscii`, `minst`, `miSave`.

Example

```
miSaveAscii mi.mtx "miFile.txt".mtx
```

Details

If the file already exists, it is written over. The ascii version of a model instance is platform independent and human readable, but long and inefficient.

3.4.14 The minst Command

Description

Creates an instance of a particular model specification.

Usage

```
minst modelSpecName.mtx unknownNames.mtx knowns.mtx
```

Inputs

`modelSpecName` String: name of model specification
`unknownNames` List of strings: user provided names for unknown quantities
`knowns` List of matrices: user provided matrices of known quantities

Outputs

`modelInst` Integer: model instance identifier.

Example

```
echo -e "1 1 2" > a.mtx
echo -e "1 1
n2 3
n3 3" > s.mtx
minst 'iidfs'.mtx 'pi'.mtx a.mtx s.mtx s.mtx
```

Details

The available model specifications are described in Chapter 2. For each model specification, there is a section “Creating a Model Instance” with the relevant information, namely

- The name of the model specification.
- The order in which the user specifies the names for the unknown quantities of the model.
- The order in which the user provides the matrices giving the values of the known quantities of the model.

3.4.15 The `mlike` Command

Description

Computes various estimates of the marginal likelihood for a model instance, with numerical standard errors.

Usage

```
mlike modelInst.mtx
mlike modelInst.mtx p.mtx taper.mtx
```

Inputs

<code>modelInst</code>	Integer: model instance identifier.
<code>p</code>	Vector of length L : truncation parameters (optional)
<code>taper</code>	Vector of length K : taper half-widths (optional)

Outputs

<code>ml</code>	Vector of length L : marginal likelihood estimates
<code>mlnse</code>	L by $K + 1$ matrix: numerical standard error estimates

See Also

`postsim`, `postsimHM`.

Example

```
# Use default truncation and taper values
mlike mi.mtx mi.mtx mi.mtx mi.mtx mi.mtx
# Use alternate truncation values
echo -e "0.1 0.3 0.5 0.7 0.9" > p.mtx
mlike mi.mtx p.mtx p.mtx p.mtx p.mtx p.mtx
# Use alternate truncation and taper values
echo -e "4.0 8.0" > taper.mtx
mlike mi.mtx p.mtx taper.mtx taper.mtx taper.mtx taper.mtx taper.mtx
```

Details

The method used is a modification described in Geweke [5] of the method proposed in Gelfand and Dey [2].

The truncation parameters $p_l \in [0, 1]$ index the truncated multivariate normal distribution $f(\cdot)$ discussed in Geweke [5]. For each p_l , `mlike` generates (internally) an unweighted vector (z_1^l, \dots, z_M^l) , where M is the number of posterior samples in the given model instance.

For each l , `mlike` calculates the sample mean \bar{z}_l and numerical standard errors $(\tilde{\tau}_1^l, \dots, \tilde{\tau}_M^l)$ from (z_1^l, \dots, z_M^l) and $(\lambda_1, \dots, \lambda_K)$ in the same way that `expect1` calculates (τ_0, \dots, τ_K) from (z_1, \dots, z_M) , a vector of equal log weights, and $(\lambda_1, \dots, \lambda_K)$. Then for all l , the estimate of the log marginal likelihood is given by

$$\mu_l = -\log \bar{z}_l$$

and for all l and k , the estimate of the numerical standard error for the log marginal likelihood is given by

$$\tau_{kl} = \frac{\tilde{\tau}_{kl}}{\bar{z}_l}$$

When numerical standard error is small, results are not sensitive to the choice of p . In these cases $L = 1$ and $p_1 = 0.5$ will suffice. However the additional computational burden of increasing L is negligible. If you are concerned about standard errors, it is best to use several values of p_l , for example, $p = (0.1, 0.2, \dots, 0.9)$.

3.4.16 The `paretoSim` Command

Description

Generates a sample from a Pareto distribution.

Usage

```
paretoSim alpha.mtx beta.mtx n.mtx
```

Inputs

<code>alpha</code>	Real scalar: tail parameter of Pareto distribution
<code>beta</code>	Real scalar: location parameter of Pareto distribution
<code>n</code>	Integer: number of draws to generate

Outputs

<code>sample</code>	n by 1 matrix: sample generated from Pareto distribution
---------------------	--

See Also

`dirichletSim`, `gaussianSim`, `gammaSim`, `wishartSim`.

Example

```
echo 1.0 > alpha.mtx
echo 4.0 > beta.mtx
gammaSim alpha.mtx beta.mtx 1000.mtx 1000.mtx
```

Details

Each of the n draws of the sample is a scalar with a pareto distribution. The result is given as a n by 1 matrix.

See Appendix A for the parametrization of the Pareto distribution.

3.4.17 The postfilter Command

Description

Filters out previously generated draws from the posterior simulation matrix of a given model instance.

Usage

```
postfilter modelInst.mtx filter.mtx
```

Inputs

<code>modelInst</code>	Integer: model instance identifier.
<code>filter</code>	Vector of integers of length n : indices of existing draws to keep

Outputs

None.

Example

```
# Create a file filter.ftr with the lines 101, ... 1000  
  
postfilter mi.mtx filter.mtx
```

Details

The i th draw of the posterior simulation matrix is kept if and only if $i = f_j$ for some j from 1 to n .

3.4.18 The `postsim` Command

Description

Generates or appends to the posterior simulation matrix of a given model instance.

Usage

```
postsim modelInst.mtx m.mtx n.mtx
```

Inputs

<code>modelInst</code>	Integer: model instance identifier.
<code>m</code>	Integer: number of posterior draws to record
<code>n</code>	Integer: number of posterior draws to generate for each one recorded

Outputs

None.

See Also

`minst`, `postfilter`, `mlike`, `priorsim`, `postsimHM`, `extract`.

Example

```
postsim mi.mtx 1000.mtx 1.mtx
```

Details

Generates draws of unknown quantities from their posterior distribution. Generates mn new posterior draws, and appends every n th draw to the posterior simulation matrix. If there are any draws from a previous invocation of `postsim`, the first new draw comes from the transition kernel of the Markov chain used for posterior simulation. Otherwise, it comes from the initial distribution of the Markov chain.

Use the `extract` command to obtain the posterior draws.

3.4.19 The `postsimHM` Command

Description

Generates or appends to the posterior HM simulation matrix of a given model instance.

Usage

```
postsimHM modelInst.mtx m.mtx n.mtx scalePrecision.mtx
```

Inputs

<code>modelInst</code>	Integer: model instance identifier.
<code>m</code>	Integer: number of posterior draws to record
<code>n</code>	Integer: number of posterior draws to generate for each one recorded
<code>scalePrecision</code>	Real scalar: factor used to rescale the precision matrix of the random walk innovation

Outputs

None.

See Also

`minst`, `mlike`, `postsim`, `extract`.

Example

```
postsimHM mi.mtx 1000.mtx 1.mtx 10.0.mtx
```

Details

Generates draws of unknown quantities from their posterior distribution using a Gaussian random walk Metropolis chain with proposal covariance proportional to the sample covariance of draws from the posterior simulation matrix. Generates mn new posterior draws, and appends every n th draw to the posterior simulation matrix. If there are any draws from a previous invocation of `postsimHM`, the first new draw comes from the transition kernel of the Markov chain used for posterior simulation. Otherwise, it comes from the initial distribution of the Markov chain.

Use the `extract` command to obtain the posterior HM draws.

3.4.20 The priorRobust Command

Description

Calculates upper and lower bounds on the mean of a posterior function of interest, as the prior distribution is varied from its original specification.

Usage

```
priorRobust logWeight.mtx sample.mtx factors.mtx
```

Inputs

<code>logWeight</code>	Vector of length m : log weights
<code>sample</code>	Vector of length m : posterior sample of some scalar function of interest
<code>factors</code>	Vector of length n : bound factors for robustness analysis

Outputs

<code>mean</code>	Real scalar: posterior sample mean for original prior specification
<code>std</code>	Real scalar: posterior sample standard deviation for original prior specification
<code>U</code>	Vector of length n : exact upper bounds
<code>L</code>	Vector of length n : exact lower bounds
<code>Ut</code>	Vector of length n : asymptotic upper bounds
<code>Lt</code>	Vector of length n : asymptotic lower bounds

Example

```
echo -e "5.0 10.0 20.0" > K.mtx
priorRobust lw.mtx beta.mtx K.mtx K.mtx K.mtx K.mtx K.mtx K.mtx K.mtx
```

Details

For each bound factor, calculates exact lower and upper bounds and asymptotic lower and upper bounds for the posterior mean. For each bound parameter k_i , `priorRobust` calculates exact lower and upper bounds L_i and U_i for the posterior mean of the function of interest g , for the following set of prior density kernels.

$$\left\{ p^*(\cdot) : \frac{1}{k_i} p(\theta) \leq p^*(\theta) \leq k_i p(\theta) \quad \forall \theta \in \Theta \right\}$$

where $p(\cdot)$ is the actual prior density. It uses the algorithm described in Geweke and Petrella [6]. Also for each k_i , `priorRobust` calculates asymptotically valid

lower and upper bounds \tilde{L}_i and \tilde{U}_i , using the results of DeRobertis and Hartigan [1].

3.4.21 The `priorfilter` Command

Description

Filters out previously generated draws from the prior simulation matrix of a given model instance.

Usage

```
priorfilter modelInst.mtx filter.mtx
```

Inputs

<code>modelInst</code>	Integer: model instance identifier.
<code>filter</code>	Vector of integers of length n : indices of existing draws to keep

Outputs

None.

Example

```
# Create a file filter.ftr with the lines 101, ... 1000  
  
priorfilter mi.mtx filter.mtx
```

Details

The i th draw of the prior simulation matrix is kept if and only if $i = f_j$ for some j from 1 to n .

3.4.22 The priorsim Command

Description

Generates or appends to the prior simulation matrix of a given model instance.

Usage

```
priorsim modelInst.mtx m.mtx n.mtx
```

Inputs

<code>modelInst</code>	Integer: model instance identifier.
<code>m</code>	Integer: number of prior draws to record
<code>n</code>	Integer: number of prior draws to generate for each one recorded

Outputs

None.

See Also

`minst`, `priorfilter`, `postsim`, `extract`.

Example

```
priorsim mi.mtx 1000.mtx 1.mtx
```

Details

Generates draws of unknown quantities from their prior distribution. Generates mn new prior draws, and appends every n th draw to the prior simulation matrix. If there are any draws from a previous invocation of `priorsim`, the first new draw comes from the transition kernel of the Markov chain used for prior simulation. Otherwise, it comes from the initial distribution of the Markov chain.

Use the `extract` command to obtain the prior draws.

3.4.23 The `setseedconstant` Command

Description

Sets the seeds of the random number generators to a constant value.

Usage

```
setseedconstant
```

Inputs

None.

Outputs

None.

See Also

```
setseedtime.
```

Example

```
setseedconstant
```

Details

This is useful for ensuring that repeated invocations of a command generating random values lead to the same results.

3.4.24 The `setseedtime` Command

Description

Sets the seeds of the random number generators to the number of seconds since the beginning of 1970.

Usage

```
setseedtime
```

Inputs

None.

Outputs

None.

See Also

```
setseedconstant.
```

Example

```
setseedtime
```

Details

This is useful for ensuring that repeated invocations of a command generating random values lead to different results.

3.4.25 The weightedSmooth Command

Description

Estimates a univariate density function for a weighted random sample, using a kernel smoothing algorithm adapted to weighted samples.

Usage

```
weightedSmooth logWeight.mtx sample.mtx
weightedSmooth logWeight.mtx sample.mtx ktype.mtx krange.mtx wwf.mtx
nplot.mtx range_a1.mtx range_a2.mtx
```

Inputs

<code>logWeight</code>	Vector of length M : log weights
<code>sample</code>	Vector of length M : a posterior sample of some function of interest
<code>ktype</code>	String: kernel type (optional)
<code>krange</code>	String: kernel range type (optional)
<code>wwf</code>	Real scalar: window width fraction (optional)
<code>nplot</code>	Integer: number of ordered pairs to generate (optional)
<code>range_a1</code>	Real scalar: left bound range parameter (optional)
<code>range_a2</code>	Real scalar: right bound range parameter (optional)

Outputs

<code>x</code>	Vector of length N : ordinate values
<code>y</code>	Vector of length N : abscissa values

Example

```
weightedSmooth lw.mtx z.mtx z.mtx z.mtx
echo 2000 > nplot.mtx
echo triangular > ktype.mtx
weightedSmooth lw.mtx z.mtx z.mtx z.mtx
```

Details

The estimated density at a point z is

$$f(z) = \frac{\sum_{m=1}^M w_m K\left(\frac{z-z_m}{h}\right)}{h \sum_{m=1}^M w_m}$$

Table 3.1: Values of `Ktype`

<code>Ktype</code>	K
<code>uniform</code>	$K(t) = \frac{1}{2}\chi_{(-1,1)}(t)$
<code>triangle</code>	$K(t) = (1 - t)\chi_{(-1,1)}(t)$
<code>biweight</code>	$K(t) = \frac{15}{16}(1 - z^2)^2\chi_{(-1,1)}(t)$

The functional form of the kernel function K depends on the value of `ktype` according to Table 3.1.

For any set S , the function $\chi_S(\cdot)$ is a set membership indicator function.

The value h is given by

$$h = \lambda(q_{\frac{3}{4}} - q_{\frac{1}{4}})$$

where q_α denotes the α 'th sample quantile of z .

The `weightedSmooth` command generates N ordered pairs (x_i, y_i) . The values x_i are evenly spaced between x_{min} and x_{max} , determined by `Krange` according to Table 3.2. The values y_i satisfy $y_i = f(x_i)$.

Table 3.2: Values of `Krange`

<code>Krange</code>	x_{min}	x_{max}
<code>quantile</code>	q_{a_1}	q_{a_2}
<code>absolute</code>	a_1	a_2

For most plotting routines, N should be in the range of 200 to 400. The choice of λ depends on how smooth the resulting plot is desired to be. As with all kernel smoothing methods, some experimentation will probably be necessary. The greater the number of simulations available, the smaller λ can be and still retain visual smoothness. It is generally easier to use the `Krange=quantile` option and specify a_1 in the range .001 to .01 and a_2 in the range .99 to .999; this will include the important part of the estimated density while not wasting space on the plot for points where the density is small.

3.4.26 The wishartSim Command

Description

Generates a sample from a Wishart distribution.

Usage

```
wishartSim A.mtx nu.mtx n.mtx
```

Inputs

A	m by m matrix: inverse scale parameter of Wishart distribution
nu	Real scalar: degrees of freedom parameter of Wishart distribution
n	Integer: number of draws to generate

Outputs

sample	n by m^2 matrix: sample generated from Wishart distribution
--------	---

See Also

paretoSim, gaussianSim, gammaSim, dirichletSim.

Example

```
echo -e "1.0 0.0
n0.0 1.0" > A.mtx
echo 100 > nu.mtx
wishartSim A.mtx nu.mtx 1000.mtx 1000.mtx
```

Details

Each of the n draws of the sample is an m by m matrix with a Wishart distribution. The result is given as a m by m^2 matrix.

See Appendix A for the parametrization of the Wishart distribution.

Chapter 4

A BACC Tutorial

In order to answer commonly asked questions, this chapter contains a step-by-step tutorial with explanations of what each step is doing and what each term means.

4.1 Creating a Model Instance

The first step is to create a “model instance”. A model instance can be thought of as a box containing everything that BACC needs to know about a particular model. This includes the type of the model, i.e., normal linear model, a list of names for the unknowns, and the names of the matrix files containing the known quantities. The next line creates an instance of the normal linear model and calls it `simhd`. References to `simhd` will now refer back to this model instance.

```
minst simhd nlm beta h betahd.mtx Hhd.mtx nuhd.mtx shd.mtx Xhd.mtx yhd.mtx
```

4.2 Simulating the Model

Simulation can be performed from the prior or the posterior of a model using the commands `priorsim` and `postsim`. After posterior simulation has been done, it is also possible to simulate the posterior using a random walk Hastings-Metropolis algorithm. This is done using the command `postsimhm`. If you wish to compare output across runs, it is a good idea to set the seeds of the random number generators. This is done using the `setSeedConstant` command. Strictly speaking, this sets the seeds of another routine that then sets the seeds of the random number generators.

First, generate 5000 samples from the prior.

```
setSeedConstant
priorsim simhd 5000 1
```

Now generate 1000 samples from the posterior.

```
setSeedConstant
postsim simhd 1000 1
```

To filter out posterior samples you can specify the numbers of the samples you wish to keep. For example, to drop the first 100 samples you would create a vector containing the numbers 101 to 1000 and store it in a filter file such as `b.ftr`. Then pass this list of samples you wish to keep to the `postfilter` routine.

```
postfilter simhd b.ftr
```

Now you are left with 900 posterior samples. Add another 4100 samples.

```
setSeedConstant
postsim simhd 4100 1
```

We can also simulate 10000 samples from the posterior using a random walk Hastings-Metropolis algorithm.

```
setSeedConstant
postsimHM simhd 10000 1 1
```

4.3 Computing Marginal likelihoods

The `mlike` command can report several log marginal likelihood values for a model instance. First, if the model has an analytic marginal likelihood, then the exact marginal likelihood is printed. Then, for either a default truncation parameter or for a user specified vector of truncation parameters, `mlike` computes an estimate of the log marginal likelihood for each truncation parameter. Finally, if there are `postsimhm` samples as well, then `mlike` also reports estimates of the log marginal likelihood based on those samples. To use the default truncation parameter and compute the log marginal likelihoods type

```
mlike simhd
```

or to specify a vector of truncation parameters enter the numbers

```
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
```

in a file called `p.mtx` (for example). Then type

```
mlike simhd p.mtx
```

4.4 Computing Moments of Functions of Interest

To compute the posterior mean of the first element of `beta` we first need to extract the vector of logweights and the vector corresponding to the first element of `beta`. This is done with the following commands:

```
getUnknownVector simhd beta 1 beta1.mtx
```

```
getLogWeight simhd lw.mtx
```

The matrix files `beta1.mtx` and `lw.mtx` containing the parameter and logweight values that we are interested in. The mean can then be calculated by

```
expect1 lw.mtx beta1.mtx
```

To specify taper parameters other than the default values, you can do the following

```
expect1 lw.mtx beta1.mtx "4.0 8.0"
```

More generally, `expect1` can compute the mean and standard deviation of any function

of interest. The first argument is a vector of log weight values and the second argument is a vector of function of interest values. In some cases it may be sufficient to simply change the weight vector. Unfortunately, this is difficult to do in the console version of BACC. Please see the Matlab or Gauss tutorials (in the BACC Matlab or Gauss manuals) for examples of the power of BACC when linked with matrix oriented software packages such as Matlab and Gauss.

To compute combined sample means and test for the equality of the individual population means use the `expectN` command. For example, to compare the posterior and prior means for the first element of the `beta` vector, we first extract this element

```
getUnknownVector simhd betaprior 1 beta1prior.mtx
```

and then type

```
expectN lw.mtx beta1.mtx lw.mtx beta1prior.mtx
```

or

```
expectN lw.mtx beta1.mtx lw.mtx beta1prior.mtx "4.0 8.0"
```

4.5 Sensitivity to the Prior

To calculate upper and lower bounds on the mean of a posterior function of interest, as the prior distribution is varied from its original specification type

```
priorRobust lw.mtx beta1.mtx
```

or

```
priorRobust lw.mtx beta1.mtx "2 10"
```

where the second version specifies values of the `bounds` parameter other than the default values.

4.6 Kernel Smoothing of Simulations

To estimate a univariate density function for a weighted random sample, using kernel smoothing, type

```
weightedSmooth lw.mtx beta1.mtx
```

This generates two files, `smoothX.mtx` and `smoothY.mtx` which can be loaded into standard software for plotting. Again, there are advantages to using BACC within software such as Matlab or Gauss as the plotting tools are at your immediate disposal.

The `weightedSmooth` command has various settings that can be changed. For example, to generate more samples and use a different kernel you might type

```
setenv nplot 2000
```

or

```
setenv ktype uniform
```

prior to using the the `weightedSmooth` command.

4.7 Saving and Loading Model Instances

4.7.1 To/From Text Files

To save a model instance called `simhd` to an ascii file called `simhd.txt` use the command

```
miSaveAscii simhd simhd.txt
```

To load a model instance stored in an ascii file called `simhd.txt` and to give it the name `simhd`, use the command

```
miLoadAscii simhd simhd.txt
```

4.8 Drawing from Various Distributions

The following code demonstrates how to draw from various distributions. See Appendix A for the parameterization of these distributions.

4.8.1 Dirichlet

```
dirichletSim aD.mtx 1000 simDirichlet.mtx
```

4.8.2 Gamma

```
gammaSim a.mtx b.mtx 1000 simGamma.mtx
```

4.8.3 Gaussian

```
gaussianSim beta.mtx H.mtx 1000 simGaussian.mtx
```

4.8.4 Pareto

```
paretoSim a.mtx b.mtx 1000 simPareto.mtx
```

4.8.5 Wishart

```
wishartSim Hhd.mtx df.mtx 1000 simWishart.mtx
```

Appendix A

Distributions

This appendix gives the density and mass functions for the distributions used in this document.

A.1 The Dirichlet Distribution

A random vector π of length n has the Dirichlet distribution with parameter vector $\alpha \in \mathfrak{R}_+^n$, denoted $\pi \sim \text{Di}(\alpha)$, if its probability density function is

$$p(\pi|\alpha) = \begin{cases} m^{-1/2} \frac{\Gamma(\sum_{i=1}^m \alpha_i)}{\prod_{i=1}^m \Gamma(\alpha_i)} \prod_{i=1}^m \pi_i^{\alpha_i-1} & \pi \in \Delta_n \equiv \{p \in \mathfrak{R}_+^n : \sum_{i=1}^n p_i = 1\} \\ 0 & \text{otherwise} \end{cases}$$

The mean and variance are given by

$$\begin{aligned} \text{E}[\pi_i|\alpha] &= \frac{\alpha_i}{\sum_{j=1}^n \alpha_j} \\ \text{Var}[\pi_i|\alpha] &= \frac{\text{E}[\pi_i|\alpha](1 - \text{E}[\pi_i|\alpha])}{1 + \sum_{j=1}^n \alpha_j} \\ \text{Cov}[\pi_i, \pi_j|\alpha] &= \frac{-\text{E}[\pi_i|\alpha]\text{E}[\pi_j|\alpha]}{1 + \sum_{k=1}^n \alpha_k} \end{aligned}$$

A.2 The Gamma Distribution

A random scalar λ has the Gamma distribution with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$, denoted $\lambda \sim \text{Ga}(\underline{\alpha}, \underline{\beta})$, if its probability density function is

$$p(\lambda|\alpha, \beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} & \lambda > 0 \\ 0 & \text{otherwise} \end{cases}$$

The mean and variance are given by

$$\begin{aligned} \mathbb{E}[\lambda|\alpha, \beta] &= \frac{\alpha}{\beta} \\ \text{Var}[\lambda|\alpha, \beta] &= \frac{\alpha}{\beta^2} \end{aligned}$$

A.3 The Normal Distribution

A random vector x has the Normal Distribution with mean parameter vector $\mu \in \mathfrak{R}^k$ and positive definite $k \times k$ variance parameter matrix Σ , denoted $x \sim \mathcal{N}(\mu, \Sigma)$, if its probability density function is

$$p(x|\mu, \Sigma) = |\Sigma|^{-\frac{1}{2}} (2\pi)^{-\frac{k}{2}} \exp \left[-\frac{1}{2} (x - \mu)' \Sigma^{-1} (x - \mu) \right] \quad \forall x \in \mathfrak{R}^n$$

The mean and variance are given by

$$\begin{aligned} \mathbb{E}[x|\mu, \Sigma] &= \mu \\ \text{Var}[x|\mu, \Sigma] &= \Sigma \end{aligned}$$

A.4 The Pareto Distribution

A random scalar x has the Pareto Distribution with parameters $\alpha > 0$ and $\beta \geq 0$, denoted $\theta \sim \text{Pa}(\alpha, \beta)$, if its probability density function is

$$p(\theta|\alpha, \beta) = \begin{cases} \alpha\beta^\alpha \theta^{-(\alpha+1)} & \theta \geq \beta \\ 0 & \text{otherwise} \end{cases}$$

The mean and variance are given by

$$\begin{aligned} \mathbb{E}[\theta|\alpha, \beta] &= \frac{\alpha\beta}{\alpha - 1} \\ \text{Var}[\theta|\alpha, \beta] &= \frac{\alpha\beta^2}{(\alpha - 1)^2(\alpha - 2)} \end{aligned}$$

A.5 The Poisson Distribution

A discrete random variable x has the Poisson distribution with mean parameter $\lambda > 0$, denoted $x \sim \text{Po}(\lambda)$, if its probability mass function is

$$p(x) = \begin{cases} e^{-\lambda} \frac{\lambda^x}{x!} & x \in \{0, 1, \dots\} \\ 0 & \text{otherwise} \end{cases}$$

The mean and variance are given by

$$\begin{aligned} E[x|\lambda] &= \lambda \\ \text{Var}[x|\lambda] &= \lambda \end{aligned}$$

A.6 The Wishart Distribution

An $m \times m$ random matrix H has the Wishart distribution with positive definite $m \times m$ scale parameter matrix A and degrees of freedom parameter $\nu > m$, denoted $H \sim \text{Wi}(A, \nu)$, if its probability density function is

$$p(H|A, \nu) = \begin{cases} \frac{\pi^{-m(m-1)/4} |A|^{-\nu/2}}{2^m \nu^{m/2} \prod_{i=1}^m \Gamma(\frac{\nu-i+1}{2})} \cdot |H|^{(\nu-m-1)/2} \exp\left[-\frac{1}{2} \text{tr}(A^{-1}H)\right] & H \text{ p.d.} \\ 0 & \text{otherwise} \end{cases}$$

The mean, and mean of the matrix H^{-1} are given by

$$\begin{aligned} E[H|A, \nu] &= \nu A \\ E[H^{-1}|A, \nu] &= \frac{1}{\nu - m - 1} A^{-1} \end{aligned}$$

Bibliography

- [1] DeRobertis, L., and J. A. Hartigan, 1981, "Bayesian Inference Using Intervals of Measures," *The Annals of Statistics* 9: 235-244.
- [2] Gelfand, A.E. and D.K. Dey, 1994, "Bayesian Model Choice: Asymptotics and Exact Calculations," *Journal of the Royal Statistical Society Series B* **56**: 501-514.
- [3] Geweke, J., 1989, "Bayesian Inference in Econometric Models Using Monte Carlo Integration" *Econometrica*, 57, 1317-1340.
- [4] Geweke, J., 1992, "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments," in J.O. Berger, J.M. Bernardo, A.P. Dawid, and A.F.M. Smith (eds.), *Proceedings of the Fourth Valencia International Meeting on Bayesian Statistics*, 169-194. Oxford: Oxford University Press.
- [5] Geweke, J. 1999, "Simulation-Based Bayesian Inference for Economic Time Series," in R.S. Mariano, T. Schuermann and M. Weeks (eds.), *Simulation-Based Inference in Econometrics: Methods and Applications*. Cambridge: Cambridge University Press, forthcoming.
- [6] Geweke, J. and L. Petrella, 1999, "Prior Density Ratio Class Robustness in Econometrics," *Journal of Business and Economic Statistics*, forthcoming.
- [7] Raftery, A.E., 1995, "Hypothesis testing and model selection via posterior simulation," University of Washington working paper.