The Pearson correlation example

OpenBUGS

The Pearson Correlation Definition

The **Pearson linear correlation coefficient**, denoted as ρ or r, is defined as:

$$\rho = \frac{S_{xy}}{\sqrt{S_{xx} \quad S_{yy}}} = \frac{\sum_{i=1}^{n}(x_i - \mu_X)(y_i - \mu_Y)}{\sqrt{\sum_{i=1}^{n}(x_i - \mu_X)^2 \sum_{i=1}^{n}(y_i - \mu_Y)^2}} \tag{7}$$

which **always takes values in the interval** $[-1, 1]$, that is, $-1 \leq \rho \leq 1$, and **the equalities hold only when the values of the two variables lie on a straight line**, getting apart from the equality when they separate from the linear trend.

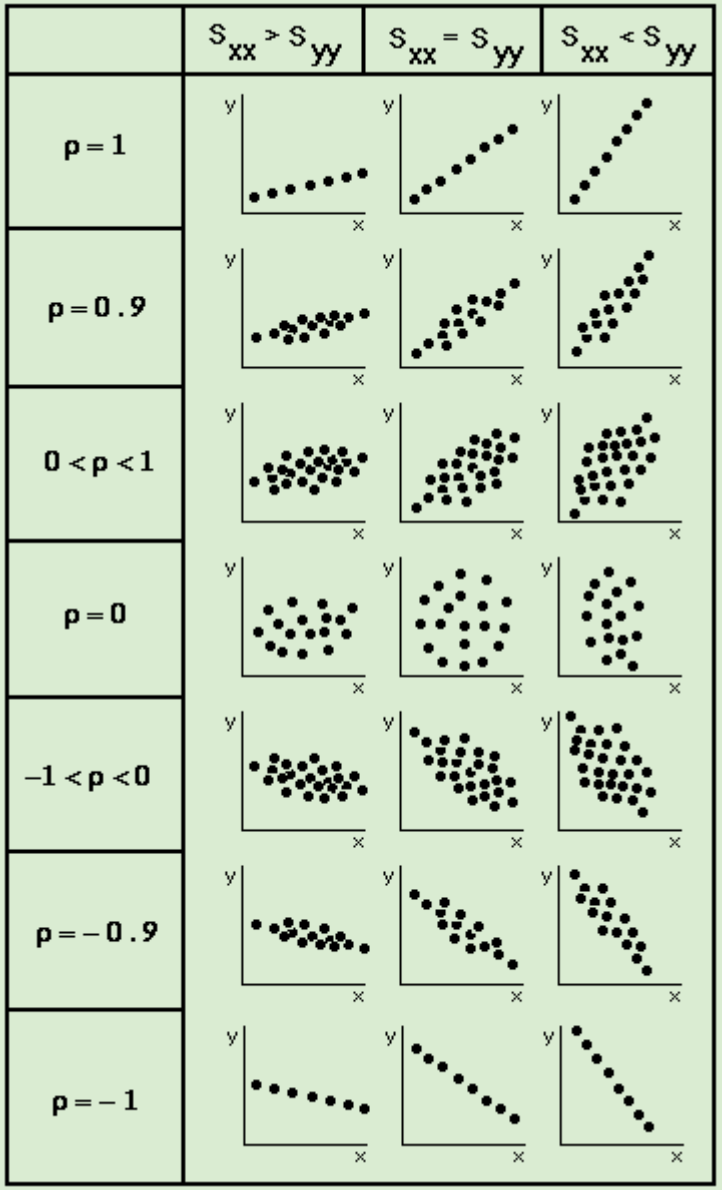
It can be easily demonstrated that the following relation holds:

$$Var[Y - aX - b] = Var[Y](1 - \rho^2). \tag{8}$$

Due to this property, **the correlation coefficient is used to measure the proportion of the variance of Y gained when predicting it with a linear law of the form $Y = aX + b$**.

It **is used to measure the convenience of adjusting or not, a straight line a the cloud of data points**. It is an alternative to the covariance, but **has the advantage of being dimensionless**.

Graphical Illustration Of The Pearson Correlation



Multivariate normal

We consider a two-dimensional normal variable, with mean

$$\mu = (\mu_1, \mu_2),$$

and covariance matrix

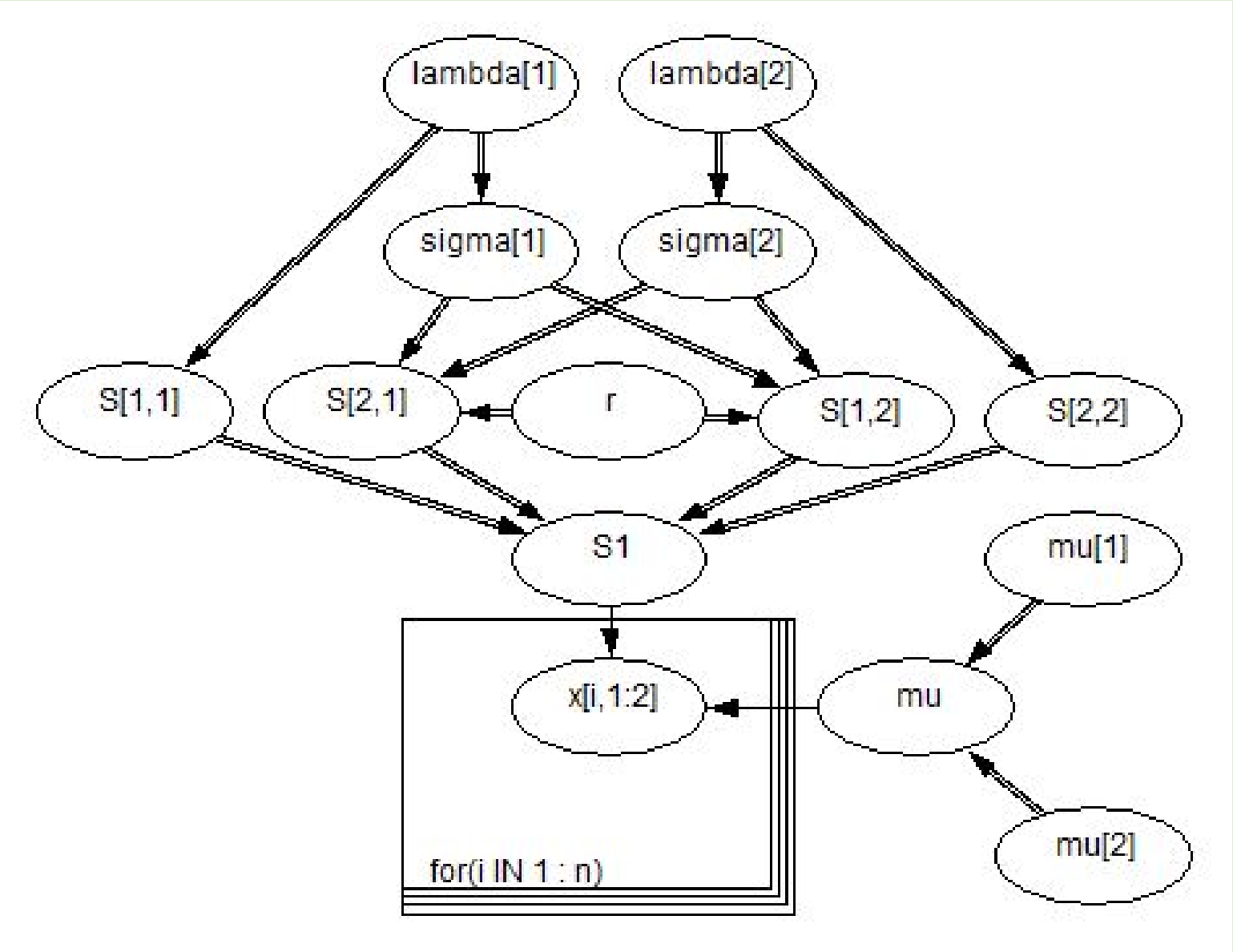
$$Covar = \begin{pmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

where μ_1 and μ_2 are the marginal means of the two variables, σ_1 and σ_2 are their standard deviations, and r is the correlation coefficient.

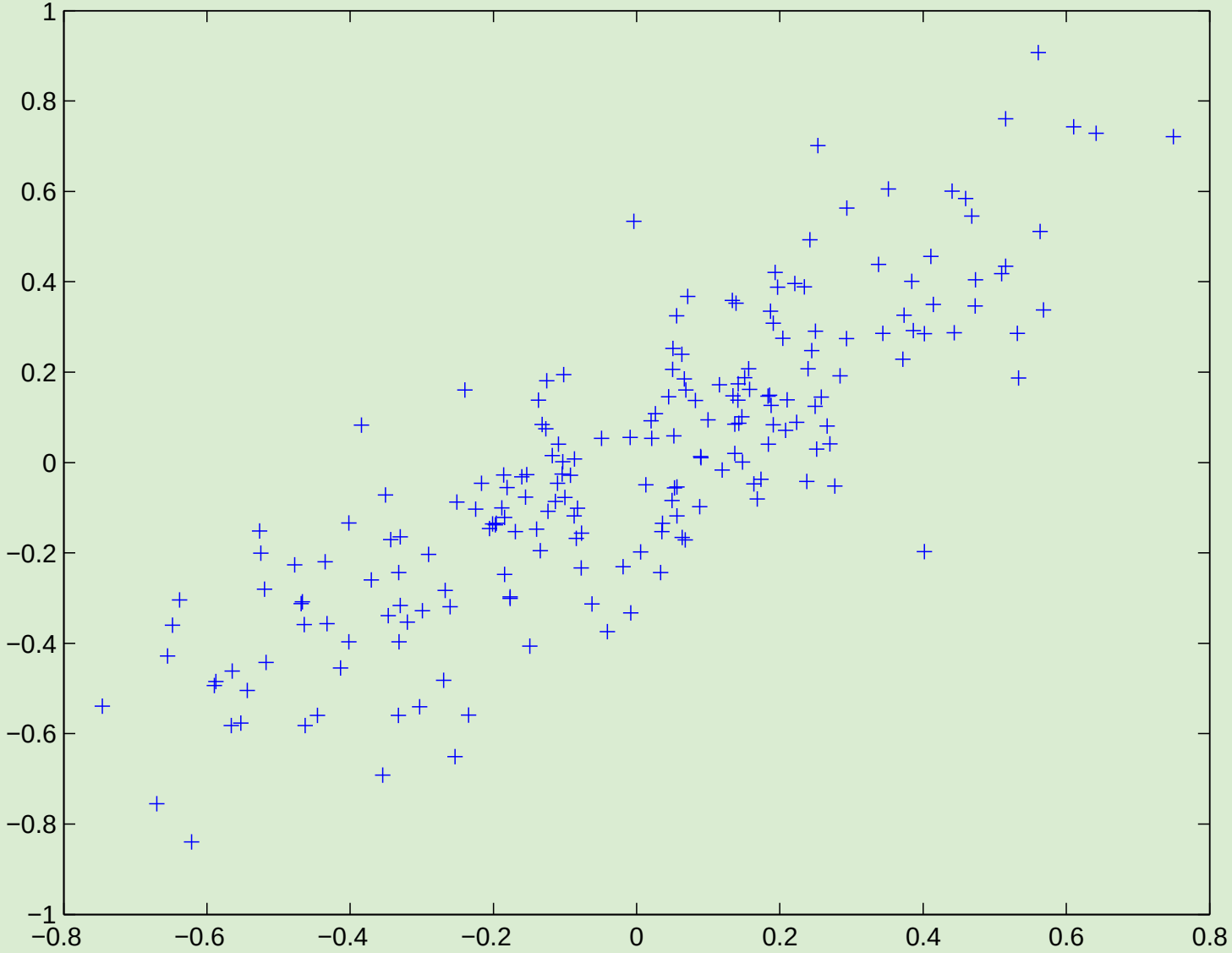
The Pearson Correlation Commands

```
# Pearson Correlation
model{
# Data
    for (i in 1:n){x[i,1:2] ~ dmnorm(mu[],S1[,])}
# Priors
    mu[1] ~ dnorm(0,.001)
    mu[2] ~ dnorm(0,.001)
    lambda[1] ~ dgamma(.001,.001)
    lambda[2] ~ dgamma(.001,.001)
    r ~ dunif(-1,1)
# Reparameterization
    sigma[1] <- 1/sqrt(lambda[1])
    sigma[2] <- 1/sqrt(lambda[2])
    S[1,1] <- 1/lambda[1]
    S[1,2] <- r*sigma[1]*sigma[2]
    S[2,1] <- r*sigma[1]*sigma[2]
    S[2,2] <- 1/lambda[2]
    S1[1:2,1:2] <- inverse(S[1:2,1:2])
}
```

The Pearson Correlation Doodle Model



The Pearson Correlation Example Data



The Pearson Correlation Commands, Data and Initial Files

OpenBUGS

File Edit Attributes Tools Info Model Inference Doodle Map Text Window Examples Manuals Help

Commands.txt

```
# Pearson Correlation
model{
# Data
for (i in 1:n){
x[i,1:2] ~ dmnorm(mu[],S1[,])
}
# Priors
mu[1] ~ dnorm(0,.001)
mu[2] ~ dnorm(0,.001)
lambda[1] ~ dgamma(.001,.001)
lambda[2] ~ dgamma(.001,.001)
r ~ dunif(-1,1)
# Reparameterization
sigma[1] <- 1/sqrt(lambda[1])
sigma[2] <- 1/sqrt(lambda[2])
S[1,1] <- 1/lambda[1]
S[1,2] <- r*sigma[1]*sigma[2]
S[2,1] <- r*sigma[1]*sigma[2]
S[2,2] <- 1/lambda[2]
S1[1:2,1:2] <- inverse(S[1:2,1:2])
```

Sample Monitor Tool

node | chains 1 to 3 | percentiles: 2.5, 5, 10, 25, median, 75, 90, 95, 97.5

beg 1 | end 10000000 | thin 1

clear | set | diagnostics | trace | jump | stats | density | bgr diag | history | accept | coda | quantiles | auto cor

Specification Tool

check model | load data | compile | num of chains 3 | load inits | for chain 1 | gen inits

Data.txt

```
list(n=200,
x=structure(
.Data=c( -0.2052362, -0.1458369,
  0.3735175,  0.3261632,
 -0.2398440,  0.1603157,
 -0.3508905, -0.0719385,
 -0.2673868, -0.2829939,
 -0.1810925, -0.0553196,
 -0.1766704, -0.3005456,
  0.0564088, -0.0542254,
 -0.0622531, -0.3133023,
  0.1854494,  0.1491656,
 -0.2693903, -0.4820510,
  0.2530836,  0.7012880,
 -0.4773157, -0.2263578,
  0.2769757, -0.0522873,
 -0.0767768, -0.1562384,
  0.0527510, -0.0560750,
 -0.6215200, -0.8392631,
 -0.4016318, -0.1337631,
  0.3716218,  0.2288494,
  0.6416768,  0.7282485,
 -0.0870124,  0.0079959,
```

Initials1.txt

```
list(mu=c(    0.000000000000,    0.000000000000),
lambda=c(    10.000000000000,   10.000000000000),
r=0.8)
```

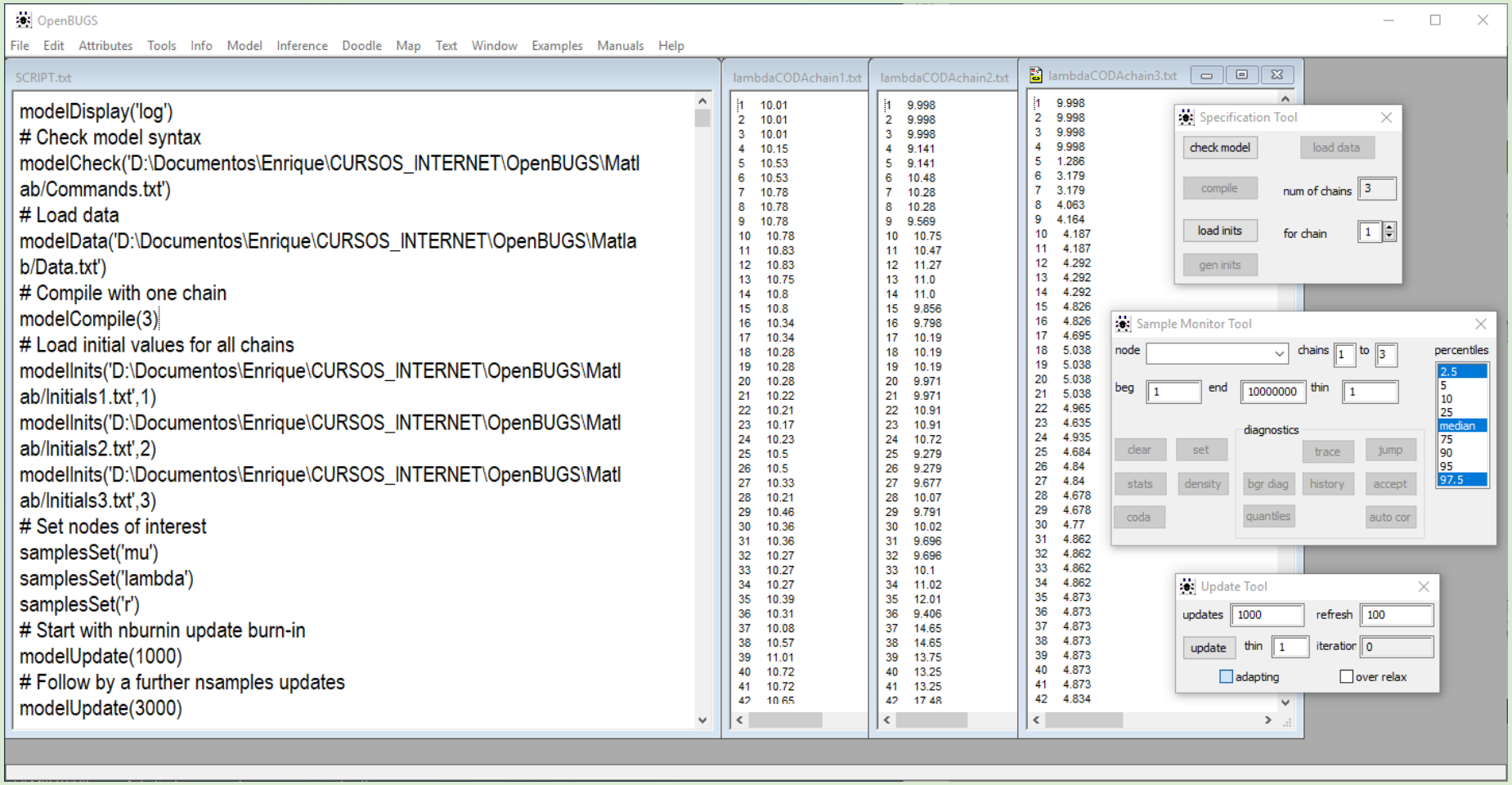
Initials2.txt

```
list(mu=c(    0.000000000000,    0.000000000000),
lambda=c(    10.000000000000,   10.000000000000),
r=0)
```

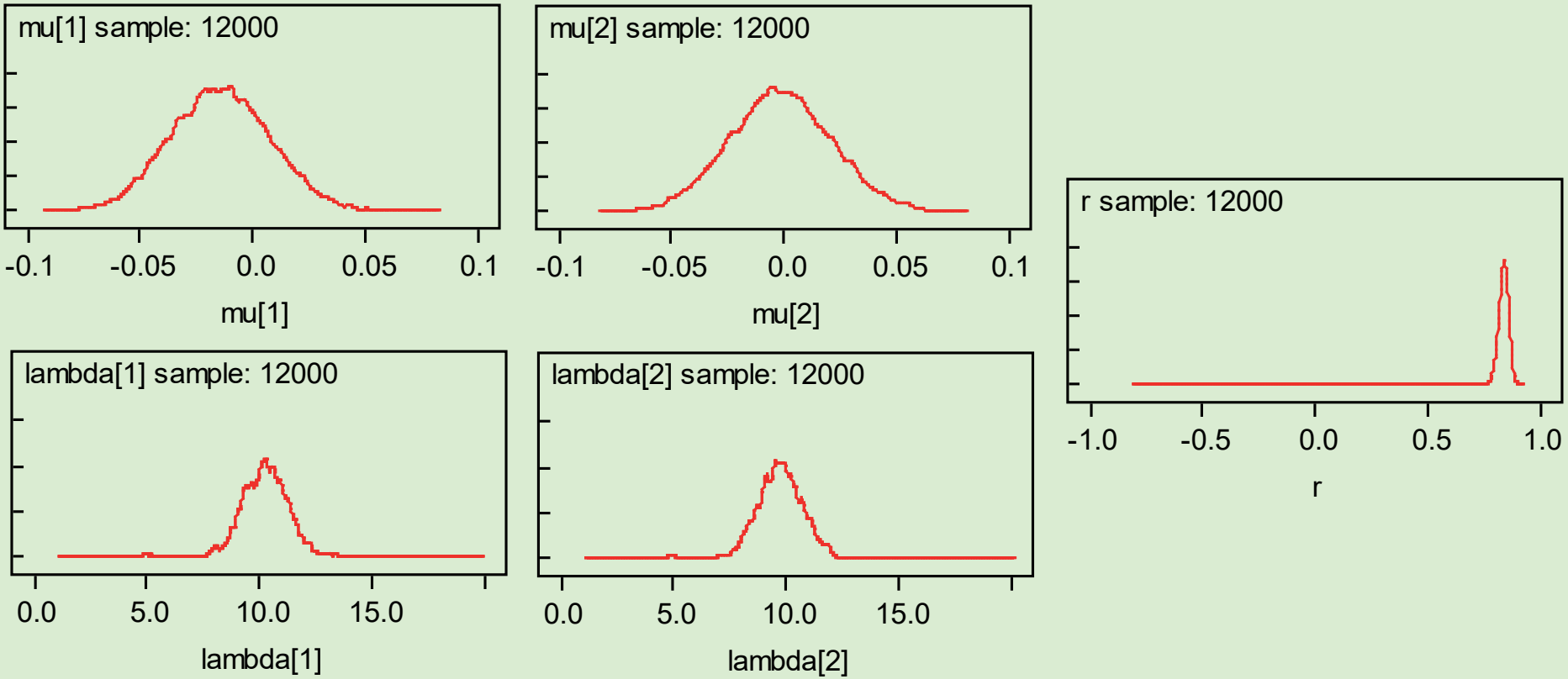
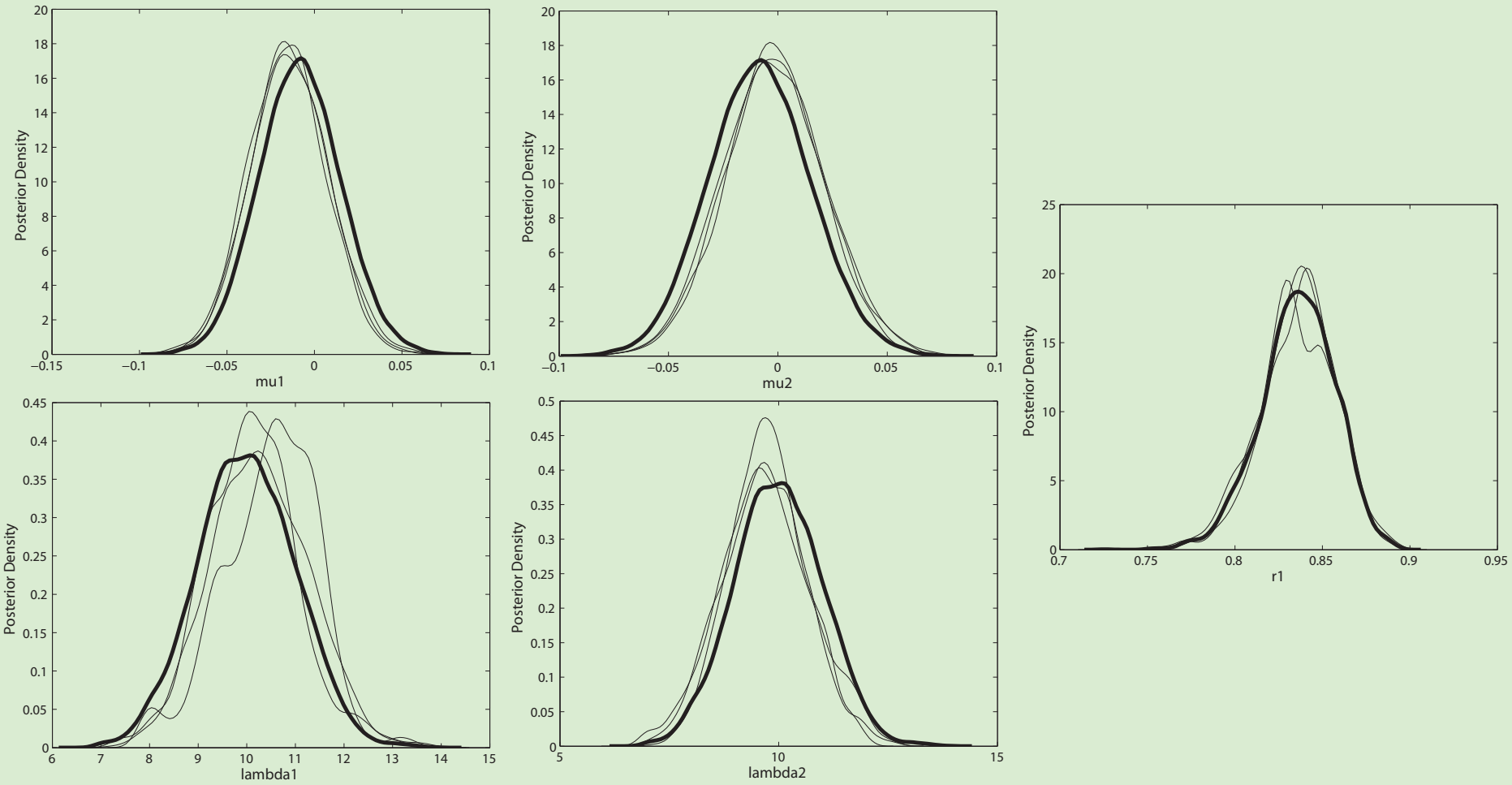
Initials3.txt

```
list(mu=c(    0.000000000000,    0.000000000000),
lambda=c(    10.000000000000,   10.000000000000),
r=-0.8)
```

The Pearson Correlation Script and Coda Files



The Pearson Correlation Posterior Densities of $\mu_1, \mu_2, \lambda_1, \lambda_2$ and r



The Pearson Correlation Posterior Densities of $\mu_1, \mu_2, \lambda_1, \lambda_2$ and r