*International Journal of Scientific World*, 2 (2) (2014) 48-55
©Science Publishing Corporation
www.sciencepubco.com/index.php/IJSW
doi: 10.14419/ijsw.v2i2.2967
Research Paper

# Using MCMC methods in some application problems

**P. Azhdari [1]*, N. Jafarpanahi [1], A. Beitollahi [2]**

[1] *Department of Statistics, Tehran North Branch, Islamic Azad university, Tehran, Iran*
[2] *Department of Statistics, Roudehen Branch, Islamic Azad university, Roudehen, Iran*
**Corresponding author E-mail: par_azhdari@yahoo.com*

**Copyright © 2014 Azhdari et. al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.**

**Abstract**

MCMC methods are very important tools for estimating unknown parameters in Bayesian models. Especially in the case of high dimensions. Gaussian mixture model is one of the applications of estimating hyper parameters by MCMC method.

***Keywords***: *Gibbs Sampling, Slice Sampling, Metropolis-Hastings Algorithm, Gaussian, Mixture Model.*

## 1. Introduction

In recent years MCMC methods have been applied for estimating of maximum likelihood of multivariate distribution especially in the case of lost or censored data [10], [11]. We will evaluate in this paper some applied examples that in them have used various methods of MCMC. In section 2, we will present brief explanation about various methods of MCMC. In section 3, we will explain role of MCMC in Gaussian Process Regression. In section 4, we will present example of applied problems that implemented in R & S-plus.

## 2. Different methods of MCMC

Gibbs sampler and Metropolis-Hastings algorithm are the most famous methods of MCMC. Although Gibbs sampler algorithm is simple and speed in comparison with Metropolis-Hastings, but it needs the calculation of marginal conditional densities [2], [9].

Another MCMC method is slice sampling [3]. Assume we can decompose density of $f(x)$ to $\prod_{i=1}^{k} f_i(x)$ For each $x$ , $f_i(x) > 0$ . This algorithm has following steps:

Having start point $x^t$ in time t:

1) $\omega_1^{t+1} \sim uniform\left[0, f_1(x^{(t)})\right]$

2) $\omega_2^{t+1} \sim uniform\left[0, f_2(x^{(t)})\right]$

And in kth step $x^{t+1} \sim uniform(A^{(t+1)})$, $A^{(t+1)} = \left\{ y : f(y) \geq \omega_i^{t+1} \right\}$.

Let $\omega_i$ ,i= 1… k, are uniform random variables in $[0, f_i(x)]$. Using joint density of $(\omega_1, ..., \omega_k, x)$ and generalization of Gibbs sampling methods, we will reach samples of expected density. Let joint density define on S, where

$$S = \{(\omega_1, ..., \omega_k, x), 0 \leq \omega_i \leq f_i(x) \ \forall x \in \sup port(f) \ \} \quad (1)$$

And assume

$$g(\omega_1, ..., \omega_k, x) = \mathrm{B} \quad (2)$$

Marginal density of X is

$$g_X(x) = B \int_0^{f_k(x)} \mathrm{L} \int_0^{f_k(x)} d\omega_1 \mathrm{L}\, d\omega_k = B f_1(x) \mathrm{L}\, f_k(x) \tag{3}$$

Then $g_X(x) = f(x)$. Now we generate random samples by Gibbs sampling of marginal density of X. Conditional distributions for each $\omega_i$ , and X are as follows:

$$g\omega_i \left|(\omega_i \right| \omega_1, \mathrm{K}, \omega_{i-1}, \omega_{i+1}, \mathrm{K}, \omega_k, x) \propto B \qquad \omega_i \in [0, f_i(x)]$$
$$gx \left|(x \right| \omega_1, \mathrm{K}, \omega_k) \propto B \tag{4}$$

Sample selection for $\omega_i$ with above conditional densities, is simple. If we suppose $\omega_i$ ,i= 1… k are obtained samples, we take sample of x from its correspondent conditional density and on D where $D = \{x : \omega_i \le f_i(x) \ \forall i, x \in Support(f)\}$ [3], [5].

# 3. The role of MCMC methods in Gaussian process regression

In a Gaussian Process Regression model that we demonstrate it by:

$y_i = f(x_i) + \varepsilon_i$ (5)

Where $\varepsilon_i \sim N(0, \sigma_D^2)$, is an error term? Errors of different data points are indipendent. For each i, $x_i$ has a normal distribution in a d-dimension space [14]. We demonstrate normal density function for a d-dimension space by:

$$\emptyset(x, \theta) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left\{\frac{1}{2}(x-\mu)^t \Sigma^{-1}(x-\mu)\right\} \tag{6}$$

Where $\sum = E\left[(x-\mu)^T (x-\mu)\right]$ , $\mu = E[x]$ and θ the vector of the mean μ and covariance matrix Σ. We suppose f ($x_i$) is a nonlinear function of $x_i$, which $x_i$ has normal distribution. So we can consider f ($x_i$) as a nonlinear combination of normal models.We show Gaussian mixture models k-trivial, GMM k-trivial, by:

$f_k(x) = \sum_{j=1}^{k} \pi_j \emptyset(x, \theta_j)$ (7)

Where$\pi_j$'s are weight of $\emptyset(x, \theta_j)$. Obviously f ($x_i$) consists of a Gaussian mixture models. For example for each i, we can consider f($x_i$) as a multivariate normal distribution with mean zero and covariance function C ($x_i$ , $x_j$)=COV(f($x_i$) , f($x_j$)) . Such as:

$$C(x_i, x_j) = C(x_i, x_j, \theta) = \upsilon_0 \exp\left(-\frac{1}{2}\sum_{q=1}^{Q} \omega_q (x_{iq} - x_{jq})^2\right) + a_0 + a_1 \sum_{q=1}^{Q} x_{iq}\, x_{jq} \tag{8}$$

Where $\theta = (\omega_1, \ldots, \omega_q, \upsilon_0, a_0, a_1, \sigma_\upsilon^2)$ denotes the set of unknown parameters. Since the vector of output variables are function of input variables, ($y_i = f(x_i) + \varepsilon_i$ ), $Y = (y_1, \cdots, y_n)^/$, and input variables have normal distribution, therefore $Y = (y_1, \cdots, y_n)^/$, has a normal distribution with mean zero and covariance function $\Psi(\theta) = C(\theta) + \sigma_\upsilon^2 I$, where I is an identity matrix, C(θ) is an N×N covariance function with elements as given in (8). So we can say Ψ(θ) is an N×N covariance function for output variables too. The set of data is D={y, x} = {($y_1$, x1) ($y_n$,$x_n$)}. Now, we should estimate parameters. The likelihood function is as follows:

$$L(\theta) = -\frac{1}{2}\log|\Psi(\theta)| - \frac{1}{2} y^{\mathrm{T}} \Psi^{-1}(\theta) y - \frac{\mathrm{N}}{2}\log(2\pi) \tag{9}$$

# 4. Applied examples

In this section, we will evaluate some applied examples using methods of MCMC that are presented in section 2. In first example, we will generate random variable of Beta density by random walks method. In second example, we will evaluate defalt rate estimation related to data of pomp that is performed by Gibbs sampling [5]. Third example is related to resolve second example using slice sampling. In forth example we will simulate bivariate normal distribution by using Gibbs sampling. Fifth example is related to resolve forth example by Metropolis-Hastings algorithm. Sixth example is related to simulate means of two population of normal mixture using Gibbs sampling.

## 4.1. Example 1

We want to generate random sample for posterior density $f(\theta \left| x \right.) \propto \theta^{x+5}(1-\theta)^{n+5-x}$ , n=12, x=8 and random walks model $y = x + z$ , $z \sim N(0,1)$. To solve this example we will use Metropolis method.

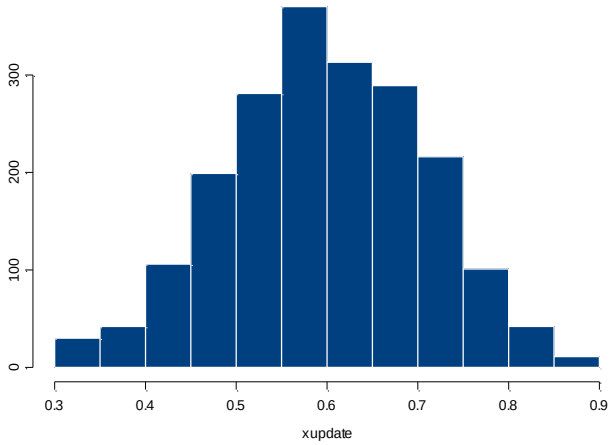



**Fig. 1:** The Histogram by S-Plus Is As Follows

## 4.2. Example 2

In 1987 obtained set of data for some pomp [7] that had defalts in t-time. Suppose we have number of defalts, $y_i$, for each 10 pomp in a nuclear plant. So we have time, $t_i$, that in it each pomp has a defalt:
Y=(5,1,5,14,3,19,1,1,4,22) t= (94,16,63,126,5,31,1,1,2,10)
We want to model number of defalts by Gibbs sampling with a poisson likelihood function that in it differs expected mean of number of defalt, $\lambda_i$. The likelihood function is:

$$\prod_{i=1}^{10} P(\lambda_i t_i)$$

Let for each $\lambda_i$, we have prior distribution $)\beta, \alpha)\Gamma$ with $\alpha = 1.8$, and $\beta$ has prior distribution $)\delta,\gamma)\Gamma$ with $\gamma = 0.01$ and $\delta = 1)$ The posterior distribution is as follows:

$$p(\lambda,\beta \mid y,t) \propto \left( \prod_{i=1}^{10} \lambda_i^{y_i+\alpha-1} e^{-(t_i+\beta)\lambda_i} \right) \beta^{10\alpha+\gamma-1} e^{-\delta\beta} \tag{10}$$

Finding fully conditionals:

$$P(\lambda_i \mid \lambda_{-i},\beta,y,t) \sim \Gamma(y_i+\alpha, t_i+\beta)$$
$$P(\beta \mid \lambda,y,t) \sim \Gamma(10\alpha+\gamma, \delta+\Sigma_{i=1}^{10}\lambda_i) \tag{11}$$

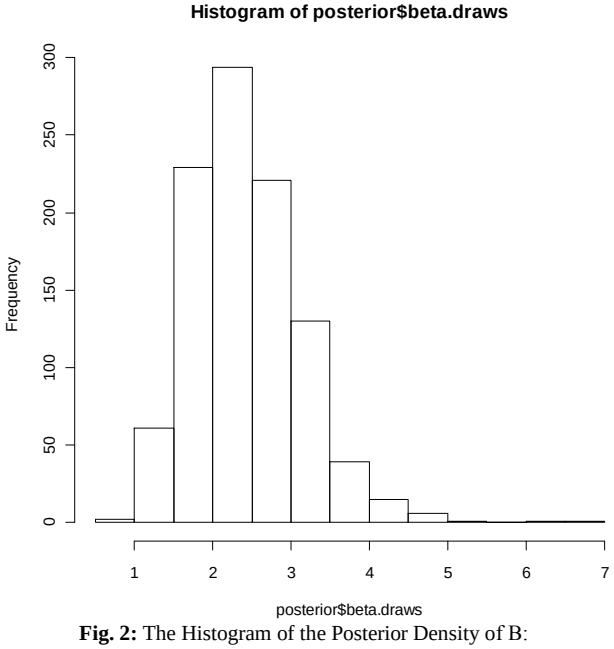



**Fig. 2:** The Histogram of the Posterior Density of B:

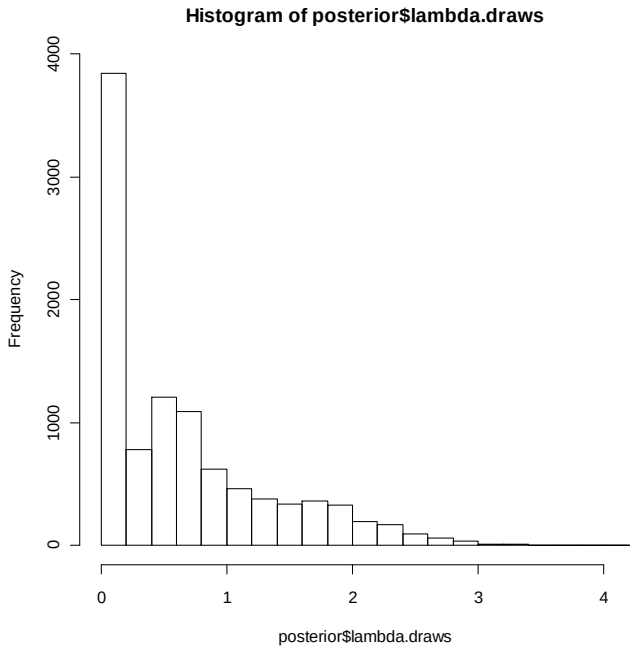
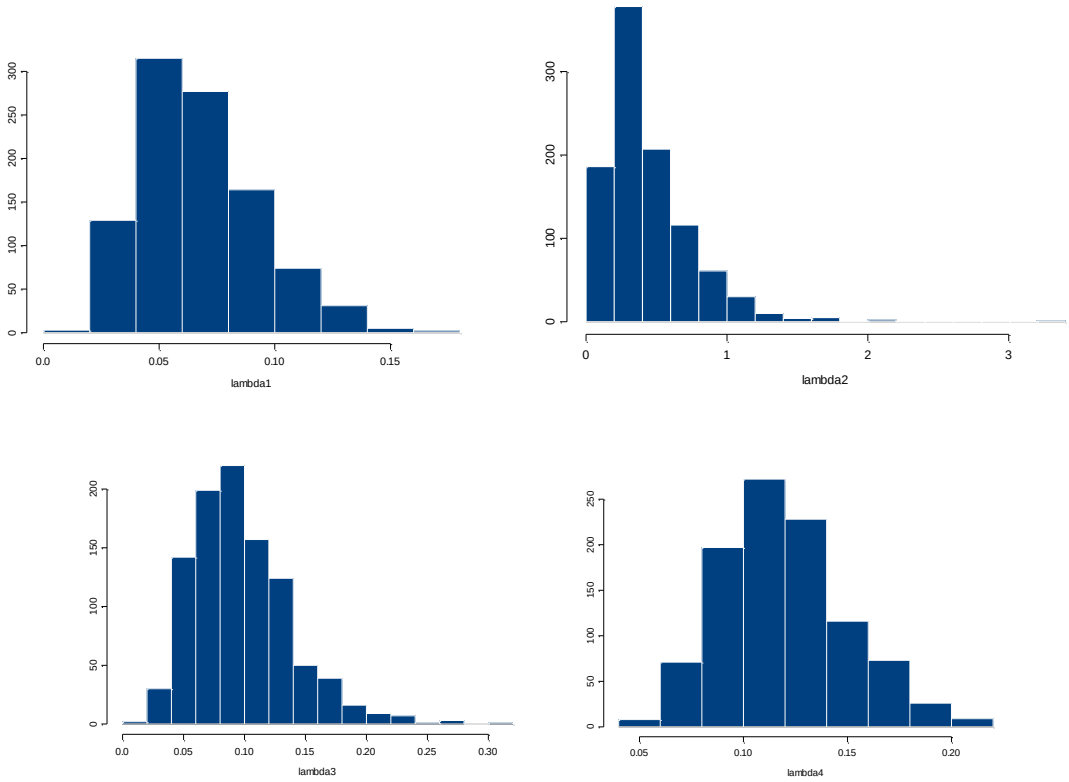



**Fig. 3:** The Histogram of the Posterior Density of Number of Defalts:

## 4.3. Example 3

In this example, we will have evaluated data related to pomp in second example again. Histogram of produced data from posterior distribution that they are obtained in Slice method by S-plus as follows:



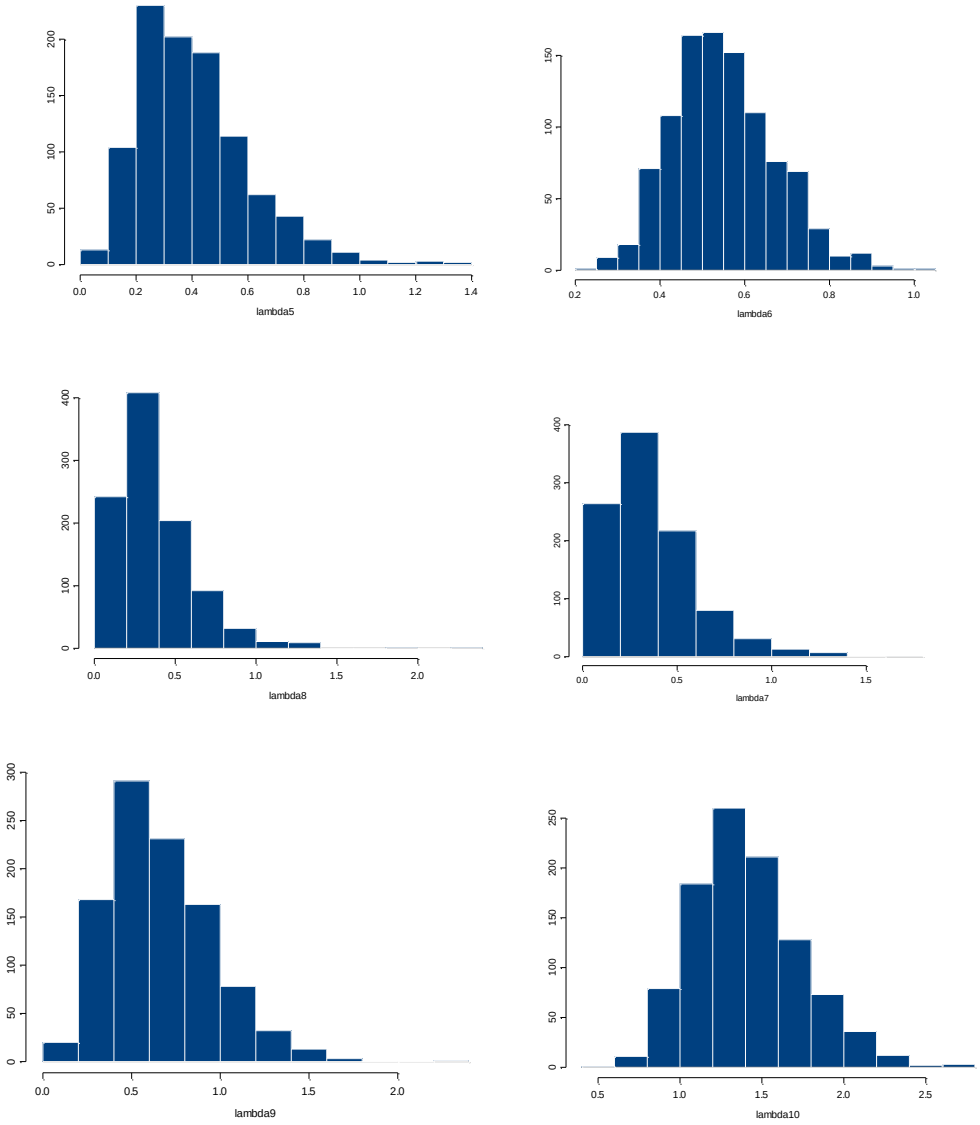



**Fig. 4:** Output Example 3: Histogram of Posterior Densities:

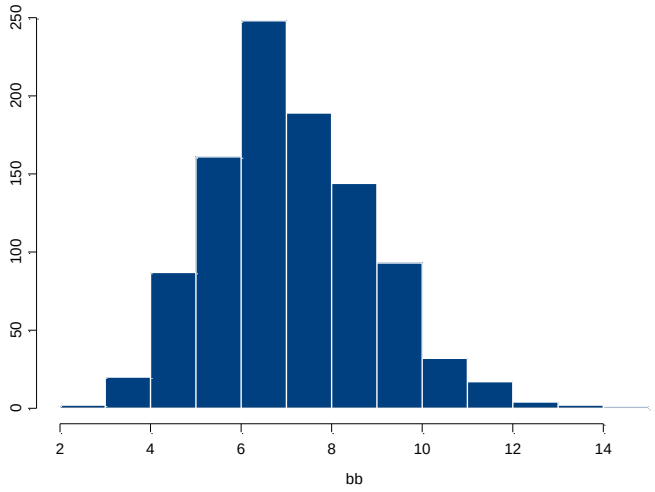



**Fig. 5:** Histogram of Posterior Density of B

## 4.4. Example 4

In this example, we will simulate N (0, 0, 1, 1, ρ) by using Gibbs sampling.

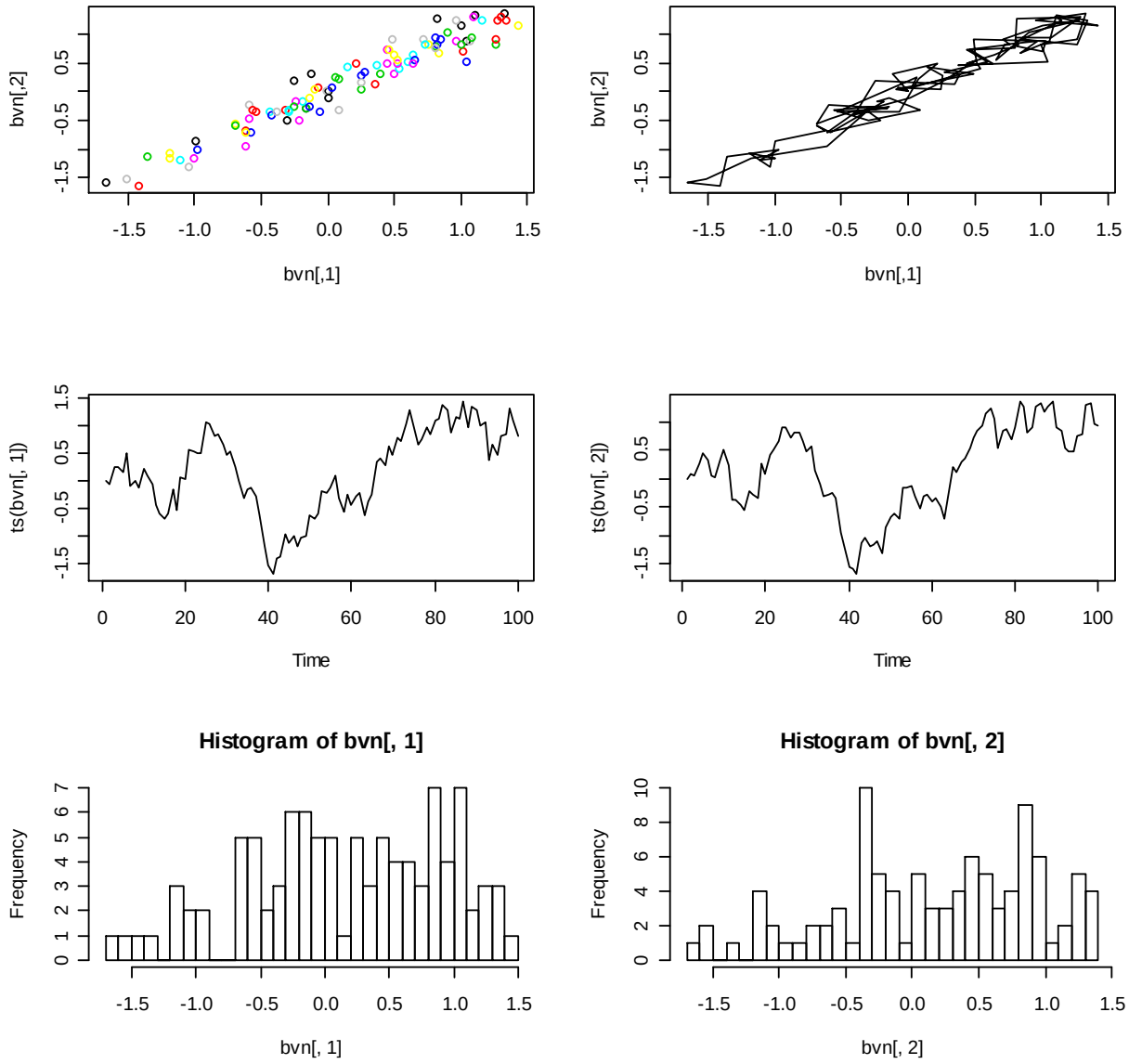



**Fig. 5:** The Histogram of Marginal Densities in R Software is as Follows

## 4.5. Example 5

This example is related to resolve 4th example by using Metropolis-Hastings algorithm.

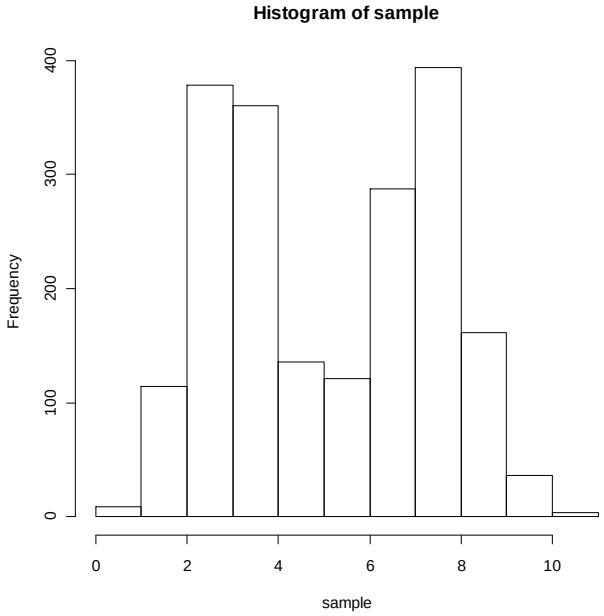



**Fig. 6:** Histogram of this Example by R Software is as Follows

## 4.6. Example 6

In this example, we will generate 500 samples to simulate means of normal mixture models by Gibbs sampler and with means 2.9, 4.5 and standard deviation 0.1.
So in this example, each of normal models has weight 1/2, that in fact show share of any normal model in Gaussian Mixture Models.

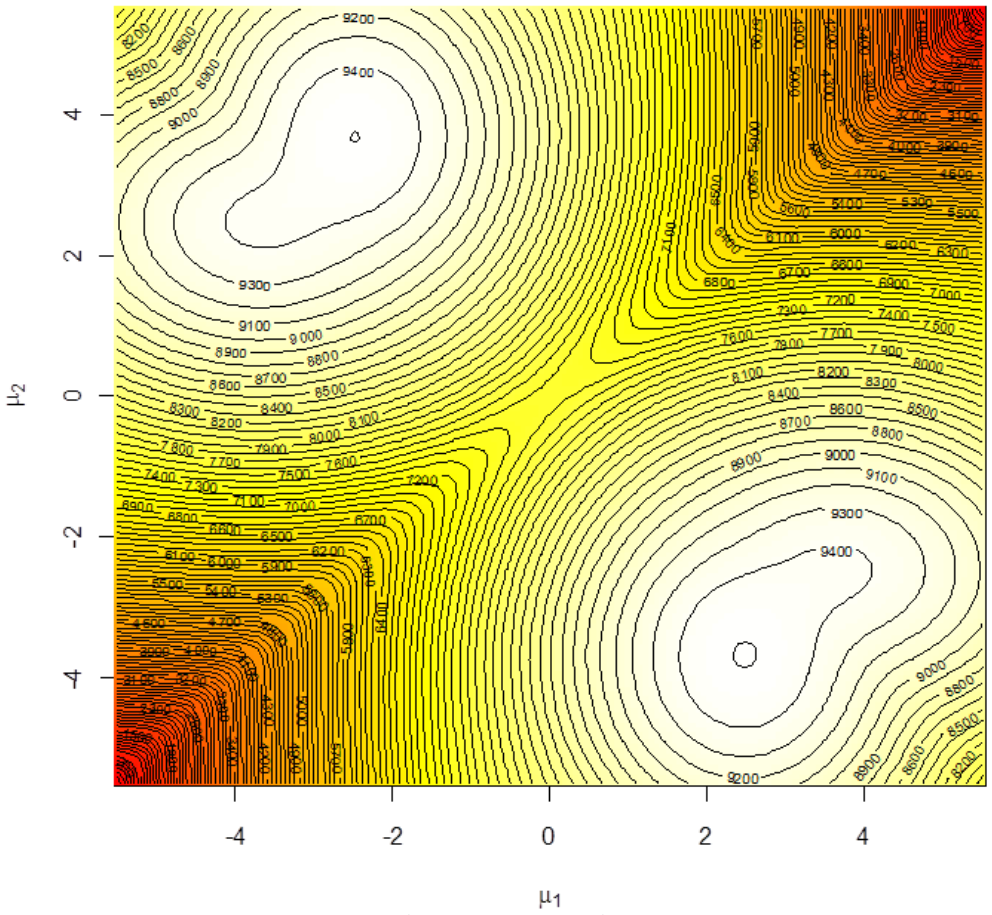



**Fig. 7:** Histogram of this Example in R Software is as Follows

# References


[1] Geman,strut & Geman,Donald," Stochastic Relaxation,gibbs distribution,and the Bayesian Restoration of images", IEEE Transactions on pattern analysis and machin intelligence, November, vol.6, (1984), p:721-741.

[2] Gelfand,A.and Smith,A.F,"Sampling Based approaches to calculating marginal densities", Jornal of the American Statistical Association,vol85, (1990), pp:398-409. http://dx.doi.org/10.1080/01621459.1990.10476213.

[3] Neal,R.M," Slice sampling (with discution)", Annals of Statistics,vol31, (2003), pp:705-767. http://dx.doi.org/10.1214/aos/1056562461.

[4] Metropolice,N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, E, "Equations of state calculations by fast computing Machienes", Jornal of Chemical Physics, (1953), pp:1087-1092. http://dx.doi.org/10.1063/1.1699114.

[5] Robert, C.P. and Casella, G., "Monte Carlo Statistical Methods", 2nd ed, Springer, Newyork, (2005).

[6] Hastings, D." Monte Carlo methods using Markov Chains and their application", Biometrika, (1970), pp: 97-109. http://dx.doi.org/10.1093/biomet/57.1.97.

[7] Gaver,D. and O'Muircheartaigh, I.," Robust empirical bayes analyses of eventrates, Technometrics",(1987), pp:1-15. http://dx.doi.org/10.1080/00401706.1987.10488178.

[8] Rubinstein Reuven.y.," Simulation and the Monte Carlo Method", wiley New York, (1981).

[9] Henry Horng-Shing Lu.," Bayesian Methods with Monte Carlo Markov Chains III", Institute of Statistics, National Chiao Tung University.

[10] Casella, G. and George, E.," Explaining the Gibbs Sampler", The Jornal of American statistical Association (JASA), Auguest, vol. 46. No.3, (1992), pp: 167-174.

[11] Robert, G. and Smith AF.M.,"Simple conditions the convergence of the Gibbs sampler and Metropolis Hastings algorithms", Stochastics Processes and their Application, vol 49, (1994), pp:207-216.

[12] Patrick Lam.," MCMC Methods: Gibbs Sampling and the Metropolis-Hastings Algorithm".

[13] Prof. Darren Wilkinson.," A simple Metropolis-Hastings independence sampler", School of Mathematics & Statistics, Newcastle University.

[14] J.Q.shi, R.Murray-Smith and D.M.titerington.," Hierarchical Gaussian Process Mixtures for regression", Springer Science + Business Media,Inc.Manufactured in the Netherlands.Statistics and Computing vol 15, .(2005), pp:31-41.