

# A Technique of Bayesian Parameter Estimation Based on AEPD

Dongsheng Li<sup>1</sup>, Youfu Wu<sup>2\*</sup>

<sup>1</sup>School of Data Science and Information Engineering, Guizhou Minzu University, Guiyang Guizhou

<sup>2</sup>Guizhou Vocational and Technical College of Communications, Guiyang Guizhou

Email: 465164730@qq.com, \*2737873590@qq.com

Received: Jul. 13<sup>th</sup>, 2020; accepted: Jul. 27<sup>th</sup>, 2020; published: Aug. 4<sup>th</sup>, 2020

---

## Abstract

In this paper, based on the asymmetric exponential power distribution (AEPD), the M-H algorithm in MCMC is used to estimate the Bayesian parameters. For the complexity of the posterior density, the recommended distribution of the random walk chain is the center symmetric distribution. Through numerical simulation, N iterations are performed under M large samples. The simulation results show that the algorithm is practical for Bayesian parameter estimation with asymmetric exponential power distribution.

## Keywords

Asymmetric Exponential Power Distribution, Bayesian Parameter Estimation, Metropolis-Hastings Algorithm

---

# 一种基于AEPD的贝叶斯参数估计方法

李东升<sup>1</sup>, 吴有富<sup>2\*</sup>

<sup>1</sup>贵州民族大学, 数据科学与信息工程学院, 贵州 贵阳

<sup>2</sup>贵州交通职业技术学院, 贵州 贵阳

Email: 465164730@qq.com, \*2737873590@qq.com

收稿日期: 2020年7月13日; 录用日期: 2020年7月27日; 发布日期: 2020年8月4日

---

## 摘要

本文基于非对称指数幂分布(AEPD), 采用MCMC中的M-H算法进行贝叶斯参数估计, 针对后验密度的复杂度。  
\*通讯作者。

杂性, 采取随机游走链的建议分布为中心对称分布。通过数值模拟, 在M个大样本下进行N次迭代, 模拟结果显示, 该算法对具有非对称指数幂分布的贝叶斯参数估计具有实用性。

## 关键词

非对称指数幂分布, 贝叶斯参数估计, Metropolis-Hastings算法

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着社会的发展, 数据的复杂性越来越大。在经济领域, 已经很难找到服从正态分布的数据集, 基于正态分布所建立的一些模型方法也不再适用于对复杂数据的分析。2009年由 Zhu Dongming [1]等提出的非对称指数幂分布(Asymmetric Exponential Power Distribution, AEPD)对数据的适应性更强, 该分布是一个新的分布族, 比如正态分布、均匀分布都是 AEPD 的特例。

基于此, 针对 AEPD 的贝叶斯参数估计采用 MCMC 中的 M-H 算法作进一步探讨。首先, 1995年 Chib [2]详细介绍了 Metropolis-Hastings 算法理论, 并对该算法在应用上做了讨论, 即使在特殊情况下, 该算法也能体现出它的优良性; 其次, 2002年 Aans Altaleb [3]使用 MCMC 中的 M-H 算法在 Logit 回归模型进行参数估计, 以及 2003年 John Geweke [4]采用独立链、随机游走链和泰勒链对贝叶斯状态空间模型进行估计; 最后, 2015年 David D.L. Minh [5]对 M-H 算法进行两种互补和直观的方式来进行推导和理解该算法, 并得出该算法具有较高的接受概率。

本文以非对称指数幂分布为基础, 利用 MCMC 中的 M-H 算法思想, 对此进行贝叶斯参数估计。

## 2. Metropolis-Hastings 算法

在统计领域中主要有两大计算方法, 一种是极大似然估计计算, 另一种是 Bayes 计算。在 Bayes 计算中, 最重要的算法就是 Markov chain Monte Carlo (MCMC)中的 Metropolis-Hastings 和 Gibbs 算法。在本文中主要研究 Metropolis-Hastings 在贝叶斯参数估计的应用。

在统计学中, 对连续变量求数学期望, 公式如下:

$$E(X) = \int_{-\infty}^{\infty} xp(x)dx \quad (1)$$

其中,  $X$  是连续随机变量,  $p(x)$  为随机变量  $X$  的密度函数。

假设我们要计算某分布函数  $F(x)$  关于概率密度函数  $\pi(x)$  的数学期望, 根据公式(1)可得:

$$E(F) = \int_x F(x)\pi(x)dx \quad (2)$$

其中  $F(x)$  和  $\pi(x)$  较为简单, 可以由上式求出。但是, 在实践中, 由分布函数或概率密度函数推出的后验分布或后验密度函数都是复杂、高维且非标准形式的, 对此就要进行 MCMC 计算。

在实际中, 假设  $\pi(x)$  是未知的, 我们可以计算归一化常数范围内的密度来替换  $\pi(x)$ , 公式如下:

$$\begin{aligned} \pi(x) &= Cg(x) \\ C &= 1/\int_x g(x)dx \end{aligned} \quad (3)$$

其中,  $g(x)$  是先验分布乘以似然函数,  $C$  是归一化常数。

在贝叶斯统计中, Metropolis-Hastings 是最常用的算法。假设  $X$  的分布依赖我们想要估计的参数  $\vartheta$ , 对于  $F(x) \geq 0$ , 该分布的似然函数为  $L(x|\vartheta)$ , 根据(1)、(2)可知, 则数学期望为:

$$E(F(X)) = \int L(x|\vartheta)F(X)dx \quad (4)$$

在贝叶斯推断中, 统计模型的观测值和参数变量都是随机的。给定样本数据  $X = (x_1, x_2, \dots, x_n)$ 、统计模型参数  $\vartheta$  和先验分布  $p(\vartheta)$ , 根据贝叶斯定理, 模型的参数联合后验密度为:

$$p(\vartheta|x) = \frac{L(x|\vartheta)p(\vartheta)}{\int L(x|s)p(s)ds} = C_x L(x|\vartheta)p(\vartheta) \quad (5)$$

其中,  $L(x|\vartheta)$  为统计模型的似然函数,  $C_x$  依赖于  $X$  的归一化常数, 在 M-H 算法中, 可以不用计算  $C_x$ , 则该参数联合后验密度可表示为:

$$p(\vartheta|x) = L(x|\vartheta)p(\vartheta) \quad (6)$$

在 M-H 算法中, 选择一个建议分布函数对(6)进行参数更新, 建议分布如下:

$$q(\vartheta, x) = \pi(x) \quad (7)$$

$q(\vartheta, y)$  是与目标分布  $p(\vartheta|x)$  较接近的已知分布, 该建议分布称为备选生成密度, 应满足:

$$\int q(\vartheta, x)dx = 1 \quad (8)$$

在进行 M = H 算法时, 应选择一个接受概率  $\alpha(\vartheta, \vartheta')$  使相应的  $q(\vartheta, y)$  以  $p(\vartheta|x)$  为其平稳分布, 最常用的选择是:

$$\alpha(\vartheta, \vartheta') = \min \left\{ \frac{p(\vartheta'|x)q(\vartheta', x)}{p(\vartheta|x)q(\vartheta, x)}, 1 \right\} \quad (9)$$

### 3. 非对称指数幂分布贝叶斯参数估计

在经济领域, 很难找到服从正态分布的数据, 往往都是尾重或者带有偏度的数据。如果采取假设正态分布的方法进行研究, 那么对数据的分析就会稍差一些, 很难完全解读数据带给我们的信息。2009年 Zhu Dongming [1]提出了非对称指数幂分布(Asymmetric Exponential Power Distribution, AEPD), 该分布对于数据的适应性很强, 其概率密度函数为:

$$f_{\text{AEP}}(y|\mu, \sigma, \alpha, p_1, p_2) = \begin{cases} \frac{1}{\sigma} \exp \left( - \left| \frac{y-\mu}{\alpha\sigma/\Gamma(1+p_1)} \right|^{p_1} \right), & y \leq \mu \\ \frac{1}{\sigma} \exp \left( - \left| \frac{y-\mu}{(1-\alpha)\sigma/\Gamma(1+p_2)} \right|^{p_2} \right), & y > \mu \end{cases} \quad (10)$$

其中,  $\Gamma(x)$  为 gamma 函数,  $0 < \alpha < 1$  为偏度参数,  $\sigma > 0$  为尺度参数,  $p_1 > 0$  和  $p_2 > 0$  分别控制左尾和右尾参数,  $-\infty < \mu < +\infty$  为位置参数。当  $\alpha = \frac{1}{2}$  和  $p_1 = p_2$  时, AEPD 分布为对称的指数幂分布; 当  $\alpha = \frac{1}{2}$  和  $p_1 = p_2 = 2$  时, AEPD 分布为标准的正态分布; 当  $p_1 = p_2 = 1$  时便是 ALD 分布;  $p_1(p_2)$  值越小, 则左尾(右尾)越厚,  $\alpha$  值越大, 则分布呈现为左偏。

图 1 给出了非对称指数幂分布(AEPD)在参数  $\mu=0, \sigma=1, p_1=2, p_2=2$  下不同偏度参数  $\alpha$  的概率密度函数。从图中可以看出, 随着  $\alpha$  的不断增大, 该分布呈现左偏现象。当  $\alpha=0.5$  时, 该分布表现为正态分布, 可以看到, 非对称指数幂分布随着不同参数的变化, 可以呈现很多常见的分布, 比如正态分布、非对称拉普拉斯分布等。

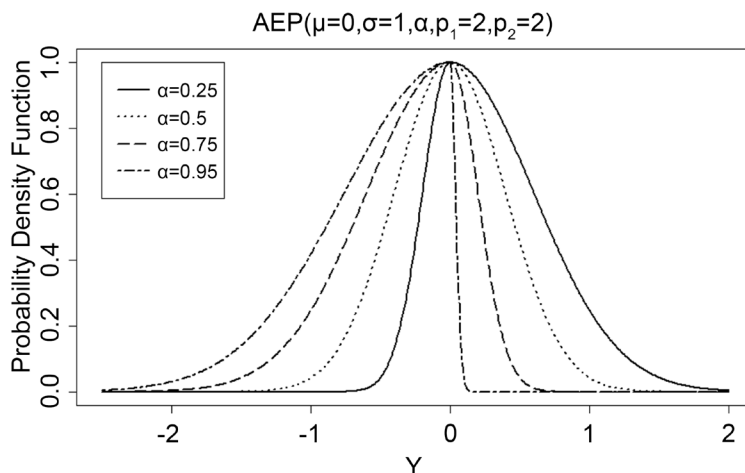


Figure 1. AEPD probability density function graph as a function of  $\alpha$   
 图 1. 随  $\alpha$  变化的 AEPD 概率密度函数图

图 2 给出了非对称指数幂分布在参数  $\mu=0, \sigma=1, \alpha=0.5, p_2=2$  下随着  $p_1$  变化的图像。可以看到当  $p_1$  不断减小时, 该分布呈现左尾厚重现象。当  $p_1=2$  时, 该分布呈现正态分布现象。

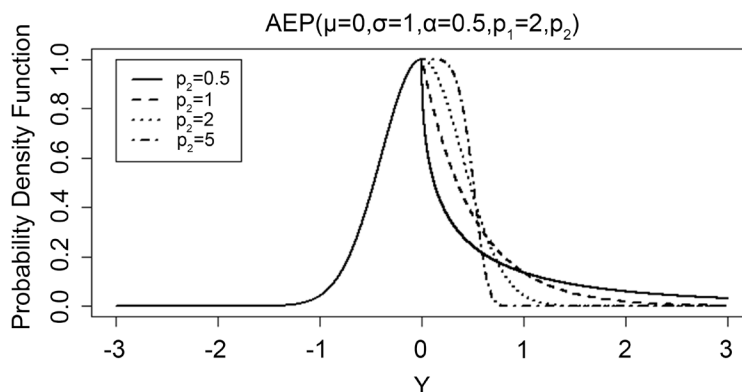


Figure 2. AEPD probability density function as a function of  $p_1$   
 图 2. 随  $p_1$  变化的 AEPD 概率密度函数图

图 3 给出了非对称指数幂分布在参数  $\mu=0, \sigma=1, \alpha=0.5, p_1=2$  下随着  $p_2$  变化的图像。可以看到当  $p_2$  不断减小时, 该分布呈现右尾厚重现象。当  $p_2=2$  时, 该分布呈现正态分布现象。

假设总体的密度函数为  $f_{AEP}(y, \vartheta)$ , 其中  $\vartheta = (\mu, \sigma, \alpha, p_1, p_2)$  为未知参数向量,  $y_i (i=1, 2, \dots, n)$  是来自该总体的样本, 由(10)可得, 该密度函数的似然函数为:

$$L(y|\vartheta) = \prod_i \frac{1}{\sigma} \left[ \exp\left(-\left(\frac{(\mu - y_i)\Gamma(1+1/p_1)}{\alpha\sigma}\right)^{p_1}\right) + \exp\left(-\left(\frac{(y_i - \mu)\Gamma(1+1/p_2)}{(1-\alpha)\sigma}\right)^{p_2}\right) \right] \quad (11)$$

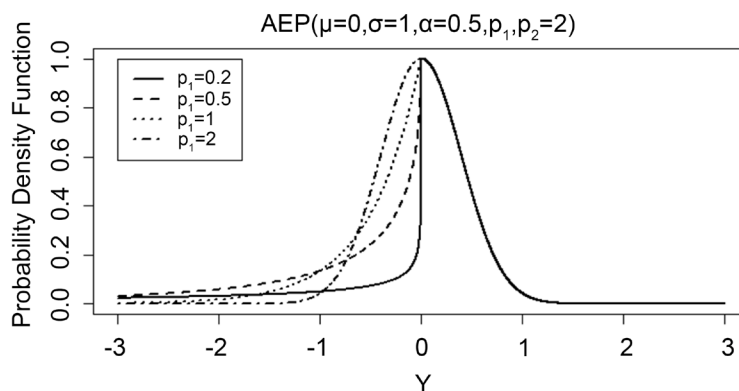


Figure 3. AEPD probability density function as a function of  $p_2$

图 3. 随  $p_2$  变化的 AEPD 概率密度函数图

对数似然函数为:

$$\ln L(y|\mathcal{G}) = -n \ln \sigma - \sum_i^n \left[ \frac{\Gamma\left(1 + \frac{1}{p_1}\right)(\mu - y_i)}{\alpha \sigma} \right]^{p_1} I(y_i \leq \mu) - \sum_i^n \left[ \frac{\Gamma\left(1 + \frac{1}{p_2}\right)(y_i - \mu)}{(1-\alpha)\sigma} \right]^{p_2} I(y_i > \mu) \quad (12)$$

在贝叶斯推断中, 选择无信息先验分布  $p(\mathcal{G}) \propto 1$ 。根据(6)可得联合后验分布为:

$$p(\mathcal{G}|y) \propto L(y|\mathcal{G})p(\mathcal{G}) \propto \frac{1}{\sigma^n} \left[ \exp\left\{-\left(\frac{\sum_i^n (\mu - y_i) \Gamma(1+1/p_1)}{\alpha \sigma}\right)^{p_1}\right\} + \exp\left\{-\left(\frac{\sum_i^n (y_i - \mu) \Gamma(1+1/p_2)}{(1-\alpha)\sigma}\right)^{p_2}\right\} \right] \quad (13)$$

其中,  $y = y_1, \dots, y_n$  为观测值。由于联合后验分布的复杂性, 要想对参数进行贝叶斯估计, 采用 MCMC 方法进行模拟。MCMC 方法的思想就是构造一个平稳分布对目标分布进行后验抽样。根据(7)、(8)、(9)进行 M-H 算法, 对参数进行贝叶斯估计。

#### 4. 数值模拟

非对称指数幂分布是一个新的分布族, 随着参数的变化, 呈现出不同的分布, 比如正态分布、均匀分布都是该分布的一个特例。对此, 利用随机模拟生成 1000 个服从正态分布的数据, 即  $y_i \sim N(0,10)$ 。利用这些模拟数据来进行非对称指数幂分布的贝叶斯参数估计。在进行贝叶斯参数估计的过程中, 用到 MCMC 中的 M-H 算法, 具体步骤如下:

对于 M-H 算法中的建议分布, 针对非对称指数幂分布, 采用正态分布作为建议分布, 在 R 语言中将从建议分布中产生的第一个值设为初始化链条的  $\mathcal{G}_1$ 。

- 选择建议分布  $q(\mathcal{G}, y_i)$  为正态分布的密度, 设  $i=2$ , 迭代 b、c、d 步;
- 从建议分布  $N(\mathcal{G}_{i-1}, 2)$  中生成  $\mathcal{G}'_i$ ;

c) 令  $\mathcal{G}'$  的接受概率为公式(9), 其中, 选择  $q(\mathcal{G}, y) = q(\mathcal{G}', y)$  为随机游走的 M 算法;

d) 产生一个服从均分布  $U(0,1)$  的随机数  $u$ , 令  $\mathcal{G}_{i+1} = \begin{cases} \mathcal{G}_i, u \leq \alpha(\mathcal{G}, \mathcal{G}') \\ \mathcal{G}_i, u > \alpha(\mathcal{G}, \mathcal{G}') \end{cases}$ ;

e) 令  $i = i + 1$ , 返回 a 步。

表 1 给出了采用 MCMC 中的随机游走 M-H 算法, 将 1000 个数据样本进行 10,000 次迭代, 剔除前 1000 个样本后的后验均值及 95% 的置信区间, 这表明 M-H 算法在贝叶斯估计中的良好表现, 解决了贝叶斯估计中高维数据处理困难的问题。从表中可以看到在不同的分位点每个待估参数不同的模拟估计值, 随着分位点的增大, 模拟的估计值也在逼近真值, 例如在分位点  $\alpha = 0.95$  时, 待估参数的真值为:  $\mu = 0$ ,  $\sigma = 1$ ,  $p_1 = 0.5$ ,  $p_2 = 2$ , 相对应的贝叶斯参数估计模拟值为:  $\mu = 0.002295$ ,  $\sigma = 1.002321$ ,  $\alpha = 0.948325$ ,  $p_1 = 0.502165$ ,  $p_2 = 2.001092$ 。在表中也可以看到每个待估参数在各个分位点的 95% 的置信区间。

图 4、图 5、图 6、图 7 和图 8 中给出了每个待估参数的模拟轨迹图、参数模拟值直方图、不同分位点轨迹图和自相关图。例如, 图 4 中给出了  $\mu$  相关的轨迹图和检验收敛图, 从左往右依次为: 剔除前 1000 个样本估计值的轨迹图, 图中可以看到, 均值  $\mu$  的马尔科夫链收敛; 第二个图为参数模拟值的直方图, 可以看到抽到的估计值基本都在真值周围; 第三个图为在 0.025、0.5 和 0.975 分位数下的轨迹图, 从图中可以看到在样本 1500 以后都趋于收敛, 其中上下两条虚线分别代表 0.025 和 0.975 分位数线的变化, 实线代表 0.5 分位数的变化; 第四个图为链的等价自相关函数图, 从图中可以看到, 最终的自相关度都趋于收敛。

**Table 1.** Simulated estimates of Bayesian parameters of AEP distributed at different quantiles

**表 1.** AEP 分布在不同分位点的贝叶斯参数模拟估计值

分位点	参数	真值	估计值	95%置信区间
0.25	$\mu$	0	0.004406	[-0.004110, 0.035680]
	$\sigma$	1	1.004956	[0.997341, 1.036673]
	$\alpha$	0.25	0.252058	[0.243971, 0.270394]
	$p_1$	0.5	0.503658	[0.495862, 0.528231]
	$p_2$	2	1.998420	[1.988583, 2.002509]
	0.5	$\mu$	0	0.003578
$\sigma$		1	1.004006	[0.997315, 1.030708]
$\alpha$		0.5	0.501049	[0.493321, 0.514576]
$p_1$		0.5	0.502986	[0.495943, 0.523567]
$p_2$		2	1.998543	[1.988976, 2.002540]
0.75		$\mu$	0	0.002511
	$\sigma$	1	1.002804	[0.997192, 1.022655]
	$\alpha$	0.75	0.749676	[0.741728, 0.756592]
	$p_1$	0.5	0.502118	[0.496191, 0.518529]
	$p_2$	2	1.999064	[1.990346, 2.004054]
	0.95	$\mu$	0	0.002295
$\sigma$		1	1.002321	[0.996092, 1.021065]
$\alpha$		0.95	0.948325	[0.939183, 0.950101]
$p_1$		0.5	0.502165	[0.496074, 0.518774]
$p_2$		2	2.001092	[1.993687, 2.013658]

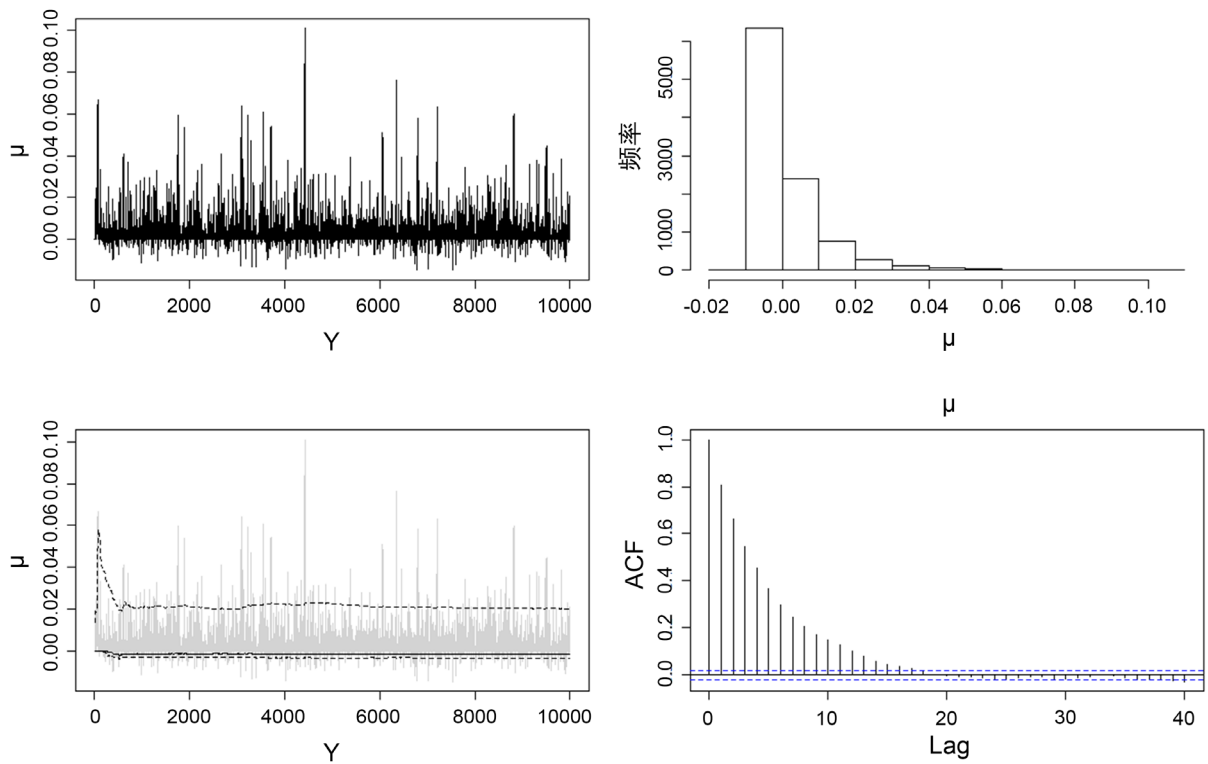


Figure 4.  $\mu$ -parameter Bayesian estimate MCMC related image

图 4.  $\mu$  参数贝叶斯估计值 MCMC 相关图像

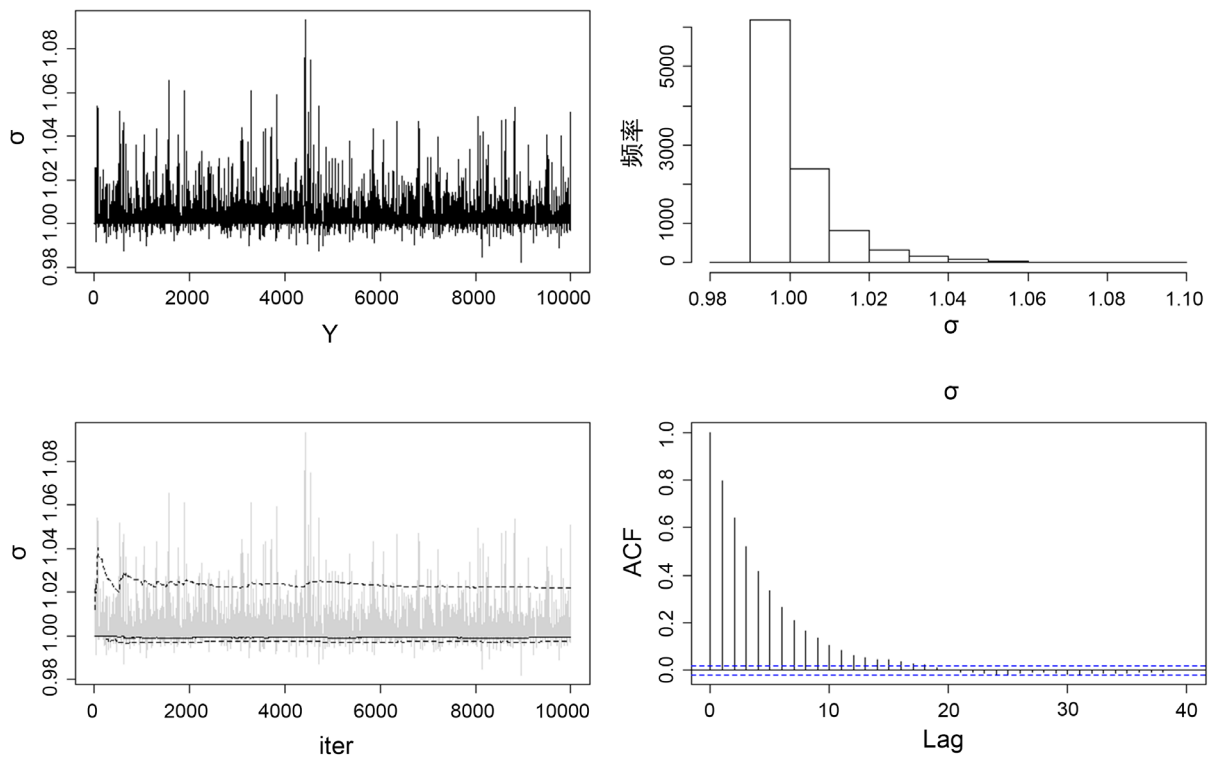


Figure 5.  $\sigma$ -parameter Bayesian estimate MCMC related image

图 5.  $\sigma$  参数贝叶斯估计值 MCMC 相关图像

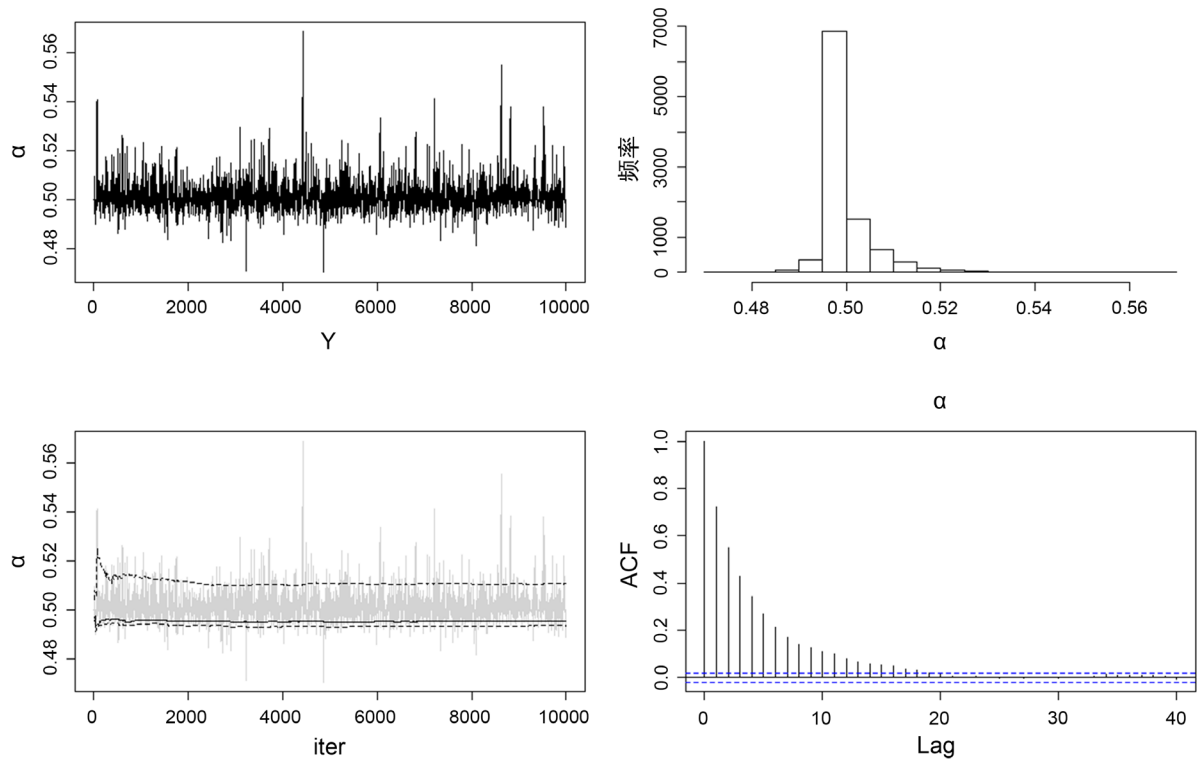


Figure 6.  $\alpha$ -parameter Bayesian estimate MCMC related image  
 图 6.  $\alpha$  参数贝叶斯估计值 MCMC 相关图像

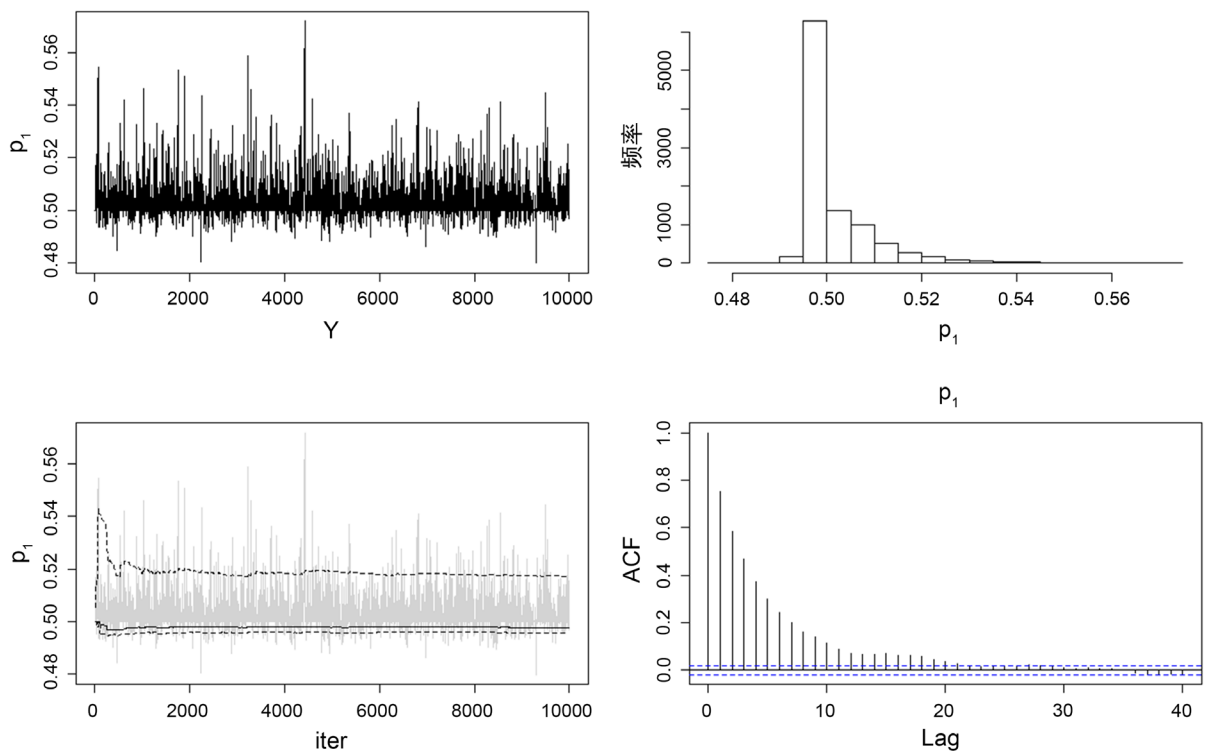


Figure 7.  $p_1$ -parameter Bayesian estimate MCMC related image  
 图 7.  $p_1$  参数贝叶斯估计值 MCMC 相关图像



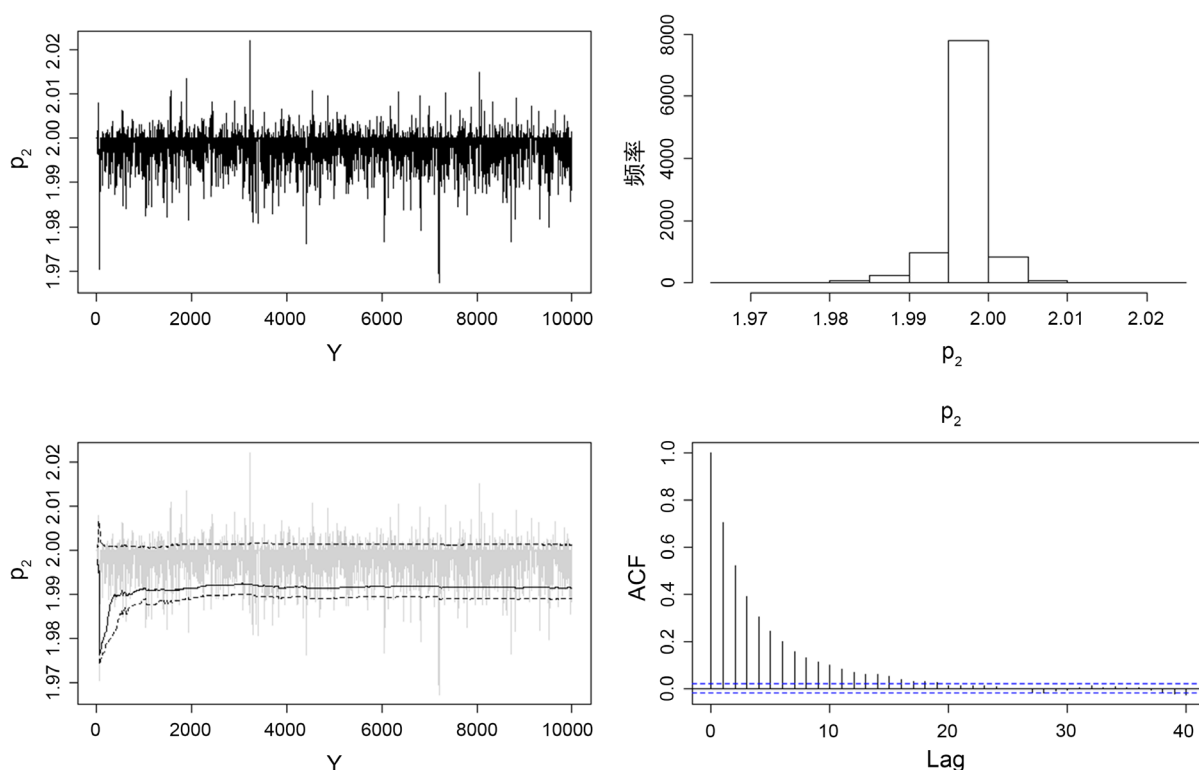


Figure 8.  $p_2$ -parameter Bayesian estimate MCMC related image

图 8.  $p_2$  参数贝叶斯估计值 MCMC 相关图像

## 5. 结论

本文针对非对称指数幂分布后验密度的复杂性, 采用 MCMC 中的 M-H 算法对参数后验密度进行抽取。对 AEP 分布进行贝叶斯参数估计, 假定先验分布为无信息先验, 利用随机游动的 M-H 算法, 对待估参数进行后验密度抽取, 从 MCMC 模拟结果可以看到, 各个分位点待估参数估计值的 MCMC 迭代轨迹图均收敛, 这表明随机游动 M-H 算法很好的模拟了非对称指数幂分布的后验密度; 得到的待估参数估计值与真值非常接近, 说明该算法对具有非对称指数幂分布的数据集来说非常适用。

## 参考文献

- [1] Zhu, D. and Zinde-Walsh, V. (2009) Properties and Estimation of Asymmetric Exponential Power Distribution. *Journal of Econometrics*, **14**, 86-99. <https://doi.org/10.1016/j.jeconom.2008.09.038>
- [2] Chib, S. and Greenberg, E. (1995) Understanding the Metropolis-Hastings Algorithm. *The American Statistician*, **49**, 327-335. <https://doi.org/10.1080/00031305.1995.10476177>
- [3] Altaleb, A. and Chauveau, D. (2002) Bayesian Analysis of the Logit Model and Comparison of Two Metropolis-Hastings Strategies. *Computational Statistics & Data Analysis*, **39**, 137-152. [https://doi.org/10.1016/S0167-9473\(01\)00055-X](https://doi.org/10.1016/S0167-9473(01)00055-X)
- [4] Geweke, J. and Tanizaki, H. (2003) Note on the Sampling Distribution for the Metropolis-Hastings Algorithm. *Communications in Stats*, **32**, 775-789. <https://doi.org/10.1081/STA-120018828>
- [5] Minh, D.D.L. and Minh, D.L. (2015) Understanding the Hastings Algorithm. *Communications in Statistics, B. Simulation and Computation*. <https://doi.org/10.1080/03610918.2013.777455>