2015-01-01

A Bayes Approach in Step-stress Accelerated Life Testings

Hao Yang Teng
University of Texas at El Paso, hyteng@miners.utep.edu

A BAYES APPROACH IN STEP-STRESS ACCELERATED LIFE TESTINGS

Hao Yang Teng

Department of Mathematical Sciences

APPROVED:

Naijun Sha, Ph.D., Chair

Ori Rosen, Ph.D.

Yirong Lin, Ph.D.

Charles Ambler, Ph.D.
Dean of the Graduate School

to my

MOTHER and FATHER

with love

A BAYES APPROACH IN STEP-STRESS ACCELERATED LIFE TESTINGS

by

Hao Yang Teng

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

Department of Mathematical Sciences

THE UNIVERSITY OF TEXAS AT EL PASO

May 2015

Acknowledgements

I would like to express my deep gratitude to my advisor, Dr Naijun Sha from the Department of Mathematical Sciences at the University of Texas at El Paso for his advice, encouragement, patience and utmost support. He has never failed to inspire me in many ways. The way he conducted our thesis meeting sessions is unique. He often asked me to present articles I read and gave me a lot of guidance to improve how I convey my thoughts to the intended party clearly and concisely. In addition, he also encouraged me to sign up for workshops which aided in my thesis writing. In short, he not only helped me to improve my analytic and theoretical aspects but also in character development to be a well-organized, punctual and disciplined person.

I also wish to thank the other members of my committee, Dr Ori Rosen from the Department of Mathematical Sciences and Dr Yirong Lin from the Department of Mechanical Engineering at University of Texas at El Paso. Dr Rosen gave me numerous advice and guidance throughout my masters courses and also on my thesis. Besides, Dr Lin suggested some improvements on conveying some concepts clearly and presenting my slides. Their guidance were invaluable to the completion of this work.

NOTE: This thesis was submitted to my Supervising Committee on the May 21, 2015.

Abstract

A Bayesian analysis for the Weibull proportional hazard (PH) model is presented. A comparison between the Weibull PH model and the Weibull cumulative exposure (CE) model is made graphically and mathematically. The PH model is as flexible as the CE model in fitting step-stress data and the mathematical form of the PH model enables researchers to do Bayesian inference much easier than the CE model. In addition, the PH model has the desirable proportional hazard property. A convex tent prior is used for Bayesian analysis. Markov chain Monte Carlo methods are used for posterior inferences. In this study, we adopt two sampling methods. The first way is to draw samples of the unknown parameters individually from each of the conditional posterior density functions. Another way is to perform joint sampling of the unknown parameters from the joint posterior density function. The performance of the two sampling methods is investigated using both simulated and real data sets.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this chapter, a brief introduction is provided on this study. This study applies survival analysis on the study of Accelerated Life Testing (ALT) which is widely used in the engineering industry. Some basic concepts regarding survival analysis will be presented in the following section to give a preliminary knowledge on the survival concepts and help understand the later parts of the study. In chapter 2, the cumulative exposure model and proportional hazard model are introduced and compared. The proportional hazard model is superior to the cumulative model in many ways. Therefore, we adopt the proportional hazard model by assuming that the failure time of the products follows a Weibull distribution in this study. In subsequent chapters, we begin our Bayesian analysis by specifying the probability density function and the likelihood function. We believe that the population indexed by θ follows a prior distribution from the convex-tent family. With the data, the prior distribution is updated to be the posterior distribution. In chapter 4, we introduce some Markov chain Monte Carlo (MCMC) sampling methods. Subsequently, the algorithms of the two sampling schemes are presented. In the last chapter, a discussion of the simulation results and real data results using the proposed model will be presented.

1.1 Survival Analysis

In this section, some concepts such as survival function, cumulative distribution function, hazard rate function and cumulative hazard function will be briefly discussed. One of the most common notations used in survival analysis is $S(t)$ which is known as the survival function (Klein and Moeschberger [16]). The survival function represents the survival probability of an individual beyond time t which is defined as

$$S(t) = Pr(T > t) = \int_t^{\infty} f(y)dy, \tag{1.1.1}$$

where the function f is the probability density function of T. $S(t)$ refers to the reliability function when it comes to products and equipment. $S(t)$ is a continuous and decreasing function if T is a continuous random variable. The survival function is related to the cumulative distribution function through $S(t) = 1 - F(t)$ where $F(t) = Pr(T \le t)$.
In other words, the probability density function can be expressed as the differentiation of the survival function which is

$$f(t) = -\frac{dS(t)}{dt}. \tag{1.1.2}$$

There are many different types of survival curves. All of them share the same basic properties such as they are monotone, non-increasing functions. In addition, the survival function is equal to one when t is zero and zero when t approaches infinity. The rate of decrease of the survival function differs according to the experience of the individual at time t.

Hazard rate $h(t)$ which is also commonly used in survival analysis to explain the failure rate of the individual is defined as

$$h(t) = \lim_{\triangle t \to 0} \frac{Pr(t \le T < t + \triangle t | T \ge t)}{\triangle t}. \tag{1.1.3}$$

If T is a continuous random variable, then

$$h(t) = \frac{f(t)}{S(t)}. \tag{1.1.4}$$

The cumulative hazard function $H(t)$ can be derived by integrating the hazard rate function up to time t given as

$$H(t) = \int_0^t h(y)dy = -\ln S(t). \quad (1.1.5)$$

From the equation (1.1.4), an individual surviving up to time t experiencing the event in the next moment can be approximated as $h(t)\triangle t$. Different shapes of hazard rate are used to describe the experience of different individuals. For instance, an individual or test unit would go through natural aging or wearing over time and it is appropriate to adopt models with increasing hazard rate. New products with high failures or patients who undergo surgery will experience higher hazard rates in the beginning and they could be fitted using models with decreasing hazard rates. The most commonly used hazard rate to describe the experience of an individual or test unit is the bathtub-shape hazard. The hazard function has a decreasing hazard rate, followed by a constant hazard rate. The later part of the function will be an increasing hazard rate. This kind of hazard rate function is suitable to describe the population starting from birth or some newly manufactured equipment and products.

Chapter 2

Literature Review

Under normal usage, researchers take a long time to collect data of a product or material. Accelerated life testing (ALT) is a process to shorten the testing period by subjecting the products to more severe conditions, resulting in shorter time in failure data collection. A model is adopted to fit the accelerated failure times obtained from the experiment and used to estimate the life distribution of the products under normal usage.

The design of an ALT is essential for data collection given limited time. For instance, the failure rate of high reliability products has to increase in order to collect more data for the experiment to be optimal or meaningful. As such, the variable which plays a vital role in ALT design is known as the "stress variable" such as voltage, temperature, etc (Dorp and Mazzuchi [36]). There are many different testing scenarios such as a constant stress, step-stress, random stress and progressive stress suitable for different environment.

If the stress is time-independent, the test units are often put under a constant stress. In other words, the test units will be tested under the same stress level throughout the whole experiment. The constant stress is displayed in Figure 2.1.

If the stress is time-dependent, the test units are put under a stress level which varies with time. Tests units under varying stress level will yield failure more quickly as compared to when the test units undergo a constant stress. The step-stress and ramp-stress are typical time-dependent stress tests. The test units will be subjected to stress level x_1 for a period of time. Then, the stress level will be varied to a different stress level and maintained for some time before it is stepped/ramped into another. The step-stress and ramp-stress are illustrated in Figure 2.2 and Figure 2.3 respectively.

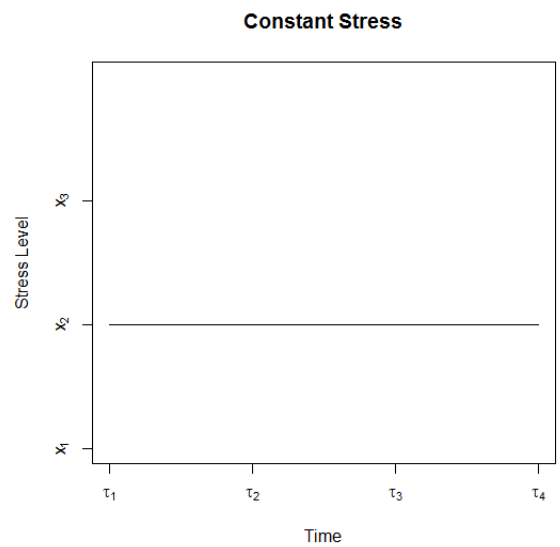

Figure 2.1: Constant Stress Loading

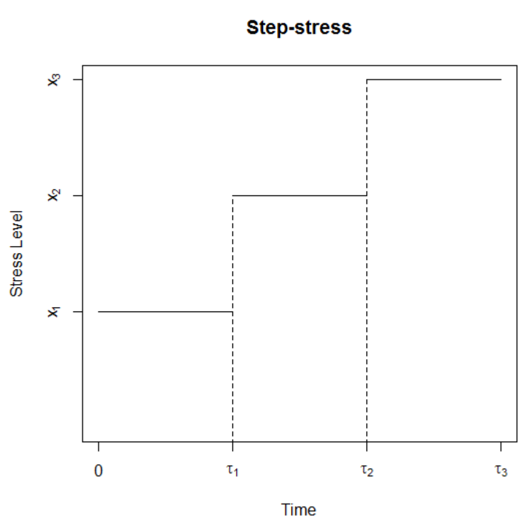

Figure 2.2: Step-stress loading

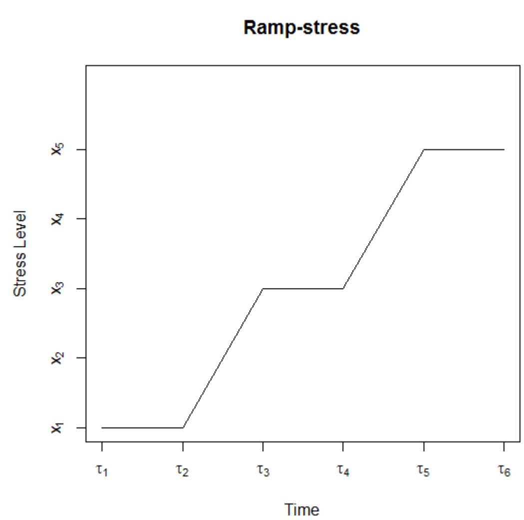

Figure 2.3: Ramp-stress loading

There are many research papers which make comparisons between the constant stress and step-stress with step-stress being more useful and advantageous (Wang and Fei [39]). For instance, in constant stress testing, it is expected that there is censored data at lower level of stress while for step-stress setting, there is few or no censoring. Khamis [15] presented a comparison between a constant stress testing and step-stress testing for Weibull models. Based on the simulation results, the step-stress testing is preferred than the constant stress when censored data occurs at lower stress level.

Our study uses step-stress testing because step-stress accelerated life testing (SSALT) is most commonly used in experiments and it has many advantages. Firstly, the use of SSALT can decrease the experiment duration without compromising on the accuracy of the estimates of the lifetime distribution (Zhao and Elsayed [43]). Secondly, in comparison with parallel constant ALT, SSALT requires fewer test units which are more practical (Tseng and Wen [35]). In addition, SSALT is chosen for its adaptability. If the information about the test stresses of a new product is not known, SSALT enables the researchers to adjust the stress level and stopping time as information is gathered over time. Besides, another attractive feature about SSALT is that it is economical. Often, experiments conducted under high stress levels will incur substantial costs. Using SSALT, the varying stress level will help save considerable resources. Tang, Yang and Xie [42] suggested an optimal SSALT degradation plan which does not require many test units and helps shorten testing time thereby saving resources.

Many researches discussed the optimal level of testing using the step-stress to conduct the experiment in a meaningful way. Bai and Chun [3] presented an optimum simple step-stress ALT with different failure causes. The life distribution assumed for each of the steps follows an exponential distribution. The mean of the distribution is related to the stress level through a log-linear function. The paper aims to create an optimum plan which minimizes the asymptotic variances of the maximum likelihood estimators of the log of mean failure time. However, the optimum simple step-stress ALT plans have some limitations such as the stress and time-to-failure are linearly related and using two stresses might not be appropriate in many test testings. An optimum 3-step step-stress plan was proposed as an extension to the linear model (Khamis and Higgins [13]). The log of the scale parameter is related to the stress level through a linear or quadratic function when the assumed life distribution follows exponential distribution. An optimum M-step step-stress test plan which has k stress variables was presented by Khamis [14]. The author used the polynomial model, proposed a goodness of fit test and discussed the relevance of adopting the asymptotic chi-square distribution when the sample size is small. The result showed that the test was conservative when the sample sizes are small but the chi-square approximation can be used for practical purposes. There are other researches presenting different optimum test plans. Nelson and Meeker [25] showed in their paper an optimum test plan which minimized the variance of the maximum likelihood estimate for Weibull and Extreme Value distributions with two test stresses. In addition, they also presented a method to determine the optimum proportion to run with optimum low test stress. The chart from Nelson and Meeker [24] suggests that the majority of the products allocated near design stress for a given stress level. They also provided an estimate for the optimum variance of the extreme value distribution and Weibull distribution. The results showed that at 50th percentile, the optimum variance was smaller than the variance obtained using the actual plan with complete data. The result is similar using Weibull distribution. Moreover, a way to determine the number of test units required for each stress level was also shown. However, the test plans are not practical if there is a misspecification of the unknown parameter values (Meeker [23]).

In general, there are a few methods to model SSALT. One of the ways is known as tampered random variable (TRV) (DeGroot and Goel [9]). The increase of the stress level lowers the future lifetime of the product by a tampering or acceleration factor of k_1 depending on the magnitude of change of the stress and vice versa. SSALT TRV was generalized when the distribution is Weibull under multiple step-stress levels (Zhao and Elsayed [43]). Another way is known as cumulative exposure (CE) model proposed by Nelson [26]. The model takes into account the worn-out conditions of the product accumulated from previous testing periods. The remaining lifetime of the product is subject to the current accumulation fractions failed under current stress level. The failure time of the product follows the cumulative distribution function (CDF) under the current stress by taking into consideration of the previous accumulation fraction failed. Using the tampered failure rate (TFR) model proposed by Bhattacharyya and Soejoeti [5] , the increase of the stress level will have a multiplicative effect on the failure rate of the product. The failure rate of the product for the next time period will increase by a factor of k_2 which is greater than 1 as the stress level increases. Linear cumulative exposure model (LCEM) suggested by Tang [32] and Tang et al. [33] could also be adopted to model SSALT. The model states that the exposure increases linearly under each stress level. Under each stress level, the exposure of a product can be derived by taking the ratio of the total time (from the start of the experiment) to the lifetime of the product at current stress. The unit is considered failed if the cumulative exposure is one

Under certain conditions, two of the models will coincide. If the survival distributions of the product under two different stresses belong to the same scale parameter family, then the TRV and CE models coincide. If the loss factor is set at $k_1 = k_2 = 1$, or the product's lifetime under first stress follows an exponential distribution, the TRV and TFR models coincide (Wang and Fei [39]) .

The limitations of CE, LCE, TRV and TFR model were discussed (Xu and Fei [40]). TRV model cannot be applied if the lifetime of the unit under previous stress cannot be described using a tampering coefficient in terms of the lifetime of the unit under current stress. For instance, if the underlying distribution is Weibull with non-constant shape parameter, the TRV model cannot be used to fit the experience. CE model is not applicable to situations where the underlying distribution has a failure-free life (FFL) with a location parameter (Tang [32]). CE model is used with the assumption that the failure-free life is less than the equivalent cumulative operating time at the current stress level. Otherwise, the cumulative distribution will be zero unless the underlying distribution has a threshold parameter ($F_i(t) = H(\frac{\ln(t-\theta_i)-\mu_i}{\sigma})), i = 1,2$ and the location parameter satisfies $\frac{\theta_2}{\theta_1} = \exp(\mu_2 - \mu_1)$. The LCE model could overcome the limitation of the CE model. However, the limitation of the model stems from the variability in the reliability component R which makes data analysis difficult. In addition, the maximum likelihood estimation provided could be incorrect when the underlying distribution is a Weibull distribution with three parameters, resulting in higher log-likelihood values. The TFR model outperforms the CE model analytically when the underlying distribution follows a Weibull distribution. However, the model has a few limitations. Firstly, the model has no explicit form for the failure rate function. It is not possible to derive the failure rate function analytically for log-normal distribution. Moreover, the TFR model is also not easy to apply as well computationally.

Another model which is also commonly used for step-stress testing is known as proportional hazard (PH) model (Bagdonavicius, Gerville-Reache and Nikulin [1]). The name is derived from having two hazard functions which are multiples of each other. In other words, the change of stress will have a multiplicative effect on the hazard rate. The hazard function for PH model presented by Cox [8] can be expressed in terms of failure time as $h(t;x) = h_0 e^{x\beta}$ where $h_0(t)$ is the baseline hazard function. The parameter β represents the stress effects, x is the stress level and the baseline hazard function does not necessary take specific form. Unlike the PH model, other lifetime distributions do not exhibit such property such as the accelerated failure time models. A generalization of multiple step-stress was presented by Madi [18]. A similar model known as K-H model based on the time-transformation of the exponential CE model was shown in [12]. K-H model is as flexible as the CE model when fitting data. The mathematical forms enables the researchers to obtain the parameter estimates and standard deviations much easier than the CE model. When the failure time follows Weibull distribution, the CE model can be re-parameterized as a PH model under constant stress. However, it is not true for other distributions and stress settings (Lawless [17]). Therefore, PH-SSALT model is essential in engineering applications with the acceleration model being a failure rate function.

In early days, the ALT inference methods were based on maximum likelihood estimation (MLE) which requires large sample sizes to make meaningful inference. In most cases, asymptotic variance plays a vital in determining the optimum design of ALT plans and the reliability of the statistical inferences. McSorley, Lu and Li [21] investigated that 100 sample size are required to adopt the large-sample inference procedures. In addition, when the model is Weibull distribution with constant shape parameter, using exponential models with large sample sizes or using Weibull models with small samples would result in poor model fitting. When the sample sizes are small, Bayesian approach is more suitable for meaningful inferences. This is more practical as it requires substantial costs to carry out the whole experiment in most cases, resulting in fewer test units.

The development of a general Bayes inference model for accelerated life testing is essential as it could cater to varying-stress tests (Van Dorp and Mazzuchi [36] and [37]). The life time distribution is assumed to follow a Weibull distribution without any assumption regarding a parametric time transformation function which is typically specified in most research papers (Van Dorp and Mazzuchi [37]). The development of a flexible likelihood function results in easier application to different test scenarios such as fixed-stress testing, regular life testing, progressive step-stress testing and profile step-stress testing. The number of test environments which consist of a combination of stress levels from the available set of stress variables is predetermined and fixed. The approach focuses on the failure rate in a constant stress environment which differs from the failure in the step-stress ALT. A general expression is derived for the ALT with different test environments illustrated below for interval data and type I censored data.

The prior information is utilized to obtain prior parameter estimates numerically. After defining the multivariate prior distribution for the parameters, Bayes point estimates for the parameters could be obtained with varying-stress tests. The inference procedures are applicable for testings with interval data and type I censored data. The

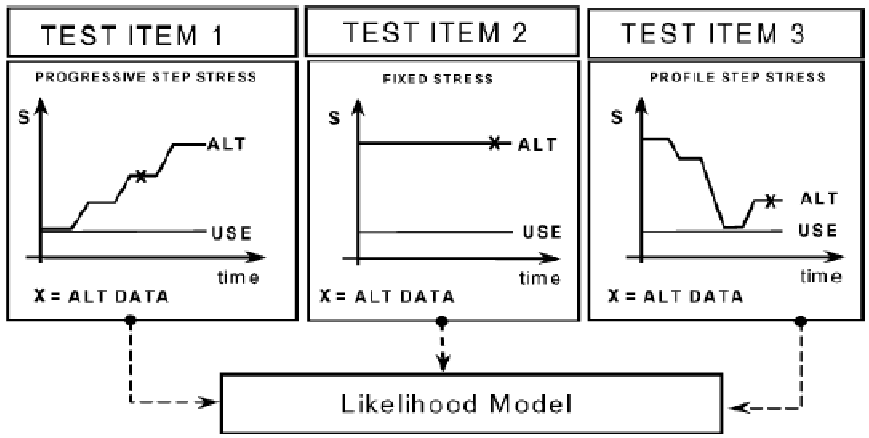

Figure 2.4: A separate ALT design for each test item

posterior distribution might not be tractable. Therefore, Markov chain Monte Carlo methods are adopted to obtain posterior approximations.

A Bayesian approach has also been adopted in various research papers such as Mazzuchi and Soyer [19] and Van Dorp et al. [38]. In those papers, the lifetime distributions under the environments are assumed to be exponentially distributed. The failure rate for the exponential model is constant. The rigid nature makes it difficult to model different failure rate for other products. The flexible nature of Weibull distribution enables it to model increasing, decreasing and constant failure rates with different parameters. Mazzuchi, Soyer and Vopatek [20] can be considered one of the pioneers to produce such model using Bayesian approach for constant stress ALT.

Thus far, parametric estimation procedures were discussed. There are a few research papers which adopt non-parametric approach. Shaked [29] and Shaked and Singpurwalla [30] presented ALT using non-parametric approach when the failure has a single cause and the test units are not censored. Basu and Ebrahimi [4] extended the work of the latter by considering censoring and competing risks. Tyoskin and Krivolapov [34] proposed a non-parametric model for the interval estimation using SSALT. They adopted a linear model of damage accumulation which is proven to satisfy the assumptions suggested in other papers. In addition, they estimated the time transformation function by using failure time data for all stress levels. A numerical example was also presented using the data simulated to investigate performance of the proposed method.

In our work, we develop a Bayesian inference procedure for PH model using SSALT which is similar to Sha and Pan [28]. We present a more general likelihood function which allows censoring during the testing period at a given stress level. In addition, we sample conditionally and jointly the posterior density function and make comparisons based on the posterior estimates and the credible intervals using simulated data and also real data.

Chapter 3

Methodology

3.1 CE model

In this section, the CE-model is presented. The model assumes the products' remaining lifetime depends on the current cumulative fraction failed and current stress level (Nelson [26]). The failure time of the test units follows the cumulative distribution function (CDF) for the current stress level with the previous worn-out conditions being considered. In other words, the previous accumulated fraction failed has to be accounted for. The process is explained with the help of some graphs as follows. Assuming that there are 3 constant stresses (x_1, x_2, x_3) with changing time $\tau_1 = 1, \tau_2 = 2, \tau_3 = 3$. The stress x_i starts at τ_j and runs to τ_{j+1} for $j = 1, 2, 3$ as in Figure 2.2. The CDF of the failure time, F for the process is obtained under a particular step-stress pattern. The cumulative fraction failed of the products under stress x_1 is

$$F(w) = F_1(w), \quad 0 \leq w \leq \tau_1. \tag{3.1.1}$$

Under stress level of x_2, let s_1 be the equivalent starting time, producing the same cumulative fraction failed as previous testing period to satisfy

$$F_2(s_1) = F_1(\tau_1). \tag{3.1.2}$$

The total cumulative failed fraction of the products by time w at stress level of x_2 is given by

$$F(w) = F_2(w - \tau_1 + s_1), \quad \tau_1 \leq w \leq \tau_2. \tag{3.1.3}$$

Similarly, under stress level of x_3, s_2 can be solved by the following equation

$$F_3(s_2) = F_2(\tau_2 - \tau_1 + s_1). \tag{3.1.4}$$

The total cumulative failed fraction is

$$F(w) = F_3(w - \tau_2 + s_2), \quad \tau_2 \leq w \leq \tau_3. \tag{3.1.5}$$

In general, the CDF can be expressed as

$$F(w) = \begin{cases} F_1(w), & \text{if } 0 \leq w < \tau_1 \\ F_2(w - \tau_1 + s_1), & \text{if } \tau_1 \leq w < \tau_2 \\ \ldots & \\ F_k(w - \tau_{k-1} + s_{k-1}), & \text{if } \tau_{k-1} \leq w < \infty. \end{cases} \tag{3.1.6}$$

with $F_{i+1}(s_i) = F_i(\tau_i - \tau_{i-1} + s_{i-1})$ for $i = 1, 2, \ldots, k$.

3.1.1 Weibull CE-Model

In the following section, the failure time w is assumed to follow Weibull distribution. The CDF can be written as

$$F(w) = 1 - \begin{cases} \exp(-\theta_1 w^{\delta}), & \text{if } 0 \leq w < \tau_1 \\ \exp(-\theta_2 (w - \tau_1 + s_1)^{\delta}), & \text{if } \tau_1 \leq w < \tau_2 \\ \ldots & \\ \exp(-\theta_k (w - \tau_{k-1} + s_{k-1})^{\delta}), & \text{if } \tau_{k-1} \leq w < \infty. \end{cases} \tag{3.1.7}$$

where s_i is the equivalent time under the ith stress level, which satisfies $F_i(\tau_i - \tau_{i-1} + s_{i-1}) = F_{i+1}(s_i)$, $i = 1, 2, \ldots, k-1$ with $s_0 = 0$ and $\tau_0 = 0$.

The hazard rate function which corresponds to each level of stress for Weibull CE-Model can be derived as the following

$$
\begin{aligned}
h_i(w) &= \frac{f_i(w-\tau_{i-1}+s_{i-1})}{1-F_i(w-\tau_{i-1}+s_{i-1})} \\
&= \frac{\delta\theta_i(w-\tau_{i-1}+s_{i-1})^{\delta-1}\exp(\theta_i(w-\tau_{i-1}+s_{i-1})^{\delta})}{\exp(\theta_i(w-\tau_{i-1}+s_{i-1})^{\delta})} \\
&= \delta\theta_i(w-\tau_{i-1}+s_{i-1})^{\delta-1}
\end{aligned}
\tag{3.1.8}
$$

The cumulative hazard rate function for each level of stress is $H_i(w) = \int_0^w h_i(t)dt = \theta_i(w-\tau_{i-1}+s_{i-1})^{\delta}$. The cumulative hazard rate function throughout the testing period is as follows

$$
H(w) = \begin{cases}
H_1(w), & \text{if } 0 \le w < \tau_1 \\
H_2(w-\tau_1+s_1), & \text{if } \tau_1 \le w < \tau_2 \\
\dots & \\
H_k(w-\tau_{k-1}+s_{k-1}), & \text{if } \tau_{k-1} \le w < \infty.
\end{cases}
\tag{3.1.9}
$$

As an illustration, Figure 3.1 shows the cumulative hazard rate function at each level of stress for over four levels of stress. The different colors plotted for different cumulative hazard rate functions represent the product's cumulative hazard rate for each testing period, corresponding to each level of stress. The composite hazard rate function which is displayed on the right in Figure 3.1 is formed by shifting the hazard rate function for each level of stress horizontally.

3.2 PH Model

The lifetime of the test units follows a general PH model. The model assumes that the stress level has a multiplicative effect on the hazard rate. If the ratio of two hazard rates from two different stress levels is computed, the ratio is constant over time.

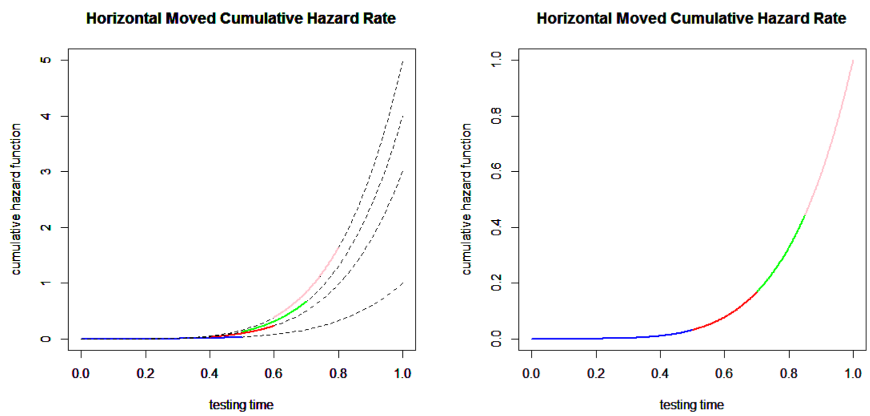

Figure 3.1: Cumulative exposure model

3.2.1 Weibull PH Model

The exponential PH model can be derived based on time transformation of the exponential CE-Model (Khamis and Higgins [12]). Subsequently, the Weibull PH model can be obtained by simple transformation from exponential CE model.

$$F(w) = 1 - \begin{cases} \exp(-\theta_1 w^{\delta}), & \text{if } 0 \le w < \tau_1 \\ \exp(-\theta_2(w^{\delta} - \tau_1^{\delta}) - \theta_1 \tau_1^{\delta}), & \text{if } \tau_1 \le w < \tau_2 \\ \ldots & \\ \exp(-\theta_k(w^{\delta} - \tau_{k-1}^{\delta}) - \ldots - \theta_2(\tau_2^{\delta} - \tau_1^{\delta}) - \theta_1 \tau_1^{\delta}), & \text{if } \tau_{k-1} \le w < \infty. \end{cases} \tag{3.2.1}$$

where the parameter θ_i is related to the stress level of x_i using a log-linear function $\log(\theta_i) = \beta_0 + \beta_1 x_i$ for $i = 1, 2, \ldots, k$ with unknown parameters β_0 and β_1. The hazard rate function of KH model can be found by

$$
\begin{aligned}
h_i(w) &= \frac{f_i(w)}{1-F_i(w)} \\
&= \frac{\delta\theta_i w^{\delta-1}\exp(-\theta_i(w^{\delta}-\tau_{i-1}^{\delta})-\theta_{i-1}\tau_{i-1}^{\delta})}{\exp(-\theta_i(w^{\delta}-\tau_{i-1}^{\delta})-\theta_{i-1}\tau_{i-1}^{\delta})} \\
&= \delta\theta_i w^{\delta-1}
\end{aligned}
\tag{3.2.2}
$$

for $i = 1, 2, ..., k$ where $\theta_0 = 0, \tau_0 = 0$.

The ratio of two hazard rates under two different stress levels in two time periods is constant which has the desirable proportional hazard property. The covariates are multiplicatively related to the hazard rates. The covariate effects can be measured using the hazard rate ratios.

$$
\frac{h_j(w)}{h_i(w)} = \frac{\delta\theta_j w^{\delta-1}}{\delta\theta_i w^{\delta-1}} = \frac{\theta_j}{\theta_i}
\tag{3.2.3}
$$

The cumulative hazard rate function of the Weibull PH model is

$$
H(w) = \begin{cases}
H_1(w), & \text{if } 0 \le w < \tau_1 \\
H_2(w) - (H_2(\tau_1) - H_1(\tau_1)), & \text{if } \tau_1 \le w < \tau_2 \\
\ldots & \\
H_k(w) - (H_k(\tau_{k-1}) - H_{k-1}(\tau_{k-1})) - \ldots - (H_2(\tau_1) - H_1(\tau_1)), & \text{if } \tau_{k-1} \le w < \infty.
\end{cases}
\tag{3.2.4}
$$

where $H_i(w) = \delta\theta_i w^{\delta-1}, w > 0$. Assuming there are four level of stresses, the cumulative hazard rate function for each level of stress on the left in Figure 3.2. The different colors plotted for different cumulative hazard rate functions represent the product's cumulative hazard rate for each testing period, corresponding to each level of stress. The plot on the right in Figure 3.2 depicts the cumulative hazard rate function of SSALT for the whole testing period. This can be formed by taking segments of $H_i(w)$ which corresponds to the level of stress and testing period and shifting the segments vertically.

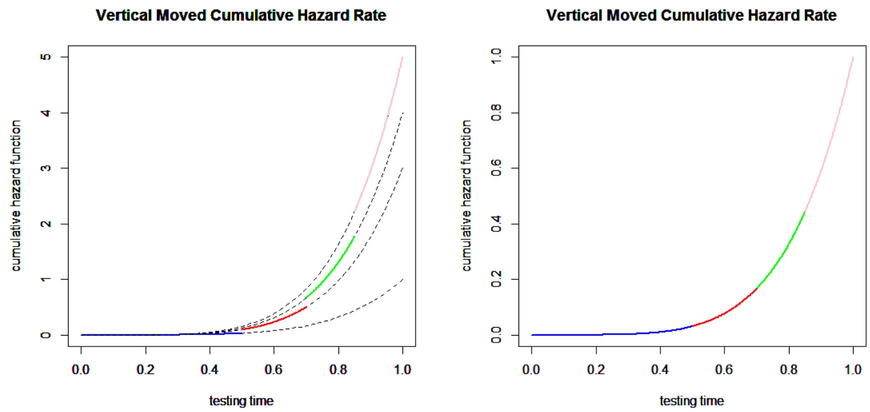

Figure 3.2: Proportional hazard model

3.3 Bayesian Analysis

In this section, probability density function (pdf) is determined using the weibull PH model. By specifying the pdf of the weibull PH model, the likelihood function for the experiment design can be formed. Subsequently, a prior distribution is defined. Lastly, the joint posterior distribution and each conditional posterior distributions are specified using the prior distribution and likelihood function.

3.3.1 Probability Density Function

The probability density function (pdf) of failure time w is needed to obtain the likelihood function and can be derived by differentiating the CDF for Weibull PH Model which is as follows

$$f(w) = \frac{dF(w)}{dw} \tag{3.3.1}$$

The pdf is given by

$$
f(w)=\begin{cases}\delta\theta_1 w^{\delta-1}\exp(-\theta_1 w^{\delta}), & \text{if } 0\le w<\tau_1\\ \delta\theta_2 w^{\delta-1}\exp(-\theta_2(w^{\delta}-\tau_1^{\delta})-\theta_1\tau_1^{\delta}), & \text{if } \tau_1\le w<\tau_2\\ \dots & \\ \delta\theta_k w^{\delta-1}\exp(-\theta_k(w^{\delta}-\tau_{k-1}^{\delta})-\dots-\theta_2(\tau_2^{\delta}-\tau_1^{\delta})-\theta_1\tau_1^{\delta}), & \text{if } \tau_{k-1}\le w<\infty.\end{cases} \tag{3.3.2}
$$

3.3.2 Likelihood Function

When constructing likelihood functions, reliability or survival experiment designs which involve censored or truncated test units have to be given due considerations (Klein and Moeschberger [16]). One of the most common assumption is that lifetimes and censoring times are independent. Otherwise, special methods have to be used. When the likelihood function is constructed, the way each test units contribute to the likelihood has to be identified. If a test unit fails, the test unit corresponding to the exact time contributes to the likelihood in terms of density function $f(x)$ of X at this time. When a test unit is right-censored, the event is not going to occur at this time. The test unit provides information in terms of the reliability or survival function $S(x)$. For left-censored test unit, the event of interest has already happened. The cumulative distribution function $F(x)$ of the test unit will be used in constructing the likelihood. For interval censored test unit, the information provided is the probability of the event of interest occurring in the time interval. Conditional probabilities are used for truncated data depending on the type of truncation. The general likelihood can then be written as

$$
L(\theta|D)\propto\prod_{i\in D}f(x_i)\prod_{i\in R}S(x_i)\prod_{i\in L}F(x_i)\prod_{i\in I}\left\{S(L_i)-S(R_i)\right\} \tag{3.3.3}
$$

where D represents the set of death time, R is the set of right-censored units, L is the set of left-censored units and I is the set of interval censored units.

For this study, we only consider failed units n_i and right-censored units m_i observed with w_{ij} and v_{ij} being the jth failure time and censored time respectively. Therefore, the likelihood function can be written as

$$L(\theta|D) \propto \prod_{i=1}^{k}\left\{\prod_{j=1}^{n_i} f(w_{ij})\prod_{j=1}^{m_i} S(v_{ij})\right\} \tag{3.3.4}$$

where the observational data is $D = \left\{w_{ij}, v_{ij}, m_i, i = 1,...,k, j = 1,...,n_i\right\}$. Using the probability density function of the PH model and the log-linear function of θ_i, $\log(\theta_i) = \beta_0 + \beta_1 x_i$, the likelihood for the whole experiment design can be expressed as

$$L(\beta_0, \beta_1, \delta|D) = \delta^n \exp\left[n\beta_0 + \left(\sum_{i=1}^{k} n_i x_i\right)\beta_1 - \sum_{i=1}^{k} e^{\beta_0+\beta_1 x_i} u_i(\delta)\right]\prod_{i=1}^{k}\prod_{j=1}^{n_i} w_{ij}^{\delta-1} \tag{3.3.5}$$

where $u_i(\delta) = \sum_{j=1}^{n_i}(w_{ij}^{\delta} - \tau_{i-1}^{\delta}) + \sum_{j=1}^{m_i}(v_{ij}^{\delta} - \tau_{i-1}^{\delta}) + (n + m - \sum_{l=1}^{i} n_l - \sum_{l=1}^{i} m_l)(\tau_i^{\delta} - \tau_{i-1}^{\delta})$ for $i = 1, 2, ..., k-1$ and $u_k(\delta) = \sum_{j=1}^{n_k}(w_{kj}^{\delta} - \tau_{k-1}^{\delta}) + \sum_{j=1}^{m_k}(v_{kj}^{\delta} - \tau_{k-1}^{\delta})$.

If $v_{ij} = \tau_i$ for $i = 1, 2, ..., k$, the likelihood reduces to the likelihood obtained from Khamis and Higgins [12].

3.3.3 Prior Distribution Function

A family of prior is introduced which is known as a convex tent family defined on a finite interval, A. The prior distribution will be updated using the likelihood to be the posterior distribution. Subsequently, the posterior distribution will be used for posterior inference. The prior density function are presented below (Sarhan [27]).

Definition The density function of the variable, θ belongs to a convex tent family if the density function has the following form

$$\pi(\theta) = \begin{cases} K(\epsilon - |\theta - \mu|)^r \theta^p \exp(q\theta), & \text{if } \theta \in [\mu - \epsilon, \mu + \epsilon] \\ 0, & \text{otherwise} \end{cases}. \tag{3.3.6}$$

where $r \in \mathbb{Z}^+ \cup \{0\}$, $p, q \in \mathbb{R}$ and K is the normalizing constant for $\pi(\theta)$. To derive the normalizing constant, the proof can be found in the appendix A1. The equation is given by

$$K = \left[\sum_{j=0}^{r} \binom{r}{j} [(\epsilon - \mu)^{r-j} G(p+j, q; \mu - \epsilon, \mu) + (-1)^j (\mu + \epsilon)^{r-j} G(p+j, q; \mu, \mu + \epsilon)] \right]^{-1} \quad (3.3.7)$$

where the function $G(m, n; z_1, z_2)$ is given by

$$G(m, n; z_1, z_2) = \int_{z_1}^{z_2} x^m \exp(nx) dx, \quad m \geq 0, n \in \mathbb{R}. \quad (3.3.8)$$

The function can be expressed in the following form

$$G(m, n; z_1, z_2) = \begin{cases} \frac{1}{m+1} \left[z_2^{m+1} - z_1^{m+1} \right], & \text{if } m \geq 0, n = 0 \\ \frac{1}{s^{m+1}} [\Gamma(m+1, sz_2) - \Gamma(m+1, sz_1)] & \text{if } m \geq 0, n = -s < 0 \\ \frac{m!}{n} [G^*(z_2; m, n) - G^*(z_1; m, n)] & \text{if } m = 0, 1, .., n \neq 0 \end{cases} . \quad (3.3.9)$$

where $\Gamma(n, z) = \int_0^z x^{n-1} \exp(-x) dx$ is the incomplete gamma function and $G^*(x; m, n)$ can be expressed as

$$G^*(x; m, n) = \left[\left(\frac{-1}{n} \right)^m + \sum_{i=0}^{m-1} \frac{(-1)^i x^{m-i}}{n^i (m-i)!} \right] \exp(nx). \quad (3.3.10)$$

The cumulative distribution function (CDF), F can be found by integrating the prior density function which is

$$F(\theta) = \int_{\mu-\epsilon}^{\theta} K(\epsilon - |y - \mu|)^r y^p \exp(qy) dy \quad (3.3.11)$$

The cdf has the following form

$$F(\theta)=\begin{cases}0, & \text{if } \theta<\mu-\epsilon\\ K\left\{\sum_{j=0}^{r}\binom{r}{j}[(\epsilon-\mu)^{r-j}G(p+j,q;\mu-\epsilon,\theta)]\right\}, & \text{if } \mu-\epsilon<\theta<\mu\\ K\left\{\sum_{j=0}^{r}\binom{r}{j}[(\epsilon-\mu)^{r-j}G(p+j,q;\mu-\epsilon,\mu)+(-1)^{j}(\mu+\epsilon)^{r-j}G(p+j,q;\mu,\theta)]\right\}, & \text{if } \mu<\theta<\mu+\epsilon\\ 1, & \text{if } \theta>\mu+\epsilon\end{cases}. \tag{3.3.12}$$

A few graphs are plotted to show the characteristics of the prior density function. When the parameters of the prior density function is set to be $p = 0, q = 0, r = 1$, the distribution is known as triangle distribution displayed in the top-left of Figure 3.3. When the parameters are set to be $p = 2, q = -4, r = 0$, the distribution is similar to the gamma distribution defined on a finite interval. The graph is in the top-right of Figure 3.3. When the scale parameter is set to be a large positive value $q = 10$ with $r = 1$ and $p = 2$, the graph is skewed to the left. The graph is skewed to the right if the scale parameter is set to be $q = -20$ with $r = 1$ and $p = 1$.

3.3.4 Bayesian Inference Procedure

Bayesian approach is generally different from the classical approach. In the classical approach the unknown parameter θ is fixed. The value of θ is estimated based on a sample data $x_1, x_2, ..., x_n$ from the population. In contrast, Bayesian approach assumes that θ follows a probability distribution. Researchers generally have some beliefs on the parameter, θ and the distribution is formulated before the data are observed. Hence, it is named as prior distribution $\pi(\theta)$. After the data is collected from the population described by θ, the prior distribution is updated to be the posterior distribution $\pi(\theta|x)$ using Bayes' Rule (Casella and Berger [6]). By denoting the likelihood function by $L(\theta|x)$, the posterior distribution can be expressed as

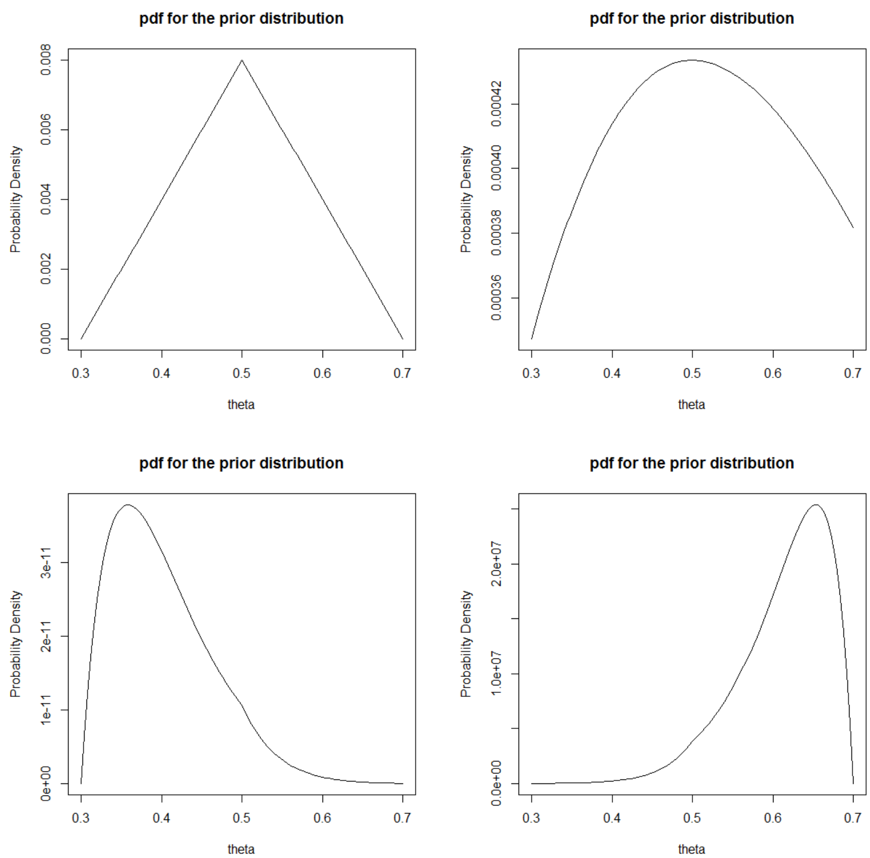

Figure 3.3: Convex tent prior

$$\pi(\theta|D) = \frac{\pi(\theta)L(\theta|D)}{m(x)} \tag{3.3.13}$$

where $m(x)$ is the marginal distribution of X given by

$$m(x) = \int \pi(\theta)L(\theta|D)d\theta \tag{3.3.14}$$

Bayesian inferences can be made on the unknown parameters θ from the posterior distribution.

A Bayesian inference for the Weibull PH model is shown. The transformation of $\alpha_0 = \exp(\beta_0)$ and $\alpha_1 = \exp(\beta_1)$ is to ensure the parameters are positive. The prior distribution used for each of the unknown parameters $\alpha_0, \alpha_1, \delta$ follows the convex tent family with positive and finite interval. Independent priors are set for the parameters β_0, β_1, δ. The joint posterior density function is proportional to the product of the priors and the likelihood function.

$$\pi(\alpha_0, \alpha_1, \delta|D) \propto L(\alpha_0, \alpha_1, \delta|D)\pi(\alpha_0)\pi(\alpha_1)\pi(\delta) \tag{3.3.15}$$

The convex tent prior is the conjugate prior of α_0. The conditional posterior density function of α_0 is given as

$$\begin{aligned}\pi(\alpha_0|\alpha_1, \delta, D) &\propto \pi(\alpha_0)L(\alpha_0|\alpha_1, \delta, D)\\ &\propto (\epsilon - |\alpha_0 - \mu|)^{r_0}\alpha_0^{p_0}\exp(\alpha_0 q_0)\delta^n \alpha_0^{n-1}\alpha_1^{\sum_{i=1}^k n_i x_i - 1}\exp\left[-\alpha_0 \sum_{i=1}^k \alpha_1^{x_i} u_i(\delta)\right]\prod_{i=1}^k \prod_{j=1}^{n_i} w_{ij}^{\delta-1}\\ &\propto (\epsilon - |\alpha_0 - \mu|)^{r_0}\alpha_0^{p_0+n-1} e^{\alpha_0(q_0 - \sum_{i=1}^k \alpha_1^{x_i} u_i(\delta))}.\end{aligned} \tag{3.3.16}$$

The distribution for conditional posterior of α_0 is $\alpha_0|\alpha_1, \delta, D \sim CVT(r_0, p_0+n, q_0-\sum_{i=1}^k \alpha_1^{x_i} u_i(\delta))$ with hyper-parameters r_0, p_0, q_0. The convex tent prior is not prior conjugate of α_1 and δ. The conditional posterior density function of α_1, δ is given as follows

$$
\begin{aligned}
\pi(\alpha_1|\alpha_0,\delta,D) &\propto \pi(\alpha_1)L(\alpha_1|\alpha_0,\delta,D)\\
&\propto (\epsilon-|\alpha_1-\mu|)^{r_1}\alpha_1^{p_1}\exp(\alpha_1 q_1)\delta^n\alpha_0^{n-1}\alpha_1^{\sum_{i=1}^k n_i x_i-1}\exp\left[-\alpha_0\sum_{i=1}^k \alpha_1^{x_i}u_i(\delta)\right]\prod_{i=1}^k\prod_{j=1}^{n_i} w_{ij}^{\delta-1}\\
&\propto (\epsilon-|\alpha_1-\mu|)^{r_1}\alpha_1^{p_1+\sum_{i=1}^k n_i x_i-1}e^{\alpha_1 q_1-\alpha_0\sum_{i=1}^k \alpha_1^{x_i}u_i(\delta)}.
\end{aligned}
\tag{3.3.17}
$$

$$
\begin{aligned}
\pi(\delta|\alpha_0,\alpha_1,D) &\propto \pi(\delta)L(\delta|\alpha_0,\alpha_1,D)\\
&\propto (\epsilon-|\delta-\mu|)^{r_2}\delta^{p_2}\exp(\delta q_2)\delta^n\alpha_0^{n-1}\alpha_1^{\sum_{i=1}^k n_i x_i-1}\exp\left[-\alpha_0\sum_{i=1}^k \alpha_1^{x_i}u_i(\delta)\right]\prod_{i=1}^k\prod_{j=1}^{n_i} w_{ij}^{\delta-1}\\
&\propto (\epsilon-|\delta-\mu|)^{r_2}\delta^{p_2+n}\exp\left\{\delta(q_2+\sum_{i=1}^k\sum_{j=1}^{n_i}\log(w_{ij}))-\alpha_0\sum_{i=1}^k \alpha_1^{x_i}u_i(\delta)\right\}.
\end{aligned}
\tag{3.3.18}
$$

where r_1, p_1, q_1 and r_2, p_2, q_2 are the hyper-parameters.

To make posterior inference, we have to draw posterior samples from their posterior distributions which are not standard. Therefore, we have to adopt a special algorithm known as Markov chain Monte Carlo method discussed in the next chapter.

Chapter 4

Sampling Techniques

Markov chain Monte Carlo (MCMC) methods such as the Metropolis-Hastings and Gibbs sampler have become increasingly popular in statistics field. In general, the MCMC sampling schemes provide an approximate sampling for the researchers when dealing with complicated probability distributions with high dimensions. With the introduction of Gibbs sampling, sampling directly from the posterior distributions of some complicated statistical models is made possible. The MCMC methods have become popular among Bayesian statisticians as most Bayesian inferences could be carried out using MCMC methods while little could be done without it (Casella and Robert [7]). The following sections discuss some concepts about MCMC and some MCMC methods .

4.1 Markov chain

A sequence of dependent random variables $\{X_t\}$ is known as a Markov chain if the probability distribution of X_{t+1} given all past variables depends only on X_t which is given by

$$P(X_{t+1}|X_1,...,X_t) = P(X_{t+1}|X_t)$$

The X_i obtains values from state space of the Markov Chain and the conditional probability distribution is known as transition kernel, K which is

$$X_{t+1}|X_1,...,X_t \sim K(X_t, X_{t+1})$$

4.1.1 Stationarity

A Markov chain $\{X_t\}$ is also a stochastic process. The stochastic process is stationary if the distribution of k-tuple $(X_{t+1}..., X_{t+k})$ does not depend on the value of t for every positive integer k. As a result, the Markov chain is also stationary. In other words, a Markov chain is stationary if and only if the marginal distribution, $\pi(x)$ of X_t does not depend on t (Casella and Robert [7]). The kernel and stationary distribution satisfy the equation

$$\int_X K(x,y)\pi(x)dx = \pi(y)$$

MCMC methods require a Markov chain on a state space to be run easily with the probability distribution π that has the desired distribution as the stationary distribution. The Markov chain has to be reversible in order to construct an MCMC algorithm. The following explains such property (Smith and Roberts [31]).

Definition. A Markov chain on as state space χ is reversible with respect to a probability distribution π on χ, if

$$\pi(dx)P(x,dy) = \pi(dy)P(y,dx), \quad x,y \in \chi$$

Proposition If a Markov chain is reversible with respect to $\pi(\cdot)$, then $\pi(\cdot)$ is stationary for the Markov chain.

Proof.

$$\int_{x\in\chi} \pi(dx)P(x,dy) = \int_{x\in\chi} \pi(dy)P(y,dx) = \pi(dy)\int_{x\in\chi} P(y,dx) = \pi(dy).$$

If a Markov chain has stationary distribution, it might not be able to converge to stationarity. The Markov chain has to contain two more properties which are irreducibility and aperiodicity. The definitions of the two terms are as follows:

Definition. A chain is ϕ-irreducible if there exists a non-zero σ- finite measure ϕ on χ such that for all set A with $A \subseteq \chi$, $\phi(A) > 0$ and for all $x \in \chi$, there exists a positive integer n such that $P^n(x,A) > 0$.

Irreducibility of a Markov chain allows free moves over the state space, which means that the probability of moving from any state to any other state in the chain is finite ($K(x,y) > 0$).

Definition. A Markov chain with stationary distribution $\pi(\cdot)$ is aperiodic if there do not exist $d \geq 2$ and disjoint subsets $\chi_1, \chi_2, ..., \chi_d \subseteq \chi$ with $P(x, \chi_{i+1}) = 1$ for all $x \in \chi_i (1 \leq i \leq d-1)$, and $P(x, \chi_1) = 1$ for all $x \in \chi_d$ such that $\pi(\chi_1) > 0$. Otherwise, the chain is periodic with period d and periodic decomposition $\chi_1, ..., \chi_d$.

The number of steps to return to any arbitrary non-negligible set must not always be a multiple of some values.

The Markov chain in MCMC algorithms are recurrent which is the expected number to return to any arbitrary nonnegligible set is finite. If the chain is recurrent, the limiting distribution is the stationary distribution. The chain is said to be ergodic. In other words, if a kernel K with an ergodic Markov chain and stationary distribution π, generating from K would produce similar results as simulating from π, that is, for integrable function g,

$$\frac{1}{N}\sum_{t=1}^{N} g(X_t) \rightarrow E_\pi(g(X)) \tag{4.1.1}$$

where N is the number of simulation. This is the Law of Large Numbers on which the concepts of MCMC methods are built.

In Bayesian statistics, when the distribution with respect to a Markov chain achieves equilibrium, the stationary distribution is the posterior distribution.

4.2 Metropolis-Hastings

Metropolis-Hastings is one of the most commonly used MCMC methods. When simulating directly from the conditional posterior density is difficult, a working conditional posterior density $q(y|x)$ is proposed which is used to simulate the samples. The ratio of the target density π and the proposed density $q(y|x)$ is constant and independent of x. In addition, $q(y|x)$ has to be allowed to explore the whole support of π. Based on the construction of the Metropolis-Hastings algorithm, the kernel K is created such that π is the limiting distribution. The Metropolis-Hastings sampling procedure is discussed below (Casella and Robert [7]).

4.2.1 Algorithm

Suppose a simulation is run to sample from the target density π which is complicated. A working conditional density q is proposed which produces a Markov chain X_t. Assuming that the current state is x_t, a new sample y_t is drawn from $q(y|x_t)$. The Metropolis-Hastings algorithm is as follows:

1) Propose a distribution, q.
2) Generate $Y_t \sim q(y|x_t)$
3) Take

$$X_{t+1} = \begin{cases} Y_t, & \text{with probability } \ p(x_t, Y_t), \\ x_t & \text{with probability } \ 1 - p(x_t, Y_t), \end{cases}.$$

where

$$p(x,y) = \min\left\{ \frac{\pi(y)}{\pi(x)} \frac{q(x|y)}{q(y|x)}, 1 \right\}.$$

The final step of the current simulation is known as Metropolis rejection. After one proposed value y_t is made, the proposed move to new state is accepted with probability $p(x_t, y_t)$, otherwise the Markov chain remains at the current state. The ratio is undefined if $\pi(x_t) = 0$. Therefore, the initial state $\pi(x) > 0$ has to be arranged. The Markov chain can never move to a new state if $\pi(x) = 0$. In addition, the conditional density $q(y|x) > 0$ with probability one must be satisfied. With $\pi(x) > 0$ and $q(y|x) > 0$, the ratio will be well defined. If $\pi(y) = 0$ and $q(x|y) > 0$ it will not cause any difficulty as the proposed value y will be accepted with probability 0. It is the case when the proposed value y if an impossible value of the desired stationary distribution or the x is not given that y is the current state.

4.3 Gibbs sampling

If the random variables $\theta_1, \theta_2, ..., \theta_n$ have joint density $\pi(\theta_1, \theta_2, ..., \theta_n)$, the Gibbs sampler draws samples directly from the full conditional distributions of each of the random variables and generate a Markov chain $\theta_1^t, \theta_2^t, ..., \theta_n^t$ as the following algorithm. If $\theta_1^t, \theta_2^t, ..., \theta_n^t$ is distributed from π, $\theta_1^{t+1}, \theta_2^{t+1}, ..., \theta_n^{t+1}$ also has the same distribution. This is because both iterations draw samples from true conditionals. For Gibbs sampling, the new sample drawn from a conditional distribution of the desired stationary distribution is accepted with probability one.

4.3.1 Algorithm

Assume that $\theta = (\theta_1, \theta_2, ..., \theta_n)$ at t-th iteration, the Gibbs sampling is displayed as follows:

1) Sample θ_1^{t+1} from the full conditional posterior distribution, $\pi(\theta_1^{t+1}|\theta_2^t, \theta_3^t, ..., \theta_n^t)$.
   Sample θ_2^{t+1} from the full conditional posterior distribution, $\pi(\theta_2^{t+1}|\theta_1^{t+1}, \theta_3^t, ..., \theta_n^t)$.
   :
   Sample θ_n^{t+1} from the full conditional posterior distribution, $\pi(\theta_n^{t+1}|\theta_1^{t+1}, \theta_2^{t+1}, ..., \theta_{n-1}^{t+1})$.
2) Repeat (1) for many times.

4.4 Conditional Sampling

In this section, the Metropolis-Hastings within Gibbs sampling method adopted to draw samples from the full conditional distributions of each of the unknown parameters is presented. Assume that $\alpha_0, \alpha_1, \delta$ at t-th iteration, α_0^{t+1} is sampled from the full conditional distribution of $\pi(\alpha_0|\alpha_1^t, \delta^t)$. Next, α_1^{t+1} is drawn from the full conditional distribution of $\pi(\alpha_1|\alpha_0^{t+1}, \delta^t)$. Due to non-standard distribution of $\alpha_1|\alpha_0^{t+1}, \delta^t$, using Metropolis-Hastings, a proposed distribution with density f_1 which follows the convex tent prior is used to draw samples for α_1^t. The new sample α_1^* is accepted with probability r_{α_1} which is shown in the algorithm. Similarly, a proposed distribution with density f_2 which follows the convex tent prior is used to draw samples for δ^{t+1}. The new sample δ^* is accepted with probability r_δ.

4.4.1 Algorithm

Assume that $\alpha_0, \alpha_1, \delta$ at t-th iteration, the Gibbs sampling is displayed as follows:

1) Sample α_0^{t+1} from $\pi(\alpha_0|\alpha_1^t, \delta^t)$.

2) Sample α_1^{t+1} from $\pi(\alpha_1|\alpha_0^{t+1}, \delta^t)$ by using Metropolis-Hastings

$$\alpha_1^{t+1} = \begin{cases} \alpha_1^*, & \text{with probability } \ r_{\alpha_1}, \\ \alpha_1^t, & \text{with probability } \ 1 - r_{\alpha_1}, \end{cases}.$$

where a proposed $\alpha_1^* \sim CVT(r_1, p_1, q_1)$ with density f_1 and
$r_{\alpha_1} = \min(\frac{\pi(\alpha_1^*|\alpha_0^{t+1}, \delta^t) f_1(\alpha_1^t|\alpha_1^*)}{\pi(\alpha_1^t|\alpha_0^{t+1}, \delta^t) f_1(\alpha_1^*|\alpha_1^t)}, 1)$.

3) Sample δ^{t+1} from $\pi(\delta|\alpha_0^{t+1}, \alpha_1^{t+1})$ by using Metropolis-Hastings

$$\delta^{t+1} = \begin{cases} \delta^*, & \text{with probability } \ r_\delta, \\ \delta^t, & \text{with probability } \ 1 - r_\delta, \end{cases}.$$

4) Repeat the process for many times.

4.5 Joint Sampling

Metropolis-Hastings sampling method is adopted to sample from the joint distribution. Let $\theta = (\alpha_0, \alpha_1, \delta)$. The proposed values of θ are drawn from $\theta^{t+1} \sim N(\theta^t, \widehat{I^{-1}(\theta^t)})$ where mean vector takes on the values from the previous joint samples and the variance-covariance matrix is obtained using the observed inverse of the Fisher information (shown in appendix A2) as it is analytically difficult to evaluate the expected value of some terms from the second derivatives. Therefore, we evaluate the expected value of all terms from the second derivatives by taking the sample mean of the function at each stress level. For instance, let h be any function, the terms can be evaluated using $E(h_i(x)) = \overline{h_i(x)}$. After the proposed joint sample is drawn, the acceptance ratio is computed to determine if the proposed value is accepted. The process is repeated until the final number of simulation is reached.

4.5.1 Algorithm

Assume that $\theta = (\alpha_0, \alpha_1, \delta)$ at t-th iteration, the algorithm is displayed as follows:

1) Sample θ^{t+1} jointly from $\pi(\alpha_0, \alpha_1, \delta|D)$ by using Metropolis-Hastings

$$\theta^{t+1} = \begin{cases} \theta^*, & \text{with probability } r_\theta, \\ \theta^t, & \text{with probability } 1 - r_\theta, \end{cases}.$$

where a proposed $\theta^* \sim N(\theta^t, \widehat{I^{-1}(\theta^t)})$ with density g_1 and $r_\theta = \min(\frac{\pi(\theta^*)g_1(\theta^t|\theta^*)}{\pi(\theta^t)g_1(\theta^*|\theta^t)}, 1)$.

2) Repeat the process for many times.

Chapter 5

Application

In this chapter, we will investigate the proposed model on both simulated data and real data using Bayesian approach. After the data is simulated, the two sampling schemes are used to obtain the posterior estimates and credible intervals for each unknown parameter. Subsequently, we apply the model to two real data sets. One of the data is obtained from Zhao and Elsayed [43] which is a step-stress of light emitting diodes (LED). Second data set is a step-stress test of cable insulation from Nelson [26].

5.1 Simulation study

The data sets are simulated from the Weibull PH model by setting the parameter values β_0 =-2.85,β_1 =-1 and δ =2.97. The data sets are simulated for each different stress level of the total number of stresses $k = 3, 4, 5, 6$. The number of right-censored units m_i are sampled from discrete uniform distribution (1,5) while the number of failures n_i, stress changing times τ_i and the terminated time τ_k are specified in Table 1. The failure time w_{ij} is sampled from the Weibull distribution for each stress level using a log-scale function $\log(\theta_i) = \beta_0 + \beta_1 x_i$ for $i = 1, 2, ..., k$ where x_i is the stress level sampled from normal distribution with mean 2 and standard deviation 1 ($x_i \sim N(2, 1)$).

Two sampling methods are performed and the performance is investigated. We first draw samples from each conditional posterior density function for each unknown parameter $\alpha_0, \alpha_1, \delta$ with hyperparameter values as $p_1 = p_2 = p_3 = 1$,$q_1 = q_2 = q_3 = -2$ and $r_1 = r_2 = r_3 = 2$. Then, we evaluate the posterior estimates by taking the mean of the remaining samples after a burn-in period. Subsequently, 95 % credible intervals for each unknown parameter are computed. The procedures are repeated for different sets of hyperparameter values where $p_1 = p_2 = p_3 = 6, q_1 = q_2 = q_3 = -2$ and $r_1 = r_2 = r_3 = 1$ and $p_1 = p_2 = p_3 = 2, q_1 = q_2 = q_3 = -8$ and $r_1 = r_2 = r_3 = 1$. The results is tabulated in Table 5.2. Next, samples for the unknown parameters are drawn jointly from the joint posterior density function using Metropolis-Hastings algorithm with the same hyperparameter values as mentioned above. The subsequent procedures are carried out in the same way to draw samples from the conditional posterior density function. The results are shown in Table 5.3.

Table 5.1: Prespecified values of simulated data

Stress number	Failures at each stress	Stress changing times	Terminated time
k	n_i	τ_i	τ_k
3	6,3,9	0.5,2.0	3.5
4	9,14,5,12	0.5,2.0,3.5	3.5
5	15,6,3,1,10	0.5,2.0,3.5,5.0	3.5
6	12,6,9,2,1,8	0.5,2.0,3.5,5.0,8.0	3.5

For each simulated data, we draw 20000 samples for each unknown parameter with different sets of hyperparameter values specified above. A burn-in period of 10000 is set based on the trace plots for each unknown parameter. In order to monitor convergence of the simulations, the scale factor estimate $\sqrt{\hat{R}} = \sqrt{\frac{Var(\psi)}{W}}$ is adopted (Gelman et al. [11]), where $Var(\psi) = \frac{n-1}{n}W + \frac{1}{n}B$. W denotes the within-sequence variances while B is the between-sequence variances and n is the iteration number. The scale reduction factors for β_0, β_1, δ fall within 0.97-1.03 for different hyperparameter values which suggests the convergence of the unknown parameters. We observe that the estimation precision improves as the number of stresses increases for both sampling schemes. In addition, the width of the 95 % credible interval decreases with the increase of the number of stresses. When there are three or four number of stress levels, the estimation precision and the width of the 95 % credible intervals are better for conditional sampling than joint sampling. However, when there are five or six number of stress levels, the latter performs better than the former in terms of estimation precision and the width of the 95 % credible intervals. These patterns are also noticeable with different hyperparameter values. It is also observed that the posterior estimates for different hyperparameter values are close to each other for different number of stress levels.

5.2 Real Data Analysis

The first SSALT data from Zhao and Elsayed [43] is shown in Table 5.4. The experiment was conducted under high temperature and humidity which can shorten the lifetime of the LED considerably to predict the product lifetime under normal temperature usage of $50°C$. The experiment was conducted at four different levels of temperature measured in Kelvin while setting the humidity level constant. The temperature was first set at 363 Kelvin and increased at certain time periods. The test was terminated at 7.2 (100 hours). Some of the test units were right censored during the experiment. Using the Arrhenius of reliability testing with temperature, the stress level is reciprocal of the temperature measured in Kelvin which is $x_i = \frac{1}{T_i}$. In this study, the stress level is set as $x_i = \frac{323}{T_i}$ to make $x = 1$ at $T = 323(50°C)$

Using the two sampling methods, the hyperparameters are set as $p_1 = p_2 = p_3 = 1$,$q_1 = q_2 = q_3 = -2$ and $r_1 = r_2 = r_3 = 2$. In this case, 20000 number of simulations is sufficient as the Gelman-Rubin statistic $\sqrt{\hat{R}}$ are near 1 for the three parameters for both sampling methods. The posterior estimates are computed by taking the means of the remaining samples after a burn-in period of 10000. Table 5.4 shows the MLE estimates, large-sample based 95 % confidence intervals from Meeker and Escobar [22] for CE and PH model, posterior estimates and 95 % credible interval from Eberly and Casella [10] for both sampling schemes.

It is observed that the estimates using Bayesian approach are close to the frequentist approach. However, the width of the interval estimations for Bayesian approach is smaller as compared to the frequentist approach. We notice that the 95 % confidence intervals of β_1 include zero which suggests the effect of the stress is not identifiable at 5 % significance level. For Bayesian approach, there is no such difficulty as the credible intervals do not include zero. Moreover, comparatively the width of the credible intervals for both sampling schemes are about the same. The trace plots and histograms for each sampling scheme is shown in Figure 5.1 and 5.2. The samples seem to converge after the burn-in period which is the vertical line at 10000.

The second step-stress test of cable insulation data is in Table 5.6 which shows the pattern of specified values (Kilovolts), final step, failure time and right-censored time of the tested specimens. The purpose of the test is to estimate the lifetime of cable at a design stress of 400 volts/mil. Each specimen was held for 10 minutes each at 5kV, 10kV, 15kV and 20kV before the test began. The specimens were then separated into four different tests according to the thickness of the specimens where the difference between changing time increases as the thickness increases.

As stated previously, the hyperparameters are set as $p_1 = p_2 = p_3 = 1$, $q_1 = q_2 = q_3 = -2$ and $r_1 = r_2 = r_3 = 2$. 20000 iteration is sufficient for convergence as the Gelman-Rubin statistic $\sqrt{\hat{R}}$ is close to 1. The posterior estimates are computed by taking the means of the remaining samples after a burn-in period of 10000. Table 5.7 shows the MLE estimates, 95 % confidence interval for CE and PH model, posterior estimates and 95 % credible interval for both sampling schemes.

The posterior estimates for both sampling methods are close to the MLE, but the credible intervals are narrower than the confidence intervals. The width of the credible intervals for both sampling methods are relatively the same. It is also observed that the confidence and credible intervals do not contain zero, suggesting that the effect of the stress is significant at 5 % significant levels. The trace plots and histograms for each sampling method is shown in Figure 5.3 and 5.4. The samples seem to converge after the burn-in period.

5.3 Conclusion

In our study, a Bayesian approach for Weibull PH model for SSALT data analysis is presented. The graphical and mathematical differences between the model and the Weibull CE model are also discussed. The Weibull PH model has one appealing proportional hazard which explains the relationship between the physical stress and the failure rate. In addition, using it will help avoid the model complexity caused by the time transformation in a cumulative way. Moreover, the model enables us to carry out the posterior inference much easier than the Weibull CE model mathematically without compromising on the flexibility of fitting data. The prior distribution we adopted is flexible in terms of fitting different shapes of probability density function which allows us to fit different scenarios. The joint posterior density function is not standard. With the help of MCMC methods, we could conduct the posterior inferences with much convenience and efficiency. Most of the time, reliability testing only gives small failure time data. This could potentially gives an extremely flat likelihood function and create large uncertainties in the parameter estimation. However, Bayesian inference could resolve this problem by using some prior information from engineering experience or other studies on the model.

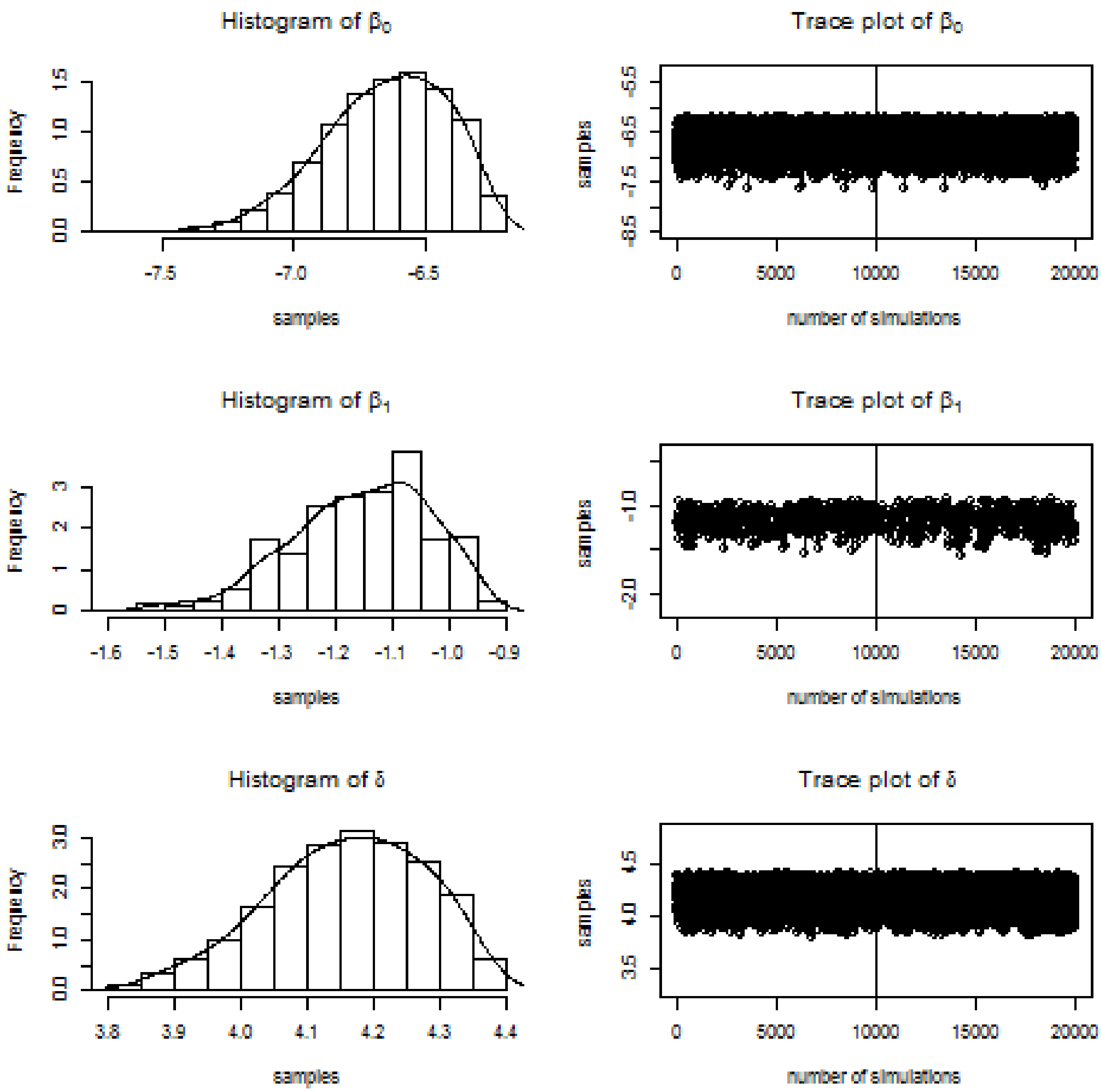

Figure 5.1: Plots for LED data using conditional sampling

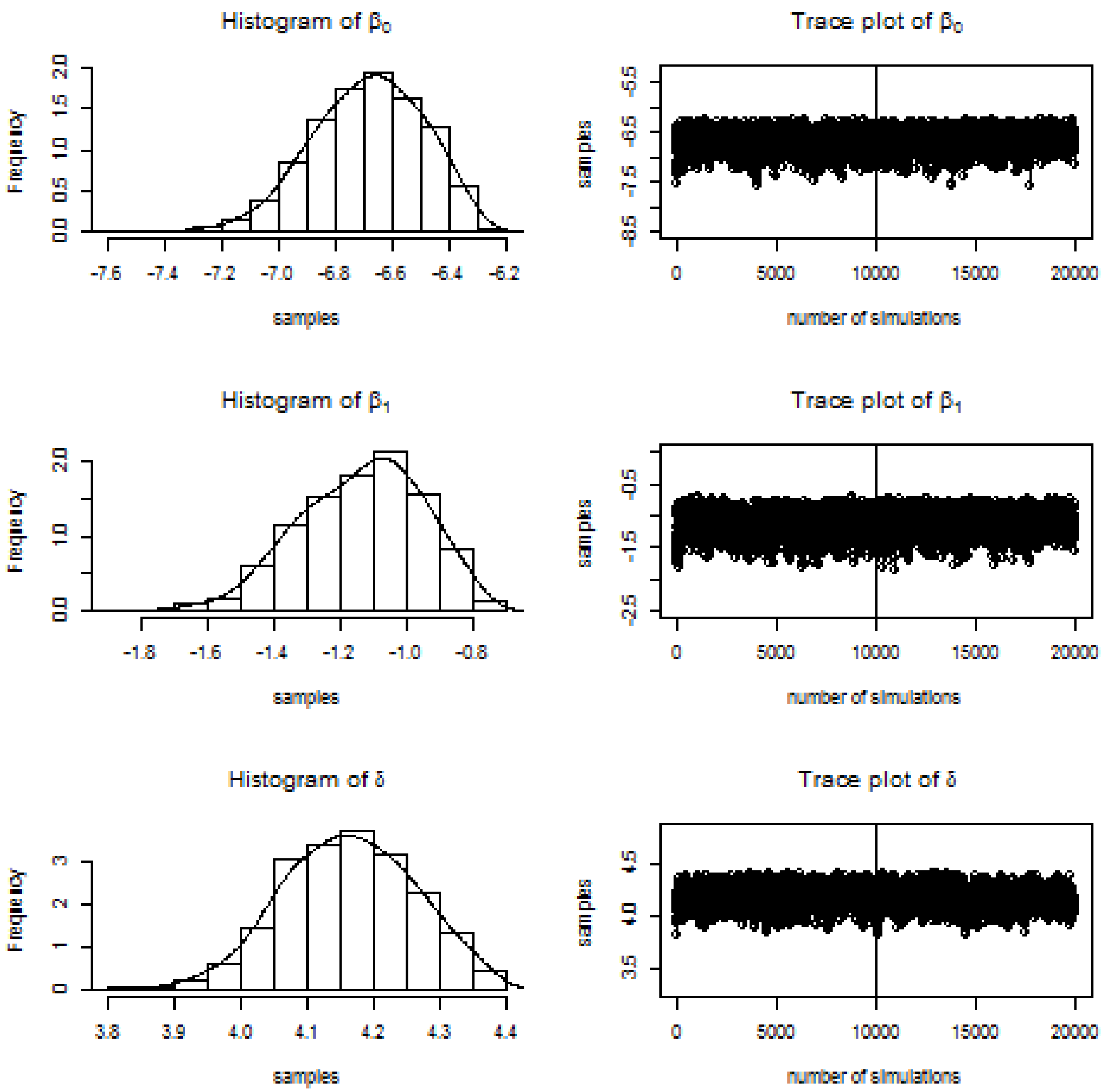

Figure 5.2: Plots for LED data using joint sampling

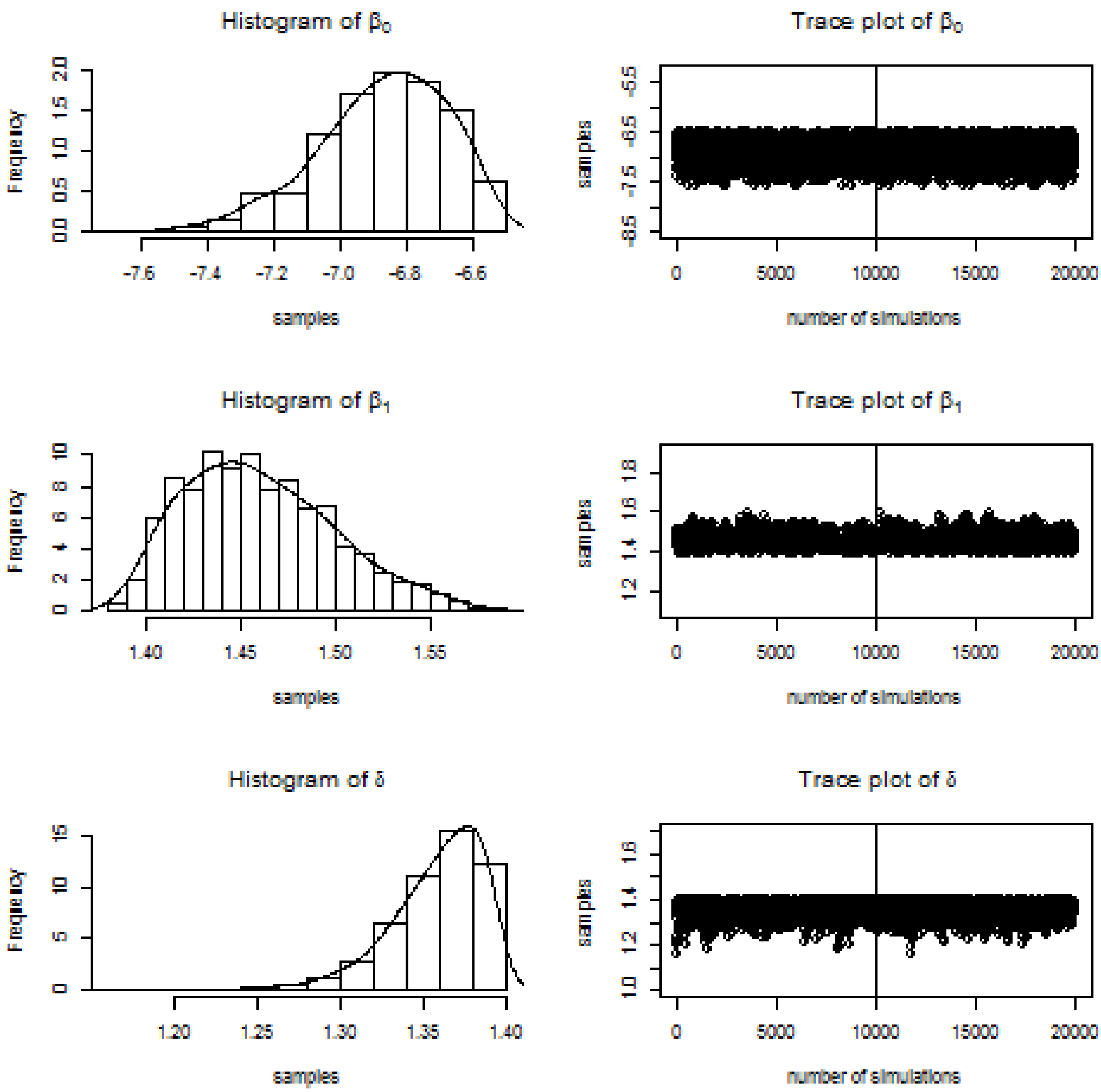

Figure 5.3: Plots for cable insulation data using conditional sampling

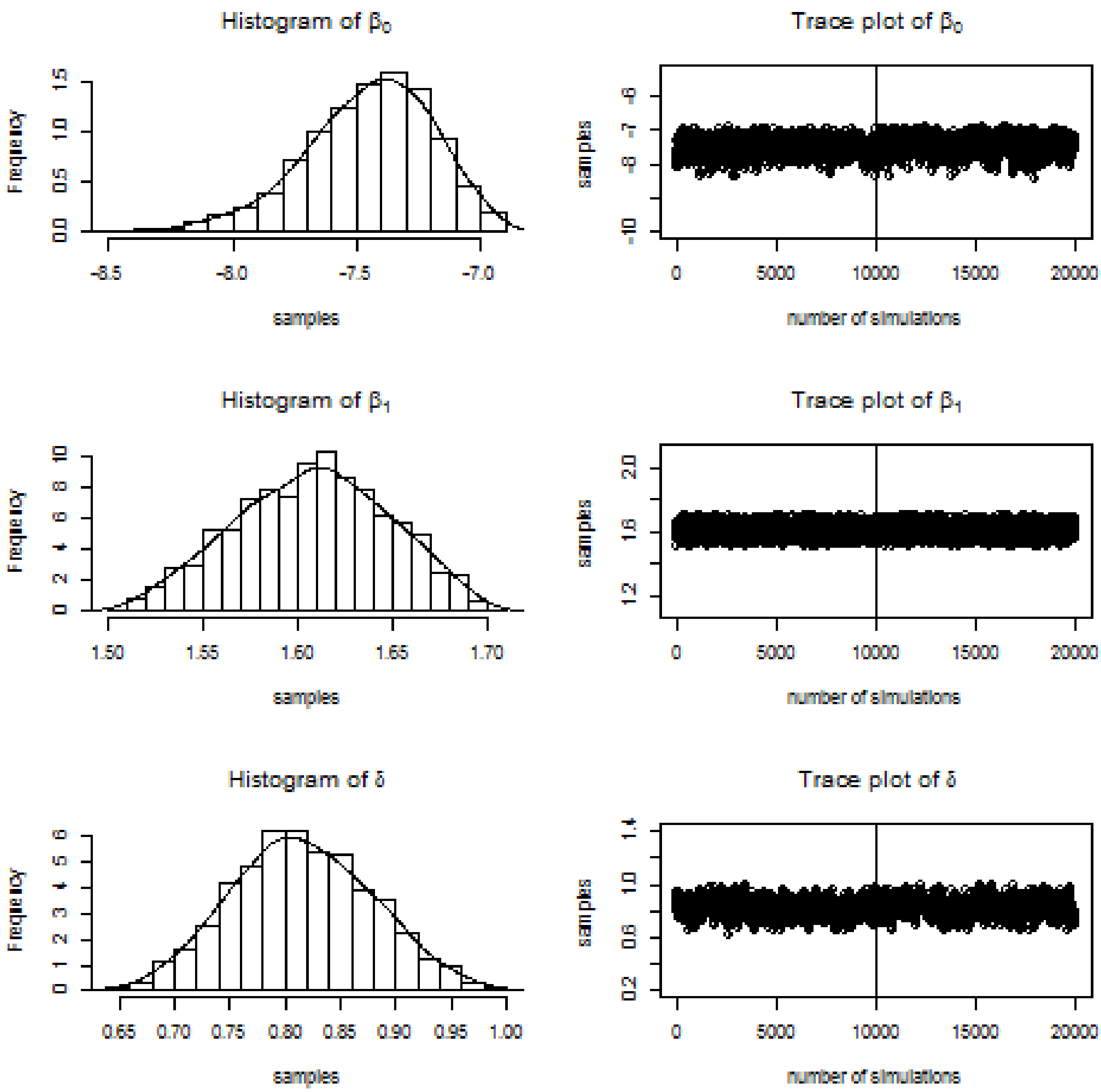

Figure 5.4: Plots for cable insulation data using joint sampling

Table 5.2: Simulated data using conditional sampling

	hyperparameter (p,q,r)		
Stress Level, k	$(1,-2,2)$ Posterior estimates (95 % credible interval)	$(6,-2,1)$ Posterior estimates (95 % credible interval)	$(2,-8,1)$ Posterior estimates (95 % credible interval)
3	$\hat{\beta}_0 = -2.4427$	$\hat{\beta}_0 = -2.5489$	$\hat{\beta}_0 = -2.5330$
	$(-2.6949, -2.3209)$	$(-2.7947, -2.4258)$	$(-2.7578, -2.4234)$
	$\hat{\beta}_1 = -0.6582$	$\hat{\beta}_1 = -0.6539$	$\hat{\beta}_1 = -0.6386$
	$(-0.7890, -0.5888)$	$(-0.7708, -0.5890)$	$(-0.7424, -0.5881)$
	$\hat{\delta} = 3.0160$	$\hat{\delta} = 3.0436$	$\hat{\delta} = 3.0195$
	$(2.7622, 3.2237)$	$(2.8040, 3.2345)$	$(2.7803, 3.2241)$
4	$\hat{\beta}_0 = -2.7366$	$\hat{\beta}_0 = -2.7101$	$\hat{\beta}_0 = -2.7104$
	$(-2.8472, -2.6751)$	$(-2.8038, -2.6650)$	$(-2.8012, -2.6659)$
	$\hat{\beta}_1 = -0.9104$	$\hat{\beta}_1 = 0.9146$	$\hat{\beta}_1 = -0.9112$
	$(-0.9449, -0.8938)$	$(-0.9563, -0.8946)$	$(-0.9465, -0.8938)$
	$\hat{\delta} = 2.9855$	$\hat{\delta} = 2.9916$	$\hat{\delta} = 2.9793$
	$(2.9022, 3.0553)$	$(2.9105, 3.0570)$	$(2.8992, 3.0514)$
5	$\hat{\beta}_0 = -2.8810$	$\hat{\beta}_0 = -2.8316$	$\hat{\beta}_0 = -2.8728$
	$(-2.9220, -2.8241)$	$(-2.894, -2.7741)$	$(-2.9312, -2.8112)$
	$\hat{\beta}_1 = -1.0225$	$\hat{\beta}_1 = -1.0216$	$\hat{\beta}_1 = -1.0238$
	$(-1.0661, -0.9797)$	$(-1.0639, -0.9792)$	$(-1.0662, -0.9800)$
	$\hat{\delta} = 2.9744$	$\hat{\delta} = 2.9745$	$\hat{\delta} = 2.9720$
	$(2.9345, 3.0107)$	$(2.9333, 3.0108)$	$(2.9319, 3.0104)$
6	$\hat{\beta}_0 = -2.8645$	$\hat{\beta}_0 = -2.8643$	$\hat{\beta}_0 = -2.8644$
	$(-2.8761, -2.8531)$	$(-2.8781, -2.8515)$	$(-2.8781, -2.8505)$
	$\hat{\beta}_1 = -1.0129$	$\hat{\beta}_1 = -1.0124$	$\hat{\beta}_1 = -1.0135$
	$(-1.0311, -0.9994)$	$(-1.0231, -0.9988)$	$(-1.0245, -0.9982)$
	$\hat{\delta} = 2.9708$	$\hat{\delta} = 2.9707$	$\hat{\delta} = 2.9704$
	$(2.9552, 2.9848)$	$(2.9549, 2.9865)$	$(2.9545, 2.9855)$

Table 5.3: Simulated data using joint sampling

	hyperparameter (p,q,r)		
Stress Level, k	$(1,-2,2)$ Posterior estimates (95 % credible interval)	$(6,-2,1)$ Posterior estimates (95 % credible interval)	$(2,-8,1)$ Posterior estimates (95 % credible interval)
3	$\hat{\beta}_0 = -2.5343$	$\hat{\beta}_0 = -2.4744$	$\hat{\beta}_0 = -2.5105$
	$(-2.8545, -2.3601)$	$(-2.7859, -2.3228)$	$(-2.7958, -2.3537)$
	$\hat{\beta}_1 = -0.3853$	$\hat{\beta}_1 = -0.3587$	$\hat{\beta}_1 = -0.4840$
	$(-0.5236, -0.3018)$	$(-2.3228, -0.2871)$	$(-0.5938, -0.4268)$
	$\hat{\delta} = 2.8899$	$\hat{\delta} = 3.0376$	$\hat{\delta} = 3.0432$
	$(2.3862, 3.4083)$	$(2.2443, 3.4247)$	$(2.4846, -3.4359)$
4	$\hat{\beta}_0 = -2.9044$	$\hat{\beta}_0 = -2.8838$	$\hat{\beta}_0 = -2.9038$
	$(-3.0145, -2.8383)$	$(-3.0025, -2.8192)$	$(-3.0106, -2.8355)$
	$\hat{\beta}_1 = -0.8371$	$\hat{\beta}_1 = -0.8232$	$\hat{\beta}_1 = -0.8177$
	$(-0.9218, -0.7891)$	$(-0.9070, -0.7837)$	$(-0.8906, -0.7822)$
	$\hat{\delta} = 2.9934$	$\hat{\delta} = 3.0983$	$\hat{\delta} = 3.0552$
	$(2.8551, 3.0820)$	$(2.8823, 3.1965)$	$(2.8864, -3.4359)$
5	$\hat{\beta}_0 = -2.8450$	$\hat{\beta}_0 = -2.8472$	$\hat{\beta}_0 = -2.8470$
	$(-2.8995, -2.7935)$	$(-2.8605, -2.8340)$	$(-2.8581, -2.8363)$
	$\hat{\beta}_1 = -1.0004$	$\hat{\beta}_1 = -1.0024$	$\hat{\beta}_1 = -1.0024$
	$(-1.0189, -0.9838)$	$(-1.0226, -0.9807)$	$(-1.0226, -0.9804)$
	$\hat{\delta} = 2.9854$	$\hat{\delta} = 2.9978$	$\hat{\delta} = 2.9983$
	$(2.9567, 3.0285)$	$(2.9837, 3.0141)$	$(2.9834, 3.0153)$
6	$\hat{\beta}_0 = -2.8475$	$\hat{\beta}_0 = -2.8475$	$\hat{\beta}_0 = -2.8473$
	$(-2.8589, -2.8368)$	$(-2.8607, -2.8338)$	$(-2.8607, -2.8340)$
	$\hat{\beta}_1 = -1.0002$	$\hat{\beta}_1 = -1.0014$	$\hat{\beta}_1 = -1.0016$
	$(-1.0100, -0.9930582)$	$(-1.0116, -0.9908)$	$(-1.0119, -0.9906)$
	$\hat{\delta} = 2.9694$	$\hat{\delta} = 2.9682$	$\hat{\delta} = 2.9687$
	$(2.9635, 2.9758)$	$(2.9603, 2.9794)$	$(2.9605, 2.9791)$

Table 5.4: LED testing conditions and lifetime data

Temperature (Kelvin)	Testing period (100 hours)	Failure or suspension (+) time (100 hours)
363	(0,3)	3.00+
413	(3,5)	3.47,3.97,4.32,4.91,5.00+,5.00+
433	(5,6)	5.12,5.67,5.74,5.88,5.97,6.00+,6.00+
448	(6,7.2)	6.03,6.05,6.15,6.33,6.34,6.37,6.44,6.53,6.75,6.84, 6.99,7.06,7.18,7.20,7.20+,7.20+,7.20+,7.20+

Table 5.5: Estimation results for the LED data

	Frequentist Method MLE (95 % confidence interval)		**Bayesian Method** Posterior mean (95 % credible interval)	
Parameter	CE model	PH model	Conditional Sampling	Joint Sampling
δ	5.9724	4.5250	4.1589	4.1567
	(4.3119,7.6359)	(2.8635,6.1865)	(3.9088,4.3540)	(3.9582,4.3368)
β_0	-6.7812	-7.2160	-6.6260	-6.6484
	(-10.0282,-3.5349)	(-10.4628,-3.9692)	(-7.1342,-6.2783)	(-7.0539,-6.3418)
β_1	-1.4326	-1.5640	-1.1489	-1.1049
	(-7.5259,4.6607)	(-7.6571,4.5291)	(-1.4376,-0.9651)	(-1.4775,-0.8194)

Table 5.6: Test data on cable insulation

Kilovolts	Final step	Failure or suspension (+) time (hours)
26.0	1	6.1+,14.9+
28.5	2	19.3
31.0	3	32.7,41.0,41.0+,45.0
33.4	4	48.7
36.0	5	1.7,1.9,1.9,18.3,18.3,69.0
38.5	6	5.75,5.75+,6.2,20.8,20.9,22.2,22.2+

Table 5.7: Estimation results for the cable insulation data

	Frequentist Method MLE (95 % confidence interval)		**Bayesian Method** Posterior mean (95 % credible interval)	
Parameter	CE model	PH model	Conditional Sampling	Joint Sampling
δ	0.7755	0.6647	1.3579	0.8156
	(0.4308,1.1203)	(0.4735,0.8560)	(1.2878,1.3949)	(0.6980,0.9442)
β_0	-8.4785	-9.0385	-6.8539	-7.4318
	(-9.7795,-7.7175)	(-10.3117,-7.7653)	(-7.2889,-6.5587)	(-8.0595,-7.0075)
β_1	2.5780	2.3002	1.4610	1.6075
	(2.1421,3.0139)	(1.9073,2.6931)	(1.3999,1.5483)	(1.5294,1.6812)

References

[1] V.B. Bagdonavicius, L. Gerville-Reache and M. Nikulin, Parametric inference for step-stress models, *IEEE Transactions on Reliability*, 51:27-31, 2002.

[2] D.S. Bai, M.S. Kim and S.H. Lee, Optimum simple step-stress accelerated life tests for Weibull distribution and type I censoring, *Naval Research Logistics*, 40:193-210, 1993.

[3] D.S. Bai and Y.R. Chun, Optimum simple step-stress accelerated life tests with competing causes of failure, *IEEE Transactions on Reliability*, 40:622-627, 1991.

[4] A.P. Basu and Nader Ebrahimi, Non-parametric Accelerated Life Testing, *IEEE Transactions on Reliability*, 31(5):432-435, 1982.

[5] G.K. Bhattacharyya and Zanzawi Soejoeti, A tampered failure rate model for step-stress accelerated life tests, *Communication in Statistics - Theory and Methods*, 18(5):1627-220, 1989.

[6] George Casella and Roger L. Berger, *Statistical Inference*, Duxbury Press, 2002.

[7] George Casella and C. Robert, *Introducing Monte Carlo Methods with R*, Springer, 2010.

[8] D. Cox, Regression models and life tables, *Journal of the Royal Statistical Society*, 34:187-220, 1972.

[9] Morris H. DeGroot and Prem K. Goel, Bayesian estimation and optimal designs in partially accelerated life testing, *Naval Research Logistics*, 26:223-235, 1979.

[10] Lynn E. Eberly and George Casella, Estimating Bayesian credible intervals, *Journal of Statistical Planning and Inference*, 112:115-132, 2003.

[11] Andrew Gelman, John B. Carlin and Donald B. Rubin, *Bayesian data analysis*, Chapman & Hill, 1996.

[12] I.H. Khamis and J.J. Higgins, A new model for step-stress testing, *IEEE Transaction of Reliability*, 47(2):131-134, 1998.

[13] I.H. Khamis and J.J. Higgins, Optimum 3-Step Step-stress Tests, *IEEE Transaction of Reliability*, 45(2):341-345, 1998.

[14] I.H. Khamis, Optimum M-step step-stress test with k stress variables, *Communication in Statistics - Simulations and Computation*, 26(4):1301-1313, 1997.

[15] I.H. Khamis, Comparison between constant and step-stress tests for Weibull models, *Journal of Quality & Reliability Management*, 14:74-81, 1997.

[16] John P. Klein and Melvin L. Moeschberger, *Survival Analysis Techniques for Censored and Truncated Data*, Springer, 2003.

[17] Jerald F. Lawless, *Statistical Models and Methods for Lifetime Data*, Wiley, 2003.

[18] M.T. Madi, Multiple step-stress accelerated life test:the tampered failure rate model, *Communication in Statistics - Theory and Methods*, 1993.

[19] Thomas A. Mazzuchi and R. Soyer, A dynamic general linear model for inference from accelerated life tests, *Naval Research Logistics*, 39:757-773, 1992.

[20] Thomas A. Mazzuchi, R. Soyer and A. Vopatek, Linear Bayesian inference for accelerated Weibull model, *Lifetime Data Analysis*, 3:63-75, 1997.

[21] E.O. McSorley, Jye-Chyi Lu and Chin-Shang Li, Performance of parameter-estimates in step-stress accelerated life-tests with various sample-sizes, *IEEE Transactions on Reliability*, 2002.

[22] William Q. Meeker and Luis A. Escobar, Teaching About Approximate Confidence Regions Based on Maximum Likelihood Estimation, *American Statistical Association*, 49:48-53, 1995.

[23] William Q. Meeker, A comparison of accelerated life test plans for Weibull and lognormal distributions and type I censored data, *Technometrics*, 26:157-172, 1984.

[24] Wayne Nelson and William Q. Meeker, Charts for optimum accelerated life tests for Weibull and extreme value distributions, *IEEE Transactions on Reliability*, 24:321-332, 1974.

[25] Wayne Nelson and William Q. Meeker, Theory for Optimum Accelerated Censored Life Tests for Weibull and Extreme Value Distributions, *American Statistical Association*, 20(2):171-177, 1978.

[26] Wayne Nelson, Accelerated Life Testing Step-Stress Models and Data Analyses, *IEEE Transaction of Reliability*, 29(3):103-108, 1980.

[27] A. Sarhan, The Bayes Procedure in Exponential Reliability Family Models using Conjugate Tent Prior Family, *Reliability Engineering System Safety*, 71:97-102, 2000.

[28] Naijun Sha and Rong Pan, Bayesian analysis for step-stress accelerated life testing using weibull proportional hazard model, *Statistical Papers*, 55:715-726, 2014.

[29] M. Shaked, An estimator for generalized hazard rate function, *Communication in Statistics - Theory and Methods*, 17-33, 1979.

[30] M. Shaked and N.D. Singpurwalla, Non-parametric estimation and goodness of fit testing hypothesis for distribution in accelerated life testing, *IEEE Transactions on Reliability*, 31:69-74, 1982.

[31] A.F.M. Smith and G.O. Roberts, Bayesian Computation Via the Gibbs Sampler and Related Markov chain Monte Carlo Methods, *Journal of the Royal Statistical Society*, 55:3-23, 1993.

[32] L.C. Tang, Y.S. Sun, T.N. Goh and H.L. Ong, Analysis of step-stress accelerated life test data: A New Approach, *IEEE Transactions on Reliability* , 45:69-74, 1996.

[33] L.C. Tang, *Multiple-step step-stress accelerated life test*, Springer, 2003.

[34] O.I. Tyoskin and S.Y. Krivolapov, Non-parametric model for step-stress accelerated life testing, *IEEE Transactions on Reliability*, 45(2):346-350, 1996.

[35] S.T. Tseng and Z.C. Wen, Step-stress accelerated degradation analysis for highly reliable products, *Journal of Quality Technology*, 2002.

[36] J. René Van Dorp and Thomas A. Mazzuchi, A general Bayes exponential inference model for accelerated life testing, *Journal of Statistical Planning and Inference*, 119:55-74, 2004.

[37] J. René Van Dorp and Thomas A. Mazzuchi, A general Bayes weibull inference model for accelerated life testing, *Reliability Engineering System Safety*, 90:140-147, 2005.

[38] J. René Van Dorp, Thomas A. Mazzuchi, G.E. Fornell and L.R. Pollock, A Bayes approach to step-stress accelerated life testing, *IEEE Transactions on Reliability*, 45(3):491-498, 1996.

[39] Ronghua Wang and Heliang Fei, Conditions for the coincidence of the TFR, TRV and CE Models, *Statistical Paperss*, 45(3):393-412, 2004.

[40] Hai-Yan Xu and Heliang Fei, Models Comparison for Step-Stress Accelerated Life Testing, *Communication in Statistics - Theory and Methods*, 41:3878-3887, 2012.

[41] Guang-Bin Yang, Optimum Constant-Stress Accelerated Life-Test Plans, *IEEE Transactions on Reliability*, 43(4):575-581, 1994.

[42] L.C. Tang, G.Y. Yang and M. Xie, Planning of step-stress accelerated degradation tests, *Proceedings of the Annual Reliability and Maintainability Symposium*, 287-292, 2004.

[43] Wenbiao Zhao and Elsayed A. Elsayed, A general accelerated life model for step-stress testing, *IIE Transactions*, 37(11):1059-1069, 2004.

Appendix A

Appendix

A.1

To solve for the normalizing constant K, we integrate the prior density function from $\mu-\epsilon$ to $\mu+\epsilon$ and equate it to one.

$$
\begin{aligned}
1 &= \int_{\mu-\epsilon}^{\mu+\epsilon} K(\epsilon-|y-\mu|)^r y^p \exp(qy)dy \\
1 &= K\left\{\int_{\mu-\epsilon}^{\mu} (\epsilon+y-\mu)^r y^p \exp(qy)dy + \int_{\mu}^{\mu+\epsilon} (\epsilon-(y-\mu))^r y^p \exp(qy)dy\right\} \\
1 &= K\left\{\int_{\mu-\epsilon}^{\mu} \sum_{j=0}^{r}\binom{r}{j}(\epsilon-\mu)^{r-j} y^j y^p \exp(qy)dy + \int_{\mu}^{\mu+\epsilon} \sum_{j=0}^{r}(-1)^j\binom{r}{j}(\epsilon+\mu)^{r-j} y^j y^p \exp(qy)dy\right\} \\
1 &= K\left\{\sum_{j=0}^{r}\binom{r}{j}(\epsilon-\mu)^{r-j} G(p+j,q;\mu-\epsilon,\mu) + \sum_{j=0}^{r}(-1)^j\binom{r}{j}(\epsilon+\mu)^{r-j} G(p+j,q;\mu,\mu+\epsilon)\right\} \\
K &= \frac{1}{\left\{\sum_{j=0}^{r}\binom{r}{j}(\epsilon-\mu)^{r-j} G(p+j,q;\mu-\epsilon,\mu) + \sum_{j=0}^{r}(-1)^j\binom{r}{j}(\epsilon+\mu)^{r-j} G(p+j,q;\mu,\mu+\epsilon)\right\}}
\end{aligned}
\tag{A.1.1}
$$

A.2

Using (3.3.5), the likelihood of β_0, β_1, Y is

$$L(\beta_0, \beta_1, Y|D) = e^{yn+1} \exp\left[n\beta_0 + \left(\sum_{i=1}^{k} n_i x_i\right)\beta_1 - \sum_{i=1}^{k} e^{\beta_0+\beta_1 x_i} u_i(y)\right] \prod_{i=1}^{k} \prod_{j=1}^{n_i} w_{ij}^{e^y-1}. \tag{A.2.1}$$

where $Y = \log(\delta)$, $u_i(y) = \sum_{j=1}^{n_i}(w_{ij}^{e^y} - \tau_{i-1}^{e^y}) + \sum_{j=1}^{m_i}(v_{ij}^{e^y} - \tau_{i-1}^{e^y}) + (n+m-\sum_{l=1}^{i} n_l - \sum_{l=1}^{i} m_l)(\tau_i^{e^y} - \tau_{i-1}^{e^y})$ for $i = 1, 2, ..., k-1$ and $u_k(\delta) = \sum_{j=1}^{n_k}(w_{kj}^{e^y} - \tau_{k-1}^{e^y}) + \sum_{j=1}^{m_k}(v_{kj}^{e^y} - \tau_{k-1}^{e^y})$.
The log-likelihood becomes

$$\begin{aligned} \log(L(\beta_0, \beta_1, Y|D)) =& ny + n\beta_0 + \left(\sum_{i=1}^{k} n_i x_i\right)\beta_1 - \sum_{i=1}^{k} e^{\beta_0+\beta_1 x_i} u_i(y) + \\ & (e^y - 1)\sum_{i=1}^{k}\sum_{j=1}^{n_i} \log(w_{ij}) + y. \end{aligned} \tag{A.2.2}$$

Let $\theta = (\beta_0, \beta_1, y)$, the first derivative of the log-likelihood is

$$\frac{\partial \log(L(\beta_0, \beta_1, Y|D))}{\partial \theta} = \begin{pmatrix} n - \sum_{i=1}^{k} e^{\beta_0+\beta_1 x_i} u_i(y) \\ \sum_{i=1}^{k} n_i x_i - \sum_{i=1}^{k} e^{\beta_0+\beta_1 x_i} x_i u_i(y) \\ n - \sum_{i=1}^{k} e^{\beta_0+\beta_1 x_i} \frac{\partial u_i(y)}{\partial y} + e^y \sum_{i=1}^{k}\sum_{j=1}^{n_i} \log(w_{ij}) + 1 \end{pmatrix}$$

where $\frac{\partial u_i(y)}{\partial y} = \sum_{j=1}^{n_i}(w_{ij}^{e^y} e^y \log(w_{ij}) - \tau_{i-1}^{e^y} e^y \log(\tau_{i-1})) + \sum_{j=1}^{m_i}(v_{ij}^{e^y} e^y \log(v_{ij}) - \tau_{i-1}^{e^y} e^y \log(\tau_{i-1})) + (n + m - \sum_{l=1}^{i} n_l - \sum_{l=1}^{i} m_l)(\tau_i^{e^y} e^y \log(\tau_i) - \tau_{i-1}^{e^y} e^y \log(\tau_{i-1}))$ for $i = 1, 2, ..., k-1$ and $\frac{\partial u_k(y)}{\partial y} = \sum_{j=1}^{n_k}(w_{kj}^{e^y} e^y \log(w_{kj}) - \tau_{k-1}^{e^y} e^y \log(\tau_{k-1})) + \sum_{j=1}^{m_k}(v_{kj}^{e^y} e^y \log(v_{kj}) - \tau_{k-1}^{e^y} e^y \log(\tau_{k-1}))$
The Hessian matrix H is defined as follows

$$H = -\begin{pmatrix} \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial \beta_0^2} & \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial \beta_0 \partial \beta_1} & \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial \beta_0 \partial y} \\ \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial \beta_1^2} & \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial \beta_1 \partial y} \\ \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial y \partial \beta_0} & \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial y \partial \beta_1} & \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial y^2} \end{pmatrix}$$

and the Fisher information matrix is given by

$$I(\theta) = -E\begin{pmatrix} \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial\beta_0^2} & \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial\beta_0\partial\beta_1} & \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial\beta_0\partial y} \\ \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial\beta_1\partial\beta_0} & \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial\beta_1^2} & \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial\beta_1\partial y} \\ \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial y\partial\beta_0} & \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial y\partial\beta_1} & \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial y^2} \end{pmatrix}$$

where

$$\frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial\beta_0^2} = -\sum_{i=1}^{k} e^{\beta_0+\beta_1 x_i} u_i(y)$$
$$\frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial\beta_1^2} = -\sum_{i=1}^{k} e^{\beta_0+\beta_1 x_i} x_i^2 u_i(y)$$
$$\frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial y^2} = -\sum_{i=1}^{k} e^{\beta_0+\beta_1 x_i} \frac{\partial^2 u_i(y)}{\partial y^2} + e^y \sum_{i=1}^{k} \sum_{j=1}^{n_i} \log(w_{ij})$$
$$\frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial\beta_0\partial\beta_1} = \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial\beta_1\partial\beta_0} = -\sum_{i=1}^{k} e^{\beta_0+\beta_1 x_i} x_i u_i(y)$$
$$\frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial\beta_0\partial y} = \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial y\partial\beta_0} = -\sum_{i=1}^{k} e^{\beta_0+\beta_1 x_i} \frac{\partial u_i(y)}{\partial y}$$
$$\frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial\beta_1\partial y} = \frac{\partial^2 \log(L(\beta_0,\beta_1,Y|D))}{\partial y\partial\beta_1} = -\sum_{i=1}^{k} e^{\beta_0+\beta_1 x_i} x_i \frac{\partial u_i(y)}{\partial y}$$

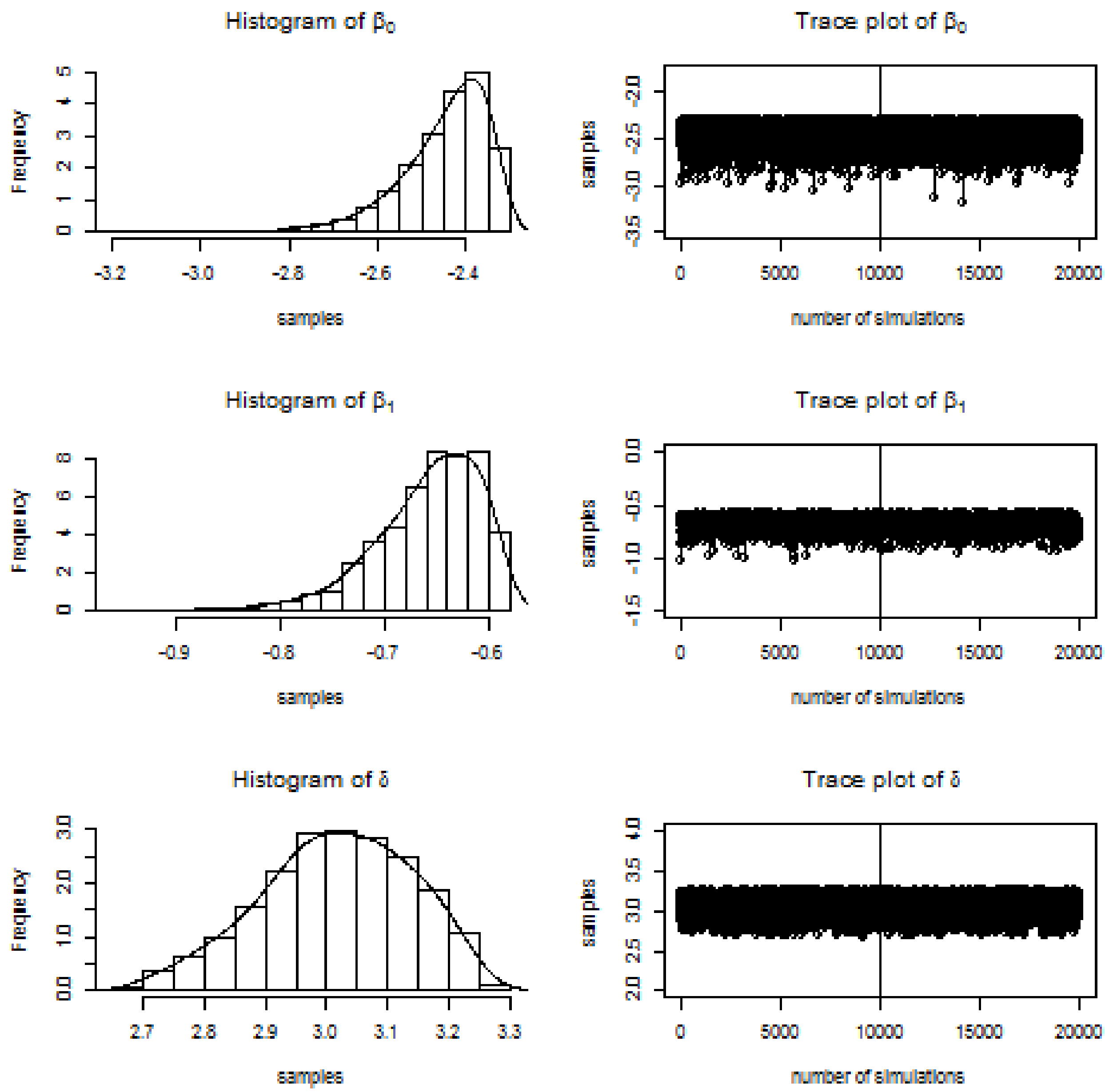

Figure A.1: Plots for conditional sampling using 3 levels of stress with hyperparameters (1,-2,2)

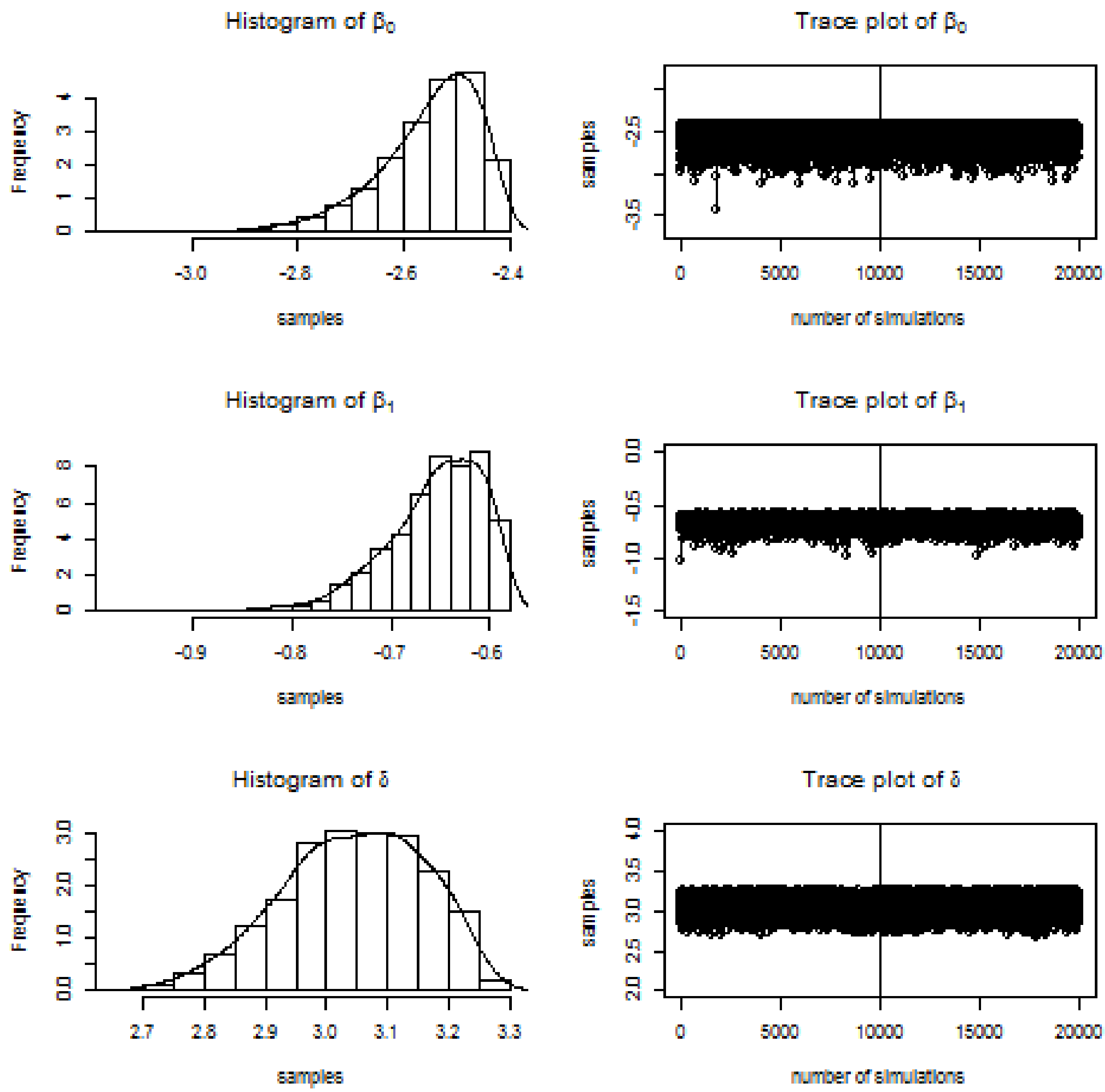

Figure A.2: Plots for conditional sampling using 3 levels of stress with hyperparameters (6,-2,1)

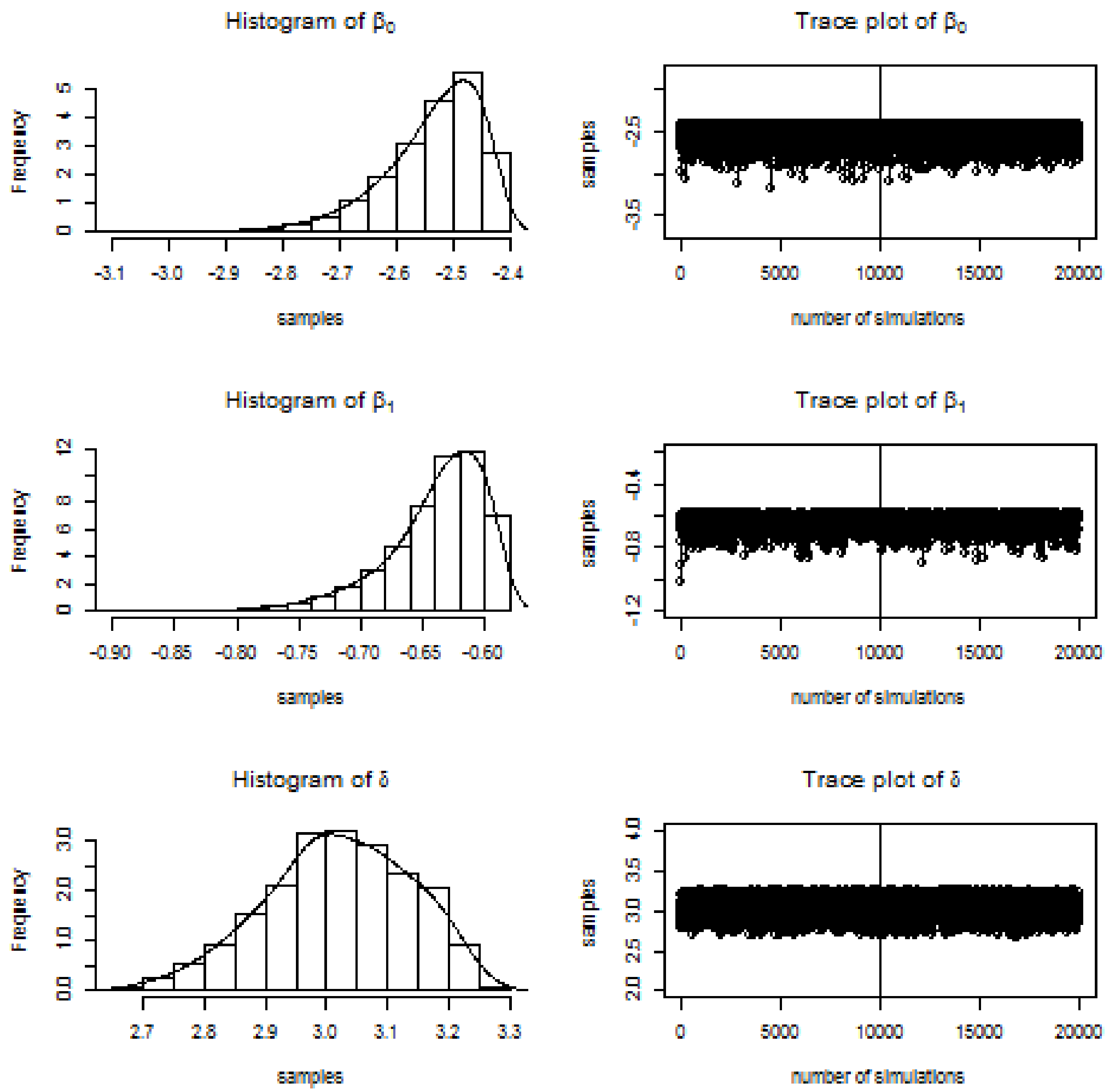

Figure A.3: Plots for conditional sampling using 3 levels of stress with hyperparameters (2,-8,1)

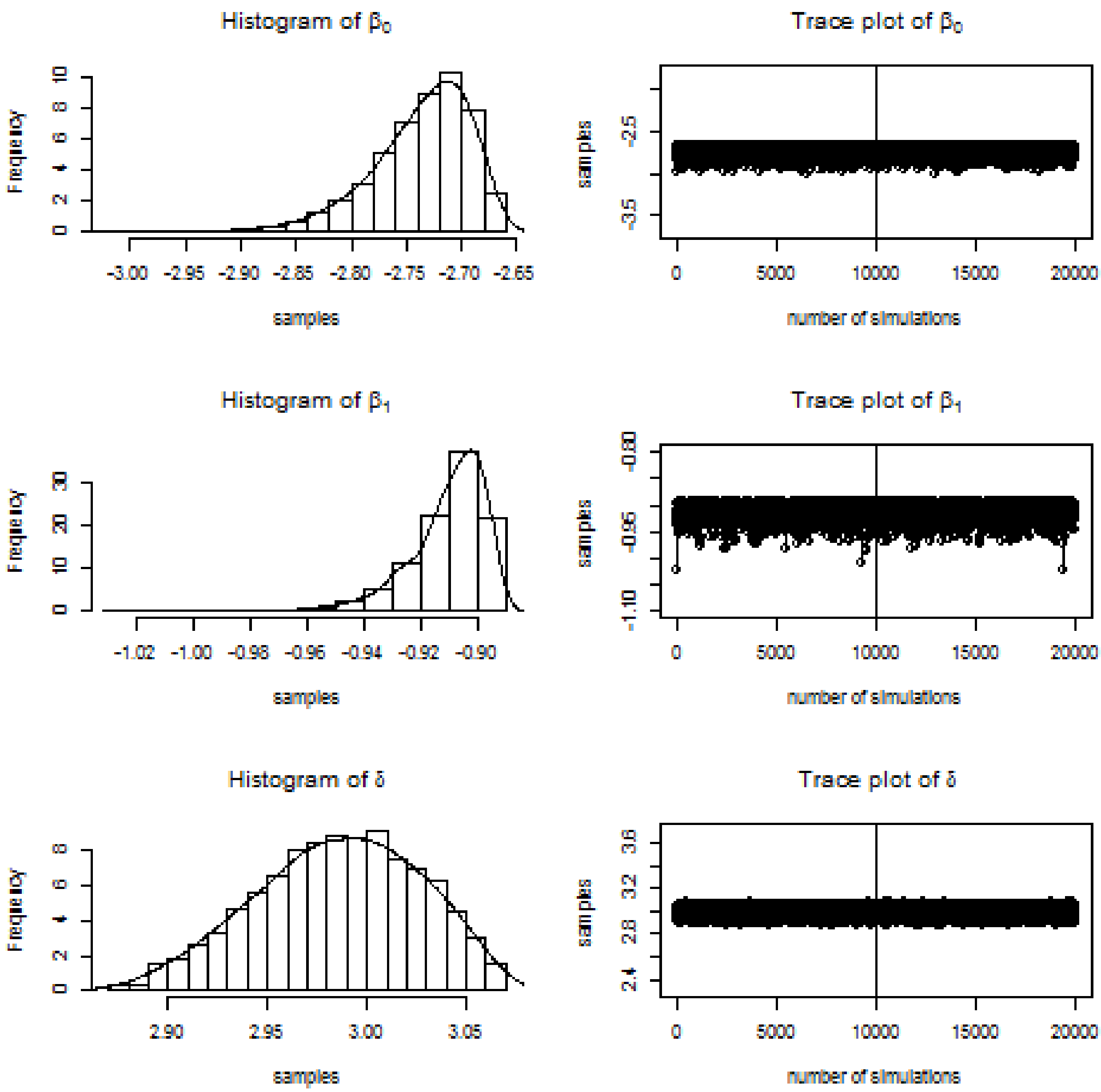

Figure A.4: Plots for conditional sampling using 4 levels of stress with hyperparameters (1,-2,2)

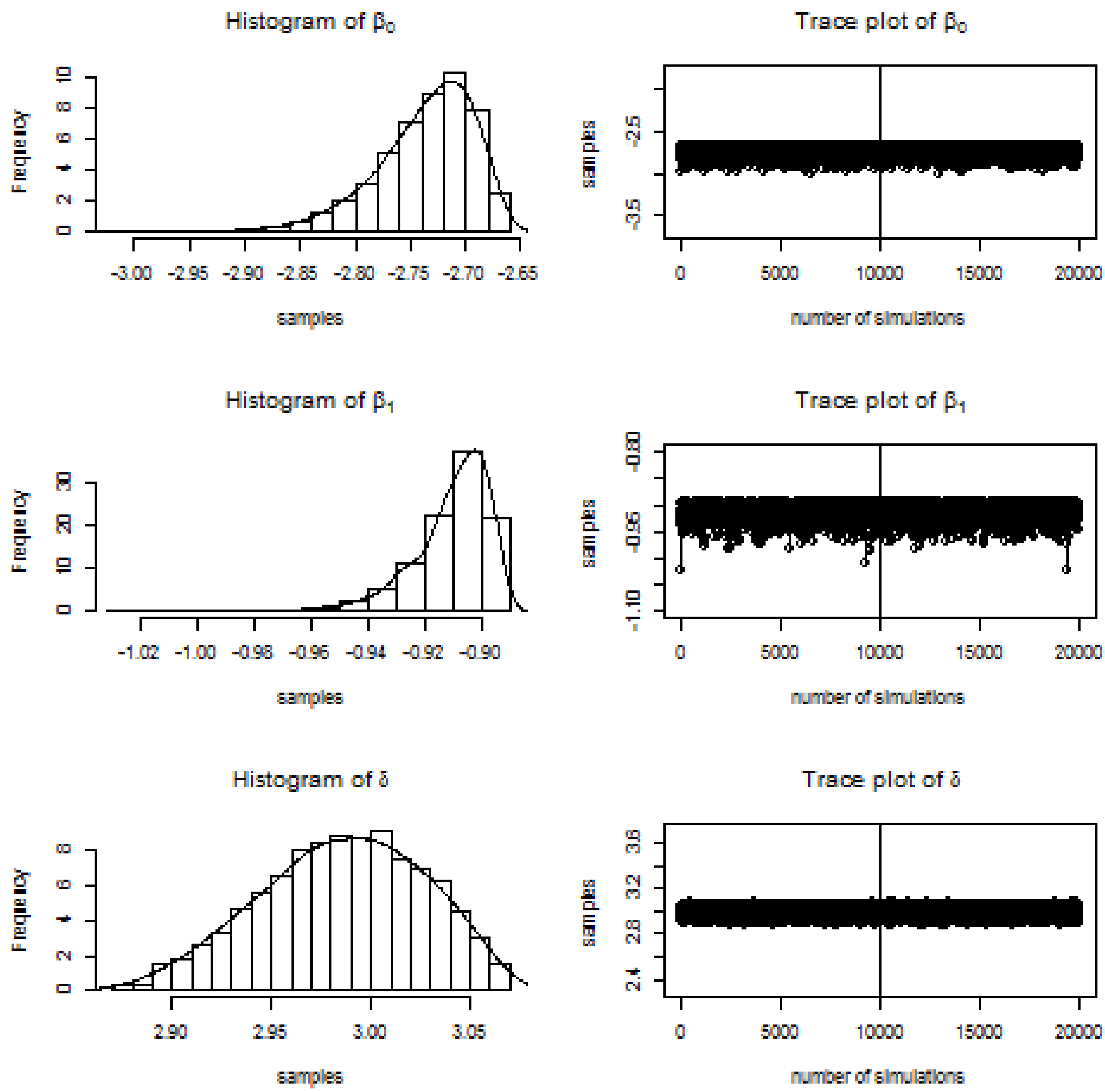

Figure A.5: Plots for conditional sampling using 4 levels of stress with hyperparameters (6,-2,1)

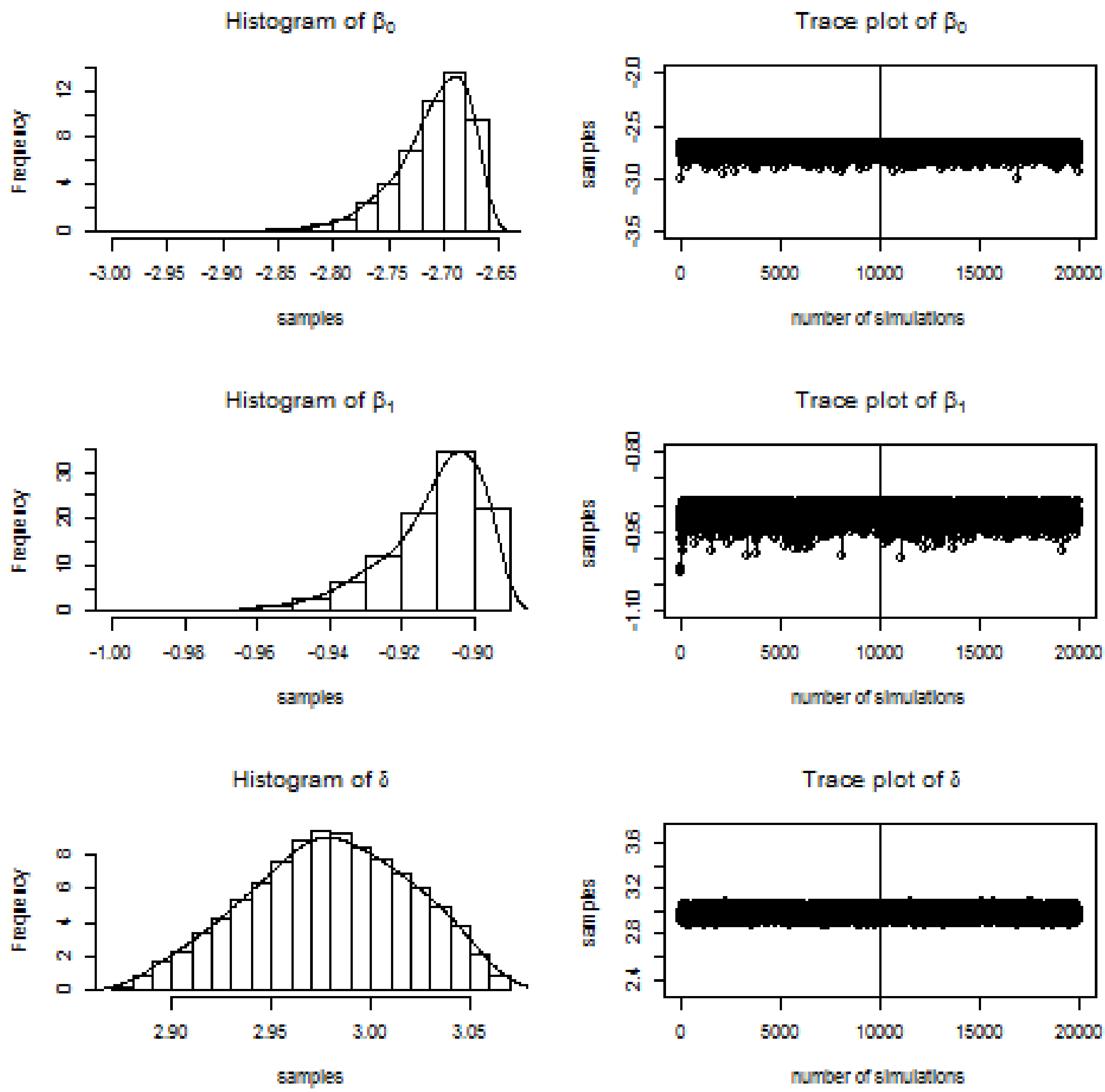

Figure A.6: Plots for conditional sampling using 4 levels of stress with hyperparameters (2,-8,1)

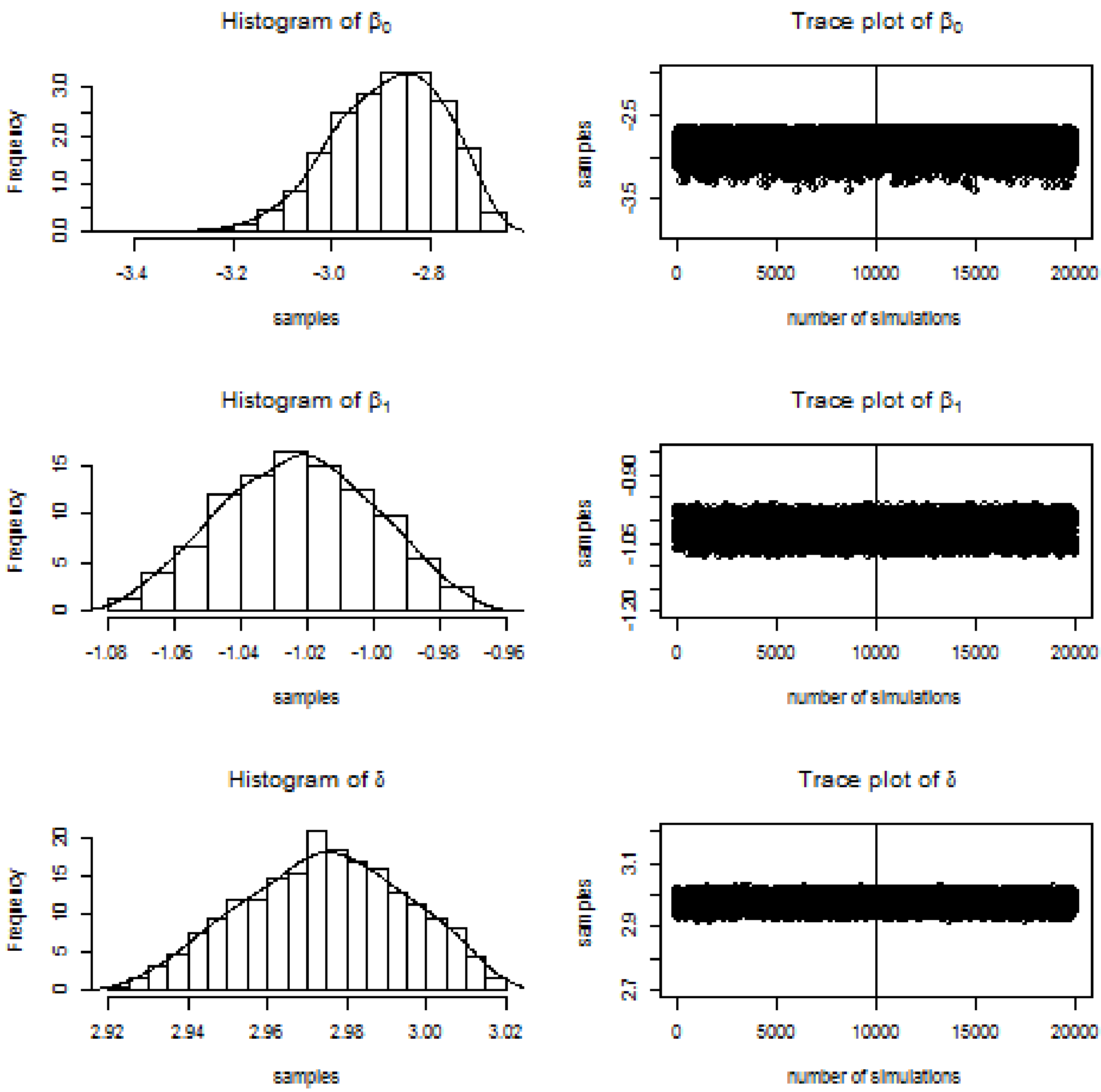

Figure A.7: Plots for conditional sampling using 5 levels of stress with hyperparameters (1,-2,2)

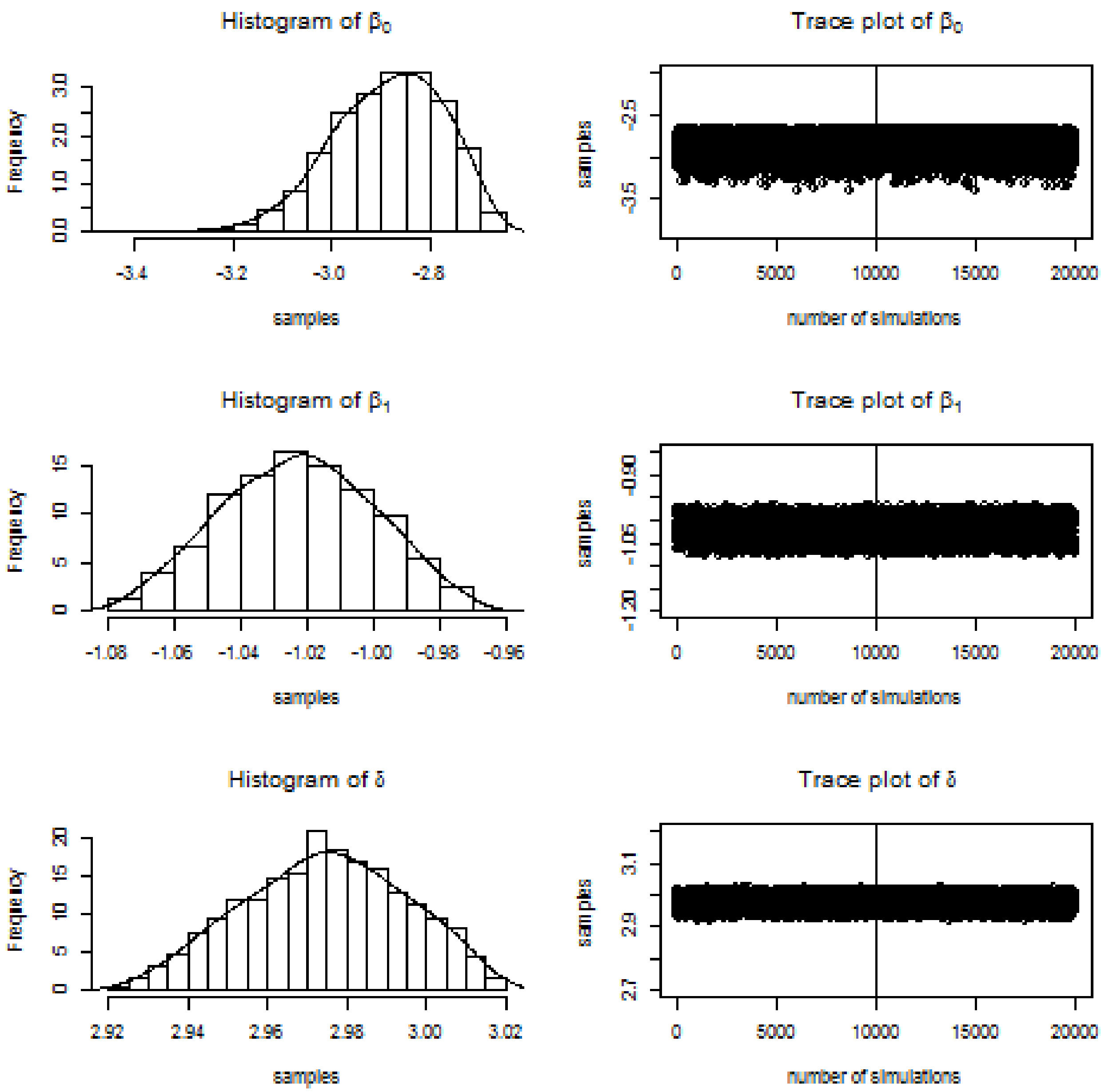

Figure A.8: Plots for conditional sampling using 5 levels of stress with hyperparameters (6,-2,1)

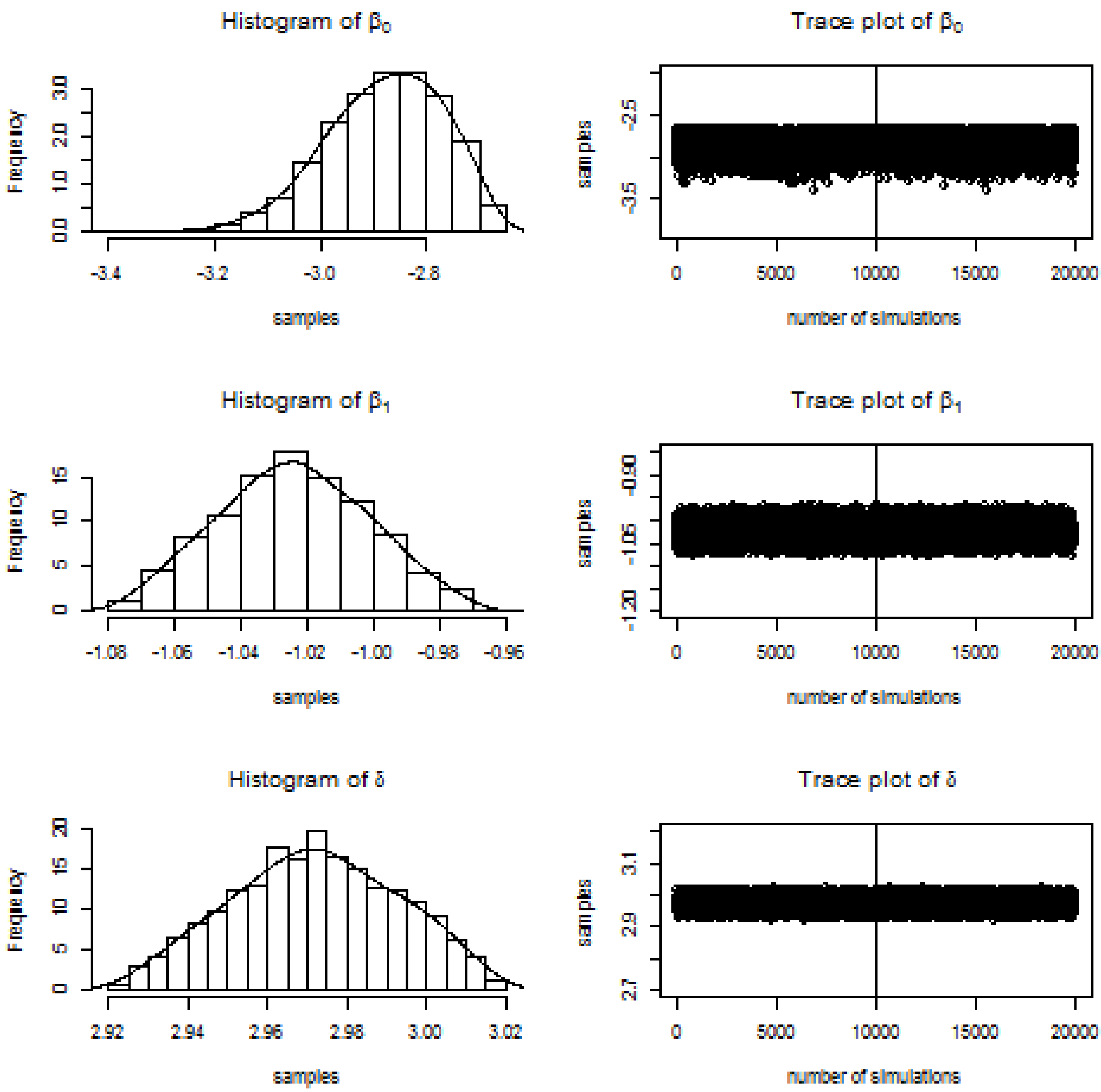

Figure A.9: Plots for conditional sampling using 5 levels of stress with hyperparameters (2,-8,1)

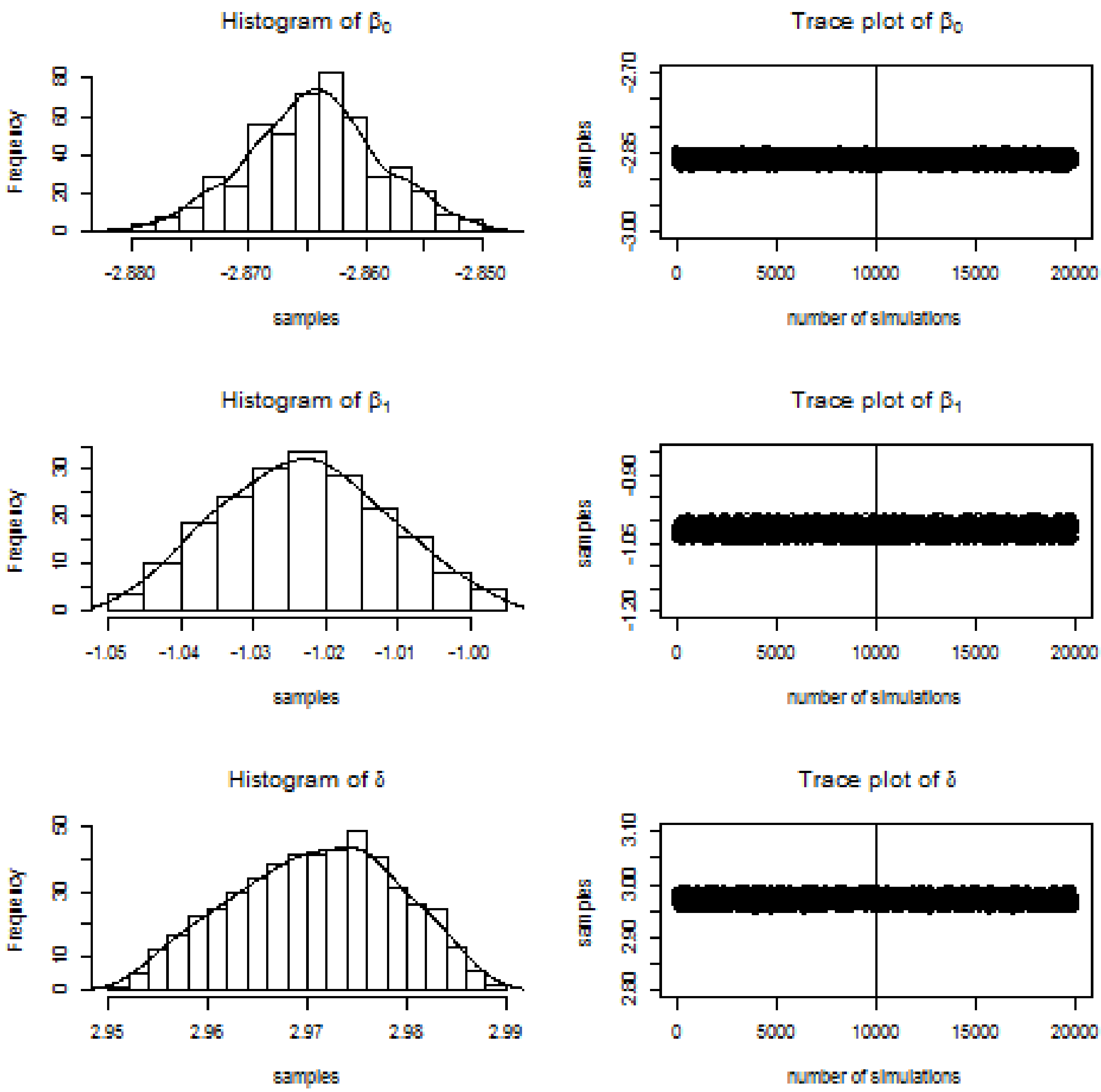

Figure A.10: Plots for conditional sampling using 6 levels of stress with hyperparameters (1,-2,2)

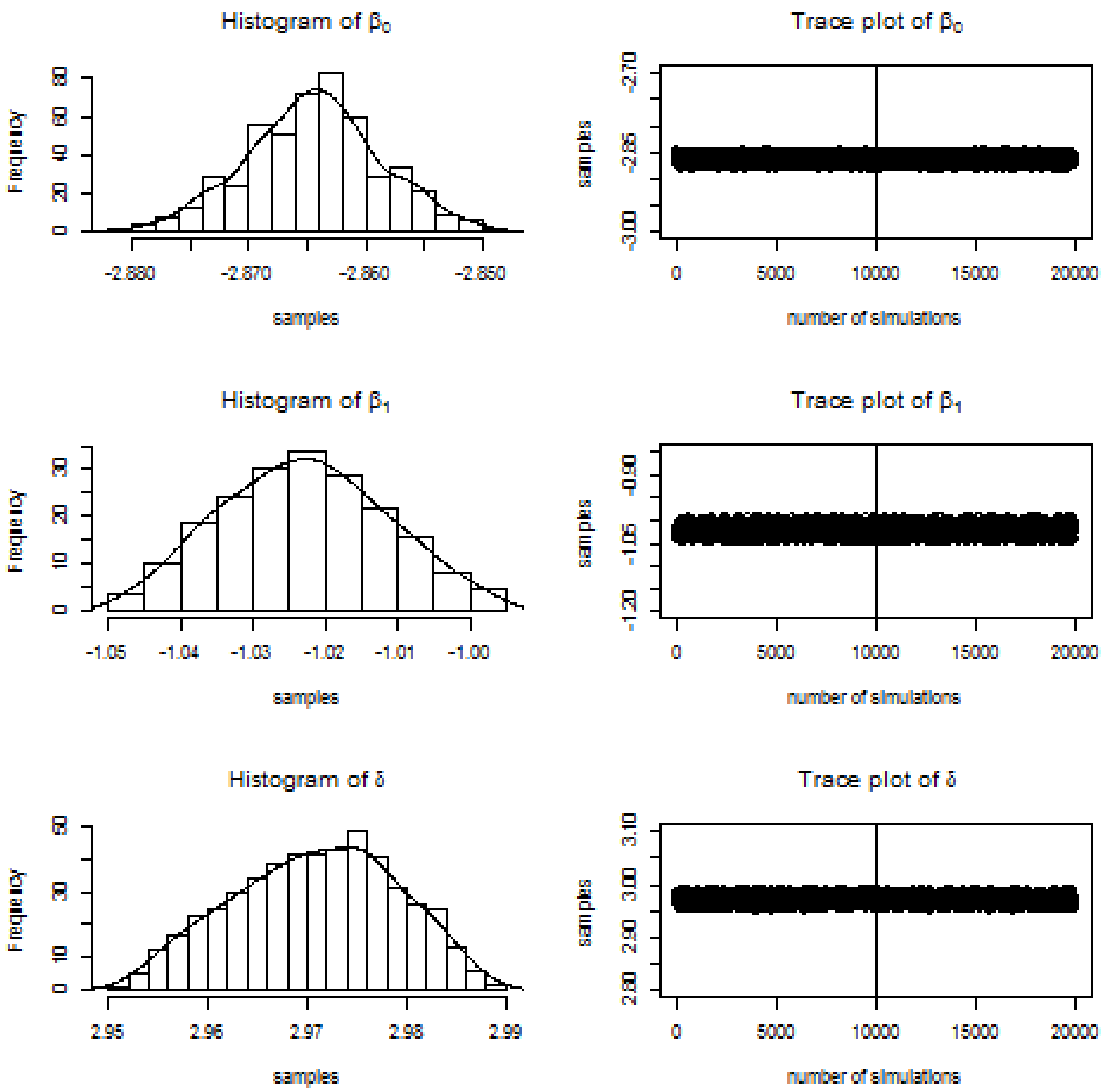

Figure A.11: Plots for conditional sampling using 6 levels of stress with hyperparameters (6,-2,1)

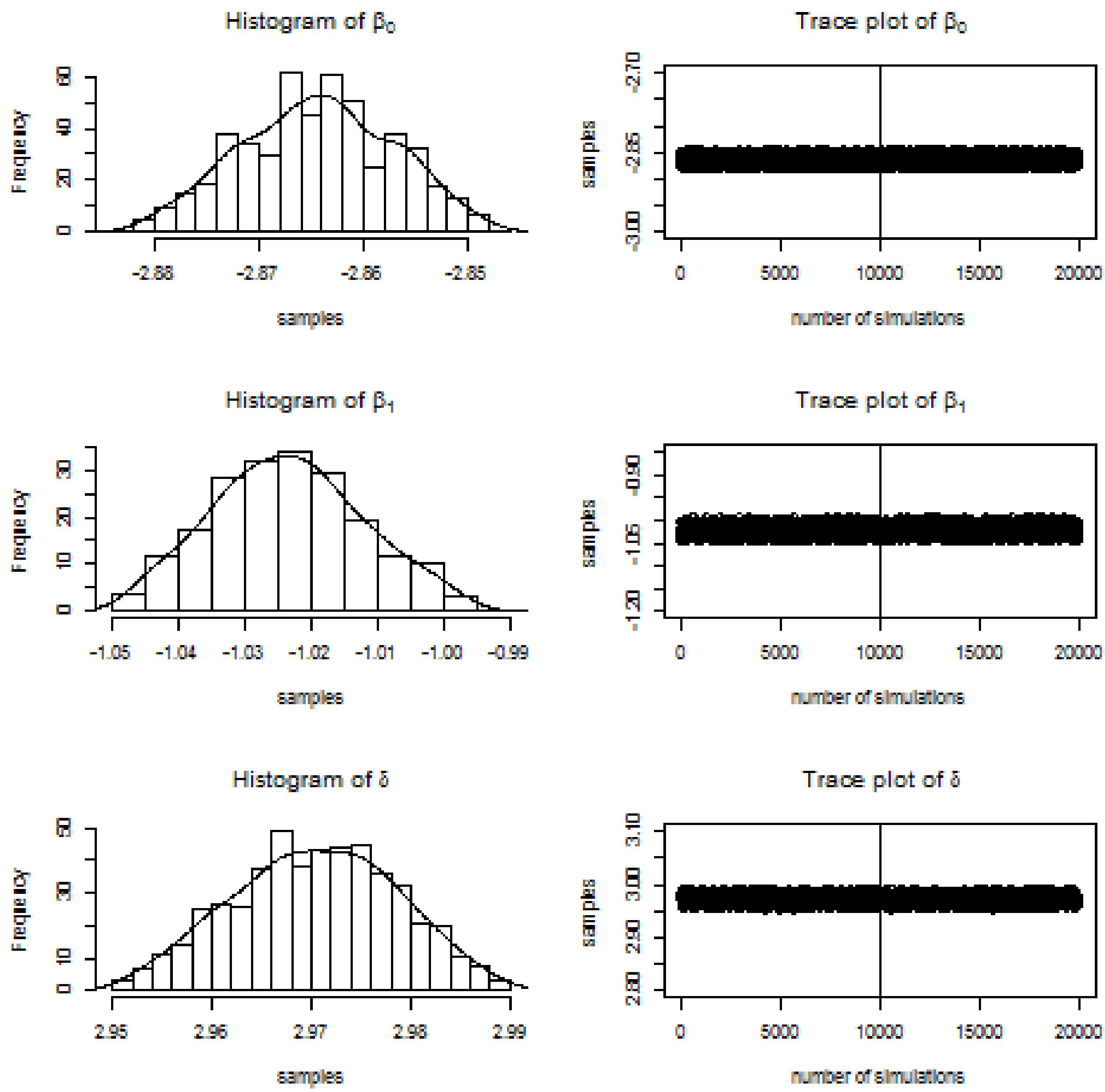

Figure A.12: Plots for conditional sampling using 6 levels of stress with hyperparameters (2,-8,1)

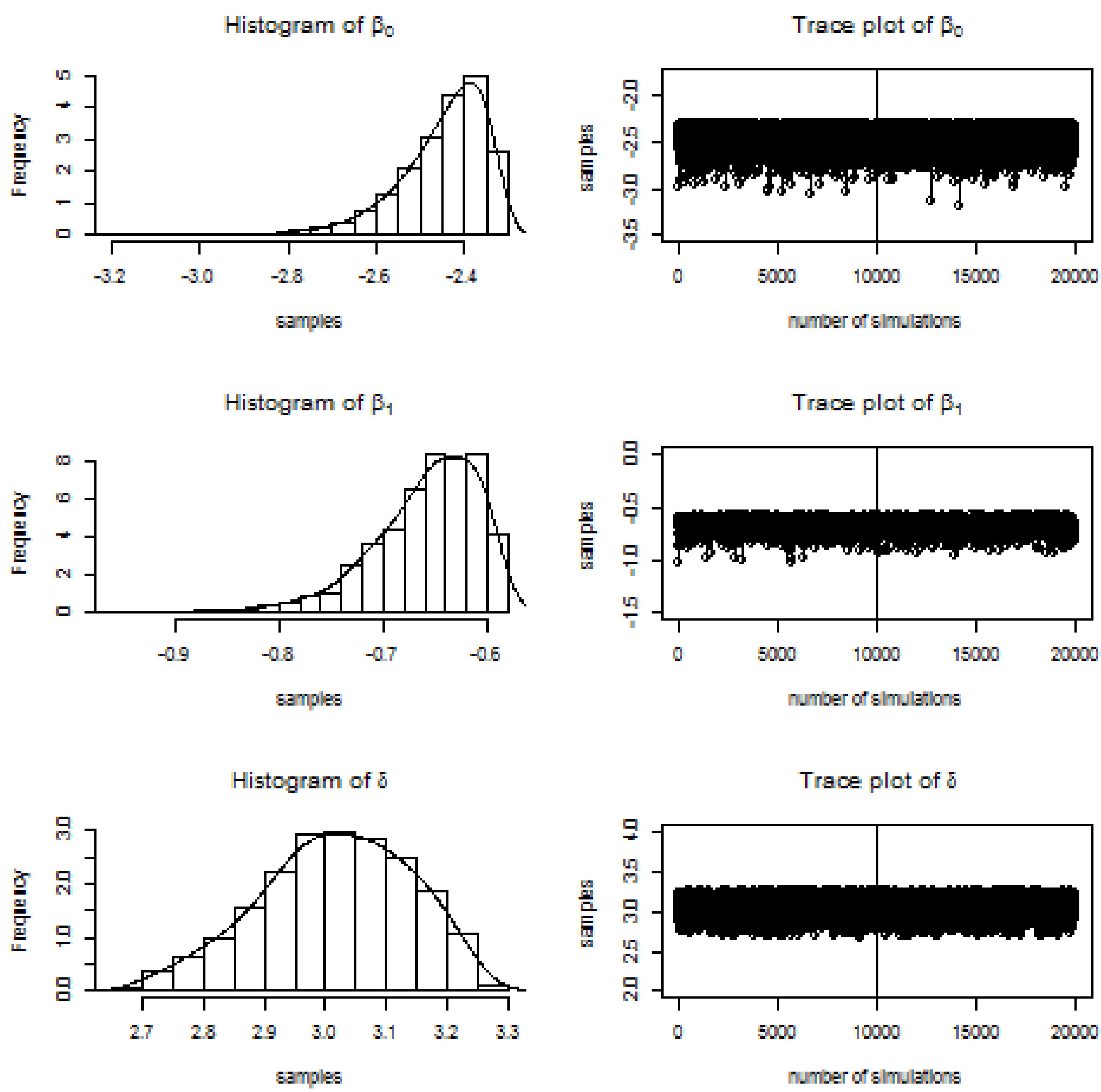

Figure A.13: Plots for joint sampling using 3 levels of stress with hyperparameters (1,-2,2)

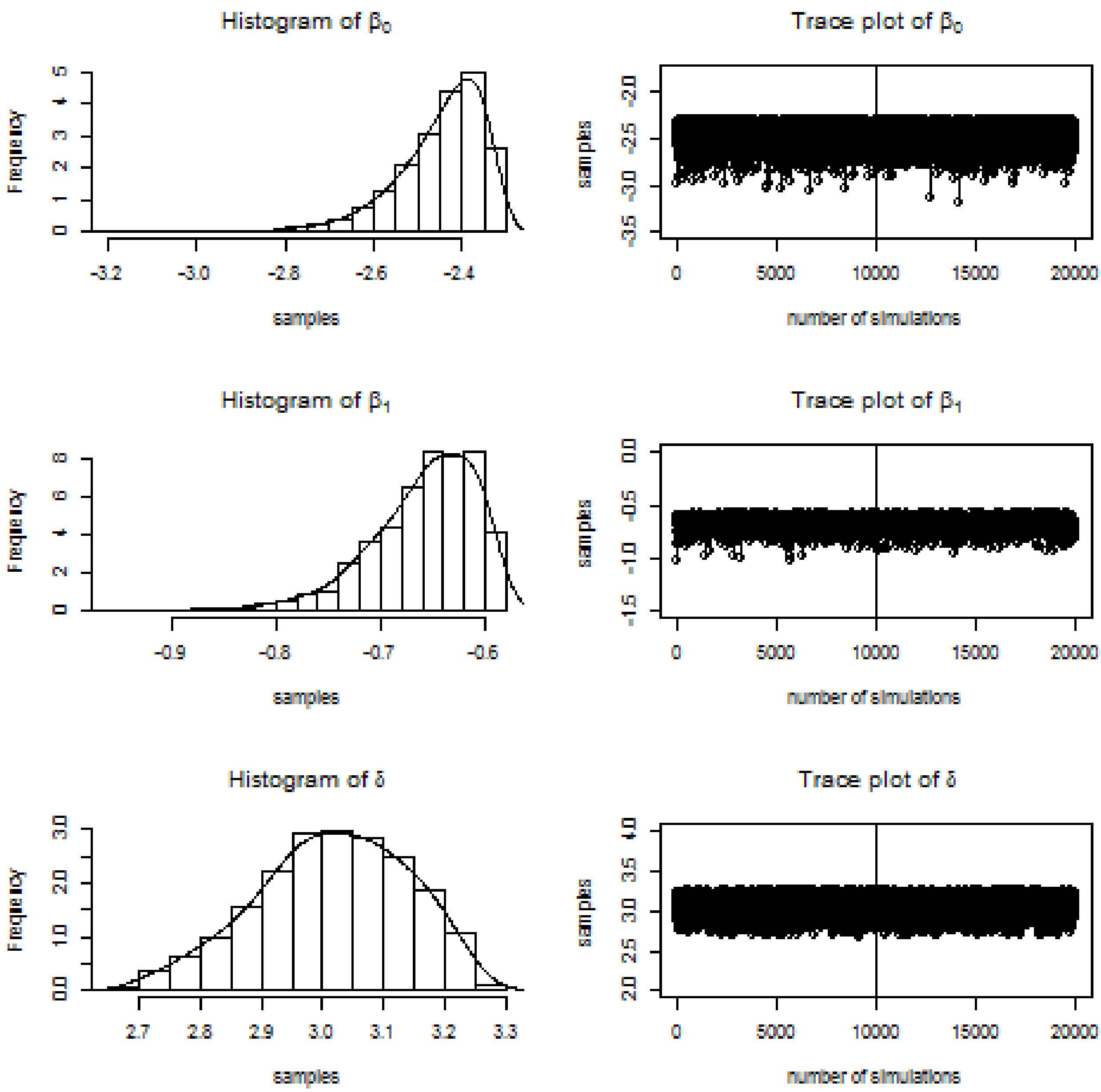

Figure A.14: Plots for joint sampling using 3 levels of stress with hyperparameters (6,-2,1)

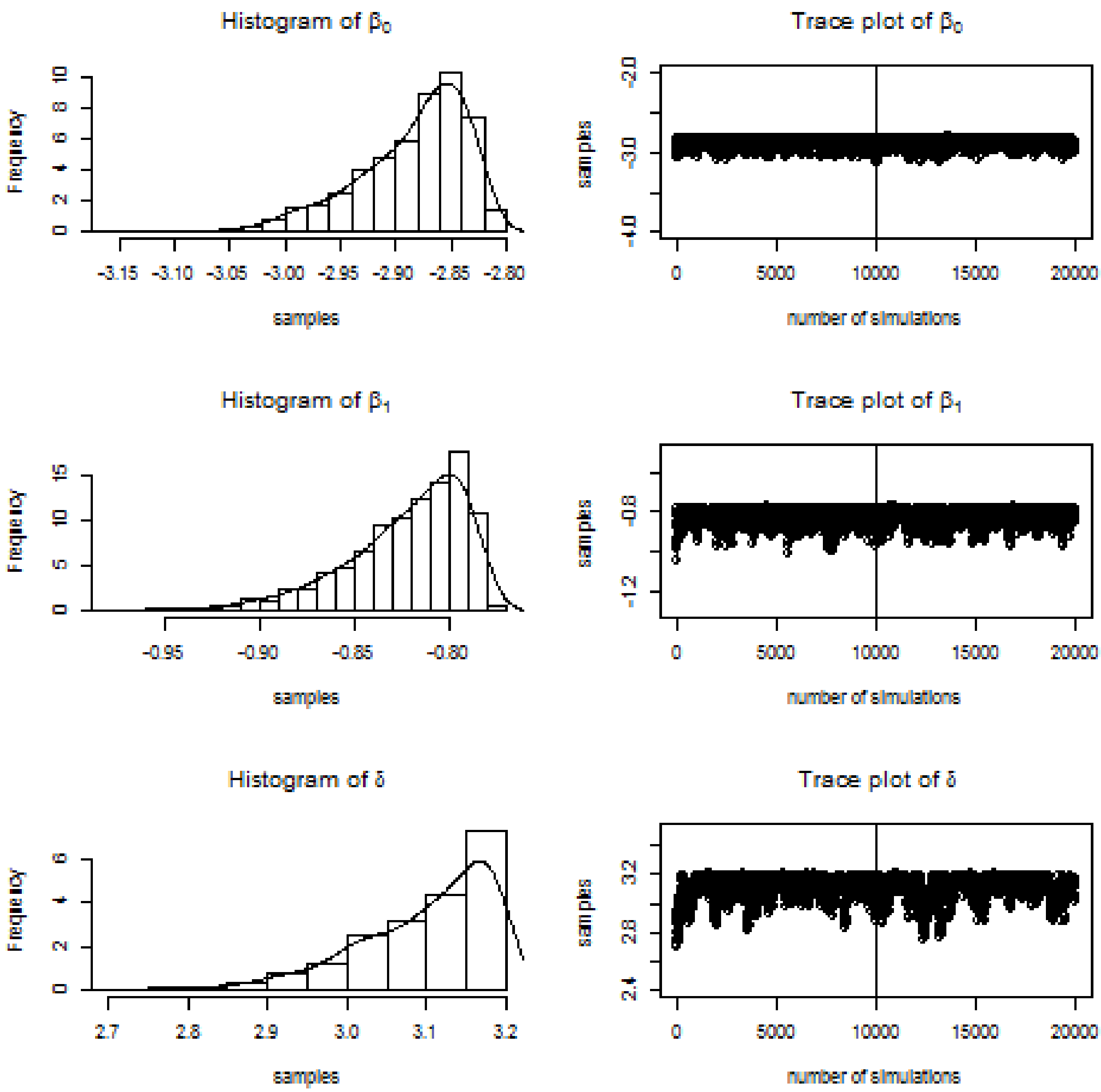

Figure A.15: Plots for joint sampling using 3 levels of stress with hyperparameters (2,-8,1)

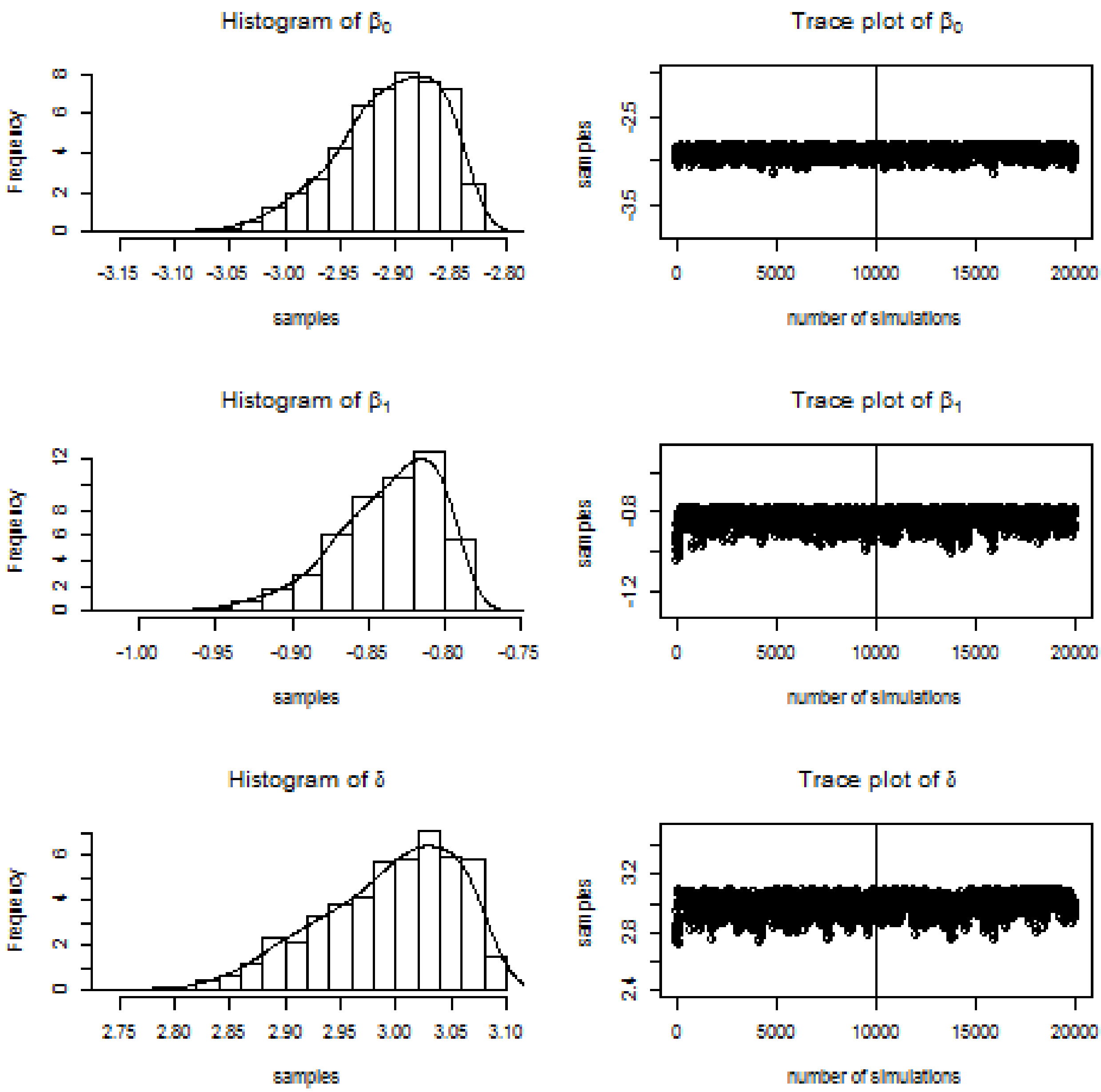

Figure A.16: Plots for joint sampling using 4 levels of stress with hyperparameters (1,-2,2)

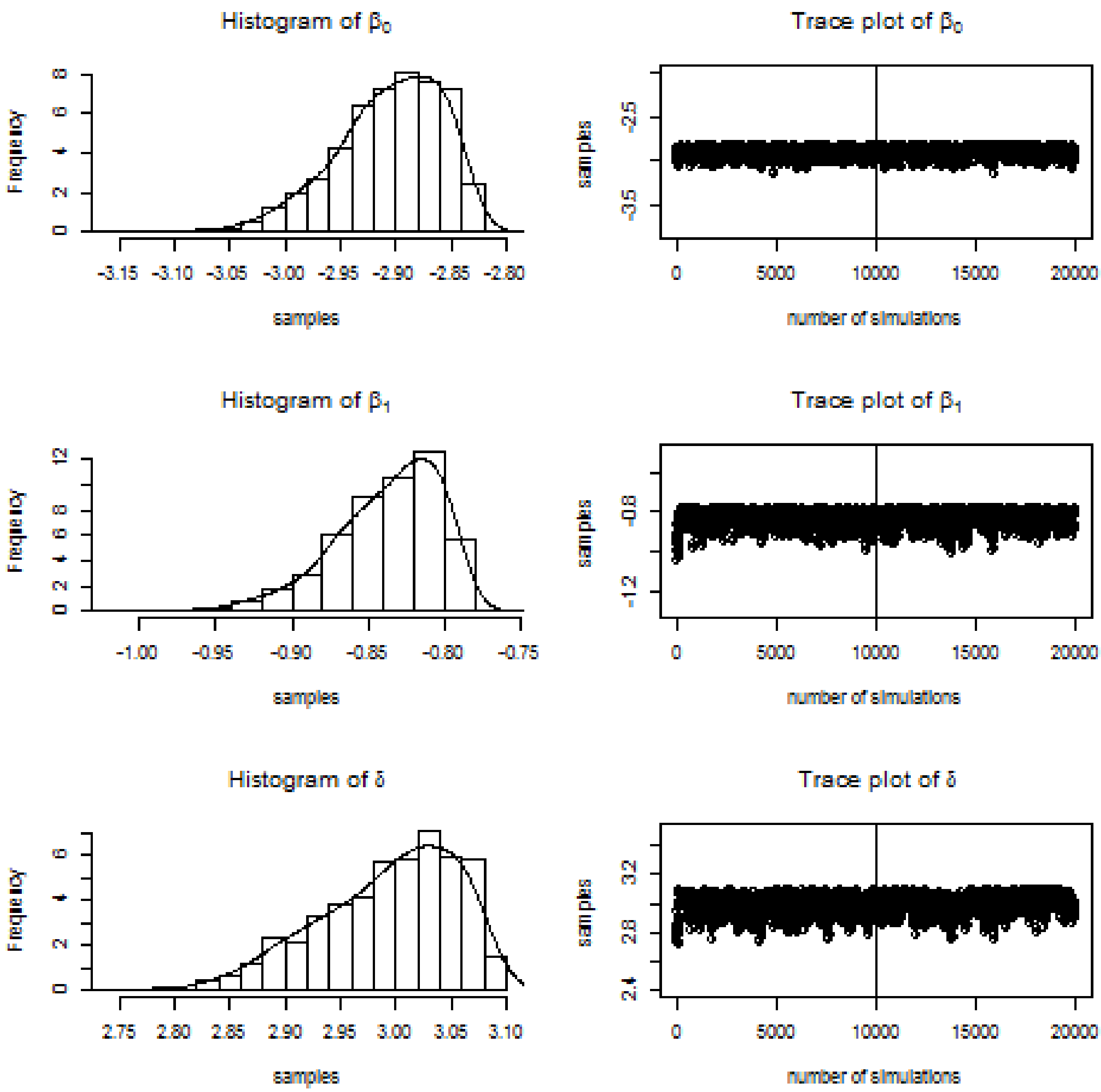

Figure A.17: Plots for joint sampling using 4 levels of stress with hyperparameters (6,-2,1)

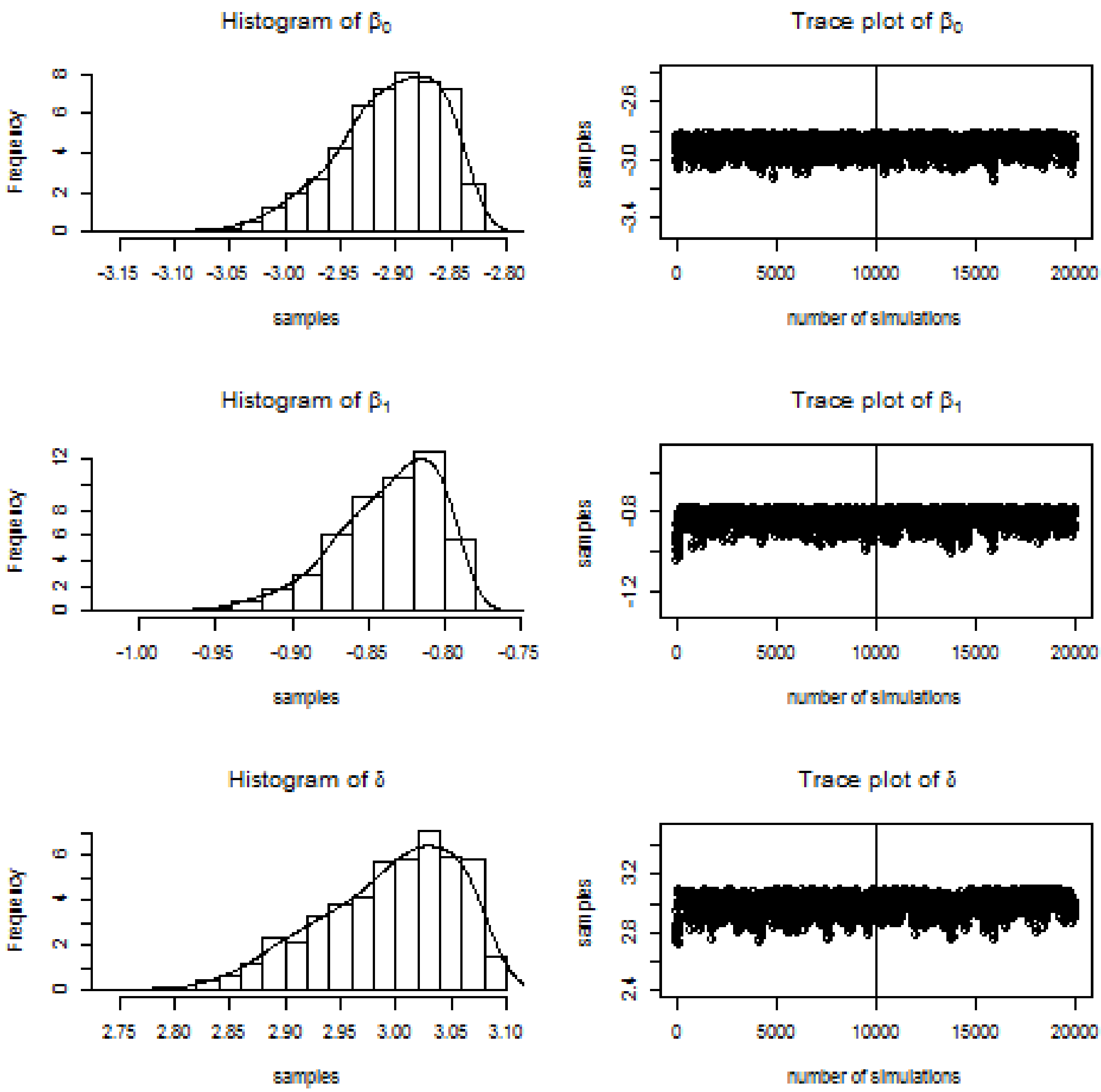

Figure A.18: Plots for joint sampling using 4 levels of stress with hyperparameters (2,-8,1)

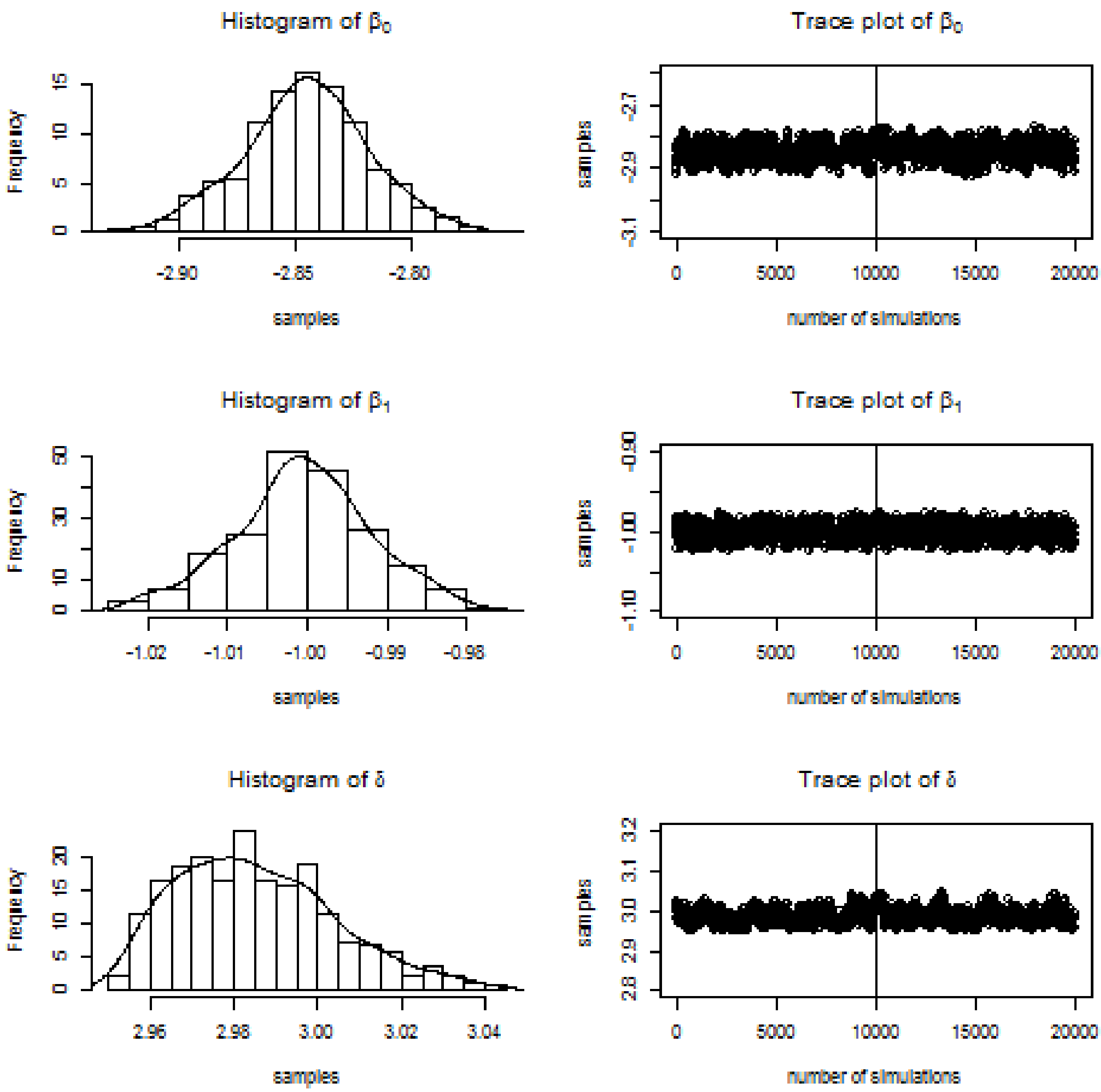

Figure A.19: Plots for joint sampling using 5 levels of stress with hyperparameters (1,-2,2)

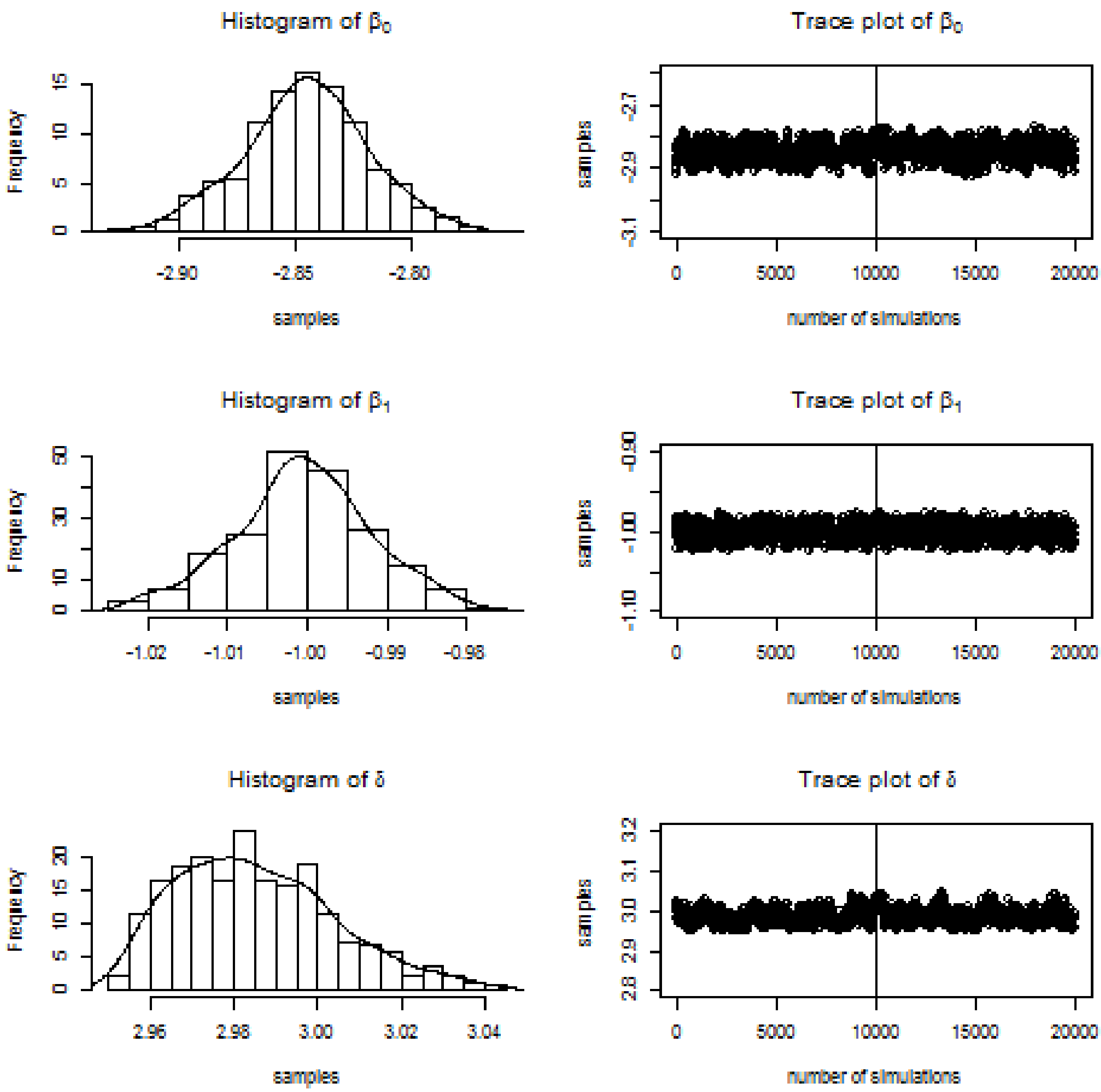

Figure A.20: Plots for joint sampling using 5 levels of stress with hyperparameters (6,-2,1)

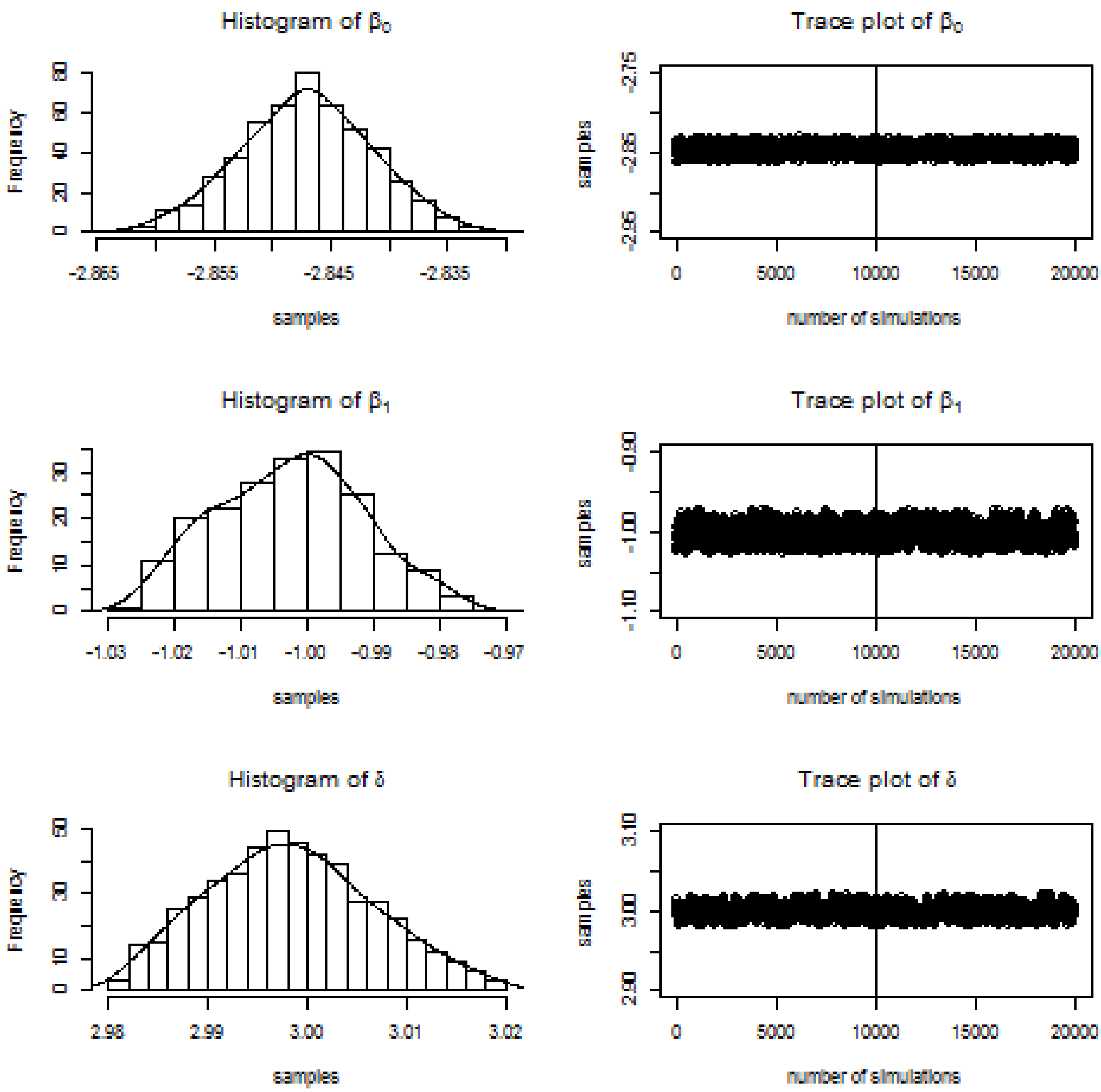

Figure A.21: Plots for joint sampling using 5 levels of stress with hyperparameters (2,-8,1)

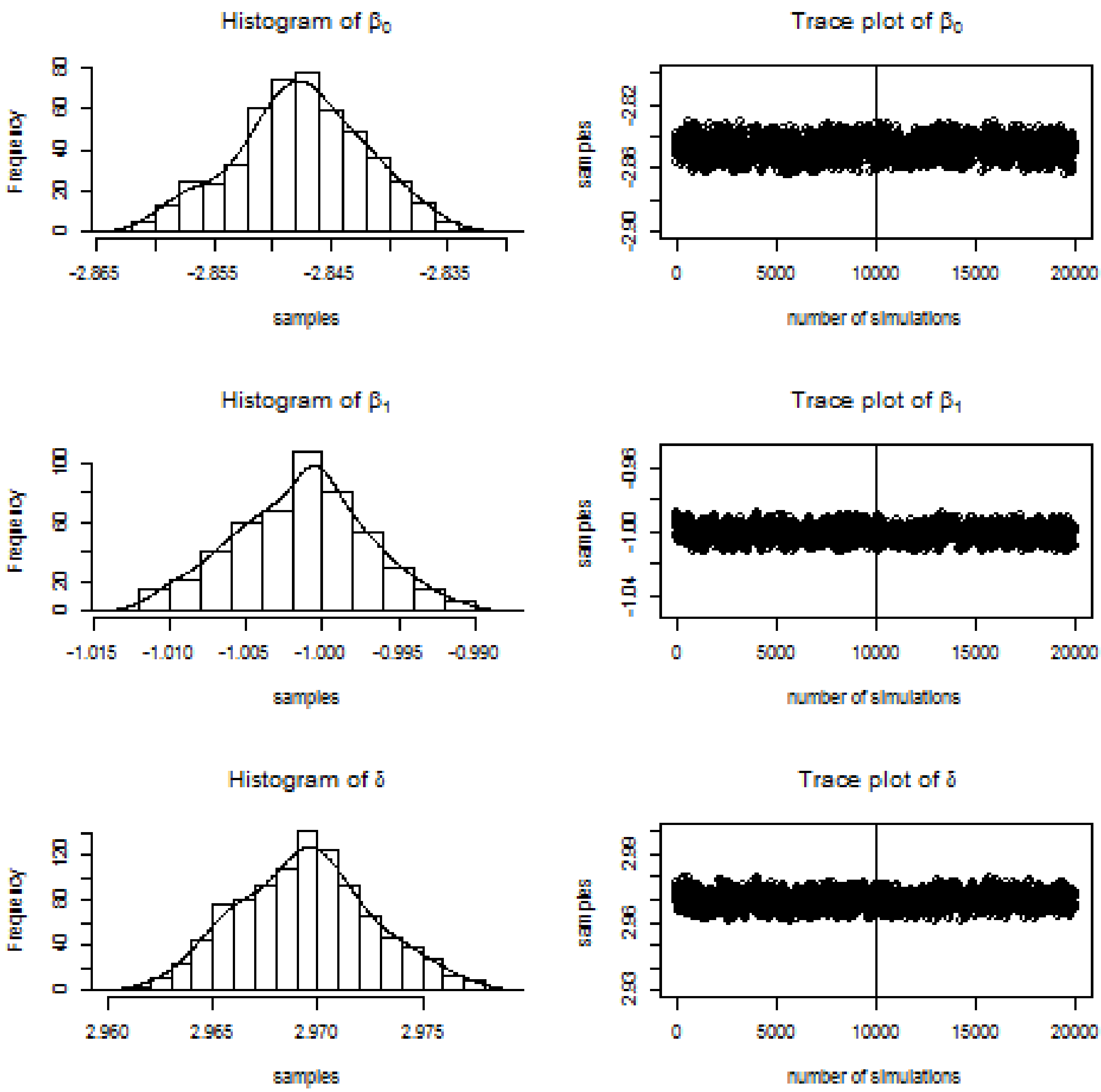

Figure A.22: Plots for joint sampling using 6 levels of stress with hyperparameters (1,-2,2)

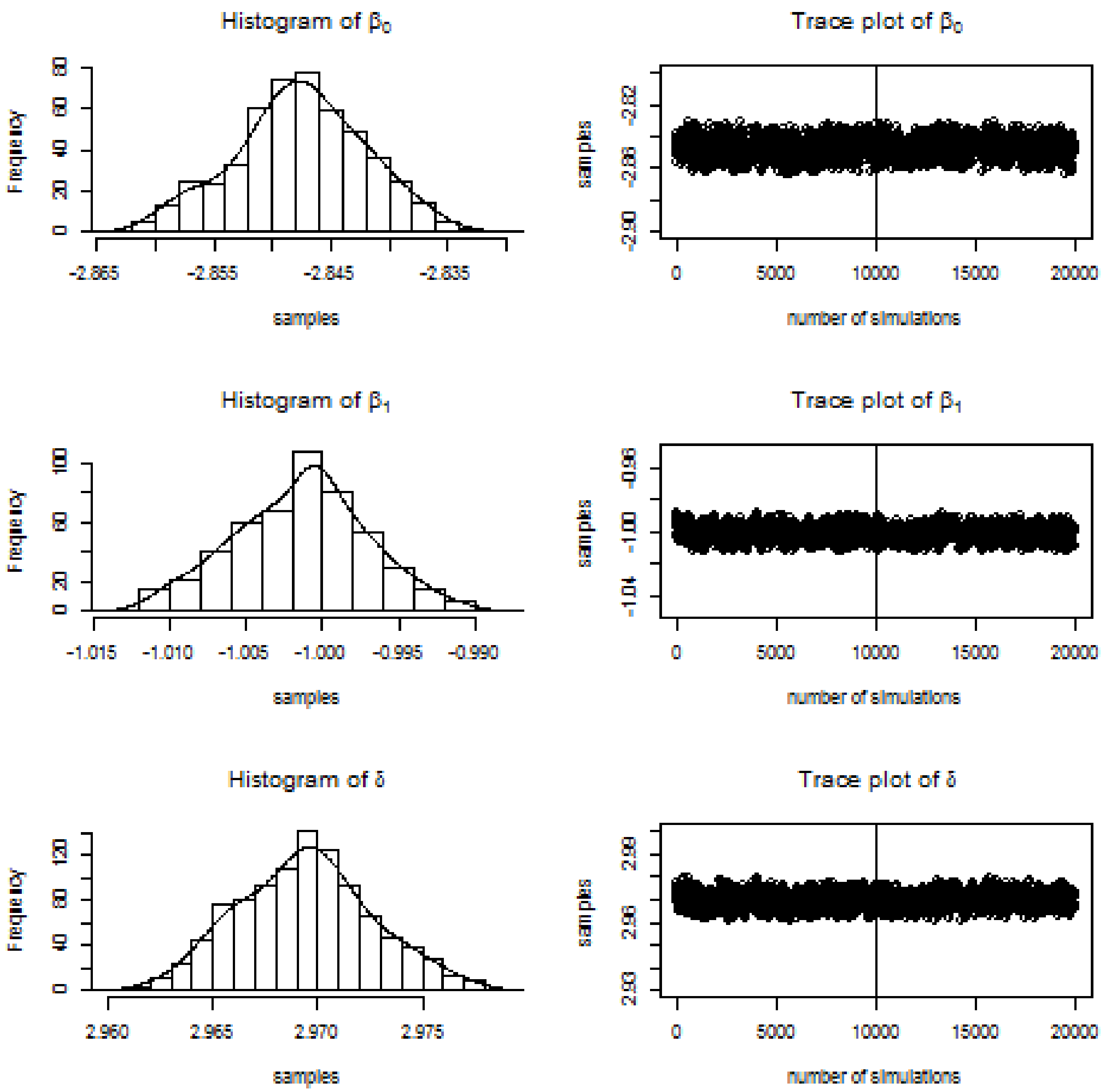

Figure A.23: Plots for joint sampling using 6 levels of stress with hyperparameters (6,-2,1)

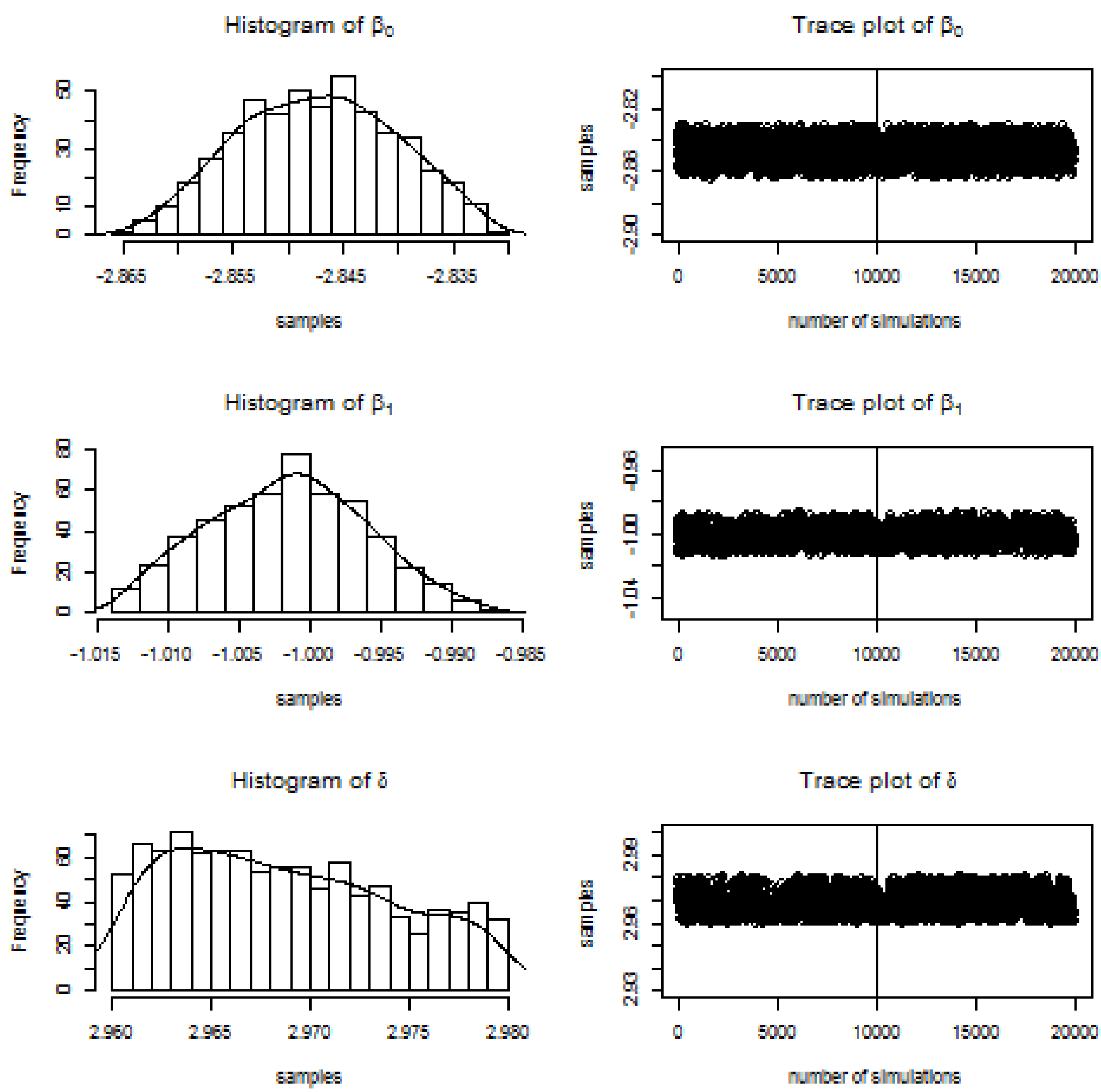

Figure A.24: Plots for joint sampling using 6 levels of stress with hyperparameters (2,-8,1)

Curriculum Vitae

Hao Yang Teng was born on October 11, 1989 in Malaysia. The first son of Say Kow Teng and Gaik Ai Goh. Both my parents pursued their education in Malaysia until college. My father worked as a high school mathematics teacher for 35 years and retired 2 years ago. My mother is a housewife. After I completed my high school education in 2006, I went to INTI international university college to pursue South Australian Matriculation. I chose to major in Actuarial studies at the University of Melbourne, Australia, as it is related to mathematics, statistics and the profession is in high demand in my country. After my third year, I was admitted into the honors program in University of Melbourne by invitation. Upon completing my studies, I was offered an actuarial position from Life Insurance Prudential Berhad Malaysia. After working for several months, I realized I had drifted away from my actual goal and aspiration. I had always wanted to pursue a Ph.D. After much thought and encouragements from my father, I decided that would be the most appropriate time. I was later admitted into the statistics masters program at University of Texas at El Paso with funding as a teaching assistant.

Permanent address: 783, Jalan Selesa 8, Kuchai Lama, Taman Gembira
58200 Kuala Lumpur, Malaysia.