Chapter 6.
EXPERIMENTAL SETTINGS AND SIMULATION STUDIES

There are two parts in this chapter. In the first section, we describe real experimental settings in our study. Since the real data set can not be revealed due to confidential regulation, in the second section, we generate a simulation data set to study the performance of parameter estimation and the degradation selection criterion. For checking model assumptions, we provide a procedure which consists of several graphic methods associated with some naive estimators, to sketch the information contained in the data set. The procedure can also be considered as a previous analysis of the data set.

6.1 Experiment Settings

As mentioned in Chapter 1, under the military standard, the lifetime of a MK71 electric detonator is typically lasting very long. Under the normal using condition, it takes long time to observe a failure occurs for the highly reliability products. Because of the time and cost constrains, it is almost impossible to collect data by only waiting for a failure occurring. Therefore, an accelerated experiment is conducted to deal with this problem. In our experiment, temperature is chosen to be the only accelerated factor, because it is easy to control and is the major effect for lifetime according to experiences.

The temperature setting in each day is cyclic. At the beginning of each cycle the temperature is increasing from the room temperature (say $25°C$) until reaching the high

temperature. It maintains at the high temperature for eight hours then decreases to the low temperature. The same, the temperature maintains at the low temperature for eight hours then return to the room temperature. It takes four hours to adjust the temperature gradually and measure (at the measuring times) between the periods of high and low temperature. Considering the cost and efficiency, 30 experimental units are applied to each three levels equally. The low temperature is set to be $-54°C$, and the high temperatures are set respectively to be $60°C$, $70°C$, and $80°C$ for level 1, 2 and 3, where $-54°C$ and $80°C$ are the lowest and highest temperature under which the experiment can be done safely.

According to engineer's experience, the probability of a MK71 electric detonator's lifetime exceeding 60 days is less than a half when the experiment is proceed at level 3. For each level, there are 10 measuring times, the $6th, 12th, \cdots, 60th$ day since the experiment begins. Thus, at the end of experiment, we have a good chance to observe failures occurring. Finally, all the experimental units fired up to see if they are still functioning in the end of the experiment.

6.2 Simulation Studies

According to the real experiment setting, we construct a simulation study. In the simulation study, 30 random samples, $T_1, \cdots, T_{30}$, are sampled according to (3.2) with covariate $W_1, \cdots, W_{30}$, respectively. The parameter (α_0, α_1) is chosen to be $(0.01, 0.05)$,

and the $W's$ are as follows:

$$W_i = \begin{cases} 35 & i = 1, \cdots, 10, \\ 45 & i = 11, \cdots, 20, \\ 55 & i = 21, \cdots, 30, \end{cases}$$

which means the data are from 3 different groups with covariates 35, 45 and 55, respectively. They denote the temperature differences of the experimental and the room temperatures. For example, in group 1, the covariate 35 is the experimental temperature 60 minus 25, which is the assuming room temperature. Under the setting, (5.4) can be solved analytically as

$$\begin{cases} \alpha_0 = \frac{30}{\sum_{i=1}^{30} E_{\boldsymbol{\theta}^{(p)}}[T_i|\boldsymbol{Y}_i, Z_i] \exp(\alpha_1 W_i)}, \\ \alpha_1 = \frac{1}{20} \log \frac{\sum_{i=1}^{10} E_{\boldsymbol{\theta}^{(p)}}[T_i|\boldsymbol{Y}_i, Z_i]}{\sum_{i=11}^{30} E_{\boldsymbol{\theta}^{(p)}}[T_i|\boldsymbol{Y}_i, Z_i]}. \end{cases}$$

The random sample is displayed in Table 1. Note that they are unobservable in reality.

group1	61.51	12.53	9.66	5.07	18.13	21.71	6.99	18.44	10.79	0.18
group2	24.67	3.48	8.03	3.70	1.12	21.63	14.90	0.83	25.68	32.96
group3	0.33	3.02	1.17	7.82	22.78	3.03	3.28	1.96	14.82	1.59

Table 1: The random sample which is used in the simulation study.

We consider 4 auxiliary measurements, Y^1, Y^2, Y^3, and Y^4. For the linear degradation the lifetime is determined by the slope and the difference between the intercept and the threshold. Considering the shifting, the intercept itself does not have impacts on the

lifetime. In the study we, therefore, consider only the effects from the slopes and the variations of the random errors. The parameter configurations for these four auxiliary measurements are in Table 2:

measurements	$(\beta_0, \beta_1, \sigma^2)$
Y^1	$(100, -10, 1)$
Y^2	$(100, -10, 4)$
Y^3	$(100, -20, 1)$
Y^4	$(100, -20, 4)$

Table 2: The parameter setting in the simulation study.

By rescaling, the measures are taken at $(t_{i1}, \cdots, t_{i10}) = (1, \cdots, 10)$ for all i. The results of generalized Y_{ij}'s as well as Z_i's (=1 if $T_i > 10 = t_{i10}$) are displayed in Figure 2.

In each graph of Figure 2, we observe that there is a linear decreasing trend and a cut point between the circles and crosses (around 90 for Y^1 and Y^2; 80 for Y^3 and Y^4, especially for Y^3, there is even no overlap). Thus, all of the $Y^1, \cdots, Y^4$ have these two important characters of an ideal degradation measurement described in section 3.1.

If the unobserved lifetimes were known, they could be used to check the assumptions about the lifetime distributions and the effects of the covariates (the proportional hazards model). To gain more information about the potential lifetime distribution, we need, at first, to estimate the unobserved T_i's. To do the implementation, we propose a 2-step method described as follows.

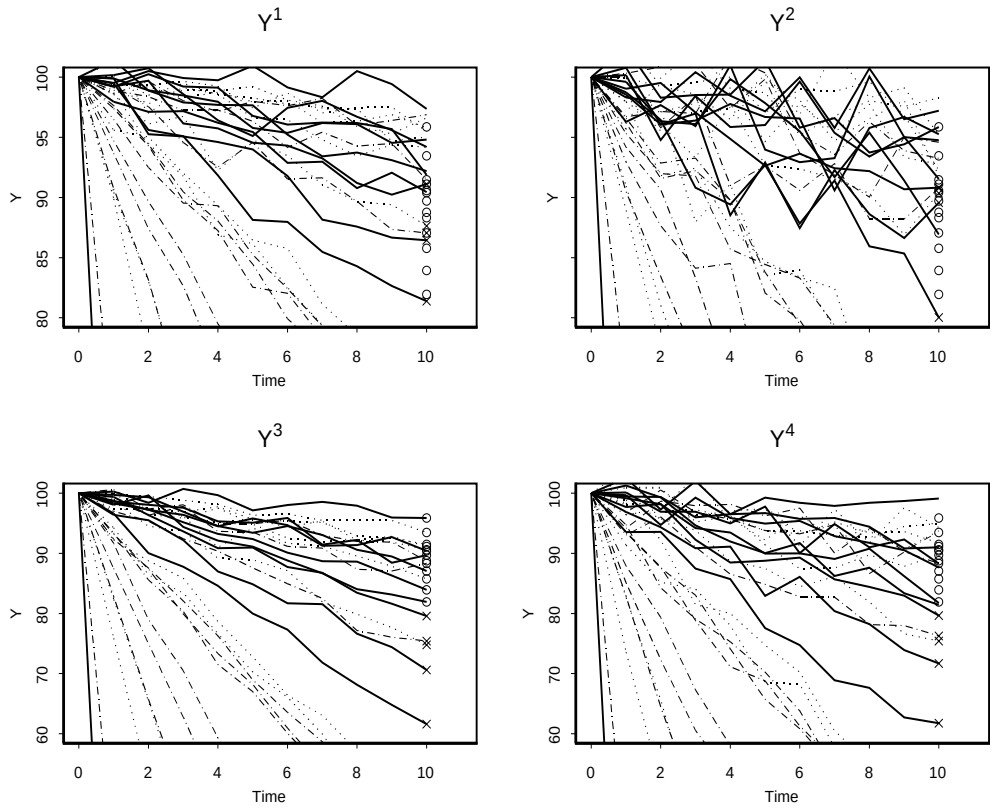

Figure 1: The simulated $Y_{ij}^1, \cdots, Y_{ij}^4$. For each i, Y_{ij} are connected with (—) for those i in group 1, $(\cdots)$ for group 2 and $(\cdot\text{—}\cdot)$ for group 3. At $t_{i10} = 10$, the circles denote those with $Z_i = 1$ and the crosses with $Z_i = 0$.

Step 1: Find a suitable threshold $\widetilde{\tau}$.

Divide Y_{i10}'s into two groups according to $Z_i = 0$ or 1. Let Y_{i10}^0 and Y_{i10}^1 denote the classification result. Then, the threshold can be estimated by

$$\widetilde{\tau} = \frac{\max Y_{i10}^0 + \min Y_{i10}^1}{2}. \tag{6.1}$$

The estimator in (6.1) usually works better than the Fisher's discriminant function,

which is $[\overline{Y^0_{i10}} + \overline{Y^1_{i10}}]/2$, where $\overline{Y^0_{i10}}$ and $\overline{Y^1_{i10}}$ are the averages of two groups, respectively. The reason is that the sample mean is affected by extreme observations more, and the variance of observations from different linear degradation paths getting larger as the observing time increasing. Hence, the Fisher's discriminant function performs poorly when an experiment takes longer time.

Step 2: Estimate the lifetime T_i.

For each unit i, consider the linear model separately

$$Y_{ij} = \beta_0 + \beta_i t_{ij} + \varepsilon_{ij}. \tag{6.2}$$

Let $(\widetilde{\beta_0}, \widetilde{\beta_i})$ be the least square estimator (LSE) of (β_0, β_i). Then we estimate T_i by

$$\widetilde{T_i} = \frac{\widetilde{\tau} - \widetilde{\beta_0}}{\widetilde{\beta_i}}.$$

The residuals can be also used to check the normality assumption of random errors and to estimate σ^2.

Once the T_i's are estimated, the rest model checking procedure can be done by using some standard graphic methods. cf. Chapter 6 of Meeker and Escobar (1998).

In our simulation study, we estimate $T_1, \cdots, T_{30}$ according to the above 2-step method. To check the exponential lifetime assumption, in Figure 3, we plot $-\log(1 - \widehat{S}(\widetilde{T}))$ over t_{ij}, where $\widehat{S}(\widetilde{T})$ is the empirical survival function.

Three parallel lines can be observed in each graph except measurement Y^2 (the circles),

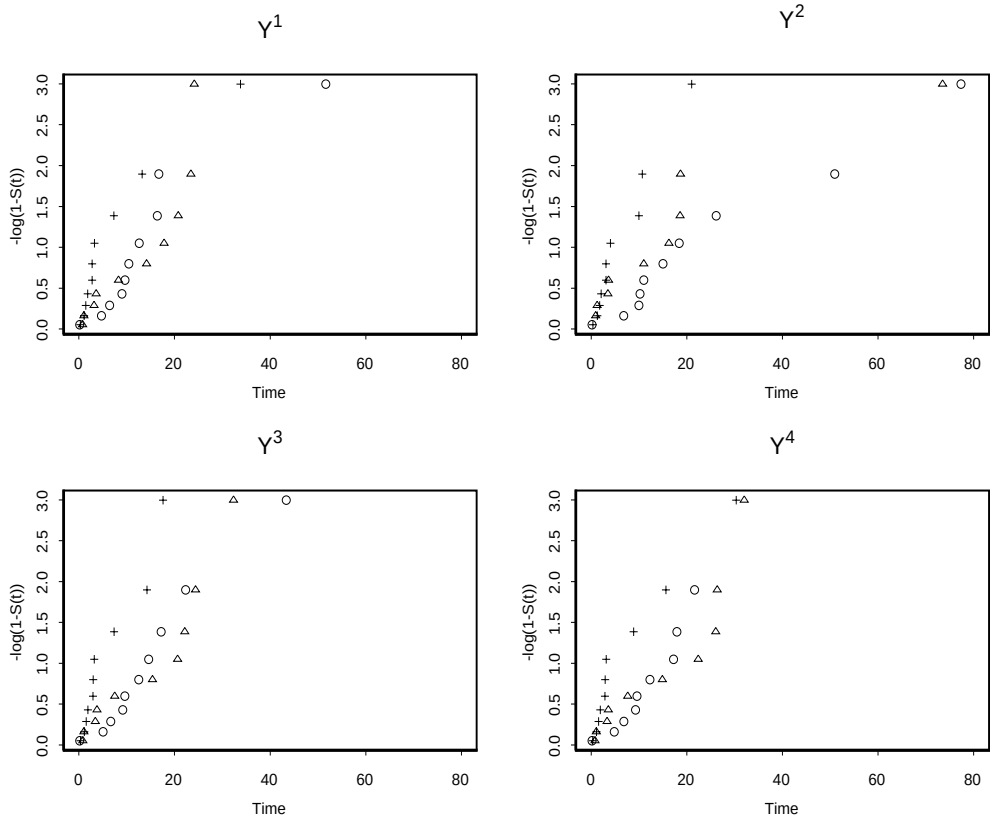

Figure 2: $-\log(1-\widehat{S}(\widetilde{T}))$ vs. t pluses are for group 1, triangles for group 2 and circles for group 3

which means the lifetime in each group follows an exponential distribution with different means.

To study covariates effect, the reciprocal of the sample mean in each group, which is the MLE of the hazard rate in an exponential distribution (α_{g1} for group 1, α_{g2} for group 2, and α_{g3} for group 3), is plotted versus the corresponding covariates as shown in Figure 4.

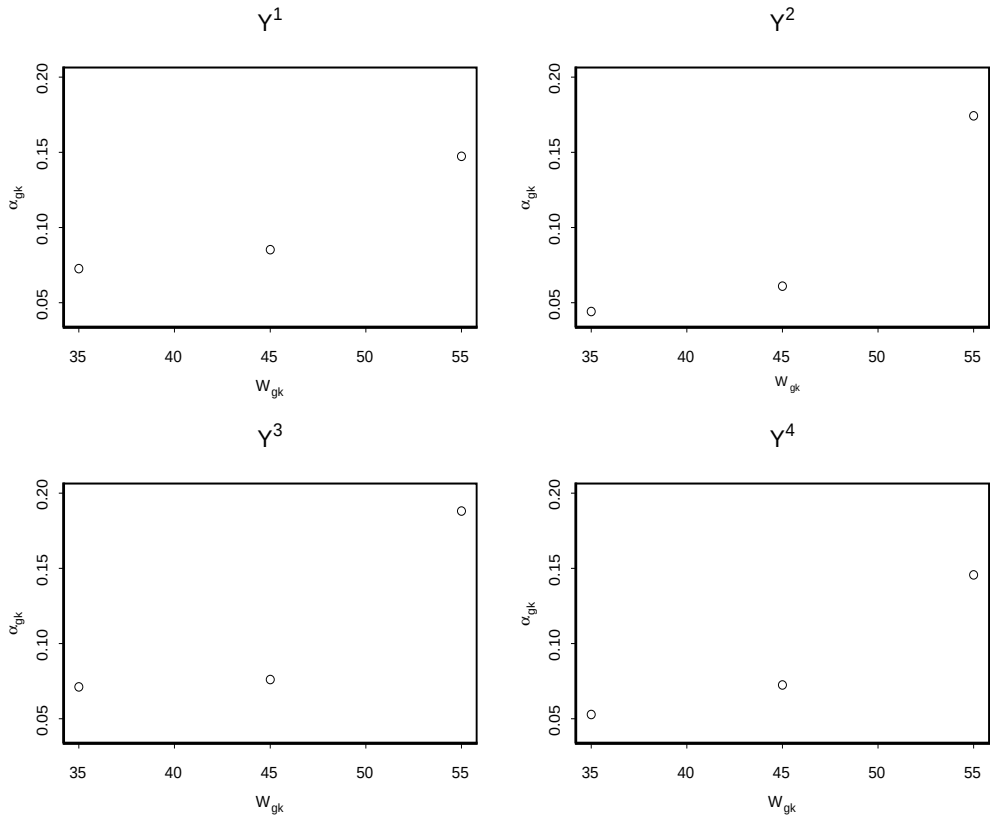

Figure 3: α_{gk} vs W_{gk}

We can observe that there is an exponential relation between the parameters and the covariates in Figure 4. Taking the logarithm over all of the α's, we fit the linear regression

$$\log \alpha_{gk} = \gamma_0 + \gamma_1 W_{gk} + \epsilon_k, \tag{6.3}$$

for $k = 1, 2, 3$, where $W_{g1} = 35, W_{g2} = 45$ and $W_{g3} = 55$. Then, the α_0 and α_1 in (3.2) can be estimated by

$$(\widetilde{\alpha_0}, \widetilde{\alpha_1}) = (\exp \widetilde{\gamma_0}, \widetilde{\gamma_1}),$$

where $(\widetilde{\gamma_0}, \widetilde{\gamma_1})$ is the LSE of (γ_0, γ_1) in (6.3).

Beside the usefulness of model checking for given assumptions, the above procedure also provides an initial estimator of the parameters, which can be used as the starting point of the EM algorithm.

In the following, we will make one modification which has mentioned in section 5.1 to implement the EM algorithm. Before that, we note the weight function is very concentrated on its maximum, which is a main reason to explain why the calculation of conditional expectations in (5.3) does not work well. To see this, the weight function is a product of several normal density functions (it is 10 in our simulation study.) Regarded it as a function of $1/T_i^{(p)}$, it is the normal random variable with mean $\mu_i^{(p)}$ and variance $\sigma_i^{2(p)}$, with

$$(\mu_i^{(p)}, \sigma_i^{2(p)}) = \Big(\frac{\sum_{j=1}^{n_i} (Y_{ij} - \beta_0^{(p)}) t_{ij}}{\beta_1^{(p)} \sum_{j=1}^{n_i} t_{ij}^2}, \frac{\sigma^{2(p)}}{\beta_1^{(p)2} \sum_{j=1}^{n_i} t_{ij}^2} \Big).$$

Under our simulation setting, the variance is 2.60×10^{-5}, 1.04×10^{-4}, 6.49×10^{-6} and 2.60×10^{-5} when $Y^1, \cdots, Y^4$ are used as an auxiliary measurement, respectively. Therefore, only few draws contribute to the calculation of the conditional expectations in (5.3), and most of them vanish. Sometimes, the weight is rounded to 0 and overflow occurs.

To make the calculation more efficient, we add a point $\widehat{T_i^{(p)}}$, to the Monte Carlo sample, which is defined as

$$\widehat{T_i^{(p)}} = \begin{cases} min[\frac{1}{\mu_i^{(p)}}, t_{10}] & \text{if} \quad Z_i = 1, \\ max[\frac{1}{\mu_i^{(p)}}, t_{10}] & \text{if} \quad Z_i = 0. \end{cases}$$

Note that this maximizes the weight function, and the modification solves the overflow

problem. The extra point contributes $\exp\{-\frac{1}{2}(\frac{\widehat{1/T^{(p)}}_i-\mu_i^{(p)}}{\sigma_i^{(p)}})^2\}$ (the normalizing constant, $\frac{1}{\sqrt{2\pi}\sigma^{(p)}{}_i}$, is omitted) to the denominator of (5.3), which is closed to 1 if $\widehat{T_i^{(p)}}$ is closed to $\frac{1}{\mu_i^{(p)}}$.

The approximation above is similar to the Laplace method. In traditional Laplace method, the integration is approximated by plugging the maximum of the integrand into the integrand. Here we plug in the maximum of the weight function. The loss of efficiency is negligible because the weight function is extremely peaked. What we gain is that once the maximum of the weight function is obtained, it can be used to calculate all the expectations which are needed in the E-step. In addition, no calculation of normalizing constant is needed. The results are reported in Table 3.

If CCP is used as the measurement selection criterion, Y^3 is the best among all the 4 measurements ($|\widehat{\beta_1}|/\widehat{\sigma}$=9.59, 5.51, 20.31 and 9.48 for $Y^1, \cdots, Y^4$, respectively). When it is used as a degradation measurement, the threshold is estimated by $80.76(\widehat{\beta_0}+\widehat{\beta_1})$.

measurements	$\widetilde{\alpha_0}, \widehat{\alpha_0}(\widehat{Std.})$	$\widetilde{\alpha_1}, \widehat{\alpha_1}(\widehat{Std.})$
Y^1	$1.97 \times 10^{-2}, 7.91 \times 10^{-3}(2.54 \times 10^{-2})$	$3.54 \times 10^{-2}, 5.62 \times 10^{-2}(2.45 \times 10^{-2})$
Y^2	$3.56 \times 10^{-3}, 9.46 \times 10^{-3}(4.60 \times 10^{-2})$	$6.86 \times 10^{-2}, 4.88 \times 10^{-2}(2.95 \times 10^{-2})$
Y^3	$1.13 \times 10^{-2}, 9.38 \times 10^{-3}(2.18 \times 10^{-2})$	$4.87 \times 10^{-2}, 5.17 \times 10^{-2}(2.36 \times 10^{-2})$
Y^4	$8.38 \times 10^{-3}, 5.08 \times 10^{-3}(1.34 \times 10^{-2})$	$5.08 \times 10^{-2}, 6.27 \times 10^{-2}(2.80 \times 10^{-2})$

$\widetilde{\beta_0}, \widehat{\beta_0}(\widehat{Std.})$	$\widetilde{\beta_1}, \widehat{\beta_1}(\widehat{Std.})$	$\widetilde{\sigma_2}, \widehat{\sigma_2}(\widehat{Std.})$	$\widetilde{r}, \widehat{r}(\widehat{Std.})$
$100.13, 100.16(0.17)$	$-9.35, -9.20(0.61)$	$1.22, 0.92(0.23)$	$8.48, 9.59(1.09)$
$99.95, 99.98(0.45)$	$-11.07, -11.05(1.49)$	$5.47, 4.02(0.52)$	$4.73, 5.51(0.96)$
$100.03, 99.99(0.14)$	$-19.27, -19.27(0.52)$	$1.23, 0.90(0.91)$	$17.38, 20.31(2.71)$
$99.72, 99.91(0.35)$	$-19.17, -18.68(1.45)$	$4.93, 3.88(0.41)$	$8.63, 9.48(1.05)$

Table 3: The results of estimation ($\widetilde{\theta}, \widehat{\theta}(\widehat{Std.})$) stands for the initial estimator, the refined estimator by EM algorithm (the standard deviations are estimated by 1000 bootstrap replicates).

To see the performance of the estimates, we repeat the procedure above 1,000 times and the means and the standard deviations of the sampling distributions are in Table 4.

measurements	α_0	α_1
Y^1	$1.93 \times 10^{-2}(3.53 \times 10^{-2})$	$5.09 \times 10^{-2}(2.54 \times 10^{-2})$
Y^2	$2.66 \times 10^{-2}(3.86 \times 10^{-2})$	$4.95 \times 10^{-2}(3.00 \times 10^{-2})$
Y^3	$1.91 \times 10^{-2}(3.67 \times 10^{-2})$	$5.04 \times 10^{-2}(2.32 \times 10^{-2})$
Y^4	$1.90 \times 10^{-2}(2.82 \times 10^{-2})$	$5.14 \times 10^{-2}(2.61 \times 10^{-2})$

β_0	β_1	σ^2	r
100.02(0.17)	−9.90(0.65)	0.96(0.28)	10.24(1.21)
100.24(0.44)	−9.33(1.36)	4.02(0.50)	4.69(0.81)
100.00(0.15)	−19.95(0.56)	0.95(0.62)	21.14(4.23)
100.05(0.34)	−19.81(1.40)	3.77(0.53)	10.34(1.66)

Table 4: The means and the standard deviations of the sampling distributions of $\widehat{\theta}$ based on 1,000 replicates.

From Table 4, we observe that the estimation of α_0 does not perform well, although, the confidence interval covers the true parameter. The poor performance is due to the variance caused by extrapolation. To increase the precision, more sample size and more sophisticated design are needed. There is only once that Y_3 is not selected as the best degradation measurement.

We generate also a data set to demonstrate the effectness of the MCMC algorithm described in section 5.2. The parameters and the prior setting are given in Table 5.

parameters	true value	prior
α	0.1	$G(4, 0.025)$
β_0	100	$N(100, 2500)$
β_1	-20	$N(-20, 100)$
σ^2	1	$IG(6, 0.2)$

Table 5: The true value and prior setting in the simulation study.

The prior distribution is chosen so that the means are the true values, and the ratios of mean to its standard deviation is 2. The generated ten latent lifetimes are

$$1.22 \quad 9.84 \quad 8.63 \quad 14.10 \quad 5.50$$

$$5.47 \quad 9.30 \quad 5.39 \quad 19.26 \quad 7.86$$

The measuring time is still set to be $(1, 2, \cdots, 10)$. The starting values are sampled from the prior distribution. We iterate the chain $50,000$ (the M_0 in section 5.2) times, and only the last $2,000$ (the M) runs are recorded. The accepted rate of the candidates from the proposal distribution is about 10% empirically. The results of the final $2,000$ runs are displayed in Figure 5 and the histograms are in Figure 6.

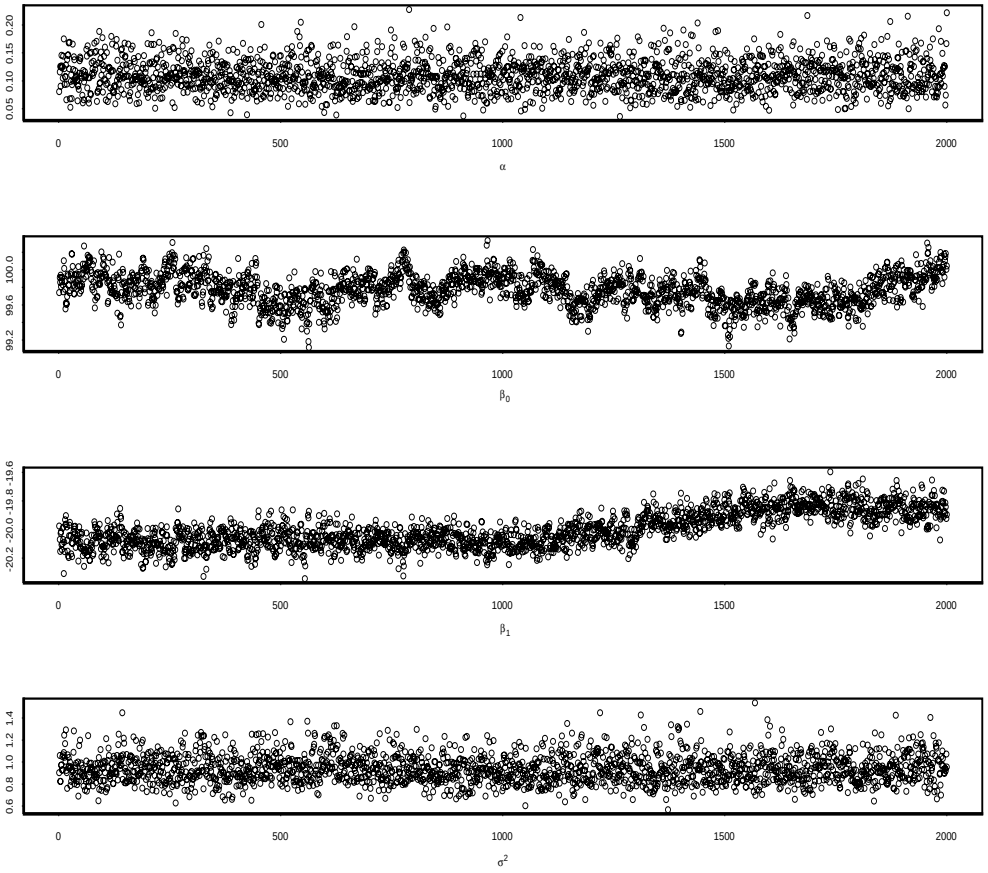

Figure 4: The final 2000 runs of the Markov chain

From the estimated posterior distribution, the posterior means and the posterior medians can be obtained as shown in Table 6.

	α	β_0	β_1	σ^2
Mean	0.1099	99.7700	-20.2019	0.9421
Median	0.1074	99.7713	-20.0218	0.9277

Table 6: The obtained posterior means and the posterior medians in the simulation study.

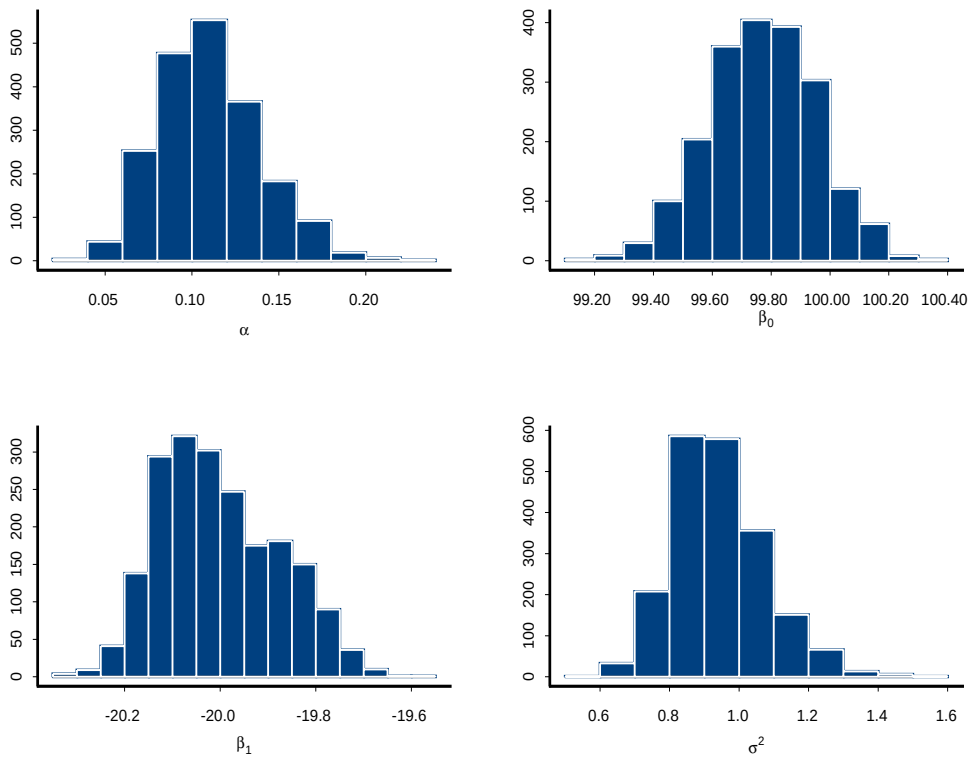

Figure 5: The histograms of the posterior distributions.