

**Figure 11.9** Sample paths for Metropolis–Hastings algorithm. The stationary density is standard normal and the proposal density  $q(u' | u)$  is  $N(u, \sigma^2)$ , with  $\sigma = 0.1, 0.5, 2.4$  and  $10$ . The initial value is  $u_0 = -10$  and the same seed is used for the random number generator in each case. Note the dependence of the acceptance rate and convergence to stationarity on  $\sigma$ . The horizontal dashed lines show the ‘usual’ range for  $\sigma$ .

**Table 11.10** Motorette data (Nelson and Hahn, 1972). Censored failure times are denoted by +.

$x$ ( $^{\circ}$ F)	Failure time (hours)									
150	8064+	8064+	8064+	8064+	8064+	8064+	8064+	8064+	8064+	8064+
170	1764	2772	3444	3542	3780	4860	5196	5448+	5448+	5448+
190	408	408	1344	1344	1440	1680+	1680+	1680+	1680+	1680+
220	408	408	504	504	504	528+	528+	528+	528+	528+

**Example 11.24 (Motorette data)** Table 11.10 contains failure times  $y_{ij}$  from an accelerated life trial in which ten motorettes were tested at each of four temperatures, with the objective of predicting lifetime at  $130^{\circ}$ F. We analyse these data using a Weibull model with

$$\Pr(Y_{ij} \leq y; x_i) = 1 - \exp\{-(y/\theta_i)^{\gamma}\}, \quad \theta_i = \exp(\beta_0 + \beta_1 x_i), \quad (11.48)$$

for  $i = 1, \dots, 4$ ,  $j = 1, \dots, 10$ , where failure time is taken in units of hundreds of hours and  $x_i$  is  $\log(\text{temperature}/100)$ .

Here we describe a simple Bayesian analysis using the Metropolis–Hastings algorithm. For illustration we take independent priors on the parameters,  $N(0, 100)$  on  $\beta_0$  and  $\beta_1$  and exponential with mean 2 on  $\gamma$ . Then the log posterior is

$$\begin{aligned} \ell_m(\beta_0, \beta_1, \gamma) &\equiv -(\beta_0^2 + \beta_1^2)/200 - \gamma/2 \\ &\quad + \sum_{i=1}^4 \sum_{j=1}^{10} d_{ij} \{\log \gamma + \gamma \log(y_{ij}/\theta_i)\} - (y_{ij}/\theta_i)^{\gamma}, \end{aligned}$$

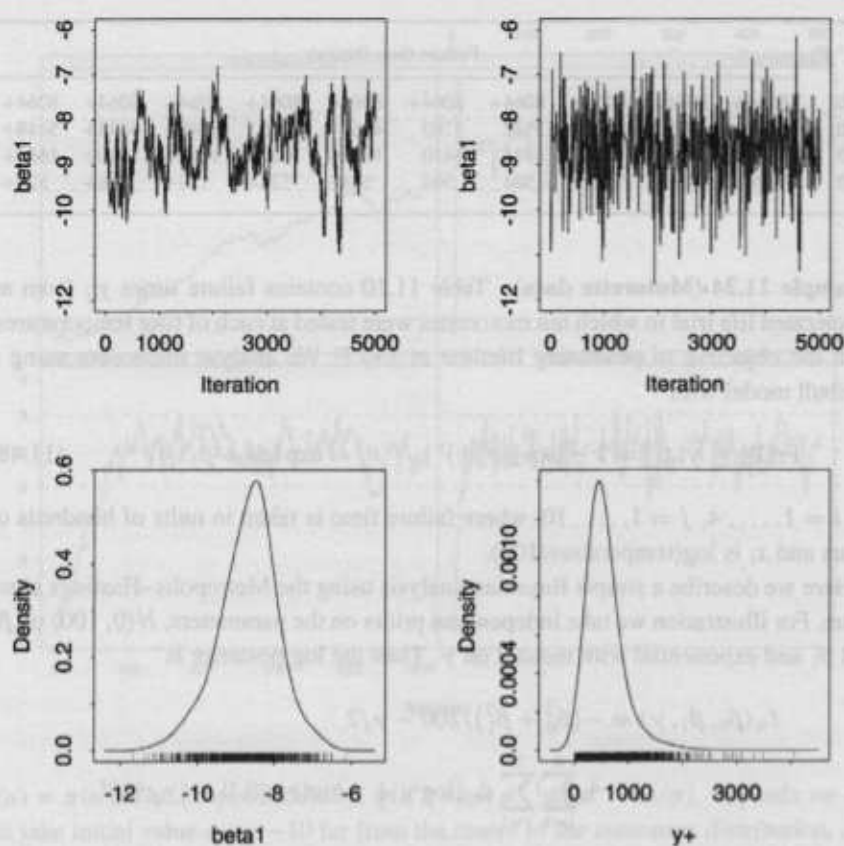
where  $d_{ij} = 0$  for uncensored  $y_{ij}$ .

For proposal distribution we update all three parameters simultaneously, by taking  $(\beta'_0, \beta'_1, \log \gamma') = (\beta_0, \beta_1, \log \gamma) + c(s_1 Z_1, s_2 Z_2, s_3 Z_3)$ , where the  $s_r$  are the standard errors of the corresponding maximum likelihood estimates,  $Z_r \stackrel{\text{iid}}{\sim} N(0, 1)$ , and  $c$  can be chosen to balance the acceptance probability and the size of the move. The ratio  $q(u | u')/q(u' | u)$  reduces to  $\gamma'/\gamma$ , so the acceptance probability equals

$$a\{(\beta'_0, \beta'_1, \gamma'), (\beta_0, \beta_1, \gamma)\} = \min[1, \exp\{\ell_m(\beta_0, \beta_1, \gamma) - \ell_m(\beta'_0, \beta'_1, \gamma')\} \gamma'/\gamma].$$

The chain is clearly irreducible and aperiodic, so the ergodic theorem applies.

We take initial values near the maximum likelihood estimates, and run the chain for 5000 iterations with  $c = 0.5$ . The sample path for  $\beta_1$  in the upper left panel of Figure 11.10 shows that despite its acceptance probability of about 0.3, the chain is not moving well over the parameter space. This is confirmed by the correlogram and partial correlogram for successive values of  $\beta_1$ , which suggest that the chain is essentially an AR(1) process with  $\rho_1 \doteq 0.99$ . In this case the variance inflation factor is  $\hat{\tau} = 199$ , so 5000 successive observations from the chain are worth about 25 independent observations. Sample paths for the other parameters are similar, and varying  $c$  does not improve matters. One reason for this is that  $\beta_0$  and  $\beta_1$  have correlation about  $-0.97$  *a posteriori*, and the proposal distribution does not respect this. It is better to



**Figure 11.10** Bayesian analysis of motorette data using Metropolis-Hastings algorithm. Upper panels: sample paths for  $\beta_1$  using two parametrizations, the right one more nearly orthogonal. Lower left: kernel density estimates of  $\pi(\beta_1 | y)$  and of  $\pi(Y_+ | y)$ , where  $Y_+$  is failure time predicted for 130 F.

**Table 11.11** Accuracy of Stirling's formula and related approximations.

reduce this correlation by replacing  $x$  by  $x - \bar{x}$ , after which  $\text{corr}(\beta_0, \beta_1 | y) \doteq -0.4$ . The sample path for  $\beta_1$  from a run of the algorithm starting near the new maximum likelihood estimates, with the new  $s_r$  and with  $c = 2$ , is shown in the upper right panel of Figure 11.10. This chain mixes much better, though its acceptance probability is about 0.2. The usual plots suggest that  $\beta_1$  follows an AR(1) process with  $\rho \doteq 0.9$ , and likewise for the other parameters, whose chains show similar good behaviour. Here  $\hat{\tau}$  has the more acceptable value 19, though 5000 iterations would remain too small in practice.

The lower panels of the figure show kernel density estimates of the posterior densities for  $\beta_1$  and for a predicted failure time  $Y_+$  for temperature 130°F. Once convergence has been verified, it is easy to obtain values for  $Y_+$ , simply by simulating a Weibull variable from (11.48) using the current parameter values at each iteration. Quantiles of the simulated distributions may be used to obtain posterior confidence intervals for the corresponding quantities.

The Metropolis-Hastings update described above changes all three parameters on each iteration, or none of them. Alternatively we may attempt to update one parameter, chosen at random. The resulting chain is also ergodic, but it does not improve on the second approach described above. ■