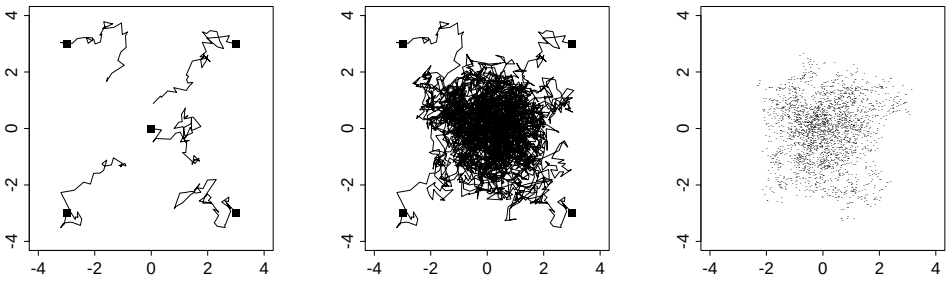

Figure 18.9 *Five independent sequences of a Metropolis algorithm, with overdispersed starting points indicated by solid squares. (a) After 50 iterations, the sequences are still far from convergence. (b) After 1000 iterations, with the sequences nearer to convergence. (c) The iterates from the second halves of the sequences, jittered so that steps in which the random walks stood still are not hidden.*

R output

```
          mean   sd  2.5%   25%   50%   75% 97.5% Rhat n.eff
a       155.3  0.6 154.1 154.9 155.3 155.7 156.3  1.0    53
b         4.8  0.2   4.5   4.7   4.8   4.9   5.1  1.0   150
sigma    26.5  0.5  25.7  26.2  26.5  26.9  27.6  1.0    69
z[3]    216.9 13.0 200.7 206.1 215.4 223.7 247.7  1.0    87
z[6]    210.9  9.3 200.3 203.5 208.9 214.9 234.1  1.0   150
z[11]   208.5  7.1 200.3 203.4 206.6 211.7 226.2  1.0   150
. . .
```

which is essentially identical to the results from the Bugs run (as it should be, given that we are fitting the same model).

Section 25.6 briefly describes a more realistic and complicated example of censoring that arises in a study of reversals of the death penalty, in which cases are censored that are still under consideration by appellate courts.

18.6 Metropolis algorithm for more general Bayesian computation

Moving to even more general models, the Gibbs sampler is a special case of a larger class of *Markov chain simulation algorithms* that can be used to iteratively estimate parameters in any statistical model. Markov chain simulation in general (and the Gibbs sampler in particular) can be thought of as iterative imputation of unknown parameters, or as a random walk through parameter space.

The Gibbs sampler updates the parameters one at a time (or in batches) using their conditional distributions. It can also be efficient to use the *Metropolis algorithm*, which takes a random walk through the space of parameters.

The Gibbs sampler and Metropolis algorithms are special cases of Markov chain simulation (also called *Markov chain Monte Carlo*, or MCMC), a general method based on drawing values of θ from approximate distributions and then correcting those draws to better approximate the target posterior distribution, $p(\theta|y)$. The samples are drawn sequentially, with the distribution of the sampled draws depending on the last value drawn; hence, the draws form a Markov chain. (As defined in probability theory, a *Markov chain* is a sequence of random variables $\theta^{(1)}, \theta^{(2)}, \ldots$, for which, for any t, the distribution of $\theta^{(t)}$ given all previous θ's depends only on the most recent value, $\theta^{(t-1)}$.) The key to the method's success, however, is not the Markov property but rather that the approximate distributions are improved at each step in the simulation, in the sense of converging to the target distribution.

Figure 18.9 illustrates a simple example of a Markov chain simulation—in this