Rejection Sampling in Bayesian Context

Problem

Suppose we want to sample from the posterior distribution of a parameter θ given data $x_{1:n}$:

$$\pi(\theta \mid x_{1:n}) \propto f(x_{1:n} \mid \theta)\, \pi(\theta),$$

where $f(x_{1:n} \mid \theta)$ is the likelihood and $\pi(\theta)$ is the prior.

1. Show how rejection sampling can be used to generate samples from this posterior distribution if we use the prior $\pi(\theta)$ as the proposal distribution.

2. Derive the acceptance probability in terms of the likelihood function.

3. Explain in words why this algorithm produces exact samples from the posterior.

Solution

1. **Proposal distribution:**
   We take the proposal to be the prior itself, $q(\theta) = \pi(\theta)$.
   Sampling from the prior is straightforward and ensures that the support of the posterior is covered.

2. **Bounding constant:**
   Define the unnormalized posterior
   $$h(\theta) = f(x_{1:n} \mid \theta)\, \pi(\theta).$$
   Since $q(\theta) = \pi(\theta)$, the rejection sampling inequality
   $$h(\theta) \leq c\, q(\theta)$$
   simplifies to
   $$f(x_{1:n} \mid \theta) \leq c \quad \text{for all } \theta.$$
   Thus, c can be chosen as the maximum of the likelihood function.

3. **Acceptance probability:**
   For a candidate $\theta^* \sim \pi(\theta)$, the acceptance probability is
   $$\alpha(\theta^*) = \frac{f(x_{1:n} \mid \theta^*)}{c}.$$

4. **Algorithm:**

- Generate $\theta^* \sim \pi(\theta)$.

- Generate $U \sim \text{Uniform}(0, 1)$.

- If $U \leq f(x_{1:n} \mid \theta^*)/c$, accept θ^*; otherwise reject and repeat.

5. **Why this works:**
   The posterior distribution is proportional to $\pi(\theta)\, f(x_{1:n} \mid \theta)$.
   By proposing from the prior, we start with $\pi(\theta)$.
   The acceptance step reweights these draws according to the likelihood: parameter values that explain the data well (high likelihood) are more likely to be accepted.
   Because c ensures that acceptance probabilities lie in $[0, 1]$, the distribution of accepted samples is exactly the posterior $\pi(\theta \mid x_{1:n})$.

Consider the Bayesian model

$$x_i \mid \theta \overset{\text{iid}}{\sim} N(\theta, \sigma^2), \quad i = 1, \dots, n,$$

with a **Laplace prior** on θ:

$$\pi(\theta) = \frac{\lambda}{2} \exp(-\lambda|\theta|).$$

The posterior distribution is

$$\pi(\theta \mid x_{1:n}) \propto \left[\prod_{i=1}^{n} \exp\left(-\frac{(x_i - \theta)^2}{2\sigma^2}\right)\right] \exp(-\lambda|\theta|),$$

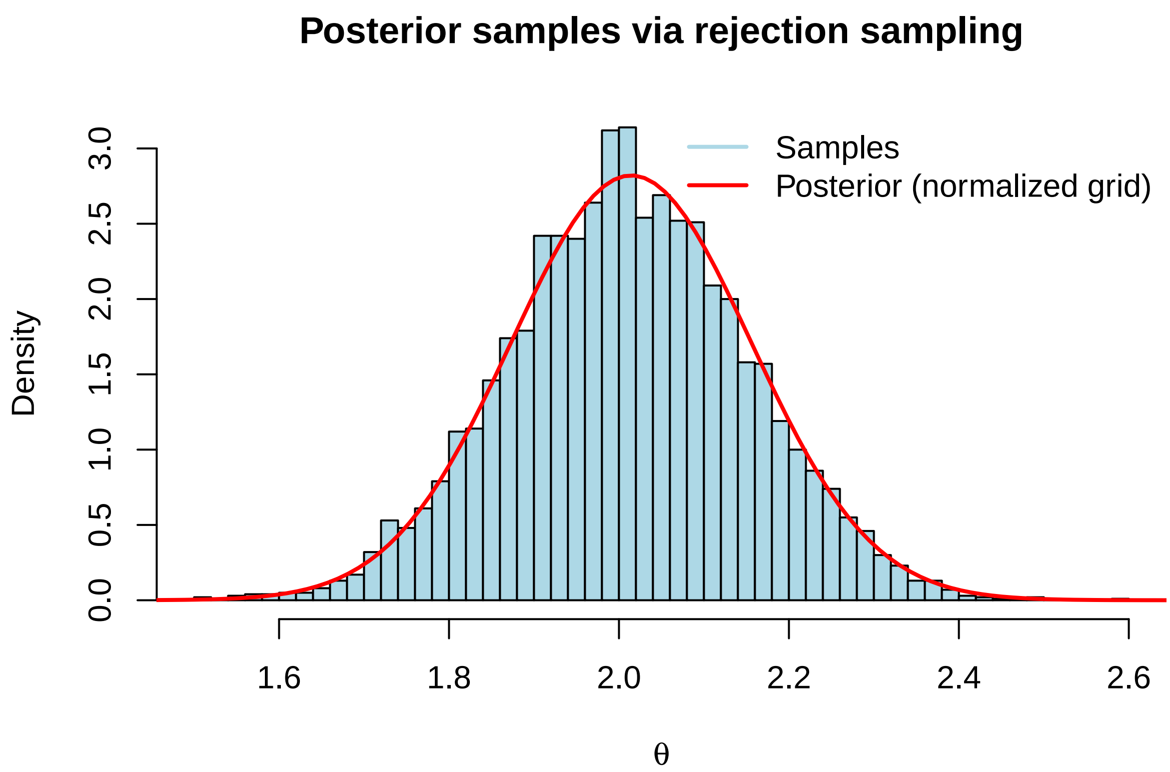
which does not have a closed-form conjugate form.

```r
set.seed(123)

# ------------------------
# Simulated data
# ------------------------
n <- 50
theta_true <- 2
sigma <- 1
x <- rnorm(n, mean = theta_true, sd = sigma)

# ------------------------
# Prior: Laplace(0, b = 1/lambda)
# density: (lambda/2) * exp(-lambda * |theta|)
# ------------------------
lambda <- 1

dprior <- function(theta) {
  (lambda / 2) * exp(-lambda * abs(theta))
}

# ------------------------
# Likelihood: Normal with known sigma
# ------------------------
dlikelihood <- function(theta, x, sigma) {
  prod(dnorm(x, mean = theta, sd = sigma))
}

# ------------------------
# Unnormalized posterior h(theta)
# ------------------------
h <- function(theta, x, sigma) {
  dlikelihood(theta, x, sigma) * dprior(theta)
}

# ------------------------
# Proposal distribution q(theta): Normal centered at MLE
# ------------------------
```

```r
theta_mle <- mean(x)
tau <- 2  # proposal variance (tune for efficiency)

rproposal <- function(N) rnorm(N, mean = theta_mle, sd = tau)
dproposal <- function(theta) dnorm(theta, mean = theta_mle, sd = tau)

# ------------------------
# Rejection sampler
# ------------------------
rejection_sampler <- function(N, x, sigma) {
  samples <- numeric(0)
  attempts <- 0

  # Find bounding constant c by overestimating
  grid <- seq(theta_mle - 5, theta_mle + 5, length.out = 1000)
  c <- max(sapply(grid, function(t) h(t, x, sigma) / dproposal(t)))

  while(length(samples) < N) {
    theta_star <- rproposal(1)
    u <- runif(1)
    accept_prob <- h(theta_star, x, sigma) / (c * dproposal(theta_star))
    if (u < accept_prob) {
      samples <- c(samples, theta_star)
    }
    attempts <- attempts + 1
  }

  list(samples = samples, attempts = attempts, efficiency = N / attempts)
}

# ------------------------
# Run the sampler
# ------------------------
res <- rejection_sampler(5000, x, sigma)

cat("Efficiency (acceptance rate):", res$efficiency, "\n")
```

```
## Efficiency (acceptance rate): 0.07088679
```

```r
# ------------------------
# Plot results
# ------------------------
hist(res$samples, breaks = 50, freq = FALSE, col = "lightblue",
     main = "Posterior samples via rejection sampling",
     xlab = expression(theta))

# Compare with grid approximation of posterior (for reference only)
posterior_grid <- function(theta) h(theta, x, sigma)
theta_vals <- seq(-1, 5, length.out = 500)
posterior_vals <- sapply(theta_vals, posterior_grid)
posterior_vals <- posterior_vals / sum(posterior_vals) / (theta_vals[2] - theta_vals[1])
lines(theta_vals, posterior_vals, col = "red", lwd = 2)
legend("topright", legend = c("Samples", "Posterior (normalized grid)"),
       col = c("lightblue", "red"), lwd = 2, bty = "n")
```


Posterior samples via rejection sampling
Samples
Posterior (normalized grid)
Density
0.0
0.5
1.0
1.5
2.0
2.5
3.0
1.6
1.8
2.0
2.2
2.4
2.6
θ