Sampling exercises

Thomas Mailund September 20, 2007

There's a lot of alchemy in Monte Carlo methods, so even if you understand the theory behind it, you might not be able to develop efficient methods. That comes down to experience, more than anything else. It is therefore important that you do as many exercises as you can, and that you experiment with it.

Below there is just a few exercises to get you started. The mandatory project, and projects in your later life as a bioinformatician, will provide plenty of opportunities for further honing your skills.

Rejection sampling

We are going to sample from the *beta distribution*, $\mathrm{Be}(\alpha, \beta)$, which is a distribution for $x \in [0, 1]$ with density given by

$$p(x \mid \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}$$

where $B(\alpha, \beta)$ is the *beta function*, but in this case just a normalising constant, so we can use just $p(x \mid \alpha, \beta) \propto \tilde{p}(x \mid \alpha, \beta) = x^{\alpha-1}(1-x)^{\beta-1}$.

The *mode* of the distribution—the value where the density takes the highest value—is $x = (\alpha-1)/(\alpha+\beta-2)$, when $\alpha > 1$ and $\beta > 1$, not defined for $\alpha = \beta = 1$ (where the distribution is uniform), or 0 or 1 (or both) otherwise. To keep the exercise simple, we only consider the case $\alpha > 1$ and $\beta > 1$.

The mean of the distribution is $\alpha/(\alpha+\beta)$ and the variance is $\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.

Proposal distribution: Since $x \in [0, 1]$ we can use a uniform distribution on this interval as our proposal distribution. How do you construct the envelope?

Sampling: Write a function that samples a single x from $\mathrm{Be}(\alpha, \beta)$. Then use the function `replicate()` to write a function that samples n values. Try plotting histograms of sampled values against the "true" distribution, which you can get in R using the function `dbeta()`.

Efficiency: Sample under various choices of α and β and count the average number of rejections per acceptance. Can you explain what you see?

Expectations: Use your sampler to calculate the expected value (the mean value) of the distribution. Compare with estimating the mean using `rbeta()`. Should you expect a difference?

A different proposal distribution: When α and β are similar, the density of the beta distribution is bell-shaped, so it seems reasonable that we can use a normal distribution as a proposal distribution (see Fig. 1 and Fig. 2). Which problems do you expect to have to solve to do this?

Importance sampling

We are not going to use the normal distribution to do rejection sampling, but instead do importance sampling. That way, at least, we do not need to worry about enveloping the beta distribution.

Sample: Program an importance sampler for calculating the mean of the beta distribution. Compare your estimates of the mean using the importance sampler with estimates using the true distribution. How does the variance of the two estimates behave?

MCMC

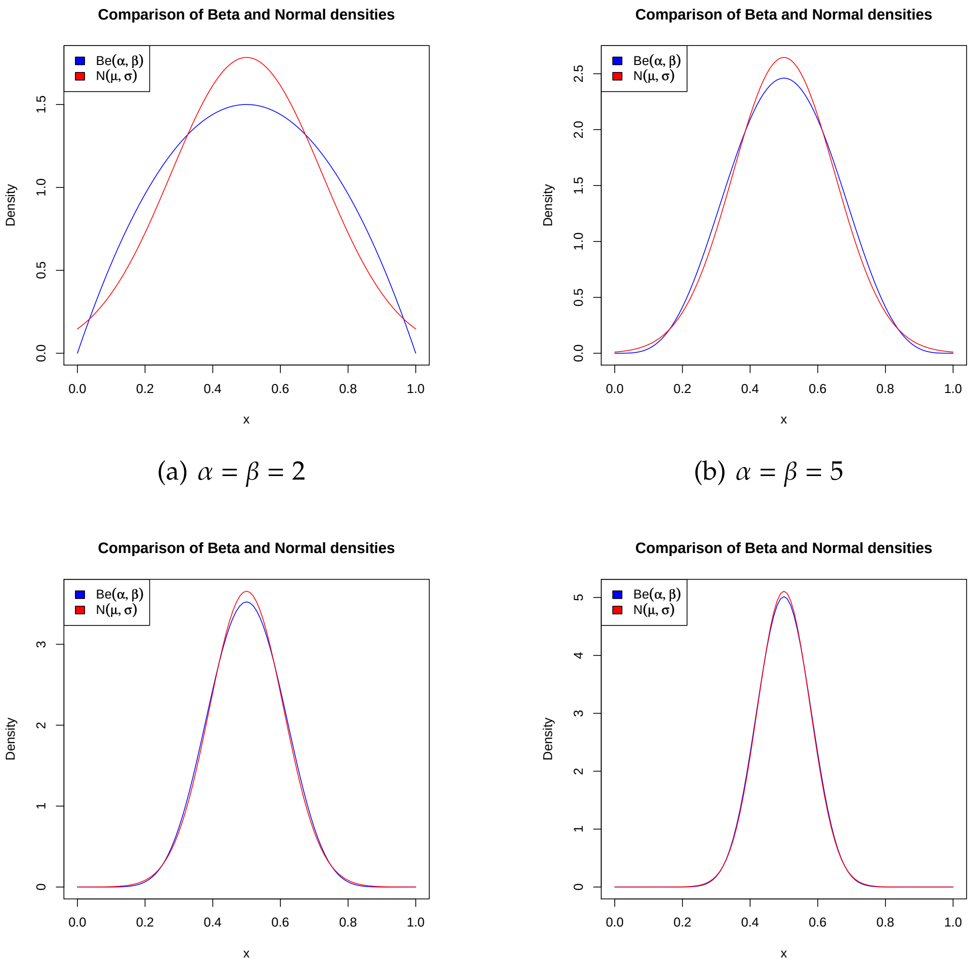
If you use a symmetric proposal distribution in the Hastings-Metropolis algorithm ($q(x \mid y) = q(y \mid x)$), your acceptance probability of a move from x to y becomes simply $\min(1, p(y)/p(x))$ (do you see why?)

One way to get a symmetric proposal distribution on an interval $[A, B]$ is to use a (conditional) symmetric distribution over the entire real line, e.g. $y \mid x \sim N(x, \sigma)$ but then "wrap around" (LUDO style) at the borders:

```
y <- rnorm(1, mean=x, sd=sigma)
while (y < A || B < y) {
  if (y < A) y = 2*A - y
  if (B < y) y = 2*B - y
}
```

(the while-loop here is typically not executed more than once; we do not expect to take such large steps that we "ping pong" back and forth between the borders, but of course it is possible).

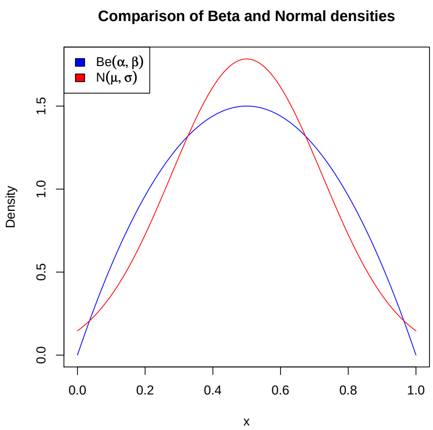
Hastings-Metropolis: Write a Hastings-Metropolis MCMC to sample from the beta distribution. As earlier, compare the distribution you get out of it with the distribution you compute using `dbeta()` and compare the expectation you calculate from it with that estimated using `rbeta()`.



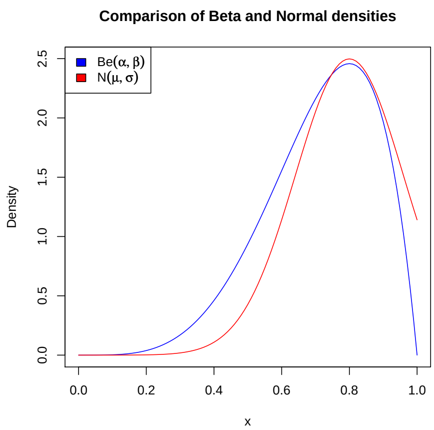
(a) $\alpha = \beta = 2$ (b) $\alpha = \beta = 5$

(c) $\alpha = \beta = 10$ (d) $\alpha = \beta = 20$

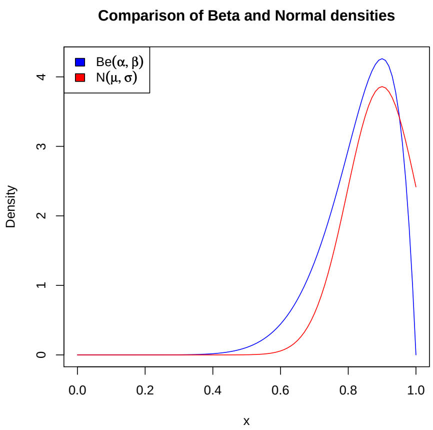
Fig. 1: Symmetric beta and normal distributions. In all cases, the Normal distribution has the beta mode as its mean (why do you think I use the mode and not the mean?) and the beta distribution standard deviation as its standard deviation.



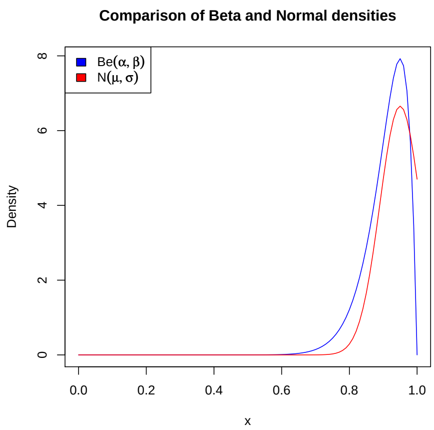
(a) $\alpha = 2, \beta = 2$



(b) $\alpha = 5, \beta = 2$



(c) $\alpha = 10, \beta = 2$



(d) $\alpha = 20, \beta = 2$

Fig. 2: Asymmetric beta distributions and (of course symmetric) normal distributions.