732A90: Computational Statistics

Computer lab4 - Group11

Sofie Jörgensen, Oriol Garrobé Guilera, David Hrabovszki

17 February 2020

Question 1: Computations with Metropolis-Hastings

In this question we define the density function $f(x)$ as:

$$f(x) \propto x^5 e^{-x}, x > 0$$

This density function is known up to some constant of proportionality, which is 120.

1.

In the first step we use the Metropolis-Hastings algorithm to generate samples from the posterior distribution $f(x)$ by using the log-normal $LN(X_t, 1)$ as the proposal distribution. We take 1 as a starting point, because we expect this value to be within the range of possible values of the distribution defined above.

The time series plot of the obtained chain of samples can be observed in Figure 1. The plot does not show that the chain converges, because it often takes the same value for many iterations and then jumps to a very different one. This implies that the sample is not close to the posterior. There does not seem to be a burn-in period, because samples from many iterations are close to the starting point (which was 1 in this case).

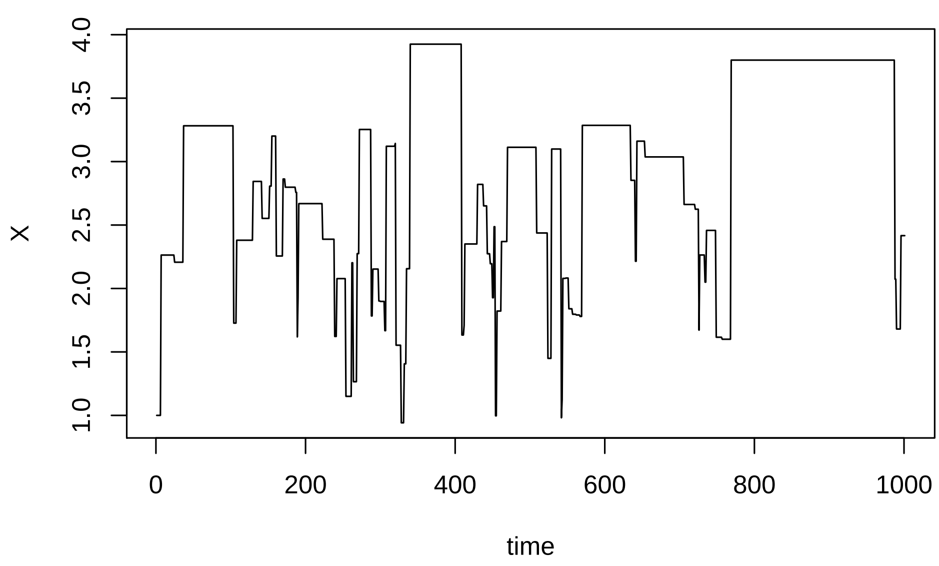

Figure 1: Metropolis-Hastings sampler with log-normal proposal

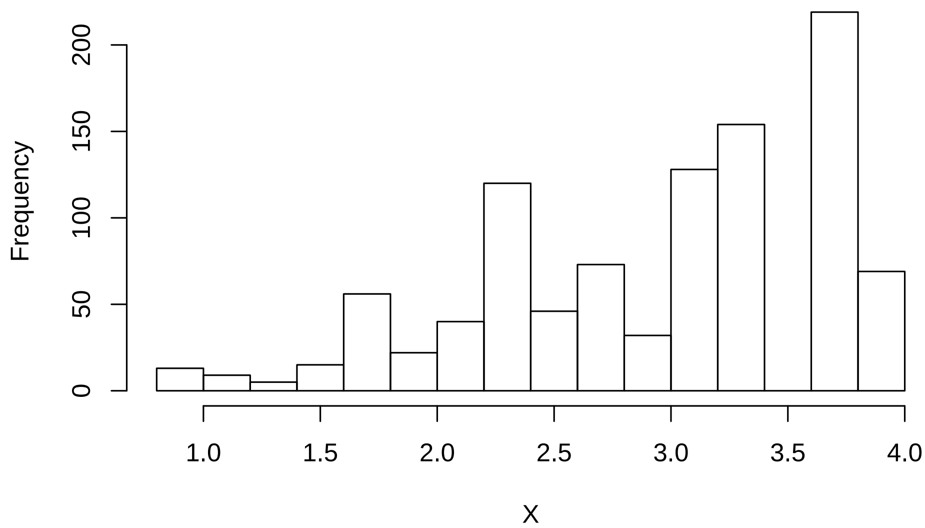

Figure 2: Histogram of generated samples with log-normal proposal (MH sampler)

2.

Now we perform Step 1 using the chi-square distribution $\chi^2(\lfloor X_t + 1 \rfloor)$ as the proposal distribution, where $\lfloor x \rfloor$ is the `floor()` function.

The time series plot is shown in Figure 3, and it can be seen that the chain converges much more than in the previous case. Again there does not seem to be a burn-in period, the starting value of 1 seems to be a probable value of the distribution, so we do not discard any samples from the beginning of the chain.

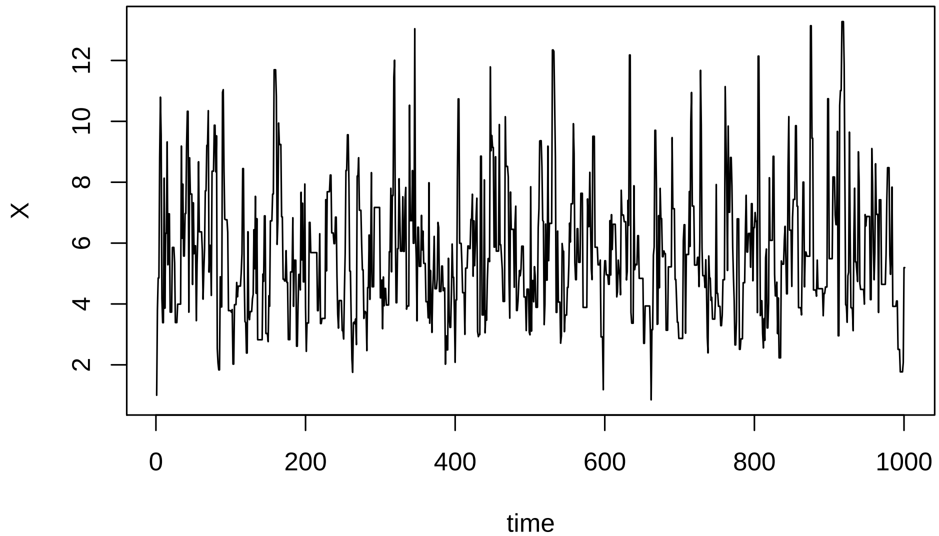

Figure 3: Metropolis-Hastings sampler with chi-square proposal

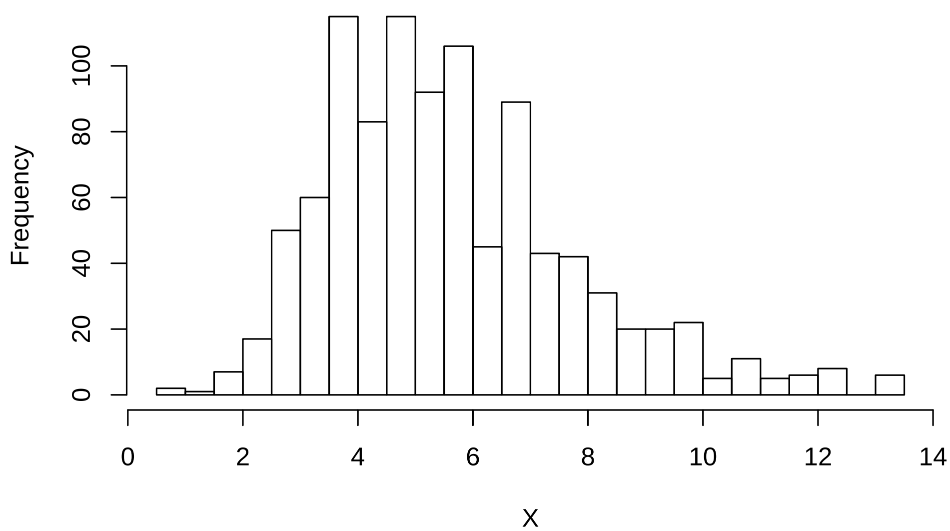

Figure 4: Histogram of generated samples with Chi-square proposal (MH sampler)

3.

We recognise that the probability distribution function $f(x)$ belongs to a gamma distribution, therefore we know that $f(x)$ must have a positive skew. This attribute is also true for the distribution of samples from

Step 2 (Figure 4), but untrue for the distribution of samples from Step 1 (Figure 2). Also, the chain with the log-normal proposal does not converge, but the chain with the Chi-square proposal does, so we conclude that the samples generated from Step 2 are a better approximation of samples from the posterior distribution.

4.

Now we generate 10 MCMC sequences using the sample generators from Step 2, and analyze their convergence using the Gelman-Rubin method. The starting points for the 10 sequences are 1,2,. . . ,10.

The output of the Gelman-Rubin's convergence diagnostic for 1000 iterations:

```
## Potential scale reduction factors:
##
##      Point est. Upper C.I.
## [1,]       1.01       1.02
```

The potential scale reduction factor is based on the comparison of within-chain and between-chain variances. For 1000 iterations, its point estimate is around 1 and its upper confidence limit is also very close to 1, so according to this diagnostics, approximate convergence has been achieved at 1000 iterations. Even for only 100 iterations, both the point estimate and the upper limit are under 1.1.

5.

In this step we estimate

$$\int_0^\infty x f(x) dx$$

using the samples generated from Steps 1 and 2.

Based on the slides from Lecture 4 we can use MCMC samples from $p(\theta|D)$ to obtain a point estimator by calculating the sample average:

$$\theta^* = \int \theta p(\theta|D) \approx \frac{1}{n}\sum_{i=1}^{n} \theta_i$$

The sample average of the chain generated in Step 1 is **2.95** and in Step 2 is **5.63**.

If we do not set the seed before the sample generation, then we get highly variable averages from Step 1, but from Step 2 it is almost always between 5.5 and 6.5.

6.

According to Definition 4.5 in Mathematical Statistics with Applications by Wackerly et al., the integral in Step 5 is equal to the expected value of random variable X, where X can take values between 0 and ∞.

The generated distribution with probability density function $f(x)$ is a gamma distribution, so we can express $f(x)$ in its general form as

$$f(x) = \left[\frac{1}{\Gamma(\alpha)\beta^\alpha}\right] x^{\alpha-1} e^{-x/\beta}$$

, where $0 < x < \infty$

Thus $E(X) = \alpha\beta$. In our case $\alpha = 6, \beta = 1$, so $E(X) = 6$.

This is the actual value of the integral in Step 5, so we conclude that our estimation using the samples generated in Step 2 with a Chi-square proposal distribution seems accurate, whereas the estimation from the samples of Step 1 with the log-normal proposal does not seem accurate at all.