

732A90: Computational Statistics

Computer lab4 - Group11

Sofie Jörgensen, Oriol Garrobé Guilera, David Hrabovszki

15 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, especially after the first 4000 iterations. 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).

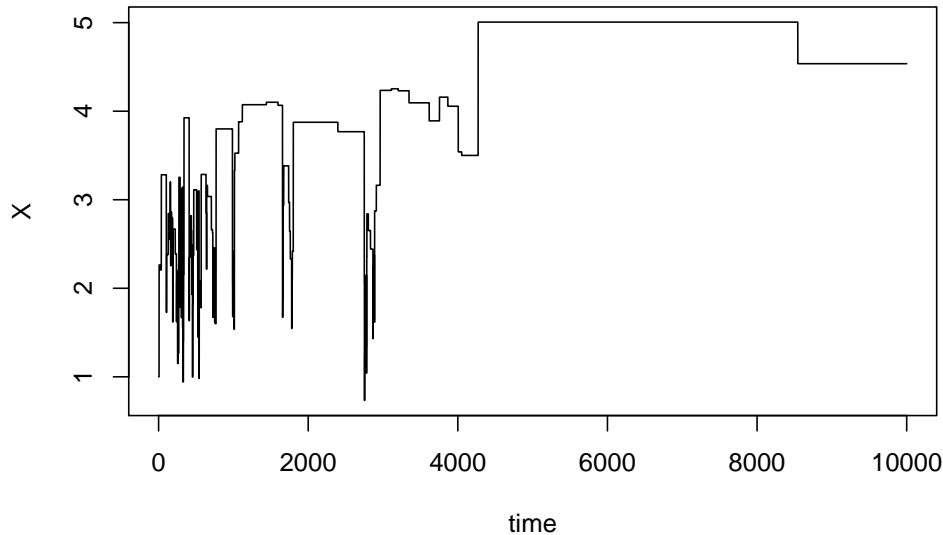


Figure 1: Metropolis-Hastings sampler with log-normal proposal

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 2, and it can be seen that already in 1000 iterations, 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 posterior distribution, so we do not discard any samples from the beginning of the chain.

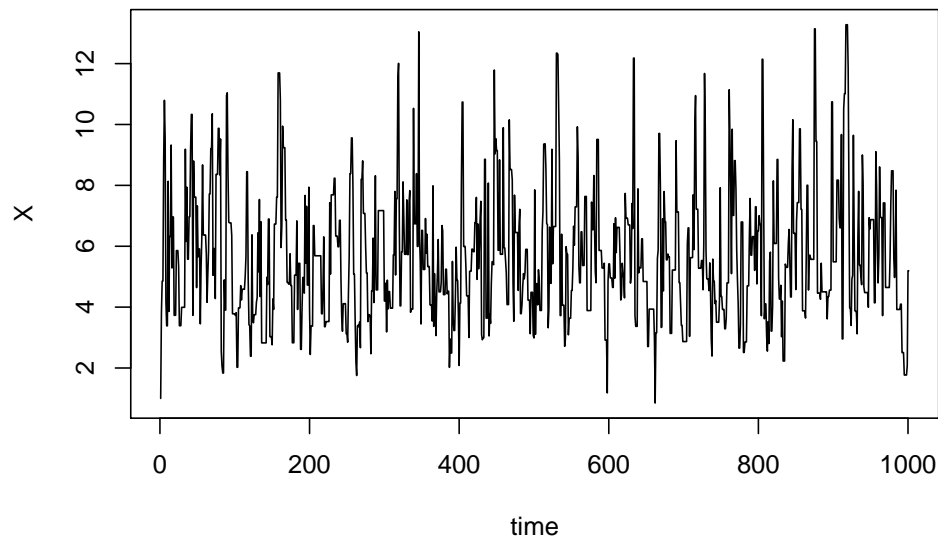


Figure 2: Metropolis-Hastings sampler with chi-square proposal

3.

COMPARE THE RESULTS FROM STEPS 2 AND 3 AND MAKE CONCLUSIONS!!!!!!

4.

Now we generate 10 MCMC sequences using the sample generators from Steps 1 and 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.

6.