Coursework Project Part 1

ST2195-Programming for Data Science

Richard Hardy Mathias

Student id: 230655968

Table of contents

1. Part 1a)

For part 1(a), we'll implement the Metropolis-Hastings algorithm to generate samples from the given probability density function $f(x)=\frac{1}{2}\exp(-|x|)$.

Steps that I will be taking to implement:

N=10000: Total number of samples.

s=1: Standard deviation for the proposal distribution.

We will define the target function f(x) and its logarithm for numerical stability and run the Metropolis-Hastings algorithm with the given proposal distribution. Then we will create a histogram and kernel density estimate (KDE) of the generated samples. Finally, we will overlay the graph of f(x) on the histogram and KDE plot.

Calculate and print the sample mean and standard deviation of the generated samples. I imported libraries such as NumPy, matplotlib and seaborn to help me with calculations and plotting of the histogram.

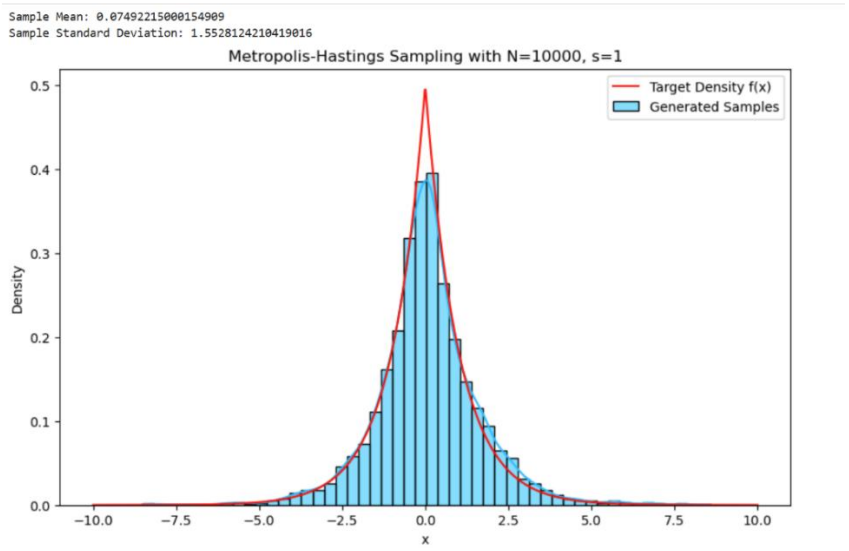

Figure 1: The result of the code(python) part a

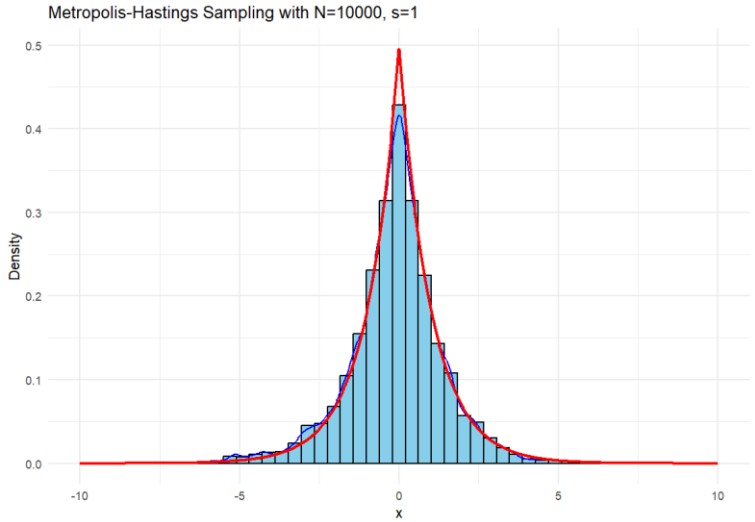

Figure 2: The result of the code (R) part a

2. Part 1b)

2.1. Computing the Convergence Diagnostic $\widehat{R}$

For part (b), we'll calculate the $\hat{R}$ diagnostic to verify convergence.

Steps that we will take includes generating multiple chains, J=4 chains, each with N=2000 samples and different initial values. Next, calculate within-chain means M_j and variances V_j. Then we will calculate the overall mean M and between-chain variance B. We will then compute $\hat{R}$ using the provided formula and plot $\hat{R}$ over a range of s values. For the coding, I used R and used libraries such as ggplot2 for the histogram

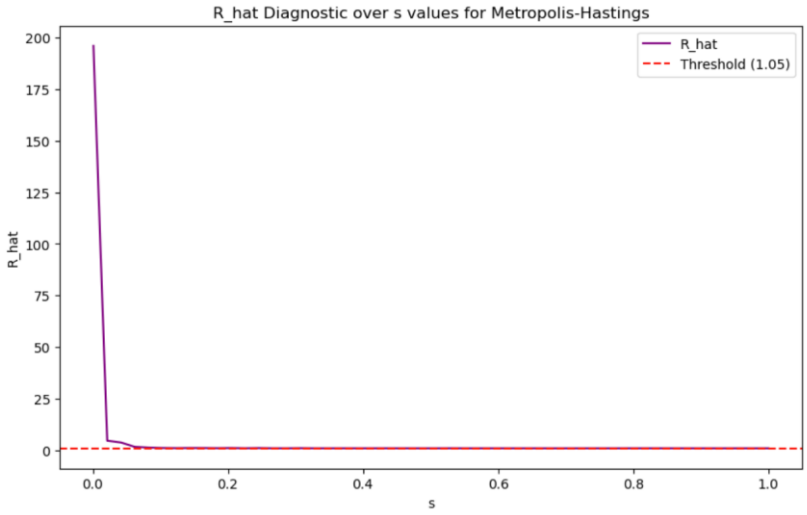

Figure 3: The result of the code(python) part b

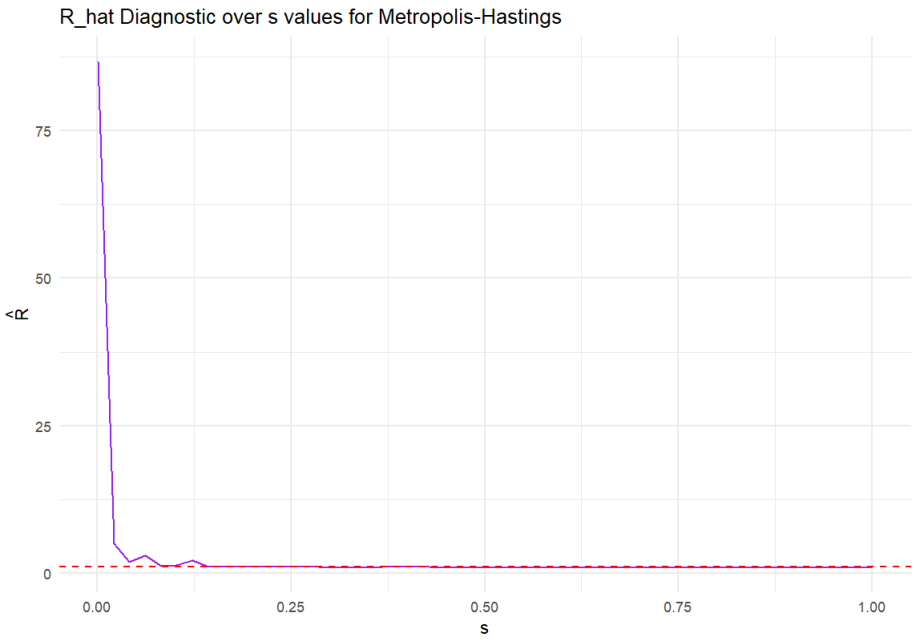

Figure 4: The result of the code(R) part b