University of London

ST2195 Programming for Data Science

Coursework (Part 1)

Prepared by:

Elizabeth Leonny Efendi

(220657172)

Table of Contents

Markov Chain Monte Carlo Simulation

This section presents the application of the Metropolis-Hastings algorithm, specifically the random walk Metropolis, to simulate random numbers for the distribution with the probability density function given by $f(x) = \frac{1}{2} e^{-|x|}$. The goal is to generate samples using the algorithm and assess the quality of the estimates through visualizations and convergence diagnostics.

1. Random Walk Metropolis Algorithm

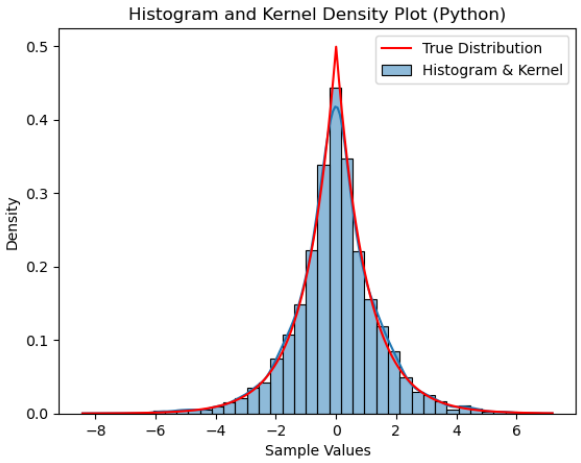

Figure 1: Histogram and Kernel Density Plot (Python)

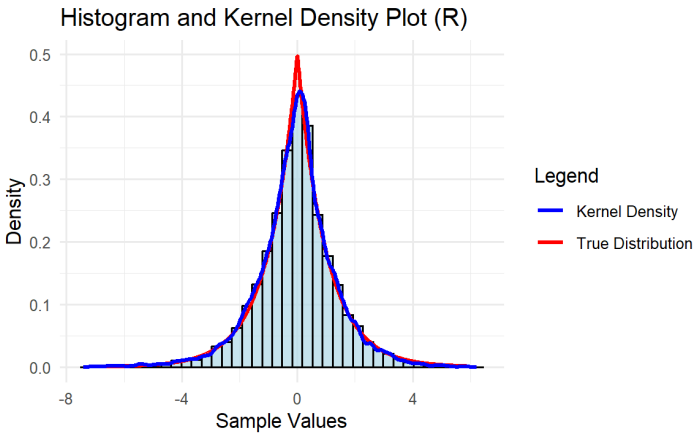

Figure 2: Histogram and Kernel Density Plot (R)

The histogram and kernel density plot demonstrate convergence, aligning well with the true distribution, indicating the effectiveness of the random walk Metropolis algorithm in approximating the target distribution. The calculated sample mean and sample standard deviation of the generated samples are 0.03 and 1.28 respectively for python, and 0.11 and 1.31 respectively for R. These information provides insights into the central tendency and spread of the generated samples.

2. Convergence Diagnostics: R-hat Value

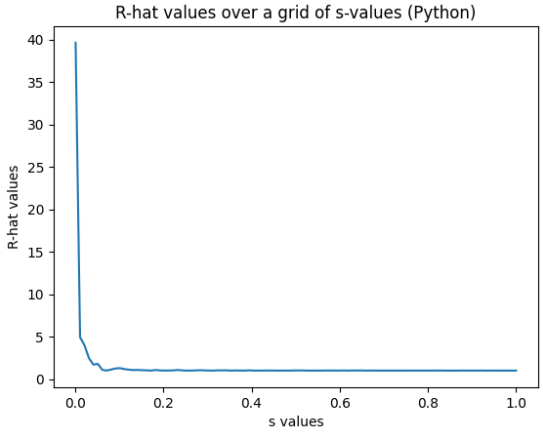

Figure 3: R-hat values over a grid of s-values (Python)

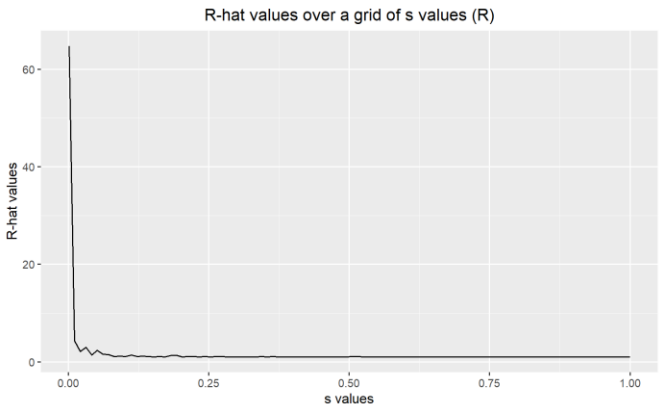

Figure 4: R-hat values over a grid of s-values (R)

To assess convergence, we computed the R-hat value using multiple chains (J = 4). The R-hat values exhibit a sharp downward trend as s increases, signifying improved convergence. For stable convergence, R-hat values should approach 1, indicating that the multiple chains are producing consistent and accurate samples from the target distributions. This downward trend in R-hat values implies the algorithm's ability to converge effectively and generate valid samples for statistical analysis.