Estimation of an Integral Using a Gaussian Copula

group 6

2025-04-04

Introduction

In this report, we statistically estimate the following integral using a copula-based Monte Carlo method, as required in Question 1 of the assignment:

$$\lim_{n\to\infty} \int_0^1 \dots \int_0^1 \frac{x_1^{101} + x_2^{101} + \dots + x_n^{101}}{x_1 + x_2 + \dots + x_n} dx_1 \dots dx_n.$$

We model the uniform marginals using a Gaussian copula to introduce dependence, and estimate:

$$I_n = E\left[\frac{\sum_{i=1}^n U_i^{101}}{\sum_{i=1}^n U_i}\right],$$

where $U_i \sim \text{Uniform}(0,1)$ and possibly correlated.

Methodology

We define a function to compute the integral estimate for a given number of variables (n) and correlation (ρ).

```
compute_integral_gaussian <- function(n, rho, samples = 200) {
  if (n == 1) {
    u <- runif(samples)
    return(mean(u^101 / u))
  }
```

```r
  gaussian_cop <- normalCopula(param = rho, dim = n)
  u_samples <- rCopula(samples, gaussian_cop)

  sum_x <- rowSums(u_samples)
  sum_x_power <- rowSums(u_samples^101)

  estimated_value <- mean(sum_x_power / sum_x)

  return(estimated_value)
}
```

Estimation for Different Correlations

We estimate the integral for different values of correlation ($\rho = 0.001, 0.0001, 0$) and varying n values.

```r
n_vals <- 1:200
correlation_values <- c(0.001, 0.0001, 0)

results_list <- lapply(correlation_values, function(rho) {
  integral_results <- sapply(n_vals, function(n) compute_integral_gaussian(n, rho, samples =
  data.frame(n = n_vals, integral = integral_results, correlation = factor(rho))
})

results_df <- do.call(rbind, results_list)
```

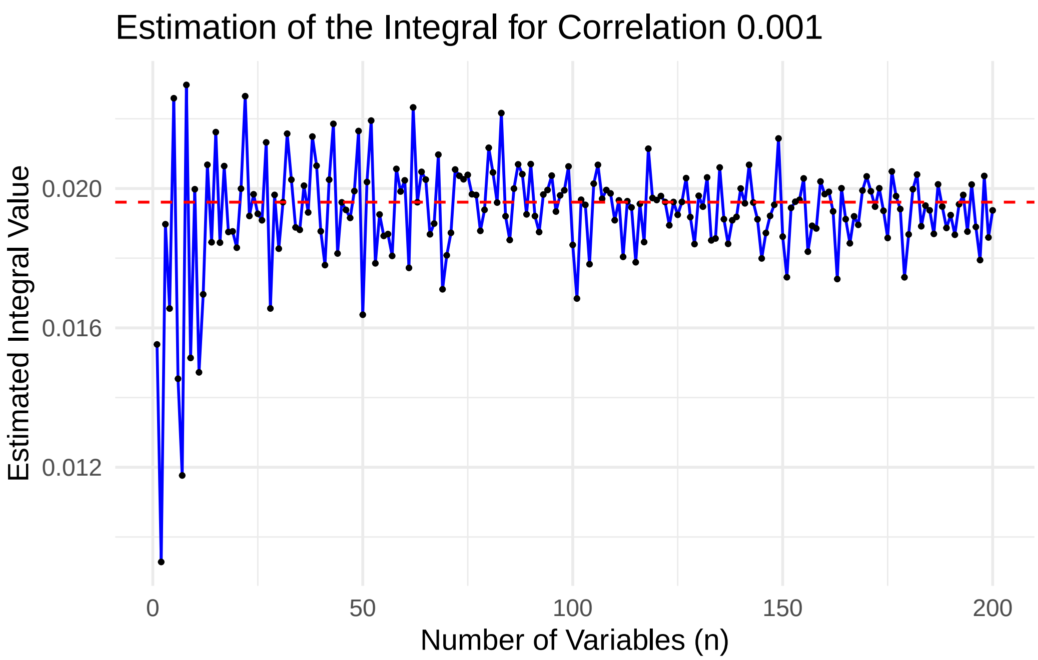
Visualization

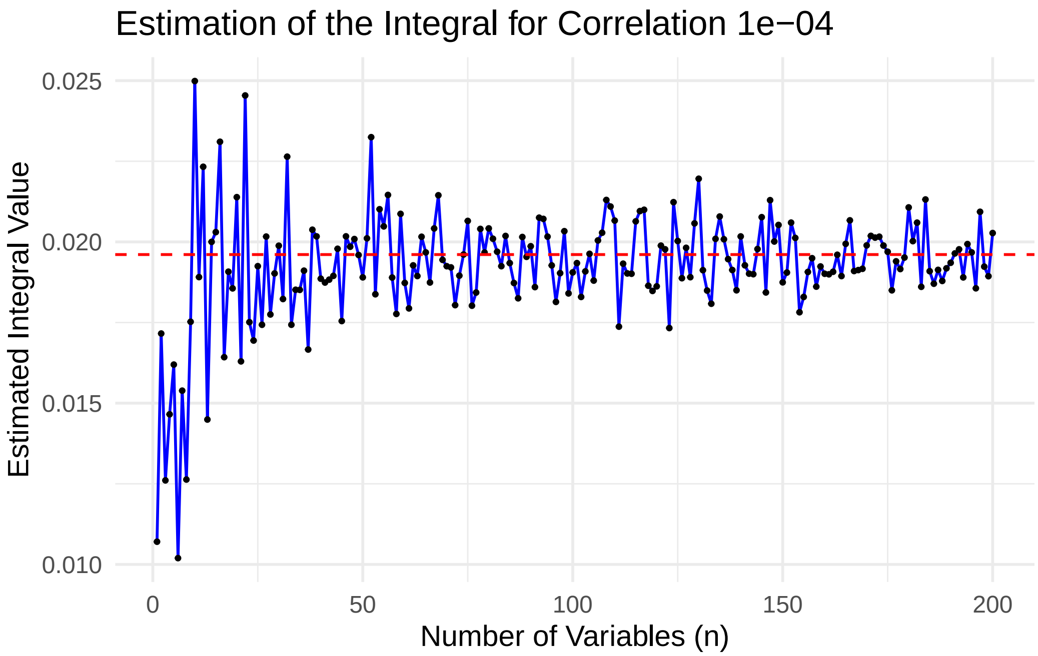
We create separate plots for each correlation value and compare with the theoretical value.

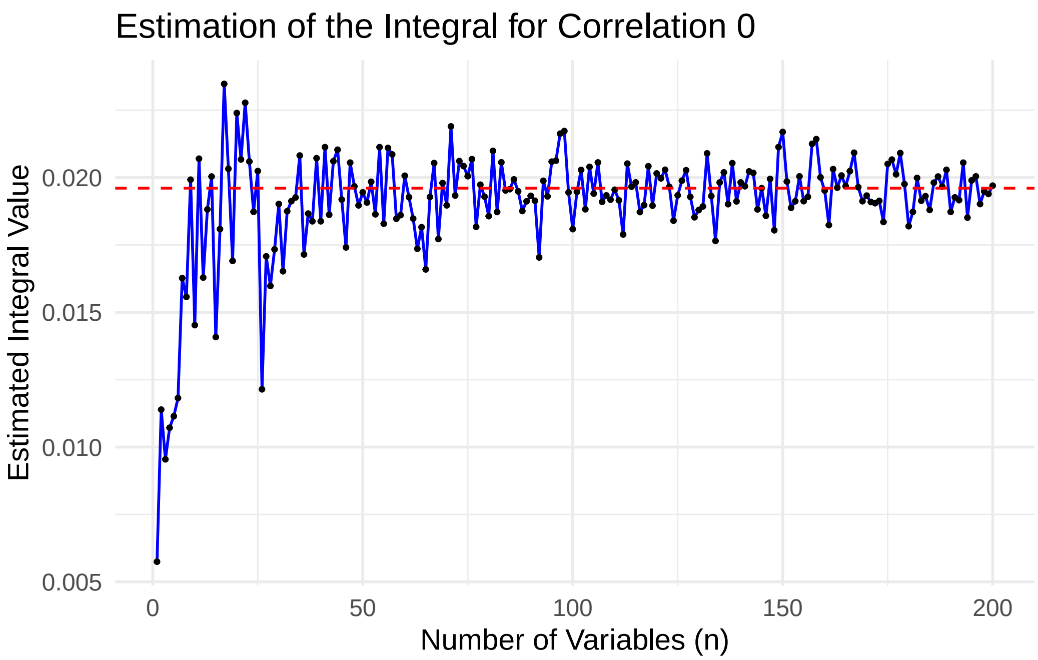
```r
theoretical_value <- 1 / 51  # From theoretical derivation

for (rho in correlation_values) {
  plot_data <- subset(results_df, correlation == rho)
  p <- ggplot(plot_data, aes(x = n, y = integral)) +
    geom_line(color = "blue") +
    geom_point(size = 0.5) +
    geom_hline(yintercept = theoretical_value, color = "red", linetype = "dashed") +
    labs(title = paste("Estimation of the Integral for Correlation", rho),
         x = "Number of Variables (n)",
```

```
          y = "Estimated Integral Value") +
    theme_minimal()
  print(p)
}
```




Estimation of the Integral for Correlation 1e−04
Estimated Integral Value
0.025
0.020
0.015
0.010
0
50
100
150
200
Number of Variables (n)


Estimation of the Integral for Correlation 0
Estimated Integral Value
0.020
0.015
0.010
0.005
0
50
100
150
200
Number of Variables (n)

Theoretical Estimation

We compare the numerical results with the theoretical approximation under the assumption of independent variables:

$$E[I_n] \approx \frac{nE[U^{101}]}{nE[U]} = \frac{\frac{1}{102}}{\frac{1}{2}} = \frac{1}{51} \approx 0.0196.$$

This serves as a benchmark to assess the convergence of the Monte Carlo estimates.

Conclusion

- The Monte Carlo estimates using the Gaussian copula converge to the theoretical value $\frac{1}{51}$ as n increases.
- The inclusion of a small correlation (ρ) affects the rate and variability of convergence.
- This method demonstrates the flexibility of copula-based modeling in high-dimensional integration problems.