Monte Carlo Simulation in R

Mohammad Saqib Ansari

January 6, 2024

Introduction to Monte Carlo Simulation

Monte Carlo Simulation is a statistical technique used to understand the impact of uncertainty and randomness in mathematical, financial, and scientific systems. It employs random sampling methods to obtain numerical results and is particularly useful when dealing with complex systems or problems for which analytical solutions are difficult to derive.

Understanding the Code

The given R code performs a Monte Carlo simulation to estimate the significance level (alpha) using a t-statistic. Here's a detailed explanation of the code:

```r
# Function to calculate t-statistics
tstatistics = function(x, y) {
  m = length(x)
  n = length(y)

  sp = sqrt(((m - 1) * sd(x)^2 + (n - 1) * sd(y)^2) / (m + n - 2))
  t.stat = (mean(x) - mean(y)) / (sp * sqrt(1 / m + 1 / n))
  return(t.stat)
}

# Set the significance level (alpha), sample sizes (m and n), and number of simulations (N)
alpha = 0.1
m = 10
n = 10
N = 1e5

# Initialize a counter for rejected hypothesis tests
n.reject = 0

# Monte Carlo Simulation
for (i in 1:N) {
  # Generate random samples from normal distributions
  x = rnorm(m, mean = 0, sd = 1)
  y = rnorm(n, mean = 0, sd = 1)

  # Calculate the t-statistic for the generated samples
  t.stat = tstatistics(x, y)
```

```
  # Check if the absolute t-statistic exceeds the critical value
  if (abs(t.stat) > qt(1 - alpha / 2, n + m - 2))
    n.reject = n.reject + 1
}

# Calculate the estimated significance level
true.sig.level = n.reject / N

# Print the estimated significance level
print(true.sig.level)
```

[1] 0.09958

Explanation of the Monte Carlo Simulation

1. **Function tstatistics:** Calculates the t-statistic for two sample sets x and y.

2. **Setting Parameters:**

   - `alpha`: Significance level.
   - `m` and `n`: Sample sizes.
   - `N`: Number of simulations.

3. **Monte Carlo Simulation Loop:**

   - Iterates `N` times.
   - Generates random samples `x` and `y` from normal distributions with means 0 and standard deviations 1.
   - Calculates the t-statistic using the `tstatistics` function.
   - Compares the absolute t-statistic to the critical value based on the significance level.
   - Counts the number of rejections.

4. **Estimation of Significance Level:**

   - Calculates the ratio of rejected hypothesis tests to the total simulations to estimate the true significance level.

The final output of `true.sig.level` represents an estimated significance level obtained through Monte Carlo simulation.

This simulation helps understand how often a hypothesis test (comparing means in this case) incorrectly rejects the null hypothesis when it is actually true.