Modul Praktikum 10 Komputasi Statistika

Simulation & Monte Carlo Integration

Tim Dosen Komputasi Statistika – Sains Data - Itera

Pendahuluan

Misalkan akan dihitung integral definit dari fungsi g(x) dalam batas tertentu, $\int_a^b g(x)dx$ (dengan asumsi integral ini ada). Dalam teori probabilitas, diketahui bahwa jika X adalah variabel acak dengan fungsi kepadatan peluang (density function) $f(x)$, maka nilai harapan matematis dari variabel acak $Y = g(X)$ didefinisikan sebagai:

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$$

Catatan:
Jika diketahui sampel acak dari distribusi X, maka estimasi tak biasa (*unbiased estimator*) untuk nilai harapan $E[g(X)]$ tersebut adalah rata-rata sampel.

Algoritma Monte Carlo

Diasumsikan bahwa integral yang ingin dihitung I, dapat ditulis ulang dalam bentuk nilai harapan $E[g(X)]$ dari suatu fungsi $g(X)$ sehubungan dengan variabel acak X pada suatu interval (a, b).

Langkah-langkah algoritma integral dengan metode monte carlo :

1. Inisialisasi Parameter:
   - Tentukan jumlah sampel yang akan digunakan, N (semakin besar N, semakin akurat hasilnya).
   - Inisialisasi variabel total/akumulator: S=0
2. Lakukan Pengambilan Sampel dan Evaluasi (Loop N Kali):
   - Ulangi langkah berikut untuk $i = 1$ hingga N:

   a. Sampel Acak: Hasilkan nilai acak, X_i, dari distribusi yang telat ditentukan pada interval $[a, b]$ (misalnya, distribusi seragam $U(a, b)$).

   b. Evalusai Fungsi: Hitung nilai fungsi $g(x)$ pada titik sampel tersebut, $Y_i = g(X_i)$.

   c. Akumulasi: Tambahkan nilai yang dihasilkan ke total: $S = S + Y_i$
3. Hitung Rata-rata Sampel:

   Hitung rata-rata sampel ($\bar{Y}$) dari hasil evaluasi:

$$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i = \frac{S}{N}$$

4. Estimasi Integral:

   Gunakan rata-rata sampel untuk mengestimasi nilai integral *I*:

$$I = (b - a) \times \bar{Y}$$

5. Hasil

   Nilai *I* adalah estimasi Monte Carlo dari integral $I = \int_a^b g(x)\ dx$

KASUS I

Hitunglah perkiraan nilai integral definit I berikut menggunakan metode **Integrasi Monte Carlo**:

$$I = \int_0^1 e^{x^2} dx$$

(Catatan: Integral ini tidak dapat diselesaikan secara analitik dengan fungsi elementer, menjadikannya kasus ideal untuk metode numerik seperti Monte Carlo).

```r
# 1. Definisikan Fungsi yang akan diintegralkan (g(x))
g <- function(x) {
return(exp(x^2))
}

# 2. Tentukan Parameter Monte Carlo
a <- 0     # Batas bawah integral
b <- 1     # Batas atas integral
N <- 10000 # Jumlah sampel acak (semakin besar semakin baik)

# 3. Langkah Implementasi Monte Carlo

# a. Ambil sampel acak dari distribusi seragam U(a, b)
# Karena a=0 dan b=1, kita bisa menggunakan runif(N)
X <- runif(N, min = a, max = b)

# b. Evaluasi fungsi g(x) pada setiap titik sampel
Y <- g(X)

# c. Hitung rata-rata sampel (Sample Mean)
rata_rata_sampel <- mean(Y)

# d. Estimasi Integral (I_hat = (b - a) * Sample Mean)
I_estimasi_MC <- (b - a) * rata_rata_sampel

# 4. Tampilkan Hasil
print(paste("Jumlah Sampel (N):", N))
print(paste("Estimasi Integral Monte Carlo (I_hat):", I_estimasi_MC))
```

Penjelasan :

Kode R ini mengimplementasikan Integrasi Monte Carlo untuk mengestimasi integral $\int_0^1 e^{x^2} dx$ dengan langkah-langkah yang jelas: pertama, fungsi yang akan diintegralkan, $g(x) = e^{x^2}$, didefinisikan. Selanjutnya, parameter seperti batas integral ($a = 0$, $b = 1$) dan jumlah sampel ($N = 10.000$) ditetapkan. Inti dari implementasi adalah membangkitkan N bilangan acak seragam ($U(0,1)$) ke dalam variabel X menggunakan fungsi runif(). Nilai-nilai ini kemudian dievaluasi menggunakan fungsi $g(x)$ untuk mendapatkan Y. Akhirnya, rata-rata sampel (mean(Y)) dihitung, dan rata-rata ini dikalikan dengan rentang integral ($(b - a)$) untuk menghasilkan estimasi Monte Carlo akhir dari integral tersebut, $\hat{I}$.

KASUS II

Sebuah perusahaan kehutanan perlu mengestimasi volume rata-rata kayu bakar yang dihasilkan dari pohon pinus di suatu area tertentu. Volume batang pohon (dari tanah hingga ketinggian 10 meter) dapat dimodelkan menggunakan fungsi non-linier $V(h)$, di mana h adalah ketinggian dalam meter:

$$V(h) = 0.5 \times \sqrt{h} \times e^{-0.1h}$$

Perusahaan ingin mengetahui **volume total rata-rata** dari batang pohon tersebut (diasumsikan fungsi ini merepresentasikan luas penampang dalam m^2) ketika diintegrasikan dari permukaan tanah ($h = 0$) hingga ketinggian 10 meter ($h = 10$).

Hitunglah perkiraan nilai integral definit I berikut menggunakan **Metode Integrasi Monte Carlo** dengan $N = 50.000$ sampel:

$$I = \int_0^{10} \left(0.5 \times \sqrt{h} \times e^{-0.1h}\right) dh$$

```r
# 1. Definisikan Fungsi (Luas Penampang Kayu Bakar)
# h = ketinggian (variabel integrasi)
g <- function(h) {
  return(0.5 * sqrt(h) * exp(-0.1 * h))
}
# 2. Tentukan Parameter Monte Carlo
a <- 0      # Batas bawah integral (tanah)
b <- 10     # Batas atas integral (ketinggian potong)
N <- 50000 # Jumlah sampel acak

# 3. Implementasi Monte Carlo

# a. Ambil sampel acak dari distribusi seragam U(a, b)
H <- runif(N, min = a, max = b)

# b. Evaluasi fungsi g(h) pada setiap titik sampel
Y <- g(H)

# c. Hitung rata-rata sampel (Sample Mean)
rata_rata_sampel <- mean(Y)

# d. Estimasi Integral (I_hat = (b - a) * Sample Mean)
I_estimasi_MC <- (b - a) * rata_rata_sampel

# 4. Tampilkan Hasil
print("--- Hasil Estimasi Volume Kayu Bakar ---")
print(paste("Jumlah Sampel (N):", N))
print(paste("Rentang Ketinggian (b-a):", (b-a), "meter"))
print(paste("Estimasi Volume Rata-rata (Integral I_hat):", round(I_estimasi_MC,
4), "m^3"))
```

Penjelasan :

Kode R ini mengaplikasikan Integrasi Monte Carlo untuk mengestimasi volume rata-rata kayu bakar ($\int_0^{10} g(h)dh$), di mana $g(h) = 0.5\sqrt{h}e^{-0.1h}$ mewakili luas penampang. Pertama, fungsi $g(h)$ dan parameter simulasi ($a = 0, b = 10, N = 50.000$) ditetapkan. Selanjutnya, program membangkitkan N titik acak seragam dalam rentang [0,10] menggunakan runif(). Nilai-nilai ini dievaluasi oleh fungsi $g(h)$ untuk mendapatkan Y. Akhirnya, rata-rata sampel (mean(Y)) dihitung dan dikalikan dengan rentang ketinggian $(b - a)$, menghasilkan estimasi Monte Carlo $\hat{I}$, yang diinterpretasikan sebagai perkiraan volume rata-rata dalam m³.

KASUS III

Fungsi $f(x) = \sqrt{1 - x^2}$ merepresentasikan setengah dari lingkaran dengan jari-jari 1. Integral fungsi ini dari 0 sampai 1 memberikan luas seperempat lingkaran, yang nilai eksaknya adalah $\frac{\pi}{4} \approx 0.7854$.
Estimasi integral ini menggunakan metode Monte Carlo dengan berbagai ukuran sampel $N = 10,100,1000,10000$.

```r
r
# Definisi fungsi
f <- function(x) sqrt(1 - x^2)

# Nilai eksak (luas seperempat lingkaran)
nilai_eksak <- pi/4

# Variasi ukuran sampel
N_vals <- c(10, 100, 1000, 10000)

# Estimasi Monte Carlo untuk setiap N
hasil <- sapply(N_vals, function(N){
    U <- runif(N)  # Generate sampel acak U(0,1)
    mean(f(U))     # Rata-rata nilai fungsi
})

# Hitung error absolut
error <- abs(nilai_eksak - hasil)

# Tampilkan hasil
hasil_df <- data.frame(
    N = N_vals,
    Estimasi = round(hasil, 4),
    Nilai_Eksak = round(nilai_eksak, 4),
    Error = round(error, 4)
)

print(hasil_df)
```

Penjelasan :
Kode R ini mengimplementasikan Integrasi Monte Carlo untuk mengestimasi luas seperempat lingkaran dengan jari-jari 1. Fungsi sqrt(1 - x^2) mendefinisikan kurva lingkaran pada kuadran pertama. Program menghitung estimasi integral menggunakan berbagai ukuran sampel (10, 100, 1.000, 10.000) dengan mengambil titik acak seragam antara 0 dan 1, mengevaluasi fungsi di titik-titik tersebut, dan merata-ratakan hasilnya. Selisih antara estimasi dan nilai eksak π/4 dihitung untuk menunjukkan konvergensi metode Monte Carlo seiring dengan meningkatnya jumlah sampel.

KASUS IV

Lakukan analisis galat dan konvergensi metode Monte Carlo untuk mengestimasi integral $\int_0^1 \operatorname{sinc}(x)dx$, dimana $\operatorname{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$. Fungsi sinc sangat penting dalam teori sinyal dan pemrosesan gambar.

```
r

# Analisis Galat dan Konvergensi untuk f(x) = sinc(x)

f <- function(x) {
  # Fungsi sinc(x) = sin(πx)/(πx)
  # Penanganan khusus untuk x = 0 (limitnya adalah 1)
  ifelse(x == 0, 1, sin(pi * x) / (pi * x))
}

# Nilai eksak: ∫₀¹ sinc(x) dx = Si(π)/π ≈ 0.902823
# dimana Si adalah integral sinus
nilai_eksak <- 0.902823  # Nilai referensi dari perhitungan analitik
N_vals <- 10^(1:5) # 10, 100, 1000, 10000, 100000
n_rep <- 100
hasil_matrix <- matrix(0, nrow = n_rep, ncol = length(N_vals))

for (i in 1:n_rep) {
    for (j in 1:length(N_vals)) {
        N <- N_vals[j]
        U <- runif(N)
        hasil_matrix[i, j] <- mean(f(U))
    }
}

rmse <- sqrt(colMeans((hasil_matrix - nilai_eksak)^2))

plot(log10(N_vals), log10(rmse),
    type = "b", pch = 19,
    xlab = "log10(N)",
    ylab = "log10(RMSE)",
    main = "Konvergensi Monte Carlo untuk f(x) = sinc(x)")
abline(a = log10(rmse[1]) + 0.5*log10(N_vals[1]),
    b = -0.5, col = "red", lty = 2)
legend("topright",
    legend = c("Observasi", "Teoritis: slope = -0.5"),
    col = c("black", "red"), lty = c(1, 2), pch = c(19, NA))
```

Penjelasan :

Kode R ini mengimplementasikan analisis galat dan konvergensi untuk integrasi Monte Carlo dari fungsi sinc(x) pada interval [0,1]. Fungsi sinc(x) = sin(πx)/(πx) merupakan fungsi penting dalam pemrosesan sinyal dengan nilai limit 1 saat x mendekati 0.

Program melakukan estimasi integral menggunakan berbagai ukuran sampel (10, 100, 1.000, 10.000, 100.000) dengan mengulangi setiap percobaan sebanyak 100 kali untuk menghitung RMSE (Root Mean Square Error). Hasilnya divisualisasikan dalam plot log-log yang menunjukkan hubungan antara jumlah sampel dan galat, dengan garis referensi slope -0.5 yang merepresentasikan konvergensi teoritis metode Monte Carlo.

Analisis ini demonstrasi bagaimana akurasi estimasi Monte Carlo meningkat seiring dengan bertambahnya jumlah sampel, sekaligus memverifikasi laju konvergensi yang diharapkan untuk fungsi sinc(x).