

隨機採樣黎曼和：定積分數值算法

武國寧·孫 娜

引言

在一些工程數學模型求解中，往往遇到下述問題：(1) 實際問題中有些函數沒有具體的表達式，僅僅知道函數在一些離散點的函數值；(2) 有些函數的原函數不能夠用基本初等函數表示，例如橢圓積分等；(3) 一些函數的原函數需要採用函數級數或者函數無限乘積來表示。還有的函數的原函數需要借助於特殊函數來表示，例如含參變量積分中的伽馬、貝塔函數等等。以上這些問題使得數值積分具有了重要的意義。常見的數值積分方法有：(1) 利用函數在一些節點的加權和來表示定積分。例如數值積分中的梯形公式、辛普森公式和高斯節點積分方法等等。這些積分方法被廣泛應用於一元函數的數值積分。若求解函數為多元函數，計算量會按指數增長。(2) 另外一種數值積分方法為蒙地卡羅方法，該方法採用隨機投點，通過計算落入積分區域與包含積分區域的一個規則區域內點的比例來計算定積分。蒙地卡羅方法提高了多元函數數值積分的計算速度，應用較為廣泛。定積分中對積分區間（或區域）分隔後，黎曼和與每個小區間上隨機取點相關。若把每個小區間上的選點過程看作一個在該小區間上服從均勻分布隨機變量的取值，則黎曼和為隨機變量的函數。大數定律表明：對於獨立重複試驗，頻率穩定於概率。受此啟發，我們可以通過重複試驗，研究黎曼和的平均值，以此研究定積分。數值試驗表明：(1) 當分割給定，黎曼和平均值隨著試驗次數的增加收斂於函數的定積分，且隨機黎曼和直方圖呈常態分布；(2) 當試驗次數給定，隨機黎曼和平均值隨著分割份數的增加收斂於積分值，方差隨著分割份數的增加趨於零。

隨機黎曼和：定積分數值算法

在微積分教材中，定積分定義如下：

定義 1: 設 $f(x)$ 為定義在 $[a, b]$ 上的有界函數，在 $[a, b]$ 中任取分點 $\{x_i\}_0^n$ ，做成一種劃分：

$$P : a = x_0 < x_1 < x_2 < \cdots < x_n = b$$

並任意取點 $\xi_i \in [x_{i-1}, x_i]$ 。記小區間的長度為 $\Delta x_i = x_i - x_{i-1}$ ，令 $\lambda(P) = \max_{1 \leq i \leq n} \Delta x_i$ ，

若極限

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i) \Delta x_i,$$

存在, 且極限值與分割 P 和 ξ_i 的取法無關, 則稱 $f(x)$ 在 $[a, b]$ 上黎曼可積分。其極限值 I 為 $f(x)$ 在 $[a, b]$ 上的定積分, 記為:

$$\int_a^b f(x) dx.$$

在上述定義中, 任意取點 $\xi_i \in [x_{i-1}, x_i]$, 可以認為 $[x_{i-1}, x_i]$ 上服從均勻分布隨機變量 X_i 的取值。當函數 $f(x)$ 和劃分 P 給定時, 函數的黎曼和

$$\sigma(f; \boldsymbol{\xi}, P) = \sum_{i=1}^n f(\xi_i) \Delta x_i,$$

可以看作隨機變量 $X_i = \xi_i$ 的表達式 (求一次黎曼和, 對應一次隨機試驗), 這裡 $\boldsymbol{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ 。為分析隨機黎曼和隨著試驗次數和分割份數增加的統計規律, 我們從兩個方面展開數值試驗:

- (1) 當分割給定, 通過增加重複試驗的次數研究黎曼和的穩定性 (頻率穩定於概率);
- (2) 當分割加細, 隨機變量增多時, 黎曼和是否近似於常態分布 (中央極限定理)。

數值試驗設計如下: 討論積分 $\int_0^1 x^2 dx$, 這裡我們把 $[0, 1]$ 區間 n 等份:

$$P: \frac{0}{n} < \frac{1}{n} < \frac{2}{n} < \dots < \frac{i-1}{n} < \frac{i}{n} < \dots < \frac{n}{n} = 1$$

假設 X_i 為第 i 個區間 $\left[\frac{i-1}{n}, \frac{i}{n}\right]$ 上服從均勻分布的隨機變量: $X_i \sim U\left(\left[\frac{i-1}{n}, \frac{i}{n}\right]\right)$, 黎曼和

$$\sigma(f; \mathbf{X} = \boldsymbol{\xi}^1, P) = \sum_{i=1}^n f(\xi_i^1) \Delta x_i$$

為一次隨機試驗, 這裡 $\boldsymbol{\xi}^1 = (\xi_1^1, \xi_2^1, \dots, \xi_n^1)$ 。我們定義平均值如下:

$$p_m = \frac{1}{m} \sum_{j=1}^m \sigma(f; \boldsymbol{\xi}^j, P) = \frac{1}{m} \sum_{j=1}^m \sum_{i=1}^n f(\xi_i^j) \Delta x_i \quad (1)$$

這裡 $\boldsymbol{\xi}^j = (\xi_1^j, \xi_2^j, \dots, \xi_n^j)$ 表示隨機向量 \mathbf{X} 第 j 次取值, p_m 表示 $[0, 1]$ 區間 n 等份後 m 次隨機試驗黎曼和的平均值。

圖 1 為 $n = 10$ (10等份), 10^2 次試驗, 黎曼和平均值 p_m 的變化趨勢; 圖 2 為 $n = 10$ (10等份), 10^4 次試驗, 黎曼和平均值 p_m 的變化趨勢。從圖中可以看出: 隨著試驗次數的增加, 黎曼和平均值趨於 $1/3$, 這驗證了頻率穩定於概率這一規律。

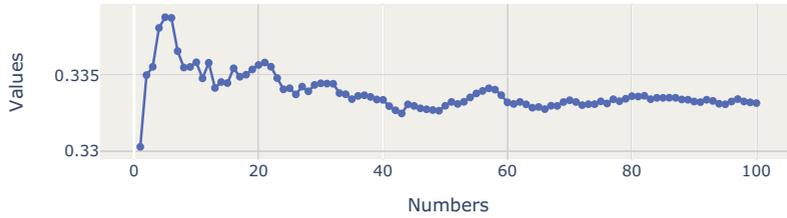


圖 1: 黎曼和平均值隨著試驗次數的變化趨勢, 試驗次數 10^2 次。

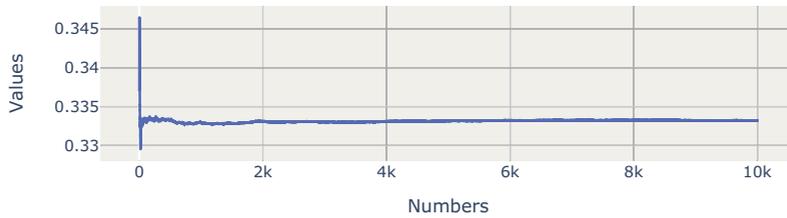


圖 2: 黎曼和平均值隨著試驗次數的變化趨勢, 試驗次數 10^4 次。

圖 3 (a) $n = 10$ (10 等份), 10^4 次隨機試驗, 黎曼和直方分佈圖; (b) $n = 10$ (10 等份), 10^4 次隨機試驗。上曲線為每次隨機試驗黎曼和的變化趨勢, 下曲線為每次試驗黎曼和與真實值之間的誤差; (c) $n = 10$ (10 等份), 10^5 次隨機試驗, 黎曼和直方分佈圖; (d) $n = 10$ (10 等份), 10^5 次隨機試驗。上曲線為每次隨機試驗黎曼和的變化趨勢, 下曲線為每次試驗黎曼和與真實值之間的誤差; (e) $n = 100$ (100 等份), 10^4 次隨機試驗, 黎曼和直方分佈圖; (f) $n = 100$ (100 等份), 10^4 次隨機試驗。上曲線為每次隨機試驗黎曼和的變化趨勢, 下曲線為每次試驗黎曼和與真實值之間的誤差。從圖中可以看出隨機黎曼和直方圖近似於常態分布, 其中心為積分的真實值。當分割的份數增加為 100 份時, 其精準度得到了較大的提高。

下面分析隨機黎曼和平均值 p_m 隨分割份數增加的變化趨勢。圖 4 為等份份數分別為 10, 20, 30, ..., 500 份, 對於每個分割份數試驗次數為 10^4 次。從圖中可以看出隨着分割份數的增加, 黎曼和平均值收斂於真實的積分值。圖 5 分析了分割份數為 10, 20, 30, ..., 500 份, 10^4 次試驗隨機黎曼和的標準差。從圖中可以看出, 隨著分割份數的增加, 黎曼和的方差趨於零。

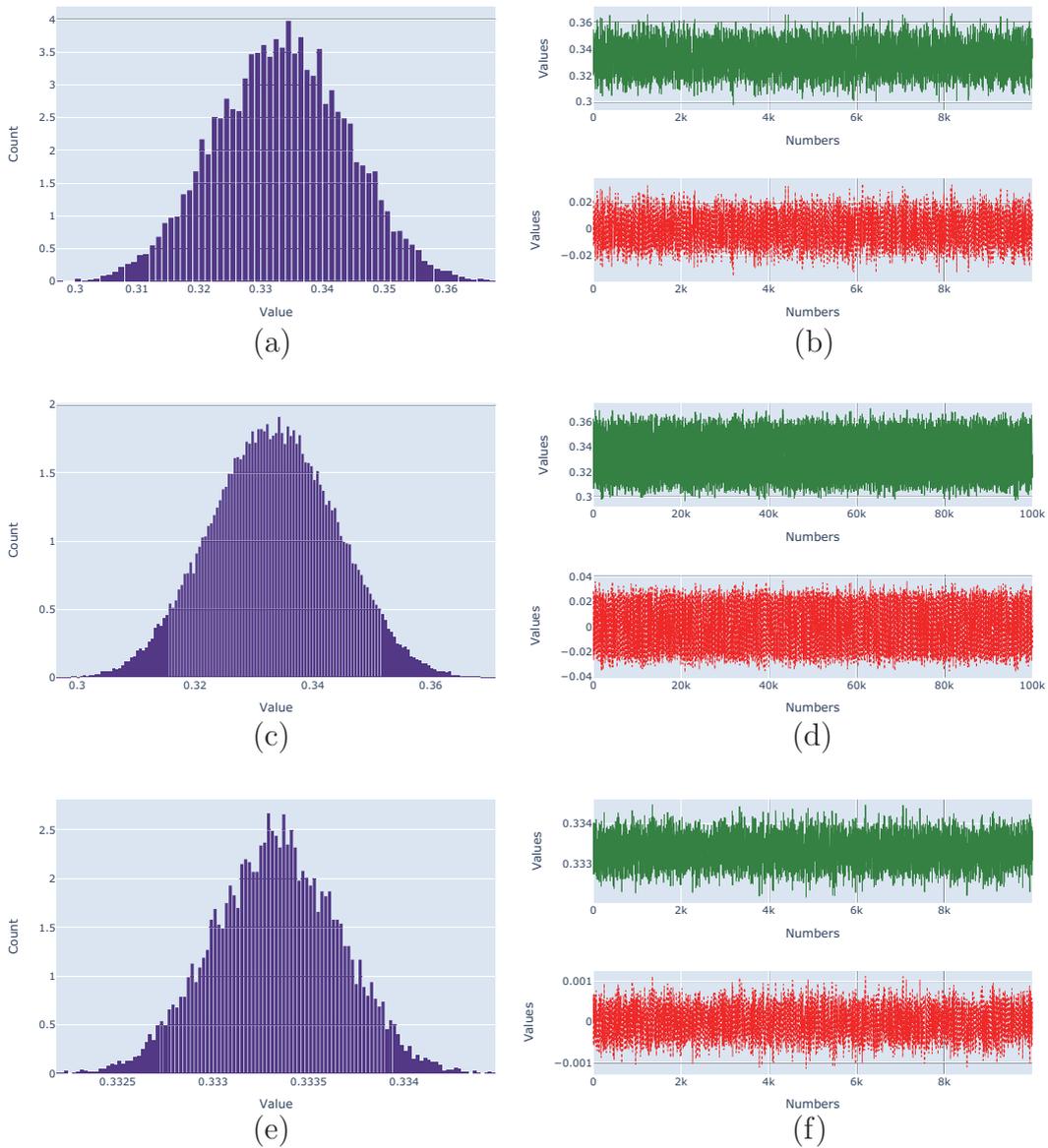


圖 3: (a) 隨機黎曼和分布直方圖, 等分份數 $n = 10$, 試驗次數 10^4 次; (b) 上曲線為隨機黎曼和, 下曲線為黎曼和與真實積分值的誤差, 等分份數 $n = 10$, 試驗次數 10^4 次; (c) 隨機黎曼和分布直方圖, 等分份數 $n = 10$, 試驗次數 10^5 次; (d) 上曲線為隨機黎曼和, 下曲線為黎曼和與真實積分值的誤差, 等分份數 $n = 10$, 試驗次數 10^5 次; (e) 隨機黎曼和分布直方圖, 等分份數 $n = 100$, 試驗次數 10^4 次; (f) 上曲線為隨機黎曼和, 下曲線為黎曼和與真實積分值的誤差, 等分份數 $n = 100$, 試驗次數 10^4 次。

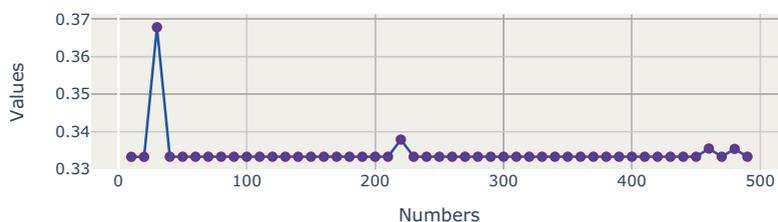


圖 4: 隨機黎曼和平均值的變化趨勢, 分割份數 $n = 10, 20, 30, \dots, 500$, 試驗次數 $n = 10^4$ 。

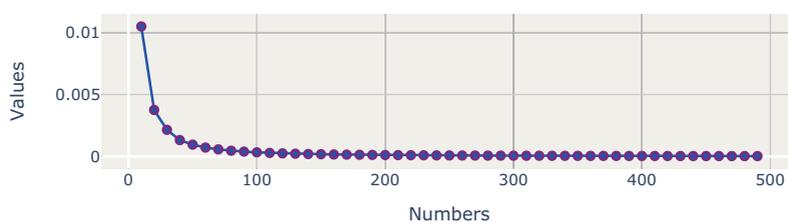


圖 5: 隨機黎曼和標準差的變化趨勢, 分割份數 $n = 10, 20, 30, \dots, 500$, 試驗次數 $n = 10^4$ 。

圖 6 為採用蒙地卡羅方法計算積分 $\int_0^1 x^2 dx$ 得到的圖形。這裡投點個數為 10^4 個, 藍色點為落在函數和座標軸所圍成區域的內部, 黃色點表示落在函數和座標軸所圍成區域的外部。該方法得到的積分近似值為 0.3300。

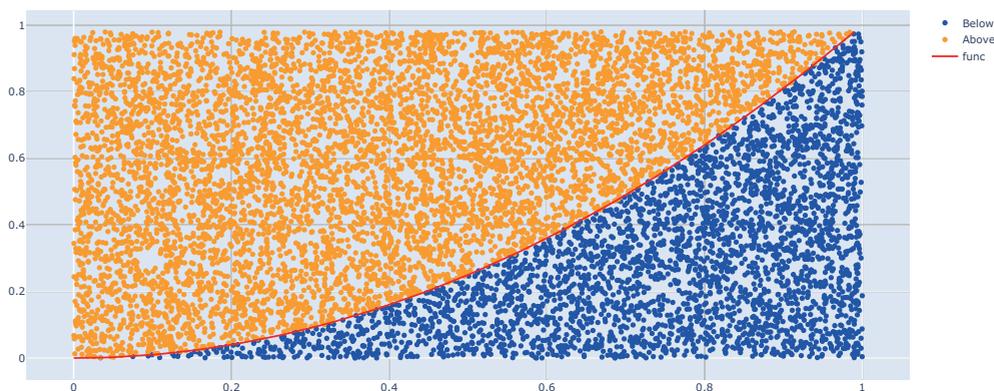


圖 6: 蒙地卡羅隨機投點試驗, 投點個數為 10^4 個, 積分值近似為 0.3300。

為了對比隨機黎曼和平均值與蒙地卡羅方法的收斂精準度, 我們做了以下對比試驗: 對於

隨機黎曼和, 分割份數為 10 份 ($n = 10$), 試驗次數分別為 10, 20, 30, ..., 10^3 次, 計算其平均值 $p_{10}, p_{20}, \dots, p_{10^3}$ 。對於蒙地卡羅方法, 投擲點的個數分別為: 100, 200, 300, ..., 10^4 , 計算其近似值。這樣保證了每次試驗兩種方法取點的個數相同。圖 7 (a), (b), (c), (d) 分別為定積分 $\int_0^1 x^2 dx$, $\int_2^{10} \frac{1}{\ln x} dx$, $\int_0^1 \frac{\sin x}{x} dx$, $\int_0^1 e^{-x^2} dx$ 採用以上兩種方法結果對比圖。圖中藍色曲線代表蒙地卡羅方法, 紅色曲線代表隨機黎曼和方法。試驗結果表明: 在相同點數的情況下, 隨機黎曼和方法具有較快的收斂速度。

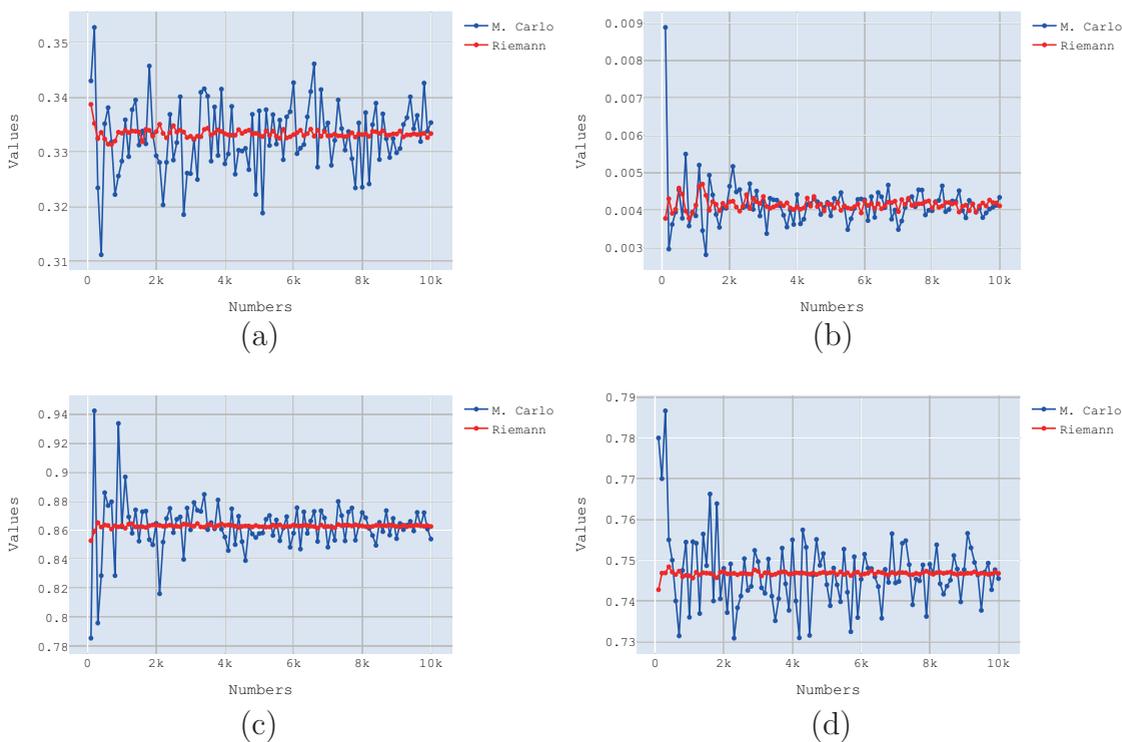


圖 7: 數值積分, 從上到下, 從左到右函數依次為 $\int_0^1 x^2 dx$, $\int_2^{10} \frac{1}{\ln x} dx$, $\int_0^1 \frac{\sin x}{x} dx$, $\int_0^1 e^{-x^2} dx$ 。積分方法: 隨機黎曼和 (紅色); 蒙地卡羅方法 (藍色)。

結語

本文將定積分中黎曼和在每個小區間的取值看作均勻分布的隨機變量, 從數值試驗的角度探討了黎曼和的分佈。試驗表明: 當分割給定, 隨機黎曼和的平均值隨著試驗次數增加收斂於函數的定積分, 其方差隨著分割份數的增加而減小。對比蒙地卡羅方法, 說明了隨機黎曼和方法具有較好的收斂速度。不足之處在於: 沒有從理論上證明隨機黎曼和隨著分割份數的增加趨於常態分布。

參考文獻

1. 王能超。數值分析簡明教程[M], 2版。武漢: 華中科技大學出版社, 43-59, 2002。
2. 克萊因。古今數學思想[M]。上海: 上海科學技術出版社, 1979。
3. Stewart, J. M., *Python for Scientists* [M], Cambridge University Press, 2017.
4. Robert, C. P., Casella, G., *Monte Carlo Statistical Methods* [M], Springer, 2009.
5. 陳紀修, 於崇華, 近路。數學分析, 上冊[M], 2版。北京: 高等教育出版社, 272-280, 2004。

—本文作者武國寧、孫娜任教中國石油大學數學系—