

Università degli Studi di Palermo
Facoltà di Economia
Dipartimento di Scienze Economiche, Aziendali e Statistiche

Appunti del corso di Matematica

14 - Integrazione numerica

Anno Accademico 2015/2016

M. Tumminello, V. Lacagnina, A. Consiglio

1. Introduzione

Sebbene la **formula fondamentale del calcolo integrale** fornisca uno strumento prezioso per valutare un integrale definito, vi sono casi in cui questa formula non può essere sfruttata, in quanto l'antiderivata della funzione integranda non può essere espressa tramite una combinazione finita di funzioni elementari (seno, logaritmo, potenze, ecc.). Un esempio classico in statistica è il calcolo della probabilità di eventi che seguono una distribuzione Gaussiana (o Normale). Infatti, considerando anche la forma più semplice di questa distribuzione (media 0 e varianza 1), la funzione densità di probabilità è:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Data questa funzione densità, la probabilità che la variabile aleatoria x assuma valori compresi tra a e b si calcola come:

$$\int_a^b f(x)dx = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

Sebbene la funzione $e^{-\frac{x^2}{2}}$ sia integrabile (la funzione è continua in tutto il suo dominio), non è possibile esprimere la sua antiderivata tramite una combinazione finita di funzioni elementari. In questo caso è necessario ricorrere a metodi numerici. Studieremo (tra i tanti) tre metodi numerici per il calcolo di integrali:

- (1) Metodo del punto medio o del rettangolo
- (2) Metodo di Simpson
- (3) Metodo Montecarlo

Vedremo che ciò che differenzia i tre metodi è principalmente l'accuratezza, ossia l'errore di approssimazione dell'integrale. Altri fattori da considerare sono tuttavia la semplicità di applicazione e la velocità computazionale.

2. Metodo del punto medio

Sia f una funzione continua sull'intervallo $[a, b]$ e sia $N > 0$ un numero naturale. Si consideri una partizione uniforme di $[a, b]$ in N sottointervalli di ampiezza omogenea: $\Delta x = \frac{b-a}{N}$.

L'interale definito secondo Riemann di f su $[a, b]$ è dato dal limite per N tendente all'infinito della somma di Riemann:

$$\lim_{N \rightarrow \infty} \sum_{i=1}^N f(x_i^*) \Delta x,$$

dove x_i^* è un punto appartenente all' i -esimo intervallo $[x_{i-1}, x_i]$. Nel metodo del punto medio si sceglie x_i^* come il punto medio dell' i -esimo

intervallo

$$x_i^* = \frac{1}{2}(x_{i-1} + x_i).$$

Come si può osservare dalla Figura 1, questa sembra essere una buona scelta. La somma di Riemann con x_i^* calcolato nel punto medio prende il nome di **approssimazione del punto medio di ordine N** :

$$\sigma_M(N, f, [a, b]) = \Delta x \sum_{i=1}^N f(x_i^*) = \frac{(b-a)}{N} \sum_{i=1}^N f(x_i^*).$$

In generale, l'approssimazione diventerà più accurata al crescere di N .

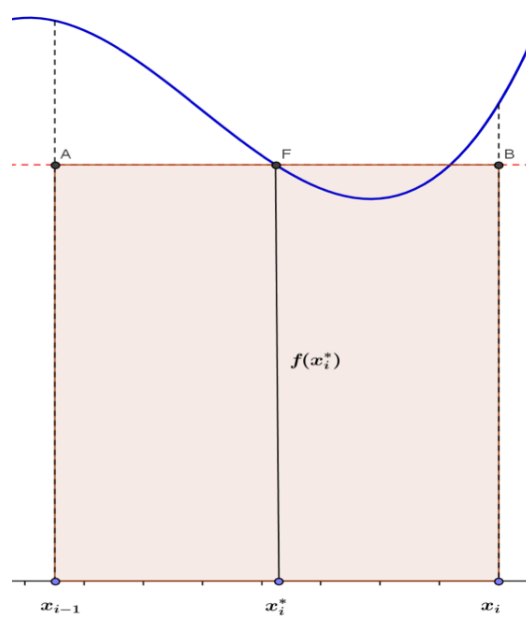


FIGURA 1. Metodo del punto medio.

D'altro canto un N troppo grande renderebbe il calcolo impraticabile. Quindi il problema è: qual è il più piccolo valore di N tale da garantire che $\sigma_M(N, f, [a, b])$ sia un'approssimazione "accettabile" dell'integrale di f su $[a, b]$? Il seguente teorema risponde a questa domanda.

TEOREMA 2.1 (Metodo del punto medio). *Sia f una funzione continua nell'intervallo $[a, b]$ e sia N un intero positivo. Per ogni intero i con $1 \leq i \leq N$, sia x_i^* il punto medio dell' i -esimo sotto-intervallo, $[x_{i-1}, x_i]$, di $[a, b]$. L'approssimazione di ordine N dell'integrale $\int_a^b f(x)dx$ è data da:*

$$\sigma_M(N, f, [a, b]) = \Delta x \sum_{i=1}^N f(x_i^*).$$

Inoltre se C è una costante tale che $|f''(x)| \leq C \forall x \in [a, b]$, allora l'errore di approssimazione è dato da:

$$\left| \int_a^b f(x) dx - \sigma_M(N, f, [a, b]) \right| \leq \frac{C}{24} \cdot \frac{(b-a)^3}{N^2} = \frac{C}{24} \cdot (b-a) \cdot \Delta x^2.$$

Questo teorema mostra, quindi, che l'errore che si commette è di secondo ordine in Δx .

Esempio 2.1

Valutare il seguente integrale definito:

$$I = \int_0^1 \frac{1}{1+x^2} dx.$$

Si osservi che, in questo caso, come abbiamo visto in classe, l'antiderivata di $\frac{1}{1+x^2}$ esiste in forma chiusa ed è $\arctan(x)$. Quindi sappiamo calcolare l'integrale in modo esatto:

$$\int_0^1 \frac{1}{1+x^2} dx = \arctan(1) - \arctan(0) = \frac{\pi}{4} - 0 = \frac{\pi}{4} \cong 0.7854$$

Proviamo ad ottenere comunque un'approssimazione numerica di questo integrale attraverso il metodo del punto medio. In figura 2 (pannello a sinistra), sono riportati la funzione integranda nell'intervallo $[0, 1]$ ed $N = 4$ sub-intervalli di $[0, 1]$. L'approssimazione dell'integrale tramite il metodo del punto medio sarà quindi data da:

$$\sigma_M(4, f, [0, 1]) = 0.25 \left[\frac{1}{1+(0.125)^2} + \frac{1}{1+(0.375)^2} + \frac{1}{1+(0.625)^2} + \frac{1}{1+(0.875)^2} \right] \cong 0.7867$$

Il valore di $\sigma_M(4, f, [0, 1])$ rappresenta l'area dei rettangoli mostrati nella figura sopra. Per ottenere l'intervallo in cui il valore esatto dell'integrale è contenuto, è necessario stimare l'errore che si commette nell'approssimazione. Nella figura 2 (pannello a destra) è riportato il grafico di $|f''(x)|$ nell'intervallo $[0, 1]$ (si ricorda che nel nostro caso $f(x) = 1/(1+x^2)$). Come si può notare

$$\max_{x \in [0, 1]} |f''(x)| = 2.$$

Ponendo quindi $C = 2$ nella formula dell'errore fornita nel teorema precedente otteniamo che:

$$\frac{C}{24} \cdot \frac{(b-a)^3}{N^2} = \frac{2}{24} \cdot \frac{(1-0)^3}{4^2} = \frac{1}{24} \cdot \frac{1}{8} = \frac{1}{192} \cong 0.0052$$

Pertanto il valore esatto dell'integrale sarà compreso nell'intervallo fra $0.7867 - 0.0052$ e $0.7867 + 0.0052$, ossia

$$0.7815 \leq I \leq 0.7919.$$

Come si può facilmente osservare, il valore esatto dell'integrale ($\pi/4$) cade effettivamente in questo intervallo. In alcuni casi viene richiesto di determinare un integrale con un ordine di errore predefinito, ad esempio, con una precisione a terza cifra decimale (quindi con un errore di 0.0005 o inferiore). In questo caso la formula per il calcolo dell'errore va utilizzata per ricavare il numero di sub-intervalli N da usare. Nel nostro caso avremmo:

$$\text{Errore} = \frac{C}{24} \cdot \frac{(b-a)^3}{N^2} = \frac{1}{12} \cdot \frac{1}{N^2} \leq 0.0005 = \frac{1}{2}10^{-3},$$

da cui

$$N \geq \sqrt{\frac{1}{6}10^3} \cong 13.$$

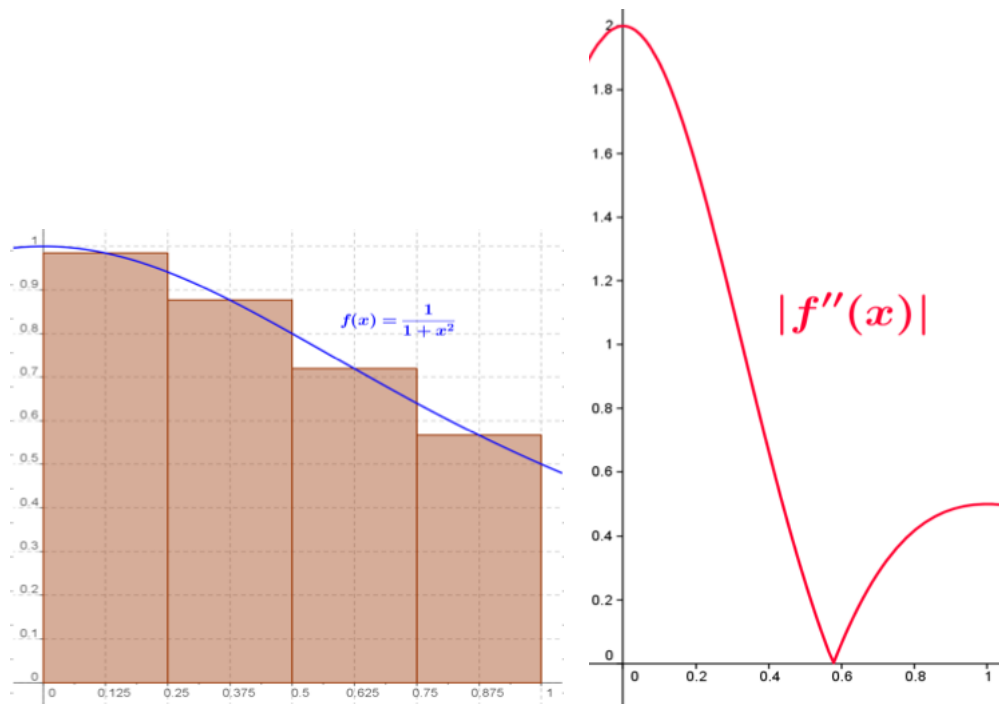


FIGURA 2. Pannello sinistro: metodo del punto medio applicato al calcolo dell'integrale definito di cui all'esercizio 2.1. Pannello destro: valore assoluto della derivata seconda della funzione integranda di cui all'esercizio 2.1.

3. Metodo di Simpson

La regola del punto medio approssima la funzione su Δx tramite una retta. Una migliore approssimazione si può ottenere se, invece della retta si utilizza una parabola, ossia un polinomio di secondo grado, come mostrato in figura 3.

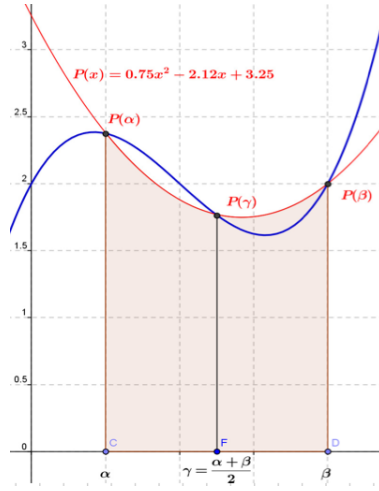


FIGURA 3. Metodo di Simpson.

Data una generica parabola

$$P(x) = Ax^2 + Bx + C$$

passante per tre punti (si ricorda che per tre punti distinti passa una ed una sola parabola) di cui due punti sono $(\alpha, f(\alpha))$ e $(\beta, f(\beta))$, con α e β estremi del sub-intervallo e l'ultimo punto è $(\gamma, f(\gamma))$ con $\gamma = \frac{\alpha + \beta}{2}$, l'area sottostante la parabola è data dal seguente teorema:

TEOREMA 3.1. *Sia $P(x) = Ax^2 + Bx + C$ definita sull'intervallo $[\alpha, \beta]$. Allora:*

$$\int_{\alpha}^{\beta} P(x) dx = \frac{\beta - \alpha}{6} \cdot [P(\alpha) + 4P(\gamma) + P(\beta)],$$

dove $\gamma = (\alpha + \beta)/2$.

La dimostrazione consiste nell'integrare il polinomio di secondo grado $P(x) = Ax^2 + Bx + C$ tra α e β , e si lascia per esercizio.

La formula di approssimazione di Simpson per l'integrale di una funzione in un certo intervallo $[a, b]$ si ottiene suddividendo questo intervallo in un numero pari $N = 2M$, con $M \in \mathbb{N}$ intervalli e approssimando, per ogni coppia di intervalli successivi la funzione con una parabola passante per i tre punti determinati dai due intervalli successivi, come mostrato in figura 3. In particolare, siano $[x_{2i-1}, x_{2i}]$

e $[x_{2i}, x_{2i+1}]$ due sub-intervalli consecutivi. L'area della parabola che insiste su questi due generici sub-intervalli, passando per i punti $(x_{2i-1}, f(x_{2i-1}))$, $(x_{2i}, f(x_{2i}))$ e $(x_{2i+1}, f(x_{2i+1}))$, è, in base al teorema precedente:

$$\begin{aligned}\int_{x_{2i-1}}^{x_{2i+1}} P(x)dx &= \frac{x_{2i+1} - x_{2i-1}}{6} \cdot [P(x_{2i-1}) + 4P(x_{2i}) + P(x_{2i+1})] = \\ &= \frac{2\Delta x}{6} \cdot [f(x_{2i-1}) + 4f(x_{2i}) + f(x_{2i+1})] = \\ &= \frac{\Delta x}{3} \cdot [f(x_{2i-1}) + 4f(x_{2i}) + f(x_{2i+1})]\end{aligned}$$

Sommando l'area ora ottenuta su tutte le coppie di intervalli adiacenti e disgiunti otteniamo l'approssimazione di Simpson dell'integrale di f su $[a, b]$:

$$\int_a^b f(x)dx = \frac{\Delta x}{3} \cdot \sum_{i=0}^{M-1} [f(x_{2i}) + 4f(x_{2i+1}) + f(x_{2i+2})].$$

Il metodo di Simpson appare un poco più complicato di quello del valor medio. Tuttavia, questa maggiore difficoltà è compensata dal grado di approssimazione che esso garantisce, come mostra il seguente teorema:

TEOREMA 3.2. *Sia f una funzione continua sull'intervallo $[a, b]$ e sia N un intero positivo pari. Sia $\sigma_S(N, f)$ l'approssimazione di Simpson si ordine N dell'integrale di f su $[a, b]$:*

$$\sigma_S(N, f) = \frac{\Delta x}{3} \cdot \sum_{i=0}^{M-1} [f(x_{2i}) + 4f(x_{2i+1}) + f(x_{2i+2})], \text{ con } N = 2M.$$

Se C è una costante tale che $|f^{IV}(x)| \leq C \forall x \in [a, b]$ allora l'errore di approssimazione è dato da:

$$\left| \int_a^b f(x)dx - \sigma_S(N, f) \right| \leq \frac{C}{180} \cdot \frac{(b-a)^5}{N^4} = \frac{C}{180} \cdot (b-a) \cdot \Delta x^4$$

Quindi, l'errore di approssimazione è di ordine 4, nettamente migliore dell'ordine 2 ottenuto con il metodo del punto medio.

3.1. Osservazioni.

- (1) La stima dell'errore fatta per i due metodi considerati è una stima dell'errore massimo, ossia del caso peggiore (**worst case scenario**), pertanto è possibile che in alcuni casi l'errore di approssimazione effettivamente commesso sia minore di quello indicato.
- (2) Per alcune forme funzionali, ad esempio le funzioni periodiche, non è detto che il metodo di Simpson ottenga approssimazioni migliori di quelle ottenute con il metodo del punto medio.

Esercizio 3.1

Lo studente applichi il metodo di Simpson per valutare l'integrale definito considerato nell'esempio 2.1.

4. Metodo Monte Carlo

Il Metodo Monte Carlo è un metodo statistico per approssimare il valore dell'integrale di una funzione limitata f su un intervallo $[a, b]$:

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

Per descrivere il metodo, assumiamo che la funzione $f(x)$ sia non negativa $\forall x \in [a, b]$. Il metodo tuttavia può essere facilmente generalizzato al caso in cui la funzione assuma anche valori negativi nell'intervallo di integrazione. Sia M un maggiorante (o anche il massimo della funzione f sull'intervallo $[a, b]$). Si costruisce il rettangolo di base $[a, b]$ e altezza M (si veda la figura 4, dove $[a, b] = [-1, 2]$). In questo modo l'area sottesa dal grafico di f , ossia l'integrale di f su $[a, b]$, sarà contenuta nell'area del suddetto rettangolo. Sia ora x un numero casuale compreso nell'intervallo $[a, b]$ e y un numero casuale compreso nell'intervallo $[0, M]$. Tutti i linguaggi di programmazione evoluti (R, Mathematica, Matlab, Maple, ecc.) consentono di generare numeri (pseudo)casuali siffatti. Si consideri ora il punto di coordinate (x, y) . Qual è la probabilità che esso si trovi al di sotto del grafico della funzione f ? In altre parole, qual è la probabilità che $y \leq f(x)$? Risulta intuitivo che questa probabilità coincida con il rapporto tra I (il valore dell'integrale che vogliamo calcolare) e l'area del rettangolo sopra descritto, $A = (b - a) \cdot M$. In formule scriveremo:

$$p(y \leq f(x)) = \frac{I}{A} = \frac{I}{(b - a) \cdot M}.$$

Questa formula ci consente di vedere il problema di approssimare I da una prospettiva completamente diversa rispetto a quelle considerate finora. Infatti, supponiamo di avere un metodo per approssimare il valore della probabilità $p(y \leq f(x))$. Questa ci consentirebbe di approssimare il nostro integrale leggendo la formula precedente come un'equazione nella variabile I :

$$I = p(y \leq f(x)) \cdot A = p(y \leq f(x)) \cdot (b - a) \cdot M.$$

La probabilità $p(y \leq f(x))$ può essere calcolata (in modo approssimato) per via **frequenzistica**. In particolare, si supponga di generare N coppie di numeri casuali (x_i, y_i) , con $i = 1, \dots, N$, con $x_i \in [a, b]$ e $y_i \in [0, M]$. Sia ora U_N il numero di coppie (x_i, y_i) tali che $y_i \leq f(x_i)$. Possiamo dunque approssimare la quantità $p(y \leq f(x))$ con la quantità:

$$\bar{p}_N = \frac{U_N}{N}.$$

In questo modo otteniamo un'approssimazione dell'integrale cercato:

$$I \cong I_N = \bar{p}_N \cdot (b - a) \cdot M = \frac{U_N}{N} \cdot (b - a) \cdot M = \frac{U_N}{N} \cdot A.$$

E' possibile inoltre dimostrare (vedrete la dimostrazione nel corso di Calcolo delle Probabilità) che, per la legge dei grandi numeri, il valore atteso dell'errore di approssimazione:

$$\sqrt{E[(I - I_N)^2]} = \sqrt{\frac{I(A - I)}{N}} \leq \frac{1}{\sqrt{N}} \sqrt{\frac{A}{2} \left(A - \frac{A}{2}\right)} = \frac{1}{\sqrt{N}} \frac{A}{2}$$

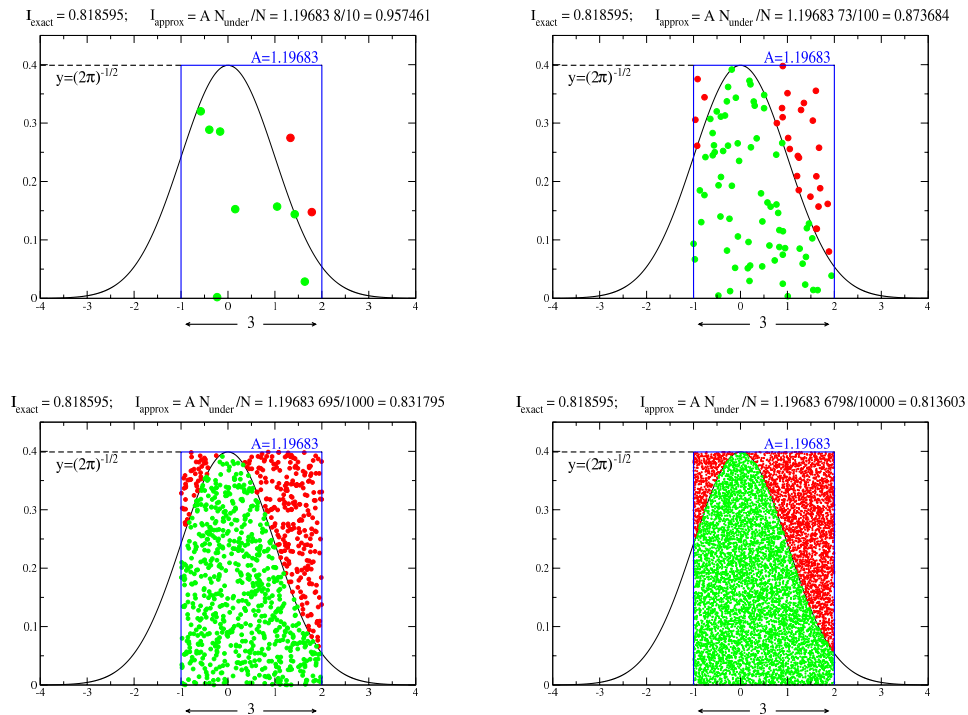


FIGURA 4. Metodo Monte Carlo applicato al calcolo dell'integrale di cui all'esempio 4.1. I diversi pannelli mostrano il grado di approssimazione dell'integrale cercato per diversi numeri di punti casuali generati: 10 (in alto a destra), 100 (in alto a destra), 1000 (in basso a sinistra), 10000 (in basso a destra).

Esempio 4.1

Si valuti il seguente integrale definito:

$$\int_{-1}^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

La funzione integranda è la funzione densità di probabilità della distribuzione Gaussiana (o Normale) con valore atteso 0 e varianza 1. In figura 4 sono riportati i risultati di simulazioni Montecarlo con i relativi valori dell'integrale approssimato al variare del numero di punti simulati, N . Si vede che l'approssimazione migliora al crescere di N . Questo fatto è espresso in modo più chiaro in figura 5, dove è riportato il valore dell'integrale approssimato rispetto al valore "esatto" dell'integrale al variare di N . In figura 5 sono anche riportate le curve: $I + \frac{1}{\sqrt{N}} \frac{A}{2}$ e $I - \frac{1}{\sqrt{N}} \frac{A}{2}$, da cui si vede che buona parte dei punti risulta compresa tra questi due valori.

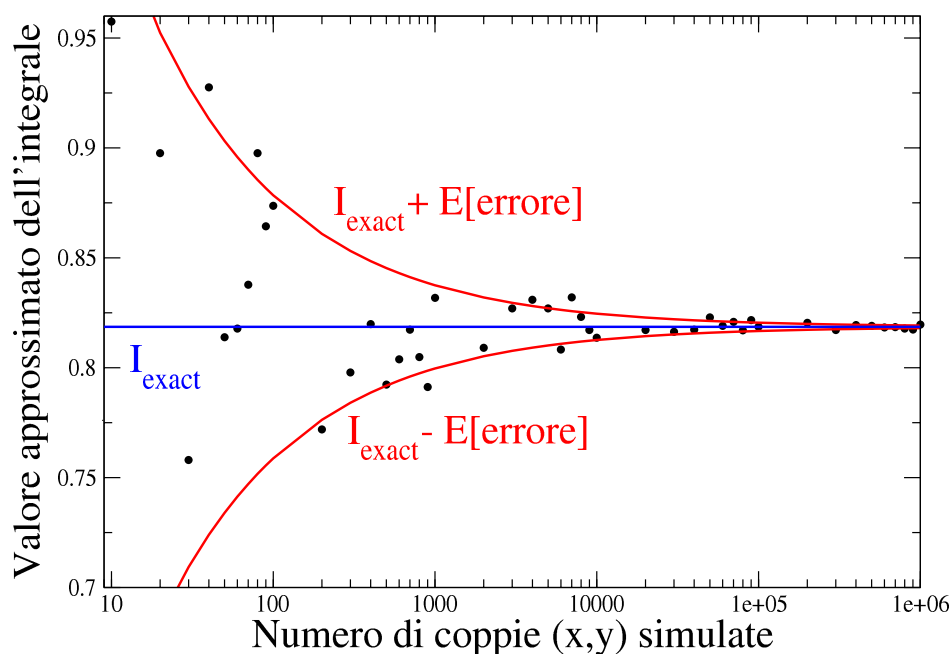


FIGURA 5. Precisione del Metodo Montecarlo per l'approssimazione dell'integrale dell'esempio 4.1.