**Esercizio 5**

**“**Si calcoli con un metodo MC l’integrale e si discutano in dettaglio le possibili tecniche di riduzione della varianza”

**Introduzione**

Un metodo Monte Carlo è una tecnica che fa uso di numeri casuali per risolvere un problema. Questi numeri verranno generati attraverso un apposito algoritmo pseudocasuale.

Posso sfruttare un metodo MC per risolvere numericamente un integrale. Con un numero di dimensioni basso è possibile utilizzare un metodo come quello di Simpson o dei Trapezi. Ma all’aumentare delle dimensioni conviene utilizzare invece un metodo MC. Questi hanno infatti la caratteristica di avere un errore sempre proporzionale a N-1/2 (ove N è il numero di campionamenti), indipendente quindi dal numero di dimensioni in cui si sta operando.

L’integrazione tramite metodo MC si basa sul teorema del limite centrale, che afferma: “date N variabili aleatorie {xi} con distribuzione di probabilità qualsiasi, al limite N🡪∞ la variabile somma sarà distribuita in modo gaussiano. Supponiamo di volere calcolare l'integrale tra 0 e 1 di una generica funzione f(x). Avendo a disposizione un generatore di numeri casuali a probabilità piatta, ed estraendo le x da questo generatore, l'integrale sarà dato dalla seguente espressione:

L'errore su questo integrale discende direttamente dalla sua natura statistica. Il teorema del limite centrale ci garantisce che, dato un campionamento sufficientemente grande, y sia distribuito gaussianamente. In particolare l'errore su I\* sarà dato da:

Si può usare un metodo HIT-MISS di Von Neumann (descritto negli esercizi precedenti) per aumentare la precisione del calcolo dei termini della sommatoria, ma è troppo dispersivo. Per questo motivo si preferisce andare direttamente con un “CRUDE MC ove i numeri sono semplicemente generati tra gli estremi dell’integrale e utilizzati per il calcolo senza controlli intermedi. Successivamente si useranno dei metodi di riduzione della varianza per aumentare la precisione della stima del calcolo numerico dell’integrale.

Le tecniche utilizzabili per ridurre la varianza sono tre:

* Sampling stratificato:
* Si scelgono i punti xi in maniera più uniforme
* Per farlo si divide l’intervallo di integrazione [a,b] in due parti [a,c] e [c,b] con metà della statistica in ogni parte, si avrà allora:
* Importance sampling:
* Questo metodo consiste nel trovare una funzione che bene approssimi il modulo della funzione f(x) in modo da campionare più frequentemente i punti che danno maggiore contributo all'integrale. Data l’ovvia identità

con .

Scegliendo g(x) simile a f(x) si può ridurre la varianza perché vale:

* Metodo delle variabili antitetiche:
* Per ogni punto xi se ne considera un secondo xi’, secondo una procedura specifica, e si usa come contributo all’ integrale I
* Sapendo che vale

Se la correlazione è negativa si riduce la varianza.

**Risultati e discussione**

L’integrale è stato calcolato attraverso il metodo MC con le 4 modalità descritte in precedenza, quella Crude MC e le 3 con le tecniche della riduzione della varianza. Per i 4 casi l’integrale è stato calcolato 1000 volte, come risultato verrà presa la media e la varianza.

I risultati sono i seguenti:

|  |  |  |
| --- | --- | --- |
|  | Media | Varianza (σ2) |
| MC (crude) | 1.71805 | 0.00243702 |
| Sampling Stratificato | 1.71787 | 0.000600326 |
| Importance sampling | 1.71827 | 1.54109\*10-10 |
| Metodo delle var. antitetiche | 1.71796 | 7.78679\*10-5 |

Per il sampling stratificato l’intervallo è stato diviso tra [0;0.5] e [0.5;1] e l’integrale è stato calcolato tra i due estremi dividendo a metà il numero di eventi generato per intervallo.

Per il metodo delle variabili antitetiche essendo la funzione monotona crescente, i valori xi random sono stati scelti tra [a,b] (nel nostro caso [0,1]) ed è stato sfruttato il fatto che prendendo

si avrà una correlazione negativa.

L’importance sampling si è dimostrato il metodo migliore. La funzione scelta è ed è stata integrata la funzione (opportunamente normalizzata) . Avendo fatto un cambio di variabile ora i numeri saranno distribuiti secondo ove r è un generato casualmente tra [0;1].