3. Übung - Zufallszahlen

(Abgabe und Besprechung am 22.11.2018 um 14:00 im CIP-Pool)

1. [*3 Punkte*] **Pseudo-Zufallszahlengeneratoren**

In dieser Aufgabe sollen Sie selbst Zufallszahlen mittels zweier algorithmischer Pseudozufallszahlengeneratoren erzeugen und deren Qualität untersuchen.

(a) Implementieren Sie zunächst den multiplikativen Kongruenzgenerator (MKG)

$$x_n = (ax_{n-1} + c) \mod m$$

mit den Parametern $a = 16807$, $c = 0$ und $m = 2^{31} - 1$. Wählen Sie einen beliebigen *Seed* x_0. Programmieren Sie außerdem einen additiven lagged Fibonacci-Generator (LFG)

$$x_n = (x_{n-p} + x_{n-q}) \mod m$$

mit $n \geq p > q > 0$. Die *Seed*-Sequenz $x_0, x_1, \ldots, x_{p-1}$ können Sie dabei mit dem multiplikativen Kongruenzgenerator erzeugen. Benutzen Sie folgende Parameter: $p = 2281$, $q = 1252$ und $m = 2^{31} - 1$.
Hinweis: Im Folgenden benötigen wir Zufallszahlen im Intervall $u_n \in [0, 1)$, modifizieren Sie die beiden Methoden entsprechend.

(b) Untersuchen Sie die Qualität der Zufallszahlen vom MKG und LFG.

i. Produzieren Sie dazu eine große Menge (z.B. $N = 10^7$) von Zufallszahlen $\{u_1, u_2, \ldots, u_N\}$ mit $u_i \in [0, 1)$ und tragen Sie diese in einem Histogramm auf.

ii. Berechnen Sie die Autokorrelationsfunktion

$$\rho(\tau) = \frac{\sum_{i=1}^{N-\tau}(u_{i+\tau} - \langle u\rangle)(u_i - \langle u\rangle)}{\sum_{i=1}^{N}(u_i - \langle u\rangle)^2},$$

wobei $\langle u\rangle = \frac{1}{N}\sum_{i=1}^{N} u_i$ den Mittelwert der erzeugten Zufallszahlen bezeichnet, und plotten Sie ρ als Funktion von τ. Können Sie mit diesem Maß die Qualität der beiden Generatoren unterscheiden?

iii. Führen Sie nun einen sogenannten Spektraltest durch. Fassen Sie dazu 3er-Tupel von generierten Zufallszahlen $p_i = (u_i, u_{i+1}, u_{i+2})$ für $i = 1, 2, \ldots, N - 2$ zusammen und stellen Sie die Tupel p_i als Punkte im $\mathbb{R}^3$ dar. Untersuchen Sie die beiden erzeugten Datensets von unterschiedlichen Blickwinkeln. Welche Unterschiede lassen Sie schließen, daß die Daten des MKG erheblich schlechtere Qualität besitzen als die des LFG?

2. [*1.5 Punkte*] **Zufallszahlen-Verteilungen**

In dieser Aufgabe sollen Zufallszahlen gemäß einer gegebenen Verteilung generiert werden. Verwenden Sie einen Generator ihrer Wahl zur Erzeugung von gleichverteilten Pseudo-Zufallszahlen im Intervall $[0, 1)$.

(a) Schreiben Sie ein Programm, welches Zufallszahlen

i. auf dem Intervall $[0, a)$ mit der Wahrscheinlichkeitsdichte $f(x) = \frac{2}{a^2}x$

ii. in $\mathbb{R}$ mit der Wahrscheinlichkeitsdichte $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$

liefert.

(b) Prüfen Sie Ihren Algorithmus für $a = 2$ sowie $\mu = 5$ und $\sigma = 2$, indem Sie 10^9 Samples generieren und den Mittelwert sowie die Varianz mit den analytisch berechenbaren Werten vergleichen. Erzeugen Sie außerdem ein Dichtehistogramm Ihrer Samples mit geeigneter Schrittweite und zeichnen Sie zum Vergleich die Dichtefunktion darüber.

3. [*2.5 Punkte*] **Bestimmung von π via** *direct sampling*

Die Zahl π lässt sich analytisch auf verschiedenen Wegen im Prinzip beliebig genau berechnen. Es existieren aber auch ungewöhnliche Methoden die auf Wahrscheinlichkeiten beruhen. Eine einfache Methode zur statistische Berechnung von π besteht darin, daß man N zufällige Punkte auf ein Quadrat "regnen" lässt und berechnet, ob sie innerhalb oder außerhalb eines einbeschriebenen Kreises liegen. Der Anteil der innen liegenden Punkte $\frac{N_{in}}{N}$ ist gleich $\frac{\pi}{4}$ für $N \to \infty$.

(a) Implementieren Sie die Methode.

(b) Erzeugen Sie für ein festes $N = 10^1, 10^2, \ldots, 10^8$ 20 statistisch unabhängige Realisationen des Quotienten $\frac{N_{in}}{N}$. Überzeugen Sie sich, daß $\frac{N_{in}}{N}$ gegen $\frac{\pi}{4}$ konvergiert.

(c) Berechnen Sie ausserdem die mittlere quadratische Abweichung

$$\mathrm{MSE} = \left\langle \left(\frac{N_{in}}{N} - \frac{\pi}{4} \right)^2 \right\rangle$$

und plotten Sie MSE als Funktion von N. Wie skaliert MSE mit N?

4. [*3 Punkte*] **Auch ein Trunkenbold kann π berechnen.**

Sei ein quadartisches Feld mit einem einbeschriebenen Kreis von einer harten Mauer umgeben. Ein Trunkenbold befinde sich zum Zeitpunkt i auf dem Feld an der Position (x_i, y_i) mit $-1 \leq x_i, y_i \leq 1$. Er versucht zu einem neuen Ort $(x^*, y^*) = (x_i + \Delta_x, y_i + \Delta_y)$, der um eine gleichförmige Zufallsvariable $\Delta_x, \Delta_y \in [-\delta, \delta]$ gegenüber dem alten Ort verschoben ist, zu gelangen. Ist die neue Position innerhalb des Quadrats, d.h. $-1 \leq x^*, y^* \leq 1$, dann macht er den Schritt und $(x_{i+1}, y_{i+1}) = (x^*, y^*)$. Falls sich die neue Position jedoch außerhalb des Quadrats befinden würde, meidet er den schmerzhaften Zusammenstoß mit der Mauer und verbleibt länger am jetzigem Ort, d.h. $(x_{i+1}, y_{i+1}) = (x_i, y_i)$. Ein solcher stochastischer Prozess wird Markow-Kette genannt, dabei hängt der neue Zustand des Systems (x_{i+1}, y_{i+1}) nur vom vorherigen Zustand (x_i, y_i) ab. Der Spaziergang dauert insgesamt N Zeitschritte und der Trunkenbold zählt die Anzahl der Zeitschritte N_{in} die er innerhalb des einbeschriebenen Kreises verbracht hat.

(a) Implementieren Sie das Markov-Ketten-Monte-Carlo-Verfahren.

(b) Überzeugen Sie sich für $\delta = 0.3$, daß $\frac{N_{in}}{N}$ gegen $\frac{\pi}{4}$ konvergiert.

(c) Halten Sie ein großes N fest und berechnen Sie die mittlere quadratische Abweichung

$$\mathrm{MSE} = \left\langle \left(\frac{N_{in}}{N} - \frac{\pi}{4} \right)^2 \right\rangle$$

und die Ablehnungsrate für unterschiedliche δ im Bereich $\delta \in [0, 3]$. Plotten Sie MSE und die Ablehnungsrate als Funktion von δ. Welche Ablehnungsrate liefert die höchste Genauigkeit?

Hinweis: Überlegen Sie wie man möglichst statistisch unabhängige Realisationen des Prozesses erzeugt.

Bei Fragen E-Mail an: a.wysocki@lusi.uni-sb.de