Übungsblatt 2

Aufgabe 3: Erzeugung Standardnormalverteilter Zufallszahlen

Die Box-Muller-Methode ist ein Verfahren, mit dem man aus jeweils zwei unabhängigen Standardzufallszahlen z_1 und z_2 zwei unabhängige standardnormalverteilte Zufallszahlen u_1 und u_2 erzeugen kann. Diese sind gegeben durch

$$u_1 = \sqrt{-2 \ln z_1} \cos(2\pi z_2) , \tag{1}$$

$$u_2 = \sqrt{-2 \ln z_1} \sin(2\pi z_2) . \tag{2}$$

Eine Alternative ist die Polarmethode nach G. Marsaglia. Dabei definiert man zunächst die Variablen

$$a_i = 2z_i - 1 , \quad i = 1, 2 \quad \text{und} \tag{3}$$

$$q = a_1^2 + a_2^2 . \tag{4}$$

Falls $q = 0$ oder $q > 1$, werden die Zahlen verworfen und neue Standardzufallszahlen gezogen. Andernfalls definiert man

$$p = \sqrt{\frac{-2 \ln q}{q}} \tag{5}$$

und erhält die standardnormalverteilten Zufallszahlen durch

$$u_i = a_i \cdot p , \quad i = 1, 2 . \tag{6}$$

a) Beweisen Sie, dass die Zufallszahlen u_1 und u_2 aus der Box-Muller-Methode gaußverteilt sind.

b) Schreiben Sie für jede der beiden Methoden ein Unterprogramm, das zwei standardnormalverteilte Zufallszahlen erzeugt.

c) Testen Sie beide Methoden, indem sie jeweils 200 000 Zufallszahlen erzeugen, ein Histogramm berechnen und mit der Normalverteilung vergleichen.

d) Vergleichen Sie die Effizienz beider Methoden, indem Sie die Zeit messen, die für 10^7 Funktionsaufrufe benötigt wird (unter Linux den Befehl `time` vor den Programmaufruf setzen).

Aufgabe 4: Kerndichteschätzung

Die neben dem Histogramm am weitesten verbreitete Methode zur Schätzung von Wahrscheinlichkeitsdichten ist die Kerndichteschätzung. Der sogenannte Kernschätzer für eine Stichprobe von N Realisationen $X_1, X_2, \ldots, X_N$ einer stochastischen Variable X ist definiert als

$$\hat{f}(x) = \frac{1}{Nh} \sum_{i=1}^{N} K\left(\frac{x - X_i}{h}\right) , \tag{7}$$

wobei $K(y)$ eine zu wählende Kernfunktion und h die Bandbreite ist. Schreiben Sie ein entsprechendes Programm, das einen Datensatz einliest und eine geschätzte PDF ausgibt. Verwenden Sie als Kernfunktion den Epanechnikov-Kern

$$K(x) = \begin{cases} \frac{3}{4}(1 - x^2) ; & |x| < 1 \\ 0 ; & \text{sonst} \end{cases} . \tag{8}$$

Schätzen Sie mit Ihrem Programm die PDF des Datensatzes `1m_radio.dat`. Welche Bandbreite erscheint Ihnen sinnvoll? Überprüfen Sie, ob die geschätzte Verteilung korrekt normiert ist.