Stochastik-Praktikum **WS 2016/17**

Prof. Dr. Matthias Birkner

Dr. Iulia Dahmer/ M.Sc. Frederik Klement

Blatt 8

Aufgabe 1 (1 + 2 Punkte)

Wir wollen mit Hilfe eines Monte-Carlo-Verfahrens folgendes Integral bestimmen:

$$\int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1}\frac{1}{\sqrt{x^2+y^2+z^2}}\,dx\,dy\,dz$$

.

a) Verwenden Sie zunächst ein naives Monte-Carlo-Verfahren, indem Sie für x, y und z drei unabhängige U_x, U_y, U_z auf $[-1, 1]$ uniformverteilte Zufallsvariablen verwenden. Simulieren Sie

$$X = f(U_x, U_y, U_z) = \frac{1}{\sqrt{U_x^2 + U_y^2 + U_z^2}}$$

$n = 10^5$ mal und schätze Sie den Wert des Integrals mit dem empirischen Mittelwert Ihrer Simulationen. Schätzen Sie die den Standardabweichung ihres Schätzers indem Sie `sd` verwenden und das Ergebnis durch $\sqrt{n}$ teilen.

b) Die Varianz unseres Monte-Carlo-Schätzers ist gegeben durch $\frac{1}{n}Var(X)$. Wir können also die Varianz unseres Schätzer verringern, indem wir X durch eine neue Zufallsvariable Y mit $E[Y] = E[X]$ und $Var(Y) < Var(X)$ ersetzen. Eine Methode hierfür ist die der **Kontrollvariable**.
Für diese benötigen wir eine mit X positiv korrelierte Zufallsvariable Z. Setzen wir

$$Y := X - \frac{cov(X, Z)}{Var(Z)}(Z - E[Z]) \tag{1}$$

dann gilt, $E[Y] = E[X]$ und $Var(Y) = Var(X) - cov(X, Z)/Var(Z)$. Wegen $cov(X, Z) > 0$ ist $Var(Y)$ dann kleiner als $Var(X)$.
Verwenden Sie die Methode der **Kontrollvariable** mit

$$Z = exp(-U_x^2 - U_y^2 - U_z^2), \tag{2}$$

wobei U_x, U_y, U_z jeweils dieselben uniformverteilten Zufallsvariablen sind mit denen Sie auch X erzeugt hatten. Erzeugen Sie $n = 10^5$ Simulationen von (X, Z). Schätzen Sie mit Hilfe ihrer Simulationen $E[Z]$, $Var(Z)$ und $cov(X, Z)$ empirisch und verwenden Sie diese Schätzungen, um Y zu berechnen. Bestimmen Sie den so gewonnen empirischen Mittelwert von Y und schätzen Sie Standardabweichung wie in **a)**.

Aufgabe 2 (1 + 1 + 1 Punkte)
In der Vorlesung haben Sie das starke Gesetz der großen Zahlen für eine Folge von unabhängigen Zufallsvariablen kennen gelernt. Aber das starke Gesetz kann auch dann noch gelten, wenn die Zufallsvariablen gar nicht unabhängig sind. Um dies zu demonstrieren, betrachten wir zuerst eine Folge von unabhängigen bernoulliverteilten Zufallsvariablen $(B_i)_{i\in\mathbb{N}}$ mit Erfolgswahrscheinlichkeit $p \in (0,1)$. Sei nun $x = (x_1, x_2, .., x_k) \in \{0,1\}$ ein festes Muster von Nullen und Einsen. Wir definieren die Folge $(M_i^x)_{i\in\mathbb{N}}$ durch

$$M_i^x = \begin{cases} 1, \text{falls } B_{i+j-1} = x_j \text{für } i \in \{1,2,...,k\} \\ 0, \text{sonst.} \end{cases} \tag{3}$$

M_i^x ist also genau dann 1, wenn die $B_i, ..., B_{i+k}$ mit $x_1, ..., x_k$ übereinstimmen. Offensichtlich besitzen die M_x^i dieselbe Verteilung, Sie sind aber nicht alle unabhängig voneinander.

a) Schreiben Sie eine Funktion, welche für ein gegebenes Muster x die Wahrscheinlichkeit exakt $E[M_x^1] = P(M_x^1 = 1)$ berechnet, nicht schätzt.

b) Schreiben Sie eine Funktion, welche n bernoulliverteilten Zufallsvariablen $(B_i)_{i=1}^n$ mit Erfolgswahrscheinlichkeit $p \in (0,1)$ simuliert und dann für ein gegebenes Muster $x \in \{0,1\}^k$ zählt, wie häufig dieses auftritt, bzw. die Zahl der Einsen in $(M_i^x)_{i=1}^{n-k}$ zählt.

c) Schätzen Sie für $n = 10^5$ die relative Häufigkeit der folgenden Muster mit Hilfe von **b)**:

1.) x=(1,1)
2.) x=(1,1,1)
3.) x=(0,0,0,0,1,1,1,1)
4.) x=(1,0,1,0,1,0,1,0)

für die Erfolgswahrscheinlichkeiten $p = 0.5$ und $p = 0.75$. Vergleichen Sie die relative Häufigkeit der Muster mit den Wahrscheinlichkeiten aus **a)**.

Aufgabe 3 (1 + 1 Punkte)
Eine einfaches Populationsmodell ist durch einen sogenannten Galton-Watson-Prozess $(X_k)_{k\in\mathbb{N}}$ gegeben, wobei X_k die Zahl der Individuen in der k-ten Generation angibt. Jedes Individuen einer Generation besitzt eine zufällige Anzahl von Nachkommen und die Gesamtzahl aller Nachkommen bildet die nächste Generation. Wenn wir mit $X_0 = k$ startet, können wir dies mathematisch so ausdrücken:

$$X_{k+1} = \sum_{i=1}^{X_k} N_{i,k}. \tag{4}$$

Zusätzlich nehmen wir an, dass es sich bei $(N_{i,k})_{i,k\in\mathbb{N}}$ um unabhängig, identisch verteilte Zufallsvariablen handelt mit $N_{1,1} \sim Poi(\lambda)$. Die Zahl der Nachkommen eines Individuums ist somit jeweils immer unabhängig vom Rest der Population.

a) Schreiben Sie eine Funktion, welche einen Galton-Watson-Prozess der mit k Individuen startet für eine feste Zahl von Generationen N simuliert. Simulieren Sie 50 Galton-Watson-Prozesse jeweils für 40 Generationen und plotten Sie die zeitlichen Verläufe der Populationsgrößen gemeinsam in einen Plot. Tuen Sie dies jeweils für $\lambda = 0.9, 1, 1.1, 1.5$ und $k = 20, 10, 1, 1$, und zeichnen Sie die vier Plots auf eine Seite. Verwenden Sie für die Plots `ylim=c(0,50)`.

b) Der Galton-Watson-Prozess besitzt eine Art zufälliges Gesetz der großen Zahlen. Wenn wir $W_k = \frac{1}{\lambda^k} X_k$ setzen, so konvergiert W_k fast sicher gegen eine zufällige Zahl, welche wir mit W_∞ bezeichnen wollen. Simulieren Sie dazu 1000 mal $(W_k)_{k \in \mathbb{N}}$ für 80 Generationen. Plotten Sie zuerst den zeitlichen Verlauf von $(W_k)_{k=1}^{80}$ der ersten 30 Simulationen in ein Plot. Plotten Sie dann ein Histogramm mit den Werten von W_{80} aller 1000 Simulationen. Schlussendlich plotten Sie in einen dritten Plot nochmal den zeitlichen Verlauf von $(W_k)_{k=1}^{80}$ aller 1000 Simulationen. Tun Sie dies jeweils für die Parameter $\lambda = 0.9, 1, 1.2$ und $k = 1, 1, 1$. Zeichnen Sie alle 9 Plots auf eine Seite mit `par(mfrow=c(3,3))`.

Hinweis: Sollten Sie alles richtig gemacht haben, sollten Sie sehen, dass auf langer Sicht der Prozess für $E[N_{1,1}] = \lambda < 1$ ausstirbt. Dasselbe passiert für $\lambda = 1$, nur etwas langsamer. Für $\lambda > 1$ jedoch gibt es eine positive Wahrscheinlichkeit, dass der Prozess niemals ausstirbt. Die Phänomen hängt übrigens nicht von der Art der Nachkommensverteilung ab, $N_{1,1}$ könnte z.B. geometrisch verteilt sein. Entscheidend ist der Erwartungswert von $E[N_{1,1}]$. Für $E[N_{1,1}]$ stirbt der Prozess immer aus. Auch die Konvergenz von $(W_k)_{k \in \mathbb{N}}$ aus **b)** tritt für beliebige Verteilungen von $N_{1,1}$ auf.

Abgabe:19.12.16