Stochastik-Praktikum **WS 2016/17**

Prof. Dr. Matthias Birkner

Dr. Iulia Dahmer / M.Sc. Frederik Klement

Blatt 7

Aufgabe 1 (1 + 1 Punkte + 1 Extrapunkt) (Box-Muller-Methode)

Die Box-Muller-Methode, ist ein Verfahren, um aus auf $[0, 1]$ gleichverteilten, unabhängiger Zufallsvariablen U, V einen zweidimensionalen standardnormalverteilte Zufallsvektor $Z = (X, Y)$ zu erzeugen, wobei

$$X = \sqrt{-2\log(U)}\cos(2\pi V) \text{ und } Y = \sqrt{-2\log(U)}\sin(2\pi V).$$

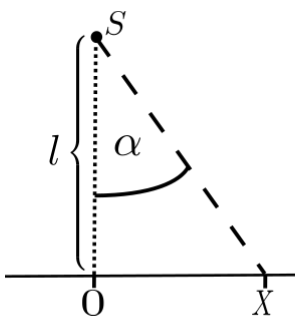
a) Erstellen Sie eine Funktion `box_muller(n)`, die unter Verwendung der Box-Muller-Methode n unabhängige zweidimensionale standardnormalverteilte Zufallsvektoren Z erzeugt.

b) Simulieren Sie mit Hilfe der Funktion aus a) 10^5 Paare (X, Y) unabhängiger normalverteilter Zufallsvariablen. Erstellen Sie ein Histogramm mit den Simulationen von X und ein weiteres mit denen von Y. Um die Abhängigkeit zwischen X und Y zu untersuchen, erstellen Sie ein Histogramm aller Simulationen von X mit $Y < 0$ und ein Histogramm aller Simulationen von X mit $Y \geq 0$. Zeichnen Sie alle vier Histogramme auf eine Seite.

Extra: Erstellen Sie einen Konturplot Ihrer Simulationen. Verwenden Sie hierfür ein Gitter mit Seitenlänge 0.3 und zählen Sie, wie viele Stichproben sich jeweils in den Quadraten zwischen den Gitterpunkten befinden. Speichern Sie Ihr Ergebnis in eine Matrix und übergeben Sie diese Matrix an `contour`. Bei der Erstellung der Matrix sind die Befehle `findInterval` und `table` sehr nützlich.

Aufgabe 2 (2 Punkte)

Wir wollen folgendes Zufallsexperiment simulieren. Wir denken uns die reellen Zahlen als eine horizontale Linie. Weiter denken wir uns einen Punkt S, der sich senkrecht über der 0 mit einem vertikalen Abstand von $l > 0$ befindet. Von diesem Punkt geht ein Strahl aus, wobei der Winkel α des Strahls zum Lot auf die 0 zufällig und uniform über $(-\frac{\pi}{2}, \frac{\pi}{2})$ verteilt ist. Mit X bezeichnen wir die Position des Schnittpunktes zwischen Strahls und der Linie der reellen Zahlen (siehe Bild).



Simulieren Sie das obige Experiment 10^6 mal für den Fall $l = 0.5$. Erstellen Sie ein Histogramm mit den relativen Häufigkeiten, wobei nur die Werte im Intervall $[-10, 10]$ betrachtet werden sollen. Verwenden Sie `breaks=200`. Mit der Transformationsformel für Dichten **(Beobachtung 1.36)** aus der Vorlesung kann man zeigen, dass der Schnittpunkt X bei einer Lotlänge von $l > 0$ eine stetige Zufallsvariable mit folgender Dichte ist

$$f_X(x) = \frac{1}{\pi(l + \frac{x^2}{l})}.$$

Durch das Abschneiden verändert sich die Dichte zu

$$f_{\tilde{X}}(x) = \frac{1}{M} \frac{1}{\pi(l + \frac{x^2}{l})} \mathbf{1}_{[-10,10]}(x),$$

mit $M := \int_{-10}^{10} f_X(x)dx$. Ergänzen Sie das Histogramm um die passende Dichte. Sie können die Funktionen `dcauchy` und `pcauchy` verwenden.

Hinweis: Die Cauchy-Verteilung ist eine Heavy-Tailed-Verteilung, dies bedeutet, dass die Dichte für extreme hohe bzw. niedrige Werte nur sehr langsam gegen 0 konvergiert. Wenn Sie also 10^6 Stichproben erzeugen, befinden sich darunter mit hoher Wahrscheinlichkeit einige wenige sehr starke Ausreißer. Unter den Standardeinstellungen von `hist` führt dies zu unschönen Histogrammen, daher schneiden wir die Ausreißer ab. An dieser Stelle laden wir Sie dazu ein, sich Gedanken zumachen, warum man die theoretische Dichte in der obigen Art und Weise modifizieren muss, damit sie wieder zu unserem Histogramm passt.

Aufgabe 3 (2 Punkte + 1 Extrapunkt)
In Aufgabe 1 von Blatt 3 haben wir die Inversionmethode kennengelernt, um Zufallsvariablen zu erzeugen. Hierbei mussten wir zuvor die entsprechende Verteilungsfunktion berechnen. Letzteres ist nicht immer möglich bzw. sehr aufwendig. In Situationen, in denen wir zwar nicht die Verteilungsfunktion F kennen jedoch die Dichte f, können wir auf die Verwerfungsmethode zurückgreifen.
Für die Verwerfungsmethode brauchen wir eine zweite Dichte g und eine Konstante $C \geq 1$, so dass für alle $x \in \mathbb{R}$ gilt

$$f(x) \leq Cg(x).$$

Um nun eine Zufallsvariable mit Dichte f zu erhalten, erzeugen wir zunächst eine Zufallsvariable Y mit Dichte g und eine auf $[0, 1]$ uniformverteilte Zufallsvariable U. Wir akzeptieren die Stichprobe und geben den Wert Y zurück, falls gilt

$$U \leq \frac{f(Y)}{Cg(Y)},$$

ansonsten starten wir die Prozedure von vorne.

Verwenden Sie die Verwerfungsmethode um Stichproben der Verteilung $\beta(4, 7)$ zu erzeugen. Verwenden Sie für g hierbei die Uniformverteilung auf $[0, 1]$ und die Konstanten $C_1 = 3$ und $C_2 = 5$. Erzeugen Sie für beide Konstanten jeweils 10^4 Zahlen und erstellen Sie ein Histogramm, ergänzen Sie die passende theoretische Dichte, benutzen Sie `dbeta`, und bestimmen Sie, wie viele uniformverteilte Zufallszahlen U Sie im Schnitt pro Simulation erzeugen müssen.

Extra: Eine andere Möglichkeit Stichproben mit Verteilung $\beta(n, k+1)$ wobei $n, k \in \mathbb{N}$ zu erzeugen ist die Ordnungsstatistik uniformverteilte Zufallsvariablen. Wenn $U_1, ..., U_{n+k}$ unabhängige auf $[0, 1]$ uniformverteilte Zufallsvariablen sind, und es sich bei $U_{[1]}, ..., U_{[n+k]}$ um diesselben Zahlen aufsteigend nach der Größe geordnet handelt, dann besitzt $U_{[n]}$ die Verteilung $\beta(n, k+1)$. Nutzen Sie diesen Fakt, um 10^4 Stichproben zu erzeugen und erstellen Sie ein Histogramm mit der passenden theoretischen Dichte für $\beta(3, 2)$.

Abgabe: 10 Uhr, 12.11.16