Wahrscheinlichkeitstheorie I WS 06/07 Serie 13

1. (Charakteristische Funktionen).

a) Die *Doppelexponentialverteilung* ist die Wahrscheinlichkeitsverteilung μ auf $\mathbb{R}$ mit Dichte $f(x) = \frac{1}{2}e^{-|x|}$. Zeige : Die charakteristische Funktion von μ ist $\varphi(t) = \frac{1}{1+t^2}$.

b) Folgere, daß $\varphi(t) = e^{-|t|}$ die charakteristische Funktion der Cauchyverteilung ist (Dichte $f(x) = \frac{1}{\pi(1+x^2)}$).

c) Zeige : Sind $X_1, X_2, \ldots$ unabhängige Cauchyverteilte Zufallsvariablen, dann sind die empirischen Mittelwerte $\frac{1}{n}\sum_{i=1}^{n} X_i$, $n \in \mathbb{N}$, auch Cauchyverteilt. Warum ist dies kein Widerspruch zum Gesetz der großen Zahlen ?

2. Unter 3000 Geburten werden in einer Klinik 1578 Jungen gezählt. Würden Sie aufgrund dieses Ergebnisses mit einer Sicherheit von 95% an der Hypothese festhalten wollen, daß die Wahrscheinlichkeit für die Geburt eines Jungen gleich $\frac{1}{2}$ ist?

3. (Abfall der Entropie bei Markovketten). Sei P eine stochastische Matrix mit invarianter Verteilung ν über einem abzählbaren Zustandsraum S. Sei μ eine beliebige Wahrscheinlichkeitsverteilung auf S. Zeige :

$$H(\mu P|\nu) \leq H(\mu|\nu) \ .$$

Folgere, daß die relative Entropie der Verteilung einer Markovkette zur Zeit n bezüglich einer invarianten Verteilung monoton fällt.

4. (Ratenfunktionen für große Abweichungen). Seien $X_1, X_2, \ldots$ i.i.d. Zufallsvariablen mit den unten angegebenen Verteilungen. Berechne jeweils die momentenerzeugenden Funktionen $M(t)$, und skizziere die Graphen von $\Lambda(t) = \log M(t)$, und von der Ratenfunktion I im Satz von Chernoff. Zeige, daß I die angegebene Form hat. Erkläre den Verlauf von I anschaulich.

a) Bernoulliverteilung mit Parameter $p \in (0,1)$:
$I(a) = a\log\frac{a}{p} + (1-a)\log\frac{1-a}{1-p}$ für $a \in [0,1]$, $\quad I(a) = \infty$ für $a \notin [0,1]$.

b) Exponentialverteilung mit Parameter $\lambda > 0$:
$I(a) = \lambda a - 1 - \log(\lambda a)$ für $a > 0$, $\quad I(a) = \infty$ für $a \leq 0$.

5. (Acceptance-Rejection Sampling). Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Seien μ, ν Wahrscheinlichkeitsverteilungen auf $(S, \mathcal{S})$, $\mu \ll \nu$ mit beschränkter Dichte ρ:

$$\mu(B) = \int_B \rho(x)\nu(dx), \quad \rho(x) \leq C, \quad C \in (1, \infty).$$

a) Sei $X : \Omega \to S$ eine Zufallsvariable mit Verteilung ν und $U : \Omega \to (0,1)$ eine hiervon unabhängige auf $(0,1)$ gleichverteilte Zufallsvariable. Zeige, dass dann gilt

$$\mu(B) = P[X \in B | U \leq \frac{\rho(X)}{C}] \quad \forall\, B \in \mathcal{S}\,,$$

d.h. μ ist bedingte Verteilung von X gegeben $U \leq \frac{\rho(X)}{C}$.

b) Seien $X_1, X_2, \ldots : \Omega \to S$, $U_1, U_2, \ldots : \Omega \to (0,1)$ unabhängig unter P mit Verteilung ν bzw. $\lambda_{(0,1)}$. Zeige: Dann ist

$$T \;:=\; \min \left\{ n \in \mathbb{N} \,\middle|\, U_n \leq \frac{\rho(X_n)}{C} \right\} \qquad (\textit{Abbruchzeit})$$

P-f.s. endlich, und die (für P-f.a. ω eindeutig definierte) Zufallsvariable

$$Y(\omega) := X_{T(\omega)}(\omega)$$

hat Verteilung μ.

c) Beschreibe einen Algorithmus zur Simulation von Stichproben von μ, wenn wir bereits Stichproben von ν simulieren können. Wie groß ist die mittlere Laufzeit des Algorithmus ? Wie sollte man C wählen ?

6. Seien μ, ν Wahrscheinlichkeitsverteilungen auf $(S, \mathcal{S})$. Betrachte eine Folge von Ereignissen $A_n \subseteq S^n$ mit $\liminf \mu^n(A_n) > 0$. Zeige :

$$\liminf \frac{1}{n} \log(\nu^n(A_n)) \geq -H(\mu|\nu).$$