Prof. Dr. Uwe Küchler
Dr. Renate Winkler
Institut für Mathematik

Sommersemester 2007

Stochastik I

Lösungsansätze zur 5. Zusatzübung

1) Es sei $x = (x_1, x_2, \ldots, x_n)^T \in R_n$ und $p = (p_1, \ldots, p_n)^T \in R_n^+$, $\sum_{k=1}^{n} p_k = 1$. Ist f eine konvexe Funktion auf einem Intervall (a, b) mit $x_k \in (a, b)$, $k = 1, \ldots, n$, so gilt

$$f\left(\sum_{k=1}^{n} p_k x_k\right) \leq \sum_{k=1}^{n} p_k f(x_k). \qquad (*)$$

Man beweise $(*)$.

Lösung: Der Beweis erfolgt induktiv über n. Der Induktionsanfang ist bei $n = 2$ per Definition der Konvexität gegeben. Nun gelte $(*)$ für n, dann gilt $(*)$ auch für $n + 1$, denn mit $q_n = p_n + p_{n+1}$, $\quad y_n = (p_n x_n + p_{n+1} x_{n+1})/q_n$ ergibt sich

$$f(\sum_{k=1}^{n+1} p_k x_k) = f(\sum_{k=1}^{n-1} p_k x_k + q_n y_n) \leq_{IV} \sum_{k=1}^{n-1} p_k f(x_k) + q_n f(y_n) \leq_{IA} \sum_{k=1}^{n+1} p_k f(x_k)$$

2) Es sei Ω eine abzählbare Menge und P eine Wahrscheinlichkeitsverteilung auf Ω. Durch

$$H(P) := -\sum_{\omega \in \Omega} P(\{\omega\}) ln P(\{\omega\})$$

ist die sogenannte Entropie von P definiert.

a) Man zeige, dass $H(P) \geq 0$ gilt. Wann gilt $H(P) = 0$?

b) Es sei $\Omega = \{1, 2, \ldots, n\}$ und Q die gleichmäßige Verteilung auf Ω. Man zeige, dass für jede andere Verteilung P auf Ω gilt $H(P) < H(Q)$.

Lösung: a) Sei $\Omega = \{\omega_1, \omega_2, \ldots\}$ und bezeichne $p_i = P(\{\omega_i\})$. O.B.d.A. gelte $p_i > 0$ für alle i. Dann ist

$H(P) = -\sum_i p_i \ln p_i = \sum_i p_i(-\ln p_i) \geq 0 \quad$ da $\quad p_i \in (0, 1]$ und

$H(P) = 0$ gdw. $p_i(-\ln p_i) = 0$ für alle i gdw. für alle i entweder $p_i = 0$ oder $p_i = 1$ gilt gdw. nur ein Element aus Ω positive Wahrscheinlichkeit, und zwar Wahrscheinlichkeit 1 hat.

b) Es gilt $Q(\{i\}) = \frac{1}{n}$ für $i = 1, \ldots, n$ und $H(Q) = \sum_{i=1}^{n} \frac{1}{n}(-\ln \frac{1}{n}) = n\frac{1}{n} \ln n = \ln n$.

Unter Benutzung des Fakts, dass ln eine konkave Funktion ist (-ln ist konvex) schließt man mit Hilfe von Übungsaufgabe 1 für jede Verteilung P mit $P(i) = p_i$:

$$H(P) = \sum_{i=1}^{n} p_i(-\ln p_i) = \sum_{i=1}^{n} p_i \ln \frac{1}{p_i} \leq \ln(\sum_{i=1}^{n} p_i \frac{1}{p_i}) = \ln n = H(Q)$$

3) Es seien $n \geq 1$ und $p \in (0,1)$ Die Zufallsgröße S_n besitze eine Binomialverteilung mit den Parametern n und p.

a) Man zeige, dass für alle $a > 0$ gilt

$$P\left(|\frac{S_n}{n} - p| \geq a \right) \leq \frac{1}{4n\, a^2}$$

b) Es sei f eine stetige Funktion von $[0,1]$ in R_1. Man zeige, dass die Funktion $f_n(p) := \mathrm{E} f(\frac{S_n}{n})$ erweitert durch $f_n(0) = f(0), f_n(1) = f(1)$ für jedes n ein Polynom in p ist. Beweisen Sie, dass (f_n) gleichmäßig bez. $p \in [0,1]$ gegen f konvergiert (Weierstraßsches Approximationstheorem).

Lösung: a) Da S_n binomialverteilt mit den Parametern n, p ist, gilt $\mathrm{E}\, S_n = n\, p$ und $\mathrm{Var}\, S_n = n\, p(1-p)$. Für die Zufallsgröße $Y_n := \frac{1}{n} S_n$ ergibt sich $\mathrm{E}\, Y_n = p$ und $\mathrm{Var}\, Y_n = \frac{1}{n}\, p(1-p)$. Mit der Tschebyschevschen Ungleichung folgt

$$P\left(|Y_n - p| \geq a \right) \leq \frac{\mathrm{Var} Y_n}{a^2} = \frac{p(1-p)}{n\, a^2} \leq \frac{1}{4n\, a^2}\,.$$

Lösung: b) Wir schreiben der Deutlichkeit halber P_p, E_p für die Wahrscheinlichkeitsverteilung mit dem Parameter p und die Erwartungswerte unter dieser Verteilung. $Y_n := \frac{1}{n} S_n$ nimmt die Werte $\frac{i}{n}$ jeweils mit Wahrscheinlichkeit $\binom{i}{n} p^i (1-p)^{n-i}$ an. Demzufolge ist

$$f_n(p) := \mathrm{E}_p f(Y_n) = \sum_{i=0}^{n} f\left(\frac{i}{n}\right) \binom{i}{n} p^i (1-p)^{n-i}$$

ein Polynom n-ten Grades in p, genannt Bernstein-Polynom.
f ist stetig auf dem abgeschlossenen Intervall $[0,1]$, also ist f gleichmäßig stetig und beschränkt durch eine Konstante $C > 0$.
Zu beliebigem $\varepsilon > 0$ existiert also ein $\delta = \delta(\varepsilon/2)$, so dass $|f(x) - f(y)| < \varepsilon/2$ für alle x, y mit $|x - y| < \delta$. Wir teilen Ω in $\Omega_\delta := \{|Y_n - p| < \delta\}$ und $\Omega \setminus \Omega_\delta$. Nach a) gilt $P(\Omega \setminus \Omega_\delta) = P(|Y_n - p| \geq \delta) \leq \frac{1}{4n\, \delta^2}$ und wir schätzen ab:

$$\begin{aligned} |f_n(p) - f(p)| &= |\mathrm{E}_p f(Y_n) - f(p)| \ \leq\ \mathrm{E}_p |f(Y_n) - f(p)| \\ &= \underbrace{\int_{\Omega_\delta} |f(Y_n) - f(p)| dP_p}_{\leq \varepsilon/2} + \underbrace{\int_{\Omega \setminus \Omega_\delta} |f(Y_n) - f(p)| dP_p}_{\leq 2C/(4n\, \delta^2)} \\ &\leq \varepsilon \qquad \text{für hinreichend großes } n \geq n(\varepsilon). \end{aligned}$$

Diese Abschätzung gilt gleichmäßig für alle $p \in [0,1]$.