Institut für angewandte Mathematik
Wintersemester 2013/14
Andreas Eberle, Lisa Hartung / Patrick Müller

universitätbonn

Musterlösung Klausur „Einführung in die W'theorie"

1. (Zufallsvariablen und ihre Verteilung)

a) Was ist eine reellwertige Zufallsvariable ? Wie sind die Verteilung und die Verteilungsfunktion definiert ? [*5 Pkt.*]

b) Skizzieren Sie die Graphen der Verteilungsfunktionen von Zufallsvariablen mit den folgenden Verteilungen:

$$\text{(i) } X \sim \mathrm{N}(3,9), \qquad \text{(ii) } T \sim \mathrm{Exp}(5), \qquad \text{(iii) } \max(Z,0) \text{ mit } Z \sim \mathrm{N}(0,1).$$

[*6*]

c) Seien S und T unabhängige, zum Parameter $\lambda = 1$ exponentialverteilte Zufallsvariablen. Berechnen Sie die Verteilungsfunktionen und Dichten der Zufallsvariablen

$$\text{(i) } 1-\exp(-T), \qquad \text{(ii) } -T, \qquad \text{(iii) } S-T.$$

[*11*]

d) Zeigen Sie: Für eine nicht-negative Zufallsvariable X gilt:

$$E[X] = \int_0^\infty (1-F_X(c))\,dc.$$

[*5*]

e) Die gemeinsame Verteilung der reellwertigen Zufallsvariablen X und Y sei absolutstetig mit Dichte

$$f(x,y) = \begin{cases} ax & \text{für } 0<y<x<1, \\ 0 & \text{sonst.} \end{cases}$$

Bestimmen Sie die Konstante a, und berechnen Sie die Dichte von X, sowie die bedingte Dichte von Y gegeben X. Sind X und Y unabhängig ? [*9*]

Lösung:

a) **Reellwertige Zufallsvariable:** Eine Abbildung $X : \Omega \to \mathbb{R}$ heisst *messbar*, falls

$$(M) \qquad X^{-1}(B) \in \mathcal{A} \qquad \text{für alle } B \in \mathcal{B}(\mathbb{R}).$$

Eine *reellwertige Zufallsvariable* ist eine auf einem Wahrscheinlichkeitsraum definierte messbare Abbildung nach $\mathbb{R}$.

Sei (Ω, A, P) ein Wahrscheinlichkeitsraum. Eine Abbildung $X : \Omega \to \mathbb{R}$ ist genau dann eine Zufallsvariable bzgl. der Borelschen σ-Algebra, wenn

$$\{X \leq c\} = \{\omega \in \Omega \mid X(\omega) \leq c\} \in \mathcal{A} \quad \forall\, c \in \mathbb{R}, \qquad \text{bzw. wenn}$$
$$\{X < c\} = \{\omega \in \Omega \mid X(\omega) < c\} \in \mathcal{A} \quad \forall\, c \in \mathbb{R}.$$

Verteilung: μ_X ist die Wahrscheinlichkeitsverteilung auf $\mathcal{B}(\mathbb{R})$ mit

$$\mu_X[B] \;=\; P[X \in B] \;=\; P[X^{-1}(B)].$$

Verteilungsfunktion: Die Funktion $F_X : \mathbb{R} \to [0,1]$,

$$F_X(c) := P[X \leq c] = \mu_X[(-\infty, c]]$$

heisst *Verteilungsfunktion (distribution function)* der Zufallsvariable $X : \Omega \to \mathbb{R}$ bzw. der Wahrscheinlichkeitsverteilung μ_X auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

b) (i) Verteilungsfunktion von $N(3,9)$: Wichtig: F(3)=1/2, Standardabweichung $\sigma = 3$, schnelle Konvergenz gegen 0 bzw. 1 für $x \to -\infty$ bzw. $x \to \infty$.

(ii) Verteilungsfunktion von Exp(5): Wichtig: $F(c) = 0$ für $c \leq 0$, $F(1/5) = 1 - 1/e$, exponentielle Konvergenz gegen 1 für $x \to \infty$.

(iii) Verteilungsfunktion von $\max(Z, 0)$, $Z \sim N(0,1)$: Wichtig: $F(c) = 0$ für $c \leq 0$, $F(0) = 1/2$, $F(x) = 1/2 + \Phi(c)/2$ für $c > 0$, d.h. schnelle Konvergenz gegen 1 für $x >> 1$.

c) (i) Für $x \in [0,1]$ gilt

$$F(x) = P(1 - e^{-T} \leq x) = P(e^{-T} \geq 1 - x) \tag{1}$$
$$= P(T \leq -\log(1-x)) = 1 - (1-x) = x \tag{2}$$

Für $x < 0$ ist $F(x) = 0$, und für $x > 1$ gilt $F(x) = 1$, also

$$F(x) \;=\; x^+ \wedge 1 \qquad \text{(Gleichverteilung).}$$

Insbesondere ist F absolutstetig mit Dichte

$$f(x) \;=\; F'(x) \;=\; 1_{[0,1]}(x) \qquad \text{für fast alle } x. \tag{3}$$

(ii) Für $x \leq 0$ gilt:

$$F(x) = P(-T \leq x) = P(T \geq -x) = e^x. \tag{4}$$

Für $x > 0$ ist $F(x) = 1$, also

$$F(x) \;=\; \exp(x) \wedge 1.$$

Insbesondere ist F absolutstetig mit Dichte

$$f(x) = e^x 1_{(-\infty,0)}(x) \qquad \text{f.ü.} \tag{5}$$

(iii) Für $x \in \mathbb{R}$ ergibt sich die Dichte von $S - T$ als Faltung der Dichten der **unabhängigen** Zufallsvariablen S und $-T$:

$$
\begin{aligned}
f(x) &= \int_{-\infty}^{\infty} f_S(s)\, f_{-T}(x-s)\, ds \\
&= \int_{-\infty}^{\infty} e^{-s}\, e^{x-s} 1_{(0,\infty)}(s) 1_{(-\infty,0)}(x-s)\, ds \\
&= \int_{\max(x,0)}^{\infty} e^{x-2s}\, ds \;=\; \frac{1}{2} e^{x-2x^+} \;=\; \frac{1}{2} e^{-|x|}.
\end{aligned}
$$

Die Verteilungsfunktion ergibt sich durch Integrieren:

$$
F(c) = \int_{-\infty}^{c} f(x)\, dx \;=\; \begin{cases} \frac{1}{2} e^{-|c|} & \text{für } c \le 0, \\ 1 - \frac{1}{2} e^{-c} & \text{für } c \ge 0. \end{cases} \tag{6}
$$

d) Mithilfe des **Satzes von Fubini** erhalten wir

$$
\int_0^{\infty} (1 - F(c))\, dc = \int_0^{\infty} P(X > c) dc \;=\; \int_0^{\infty} E\left[1_{\{X>c\}}\right]\, dc \tag{7}
$$

$$
= E\left[\int_0^{\infty} 1_{\{X>c\}}\, dc\right] \;=\; E\left[\int_0^{X} dc\right] \;=\; E[X], \tag{8}
$$

wobei wir $X \ge 0$ benutzt haben !

e) Wegen

$$
\int\int f(x,y)\, dy\, dx \;=\; \int_0^1 \int_0^x ax\, dy\, dx \;=\; \int_0^1 ax^2\, dx \;=\; \frac{a}{3} \tag{9}
$$

muss man $a = 3$ wählen, um eine W'dichte zu erhalten.

Dichte von X: Es gilt fast überall

$$
f_X(x) \;=\; \int f(x,y)\, dy \;=\; \int_0^x ax\, dy \;=\; ax^2 \;=\; 3x^2. \tag{10}
$$

für $0 < x < 1$, und $f_X(x) = 0$ sonst.

Bedingte Dichte von Y gegeben X:

$$
f_{Y|X}(y|x) \;=\; \frac{f(x,y)}{f_X(x)} \;=\; \frac{3x\, 1_{(0,x)}(y)}{3x^2} \;=\; \frac{1}{x} 1_{(0,x)}(y) \tag{11}
$$

für $0 < x < 1$, andernfalls willkürlich definiert (verschiedene Versionen der bedingten Dichte).

Die Zufallsvariablen sind nicht unabhängig da die bedingte Dichte von x abhängt.

2. (Zentraler Grenzwertsatz)

a) Was bedeutet Konvergenz in Verteilung einer Folge $(X_n)_{n\in\mathbb{N}}$ von reellen Zufallsvariablen (Definition) ? Zeigen Sie anhand eines Beispiels, dass aus Konvergenz in Verteilung nicht die Konvergenz der Verteilungen in Variationsdistanz folgt. *[5 Pkt.]*

b) Formulieren und beweisen Sie den zentralen Grenzwertsatz für unabhängige, identisch verteilte, quadratintegrierbare Zufallsvariablen. Nötige Aussagen aus der Analysis sowie der Konvergenzsatz von Lévy können ohne Beweis vorausgesetzt werden - die verwendeten Aussagen sollten aber vollständig mit Voraussetzungen angegeben werden. *[15]*

c) Seien U eine auf $(0,1)$ gleichverteilte, und Z eine standardnormalverteilte Zufallsvariable.

(i) Zeigen Sie, dass $E[U] = 1/2$ und $\mathrm{Var}[U] = 1/12$ gilt. *[5]*

(ii) Berechnen Sie den Erwartungswert und die Varianz von $|Z|$. *[5]*

d) Sei S die Summe von 1200 unabhängigen, auf $(0,1)$ gleichverteilten Zufallsvariablen U_i, $i = 1, 2, \ldots, 1200$. Berechnen Sie die Wahrscheinlichkeit $P[S > 610]$ näherungsweise.

Hinweis: Für eine standardnormalverteilte Zufallsvariable Z gilt:

$$P[|Z| > 1] \approx 0,317, \qquad P[|Z| > 2] \approx 0,046, \quad \text{und} \quad P[|Z| > 3] \approx 0,0026.$$

[6]

Lösung:

a) **Konvergenz in Verteilung von Zufallsvariablen:** Eine Folge $(X_n)_{n\in\mathbb{N}}$ von Zufallsvariablen mit Werten in S *konvergiert in Verteilung* gegen eine Zufallsvariable X bzw. gegen die Verteilung von X, falls

$$\text{Verteilung}\,(X_n) \quad \xrightarrow{w} \quad \text{Verteilung}\,(X),$$

d.h. falls

$$E[f(X_n)] \quad \longrightarrow \quad E[f(X)] \qquad \text{für alle } f \in C_b(S) \text{ gilt.}$$

Gegenbeispiel: Seien $X_n = 1/n$ mit Wahrscheinlichkeit 1. Die Verteilungen sind $\delta_{1/n}$ und δ_0. Es gilt

$$\|\delta_{1/n} - \delta_0\|_{TV} \;\geq\; |\delta_{1/n}(0) - \delta_0(0)| = 1$$

für alle n, aber für $f \in C_b(S)$ gilt

$$E[f(X_n)] \;=\; f(1/n) \;\to\; f(0) \;=\; E[f(X)].$$

b) **Zentraler Grenzwertsatz:** Seien $X_1, X_2, ... \in \mathcal{L}^2(\Omega, \mathcal{A}, P)$ unabhängige, identisch verteilte Zufallsvariablen mit Varianz σ^2 und sei

$$S_n \;=\; X_1 + ... + X_n \;.$$

Dann konvergieren die Verteilungen der standardisierten Summen

$$\hat{S}_n \;=\; \frac{S_n - E[S_n]}{\sqrt{n}} \;=\; \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - E[X_i])$$

schwach gegen $N(0, \sigma^2)$.

Beweis des Satzes:
O.B.d.A. sei $E[X_i] = 0$, ansonsten betrachten wir die zentrierten Zufallsvariablen $\widetilde{X}_i := X_i - E[X_i]$. Nach dem Konvergenzsatz von Lévy genÃ¼gt es zu zeigen, dass die charakteristischen Funktionen der standardisierten Summen $\widehat{S}_n$ punktweise gegen die charakteristische Funktion der Normalverteilung $N(0, \sigma^2)$ konvergieren, d.h.

$$\phi_{\hat{S}_n}(t) \;\to\; \phi_{N(0,\sigma^2)}(t) \;=\; e^{-\frac{\sigma^2 t^2}{2}} \qquad \forall\, t \in \mathbb{R}. \tag{12}$$

Da die Zufallsvariablen X_i unabhängig, identisch verteilt und zentriert sind, gilt für $t \in \mathbb{R}$:

$$\phi_{\hat{S}_n}(t) \;\overset{E[S_n]=0}{=}\; \phi_{S_n}\left(\frac{t}{\sqrt{n}}\right) \;\overset{X_i \text{ iid}}{=}\; \left(\phi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n .$$

Aus $X_1 \in \mathcal{L}^2$ folgt $\phi_{X_1} \in C^2(\mathbb{R})$, und

$$\phi_{X_1}(t) \;=\; E[e^{itX_1}] \;=\; 1 + itE[X_1] + \frac{(it)^2}{2} E[X_1^2] + o(t^2) \;=\; 1 - \frac{t^2\sigma^2}{2} + o(t^2),$$

wobei o für eine Funktion $o : \mathbb{R}^+ \to \mathbb{C}$ mit $\lim_{\epsilon \downarrow 0} \frac{|o(\epsilon)|}{\epsilon} = 0$ steht. Damit erhalten wir:

$$\phi_{\hat{S}_n}(t) \;=\; \left(1 - \frac{t^2\sigma^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n .$$

Dieser Ausdruck konvergiert für $n \to \infty$ gegen $e^{-\frac{t^2\sigma^2}{2}}$, denn für komplexe Zahlen $z_i, w_i \in \mathbb{C}$ mit $|z_i|, |w_i| \leq 1$ gilt nach der Dreiecksungleichung

$$\begin{aligned} \left|\prod_{i=1}^n z_i - \prod_{i=1}^n w_i\right| \;&=\; |(z_1 - w_1)z_2 z_3 \cdots z_n + w_1(z_2 - w_2)z_3 z_4 \cdots z_n + \ldots + w_1 \cdots w_{n-1}(z_n - w_n)| \\ &\leq\; \sum_{i=1}^n |z_i - w_i|, \end{aligned}$$

und damit

$$\begin{aligned} \left|\phi_{\widehat{S}_n}(t) - \exp\left(-\frac{t^2\sigma^2}{2}\right)\right| \;&=\; \left|\left(1 - \frac{t^2\sigma^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n - \exp\left(-\frac{t^2\sigma^2}{2n}\right)^n\right| \\ &\leq\; n \cdot \left|1 - \frac{t^2\sigma^2}{2n} + o\left(\frac{t^2}{n}\right) - \exp\left(-\frac{t^2\sigma^2}{2n}\right)\right| . \end{aligned}$$

c) (i) Der Erwartungswert ist gegeben durch

$$E(U) = \int_0^1 x\,dx = \left[\frac{1}{2}x^2\right]_0^1 = \frac{1}{2}.$$

Für das zweite Moment erhält man

$$E(U^2) = \int_0^1 x^2\,dx = \left[\frac{1}{3}x^3\right]_0^1 = \frac{1}{3}.$$

Daher gilt für die Varianz

$$\operatorname{Var}(U) = E(X^2) - E(X)^2 = \frac{1}{12}.$$

(ii) Der Erwartungswert ist

$$E(|Z|) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |z| e^{-z^2/2} dz = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} z e^{-z^2/2} dz = \frac{2}{\sqrt{2\pi}} \left[-e^{-z^2/2}\right]_0^{\infty} = \sqrt{\frac{2}{\pi}}.$$

Das zweite Moment ist gleich

$$E(|Z|^2) = E(Z^2) = 1.$$

Daher ist die Varianz gleich

$$\operatorname{Var}(|Z^2|) = E(|Z|^2) - E(|Z|)^2 = 1 - \left(\sqrt{\frac{2}{\pi}}\right)^2 = 1 - \frac{2}{\pi}.$$

d) Nach Aufgabenteil c) (i) gilt

$$E(S) = 1200 \cdot E(U_1) = 600$$

und

$$\sigma^2 \;:=\; \operatorname{Var}(U_i) = \frac{1}{12}$$

für alle i. Damit gilt

$$P[S > 610] \;=\; P[S - E[S] > 10] \;=\; P\left[\frac{S - E[S]}{\sqrt{1200\sigma^2}} > 1\right].$$

Nach dem ZGS ist die Zufallsvariable auf der rechten Seite approximativ standardnormalverteilt. Daher erhalten wir näherungsweise

$$P[S > 610] \;\approx\; P(Z > 1) \;=\; \frac{1}{2}P(|Z| > 1) \;\approx\; 0.1585.$$

3. (Konvergenzbegriffe für Zufallsvariablen)

a) Seien $Z_n : \Omega \to \mathbb{R}$ $(n \in \mathbb{N})$ und $Z : \Omega \to \mathbb{R}$ reelle Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Wie sind die fast sichere Konvergenz, die stochastische Konvergenz, und die L^2-Konvergenz von $(Z_n)_{n\in\mathbb{N}}$ gegen Z definiert ? *[3 Pkt.]*

b) Sei $\alpha \in \mathbb{R}$. Es gelte

$$P[Z_n = n^\alpha] \;=\; \frac{1}{n}, \qquad P[Z_n = 0] \;=\; 1 - \frac{1}{n}$$

Zeigen Sie, dass (Z_n) genau dann in L^2 gegen 0 konvergiert, wenn $\alpha < 1/2$ ist. Wann konvergiert (Z_n) stochastisch gegen 0 ? *[6]*

c) Seien A_n $(n \in \mathbb{N})$ Ereignisse mit $\sum_{n=1}^\infty P[A_n] < \infty$. Zeigen Sie, dass mit Wahrscheinlichkeit 1 nur endlich viele der Ereignisse eintreten. *[5]*

d) Folgern Sie: Gilt für eine Folge (Z_n) von Zufallsvariablen

$$\sum_{n=1}^{\infty} P[|Z_n - Z| > \varepsilon] \;<\; \infty \qquad \text{für alle } \varepsilon > 0,$$

dann konvergiert (Z_n) fast sicher gegen Z. *[4]*

e) Seien $U_1, U_2, \ldots$ unabhängige, auf $(0,1)$ gleichverteilte Zufallsvariablen, und sei

$$Z_n \;=\; \min(U_1, \ldots, U_n).$$

Zeigen Sie:

(i) $Z_n \to 0$ P-fast sicher. *[5]*

(ii) Für die Zufallsvariablen $W_n := nZ_n$ gilt

$$P[W_n \le c] \;\longrightarrow\; \begin{cases} 1 - e^{-c} & \text{für } c > 0, \\ 0 & \text{für } c \le 0. \end{cases}$$

[5]

f) In einem Glücksspiel verdoppelt bzw. halbiert sich das Kapital des Spielers in jeder Runde mit Wahrscheinlichkeit p bzw. $1-p$, $p \in (0,1)$. Sei Z_n das Kapital nach n Runden bei Startkapital $Z_0 = 1$.

(i) Beschreiben Sie Z_n $(n = 0, 1, 2, \ldots)$ als Zufallsvariablen in einem geeigneten Modell. *[2]*

(ii) Für welche Werte von p konvergiert (Z_n) in L^2 gegen 0 ? *[3]*

(iii) Zeigen Sie, dass (Z_n) für $p < 1/2$ fast sicher gegen 0 konvergiert. *[3]*

Lösungen:

a) – *Fast sichere Konvergenz:*
Die Folge $(Y_n)_{n\in\mathbb{N}}$ konvergiert P-fast sicher gegen Y, falls gilt:

$$P\left[\lim_{n\to\infty} Y_n = Y\right] \quad = \quad P[\{\omega\in\Omega | Y_n(\omega)\to Y(\omega)\}] \quad = \quad 1.$$

– *Stochastische Konvergenz*:
Die Folge $(Y_n)_{n\in\mathbb{N}}$ konvergiert P-stochastisch gegen Y (Notation $Y_n \stackrel{P}{\to} Y$), falls

$$\lim_{n\to\infty} P[|Y_n - Y| > \epsilon] \quad = \quad 0 \qquad \text{für alle } \epsilon > 0 \text{ gilt.}$$

– *$\mathcal{L}^2$-Konvergenz* : $(Y_n)_{n\in\mathbb{N}}$ konvergiert in $\mathcal{L}^2(\Omega,\mathcal{A},P)$ gegen Y, falls

$$\lim_{n\to\infty} E[|Y_n - Y|^2] \quad = \quad 0.$$

b) *$\mathcal{L}^2$-Konvergenz:*

$$E[|Z_n - 0|^2] \;=\; n^{2\alpha}\frac{1}{n} \;=\; n^{2\alpha-1} \;\to\; 0 \qquad \text{gdw.} \qquad \alpha < \frac{1}{2} \tag{13}$$

Stochastische Konvergenz gegen 0 *gilt immer:*

$$P(|Z_n - 0| > \epsilon) \;=\; \begin{cases} 0 \text{ falls } \epsilon > n^\alpha \\ \frac{1}{n} \text{ sonst} \end{cases} \qquad \to 0 \text{ für } n\to\infty. \tag{14}$$

c) **1. Borel - Cantelli - Lemma:** Für Ereignisse $A_1, A_2, \ldots \in \mathcal{A}$ mit $\sum P[A_n] < \infty$ gilt:

$$P[\text{„unendlich viele der } A_n \text{ treten ein“}] = P\left[\bigcap_m \bigcup_{n\ge m} A_n\right] = 0.$$

Beweis: Da die Folge $\bigcup\limits_{n\ge m} A_n =: B_m$ von Ereignissen aus $\mathcal{A}$ **monoton fallend** ist, ergibt sich

$$\begin{aligned} P\left[\bigcap_m \bigcup_{n\ge m} A_n\right] &= P\left[\bigcap_m B_m\right] \\ &= \lim_{m\to\infty} P[B_m] \\ &= \lim_{m\to\infty} \underbrace{P\left[\bigcup_{n\ge m} A_n\right]}_{\le \sum_{n=m}^{\infty} P[A_n]} \\ &\le \liminf_{m\to\infty} \underbrace{\sum_{n=m}^{\infty} P[A_n]}_{\stackrel{m\to\infty}{\to} 0} = 0, \end{aligned}$$

da die Summe $\sum P[A_n]$ nach Voraussetzung konvergiert.

d) Aus *schneller stochastischer Konvergenz* folgt *fast sichere Konvergenz.*

Beweis: Wir können o.B.d.A. $Z = 0$ annehmen. Konvergiert Z_n schnell stochastisch gegen 0, dann gilt für alle $\varepsilon > 0$:

$$P[\limsup |Z_n| \leq \epsilon] \quad \geq \quad P[|Z_n| \geq \epsilon \text{ nur endlich oft}] \quad = \quad 1.$$

Es folgt

$$P[\limsup |Z_n| \neq 0] \quad = \quad P\left[\bigcup_{\epsilon \in \mathbb{Q}_+} \{\limsup |Z_n| > \epsilon\}\right] \quad = \quad 0.$$

e) (i) Wegen der **Unabhängigkeit** der U_i erhalten wir für $\varepsilon > 0$:

$$\begin{aligned} P[|Z_n| > \varepsilon] \quad &= \quad P[\min(U_1, \ldots, U_n) > \varepsilon] \\ &= \quad P[U_i > \varepsilon \ \forall\ 1 \leq i \leq n] \quad = \quad \prod_{i=1}^{n} P[U_i > \varepsilon] \quad = \quad (1-\varepsilon)^n \end{aligned}$$

Nach d) folgt die fast sichere Konvergenz von Z_n gegen 0, da die geometrische Reihe summierbar ist.

(ii) Für $c > 0$ gilt

$$P[W_n \leq c] = P[Z_n \leq \frac{c}{n}] = 1 - (1 - \frac{c}{n})^n \to 1 - e^{-c}, \tag{15}$$

für $c \leq 0$ ist $P[W_n \leq c] = 0$ für alle n.

f) (i) Es seien $A_k = \begin{cases} 2 \text{ mit W'keit } p \\ \frac{1}{2} \text{ mit W'keit } 1-p \end{cases}$ unabhängig, und $Z_n = \prod_{k=1}^n A_k$.

ODER

Es seien $X_k = \begin{cases} 1 \text{ mit WSK } p \\ -1 \text{ mit WSK } 1-p \end{cases}$ unabhängig, und $Z_n = \prod_{k=1}^n 2^{X_n}$

(ii) Wegen der Unabhängigkeit der A_i gilt

$$E[|Z_n - 0|^2] = \left(E[A_1^2]\right)^n = \left(2^2 p + \frac{1}{2^2}(1-p)\right)^n = \left(\frac{15}{4}p + \frac{1}{4}\right)^n \tag{16}$$

Dies konvergiert genau dann gegen Null, wenn $p \leq \frac{1}{5}$.

(iii) Mit der zweiten Darstellung und dem Wissen dass $S_n = \sum_{k=1}^n X_k$ ein klassischer Random walk ist, der für $p < 1/2$ fast sicher gegen $-\infty$ konvergiert, folgt

$$Z_n \quad = \quad 2^{S_n} \quad \to \quad 0$$

fast sicher für alle $p < 1/2$.

4. (Entropie)
Seien μ und ν Wahrscheinlichkeitsverteilungen auf einer abzählbaren Menge S.

a) Wie ist die Entropie $H(\mu)$ definiert ? Zeigen Sie:

$$H(\mu \otimes \nu) \;=\; H(\mu) + H(\nu) \qquad \text{und} \qquad H(\mu^n) = nH(\mu) \;\text{ für } n \in \mathbb{N}.$$

[8 Pkt.]

b) Berechnen Sie die Entropie der folgenden Wahrscheinlichkeitsverteilungen:

$$\text{(i) } \mu = \delta_a \;\; (a \in S), \qquad \text{(ii) } \mu = \mathcal{U}_S, \qquad \text{(iii) } \mu = \mathrm{Geom}(p) \;\; (0 < p < 1).$$

[9]

c) Seien $X_1, X_2, \ldots$ unabhängige Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ mit Verteilung μ. Geben Sie die Massenfunktion p_n von $(X_1, \ldots, X_n)$ explizit an, und zeigen Sie

$$\frac{1}{n} \log p_n(X_1, \ldots, X_n) \;\longrightarrow\; -H(\mu) \qquad P\text{-fast sicher.}$$

[6]

d) Sei $\varepsilon > 0$. Folgern Sie, dass die Wahrscheinlichkeiten der Mengen

$$B_{n,\varepsilon} \;:=\; \left\{ x \in S^n \;:\; e^{-n(H(\mu)+\varepsilon)} \leq p_n(x_1, \ldots, x_n) \leq e^{-n(H(\mu)-\varepsilon)} \right\}$$

unter dem Produktmaß μ^n für $n \to \infty$ gegen 1 konvergieren. *[4]*

e) Seien $A_n \subset S^n$ Mengen mit

$$\lim_{n\to\infty} P[(X_1, \ldots, X_n) \in A_n] \;=\; 1.$$

Zeigen Sie, dass für alle $\varepsilon > 0$

$$\lim_{n\to\infty} \sum_{A_n \cap B_{n,\varepsilon}} p_n(x_1, \ldots, x_n) \;=\; 1$$

gilt, und folgern Sie

$$\liminf_{n\to\infty} \frac{1}{n} \log |A_n| \;\geq\; H(\mu).$$

[9]

Lösung:

a) **Entropie einer diskreten Wahrscheinlichkeitsverteilung** μ auf S:

$$H(\mu) \quad := \quad - \sum_{\substack{x \in S \\ \mu(x) \neq 0}} \mu(x) \log \mu(x) \quad \in \; [0, \infty]$$

Faktorisierungseigenschaft:

$$H(\nu \otimes \mu) \;=\; H(\nu) + H(\mu).$$

Beweis: Nach Definition der Entropie gilt:

$$\begin{aligned} H(\nu \otimes \mu) &= \sum_{\substack{x,y \\ \nu(x)\mu(y)\neq 0}} \nu(x)\mu(y)\log(\nu(x)\mu(y)) \\ &= -\sum_{x:\nu(x)\neq 0} \nu(x)\log(\nu(x)) - \sum_{y:\mu(y)\neq 0} \mu(y)\log(\mu(y)) \\ &= H(\nu) + H(\mu). \end{aligned}$$

Beweis von $H(\mu^n) = nH(\mu)$: Induktion über n. Für $n = 1$ ist die Aussage klar. Nehme nun an das die Behauptung für n richtig ist. Dann gilt sie auch für $n+1$, da nach der eben gezeigten Faktorisierungseigenschaft und Induktionsvoraussetzung:

$$H(\mu^{n+1}) \;=\; H(\mu^n \otimes \mu) \;=\; H(\mu^n) + H(\mu) \;=\; (n+1)H(\mu).$$

b) (i) Nach Definition

$$H(\delta_a) = -\mu(a)\log(\mu(a)) = 0.$$

(ii) Nach Definition

$$H(\mathcal{U}_S) = -\sum_{x\in\mathcal{S}} \frac{1}{|S|}\log\left(\frac{1}{|S|}\right) = -\log\left(\frac{1}{|S|}\right) = \log(|S|).$$

(iii) Nach Definiton

$$\begin{aligned} H(\mathrm{Geom}(p)) &= -\sum_{n=1}^{\infty} p(1-p)^{n-1}\log\left(p(1-p)^{n-1}\right) \\ &= -p\log(p)\sum_{n=1}^{\infty}(1-p)^{n-1} - \log(1-p)\sum_{n=1}^{\infty}(n-1)p(1-p)^{n-1} \\ &= -\log(p) - \log(1-p)\left(\frac{1}{p}-1\right) \;=\; -\log(p) - \frac{1-p}{p}\log(1-p). \end{aligned}$$

Beim Berechnen der zweiten Summe haben wir benutzt, dass der Erwartungswert einer Geom(p)-verteilten Zufallsvariable T gleich $1/p$ ist - die Summe ist der Erwartungswert von $T-1$.

c) Die Massenfunktion p_n ist die Wahrscheinlichkeit einer Folge von Ausgängen $x_1, \ldots, x_n$ bei Entnehmen einer Stichprobe von n unabhängigen Zufallsgrößen mit Verteilung μ und ist gegeben durch

$$p_n(x_1, ..., x_n) = \prod_{i=1}^{n} \mu(x_i).$$

Mit Wahrscheinlichkeit 1 gilt $0 < \mu(X_i) \leq 1$ für alle i, also $p_n(X_1, \dots, X_n) > 0$ und $-\log \mu(X_i) \geq 0$. Nach dem Gesetz der Großen Zahlen (ohne Integrierbarkeit) folgt fast sicher

$$-\frac{1}{n} \log p_n(X_1, \dots, X_n) \quad = \quad -\frac{1}{n} \sum_{i=1}^{n} \log \mu(X_i) \quad \stackrel{n\to\infty}{\longrightarrow} \quad -\int \log \mu \, d\mu \quad = \quad H(\mu).$$

d) Wir schreiben die Mengen $B_{n,\epsilon}$ um:

$$\begin{aligned} B_{n,\epsilon} &= \{x \in \mathcal{S}^n : -n(H(\mu)+\epsilon) \leq \log p_n(x_1, \dots, x_n) \leq -n(H(\mu)-\epsilon)\} \\ &= \left\{ x \in \mathcal{S}^n : -H(\mu) - \epsilon \leq \frac{1}{n} \log p_n(x_1, \dots, x_n) \leq -H(\mu) + \epsilon \right\} \\ &= \left\{ x \in \mathcal{S}^n : \left| \frac{1}{n} \log p_n(x_1, \dots, x_n) + H(\mu) \right| \leq \epsilon \right\} \end{aligned}$$

Nun folgt die Aussage aus Aufgabenteil c), da aus der fast sicheren Konvergenz die stochastische Konvergenz folgt.

e) Es gilt

$$\sum_{A_n \cap B_{n,\epsilon}} p_n(x_1, \dots, x_n) \; = \; P[(X_1, \dots, X_n) \in A_n \cap B_{n,\epsilon})]. \tag{17}$$

Nach Voraussetzung konvergiert $P[(X_1, \dots, X_n) \notin A_n]$ gegen Null. Außerdem gilt nach d):

$$P[(X_1, \dots, X_n) \notin B_{n,\varepsilon}] \; = \; \mu^n[B_{n,\varepsilon}^C] \; \to \; 0.$$

Also gilt nach (17):

$$\lim_{n\to\infty} \sum_{A_n \cap B_{n,\epsilon}} p(x_1, \dots, x_n) = 1.$$

Da für $x \in B_{n,\epsilon}$ gilt $p_n(x_1, \dots, x_n) \leq e^{-n(H(\mu)-\epsilon)}$, erhalten wir

$$\sum_{A_n \cap B_{n,\epsilon}} p(x_1, \dots, x_n) \leq |A_n \cap B_{n,\epsilon}| e^{-n(H(\mu)-\epsilon)} \leq |A_n| e^{-n(H(\mu)-\epsilon)}.$$

Daraus folgt für alle $\epsilon > 0$:

$$\frac{1}{n} \log |A_n| \geq H(\mu) - \epsilon + \frac{1}{n} \log \left(\sum_{A_n \cap B_{n,\epsilon}} p(x_1, \dots, x_n) \right)$$

Folglich

$$\frac{1}{n} \log |A_n| \geq H(\mu) + \frac{1}{n} \log \left(\sum_{A_n \cap B_{n,\epsilon}} p(x_1, \dots, x_n) \right).$$

Nun lassen wir n gegen unendlich laufen und erhalten wegen der eben gezeigten Aussage:

$$\liminf_{n\to\infty} \frac{1}{n} \log |A_n| \geq H(\mu).$$