Institut für angewandte Mathematik
Wintersemester 2017/18
Andreas Eberle, Maximilian Fels

Klausur „Einführung in die W'theorie"

Bitte diese Felder in Druckschrift ausfüllen

Name:		Vorname:	
Matrikelnr.:		Studiengang:	

Wichtige Hinweise:

- **Es sind keine eigenen Unterlagen, Handys, Taschenrechner u.ä. zugelassen!**
- Die Klausur enthält 3 Aufgaben, die Sie alle bearbeiten sollten.
- Dieses Deckblatt ist vollständig ausgefüllt zusammen mit den Lösungen abzugeben. Jedes abgegebene Blatt ist zudem mit Namen und Matrikelnummer zu versehen.
- Bitte den Studentenausweis und einen amtlichen Lichtbildausweis bereithalten!
- Abgabe bis spätestens 11.20 Uhr.

Viel Erfolg!

Diese Felder NICHT ausfüllen:

Aufgabe	1	2	3				**Summe**	**Note**
Punkte								

1. (Zufallsvariablen und ihre Verteilung [32 Punkte])

a) Sind die Verteilungen der folgenden Zufallsvariablen absolutstetig ? Berechnen Sie die Verteilungsfunktionen und im absolutstetigen Fall die Dichten.

(i) $U \sim \text{Unif}(-1, 1)$,

(ii) $X = |U|$,

(iii) $Y = U^2$,

(iv) $Z = U^+$. *[15]*

b) Seien $X, Y : \Omega \to \mathbb{R}$ Zufallsvariablen. Wie ist die gemeinsame Verteilung von X und Y definiert ? Wann nennt man X und Y unabhängig ? Wie erkennt man Unabhängigkeit an der Dichte der gemeinsamen Verteilung, wenn diese absolutstetig ist ? *[6]*

c) Seien nun W und Z unabhängige standardnormalverteilte Zufallsvariablen. Zeigen Sie, dass die Zufallsvariablen $X = W + Z$ und $Y = W - Z$ wieder unabhängig sind mit Verteilung $X, Y \sim N(0, 2)$. *[7]*

d) Bestimmen Sie die Verteilung der durch

$$V = \begin{cases} |W| & \text{falls } Z \geq 0, \\ -|W| & \text{falls } Z < 0, \end{cases}$$

definierten Zufallsvariable. *[4]*

Lösung.

a) (i) Die Verteilung ist absolutstetig. Die Dichte ist gegeben durch $f(x) = \frac{1}{2}\mathbb{1}_{(-1,1)}$. Die Verteilungsfunktion lässt sich entsprechend wie folgt berechnen:

$$F(c) = \begin{cases} \int_{-\infty}^{c} f(x)\mathrm{d}x = \int_{-1}^{c} \frac{1}{2}\mathrm{d}x = \frac{1+c}{2}, & \text{für } c \in [-1, 1], \\ 0, \quad \text{für } c \leq -1, \\ 1, \quad \text{für } c \geq 1. \end{cases}$$

(ii) Für $c \in [0, 1]$ gilt

$$F(c) = \mathbb{P}\left[|U| \leq c\right] = \mathbb{P}\left[-c \leq U \leq c\right] = \frac{1+c}{2} - \frac{1-c}{2} = c.$$

Da die Betragsfunktion nicht-negativ und $F(1) = 1$ gilt insgesamt:

$$F(c) = \begin{cases} 0, & \text{für } c \leq 0, \\ c, & \text{für } c \in [0, 1], \\ 1, & \text{für } c \geq 1. \end{cases}$$

Außerdem ist die Verteilung absolutstetig mit Dichte

$$f(x) = F'(x) = \mathbb{1}_{(0,1)}(x), \text{ für fast alle } x.$$

(iii) Die Verteilungsfunktion lässt sich wie folgt berechnen:

$$F(c) = \begin{cases} \mathbb{P}\left[U^2 \le c\right] = \mathbb{P}\left[|U| \le \sqrt{c}\right] = \sqrt{c}, \text{ für } c \in [0,1] \\ 0, \text{ für } c \le 0 \\ 1, \text{ für } c \ge 1. \end{cases}$$

Die Verteilung ist zudem absolutstetig und die Dichte lautet

$$f(x) = F'(x) = \frac{1}{2\sqrt{x}} \mathbb{1}_{(0,1)}(x).$$

(iv) Die Verteilungsfunktion für $c \ge 0$ ist mit (i) gegeben durch

$$F(c) = \mathbb{P}\left[U^+ \le c\right] = \mathbb{P}\left[U \le c\right] = \frac{1+c}{2}.$$

Insbesondere ist $F(0) = \frac{1}{2}$ und $F(c) = 0$ für $c < 0$. Somit ist $\mathbb{P}\left[U^+ = 0\right] = \frac{1}{2}$ und die Verteilung nicht absolutstetig an der Stelle $c = 0$.

b) Gemeinsame Verteilung: $\mu : \mathcal{B}(\mathbb{R}^2) \to [0,1]$

$$\mu\left[B\right] = \mathbb{P}\left[(X,Y) \in B\right], \quad B \in \mathcal{B}(\mathbb{R}^2).$$

$$\begin{aligned} X, Y \text{unabh.} &\iff \mathbb{P}\left[X \in A, Y \in B\right] = \mathbb{P}\left[X \in A\right]\mathbb{P}\left[Y \in B\right] \ \ \forall A, B \in \mathcal{B}(\mathbb{R}) \\ &\iff \mu\left[A \times B\right] = \mu\left[A\right]\mu\left[B\right] \ \ \forall A, B \in \mathcal{B}(\mathbb{R}) \end{aligned}$$

μ absolutstetig mit Dichte f:

X, Y unabh. $\iff f(x,y) = g(x)h(y)$ mit messbaren Funktionen $g, h : \mathbb{R} \to [0,1]$.

c)

$$\begin{pmatrix} W+Z \\ W-Z \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} W \\ Z \end{pmatrix} \qquad \begin{matrix} X = W + Z \\ \\ Y = W - Z \end{matrix} \quad \iff \quad \begin{matrix} W = \frac{X+Y}{2} \\ \\ Z = \frac{X-Y}{2} \end{matrix}$$

$$\begin{aligned} f_{(X,Y)}(x,y) &= f_{(W,Z)}\left(\frac{x+y}{2}, \frac{x-y}{2}\right) \left| \begin{matrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{matrix} \right| = \frac{1}{2\pi} e^{-\frac{1}{2}\left(\frac{x+y}{2}\right)^2 - \frac{1}{2}\left(\frac{x-y}{2}\right)^2} \frac{1}{2} \\ &= \frac{1}{4\pi} e^{-\frac{x^2}{4} - \frac{y^2}{4}} = \frac{1}{\sqrt{2\pi \cdot 2}} \frac{1}{\sqrt{2\pi \cdot 2}} e^{-\frac{x^2}{2\cdot 2}} e^{-\frac{y^2}{2\cdot 2}} = \Phi_{N(0,2)}(x)\Phi_{N(0,2)}(y). \end{aligned}$$

$\Rightarrow X, Y$ unabh. und $\sim N(0,2)$

d)

$$\begin{aligned}
\mathbb{P}[V \le c] &= \mathbb{P}[|W| \le c | Z \ge 0]\,\mathbb{P}[Z \ge 0] + \mathbb{P}[-|W| \le c | Z < 0]\,\mathbb{P}[Z < 0] \\
&\underset{\text{unabh.}}{\overset{Z\sim N(0,1)}{=}} \frac{1}{2}\mathbb{P}[|W| \le c] + \frac{1}{2}\mathbb{P}[|W| \ge -c]
\end{aligned}$$

Für $c \ge 0$ erhalten wir

$$\mathbb{P}[V \le c] = \frac{1}{2}\mathbb{P}[-c \le W \le c] + \frac{1}{2} = \mathbb{P}[W \in [0,c]] + \frac{1}{2} = \mathbb{P}[W \le c].$$

Für $c < 0$ folgt entsprechend

$$\mathbb{P}[V \le c] = 0 + \frac{1}{2}\mathbb{P}[W \notin [c,-c]] = \mathbb{P}[W \le c].$$

Also gilt $V \sim W \sim N(0,1)$

2. (Erwartungswerte und erzeugende Funktionen [24 Punkte])

a) Sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable. Wie sind die Varianz, die momentenerzeugende und die charakteristische Funktion von X definiert ? *(Die Definition des Erwartungswerts kann ohne Erläuterung vorausgesetzt werden)* *[6]*

b) Sei $X \sim N(m, v)$ mit $m \in \mathbb{R}$ und $v \in [0, \infty)$. Berechnen Sie die momentenerzeugende Funktion von X. *[5]*

c) Eine Zufallsvariable Y heisst *log-normalverteilt*, wenn $X = \log(Y)$ normalverteilt ist (log bezeichnet hier den natürlichen Logarithmus). Berechnen Sie den Erwartungswert und die Varianz einer log-normalverteilten Zufallsvariable. *[4]*

d) Zeigen Sie: Für eine Zufallsvariable $X : \Omega \to \mathbb{R}$ gilt $X \sim -X$ genau dann, wenn der Imaginärteil der charakteristischen Funktion verschwindet. *[6]*

e) Nennen Sie ein Beispiel einer Folge $(\mu_n)_{n\in\mathbb{N}}$ von Wahrscheinlichkeitsverteilungen auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, deren charakteristische Funktionen punktweise gegen eine Funktion $\varphi(t)$ konvergieren, die selbst keine charakteristische Funktion einer Wahrscheinlichkeitsverteilung auf $\mathbb{R}$ ist. *(Hier ist kein Beweis verlangt, das Angeben eines Beispiels genügt)* *[3]*

Lösung.

a)

$$\mathrm{Var}[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 \qquad \text{falls } X \in \mathcal{L}^1.$$

momenterzeugende Funktion: $M(t) = \mathbb{E}\left[e^{tX}\right]$, $t \in \mathbb{R}$.

charakteristische Funktion: $\phi(t) = \mathbb{E}\left[e^{itX}\right]$, $t \in \mathbb{R}$.

b)

$$\begin{aligned}
\mathbb{E}\left[e^{tX}\right] &= \frac{1}{\sqrt{2\pi v}} \int_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2v}} e^{tx} \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi v}} \int_{-\infty}^{\infty} e^{-\frac{1}{2v}\left(x^2 - 2x(m+tv) + m^2\right)} \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi v}} \int_{-\infty}^{\infty} e^{-\frac{1}{2v}(x-(m+tv))^2} e^{-\frac{1}{2v}\left(m^2 - (m+tv)^2\right)} \mathrm{d}x \\
&= e^{mt + \frac{t^2 v}{2}} \underbrace{\frac{1}{\sqrt{2\pi v}} \int_{-\infty}^{\infty} e^{-\frac{1}{2v}(x-(m+tv))^2} \mathrm{d}x}_{=1} \\
&= e^{mt + \frac{t^2 v}{2}}.
\end{aligned}$$

c)

$$
\begin{aligned}
\mathbb{E}[Y] &= \mathbb{E}\left[e^X\right] = M(1) \\
\mathbb{E}\left[Y^2\right] &= \mathbb{E}\left[e^{2X}\right] = M(2) \\
\mathrm{Var}[Y] &= M(2) - M(1)^2
\end{aligned}
$$

d)

$$
X \sim -X \underset{\substack{\text{Verteilung ist durch} \\ \text{char. Fkt. festgelegt}}}{\iff} \mathbb{E}\left[e^{itX}\right] = \mathbb{E}\left[e^{it(-X)}\right] \ \forall t \in \mathbb{R}.
$$

Weiter ist

$$
\mathbb{E}\left[e^{-itX}\right] = \mathbb{E}[\cos(-tX)] + i\mathbb{E}[\sin(-tX)] = \mathbb{E}[\cos(tX)] - i\mathbb{E}[\sin(tX)].
$$

Also folgt

$$
X \sim -X \iff \underbrace{\mathbb{E}[\sin(tX)]}_{=\mathrm{Im}(\mathbb{E}[e^{itX}])=\mathrm{Im}(\phi(t))} = 0 \qquad \forall t \in \mathbb{R}
$$

e) Für $\mu_n = N(0,n)$ gilt $\phi_n(t) = e^{-nt^2/2} \overset{n\to\infty}{\to} \mathbb{1}_{\{0\}}(t)$. Da der Limes nicht stetig ist, ist dieser keine charakteristische Funktion einer WV auf $\mathbb{R}$ (Die Masse wandert ins unendliche ab).

3. (Grenzwertsätze [44 Punkte])

a) Seien X_n $(n \in \mathbb{N})$ unabhängige, identisch verteilte Zufallsvariablen in $\mathcal{L}^2(\Omega, \mathcal{A}, P)$ mit $E[X_n] = 0$. Formulieren und beweisen Sie den zentralen Grenzwertsatz für die Summen $X_1 + \ldots + X_n$. Dabei können Sie auf grundlegende Sätze aus der Vorlesung verweisen, ohne diese separat zu beweisen. Lässt sich die Aussage auf unabhängige, identisch verteilte Zufallsvariablen in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$ erweitern ? *[16]*

b) Seien X_n und Y_n $(n \in \mathbb{N})$ unabhängige, identisch verteilte Zufallsvariablen in $\mathcal{L}^2(\Omega, \mathcal{A}, P)$ mit Erwartungswert m und Varianz $v > 0$. Zeigen Sie, dass die Zufallsvariablen

$$U_n \;=\; \frac{1}{\sqrt{2nv}} \left(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} Y_i \right)$$

in Verteilung gegen eine standardnormalverteilte Zufallsvariable Z konvergieren. *[6]*

c) Sei A_n $(n \in \mathbb{N})$ eine Folge von Ereignissen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Zeigen Sie

$$P[A_n \text{unendlich oft}] \;\geq\; \limsup P[A_n].$$

[5]

d) Seien $X_n : \Omega \to \mathbb{R}$ $(n \in \mathbb{N})$ *unabhängige* Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Beweisen oder widerlegen Sie:

(i) $X_n \to X$ P-stochastisch $\quad \Rightarrow \quad E[X_n] \to E[X]$,

(ii) $X_n \to X$ in $L^2(P)$ $\quad \Rightarrow \quad X_n \to X$ in $L^1(P)$.

(iii) $X_n \to X$ schnell stochastisch $\quad \Leftrightarrow \quad X_n \to X$ P-fast sicher.

Dabei können Sie wieder auf grundlegende Sätze aus der Vorlesung verweisen ohne diese separat zu beweisen. *[13]*

e) Sei S eine nichtleere endliche Menge. Wie ist die Entropie einer Wahrscheinlichkeitsverteilung auf S definiert ? Zeigen Sie, dass die Gleichverteilung die Entropie unter allen Wahrscheinlichkeitsverteilungen auf S maximiert. *[4]*

Lösung.

a) ZGS: Es sei $S_n = X_1 + \ldots X_n$ und $v = \mathrm{Var}\,[X_1]$. Dann gilt,

$$\frac{S_n}{\sqrt{n}} \stackrel{\mathcal{D}}{\to} N(0, v),$$

d.h.

$$\mu_n = \mathrm{Vert}(\frac{S_n}{\sqrt{n}}) \stackrel{w}{\to} N(0, v).$$

Beweis:
Nach Konvergenzsatz von Lévy reicht es zu zeigen, dass

$$\phi_n(t) \to \phi_{N(0,v)}(t) \quad \forall t \in \mathbb{R}.$$

Es ist,

$$\phi_n(t) = \mathbb{E}\left[e^{itS_n/\sqrt{n}}\right] \overset{\text{iid}}{=} \mathbb{E}\left[e^{itX_1/\sqrt{n}}\right]^n = \phi_{X_1}(t/\sqrt{n})^n.$$

Nun ist nach Voraussetzung $X_1 \in \mathcal{L}^2$, also folgt

$$\phi_{X_1}(t) = 1 + i\mathbb{E}\left[X_1\right]t - \frac{1}{2}\mathbb{E}\left[X_1^2\right]t^2 + o(t^2) \overset{\text{hier}}{=} 1 - \frac{v}{2}t^2 + o(t^2).$$

$$\Rightarrow \phi_n(t) = \left(1 - \overbrace{\frac{vt^2}{2n} + o\left(\frac{t^2}{n}\right)}^{|\cdot|\leq 1 \text{ für große } n}\right)^n \overset{n\to\infty}{\to} e^{-\frac{v}{2}t^2} = \phi_{N(0,v)}(t) \quad \forall t \in \mathbb{R},$$

da $|\prod_{i=1}^n z_i - \prod_{i=1}^n w_i| \leq \sum_{i=1}^n |z_i - w_i|$, falls $|z_i|, |w_i| \leq 1$. Insgesamt folgt,

$$\mathrm{Vert}(S_n/\sqrt{n}) \overset{w}{\to} N(0,v)$$

Ist die Voraussetzung $X_n \in \mathcal{L}^2$ nicht erfüllt, dann können sich andere Skalierungslimiten ergeben, siehe zum Beispiel den Grenzwertsatz für α-stabile Verteilungen.

b) Zunächst gilt

$$U_n = \frac{1}{\sqrt{2nv}}\left(\sum_{i=1}^n X_i - \sum_{i=1}^n Y_i\right) = \frac{1}{\sqrt{n}} \underbrace{\sum_{i=1}^n \frac{X_i - Y_i}{\sqrt{2v}}}_{=S_n}.$$

Nach dem Gruppierungssatz sind die Zufallsvariablen $\frac{X_i - Y_i}{\sqrt{2v}}$ wieder i.i.d. mit

$$\begin{aligned}
\mathbb{E}\left[\frac{X_i - Y_i}{\sqrt{2v}}\right] &= \frac{1}{\sqrt{2v}}(\mathbb{E}[X_i] + \mathbb{E}[Y_i]) = 0, \\
\mathrm{Var}\left[\frac{X_i - Y_i}{\sqrt{2v}}\right] &= \frac{1}{2v}\left(\mathrm{Var}\left[X_i\right] + \mathrm{Var}\left[Y_i\right]\right) = \frac{\mathrm{Var}\left[X_i\right]}{v} = 1.
\end{aligned}$$

Aus dem ZGS folgt daher

$$U_n = \frac{S_n}{\sqrt{n}} \overset{\mathcal{D}}{\to} N(0,1).$$

c)

$$\begin{aligned}
\mathbb{P}\left[A_n \text{ unendlich oft}\right] &= \mathbb{P}[\bigcap_n \underbrace{\bigcup_{m\geq n} A_m}_{\text{monoton fallend}}] \overset{\text{mon. Stetigkeit}}{=} \lim_{n\to\infty} \overbrace{\mathbb{P}[\underbrace{\bigcup_{m\geq n} A_m}_{\supseteq A_m\, \forall m\geq n}]}^{\geq \sup_{m\geq n} \mathbb{P}[A_m]} \\
&\geq \limsup_{n\to\infty} \mathbb{P}\left[A_n\right].
\end{aligned}$$

d) (i) falsch: Betrachte eine Folge unabhängiger Zufallsvariablen mit Verteilung

$$P[X_n = n] = 1/n \qquad \text{und} \qquad P[X_n = 0] = 1 - 1/n.$$

Dann gilt $X_n \to 0$ P-stochastisch, denn $P[|X_n| \geq \epsilon] \leq P[X_n \neq 0] = 1/n$ für alle $\epsilon > 0$. Andererseits gilt aber $E[X_n] = 1$ für alle n, also konvergieren die Erwartungswerte nicht gegen 0.

(ii) wahr nach Cauchy-Schwarz:

$$\mathbb{E}\left[|X_n - X|\right] \leq \mathbb{E}\left[|X_n - X|^2\right]^{1/2}.$$

(iii) wahr: Einerseits gilt nach dem ersten Borel-Cantelli-Lemma:

$$\begin{aligned} X_n \to X \text{ schnell stoch.} &\Rightarrow \sum_{n=1}^{\infty} \mathbb{P}\left[|X_n - X| > \epsilon\right] < \infty \, \forall \epsilon > 0 \\ &\overset{\text{BC}}{\Rightarrow} \quad \mathbb{P}\left[|X_n - X| > \epsilon \text{ nur endl. oft}\right] = 1 \, \forall \epsilon > 0 \\ &\Rightarrow \quad \mathbb{P}\left[\forall \epsilon \in \mathbb{Q}_+ \, \exists n_0 \in \mathbb{N} \, \forall n \geq n_0 : |X_n - X| \leq \epsilon\right] = 1 \\ &\Rightarrow \quad X_n \to X \quad \mathbb{P}\text{-fast sicher.} \end{aligned}$$

Umgekehrt folgt nach dem Kolmogorovschen 0-1-Gesetz aus $X_n \to X$ fast sicher, dass X fast sicher konstant ist. Also sind auch $X_n - X$ $(n \in \mathbb{N})$ wieder unabhängige Zufallsvariablen. Damit können wir dass zweite Borel-Cantelli-Lemma anwenden, und so schließen, dass die Aussagen oben alle äquivalent sind.

e) Die Entropie einer Wahrscheinlichkeitsverteilung μ auf dem Zustandsraum S ist

$$H(\mu) = -\sum_{\substack{x \in S \\ \mu(x) \neq 0}} \mu(x) \log(\mu(x)).$$

Die Funktion $u(t) = t \log(t)$ ist konvex. Also gilt mit der Jensen'schen Ungleichung:

$$\begin{aligned} H(\mu) &= -\sum_{\substack{x \in S \\ \mu(x) \neq 0}} \mu(x) \log(\mu(x)) = -|S| \int_S \mu(x) \log(\mu(x)) \mathrm{Unif}_S(\mathrm{d}x) \\ &\overset{\text{Jensen}}{\leq} -|S| \underbrace{\int_S \mu(x) \mathrm{Unif}_S(\mathrm{d}x)}_{=1/|S|} \log\left(\int_S \mu(x) \mathrm{Unif}_S(\mathrm{d}x)\right) = \log(|S|). \end{aligned}$$

Insbesondere maximiert die Gleichverteilung auf S die Entropie:

$$H(\mathrm{Unif}_S) = -\sum_{x \in S} \frac{1}{|S|} \log(1/|S|) = \log(|S|).$$