

2 Grundlagen (Version Dezember 2017)

2.1 Charakteristische Funktionen

Wir hatten in der Vorlesung Wahrscheinlichkeitstheorie 1 schon charakteristische Funktionen von reellen Zufallsvariablen oder Wahrscheinlichkeitsmaßen auf \mathbb{R} kennengelernt. Wir verallgemeinern nun den Begriff auf allgemeine Dimensionen.

Definition 2.1. Sei $n \in \mathbb{N}$ und μ ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}^n, \mathcal{B}^n)$ (wobei \mathcal{B}^n die Borel- σ -Algebra auf \mathbb{R}^n ist). Die *charakteristische Funktion* $\hat{\mu} : \mathbb{R}^n \rightarrow \mathbb{C}$ ist definiert als

$$\hat{\mu}(t) = \int_{\mathbb{R}^n} e^{i\langle t, x \rangle} \mu(dx) := \int_{\mathbb{R}^n} \cos(\langle t, x \rangle) \mu(dx) + i \int_{\mathbb{R}^n} \sin(\langle t, x \rangle) \mu(dx),$$

wobei $\langle \cdot, \cdot \rangle$ das kanonische Skalarprodukt auf \mathbb{R}^n ist.

Ist X eine $(\mathbb{R}^n, \mathcal{B}^n)$ -wertige Zufallsvariable auf einem Wahrscheinlichkeitsraum, so nennt man die charakteristische Funktion der Verteilung von X auch *charakteristische Funktion von X* .

Man sieht sofort, dass immer $|\hat{\mu}(t)| \leq 1$ gilt und dass (wegen dominierter Konvergenz) die Funktion $t \mapsto \hat{\mu}(t)$ stetig ist.

Wichtig ist der folgende Eindeutigkeitsatz.

Satz 2.2. Seien μ und ν zwei Wahrscheinlichkeitsmaße auf $(\mathbb{R}^n, \mathcal{B}^n)$. Gilt $\hat{\mu}(t) = \hat{\nu}(t)$ für alle $t \in \mathbb{R}^n$, so folgt $\mu = \nu$.

Beweis. Sei \mathcal{K} die Familie aller Mengen der Form $[a_1, b_1] \times \cdots \times [a_n, b_n]$ in \mathbb{R}^n (wobei $[a_j, b_j] := \emptyset$ falls $a_j > b_j$). Da \mathcal{K} die Borel- σ -Algebra \mathcal{B}^n erzeugt und durchschnittsstabil ist, genügt es zu zeigen, dass $\mu(K) = \nu(K)$ für alle $K \in \mathcal{K}$ gilt. Für $K = \emptyset$ ist die Aussage klar. Sei daher $K = [a_1, b_1] \times \cdots \times [a_n, b_n] \neq \emptyset$. Wir definieren für $m \in \mathbb{N}$ und $j \in \{1, \dots, n\}$

$$f_{m,j}(s) := (1 - md(s, [a_j, b_j]))^+, \quad s \in \mathbb{R}; \quad f_m(x) := \prod_{j=1}^n f_{m,j}(x_j), \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n,$$

wobei $d(s, M) := \inf\{|s - y| : y \in M\}$ für $M \subset \mathbb{R}$. Wenn wir zeigen können, dass für jedes $m \in \mathbb{N}$ gilt, dass $\int f_m d\mu = \int f_m d\nu$, dann folgt wegen $f_m \downarrow 1_K$ mit dominierter Konvergenz $\mu(K) = \nu(K)$. Wir fixieren nun neben der Menge K auch $m \in \mathbb{N}$ und zeigen $\int f d\mu = \int f d\nu$ für $f := f_m$.

Sei $\varepsilon > 0$ und $N \in \mathbb{N}$ so groß, dass $B_N := [-N, N]^n$ den Träger von f_m enthält und ausserdem $\mu(B_N^c), \nu(B_N^c) < \varepsilon$ gelten. Nach dem Approximationssatz von Weierstraß existiert für jedes $j \in \{1, \dots, n\}$ eine Funktion der Form $g_j(s) = \sum_{k=1}^{M_j} c_{j,k} \exp\{i \frac{\pi}{N} t_{j,k} s\}$ mit $t_{j,k} \in \mathbb{Z}$, die $f_{m,j}$ auf $[-N, N]$ bis auf ε approximiert. Es folgt daher für $g(x) := \prod_{j=1}^n g_j(x_j)$, dass $\sup_x |g(x)| \leq (1 + \varepsilon)^n$ und

$$\left| \int f d\mu - \int f d\nu \right| \leq \left| \int f d\mu - \int g d\mu \right| + \left| \int g d\mu - \int g d\nu \right| + \left| \int g d\nu - \int f d\nu \right|.$$

Der zweite Summand ist wegen der Voraussetzung $\hat{\mu}(t) = \hat{\nu}(t)$ Null. Der erste Summand ist kleiner oder gleich

$$|\int_{B_N} f d\mu - \int_{B_N} g d\mu| + |\int_{B_N^c} f d\mu| + |\int_{B_N^c} g d\mu| \leq n\varepsilon(1 + \varepsilon)^{n-1} + 0 + (1 + \varepsilon)^n \varepsilon$$

und der dritte Summand wird analog abgeschätzt. Da $\varepsilon > 0$ beliebig war, folgt $\int f d\mu = \int f d\nu$. \square

Beispiele 2.3. a) Sei $\mu = \mathcal{N}(0, 1)$. Dann gilt $\hat{\mu}(t) = \exp\{-\frac{t^2}{2}\}$ (Beweis mit Funktionentheorie).

b) Sei μ die Standardnormalverteilung auf dem \mathbb{R}^n . Dann folgt aus a) sofort $\hat{\mu}(t) = \exp\{-\frac{1}{2} \sum_{j=1}^n t_j^2\}$, $t = (t_1, \dots, t_n)$.

c) Sei A eine reelle $n \times m$ Matrix, $b \in \mathbb{R}^n$ und $\varphi : \mathbb{R}^m \rightarrow \mathbb{R}^n$ definiert durch $\varphi(x) := Ax + b$. Sei μ die Standardnormalverteilung auf \mathbb{R}^m und ν das Bildmaß von μ unter φ . Dann gilt für $t \in \mathbb{R}^n$

$$\begin{aligned} \hat{\nu}(t) &= \int e^{i\langle t, y \rangle} \nu(dy) = \int e^{i\langle t, \varphi(x) \rangle} \mu(dx) \\ &= e^{i\langle t, b \rangle} \int e^{i\langle A^T t, x \rangle} \mu(dx) \\ &= e^{i\langle t, b \rangle} \hat{\mu}(A^T t) = e^{i\langle t, b \rangle} e^{-\frac{1}{2} \langle A^T t, A^T t \rangle} \\ &= \exp\{i\langle t, b \rangle - \frac{1}{2} \langle t, \Sigma t \rangle\}, \end{aligned}$$

wobei $\Sigma = AA^T$ die Kovarianzmatrix von ν ist.

Man beachte, dass wegen Satz 2.2 die Verteilung von $\mu\varphi^{-1}$ nur von Σ und b , nicht aber von der speziellen Wahl der Matrix A abhängt, solange $AA^T = \Sigma$ gilt. Daher ist es sinnvoll, von *der* $\mathcal{N}(b, \Sigma)$ -Verteilung zu sprechen (diese Aussage konnten wir in der Vorlesung Wahrscheinlichkeitstheorie 1 nicht beweisen). Man beachte auch, dass für jede positiv semidefinite Matrix Σ (mindestens) eine Matrix A existiert, so dass $\Sigma = AA^T$ gilt. Also existiert zu jedem $n \in \mathbb{N}_0$, $b \in \mathbb{R}^n$ und jeder positiv semidefiniten $n \times n$ Matrix Σ genau eine $\mathcal{N}(b, \Sigma)$ -Verteilung.

2.2 Terminale- σ -Algebren und Kolmogorovs 0-1-Gesetz

Satz 2.4 (Kolmogorovs 0-1-Gesetz). Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und \mathcal{F}_n , $n \in \mathbb{N}$ eine unabhängige Folge von Teil- σ -Algebren von \mathcal{F} . Sei $\check{\mathcal{F}}_n := \bigvee_{k=n}^{\infty} \mathcal{F}_k$ und $\mathcal{T}_{\infty} := \bigcap_{n=1}^{\infty} \check{\mathcal{F}}_n$ die terminale σ -Algebra. Dann gilt $\mathbb{P}(A) \in \{0, 1\}$ für jedes $A \in \mathcal{T}_{\infty}$.

Beweis. Aus der Vorlesung Wahrscheinlichkeitstheorie 1 (Korollar 4.12; oder Skript von W. König, 7.2.6) folgt, dass für jedes $n \in \mathbb{N}$ die beiden σ -Algebren $\check{\mathcal{F}}_{n+1}$ und $\bigvee_{k=1}^n \mathcal{F}_k$ unabhängig sind. Wegen $\mathcal{T}_\infty \subset \check{\mathcal{F}}_{n+1}$ sind daher (nach 4.9.a bzw. 7.2.3.a) auch \mathcal{T}_∞ und $\bigvee_{k=1}^n \mathcal{F}_k$ unabhängig für jedes $n \in \mathbb{N}$ und somit sind auch \mathcal{T}_∞ und $\bigcup_{n=1}^\infty \bigvee_{k=1}^n \mathcal{F}_k$ unabhängig (4.9.b bzw. 7.2.3.b). Da das letzte Mengensystem durchschnittsstabil ist, sind auch \mathcal{T}_∞ und $\bigvee_{k=1}^\infty \mathcal{F}_k$ unabhängig (4.10 bzw. 7.2.4.). Nun gilt $\mathcal{T}_\infty \subset \bigvee_{k=1}^\infty \mathcal{F}_k$ und daher sind auch \mathcal{T}_∞ und \mathcal{T}_∞ unabhängig (4.9.a bzw. 7.2.3.a). Dies bedeutet, dass für jedes $A \in \mathcal{T}_\infty$ gilt: $\mathbb{P}(A) = \mathbb{P}(A \cap A) = (\mathbb{P}(A))^2$, also $\mathbb{P}(A) \in \{0, 1\}$. \square

Wir betrachten zur Illustration folgendes Beispiel (siehe auch Skript von W. König, Beispiel 7.2.9).

Beispiel 2.5. (Perkolaton). Sei \mathcal{K} die Menge aller Kanten zwischen je zwei Nachbarn auf dem Gitter \mathbb{Z}^d . Für gegebenes $p \in (0, 1)$ sei $(X_i)_{i \in \mathcal{K}}$ eine Familie von unabhängigen Zufallsvariablen mit Werten in $\{0, 1\}$ und $\mathbb{P}(X_i = 1) = p$. Wir interpretieren $X_i = 1$ als das Ereignis, dass die Kante i *offen* oder *durchlässig* ist und $X_i = 0$, dass i *geschlossen* oder *undurchlässig* ist. Wir fragen nach der Wahrscheinlichkeit $\Theta_d(p)$, dass ein unendlich großes Cluster offener Kanten existiert, also ein unendlicher zusammenhängender Teilgraph von \mathbb{Z}^d , bei dem alle Kanten offen sind. Eine einfache Anwendung des letzten Satzes zeigt, dass $\Theta_d(p)$ nur die Werte 0 und 1 annehmen kann. Wir werden das in der Vorlesung diskutieren. Die Frage, für welche d und p der Wert 1 angenommen wird, ist allgemein nach wie vor ungelöst.

Nützlich ist das folgende Lemma. Dazu nennen wir einen Messraum (E, \mathcal{E}) *punktetrennend*, wenn zu je zwei verschiedenen $e_1, e_2 \in E$ disjunkte Mengen $E_1, E_2 \in \mathcal{E}$ existieren, so dass $e_1 \in E_1$ und $e_2 \in E_2$ gilt. Weiter heißt (E, \mathcal{E}) *abzählbar erzeugt*, wenn es einen abzählbaren Erzeuger von \mathcal{E} gibt. Offensichtlich ist $(\mathbb{R}^n, \mathcal{B}^n)$ punktetrennend und abzählbar erzeugt, aber auch $(\bar{\mathbb{R}}, \bar{\mathcal{B}})$.

Lemma 2.6. *Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum mit der Eigenschaft, dass $\mathbb{P}(A) \in \{0, 1\}$ für alle $A \in \mathcal{F}$ gilt. Weiter sei (E, \mathcal{E}) ein punktetrennender und abzählbar erzeugter Messraum und Z eine E -wertige Zufallsvariable auf $(\Omega, \mathcal{F}, \mathbb{P})$. Dann existiert ein $c \in E$ mit $\mathbb{P}(Z = c) = 1$ (insbesondere ist die Menge $\{Z = c\}$ messbar).*

Beweis. Sei G_1, G_2, \dots ein abzählbarer Erzeuger von \mathcal{E} . Dann gilt $\mathbb{P}(Z \in G_i) \in \{0, 1\}$. Für alle i mit $\mathbb{P}(Z \in G_i) = 0$ ersetzen wir die Menge G_i durch ihr Komplement. Also gilt dann $\mathbb{P}(Z \in G_i) = 1$ für alle i . Sei $H := \bigcap_{i=1}^\infty G_i$. Dann gilt $\mathbb{P}(Z \in H) = 1$. Weiter ist H einelementig, d.h. $H = \{c\}$. Um das einzusehen, betrachten wir die Familie \mathcal{G} aller Mengen der Form $H \cup U$, wobei $U \subset H^c$ gilt. \mathcal{G} ist eine σ -Algebra und enthält \mathcal{F} . Da \mathcal{F} punktetrennend ist, gilt dies für \mathcal{G} erst recht. Also ist H einelementig. \square

Bemerkung 2.7. Ohne die Voraussetzungen “punktetrennend” und “abzählbar erzeugt” ist die Aussage des letzten Lemmas allgemein falsch. Als Beispiel wählen wir $\Omega = E = [0, 1]$ ausgestattet mit der σ -Algebra \mathcal{E} , die von den abzählbaren Teilmengen von E erzeugt wird. Das Wahrscheinlichkeitsmaß \mathbb{P} auf Ω nehme den Wert Null auf abzählbaren und den Wert 1 auf den Komplementen abzählbarer Mengen an. Schließlich sei $Z : \Omega \rightarrow E$ die Identität. Dann sind alle Voraussetzungen des Lemmas außer der abzählbaren Erzeugtheit erfüllt und es existiert offensichtlich kein $c \in [0, 1]$ mit $\mathbb{P}(Z = c) = 1$ (obwohl die Menge $\{Z = c\}$ für jedes c messbar ist).

Als Anwendung betrachten wir eine Folge $X_n, n \in \mathbb{N}$ unabhängiger reeller Zufallsgrößen und definieren $S_n := \sum_{i=1}^n X_i$. Sei $\mathcal{F}_n := X_n^{-1}(\mathcal{B})$ und \mathcal{T}_∞ wie in Satz 2.4 definiert.

Lemma 2.8. Sei $a_n, n \in \mathbb{N}$ eine Folge positiver Zahlen mit $a_n \rightarrow \infty$. Dann sind $\bar{Y} := \limsup_{n \rightarrow \infty} \frac{S_n}{a_n}$ und $\underline{Y} := \liminf_{n \rightarrow \infty} \frac{S_n}{a_n}$ $(\mathcal{T}_\infty, \bar{\mathcal{B}})$ -messbare Zufallsgrößen.

Beweis. Für jedes $m \in \mathbb{N}$ gilt

$$\bar{Y} = \limsup_{n \rightarrow \infty} \frac{S_n}{a_n} = \limsup_{n \rightarrow \infty} \frac{1}{a_n} \left(\sum_{j=1}^m X_j + \sum_{j=m+1}^n X_j \right) = \limsup_{n \rightarrow \infty} \frac{1}{a_n} \sum_{j=m+1}^n X_j.$$

Der letzte Ausdruck ist $\tilde{\mathcal{F}}_m$ -messbar. Also ist \bar{Y} \mathcal{T}_∞ -messbar. Der Beweis für \underline{Y} ist analog. \square

Als Folgerung dieses Lemmas, Lemma 2.6 und des Kolmogorovschen 0-1-Gesetzes erhalten wir die folgende Aussage.

Satz 2.9. Unter obigen Voraussetzungen sind \bar{Y} und \underline{Y} fast sicher konstant.

Eine interessante Frage ist, wie man unter obigen Voraussetzungen eine Folge (a_n) findet, so dass zum Beispiel $\bar{Y} = 1$ fast sicher ist. Wir werden darauf kurz in der Vorlesung eingehen.

2.3 Stochastische Prozesse, der Existenzsatz von Kolmogorov und der Satz von Ionescu Tulcea

Dieser Abschnitt ist eine Kurzfassung von Kapitel 10 im Skript von W. König. Für einige Beweise verweisen wir auf jenes Skript.

Sei (E, \mathcal{E}) ein Messraum. Mit $\mathcal{E}^{\mathbb{N}_0}$ bezeichnen wir die kleinste σ -Algebra auf $E^{\mathbb{N}_0}$, die alle Mengen der Form $A_0 \times A_1 \times \dots \times A_m \times E \times E \times \dots$ mit $m \in \mathbb{N}_0$ und $A_0, \dots, A_m \in \mathcal{E}$ enthält. Äquivalent ist $\mathcal{E}^{\mathbb{N}_0}$ die kleinste σ -Algebra auf $E^{\mathbb{N}_0}$, die alle Mengen der Form $E \times \dots \times E \times A_m \times E \times \dots$ mit $m \in \mathbb{N}_0$ und $A_m \in \mathcal{E}$ enthält. Für $n \in \mathbb{N}_0$ sei die Projektion $\pi_n : E^{\mathbb{N}_0} \rightarrow E$ definiert als

$$\pi_n((x_k)_{k \in \mathbb{N}_0}) = x_n.$$

Dann ist π_n ($\mathcal{E}^{\mathbb{N}_0}, \mathcal{E}$)-messbar und $\mathcal{E}^{\mathbb{N}_0}$ ist die kleinste σ -Algebra auf $E^{\mathbb{N}_0}$, für die alle π_n messbar sind.

Definition 2.10. Eine Folge von E -wertigen Zufallsvariablen $X_n, n \in \mathbb{N}_0$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ heißt E -wertiger *stochastischer Prozess*.

Lemma 2.11. Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und X_0, X_1, X_2, \dots Abbildungen von Ω nach E . Dann ist $\mathbb{X} = (X_n)_{n \in \mathbb{N}_0}$ ein E -wertiger stochastischer Prozess genau dann, wenn $\mathbb{X} : \Omega \rightarrow E^{\mathbb{N}_0}$ eine $(E^{\mathbb{N}_0}, \mathcal{E}^{\mathbb{N}_0})$ -wertige Zufallsvariable ist.

Beweis. Wenn \mathbb{X} eine $(E^{\mathbb{N}_0}, \mathcal{E}^{\mathbb{N}_0})$ -wertige Zufallsvariable ist, dann ist $X_n = \pi_n \circ \mathbb{X}$ als Komposition messbarer Abbildungen (\mathcal{F}, \mathcal{E})-messbar, also ist $(X_n)_{n \in \mathbb{N}_0}$ ein stochastischer Prozess.

Ist umgekehrt \mathbb{X} ein stochastischer Prozess und $A \in \mathcal{E}^{\mathbb{N}_0}$ von der Form $A = A_0 \times A_1 \times \dots \times A_m \times E \times E \times \dots$, so ist $\mathbb{X}^{-1}(A) = \{\omega \in \Omega : X_1(\omega) \in A_1, \dots, X_m(\omega) \in A_m\} \in \mathcal{F}$. Da $\mathcal{E}^{\mathbb{N}_0}$ von Mengen dieser Form erzeugt wird, ist \mathbb{X} eine $(E^{\mathbb{N}_0}, \mathcal{E}^{\mathbb{N}_0})$ -wertige Zufallsvariable. \square

Im folgenden werden wir neben π_n noch weitere Projektionsabbildungen benötigen. Für $n \in \mathbb{N}_0$ seien $\rho_n : E^{\mathbb{N}_0} \rightarrow E^{n+1}$ und $\varphi_{n+1} : E^{n+2} \rightarrow E^{n+1}$ definiert als

$$\rho_n(e_0, e_1, \dots) = (e_0, \dots, e_n); \quad \varphi_{n+1}(e_0, \dots, e_{n+1}) := (e_0, \dots, e_n).$$

Ist Q ein Wahrscheinlichkeitsmaß auf $(E^{\mathbb{N}_0}, \mathcal{E}^{\mathbb{N}_0})$, so nennen wir die Bildmaße $Q_n := Q\rho_n^{-1}$ die *endlich dimensionalen Verteilungen* von Q . Da es meist einfacher ist, die Wahrscheinlichkeitsmaße Q_n anzugeben als Q stellt sich folgende wichtige Frage:

Welche Bedingungen muss man an eine Folge Q_n von Wahrscheinlichkeitsmaßen auf E^{n+1} stellen, damit ein Q existiert mit der Eigenschaft $Q_n := Q\rho_n^{-1}$ für alle $n \in \mathbb{N}_0$? Wenn ein solches Q existiert, ist es dann eindeutig?

Die Frage nach der Eindeutigkeit ist leicht zu beantworten, die nach der Existenz dagegen nicht.

Proposition 2.12. Für jedes $n \in \mathbb{N}_0$ sei Q_n ein Wahrscheinlichkeitsmaß auf $(E^{n+1}, \mathcal{E}^{n+1})$. Dann existiert höchstens ein Wahrscheinlichkeitsmaß Q auf $(E^{\mathbb{N}_0}, \mathcal{E}^{\mathbb{N}_0})$ so dass $Q_n := Q\rho_n^{-1}$ für alle $n \in \mathbb{N}_0$ gilt.

Beweis. Sind Q und \bar{Q} zwei solche Wahrscheinlichkeitsmaße, so stimmen sie nach Voraussetzung auf allen Mengen der Form $A_0 \times A_1 \times \dots \times A_m \times E \times E \times \dots$ mit $m \in \mathbb{N}_0$ und $A_0, \dots, A_m \in \mathcal{E}$ überein. Da dieses Mengensystem ein durchschnittsstabiler Erzeuger von $\mathcal{E}^{\mathbb{N}_0}$ ist, folgt $Q = \bar{Q}$. \square

Damit man überhaupt eine Chance hat, dass ein solches Q existiert, müssen die Q_n offensichtlich konsistent im Sinne der folgenden Definition sein.

Definition 2.13. Die Folge von Wahrscheinlichkeitsmaßen Q_n auf $(E^{n+1}, \mathcal{E}^{n+1})$, $n \in \mathbb{N}_0$ heißt *konsistent*, wenn für jedes $n \in \mathbb{N}$ gilt: $Q_{n-1} = Q_n\varphi_n^{-1}$.

Nun folgt (zunächst ohne Beweis) die Aussage des berühmten Kolmogorovschen Existenzsatzes.

Satz 2.14 (Existenzsatz von Kolmogorov). *Ist $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B})$ und ist $Q_n, n \in \mathbb{N}_0$ eine konsistente Folge von Wahrscheinlichkeitsmaßen (wie in der letzten Definition), so existiert (genau) ein Wahrscheinlichkeitsmaß Q auf $(E^{\mathbb{N}_0}, \mathcal{E}^{\mathbb{N}_0})$ so dass $Q_n = Q\rho_n^{-1}$ für alle $n \in \mathbb{N}_0$ gilt.*

Interessanterweise gilt die Aussage nicht nur für $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B})$, sondern deutlich allgemeiner (nämlich für separable vollständige metrische Räume ausgestattet mit ihrer Borel- σ -Algebra \mathcal{E}), aber - ebenfalls interessanterweise - nicht für beliebige Messräume (E, \mathcal{E}) . Diese negative Aussage kann man so interpretieren, dass es keine Möglichkeit gibt, Realisierungen der Wahrscheinlichkeitsmaße Q_n rekursiv zu simulieren (obwohl sich natürlich jedes einzelne Q_n simulieren lässt und man damit durch Restriktion automatisch auch eine Simulation aller Q_k mit $k < n$ hat).

Satz 2.14 lässt sich mit dem Satz von Ionescu-Tulcea zeigen, den wir ohne Beweis formulieren werden (ein Beweis findet sich z.B. im Skript von W. König).

Als Vorbereitung brauchen wir folgende Definition.

Definition 2.15. Es seien (E_1, \mathcal{E}_1) und (E_2, \mathcal{E}_2) Messräume und K ein Markovkern von E_1 nach E_2 , das heißt eine Abbildung $K : E_1 \times \mathcal{E}_2 \rightarrow [0, 1]$, so dass für jedes $e \in E_1$ $K(e, \cdot)$ ein Wahrscheinlichkeitsmaß auf (E_2, \mathcal{E}_2) ist und für jedes $A \in \mathcal{E}_2$ die Abbildung $e \mapsto K(e, A)$ messbar ist. Weiter sei μ ein Wahrscheinlichkeitsmaß auf (E_1, \mathcal{E}_1) . Dann ist das Wahrscheinlichkeitsmaß $\mu \otimes K$ auf $(E_1 \times E_2, \mathcal{E}_1 \otimes \mathcal{E}_2)$ definiert als

$$(\mu \otimes K)(A) = \int_{E_1} K(\omega_1, A_{\omega_1}^{(1)})\mu(d\omega_1),$$

wobei $A_{\omega_1}^{(1)} := \{\omega_2 \in E_2 : (\omega_1, \omega_2) \in A\}$ der ω_1 -Schnitt von A ist.

Dass diese Definition, die eine Verallgemeinerung des Produkts von Wahrscheinlichkeitsmaßen darstellt, sinnvoll ist und in der Tat ein Wahrscheinlichkeitsmaß definiert zeigt man so, wie man die Existenz von Produktmaßen zeigt (siehe Bemerkung 10.1.10. im Skript von W. König).

Wir betrachten nun wieder einen allgemeinen Messraum (E, \mathcal{E}) . Für jedes $n \in \mathbb{N}_0$ sein ein Markovkern K_n von $(E^{n+1}, \mathcal{E}^{n+1})$ nach (E, \mathcal{E}) gegeben. Weiter sei als "Startverteilung" das Wahrscheinlichkeitsmaß Q_0 auf (E, \mathcal{E}) gegeben. Wir definieren rekursiv die Wahrscheinlichkeitsmaße $Q_{n+1} := Q_n \otimes K_n$ für $n \in \mathbb{N}_0$.

Satz 2.16 (Satz von Ionescu Tulcea). *In der eben beschriebenen Situation existiert ein (eindeutiges) Wahrscheinlichkeitsmaß Q auf $(E^{\mathbb{N}_0}, \mathcal{E}^{\mathbb{N}_0})$ so dass $Q_n = Q\rho_n^{-1}$ für alle $n \in \mathbb{N}_0$ gilt.*

Interessanterweise gilt diese Aussage ohne Restriktion an den Messraum (E, \mathcal{E}) . Wir werden die Aussage nicht in der Vorlesung beweisen, aber eine Beweisidee sehen.

Man kann nun den Satz von Kolmogorov (Satz 2.14) mit Hilfe des Satzes von Ionescu Tulcea beweisen. Man muss dazu “nur” zeigen, dass sich im Fall $E = \mathbb{R}$ jedes der vorgegebenen (konsistenten) Wahrscheinlichkeitsmaße Q_n auf \mathbb{R}^{n+1} in der Form $Q_{n-1} \otimes K_{n-1}$ mit einem geeigneten Markovkern K_{n-1} darstellen lässt. Dies wurde in Wahrscheinlichkeitstheorie I im Kapitel 8 im Abschnitt über die Existenz regulärer bedingter Verteilungen gezeigt.