ETH zürich

Departement Mathematik

Wahrscheinlichkeitstheorie und Statistik

Philipp Harms, Josef Teichmann

Frühling 2025

Auf Basis von Skripten von Hans Föllmer, Hansruedi Künsch, Martin Schweizer und Sara van der Geer

Outline

1. Wahrscheinlichkeitsräume

Die Axiome von Kolmogorov

Sei $\Omega \neq \emptyset$ eine nichtleere Menge, $\mathcal{F}$ eine Menge von Teilmengen $A \subseteq \Omega$ und $\mathbb{P} : \mathcal{F} \to [0,1]$ eine Abbildung von $\mathcal{F}$ in das Einheitsintervall.

Die Elemente $\omega \in \Omega$ interpretieren wir als die (im Modell in Betracht gezogenen) **möglichen Fälle**, die Teilmengen $A \in \mathcal{F}$ als die (im Modell zugelassenen) **Ereignisse**, und für $A \in \mathcal{F}$ interpretieren wir die Zahl $\mathbb{P}(A)$ als die (im Modell angenommene) **Wahrscheinlichkeit des Ereignisses** A.

Die folgenden **Axiome von Kolmogorov** (1933) verlangen nun, dass die Kollektion $\mathcal{F}$ der Ereignisse abgeschlossen ist unter abzählbaren Mengenoperationen, und dass die Zuordnung $A \to \mathbb{P}(A)$ die Rechenregeln für Wahrscheinlichkeiten erfüllt, die uns von diskreten Modellen schon vertraut sind.

Axiome von Kolmogorov

Das Tripel $(\Omega, \mathcal{F}, \mathbb{P})$ heisst ein **Wahrscheinlichkeitsraum**, wenn gilt:

1) $\mathcal{F}$ ist eine σ-**Algebra**, d.h.

$$\begin{aligned} \Omega \ &\in \ \mathcal{F} \\ A \in \mathcal{F} \ &\Rightarrow \ A^c \in \mathcal{F} \\ A_1, A_2, \dots \in \mathcal{F} \ &\Rightarrow \ \bigcup_i A_i \in \mathcal{F} \end{aligned}$$

2) $\mathbb{P}$ ist ein **Wahrscheinlichkeitsmass**, d.h.

$$\begin{aligned} \mathbb{P}(\Omega) \ &= \ 1 \\ A_1, A_2, \dots \in \mathcal{F},\ A_i \cap A_j = \emptyset \ (i \neq j) \ &\Rightarrow \ \mathbb{P}\left(\bigcup_i A_i\right) = \sum_i \mathbb{P}(A_i) \end{aligned}$$

Frequentistische Interpretation

Wir zählen, wie oft in einem **wiederholten Experiment** mit möglichen Ausgängen Ω das Ereignis A eintritt:

n :	Anzahl der Wiederholungen des Experiments,
n_A :	Anzahl des Auftretens des Ereignisses A

Empirisch gesehen ist (unter gewissen Bedingungen) das Verhältnis $\frac{n_A}{n}$ für grossen n annähernd konstant, also annähernd eine feste Zahl $\mathbb{P}(A) \in [0,1]$. Die so definierte Abbildung $\mathbb{P}$ ist normiert und additiv:

$$\mathbb{P}(\Omega) \approx \frac{n_\Omega}{n} = 1, \qquad \mathbb{P}(A \cup B) \approx \frac{n_A}{n} + \frac{n_B}{n} \approx \mathbb{P}(A) + \mathbb{P}(B),$$

für disjunkte Ereignisse A und B.

Mass- und Integrationstheorie

Die Nützlichkeit von Kolmogorovs Axiomen ist unter anderem den guten Eigenschaften des entsprechenden Mass-Integrals zu verdanken.

Ohne auf die Definition von Messbarkeit einzugehen, sei vorausgeschickt: Das **Mass-Integral**

$$\int_\Omega f \; \mathbb{P} \equiv \int_\Omega f(\omega) \; \mathbb{P}(d\omega)$$

ist für jede messbare Funktion $f : \Omega \to \mathbb{R}$ definiert, die nicht-negativ ist oder $\int_\Omega |f| \; \mathbb{P}$ erfüllt.

In der Vorlesung brauchen wir dieses allgemeine Integral selten; meist reichen **Summen und Riemann-Integrale** als Spezialfälle aus. Trotzdem ist gut zu wissen, dass diese Spezialfälle ein kohärentes Ganzes ergeben.

2. Wahrscheinlichkeitsverteilungen

Gewichtsfunktionen

Eine **Gewichtsfunktion** ist eine Funktion $f : \Omega \to [0, 1]$ auf einem nicht-leeren **abzählbaren** Grundraum Ω, die normiert ist auf

$$\sum_{\omega \in \Omega} f(\omega) = 1.$$

In Folge interpretieren wir $f(\omega)$ als die Wahrscheinlichkeit, dass das Ereignis $\{\omega\}$ eintritt. Dazu muss die σ-Algebra alle diese Ereignisse enthalten. Wir nehmen also an, dass $\mathcal{F} = 2^{\Omega}$.

Diskrete Wahrscheinlichkeitsräume und Zufallsvariablen

Jede Gewichtsfunktion f bestimmt ein Wahrscheinlichkeitsmass

$$\mathbb{P} : 2^{\Omega} \to [0,1], \qquad \mathbb{P}(A) = \sum_{\omega \in A} f(\omega).$$

Umgekehrt ist jedes Wahrscheinlichkeitsmass $\mathbb{P} : 2^{\Omega} \to [0,1]$ von dieser Form mit

$$f : \Omega \to [0,1], \qquad f(\omega) = \mathbb{P}(\{\omega\}).$$

Wahrscheinlichkeitsmasse mit Gewichtsfunktion heissen **diskret**.

Diskrete Gleichverteilung

Sei Ω *endlich*. In vielen Situationen ist es sinnvoll, eine **konstante Gewichtsfunktion** f zu wählen (Indifferenzprinzip, Prinzip vom unzureichenden Grund, Symmetrie eines Würfels etc.). Es folgt

$$f(\omega) = \frac{1}{|\Omega|},$$

und das zugehörige Wahrscheinlichkeitsmass erfüllt

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl der günstigen Fälle}}{\text{Anzahl der möglichen Fälle}},$$

für jedes Ereignis $A \subseteq \Omega$. Man nennt $\mathbb{P}$ auch die **diskrete Gleichverteilung** oder das **Laplace-Modell**.

Bernoulli-Verteilung

Die **Bernoulli-Verteilung** mit Parameter $p \in [0,1]$ ist gegeben durch die Gewichtsfunktion

$$f : \{0,1\} \to [0,1], \qquad f(1) = p, \qquad f(0) = 1 - p.$$

Das entsprechende Wahrscheinlichkeitsmass $\mathbb{P}$ erfüllt

$$\mathbb{P}(\{1\}) = p, \qquad \mathbb{P}(\{0\}) = 1 - p.$$

Ereignisse in Euklidischen Räumen

Die **Borel σ-Algebra** auf $\mathbb{R}^d$ ist die kleinste σ-Algebra, welche alle Quader folgender Form enthält:

$$(a_1, b_1] \times \cdots \times (a_d, b_d].$$

Sie wird mit $\mathcal{B}(\mathbb{R}^d)$ bezeichnet und ihre Elemente heissen **Borel-Mengen** oder einfach Ereignisse in $\mathbb{R}^d$.

Standardmässig trägt $\mathbb{R}^d$ immer die Borel σ-Algebra. Probabilistisch gesprochen gilt somit jeder Quader, und allgemeiner auch jede Kugel und jede offene Menge, als Ereignis in $\mathbb{R}^d$. Ob offene, halb-offene oder geschlossene Intervalle verwendet werden, ist irrelevant.

Dichtefunktionen

Das kontinuierliche Analogon der Gewichtsfunktion ist die Dichtefunktion.

Eine **Dichtefunktion** ist eine messbare Funktion $f : \mathbb{R}^d \to [0, \infty)$ mit der Normiertheitseigenschaft

$$\int_{\mathbb{R}^d} f(x)\, dx = 1.$$

(Die Definition von Messbarkeit folgt in Abschnitt 4. Messbarkeit garantiert die Existenz des Integrals.)

Absolut stetige Verteilungen und Zufallsvariablen

Jede Dichtefunktion f definiert ein **Wahrscheinlichkeitsmass** auf der Borel σ-Algebra $\mathcal{B}(\mathbb{R}^d)$:

$$\mathbb{P} : \mathcal{B}(\mathbb{R}^d) \to [0,1], \qquad \mathbb{P}(A) = \int_A f(x)\,dx$$

Wahrscheinlichkeitsmasse mit Dichtefunktion heissen **absolut stetig**.

Kontinuierliche Gleichverteilung

Als Dichtefunktion f betrachten wir die Indikatorfunktion des **Einheitswürfels**,

$$f : \mathbb{R}^d \to [0, \infty), \qquad f(x) = \begin{cases} 1, & x \in [0,1]^d, \\ 0, & \text{sonst.} \end{cases}$$

Das zugehörige Wahrscheinlichkeitsmass $\mathbb{P} : \mathcal{B}(\mathbb{R}^d) \to [0,1]$ bemisst Quader mit deren **Volumen**:

$$\mathbb{P}\left([a_1, b_1] \times \cdots \times [a_d, b_d]\right) = |b_1 - a_1| \cdots |b_d - a_d|, \qquad a, b \in [0,1]^d.$$

Durch Einschränken des Definitionsbereichs von $\mathbb{P}$ auf Teilmengen von $[0,1]^d$ erhält man ein Wahrscheinlichkeitsmass auf $[0,1]^d$, auch **kontinuierliche Gleichverteilung** genannt. Analog erhält man auch Gleichverteilungen auf anderen Borel-Mengen als $[0,1]^d$, unter geeigneter Normierung der Dichtefunktion.

Exkurs: Nicht-messbare Mengen

Man kann $\mathbb{P}$ nicht auf die Potenzmenge von $[0,1]^d$ erweitern, ohne dass σ-Additivität verloren ginge.

Ab Dimension $d = 3$ geht sogar Additivität verloren.

Es braucht also eine Unterscheidung zwischen messbaren Ereignissen und allgemeineren nicht-messbaren Mengen.

(Diese Aussagen verwenden das Auswahlaxiom. Ohne Auswahlaxiom lässt sich $\mathbb{P}$ zu einem Wahrscheinlichkeitsmass auf der Potenzmenge erweitern.)

Nullmengen

Eine **Nullmenge** ist ein Ereignis mit Wahrscheinlichkeit 0, oder eine Teilmenge eines solchen Ereignisses, welcher ebenfalls Wahrscheinlichkeit 0 zugeschrieben wird.

Unter absolut stetigen Verteilungen sind alle ein-elementigen Mengen Nullmengen, und wegen σ-Additivität auch alle **abzählbaren Mengen**:

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{\omega \in A} \{\omega\}\right) = \sum_{\omega \in A} \mathbb{P}\big(\{\omega\}\big) = 0, \qquad \text{für } A \text{ abzählbar.}$$

Insbesondere hat die Menge aller Punkte mit rationalen Koordinaten Wahrscheinlichkeit 0. Bei der **Simulation** von ‘Zufallszahlen’ auf einem Computer sieht das aber natürlich anders aus!

Verteilungsfunktion und Quantilfunktion

Sei $\mathbb{P}$ ein Wahrscheinlichkeitsmass auf den reellen Zahlen.

Die **Verteilungsfunktion** von $\mathbb{P}$ ist

$$F : \mathbb{R} \to [0,1], \qquad F(x) = \mathbb{P}\left((-\infty, x]\right).$$

Die **Quantilfunktion** von $\mathbb{P}$ ist

$$Q : (0,1) \to \mathbb{R}, \qquad Q(p) = \inf\{x \in \mathbb{R} \colon F(x) \geq p\}.$$

Charakterisierung von Verteilungsfunktionen

Jede Verteilungsfunktion $F : \mathbb{R} \to [0, 1]$ ist nicht-fallend, rechts-stetig und erfüllt

$$\lim_{x \to -\infty} F(x) = 0, \quad \lim_{x \to +\infty} F(x) = 1.$$

Umgekehrt ist jede solche Funktion die Verteilungsfunktion eines Wahrscheinlichkeitsmasses.

Stetigkeitseigenschaften

Wahrscheinlichkeitsmasse sind stetig entlang monoton wachsender Folgen von Ereignissen:

$$A_1 \subseteq A2 \subseteq \cdots \Longrightarrow \lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcup_n A_n\right).$$

Dies folgt aus der σ-Additivität. Durch Komplement-Bildung erhält man analog:

$$A_1 \supseteq A2 \supseteq \cdots \Longrightarrow \lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_n A_n\right),$$

Für Verteilungsfunktionen folgt daraus

$$\lim_{n\to\infty} F(x+1/n) = \lim_{n\to\infty} \mathbb{P}\left((-\infty, x+1/n]\right) = \mathbb{P}\left(\bigcap_n (-\infty, x+1/n]\right) = \mathbb{P}\left((-\infty, x]\right) = F(x),$$

$$\lim_{n\to\infty} F(x-1/n) = \lim_{n\to\infty} \mathbb{P}\left((-\infty, x-1/n]\right) = \mathbb{P}\left(\bigcup_n (-\infty, x-1/n]\right) = \mathbb{P}\left((-\infty, x)\right) \leqslant F(x).$$

Absolutstetige Verteilungsfunktionen

Das Wahrscheinlichkeitsmass $\mathbb{P}$ ist **absolut stetig** genau dann, wenn die Verteilungsfunktion F absolut stetig ist. Dann ist F die Stammfunktion der Dichtefunktion f von $\mathbb{P}$:

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(x)\,dx && \text{für alle } x, \\ f(x) &= F'(x), && \text{für fast alle } x. \end{aligned}$$

Zusammenhang zwischen Verteilungs- und Quantilfunktionen

Die Verteilungs- und Quantilfunktionen **bestimmen sich gegenseitig** eindeutig durch die Beziehung

$$p \leq F(x) \Longleftrightarrow Q(p) \leq x, \qquad \text{für alle } x \in \mathbb{R}, p \in (0,1).$$

Wenn F stetig und strikt monoton wachsend ist, dann ist Q die **Umkehrfunktion** von F, d.h.

$$F(Q(p)) = p, \qquad Q(F(x)) = x, \qquad \text{für alle } p \in (0,1), x \in \mathbb{R}.$$

Dirac-Verteilung

Sei Ω eine beliebige nicht-leere Menge mit beliebiger σ-Algebra $\mathcal{F}$.

Das **Dirac-Mass** mit Parameter $\omega \in \Omega$ ist definiert als

$$\mathbb{P}(A) = \mathbb{1}_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$$

Im diskreten Fall Ω abzählbar und $\mathcal{F} = 2^\Omega$ entspricht das der Indikatorfunktion $\mathbb{1}_{\{\omega\}}$ als Gewichtsfunktion. Im kontinuierlichen Fall $\Omega = \mathbb{R}^d$ und $\mathcal{F} = \mathcal{B}(\mathbb{R}^d)$ ist das Dirac-Mass ein Beispiel eines **nicht-absolut stetigen** Wahrscheinlichkeitsmasses (sonst wäre $\mathbb{P}(\{\omega\}) = 0$).

Die **Verteilungsfunktion** der Dirac-Verteilung auf $\Omega = \mathbb{R}$ springt bei ω.

3. Kombinatorik

Stichproben

Wir zählen verschiedene Arten von k-elementigen **Stichproben** aus einer n-elementigen Menge M:

Anzahl k-elementiger Stichprobem	geordnet (Variation)	ungeordnet (Kombination)
Ohne Wiederholung	$\frac{n!}{(n-k)!}$ für $k \leq n$	$\binom{n}{k}$ für $k \leq n$
Mit Wiederholung	n^k	$\binom{n+k-1}{k}$

Über das Prinzip *'günstige Fälle durch mögliche Fälle'* erhält man daraus Wahrscheinlichkeiten von verschiedenen Stichproben-Mengen unter dem **Laplace-Modell**.

Dasselbe Prinzip kommt auch zur Anwendung, wenn man den Computer zählen lässt, allerdings braucht man dazu keine mathematische Analysis. Daher belassen wir es hier bei einem Überblick.

Geordnete Stichproben ohne Wiederholung

Eine **geordnete Stichprobe ohne Wiederholung** ist ein Tupel

$$(x_1, \ldots, x_k) \quad \text{aus paarweise verschiedenen Elementen } x_i \in M,$$

oder, äquivalent, eine injektive Abbildung $x : \{1, \ldots k\} \to M$.

Die **Anzahl** solcher Stichproben ist

$$n(n-1)\ldots(n-k+1) = \frac{n!}{(n-k)!}, \qquad \text{falls } k \leq n, \text{ und } 0 \text{ sonst.}$$

Beweis: Zuerst n Möglichkeiten für x_1, dann $n-1$ Möglichkeiten für x_2, etc.

Spezialfall: Für $M = \{1, \ldots, k\}$ erhält man **Permutationen** und es gibt $k!$ davon.

Beispiele: Geordnete Stichproben ohne Wiederholung

- Die Anzahl möglicher Podiumsplatzierungen (gold, silber, bronze) in einem Rennen mit 10 Bewerbern ist $10 * 9 * 8$.
- Wenn das Pauli-Prinzip gilt (keine zwei Elektronen im selben Orbital), können 2 Elektronen in 5 Orbitalen auf $5 * 4$ Möglichkeiten angeordnet werden.

Geordnete Stichproben mit Wiederholung

Eine **geordnete Stichprobe mit Wiederholung** ist ein Tupel

$$(x_1, \ldots, x_k) \quad \text{aus Elementen } x_i \in M,$$

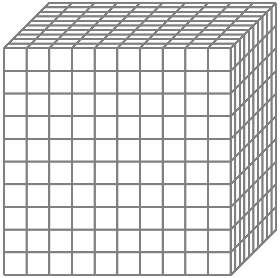
oder, äquivalent, eine Abbildung $x : \{1, \ldots k\} \to M$.

Die **Anzahl** solcher Stichproben ist n^k.

Beweis: Zuerst n Möglichkeiten für x_1, dann n Möglichkeiten für x_2, etc.

Beispiele: Geordnete Stichproben mit Wiederholung

- Es gibt $(2 * 26 + 10)^4$ alphanumerische Passwörter der Länge 4.
- Die Potenzmenge einer k-elementigen Menge hat 2^k Elemente.
- Die Finite-Elemente-Methode auf $[0, 1]^d$ mit Rechtecksgitter von Feinheit $1/n$ braucht n^d Elemente. Das skaliert schlecht in n und noch schlechter in d (curse of dimensionality).



Ungeordnete Stichproben ohne Wiederholung

Eine **ungeordnete Stichprobe ohne Wiederholung** ist eine Menge

$$\{x_1, \dots, x_k\} \quad \text{mit paarweise verschiedenen } x_i \in M$$

oder, äquivalent, eine Abbildung $w : M \to \{0, 1\}$ mit $\sum_{x \in M} w(x) = k$.

Die **Anzahl** solcher Stichproben ist

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \qquad \text{falls } k \leq n, \text{ und } 0 \text{ sonst.}$$

Beweis: Das sind $1/k!$ mal weniger als bei geordneten Stichproben ohne Wiederholungen, weil es $k!$ Permutationen von $\{1, \dots, k\}$ gibt.

Beispiele: Ungeordnete Stichproben ohne Wiederholungen

- Beim Schweizer Lotto '6 aus 42' ist die Wahrscheinlichkeit für einen 6er der Kehrwert von $\binom{42}{6}$, falls wir wie im Laplace-Modell annehmen, dass alle Tipps gleich wahrscheinlich sind.
- Eine genomweite Assoziationsstudie untersucht die Assoziation von Brustkrebs mit 10^6 genetischen Variationen (SNPs). Wenn Kombinationen von k SNPs betrachtet werden, wächst die Anzahl der zu testenden Hypothesen auf $\binom{10^6}{k}$. (Achtung: Multiple Testing!)

Ungeordnete Stichproben mit Wiederholung

Eine **ungeordnete Stichprobe mit Wiederholung** ist eine Multimenge[1]

$$\{\!|x_1, \ldots, x_k|\!\} \quad \text{mit } x_i \in M$$

oder, äquivalent, eine Abbildung $w : M \to \mathbb{N}$ mit $\sum_{x \in M} w(x) = k$.

Die **Anzahl** solcher Stichproben ist

$$\binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-1)!}.$$

Beweis: Jede Multimenge entspricht genau einem Tupel aus k Platzhaltern • und $n-1$ Trennstrichen $|$. Zum Beispiel bedeutet $\bullet\bullet||\bullet$ zwei Mal das erste Element plus ein Mal das dritte Element von M. Nun verwendet man das Ergebnis der vorherigen Folie.

[1] https://de.wikipedia.org/wiki/Multimenge

Beispiele: Ungeordnete Stichproben mit Wiederholung

- Wenn 10 Eissorten zur Auswahl stehen, gibt es $\binom{12}{3}$ verschiedene Eisbecher mit 3 Kugeln Eis.
- Die Gleichung $x_1 + x_2 + x_3 + x_4 = 18$ hat $\binom{21}{18}$ Lösungen in den natürlichen Zahlen. (Hier ist $n = 4$ und $k = 18$.)

4. Zufallsvariablen

Messbarkeit

Wir betrachten σ-Algebren $\mathcal{F}$ und $\mathcal{F}'$ auf nicht-leeren Mengen Ω bzw. Ω'. Ereignisse, also Mengen in $\mathcal{F}$ bzw. $\mathcal{F}'$, werden in der Masstheorie auch **messbar** genannt.

Eine Funktion $X : \Omega \to \Omega'$ heisst **messbar**, wenn Urbilder messbarer Mengen wieder messbar sind, d.h.

$$X^{-1}(A') \equiv \left\{\omega \in \Omega : X(\omega) \in A'\right\} \in \mathcal{F}, \qquad \text{für alle } A' \in \mathcal{F}'.$$

Die **Komposition** messbarer Abbildungen ist messbar. Das Kartesische **Produkt** messbarer Abbildungen ist messbar genau dann, wenn jede einzelne Abbildung messbar ist.

Messbarkeit in Euklidischen Räumen

Für Funktionen mit Werten in Euklidischen Räumen ist äquivalent:

- Urbilder messbarer Mengen sind messbar
- Urbilder offener Mengen sind messbar
- Urbilder von Quadern sind messbar

Für Funktionen zwischen Euklidischen Räumen gilt:

- Stetigkeit impliziert Messbarkeit
- Die Umkehrung gilt nicht

Messbarkeit bleibt erhalten unter Linearkombinationen und punktweise Limiten.

Zufallsvariablen

Eine **Zufallsvariable** ist eine messbare Funktion auf einem Wahrscheinlichkeitsraum. Oft wird implizit angenommen, sofern nicht anders angegeben, dass die Zielmenge $\mathbb{R}$ ist.

Notation: Für Ereignisse B in der Zielmenge einer Zufallsvariable X schreibt man kurz

$$\begin{aligned} \{X \in B\} &\text{ für } \{\omega \in \Omega : X(\omega) \in B\}, \\ \mathbb{P}(X \in B) &\text{ für } \mathbb{P}(\{\omega \in \Omega : X(\omega) \in B\}), \qquad \text{etc.} \end{aligned}$$

Wir interpretieren Zufallsvariablen als zufallsabhängige **Beobachtungen**. Die Indikatorfunktion eines Ereignisses ist eine Zufallsvariable. Somit gelten auch Ereignisse als beobachtbar.

Verteilung von Zufallsvariablen

Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $(\Omega', \mathcal{F}')$ eine nicht-leere Menge mit einer σ-Algebra.

Die **Verteilung einer Zufallsvariable** $X : \Omega \to \Omega'$ ist das transformierte Wahrscheinlichkeitsmass, auch Bildmass genannt,

$$\mathbb{P}_X := \mathbb{P} \circ X^{-1} : \mathcal{F}' \to [0,1], \qquad \mathbb{P}_X(A') = \mathbb{P}(X^{-1}(A')).$$

Ein Zufallsvariable heisst **diskret** bzw. **absolut stetig**, wenn ihre Verteilung diskret bzw. absolut stetig ist. Analog ist die **Gewichtsfunktion**, **Dichtefunktion**, **Verteilungsfunktion** bzw. **Quantilsfunktion** einer Zufallsvariablen jene ihrer Verteilung.

In der Statistik interessiert man sich häufig mehr für die Verteilung einer Zufallsvariable als für den Grundraum und die Zufallsvariable selbst.

Gemeinsame Verteilung und Randverteilungen

Seien X und Y Zufallsvariablen. Die Verteilung von (X, Y) heisst **gemeinsame Verteilung** und die Verteilungen von X und Y heissen **Randverteilungen**.

Sind X und Y diskret, dann auch (X, Y) und es gilt für die jeweiligen **Gewichtsfunktionen**:

$$f_X(x) = \sum_y f_{X,Y}(x, y), \qquad f_Y(y) = \sum_x f_{X,Y}(x, y).$$

Ist (X, Y) absolut stetig, dann auch X und Y, und es gilt für die jeweiligen **Dichtefunktionen**:

$$f_X(x) = \int f_{X,Y}(x, y)\, dy, \qquad f_Y(y) = \int f_{X,Y}(x, y)\, dx.$$

Transformationen von Zufallsvariablen

Die **Transformation** $Y = g \circ X$ einer Zufallsvariablen X durch eine messbare Funktion g ist wieder eine Zufallsvariable.

$$\begin{tikzcd} \Omega \arrow[r, "X"] \arrow[dr, "Y"'] & \Omega' \arrow[d, "g"] \\ & \Omega'' \end{tikzcd}$$

Sei $\sigma(X)$ die von X **erzeugte** σ-Algebra, d.h. die kleinste σ-Algebra auf Ω bezüglich der X messbar ist:

$$\sigma(X) = \left\{ X^{-1}(A') : A' \in \mathcal{F}' \right\}.$$

Alle Transformationen von X sind $\sigma(X)$-messbar. Für Ω'' Euklidisch gilt auch die Umkehrung.

Beispiel

Die **Parität** der Augenzahl eines Würfels ist die Zufallsvariable

$$X : \{1, \ldots, 6\} \to \{0, 1\}, \qquad X(\omega) = \omega \mod 2.$$

Die von X erzeugte σ-Algebra ist

$$\sigma(X) = \big\{\emptyset, \{1, 3, 5\}, \{2, 4, 6\}, \{\Omega\}\big\}.$$

Die Transformationen von X sind genau jene Funktionen, die auf den sogenannten **Atomen** $\{1, 3, 5\}$ und $\{2, 4, 6\}$ konstant sind.

Jede solche Funktion ist $\sigma(X)$-messbar, und jede $\sigma(X)$-messbare Funktion mit Werten in einem Euklidischen Raum ist konstant auf diesen Atomen.

Transformationen von Verteilungen

Wenn X diskret ist, dann ist auch $Y := g \circ X$ diskret und die **Gewichtsfunktionen** erfüllen

$$f_Y(y) = \mathbb{P}(Y = y) = \mathbb{P}(X = g^{-1}(y)) = f_X(g^{-1}(y)).$$

Wenn X absolut stetig ist, $g : \mathbb{R}^n \to \mathbb{R}^n$ eine Umkehrfunktion g^{-1} hat und $\det g' > 0$, dann ist auch $Y := g \circ X$ absolut stetig und die **Dichtefunktionen** erfüllen

$$f_Y(y) = \frac{f_X(g^{-1}(y)}{\det g'(g^{-1}(y))}.$$

Wenn $g : \mathbb{R} \to \mathbb{R}$ eine Umkehrfunktion g^{-1} hat, dann erfüllen die **Verteilungsfunktionen** von X und Y

$$F_Y(b) = \mathbb{P}\left(g(X) \leq b\right) = \mathbb{P}\left(X \leq g^{-1}(b)\right) = F_X(g^{-1}(b)).$$

Beispiel: Verteilung von Summen

Sei $X = (X_1, X_2)$ eine zweidimensionale absolut stetige Zufallsvariable und $Y = g \circ X$ mit

$$g : \mathbb{R}^2 \to \mathbb{R}^2, \qquad g(x_1, x_2) = (x_1, x_1 + x_2), \qquad g^{-1}(y_1, y_2) = (y_1, y_2 - y_1).$$

Aufgrund des Transformationssatzes für Dichten gilt

$$f_Y(y_1, y_2) = f_X(y_1, y_2 - y_1)$$

Die Verteilung von $X_1 + X_2 = Y_2$ erhält man als Randverteilung von Y:

$$f_{X_1+X_2}(x) = f_{Y_2}(x) = \int_{-\infty}^{\infty} f_X(y, x - y)\, dy.$$

Wenn X_1 und X_2 unabhängig sind (siehe Abschnitt 5), dann ist f_X das Produkt der Randdichten f_{X_1} und f_{X_2}, und folglich ist $f_{X_1+X_2}$ die Faltung dieser Randdichten:

$$f_{X_1+X_2}(x) = \int_{-\infty}^{\infty} f_{X_1}(y) f_{X_2}(x - y) dx_1.$$

5. Unabhängigkeit

Philipp Harms, Josef Teichmann
philipp.harms@math.ethz.ch , josef.teichmann@math.ethz.ch

ETH Zürich
Departement Mathematik