Stochastik für Informatiker

– 6. Vorlesung –

Prof. Dr. Holger Kösters

Institut für Mathematik, MNF, Universität Rostock

28. November 2019

Kapitel 2
Allgemeine Wahrscheinlichkeitsräume

Kapitel 2.1
Allgemeine Wahrscheinlichkeitsräume

Motivation

Wir wollen uns nun zufallsabhängigen Vorgängen zuwenden, bei denen sich die Menge der möglichen Versuchsausgänge durch ein Intervall $(a, b) \subseteq \mathbb{R}$ mit $a < b$, also durch eine überabzählbare Menge beschreiben lässt.
Diese Situation tritt häufig dann auf, wenn kontinuierliche Größen (Zeiten, Längen, . . .) gemessen werden, die vom Zufall beeinflusst sind.

Motivation

Wir wollen uns nun zufallsabhängigen Vorgängen zuwenden, bei denen sich die Menge der möglichen Versuchsausgänge durch ein Intervall $(a, b) \subseteq \mathbb{R}$ mit $a < b$, also durch eine <u>überabzählbare</u> Menge beschreiben lässt.
Diese Situation tritt häufig dann auf, wenn kontinuierliche Größen (Zeiten, Längen, ...) gemessen werden, die vom Zufall beeinflusst sind.

Ansatz: Diskretisierung

$$\begin{array}{ccccccc} \text{kontinuierliches} & & \text{diskretes} & & \text{diskretes} & & \text{kontinuierliches} \\ \text{Problem} & \rightsquigarrow & \text{Problem} & \rightsquigarrow & \text{Modell} & \rightsquigarrow & \text{Modell} \end{array}$$

Motivation

Wir wollen uns nun zufallsabhängigen Vorgängen zuwenden, bei denen sich die Menge der möglichen Versuchsausgänge durch ein Intervall $(a, b) \subseteq \mathbb{R}$ mit $a < b$, also durch eine überabzählbare Menge beschreiben lässt. Diese Situation tritt häufig dann auf, wenn kontinuierliche Größen (Zeiten, Längen, ...) gemessen werden, die vom Zufall beeinflusst sind.

Ansatz: Diskretisierung

$$\begin{array}{c} \text{kontinuierliches} \\ \text{Problem} \end{array} \rightsquigarrow \begin{array}{c} \text{diskretes} \\ \text{Problem} \end{array} \rightsquigarrow \begin{array}{c} \text{diskretes} \\ \text{Modell} \end{array} \rightsquigarrow \begin{array}{c} \text{kontinuierliches} \\ \text{Modell} \end{array}$$

(*Bemerkung:* Dieser Ansatz dient hier nur der Motivation. Anschließend werden wir immer „direkt" ein kontinuierliches Modell wählen.)

Beispiel 2.1 (a) : Zufälliger Punkt

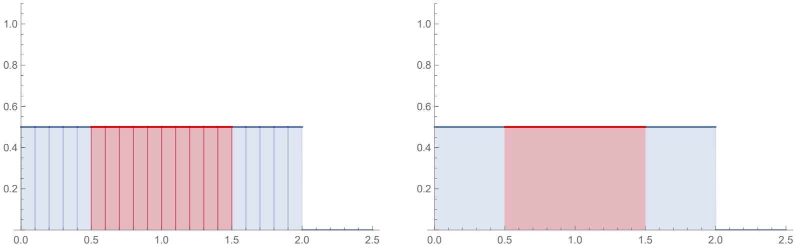
Es soll ein zufälliger Punkt aus dem Intervall $]0, T]$ ausgewählt werden, wobei $T \in \mathbb{N}$ fest und Intervalle $I \subseteq]0, T]$ derselben Länge dieselbe W. besitzen sollen.

Beispiel 2.1 (a) : Zufälliger Punkt

Es soll ein zufälliger Punkt aus dem Intervall $]0, T]$ ausgewählt werden, wobei $T \in \mathbb{N}$ fest und Intervalle $I \subseteq]0, T]$ derselben Länge dieselbe W. besitzen sollen. Wir unterteilen das Intervall $]0, T]$ (für großes $n \in \mathbb{N}$) in Abschnitte der Länge $\frac{1}{n}$. Dann bietet sich zur Beschreibung der diskrete W.raum $(\Omega_n, \mathfrak{P}(\Omega_n), \mathbb{P}_n)$ mit $\Omega_n := \{1, \ldots, Tn\}$ und $\mathbb{P}_n := \mathcal{U}(\Omega_n)$ (d.h. $f_n(k) := \frac{1}{Tn}\ \forall\, k \in \Omega_n$) an.

Beispiel 2.1 (a) : Zufälliger Punkt

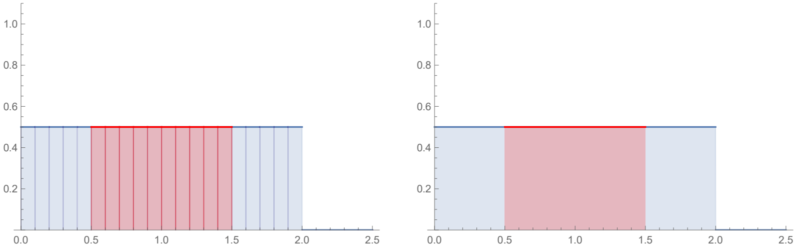
Es soll ein zufälliger Punkt aus dem Intervall $]0, T]$ ausgewählt werden, wobei $T \in \mathbb{N}$ fest und Intervalle $I \subseteq]0, T]$ derselben Länge dieselbe W. besitzen sollen.
Wir unterteilen das Intervall $]0, T]$ (für großes $n \in \mathbb{N}$) in Abschnitte der Länge $\frac{1}{n}$.
Dann bietet sich zur Beschreibung der diskrete W.raum $(\Omega_n, \mathfrak{P}(\Omega_n), \mathbb{P}_n)$ mit $\Omega_n := \{1, \dots, Tn\}$ und $\mathbb{P}_n := \mathcal{U}(\Omega_n)$ (d.h. $f_n(k) := \frac{1}{Tn} \; \forall\, k \in \Omega_n$) an.
Die W. eines Intervalls $I \subseteq]0, T]$ entspricht dann *näherungsweise* der Summe $\sum_{k \in \Omega_n : k/n \in I} f_n(k)$ bzw. (für $n \to \infty$) dem Integral $\int_I \frac{1}{T}\, dx$:



Veranschaulichung der Approximation für $T = 2,\ n = 10,\ I =]0.5, 1.5]$

Beispiel 2.1 (a) : Zufälliger Punkt

Es soll ein zufälliger Punkt aus dem Intervall $]0, T]$ ausgewählt werden, wobei $T \in \mathbb{N}$ fest und Intervalle $I \subseteq]0, T]$ derselben Länge dieselbe W. besitzen sollen.
Wir unterteilen das Intervall $]0, T]$ (für großes $n \in \mathbb{N}$) in Abschnitte der Länge $\frac{1}{n}$.
Dann bietet sich zur Beschreibung der diskrete W.raum $(\Omega_n, \mathfrak{P}(\Omega_n), \mathbb{P}_n)$ mit $\Omega_n := \{1, \ldots, Tn\}$ und $\mathbb{P}_n := \mathcal{U}(\Omega_n)$ (d.h. $f_n(k) := \frac{1}{Tn} \; \forall \, k \in \Omega_n$) an.
Die W. eines Intervalls $I \subseteq]0, T]$ entspricht dann *näherungsweise* der Summe $\sum_{k \in \Omega_n : k/n \in I} f_n(k)$ bzw. (für $n \to \infty$) dem Integral $\int_I \frac{1}{T} \, dx$:



Veranschaulichung der Approximation für $T = 2, \; n = 10, \; I =]0.5, 1.5]$

Damit können wir für $n \to \infty$ die Grundmenge $\Omega =]0, T]$ und die Funktion $f(x) = \frac{1}{T}$ wählen und die W. von I durch Integration von f über I berechnen.

Beispiel 2.1 (b) : Wartezeitproblem

Wir warten an einer (wenig befahrenen) Straße auf das nächste Auto, wobei wir die Zeit in Stunden messen.

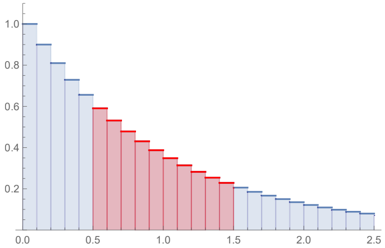
Beispiel 2.1 (b) : Wartezeitproblem

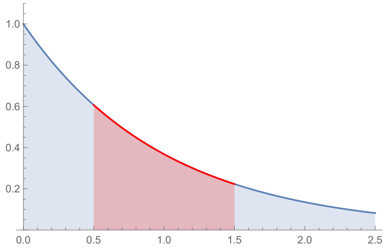
Wir warten an einer (wenig befahrenen) Straße auf das nächste Auto, wobei wir die Zeit in Stunden messen. Wir unterteilen das Intervall $(0, \infty)$ (für großes $n \in \mathbb{N}$) in Abschnitte der Länge $\frac{1}{n}$.

Beispiel 2.1 (b) : Wartezeitproblem

Wir warten an einer (wenig befahrenen) Straße auf das nächste Auto, wobei wir die Zeit in Stunden messen. Wir unterteilen das Intervall $(0, \infty)$ (für großes $n \in \mathbb{N}$) in Abschnitte der Länge $\frac{1}{n}$. Dann bietet sich zur Beschreibung (unter den Annahmen aus Bem. 1.28) der diskrete W.raum $(\Omega_n, \mathfrak{P}(\Omega_n), \mathbb{P}_n)$ mit $\Omega_n := \{1, 2, 3, \dots\}$ und $\mathbb{P}_n := \mathcal{G}(\frac{\lambda}{n})$ (d.h. $f_n(k) = (1 - \frac{\lambda}{n})^{k-1} \frac{\lambda}{n} \; \forall \, k \in \Omega_n$) an.

Wir warten an einer (wenig befahrenen) Straße auf das nächste Auto, wobei wir die Zeit in Stunden messen. Wir unterteilen das Intervall $(0, \infty)$ (für großes $n \in \mathbb{N}$) in Abschnitte der Länge $\frac{1}{n}$. Dann bietet sich zur Beschreibung (unter den Annahmen aus Bem. 1.28) der diskrete W.raum $(\Omega_n, \mathfrak{P}(\Omega_n), \mathbb{P}_n)$ mit $\Omega_n := \{1, 2, 3, \dots\}$ und $\mathbb{P}_n := \mathcal{G}(\frac{\lambda}{n})$ (d.h. $f_n(k) = (1 - \frac{\lambda}{n})^{k-1} \frac{\lambda}{n} \; \forall\, k \in \Omega_n$) an. Die W. eines Intervalls $I \subseteq (0, \infty)$ entspricht dann *näherungsweise* der Summe $\sum_{k \in \Omega_n : k/n \in I} f_n(k)$ bzw. (für $n \to \infty$) dem Integral $\int_I \lambda e^{-\lambda x} \, dx$:

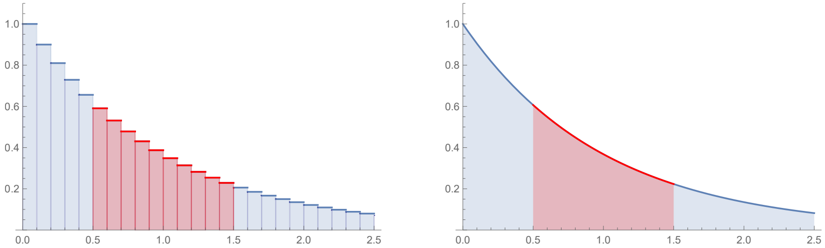




Veranschaulichung der Approximation für $\lambda = 1$, $n = 10$, $I =]0.5, 1.5]$

Beispiel 2.1 (b) : Wartezeitproblem

Wir warten an einer (wenig befahrenen) Straße auf das nächste Auto, wobei wir die Zeit in Stunden messen. Wir unterteilen das Intervall $(0, \infty)$ (für großes $n \in \mathbb{N}$) in Abschnitte der Länge $\frac{1}{n}$. Dann bietet sich zur Beschreibung (unter den Annahmen aus Bem. 1.28) der diskrete W.raum $(\Omega_n, \mathfrak{P}(\Omega_n), \mathbb{P}_n)$ mit $\Omega_n := \{1, 2, 3, \ldots\}$ und $\mathbb{P}_n := \mathcal{G}(\frac{\lambda}{n})$ (d.h. $f_n(k) = (1 - \frac{\lambda}{n})^{k-1} \frac{\lambda}{n} \;\forall\, k \in \Omega_n$) an. Die W. eines Intervalls $I \subseteq (0, \infty)$ entspricht dann *näherungsweise* der Summe $\sum_{k \in \Omega_n : k/n \in I} f_n(k)$ bzw. (für $n \to \infty$) dem Integral $\int_I \lambda e^{-\lambda x}\, dx$:



Veranschaulichung der Approximation für $\lambda = 1$, $n = 10$, $I =]0.5, 1.5]$

Damit können wir für $n \to \infty$ die Grundmenge $\Omega = (0, \infty)$ und die Funktion $f(x) = \lambda e^{-\lambda x}$ wählen und die W. von I durch Integration von f über I berechnen.

Beziehung zwischen Riemann-Integralen und Summen:
Ist f auf $]a,b]$ Riemann-integrierbar, so gilt für großes $n \in \mathbb{N}$

$$\int_a^b f(x)\,dx \approx \sum_{k \in \mathbb{Z}:\, \frac{k}{n} \in]a,b]} f\left(\tfrac{k}{n}\right) \tfrac{1}{n}\,.$$

Beziehung zwischen Riemann-Integralen und Summen:
Ist f auf $]a, b]$ Riemann-integrierbar, so gilt für großes $n \in \mathbb{N}$

$$\int_a^b f(x)\,dx \approx \sum_{k \in \mathbb{Z}\,:\, \frac{k}{n} \in]a,b]} f\left(\tfrac{k}{n}\right) \tfrac{1}{n}\,.$$

Da sich mit Integralen i. d. R. einfacher rechnen lässt als mit Summen, wollen wir von nun an immer direkt ein kontinuierliches Modell wählen.

Merkregel:
Gegeben ein Intervall $\Omega \subseteq \mathbb{R}$ und eine Funktion $f : \Omega \to \mathbb{R}$ mit

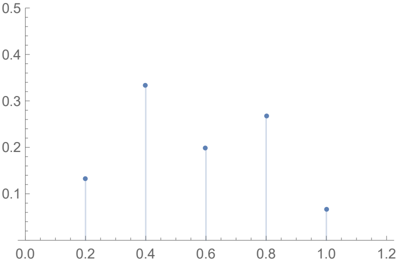
$$f(x) \geq 0 \text{ für alle } x \in \Omega \quad \text{und} \quad \int_\Omega f(x)\,dx = 1$$

(wobei die Existenz des Riemann-Integrals vorauszusetzen ist) können wir das stetige W.maß $\mathbb{P}$ betrachten, das jedem Intervall $I \subseteq \Omega$ die W.

$$\mathbb{P}(I) = \int_I f(x)\,dx$$

zuordnet. Die Funktion f wird auch als W.dichte von $\mathbb{P}$ bezeichnet.

Diskrete Modelle

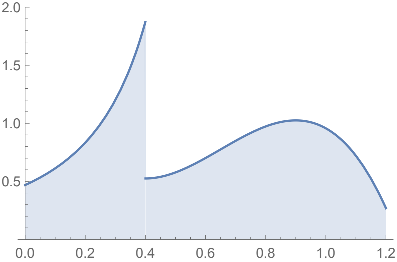


Ω abzählbar, $f : \Omega \to \mathbb{R}$,
$f(\omega) \geq 0 \ \forall \omega \in \Omega$, $\sum_{\omega \in \Omega} f(\omega) = 1$.

$$\mathbb{P}(A) := \sum_{\omega \in A} f(\omega) \quad (A \subseteq \Omega \text{ abzählbar})$$

diskretes W.maß (mit der W.dichte f)

Stetige Modelle



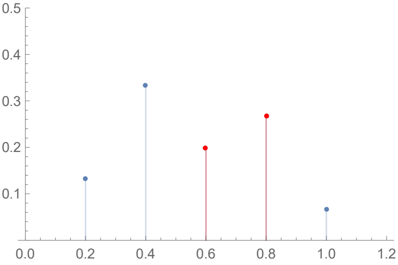
Ω Intervall, $f : \Omega \to \mathbb{R}$,
$f(x) \geq 0 \ \forall x \in \Omega$, $\int_\Omega f(x)\,dx = 1$.

$$\mathbb{P}(A) := \int_A f(x)\,dx \quad (A \subseteq \Omega \text{ Intervall})$$

stetiges W.maß (mit der W.dichte f)

Diskrete Modelle vs. Stetige Modelle

Diskrete Modelle

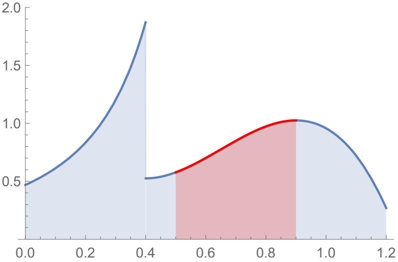


Ω abzählbar, $f : \Omega \to \mathbb{R}$,
$f(\omega) \geq 0 \; \forall \omega \in \Omega$, $\sum_{\omega \in \Omega} f(\omega) = 1$.

$$\mathbb{P}(A) := \sum_{\omega \in A} f(\omega) \quad (A \subseteq \Omega \text{ abzählbar})$$

diskretes W.maß (mit der W.dichte f)

Stetige Modelle



Ω Intervall, $f : \Omega \to \mathbb{R}$,
$f(x) \geq 0 \; \forall x \in \Omega$, $\int_\Omega f(x)\, dx = 1$.

$$\mathbb{P}(A) := \int_A f(x)\, dx \quad (A \subseteq \Omega \text{ Intervall})$$

stetiges W.maß (mit der W.dichte f)

Merkregel:
Gegeben ein Intervall $\Omega \subseteq \mathbb{R}$ *und eine Funktion* $f : \Omega \to \mathbb{R}$ *mit*

$$f(x) \geq 0 \text{ für alle } x \in \Omega \quad \text{und} \quad \int_\Omega f(x)\, dx = 1$$

(wobei die Existenz des Riemann-Integrals vorauszusetzen ist) können wir das *stetige W.maß* $\mathbb{P}$ *betrachten, das jedem Intervall* $I \subseteq \Omega$ *die W.*

$$\mathbb{P}(I) = \int_I f(x)\, dx$$

zuordnet. Die Funktion f *wird auch als* *W.dichte* *von* $\mathbb{P}$ *bezeichnet.*

Wichtige Anmerkungen:

- Das Riemann-Integral „misst" den Flächeninhalt unter dem Graphen von f.
- Die Werte $f(x)$ dürfen hier nicht als Elementar-W.'en interpretiert werden: Für jede Einpunktmenge $\{x\} \subseteq \Omega$ gilt nämlich $\mathbb{P}(\{x\}) = \int_{\{x\}} f(y)\, dy = 0$, was i. d. R. $\neq f(x)$ ist. Es kann sogar $f(x) > 1$ gelten, vgl. obiges Beispiel!
- Man muss hier „in Intervallen" statt „in Einpunktmengen" denken: Für jedes Intervall $I \subseteq \Omega$ gilt $\mathbb{P}(I) = \int_I f(y)\, dy$.

Beispiele 2.2: Stetige W.verteilungen auf $\mathbb{R}$

Dichte	**Name der W.verteilung**	**Parameter**
$f(x) = \frac{1}{b-a}\,\mathbf{1}_{\langle a,b\rangle}(x)$	*Gleichverteilung auf* $\langle a, b\rangle$ $\mathcal{U}(a, b)$	$-\infty < a < b < +\infty$
$f(x) = \lambda e^{-\lambda x}\,\mathbf{1}_{(0,\infty)}(x)$	*Exponentialverteilung* $\mathcal{E}(\lambda)$	$\lambda > 0$
$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}$	*Normalverteilung* $\mathcal{N}(\mu, \sigma^2)$	$\mu \in \mathbb{R},\ \sigma \in (0, \infty)$

Dichte	Name der W.verteilung	Parameter
$f(x) = \frac{1}{b-a} \mathbf{1}_{\langle a,b \rangle}(x)$	*Gleichverteilung auf* $\langle a, b \rangle$ $\mathcal{U}(a, b)$	$-\infty < a < b < +\infty$
$f(x) = \lambda e^{-\lambda x} \mathbf{1}_{(0,\infty)}(x)$	*Exponentialverteilung* $\mathcal{E}(\lambda)$	$\lambda > 0$
$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}$	*Normalverteilung* $\mathcal{N}(\mu, \sigma^2)$	$\mu \in \mathbb{R},\ \sigma \in (0, \infty)$

Beachte, dass wir hier als Grundmenge stets $\Omega = \mathbb{R}$ wählen und die Dichte bei Bedarf mit einer Indikatorfunktion $\mathbf{1}_I$ multiplizieren; alternativ könnten wir als Grundmenge $\Omega = I$ wählen und die Indikatorfunktion in der Dichte weglassen. Gleichverteilung und Exponentialverteilung sind uns schon in Bsp. 2.1 (a) + (b) begegnet; die Exponentialverteilung eignet sich zur Modellierung von Wartezeiten.

Bemerkung 2.3

Frage: Welchen Mengen können wir Wahrscheinlichkeiten zuordnen?

Bemerkung 2.3

Frage: Welchen Mengen können wir Wahrscheinlichkeiten zuordnen?

In Beispiel 2.1 (a) speziell für $T = 1$ wählen wir als Grundmenge $\Omega =]0, 1]$.

Bemerkung 2.3

Frage: Welchen Mengen können wir Wahrscheinlichkeiten zuordnen?

In Beispiel 2.1 (a) speziell für $T = 1$ wählen wir als Grundmenge $\Omega =]0, 1]$.
Als W.maß hätten wir gern eine Abbildung $\mathbb{P} : \mathfrak{P}(\Omega) \to \mathbb{R}$, so dass gilt:

(i) $\forall A \in \mathfrak{P}(\Omega) : \mathbb{P}(A) \geq 0$ (*Nicht-Negativität*)

(ii) $\mathbb{P}(\Omega) = 1$ (*Normiertheit*)

(iii) $\forall A_1, A_2, \ldots \in \mathfrak{P}(\Omega) : A_1, A_2, \ldots$ paarweise disjunkt
$\Rightarrow \mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$ (*σ-Additivität*)

(iv) $\forall A \in \mathfrak{P}(\Omega)\ \forall t \in \Omega : A \subseteq (0, 1) \wedge A + t \subseteq (0, 1)$
$\Rightarrow \mathbb{P}(A) = \mathbb{P}(A + t)$ (*„Symmetrie“*)

Bemerkung 2.3

Frage: Welchen Mengen können wir Wahrscheinlichkeiten zuordnen?

In Beispiel 2.1 (a) speziell für $T = 1$ wählen wir als Grundmenge $\Omega =]0, 1]$.
Als W.maß hätten wir gern eine Abbildung $\mathbb{P} : \mathfrak{P}(\Omega) \to \mathbb{R}$, so dass gilt:

(i) $\forall A \in \mathfrak{P}(\Omega) : \mathbb{P}(A) \geq 0$ (*Nicht-Negativität*)

(ii) $\mathbb{P}(\Omega) = 1$ (*Normiertheit*)

(iii) $\forall A_1, A_2, \ldots \in \mathfrak{P}(\Omega) : A_1, A_2, \ldots$ paarweise disjunkt
$\Rightarrow \mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$ (*σ-Additivität*)

(iv) $\forall A \in \mathfrak{P}(\Omega)\ \forall t \in \Omega : A \subseteq (0, 1) \wedge A + t \subseteq (0, 1)$
$\Rightarrow \mathbb{P}(A) = \mathbb{P}(A + t)$ (*„Symmetrie"*)

Problem:

Es lässt sich (mit Hilfe des Auswahlaxioms) zeigen, dass eine Abbildung $\mathbb{P} : \mathfrak{P}(\Omega) \to \mathbb{R}$ mit den Eigenschaften (i) – (iv) nicht existiert!

Bemerkung 2.3

Frage: Welchen Mengen können wir Wahrscheinlichkeiten zuordnen?

In Beispiel 2.1 (a) speziell für $T = 1$ wählen wir als Grundmenge $\Omega =]0,1]$.
Als W.maß hätten wir gern eine Abbildung $\mathbb{P} : \mathfrak{P}(\Omega) \to \mathbb{R}$, so dass gilt:

(i) $\forall A \in \mathfrak{P}(\Omega) : \mathbb{P}(A) \geq 0$ (*Nicht-Negativität*)

(ii) $\mathbb{P}(\Omega) = 1$ (*Normiertheit*)

(iii) $\forall A_1, A_2, \ldots \in \mathfrak{P}(\Omega)$: $A_1, A_2, \ldots$ paarweise disjunkt
$\Rightarrow \mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$ (*σ-Additivität*)

(iv) $\forall A \in \mathfrak{P}(\Omega)\ \forall t \in \Omega : A \subseteq (0,1) \land A + t \subseteq (0,1)$
$\Rightarrow \mathbb{P}(A) = \mathbb{P}(A + t)$ (*„Symmetrie“*)

Problem:

Es lässt sich (mit Hilfe des Auswahlaxioms) zeigen, dass eine Abbildung $\mathbb{P} : \mathfrak{P}(\Omega) \to \mathbb{R}$ mit den Eigenschaften (i) – (iv) nicht existiert!

Lösung:

Wir ersetzen den Def.bereich $\mathfrak{P}(\Omega)$ durch ein Mengensystem $\mathcal{A} \subseteq \mathfrak{P}(\Omega)$ und suchen eine Abbildung $\mathbb{P} : \mathcal{A} \to \mathbb{R}$, so dass die Eigenschaften (i) – (iv) für alle Mengen aus $\mathcal{A}$ gelten.

σ-Algebren

Definition 2.4

Sei $\Omega \neq \emptyset$. Ein Mengensystem $\mathcal{A} \subseteq \mathfrak{P}(\Omega)$ heißt *σ-Algebra* über Ω, falls gilt:

(i) $\Omega \in \mathcal{A}$

(ii) $\forall A \in \mathcal{A} : A^c \in \mathcal{A}$

(iii) $\forall A_1, A_2, \ldots \in \mathcal{A} : \bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$

σ-Algebren

Definition 2.4

Sei $\Omega \neq \emptyset$. Ein Mengensystem $\mathcal{A} \subseteq \mathfrak{P}(\Omega)$ heißt *σ-Algebra* über Ω, falls gilt:

(i) $\Omega \in \mathcal{A}$

(ii) $\forall A \in \mathcal{A} : A^c \in \mathcal{A}$

(iii) $\forall A_1, A_2, \ldots \in \mathcal{A} : \bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$

Bemerkung 2.5

Jede σ-Algebra ist abg. unter den Operationen ^{c}, $\cup$, $\cap$, $\setminus$, Δ, $\bigcup_{n=1}^{\infty}$, $\bigcap_{n=1}^{\infty}$.

$\boldsymbol{\sigma}$-Algebren

Definition 2.4

Sei $\Omega \neq \emptyset$. Ein Mengensystem $\mathcal{A} \subseteq \mathfrak{P}(\Omega)$ heißt *σ-Algebra* über Ω, falls gilt:

(i) $\Omega \in \mathcal{A}$

(ii) $\forall A \in \mathcal{A} : A^c \in \mathcal{A}$

(iii) $\forall A_1, A_2, \ldots \in \mathcal{A} : \bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$

Bemerkung 2.5

Jede σ-Algebra ist abg. unter den Operationen ^{c}, $\cup$, $\cap$, $\setminus$, Δ, $\bigcup_{n=1}^{\infty}$, $\bigcap_{n=1}^{\infty}$.

Bemerkung 2.6

σ-Algebren werden i. d. R. angegeben, indem man ein System $\mathcal{E} \subseteq \mathfrak{P}(\Omega)$ vorgibt und dann die kleinste σ-Algebra über Ω betrachtet, die $\mathcal{E}$ enthält; diese σ-Algebra wird mit $\sigma(\mathcal{E})$ bzw. als die von $\mathcal{E}$ erzeugte σ-Algebra bezeichnet.

Wahrscheinlichkeitsmaße

Definition 2.7

Sei $\Omega \neq \emptyset$ und $\mathcal{A}$ eine σ-Algebra über Ω. Eine Abbildung $\mathbb{P} : \mathcal{A} \to \mathbb{R}$ heißt *Wahrscheinlichkeitsmaß* auf $\mathcal{A}$, falls gilt:

(i) $\forall A \in \mathcal{A} : \mathbb{P}(A) \geq 0$ (*Nicht-Negativität*)

(ii) $\mathbb{P}(\Omega) = 1$ (*Normiertheit*)

(iii) $\forall A_1, A_2, \ldots \in \mathcal{A} : A_1, A_2, \ldots$ paarweise disjunkt $\Rightarrow \mathbb{P}(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mathbb{P}(A_k)$ (σ-*Additivität*)

Wahrscheinlichkeitsmaße

Definition 2.7

Sei $\Omega \neq \emptyset$ und $\mathcal{A}$ eine σ-Algebra über Ω. Eine Abbildung $\mathbb{P} : \mathcal{A} \to \mathbb{R}$ heißt *Wahrscheinlichkeitsmaß* auf $\mathcal{A}$, falls gilt:

(i) $\forall A \in \mathcal{A} : \mathbb{P}(A) \geq 0$ (*Nicht-Negativität*)

(ii) $\mathbb{P}(\Omega) = 1$ (*Normiertheit*)

(iii) $\forall A_1, A_2, \ldots \in \mathcal{A} : A_1, A_2, \ldots$ paarweise disjunkt
$\Rightarrow \mathbb{P}(\bigcup_{k=1}^\infty A_k) = \sum_{k=1}^\infty \mathbb{P}(A_k)$ (*σ-Additivität*)

Bemerkung 2.8

Sei $\Omega \neq \emptyset$, $\mathcal{A}$ eine σ-Algebra über Ω und $\mathbb{P}$ ein W.maß auf $\mathcal{A}$. Satz 1.6 (Rechenregeln für W.maße) und Lemma 1.8 (Charakterisierung der σ-Additivität) gelten entsprechend, sofern man überall nur Mengen aus $\mathcal{A}$ zulässt.

Bemerkung 2.9

Wahrscheinlichkeitsmaße

Definition 2.7

Sei $\Omega \neq \emptyset$ und $\mathcal{A}$ eine σ-Algebra über Ω. Eine Abbildung $\mathbb{P} : \mathcal{A} \to \mathbb{R}$ heißt *Wahrscheinlichkeitsmaß* auf $\mathcal{A}$, falls gilt:

(i) $\forall A \in \mathcal{A} : \mathbb{P}(A) \geq 0$ (*Nicht-Negativität*)

(ii) $\mathbb{P}(\Omega) = 1$ (*Normiertheit*)

(iii) $\forall A_1, A_2, \ldots \in \mathcal{A} : A_1, A_2, \ldots$ paarweise disjunkt
$\Rightarrow \mathbb{P}(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mathbb{P}(A_k)$ (σ-*Additivität*)

Bemerkung 2.8

Bemerkung 2.9

W.maße werden i. d. R. dadurch konstruiert, dass man ihre Werte auf einem (geeigneten) Erzeuger $\mathcal{E}$ von $\mathcal{A}$ vorschreibt und sich überlegt, dass genau eine Fortsetzung zu einem W.maß auf $\mathcal{A}$ existiert.

Definition 2.10 (Kolmogorov 1933)

Ein *W.raum* ist ein Tripel $(\Omega, \mathcal{A}, \mathbb{P})$, wobei $\Omega \neq \emptyset$, $\mathcal{A}$ σ-Algebra über Ω und $\mathbb{P}$ W.maß auf $\mathcal{A}$.

Definition 2.10 (Kolmogorov 1933)

Ein *W.raum* ist ein Tripel $(\Omega, \mathcal{A}, \mathbb{P})$, wobei $\Omega \neq \emptyset$, $\mathcal{A}$ σ-Algebra über Ω und $\mathbb{P}$ W.maß auf $\mathcal{A}$.

Wir bezeichnen die Elemente von Ω als *Ergebnisse*, die Elemente von $\mathcal{A}$ (und nur diese!) als *Ereignisse* (oder *messbare Mengen*) und die Werte von $\mathbb{P}$ als *Wahrscheinlichkeiten*. Ferner bezeichnen wir $\mathbb{P}$ auch als W.maß auf dem *messbaren Raum* $(\Omega, \mathcal{A})$ oder als W.maß auf Ω.

Definition 2.10 (Kolmogorov 1933)

Ein *W.raum* ist ein Tripel $(\Omega, \mathcal{A}, \mathbb{P})$, wobei $\Omega \neq \emptyset$, $\mathcal{A}$ σ-Algebra über Ω und $\mathbb{P}$ W.maß auf $\mathcal{A}$.

Wir bezeichnen die Elemente von Ω als *Ergebnisse*, die Elemente von $\mathcal{A}$ (und nur diese!) als *Ereignisse* (oder *messbare Mengen*) und die Werte von $\mathbb{P}$ als *Wahrscheinlichkeiten*. Ferner bezeichnen wir $\mathbb{P}$ auch als W.maß auf dem *messbaren Raum* $(\Omega, \mathcal{A})$ oder als W.maß auf Ω.

Beispiele für Wahrscheinlichkeitsräume:

- **diskrete W.räume:** Ω abzählbar
  $\mathcal{A} = \sigma(\{\text{ Einpunktmengen }\}) = \mathfrak{P}(\Omega)$
  $\mathbb{P} =$ diskretes W.maß über Ω
- **stetige W.räume:** $\Omega \subseteq \mathbb{R}^n$
  $\mathcal{A} = \sigma(\{\text{ Intervalle}|_\Omega \}) =: \mathbb{B}^n|_\Omega$
  $\mathbb{P} =$ stetiges W.maß über Ω

Beispiele für W.räume I

(Diskrete W.räume)

$\Omega \neq \emptyset$ abzählbar

$\mathcal{A}$ = kleinste σ-Algebra über Ω, die alle Einpunktmengen enthält $= \mathfrak{P}(\Omega)$ (Potenzmenge über Ω)

$f : \Omega \to \mathbb{R}$ mit $f(\omega) \geq 0$ für alle $\omega \in \Omega$ und $\sum_{\omega\in\Omega} f(\omega) = 1$
$\rightsquigarrow$ Es exist. genau ein W.maß $\mathbb{P}$ auf $\mathcal{A}$ mit $\mathbb{P}(\{x\}) = f(x)$ für alle $x \in \Omega$, nämlich das W.maß aus Definition 1.1.

$\mathbb{P}$ heißt <u>diskretes W.maß mit der (W.-)Dichte f</u>.

Genauer ist jedes W.maß auf $\mathcal{A}$ von dieser Gestalt. (*Übung!*)

Beispiele für W.räume II

(Stetige W.räume über $\mathbb{R}$)

$\Omega = \mathbb{R}$

$\mathcal{A}$ = *kleinste* σ*-Algebra über* $\mathbb{R}$, *die alle Intervalle enthält* $=: \mathbb{B}\ [\neq \mathfrak{P}(\Omega)]$
(*Borel-*σ*-algebra über* $\mathbb{R}$)
Es ist nicht möglich (aber auch nicht nötig), $\mathbb{B}$ *explizit anzugeben.*

$f : \mathbb{R} \to \mathbb{R}$ *mit* $f(x) \geq 0$ *für alle* $x \in \mathbb{R}$ *und* $\int_{\mathbb{R}} f(x)\, dx = 1$
$\rightsquigarrow$ *Es ex. genau ein W.maß* $\mathbb{P}$ *auf* $\mathcal{A}$ *mit* $\mathbb{P}(]a,b]) = \int_{]a,b]} f(x)\, dx \ \forall\,]a,b]$.

$\mathbb{P}$ *heißt* *stetiges W.maß mit der (W.-)Dichte* f.

Allerdings ist nicht jedes W.maß auf $\mathcal{A}$ von dieser Gestalt. (*Übung!*)

Beispiele für W.räume II

Bemerkung 2.11

Sei $\mathbb{P}$ *ein stetiges W.maß mit der Dichte* f.

Beispiele für W.räume II

Bemerkung 2.11

Sei $\mathbb{P}$ *ein stetiges W.maß mit der Dichte* f.

(i) *Es gilt* $\mathbb{P}(\{x\}) = 0$ *für alle* $x \in \mathbb{R}$.

Beispiele für W.räume II

Bemerkung 2.11

Sei $\mathbb{P}$ *ein stetiges W.maß mit der Dichte* f.

(i) *Es gilt* $\mathbb{P}(\{x\}) = 0$ *für alle* $x \in \mathbb{R}$.

(ii) *Weiter gilt damit wegen der Additivität / σ-Additivität von* $\mathbb{P}$ *auch* $\mathbb{P}(A) = 0$ *für jede endliche / abzählbare Teilmenge* A *von* $\mathbb{R}$*; insbesondere kann* $\mathbb{P}$ *also kein diskretes W.maß sein!*

Beispiele für W.räume II

Bemerkung 2.11

Sei $\mathbb{P}$ ein stetiges W.maß mit der Dichte f.

(i) *Es gilt $\mathbb{P}(\{x\}) = 0$ für alle $x \in \mathbb{R}$.*

(ii) *Weiter gilt damit wegen der Additivität / σ-Additivität von $\mathbb{P}$ auch $\mathbb{P}(A) = 0$ für jede endliche / abzählbare Teilmenge A von $\mathbb{R}$; insbesondere kann $\mathbb{P}$ also kein diskretes W.maß sein!*

(iii) *$f(x)$ lässt sich hier nicht als (Elementar-)Wahrscheinlichkeit interpretieren; es kann sogar $f(x) > 1$ gelten!*

Bemerkung 2.11

Sei $\mathbb{P}$ ein stetiges W.maß mit der Dichte f.

(i) *Es gilt $\mathbb{P}(\{x\}) = 0$ für alle $x \in \mathbb{R}$.*

(ii) *Weiter gilt damit wegen der Additivität / σ-Additivität von $\mathbb{P}$ auch $\mathbb{P}(A) = 0$ für jede endliche / abzählbare Teilmenge A von $\mathbb{R}$; insbesondere kann $\mathbb{P}$ also kein diskretes W.maß sein!*

(iii) *$f(x)$ lässt sich hier nicht als (Elementar-)Wahrscheinlichkeit interpretieren; es kann sogar $f(x) > 1$ gelten!*

(iv) *Man kann f (z. B. an endlich vielen Stellen) abändern, ohne $\mathbb{P}$ zu verändern. Die Dichte f ist also durch das W.maß $\mathbb{P}$ nicht (ganz) eindeutig bestimmt.*

Beispiele für W.räume II

Bemerkung 2.11

Sei $\mathbb{P}$ ein stetiges W.maß mit der Dichte f.

(i) *Es gilt $\mathbb{P}(\{x\}) = 0$ für alle $x \in \mathbb{R}$.*

(ii) *Weiter gilt damit wegen der Additivität / σ-Additivität von $\mathbb{P}$ auch $\mathbb{P}(A) = 0$ für jede endliche / abzählbare Teilmenge A von $\mathbb{R}$; insbesondere kann $\mathbb{P}$ also kein diskretes W.maß sein!*

(iii) *$f(x)$ lässt sich hier nicht als (Elementar-)Wahrscheinlichkeit interpretieren; es kann sogar $f(x) > 1$ gelten!*

(iv) *Man kann f (z. B. an endlich vielen Stellen) abändern, ohne $\mathbb{P}$ zu verändern. Die Dichte f ist also durch das W.maß $\mathbb{P}$ nicht (ganz) eindeutig bestimmt.*

(v) *Es ist im Allgemeinen schwierig zu entscheiden, ob eine Menge $A \subseteq \mathbb{R}$ in $\mathbb{B}$ liegt oder ob für eine Menge $A \in \mathbb{B}$ die W. $\mathbb{P}(A)$ durch Integration bestimmt werden kann. Da dies bei den in Anwendungen auftretenden Mengen i. d. R. der Fall ist, wollen wir dies immer stillschweigend voraussetzen:*

$$\boxed{\mathbb{P}(A) = \int_A f(x)\,dx = \int_{\mathbb{R}} \mathbf{1}_A(x)\,f(x)\,dx}$$

Beispiele für W.räume II

Bemerkung 2.11

Sei $\mathbb{P}$ ein stetiges W.maß mit der Dichte f.

(i) *Es gilt $\mathbb{P}(\{x\}) = 0$ für alle $x \in \mathbb{R}$.*

(ii) *Weiter gilt damit wegen der Additivität / σ-Additivität von $\mathbb{P}$ auch $\mathbb{P}(A) = 0$ für jede endliche / abzählbare Teilmenge A von $\mathbb{R}$; insbesondere kann $\mathbb{P}$ also kein diskretes W.maß sein!*

(iii) *$f(x)$ lässt sich hier nicht als (Elementar-)Wahrscheinlichkeit interpretieren; es kann sogar $f(x) > 1$ gelten!*

(iv) *Man kann f (z. B. an endlich vielen Stellen) abändern, ohne $\mathbb{P}$ zu verändern. Die Dichte f ist also durch das W.maß $\mathbb{P}$ nicht (ganz) eindeutig bestimmt.*

(v) *Es ist im Allgemeinen schwierig zu entscheiden, ob eine Menge $A \subseteq \mathbb{R}$ in $\mathbb{B}$ liegt oder ob für eine Menge $A \in \mathbb{B}$ die W. $\mathbb{P}(A)$ durch Integration bestimmt werden kann. Da dies bei den in Anwendungen auftretenden Mengen i. d. R. der Fall ist, wollen wir dies immer stillschweigend voraussetzen:*

$$\boxed{\mathbb{P}(A) = \int_A f(x)\,dx = \int_{\mathbb{R}} \mathbf{1}_A(x)\,f(x)\,dx}$$

(vi) *Man kann auch stetige W.maße über Teilmengen $\Omega \subseteq \mathbb{R}$ einführen.*

Beispiele für W.räume II (Beispiel 2.12)

Beispiele für W.räume II (Beispiel 2.12)

Bei der Berechnung von Wahrscheinlichkeiten ist es gelegentlich nützlich, zuerst die Rechenregeln für W.maße anzuwenden und anschließend die Wahrscheinlichkeiten der verbleibenden (einfacheren) Mengen durch Integration zu bestimmen:

Wir betrachten Beispiel 2.1 (b). Mit welcher W. fährt das erste Auto innerhalb der ersten Stunde oder nach der vierten Stunde vorbei?

Beispiele für W.räume II (Beispiel 2.12)

Bei der Berechnung von Wahrscheinlichkeiten ist es gelegentlich nützlich, zuerst die Rechenregeln für W.maße anzuwenden und anschließend die Wahrscheinlichkeiten der verbleibenden (einfacheren) Mengen durch Integration zu bestimmen:

Wir betrachten Beispiel 2.1 (b). Mit welcher W. fährt das erste Auto innerhalb der ersten Stunde oder nach der vierten Stunde vorbei?

Lösung:

Wir wählen $\Omega := \mathbb{R}$, $\mathcal{A} := \mathbb{B}$, $\mathbb{P} :=$ stetiges W.maß über $\mathbb{R}$ mit der Dichte $f(x) = \lambda e^{-\lambda x}\,\mathbf{1}_{(0,\infty)}(x)$ und $A := (0,1) \cup (4,\infty)$. Damit ergibt sich wegen $A = (0,\infty) \setminus [1,4]$

$$\mathbb{P}(A) \underset{\text{Subtraktivität}}{=} \mathbb{P}((0,\infty)) - \mathbb{P}([1,4]) = 1 - \int_1^4 \lambda e^{-\lambda x}\,dx = 1 - e^{-\lambda} + e^{-4\lambda}\,.$$

(Alternative: $\Omega := (0,\infty)$, $\mathcal{A} := \mathbb{B}|_\Omega$, $\mathbb{P} :=$ stetiges W.maß über $(0,\infty)$ mit der Dichte $f(x) = \lambda e^{-\lambda x}$.)

Beispiele für W.räume III

(Stetige W.räume über $\mathbb{R}^n$)

$\Omega = \mathbb{R}^n$

$\mathcal{A} =$ *kleinste* σ*-Algebra über* $\mathbb{R}^n$*, die alle n-dimensionalen Intervalle* $]a, b] =]a_1, b_1] \times \cdots \times]a_n, b_n]$ *enthält* $=: \mathbb{B}^n$ *(Borel-*σ*-algebra über* $\mathbb{R}^n$*)*

$f : \mathbb{R}^n \to \mathbb{R}$ *mit* $f(x) \geq 0$ *für alle* $x \in \mathbb{R}^n$ *und* $\int_{\mathbb{R}^n} f(x)\,dx = 1$
$\rightsquigarrow$ *Es ex. genau ein W.maß* $\mathbb{P}$ *auf* $\mathcal{A}$ *mit* $\mathbb{P}(]a, b]) = \int_{]a,b]} f(x)\,dx \;\forall\,]a, b]$.

$\mathbb{P}$ *heißt stetiges W.maß mit der (W.-)Dichte f.*

Erläuterung:
Dabei wollen wir eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ als „integrierbar" bezeichnen, wenn für jedes n-dim. Intervall $]a, b]$ das iterierte Riemann-Integral exist. und jede Integrationsreihenfolge dasselbe Ergebnis liefert, z. B. für $n = 2$:

$$\int_{]a,b]} f(x)\,dx = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x_1, x_2)\,dx_2\,dx_1 = \int_{a_2}^{b_2} \int_{a_1}^{b_1} f(x_1, x_2)\,dx_1\,dx_2\,.$$

Beispiele für W.räume III

Beispiel 2.13 (Stetige Gleichverteilung)

Es sei $B \subseteq \mathbb{R}^n$ *eine Menge mit* $\operatorname{vol}(B) := \int_{\mathbb{R}^n} \mathbf{1}_B(x)\,dx \in (0, \infty)$

(d. h. insbesondere muss $\mathbf{1}_B$ *integrierbar sein).*

Beispiele für W.räume III

Beispiel 2.13 (Stetige Gleichverteilung)

Es sei $B \subseteq \mathbb{R}^n$ *eine Menge mit* $\operatorname{vol}(B) := \int_{\mathbb{R}^n} \mathbf{1}_B(x)\,dx \in (0,\infty)$

(d. h. insbesondere muss $\mathbf{1}_B$ *integrierbar sein).*

Dann heißt das stetige W.maß $\mathbb{P}$ *auf* $\mathbb{B}^n$ *mit der Dichte* $f(x) := \frac{1}{\operatorname{vol}(B)}\,\mathbf{1}_B(x)$

(stetige) Gleichverteilung auf B*, kurz* $\mathcal{U}_B$*,*

Beispiele für W.räume III

Beispiel 2.13 (Stetige Gleichverteilung)

Es sei $B \subseteq \mathbb{R}^n$ *eine Menge mit* $\mathrm{vol}(B) := \int_{\mathbb{R}^n} \mathbf{1}_B(x)\, dx \in (0, \infty)$

(d. h. insbesondere muss $\mathbf{1}_B$ *integrierbar sein).*

Dann heißt das stetige W.maß $\mathbb{P}$ *auf* $\mathbb{B}^n$ *mit der Dichte* $f(x) := \frac{1}{\mathrm{vol}(B)} \mathbf{1}_B(x)$

(stetige) Gleichverteilung auf B, *kurz* $\mathcal{U}_B$,

und es gilt

$$\mathbb{P}(A) = \frac{\mathrm{vol}(A \cap B)}{\mathrm{vol}(B)}$$

für jedes n-dimensionale Intervall $A = \,]a, b]$

(bzw. allgemeiner sogar für jede Menge $A \in \mathbb{B}^n$, *so dass* $\mathrm{vol}(A \cap B)$ *definiert ist).*

Beispiele für W.räume III

Beispiel 2.13 (Stetige Gleichverteilung)

Es sei $B \subseteq \mathbb{R}^n$ *eine Menge mit* $\mathrm{vol}(B) := \int_{\mathbb{R}^n} \mathbf{1}_B(x)\,dx \in (0,\infty)$

(d. h. insbesondere muss $\mathbf{1}_B$ *integrierbar sein).*

Dann heißt das stetige W.maß $\mathbb{P}$ *auf* $\mathbb{B}^n$ *mit der Dichte* $f(x) := \frac{1}{\mathrm{vol}(B)}\,\mathbf{1}_B(x)$

(stetige) Gleichverteilung auf B*, kurz* $\mathcal{U}_B$*,*

und es gilt

$$\mathbb{P}(A) = \frac{\mathrm{vol}(A \cap B)}{\mathrm{vol}(B)}$$

für jedes n-dimensionale Intervall $A = \,]a,b]$

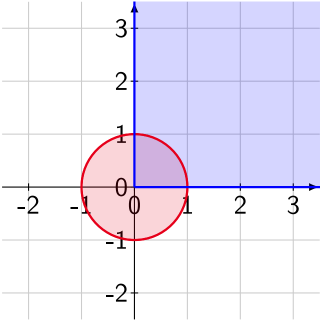
(bzw. allgemeiner sogar für jede Menge $A \in \mathbb{B}^n$*, so dass* $\mathrm{vol}(A \cap B)$ *definiert ist).*

Damit lässt sich die Bestimmung von W.en auf die Bestimmung von Volumina zurückführen.

Für $n = 1$, $B = (a,b)$ erhalten wir gerade die Gleichverteilung auf (a,b) zurück.
Für einfache geometrische Figuren ist es i. d. R. einfacher, die Volumina mit den Formeln aus der Schule als mit dem Integral zu bestimmen.

Beispiele für W.räume III (Beispiel 2.14)

Mit welcher W. liegt ein „rein zufällig“ gewählter Punkt im Einheitskreis $K := \{x \in \mathbb{R}^2 : \|x\|_2^2 \leq 1\}$ *im 1. Quadranten* (*d.h. in* $A := (0,\infty)^2$) *?*

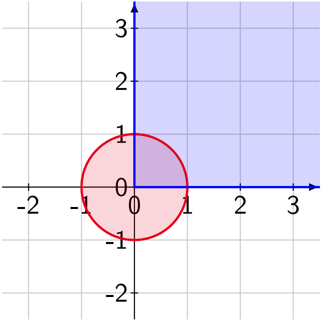


Skizze

Beispiele für W.räume III (Beispiel 2.14)

Mit welcher W. liegt ein „rein zufällig“ gewählter Punkt im Einheitskreis $K := \{x \in \mathbb{R}^2 : \|x\|_2^2 \leq 1\}$ im 1. Quadranten (d.h. in $A := (0,\infty)^2$) ?

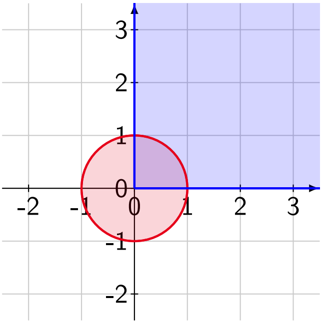
$(\Omega, \mathcal{A}, \mathbb{P}) := (\mathbb{R}^2, \mathbb{B}^2, \mathcal{U}_K)$ (da „rein zufällig“)



Skizze

Mit welcher W. liegt ein „rein zufällig" gewählter Punkt im Einheitskreis $K := \{x \in \mathbb{R}^2 : \|x\|_2^2 \leq 1\}$ im 1. Quadranten (d.h. in $A := (0,\infty)^2$) ?

$$(\Omega, \mathcal{A}, \mathbb{P}) := (\mathbb{R}^2, \mathbb{B}^2, \mathcal{U}_K) \quad \text{(da „rein zufällig")}$$



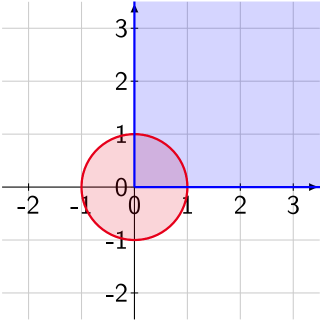
Skizze

Lösung Nr. 1:

$$\mathbb{P}(A) \underset{\substack{\uparrow \\ \text{Gleichvtlg.}}}{=} \frac{\mathrm{vol}(A \cap K)}{\mathrm{vol}(K)} = \frac{\frac{1}{4}\,\mathrm{vol}(K)}{\mathrm{vol}(K)} = \frac{1}{4}\,.$$

Beispiele für W.räume III (Beispiel 2.14)

Mit welcher W. liegt ein „rein zufällig“ gewählter Punkt im Einheitskreis $K := \{x \in \mathbb{R}^2 : \|x\|_2^2 \leq 1\}$ im 1. Quadranten (d.h. in $A := (0,\infty)^2$) ?



Skizze

$$(\Omega, \mathcal{A}, \mathbb{P}) := (\mathbb{R}^2, \mathbb{B}^2, \mathcal{U}_K) \quad \text{(da „rein zufällig“)}$$

Lösung Nr. 1:

$$\mathbb{P}(A) \underset{\substack{\uparrow \\ \text{Gleichvtlg.}}}{=} \frac{\mathrm{vol}(A \cap K)}{\mathrm{vol}(K)} = \frac{\frac{1}{4}\,\mathrm{vol}(K)}{\mathrm{vol}(K)} = \frac{1}{4}\,.$$

Lösung Nr. 2:

$$\mathbb{P}(A) \underset{\substack{\uparrow \\ \text{Dichte}}}{=} \int_A \frac{1}{\pi}\,\mathbf{1}_K(x)\,dx$$

$$\underset{\substack{\uparrow \\ \text{iteriertes} \\ \text{Integral}}}{=} \int_0^1 \int_0^{\sqrt{1-x_1^2}} \frac{1}{\pi}\,dx_2\,dx_1 = \cdots = \frac{1}{4}\,.$$

Zusammenfassung:
Gegeben eine Teilmenge $\Omega \subseteq \mathbb{R}^n$ *und eine Funktion* $f : \Omega \to \mathbb{R}$ *mit*

$$f(x) \geq 0 \text{ für alle } x \in \Omega \quad \text{und} \quad \int_\Omega f(x)\,dx = 1$$

(wobei die Existenz des Riemann-Integrals vorauszusetzen ist)
können wir das <u>stetige W.maß</u> $\mathbb{P}$ *auf* $\mathbb{B}^n|_\Omega$ *mit der W.dichte* f *betrachten.*
Für die in Anwendungen auftretenden Mengen $A \subseteq \Omega$ *können wir dann die W.* $\mathbb{P}(A)$ *bestimmen, indem wir die Dichte* f *über die Menge* A *integrieren:*

$$\mathbb{P}(A) = \int_A f(x)\,dx$$

Stetige W.verteilungen

Zusammenfassung:
Gegeben eine Teilmenge $\Omega \subseteq \mathbb{R}^n$ *und eine Funktion* $f : \Omega \to \mathbb{R}$ *mit*

$$f(x) \geq 0 \text{ für alle } x \in \Omega \quad \text{und} \quad \int_\Omega f(x)\,dx = 1$$

(wobei die Existenz des Riemann-Integrals vorauszusetzen ist)
können wir das stetige W.maß $\mathbb{P}$ *auf* $\mathbb{B}^n|_\Omega$ *mit der W.dichte* f *betrachten.*
Für die in Anwendungen auftretenden Mengen $A \subseteq \Omega$ *können wir dann die W.* $\mathbb{P}(A)$ *bestimmen, indem wir die Dichte* f *über die Menge* A *integrieren:*

$$\mathbb{P}(A) = \int_A f(x)\,dx$$

Warnungen:

- Wir verwenden die Bezeichnung „Dichte“ sowohl bei diskreten als auch bei stetigen W.maßen.
- Der Begriff „Gleichverteilung“ kann entweder ein diskretes oder ein stetiges W.maß bezeichnen.

Kapitel 1.3
Allgemeine W.räume

Beschreibung aller W.maße auf $(\mathbb{R}, \mathbb{B})$

Beschreibung aller W.maße auf $(\mathbb{R}, \mathbb{B})$

Satz 2.15

Ist $\mathbb{P}$ ein W.maß auf $(\mathbb{R}, \mathbb{B})$, so ist $F : \mathbb{R} \to \mathbb{R}$, $F(x) := \mathbb{P}(]-\infty, x])$ eine Abbildung mit den folgenden Eigenschaften:

(i) *F ist monoton wachsend.*
(ii) *F ist rechtsseitig stetig.*
(iii) $\lim_{x \to -\infty} F(x) = 0$.
(iv) $\lim_{x \to +\infty} F(x) = 1$.

Umgekehrt gibt es zu jeder solchen Abbildung F (genau) ein W.maß $\mathbb{P}$ auf $(\mathbb{R}, \mathbb{B})$, so dass $\mathbb{P}(]-\infty, x]) = F(x)$ für alle $x \in \mathbb{R}$ gilt.

Definition 2.16 (Verteilungsfunktion)

Eine Funktion $F : \mathbb{R} \to \mathbb{R}$ mit den Eigenschaften (i) – (iv) in Satz 2.15 heißt *Verteilungsfunktion* auf $\mathbb{R}$.

Verteilungsfunktionen

Satz 2.15

Ist $\mathbb{P}$ ein W.maß auf $(\mathbb{R}, \mathbb{B})$, so ist $F : \mathbb{R} \to \mathbb{R}$ mit $F(x) := \mathbb{P}(]-\infty, x])$ eine Abbildung mit den folgenden Eigenschaften:

(i) monoton wachsend (ii) rechtsseitig stetig (iii) $\lim_{x \to -\infty} F(x) = 0$ *(iv)* $\lim_{x \to +\infty} F(x) = 1$

Umgekehrt gibt es zu jeder solchen Abbildung F (genau) ein W.maß $\mathbb{P}$ auf $(\mathbb{R}, \mathbb{B})$, so dass $\mathbb{P}(]-\infty, x]) = F(x)$ für alle $x \in \mathbb{R}$ gilt.

Definition 2.16 (Verteilungsfunktion)

Eine Funktion $F : \mathbb{R} \to \mathbb{R}$ mit den Eigenschaften (i) – (iv) in Satz 2.15 heißt *Verteilungsfunktion* auf $\mathbb{R}$.

Genauer heißt bei gegebenem W.maß $\mathbb{P}$ auf $(\mathbb{R}, \mathbb{B})$ die Funktion $F : \mathbb{R} \to \mathbb{R}$ mit $F(x) = \mathbb{P}(]-\infty, x])$ für alle $x \in \mathbb{R}$ die Verteilungsfunktion zum W.maß $\mathbb{P}$ (oder Verteilungsfunktion von $\mathbb{P}$) und umgekehrt bei gegebener Verteilungsfunktion F auf $\mathbb{R}$ das W.maß $\mathbb{P}$ auf $(\mathbb{R}, \mathbb{B})$ mit $\mathbb{P}(]a, b]) = F(b) - F(a)$ für alle $a, b \in \mathbb{R}$ mit $a < b$ das W.maß zur Verteilungsfunktion F.

Beschreibung aller W.maße auf $(\mathbb{R}, \mathbb{B})$

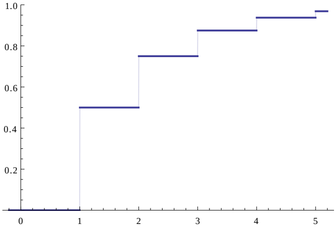
Bemerkungen 2.17

(i) *Nach Satz 2.15 besteht eine 1:1-Beziehung zwischen W.maßen auf* $(\mathbb{R}, \mathbb{B})$ *und Verteilungsfunktionen auf* $\mathbb{R}$.

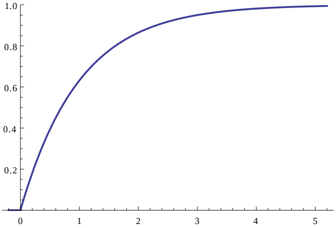
(ii) *Die diskreten W.maße entsprechen den „Sprung-Verteilungsfunktionen".*
(Bemerkung: Die Menge der Sprungstellen ist abzählbar, kann aber dicht in $\mathbb{R}$ *liegen.)*

(iii) *Die stetigen W.maße mit (stückweise) stetigen Dichten entsprechen den stetigen und (stückweise) stetig differenzierbaren Verteilungsfunktionen.*

(iv) *Es gibt weitere Typen, z. B. „Mischtypen" oder die Cantor-Verteilung.*

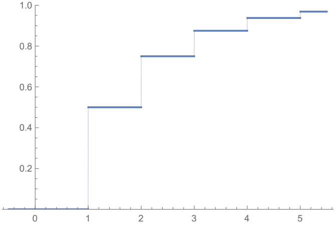


geometrische Verteilung ($p = 0.5$)

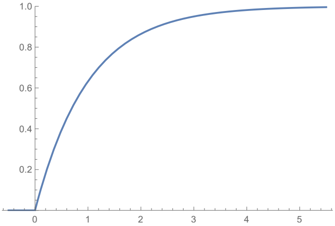


Exponentialverteilung ($\lambda = 1$)

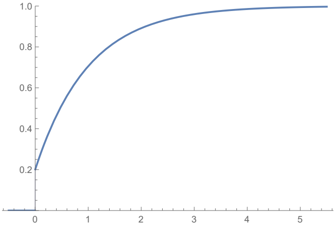
Beschreibung aller W.maße auf $(\mathbb{R}, \mathbb{B})$



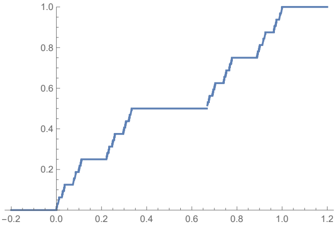
Verteilungsfunktion der $\mathcal{G}(\frac{1}{2})$-Vtlg.



Verteilungsfunktion der $\mathcal{E}(1)$-Vtlg.

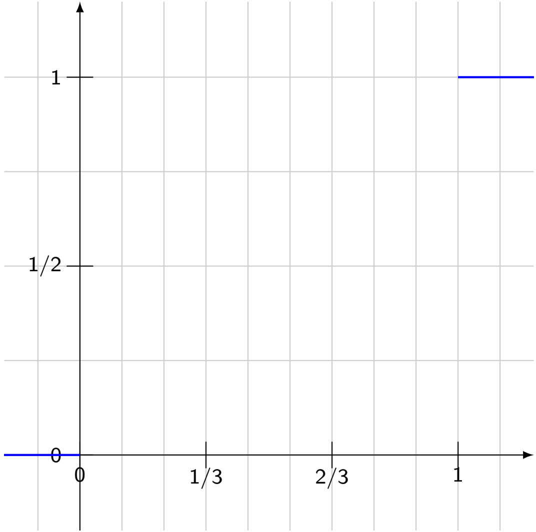


Verteilungsfunktion einer „Mischung“

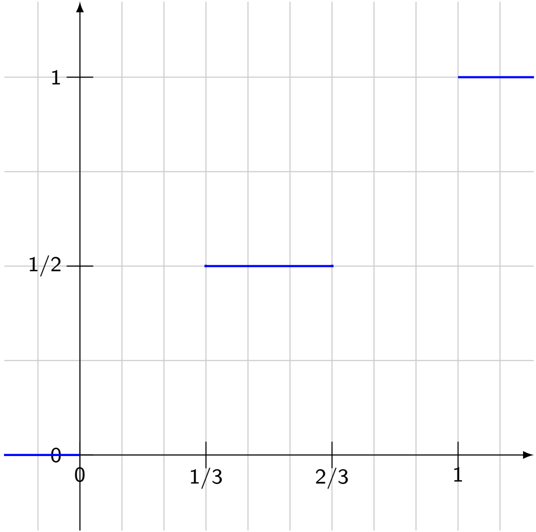


Verteilungsfunktion der Cantor-Vtlg.

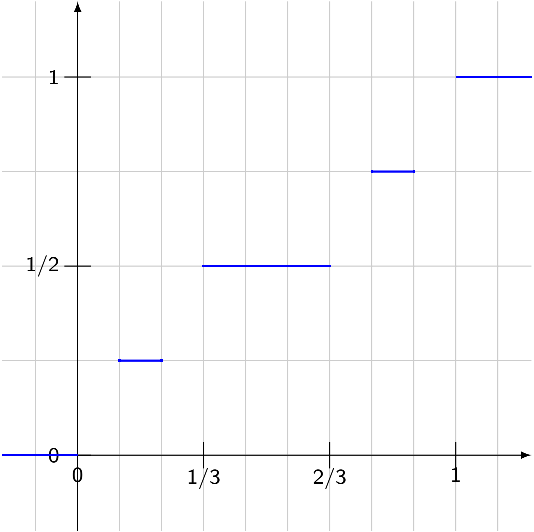
Konstruktion der Cantor-Funktion



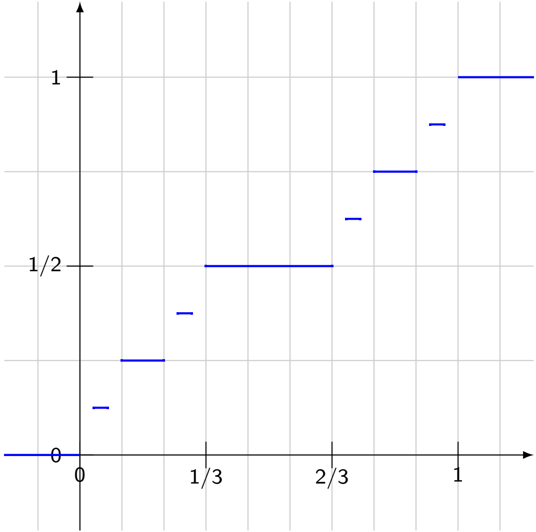
Konstruktion der Cantor-Funktion



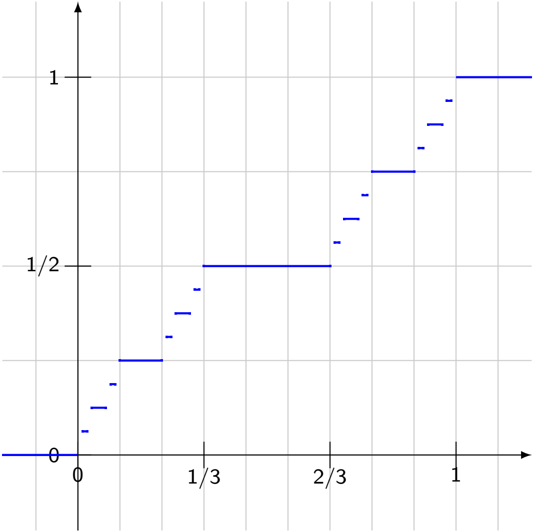
Konstruktion der Cantor-Funktion



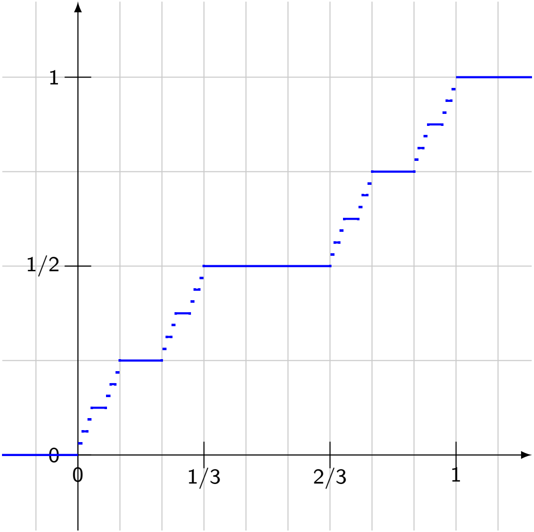
Konstruktion der Cantor-Funktion



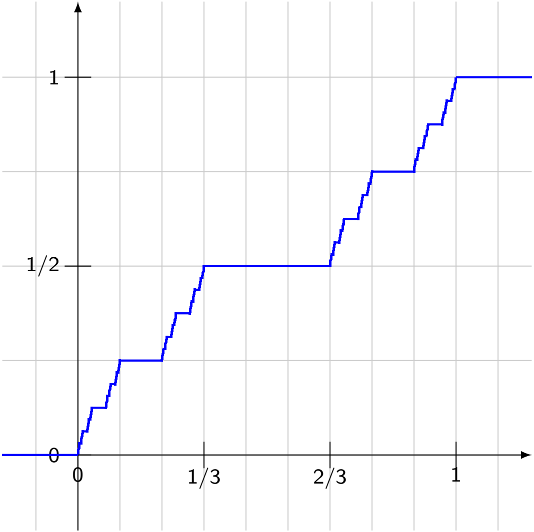
Konstruktion der Cantor-Funktion



Konstruktion der Cantor-Funktion



Konstruktion der Cantor-Funktion



Arbeiten mit Verteilungsfunktionen

Bemerkung (Berechnung von Wahrscheinlichkeiten)

Ist ein W.maß $\mathbb{P}$ durch seine Verteilungsfunktion F gegeben, so kann man die W. eines Intervalls $]a, b]$ mit der Formel

$$\mathbb{P}(]a, b]) = F(b) - F(a)$$

bestimmen und die W. eines allgemeinen Ereignisses $B \in \mathbb{B}$ dadurch bestimmen, dass man dieses auf Intervalle zurückführt.

Ist $\mathbb{P}$ ein diskretes oder stetiges W.maß, so ist es auch möglich, zunächst die Zähldichte bzw. die Riemann-Dichte zu bestimmen und anschließend eine Summe bzw. ein Integral über diese Dichte zu berechnen. Dieser Zugang ist (im Prinzip) auch anwendbar, wenn $\mathbb{P}$ ein „Mischtyp" einer diskreten und einer stetigen W.verteilung ist.

Kapitel 2.2
Zufallsgrößen auf allgemeinen W.räumen

Zufallsgrößen auf allgemeinen W.räumen I

Da auf allgemeinen W.räumen nur noch Mengen in $\mathcal{F}$ eine W. zugeordnet wird, müssen wir sicherstellen, dass nur solche Mengen als Urbildmengen auftreten.

Definition 2.18 (Zufallsgröße)

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein allgemeiner W.raum, $(\mathcal{X}, \mathcal{B})$ ein messbarer Raum und $X : \Omega \to \mathcal{X}$ eine $\mathcal{A}$-$\mathcal{B}$-messbare Abbildung (kurz: $X : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{B})$), d. h. es gelte $X^{-1}(B) \in \mathcal{A}$ für alle $B \in \mathcal{B}$.
Dann heißt X *Zufallsgröße* oder *Zufallsvariable* (mit Werten in $(\mathcal{X}, \mathcal{B})$), und die Abbildung $\mathbb{P}_X : \mathcal{B} \to \mathbb{R}$ mit $\mathbb{P}_X(B) := \mathbb{P}(X^{-1}(B))$ für alle $B \in \mathcal{B}$ heißt *induzierte Verteilung* (oder auch nur *Verteilung*) von X unter $\mathbb{P}$.

Zufallsgrößen auf allgemeinen W.räumen I

Da auf allgemeinen W.räumen nur noch Mengen in $\mathcal{F}$ eine W. zugeordnet wird, müssen wir sicherstellen, dass nur solche Mengen als Urbildmengen auftreten.

Definition 2.18 (Zufallsgröße)

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein allgemeiner W.raum, $(\mathcal{X}, \mathcal{B})$ ein messbarer Raum und $X : \Omega \to \mathcal{X}$ eine $\mathcal{A}$-$\mathcal{B}$-messbare Abbildung (kurz: $X : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{B})$), d. h. es gelte $X^{-1}(B) \in \mathcal{A}$ für alle $B \in \mathcal{B}$.
Dann heißt X *Zufallsgröße* oder *Zufallsvariable* (mit Werten in $(\mathcal{X}, \mathcal{B})$), und die Abbildung $\mathbb{P}_X : \mathcal{B} \to \mathbb{R}$ mit $\mathbb{P}_X(B) := \mathbb{P}(X^{-1}(B))$ für alle $B \in \mathcal{B}$ heißt *induzierte Verteilung* (oder auch nur *Verteilung*) von X unter $\mathbb{P}$.

Lemma 2.19

In der Situation von Definition 2.18 ist $\mathbb{P}_X$ ein W.maß auf $(\mathcal{X}, \mathcal{B})$.

Zufallsgrößen auf allgemeinen W.räumen I

Da auf allgemeinen W.räumen nur noch Mengen in $\mathcal{F}$ eine W. zugeordnet wird, müssen wir sicherstellen, dass nur solche Mengen als Urbildmengen auftreten.

Definition 2.18 (Zufallsgröße)

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein allgemeiner W.raum, $(\mathcal{X}, \mathcal{B})$ ein messbarer Raum und $X : \Omega \to \mathcal{X}$ eine $\mathcal{A}$-$\mathcal{B}$-messbare Abbildung (kurz: $X : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{B})$), d. h. es gelte $X^{-1}(B) \in \mathcal{A}$ für alle $B \in \mathcal{B}$.
Dann heißt X *Zufallsgröße* oder *Zufallsvariable* (mit Werten in $(\mathcal{X}, \mathcal{B})$), und die Abbildung $\mathbb{P}_X : \mathcal{B} \to \mathbb{R}$ mit $\mathbb{P}_X(B) := \mathbb{P}(X^{-1}(B))$ für alle $B \in \mathcal{B}$ heißt *induzierte Verteilung* (oder auch nur *Verteilung*) von X unter $\mathbb{P}$.

Lemma 2.19

In der Situation von Definition 2.18 ist $\mathbb{P}_X$ ein W.maß auf $(\mathcal{X}, \mathcal{B})$.

Lemma 2.20

Sind $(\Omega, \mathcal{A}, \mathbb{P})$ ein W.raum, $X : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{B})$ eine Zufallsgröße und $h : (\mathcal{X}, \mathcal{B}) \to (\mathcal{Y}, \mathcal{C})$ eine messbare Abbildung, so ist auch $h(X) := h \circ X : (\Omega, \mathcal{A}) \to (\mathcal{Y}, \mathcal{C})$ eine Zufallsgröße, und es gilt $\mathbb{P}_{h(X)} = (\mathbb{P}_X)_h$.

Zufallsgrößen auf allgemeinen W.räumen II

Da auf allgemeinen W.räumen nur noch Mengen in $\mathcal{A}$ eine W. zugeordnet wird, müssen wir sicherstellen, dass nur solche Mengen als Urbildmengen auftreten.

Definition 2.18 (Zufallsgröße)

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein allgemeiner W.raum, $(\mathcal{X}, \mathcal{B})$ ein messbarer Raum und $X : \Omega \to \mathcal{X}$ eine $\mathcal{A}$-$\mathcal{B}$-messbare Abbildung (kurz: $X : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{B})$), d. h. es gelte $X^{-1}(B) \in \mathcal{A}$ für alle $B \in \mathcal{B}$.
Dann heißt X *Zufallsgröße* oder *Zufallsvariable* (mit Werten in $(\mathcal{X}, \mathcal{B})$), und die Abbildung $\mathbb{P}_X : \mathcal{B} \to \mathbb{R}$ mit $\mathbb{P}_X(B) := \mathbb{P}(X^{-1}(B))$ für alle $B \in \mathcal{B}$ heißt *induzierte Verteilung* (oder auch nur *Verteilung*) von X unter $\mathbb{P}$.

Ist der Grundraum Ω „diskret" bzw. „kontuinierlich", so verzichten wir häufig auf die Angabe der σ-Algebra; es ist dann stets die Potenzmenge $\mathfrak{P}(\Omega)$ bzw. die Borel-σ-algebra $\mathbb{B}|_\Omega$ zugrunde zu legen.

Wir wollen nicht näher darauf eingehen, wann eine Funktion messbar ist, da dies bei in Anwendungen auftretenden Funktionen i. d. R. der Fall ist. Man kann z. B. zeigen, dass jede (stückweise) stetige oder monotone Funktion $f : \mathbb{R} \to \mathbb{R}$ messbar (d. h. $\mathbb{B}$-$\mathbb{B}$-messbar) ist.

Zufallsgrößen auf allgemeinen W.räumen II

Da auf allgemeinen W.räumen nur noch Mengen in $\mathcal{A}$ eine W. zugeordnet wird, müssen wir sicherstellen, dass nur solche Mengen als Urbildmengen auftreten.

Definition 2.18 (Zufallsgröße)

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein allgemeiner W.raum, $(\mathcal{X}, \mathcal{B})$ ein messbarer Raum und $X : \Omega \to \mathcal{X}$ eine $\mathcal{A}$-$\mathcal{B}$-messbare Abbildung (kurz: $X : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{B})$), d. h. es gelte $X^{-1}(B) \in \mathcal{A}$ für alle $B \in \mathcal{B}$.
Dann heißt X *Zufallsgröße* oder *Zufallsvariable* (mit Werten in $(\mathcal{X}, \mathcal{B})$), und die Abbildung $\mathbb{P}_X : \mathcal{B} \to \mathbb{R}$ mit $\mathbb{P}_X(B) := \mathbb{P}(X^{-1}(B))$ für alle $B \in \mathcal{B}$ heißt *induzierte Verteilung* (oder auch nur *Verteilung*) von X unter $\mathbb{P}$.

Ist X eine Zufallsgröße mit Werten in $\mathbb{R}$, so wird die Verteilungsfunktion F_X von $\mathbb{P}_X$ auch als *Verteilungsfunktion von X* bezeichnet:

$$F_X : \mathbb{R} \to \mathbb{R} \qquad F_X(x) = \mathbb{P}_X(]-\infty, x]) = \mathbb{P}(X \in]-\infty, x]) =: \mathbb{P}(X \leq x)\,.$$

Es sei daran erinnert, dass $\mathbb{P}_X$ durch F_X eindeutig bestimmt ist.
Entsprechend wird (im Falle der Existenz!) die Riemann-Dichte f_X von $\mathbb{P}_X$ auch als *Riemann-Dichte von X* bezeichnet.

Bemerkung 2.21 (Bestimmung von Verteilungen von Zufallsgrößen)

Es seien $(\Omega, \mathcal{A}, \mathbb{P})$ ein W.raum, X eine Zufallsgröße mit Werten in $\mathcal{X} \subseteq \mathbb{R}$ und $h : \mathcal{X} \to \mathbb{R}$ eine messbare Funktion. Dann gibt es die folgenden Ansätze zur Bestimmung der Verteilung von $Y := h(X)$:

(a) *(über Elementar-W.'en)*

Gibt es eine abzählbare Menge $\mathcal{Y} \subseteq \mathbb{R}$ mit $\mathbb{P}(Y \in \mathcal{Y}) = 1$, so ist $\mathbb{P}_Y$ ein diskretes W.maß und kann am einfachsten über seine Elementar-W.en bestimmt werden. Dies ist insbesondere (aber nicht nur) der Fall, wenn $\mathbb{P}_X$ ein diskretes W.maß ist.

(b) *(über Verteilungsfunktion)*

Allgemein kann man versuchen, die Verteilung $\mathbb{P}_Y$ über die Verteilungsfunktion F_Y zu bestimmen, indem man $\mathbb{P}(Y \leq y)$ berechnet. Dies bietet sich vor allem an, wenn Y alle Werte in einem (echten) Intervall annimmt.

(c) *(über Dichte)*

...

Beispiel 2.22 (a)
Ist $\mathbb{P}_X = \mathcal{E}(\lambda)$ und $Y := \lceil X \rceil$, so gilt $\mathbb{P}_Y = \mathcal{G}(1 - e^{-\lambda})$.

denn:

Beispiel 2.22 (a)
Ist $\mathbb{P}_X = \mathcal{E}(\lambda)$ und $Y := \lceil X \rceil$, so gilt $\mathbb{P}_Y = \mathcal{G}(1 - e^{-\lambda})$.

denn: Hier nimmt Y m. W. 1 Werte in $\mathbb{N}$ an, so dass wir die Elementar-W.en bestimmen:

Beispiel 2.22 (a)
Ist $\mathbb{P}_X = \mathcal{E}(\lambda)$ und $Y := \lceil X \rceil$, so gilt $\mathbb{P}_Y = \mathcal{G}(1 - e^{-\lambda})$.

denn: Hier nimmt Y m. W. 1 Werte in $\mathbb{N}$ an, so dass wir die Elementar-W.en bestimmen: Für jedes $k \in \mathbb{N}$ gilt

$$\begin{aligned}\mathbb{P}(Y = k) = \mathbb{P}(k-1 < X \leq k) &= \int_{k-1}^{k} \lambda e^{-\lambda t}\, dt \\ &= e^{-\lambda(k-1)} - e^{-\lambda k} = (e^{-\lambda})^{k-1}(1 - e^{-\lambda})\,.\end{aligned}$$

Beispiel 2.22 (a)
Ist $\mathbb{P}_X = \mathcal{E}(\lambda)$ und $Y := \lceil X \rceil$, so gilt $\mathbb{P}_Y = \mathcal{G}(1 - e^{-\lambda})$.

denn: Hier nimmt Y m. W. 1 Werte in $\mathbb{N}$ an, so dass wir die Elementar-W.en bestimmen: Für jedes $k \in \mathbb{N}$ gilt

$$\mathbb{P}(Y = k) = \mathbb{P}(k - 1 < X \leq k) = \int_{k-1}^{k} \lambda e^{-\lambda t}\, dt$$
$$= e^{-\lambda(k-1)} - e^{-\lambda k} = (e^{-\lambda})^{k-1}(1 - e^{-\lambda})\,.$$

Dies sind gerade die Elementar-W.en der geom. Verteilung zum Parameter $1 - e^{-\lambda}$. $\Rightarrow$ Beh.

Beispiel 2.22 (b)
Ist $\mathbb{P}_X = \mathcal{N}(0,1)$ und $Y := \mu + \sigma X$ (wobei $\sigma \neq 0$), so gilt $\mathbb{P}_Y = \mathcal{N}(\mu, \sigma^2)$.

denn:

Beispiel 2.22 (b)
Ist $\mathbb{P}_X = \mathcal{N}(0,1)$ und $Y := \mu + \sigma X$ (wobei $\sigma \neq 0$), so gilt $\mathbb{P}_Y = \mathcal{N}(\mu, \sigma^2)$.

denn: Hier nimmt Y m. W. 1 Werte in $\mathbb{R}$ an, so dass wir F_Y auf $\mathbb{R}$ bestimmen.

Beispiel 2.22 (b)
Ist $\mathbb{P}_X = \mathcal{N}(0,1)$ und $Y := \mu + \sigma X$ (wobei $\sigma \neq 0$), so gilt $\mathbb{P}_Y = \mathcal{N}(\mu, \sigma^2)$.

denn: Hier nimmt Y m. W. 1 Werte in $\mathbb{R}$ an, so dass wir F_Y auf $\mathbb{R}$ bestimmen.
Wir betrachten nur den Fall $\sigma > 0$; der Fall $\sigma < 0$ kann analog behandelt werden.

Bestimmung von Verteilungen von Zufallsgrößen II

Beispiel 2.22 (b)
Ist $\mathbb{P}_X = \mathcal{N}(0,1)$ und $Y := \mu + \sigma X$ (wobei $\sigma \neq 0$), so gilt $\mathbb{P}_Y = \mathcal{N}(\mu, \sigma^2)$.

denn: Hier nimmt Y m. W. 1 Werte in $\mathbb{R}$ an, so dass wir F_Y auf $\mathbb{R}$ bestimmen. Wir betrachten nur den Fall $\sigma > 0$; der Fall $\sigma < 0$ kann analog behandelt werden. Es gilt

$$F_Y(y) = \mathbb{P}(\mu + \sigma X \leq y) \underset{\sigma>0}{=} \mathbb{P}(X \leq \tfrac{y-\mu}{\sigma}) = \Phi(\tfrac{y-\mu}{\sigma}) \qquad (y \in \mathbb{R})$$

wobei $\Phi(x) := \int_{-\infty}^{x} \varphi(t)\,dt$ die Verteilungsfunktion der $\mathcal{N}(0,1)$-Verteilung ist.

Beispiel 2.22 (b)
Ist $\mathbb{P}_X = \mathcal{N}(0,1)$ und $Y := \mu + \sigma X$ (wobei $\sigma \neq 0$), so gilt $\mathbb{P}_Y = \mathcal{N}(\mu, \sigma^2)$.

denn: Hier nimmt Y m. W. 1 Werte in $\mathbb{R}$ an, so dass wir F_Y auf $\mathbb{R}$ bestimmen. Wir betrachten nur den Fall $\sigma > 0$; der Fall $\sigma < 0$ kann analog behandelt werden. Es gilt

$$F_Y(y) = \mathbb{P}(\mu + \sigma X \leq y) \underset{\sigma>0}{=} \mathbb{P}(X \leq \tfrac{y-\mu}{\sigma}) = \Phi(\tfrac{y-\mu}{\sigma}) \qquad (y \in \mathbb{R})$$

wobei $\Phi(x) := \int_{-\infty}^{x} \varphi(t)\, dt$ die Verteilungsfunktion der $\mathcal{N}(0,1)$-Verteilung ist. Da F_Y stetig differenzierbar ist, besitzt $\mathbb{P}_Y$ damit die Riemann-Dichte

$$f_Y(y) = F_Y'(y) = \Phi'(\tfrac{y-\mu}{\sigma})\tfrac{1}{\sigma} = \varphi(\tfrac{y-\mu}{\sigma})\tfrac{1}{|\sigma|} = \varphi_{\mu,\sigma^2}(y) \qquad (y \in \mathbb{R})$$

wobei φ_{μ,σ^2} die Dichte der $\mathcal{N}(\mu, \sigma^2)$-Verteilung ist. $\Rightarrow$ Beh.

Bestimmung von Verteilungen von Zufallsgrößen II

Beispiel 2.22 (b)
Ist $\mathbb{P}_X = \mathcal{N}(0,1)$ und $Y := \mu + \sigma X$ (wobei $\sigma \neq 0$), so gilt $\mathbb{P}_Y = \mathcal{N}(\mu, \sigma^2)$.

denn: Hier nimmt Y m. W. 1 Werte in $\mathbb{R}$ an, so dass wir F_Y auf $\mathbb{R}$ bestimmen. Wir betrachten nur den Fall $\sigma > 0$; der Fall $\sigma < 0$ kann analog behandelt werden. Es gilt

$$F_Y(y) = \mathbb{P}(\mu + \sigma X \leq y) \underset{\sigma>0}{=} \mathbb{P}(X \leq \tfrac{y-\mu}{\sigma}) = \Phi(\tfrac{y-\mu}{\sigma}) \qquad (y \in \mathbb{R})$$

wobei $\Phi(x) := \int_{-\infty}^{x} \varphi(t)\,dt$ die Verteilungsfunktion der $\mathcal{N}(0,1)$-Verteilung ist. Da F_Y stetig differenzierbar ist, besitzt $\mathbb{P}_Y$ damit die Riemann-Dichte

$$f_Y(y) = F_Y'(y) = \Phi'(\tfrac{y-\mu}{\sigma})\tfrac{1}{\sigma} = \varphi(\tfrac{y-\mu}{\sigma})\tfrac{1}{|\sigma|} = \varphi_{\mu,\sigma^2}(y) \qquad (y \in \mathbb{R})$$

wobei φ_{μ,σ^2} die Dichte der $\mathcal{N}(\mu, \sigma^2)$-Verteilung ist. $\Rightarrow$ Beh.

Allgemeiner gilt: Ist $\mathbb{P}_X = \mathcal{N}(\mu, \sigma^2)$ und $Y := a + bX$ (wobei $b \neq 0$), so gilt $\mathbb{P}_Y = \mathcal{N}(a + b\mu, b^2\sigma^2)$.

Beispiel 2.22 (b)
Ist $\mathbb{P}_X = \mathcal{N}(0,1)$ und $Y := \mu + \sigma X$ (wobei $\sigma \neq 0$), so gilt $\mathbb{P}_Y = \mathcal{N}(\mu, \sigma^2)$.

denn: Hier nimmt Y m. W. 1 Werte in $\mathbb{R}$ an, so dass wir F_Y auf $\mathbb{R}$ bestimmen. Wir betrachten nur den Fall $\sigma > 0$; der Fall $\sigma < 0$ kann analog behandelt werden. Es gilt

$$F_Y(y) = \mathbb{P}(\mu + \sigma X \leq y) \underset{\sigma>0}{=} \mathbb{P}(X \leq \tfrac{y-\mu}{\sigma}) = \Phi(\tfrac{y-\mu}{\sigma}) \qquad (y \in \mathbb{R})$$

wobei $\Phi(x) := \int_{-\infty}^{x} \varphi(t)\, dt$ die Verteilungsfunktion der $\mathcal{N}(0,1)$-Verteilung ist. Da F_Y stetig differenzierbar ist, besitzt $\mathbb{P}_Y$ damit die Riemann-Dichte

$$f_Y(y) = F_Y'(y) = \Phi'(\tfrac{y-\mu}{\sigma})\tfrac{1}{\sigma} = \varphi(\tfrac{y-\mu}{\sigma})\tfrac{1}{|\sigma|} = \varphi_{\mu,\sigma^2}(y) \qquad (y \in \mathbb{R})$$

wobei φ_{μ,σ^2} die Dichte der $\mathcal{N}(\mu, \sigma^2)$-Verteilung ist. $\Rightarrow$ Beh.

Allgemeiner gilt: Ist $\mathbb{P}_X = \mathcal{N}(\mu, \sigma^2)$ und $Y := a + bX$ (wobei $b \neq 0$), so gilt $\mathbb{P}_Y = \mathcal{N}(a + b\mu, b^2\sigma^2)$.

denn: Dieser Fall lässt sich auf den bereits behandelten Spezialfall zurückführen. Nach Lemma 2.20 können wir o. E. annehmen, dass $X = \mu + \sigma Z$ mit $Z \sim \mathcal{N}(0,1)$, und dann folgt $a + bX = a + b(\mu + \sigma Z) = (a + b\mu) + (b\sigma)Z \sim \mathcal{N}(a + b\mu, b^2\sigma^2)$.

Beispiel 2.22 (c)
Ist $\mathbb{P}_X = \mathcal{N}(0,1)$ und $Y := X^2$, so gilt $\mathbb{P}_Y = \chi_1^2$.

Dabei sei χ_1^2 das stetige W.maß zur Dichte $f_1(x) = \frac{1}{\sqrt{2\pi}}\, x^{-1/2}\, e^{-x/2}\, \mathbf{1}_{(0,\infty)}(x)$;
χ_1^2 heißt *Chi-Quadrat-Verteilung zum Parameter* 1.

denn:

Bestimmung von Verteilungen von Zufallsgrößen III

Beispiel 2.22 (c)
Ist $\mathbb{P}_X = \mathcal{N}(0,1)$ und $Y := X^2$, so gilt $\mathbb{P}_Y = \chi_1^2$.

Dabei sei χ_1^2 das stetige W.maß zur Dichte $f_1(x) = \frac{1}{\sqrt{2\pi}}\, x^{-1/2}\, e^{-x/2}\, \mathbf{1}_{(0,\infty)}(x)$;
χ_1^2 heißt *Chi-Quadrat-Verteilung zum Parameter* 1.

denn: Hier nimmt Y m. W. 1 Werte in $(0,\infty)$ an, so dass wir F_Y auf $(0,\infty)$ bestimmen.

Beispiel 2.22 (c)
Ist $\mathbb{P}_X = \mathcal{N}(0,1)$ und $Y := X^2$, so gilt $\mathbb{P}_Y = \chi_1^2$.

Dabei sei χ_1^2 das stetige W.maß zur Dichte $f_1(x) = \frac{1}{\sqrt{2\pi}}\, x^{-1/2}\, e^{-x/2}\, \mathbf{1}_{(0,\infty)}(x)$;
χ_1^2 heißt *Chi-Quadrat-Verteilung zum Parameter* 1.

denn: Hier nimmt Y m. W. 1 Werte in $(0,\infty)$ an, so dass wir F_Y auf $(0,\infty)$ bestimmen. Für alle $y > 0$ gilt

$$F_Y(y) = \mathbb{P}(X^2 \leq y) = \mathbb{P}(0 \leq X \leq \sqrt{y}) + \mathbb{P}(-\sqrt{y} \leq X \leq 0) \underset{\text{Symmetrie}}{=} 2(\Phi(\sqrt{y}) - \Phi(0))\,.$$

Beispiel 2.22 (c)
Ist $\mathbb{P}_X = \mathcal{N}(0,1)$ und $Y := X^2$, so gilt $\mathbb{P}_Y = \chi_1^2$.

Dabei sei χ_1^2 das stetige W.maß zur Dichte $f_1(x) = \frac{1}{\sqrt{2\pi}}\, x^{-1/2}\, e^{-x/2}\, \mathbf{1}_{(0,\infty)}(x)$;
χ_1^2 heißt *Chi-Quadrat-Verteilung zum Parameter* 1.

denn: Hier nimmt Y m. W. 1 Werte in $(0,\infty)$ an, so dass wir F_Y auf $(0,\infty)$ bestimmen. Für alle $y > 0$ gilt

$$F_Y(y) = \mathbb{P}(X^2 \leq y) = \mathbb{P}(0 \leq X \leq \sqrt{y}) + \mathbb{P}(-\sqrt{y} \leq X \leq 0) \underset{\text{Symmetrie}}{=} 2(\Phi(\sqrt{y}) - \Phi(0))\,.$$

Da F_Y auf $(0,\infty)$ stetig differenzierbar ist, besitzt $\mathbb{P}_Y$ auf $(0,\infty)$ die Riemann-Dichte $f_Y(y) = F_Y'(y) = 2\varphi(\sqrt{y})\frac{1}{2\sqrt{y}}$. $\Rightarrow$ Beh.

Bestimmung von Randverteilungen

Häufig tritt auch die Situation auf, dass die Verteilung einer „zusammengesetzten“ Zufallsgröße $(X, Y) : \Omega \to \mathbb{R}^2$ gegeben ist und die Verteilungen von X und von Y (die sog. *Randverteilungen*) gesucht sind. Besitzt (X, Y) eine diskrete Verteilung, kann man wie in Lemma 1.19 vorgehen (auch wenn der W.raum selbst nicht diskret ist).

Bestimmung von Randverteilungen

Häufig tritt auch die Situation auf, dass die Verteilung einer „zusammengesetzten“ Zufallsgröße $(X, Y) : \Omega \to \mathbb{R}^2$ gegeben ist und die Verteilungen von X und von Y (die sog. *Randverteilungen*) gesucht sind. Besitzt (X, Y) eine diskrete Verteilung, kann man wie in Lemma 1.19 vorgehen (auch wenn der W.raum selbst nicht diskret ist). Besitzt (X, Y) eine stetige Verteilung, so gilt:

Lemma 2.23 (Bestimmung von Randverteilungen)

Ist $(X, Y) : \Omega \to \mathbb{R}^2$ eine Zufallsgröße mit einer 2-dim. Riemann-Dichte $f_{(X,Y)}$, so besitzt X die 1-dim. Riemann-Dichte f_X mit

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

für alle $x \in \mathbb{R}$. (Insbesondere besitzt X also überhaupt eine Riemann-Dichte!)

Eine analoge Aussage gilt natürlich für Y statt X.

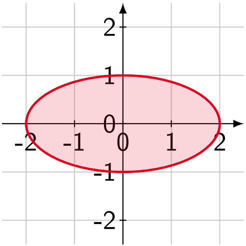
Bestimmung von Randverteilungen

Allgemeiner kann man den Fall betrachten, dass die *n*-dimensionale Riemann-Dichte einer Zufallsgröße $(X_1, \ldots, X_n)$ mit Werten in $\mathbb{R}^n$ gegeben ist und eine *Randverteilung* $\mathbb{P}_{X_i}$ $(i = 1, \ldots, n)$ bzw. sogar eine *mehrdimensionale Randverteilung* $\mathbb{P}_{(X_{i_1}, \ldots, X_{i_m})}$ $(1 \leq i_1 < \cdots < i_m \leq n)$ gesucht ist. Hier gilt:

Ist die W.dichte einer zusammengesetzten Zufallsgröße $(X_1, \ldots, X_n)$ mit Werten in $\mathbb{R}^n$ gegeben, so erhalten wir die W.dichten der Randverteilungen, indem wir über die „freien" Variablen ausintegrieren.

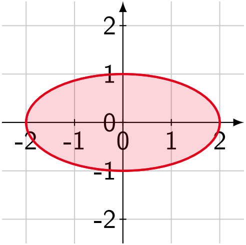
Beispiel 2.24
Die Zufallsgröße (X, Y) sei gleichverteilt auf der Menge $\Omega := \{(x, y) \in \mathbb{R}^2 : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1\}$ (wobei $a, b > 0$), d. h. die Zufallsgröße besitze die 2-dim. Riemann-Dichte

$$f_{(X,Y)}(x, y) = \frac{1}{\mathrm{vol}(\Omega)} \mathbf{1}_\Omega(x, y) = \frac{1}{\pi ab} \mathbf{1}_{\{\frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1\}} .$$



Skizze $(a = 2,\ b = 1)$

Beispiel 2.24
Die Zufallsgröße (X, Y) sei gleichverteilt auf der Menge $\Omega := \{(x,y) \in \mathbb{R}^2 : \frac{x^2}{a^2} + \frac{y^2}{b^2} \le 1\}$ (wobei $a, b > 0$), d. h. die Zufallsgröße besitze die 2-dim. Riemann-Dichte



Skizze $(a = 2,\ b = 1)$

$$f_{(X,Y)}(x,y) = \frac{1}{\mathrm{vol}(\Omega)} \mathbf{1}_\Omega(x,y) = \frac{1}{\pi ab} \mathbf{1}_{\{\frac{x^2}{a^2}+\frac{y^2}{b^2}\le 1\}} \,.$$

Dann sind $\mathbb{P}_X$ und $\mathbb{P}_Y$ durch die 1-dim. Riemann-Dichten

$$f_X(x) = \int_{\mathbb{R}} f_{(X,Y)}(x,y)\,dy \begin{cases} = \displaystyle\int_{-b\sqrt{1-(x^2/a^2)}}^{+b\sqrt{1-(x^2/a^2)}} \frac{1}{\pi ab}\,dy = \frac{2}{\pi a^2}\sqrt{a^2-x^2} & ; |x| \le a \\ = 0 & ; |x| > a \end{cases}$$

$$f_Y(y) = \int_{\mathbb{R}} f_{(X,Y)}(x,y)\,dx \begin{cases} = \displaystyle\int_{-a\sqrt{1-(y^2/b^2)}}^{+a\sqrt{1-(y^2/b^2)}} \frac{1}{\pi ab}\,dx = \frac{2}{\pi b^2}\sqrt{b^2-y^2} & ; |y| \le b \\ = 0 & ; |y| > b \end{cases}$$

gegeben (*Halbkreisverteilung*).

Anwendung: Simulation von Verteilungen

- Bei der *Simulation* versucht man, dass Verhalten realer Systeme anhand eines Modells (etwa auf dem Computer) nachzuvollziehen.
- Bei der Simulation zufälliger Systeme ist man oft darauf angewiesen, Realisierungen von Zufallsgrößen mit gegebenen Verteilungen zu erzeugen. (Man spricht dann auch von der *Simulation von Verteilungen*.)
- Routine `rand()` $\rightsquigarrow$ $\mathcal{U}_{(0,1)}$-verteilte (Pseudo-)Zufallszahlen
- Ausdruck `phi(rand())` $\rightsquigarrow$ Q-verteilte (Pseudo-)Zufallszahlen (wobei Q eine vorgegebene Verteilung ist)

Anwendung: Simulation von Verteilungen

- Bei der *Simulation* versucht man, dass Verhalten realer Systeme anhand eines Modells (etwa auf dem Computer) nachzuvollziehen.
- Bei der Simulation zufälliger Systeme ist man oft darauf angewiesen, Realisierungen von Zufallsgrößen mit gegebenen Verteilungen zu erzeugen. (Man spricht dann auch von der *Simulation von Verteilungen*.)
- Routine `rand()` $\rightsquigarrow$ $\mathcal{U}_{(0,1)}$-verteilte (Pseudo-)Zufallszahlen
- Ausdruck `phi(rand())` $\rightsquigarrow$ Q-verteilte (Pseudo-)Zufallszahlen (wobei Q eine vorgegebene Verteilung ist)

Problem:
Finde zu einer vorgegebenen Verteilung Q eine geeignete Transformation φ!

Lösung: (zumindest im Prinzip)

Satz 2.25 (Inversionsmethode)

Ist U eine $\mathcal{U}_{(0,1)}$-verteilte Zufallsgröße, ist F eine Verteilungsfunktion und ist $G(u) := \inf\{x \in \mathbb{R} : F(x) \geq u\}$ die sog. Pseudo-Inverse von F, so ist $\mathbb{P}_{G \circ U}$ das W.maß zur Verteilungsfunktion F.

Zusammenfassung

- *Eine Zufallsgröße auf einem allgemeinen W.raum ist eine* (messbare) *Abbildung* $X : \Omega \to \mathcal{X}$.
- *Oft interessiert uns vor allem die Verteilung einer Zufallsgröße; diese lässt sich für eine* reellwertige *Zufallsgröße* X *z. B. über die Verteilungsfunktion* $F_X(x) = \mathbb{P}(X \leq x)$ *beschreiben.*

Kapitel 3
Bedingte Wahrscheinlichkeiten und Unabhängigkeit

Kapitel 3.1
Bedingte Wahrscheinlichkeiten

Bedingte Wahrscheinlichkeiten

Häufig stehen gewisse Teilinformationen über den Ausgang eines Zufalls-experiments zur Verfügung.

Beispiel 3.1 (Würfel)

Ein gezinkter Würfel liefert die Ergebnisse $1-6$ *mit den folgenden W.en:*

ω	1	2	3	4	5	6
$f(\omega)$	$\frac{1}{25}$	$\frac{3}{25}$	$\frac{3}{25}$	$\frac{4}{25}$	$\frac{4}{25}$	$\frac{10}{25}$

Wir würfeln, ohne das Ergebnis sehen zu können, und erfahren, dass wir keine Sechs gewürfelt haben. Mit welcher W. ist das Ergebnis eine gerade Zahl?

Bedingte Wahrscheinlichkeiten

Häufig stehen gewisse Teilinformationen über den Ausgang eines Zufallsexperiments zur Verfügung.

Beispiel 3.1 (Würfel)

Ein gezinkter Würfel liefert die Ergebnisse $1-6$ *mit den folgenden W.en:*

ω	1	2	3	4	5	6
$f(\omega)$	$\frac{1}{25}$	$\frac{3}{25}$	$\frac{3}{25}$	$\frac{4}{25}$	$\frac{4}{25}$	$\frac{10}{25}$

Wir würfeln, ohne das Ergebnis sehen zu können, und erfahren, dass wir keine Sechs gewürfelt haben. Mit welcher W. ist das Ergebnis eine gerade Zahl?

Sei B das Ereignis „keine Sechs" und A das Ereignis „gerade Zahl".
Wir interessieren uns dann für die W. von A unter der *Bedingung*, dass B eintritt.
Um diese W. zu bestimmen, orientieren wir uns an relativen Häufigkeiten.
Wenn wir n-mal würfeln und nur noch diejenigen Durchführungen betrachten, bei denen B eintritt, ...

Bedingte Wahrscheinlichkeiten

Häufig stehen gewisse Teilinformationen über den Ausgang eines Zufallsexperiments zur Verfügung.

Beispiel 3.1 (Würfel)

Ein gezinkter Würfel liefert die Ergebnisse $1-6$ *mit den folgenden W.en:*

ω	1	2	3	4	5	6
$f(\omega)$	$\frac{1}{25}$	$\frac{3}{25}$	$\frac{3}{25}$	$\frac{4}{25}$	$\frac{4}{25}$	$\frac{10}{25}$

Wir würfeln, ohne das Ergebnis sehen zu können, und erfahren, dass wir keine Sechs gewürfelt haben. Mit welcher W. ist das Ergebnis eine gerade Zahl?

Sei B das Ereignis „keine Sechs" und A das Ereignis „gerade Zahl".
Wir interessieren uns dann für die W. von A unter der *Bedingung*, dass B eintritt.
Um diese W. zu bestimmen, orientieren wir uns an relativen Häufigkeiten.
Wenn wir n-mal würfeln und nur noch diejenigen Durchführungen betrachten, bei denen B eintritt, so ist die *bedingte* relative Häufigkeit von A

$$\frac{H_n(A\cap B)}{H_n(B)} = \frac{\frac{1}{n}H_n(A\cap B)}{\frac{1}{n}H_n(B)} \approx \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{2,4\})}{\mathbb{P}(\{6\}^c)} = \frac{\frac{3}{25}+\frac{4}{25}}{1-\frac{10}{25}} = \frac{7}{15};$$

die *bedingte* W. von A sollte also $\frac{7}{15}$ sein.

Bedingte Wahrscheinlichkeiten

Wir definieren allgemein:

Definition 3.2 (Bedingte Wahrscheinlichkeit)

Ist $(\Omega, \mathcal{A}, \mathbb{P})$ ein W.raum und ist $B \in \mathcal{A}$ ein Ereignis mit $\mathbb{P}(B) > 0$, so heißt für jedes Ereignis $A \in \mathcal{A}$

$$\mathbb{P}(A \,|\, B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

(elementare) bedingte Wahrscheinlichkeit von A unter B / gegeben B.

Anschaulich beschreibt $\mathbb{P}(A \,|\, B)$ die Chance des Eintretens von A, wenn schon feststeht, dass B eintritt.

Beispiel 3.3 (Kugelziehung)

Aus einer Urne mit r roten und s schwarzen Kugeln wird $2 \times$ nacheinander gezogen. Wie groß ist die W., dass in der 2. Ziehung „rot“ gezogen wird, wenn in der 1. Ziehung „schwarz“ gezogen worden ist?

Beispiel 3.3 (Kugelziehung)

Aus einer Urne mit r roten und s schwarzen Kugeln wird $2\times$ nacheinander gezogen. Wie groß ist die W., dass in der 2. Ziehung „rot“ gezogen wird, wenn in der 1. Ziehung „schwarz“ gezogen worden ist?

Ziehen mit Zurücklegen:
Wähle $\Omega := \{1, \dots, r+s\}^2$, $\mathbb{P} := \mathcal{U}_\Omega$ (wegen Symmetrie auf der Ebene der Kugeln), $A := \{\omega \in \Omega : \omega_2 \leq r\}$, $B := \{\omega \in \Omega : \omega_1 > r\}$. Dann gilt

$$\mathbb{P}(A \,|\, B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\frac{rs}{(r+s)^2}}{\frac{s}{r+s}} = \frac{r}{r+s}\,.$$

Beispiel 3.3 (Kugelziehung)

Aus einer Urne mit r roten und s schwarzen Kugeln wird $2\times$ nacheinander gezogen. Wie groß ist die W., dass in der 2. Ziehung „rot" gezogen wird, wenn in der 1. Ziehung „schwarz" gezogen worden ist?

Ziehen mit Zurücklegen:
Wähle $\Omega := \{1,\dots,r+s\}^2$, $\mathbb{P} := \mathcal{U}_\Omega$ (wegen Symmetrie auf der Ebene der Kugeln), $A := \{\omega \in \Omega : \omega_2 \le r\}$, $B := \{\omega \in \Omega : \omega_1 > r\}$. Dann gilt

$$\mathbb{P}(A \,|\, B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\frac{rs}{(r+s)^2}}{\frac{s}{r+s}} = \frac{r}{r+s}\,.$$

Ziehen ohne Zurücklegen:
Wähle $\Omega := \{(\omega_1,\omega_2) \in \{1,\dots,r+s\}^2 : \omega_1 \neq \omega_2\}$, $\mathbb{P} := \mathcal{U}_\Omega$ (wegen Symmetrie auf der Ebene der Kugeln), $A := \{\omega \in \Omega : \omega_2 \le r\}$, $B := \{\omega \in \Omega : \omega_1 > r\}$. Dann gilt

$$\mathbb{P}(A \,|\, B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\frac{rs}{(r+s)(r+s-1)}}{\frac{s}{r+s}} = \frac{r}{r+s-1}\,.$$

Bedingte Wahrscheinlichkeiten

Definition 3.2 (Bedingte Wahrscheinlichkeit)

Ist $(\Omega, \mathcal{A}, \mathbb{P})$ ein W.raum und ist $B \in \mathcal{A}$ ein Ereignis mit $\mathbb{P}(B) > 0$, so heißt für jedes Ereignis $A \in \mathcal{A}$

$$\mathbb{P}(A \,|\, B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

(elementare) bedingte Wahrscheinlichkeit von A unter B / gegeben B.

Bedingte Wahrscheinlichkeiten

Definition 3.2 (Bedingte Wahrscheinlichkeit)

Ist $(\Omega, \mathcal{A}, \mathbb{P})$ ein W.raum und ist $B \in \mathcal{A}$ ein Ereignis mit $\mathbb{P}(B) > 0$, so heißt für jedes Ereignis $A \in \mathcal{A}$

$$\mathbb{P}(A \,|\, B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

<u>(elementare) bedingte Wahrscheinlichkeit von A unter B / gegeben B</u>.

Lemma 3.4

Ist $(\Omega, \mathcal{A}, \mathbb{P})$ ein W.raum und ist $B \in \mathcal{A}$ mit $\mathbb{P}(B) > 0$, so ist die Abbildung

$$\mathbb{P}(\,\cdot\,|\,B) : \mathcal{A} \to \mathbb{R} \qquad A \mapsto \mathbb{P}(A \,|\, B)$$

ein W.maß auf $\mathcal{A}$ mit $P(B \,|\, B) = 1$ und $\mathbb{P}(B^c \,|\, B) = 0$.

Bedingte Wahrscheinlichkeiten

Definition 3.2 (Bedingte Wahrscheinlichkeit)

Ist $(\Omega, \mathcal{A}, \mathbb{P})$ ein W.raum und ist $B \in \mathcal{A}$ ein Ereignis mit $\mathbb{P}(B) > 0$, so heißt für jedes Ereignis $A \in \mathcal{A}$

$$\mathbb{P}(A \,|\, B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

(elementare) bedingte Wahrscheinlichkeit von A unter B / gegeben B.

Lemma 3.4

Ist $(\Omega, \mathcal{A}, \mathbb{P})$ *ein W.raum und ist* $B \in \mathcal{A}$ *mit* $\mathbb{P}(B) > 0$*, so ist die Abbildung*

$$\mathbb{P}(\,\cdot\,|\,B) : \mathcal{A} \to \mathbb{R} \qquad A \mapsto \mathbb{P}(A \,|\, B)$$

ein W.maß auf $\mathcal{A}$ *mit* $P(B \,|\, B) = 1$ *und* $\mathbb{P}(B^c \,|\, B) = 0$.

Bemerkung

Damit bleiben alle Rechenregeln auch für bedingte W.en gültig.