



$$\Omega = \{B, W\}$$

$$P(W) = P(B) = \frac{1}{2}$$

$$1_B(\omega) = \begin{cases} 1 & \text{wenn } \omega = B, \\ 0 & \text{wenn } \omega = W. \end{cases}$$

$$\frac{1}{n} \sum_{i=1}^n 1_B(\omega_i) = ?$$

### Maßraum

Ein Maßraum  $(\Omega, \mathcal{A})$  ist ein Tupel bestehend aus der Grundmenge  $\Omega$  und einer  $\sigma$ -Algebra  $\mathcal{A} \subset \mathcal{P}(\Omega)$

### Maß

Ein Maß auf einem Maßraum  $(\Omega, \mathcal{A})$  ist Abbildung  $\mu : \mathcal{A} \rightarrow \mathbb{R}_{\geq 0}$

$$\mu\left(\bigcup_i A_i\right) = \sum_i \mu(A_i), \text{ mit } A_i \cap A_j = \emptyset \text{ für } i \neq j$$

### Wahrscheinlichkeitsmaß

Ein Maß mit  $\mu(\Omega) = 1$  ist ein Wahrscheinlichkeitsmaß.

### erzeugte Sigma-Algebra

Sei  $X$  eine Menge und  $\mathcal{C} \subseteq \mathcal{P}(X)$  ein Mengensystem.

Die von  $\mathcal{C}$  erzeugte  $\sigma$ -Algebra ist definiert als

$$\sigma(\mathcal{C}) = \bigcap \{ \mathcal{F} \subseteq \mathcal{P}(X) \mid \mathcal{F} \text{ ist eine } \sigma\text{-Algebra und } \mathcal{C} \subseteq \mathcal{F} \}.$$

- $\mathcal{P}(X)$  selbst ist eine  $\sigma$ -Algebra und enthält  $\mathcal{C}$ .
- Daher ist die Menge aller  $\sigma$ -Algebren, die  $\mathcal{C}$  enthalten, nicht leer.
- Der Schnitt beliebig vieler  $\sigma$ -Algebren ist wieder eine  $\sigma$ -Algebra.
- Damit existiert und ist eindeutig die kleinste  $\sigma$ -Algebra mit  $\mathcal{C} \subseteq \sigma(\mathcal{C})$ .
- $\sigma(\mathcal{C})$  enthält genau diejenigen Mengen, die aus  $\mathcal{C}$  gewonnen werden können durch
  - Komplementbildung,
  - abzählbare Vereinigungen,
  - (und folglich abzählbare Durchschnitte).

### Borellsche Sigma-Algebra auf $\mathbb{R}$

Die *Borel- $\sigma$ -Algebra*  $\mathcal{B}(\mathbb{R})$  ist definiert als

$$\mathcal{B}(\mathbb{R}) = \sigma(\{\text{offene Teilmengen von } \mathbb{R}\}).$$

Äquivalent:

$$\mathcal{B}(\mathbb{R}) = \sigma(\{(a, b) : a < b\}) = \sigma(\{(-\infty, a) : a \in \mathbb{R}\}).$$

Jede offene Menge  $U$  lässt sich darstellen als

$$U = \bigcup_{k=1}^{\infty} (a_k, b_k), \quad a_k, b_k \in \mathbb{Q}.$$

$$(a, b) = (-\infty, b) \cap ((-\infty, a])^c$$

$$(-\infty, a] = \bigcap_{n=1}^{\infty} (-\infty, a + \frac{1}{n}) \in \sigma(\mathcal{H})$$

### Meßbare Abbildung

Eine Abbildung  $f : \Omega \rightarrow \Omega'$  zwischen zwei Maßräumen  $(\Omega, \mathcal{A})$  und  $(\Omega', \mathcal{A}')$  heißt meßbar, falls

$$f^{-1}(A') \in \mathcal{A} \text{ für alle } A' \in \mathcal{A}'$$

### Meßbare Abbildung

Das Urbild jedes Ereignisses ist ein Ereignis

### Beispiel

Bei endlichen Mengen mit der Potenzmenge als Sigma-Algebra ist jede Funktion Meßbar.

## Meßbare Funktionen

Die Menge der meßbaren Funktionen  $f : \Omega \rightarrow \mathbb{R}$  bezeichnen wir mit  $\mathcal{M}_\Omega$  oder einfach  $\mathcal{M}$  wenn der Kontext klar ist. Mit  $\mathcal{M}^+$  bezeichnen wir die meßbaren Funktionen  $f : \Omega \rightarrow \mathbb{R}$  mit  $f(\omega) \geq 0$ .



## Messbarkeit von Abbildungen

Sei  $(X, \Sigma)$  ein Messraum,  $Y$  eine Menge und

$$\mathcal{E} \subseteq \mathcal{P}(Y), \quad \mathcal{T} = \sigma(\mathcal{E})$$

die von  $\mathcal{E}$  erzeugte  $\sigma$ -Algebra auf  $Y$ . Dann ist

$$f: (X, \Sigma) \longrightarrow (Y, \mathcal{T})$$

genau dann messbar, wenn

$$f^{-1}(E) \in \Sigma \quad \text{für alle } E \in \mathcal{E}.$$

## Indikatorfunktion

Für eine Teilmenge  $A \subset \Omega$  heißt

$$1_A(x) := \begin{cases} 1 & \text{falls } x \in A \\ 0 & \text{sonst} \end{cases}$$

Indikatorfunktion.

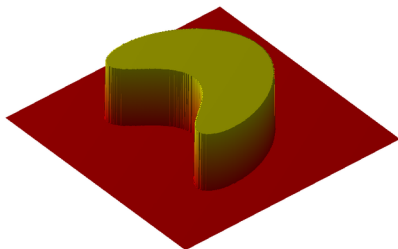


Figure: Quelle: Wikipedia:

[https://commons.wikimedia.org/wiki/File:Indicator\\_function\\_illustration.png](https://commons.wikimedia.org/wiki/File:Indicator_function_illustration.png)

## Treppenfunktion

Eine meßbare Funktion  $u : \Omega \rightarrow \mathbb{R}$  heißt Treppenfunktion, falls sie nur endlich viele verschiedene Werte annimmt.

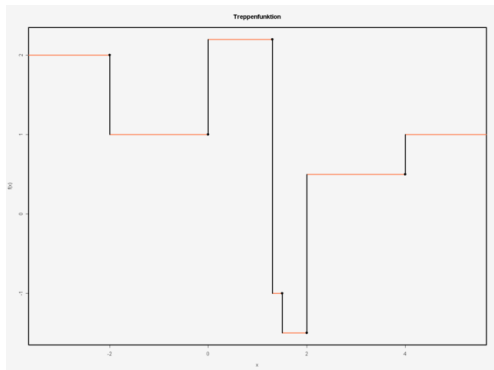


Figure: Quelle: Wikipedia:

<https://commons.wikimedia.org/wiki/File:Stepfunction1.png>

## Beispiel einer Treppenfunktion

$u(x) = \sum_{i=1}^n a_i \cdot 1_{A_i}(x)$  mit  $A_i \cap A_j = \emptyset$  für  $i \neq j$  ist eine Treppenfunktion.

## Treppenfunktion

Eine Treppenfunktion  $u$  hat eine Darstellung

$u(x) = \sum_{i=1}^n a_i \cdot 1_{A_i}(x)$  mit  $A_i \cap A_j = \emptyset$  für  $i \neq j$  ist eine Treppenfunktion.

## Treppenfunktion

Die Menge der Treppenfunktionen bezeichnen wir mit  $\mathcal{T}$  und die Treppenfunktionen mit  $a_i > 0$  mit  $\mathcal{T}^+$ .

## Eindeutigkeit der Darstellung

Sind  $u = \sum_{i=1}^n a_i 1_{A_i} = \sum_{j=1}^m b_j 1_{B_j}$  zwei verschiedene Darstellungen einer Treppenfunktion  $u \in \mathcal{T}^+$  so ist  
$$\sum_{i=1}^n a_i \mu(A_i) = \sum_{j=1}^m b_j \mu(B_j)$$

## Integral einer Treppenfunktion

Für eine Treppenfunktion  $u \in \mathcal{T}^+$  definieren wir

$$\int_{\Omega} u \, d\mu = \sum_{i=1}^n a_i \mu(A_i)$$

Diese ist unabhängig von der Darstellung.

## Eigenschaften des Integrals von Treppenfunktionen

Sind  $u$  und  $v$  zwei Treppenfunktionen, dann gilt:

- $\int_{\Omega} 1_A d\mu = \mu(A)$
- $\int_{\Omega} \alpha u + \beta v d\mu = \alpha \int_{\Omega} u d\mu + \beta \int_{\Omega} v d\mu$
- Ist  $u(x) \leq v(x)$  für alle  $x$ , so ist  $\int_{\Omega} u d\mu \leq \int_{\Omega} v d\mu$

## Integral nicht negativer meßbarer Funktionen

Für eine Funktion  $f \in \mathcal{M}^+$  definieren wir

$$\int_{\Omega} f \, d\mu := \sup \left\{ \int_{\Omega} u \, d\mu \mid u \in \mathcal{T}^+, u(x) \leq f(x) \text{ für alle } x \in \Omega \right\}$$

## Meßbare Abbildungen

Eine nicht negative messbare Funktion  $f : \Omega \rightarrow \mathbb{R}_{\geq 0}$  ist genau dann meßbar, wenn es eine Folge  $f_n \in \mathcal{T}^+$  gibt mit  $f_n \uparrow f$ .



Sei  $f \in \mathcal{M}^+$  meßbar:  
definiere

$$A_{j,n} := \begin{cases} \{ \frac{j}{2^n} \leq f \leq \frac{j+1}{2^n} \} & \text{für } j = 0, \dots, n \cdot 2^n - 1 \\ \{ f \geq n \} & \text{für } j = n \cdot 2^n \end{cases}$$

und damit

$$f_n := \sum_{j=0}^{n2^n} \frac{j}{2^n} 1_{A_{j,n}}$$

Damit gilt  $f_n(x) \leq f(x) \leq f_n(x) + 2^{-n}$ .

Sei

$$f_n(x) = \sum_{i=1}^{k_n} a_{n,i} \mathbf{1}_{A_{n,i}}(x),$$

eine punktweise konvergente Folge  $f_n \uparrow f$  von Treppenfunktionen.

**1. Treppenfunktionen sind messbar.**

Für feste  $n$  und  $\alpha \in \mathbb{R}$  gilt

$$\{x : f_n(x) < \alpha\} = \bigcup_{\substack{1 \leq i \leq k_n \\ a_{n,i} < \alpha}} A_{n,i} \in \Sigma,$$

da endliche und abzählbare Vereinigungen sowie Komplemente in  $\Sigma$  liegen. Also ist  $f_n$  messbar.

**2. Charakterisierung der Messbarkeit.**

Eine Funktion  $g: X \rightarrow \mathbb{R}$  ist genau dann messbar, wenn für alle  $\alpha \in \mathbb{R}$

$$\{x : g(x) < \alpha\} \in \Sigma.$$

Da  $f_n(x) \rightarrow f(x)$  punktweise, gilt für jedes  $x \in X$  und jedes  $\alpha \in \mathbb{R}$ :

$$f(x) < \alpha \iff \exists k \forall n \geq k : f_n(x) < \alpha.$$

Daher

$$\{x : f(x) < \alpha\} = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \{x : f_n(x) < \alpha\}.$$

## Integral für meßbare Funktionen

Für eine meßbare Funktion  $f \in \mathcal{M}$  setzen wir

$$\int_{\Omega} f \, d\mu = \int_{\Omega} f^{+} \, d\mu - \int_{\Omega} -f^{-} \, d\mu$$

wobei  $f^{+}(x) := \max(0, f(x))$  und  $f^{-}(x) := \min(0, f(x))$

## Integral für meßbare Funktionen

Eine meßbare Funktion heißt integrierbar, falls ihr Integral endlich ist.

## Eigenschaften des Integrals

Sind  $f$  und  $g$  zwei meßbare Funktionen, dann gilt:

- $\int_{\Omega} 1_A d\mu = \mu(A)$
- $\int_{\Omega} \alpha f + \beta g d\mu = \alpha \int_{\Omega} f d\mu + \beta \int_{\Omega} g d\mu$
- Ist  $f(x) \leq g(x)$  für alle  $x$ , so ist  $\int_{\Omega} f d\mu \leq \int_{\Omega} g d\mu$
- $\left| \int_{\Omega} f d\mu \right| \leq \int_{\Omega} |f| d\mu$

## Zufallsvariable

Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum,  $(R, \mathcal{B})$  ein Messraum.  
Eine Zufallsvariable ist eine messbare Abbildung  $X : \Omega \rightarrow R$ .

## Reelle Zufallsvariable

## Verteilung und Unabhängigkeit

Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum,  $(R, \mathcal{B})$  ein Messraum und  $\{X_i\}_{i=1}^n$  ein Folge von Zufallsvariablen  $X_i : \Omega \rightarrow R$ . Die Zufallsvariablen heißen identisch verteilt, falls  $P_{X_i} = P_{X_j}$  für alle  $i, j$  und stochastisch unabhängig, falls  $P_{(X_1, \dots, X_n)} = \prod_{i=1}^n P_{X_i}$  gilt.

## Erwartungswert

Für eine reelle integrierbare Zufallsvariable ist ihr Erwartungswert definiert durch

$$\mathbb{E}(X) := \int_{\Omega} X \, dP .$$

## Erwartungswert

Ist  $(\Omega, \mathcal{A}, P)$  ein diskreter Wahrscheinlichkeitsraum und  $X : \Omega \rightarrow \mathbb{R}$  eine reelle Zufallsvariable, so ist

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega) \cdot P(\omega)$$



## Eigenschaften

Sind  $X, Y : \Omega \rightarrow \mathbb{R}^n$  reelle, integrierbare Zufallsvariablen und  $a, b \in \mathbb{R}$  konstant, so gilt:

$$\mathbb{E}(a \cdot X + b \cdot Y) = a \cdot \mathbb{E}(X) + b \cdot \mathbb{E}(Y)$$

$$X(x) \leq Y(x) \quad \forall x \in \Omega \Rightarrow \mathbb{E}(X) \leq \mathbb{E}(Y)$$

$$X, Y \text{ stoch. unabhängig} \Rightarrow \mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$$

$$\mathbb{E}(1_A) = P(A)$$

## Varianz

Für eine reelle Zufallsvariable ist die Varianz definiert durch

$$\mathbb{V}(X) := \mathbb{E}((X - \mathbb{E}(X))^2) .$$

## Verschiebungssatz

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{E}(X^2 - 2X\mathbb{E}(X) + \mathbb{E}(X)^2) = \mathbb{E}(X^2) - 2\mathbb{E}(X)^2 + \mathbb{E}(X)^2 \\ &= \mathbb{E}(X^2) - \mathbb{E}(X)^2\end{aligned}$$

## Kovarianz

Für reelle Zufallsvariable  $X, Y$  ist die Kovarianz definiert durch

$$\mathcal{C}(X, Y) := \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) .$$

## Kovarianz

Per Definition ist

$$\mathcal{C}(X, X) := \mathbb{V}(X).$$

## Beispiel

$\Omega = \{\text{Kopf}, \text{Zahl}\}, P(\text{Kopf}) = P(\text{Zahl}) = \frac{1}{2},$   
 $X(\text{Kopf}) = 0, X(\text{Zahl}) = 1$

$$\begin{aligned}\mathbb{E}(X) &= 0 \cdot P(X^{-1}(0)) + 1 \cdot P(X^{-1}(1)) \\ &= 0 \cdot P(\text{Kopf}) + 1 \cdot P(\text{Zahl}) = \frac{1}{2}\end{aligned}$$

## Markov Ungleichung

Sei  $Y : \Omega \rightarrow \mathbb{R}$  eine reelle, integrierbare Zufallsvariable und  $f : [0, \infty) \rightarrow [0, \infty)$  monoton wachsend. Dann gilt für alle  $\epsilon > 0$  mit  $f(\epsilon) > 0$

$$P(|Y| \geq \epsilon) \leq \frac{\mathbb{E}(f \circ |Y|)}{f(\epsilon)}$$

## Beweis

Da  $f(\epsilon)1_{\{|Y| \geq \epsilon\}} \leq f \circ |Y|$  folgt

$$\begin{aligned} f(\epsilon)P(|Y| \geq \epsilon) &= f(\epsilon)\mathbb{E}(1_{\{|Y| \geq \epsilon\}}) = \mathbb{E}(f(\epsilon)1_{\{|Y| \geq \epsilon\}}) \\ &\leq \mathbb{E}(f \circ |Y|) \end{aligned}$$

## Tschebyscheff-Ungleichung

Für eine reelle, integrierbare und quadratintegrierbare Zufallsvariable  $Y : \Omega \rightarrow \mathbb{R}$  gilt:

$$P(|Y - \mathbb{E}(Y)| \geq \epsilon) \leq \frac{\mathbb{V}(Y)}{\epsilon^2}$$

## Beweis

Folgt direkt aus der Markov-Ungleichung mit  $Y' = Y - \mathbb{E}(Y)$  und  $f(x) = x^2$



## Schwaches Gesetz der großen Zahlen

Seien  $X_i : \Omega \rightarrow \mathbb{R}$  unabhängige, reelle Zufallsvariablen (iiv, iid(englisch)) mit  $\mathbb{E}(X_i) = \mu < \infty$  und  $\mathbb{V}(X_i) = \sigma < \infty$ , dann gilt

$$P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \mu\right| \geq \epsilon\right) \leq \frac{\sigma}{n \cdot \epsilon^2} \xrightarrow{n \rightarrow \infty} 0$$

(stochastische Konvergenz).



## Beweis

Mit  $Y_n = \frac{1}{n} \sum_{i=1}^n X_i - \mu$  ist  $\mathbb{E}(Y_n) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i - \mu) = 0$  und  $\mathbb{V}(Y_n) = \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}(X_i) = \frac{\sigma}{n}$ . Aus der Tschebyscheff-Ungleichung folgt die Behauptung.

# Erwartungswert

