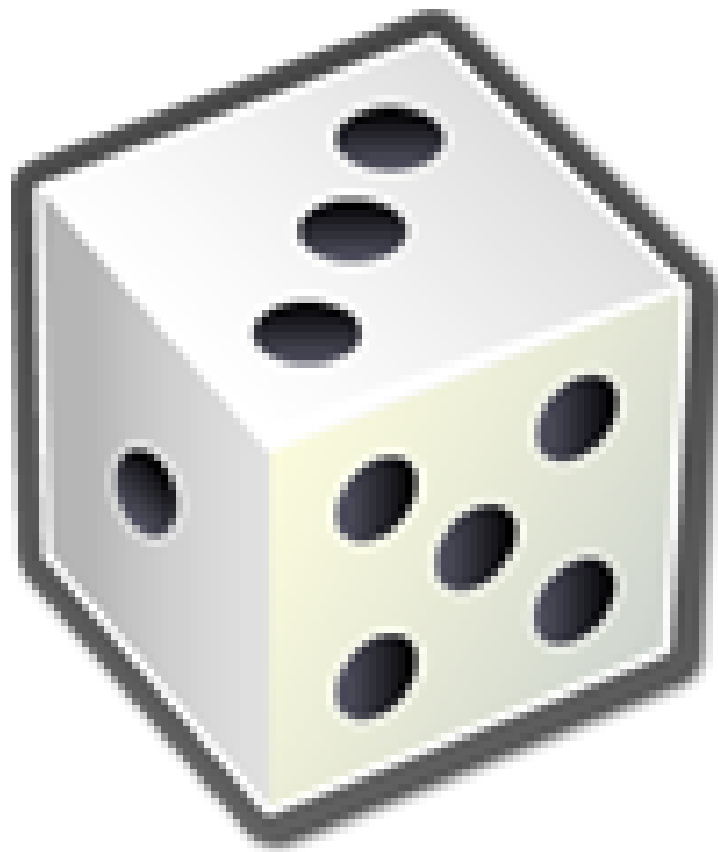


# Stochastik

Dr. Johannes Riesterer

25. April 2021



©Johannes Riesterer

## Vorwort

Kann jeder Mathematik lernen? Als Antwort auf diese Frage möchte ich auf den interessanten Lebenslauf einer der bedeutendsten Mathematikerinnen aller Zeiten eingehen (Auszug aus Wikipedia):

Emmy Noether war eine deutsche Mathematikerin, die grundlegende Beiträge zur abstrakten Algebra und zur theoretischen Physik lieferte. Insbesondere hat Noether die Theorie der Ringe, Körper und Algebren revolutioniert. Das nach ihr benannte Noether-Theorem gibt die Verbindung zwischen Symmetrien von physikalischen Naturgesetzen und Erhaltungsgrößen an.

Sie zeigte in mathematischer Richtung keine besondere Frühreife, sondern hatte in ihrer Jugend Interesse an Musik und Tanzen. Sie besuchte die Städtische Höhere Töchterschule – das heutige Marie-Therese-Gymnasium – in der Schillerstraße in Erlangen. Mathematik wurde dort nicht intensiv gelehrt. Im April 1900 legte sie die Staatsprüfung zur Lehrerin der englischen und französischen Sprache an Mädchenschulen in Ansbach ab. 1903 holte sie in Nürnberg die externe Abiturprüfung am Königlichen Realgymnasium – dem heutigen Willstätter-Gymnasium – nach.

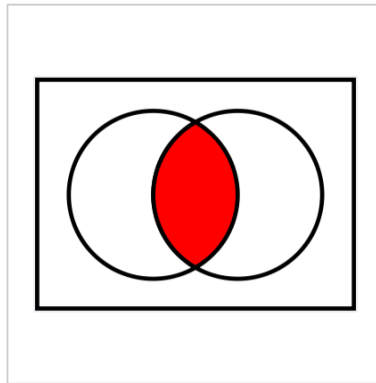
# Inhaltsverzeichnis

<b>1</b>	<b>Notationen</b>	<b>6</b>
<b>2</b>	<b>Diskrete Modelle</b>	<b>7</b>
2.1	Naiver Bayes'scher Spam Filter . . . . .	9
<b>3</b>	<b>Zufallsvariablen</b>	<b>10</b>
3.1	Integration bezüglich eines Wahrscheinlichkeitsmaßes . . . . .	10
3.2	Reelle Zufallsvariablen . . . . .	11
3.3	Erwartungswert . . . . .	12
3.4	Erwartungswert . . . . .	14
3.5	Varianz . . . . .	15
	<b>Tabellenverzeichnis</b>	<b>17</b>
	<b>Abbildungsverzeichnis</b>	<b>18</b>

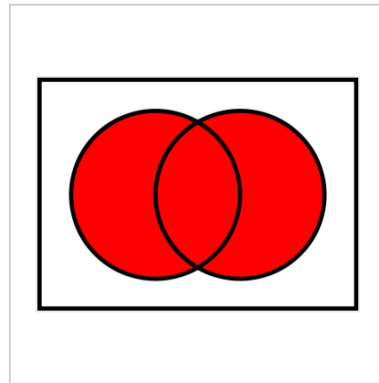


## 1 Notationen

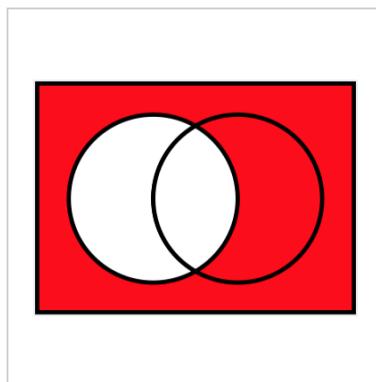
Für eine Menge  $\Omega$  bezeichnet  $\#\Omega$  die Anzahl ihrer Elemente und  $\mathcal{P}(\Omega) := \{A \mid A \subseteq \Omega\}$  die Potenzmenge von  $\Omega$ .



Intersection of two sets  $A \cap B$



Union of two sets  $A \cup B$



Absolute complement of A in U

$$A^c = U \setminus A$$

Abbildung 1: Quelle: Wikipedia

Schnitt	Vereinigung	
$A \cap B = B \cap A$	$A \cup B = B \cup A$	Kommutativgesetze
$A \cap (B \cap C) = (A \cap B) \cap C$	$A \cup (B \cup C) = (A \cup B) \cup C$	Assoziativgesetze
$A \cap A = A$	$A \cup A = A$	Idempotenzgesetze
$A \cap G = A$	$A \cup \{\} = A$	Neutralitätsgesetze
$A \cap \{\} = \{\}$	$A \cup G = G$	Extremalgesetze
$A \cap \bar{A} = \{\}$	$A \cup \bar{A} = G$	Komplementärsgesetze
$A \cap \bar{B} = \bar{A} \cup \bar{B}$	$\bar{A} \cup \bar{B} = \overline{A \cap B}$	De Morgansche Gesetze
$A \cap (A \cup B) = A$	$A \cup (A \cap B) = A$	Absorptionsgesetze

$G$ : Grundmenge

Distributivgesetze:

1.  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
2.  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Involution:

$$\overline{\overline{A}} = A$$

Abbildung 2: Quelle: Wikipedia

## 2 Diskrete Modelle

**Definition 1** (Laplace Wahrscheinlichkeit). Sei  $\Omega = \{\omega_1, \dots, \omega_n\}$  eine endliche Menge. Für  $A \subseteq \Omega$  definieren wir die Wahrscheinlichkeit durch

$$P(A) := \frac{\#A}{\#\Omega}.$$

Die Elemente  $\omega_1, \dots, \omega_n \in \Omega$  nennen wir Elementarereignisse und Teilmengen  $A \subseteq \Omega$  Ereignisse. Die Abbildung

$$P : \mathcal{P}(\Omega) \rightarrow [0, 1]$$

nennen wir die Laplace-Verteilung bzw. Gleichverteilung auf  $\Omega$ .

**Definition 2** (Variationen und Kombinationen).

- $Var_k^n(\Omega, m.W.) := \{(\omega_1, \dots, \omega_k) \mid \omega_i \in \Omega\}$  Menge aller Variationen mit Wiederholung.
- $Var_k^n(\Omega, o.W.) := \{(\omega_1, \dots, \omega_k) \mid \omega_i \in \Omega; \omega_i \neq \omega_j\}$  Menge aller Variationen ohne Wiederholung.
- $Kom_k^n(\Omega, m.W.) := \{(\omega_{i_1}, \dots, \omega_{i_k}) \mid \omega_{i_l} \in \Omega; 1 \leq i_1 \leq \dots \leq i_k\}$  Menge aller Kombinationen mit Wiederholung.
- $Kom_k^n(\Omega, o.W.) := \{(\omega_{i_1}, \dots, \omega_{i_k}) \mid \omega_{i_l} \in \Omega; 1 \leq i_1 \leq \dots \leq i_k; \omega_{i_l} \neq \omega_{i_j}\}$  Menge aller Kombinationen ohne Wiederholung.

**Lemma 1.**

- $\#Var_k^n(\Omega, m.W.) = n^k = \underbrace{n \cdot n \cdots n}_{k\text{-mal}}$
- $\#Var_k^n(\Omega, o.W.) = n_k = n \cdot (n-1) \cdots (n-(k-1)) = \frac{n!}{(n-k)!}$
- $\#Kom_k^n(\Omega, o.W.) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$
- $\#Kom_k^n(\Omega, m.W.) = \binom{n+k-1}{k}$

**Definition 3** (Diskrete Wahrscheinlichkeitsverteilung). Es sei  $\Omega$  eine (nicht leere) abzählbare Menge. Eine Abbildung  $P : \mathcal{P}(\Omega) \rightarrow [0, 1]$  heißt diskrete Wahrscheinlichkeitsverteilung (Wahrscheinlichkeitsmaß), falls gilt:

$$P(\Omega) = 1$$

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n), \text{ mit } A_i \cap A_j = \emptyset \text{ für } i \neq j$$

**Beispiel 1** (Laplace Wahrscheinlichkeit).  $\Omega$  endlich und  $P(A) = \frac{\#A}{\#\Omega}$ .

**Definition 4** (Bedingte Wahrscheinlichkeit). Für  $A, B \in \mathcal{A}$  und  $P(B) > 0$  heißt

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

die bedingte Wahrscheinlichkeit (von  $A$  unter  $B$ ).

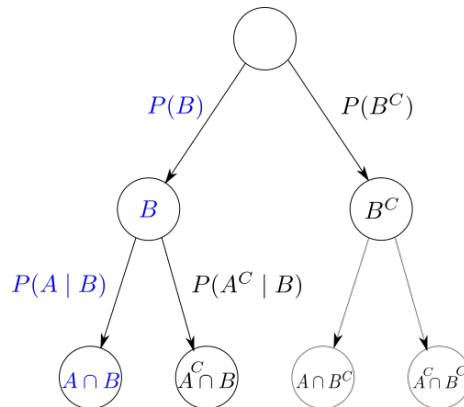


Abbildung 3: Quelle: Wikipedia

**Satz 1** (Satz der totalen Wahrscheinlichkeit). Für eine Zerlegung  $\Omega = \bigcup_{j=1}^n B_j$ , mit  $B_i \cap B_k = \emptyset$  für  $i \neq k$

$$P(A) = \sum_{j=1}^n P(A | B_j) \cdot P(B_j)$$



*Beweis.*  $A = A \cap B \cup A \cap \bar{B}$  und  $A \cap B \cap A \cap \bar{B} = \emptyset$ .

$$P(A) = P(A \cap B) + P(A \cap \bar{B}) = P(B) \cdot P(A | B) + P(\bar{B}) \cdot P(A | \bar{B})$$

□

**Satz 2** (Satz von Bayes). Für  $A, B \in \mathcal{A}$  mit  $P(B) > 0$  gilt

$$P(A | B) = \frac{P(B | A) \cdot P(A)}{P(B)}$$

*Beweis.*

$$P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{P(A \cap B) \cdot P(A)}{P(A)}}{P(B)} = \frac{P(B | A) \cdot P(A)}{P(B)}$$

□

**Definition 5** (Stochastische Unabhängigkeit). Zwei Ereignisse  $A, B$  heißen stochastisch unabhängig, falls

$$P(A \cap B) = P(A) \cdot P(B)$$

gilt. Somit ist auch  $P(A|B) = P(A)$  und  $P(B|A) = P(B)$ .

## 2.1 Naiver Bayes'scher Spam Filter

Gegeben ist eine E-Mail  $E$ . Wir möchten anhand des Vorkommens bestimmter Wörter  $A_1, \dots, A_n$  in der Mail entscheiden, ob es sich um eine erwünschte Mail  $H$  oder eine unerwünschte Mail  $S$  (Ham or Spam) handelt. (Typische Wörter wären zum Beispiel "reichwerden", "onlinecasino"...). Aus einer Datenbank kann man das Vorkommen dieser Wörter in Spam und Ham Mails zählen und damit empirisch die Wahrscheinlichkeiten  $P(A_i|S)$  und  $P(A_i|H)$  des Vorkommens dieser Wörter in Spam und Ham Mails ermitteln. Wir gehen davon aus, dass es sich bei der Mail prinzipiell mit Wahrscheinlichkeit  $P(E = S) = P(E = H) = \frac{1}{2}$  um eine erwünschte Mail  $H$  oder eine unerwünschte Mail  $S$  handeln kann. Wir machen zudem die (naive) Annahme, dass das Vorkommen der Wörter stochastisch unabhängig ist, also

$$\begin{aligned} P(A_1 \cap \dots \cap A_n | S) &= P(A_1 | S) \cdots P(A_n | S) \\ P(A_1 \cap \dots \cap A_n | H) &= P(A_1 | H) \cdots P(A_n | H) \end{aligned}$$

gilt.

Mit der Formel von Bayes und der totalen Wahrscheinlichkeit können wir somit berechnen

$$\begin{aligned} P(E = S | A_1 \cap \dots \cap A_n) &= \frac{P(A_1 \cap \dots \cap A_n | S) \cdot P(S)}{P(A_1 \cap \dots \cap A_n)} \\ &= \frac{P(A_1 | S) \cdots P(A_n | S) \cdot P(S)}{P(A_1 \cap \dots \cap A_n | H) + P(A_1 \cap \dots \cap A_n | S)} \\ &= \frac{P(A_1 | S) \cdots P(A_n | S) \cdot P(S)}{P(A_1 | H) \cdots P(A_n | H) + P(A_1 | S) \cdots P(A_n | S)} \end{aligned}$$

Bemerkung:  $P(E = H | A_1 \cap \dots \cap A_n) = 1 - P(E = S | A_1 \cap \dots \cap A_n)$

### 3 Zufallsvariablen

**Definition 6.** Ein Messraum ist ein Paar  $(\Omega, \mathcal{A})$  bestehend aus einer Menge  $\Omega$  und einer Sigma-Algebra  $\mathcal{A} \subset \mathcal{P}(\Omega)$ .

**Definition 7.** Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum und  $(\Omega', \mathcal{A}')$  ein Messraum. Eine Zufallsvariable ist eine Abbildung

$$X : \Omega \rightarrow \Omega'$$

so dass für alle Ereignisse  $A' \in \mathcal{A}'$

$$X^{-1}(A') \in \mathcal{A}$$

ein Ereignis in  $\mathcal{A}$  ist. Urbilder von Ereignissen sind also Ereignisse.

**Definition 8.** Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum,  $(\Omega', \mathcal{A}')$  ein Messraum und  $X : \Omega \rightarrow \Omega'$  Eine Zufallsvariable. Durch

$$P_X(A') := P(X^{-1}(A'))$$

für  $A' \in \mathcal{A}'$  wird ein Wahrscheinlichkeitsmaß auf  $(\Omega', \mathcal{A}')$  definiert. Es wird Bildmaß genannt. Anstelle von  $P_X(A')$  wird auch die Schreibweise  $P(X \in A') := P_X(A')$  verwendet.

**Definition 9.** Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum,  $(R, \mathcal{B})$  ein Messraum und  $\{X_i\}_{i=1}^n$  eine 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.

#### 3.1 Integration bezüglich eines Wahrscheinlichkeitsmaßes

in diesem Abschnitt ist  $(\Omega, \mathcal{A}, P)$  ein fest gewählter Wahrscheinlichkeitsraum.

**Definition 10.** Für eine Teilmenge  $A \in \mathcal{A}$  heißt

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

Indikatorfunktion.

**Definition 11.** Eine Funktion

$$\phi(x) := \sum_{k=1}^m c_k 1_{A_k}(x)$$

mit  $c_k \in \mathbb{R}$  und  $A_k \in \mathcal{A}$  mit  $A_i \cap A_j = \emptyset$  für  $i \neq j$  heißt Treppenfunktion.

**Definition 12.** Eine Hüllreihe zu einer Funktion  $f : \Omega \rightarrow \mathbb{R}$  ist eine Reihe  $\phi(x) := \sum_{k=1}^{\infty} c_k 1_{A_k}(x)$  mit den folgenden Eigenschaften:

- $c_k \in \mathbb{R}$  sind positive reelle Zahlen  $c_k > 0$ .
- $A_k \in \mathcal{A}$ .
- Für alle  $x \in \Omega$  gilt  $|f(x)| \leq \phi(x)$ .

**Definition 13.** Für eine Treppenfunktion  $\phi(x) := \sum_{k=1}^m c_k 1_{A_k}(x)$  definieren wir das Integral durch

$$\int_{\Omega} \phi dP := \sum_{k=1}^m c_k P(A_k).$$

**Definition 14.** Der Inhalt einer Hüllreihe  $\phi(x) := \sum_{k=1}^{\infty} c_k 1_{A_k}(x)$  ist definiert durch

$$I_P(\phi) := \sum_{k=1}^{\infty} c_k P(A_k).$$

**Definition 15.** Die  $L_{P^1}$ -Halbnorm einer Funktion  $f : \Omega \rightarrow \mathbb{R}$  ist definiert durch das Infimum der Inhalte der Hüllreihen zu  $f$

$$\|f\|_{P^1} := \inf \left\{ I(\phi) \mid \phi \text{ ist Hüllreihe zu } f \right\}.$$

**Definition 16.** Eine Funktion  $f : \Omega \rightarrow \mathbb{R}$  heißt integrierbar, falls eine Folge von Treppenfunktionen  $\phi_k$  existiert mit

$$\|f - \phi_k\|_{P^1} \rightarrow 0 \text{ für } k \rightarrow \infty.$$

In diesem Fall heißt

$$\int_{\Omega} f(x) dP := \lim_{k \rightarrow \infty} \int_{\Omega} \phi_k dP$$

das Integral von  $f$  über  $\Omega$ .

### 3.2 Reelle Zufallsvariablen

**Definition 17.** Eine Menge  $U \subset \mathbb{R}^n$  heißt offen, falls für jeden Punkt  $x \in U$  ein Radius  $\epsilon > 0$  existiert, so dass der Ball  $B_{\epsilon}(x)$  in  $U$  enthalten ist, also  $B_{\epsilon}(x) \subset U$  gilt.

**Definition 18.** Die Borel'sche  $\sigma$ -Algebra  $\mathcal{B}(\mathbb{R}^n)$  über  $\mathbb{R}^n$  ist die kleinste  $\sigma$ -Algebra, die alle offenen Mengen  $U$  enthält, also

$$\mathcal{A}_{\sigma}(U) := \bigcap \{ \mathcal{A} \subset \mathcal{P}(\mathbb{R}^n); U \in \mathcal{A}, \mathcal{A} \text{ ist } \sigma\text{-Algebra} \}$$

**Satz 3.** Die Borel'sche  $\sigma$ -Algebra existiert, da die Potenzmenge eine  $\sigma$ -Algebra ist.

**Satz 4.** Die Borel'sche  $\sigma$ -Algebra ist in der  $\sigma$ -Algebra der Lebesgue messbaren Mengen enthalten.

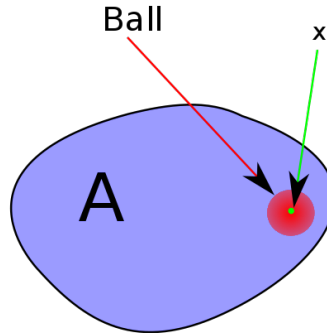


Abbildung 4: Quelle: Wikipedia

**Definition 19.** *Unter einer reellen Zufallsvariable verstehen wir eine Zufallsvariable*

$$X : \Omega \rightarrow \mathbb{R}^n$$

$$X(\omega) := \left( X_1(\omega), \dots, X_n(\omega) \right),$$

wobei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum ist und  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$  der  $\mathbb{R}^n$  zusammen mit der Borell'schen Sigma-Algebra ist. Das Integral ist komponentenweise definiert durch

$$\int_{\Omega} X dP := \left( \int_{\Omega} X_1 dP, \dots, \int_{\Omega} X_n dP \right)$$

### 3.3 Erwartungswert

**Satz 5.** *Eine reelle Zufallsvariable ist integrierbar.*

**Definition 20.** *Für eine reelle Zufallsvariable heißt*

$$F_X : \Omega \rightarrow [0, 1]$$

$$F_X(x) := P(X \leq x) := P_X((-\infty, x]) = P(X^{-1}((-\infty, x]))$$

*Verteilungsfunktion von  $X$ .*

**Definition 21.** *Sei  $\Omega \subset \mathbb{R}^n$  und  $(\Omega, \mathcal{A})$  ein Messraum wobei alle  $A \in \mathcal{A}$  Lebesgue-messbar sind. Eine Funktion  $f : \Omega \rightarrow \mathbb{R}$  heißt Dichte, falls für ihr Lebesgue-Integral  $\int_{\Omega} f d\mu = 1$  gilt.*

**Beispiel 2.** Die Funktion  $f(x) = \frac{1}{\sqrt{\pi}}e^{-x^2}$  ist eine Dichte auf  $\mathbb{R}$ .

$$I := \int_0^\infty e^{-x^2} dx$$

$$I^2 = \int_0^\infty \int_0^\infty e^{-(x^2+y^2)} dx dy$$

$$x = r \cos \varphi, y = r \sin \varphi, r^2 = x^2 + y^2 \text{ (da } \cos^2 + \sin^2 = 1)$$

LINK: Polarkoordinatentransformation

$$= \int_0^{\frac{\pi}{2}} \int_0^\infty r \cdot e^{-r^2} dr d\varphi$$

$$= \frac{\pi}{2} \int_0^\infty r \cdot e^{-r^2} dr$$

$$= -\frac{\pi}{4} [e^{-r^2}]_0^\infty = \frac{\pi}{4} \Rightarrow I = \frac{\sqrt{\pi}}{2}$$

Analog beweist man, dass für alle  $\mu \in \mathbb{R}, \sigma > 0$  die Funktion  $f(x) := \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx$  eine Dichte auf  $\mathbb{R}$  ist.

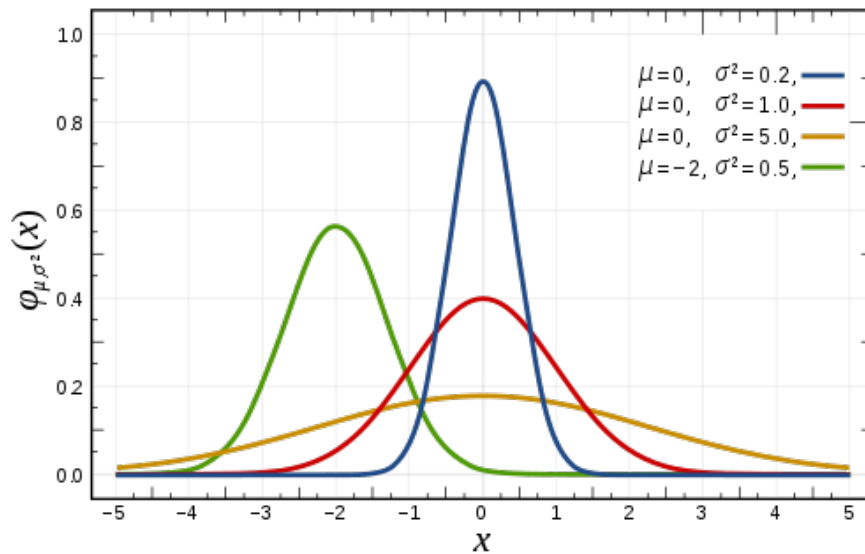


Abbildung 5: Quelle: Wikipedia

**Definition 22** (Normalverteilung). Eine reelle Zufallsvariable  $X : \Omega \rightarrow \mathbb{R}$  heisst Normalverteilt, wenn  $F_X(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx$  mit  $\mu \in \mathbb{R}, \sigma > 0$  gilt. Man schreibt auch  $X \sim \mathcal{N}(\mu, \sigma^2)$ .

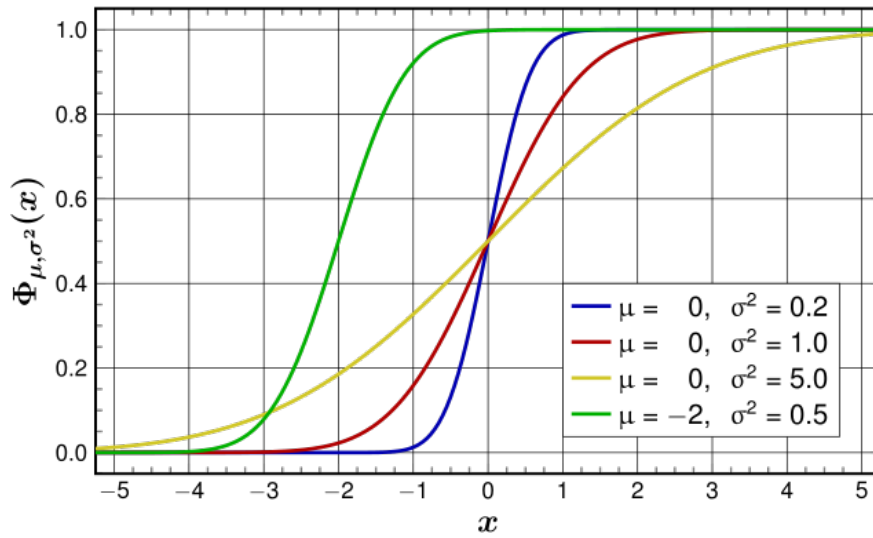


Abbildung 6: Quelle: Wikipedia

### 3.4 Erwartungswert

**Definition 23.** Für eine reelle Zufallsvariable ist der Erwartungswert definiert durch

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

**Bemerkung 1.** Ist  $(\Omega, \mathcal{A}, P)$  ein diskreter Wahrscheinlichkeitsraum und  $X : \Omega \rightarrow \mathbb{R}$  eine eindimensionale reelle Zufallsvariable, so ist

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

**Satz 6.** Für eine reelle Zufallsvariable  $X : \Omega \rightarrow \mathbb{R}^n$  und eine integrierbare Funktion  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  gilt

$$\mathbb{E}(g \circ X) := \int_{\Omega} g \circ X dP = \int_{\mathbb{R}^n} g dP_X.$$

Ist insbesondere  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$  eine Dichte für  $P_X$ , so ist

$$\mathbb{E}(X) = \int_{\mathbb{R}^n} x \cdot f(x) d\mu$$

das Lebesgue-Integral der Funktion  $x \cdot f(x)$ . Für eine Treppenfunktion  $g = \sum_{i=1}^n c_i \mathbf{1}_{A_i}$  folgt das Ergebnis aus der Linearität des Integrals. Für integrierbares  $g$  folgt das Resultat mit Hilfe von Konvergenzsätzen für das Integral.

*Beweis.* Für  $g = \mathbf{1}_A$  mit  $A \in \mathcal{B}(\mathbb{R}^n)$  ist

$$\begin{aligned} \int \mathbf{1}_A dP_X &= P_X(A) = P(X^{-1}(A)) = \int \mathbf{1}_{X^{-1}(A)} dP \\ &= \int \mathbf{1}_A \circ X dP \end{aligned}$$

□

**Beispiel 3.**  $\Omega = \{\text{Kopf}, \text{Zahl}\}$ ,  $P(\text{Kopf}) = P(\text{Zahl}) = \frac{1}{2}$ ,  $X(\text{Kopf}) = 0$ ,  $X(\text{Zahl}) = 1$

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

**Beispiel 4.** Sei  $X \sim \mathcal{N}(\mu, \sigma^2)$ .

$$\begin{aligned}\mathbb{E}(X) &:= \int_{\mathbb{R}} x \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\ &= \int_{\mathbb{R}} (y+\mu) \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}y^2} dy \\ &= \mu \int_{\mathbb{R}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}y^2} dy + \int_{\mathbb{R}} y \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}y^2} dy = \mu\end{aligned}$$

**Satz 7.** Sind  $X, Y : \Omega \rightarrow \mathbb{R}$  reelle, integrierbare Zufallsvariablen und  $a, b \in \mathbb{R}$  konstant, so gilt:

$$\begin{aligned}\mathbb{E}(a \cdot X + b \cdot Y) &= a \cdot \mathbb{E}(X) + b \cdot \mathbb{E}(Y) \\ X(x) \leq Y(x) \quad \forall x \in \Omega &\Rightarrow \mathbb{E}(X) \leq \mathbb{E}(Y) \\ X, Y \text{ stoch. unabhängig} &\Rightarrow \mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y) \\ \mathbb{E}(1_A) &= P(A)\end{aligned}$$

**Satz 8.** Markov-Ungleichung Sei  $Y : \Omega \rightarrow \mathbb{R}$  eine reelle, integrierbare Zufallsvariablen 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

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

□

### 3.5 Varianz

**Definition 24.** Für eine reelle Zufallsvariable ist die Varianz definiert durch

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

**Satz 9.** Verschiebungssatz

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

**Beispiel 5.** Sei  $X \sim \mathcal{N}(\mu, \sigma^2)$ .

$$\begin{aligned}\mathbb{V}(X) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x(xe^{-\frac{x^2}{2}}) dx \\ &= \frac{1}{\sqrt{2\pi}} \left( \left[ x(e^{-\frac{x^2}{2}}) \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} -e^{-\frac{x^2}{2}} dx \right) = 0 + 1 = 1\end{aligned}$$

*LINK: Partielle Integration. Mit "Verschiebungstrick"  $\Rightarrow \mathbb{V}(X) = \sigma^2$ .*

**Satz 10.** *Tschebyscheff-Ungleichung Für eine reelle, integrierbare und quadratintegrierbare Zufallsvariablen  $Y : \Omega \rightarrow \mathbb{R}$  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  $f(x) = x^2$  □

**Satz 11.** *Schwaches Gesetz der großen Zahlen Seien  $\{X_i\}_i : \Omega \rightarrow \mathbb{R}$  unabhängige, reelle Zufallsvariablen (iid, 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. □



## Tabellenverzeichnis

## Abbildungsverzeichnis

1	Quelle: Wikipedia . . . . .	6
2	Quelle: Wikipedia . . . . .	7
3	Quelle: Wikipedia . . . . .	8
4	Quelle: Wikipedia . . . . .	12
5	Quelle: Wikipedia . . . . .	13
6	Quelle: Wikipedia . . . . .	14