

---

---

# CHAPTER 1

---

## WAHRSCHEINLICHKEIT UND STATISTIK

1	Kombinatorik . . . . .	2
2	Ereignisse . . . . .	3
3	Wahrscheinlichkeit . . . . .	5
4	Zufallsvariable . . . . .	7
5	Erwartungswert . . . . .	9
6	Varianz . . . . .	12
7	Bewertung vom Mittelwert / Erwartungswert . . . . .	13
8	Lineare Regression . . . . .	14
9	Verteil- und Dichtefunktionen	14
10	Gleichverteilung . . . . .	18

# 1 Kombinatorik

Die Kombinatorik ist eine Teildisziplin der Mathematik, die sich mit endlichen oder abzählbar unendlichen diskreten Strukturen beschäftigt und deshalb auch dem Oberbegriff Diskrete Mathematik zugerechnet wird.

## Definition 1.1: Produktregel (Für-jede-gibt-es-Regel)

Wenn es für jede von  $n$  Wahlmöglichkeiten  $m$  mögliche weitere Wahlmöglichkeiten gibt, dann gibt es insgesamt  $n \cdot m$  Möglichkeiten.

$$n_1 \cdot n_2 \dots n_k = \prod_{i=1}^k n_i \quad (1.1)$$

## Definition 1.2: Permutation (Anordnung)

Gegeben sind  $n$  anzuordnende Objekte. Die Fakultät  $n!$  beschreibt dabei die Anzahl Arten wie dies geschehen kann.

$$\begin{aligned} P_n &= \#\{\text{Plätze für 1. Objekt (n)}\} \cdot \#\{\text{Plätze für 2. Objekt (n-1)}\} \cdot \\ &\quad \dots \cdot \#\{\text{Plätze für n. Objekt (1)}\} \\ &= n \cdot (n-1) \dots 1 \\ &= n! \end{aligned} \quad (1.2)$$

## Definition 1.3: Kombination (Auswahl)

Die Anzahl Objekte welche aus  $n$  von  $k$  möglichen ausgewählt werden (ohne zurücklegen) wird beschrieben durch:

$$\begin{aligned} \frac{n!}{k!(n-k)!} &= \binom{n}{k} \\ \text{Binomialkoeffizient } (a+b)^n &= \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \\ \text{oder auch } (1+z)^n &= \sum_{k=0}^n \binom{n}{k} z^k \end{aligned} \quad (1.3)$$

## Definition 1.4: Variation (Kerlenketten)

Die Arten auf welche man  $k$  mal unter  $n$  verschiedenen Objekten auswählen kann wird beschrieben durch:

$$= n \cdot n^{k-1} = n^k \quad (1.4)$$

## 1.1 Erzeugende Funktion

Ist die formale Potenzreihe einer Folge  $a_n$ . Diese kann verwendet werden um Zahlen  $a_n$  in Funktionen zu codieren. Damit lassen sich kombinatorische Probleme mit algebraischer Manipulation von erzeugenden Funktionen lösen. Z.B. Gegeben 3 Einfränkler, 4 Zweifränkler und 7 Fünfliber, auf wieviele Arten kann man 39 Franken bezahlen?

$$\begin{aligned}
 1\text{Fr} &= z \\
 0-3 \text{ Einfränkler} &= f_1(z) = 1 + z + z^2 + z^3 \\
 0-4 \text{ Zweifränkler} &= f_2(z) = 1 + z^2 + z^4 + z^6 + z^8 \\
 0-7 \text{ Fünfliber} &= f_3(z) = 1 + z^5 + \dots + z^{35} \\
 \text{Alle } f_1(z) \cdot f_2(z) \cdot f_3(z) &= \\
 &= 1 + z + 2z^2 + 2z^3 + \dots + 4z^{23} + \dots
 \end{aligned} \tag{1.5}$$

## 2 Ereignisse

### Definition 2.1: Elementarereignis

Der Ausgang eines Experimentes heisst Elementarereignis, die Menge aller Elementarereignisse wird mit  $\Omega$  bezeichnet. Ein Elementarereignis  $\omega$  ist also ein Element von  $\Omega$ ,  $\omega \in \Omega$ .

### Definition 2.2: Ereignis

Ist  $\Omega$  eine Menge von Versuchsausgängen, dann heisst eine Teilmenge  $A \subset \Omega$  ein Ereignis. Man sagt, das Ereignis  $A$  ist eingetreten, wenn bei einer Durchführung des Experimentes ein Versuchsausgang  $\omega \in A$  aufgetreten ist.

### Definition 2.3: Ereignisalgebra

Eine Ereignisalgebra  $(\Omega, A)$  ist eine Menge  $\Omega$  mit einer Menge  $A \subset \text{Potenzmenge}(\Omega)$  von Teilmengen von  $\Omega$ , die folgende Bedingungen erfüllen:

1. Vereinigungen von Elementen von  $A$  sind ebenfalls in  $A$ , also

$$B, C \in A \Rightarrow B \cup C \in A$$

2. Differenzen von Elementen von  $A$  sind in  $A$ , also

$$B, C \in A \Rightarrow B \setminus C \in A$$

3.  $\Omega \in A$  d.h. es gibt das sichere Ereignis.

Sind nur die Bedingungen 1 und 2 erfüllt, spricht man auch von einem Mengen-Ring. Eine Ereignisalgebra heisst manchmal auch ein Mengenkörper. Aus den Axiomen folgt direkt:

1. Es gibt das unmögliche Ereignis:  $\emptyset$
2. Das Komplement eines Ereignisses ist ebenfalls ein Ereignis
3. Der Durchschnitt zweier Ereignisse ist ebenfalls ein Ereignis.

$$\begin{aligned} A \cap (B \cup C) &= (A \cap B) \cup (A \cap C) \\ A \cup (B \cap C) &= (A \cup B) \cap (A \cup C) \\ \overline{A \cap B} &= \bar{A} \cup \bar{B} \text{ and vice versa} \end{aligned} \tag{1.6}$$

### 3 Wahrscheinlichkeit

Die Wahrscheinlichkeit soll ein Mass dafür sein, dass ein Ereignis eintritt. Es gibt verschiedene Ansätze, wie wir zu einem solchen Mass kommen könnten:

1. Je häufiger ein Ereignis eintritt, desto grösser sollte die Wahrscheinlichkeit sein. Dies setzt voraus, dass das Experiment im Prinzip beliebig oft wiederholbar ist. Man nennt dies den frequentistischen Ansatz.
2. Die Wahrscheinlichkeit ist ein Mass für die persönliche Überzeugung, dass ein Ereignis eintreten wird. Im Unterschied zum frequentistischen Ansatz sollte man bei diesem Ansatz einem Ereignis auch dann eine Wahrscheinlichkeit geben können, wenn sich ein Experiment nicht wirklich wiederholen lässt. Dies ist der Bayessche Ansatz.
3. Wir könnten eine Reihe von plausiblen Axiomen postulieren, nach denen sich die Wahrscheinlichkeit zu verhalten hat, und dann zu untersuchen, ob ein solches Objekt tatsächlich existiert. Dabei ist uns egal, was der Wahrscheinlichkeitswert genau bedeutet. Dieser axiomatische Ansatz hat den Vorteil, logisch konsistent zu sein, was bei den anderen Ansätzen nicht von vornherein garantiert ist.

In allen Fällen ergibt sich eine Reihe von Gesetzmässigkeiten oder Formeln, welche die jeweiligen Wahrscheinlichkeitsbegriffe erfüllen müssen. Soll die Wahrscheinlichkeit eine objektive Grösse sein, dann muss für wiederholbare Experimente die für den Bayesschen Ansatz nötige persönliche Überzeugung direkt mit der Häufigkeit des Eintretens zusammenhängen. Der frequentistische und der Bayessche Ansatz werden also in diesem Fall übereinstimmen. Die Axiome im axiomatischen Ansatz sind natürlich genau die Rechenregeln, die man sowohl im frequentistischen Ansatz wie auch im Bayesschen Ansatz von der Wahrscheinlichkeit erwartet. Man darf daher davon ausgehen, dass die drei Ansätze die gleichen numerischen Resultate liefern, sie unterscheiden sich höchstens in der Interpretation der Resultate.

#### Definition 3.1: Wahrscheinlichkeit als relative Häufigkeit

$$P(A) = \lim_{N \rightarrow \infty} \frac{n}{N} \quad (1.7)$$

**Definition 3.2: Axiome eines Wahrscheinlichkeitsraumes****Wertebereich**

$$0 \leq P(A) \leq 1 \quad (1.8)$$

**Sicheres Ereignis**

$$P(\Omega) = 1 \quad (1.9)$$

**Vereinigung.** Sind die Ereignisse  $A_1, A_2, \dots$  paarweise disjunkt, also  $A_i \cap A_j = \emptyset$  für  $i \neq j$ , dann gilt

$$P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots \quad (1.10)$$

Aus den Axiomen folgt:

1. Die Wahrscheinlichkeit des unmöglichen Ereignisses ist

$$P(\emptyset) = 0$$

2. Die Wahrscheinlichkeit des komplementären Ereignisses ist

$$P(\bar{A}) = P(\Omega \setminus A) = 1 - P(A)$$

3. Die Wahrscheinlichkeit der Differenz der Ereignisse A und B ist

$$P(A \setminus B) = P(A) - P(A \cap B)$$

4. Die Wahrscheinlichkeit der Vereinigung zweier beliebiger Ereignisse ist (Ein-/Ausschaltformel)

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

**Definition 3.3: Laplace Experiment**

Ein Experiment mit  $n = |\Omega|$  Ausgängen heisst ein Laplace-Experiment, wenn jeder Versuchsausgang gleich wahrscheinlich mit Wahrscheinlichkeit

$$P(\omega) = \frac{1}{n}, \omega \in \Omega \quad (1.11)$$

$$P(A) = \frac{|A|}{|\Omega|} \quad (1.12)$$

Nur das unmögliche Ereignis hat Wahrscheinlichkeit 0.

**Definition 3.4: Bernoulli Experiment**

Genau zwei Versuchsausgänge mit Wahrscheinlichkeiten  $p$  und  $1 - p$ .

**Definition 3.5: Bedingte Wahrscheinlichkeit**

Die bedingte Wahrscheinlichkeit eines Ereignisses  $A$  unter der Bedingung  $B$  ist

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (1.13)$$

Man liest dies auch als “Wahrscheinlichkeit von  $A$  bedingt  $B$ ”.

**Definition 3.6: Unabhängigkeit**

Die Ereignisse  $A$  und  $B$  heißen unabhängig, wenn gilt:

$$P(A \cap B) = P(A) \cdot P(B) \quad (1.14)$$

**Definition 3.7: Satz der totalen Wahrscheinlichkeit**

Ist  $B_i$  eine Folge paarweise disjunkter Mengen mit  $\bigcup_{i=0}^n B_i = \Omega$ , dann gilt für jedes Ereignis  $A$

$$P(A) = \sum_{i=0}^n P(A|B_i) \cdot P(B_i) \quad (1.15)$$

**Definition 3.8: Satz von Bayes**

Für zwei beliebige Ereignisse mit  $A$  und  $B$  mit nicht verschwindender Wahrscheinlichkeit  $P(B) \neq 0$  gilt

$$\begin{aligned} P(A|B) \cdot P(B) &= P(A \cap B) = P(B|A) \cdot P(A), \\ P(A|B) &= \frac{P(B|A) \cdot P(A)}{P(B)} \end{aligned} \quad (1.16)$$

## 4 Zufallsvariable

In der **Stochastik** ist eine **Zufallsvariable** (auch **zufällige Variable**, **zufällige Größe**, **zufällige Veränderliche**, **zufälliges Element**, **Zufallselement**, **Zufallsveränderliche**) eine Größe, deren Wert vom Zufall abhängig ist. Formal ist eine Zufallsvariable eine **Funktion**, die jedem **möglichen Ergebnis** eines **Zufallsexperiments** eine Größe zuordnet. Ist diese Größe eine **reelle Zahl**, so spricht man von einer **reellen Zufallsvariablen** oder **Zufallsgröße**. Beispiele für reelle Zufallsvariablen sind die Augensumme von zwei geworfenen Würfeln und die Gewinnhöhe in einem **Glücksspiel**. Zufallsvariablen können aber auch komplexere mathematische Objekte sein, wie **Zufallsfelder**, **Zufallsbewegungen**, **Zufallspermutationen** oder **Zufallsgraphen**. Über verschiedene Zuordnungsvorschriften können einem Zufallsexperiment auch verschiedene Zufallsvariablen zugeordnet werden.

**Definition 4.1: Zufallsvariable**

Eine Zufallsvariable ist eine Funktion  $Q \rightarrow \mathbb{R}$ . Eine diskrete Zufallsvariable nimmt nur diskrete Werte in  $\mathbb{R}$  an. Bei einer stetigen Zufallsvariable sind beliebige Werte  $X(\omega) \in (R)$  möglich. Mit der zusätzlichen Eigenschaft, dass die Mengen  $\{\omega | X(\omega) < x\}$  Ereignisse sind.

Mit einer Zufallsvariablen  $X$  kann man wieder neue Ereignisse definieren.

$$\begin{aligned} \{X = a\} &= \{\omega \in \Omega | X(\omega) = a\} \\ \{a < X \leq b\} &= \{\omega \in \Omega | a < X(\omega) \leq b\} \end{aligned} \tag{1.17}$$

Bei einer stetigen Zufallsvariable sind die Ereignisse der Form  $\{X = a\}$  nur von beschränktem Nutzen, da nur ganz wenige Versuche dazu führen werden, dass das Ereignis eintritt.



## 5 Erwartungswert

Der **Erwartungswert** (selten und doppeldeutig *Mittelwert*) ist ein Grundbegriff der *Stochastik*. Der Erwartungswert ist eine Kennzahl einer *Zufallsvariablen*. Bei einer engeren Definition ist der Erwartungswert einer Zufallsvariablen eine reelle Zahl und damit endlich; bei einer weiteren Definition sind für den Erwartungswert einer Zufallsvariablen auch die Werte  $\pm\infty$  zugelassen. Es gibt Zufallsvariablen, für die kein Erwartungswert definiert ist.

### Definition 5.1: Erwartungswert

Sei  $X$  eine Funktion auf  $\Omega$ , und lasse sich  $\Omega$  in endlich viele Ereignisse  $A_i$  zerlegen, auf denen  $X(\omega)$  konstant ist, dann ist der Erwartungswert von  $X$

$$E(X) = \sum_{i=0}^n P(A_i) \cdot X(A_i) \quad (1.18)$$

**Der Erwartungswert einer reellen Zufallsvariable ist diejenige Zahl, für die die Varianz minimal wird.**

Sind  $X$  und  $Y$  Zufallsvariable mit Werten in  $\mathbb{R}$ , und  $\lambda$ , dann gilt

1.  $E(X + Y) = E(X) + E(Y)$
2.  $E(\lambda X) = \lambda E(X)$
3. Sei  $X_A$  die charakteristische Funktion des Ereignisses  $A \in \mathbb{A}$ , welche definiert ist durch

$$X_A : \Omega \rightarrow \mathbb{R} : \omega \mapsto \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

dann gilt  $E(X_A) = P(A)$ .

### Definition 5.2: Unabhängigkeit

Zwei Zufallsvariable  $X$  und  $Y$  heissen unabhängig, wenn die Ereignisse  $\{\omega | X(\omega) \leq x\}$  und  $\{\omega | Y(\omega) \leq y\}$  für alle  $x, y \in \mathbb{R}$  unabhängig sind, also

$$P((X \leq x) \wedge (Y \leq y)) = P(X \leq x)P(Y \leq y) \quad \forall x, y \in \mathbb{R} \quad (1.19)$$

### Definition 5.3: Erwartungswert Produktregel

Sind  $X$  und  $Y$  unabhängige Zufallsvariablen, dann gilt

$$E(XY) = E(X) \cdot E(Y)$$

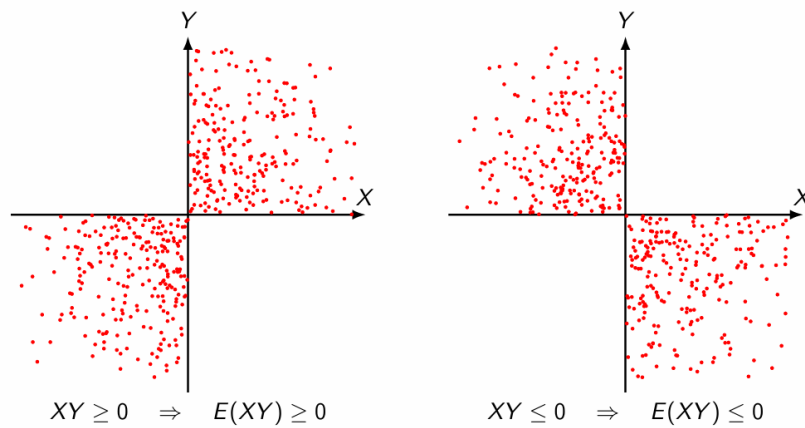
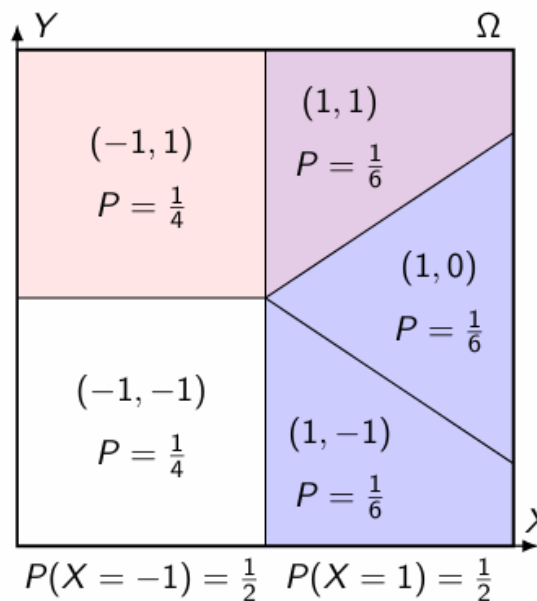


Figure 1.1: Unabhängigkeit

unkorreliert  $\nRightarrow$  unabhängig



unkorreliert:

$$E(X) = (-1) \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 0$$

$$E(Y) = (-1) \cdot \frac{5}{12} + 1 \cdot \frac{5}{12} = 0$$

$$E(XY) = (-1) \cdot \frac{5}{12} + 1 \cdot \frac{5}{12} = 0 \\ = E(X) \cdot E(Y)$$

nicht unabhängig:

$$P(X = 1) = \frac{1}{2}$$

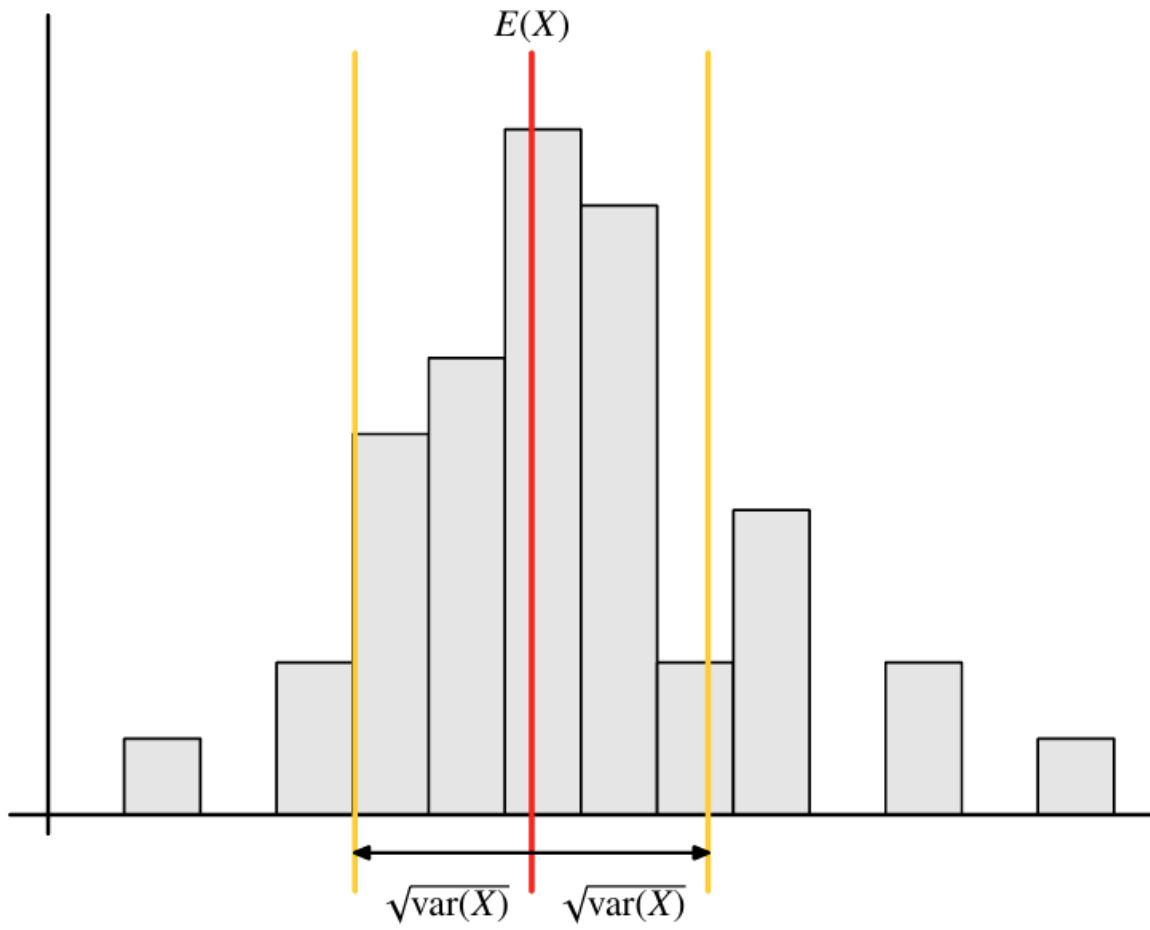
$$P(Y = 1) = \frac{5}{12}$$

$$P(X = 1) \cdot P(Y = 1) = \frac{5}{24}$$

$$P(X = 1 \wedge Y = 1) = \frac{1}{6} = \frac{4}{24}$$

$$P(X = 1 \wedge Y = 1) \neq P(X = 1) \cdot P(Y = 1)$$

Figure 1.2: Unkorreliert - Unabhängig



**Figure 1.3:** Darstellung von Erwartungswert und Varianz

## 6 Varianz

Die **Varianz** (lateinisch *variantia* „Verschiedenheit“ bzw. *variare* „(ver)ändern, verschieden sein“) ist ein Maß für die Streuung einer Wahrscheinlichkeitsdichte um ihren Schwerpunkt. Mathematisch wird sie definiert als die mittlere quadratische Abweichung einer reellen Zufallsvariablen von ihrem Erwartungswert. Sie ist das zentrale Moment zweiter Ordnung einer Zufallsvariablen.

### Definition 6.1: Varianz

Sei  $X : \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable, dann heisst die durch  $\text{var}(X)^2 = E((X - E(X))^2)$  definierte Grösse  $\text{var}(X)$  die Varianz von  $X$ . Oder auch die mittlere quadratische Abweichung vom Erwartungswert. Es ist insbesondere

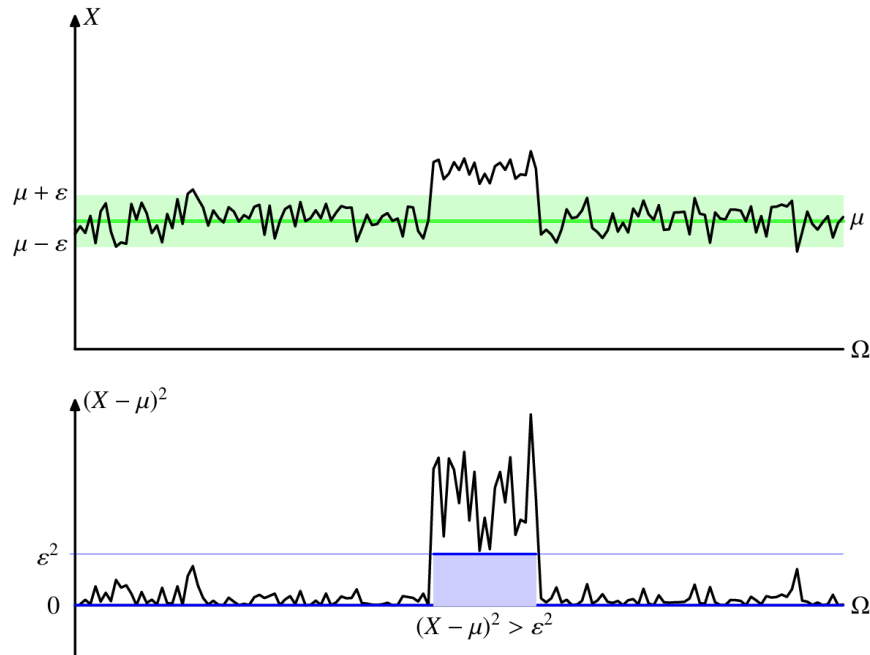
$$\text{var}(X) = E((X - E(X))^2) = E(X^2) - E(X)^2 \quad (1.20)$$

Man beachte, dass  $\text{var}(X)$  die Masseinheit des Quadrates von  $X$  hat. Wenn also  $X$  als Masseinheit eine Länge hat, dann hat  $\text{var}(X)$  die Masseinheit einer Fläche. Insbesondere kann man  $\text{var}(X)$  nicht in der gleichen Zeichnung visualisieren wie  $E(X)$ . Aber die Grösse  $\sqrt{\text{var}(X)}$  hat die gleiche Masseinheit, sie drückt die „Streubreite“ der Werte von  $X$ .

### Definition 6.2: Rechenregeln

Seien  $X$  und  $Y$  unabhängige Zufallsvariable, dann haben Summe und Produkt folgende Varianz

$$\begin{aligned} \text{var}(\lambda X) &= \lambda^2 \text{var}(X) \\ \text{var}(X + Y) &= \text{var}(X) + \text{var}(Y) \\ \text{var}(XY) &= \text{var}(X)\text{var}(Y) + \text{var}(Y)E(X)^2 + \text{var}(X)E(Y)^2 \end{aligned} \quad (1.21)$$



**Figure 1.4:** Herleitung der Tschebyscheff-Ungleichung, charakteristische Funktion des Ereignisses  $|X - \mu| > \epsilon$  ist blau eingezeichnet.

## 7 Bewertung vom Mittelwert / Erwartungswert

### Definition 7.1: Tschebyscheff Ungleichung

$X$  eine Zufallsvariable mit Erwartungswert  $\mu = E(X)$ , dann lässt sich die Wahrscheinlichkeit, dass  $X$  um mehr als  $\epsilon$  vom Erwartungswert abweicht, wie folgt abschätzen:

$$P(|X - \mu| > \epsilon) \leq \frac{\text{var}(x)}{\epsilon^2} \quad (1.22)$$

Lässt man  $\epsilon$  wachsen findet man, dass Abweichungen deutlich grösser als  $\sqrt{\text{var}(X)}$  unwahrscheinlich sind. Leider sind die Vorhersagen wie auch bei der Tschebyscheff-Ungleichung nur beschränkt nützlich. Wenn die Wahrscheinlichkeit, dass der Mittelwert um mehr als  $\sigma$  vom Erwartungswert abweicht, kleiner als 1% sein soll, sind dafür nach dem Gesetz der grossen Zahlen mindestens 100 Summanden nötig.

### Definition 7.2: Bernoulli Gesetz der grossen Zahlen

Die Wahrscheinlichkeit, dass der Mittelwert von  $n$  unabhängigen Zufallsvariablen mit Erwartungswert  $\mu$  und Varianz  $\sigma^2$  mehr als  $\varepsilon$  von  $\mu$  abweicht, ist :

$$P(|M_n - \mu| > \varepsilon) \leq \frac{\text{var}(X)}{n\varepsilon^2} \quad (1.23)$$

Insbesondere gilt

$$\lim_{n \rightarrow \infty} P(|M_n - \mu| > \varepsilon) = 0 \quad (1.24)$$

### Definition 7.3: Genauigkeit der empirischen Häufigkeit

Wird ein Experiment  $n$  mal durchgeführt, und tritt dabei das Ereignis  $A$  mit der relativen Häufigkeit  $h$  ein, dann ist die Wahrscheinlichkeit, dass  $h$  um mehr als  $\varepsilon$  von  $P(A)$  abweicht

$$P(|h - P(A)| > \varepsilon) \leq \frac{P(A)(1 - P(A))}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2} \quad (1.25)$$

## 8 Lineare Regression

## 9 Verteil- und Dichtefunktionen

### Definition 9.1: Verteilfunktion

Die Verteilungsfunktion beschreibt die Wahrscheinlichkeiten der Werte einer Zufallsvariable:

$$F(x) = P(X \leq x) \quad (1.26)$$

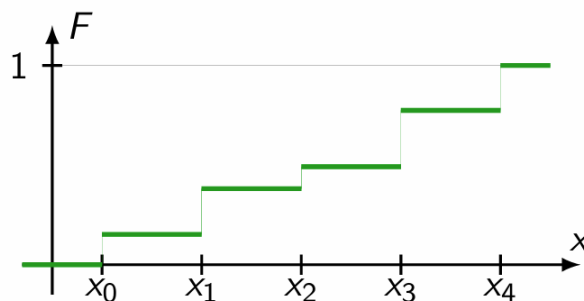


Figure 1.5: Diskrete Verteilfunktion

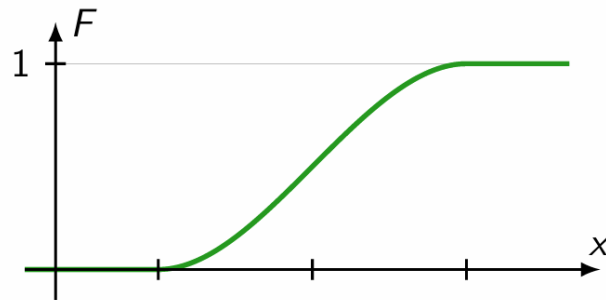


Figure 1.6: Stetige Verteilfunktion

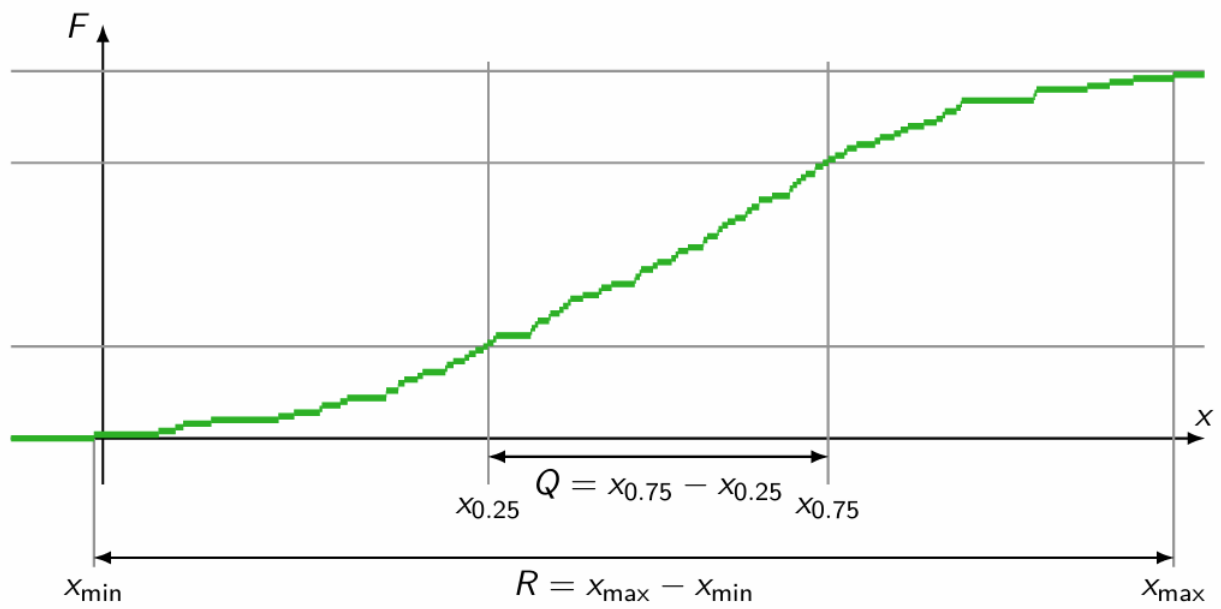


Figure 1.7: Spannweite R und Quartilabstand Q

## 9.1 Wahrscheinlichkeitsdichte

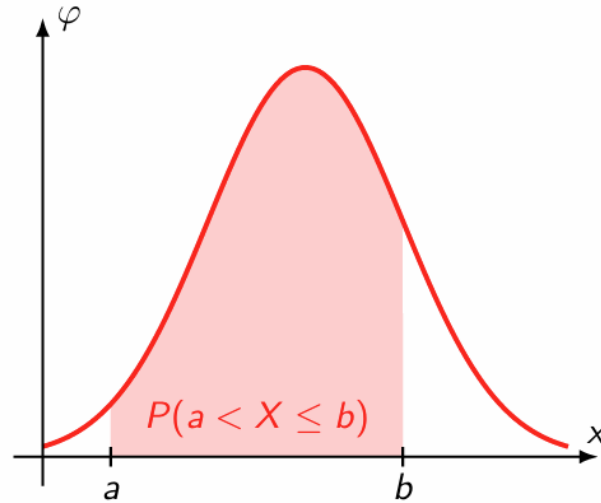


Figure 1.8: Wahrscheinlichkeitsdichte

### Definition 9.2: Wahrscheinlichkeitsdichte Funktion

Die Ableitung der Verteilungsfunktion heisst Wahrscheinlichkeitsdichte:

$$\varphi(x) = \frac{d}{dx}F(x) = F'(X) \quad (1.27)$$

Wenn  $F(x)$  (Verteilfunktion) die Stammfunktion ist bedeutet dies:

$$P(a < X \leq b) = \int_a^b \varphi(x)dx \quad (1.28)$$

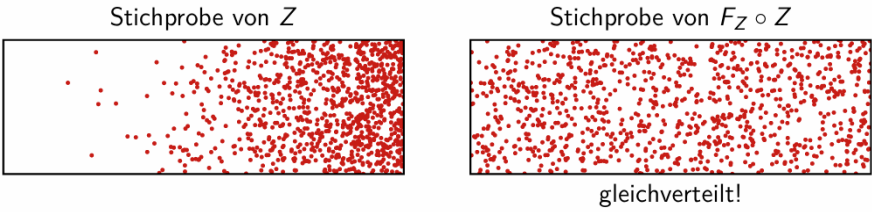
Mit  $\varphi(x)$  kann man nun auch  $E(X)$  berechnen:

$$\begin{aligned} E(X) &= \sum x \cdot p(x) \\ &= \int_{-\infty}^{\infty} x \cdot \varphi(x)dx \end{aligned} \quad (1.29)$$

## 9.2 Gleichverteilung kreieren

Durch erneute Anwendung der Verteilfunktion auf  $X$  wird die Zufallsvariable gleichverteilt.





**Figure 1.9:** Enter Caption

## 10 Gleichverteilung

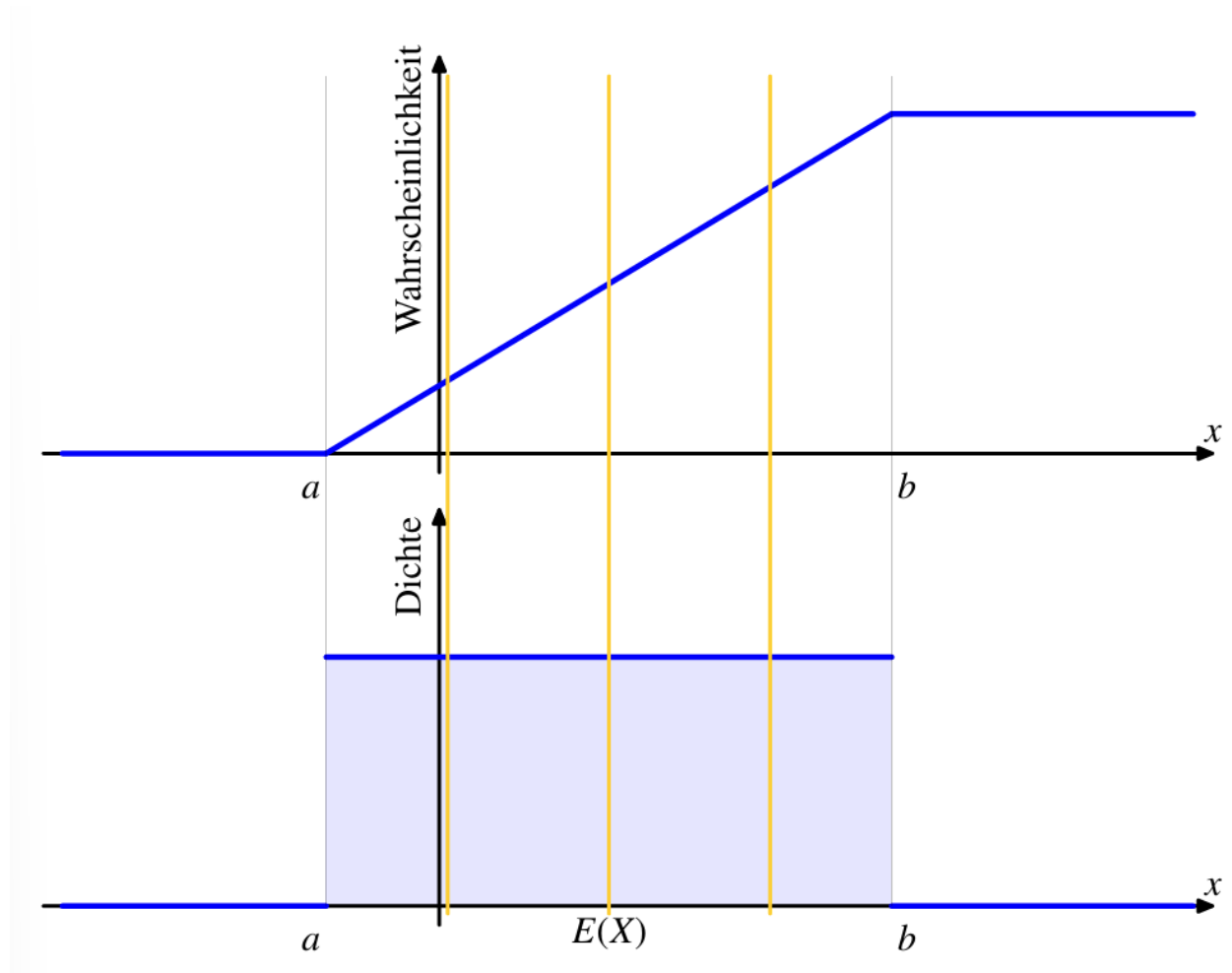


Figure 1.10: Gleichverteilung

### Definition 10.1: Dichtefunktion

$$\varphi(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{sonst} \end{cases} \quad (1.30)$$

### Definition 10.2: Verteilfunktion

$$F(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a < x < b \\ 1 & x \geq b \end{cases} \quad (1.31)$$

**Definition 10.3: Erwartungswert**

$$E(X) = \frac{a+b}{2} \quad (1.32)$$

**Definition 10.4: Varianz**

$$\sigma^2 = \frac{(b-a)^2}{12} \quad (1.33)$$

**Definition 10.5: Median**

$$x_{\frac{1}{2}} = \frac{a+b}{2} \quad (1.34)$$

**Definition 10.6: Wahrscheinlichkeit für Abweichung**

$$P(|X - E(X)| > \epsilon) = 1 \frac{2\epsilon}{b-a} \text{ für } \epsilon < \frac{b-a}{2} \quad (1.35)$$