Wahrscheinlichkeitstheorie und Mathematische Statistik

Prof. Dr. rer. nat. Martin Redmann
Sommersemester 2025

Inhaltsverzeichnis

Diese Mitschrift basiert auf der gleichnamigen Vorlesung *Wahrscheinlichkeitstheorie und Mathematische Statistik*, gehalten im Sommersemester 2025 an der Universität Rostock.

0 Einleitung

Motivation

In der Praxis sind viele Größen, mit denen wir arbeiten, nicht exakt bekannt. Dazu zählen etwa die Wahrscheinlichkeit $\mathbb{P}(A)$ eines Ereignisses A, die Verteilungsfunktion F_X einer Zufallsgröße X, ihr Erwartungswert $\mathbb{E}[X]$ oder die Varianz $D^2[X]$. (Letztere misst die durchschnittliche Abweichung von X gegenüber ihrem Erwartungswert und dient damit als Maß für das Risiko)

Wenn wir diese Kenngrößen nicht wissen, bleibt uns nichts anderes übrig, als sie anhand von Beobachtungen zu schätzen. Es stellt sich also die grundlegende Frage: Was genau sind Beobachtungen im mathematischen Sinn, und wie lassen sich daraus fundierte Erkenntnisse ableiten?

Beispiel 0.1 (Versicherung und Produktion).

- $X \mathrel{\widehat{=}}$ Schadenshöhe eines Autos in der Kfz-Versicherung
- $X \mathrel{\widehat{=}}$ Anzahl fehlerhafter Teile pro Tag in einer Produktionsfirma

In beiden Fällen kann das jeweilige Unternehmen konkrete Daten beobachten oder messen: Wie hoch sind die Schäden? Wie viele fehlerhafte Teile wurden produziert?

Diese Beobachtungen erlauben Rückschlüsse auf die Verteilung F, den Erwartungswert $\mathbb{E}[X]$ und die Varianz $\mathbb{D}^2[X]$ von X, Größen welche bei Planung und Vorhersagen essenziell sind.

Was sind Beobachtungen im mathematischen Sinne?

Formal gesprochen handelt es sich bei Beobachtungen um Zufallsgrößen $X_1, \ldots, X_n$, die unabhängig und identisch verteilt sind mit gemeinsamer Verteilungsfunktion F. Diese heißen dann *unabhängige Beobachtungen* einer Zufallsgröße X bzw. eines zugrunde liegenden Zufallsvorgangs.

Die zugehörigen Realisierungen $x_1, \ldots, x_n \in \mathbb{R}$, also konkrete numerische Ergebnisse wie $X_i(\omega) = x_i$ für ein festes, aber unbekanntes ω, nennen wir *konkrete unabhängige* Beobachtungen von X.

Beispiel 0.2 (Schadensbeobachtung).
Eine Versicherung beobachtet n Schäden aus n Unfällen. Die Zufallsgrößen $X_1, \ldots, X_n$ sind a priori ungewiss, es ist also im Voraus unbekannt wie hoch ein Schaden sein wird. Sobald alle Unfälle eingetreten sind, erhalten wir dann konkrete Werte $x_1, \ldots, x_n$.

Die x_i können als Realisierungen der Schadenshöhen von den Zufallsgrößen X_i interpretiert werden.

Wann sind Beobachtungen unabhängig?

Unabhängigkeit ist dann gegeben, wenn zum Beispiel eine der folgenden Bedingungen erfüllt sind:

- Stichprobenziehung mit Zurücklegen: Der Auswahlvorgang hat keinen Einfluss auf die Grundgesamtheit.
- Jedes Element der Grundgesamtheit hat die gleiche Wahrscheinlichkeit, in die Stichprobe aufgenommen zu werden (z. B. bei einer telefonischen Umfrage).
- Es liegt kein Clustering vor, z. B. nicht nur Fahrgäste einer einzelnen Straßenbahn kontrollieren, da die i-te Ziehung sonst Einfluss auf die $(i + 1)$-te Ziehung haben könnte.

Weitere praktische Beispiele

Beispiel 0.3 (Glücksspiel).
Ein Spieler wirft eine möglicherweise verzerrte Münze. Abhängig vom Ergebnis erhält er eine Auszahlung X:

- „Kopf“ $\to$ Gewinn von 2 €
- „Zahl“ $\to$ Verlust von 1 €

Um zu entscheiden, ob sich die Teilnahme lohnt, müssen wir die Wahrscheinlichkeit p für „Kopf“, sowie den Erwartungswert des Spiels $\mathbb{E}[X]$ und die zugehörige Varianz $D^2[X]$ kennen.

Beispiel 0.4 (Schraubenproduktion).
In einem Betrieb werden Schrauben mit Soll-Länge 5 cm produziert. Zur Qualitätskontrolle entnimmt man n Schrauben aus dem laufenden Prozess. Die i-te Messung wird durch die Zufallsgröße X_i beschrieben.

Ziel ist es nun, auf Grundlage dieser Messungen Aussagen über die Verteilung, den Erwartungswert oder die Varianz der Länge X zu treffen.

Verknüpfung zur Wahrscheinlichkeitstheorie

Die mathematische Statistik baut in vielen Aspekten direkt auf Ergebnissen der Wahrscheinlichkeitstheorie auf. Zu den wichtigsten theoretischen Sätzen gehören:

- **Gesetz der großen Zahlen**: garantiert Konvergenz des Stichprobenmittels,
- **Zentrale Grenzwertsatz**: Normalverteilung als Grenzverteilung.

Darüber hinaus benötigen wir weitere wichtige Begriffe der Wahrscheinlichkeitsrechnung:

- Ereignisse und Wahrscheinlichkeiten
- Zufallsgrößen und deren Verteilungen
- Konvergenzbegriffe für Zufallsgrößen
- charakteristische Funktionen
- Erwartungswert und Varianz
- stochastische Unabhängigkeit

Wir beginnen deshalb mit einer gezielten Wiederholung bzw. Erweiterung dieser Konzepte.

Zufallsgrößen als messbare Abbildungen

Eine Zufallsgröße X ist eine Abbildung

$$X : \Omega \to \mathbb{R}$$

auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$, wobei X messbar sein muss. Über diese Abbildung lässt sich eine Verteilung auf $\mathbb{R}$ induzieren.

Verteilung einer Zufallsgröße

Die Verteilung μ_X einer Zufallsgröße X ist das durch X induzierte Maß auf $\mathbb{R}$ definiert:

$$\mu_X(B) := \mathbb{P}\big(\{\omega \in \Omega \mid X(\omega) \in B\}\big) \quad \text{für } B \in \mathcal{B}(\mathbb{R}).$$

Dies ist auch als *Bildmaß* von $\mathbb{P}$ unter X bekannt.

Verteilungsfunktion

Die Verteilungsfunktion $F_X \colon \mathbb{R} \to [0,1]$ ist gegeben durch

$$F_X(x) := \mathbb{P}(X \leq x) = \mu_X\big((-\infty, x]\big).$$

Erwartungswert als Integral

Der Erwartungswert einer Zufallsgröße X wird als Lebesgue-Integral definiert:

$$\mathbb{E}[X] := \int_\Omega X(\omega)\ \mathrm{d}\mathbb{P}(\omega) = \int_\mathbb{R} x\, \mu_X(\mathrm{d}x).$$

Dies kann auch als *Stieltjes-Integral* über F_X geschrieben werden kann:

$$\mathbb{E}[X] = \int_\mathbb{R} x\ \mathrm{d}F_X(x).$$

Dieser Zugang erlaubt es, Erwartungswerte und andere Kennzahlen auch für diskrete, stetige oder allgemeinere Verteilungen einheitlich zu behandeln:

Definition 0.5 (Stieltjes-Integral).
Sei $F \colon [a,b] \to \mathbb{R}$ eine monotone, rechtsseitig stetige Funktion (z. B. eine Verteilungsfunktion) und $g \colon [a,b] \to \mathbb{R}$ eine beschränkte, stetige Funktion. Dann definiert man das *Stieltjes-Integral* von g bezüglich F durch

$$\int_a^b g(x)\ \mathrm{d}F(x)$$

als den Grenzwert geeigneter Riemann-Stieltjes-Summen:
Sei $\big((x_i^n)_{i=0}^{N_n}\big)_{n\in\mathbb{N}}$ so gewählt, dass

$$a = x_0^n < x_1^n < \cdots < x_{X_N}^n = b \quad \text{und} \quad \max_{i=0,\ldots,N_n-1} |x_{i+1}^n - x_i^n| \xrightarrow{n\to\infty} 0$$

dann definieren wir im Falle der Existenz

$$\int_a^b g(x)\ \mathrm{d}F(x) := \lim_{n\to\infty} \sum_{i=0}^{N_n-1} g(x_i^n) \cdot \big(F(x_{i+1}^n) - F(x_i)^n)\big)$$

und das *uneigentliche Stieltjes-Integral* entsprechend durch

$$\int_\mathbb{R} g(x)\ \mathrm{d}F(x) := \lim_{\substack{a\to-\infty \\ b\to\infty}} \int_a^b \int_a^b g(x)\ \mathrm{d}F(x)$$

Dieses Integral verallgemeinert das gewöhnliche Riemann-Integral, indem nicht über ein Maß bezüglich der Länge (wie bei dx), sondern bezüglich der Änderung einer Funktion F integriert wird.

1 Ergänzugen/Wiederholung Wahrscheinlichkeitstheorie

1.1 Gesetz der großen Zahlen

Lemma 1.1 (Ungleichung von Markov).
Es sei X eine integrierbare Zufallsgröße, d. h. $\mathbb{E}|X| < \infty$, dann gilt für alle $\varepsilon > 0$:

$$\mathbb{P}\big(\{\omega : |X(\omega)| \geq \varepsilon\}\big) \leq \frac{1}{\varepsilon}\mathbb{E}|X|$$

Beweis.
Verwenden wir die Definition des Erwartungswertes mittels Stieltjes-Integral so ergibt sich die Ungleichung durch Anwendung der Linearität und Montonie des Integrals:

$$\begin{aligned}
\mathbb{E}|X| &= \int_{\mathbb{R}} |x| \; \mathrm{d}F_X(x) \\
&= \underbrace{\int_{\{x:|x|<\varepsilon\}} |x| \; \mathrm{d}F_X(x)}_{\geq 0} + \int_{\{x:|x|\geq\varepsilon\}} |x| \; \mathrm{d}F_X(x) \\
&\geq \int_{\{x:|x|\geq\varepsilon\}} |x| \; \mathrm{d}F_X(x) \\
&\geq \int_{\{x:|x|\geq\varepsilon\}} \varepsilon \; \mathrm{d}F_X(x) \\
&= \varepsilon \cdot \int_{\{x:|x|\geq\varepsilon\}} \mathrm{d}\mu_X(x) \\
&= \varepsilon \cdot \mu_X\big(\{x : |x| \geq \varepsilon\}\big) \\
&= \varepsilon \cdot \mathbb{P}\Big(\big\{\omega : X(\omega) \in \{x : |x| \geq \varepsilon\}\big\}\Big) \\
&= \varepsilon \cdot \mathbb{P}\big(\{\omega : |X(\omega)| \geq \varepsilon\}\big)
\end{aligned}$$

□

Bemerkung 1.2. Aus Lemma 1.1 folgt

$$\begin{aligned}
\mathbb{P}(|X - \mathbb{E}[X]| \geq \varepsilon) &= \mathbb{P}\big(\underbrace{|X - \mathbb{E}[X]|^2}_{\text{neue Zufallsgröße}} \geq \varepsilon^2\big) \\
&\leq \frac{1}{\varepsilon^2}\mathbb{E}|X - \mathbb{E}[X]|^2 \\
&= \frac{1}{\varepsilon^2} D^2[X]
\end{aligned}$$

Analog lässt sich auch die *Chebyshev Ungleichung* herleiten, für ein beliebiges $p > 0$ gilt

$$\mathbb{P}(|X| \geq \varepsilon) \leq \frac{E|X|^p}{\varepsilon^p}$$

Definition 1.3 (Stochastische Konvergenz).
Eine Folge (X_n) von Zufallsgrößen *konvergiert stochastisch* (oder in Wahrscheinlichkeit) gegen eine Zufallsgröße X, falls

$$\mathbb{P}\big(\{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\}\big) \xrightarrow{n\to\infty} 0 \quad \text{für alle } \varepsilon \geq 0$$

oder dazu äquivalent, falls

$$\mathbb{P}\big(\{\omega : |X_n(\omega) - X(\omega)| \leq \varepsilon\}\big) \xrightarrow{n\to\infty} 1 \quad \text{für alle } \varepsilon > 0$$

Wir schreiben im folgenden auch $X_n \xrightarrow{\mathbb{P}} X$ mit $n \to \infty$.

Definition 1.4 (Definition des schwachen Gesetzes der großen Zahlen).
Eine Folge integrierbarer Zufallsgrößen (X_n), d. h. $\mathbb{E}|X_n| < \infty \ \forall n$ genügt genau dann dem schwachen Gesetz der großen Zahlen, wenn

$$\frac{1}{n}\sum_{k=1}^{n}(X_k - \mathbb{E}[X_k]) \xrightarrow{\mathbb{P}} 0, \quad \text{für } n \to \infty.$$

Ein Spezialfall bilden hier Beobachtungen: Seien $X_1, \ldots, X_n$ unabhängige Beobachtungen einer Zufallsgröße X, so gilt $\mathbb{E}[X_k] = \mathbb{E}[X]$ und der Term vereinfacht sich durch

$$\frac{1}{n}\sum_{k=1}^{n}(X_k - \mathbb{E}[X_k]) = \frac{1}{n}\sum_{k=1}^{n} X_k - \mathbb{E}[X]$$

Satz 1.5 (Schwaches Gesetz der großen Zahlen, $L^1(\Omega, \mathcal{F}, \mathbb{P})$ Version).
Es sei (X_n) eine Folge von Zufallsgrößen X_n, welche unabhängig und gleichverteilt wie eine integrierbare Zufallsgröße X sind. Dann genügt (X_n) dem schwachen Gesetz der großen Zahlen (gemäß Definition 1.4), d. h. konkret, dass

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{\mathbb{P}} \mathbb{E}[X], \quad \text{für } n \to \infty.$$

Beweis.

1.2 Charakteristische Funktionen

1.3 Zentraler Grenzwertsatz und Konvergenz

2 Stichproben und Stichprobenfunktion

3 Parameterschätzer/Punktschätzer

4 Signifikanztests

4.1 Konfidenzschätzungen

5 Lineare Regression