# Wahrscheinlichkeitstheorie für Informatikstudien (Kapitel #9)

Siegfried Hörmann

TU Graz

WS 2020/21

# Kursaufbau

1. Wahrscheinlichkeitsräume.
2. Laplace-Experimente.
3. Wichtige diskrete Verteilungen.
4. Zufallsvariablen.
5. Zufallsvektoren.
6. Momente.
7. Bedingte Wahrscheinlichkeiten.
8. Unabhängigkeit.
9. Konvergenz von Zufallsvariablen.
10. Der zentrale Grenzwertsatz.

# Konvergenz von Zufallsvariablen

9.1. Konvergenz im quadratischen Mittel.

9.2. Gesetz der großen Zahlen (LLN).

9.3. Konvergenz in Wahrscheinlichkeit.

9.4. Anwendungen.

# Konvergenz im quadratischen Mittel

Wir betrachten eine Folge von ZVen $(X_n \colon n \geq 1)$. In diesem Kapitel wollen wir uns überlegen, wie wir Konvergenz von $X_n$ zu einem Grenzwert $X$ (der womöglich auch zufällig ist) definieren können.

## Beispiel

Sei $X_n$ der Mittelwert von $n$ (unabhängigen) Würfelresultaten $\xi_k$:

$$X_n := \frac{1}{n}(\xi_1 + \cdots + \xi_n).$$

Konvergiert $X_n$? Und wie sollten wir Konvergenz überhaupt definieren?

## Konvergenz im quadratischen Mittel

Das Problem ist nicht ganz einfach. Im letzten Beispiel ist $\Omega = \{1, \ldots, 6\}^\infty$ und $\xi_k(\omega) = \omega_k$, wobei $\omega = (\omega_1, \omega_2, \ldots)$.

Offensichtlich ist der Grenzwert von $X_n(\omega)$—sofern dieser überhaupt existiert—nicht für alle $\omega$ gleich:

$$\lim X_n(\omega) = \begin{cases} 1 & \omega = (1, 1, 1, \ldots); \\ 1.5 & \omega = (1, 2, 1, 2, \ldots); \\ 6 & \omega = (6, 6, 6, \ldots); \\ \text{existiert nicht} & \omega = (1, 6, \ldots, 6, \underbrace{1, 1, \ldots, 1, 1}_{2^{k^2}}, \ldots); \\ \ldots & \end{cases}$$

## Konvergenz im quadratischen Mittel

Anstelle sich die Konvergenz in jedem $\omega$ einzeln anzusehen, schauen wir zuerst die erwartete Abweichung vom vermeintlichen Grenzwert an. So wie bei der Varianz, sind quadratische Abweichungen einfacher.

Sei $(X_n \colon n \geq 1)$ eine Folge von ZVen in $L^2$ (also $E(X_n^2) < \infty$). Wir sagen $X_n$ konvergiert zu $X$ im quadratischen Mittel falls

$$E(|X_n - X|^2) \to 0 \quad (n \to \infty).$$

Kurz: $X_n \overset{L^2}{\to} X$.

# Konvergenz im quadratischen Mittel

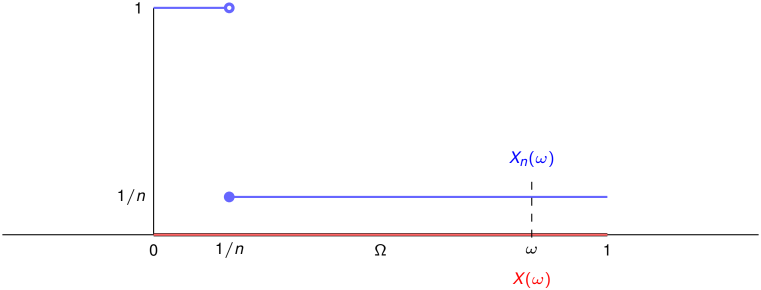
## Beispiel (ZVen auf Einheitsintervall)

Am besten veranschaulicht man das Problem mit $\Omega = [0, 1]$. Dann sind $X_n \colon [0, 1] \to \mathbb{R}$ und können graphisch skizziert werden.

Nehmen wir weiters $P((a, b]) = b - a$ für $0 \leq a < b \leq 1$.

Seien

$$X_n(\omega) = I\{\omega < 1/n\} + \frac{1}{n} I\{\omega \geq 1/n\} \quad \text{und} \quad X(\omega) \equiv 0.$$



## Konvergenz im quadratischen Mittel

Wir sehen:

$$X_n(\omega) \to \begin{cases} 0 & \text{wenn } \omega \in (0,1]; \\ 1 & \text{wenn } \omega = 0. \end{cases}$$

Also $X_n(\omega)$ konvergiert gegen $X(\omega)$ für alle $\omega \in \Omega$, außer $\omega = 0$.

Jetzt sehen wir uns $E((X_n - X)^2)$ an. Es gilt in unserem Beispiel $X_n - X = X_n \in \{1/n, 1\}$. Also ist

$$\begin{aligned} E((X_n - X)^2) &= E(X_n^2) \\ &= \frac{1}{n^2} \times P(X_n = 1/n) + 1^2 \times P(X_n = 1) \\ &= \frac{1}{n^2} \times P(\{\omega \geq 1/n\}) + 1 \times P(\{\omega < 1/n\}) \\ &= \frac{1}{n^2}\left(1 - \frac{1}{n}\right) + \frac{1}{n} \to 0. \end{aligned}$$

# Konvergenz im quadratischen Mittel

### Beispiel (Gleicher Erwartungswert)

Nehmen wir den wichtigen Spezialfall, wo alle ZVen $X_n$ den gleichen Erwartungswert $\mu$ haben. Dann konvergiert $X_n$ definitionsgemäß im quadratischen Mittel gegen $\mu$, falls

$$E((X_n - \mu)^2) = \mathrm{Var}(X_n) \to 0.$$

Mit dieser einfachen Beobachtung erhalten wir fast unmittelbar den folgenden fundamentalen Satz.

# Gesetz der großen Zahlen (LLN)

## Satz ($L^2$ Version des Gesetzes der großen Zahlen)

*Sei* $(X_n\colon n \geq 1)$ *eine i.i.d. Folge in* $L^2$ *mit* $E(X_1) = \mu$*. Dann gilt:*

$$\overline{X}_n = \frac{1}{n}(X_1 + \cdots + X_n) \xrightarrow{L^2} \mu.$$

*Genauer gesagt, haben wir*

$$E(|\bar{X}_n - \mu|^2) = \frac{1}{n}\mathrm{Var}(X_1).$$

# Gesetz der großen Zahlen (LLN)

Beweis. Da $E(\overline{X}_n) = \mu$ gilt

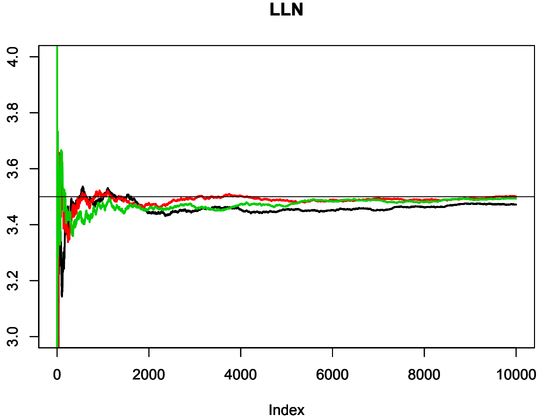
$$
\begin{aligned}
E(|\overline{X}_n - \mu|^2) &= \mathrm{Var}(\overline{X}_n) \\
&= \frac{1}{n^2}\mathrm{Var}(X_1 + \cdots + X_n) \\
&= \frac{1}{n}\mathrm{Var}(X_1).
\end{aligned}
$$

$\square$

Bemerkung: Beachte, dass wir in diesem Argument nur $\mathrm{Cov}(X_i, X_j) = 0$ $(i \neq j)$ benutzt haben. Damit bleibt das LLN auch für unkorrelierte ZV erhalten.

## Gesetz der großen Zahlen (LLN)

Illustration mit $10000\times$ würfeln. Die Graphen zeigen jeweils $\overline{X}_n$ für $n = 1, \ldots, 10000$. Wir zeigen drei Versuche dieses Experiments.



# Konvergenz in Wahrscheinlichkeit

Die folgende Art der Konvergenz wird üblicherweise in der Statistik verwendet.

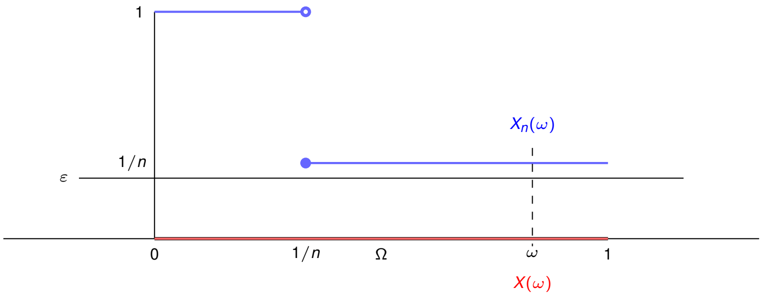
Sei $(X_n \colon n \geq 1)$ eine Folge von ZVen. Wir sagen, dass $X_n$ zu $X$ in Wahrscheinlichkeit konvergiert falls für alle $\varepsilon > 0$ gilt

$$P(|X_n - X| > \varepsilon) \to 0 \quad (n \to \infty).$$

Kurz: $X_n \xrightarrow{P} X$.

# Konvergenz in Wahrscheinlichkeit

Wir betrachten wieder das Beispiel von oben:



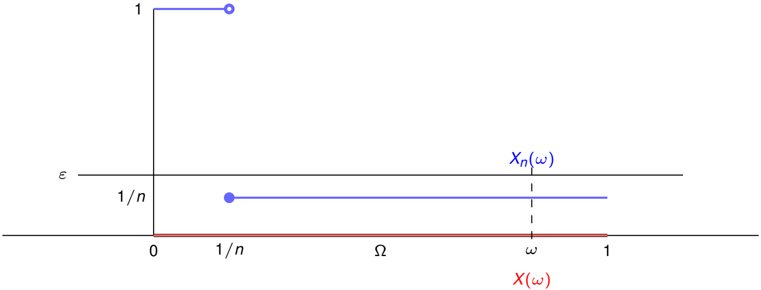

Das $n$ ist noch recht klein und es gilt

$$\{\omega\colon |X_n - X| > \varepsilon\} = [0, 1].$$

Somit ist $P(|X_n - X| > \varepsilon) = 1$.

# Konvergenz in Wahrscheinlichkeit

Wir machen $n$ größer:




Jetzt ist

$$\{\omega\colon |X_n - X| > \varepsilon\} = [0, 1/n).$$

Somit ist $P(|X_n - X| > \varepsilon) = 1/n$.

## Konvergenz in Wahrscheinlichkeit

Bevor wir fortfahren, untersuchen wir den Zusammenhang dieser beiden Konzepte.

### Proposition

*Wenn $X_n \stackrel{L^2}{\to} X$ so gilt auch $X_n \stackrel{P}{\to} X$. Die Gegenrichtung gilt i.A. nicht.*

Beweis. Wegen der Markov-Ungleichung haben wir

$$P(|X_n - X| > \varepsilon) \leq \frac{E(|X_n - X|^2)}{\varepsilon^2} \to 0 \quad (n \to \infty).$$

Insbesonders gilt:

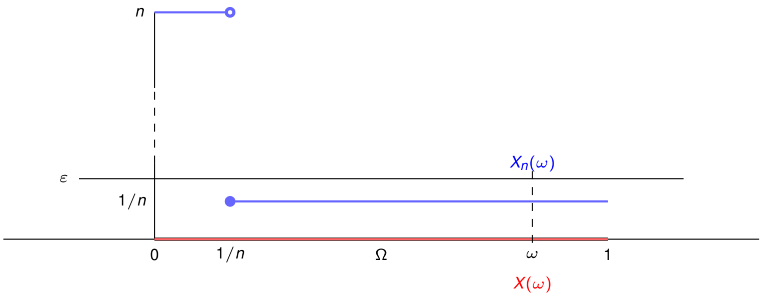
### Satz (Schwaches Gesetz der großen Zahlen)

*Sei $(X_n \colon n \geq 1)$ eine i.i.d. Folge in $L^2$ mit $E(X_1) = \mu$. Dann gilt:*

$$P(|\overline{X}_n - \mu| > \varepsilon) \to 0 \quad \forall \varepsilon > 0, n \to \infty.$$

# Konvergenz in Wahrscheinlichkeit

Um zu sehen, dass die Gegenrichtung i.A. falsch ist, betrachten wir wieder unser Beispiel, aber diesmal ist $X_n$ sehr groß auf $[0, 1/n)$:




Jetzt ist wieder $\{\omega\colon |X_n - X| > \varepsilon\} = [0, 1/n)$ und somit $P(|X_n - X| > \varepsilon) = 1/n$. Aber:

$$E(|X_n - X|^2) = n^2 \times \frac{1}{n} + \frac{1}{n^2} \times \left(1 - \frac{1}{n}\right) \to \infty.$$

### Beispiel (Monte-Carlo-Integration)

Sei $f : [0, 1] \to \mathbb{R}$ eine integrierbare Funktion. Wir sind daran interessiert, $\mathcal{I} = \int_0^1 f(x)dx$ zu berechnen. Oft ist das nicht so einfach. Für eine komplizierte Funktion $f$ kann man oft keine explizite Form finden.

Wenn wir allerdings i.i.d. ZV $U_i \sim \mathrm{Unif}([0, 1])$ generieren, dann gilt

$$E(f(U_i)) = \mathcal{I},$$

und mit dem LLN

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n} f(U_k) - \mathcal{I}\right| > \varepsilon\right) \leq \frac{\mathrm{Var}(f(U_1))}{n\varepsilon^2}.$$

### Beispiel (Der Fundamentalsatz der Statistik)

Angenommen wir erheben i.i.d. Daten $(X_n \colon n \geq 1)$ wobei die $X_i$ der Verteilungsfunktion $F(x) = P(X \leq x)$ folgen. In der Praxis kennen wir $F(x)$ nicht. Ist es möglich, $F(x)$ basierend auf unseren Beobachtungen zu bestimmen?

Wir definieren die empirische Verteilungsfunktion

$$F_n(x) := \frac{1}{n} \sum_{k=1}^{n} I\{X_k \leq x\}.$$

Die Variablen $Y_k := I\{X_k \leq x\}$ sind wieder i.i.d. und $Y_k \sim B_{1,F(x)}$. Daher folgt mit dem Gesetz der großen Zahlen

$$P(|F_n(x) - F(x)| > \varepsilon) \leq \frac{F(x)(1 - F(x))}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}.$$

### Beispiel (Casino)

Wir haben gesehen: egal wie man setzt, der erwartete Gewinn des Spielers beim Roulette ist $\frac{-\text{Einsatz}}{37}$, und somit der erwartete Gewinn des Casinos $\frac{\text{Einsatz}}{37}$.

Sagen wir in einem Casino wird im Monat $n$ Mal ($n$ ist sehr groß) Roulette gespielt und der Einsatz ist jeweils 5 Euro. Sei $X_i$ der Gewinn beim $i$-ten Spiel. Dann ist der durchschnittliche Gewinn des Casinons:

$$\frac{1}{n}(X_1 + \cdots + X_n) \approx \frac{5}{37}.$$

Und somit ist der Gesamtgewinn $X_1 + \cdots + X_n \approx n \times \frac{5}{37}$.