Einführung in die Stochastik

Konvergenz von Zufallsvariablen und Zentraler Grenzwertsatz

Prof. Dr. Nejjar

June 25, 2025

Verschiedene Konvergenzbegriffe

- In der Stochastik gibt es mehrere Konvergenzbegriffe, die unterschiedliche "Stärken" haben und für verschiedene Fragestellungen relevant sind:
    - Konvergenz in Wahrscheinlichkeit ($X_n \xrightarrow{i.W.} X$)
    - Konvergenz in Verteilung ($X_n \xrightarrow{i.V.} X$)
    - **Fast sichere Konvergenz** ($X_n \xrightarrow{f.s.} X$)
- Wir werden uns heute auf die fast sichere Konvergenz und die Konvergenz in Verteilung konzentrieren.

Intuition hinter "Fast Sicher"

- Stellen Sie sich vor, Sie wiederholen ein Experiment unendlich oft. Für jede Ausführung (jedes Ergebnis ω im Wahrscheinblichkeitsraum Ω) erhalten Sie eine Sequenz von reellen Zahlen $X_1(\omega), X_2(\omega), \ldots$.
- Fast sichere Konvergenz bedeutet, dass für **fast alle** dieser Sequenzen (d.h. für alle ω bis auf eine Menge von ω mit Wahrscheinlichkeit Null) die Zahlenfolge $X_n(\omega)$ im üblichen Sinne der reellen Analysis gegen $X(\omega)$ konvergiert.
- Man kann sich dies als "pfadweise" Konvergenz vorstellen.
- *Beispiel:* Münzwurf. Wenn Sie eine faire Münze unendlich oft werfen, konvergiert der Anteil der Köpfe fast sicher gegen 0.5. Nur in einer verschwindend kleinen Menge von Fällen (z.B. immer nur Kopf) würde es nicht konvergieren.

Formale Definition der Fast Sicheren Konvergenz

Definition (Fast Sichere Konvergenz)

Eine Folge von Zufallsvariablen $(X_n)_{n\in\mathbb{N}}$, definiert auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$, konvergiert **fast sicher** gegen eine Zufallsvariable X auf demselben Raum, bezeichnet als $X_n \xrightarrow{f.s.} X$, wenn

$$\mathbb{P}\left(\left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = 1$$

Äquivalent dazu:

$$\mathbb{P}\left(\omega \in \Omega : \forall\epsilon > 0, \exists N \in \mathbb{N} \text{ s.t. } \forall n \geq N, |X_n(\omega) - X(\omega)| < \epsilon\right) = 1$$

Eigenschaften der Fast Sicheren Konvergenz

- **Starke Konvergenzform**: Sie impliziert Konvergenz in Wahrscheinlichkeit. Wenn $X_n \xrightarrow{f.s.} X$, dann $X_n \xrightarrow{i.W.} X$. Die Umkehrung gilt im Allgemeinen nicht.
- **Bewahrt arithmetische Operationen**: Wenn $X_n \xrightarrow{f.s.} X$ und $Y_n \xrightarrow{f.s.} Y$, dann gilt:
  - $X_n + Y_n \xrightarrow{f.s.} X + Y$
  - $X_n Y_n \xrightarrow{f.s.} XY$ (falls X, Y f.s. endlich sind)
  - $X_n / Y_n \xrightarrow{f.s.} X/Y$ (falls $Y_n \neq 0$ und $Y \neq 0$ f.s.)
- **Continuous Mapping Theorem**: Wenn $X_n \xrightarrow{f.s.} X$ und g eine stetige Funktion ist, dann $g(X_n) \xrightarrow{f.s.} g(X)$.

Fast Sicher vs. In Wahrscheinlichkeit

- **Fast Sichere Konvergenz** ($X_n \xrightarrow{f.s.} X$): Für jedes ω (außer einer Nullmenge) konvergiert die Zahlenfolge $X_n(\omega)$. Dies betrifft die pfadweise Konvergenz.
- **Konvergenz in Wahrscheinlichkeit** ($X_n \xrightarrow{i.W.} X$): $\mathbb{P}(|X_n - X| > \epsilon) \to 0$ für $n \to \infty$ für jedes $\epsilon > 0$. Dies betrifft die Wahrscheinlichkeit, dass große Abweichungen gegen Null gehen.

Fast Sicher vs. In Wahrscheinlichkeit

- **Fast Sichere Konvergenz** ($X_n \xrightarrow{f.s.} X$): Für jedes ω (außer einer Nullmenge) konvergiert die Zahlenfolge $X_n(\omega)$. Dies betrifft die pfadweise Konvergenz.
- **Konvergenz in Wahrscheinlichkeit** ($X_n \xrightarrow{i.W.} X$): $\mathbb{P}(|X_n - X| > \epsilon) \to 0$ für $n \to \infty$ für jedes $\epsilon > 0$. Dies betrifft die Wahrscheinlichkeit, dass große Abweichungen gegen Null gehen.

Wichtige Erkenntnis: Fast sichere Konvergenz impliziert Konvergenz in Wahrscheinlichkeit, aber nicht umgekehrt. Fast sichere Konvergenz ist eine stärkere Bedingung, da sie die Konvergenz für fast jede Realisierung des Zufallsprozesses garantiert, nicht nur, dass die Wahrscheinlichkeit der Abweichung gering wird.

Ein kurzes Gegenbeispiel: Konvergenz in Wahrscheinlichkeit, aber nicht fast sicher

- Betrachten Sie eine Folge von Zufallsvariablen X_n, die auf dem Intervall $[0, 1]$ definiert sind.
- Teile das Intervall in immer kleinere Stücke auf, und definiere $X_n = 1$ auf einem dieser Stücke und $X_n = 0$ sonst.
- Beispiel:
    - $X_1 = \mathbf{1}_{[0,1]}$
    - $X_2 = \mathbf{1}_{[0,1/2]}$, $X_3 = \mathbf{1}_{[1/2,1]}$
    - $X_4 = \mathbf{1}_{[0,1/4]}$, $X_5 = \mathbf{1}_{[1/4,1/2]}$, $X_6 = \mathbf{1}_{[1/2,3/4]}$, $X_7 = \mathbf{1}_{[3/4,1]}$
    - Und so weiter, die Intervalle werden immer kleiner und wandern durch $[0, 1]$.
- Für jedes $\omega \in [0, 1]$ springt $X_n(\omega)$ unendlich oft zwischen 0 und 1 hin und her. Die Folge $X_n(\omega)$ konvergiert für **kein** ω.
- Aber $\mathbb{P}(|X_n - 0| > \epsilon) = P(X_n = 1)$ geht gegen Null, da die Länge der Intervalle schrumpft.
- Dies ist ein klassisches Beispiel für Konvergenz in Wahrscheinlichkeit zu 0, aber nicht fast sichere Konvergenz.

Intuitives Verständnis des SLLN

- Das Starke Gesetz der Großen Zahlen ist eine der wichtigsten Anwendungen der fast sicheren Konvergenz.
- Es untermauert die Idee, dass sich die empirischen Mittelwerte von Zufallsvariablen auf lange Sicht dem theoretischen Erwartungswert annähern.
- Im Gegensatz zum Schwachen Gesetz (das Konvergenz in Wahrscheinlichkeit nutzt), garantiert das Starke Gesetz eine stärkere Form der Konvergenz: Es sagt aus, dass der Stichprobenmittelwert **fast sicher** gegen den Erwartungswert konvergiert.
- Dies bedeutet, dass für fast jede unendliche Folge von Experimenten der beobachtete Durchschnitt der Ergebnisse dem wahren Durchschnitt der Population entsprechen wird.

Formulierung des Starken Gesetzes der Großen Zahlen

Satz (Starkes Gesetz der Großen Zahlen (SLLN))

Sei $(X_i)_{i \in \mathbb{N}}$ eine Folge von unabhängigen und identisch verteilten (i.i.d.) Zufallsvariablen mit endlichem Erwartungswert $\mathbb{E}[X_1] = \mu$. Dann konvergiert der Stichprobenmittelwert $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$ fast sicher gegen den Erwartungswert μ:

$$\bar{X}_n \xrightarrow{f.s.} \mu$$

- **Wichtige Bemerkung**: Im Gegensatz zum WLLN erfordert das SLLN unter der Annahme von i.i.d.-Variablen nur die Existenz des ersten Moments ($\mathbb{E}[|X_1|] < \infty$). Eine endliche Varianz ist nicht zwingend notwendig, macht den Beweis aber einfacher
- Das SLLN ist die mathematische Rechtfertigung für die Verwendung von Stichprobenmittelwerten als Schätzer für Populationserwartungswerte im Langzeitverhalten.

Intuition hinter "Konvergenz in Verteilung"

- Bei der Konvergenz in Verteilung interessieren wir uns nicht für die Konvergenz der Zufallsvariablen selbst, sondern für die Konvergenz ihrer **Verteilungsfunktionen**.
- Es geht darum, dass die Wahrscheinlichkeitsverteilung von X_n sich der Wahrscheinlichkeitsverteilung von X annähert, wenn n groß wird.
- Die Zufallsvariablen X_n und X müssen nicht einmal auf demselben Wahrscheinlichkeitsraum definiert sein.
- *Beispiel:* Der Zentraler Grenzwertsatz. Die normierte Summe von vielen unabhängigen Zufallsvariablen konvergiert gegen eine Normalverteilung, egal welche Verteilung die einzelnen Summanden hatten.

Formale Definition der Konvergenz in Verteilung

Definition (Konvergenz in Verteilung)

Eine Folge von Zufallsvariablen $(X_n)_{n\in\mathbb{N}}$ konvergiert **in Verteilung** gegen eine Zufallsvariable X, bezeichnet als $X_n \xrightarrow{i.V.} X$, wenn die Verteilungsfunktionen $F_{X_n}(x) = P(X_n \leq x)$ konvergieren gegen die Verteilungsfunktion $F_X(x) = P(X \leq x)$ für alle Punkte $x \in \mathbb{R}$, an denen F_X stetig ist.

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \quad \text{für alle } x \text{ an denen } F_X \text{ stetig ist.}$$

Einfaches Beispiel: Konvergenz einer Gleichverteilung

- Betrachten wir eine Folge von Zufallsvariablen $X_n \sim U(0, 1/n)$, also gleichverteilt auf dem Intervall $[0, 1/n]$.
- Die Verteilungsfunktion von X_n ist gegeben durch:

$$F_{X_n}(x) = P(X_n \leq x) = \begin{cases} 0 & \text{für } x < 0 \\ nx & \text{für } 0 \leq x < 1/n \\ 1 & \text{für } x \geq 1/n \end{cases}$$

- Mit $n \to \infty$, wird das Intervall $[0, 1/n]$ immer kleiner und zieht sich auf den Punkt 0 zusammen.
- Die Grenzverteilungsfunktion $F_X(x)$ wird die Verteilungsfunktion einer Zufallsvariablen X sein, die mit Wahrscheinlichkeit 1 den Wert 0 annimmt (eine Dirac-Verteilung oder eine Punktmasse bei 0).

$$F_X(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 & \text{für } x \geq 0 \end{cases}$$

Einfaches Beispiel: Konvergenz einer Gleichverteilung

- Betrachten wir eine Folge von Zufallsvariablen $X_n \sim U(0, 1/n)$, also gleichverteilt auf dem Intervall $[0, 1/n]$.
- Die Verteilungsfunktion von X_n ist gegeben durch:

$$F_{X_n}(x) = P(X_n \leq x) = \begin{cases} 0 & \text{für } x < 0 \\ nx & \text{für } 0 \leq x < 1/n \\ 1 & \text{für } x \geq 1/n \end{cases}$$

- Mit $n \to \infty$, wird das Intervall $[0, 1/n]$ immer kleiner und zieht sich auf den Punkt 0 zusammen.
- Wir haben

$$F_X(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 & \text{für } x \geq 0 \end{cases}$$

- Für alle $x \neq 0$ (also alle Stetigkeitspunkte von $F_X(x)$) gilt $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$.
- Daher konvergiert X_n in Verteilung gegen $X \sim \delta_0$ (eine Punktmasse bei 0).

Eigenschaften der Konvergenz in Verteilung

- **Schwächste Form der Konvergenz**: Konvergenz in Wahrscheinlichkeit impliziert Konvergenz in Verteilung ($X_n \xrightarrow{i.W.} X \implies X_n \xrightarrow{i.V.} X$), aber nicht umgekehrt.
- **Continuous Mapping Theorem (Erweitert)**: Wenn $X_n \xrightarrow{i.V.} X$ und g eine stetige Funktion ist, dann $g(X_n) \xrightarrow{i.V.} g(X)$.

Beispiel: Der Zentrale Grenzwertsatz (ZGS)

- Der Zentrale Grenzwertsatz ist das Herzstück vieler statistischer Anwendungen. Er begründet, warum die Normalverteilung so prominent ist.

Satz (Zentraler Grenzwertsatz (ZGS))

Sei $(X_i)_{i=1}^n$ eine Folge von unabhängigen und identisch verteilten (i.i.d.) Zufallsvariablen mit endlichem Erwartungswert $\mu = \mathbb{E}[X_1]$ und endlicher Varianz $\sigma^2 = Var(X_1) > 0$. Dann konvergiert die standardisierte Summe der Zufallsvariablen in Verteilung gegen eine Standardnormalverteilung $\mathcal{N}(0,1)$:

$$\frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \xrightarrow{i.V.} \mathcal{N}(0,1)$$

Oder äquivalent, wenn $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ der Stichprobenmittelwert ist:

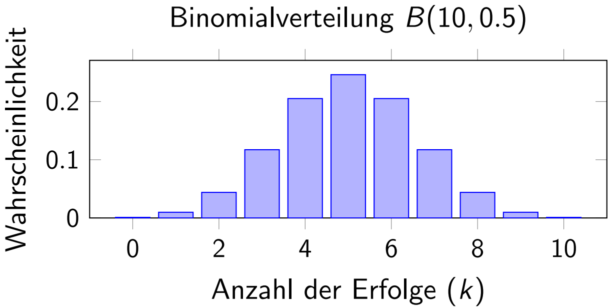
$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{i.V.} \mathcal{N}(0,\sigma^2)$$

Die Binomialverteilung: Eine kurze Einführung

- Die Binomialverteilung beschreibt die Anzahl der Erfolge in einer festen Anzahl von unabhängigen Versuchen, von denen jeder nur zwei mögliche Ausgänge hat (Erfolg oder Misserfolg).
- Solche Versuche werden **Bernoulli-Versuche** genannt.
- Die Binomialverteilung $B(n, p)$ hat zwei Parameter:
  - n: Die Anzahl der Versuche (z.B. wie oft eine Münze geworfen wird).
  - p: Die Wahrscheinlichkeit eines Erfolges in einem einzelnen Versuch (z.B. die Wahrscheinlichkeit, Kopf zu werfen).
- Die Wahrscheinlichkeitsfunktion für k Erfolge in n Versuchen ist gegeben durch:

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

- Ist $X = X_1 + \cdots + X_n$, mit $X1, X_2, \ldots$ i.i.d. Bernoulli(p) -verteilt auf $\{0, 1\}$, so ist X binomialverteilt mit Parametern n, p
- Der Erwartungswert einer binomialverteilten Zufallsvariable ist $\mathbb{E}[X] = np$ und die Varianz ist $\mathrm{Var}(X) = np(1 - p)$.


Binomialverteilung $B(10, 0.5)$
Wahrscheinlichkeit
0.2
0.1
0
0
2
4
6
8
10
Anzahl der Erfolge (k)

Spezialfall des ZGS: Der Satz von De Moivre-Laplace

- Ein klassisches und sehr anschauliches Beispiel für den Zentralen Grenzwertsatz ist die Normalapproximation der Binomialverteilung.
- Betrachten Sie eine Folge von n unabhängigen Bernoulli-Versuchen (z.B. Münzwürfe), bei denen die Erfolgswahrscheinlichkeit p ist.
- Die Anzahl der Erfolge S_n ist binomialverteilt: $S_n \sim B(n, p)$.

Satz (Satz von De Moivre-Laplace)

Sei $S_n \sim B(n, p)$ die Anzahl der Erfolge in n unabhängigen Bernoulli-Versuchen mit Erfolgswahrscheinlichkeit $p \in (0, 1)$. Dann konvergiert die standardisierte Zufallsvariable

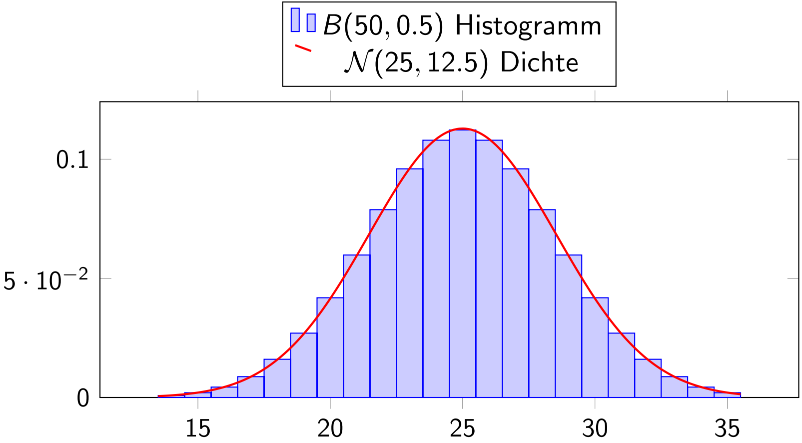
$$Z_n = \frac{S_n - np}{\sqrt{np(1-p)}}$$

in Verteilung gegen die Standardnormalverteilung $\mathcal{N}(0, 1)$:

$$Z_n \xrightarrow{i.V.} \mathcal{N}(0, 1)$$

Visualisierung: Binomialverteilung nähert sich der Normalverteilung an

- Betrachten wir die Binomialverteilung $B(n, p = 0.5)$ für $n = 50$.



Beziehungen zwischen den Konvergenzarten

$$\textbf{Fast Sichere Konvergenz } (\mathbf{X_n} \xrightarrow{\text{f.s.}} \mathbf{X})$$
$$\Downarrow$$
$$\textbf{Konvergenz in Wahrscheinlichkeit } (\mathbf{X_n} \xrightarrow{\text{i.W.}} \mathbf{X})$$
$$\Downarrow$$
$$\textbf{Konvergenz in Verteilung } (\mathbf{X_n} \xrightarrow{\text{i.V.}} \mathbf{X})$$

- Jede obere Konvergenzart impliziert die darunter liegende.
- Die Umkehrungen gelten im Allgemeinen nicht (wie am Gegenbeispiel gezeigt).
- Fast sichere Konvergenz ist die stärkste Form und garantiert "pfadweise" Konvergenz.
- Konvergenz in Verteilung ist die schwächste, aber extrem nützlich für die asymptotische Approximation der Verteilung einer Zufallsvariable.