Mathe III

Konvergenzen

Ralf Herbrich

Hasso Plattner Institut
Digital Engineering · Universität Potsdam

Überblick

1. Stochastische Konvergenz
2. Gesetz der großen Zahlen
3. Zentraler Grenzwertsatz

Überblick

1. **Stochastische Konvergenz**
2. Gesetz der großen Zahlen
3. Zentraler Grenzwertsatz

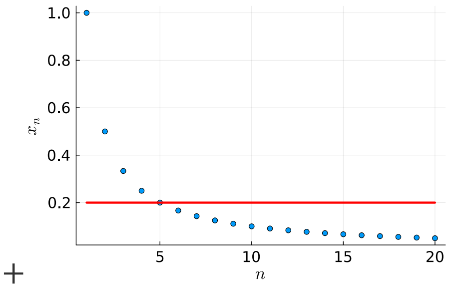
Konvergenzen

- **Definition (Klassische Konvergenz)**. Gegeben eine Folge $(x_n)_{n\in\mathbb{N}}$ von Zahlen in $\mathbb{R}$, sagen wir, dass die Folge gegen eine Zahl x konvergiert, $x_n \to x$, genau dann wenn für alle $\varepsilon > 0$ ein Index $N(\varepsilon)$ existiert, so dass

$$\forall j \geq N(\varepsilon)\colon \left|x_j - x\right| < \varepsilon$$

- **Beispiel (Konvergenz)**. Gegeben sei die Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n = \frac{1}{n}$. Gegen welchen Wert konvergiert diese Folge?
  - **Lösung**: Die Folge konvergiert gegen 0. Für ein beliebiges $\varepsilon > 0$ wählen wir $N(\varepsilon) = \frac{1}{\varepsilon} + 1$. Dann gilt für alle $j \geq N(\varepsilon)$, dass

$$j \geq \frac{1}{\varepsilon} + 1 \quad\Leftrightarrow\quad \varepsilon \geq \frac{1}{j-1} > \frac{1}{j} \quad\Leftrightarrow\quad \left|\frac{1}{j} - 0\right| < \varepsilon \quad\Leftrightarrow\quad \left|x_j - x\right| < \varepsilon$$



- **Aber**: Im Gegensatz zur Analysis kann eine Zufallsvariable alle Werte in der Ergebnismenge annehmen ($\mathbb{N}$ oder $\mathbb{R}$). Wie definiert man dann *Konvergenzen*?

Stochastische Konvergenz

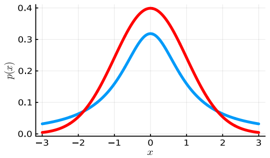
- **Definition (Stochastische Konvergenz).** Gegeben eine Folge $(X_n)_{n\in\mathbb{N}}$ von Zufallsvariablen, sagen wir, dass die Folge gegen eine Zufallsvariable X konvergiert, $X_n \to X$, wenn sie entweder
  - in Verteilung konvergiert: $X_n \xrightarrow{d} X$
  - in Wahrscheinlichkeit konvergiert: $X_n \xrightarrow{P} X$
  - fast sicher konvergiert: $X_n \xrightarrow{f.s.} X$
- **Bemerkungen (Stochastische Konvergenz)**
  - Bei diesen Konvergenzarten wird durch die darüberstehende Konvergenzart impliziert.
  - Die **Konvergenz in Verteilung** prüft lediglich, ob die Verteilungsfunktion der Folge gegen die Verteilungsfunktion der Zufallsvariable konvergiert.
  - Bei der **Konvergenz in Wahrscheinlichkeit** muss gelten, dass der Abstand zwischen der Folge und der Zufallsvariable „fast sicher“ klein ist (gegen Wahrscheinlichkeit 0 konvergiert).
  - Bei der **fast sicheren Konvergenz** muss die Wahrscheinlichkeit für eine Konvergenz bei 1 liegen.

Konvergenz in Verteilung

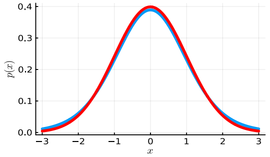
- **Intuitiv**: Die Konvergenz in Verteilung prüft, ob die Verteilungsfunktion der Folge gegen die Verteilungsfunktion der Zufallsvariable konvergiert.
- **Definition (Konvergenz in Verteilung)**. Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge reeller Zufallsvariablen und X eine reelle Zufallsvariable. Dann konvergiert $(X_n)_{n\in\mathbb{N}}$ in Verteilung gegen X, $X_n \xrightarrow{d} X$, genau dann, wenn für jedes $c \in \mathbb{R}$, bei welchem F_X stetig ist, gilt, dass
$$\lim_{n\to\infty} F_{X_n}(c) = F_X(c)$$
- **Bemerkungen (Konvergenz in Verteilung)**
  - Es ist wichtig, dass dies nur erfüllt sein muss an Punkten, bei welchem F_X stetig ist!
  - Wir haben schon eine Konvergenz in Verteilung kennengelernt: Die t-Verteilung konvergiert mit steigenden Freiheitsgraden gegen die Standardnormalverteilung!
- **Beispiel (Konvergenz in Verteilung)**. Sei $X \sim \mathrm{Bern}(0.5)$ und $(X_n)_{n\in\mathbb{N}}$ mit $X_n = 1 - X$. Dann gilt offensichtlich $X_n \sim \mathrm{Bern}(0.5)$ und offensichtlich konvergiert $(X_n)_{n\in\mathbb{N}}$ in Verteilung gegen X, da die Verteilungen identisch sind.

  Allerdings gilt für alle $n \in \mathbb{N}$ stets, dass $|X_n - X| = 1$, also keine Konvergenz in Wahrscheinlichkeit besteht!

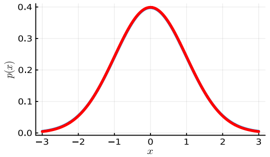
$t(1)$



$t(10)$



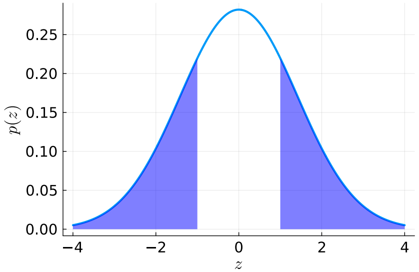
$t(100)$



Konvergenz in Wahrscheinlichkeit

- **Intuitiv**: Bei der Konvergenz in Wahrscheinlichkeit muss gelten, dass der Abstand zwischen der Folge und der Zufallsvariable fast sicher klein ist.
- **Definition (Konvergenz in Wahrscheinlichkeit)**. Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge reeller Zufallsvariablen und X eine reelle Zufallsvariable. Dann konvergiert $(X_n)_{n\in\mathbb{N}}$ in Wahrscheinlichkeit gegen X, $X_n \xrightarrow{P} X$, genau dann, wenn für jedes $\varepsilon > 0$ gilt, dass
$$\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$$
- **Bemerkungen (Konvergenz in Wahrscheinlichkeit)**
  - Die Konvergenz in Wahrscheinlichkeit bedeutet nicht, dass die Verteilungen sich mit steigendem n annähern, sondern dass sich die *Werte* der Stichproben annähern!
  - Obwohl $X_n \xrightarrow{d} X$ für $X_n \sim t(n)$ und $X \sim \mathcal{N}(0,1)$, hat das Ereignis $|X_n - X| \geq 1$ die Wahrscheinlichkeit von $\approx 48\%$ egal wie groß n wird!

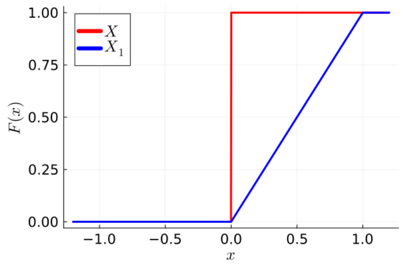
$$\lim_{n\to\infty} P(|X_n - X| \geq 1) = P(\underbrace{|Z|}_{Z \sim \mathcal{N}(0,2)} \geq 1) = P(Z \leq -1) + \underbrace{P(Z \geq 1)}_{P(Z \leq -1)} = 2 \cdot \Phi\left(-\frac{1}{\sqrt{2}}\right) \approx 0.4795$$



Fast sichere Konvergenz

- **Intuitiv**: Die Wahrscheinlichkeit für eine Konvergenz muss bei 1 liegen.
- **Definition (Fast sichere Konvergenz).** Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge reeller Zufallsvariablen und X eine reelle Zufallsvariable auf einem Ereignisraum $(\Omega, \mathcal{F}, P)$. Dann konvergiert $(X_n)_{n\in\mathbb{N}}$ fast sicher gegen X, $X_n \xrightarrow{f.s.} X$, genau dann, wenn gilt, dass

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



- **Bemerkung (Fast sichere Konvergenz)**
  - Fast sichere Konvergenz ist äquivalent zu: Wenn für jedes $\varepsilon > 0$ gilt, dass

$$\lim_{n\to\infty} P\left(\max_{m\geq n} |X_m - X| > \varepsilon\right) = 0$$

Im Vergleich zu: $\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$

- **Beispiel (Fast sichere Konvergenz).** Sei X eine diskrete Zufallsvariable mit $P(X = 0) = 1$. Seien $(X_n)_{n\in\mathbb{N}}$ eine Folge reeller Zufallsvariablen mit $X_n \sim \mathcal{U}\left(0, \frac{1}{n}\right)$. Dann gilt

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

womit die Folge fast sicher konvergiert.

Eigenschaften der Konvergenzen

- **Satz (Stabilitätseigenschaften der Konvergenzen).** Seien $(X_n)_{n\in\mathbb{N}}$ und $(Y_n)_{n\in\mathbb{N}}$ Folgen reeller Zufallsvariablen. Dann gilt für die fast sichere Konvergenz:

  1. Falls $X_n \xrightarrow{f.s.} X$ und $Y_n \xrightarrow{f.s.} Y$, so gilt auch $a \cdot X_n + b \cdot Y_n \xrightarrow{f.s.} a \cdot X + b \cdot Y$.
  2. Falls $X_n \xrightarrow{f.s.} X$ und $Y_n \xrightarrow{f.s.} Y$, so gilt auch $X_n \cdot Y_n \xrightarrow{f.s.} X \cdot Y$.

  Analog gilt für die Konvergenz in Wahrscheinlichkeit:

  1. Falls $X_n \xrightarrow{P} X$ und $Y_n \xrightarrow{P} Y$, so gilt auch $a \cdot X_n + b \cdot Y_n \xrightarrow{P} a \cdot X + b \cdot Y$.
  2. Falls $X_n \xrightarrow{P} X$ und $Y_n \xrightarrow{P} Y$, so gilt auch $X_n \cdot Y_n \xrightarrow{P} X \cdot Y$.

- **Bemerkungen (Konvergenzeigenschaften)**
  - Diese Konvergenzen lassen sich ähnlich wie die Grenzwertgesetze der reellen Zahlen benutzen.
  - **Achtung**: Keine dieser Eigenschaften gilt für die Konvergenz in Verteilung!

Überblick

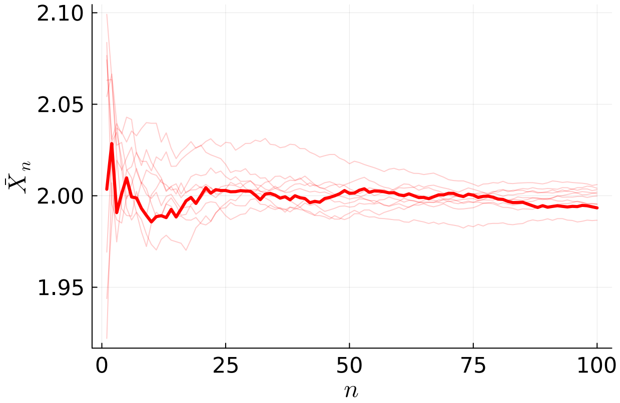
1. Stochastische Konvergenz
2. **Gesetz der großen Zahlen**
3. Zentraler Grenzwertsatz

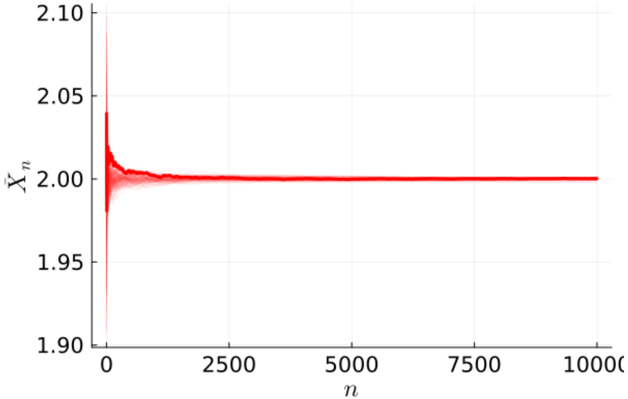
Gesetz der großen Zahlen

- **Beispiel (Gesetz der großen Zahlen)**. Eine Maschine produziert Schrauben, deren Gewicht um einen Sollwert schwankt. Dieser Sollwert ist nicht bekannt, jedoch kann beliebig oft eine Schraube als Stichprobe gezogen und vermessen werden. Wie verhält sich das Schraubengewicht im Mittel über viele Beobachtungen?

- **Mathematische Formulierung:** Wir betrachten ein n-stufiges Zufallsexperiment. Für $i = 1, \ldots, n$ gibt X_i das Gewicht der Schraube beim i-ten Zufallsexperiment an. Wir bilden eine neue Zufallsvariable $\bar{X}_n$ mit

$$\bar{X}_n = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$$





Schwaches Gesetz der großen Zahlen

- **Satz (Schwaches Gesetz der großen Zahlen).** Seien $(X_n)_{n\in\mathbb{N}}$ paarweise unabhängige, identisch verteilte, reelle Zufallsvariablen mit dem Erwartungswert $E[X_n] = m$ für alle $n \in \mathbb{N}$. Dann gilt

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} m$$

- **Bemerkungen (Schwaches Gesetz der großen Zahlen)**
  - Die konvergierende Folge ist hier die Folge der Mittelwerte $(\bar{X}_n)_{n\in\mathbb{N}}$ mit

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

    und nicht die Folge $(X_n)_{n\in\mathbb{N}}$!
  - Man kann die Bedingung der **Unabhängigkeit** noch abschwächen!

Schwaches Gesetz der großen Zahlen

- **Satz (Schwaches Gesetz der großen Zahlen – Alternative Formulierung).** Seien $(X_n)_{n\in\mathbb{N}}$ paarweise *unkorrelierte*, reelle Zufallsvariablen mit dem Erwartungswert $E[X_n] = m$ und Varianz $V[X_n] = v$ für alle $n \in \mathbb{N}$. Dann gilt
$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} m$$
- **Beweis**: Wir benutzen die Ungleichung von Tschebyscheff und Eigenschaften des Erwartungswertes und der Varianz für $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Dann gilt

$$E[\bar{X}_n] = \frac{1}{n}\cdot\sum_{i=1}^{n} E[X_i] = \frac{n}{n}\cdot m = m$$

$$V[\bar{X}_n] = \frac{1}{n^2}\cdot\sum_{i=1}^{n} V[X_i] = \frac{n}{n^2}\cdot v = \frac{v}{n}$$

weil unkorreliert!

Daher gilt für alle $\varepsilon > 0$:

$$P(|\bar{X}_n - E[\bar{X}_n]| \geq \varepsilon) \leq \frac{V[\bar{X}_n]}{\varepsilon^2} \quad \Rightarrow \quad \lim_{n\to\infty} P(|\bar{X}_n - m| \geq \varepsilon) \leq \lim_{n\to\infty} \frac{v}{n\cdot\varepsilon^2} = 0$$

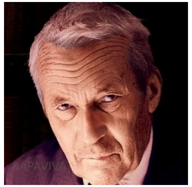
Tschebyscheff. Für alle $\varepsilon > 0$

$$P(|X - E[X]| \geq \varepsilon) \leq \frac{V[X]}{\varepsilon^2}$$

Starkes Gesetz der großen Zahlen

- **Satz (Starkes Gesetz der großen Zahlen).** Seien $(X_n)_{n\in\mathbb{N}}$ paarweise unabhängige, identisch verteilte, reelle Zufallsvariablen mit dem Erwartungswert $E[X_n] = m$ für alle $n \in \mathbb{N}$. Dann gilt

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{f.s.} m$$



Andrey Kolmogorov (1903 – 1987)

- **Satz (Starkes Gesetz der großen Zahlen – Alternative Formulierung).** Seien $(X_n)_{n\in\mathbb{N}}$ paarweise *unkorrelierte*, reelle Zufallsvariablen mit dem Erwartungswert $E[X_n] = m$ und Varianz $V[X_n] = v$ für alle $n \in \mathbb{N}$. Dann gilt

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{f.s.} m$$

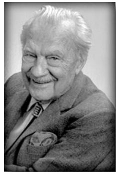
- **Bemerkungen (Starkes Gesetz der großen Zahlen)**
  - Alles was sich ändert ist die Konvergenz in Wahrscheinlichkeit!
  - Der Beweis ist deutlich schwerer und von Andrey Kolmogorov im Jahr 1930 erbracht.

Anwendung: Monte-Carlo-Integration

- **Idee**: Können wir das Gesetz der großen Zahlen zur Integration von komplizierten Funktionen benutzen?
- **Monte-Carlo Integration**:
  - **Gegeben**: Eine nach oben beschränkte, positive Funktion $h: [0,1] \to [0,c]$
  - **Gesucht**: Der numerische Wert für $\int_0^1 h(x)\,dx$
  - **Algorithmus**:
    1. Simuliere n gleichverteilte Zufallsvariablen $(X_i)_{1\leq i\leq n}$ mit $X_i \sim \mathcal{U}(0,1)$
    2. Berechne $\frac{1}{n}\sum_{i=1}^n h(X_i)$ als Approximation von $\int_0^1 h(x)\,dx$
  - **Korrektheit**:
    - Es gilt $E[h(X)] = \int_{\mathbb{R}} h(x) \cdot p_X(x)\,dx = \int_0^1 h(x) \cdot 1\,dx = \int_0^1 h(x)dx$
    - Ferner gilt $V[h(X)] \leq c^2$ (da $h(X_i)$ nur Werte in $[0,c]$ liefert)
    - Mit der Tschebyscheff-Ungleichung folgt für alle $\varepsilon > 0$

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} h(X_i) - \int_0^1 h(x)\,dx\right| \geq \varepsilon\right) \leq \frac{c^2}{n \cdot \varepsilon^2}$$

Stanisław Ulam
(1909 – 1984)



Nicholas Metropolis
(1915 – 1999)

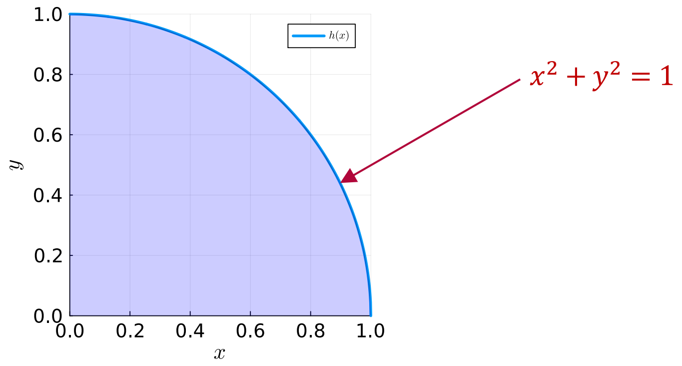
Monte-Carlo-Integration: Beispiel

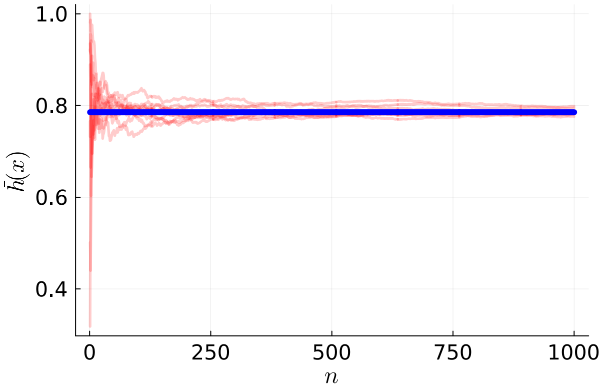
- **Beispiel (Monte-Carlo-Integration).** Der Einheitskreis hat eine Fläche von $\pi \cdot 1^2 = \pi$.

  Wir wollen ein Viertel des Einheitskreises mittels numerischer Integration ermitteln.

  $$h: x \mapsto \sqrt{1 - x^2}$$

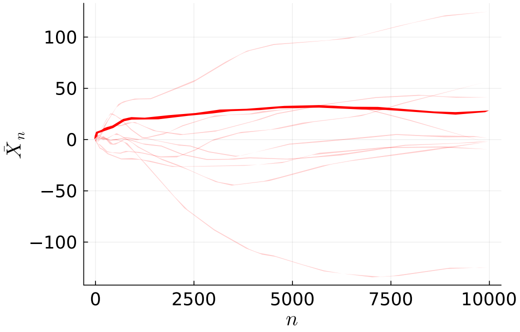
  Der Mittelwert konvergiert gegen 0.785, was in etwa $\frac{\pi}{4}$ entspricht.

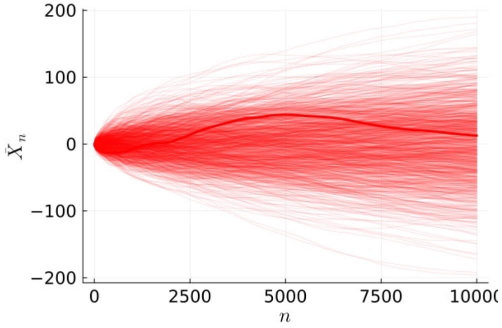


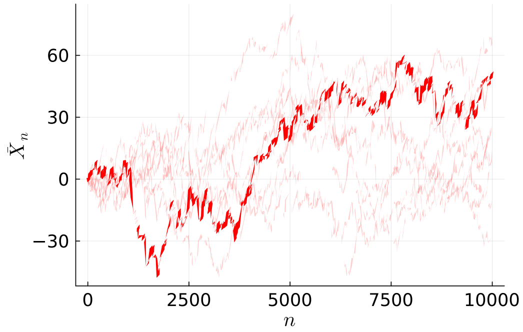


Grenzen des Gesetzes der großen Zahlen

- **Beispiel (Korrelation)**. Wir betrachten $X_n | X_{n-1} \sim \mathcal{N}(X_{n-1}, 1)$
  - Korrelation führt zwar zur Konvergenz, aber nicht gegen den Mittelwert.
  - Außerdem konvergiert jede Realisierung gegen einen anderen Wert.
  - Daher ist das Gesetz der großen Zahlen hierfür nicht anwendbar.
- **Beispiel (Varianz)**. Wir betrachten $X_n \sim \mathcal{U}(-n, n)$, aber paarweise unabhängig
  - Offenbar konvergiert der Mittelwert überhaupt nicht!
  - Daher ist das Gesetz der großen Zahlen hierfür nicht anwendbar.





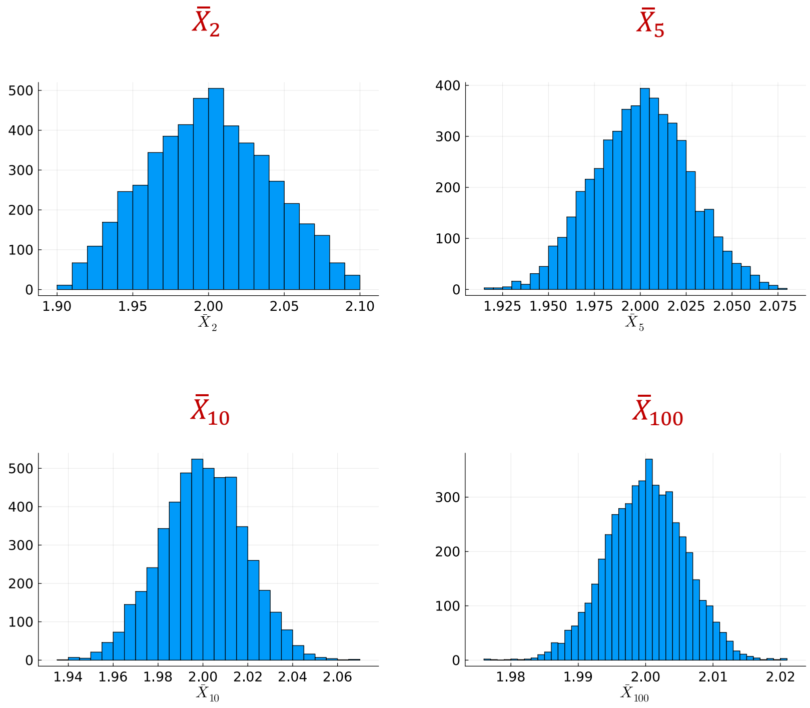


Überblick

1. Stochastische Konvergenz
2. Gesetz der großen Zahlen
3. **Zentraler Grenzwertsatz**

Verteilung des Mittelwertes

- **Wir wissen nach dem Gesetz der großen Zahlen:**
  - Bei wachsender Anzahl von Beobachtungen konvergiert der Mittelwert aller Stichproben fast sicher gegen den Erwartungswert.
  - Ergo: Für sehr viele Beobachtungen hat der Mittelwert den gleichen Erwartungswert wie eine Stichprobe und eine sehr niedrige Varianz.
- **Was wir noch nicht wissen:**
  - Wie entwickelt sich diese Konvergenz?
  - Welcher Verteilung folgen die Mittelwerte?



$$\bar{X}_n = \sum_{i=1}^{n} X_i\,, \qquad X_i \sim \mathcal{U}(1.9, 2.1)$$

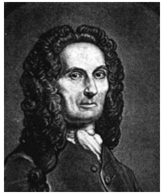
Zentraler Grenzwertsatz

- **Satz (Zentraler Grenzwertsatz).** Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von unabhängigen, identisch verteilten, reellen Zufallsvariablen mit Erwartungswert $E[X_i] = m$ und positiver Varianz $V[X_i] = v > 0$. Dann gilt

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{X_i - m}{\sqrt{v}} \xrightarrow{d} \mathcal{N}(0,1)$$

- **Bemerkungen (Zentraler Grenzwertsatz)**
  - Der zentrale Grenzwertsatz ist in seinen Voraussetzungen strenger als das Gesetz der großen Zahlen.
    - Es ist nicht möglich, auf die endliche Varianz zu verzichten.
    - Es ist nicht möglich, die Unabhängigkeit durch Unkorreliertheit zu ersetzen.
  - Die asymptotische Verteilung des Mittelwerts $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ ist gegeben durch

$$\bar{X}_n \xrightarrow{d} \mathcal{N}\left(m, \frac{v}{n}\right) \longleftarrow \frac{1}{n}\sum_{i=1}^{n}(X_i - m) \xrightarrow{d} \mathcal{N}\left(0, \frac{v}{n}\right)$$


Abraham de Moivre (1667 – 1754)


George Pólya (1887 – 1985)

Zentraler Grenzwertsatz: Beispiel

- **Beispiel (Zentraler Grenzwertsatz)**. Eine Maschine stellt Schrauben her, deren Gewicht mit einer Varianz von 0.2 um den Erwartungswert von 2 schwankt. Welcher Verteilung folgt der Mittelwert von 100 gezogenen Schrauben?

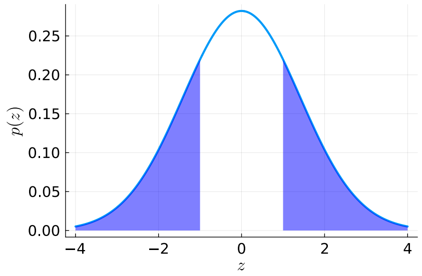
- **Lösung**: Aufgrund des zentralen Grenzwertsatzes folgt

$$\frac{1}{100}\sum_{i=1}^{100} X_i \sim \mathcal{N}\left(2, \frac{0.2}{100}\right)$$

- **Bemerkungen (Zentraler Grenzwertsatz)**
  - Offensichtlich geht $\frac{v}{n}$ für $n \to \infty$ gegen 0.
  - Die Varianz der Normalverteilung nimmt damit für größere n ab und streut immer weniger um den konstanten Erwartungswert m.
  - Der zentrale Grenzwertsatz beweist das schwache Gesetz der großen Zahlen.

Zentraler Grenzwertsatz: Beispiel

- **Beispiel (Zentraler Grenzwertsatz).** Ein Astronom will die Distanz d von der Erde bis zu einem neuen Stern messen. Durch atmosphärisches Rauschen ist die Distanzmessung verrauscht. Wenn der Astronom annimmt, dass jede Messung den Erwartungswert d und eine Standardabweichung von 2 Lichtjahren hat, wie viele Messungen müssen gemacht werden, um sich mit 95% sicher zu sein, dass der Mittelwert bis auf ± 0.5 Lichtjahre akkurat ist?



- **Lösung**: Für n Messungen ist der Mittelwert approximativ normalverteilt mit

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathcal{N}\left(d, \frac{4}{n}\right)$$

Wir suchen also n so, dass

$$P(|\bar{X}_n - d| < 0.5) \geq 0.95 \quad \Leftrightarrow \quad P\left(\underbrace{\left|\frac{\bar{X}_n - d}{2/\sqrt{n}}\right|}_{\mathcal{N}(0,1)} < \frac{\sqrt{n}}{4}\right) \geq 0.95 \quad \Leftrightarrow \quad \Phi\left(\frac{\sqrt{n}}{4}\right) - \Phi\left(-\frac{\sqrt{n}}{4}\right) \geq 0.95$$

$$\Leftrightarrow \quad 2\Phi\left(\frac{\sqrt{n}}{4}\right) - 1 \geq 0.95 \quad \Leftrightarrow \quad \Phi\left(\frac{\sqrt{n}}{4}\right) \geq 0.975$$

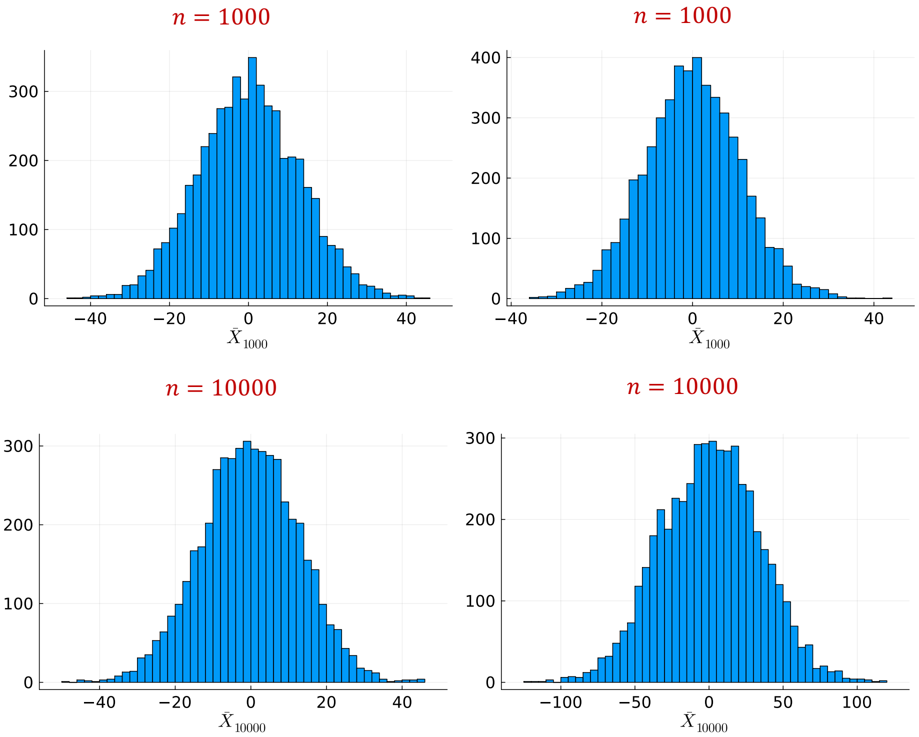
`qnorm(0.975)`

$$\Leftrightarrow \quad \frac{\sqrt{n}}{4} \geq \Phi^{-1}(0.975) \quad \Leftrightarrow \quad n \geq (1.96 \cdot 4)^2 \approx 61.4$$

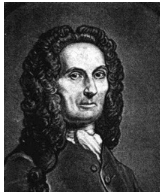
Grenzen des zentralen Grenzwertsatzes

- **Beispiel (Korrelation)**. Welcher Verteilung folgt der Mittelwert aus normalverteilten Stichproben mit Erwartungswert 0, welche miteinander korreliert sind?
  - Korrelation führt dazu, dass auch bei vielen Realisierungen unsaubere Glockenkurven entstehen.
  - Offenbar sinkt auch die Varianz kaum.
- **Beispiel (Varianz)**. Sei $X_n \sim \mathcal{U}(-n, n)$. Welcher Verteilung folgt der Mittelwert der ersten n Werte der Folge?
  - Es entstehen keine sauberen Glockenkurven.
  - Offenbar nimmt die Streuung sogar zu statt ab bei steigendem n.
  - Auch hier ist der zentrale Grenzwertsatz daher nicht anwendbar.

$$X_n | X_{n-1} \sim \mathcal{N}\left(X_{n-1}, \frac{10}{n}\right) \qquad X_n \sim \mathcal{U}(-n, n)$$



Exkurs: Beweis des Zentralen Grenzwertsatzes



Abraham de Moivre
(1667 – 1754)

- Das zentrale Konzept für den Beweis ist das Konzept der **momenterzeugenden Funktion** *(moment-generating function)* einer Zufallsvariablen.
- **Definition (Momenterzeugende Funktion).** Sei X eine reelle Zufallsvariable. Dann nennen wir die Funktion $M_X\colon \mathbb{R} \to \mathbb{R}$ die momenterzeugende Funktion (MGF) mit

$$M_X(t) = E[\exp(tX)]$$

- **Bemerkungen (Momenterzeugende Funktion)**
  - Die Funktion erhält ihren Namen, weil man aus ihr alle Momente $(E[X^n])_{n\in\mathbb{N}}$ der Verteilung bestimmen kann (und damit die gesamte Verteilungsfunktion!)
  - Das n-te Moment der Verteilung von X ist der Wert der n-ten Ableitung an der Stelle 0

Leibniz Notation für die n-te Ableitung von $M_X(t)$

$$\frac{d^n M_X(t)}{dt^n} = M_X^{(n)}(t) = E[\exp(tX) \cdot X^n]$$

$= 1$ für $t = 0$

  - Leicht per Induktion über n zu beweisen: Gilt offensichtlich für $n = 0$. Wenn die Aussage für n gilt, dann ist

Induktionsannahme

$$\frac{d^{n+1} M_X(t)}{dt^{n+1}} = \frac{dE[\exp(tX) \cdot X^n]}{dt} = E[\exp(tX) \cdot X \cdot X^n] = E[\exp(tX) \cdot X^{n+1}]$$

Mathe III

Unit 7 – Konvergenzen

Exkurs: Eigenschaften der MGF

- **Satz (Momenterzeugende Funktion der Summe).** Seien X und Y zwei unabhängige reelle Zufallsvariable. Dann gilt

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$$

- **Beweis**: Wir beweisen den Fall von stetigen Zufallsvariablen. Sei $Z = X + Y$. Dann gilt

$$\begin{aligned}
M_{X+Y}(t) &= M_Z(t) = E[\exp(tZ)] \\
&= \int_{-\infty}^{+\infty} \exp(tz) \cdot p_Z(z)\, dz = \int_{-\infty}^{+\infty} \exp(tz) \cdot \boxed{\left[\int_{-\infty}^{+\infty} p_X(x) \cdot p_Y(z-x)\, dx\right]} dz \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp\big(t(x+y)\big)\, p_X(x) \cdot p_Y(y)\, dx\, dy \\
&= \int_{-\infty}^{+\infty} \boxed{\left[\int_{-\infty}^{+\infty} \exp(tx)\, p_X(x)\, dx\right]} \cdot \exp(ty)\, p_Y(y) \cdot dy \\
&= \boxed{\left[\int_{-\infty}^{+\infty} \exp(tx)\, p_X(x)\, dx\right]} \cdot \boxed{\left[\int_{-\infty}^{+\infty} \exp(ty) \cdot p_Y(y)\, dy\right]}
\end{aligned}$$

weil X und Y unabhängig sind

Variablensubstitution $z = x + y$

feste Zahl für alle $y \in \mathbb{R}$

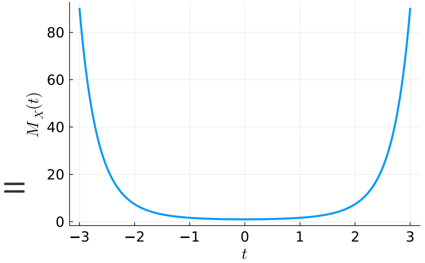
$M_X(t)$

$M_Y(t)$

Exkurs: MGF der Normalverteilung

- **Satz (Momenterzeugende Funktion der Normalverteilung).** Seien X eine standard-normalverteilte Zufallsvariable, $X \sim \mathcal{N}(0,1)$. Dann gilt

$$M_X(t) = \exp\left(\frac{t^2}{2}\right)$$



- **Beweis**: Wir benutzen die Dichte der Standardnormalverteilung und $M_X(t) = E[\exp(tX)]$

$$M_X(t) = \int_{-\infty}^{+\infty} \exp(tx) \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx$$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{x^2}{2} + tx \underbrace{- \frac{t^2}{2} + \frac{t^2}{2}}_{=0}\right) dx$$

$$= \exp\left(\frac{t^2}{2}\right) \cdot \underbrace{\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-t)^2}{2}\right) dx}_{=1}$$

Integral über die Dichte der Normalverteilung $\mathcal{N}(t, 1)$

Exkurs: Beweis des Zentralen Grenzwertsatzes

- **Satz (Zentraler Grenzwertsatz).** Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von paarweise unabhängigen, identisch verteilten, reellen Zufallsvariablen mit $E[X_i]=0$ und $V[X_i]=1$. Dann gilt

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i \xrightarrow{d} \mathcal{N}(0{,}1)$$

- **Beweis**: Sei $Z_n = \frac{1}{\sqrt{n}}(X_1+\cdots+X_n)$. Dann zeigen wir, dass $\lim\limits_{n\to\infty} M_{Z_n}(t) = \exp\left(\frac{t^2}{2}\right)$, was die Konvergenz in Verteilung beweist. Wir bezeichnen die MGF von allen X_i mit M_X und benutzen $M_{\frac{X}{k}}(t) = M_X\left(\frac{t}{k}\right)$.

$$\lim_{n\to\infty}\log\left(M_{Z_n}(t)\right) = \lim_{n\to\infty}\log\left(\prod_{i=1}^{n} M_{\frac{X_i}{\sqrt{n}}}(t)\right) = \lim_{n\to\infty}\log\left(M_X\left(\frac{t}{\sqrt{n}}\right)\right)^n$$

L'Hôpital's Regel

$\varepsilon^2 = \frac{1}{n}$

$$= \boxed{\lim_{\varepsilon\to 0}\frac{\log\left(M_X(\varepsilon t)\right)}{\varepsilon^2}} = \lim_{\varepsilon\to 0}\frac{\frac{d}{dt}M_X(\varepsilon t)\cdot t}{\boxed{M_X(\varepsilon t)}\cdot 2\cdot\varepsilon} = \frac{t}{2}\cdot\boxed{\lim_{\varepsilon\to 0}\frac{\frac{d}{dt}M_X(\varepsilon t)}{\varepsilon}}$$

L'Hôpital's Regel

$= 1$ da nulltes Moment

$$= \frac{t}{2}\cdot\lim_{\varepsilon\to 0}\boxed{\frac{d^2}{dt^2}M_X(\varepsilon t)}\cdot t$$

$$= \frac{t^2}{2}$$

$= 1$ da zweites Moment ($V[X_i]$ bei $E[X_i]=0$)

Viel Spaß bis zur nächsten Vorlesung!