Kapitel 5 Folgen stochastischer Größen und Grenzwertsätze.

In vielen Anwendungen wird die Summe einer durchaus großen Anzahl statistisch unabhängiger, identisch verteilter Zufallsvariablen betrachtet:

$$Z = \sum_{i=1}^{n} X_i$$

Diese Aufgabe tritt zum Beispiel bei der Berechnung relativer Häufigkeiten in der Analyse eines Zufallsexperiments auf. Das statistische Verhalten dieser so gebildeten Zufallsvariablen Z soll in diesem Abschnitt analytisch untersucht werden.

In diesem Zusammenhang steht auch der Begriff des *zentralen Grenzwertsatzes*. Dieser ist, anders als andere Sätze in der Mathematik, ein Sammelbegriff für eine Reihe von mathematischen Aussagen und Sätzen, welche alle die Konvergenz einer Verteilungsfunktion betrachten, die sich aus der Summierung statistisch unabhängiger identisch verteilter Zufallsvariablen ergibt. Als ein wirklich interessantes Ergebnis kann nachgewiesen werden, dass die Summe bzw. das arithmetische Mittel dieser Zufallsvariablen für große Werte n stets gegen eine Standardnormalverteilung konvergiert. Es werden Bedingungen angegeben, unter denen die Konvergenz der Zufallsvariablen Z gegen die Normalverteilung eintritt.

Man betrachtet also im Folgenden eine Zufallsvariable Z_n, die sich als Summe von statistisch unabhängigen, identisch verteilten Zufallsvariablen X_i ergibt:

$$Z_n = \sum_{i=1}^{n} X_i$$

Der Erwartungswert für diese neue Zufallsvariable Z_n kann bereits durch Anwendung der Linearität der Erwartungswertbildung wie folgt berechnet werden:

$$E\{Z_n\} = \sum_{i=1}^{n} E\{X_i\}$$

Wegen der angenommenen statistischen Unabhängigkeit kann auch die Varianz der Zufallsvariablen Z bereits mit dem Satz von Bienaymé wie folgt berechnet werden:

$$Var\{Z_n\} = \sum_{i=1}^{n} Var\{X_i\}$$

Zusätzlich wurde bereits in den vorangegangen Kapiteln gezeigt, dass die resultierende Wahrscheinlichkeitsdichtefunktion $f_Z(z)$ durch Faltung der einzelnen Wahrscheinlichkeitsdichtefunktionen der Zufallsvariablen X_i vollständig berechnet werden kann:

$$f_Z(x) = f_{X_1}(x) * f_{X_2}(x) * \cdots * f_{X_n}(x)$$

Die eigentliche Zielsetzung dieses Kapitels liegt aber darin, nicht nur Erwartungswert und Varianz sondern auch die Wahrscheinlichkeitsdichtefunktion der Zufallsvariablen Z_n insbesondere für große Werte n explizit zu bestimmen bzw. geeignet abzuschätzen.

Grenzwertsätze der Binomialverteilung

Im ersten Schritt soll ein BERNOULLI-Experiment auf sein Grenzwertverhalten untersucht werden. D. h. die Zufallsvariable Z_n setzt sich in diesem Fall aus einer Summe von binärwertigen Zufallsvariablen X_i zusammen:

$$Z_n = \sum_{i=1}^{n} X_i$$

BERNOULLI-Experiment

- Es wird ein Experiment mit binärem Ausgang (Alternativverteilung) n-mal nacheinander durchgeführt oder es werden alternativ n identische Experimente gleichzeitig durchgeführt.

- Die Wahrscheinlichkeit, mit der die Zufallsvariable X_i eine Eins annimmt, soll p betragen:

$$P[X_i = 1] = p \qquad P[X_i = 0] = 1 - p$$

- Die zufälligen Ereignisse seien unabhängig von vorhergehenden oder nachfolgenden Versuchen, d.h., die Zufallsvariablen X_i seien statistisch unabhängig und identisch verteilt.

- $P_n(k)$ beschreibt die Wahrscheinlichkeit, mit der k Einsen in einer Folge von n Versuchen auftreten, d.h. mit der die Zufallsvariable Z_n den Wert k annimmt $[Z_n = k]$:

$$P_n(k) = P[Z = k]$$

- Die Wahrscheinlichkeit $P_n(k)$, dass in n Versuchen k-mal eine Eins auftritt kann explizit und analytisch wie folgt angegeben werden:

$$P_n(k) = P[Z = k] = \binom{n}{k} p^k (1-p)^{n-k}$$

Damit ist das Zufallsexperiment eigentlich vollständig beschrieben. Im Folgenden wird aber die spezielle Frage diskutiert, wie das wahrscheinlichkeitstheoretische Verhalten des Bernoulli-Experiments sich für große Werte von n entwickelt und wie die Berechnung der relevanten Wahrscheinlichkeiten vereinfacht werden kann.

Beispiel 5.1 (Paketfehlerwahrscheinlichkeit)

Verwendet man zum Fehlerschutz einer digitalen Übertragung einen *Blockcode* der Länge n, so können durch diesen beispielsweise *maximal m Bitfehler korrigiert* werden. Treten mehr Bitfehler auf, so ist das komplette Datenpaket fehlerhaft.

Es soll die Wahrscheinlichkeit berechnet werden, mit der *Paketfehler* am Ausgang der Fehlerkorrektur auftreten. Dazu müssen die Wahrscheinlichkeiten für die Fälle summiert werden, in denen $m + 1$ oder mehr Bitfehler auftreten:

$$\begin{aligned} P(\text{Paketfehler}) &= 1 - P(\text{kein Paketfehler}) \\ &= 1 - P(\text{weniger als } m + 1 \text{ Bitfehler}) \\ &= 1 - \sum_{k=0}^{m} P(k \text{ Bitfehler}) \end{aligned}$$

Die Wahrscheinlichkeit für k Bitfehler wird bestimmt aus der Bitfehlerwahrscheinlichkeit p und der Zahl der Anordnungsmöglichkeiten der Bitfehler:

$$P(k \text{ Bitfehler}) = \binom{n}{k} p^k (1-p)^{n-k}$$

Die *Bitfehlerwahrscheinlichkeit* p hängt vom eingesetzten Modulationsverfahren und dem Zustand des Übertragungskanals ab. Zur Übertragung werden mehrere Bits zu einem Datenpaket der Länge n zusammengefasst. Im Folgenden sind drei bekannte Blockcodes und deren Datenpaketlänge n sowie deren Korrigierfähigkeit beispielhaft angegeben.

Code	Blocklänge n	Korrigierbare Fehler m
Hamming	31	1
Reed-Muller	64	15
BCH	127	14

Bei der praktischen Berechnung der Paketfehlerrate tritt bei großen Blocklängen n eine rein rechentechnische Besonderheit auf. Die Binomialkoeffizienten nehmen sehr große Werte an und es stellt sich die Frage der Rechengenauigkeit. Dies zeigt das folgende einfache Beispiel:

$$\binom{64}{15} \approx 1.6 \cdot 10^{14}, \; \binom{127}{14} \approx 1.6 \cdot 10^{18}$$

Daher gestaltet sich auch die Berechnung der Summe über die Einzelwahrscheinlichkeiten zur Berechnung der Paketfehlerwahrscheinlichkeit um so schwieriger:

$$P(\text{Kein Paketfehler}) = \sum_{k=0}^{15} \binom{64}{k} p^k (1-p)^{(64-k)}$$

Zur Lösung dieser Aufgabe bedient man sich eines *Grenzwertsatzes*, der im folgenden Abschnitt erläutert werden soll.

Lokaler Grenzwertsatz von Moivre-Laplace

Für ein BERNOULLI-Experiment mit insgesamt n Versuchen können die einzelnen Auftrittswahrscheinlichkeiten wie folgt angegeben werden:

$$P_n(k) = P[Z_n = k] = \binom{n}{k} p^k (1-p)^{n-k}.$$

Die Berechnung der Verteilungsfunktion eines solchen BERNOULLI-Versuchs

$$P[Z_n \leq k] = \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i}$$

kann unter praktischen Gesichtspunkten zum Teil erhebliche rechentechnische Schwierigkeiten bereiten. Aus diesem Grund ist eine Approximation, mit der die Wahrscheinlichkeitsdichte- oder Verteilungsfunktion berechnet werden kann, sehr gefragt. Diese Approximation kann durch den **lokalen Grenzwertsatz von Moivre-Laplace** hergeleitet werden.

lokaler Grenzwertsatz von Moivre-Laplace

Bei diesem Experiment mit binärem Ausgang ist der Erwartungswert der Zufallsvariablen Z_n (also beispielsweise die mittlere Fehleranzahl pro Paket) wie folgt zu berechnen:

$$\mu_n = E\{Z_n\} = n \cdot p$$

Die Varianz von Z_n, aus der das Streuverhalten der Zufallsvariablen Z_n abgelesen werden kann, ist

$$Var\{Z_n\} = n \cdot p \cdot (1-p)$$

bzw. die Standardabweichung ist:

$$\sigma_n = \sqrt{n \cdot p \cdot (1-p)}.$$

Wenn die Bedingungen des BERNOULLI-Experiments erfüllt sind, dann können die Wahrscheinlichkeiten $P_n(k)$ für große Werte n wie folgt berechnet bzw. approximiert werden:

$$\boxed{\lim_{n\to\infty} \{P_n(k)\} = \lim_{n\to\infty} \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{1}{2}\left(\frac{k-np}{\sqrt{np(1-p)}}\right)^2}}$$

Aus dieser analytischen Gleichung ist zu erkennen, dass die Paketfehlerwahrscheinlichkeiten gegen eine Gauß'sche Normalverteilung mit dem obigen Erwartungswert $E\{Z_n\} = np$ und der Varianz $Var\{Z_n\} = np(1-p)$ konvergieren. Damit sind diese Wahrscheinlichkeiten $P_n(k)$ wesentlich leichter berechenbar und direkt aus den Werten der Gauß'schen Glockenkurve ablesbar.

Beispiel 5.2 (Werfen einer Münze)

Mit Hilfe des Theorems von DeMoivre-Laplace kann man eine gute Abschätzung der Wahrscheinlichkeiten $P_n(k)$ auch bei endlicher Blocklänge n angeben:

Eine Münze wird $n = 1000$ mal geworfen. Man ermittle die Wahrscheinlichkeit, mit der bei diesem Experiment das Ereignis „Kopf" genau $k = 510$ mal auftritt. Mit

$$p = 0.5, \quad n = 1000, \quad k = 510, \quad \sqrt{np(1-p)} = 5\sqrt{10}, \quad n \cdot p = 500$$

erhält man die folgende Abschätzung

$$P_n(k) = P_{1000}(510) \approx \frac{1}{\sqrt{2\pi np(1-p)}} \cdot e^{-\frac{(510-np)^2}{2np((1-p)}} = 0,0207$$

$P_n(k) = 0.0206564$

Integralgrenzwertsatz von MOIVRE - LAPLACE

Zur Bestimmung der Wahrscheinlichkeit, mit der in einem Bernoulli Experiment die Zufallsvariable Z_n Werte in einem Interval zwischen k_1 und k_2 annimmt

$$P[k_1 \leq Z_n \leq k_2]$$

müssen die Wahrscheinlichkeiten $P_n(k)$ für die Werte k von k_1 bis k_2 (also insgesamt $k_2 - k_1 + 1$ Werte) aufsummiert werden, da die Ereignisse unabhängig voneinander sind.

$$\boxed{P[k_1 \leq Z_n \leq k_2] = \sum_{k=k_1}^{k_2} P_n(k) = \sum_{k=k_1}^{k_2} \binom{n}{k} p^k (1-p)^{n-k}}$$

Zur Lösung der Aufgabe und der Bestimmung der Wahrscheinlichkeit kann man den lokalen Grenzwertsatz erweitern, so dass unter den gleichen Voraussetzungen eines BERNOULLI Experiments für den *Integralgrenzwertsatz* gilt:

Integralwertsatz von Moivre-Laplace

$$\boxed{P[k_1 \leq Z_n \leq k_2] \approx \frac{1}{\sqrt{2\pi np(1-p)}} \int_{k_1}^{k_2} e^{-\frac{1}{2}\left(\frac{x-np}{\sqrt{np(1-p)}}\right)^2} dx} \qquad (*)$$

Die Wahrscheinlichkeit für den in Gleichung (*) gesuchten Intervalwert ergibt sich damit dann zu

$$\boxed{P[k_1 \leq Z_n \leq k_2] \approx \Phi\left(\frac{k_2 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{np(1-p)}}\right)}$$

Beispiel 5.3 (Telefonanrufe)

An einer Telefonzentrale sind 180 Telefone angeschlossen. Für jedes Telefon beträgt die Wahrscheinlichkeit, dass von ihm innerhalb von vier Stunden ein Telefonat geführt wird $p = \frac{1}{3}$. Mit dieser Angabe kann die Wahrscheinlichkeit dafür, dass insgesamt ein einziges Gespräch in 4 Stunden in der Telefonzentrale auftritt, wie folgt berechnet werden:

$$P_{180}(1) = \binom{180}{1} \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^{179} = 1.81 \cdot 10^{-30}$$

Ganz allgemein gilt: k Anrufe werden in 4 Stunden mit folgender Wahrscheinlichkeit getätigt:

$$P_{180}(k) = \binom{180}{k} \left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{180-k}.$$

Die Wahrscheinlichkeit für 50 bis 70 Anrufe innerhalb von 4 Stunden kann nach der obigen Abschätzung wie folgt berechnet werden:

$\sqrt{np(1-p)}$

$$\sum_{k=50}^{70} P_{180}(k) \approx \Phi\left(\frac{70 - \overset{n}{180} \cdot \overset{p}{\frac{1}{3}}}{\sqrt{180 \cdot \frac{1}{3} \cdot \frac{2}{3}}}\right) - \Phi\left(\frac{50-60}{\sqrt{40}}\right)$$

$$= \Phi(\sqrt{2,5}) - \Phi(-\sqrt{2,5}) = 2\Phi(\sqrt{2,5}) - 1 \approx 0,886$$

Stetigkeitskorrektur

Eine noch bessere Approximation (insbesondere für kleine Werte von n) erhält man im Integralwertsatz von Moivre-Laplace, wenn die ursprünglichen diskreten Grenzen k_1 und k_2 durch zwei neue Grenzen $(k_1 - 0,5)$ und $(k_2 + 0,5)$ ersetzt werden. Dadurch erhält das Integrationsintervall die gewollte Länge von $k_2 - k_1 + 1$. Dieser Vorgang wird auch als *Stetigkeitskorrektur* bezeichnet. Hierzu ersetzt man im Argument der Normalverteilung die Grenzen k_1 und k_2 durch $(k_1 - 0,5)$ und $(k_2 + 0,5)$:

$$P[k_1 \leq Z_n \leq k_2] \approx \Phi\left(\frac{k_2 + 0,5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k_1 - 0,5 - np}{\sqrt{np(1-p)}}\right)$$

wobei $\Phi(\cdot)$ wieder die Standard-Normalverteilung bezeichnet.

Beispiel

Wie groß ist die Wahrscheinlichkeit dafür, dass bei insgesamt 100 Würfen mit einem Würfel mindestens 10 mal und höchstens 20 mal eine Sechs auftritt? Es soll ein Näherungswert für diese Wahrscheinlichkeit angegeben werden. Die Zahl der Sechsen in 100 Würfen wird durch eine $B_{n,p}$-verteilte ($n = 100, p = 1/6$) Zufallsvariable Z_n beschrieben. Mit $k_1 = 10$ und $k_2 = 20$ ergibt sich:

1. **Approximation ohne Stetigkeitskorrektur:**

$$P[10 \leq Z_n \leq 20] \approx \Phi\left(\frac{20 - 100 \cdot \frac{1}{6}}{\sqrt{100 \cdot \frac{1}{6} \cdot \frac{5}{6}}}\right) - \Phi\left(\frac{10 - 100 \cdot \frac{1}{6}}{\sqrt{100 \cdot \frac{1}{6} \cdot \frac{5}{6}}}\right)$$
$$= \Phi(0.894) - \Phi(-1.789) = 0.777$$

2. **Approximation mit Stetigkeitskorrektur:**

$$P[10 \leq Z_n \leq 20] \approx \Phi\left(\frac{20 + 0.5 - 100 \cdot \frac{1}{6}}{\sqrt{100 \cdot \frac{1}{6} \cdot \frac{5}{6}}}\right) - \Phi\left(\frac{10 - 0.5 - 100 \cdot \frac{1}{6}}{\sqrt{100 \cdot \frac{1}{6} \cdot \frac{5}{6}}}\right)$$
$$= \Phi(1.03) - \Phi(-1.92) = 0.821$$

3. **Exakte Berechnung über Summe:**

$$P[10 \leq Z_n \leq 20] = \sum_{k=10}^{20} \binom{100}{k} \cdot \left(\frac{1}{6}\right)^k \cdot \left(\frac{5}{6}\right)^{100-k}$$
$$= 0.827$$

Anhand dieses Beispiels wird deutlich, dass sich die Güte der Approximation bei relativ kleinen n durch die Stetigkeitskorrektur deutlich verbessert.

Die hier diskutierten Ergebnisse unterstreichen die Bedeutung der Normalverteilung als Mittel zur einfachen numerischen Berechnung.

Zentraler Grenzwertsatz von Lindeberg-Lévy

Die bisherige Annahme basierte auf einer Zufallsvariablen Z_n, die einer Binomialverteilung gehorcht. In vielen anderen Anwendungsfällen ist aber lediglich bekannt, dass die Zufallsvariablen X_i, aus denen die Summe oder der arithmetische Mittelwert gebildet werden, statistisch unabhängig und identisch verteilt sind. Auch in diesem Fall stellt sich die Frage nach dem statistischen Verhalten der Zufallsvariablen Z_n:

$$Z_n = \sum_{i=1}^{n} X_i$$

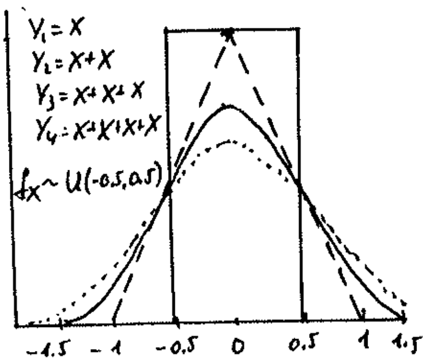
Zentraler Grenzwertsatz von LINDEBERG-LÉVY

Definition **(Zentraler Grenzwertsatz von LINDEBERG-LÉVY)** *Wenn die statistisch unabhängigen Zufallsgrößen X_i alle die gleiche (bzw. eine identische) Verteilung mit dem Mittelwert μ und der Varianz $\sigma^2 > 0$ haben, dann konvergiert die Folge der Verteilungsfunktionen der normierten Zufallsgröße*

$$Z_n = \frac{\left(\sum_{i=1}^{n} X_i\right) - n\mu}{\sqrt{n\sigma^2}}$$

gegen die Verteilungsfunktion der Standardnormalverteilung:

$$\boxed{\lim_{n\to\infty} P[Z_n \le z] = \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{\zeta^2}{2}}\, d\zeta}$$



Überlagerung identischer Verteilungen.

Zentraler Grenzwertsatz von Ljapunow

Zentraler Grenzwertsatz von LJAPUNOW

Definition **(Zentraler Grenzwertsatz von LJAPUNOW)** *Genügen die stochastisch unabhängigen Zufallsgrößen X_i jeweils einer Verteilungsfunktion mit endlichem Erwartungswert μ_i sowie einer Varianz $\sigma_i^2 > 0$ und kann man zusätzlich eine positive Zahl $\delta > 0$ mit*

$$\sigma_n^2 = \sum_{i=1}^{n} \sigma_i^2$$

so wählen, dass

$$\lim_{n\to\infty} \sum_{i=1}^{n} \frac{E\{|X_i - \mu_i|^{2+\delta}\}}{\sigma_n^{2+\delta}} = 0$$

gilt, dann konvergiert die Verteilungsfunktion der Zufallsvariablen Z_n mit:

$$Z_n = \frac{1}{\sigma_n} \sum_{i=1}^{n} (X_i - \mu_i)$$

gegen die Standardnormalverteilung.

Bedeutung des Satzes

- Die Bedingung, dass die betrachteten Zufallsvariablen X_i als identisch verteilt angenommen werden, entfällt in diesem Fall.
- Jede Zufallsvariable Z_n, die als Summe einer großen Zahl von unabhängigen Zufallsvariablen dargestellt werden kann, gehorcht also für große Werte n einer Normalverteilung.

Tschebyscheff'sche Ungleichung

In einigen Anwendungsfällen ist von einer Zufallsvariablen X nur wenig über das explizite Zufallsverhalten bekannt; insbesondere liegt die Kenntnis der Wahrscheinlichkeitsdichtefunktion nicht vor.

Wenn allerdings wenigstens der Erwartungswerts $E\{X\}$ und die Standardabweichung σ dieser Zufallsvariablen bekannt sind, dann kann man bereits eine gute Abschätzung über das wahrscheinlichkeitstheoretische Verhalten der Zufallsvariablen X berechnen. Das ist die Ausgangssituation für die Tschebyscheff'sche Ungleichung, die für alle Zufallsvariablen gilt, für die der Erwartungswert und die Standardabweichung bekannt sind. Lediglich aus der Kenntnis des Erwartungswertes und der Standarabweichung einer Zufallsvariablen X kann folgende grobe Abschätzung über das wahrscheinlichkeitstheoretische Verhalten dieser Zufallsvariablen hergeleitet werden:

Tschebyscheff'sche Ungleichung

Definition. **(Tschebyscheff'sche Ungleichung)** *Es sei X eine reelle Zufallsvariable mit dem Erwartungswert $E\{X\}$ und der Varianz σ^2. Dann gilt für alle $t > 0$*

$$\boxed{\begin{aligned} & P(|X - E\{X\}| \geq t\sigma) \leq \frac{1}{t^2} \\ \textit{bzw.}\quad & P(|X - E\{X\}| \geq t) \leq \frac{\sigma^2}{t^2} \end{aligned}}$$

Die Tschebyscheff'sche Ungleichung gilt für alle Zufallsvariablen auch bei unbekannter Wahrscheinlichkeitsdichtefunktion.

Beispiel **(Produktion von Maschinenteilen)**
In einer Fabrik für Maschinenteile hat ein bestimmtes Teil einen Solldurchmesser von 12,5 mm. Dieser Wert darf um maximal 0,2 mm über- bzw. unterschritten werden, d.h. fehlerfreie Teile haben einen Durchmesser im Intervall $[12,3; 12,7]$mm.

Von den produzierten Maschinenteilen sei bekannt, dass der Erwartungswert $E\{X\}$ sämtlicher gemessenen Durchmessers genau 12,5 mm betrüge und dass die resultierenden Abweichungen bzw. Produktionsfehler durch eine Standardabweichung von $\sigma = 0,08$mm quantitativ angegeben werden kann.

Mit Hilfe der Tschebyscheff'schen Ungleichung kann jetzt ohne explizite Kenntnis der Wahrscheinlichkeitsdichtefunktion die Wahrscheinlichkeit berechnet bzw. eine Abschätzung angegeben werden, mit denen die produzierten Maschinenteile außerhalb des geforderten Toleranzbereiches liegen.

$$P(|X - 12,5| \geq 0,2) \leq \frac{0,08^2}{0,2^2} = 0,16.$$

Maximal 16% der produzierten Maschinenteile können also außerhalb der vorgegebenen Toleranzgrenzen liegen. Aus Kenntnis des Erwartungswertes und der Standardabweichung kann aber sicher ausgesagt werden, dass mindestens 84% der Maschinenteile innerhalb der geforderten Toleranzgrenze liegen. Wäre die Wahrscheinlichkeitsdichtefunktion der Zufallsvariablen X explizit bekannt, dann könnte die Wahrscheinlichkeit, mit der die Maschinenteile innerhalb der Toleranzgrenzen liegen, genau berechnet und müsste nicht durch eine Abschätzung angegeben werden.

Beispiel . **(Anwendung der Tschebyscheff'schen Ungleichung auf die Normalverteilung)**
Für eine beispielhafte $N_{\mu,\sigma}$-normalverteilte Zufallsvariable können die in der Tschebyscheff'schen Ungleichung auftretenden Wahrscheinlichkeiten explizit wie folgt berechnet werden:

$$P(|X-\mu| \geq 2\sigma) = 2(1-\Phi(1)) \approx 0.0454 \qquad P(|X-\mu| \geq 3\sigma) = 2(1-\Phi(2)) \approx 0.0026$$

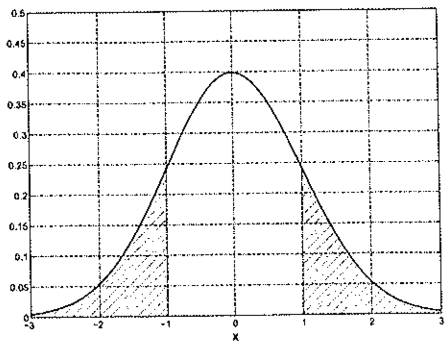


Abbildung 5.5: Beispiel: Anwendung der Tschebyscheff'schen Ungleichung auf die Normalverteilung

Mit Hilfe der Tschebyscheff'schen Ungleichung erhält man folgende für jede beliebige Zufallsvariable X gültige sowie grob genäherte Schranken (vgl. Abbildung 5.5):

$$P(|X-\mu| \geq 2\sigma) \leq \frac{\sigma^2}{4\sigma^2} = \frac{1}{4} \qquad P(|X-\mu| \geq 3\sigma) \leq \frac{\sigma^2}{49\sigma^2} = \frac{1}{9}$$

Aus diesem Beispiel ist zu erkennen, dass die Aussagen der Tschebyscheff'schen Ungleichung durchaus sehr grob sind, gemessen an der expliziten Kenntnis der Wahrscheinlichkeitsdichtefunktion. Aber die Abschätzung liefert trotzdem ein interessantes Ergebnis, weil diese Angaben für alle Zufallsvariablen X gültig sind.

Gesetz der großen Zahlen

In der Praxis ist die Verteilungsfunktion einer Zufallsvariablen oder zumindest deren Parameter (Erwartungswert, Varianz) fast immer unbekannt. Falls von einer Zufallsvariablen X mehrere Realisierungen bekannt sind, dann können die unbekannten Parameter der Verteilungsfunktion durch Anwendung der obigen Grenzwertsätze sehr gut geschätzt werden.

Um die unbekannten Parameter der Verteilungsfunktion einer Zufallsvariablen X zu bestimmen, werden mehrere Realisierungen X_i dieser Zufallsvariablen gewonnen. Aus diesen zufälligen Werten X_i sollen Parameter der Verteilungsfunktion, wie Erwartungswert und Varianz, geschätzt werden. Das Gesetz der großen Zahlen sagt beispielsweise aus, dass der Erwartungswert mit steigender Anzahl von Messwerten bzw. Realisierungen immer genauer durch das arithmetische Mittel geschätzt werden kann.

Schwaches Gesetz der großen Zahlen

Schwaches Gesetz der großen Zahlen

Definition (Schwache stochastische Konvergenz) *Eine Folge $X_n, n \in \mathbb{N}$ von Zufallsvariablen* konvergiert stochastisch *gegen eine Zufallsvariable X, falls gilt*

$$\lim_{n \to \infty} P\left(|X_n - X| \geq \epsilon\right) = 0.$$

Diese Aussage der schwachen stochastischen Konvergenz kann auf das folgende Beispiel übertragen und dort direkt angewandt werden. Es wird eine Folge $X_1, X_2, \ldots$ statistisch unabhängiger identisch verteilter Zufallsvariablen betrachtet. Daraus wird eine weitere Folge von Zufallsvariable Z_n berechnet, die jeweils den arithmetischen Mittelwert über insgesamt n Zufallsvariable X_i bildet.

$$Z_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Erwartungswert $E\{Z_n\}$ und Varianz $VAR\{Z_n\}$ dieser Zufallsfolge Z_n können wie folgt berechnet werden:

$$E\{Z_n\} = E\{X\} VAR\{Z_n\} = \frac{\sigma_X^2}{n}$$

Mit der Tschebyscheff'schen Ungleichung gilt:

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - E\{X\}\right| > \epsilon\right) = P\left(|Z_n - E\{X\}| > \epsilon\right) \leq \frac{\sigma_X^2}{n \cdot \epsilon^2}$$

Die Folge dieser Zufallsvariablen Z_n konvergiert also gegen den gemeinsamen Erwartungswert $E\{X\}$ im Sinne des obigen Konvergenzkriteriums, wenn man in der obigen Gleichung den Grenzübergang für n gegen Unendlich berechnet, d.h.

$$\boxed{\lim_{n \to \infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - E\{X\}\right| > \epsilon\right) = 0.}$$

Bei dieser stochastischen Konvergenz handelt es sich um eine schwache Form der Konvergenz, vergleichbar mit der punktweisen Konvergenz bei Funktionenfolgen in der Analysis. Auch hier sind bei großen Werten für n noch starke Ausreißer bzw. Abweichungen grundsätzlich möglich, allerdings geht deren Wahrscheinlichkeit im Grenzfall gegen Null.

Bedeutung

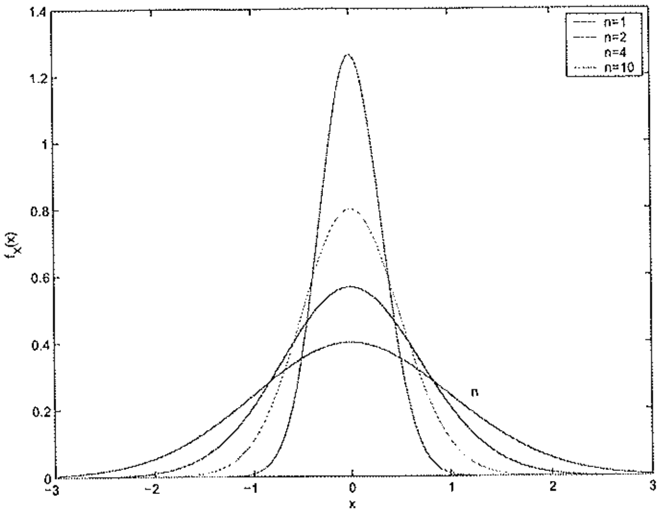


Abbildung . Mit steigender Anzahl n der Messungen sinkt die mittlere Abweichung des gemessenen arithmetischen Mittelwerts vom Erwartungswert.

Abbildung zeigt die Verteilungsfunktion der berechneten arithmetischen Mittelwert $Z_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ für unterschiedliche Werte n.
Aus dem Bild ist erkennbar, dass die resultierende Varianz mit wachsendem n gegen den Wert Null geht.
Dementsprechend sinkt auch die Wahrscheinlichkeit, mit der die Werte Z_n außerhalb einer vorgegebenen ε-Umgebung liegen. Falls der Wert n gegen unendlich geht, dann schrumpft die Verteilungsfunktion auf einen einzigen Punkt, den Erwartungswert $\mathbb{E}[X]$, der Zufallsvariablen X_i.

Die Folge der Zufallsvariablen Z_n konvergiert also stochastisch gegen den Erwartungswert $\mathbb{E}[X]$. Danach verringert sich die Wahrscheinlichkeit, mit der die Werte Z_n außerhalb einer vorgegebenen ε-Umgebung liegen kontinuierlich mit wachsendem Wert n. Für jeden Wert n kann die in Gleichung

$$\lim_{n \to \infty} P\left[\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}[X]\right| > \varepsilon\right] = 0$$

angegebene Wahrscheinlichkeit

$$P\left[\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}[X]\right| > \varepsilon\right] = P\left[\left|Z_n - \mathbb{E}[X]\right| > \varepsilon\right]$$

direkt berechnet werden. Diese so entstandene reellwertige Zahlenfolge konvergiert gegen Null.

Anschauliche Darstellung der schwachen stochastischen Konvergenz

Es werden Partikel beobachtet, die eine Blende passieren. Dazu wird ein ϵ Umgebung vorgegeben. Die Werte der Wahrscheinlichkeiten $P(|Z_n - E\{X\}| > \epsilon)$ stellen eine gegen Null konvergierende Zahlenfolge dar. Das bedeutet, die Wahrscheinlichkeit, mit der die Partikel außerhalb der ϵ Umgebung angeordnet sind, wird mit wachsendem Wert n kontinuierlich kleiner. In Abbildung sind zwei Realisierungen einer Folge Z_n zusammen mit zwei ϵ Blenden zu unterschiedlichen Zeiten n dargestellt.

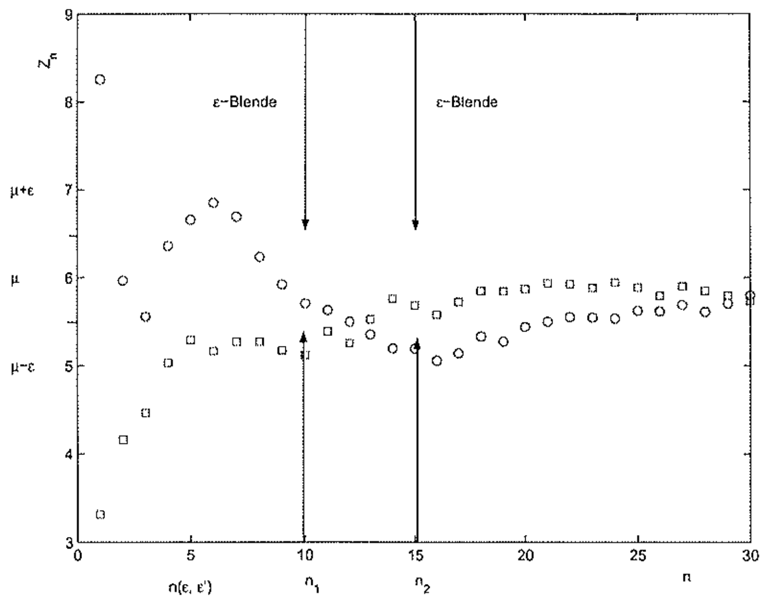


Abbildung Passieren einzelner Partikel an einer Blende

Starkes Gesetz der großen Zahlen.

Bei der schwachen stochastischen Konvergenz konnte es durchaus passieren, dass eine Zufallsvariable Z_n eine ϵ Blende zum Zeitpunkt n_1 passiert, aber zu einem späteren Zeitpunkt n_2 sich außerhalb der ϵ Blende befindet. Die Wahrscheinlichkeit für solche Ereignisse ist allerdings sehr klein.

Starkes Gesetz der großen Zahlen

Definition **(Starke stochastische Konvergenz)** *Eine Folge von Zufallsvariablen* $X_n, n \in \mathbb{N}$ konvergiert fast sicher *(mit der Wahrscheinlichkeit 1) gegen eine Zufallsvariable* X*, falls gilt*

$$P\left(\lim_{n\to\infty} |X_n - X| = 0\right) = 1.$$

Bei dieser starken stochastischen Konvergenz wird gefordert, dass ab einem gegebenen Index n *alle folgenden Zufallsvariablen* Z_n *innerhalb der* ϵ *Blende liegen müssen. Die Folge der Zufallsvariablen* Z_n *verbleibt somit in einem vorgegebenen* ϵ *Streifen.*

Dies Forderung führt auf die folgende Konvergenzbedingung:

$$P\left(\lim_{n\to\infty} |Z_n - \mu| = 0\right) = 1$$

und wird als starke stochastische Konvergenz bezeichnet.

Anschauliche Darstellung der starken Konvergenz

nächste Seite

In Abbildung (*) ist eine ϵ *Blende eingezeichnet verbunden mit der oben formulierten Forderung, dass sämtliche Zufallsvariable* Z_n *ab einem vorgegebenen Index* n *Werte innerhalb der* ϵ *Blende annehmen müssen. In Abbildung (*) sind zwei Zufallsfolgen* Z_n *basierend auf den berechneten arithmetischen Mittelwerten beispielhaft dargestellt. Mit großer Wahrscheinlichkeit verbleibt die Folge* Z_n *im Bereich eines vorgegebenen* ϵ *Streifens.*

Die aus den statistisch unabhängigen und identisch verteilten Zufallsvariablen $X_1, X_2, \ldots$ *berechnete Folge der arithmetischen Mittelwerte* Z_n,

$$Z_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

konvergiert auch nach dem Kriterium der starken stochastischen Konvergenz gegen den gemeinsamen Erwartungswert $E\{X\}$, *d.h.*

$$\boxed{P\left(\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = E\{X\}\right) = P\left(\lim_{n\to\infty} Z_n = E\{X\}\right) = 1.}$$

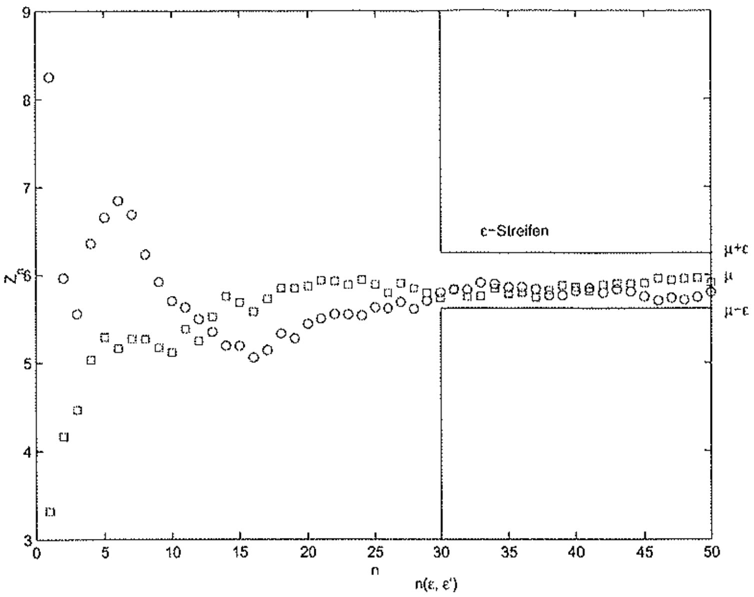


Abbildung (*) Verbleiben in einem ϵ Streifen bei der starken stochastischen Konvergenz

Zusammenfassung

- **Grenzwertsätze**

  Beschreiben die Konvergenz der Summe von unabhängigen Zufallsvariablen gegen die Normalverteilung.

  $\rightsquigarrow$ Besondere Bedeutung der Normalverteilung!

- **Gesetz der Großen Zahlen**

  Beschreibt die Konvergenz einer Zufallsfolge Z_n gegen deren Erwartungswert μ.

$$\text{schwach:} \quad \lim_{n\to\infty} P(|Z_n - \mu| < \epsilon) = 1$$

$$\text{stark:} \quad P\left(\lim_{n\to\infty} Z_n = \mu\right) = 1$$